A Level-Set Method for Magnetic Substance Simulation

XINGYU NI, CFCS, Peking University & AICFVE, Beijing Film Academy
BO ZHU, Dartmouth College
BIN WANG*, AICFVE, Beijing Film Academy
BAOQUAN CHEN*, CFCS, Peking University & AICFVE, Beijing Film Academy

Fig. 1. Our unified level-set based approach can simulate and visualize the dynamics of a broad array of magnetic phenomena including ferrofluids, deformable magnetic bodies, rigid magnetic bodies, and multi-physics interactions.

We present a versatile numerical approach to simulating various magnetic phenomena using a level-set method. At the heart of our method lies a novel two-way coupling mechanism between a magnetic field and a magnetizable mechanical system, which is based on the interfacial Helmholtz force drawn from the Minkowski form of the Maxwell stress tensor. We show that a magnetic-mechanical coupling system can be solved as an interfacial problem, both theoretically and computationally. In particular, we employ a Poisson equation with a jump condition across the interface to model the mechanical-to-magnetic interaction and a Helmholtz force on the free surface to model the magnetic-to-mechanical effects. Our computational framework can be easily integrated into a standard Euler fluid solver, enabling both simulation and visualization of a complex magnetic field and its interaction with immersed magnetizable objects in a large domain. We demonstrate the efficacy of our method through an array of magnetic substance simulations that exhibit rich geometric and dynamic characteristics, encompassing ferrofluid, rigid magnetic body, deformable magnetic body, and multi-phase couplings.

CCS Concepts: • **Computing methodologies** → **Physical simulation**; • **Applied computing** → *Physics.*

*corresponding authors

Authors' addresses: Xingyu Ni, nixy@pku.edu.cn, CFCS, Peking University & AICFVE, Beijing Film Academy; Bo Zhu, bo.zhu@dartmouth.edu, Dartmouth College; Bin Wang, binwangbuaa@gmail.com, AICFVE, Beijing Film Academy; Baoquan Chen, baoquan@pku.edu.cn, CFCS, Peking University & AICFVE, Beijing Film Academy.

0730-0301/2020/7-ART29 $15.00
https://doi.org/10.1145/3386569.3392445

Additional Key Words and Phrases: magnetic simulation, fluid simulation, ferrofluid, level-set method, immersed boundary

ACM Reference Format:
Xingyu Ni, Bo Zhu, Bin Wang, and Baoquan Chen. 2020. A Level-Set Method for Magnetic Substance Simulation. *ACM Trans. Graph.* 39, 4, Article 29 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392445

1 INTRODUCTION

The coupling between volumetric and interfacial forces acts as the fundamental mechanism for many intricate free-surface flow phenomena that are characterized by visually appealing dynamics and geometries. Among these phenomena, the surface tension flow is the most ubiquitous example, demonstrating the beauty and complexity of such interface-volume interactions. A variety of small-scale features, such as the pinched off droplets [Da et al. 2016a; O'brien and Hodgins 1995; Thürey et al. 2010; Zheng et al. 2015; Zhu et al. 2014], filaments [Bergou et al. 2010], curved thin sheets [Ando and Tsuruno 2011; Batty et al. 2012; Brochu et al. 2012; Da et al. 2014, 2015; Larionov et al. 2017; Saye and Sethian 2013], capillary waves [He et al. 2012; Jeschke and Wojtan 2015a; Saye 2016; Yang et al. 2016], and their co-dimensional combinations [Zhu et al. 2015, 2014], have been captured numerically by the invention of a broad spectrum of computational tools to accommodate the modeling of free-surface flow in computational physics and computer graphics. Among these surface-tension-driven phenomena, magnetic flow exhibits its peculiar surface geometries and dynamics featured by the emergence and evolution of arrays of uniform and sharp cone structures. These appealing features arise due to the multilateral interactions among pressure, surface tension, and magnetic forces.

A natural and immediate question to ask when extending a conventional surface tension solver to model a magnetic flow phenomenon is that, *"Is the magnetic force exerted on a physical substance*

A Massively Parallel and Scalable Multi-GPU Material Point Method

XINLEI WANG*, Zhejiang University and University of Pennsylvania
YUXING QIU*, University of California, Los Angeles and University of Pennsylvania
STUART R. SLATTERY, Oak Ridge National Laboratory
YU FANG, University of Pennsylvania
MINCHEN LI, University of Pennsylvania
SONG-CHUN ZHU, University of California, Los Angeles
YIXIN ZHU, University of California, Los Angeles
MIN TANG, Zhejiang University
DINESH MANOCHA, University of Maryland
CHENFANFU JIANG, University of Pennsylvania

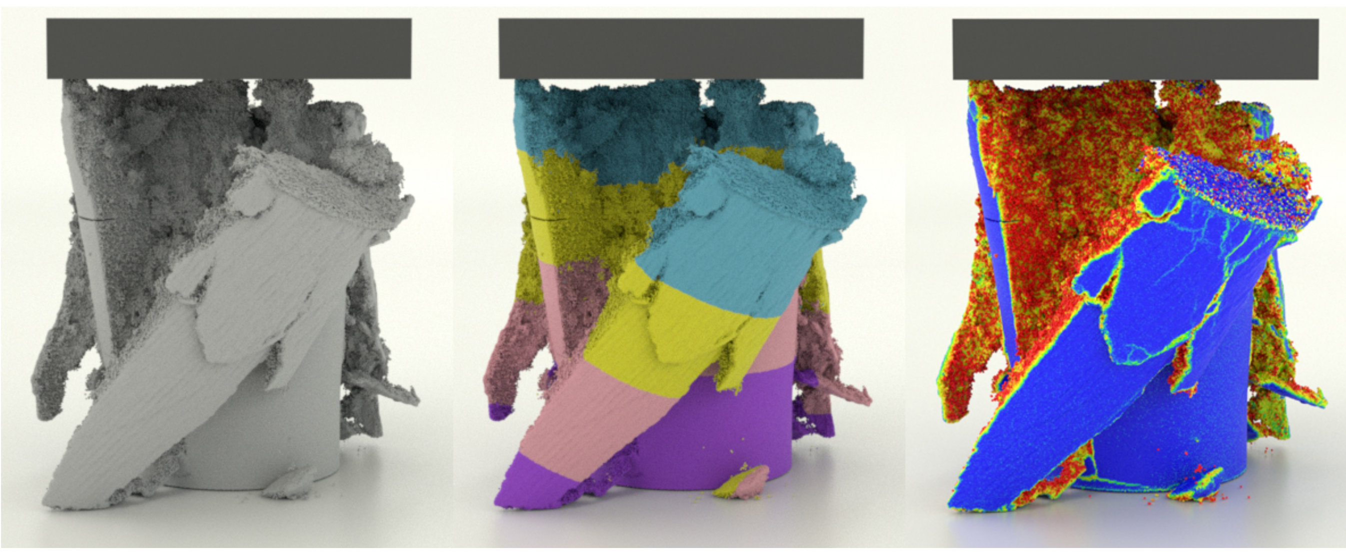


Fig. 1. **Crushing concrete.** Our system enables this concrete crushing simulation (inspired by the hydraulic press) on a single workstation with 4 NVIDIA Quadro P6000 GPUs. This simulation contains 93.8 million particles on a 1024^3 grid, achieving a 3.9 min/frame performance. (Left) A concrete-style render. (Middle) Coloring by GPU. (Right) Coloring by the plastic volumetric strain for visualizing the damage propagation.

Harnessing the power of modern multi-GPU architectures, we present a massively parallel simulation system based on the Material Point Method (MPM) for simulating physical behaviors of materials undergoing complex topological changes, self-collision, and large deformations. Our system makes three critical contributions. First, we introduce a new particle data structure that promotes coalesced memory access patterns on the GPU and eliminates the need for complex atomic operations on the memory hierarchy when writing particle data to the grid. Second, we propose a kernel fusion approach using a new Grid-to-Particles-to-Grid (*G2P2G*) scheme, which efficiently reduces GPU kernel launches, improves latency, and significantly reduces the amount of global memory needed to store particle data. Finally, we introduce optimized algorithmic designs that allow for efficient sparse grids in a shared memory context, enabling us to best utilize modern multi-GPU computational platforms for hybrid Lagrangian-Eulerian computational patterns. We demonstrate the effectiveness of our method with extensive benchmarks, evaluations, and dynamic simulations with elastoplasticity, granular media, and fluid dynamics. In comparisons against an open-source and heavily optimized CPU-based MPM codebase [Fang et al. 2019] on an elastic sphere colliding scene with particle counts ranging from 5 to 40 million, our GPU MPM achieves over 100× per-time-step speedup on a workstation with an Intel 8086K CPU and a single Quadro P6000 GPU, exposing exciting possibilities for future MPM simulations in computer graphics and computational science. Moreover, compared to the state-of-the-art GPU MPM method [Hu

*equal contributions

Authors' addresses: Xinlei Wang, Zhejiang University and University of Pennsylvania; Yuxing Qiu, University of California, Los Angeles and University of Pennsylvania; Stuart R. Slattery, Oak Ridge National Laboratory; Yu Fang, University of Pennsylvania; Minchen Li, University of Pennsylvania; Song-Chun Zhu, University of California, Los Angeles; Yixin Zhu, University of California, Los Angeles; Min Tang, Zhejiang University; Dinesh Manocha, University of Maryland; Chenfanfu Jiang, University of Pennsylvania.

0730-0301/2020/7-ART30 $15.00
https://doi.org/10.1145/3386569.3392442

A Model for Soap Film Dynamics with Evolving Thickness

SADASHIGE ISHIDA* and PETER SYNAK*, IST Austria
FUMIYA NARITA, Unaffiliated
TOSHIYA HACHISUKA, The University of Tokyo
CHRIS WOJTAN, IST Austria

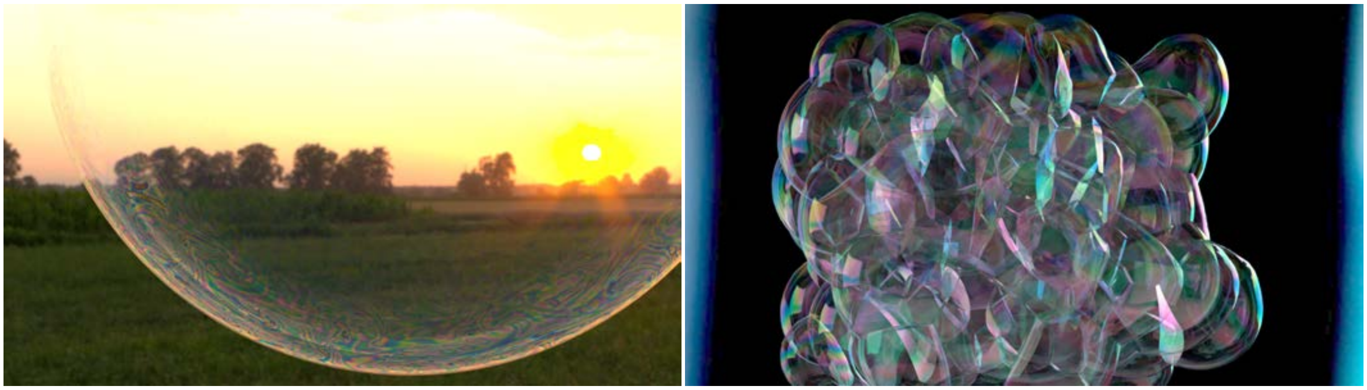


Fig. 1. Our model simulates the evolution of soap films, leading to detailed advection patterns (left) and interplays between draining, evaporation, capillary waves, and ruptures in a foam (right).

Previous research on animations of soap bubbles, films, and foams largely focuses on the motion and geometric shape of the bubble surface. These works neglect the evolution of the bubble's thickness, which is normally responsible for visual phenomena like surface vortices, Newton's interference patterns, capillary waves, and deformation-dependent rupturing of films in a foam. In this paper, we model these natural phenomena by introducing the film thickness as a reduced degree of freedom in the Navier-Stokes equations and deriving their equations of motion. We discretize the equations on a non-manifold triangle mesh surface and couple it to an existing bubble solver. In doing so, we also introduce an incompressible fluid solver for 2.5D films and a novel advection algorithm for convecting fields across non-manifold surface junctions. Our simulations enhance state-of-the-art bubble solvers with additional effects caused by convection, rippling, draining, and evaporation of the thin film.

CCS Concepts: • **Mathematics of computing** → *Geometry*; *Partial differential equations*; • **Computing methodologies** → *Physics*; *Animation.*

Additional Key Words and Phrases: Soap films, Fluid dynamics, Physical modeling

*Joint first authors.

Authors' addresses: Sadashige Ishida; Peter Synak, IST Austria; Fumiya Narita, Unaffiliated; Toshiya Hachisuka, The University of Tokyo; Chris Wojtan, IST Austria.

0730-0301/2020/7-ART31 $15.00
https://doi.org/10.1145/3386569.3392405

ACM Reference Format:
Sadashige Ishida, Peter Synak, Fumiya Narita, Toshiya Hachisuka, and Chris Wojtan. 2020. A Model for Soap Film Dynamics with Evolving Thickness. *ACM Trans. Graph.* 39, 4, Article 31 (July 2020), 11 pages. https://doi.org/10.1145/3386569.3392405

1 INTRODUCTION

This paper concerns the animation of soap films, bubbles, and foams. These natural phenomena exhibit fascinating and beautiful complexity in their geometry, dynamics, and color. Interestingly, the influence of fluid forces cascades all the way down to their surface appearance — surface tension and body forces change the film's curvature and thickness, and the film thickness in turn causes swirling interference patterns when it interacts with light. The typical approach to animating these phenomena is to first use a surface-tension solver to simulate dynamic foams while holding the thickness constant, and then retroactively model a noisy film thickness in a surface shader during the rendering step. However, we argue that several interesting phenomena are missed by not directly simulating the evolution of the film thickness. In particular, dynamic film thickness is responsible for the appearance of swirling vortices, ripple patterns, gravity-dependent thickness variation, as well as the bursting of bubbles and thus the ultimate shape of large foam structures.

In this paper, we propose to model soap film thickness within a bubble simulation. We introduce the film thickness as a reduced degree of freedom in the Navier-Stokes equations and derive the relevant equations of motion. These thickness dynamics result in a physical model for film advection, mass conservation, draining, evaporation, and surface tension ripples. We discretize the equations on a non-manifold triangle mesh surface and couple it to an existing

A Practical Octree Liquid Simulator with Adaptive Surface Resolution

RYOICHI ANDO, National Institute of Informatics, Tokyo
CHRISTOPHER BATTY, University of Waterloo, Canada

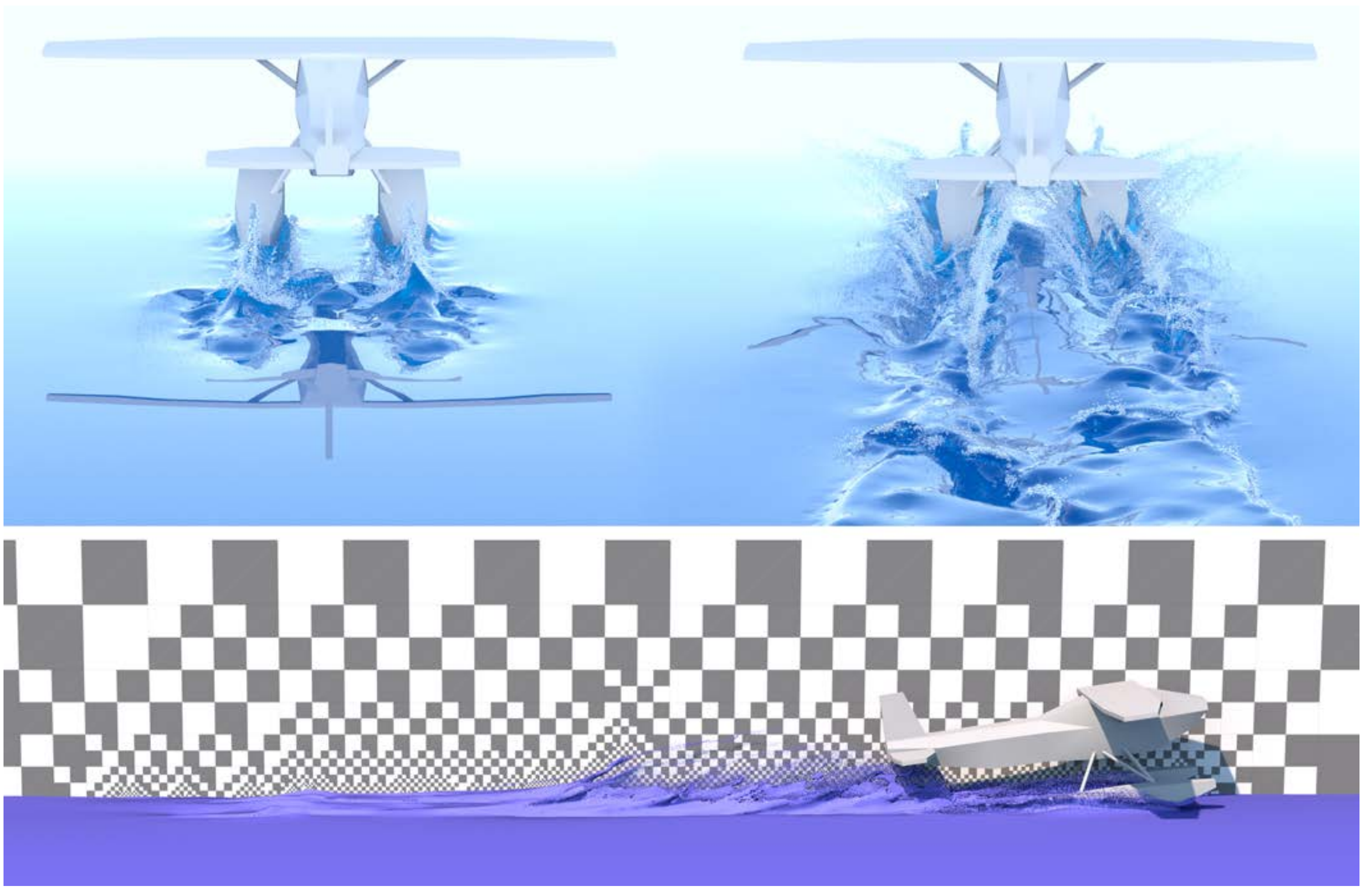


Fig. 1. Top: a seaplane (wingspan: 11m; height: 4.3m) takes off from a lake (30m depth) at high speed. Effective resolution: $1024 \times 1024 \times 512$. Compute time: 2.85 minutes per video frame. Bottom: the background octree cells, exhibiting a wide range of resolutions.

We propose a new adaptive liquid simulation framework that achieves highly detailed behavior with reduced implementation complexity. Prior work has shown that spatially adaptive grids are efficient for simulating large-scale liquid scenarios, but in order to enable adaptivity *along the liquid surface* these methods require either expensive boundary-conforming (re-)meshing or elaborate treatments for second order accurate interface conditions. This complexity greatly increases the difficulty of implementation and maintainability, potentially making it infeasible for practitioners. We therefore present new algorithms for adaptive simulation that are comparatively easy to implement yet efficiently yield high quality results. First, we develop a novel staggered octree Poisson discretization for free surfaces that is second order in pressure and gives smooth surface motions even across octree T-junctions, without a power/Voronoi diagram construction. We augment this discretization with an adaptivity-compatible surface tension force that likewise supports T-junctions. Second, we propose a moving least squares strategy for level set and velocity interpolation that requires minimal knowledge of the local tree structure while blending near-seamlessly with standard trilinear interpolation in uniform regions. Finally, to maximally exploit the flexibility of our new surface-adaptive solver, we propose several novel extensions to sizing function design that enhance its effectiveness and flexibility. We perform a range of rigorous numerical experiments to evaluate the reliability and limitations of our method, as well as demonstrating it on several complex high-resolution liquid animation scenarios.

Authors' addresses: Ryoichi Ando, National Institute of Informatics, Tokyo, rand@nii.ac.jp; Christopher Batty, University of Waterloo, Canada, christopher.batty@uwaterloo.ca.

0730-0301/2020/7-ART32 $15.00
https://doi.org/10.1145/3386569.3392460

A Scalable Approach to Control Diverse Behaviors for Physically Simulated Characters

JUNGDAM WON, Facebook AI Research
DEEPAK GOPINATH, Facebook AI Research
JESSICA HODGINS, Facebook AI Research

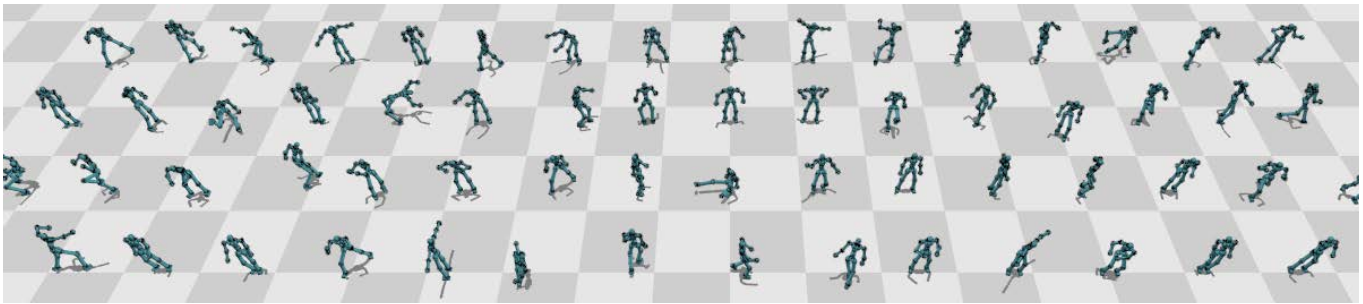


Fig. 1. Using a database with many, heterogeneous motion clips, our framework learns a single dynamic controller that generates a large variety of motions.

Human characters with a broad range of natural looking and physically realistic behaviors will enable the construction of compelling interactive experiences. In this paper, we develop a technique for learning controllers for a large set of heterogeneous behaviors. By dividing a reference library of motion into clusters of like motions, we are able to construct *experts*, learned controllers that can reproduce a simulated version of the motions in that cluster. These experts are then combined via a second learning phase, into a general controller with the capability to reproduce any motion in the reference library. We demonstrate the power of this approach by learning the motions produced by a motion graph constructed from eight hours of motion capture data and containing a diverse set of behaviors such as dancing (ballroom and breakdancing), Karate moves, gesturing, walking, and running.

CCS Concepts: • **Computing methodologies** → **Animation**; *Physical simulation*; *Reinforcement learning*; *Neural networks.*

Additional Key Words and Phrases: Character Animation, Physics-based Simulation and Control, Reinforcement Learning, Deep Learning, Neural Network, Locomotion Control

ACM Reference Format:
Jungdam Won, Deepak Gopinath, and Jessica Hodgins. 2020. A Scalable Approach to Control Diverse Behaviors for Physically Simulated Characters. *ACM Trans. Graph.* 39, 4, Article 33 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392381

Authors' addresses: Jungdam Won, Facebook AI Research; Deepak Gopinath, Facebook AI Research; Jessica Hodgins, Facebook AI Research.

0730-0301/2020/7-ART33 $15.00
https://doi.org/10.1145/3386569.3392381

1 INTRODUCTION

Human characters that can move and behave naturally in real-time are required if we are to create compelling populated virtual worlds. Because human behaviors are governed by physical laws, we can ensure physical realism in human motion and interactions with objects and the environment by incorporating the laws of physics into the motion generation process through simulation. However, physics is not a sufficient constraint to guarantee the naturalness of human behaviors. Learning natural control policies for physically simulated humanoid characters is challenging because the characters are under-actuated (there are more degrees of freedom (DOF) than controllable DOFs) and because there are many physically correct ways to perform a task or behavior that do not appear natural (for example, requiring unrealistically high torques or employing strategies that a human would not use).

Recently, imitation learning approaches that provide an effective way to generate natural looking motions for physically simulated humanoid characters have been developed. The core idea is to learn a control policy that allows a simulated character to successfully imitate selected motion capture clips. With recent advances in deep reinforcement learning, these approaches are able to reproduce realistic and diverse behaviors in a physically realistic and natural looking manner (see, for example, [Bergamin et al. 2019; Liu and Hodgins 2017, 2018; Park et al. 2019; Peng et al. 2018]). However, they are still limited in their ability to scale to large, heterogeneous datasets. Current approaches only learn from a relatively small corpus of homogeneous data, where each learned policy is only effective on a limited set of behaviors. Combining together multiple learned behaviors to create a character with a broad repertoire of behaviors is not a solved problem.

In this paper, we propose a control framework for physically simulated humanoid characters that can scale to model many hours of human motion data and generate motions that look natural. The input of our framework is a motion controller, which can be a motion

Accurate Face Rig Approximation with Deep Differential Subspace Reconstruction

STEVEN L. SONG*, Blue Sky Studios
WEIQI SHI*, Yale University
MICHAEL REED, Blue Sky Studios

Fig. 1. Our rig approximation method learns localized shape information in differential coordinates and, separately, a subspace for mesh reconstruction.

To be suitable for film-quality animation, rigs for character deformation must fulfill a broad set of requirements. They must be able to create highly stylized deformation, allow a wide variety of controls to permit artistic freedom, and accurately reflect the design intent. Facial deformation is especially challenging due to its nonlinearity with respect to the animation controls and its additional precision requirements, which often leads to highly complex face rigs that are not generalizable to other characters. This lack of generality creates a need for approximation methods that encode the deformation in simpler structures. We propose a rig approximation method that addresses these issues by learning localized shape information in differential coordinates and, separately, a subspace for mesh reconstruction. The use of differential coordinates produces a smooth distribution of errors in the resulting deformed surface, while the learned subspace provides constraints that reduce the low frequency error in the reconstruction. Our method can reconstruct both face and body deformations with high fidelity and does not require a set of well-posed animation examples, as we demonstrate with a variety of production characters.

CCS Concepts: • **Computing methodologies → Machine learning**; **Animation**.

Additional Key Words and Phrases: rigging, deep learning, facial animation

ACM Reference Format:
Steven L. Song, Weiqi Shi, and Michael Reed. 2020. Accurate Face Rig Approximation with Deep Differential Subspace Reconstruction. *ACM Trans. Graph.* 39, 4, Article 34 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392491

*Authors contributed equally.

Authors' addresses: Steven L. Song, stevens@blueskystudios.com, Blue Sky Studios, 1 American Ln, Greenwich, CT, 06831; Weiqi Shi, weiqi.shi@yale.edu, Yale University, New Haven, CT, 06520; Michael Reed, reed@blueskystudios.com, Blue Sky Studios, 1 American Ln, Greenwich, CT, 06831.

0730-0301/2020/7-ART34 $15.00
https://doi.org/10.1145/3386569.3392491

1 INTRODUCTION

Film-quality character rigs rely on a complex hierarchy of procedural deformers, driven by a large number of animation controls, that map to the deformation of the vertices of a character's surface mesh. Because the characters are subject to high aesthetic standards, and the rigs are the primary means by which the animators interact with them, the rigs themselves have strict performance requirements: the character's skin must behave predictably and precisely over the entire range of control, which for animated characters can be extreme because of the caricatured design and motion.

Adaptive Merging for Rigid Body Simulation

EULALIE COEVOET, OTMAN BENCHEKROUN, and PAUL G. KRY, McGill University

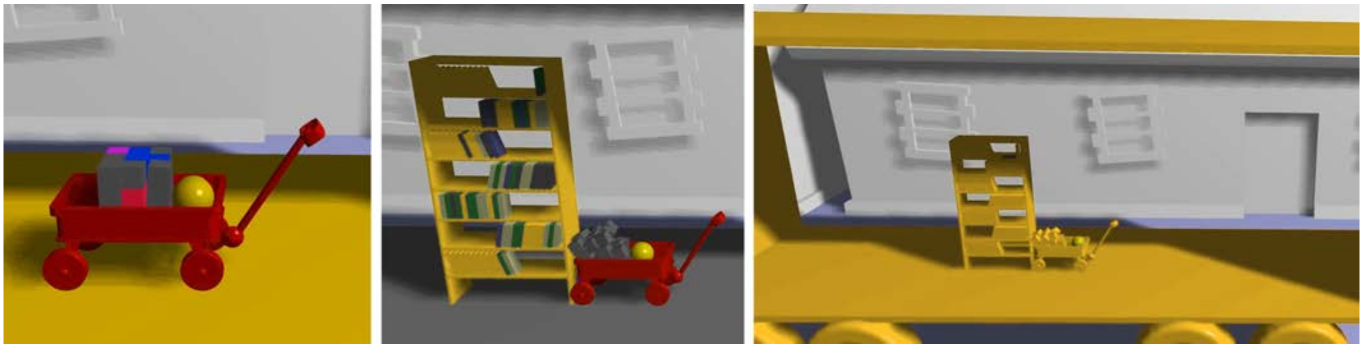


Fig. 1. Simulation of a moving wagon filled with toys, colliding with a bookcase full of books, while being transported in the back of a moving truck. Collections of bodies with common relative velocity merge and simulate as one, displayed here in different bright colors. Only a fraction of the bodies in the scene needs to be simulated at any given time step, with bodies merging into and out of collections as necessary due to external forces, and interactions (e.g., when the wagon collides with the bookcase).

We reduce computation time in rigid body simulations by merging collections of bodies when they share a common spatial velocity. Merging relies on monitoring the state of contacts, and a metric that compares the relative linear and angular motion of bodies based on their sizes. Unmerging relies on an inexpensive single iteration projected Gauss-Seidel sweep over contacts between merged bodies, which lets us update internal contact forces over time, and use the same metrics as merging to identify when bodies should unmerge. Furthermore we use a contact ordering for graph traversal refinement of the internal contact forces in collections, which helps to correctly identify all the bodies that must unmerge when there are impacts. The general concept of merging is similar to the common technique of sleeping and waking rigid bodies in the inertial frame, and we exploit this too, but our merging is in moving frames, and unmerging takes place at contacts between bodies rather than at the level of bodies themselves. We discuss the previous relative motion metrics in comparison to ours, and evaluate our method on a variety of scenarios.

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: merging, sleeping, constraints, contact, friction

ACM Reference Format:
Eulalie Coevoet, Otman Benchekroun, and Paul G. Kry. 2020. Adaptive Merging for Rigid Body Simulation. *ACM Trans. Graph.* 39, 4, Article 35 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392417

Authors' address: Eulalie Coevoet, eulalie.coevoet@gmail.com; Otman Benchekroun, otmanbench@gmail.com; Paul G. Kry, kry@cs.mcgill.ca, McGill University.

0730-0301/2020/7-ART35 $15.00
https://doi.org/10.1145/3386569.3392417

1 INTRODUCTION

In a multi-body system with contacts, the motion of each body is dependant on external forces and the contacts that exist between bodies. Determining where the contacts occur and the forces they produce are the two most computationally expensive steps of a rigid-body simulation. Methods to speed-up these computations while maintaining plausible (or correct) motion are valuable as it allows the complexity of the scene to be increased, and this is of particular interest in the context of interactive simulations.

Sleeping and waking is a popular mechanism in most physics engines to reduce the compute time by not performing collision detection between sleeping bodies, and avoiding the calculation of their resting contact forces. A sleeping body, and typically the neighboring bodies in contact, will wake upon receiving an external force, making it still possible for a user to freely interact with all bodies, sleeping or otherwise, in a large simulated world. This can be very effective for eliminating unnecessary computation. However, sleeping alone misses an important opportunity: allowing large groups of bodies to sleep relative to one another, even if they are moving in the inertial frame. We call this *adaptive merging*, and it can arise, for instance, when a vertical stack of books is pushed at its bottom-most book; the merged collection of bodies can slide together as one merged object on a flat surface in the simulation. Merging can also occur in scenarios where bodies share a common rotational velocity, for example, with objects on a swing, or stacks of equipment on a rotating space station. Figure 1 shows an elaborate scenario involving many books and toys on different moving platforms, which we simulate efficiently, with collections of objects unmerging only when necessary.

Merging allows us to bypass collision detection for bodies that are members of the same collection. Contact force computations are likewise only necessary between collections and active bodies in the scene. To make this all possible, and to allow bodies to unmerge,

An Implicit Compressible SPH Solver for Snow Simulation

CHRISTOPH GISSLER, University of Freiburg, Germany and FIFTY2 Technology GmbH, Germany
ANDREAS HENNE, FIFTY2 Technology GmbH, Germany
STEFAN BAND, University of Freiburg, Germany
ANDREAS PEER, FIFTY2 Technology GmbH, Germany
MATTHIAS TESCHNER, University of Freiburg, Germany

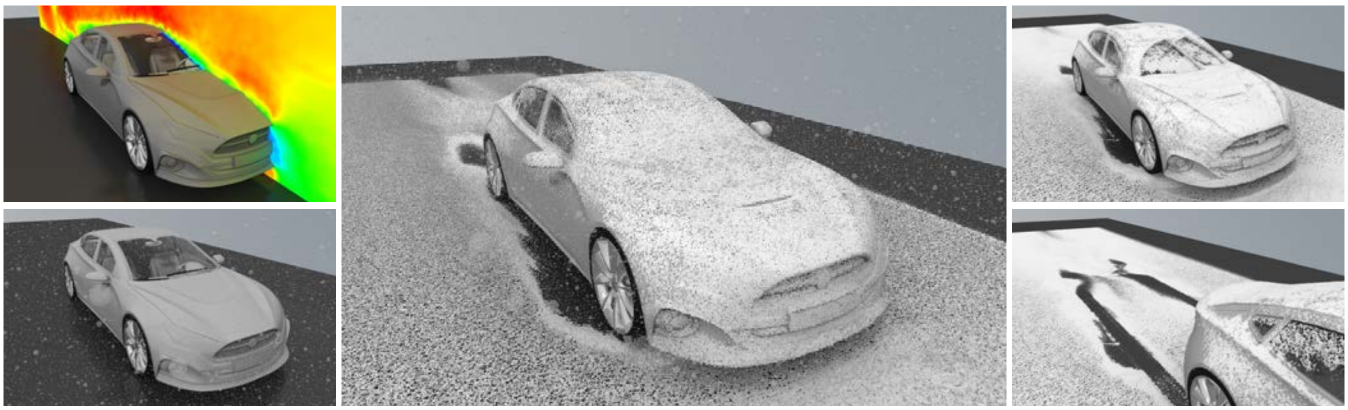


Fig. 1. Snow fall onto a car is simulated with single snow flakes modeled as single particles. The particles are coupled to a precomputed airflow field (visualized in the top left). The snow accumulates and is wiped away by moving windshield wipers. The car drives away in the end, compressing the snow below its tires.

Snow is a complex material. It resists elastic normal and shear deformations, while some deformations are plastic. Snow can deform and break. It can be significantly compressed and gets harder under compression. Existing snow solvers produce impressive results. E.g., hybrid Lagrangian/Eulerian techniques have been used to capture all material properties of snow. The auxiliary grid, however, makes it challenging to handle small volumes. In particular, snow fall and accumulation on surfaces have not been demonstrated with these solvers yet. Existing particle-based snow solvers, on the other hand, can naturally handle small snow volumes. However, existing solutions consider simplified material properties. In particular, shear deformation and the hardening effect are typically omitted.

We present a novel Lagrangian snow approach based on Smoothed Particle Hydrodynamics (SPH). Snow is modeled as an elastoplastic continuous material that captures all above-mentioned effects. The compression of snow is handled by a novel compressible pressure solver, where the typically employed state equation is replaced by an implicit formulation. Acceleration due to shear stress is computed using a second implicit formulation. The linear solvers of the two implicit formulations for accelerations due to shear and normal stress are realized with matrix-free implementations. Using implicit formulations and solving them with matrix-free solvers allows to couple the snow to other phases and is beneficial to the stability and the time step size, i.e., performance of the approach. Solid boundaries are represented with particles and a novel implicit formulation is used to handle friction at solid boundaries. We show that our approach can simulate accumulation, deformation, breaking, compression and hardening of snow. Furthermore, we demonstrate two-way coupling with rigid bodies, interaction with incompressible and highly viscous fluids and phase change from fluid to snow.

Authors' addresses: Christoph Gissler, University of Freiburg, Georges-Köhler-Allee 52, Freiburg im Breisgau, 79110, Germany, FIFTY2 Technology GmbH, Tullastraße 80, Freiburg im Breisgau, 79108, Germany, gisslerc@informatik.uni-freiburg.de; Andreas Henne, FIFTY2 Technology GmbH, Tullastraße 80, Freiburg im Breisgau, 79108, Germany, andreas.henne@fifty.eu; Stefan Band, University of Freiburg, Georges-Köhler-Allee 52, Freiburg im Breisgau, 79110, Germany, bands@informatik.uni-freiburg.de; Andreas Peer, FIFTY2 Technology GmbH, Tullastraße 80, Freiburg im Breisgau, 79108, Germany, andreas.peer@fifty.eu; Matthias Teschner, University of Freiburg, Georges-Köhler-Allee 52, Freiburg im Breisgau, 79110, Germany, teschner@informatik.uni-freiburg.de.

0730-0301/2020/7-ART36 $15.00
https://doi.org/10.1145/3386569.3392431

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: Implicit solver, physically-based animation, smoothed particle hydrodynamics, snow, two-way coupling

ACM Reference Format:
Christoph Gissler, Andreas Henne, Stefan Band, Andreas Peer, and Matthias Teschner. 2020. An Implicit Compressible SPH Solver for Snow Simulation. *ACM Trans. Graph.* 39, 4, Article 36 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392431

1 INTRODUCTION

Smoothed Particle Hydrodynamics (SPH) is mostly known in the computer graphics community for the simulation of fluids, e.g., as

AnisoMPM: Animating Anisotropic Damage Mechanics

JOSHUAH WOLPER, University of Pennsylvania
YUNUO CHEN, University of Pennsylvania
MINCHEN LI, University of Pennsylvania
YU FANG, University of Pennsylvania
ZIYIN QU, University of Pennsylvania
JIECONG LU, University of Pennsylvania
MEGGIE CHENG, University of Pennsylvania
CHENFANFU JIANG, University of Pennsylvania

Fig. 1. **Anisotropic Feast.** (Left) We tear raw meat to reveal intricate guided fracture along its grain. (Top R.) An orange slice is torn to illustrate the the beautiful radial anisotropy of the fruit. (Bot. R.) A cheese stick is peeled apart to demonstrate the underlying stringy fiber structures intrinsic to mozzarella.

Dynamic fracture surrounds us in our day-to-day lives, but animating this phenomenon is notoriously difficult and only further complicated by anisotropic materials—those with underlying structures that dictate preferred fracture directions. Thus, we present AnisoMPM: a robust and general approach for animating the dynamic fracture of isotropic, transversely isotropic, and orthotropic materials. AnisoMPM has three core components: a technique for anisotropic damage evolution, methods for anisotropic elastic response, and a coupling approach. For anisotropic damage, we adopt a non-local continuum damage mechanics (CDM) geometric approach to crack modeling and augment this with structural tensors to encode material anisotropy. Furthermore, we discretize our damage evolution with explicit and implicit integration, giving a high degree of computational efficiency and flexibility. We also utilize a QR-decomposition based anisotropic constitutive model that is inversion safe, more efficient than SVD models, easy to implement, robust to extreme deformations, and that captures all aforementioned modes of anisotropy. Our elasto-damage coupling is enforced through an additive decomposition of our hyperelasticity into a tensile and compressive component in which damage is used to degrade the tensile contribution to allow for material separation. For extremely stiff fibered materials, we further introduce a novel Galerkin weak form discretization that enables embedded directional inextensibility. We present this as a hard-constrained grid velocity solve that poses an alternative to our anisotropic elasticity that is locking-free and can model very stiff materials.

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: Numerical methods, MPM, damage mechanics, fracture, topology change, VFX

ACM Reference Format:
Joshuah Wolper, Yunuo Chen, Minchen Li, Yu Fang, Ziyin Qu, Jiecong Lu, Meggie Cheng, and Chenfanfu Jiang. 2020. AnisoMPM: Animating

Authors' addresses: Joshuah Wolper, University of Pennsylvania; Yunuo Chen, University of Pennsylvania; Minchen Li, University of Pennsylvania; Yu Fang, University of Pennsylvania; Ziyin Qu, University of Pennsylvania; Jiecong Lu, University of Pennsylvania; Meggie Cheng, University of Pennsylvania; Chenfanfu Jiang, University of Pennsylvania.

0730-0301/2020/7-ART37 $15.00
https://doi.org/10.1145/3386569.3392428

CARL: Controllable Agent with Reinforcement Learning for Quadruped Locomotion

YING-SHENG LUO*, Inventec Corp., Taiwan
JONATHAN HANS SOESENO*, Inventec Corp., Taiwan
TRISTA PEI-CHUN CHEN, Inventec Corp., Taiwan
WEI-CHAO CHEN, Skywatch Innovation Inc. and Inventec Corp., Taiwan

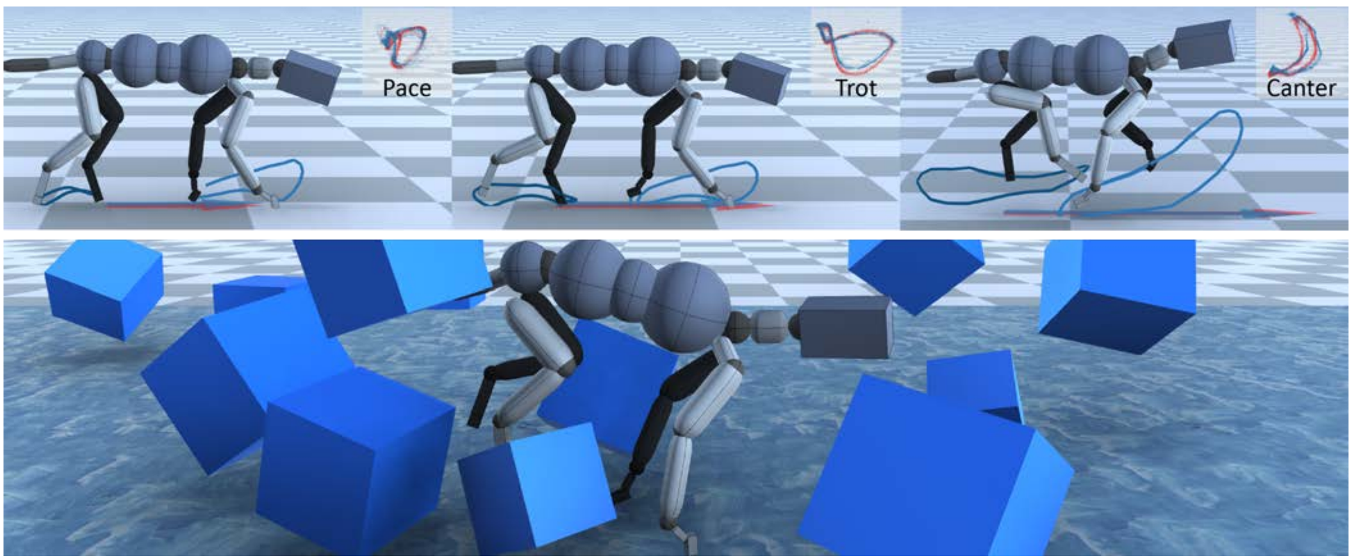


Fig. 1. Our controllable agent produces natural movements (blue) alike depicted in the reference motion clip (red) and adapt to external perturbations.

Motion synthesis in a dynamic environment has been a long-standing problem for character animation. Methods using motion capture data tend to scale poorly in complex environments because of their larger capturing and labeling requirement. Physics-based controllers are effective in this regard, albeit less controllable. In this paper, we present CARL, a quadruped agent that can be controlled with high-level directives and react naturally to dynamic environments. Starting with an agent that can imitate individual animation clips, we use Generative Adversarial Networks to adapt high-level controls, such as speed and heading, to action distributions that correspond to the original animations. Further fine-tuning through the deep reinforcement learning enables the agent to recover from unseen external perturbations while producing smooth transitions. It then becomes straightforward to create autonomous agents in dynamic environments by adding navigation modules over the entire process. We evaluate our approach by measuring the agent's ability to follow user control and provide a visual analysis of the generated motion to show its effectiveness.

CCS Concepts: • **Computing methodologies → Animation**; *Physical simulation*; *Reinforcement learning*.

Additional Key Words and Phrases: deep reinforcement learning (DRL), generative adversarial network (GAN), motion synthesis, locomotion, quadruped

ACM Reference Format:
Ying-Sheng Luo, Jonathan Hans Soeseno, Trista Pei-Chun Chen, and Wei-Chao Chen. 2020. **CARL**: **C**ontrollable **A**gent with **R**einforcement **L**earning for Quadruped Locomotion. *ACM Trans. Graph.* 39, 4, Article 38 (July 2020), 10 pages. https://doi.org/10.1145/3386569.3392433

*Joint first authors.

Authors' addresses: Ying-Sheng Luo, luo.ying-sheng@inventec.com, Inventec Corp., Taiwan; Jonathan Hans Soeseno, soeseno.jonathan@inventec.com, Inventec Corp., Taiwan; Trista Pei-Chun Chen, chen.trista@inventec.com, Inventec Corp., Taiwan; Wei-Chao Chen, weichao.chen@skywatch24.com, Skywatch Innovation Inc., Inventec Corp., Taiwan.

0730-0301/2020/7-ART38
https://doi.org/10.1145/3386569.3392433

1 INTRODUCTION

The quality of character animation in cartoons, video games, and digital special effects have improved drastically in the past decades with new tools and techniques developed by researchers in the field. Amongst various types of characters, quadrupeds are especially challenging to animate due to their wide variations of style, cadence, and gait pattern. For real-time applications such as video games, the

Catch & Carry: Reusable Neural Controllers for Vision-Guided Whole-Body Tasks

JOSH MEREL, SARAN TUNYASUVUNAKOOL, ARUN AHUJA, YUVAL TASSA, LEONARD HASENCLEVER, VU PHAM, TOM EREZ, GREG WAYNE, and NICOLAS HEESS, DeepMind

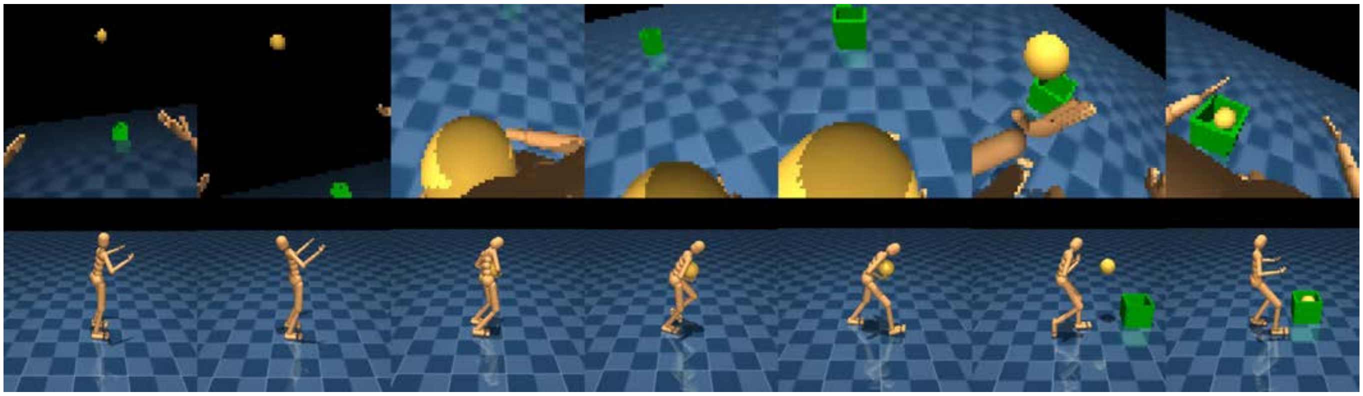


Fig. 1. A catch-carry-toss sequence (bottom) from first-person visual inputs (top). Note how the character's gaze and posture track the ball.

We address the longstanding challenge of producing flexible, realistic humanoid character controllers that can perform diverse whole-body tasks involving object interactions. This challenge is central to a variety of fields, from graphics and animation to robotics and motor neuroscience. Our physics-based environment uses realistic actuation and first-person perception – including touch sensors and egocentric vision – with a view to producing active-sensing behaviors (e.g. gaze direction), transferability to real robots, and comparisons to the biology. We develop an integrated neural-network based approach consisting of a motor primitive module, human demonstrations, and an instructed reinforcement learning regime with curricula and task variations. We demonstrate the utility of our approach for several tasks, including goal-conditioned box carrying and ball catching, and we characterize its behavioral robustness. The resulting controllers can be deployed in real-time on a standard PC.[1]

CCS Concepts: • **Computing methodologies → Artificial intelligence**; **Control methods**; **Physical simulation**; **Motion capture**.

Additional Key Words and Phrases: reinforcement learning, physics-based character, motor control, object interaction

[1]See overview video, **Video 1**. Refer to Supplementary Section F for all video captions.

Authors' address: Josh Merel, jsmerel@google.com; Saran Tunyasuvunakool; Arun Ahuja; Yuval Tassa; Leonard Hasenclever; Vu Pham; Tom Erez; Greg Wayne; Nicolas Heess, DeepMind.

0730-0301/2020/7-ART39
https://doi.org/10.1145/3386569.3392474

ACM Reference Format:
Josh Merel, Saran Tunyasuvunakool, Arun Ahuja, Yuval Tassa, Leonard Hasenclever, Vu Pham, Tom Erez, Greg Wayne, and Nicolas Heess. 2020. Catch & Carry: Reusable Neural Controllers for Vision-Guided Whole-Body Tasks. *ACM Trans. Graph.* 39, 4, Article 39 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392474

1 INTRODUCTION

Endowing embodied agents with the motor intelligence that is required for natural and flexible goal-directed interaction with their physical environment is a longstanding challenge in artificial intelligence [Pfeifer and Scheier 2001]. This is a problem of obvious practical relevance to a number of fields including robotics [Arkin et al. 1998] and motor neuroscience [Merel et al. 2019b]. But it is also a topic of longstanding interest in the graphics and animation setting [since e.g. Raibert and Hodgins 1991; Van de Panne and Fiume 1993] – the ability to control agents with physically simulated bodies and sensors that naturally behave in response to high-level instructions may reduce the effort in creating realistic animations of agent-environment interactions.

Conventional approaches such as manual editing and kinematic blending of motion capture data require specification of character movements in their entirety, including how movements conform to the surrounding scene and task context. This can be challenging, especially when fine-grained motion is required, e.g. for object interaction. However, when controlling a goal-directed agent with physically simulated body, many aspects of its movements will emerge from the physical interaction; while other features, such as gaze direction, emerge from the interaction between the agent's goals and the constraints imposed by its body and sensors.

Recent developments in Deep Reinforcement Learning (Deep RL) have enabled great strides in learning from scratch for many game

Character Controllers using Motion VAEs

HUNG YU LING, University of British Columbia, Canada
FABIO ZINNO, Electronic Arts Vancouver, Canada
GEORGE CHENG, Electronic Arts Vancouver, Canada
MICHIEL VAN DE PANNE, University of British Columbia, Canada

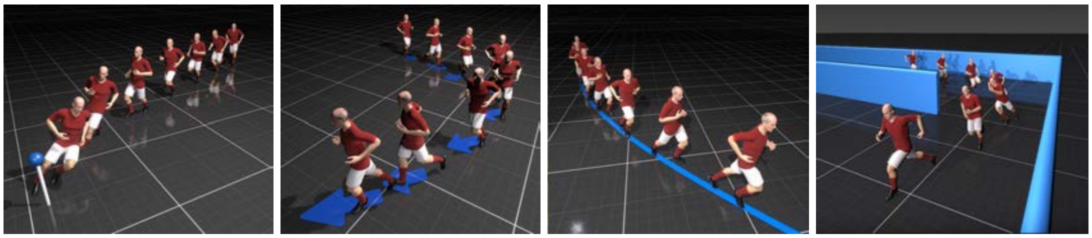


Fig. 1. Given example data, we learn an autoregressive conditional variational autoencoder that predicts the next pose one frame at a time. A variety of task-specific control policies can then be learned on top of this model.

A fundamental problem in computer animation is that of realizing purposeful and realistic human movement given a sufficiently-rich set of motion capture clips. We learn data-driven generative models of human movement using autoregressive conditional variational autoencoders, or Motion VAEs. The latent variables of the learned autoencoder define the action space for the movement and thereby govern its evolution over time. Planning or control algorithms can then use this action space to generate desired motions. In particular, we use deep reinforcement learning to learn controllers that achieve goal-directed movements. We demonstrate the effectiveness of the approach on multiple tasks. We further evaluate system-design choices and describe the current limitations of Motion VAEs.

CCS Concepts: • **Computing methodologies → Motion capture**; **Reinforcement learning**.

Additional Key Words and Phrases: motion synthesis, character control, human motion model, reinforcement learning

ACM Reference Format:
Hung Yu Ling, Fabio Zinno, George Cheng, and Michiel van de Panne. 2020. Character Controllers using Motion VAEs. *ACM Trans. Graph.* 39, 4, Article 40 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392422

1 INTRODUCTION

Given example motions, how can we generalize these to produce new purposeful motions? This problem is at the core of interactive character animation and control, with applications that include games, simulations, and virtual reality. The solutions should ideally produce high-quality motion, be compact, be efficient to compute (at runtime), and support a variety of goal-directed behaviors.

In this paper, we take a two-step approach to this problem. First, we learn a kinematic generative model of human motion given the example motion data. This is based on an *autoregressive conditional variational autoencoder*, which we refer to more simply as a *motion VAE (MVAE)*. Given the current character pose, the MVAE predicts the pose at the next time step. Importantly, the model can produce a *distribution* of next-state predictions because it is also conditioned on a set of stochastic latent variables – each sampling of these variables corresponds to a different future feasible next-state prediction. The model is autoregressive, meaning that the current predicted pose becomes the current character pose for the following prediction.

Given a trained MVAE, it can be controlled to generate desired motions in several ways. The simplest is to randomly sample from the next-state predictions at each time-step, which produces a random walk through the learned dynamics of the MVAE. More interestingly, we can treat the stochastic variables that govern the next-state predictions as the action space for a reinforcement learning problem, which is not possible for learning approaches based purely on supervised learning. Given a reward function that defines the goals of the character, a control policy can then be learned which uses this action space to guide the generative model in accordance with those goals.

We note that VAEs have been previously identified as a promising class of models for kinematic motion generation, along with RNNs. However, the stable generation of long motion sequences is commonly acknowledged as a significant challenge, as are issues of motion quality. MVAEs produce high-quality results as demonstrated with fully-skinned characters and with no additional foot-skate cleanup. Further, unlike previous approaches using a memory-enabled RNN architecture, we show that the MVAE can work well in a memoryless fashion, i.e., conditioned only on the previous pose

Authors' addresses: Hung Yu Ling, University of British Columbia, Vancouver, Canada, hyuling@cs.ubc.ca; Fabio Zinno, Electronic Arts Vancouver, Vancouver, Canada, fzinno@ea.com; George Cheng, Electronic Arts Vancouver, Vancouver, Canada, gecheng@ea.com; Michiel van de Panne, University of British Columbia, Vancouver, Canada, van@cs.ubc.ca.

0730-0301/2020/7-ART40 $15.00
https://doi.org/10.1145/3386569.3392422

Chemomechanical Simulation of Soap Film Flow on Spherical Bubbles

WEIZHEN HUANG, JULIAN ISERINGHAUSEN, and TOM KNEIPHOF, University of Bonn, Germany
ZIYIN QU and CHENFANFU JIANG, University of Pennsylvania, USA
MATTHIAS B. HULLIN, University of Bonn, Germany

Fig. 1. Spatially varying iridescence of a soap bubble evolving over time (left to right). The complex interplay of soap and water induces a complex flow on the film surface, resulting in an ever changing distribution of film thickness and hence a highly dynamic iridescent texture. This image was simulated using the method described in this paper, and path-traced in Mitsuba [Jakob 2010] using a custom shader under environment lighting.

Soap bubbles are widely appreciated for their fragile nature and their colorful appearance. The natural sciences and, in extension, computer graphics, have comprehensively studied the mechanical behavior of films and foams, as well as the optical properties of thin liquid layers. In this paper, we focus on the dynamics of material flow within the soap film, which results in fascinating, extremely detailed patterns. This flow is characterized by a complex coupling between surfactant concentration and Marangoni surface tension. We propose a novel chemomechanical simulation framework rooted in lubrication theory, which makes use of a custom semi-Lagrangian advection solver to enable the simulation of soap film dynamics on spherical bubbles both in free flow as well as under body forces such as gravity or external air flow. By comparing our simulated outcomes to videos of real-world soap bubbles recorded in a studio environment, we show that our framework, for the first time, closely recreates a wide range of dynamic effects that are also observed in experiment.

CCS Concepts: • **Computing methodologies → Physical simulation**; *Reflectance modeling*.

Additional Key Words and Phrases: soap films, fluid dynamics, physical simulation, coordinate singularity, thin-film inteference

ACM Reference Format:
Weizhen Huang, Julian Iseringhausen, Tom Kneiphof, Ziyin Qu, Chenfanfu Jiang, and Matthias B. Hullin. 2020. Chemomechanical Simulation of Soap Film Flow on Spherical Bubbles. *ACM Trans. Graph.* 39, 4, Article 41 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392094

Authors' e-mail addresses: whuang@cs.uni-bonn.de; iseringhausen@cs.uni-bonn.de; kneiphof@cs.uni-bonn.de; ziyinq@seas.upenn.edu; cffjiang@seas.upenn.edu; hullin@cs.uni-bonn.de.

0730-0301/2020/7-ART41 $15.00
https://doi.org/10.1145/3386569.3392094

1 MOTIVATION

The beauty of soap films and bubbles is of great appeal to people of all ages and cultures, and the scientific community is no exception. In the computer graphics community, it is now widely understood how films, bubbles and foams form, evolve and break. On the rendering side, it has become possible to recreate their characteristic iridescent appearance in physically based renderers. The main parameter governing this appearance, the thickness of the film, has so far only been driven using ad-hoc noise textures [Glassner 2000], or was assumed to be constant. The resulting renderings appear largely plausible but static, as they lack the rich dynamics known from real-world soap films.

With this paper, we aim to close this gap in order to achieve greater realism. We do so by contributing a chemomechanical framework targeted specifically at simulating the rich and detailed microscopic flow on spherical soap bubbles. Our framework employs a leading-order approximation for the soap film dynamics developed by [Chomaz 2001; Ida and Miksis 1998a]. A soap bubble is modeled as a two-dimensional flow on a static spherical surface with two associated scalar fields: the film thickness and the soap concentration. We are able to show that this state-of-the-art model, paired with a custom solver, is capable of expressing the intricate flows found on real-world soap bubbles (Figure 1) under the mutual influence of mechanical stress, film thickness and surfactant concentration as well as body and surface forces like gravity and air friction. Our simulation is performed on a staggered grid, using finite differences in space and time. An advection scheme based on BiMocq^2 [Qu et al. 2019] minimizes numerical dissipation in order to prevent fine details from washing out over time. The resulting thickness maps are presented in real time using a custom, very efficient polarization-aware spectral rendering scheme.

Besides the underlying physical model, our framework is enabled by the following key contributions:

Codimensional Surface Tension Flow using Moving-Least-Squares Particles

HUI WANG, Shanghai Jiao Tong University
YONGXU JIN, Stanford University
ANQI LUO, Dartmouth College
XUBO YANG*, Shanghai Jiao Tong University
BO ZHU*, Dartmouth College

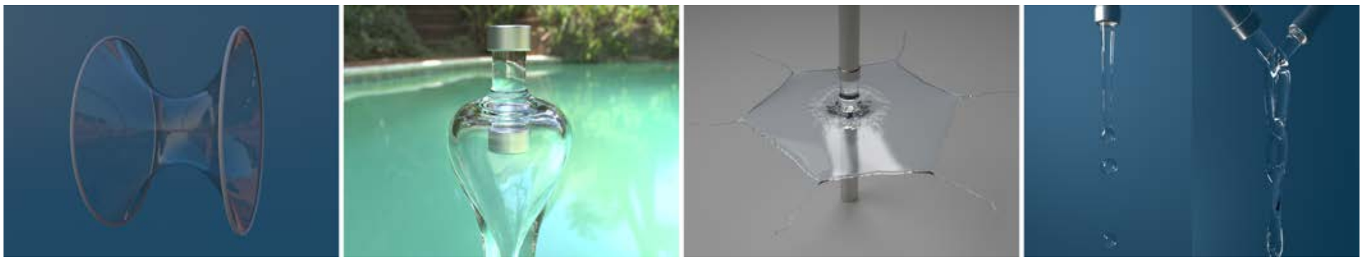


Fig. 1. Various codimensional fluid phenomena simulated using our codimensional MLS particle method. (Far Left) Film catenoid: A membrane attached to two rings contracts due to surface tension. (Middle Left) Water bell: Two vertical jets of water strike together, resulting in a bell-like water sheet. (Middle Right) Fluid polygon: Two vertical jets of water collide and form a polygonal thin sheet with pinched off filaments. (Far Right) Dripping water and fluid chain: Water flows from a pipe and forms droplets; Two jets collide and form a chain-like structure.

We propose a new Eulerian-Lagrangian approach to simulate the various surface tension phenomena characterized by volume, thin sheets, thin filaments, and points using Moving-Least-Squares (MLS) particles. At the center of our approach is a meshless Lagrangian description of the different types of codimensional geometries and their transitions using an MLS approximation. In particular, we differentiate the codimension-1 and codimension-2 geometries on Lagrangian MLS particles to precisely describe the evolution of thin sheets and filaments, and we discretize the codimension-0 operators on a background Cartesian grid for efficient volumetric processing. Physical forces including surface tension and pressure across different codimensions are coupled in a monolithic manner by solving one single linear system to evolve the surface-tension driven Navier-Stokes system in a complex non-manifold space. The codimensional transitions are handled explicitly by tracking a codimension number stored on each particle, which replaces the tedious meshing operators in a conventional mesh-based approach. Using the proposed framework, we simulate a broad array of visually appealing surface tension phenomena, including the fluid chain, bell, polygon, catenoid, and dripping, to demonstrate the efficacy of our approach in capturing the complex fluid characteristics with mixed codimensions, in a robust, versatile, and connectivity-free manner.

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: codimensional fluids, surface tension, PIC/FLIP, moving-least-squares

ACM Reference Format:
Hui Wang, Yongxu Jin, Anqi Luo, Xubo Yang, and Bo Zhu. 2020. Codimensional Surface Tension Flow using Moving-Least-Squares Particles. *ACM Trans. Graph.* 39, 4, Article 42 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392487

1 INTRODUCTION

Particle methods are inherently adaptive. With particles, a broad array of interesting interfacial fluid phenomena can be modeled effectively by deploying a sufficient number of point samples around the region of interest to produce high-quality simulations. Typical examples can be seen in particle level-set methods [Enright et al. 2002], adaptive SPH [Adams et al. 2007; Winchenbach et al. 2017], adaptive MPM [Gao et al. 2017], narrow-band PIC/FLIP [Ferstl et al. 2016], etc. All these methods build the discrete differential operators in a volumetric way, either on particles directly or on an auxiliary background grid, by treating each particle as an isotropic point sample in space. The success of these computational paradigms relies on a well-defined sample distribution over the computational domain, which needs to be sufficiently dense and regular, to carry out local approximations by smoothed particle quantities using a mollified kernel function (see [Cottet et al. 2000]). However, on the minus side, particle methods are not inherently codimensional. It is extremely challenging to simulate complex fluid systems composed of volumes and codimensional features. These feature types include

*Corresponding authors

Authors' addresses: Hui Wang, wanghehv@sjtu.edu.cn, Shanghai Jiao Tong University; Yongxu Jin, yxjin@stanford.edu, Stanford University; Anqi Luo, anqi.luo.gr@dartmouth.edu, Dartmouth College; Xubo Yang, yangxubo@sjtu.edu.cn, Shanghai Jiao Tong University; Bo Zhu, bo.zhu@dartmouth.edu, Dartmouth College.

0730-0301/2020/7-ART42 $15.00
https://doi.org/10.1145/3386569.3392487

Constraint Bubbles and Affine Regions: Reduced Fluid Models for Efficient Immersed Bubbles and Flexible Spatial Coarsening

RYAN GOLDADE, University of Waterloo, Canada
MRIDUL AANJANEYA, Rutgers University, USA
CHRISTOPHER BATTY, University of Waterloo, Canada

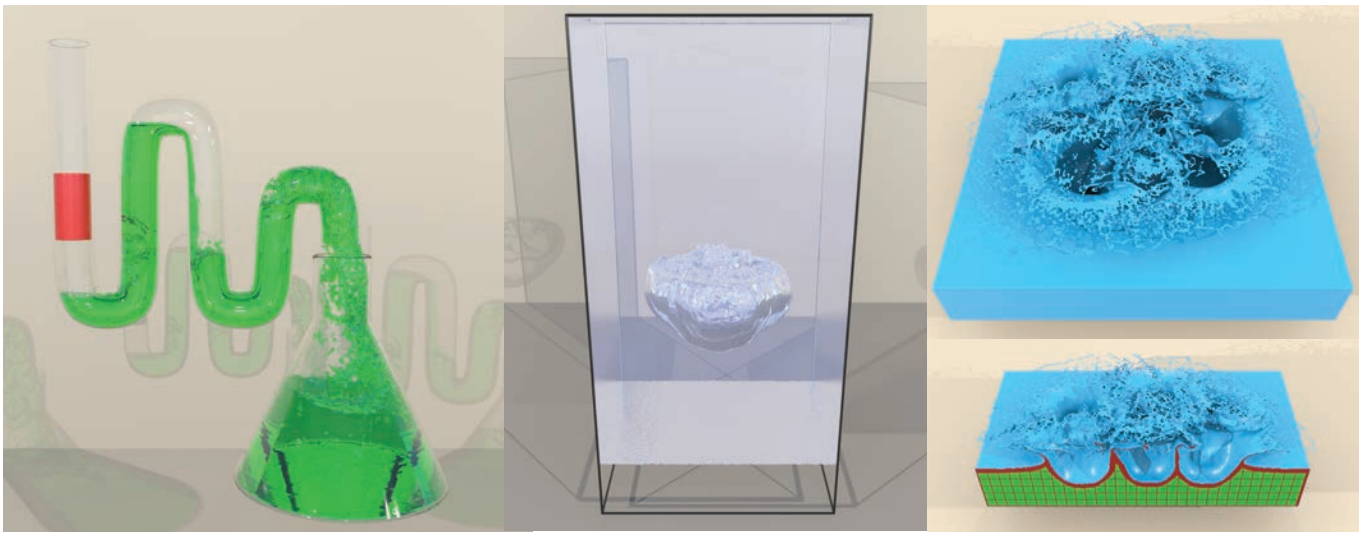


Fig. 1. (Left) Our *constraint bubble* model allows distinct liquid (green) and solid (red) bodies to physically interact across completely unsimulated air gaps. (Middle) Immersed bubbles denser than water correctly sink, despite the interior degrees of freedom being radically reduced with our *affine region* model. (Right) Our affine region model also enables a convenient and flexible approach to liquid adaptivity with irregularly shaped coarse tiles (green).

We propose to enhance the capability of standard free-surface flow simulators with efficient support for immersed bubbles through two new models: constraint-based bubbles and affine fluid regions. Unlike its predecessors, our constraint-based model entirely dispenses with the need for advection or projection inside zero-density bubbles, with extremely modest additional computational overhead that is proportional to the surface area of all bubbles. This surface-only approach is easy to implement, realistically captures many familiar bubble behaviors, and even allows two or more distinct liquid bodies to correctly interact across completely unsimulated air. We augment this model with a per-bubble volume-tracking and correction framework to minimize the cumulative effects of gradual volume drift. To support bubbles with non-zero densities, we propose a novel reduced model for an irregular fluid region with a single pointwise incompressible affine vector field. This model requires only 11 interior velocity degrees of freedom per affine fluid region in 3D, and correctly reproduces buoyant, stationary, and sinking behaviors of a secondary fluid phase with non-zero density immersed in water. Since the pressure projection step in both the above schemes is a slightly modified Poisson-style system, we propose novel Multigrid-based preconditioners for Conjugate Gradients for fast numerical solutions of our new discretizations. Furthermore, we observe that by enforcing an incompressible affine vector field over a coalesced set of grid cells, our reduced model is effectively an irregular coarse super-cell. This offers a convenient and flexible adaptive coarsening strategy that integrates readily with the standard staggered grid approach for fluid simulation, yet supports coarsened regions that are arbitrary voxelized shapes, and provides an analytically divergence-free interior. We demonstrate its effectiveness with a new adaptive liquid simulator whose interior regions are coarsened into a mix of tiles with regular and irregular shapes.

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: liquid, adaptivity, two-phase, reduced model

ACM Reference Format:
Ryan Goldade, Mridul Aanjaneya, and Christopher Batty. 2020. Constraint Bubbles and Affine Regions: Reduced Fluid Models for Efficient Immersed Bubbles and Flexible Spatial Coarsening. *ACM Trans. Graph.* 39, 4, Article 43 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392455

Authors' addresses: Ryan Goldade, University of Waterloo, 200 University Ave W., Waterloo, ON, N2L 3G1, Canada, rgoldade@uwaterloo.ca; Mridul Aanjaneya, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ, 08854, USA, mridul.aanjaneya@rutgers.edu; Christopher Batty, University of Waterloo, 200 University Ave W., Waterloo, ON, N2L 3G1, Canada, christopher.batty@uwaterloo.ca.

0730-0301/2020/7-ART43 $15.00
https://doi.org/10.1145/3386569.3392455

Directional Sources and Listeners in Interactive Sound Propagation using Reciprocal Wave Field Coding

CHAKRAVARTY R. ALLA CHAITANYA*, Microsoft Research and McGill University
NIKUNJ RAGHUVANSHI*, Microsoft Research
KEITH W. GODIN, Microsoft Mixed Reality
ZECHEN ZHANG, Microsoft Research and Cornell University
DEREK NOWROUZEZAHRAI, McGill University
JOHN M. SNYDER, Microsoft Research

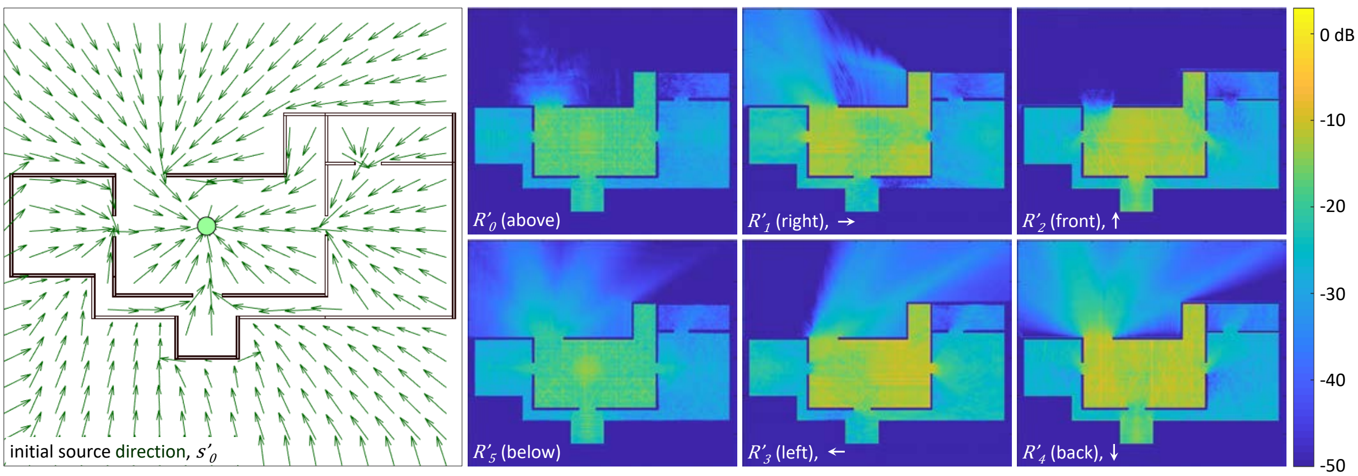


Fig. 1. Summarized parameter fields for HouseScene. For each listener position (green circle), we precompute and store directional parameter fields that vary over 3D source position. We visualize slices at listener height through these fields. Initial source direction (left) encodes the direction of radiation at each source location for the shortest (earliest arriving or "direct") path arriving at the listener. Similarly an arrival direction at listener is also encoded for each source location, not shown. Indirect energy transfer between source and listener is encoded in a 6×6 "reflections transfer matrix" (RTM) that aggregates about six signed axes in world space around both source and listener. RTM images (right) are summarized here by summing over listener directions via (16), representing source transfer anisotropy for an omnidirectional microphone. Our encoding and runtime use the full matrix, shown in Figure 11. Overall our fully reciprocal encoding captures directionality at both source and listener.

Common acoustic sources, like voices or musical instruments, exhibit strong frequency and directional dependence. When transported through complex environments, their anisotropic radiated field undergoes scattering, diffraction, and occlusion before reaching a directionally-sensitive listener. We present the first wave-based interactive auralization system that encodes and renders a complete reciprocal description of acoustic wave fields in general scenes. Our method renders directional effects at freely moving and rotating sources and listeners and supports any tabulated *source directivity function* and head-related transfer function. We represent a static scene's global acoustic transfer as an 11-dimensional *bidirectional impulse response* (BIR) field, which we extract from a set of wave simulations. We parametrically encode the BIR as a pair of radiating and arriving directions for the perceptually-salient initial (*direct*) response, and a compact 6×6 *reflections transfer matrix* capturing indirect energy transfer with scene-dependent anisotropy. We render our encoded data with an efficient and scalable algorithm – integrated in the Unreal Engine™ – whose CPU performance is agnostic to scene complexity and angular source/listener resolutions. We demonstrate convincing effects that depend on detailed scene geometry, for a variety of environments and source types.

CCS Concepts: • **Applied computing → Sound and music computing**; • **Computing methodologies → Virtual reality**.

Additional Key Words and Phrases: bidirectional impulse response, equalization, head-related transfer function (HRTF), spatial audio, sound propagation, source directivity, virtual acoustics, wave simulation

*Equal contribution.

Authors' addresses: Chakravarty R. Alla Chaitanya, chakravarty.alla@gmail.com, Microsoft Research and McGill University; Nikunj Raghuvanshi, nikunjr@microsoft.com, Microsoft Research; Keith W. Godin, kegodin@microsoft.com, Microsoft Mixed Reality; Zechen Zhang, zz335@cornell.edu, Microsoft Research and Cornell University; Derek Nowrouzezahrai, derek@cim.mcgill.ca, McGill University; John M. Snyder, johnsny@microsoft.com, Microsoft Research.

0730-0301/2020/7-ART44 $15.00
https://doi.org/10.1145/3386569.3392459

Example-driven Virtual Cinematography by Learning Camera Behaviors

HONGDA JIANG*, CFCS, Peking University & AICFVE, Beijing Film Academy
BIN WANG*, AICFVE, Beijing Film Academy
XI WANG, University Rennes, Inria, CNRS, IRISA & AICFVE, Beijing Film Academy
MARC CHRISTIE, University Rennes, Inria, CNRS, IRISA & AICFVE, Beijing Film Academy
BAOQUAN CHEN†, CFCS, Peking University & AICFVE, Beijing Film Academy

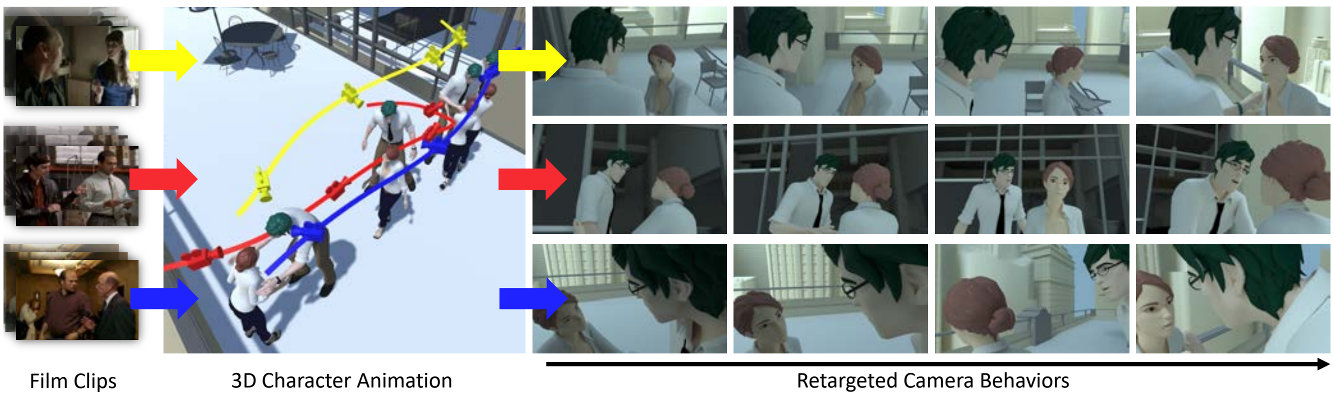


Fig. 1. We propose the design of a camera motion controller which has the ability to automatically extract camera behaviors from different film clips (on the left) and re-apply these behaviors to a 3D animation (center). In this example, three distinct camera trajectories are automatically generated (red, blue and yellow curves) from three different reference clips. Results display viewpoints at 4 specific instants along each camera trajectory demonstrating the capacity of our system to encode and reproduce camera behaviors from distinct input examples.

Designing a camera motion controller that has the capacity to move a virtual camera automatically in relation with contents of a 3D animation, in a cinematographic and principled way, is a complex and challenging task. Many cinematographic rules exist, yet practice shows there are significant stylistic variations in how these can be applied.

In this paper, we propose an example-driven camera controller which can extract camera behaviors from an example film clip and re-apply the extracted behaviors to a 3D animation, through learning from a collection of camera motions. Our first technical contribution is the design of a low-dimensional cinematic feature space that captures the essence of a film's cinematic characteristics (camera angle and distance, screen composition and character configurations) and which is coupled with a neural network to automatically extract these cinematic characteristics from real film clips. Our second technical contribution is the design of a cascaded deep-learning architecture trained to (i) recognize a variety of camera motion behaviors from the extracted cinematic features, and (ii) predict the future motion of a virtual camera given a character 3D animation. We propose to rely on a Mixture of Experts (MoE) gating+prediction mechanism to ensure that distinct camera behaviors can be learned while ensuring generalization.

We demonstrate the features of our approach through experiments that highlight (i) the quality of our cinematic feature extractor (ii) the capacity to learn a range of behaviors through the gating mechanism, and (iii) the ability to generate a variety of camera motions by applying different behaviors extracted from film clips. Such an example-driven approach offers a high level of controllability which opens new possibilities toward a deeper understanding of cinematographic style and enhanced possibilities in exploiting real film data in virtual environments.

CCS Concepts: • **Computing methodologies** → **Procedural animation**.

Additional Key Words and Phrases: Virtual Cinematography, Camera Behaviors, Machine Learning

ACM Reference Format:
Hongda Jiang, Bin Wang, Xi Wang, Marc Christie, and Baoquan Chen. 2020. Example-driven Virtual Cinematography by Learning Camera Behaviors. *ACM Trans. Graph.* 39, 4, Article 45 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392427

*denotes equal contribution
†corresponding author

Authors' addresses: Hongda Jiang, jianghd@pku.edu.cn, CFCS, Peking University & AICFVE, Beijing Film Academy; Bin Wang, binwangbuaa@gmail.com, AICFVE, Beijing Film Academy; Xi Wang, xi.wang@inria.fr, University Rennes, Inria, CNRS, IRISA & AICFVE, Beijing Film Academy; Marc Christie, marc.christie@irisa.fr, University Rennes, Inria, CNRS, IRISA & AICFVE, Beijing Film Academy; Baoquan Chen, baoquan@pku.edu.cn, CFCS, Peking University & AICFVE, Beijing Film Academy.

0730-0301/2020/7-ART45 $15.00
https://doi.org/10.1145/3386569.3392427

Fast and Flexible Multilegged Locomotion Using Learned Centroidal Dynamics

TAESOO KWON, Hanyang University
YOONSANG LEE, Hanyang University
MICHIEL VAN DE PANNE, University of British Columbia

(a) Humanoid character (biped) (b) Luxo character (monopod) (c) Cassie robot (biped) (d) ANYmal robot (quadruped)

Fig. 1. Our system generates rich variations of walking, running, and jumping for a variety of characters at interactive rates.

We present a flexible and efficient approach for generating multilegged locomotion. Our model-predictive control (MPC) system efficiently generates terrain-adaptive motions, as computed using a three-level planning approach. This leverages two commonly-used simplified dynamics models, an inverted pendulum on a cart model (IPC) and a centroidal dynamics model (CDM). Taken together, these ensure efficient computation and physical fidelity of the resulting motion. The final full-body motion is generated using a novel momentum-mapped inverse kinematics solver and is responsive to external pushes by using CDM forward dynamics. For additional efficiency and robustness, we then learn a predictive model that then replaces two of the intermediate steps. We demonstrate the rich capabilities of the method by applying it to monopeds, bipeds, and quadrupeds, and showing that it can generate a very broad range of motions at interactive rates, including banked variable-terrain walking and running, hurdles, jumps, leaps, stepping stones, monkey bars, implicit quadruped gait transitions, moon gravity, push-responses, and more.

CCS Concepts: • **Computing methodologies → Physical simulation**; *Motion path planning*.

Additional Key Words and Phrases: character animation, motion planning, rigid-body simulation

Authors' addresses: Taesoo Kwon, Department of Computer Science and Engineering, Hanyang University, taesoo@hanyang.ac.kr; Yoonsang Lee, Department of Computer Science and Engineering, Hanyang University, yslee111@gmail.com; Michiel van de Panne, Department of Computer Science, University of British Columbia, van@cs.ubc.ca.

0730-0301/2020/7-ART46 $15.00
https://doi.org/10.1145/3386569.3392432

ACM Reference Format:
Taesoo Kwon, Yoonsang Lee, and Michiel van de Panne. 2020. Fast and Flexible Multilegged Locomotion Using Learned Centroidal Dynamics. *ACM Trans. Graph.* 39, 4, Article 46 (July 2020), 24 pages. https://doi.org/10.1145/3386569.3392432

1 INTRODUCTION

Physically simulated characters can respond to user input while also adapting to environmental and external forces. However, it is not easy to create a controller which robustly generates natural locomotion because legged locomotion is inherently underactuated. It means the character cannot be directly controlled, but instead must be controlled through contact forces applied to the feet or hands. This indirect control is challenging because the contact forces can only push, the contact point should be within the reach of the character, and all contacts must occur at appropriate time intervals. Controlling locomotion skills for wide variations of terrains, turns, speeds, and accelerations is even more challenging.

Trajectory optimization is one approach to address this problem, and formulates it using an objective function and a set of constraints to define the desired motion. Model predictive control (MPC) is an online trajectory optimization approach, which is powerful for realizing motions that require significant anticipation and planning. In this approach, an optimized motion is computed for the near-term future at each planning time-step. However, trajectory optimization remains a challenging computational problem because this formulation inherently requires a high-dimensional search space that is related to the degrees of freedoms of a character, and a number of variables that increases linearly with the time horizon, and the complexity of the environment.

To mitigate the computational cost, trajectory optimization is often combined with simplified dynamic models such as a spring-loaded inverted pendulum model (SLIP) [Mordatch et al. 2010], an inverted pendulum on a cart model (IPC) [Kwon and Hodgins 2010],

Fast and Scalable Turbulent Flow Simulation with Two-Way Coupling

WEI LI, ShanghaiTech University/SIMIT/UCAS
YIXIN CHEN, ShanghaiTech University/DGene
MATHIEU DESBRUN, ShanghaiTech/Caltech
CHANGXI ZHENG, Columbia University
XIAOPEI LIU, ShanghaiTech University

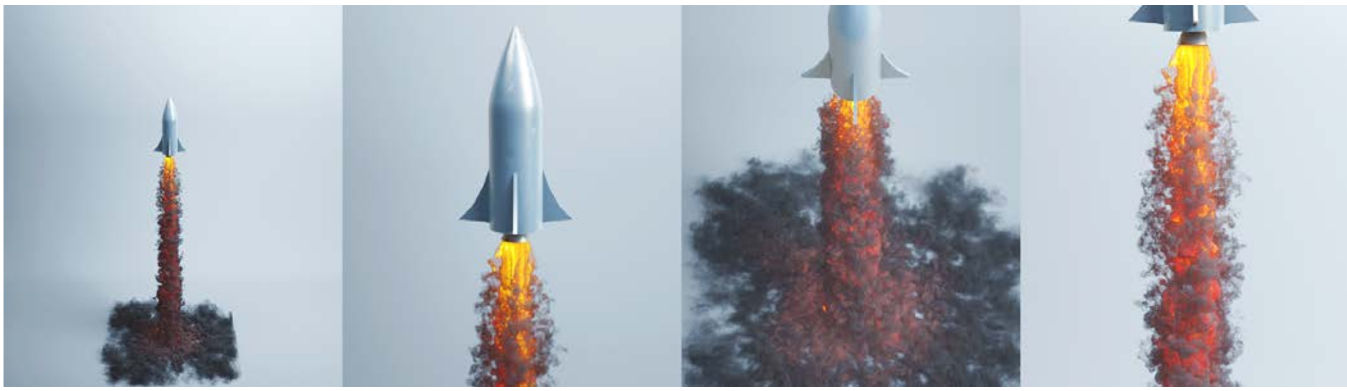


Fig. 1. **Rocket launching.** Our novel solver can accurately simulate fluid flows of either laminar or turbulent nature in an efficient and scalable manner. Here, a large-scale simulation of a rocket launch (computed with 4 GPUs on a grid of size 500×1200×500, 3.7 hours of computation for 1s of simulation); two-way coupling automatically propels the rocket in the air while generating a turbulent wake affecting the entire simulation domain (thermal effects are ignored).

Despite their cinematic appeal, turbulent flows involving fluid-solid coupling remain a computational challenge in animation. At the root of this current limitation is the numerical dispersion from which most accurate Navier-Stokes solvers suffer: proper coupling between fluid and solid often generates artificial dispersion in the form of local, parasitic trains of velocity oscillations, eventually leading to numerical instability. While successive improvements over the years have led to conservative and detail-preserving fluid integrators, the dispersive nature of these solvers is rarely discussed despite its dramatic impact on fluid-structure interaction. In this paper, we introduce a novel low-dissipation and low-dispersion fluid solver that can simulate two-way coupling in an efficient and scalable manner, even for turbulent flows. In sharp contrast with most current CG approaches, we construct our solver from a kinetic formulation of the flow derived from statistical mechanics. Unlike existing lattice Boltzmann solvers, our approach leverages *high-order moment relaxations* as a key to controlling both dissipation and dispersion of the resulting scheme. Moreover, we combine our new fluid solver with the immersed boundary method to easily handle fluid-solid coupling through time adaptive simulations. Our kinetic solver is highly parallelizable by nature, making it ideally suited for implementation on single- or multi-GPU computing platforms. Extensive comparisons with existing solvers on synthetic tests and real-life experiments are used to highlight the multiple advantages of our work over traditional and more recent approaches, in terms of accuracy, scalability, and efficiency.

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: Fluid Simulation, Kinetic Theory, Lattice Boltzmann Method, Fluid-Solid Coupling, Immersed Boundary Method.

ACM Reference Format:
Wei Li, Yixin Chen, Mathieu Desbrun, Changxi Zheng, and Xiaopei Liu. 2020. Fast and Scalable Turbulent Flow Simulation with Two-Way Coupling. *ACM Trans. Graph.* 39, 4, Article 47 (July 2020), 20 pages. https://doi.org/10.1145/3386569.3392400

Authors' addresses: Wei Li, Yixin Chen, Xiaopei Liu, School of Information Science and Technology (Shanghai Engineering Research Center of Intelligent Vision and Imaging) of ShanghaiTech University, Shanghai, China; Wei Li is also affiliated with the Shanghai Institute of Microsystem and Information Technology (SIMIT) and the University of the Chinese Academy of Sciences (UCAS); Mathieu Desbrun, California Institute of Technology, Pasadena (CA), USA, on sabbatical at SIST in ShanghaiTech University, Shanghai, China; Changxi Zheng, Columbia University, New York (NY), USA.

0730-0301/2020/7-ART47 $15.00
https://doi.org/10.1145/3386569.3392400

1 INTRODUCTION

The intricate interplay between fluid and solids is a common occurrence all around us: from the fluttering of a dry leaf in the wind to the tumbling of a paper cup, the interaction, or *coupling*, between the motions of a fluid and of a solid is all the more complex than the flow is turbulent. Automatically capturing such complex phenomena being particularly desirable for movie production, numerous approaches to fluid-solid coupling have been proposed in the CG literature, most of them combining an incompressible Navier-Stokes solver with a rigid- or deformable-body simulator. Whether they rely on uniform grids [Carlson et al. 2004; Robinson-Mosher et al. 2008; Qiu et al. 2015], body-fitted meshes [Feldman et al. 2005; Klingner et al. 2006; Elcott et al. 2007; Clausen et al. 2013], cut-cells [Batty et al. 2007; Ng et al. 2009; Azevedo et al. 2016; Liu et al. 2015], or are based on particles [Akinci et al. 2012; Ihmsen et al. 2013; Band et al. 2018],

Homogenized Yarn-Level Cloth

GEORG SPERL, IST Austria
RAHUL NARAIN, Indian Institute of Technology Delhi
CHRIS WOJTAN, IST Austria

Fig. 1. Left: A comparison between direct yarn-level simulation (YLC) and simulation with our homogenized model (HYLC); our homogenized model accurately captures the non-trivial elastic stretching and bending response of the fabric. Middle and right: Results simulated with homogenized continuum models of woven and knitted patterns; our method allows us to efficiently compute large-scale simulations where direct yarn-level simulation would be prohibitively slow.

We present a method for animating yarn-level cloth effects using a thin-shell solver. We accomplish this through numerical homogenization: we first use a large number of yarn-level simulations to build a model of the potential energy density of the cloth, and then use this energy density function to compute forces in a thin shell simulator. We model several yarn-based materials, including both woven and knitted fabrics. Our model faithfully reproduces expected effects like the stiffness of woven fabrics, and the highly deformable nature and anisotropy of knitted fabrics. Our approach does not require any real-world experiments nor measurements; because the method is based entirely on simulations, it can generate entirely new material models quickly, without the need for testing apparatuses or human intervention. We provide data-driven models of several woven and knitted fabrics, which can be used for efficient simulation with an off-the-shelf cloth solver.

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: knitted, woven, cloth simulation, yarn-level cloth, homogenization, data fitting

ACM Reference Format:
Georg Sperl, Rahul Narain, and Chris Wojtan. 2020. Homogenized Yarn-Level Cloth. *ACM Trans. Graph.* 39, 4, Article 48 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392412

Authors' addresses: Georg Sperl, IST Austria, georg.sperl@ist.ac.at; Rahul Narain, Indian Institute of Technology Delhi, narain@cse.iitd.ac.in; Chris Wojtan, IST Austria, wojtan@ist.ac.at.

0730-0301/2020/7-ART48 $15.00
https://doi.org/10.1145/3386569.3392412

1 INTRODUCTION

The simulation and analysis of yarn-level cloth has recently generated a great deal of research in the computer graphics [Cirio et al. 2014; Kaldor et al. 2008; Narayanan et al. 2018; Yuksel et al. 2012], materials science [Choi and Lo 2003; Fillep et al. 2017], and physics communities [Poincloux et al. 2018]. Woven and knitted materials can exhibit a wide array of behaviors (highly variable stretchiness, anisotropy, area-preservation effects, etc.). They can be produced from simple threads or wires, leading to their ubiquity in everyday life. Furthermore, these materials are fascinating from a theoretical point of view, because their varying material properties arise almost entirely from the *geometric structure* of the threads — subtly different stitch patterns can lead to dramatically different material behaviors. For example, Figure 2 illustrates how different knit patterns influence area-preservation and curling of the fabric.

Simulating woven and knitted materials as a collection of interacting threads can accurately reproduce highly complex behaviors, but this direct strategy tends to be computationally expensive. On the other hand, finite element-based cloth simulations are relatively computationally efficient, because they replace the simulation of individual fibers with an approximate material model based on continuum mechanics. Unfortunately, choosing a suitable material model is a nontrivial task, and Little is known about the continuum behavior of many woven and knitted fabrics in particular, so many yarn-level effects cannot be captured by existing material models in computer graphics.

In this work, we aim to determine material properties directly from yarn-level geometry using numerical homogenization. We precompute the effective material response from periodic yarn-level simulations, learn an approximate material model from the resulting

Incremental Potential Contact: Intersection- and Inversion-free, Large-Deformation Dynamics

MINCHEN LI, University of Pennsylvania & Adobe Research
ZACHARY FERGUSON and TESEO SCHNEIDER, New York University
TIMOTHY LANGLOIS, Adobe Research
DENIS ZORIN and DANIELE PANOZZO, New York University
CHENFANFU JIANG, University of Pennsylvania
DANNY M. KAUFMAN, Adobe Research

Fig. 1. **Squeeze out:** Incremental Potential Contact (IPC) enables high-rate time stepping, here with $h = 0.01$s, of extreme nonlinear elastodynamics with contact that is intersection- and inversion-free at all time steps, irrespective of the degree of compression and contact. Here a plate compresses and then forces a collection of complex soft elastic FE models (181K tetrahedra in total, with a neo-Hookean material) through a thin, codimensional obstacle tube. The models are then compressed entirely together forming a tight mush to fit through the gap and then once through they cleanly separate into a stable pile.

Contacts weave through every aspect of our physical world, from daily household chores to acts of nature. Modeling and predictive computation of these phenomena for solid mechanics is important to every discipline concerned with the motion of mechanical systems, including engineering and animation. Nevertheless, efficiently time-stepping accurate and consistent simulations of real-world contacting elastica remains an outstanding computational challenge. To model the complex interaction of deforming solids in contact we propose Incremental Potential Contact (IPC) – a new model and algorithm for variationally solving implicitly time-stepped nonlinear elastodynamics. IPC maintains an intersection- and inversion-free trajectory regardless of material parameters, time step sizes, impact velocities, severity of deformation, or boundary conditions enforced.

Constructed with a custom nonlinear solver, IPC enables efficient resolution of time-stepping problems with separate, user-exposed accuracy tolerances that allow independent specification of the physical accuracy of the dynamics and the geometric accuracy of surface-to-surface conformation. This enables users to decouple, as needed per application, desired accuracies for a simulation's dynamics and geometry.

The resulting time stepper solves contact problems that are intersection-free (and thus robust), inversion-free, efficient (at speeds comparable to or faster than available methods that lack both convergence and feasibility), and accurate (solved to user-specified accuracies). To our knowledge this is the first implicit time-stepping method, across both the engineering and graphics literature that can consistently enforce these guarantees as we vary simulation parameters.

In an extensive comparison of available simulation methods, research libraries and commercial codes we confirm that available engineering and computer graphics methods, while each succeeding admirably in custom-tuned regimes, often fail with instabilities, egregious constraint violations and/or inaccurate and implausible solutions, as we vary input materials, contact numbers and time step. We also exercise IPC across a wide range

Authors' addresses: Minchen Li, University of Pennsylvania & Adobe Research, minchernl@gmail.com; Zachary Ferguson; Teseo Schneider, New York University; Timothy Langlois, Adobe Research; Denis Zorin; Daniele Panozzo, New York University; Chenfanfu Jiang, University of Pennsylvania; Danny M. Kaufman, Adobe Research, dannykaufman@gmail.com.

0730-0301/2020/7-ART49 $15.00
https://doi.org/10.1145/3386569.3392425

Informative Scene Decomposition for Crowd Analysis, Comparison and Simulation Guidance

FEIXIANG HE, University of Leeds, United Kingdom
YUANHANG XIANG, Xi'an Jiaotong University, China
XI ZHAO*, Xi'an Jiaotong University, China
HE WANG†, University of Leeds, United Kingdom

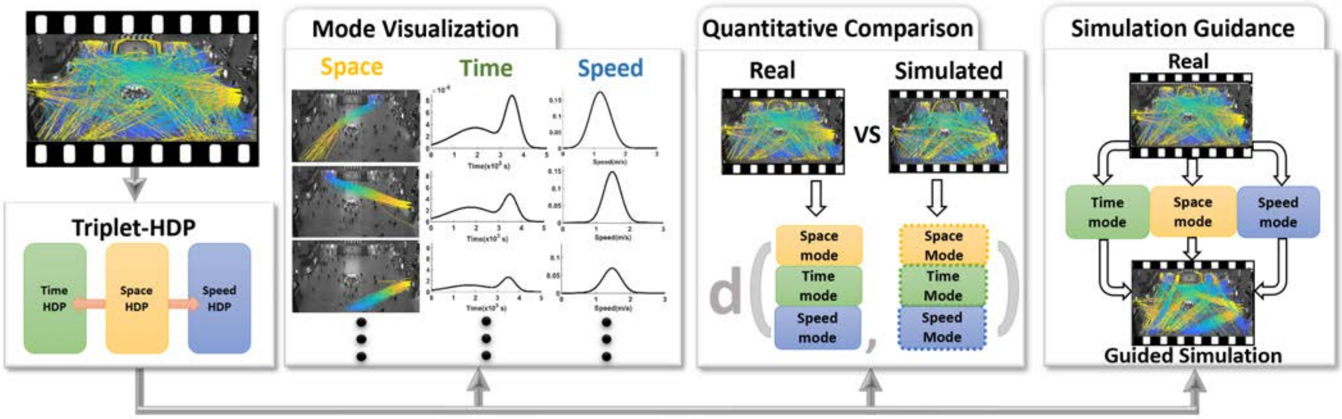


Fig. 1. Overview of our framework.

Crowd simulation is a central topic in several fields including graphics. To achieve high-fidelity simulations, data has been increasingly relied upon for analysis and simulation guidance. However, the information in real-world data is often noisy, mixed and unstructured, making it difficult for effective analysis, therefore has not been fully utilized. With the fast-growing volume of crowd data, such a bottleneck needs to be addressed. In this paper, we propose a new framework which comprehensively tackles this problem. It centers at an unsupervised method for analysis. The method takes as input raw and noisy data with highly mixed multi-dimensional (space, time and dynamics) information, and automatically structure it by learning the correlations among these dimensions. The dimensions together with their correlations fully describe the scene semantics which consists of recurring activity patterns in a scene, manifested as space flows with temporal and dynamics profiles. The effectiveness and robustness of the analysis have been tested on datasets with great variations in volume, duration, environment and crowd dynamics. Based on the analysis, new methods for data visualization, simulation evaluation and simulation guidance are also proposed. Together, our framework establishes a highly automated pipeline from raw data to crowd analysis, comparison and simulation guidance. Extensive experiments and evaluations have been conducted to show the flexibility, versatility and intuitiveness of our framework.

CCS Concepts: • **Computing methodologies → Animation**; **Topic modeling**; **Learning in probabilistic graphical models**; **Scene understanding**; **Activity recognition and understanding**; **Multi-agent planning**; • **Mathematics of computing → Probabilistic inference problems**; **Nonparametric statistics**.

Additional Key Words and Phrases: Crowd Simulation, Simulation Evaluation, Bayesian Inference

ACM Reference Format:
Feixiang He, Yuanhang Xiang, Xi Zhao, and He Wang. 2020. Informative Scene Decomposition for Crowd Analysis, Comparison and Simulation Guidance. *ACM Trans. Graph.* 39, 4, Article 50 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392407

*Corresponding author
†Corresponding author

Authors' addresses: Feixiang He, University of Leeds, School of Computing, United Kingdom, fxhe1992@gmail.com; Yuanhang Xiang, Xi'an Jiaotong University, School of Computer Science and Technology, China, xiangyuanhang@icloud.com; Xi Zhao, Xi'an Jiaotong University, School of Computer Science and Technology, China, zhaoxi.jade@gmail.com; He Wang, University of Leeds, School of Computing, United Kingdom, realcrane@gmail.com.

0730-0301/2020/7-ART50 $15.00
https://doi.org/10.1145/3386569.3392407

1 INTRODUCTION

Crowd simulation has been intensively used in computer animation, as well as other fields such as architectural design and crowd management. The fidelity or realism of simulation has been a long-standing problem. The main complexity arises from its multifaceted nature. It could mean high-level global behaviors [Narain et al. 2009], mid-level flow information [Wang et al. 2016] or low-level individual motions [Guy et al. 2012]. It could also mean perceived realism [Ennis et al. 2011] or numerical accuracy [Wang et al. 2017]. In

IQ-MPM: An Interface Quadrature Material Point Method for Non-sticky Strongly Two-Way Coupled Nonlinear Solids and Fluids

YU FANG*, University of Pennsylvania
ZIYIN QU*, University of Pennsylvania
MINCHEN LI, University of Pennsylvania
XINXIN ZHANG, Tencent
YIXIN ZHU, University of California, Los Angeles
MRIDUL AANJANEYA, Rutgers University
CHENFANFU JIANG, University of Pennsylvania

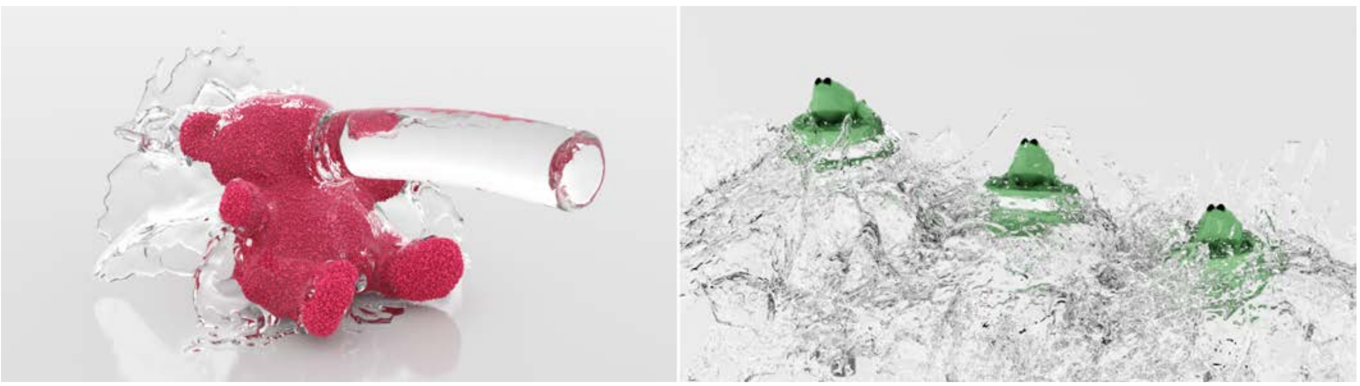


Fig. 1. (Left) A liquid jet hits a hyperelastic bear, pushing it down. (Right) Three hyperelastic frogs with different densities are lifted by liquid fountains, creating vibrant splashes. Our novel formulation properly enforces the free-slip boundary condition, while allowing for strong two-way coupled dynamics.

We propose a novel scheme for simulating two-way coupled interactions between nonlinear elastic solids and incompressible fluids. The key ingredient of this approach is a *ghost matrix* operator-splitting scheme for strongly coupled nonlinear elastica and incompressible fluids through the weak form of their governing equations. This leads to a stable and efficient method handling *large time steps* under the CFL limit while using *a single monolithic solve* for the coupled pressure fields, even in the case with highly nonlinear elastic solids. The use of the Material Point Method (MPM) is essential in the designing of the scheme, it not only preserves discretization consistency with the hybrid Lagrangian-Eulerian fluid solver, but also works naturally with our novel *interface quadrature* (IQ) discretization for free-slip boundary conditions. While traditional MPM suffers from sticky numerical artifacts, our framework naturally supports discontinuous tangential velocities at the solid-fluid interface. Our IQ discretization results in an easy-to-implement, fully particle-based treatment of the interfacial boundary, avoiding the additional complexities associated with intermediate level set or explicit mesh representations. The efficacy of the proposed scheme is verified by various challenging simulations with fluid-elastica interactions.

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: Fluids, numerical methods, MPM, Fluid-structure interaction

ACM Reference Format:
Yu Fang, Ziyin Qu, Minchen Li, Xinxin Zhang, Yixin Zhu, Mridul Aanjaneya, and Chenfanfu Jiang. 2020. IQ-MPM: An Interface Quadrature Material Point Method for Non-sticky Strongly Two-Way Coupled Nonlinear Solids and Fluids. *ACM Trans. Graph.* 39, 4, Article 51 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392438

*equal contributions

Authors' addresses: Yu Fang, University of Pennsylvania; Ziyin Qu, University of Pennsylvania; Minchen Li, University of Pennsylvania; Xinxin Zhang, Tencent; Yixin Zhu, University of California, Los Angeles; Mridul Aanjaneya, Rutgers University; Chenfanfu Jiang, University of Pennsylvania.

0730-0301/2020/7-ART51 $15.00
https://doi.org/10.1145/3386569.3392438

1 INTRODUCTION

Modern day applications thrive on the dynamic interactions between solids and fluids, such as air-filled rubber tires, hydraulics, a flying airplane, a floating ship, windmills, etc. Apart from the natural appeal of modeling rich physics-based interactions in virtual environments for more realistic visual effects, there is a growing need for fast methods for solid-fluid coupling due to emerging applications in virtual surgery [Lee et al. 2018; Mitchell et al. 2015], digital fabrication [Ma et al. 2017], and soft robotics [Hu et al. 2019].

The simplest approach to solid-fluid coupling is via the use of *partitioned* schemes, that iterate separately between the solid and

Lagrangian Neural Style Transfer for Fluids

BYUNGSOO KIM, ETH Zurich
VINICIUS C. AZEVEDO, ETH Zurich
MARKUS GROSS, ETH Zurich
BARBARA SOLENTHALER, ETH Zurich

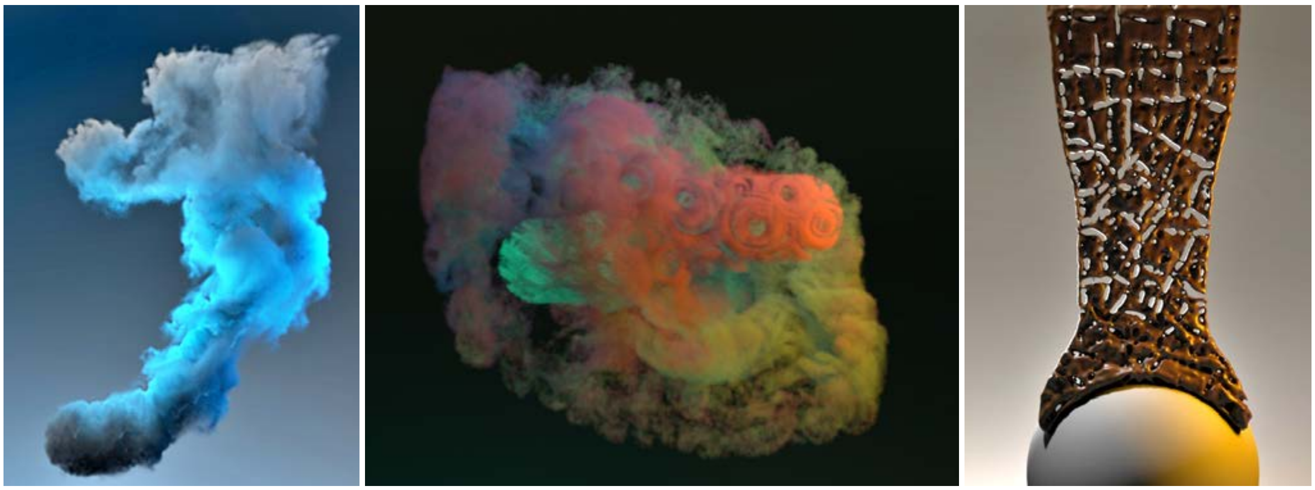


Fig. 1. Our Lagrangian neural style transfer enables novel artistic manipulations, such as time-coherent stylization of smoke, multiple fluids and liquids.

Artistically controlling the shape, motion and appearance of fluid simulations pose major challenges in visual effects production. In this paper, we present a neural style transfer approach from images to 3D fluids formulated in a Lagrangian viewpoint. Using particles for style transfer has unique benefits compared to grid-based techniques. Attributes are stored on the particles and hence are trivially transported by the particle motion. This intrinsically ensures temporal consistency of the optimized stylized structure and notably improves the resulting quality. Simultaneously, the expensive, recursive alignment of stylization velocity fields of grid approaches is unnecessary, reducing the computation time to less than an hour and rendering neural flow stylization practical in production settings. Moreover, the Lagrangian representation improves artistic control as it allows for multi-fluid stylization and consistent color transfer from images, and the generality of the method enables stylization of smoke and liquids likewise.

CCS Concepts: • **Computing methodologies → Physical simulation**; **Neural networks**.

Additional Key Words and Phrases: physically-based animation, fluid simulation, deep learning, neural style transfer

Authors' addresses: Byungsoo Kim, ETH Zurich, kimby@inf.ethz.ch; Vinicius C. Azevedo, ETH Zurich, vinicius.azevedo@inf.ethz.ch; Markus Gross, ETH Zurich, grossm@inf.ethz.ch; Barbara Solenthaler, ETH Zurich, solenthaler@inf.ethz.ch.

0730-0301/2020/7-ART52 $15.00
https://doi.org/10.1145/3386569.3392473

ACM Reference Format:
Byungsoo Kim, Vinicius C. Azevedo, Markus Gross, and Barbara Solenthaler. 2020. Lagrangian Neural Style Transfer for Fluids. *ACM Trans. Graph.* 39, 4, Article 52 (July 2020), 10 pages. https://doi.org/10.1145/3386569.3392473

1 INTRODUCTION

In visual effects production, physics-based simulations are not only used to realistically re-create natural phenomena, but also as a tool to convey stories and trigger emotions. Hence, artistically controlling the shape, motion and the appearance of simulations is essential for providing directability for physics. Specifically to fluids, the major challenge is the non-linearity of the underlying fluid motion equations, which makes optimizations towards a desired target difficult. Keyframe matching either through expensive fully-optimized simulations [McNamara et al. 2004; Pan and Manocha 2017; Treuille et al. 2003] or simpler distance-based forces [Nielsen and Bridson 2011; Raveendran et al. 2012] provide control over the shape of fluids. The fluid motion can be enhanced with turbulence synthesis approaches [Kim et al. 2008; Sato et al. 2018] or guided by coarse grid simulations [Nielsen and Bridson 2011], while patch-based texture composition [Gagnon et al. 2019; Jamriška et al. 2015] enables manipulation over appearance by automatic transfer of input 2D image patterns.

The recently introduced Transport-based Neural Style Transfer (TNST) [Kim et al. 2019a] takes flow appearance and motion control to a new level: arbitrary styles and semantic structures given by 2D input images are automatically transferred to 3D smoke simulations. The achieved effects range from natural turbulent structures

Learned Motion Matching

DANIEL HOLDEN, Ubisoft La Forge, Ubisoft, Canada
OUSSAMA KANOUN, Ubisoft La Forge, Ubisoft, Canada
MAKSYM PEREPICHKA, Concordia University, Canada
TIBERIU POPA, Concordia University, Canada

Fig. 1. Our method applied to various situations including navigating rough terrain, interaction with other characters, and using scene props.

In this paper we present a learned alternative to the Motion Matching algorithm which retains the positive properties of Motion Matching but additionally achieves the scalability of neural-network-based generative models. Although neural-network-based generative models for character animation are capable of learning expressive, compact controllers from vast amounts of animation data, methods such as Motion Matching still remain a popular choice in the games industry due to their flexibility, predictability, low preprocessing time, and visual quality - all properties which can sometimes be difficult to achieve with neural-network-based methods. Yet, unlike neural networks, the memory usage of such methods generally scales linearly with the amount of data used, resulting in a constant trade-off between the diversity of animation which can be produced and real world production budgets. In this work we combine the benefits of both approaches and, by breaking down the Motion Matching algorithm into its individual steps, show how learned, scalable alternatives can be used to replace each operation in turn. Our final model has no need to store animation data or additional matching meta-data in memory, meaning it scales as well as existing generative models. At the same time, we preserve the behavior of Motion Matching, retaining the quality, control, and quick iteration time which are so important in the industry.

CCS Concepts: • **Computing methodologies → Motion capture**.

Additional Key Words and Phrases: Motion Matching, Generative Models, Neural Networks, Character Animation, Animation,

ACM Reference Format:
Daniel Holden, Oussama Kanoun, Maksym Perepichka, and Tiberiu Popa. 2020. Learned Motion Matching. *ACM Trans. Graph.* 39, 4, Article 53 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392440

Authors' addresses: Daniel Holden, Ubisoft La Forge, Ubisoft, 5505 St Laurent Blvd, Montreal, QC, H2T 1S6, Canada, daniel.holden@ubisoft.com; Oussama Kanoun, Ubisoft La Forge, Ubisoft, 5505 St Laurent Blvd, Montreal, QC, H2T 1S6, Canada, oussama.kanoun@ubisoft.com; Maksym Perepichka, maksym@perepichka.com, Concordia University, 1493 Saint-Catherine Street West, Montreal, Quebec, Canada, H3G 2W1; Tiberiu Popa, tiberiu.popa@concordia.ca, Concordia University, 1493 Saint-Catherine Street West, Montreal, Quebec, Canada, H3G 2W1.

0730-0301/2020/7-ART53 $15.00
https://doi.org/10.1145/3386569.3392440

1 INTRODUCTION

In interactive applications such as video games, demand for larger, more immersive and dynamic worlds has steadily made it more difficult to produce characters which can respond realistically and naturally in the exponentially growing number of different situations that are presented to them. Meanwhile, the amount of data required has also slowly grown, and AAA video games now often contain tens of thousands of unique animations that all must be triggered in the correct context [Holden 2018].

Introduced by Clavet and Büttner [2015], *Motion Matching* is a method of searching a large database of animations for the animation which best fits the given context. This method has quickly been adopted by many studios due to its simplicity, flexibility, controllability, and the quality of the motion it produces [Büttner 2019; Clavet 2016; Harrower 2018; Hussain 2019; Zinno 2019]. Rather than specifying the fine-grained animation logic via a state-graph, Motion Matching allows animators to specify the properties of the animation which should be produced, and the best fitting match is selected automatically via a nearest neighbor search. When combined with large amounts of data, Motion Matching proves a simple and effective way of dealing with the vast number of possible transitions and interactions that are required by a modern AAA video game. Additionally, since Motion Matching plays back the animation data stored in the database as-is, with only simple blending and post-processing such as inverse kinematics applied, quality is generally preserved, animators retain a level of control, and the behaviour can be tracked and debugged with appropriate tools. Finally, since it has minimal training/pre-processing time, adjustments can often

Local Motion Phases for Learning Multi-Contact Character Movements

SEBASTIAN STARKE, University of Edinburgh, UK and Electronic Arts, USA
YIWEI ZHAO, Electronic Arts, USA
TAKU KOMURA, University of Edinburgh, UK
KAZI ZAMAN, Electronic Arts, USA

Fig. 1. A selection of results using our method to generate ball-dribbling movements and interaction behaviours with other characters.

Training a bipedal character to play basketball and interact with objects, or a quadruped character to move in various locomotion modes, are difficult tasks due to the fast and complex contacts happening during the motion. In this paper, we propose a novel framework to learn fast and dynamic character interactions that involve multiple contacts between the body and an object, another character and the environment, from a rich, unstructured motion capture database. We use one-on-one basketball play and character interactions with the environment as examples. To achieve this task, we propose a novel feature called local motion phase, that can help neural networks to learn asynchronous movements of each bone and its interaction with external objects such as a ball or an environment. We also propose a novel generative scheme to reproduce a wide variation of movements from abstract control signals given by a gamepad, which can be useful for changing the style of the motion under the same context. Our scheme is useful for animating contact-rich, complex interactions for real-time applications such as computer games.

CCS Concepts: • **Computing methodologies** → **Motion capture**; *Neural networks.*

Additional Key Words and Phrases: neural networks, human motion, character animation, character control, character interactions, deep learning

Authors' addresses: Sebastian Starke, University of Edinburgh, 10 Crichton Street, Edinburgh, EH8 9AB, UK, sebastian.starke@ed.ac.uk, Electronic Arts, USA, sstarke@ea.com; Yiwei Zhao, yiwzhao@ea.com, Electronic Arts, USA; Taku Komura, tkomura@ed.ac.uk, University of Edinburgh, 10 Crichton Street, Edinburgh, EH8 9AB, UK; Kazi Zaman, kzaman@ea.com, Electronic Arts, USA.

0730-0301/2020/7-ART54 $15.00
https://doi.org/10.1145/3386569.3392450

ACM Reference Format:
Sebastian Starke, Yiwei Zhao, Taku Komura, and Kazi Zaman. 2020. Local Motion Phases for Learning Multi-Contact Character Movements. *ACM Trans. Graph.* 39, 4, Article 54 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392450

1 INTRODUCTION

There is a huge demand in simulating fast and complex interactions that involve multiple contacts between a character and objects, an environment, and other characters, especially in computer games and films. For example, for basketball games, the players need to dribble the ball while making various movements with different foot-fall patterns to compete with the opponent characters.

Previous techniques to learn from unstructured motion capture database have limitations in terms of scalability, realism and variation in the movements. Firstly, most techniques require aligning the motions by a global temporal parameter such as the phase, which is often difficult when the motion involves multiple contacts that are asynchronous. Secondly, even when the motions are learned by the controller, there can be issues reproducing a wide variation of movements from low dimensional control signals such as those provided by the user through keyboards or gamepads.

In this paper, we propose a novel data-driven framework to learn fast and dynamic interactions that involve multiple contacts. We make use of a large database of one-on-one basketball play as our main example, where one player catches, dribbles and plays tricks with the ball, while avoiding the opponent player who tries to defend and intercept the ball. We design and train a neural character controller that can learn and produce realistic offense and defense actions under a unified framework, so that the players can easily switch from the offense to the defense during the play.

To let the model learn movements that involve fast and complex interactions where the contacts between the body and the ball or the ground quickly switch in an asynchronous manner, we propose

N-Dimensional Rigid Body Dynamics

MARC TEN BOSCH, mtb design works, Inc., USA

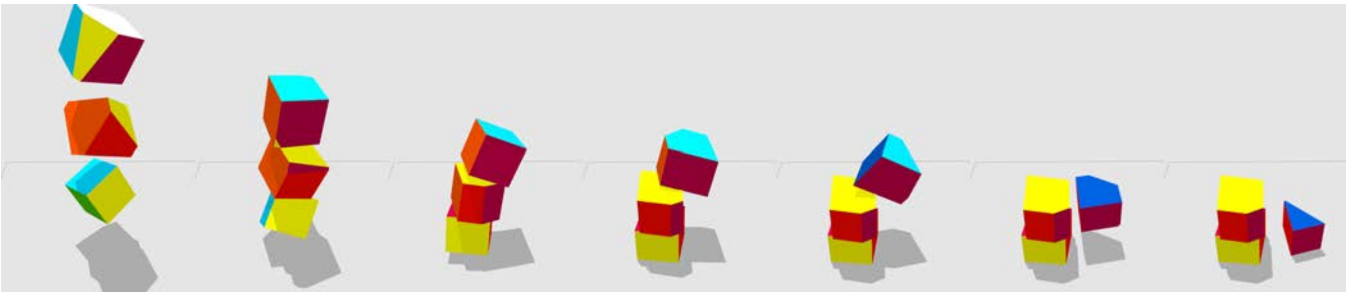


Fig. 1. A stack of three 4D hypercubes

I present a formulation for Rigid Body Dynamics that is independent of the dimension of the space. I describe the state and equations of motion of rigid bodies using geometric algebra. Using collision detection algorithms extended to nD I resolve collisions and contact between bodies. My implementation is 4D, but the techniques described here apply to any number of dimensions. I display these four-dimensional rigid bodies by taking a three-dimensional slice through them. I allow the user to manipulate these bodies in real-time.

CCS Concepts: • **Computing methodologies → Physical simulation**; **Collision detection**; *Virtual reality*.

Additional Key Words and Phrases: Rigid Bodies, N-Dimensional, Physics, Geometric Algebra, Fourth Dimension

ACM Reference Format:
Marc ten Bosch. 2020. N-Dimensional Rigid Body Dynamics. *ACM Trans. Graph.* 39, 4, Article 55 (July 2020), 6 pages. https://doi.org/10.1145/3386569.3392483

1 INTRODUCTION

Our experience of physical space is three dimensional. Consequently, physically-based simulations (physics engines) have so far been focused on and restricted to the two and three-dimensional cases. However using the appropriate formulation of the required equations it is possible to extend them to higher dimensions. Geometric algebra provides a simple dimension-independent formulation. This allows real time manipulation of n-dimensional shapes that collide with each other as if they were real objects, which makes them much less abstract, in stark contrast with most people's experience of them. While attention has been given to understanding and visualizing high-dimensional spaces and other abstract mathematical concepts, it has most often remained limited to visualizing these concepts without any physicality or object-to-object relationships.

Contributions. The contributions of this paper include:

(1) extending the geometric algebra-based formulation of classical 3D rigid body dynamics to nD. By representing geometric algebra operators as matrices, one can in particular construct, diagonalize and transform the inertia tensor for arbitrary nD simplicial meshes for any n, in a simple way. This enables formulating the Euler equation in nD, which allows e.g. studying the case of the 4D Euler equation under torque-free conditions.
(2) computing collision and contact processing in nD, including static and kinetic friction. I give an nD formulation for the Minkowski difference and separating axis theorem collision detection methods based on geometric algebra.
(3) a method of interacting with 4D objects that is akin to our 3D experience of reality.

2 RELATED WORK

Interactive Simulation of 3D Rigid Body Dynamics is a broad field. Bender et al. [2014] provide a survey.

[Cameron 1990] has formulated the 3D continuous collision detection problem as a discrete 4D collision problem by considering the extrusion of each object over time.

Visualizing 4D objects is an interesting and challenging problem of its own, with a long history [Abbott 1884; Banchoff 1990; Chu et al. 2009; Hilbert and Cohn-Vossen 1952]. Many ways of manipulating 4D shapes have been proposed [Yan et al. 2012; Zhang and Hanson 2006].

Geometric algebra operates on a space of elements called multivectors, of which vectors are a subspace. While vectors can be thought of as oriented line-segments, other elements can represent oriented areas (bivectors), volumes (trivectors), and so on. Geometric algebra also defines operations that correspond to translations and rotations of these elements. The book by Macdonald [2011]

Author's address: Marc ten Bosch, mtb design works, Inc., San Francisco, CA, USA, marc@marctenbosch.com.

0730-0301/2020/7-ART55 $15.00
https://doi.org/10.1145/3386569.3392483

Phong Deformation: A better C^0 interpolant for embedded deformation

DOUG L. JAMES, Pixar Animation Studios and Stanford University

Physics-based simulations of deforming tetrahedral meshes are widely used to animate detailed embedded geometry. Unfortunately most practitioners still use linear interpolation (or other low-order schemes) on tetrahedra, which can produce undesirable visual artifacts, e.g., faceting and shading artifacts, that necessitate increasing the simulation's spatial resolution and, unfortunately, cost.

In this paper, we propose *Phong Deformation,* a simple, robust and practical vertex-based quadratic interpolation scheme that, while still only C^0 continuous like linear interpolation, greatly reduces visual artifacts for embedded geometry. The method first averages element-based linear deformation models to vertices, then barycentrically interpolates the vertex models while also averaging with the traditional linear interpolation model. The method is a fast, robust, and easily implemented replacement for linear interpolation that produces visually better results for embedded deformation with irregular tetrahedral meshes.

CCS Concepts: • **Computing methodologies** → *Modeling and simulation*; *Physical simulation.*

Additional Key Words and Phrases: Embedded deformation, tetrahedra, interpolation, cell-to-vertex reconstruction.

ACM Reference Format:
Doug L. James. 2020. Phong Deformation: A better C^0 interpolant for embedded deformation. *ACM Trans. Graph.* 39, 4, Article 56 (July 2020), 9 pages. https://doi.org/10.1145/3386569.3392371

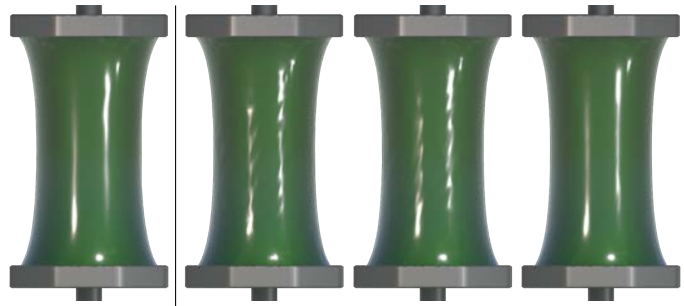


Fig. 1. **Bad embedded deformations make simulations look worse!** (FarLeft) An elastic cylinder is stretched between two plates in this Houdini FEM simulation (70,476 tetrahedra) and looks reasonable. However, embedded deformation of a high-resolution cylinder mesh (30,720 triangles) produces unsightly visual artifacts when using popular displacement interpolants: (MidLeft) Houdini FEM's default "attribute transfer" approach, or (MidRight) traditional linear interpolation. Such artifacts can lead practitioners to unnecessarily increase the FEM simulation resolution (which is very costly), or decrease the tri-mesh resolution, to hide artifacts. (FarRight) In contrast, our proposed Phong Deformation scheme provides a fast, high-quality embedded deformation with minimal artifacts.

1 INTRODUCTION

Physics-based animations of detailed geometric models are often performed by embedding the latter in simulated tetrahedral meshes, with coarse tetrahedral meshes used for speed whenever possible. By far the most common approach for transferring deformations from tetrahedral simulation meshes to high-resolution embedded meshes is linear interpolation within each tetrahedron, and for good reason: it is simple, fast, and trivial to implement. Unfortunately, linear interpolation produces only C^0 continuous displacement fields that have unsightly visual artifacts, such as faceting, in deformations and shading (see Figure 1). Popular alternatives, such as kernel-based scattered data interpolation schemes like Houdini's "attribute transfer" [SideFX 2020] can reduce faceting but introduce other artifacts. Animators seeking to avoid such artifacts are forced to increase the simulation resolution and lament the cost of FEM solvers, and/or reduce the resolution of embedded meshes. While higher continuity than C^0 is not strictly necessary when deforming embedded meshes, reducing interpolation errors for embedded deformation would benefit practitioners everywhere.

Author's address: Doug L. JamesPixar Animation Studios; Stanford University, djames@cs.stanford.edu.

0730-0301/2020/7-ART56
https://doi.org/10.1145/3386569.3392371

In this paper, we introduce *Phong Deformation,* a simple and practical interpolation scheme that, while still only C^0 continuous, greatly reduces visual artifacts for embedded deformations using coarse tetrahedral meshes (see Figures 1 and 3). Phong Deformation is, like its counterpart, linear interpolation, inherently simple, fast, robust, and easy to implement. However it is far more effective at reducing interpolation artifacts due to its quadratic deformation. Mathematically, a robust cell-to-vertex reconstruction is first performed to estimate second-order-accurate deformation gradients, and related linear deformation models, at vertices. Embedded deformation then involves barycentrically interpolating these vertex deformers across the element while also averaging with the traditional barycentric (linear) interpolation model (see Figure 2). Phong Deformation is fast, trivial to implement, affine invariant, and it produces visually better results than linear interpolation in practice, in part because the latter is only a second-order accurate interpolation scheme, whereas we show that Phong Deformation can achieve third-order accuracy, i.e., $O(h^2)$ vs $O(h^3)$ truncation error. The Phong deformer is similar to classical 10-node tetrahedral quadratic interpolants, but its novel construction is entirely vertex based and so it avoids the computation and storage of intermediate edge midpoint values, and it has a simple evaluation based on barycentric interpolation of vertex quantities. Finally, a key contribution is the robust regularized estimation of vertex gradients such that Phong Deformation gracefully degrades from cubic- to second-order accuracy (depending on what neighbor data is available), and never fails in practice.

Projective Dynamics with Dry Frictional Contact

MICKAËL LY, Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France
JEAN JOUVE, ENS Rennes and Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France
LAURENCE BOISSIEUX, Univ. Grenoble Alpes, Inria, France
FLORENCE BERTAILS-DESCOUBES, Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France

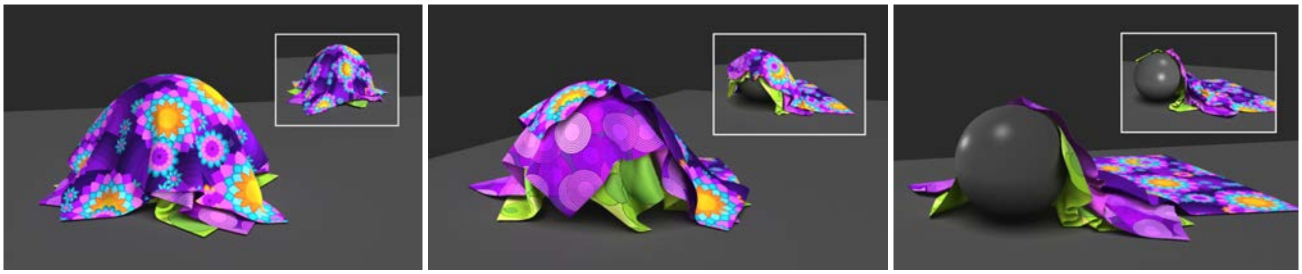


Fig. 1. Our frictional contact algorithm embedded inside Projective dynamics [Bouaziz et al. 2014] makes it possible to simulate challenging self-contacting scenarios, even at a modest number of iterations (20 iterations here). Compared to reference implicit contact solvers of the literature (here, Argus [Li et al. 2018] in inset, with fixed mesh resolution), it performs an order of magnitude faster while generating similar visual effects at the macroscopic scale.

Projective dynamics was introduced a few years ago as a fast method to yield an approximate yet stable solution to the dynamics of nodal systems subject to stiff internal forces. Previous attempts to include contact forces in that framework considered adding a quadratic penalty energy to the global system, which however broke the simple – constant matrix – structure of the global linear equation, while failing to treat contact in an implicit manner. In this paper we propose a simple yet effective method to integrate in a unified and semi-implicit way contact as well as dry frictional forces into the nested architecture of Projective dynamics. Assuming that contacts apply to nodes only, the key is to split the global matrix into a diagonal and a positive matrix, and use this splitting in the local step so as to make a good prediction of frictional contact forces at next iteration. Each frictional contact force is refined independently in the local step, while the original efficient structure of the global step is left unchanged. We apply our algorithm to cloth simulation and show that contact and dry friction can be captured at a reasonable precision within a few iterations only, hence one order of magnitude faster compared to global implicit contact solvers of the literature.

CCS Concepts: • **Computing methodologies → Animation**; **Physical simulation**.

Additional Key Words and Phrases: Projective Dynamics, dry frictional contact, cloth simulation, interactive simulation

Authors' addresses: Mickaël Ly, Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000, Grenoble, France, mickael.ly@inria.fr; Jean Jouve, ENS Rennes, Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000, Grenoble, France, jean.jouve@ens-rennes.fr; Laurence Boissieux, Univ. Grenoble Alpes, Inria, 38000, Grenoble, France, laurence.boissieux@inria.fr; Florence Bertails-Descoubes, Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000, Grenoble, France, florence.descoubes@inria.fr.

0730-0301/2020/7-ART57 $15.00
https://doi.org/10.1145/3386569.3392396

ACM Reference Format:
Mickaël Ly, Jean Jouve, Laurence Boissieux, and Florence Bertails-Descoubes. 2020. Projective Dynamics with Dry Frictional Contact. *ACM Trans. Graph.* 39, 4, Article 57 (July 2020), 8 pages. https://doi.org/10.1145/3386569.3392396

1 INTRODUCTION

Properly accounting for contact and friction is a major challenge when simulating dynamic solid objects, such as 3D soft bodies, cloth, or hair. During motion, objects should not interpenetrate, but still interact in a stable way. At contact, they should dissipate energy in a realistic manner, following the Coulomb law for dry friction, which captures the typical threshold effect between stick and slip. In the last decades, a number of implicit algorithms have been designed so as to satisfy these major requirements in a stable and robust way [Daviet et al. 2011; Li et al. 2018; Otaduy et al. 2009]. However, such solvers currently remain inaccessible to interactive scenarios. Instead, interactivity is often obtained at the price of a simplification and/or a decoupled treatment of the interaction model [Kavan et al. 2011; Macklin et al. 2014]. Our goal in this paper is to build a realistic numerical model for frictional contact, fully coupled with the dynamics, which gets closer to interactive needs.

Consistently degradable model. Simulating complex deformable objects in a both realistic and interactive fashion is more than ever a topical issue, as new applications like virtual try on or fast prototyping, requiring both faithfulness and speed, are being deployed [Bartle et al. 2016; Wang 2018]. In addition, possessing tools to design an animation sequence interactively before proceeding to (offline) high-accuracy simulation constitutes a longstanding issue in the field of feature film production [Barbič et al. 2012]. In all cases, having access to a *consistently degradable* model, that is, a discrete model which converges to the desired continuous model while generating *satisfactory* (in a sense to be defined) simulations at low cost, is truly beneficial. In our case, we will qualify a simulation of satisfactory if

RigNet: Neural Rigging for Articulated Characters

ZHAN XU, University of Massachusetts Amherst
YANG ZHOU, University of Massachusetts Amherst
EVANGELOS KALOGERAKIS, University of Massachusetts Amherst
CHRIS LANDRETH, University of Toronto
KARAN SINGH, University of Toronto

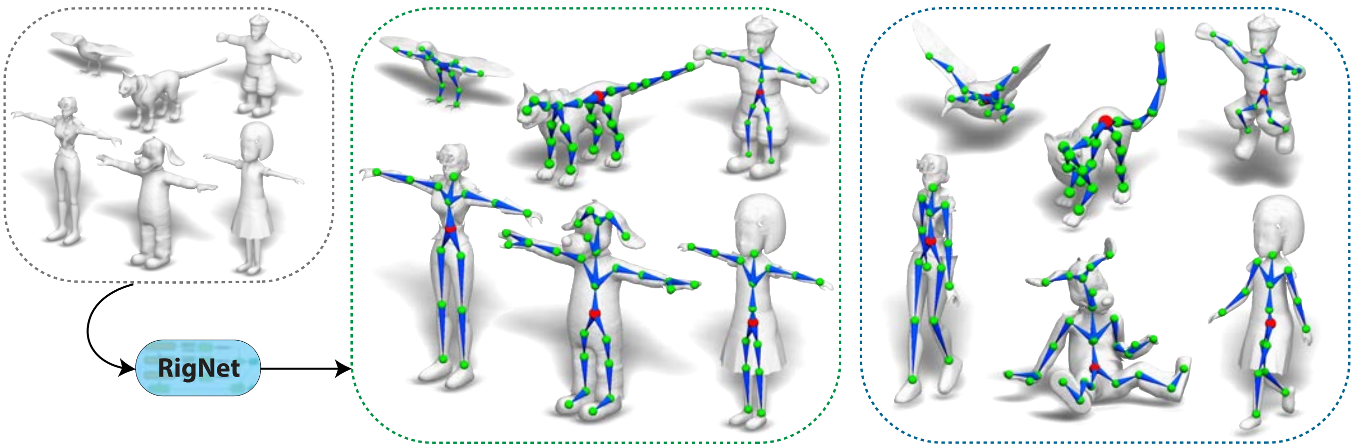


Fig. 1. Given a 3D character mesh, RigNet produces an animation skeleton and skin weights tailored to the articulation structure of the input character. From left to right: input examples of test 3D meshes, predicted skeletons for each of them (joints are shown in green and bones in blue), and resulting skin deformations under different skeletal poses. Please see also our supplementary video: https://youtu.be/J90VETgWIDg

We present *RigNet*, an end-to-end automated method for producing animation rigs from input character models. Given an input 3D model representing an articulated character, *RigNet* predicts a skeleton that matches the animator expectations in joint placement and topology. It also estimates surface skin weights based on the predicted skeleton. Our method is based on a deep architecture that directly operates on the mesh representation without making assumptions on shape class and structure. The architecture is trained on a large and diverse collection of rigged models, including their mesh, skeletons and corresponding skin weights. Our evaluation is three-fold: we show better results than prior art when quantitatively compared to animator rigs; qualitatively we show that our rigs can be expressively posed and animated at multiple levels of detail; and finally, we evaluate the impact of various algorithm choices on our output rigs. [1]

[1]Our project page with source code, datasets, and supplementary video is available at https://zhan-xu.github.io/rig-net

Authors' addresses: Zhan Xu, University of Massachusetts Amherst, zhanxu@cs.umass.edu; Yang Zhou, University of Massachusetts Amherst, yangzhou@cs.umass.edu; Evangelos Kalogerakis, University of Massachusetts Amherst, kalo@cs.umass.edu; Chris Landreth, University of Toronto, chrisl@dgp.toronto.edu; Karan Singh, University of Toronto, karan@dgp.toronto.edu.

0730-0301/2020/7-ART58 $15.00
https://doi.org/10.1145/3386569.3392379

CCS Concepts: • **Computing methodologies** → **Animation**; *Machine learning approaches*.

Additional Key Words and Phrases: character rigging, animation skeletons, skinning, neural networks

ACM Reference Format:
Zhan Xu, Yang Zhou, Evangelos Kalogerakis, Chris Landreth, and Karan Singh. 2020. RigNet: Neural Rigging for Articulated Characters. *ACM Trans. Graph.* 39, 4, Article 58 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392379

1 INTRODUCTION

There is a rapidly growing need for diverse, high-quality, animation-ready characters and avatars in the areas of games, films, mixed Reality and social media. Hand-crafted character "rigs", where users create an animation "skeleton" and bind it to an input mesh (or "skin"), have been the workhorse of articulated figure animation for over three decades. The skeleton represents the articulation structure of the character, and skeletal joint rotations provide an animator with direct hierarchical control of character pose.

We present a deep-learning based solution for automatic rig creation from an input 3D character. Our method predicts both a skeleton and skinning that match animator expectations (Figures 1, 10). In contrast to prior work that fits pre-defined skeletal templates of fixed joint count and topology to input 3D meshes [Baran and Popović 2007], our method outputs skeletons more tailored to the underlying articulation structure of the input. Unlike pose estimation approaches designed for particular shape classes, such as humans or hands [Haque et al. 2016; Huang et al. 2018; Moon et al. 2018;

Robust Eulerian-on-Lagrangian Rods

ROSA M. SÁNCHEZ-BANDERAS, Universidad Rey Juan Carlos
ALEJANDRO RODRÍGUEZ, SEDDI Labs
HÉCTOR BARREIRO, Universidad Rey Juan Carlos
MIGUEL A. OTADUY, Universidad Rey Juan Carlos

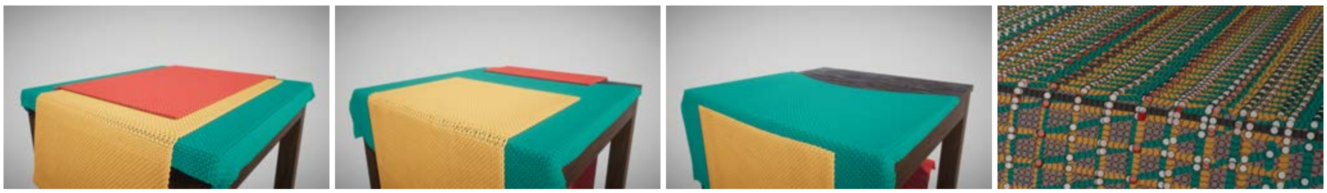


Fig. 1. Simulation of three yarn-level tablecloths stacked on top of each other. We handle both intra-fabric and inter-fabric contacts implicitly with our novel Eulerian-on-Lagrangian (EoL) simulation. Frictional contact is correctly handled, even under extreme sliding and crossing of yarns. We introduce novel EIL nodes to handle robustly pervasive degeneracies in the discretization. The rightmost image shows the soup of nodes around the edge of the table. White dots represent contacts between standard EoL nodes, red dots contacts between our novel EIL nodes, and pink dots contacts between EoL and EIL nodes.

This paper introduces a method to simulate complex rod assemblies and stacked layers with implicit contact handling, through Eulerian-on-Lagrangian (EoL) discretizations. Previous EoL methods fail to handle such complex situations, due to ubiquitous and intrinsic degeneracies in the contact geometry, which prevent the use of remeshing and make simulations unstable. We propose a novel mixed Eulerian-Lagrangian discretization that supports accurate and efficient contact as in EoL methods, but is transparent to internal rod forces, and hence insensitive to degeneracies. By combining the standard and novel EoL discretizations as appropriate, we derive mixed statics-dynamics equations of motion that can be solved in a unified manner with standard solvers. Our solution is simple and elegant in practice, and produces robust simulations on large-scale scenarios with complex rod arrangements and pervasive degeneracies. We demonstrate our method on multi-layer yarn-level cloth simulations, with implicit handling of both intra- and inter-layer contacts.

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: Cloth, yarns, rods.

ACM Reference Format:
Rosa M. Sánchez-Banderas, Alejandro Rodríguez, Héctor Barreiro, and Miguel A. Otaduy. 2020. Robust Eulerian-on-Lagrangian Rods. *ACM Trans. Graph.* 39, 4, Article 59 (July 2020), 10 pages. https://doi.org/10.1145/3386569.3392489

Authors' addresses: Rosa M. Sánchez-Banderas, Universidad Rey Juan Carlos, rosanban@gmail.com; Alejandro Rodríguez, SEDDI Labs, alejandro.rodriguez@seddi.com; Héctor Barreiro, Universidad Rey Juan Carlos, hecbarcab@gmail.com; Miguel A. Otaduy, Universidad Rey Juan Carlos, miguel.otaduy@urjc.es.

0730-0301/2020/7-ART59 $15.00
https://doi.org/10.1145/3386569.3392489

1 INTRODUCTION

The simulation of rods has been extensively studied in computer graphics. This is not surprising, as many daily-life objects are composed of rod-like structures, such as ropes [Bergou et al. 2008; Pai 2002], chains, belts, cloth [Kaldor et al. 2008], cables [Servin et al. 2010], hair [Selle et al. 2008], tendons [Sachdeva et al. 2015], spaghetti, or even fluid filaments [Bergou et al. 2010].

Rod simulations become particularly challenging under complex contact arrangements, as their small cross-section makes them vulnerable to collision handling errors. But when contacts persist over time, with rods sliding with respect to each other, Eulerian-on-Lagrangian (EoL) discretizations [Sueda et al. 2011] offer an attractive approach to gain robustness in contact handling, at the expense of a slightly more complex derivation of the equations of motion. EoL methods augment the classic Lagrangian discretization of deformable solids with Eulerian coordinates that allow nodes to move in the material domain. The power of EoL methods is the ability to track explicitly contact points both in the spatial and material domains, simply by placing nodes at contact locations, and thus reduce the complexity and increase the accuracy of contact handling.

However, we have observed that existing EoL works exploit only moderately this power. To the best of our knowledge, none of them shows multiple stacked layers of rods or shells sliding with respect to each other. As we discuss in detail later in Section 3, there is a fundamental reason for this. With sliding contacts, EoL discretizations easily become degenerate. To avoid simulation instabilities, previous EoL works use remeshing, but this strategy cannot be used under multiple stacked layers, when the geometry of the contact configuration is intrinsically degenerate, with contacts crossing each other.

We present a simulation method that handles robustly degenerate discretizations in EoL rods. Our method does not use remeshing; it relies instead on a formulation of equations of motion that is insensitive to degenerate elements. As a result, we can simulate stacks of rods sandwiched between sliding contacts, with nodes

Robust Motion In-betweening

FÉLIX G. HARVEY, Polytechnique Montreal, Canada, Mila, Canada, and Ubisoft Montreal, Canada
MIKE YURICK, Ubisoft Montreal, Canada
DEREK NOWROUZEZAHRAI, McGill University, Canada and Mila, Canada
CHRISTOPHER PAL, CIFAR AI Chair, Canada, Polytechnique Montreal, Canada, Mila, Canada, and Element AI, Canada

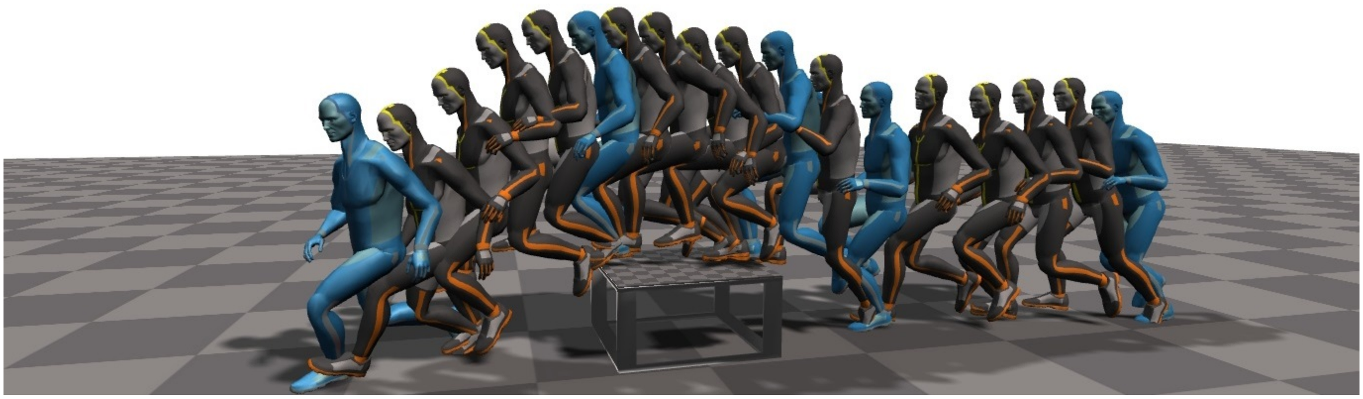


Fig. 1. Transitions automatically generated by our system between target keyframes (in blue). For clarity, only one in four generated frames is shown. Our tool allows for generating transitions of variable lengths and for sampling different variations of motion given fixed keyframes.

In this work we present a novel, robust transition generation technique that can serve as a new tool for 3D animators, based on adversarial recurrent neural networks. The system synthesises high-quality motions that use temporally-sparse keyframes as animation constraints. This is reminiscent of the job of *in-betweening* in traditional animation pipelines, in which an animator draws motion frames between provided keyframes. We first show that a state-of-the-art motion prediction model cannot be easily converted into a robust transition generator when only adding conditioning information about future keyframes. To solve this problem, we then propose two novel additive embedding modifiers that are applied at each timestep to latent representations encoded inside the network's architecture. One modifier is a *time-to-arrival embedding* that allows variations of the transition length with a single model. The other is a *scheduled target noise* vector that allows the system to be robust to target distortions and to sample different transitions given fixed keyframes. To qualitatively evaluate our method, we present a custom MotionBuilder plugin that uses our trained model to perform in-betweening in production scenarios. To quantitatively evaluate performance on transitions and generalizations to longer time horizons, we present well-defined in-betweening benchmarks on a subset of the widely used Human3.6M dataset and on LaFAN1, a novel high quality motion capture dataset that is more appropriate for transition generation. We are releasing this new dataset along with this work, with accompanying code for reproducing our baseline results.

Authors' addresses: Félix G. Harvey, Polytechnique Montreal, 2500 Chemin de la Polytechique, Montreal, QC, H3T 1J4, Canada, Mila, 6666 St-Urbain Street, #200, Montreal, QC, H2S 3H1, Canada, Ubisoft Montreal, 5505 Boul Saint-Laurent, #2000, Montreal, QC, H2T 1S6, Canada, felix.gingras-harvey@polymtl.ca; Mike Yurick, Ubisoft Montreal, 5505 Boul Saint-Laurent, #2000, Montreal, QC, H2T 1S6, Canada, mike.yurick@ubisoft.com; Derek Nowrouzezahrai, McGill University, 3480 University St, Montreal, QC, H3A 0E9, Canada, Mila, 6666 St-Urbain Street, #200, Montreal, QC, H2S 3H1, Canada, derek@cim.mcgill.ca; Christopher Pal, CIFAR AI Chair, 661 University Ave., Suite 505, Toronto, ON, M5G 1M1, Canada, Polytechnique Montreal, 2500 Chemin de la Polytechique, Montreal, QC, H3T 1J4, Canada, Mila, 6666 St-Urbain Street, #200, Montreal, QC, H2S 3H1, Canada, Element AI, 6650 St-Urbain Street, #500, Montreal, QC, H2S 3G9, Canada.

0730-0301/2020/7-ART60 $15.00
https://doi.org/10.1145/3386569.3392480

CCS Concepts: • **Computing methodologies → Motion capture**; *Neural networks*.

Additional Key Words and Phrases: animation, locomotion, transition generation, in-betweening, deep learning, LSTM

ACM Reference Format:
Félix G. Harvey, Mike Yurick, Derek Nowrouzezahrai, and Christopher Pal. 2020. Robust Motion In-betweening. *ACM Trans. Graph.* 39, 4, Article 60 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392480

1 INTRODUCTION

Human motion is inherently complex and stochastic for long-term horizons. This is why Motion Capture (MOCAP) technologies still often surpass generative modeling or traditional animation techniques for 3D characters with many degrees of freedom. However, in modern video games, the number of motion clips needed to properly animate a complex character with rich behaviors is often very large and manually authoring animation sequences with keyframes or using a MOCAP pipeline are highly time-consuming processes. Some methods to improve curve fitting between keyframes [Ciccone et al. 2019] or to accelerate the MOCAP workflow [Holden 2018] have been proposed to improve these processes. On another

Simple and Scalable Frictional Contacts for Thin Nodal Objects

GILLES DAVIET, Weta Digital, New Zealand

Fig. 1. A virtual character going through a running cycle, letting his hair impact repeatedly the back and neck of his shirt. The groom consists of 54, 450 Discrete Elastic Rods, totalling $1.2M$ degrees of freedom, while the shirt mesh contains about 27, 000 vertices. This scene induces as much as $4.5M$ contacts, which are solved implicitly and with nonlinear Coulomb friction thanks to our proposed algorithm. © Weta Digital.

Frictional contacts are the primary way by which physical bodies interact, yet they pose many numerical challenges. Previous works have devised robust methods for handling collisions in elastic bodies, cloth, or fiber assemblies such as hair, but the performance of many of those algorithms degrades when applied to objects with different topologies or constitutive models, or simply cannot scale to high-enough numbers of contacting points.

In this work we propose a unified approach, able to handle a large class of dynamical objects, that can solve for millions of contacts with unbiased Coulomb friction while keeping computation time and memory usage reasonable. Our method allows seamless coupling between the various simulated components that comprise virtual characters and their environment. Furthermore, our proposed approach is simple to implement and can be easily integrated in popular time integrators such as Projected Newton or ADMM.

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: Contact dynamics, Coulomb friction

ACM Reference Format:
Gilles Daviet. 2020. Simple and Scalable Frictional Contacts for Thin Nodal Objects. *ACM Trans. Graph.* 39, 4, Article 61 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392439

1 INTRODUCTION

The continuously increasing demand for visual richness of virtual environments has prompted the use of physical simulation to generate an ever larger portion of the final rendered frames. While the dynamics of the various objects that compose virtual environments can in themselves be of great interest, a large part of the visual complexity of their shape and motion can be attributed to the interactions that they have with surrounding components. In many cases, these interactions take the form of contact with dry friction; consider examples as diverse as a granular material, the leaves of a tree gently colliding under a breeze, or hair strands subtly entangled with neighbouring fibers, themselves lying upon layers of cloth and eventually skin.

Yet, simulation of dry friction has always proved numerically challenging, and robust approaches capable of efficiently handling these various object topologies and interactions in a unified manner are still scarce. Many previous works have focused on either methods capable of treating collisions between a very high number of "simple" components, such as granular materials, rigid bodies, or fiber assemblies, or at the other end of the spectrum a more reasonable number of large deformable bodies, such as flesh or cloth. Another dichotomy has also often been made between volumetric and thin objects, as slight amounts of inter-penetration can be tolerated for the former, but would immediately lead to hard-to-recover-from artefacts for the latter.

In this work, we propose a novel, simple yet scalable approach for the numerical simulation of hard contacts with nonlinear Coulomb friction for a large class of dynamical objects. Our method is especially compelling for assemblies of thin objects where the ratio of degrees of freedom to contact points is unfavorable, as illustrated in Fig. 1 where the garments of a character and the tens of thousands of fibers that comprise their hairstyle are simulated together, amounting to millions of degrees of freedom and contact points.

2 RELATED WORK

Due to the ubiquity of frictional contacts in physical and virtual environments, treatment of collisions in numerical simulations has been the subject of a vast amount of literature over the last decades.

2.1 Penalty forces

Inspired by molecular dynamics, the first methods developed in mechanics [Cundall and Strack 1979] and Computer Graphics [Moore and Wilhelms 1988] suggested to resolve contacts through the application of penalties, repulsive forces of intensity proportional to the inter-penetration depth of the colliding bodies. While Baraff and Witkin [1998] proposed an implicit integration scheme for elastic and collision energies that alleviated the drastic timestep restrictions

Author's address: Gilles Daviet, Weta Digital, New Zealand, gdaviet@wetafx.co.nz.

0730-0301/2020/7-ART61 $15.00
https://doi.org/10.1145/3386569.3392439

Skeleton-Aware Networks for Deep Motion Retargeting

KFIR ABERMAN*, AICFVE, Bejing Film Academy & Tel-Aviv University
PEIZHUO LI*, CFCS, Peking University & AICFVE, Beijing Film Academy
DANI LISCHINSKI, The Hebrew University of Jerusalem & AICFVE, Beijing Film Academy
OLGA SORKINE-HORNUNG, ETH Zurich & AICFVE, Beijing Film Academy
DANIEL COHEN-OR, Tel-Aviv University & AICFVE, Beijing Film Academy
BAOQUAN CHEN†, CFCS, Peking University & AICFVE, Beijing Film Academy

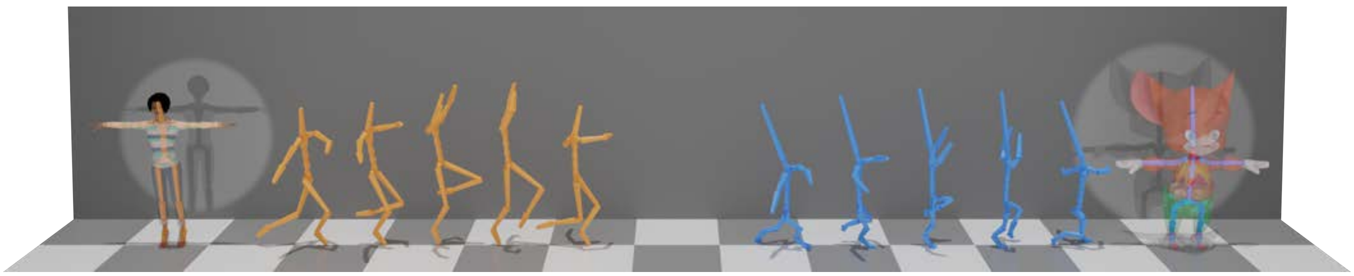


Fig. 1. Unpaired, cross-structural, motion retargeting. An input motion sequence (orange skeletons) is retargeted to a target skeleton (rightmost, blue), which has different body proportions, as well as a different number of bones (marked in red).

We introduce a novel deep learning framework for data-driven motion retargeting between skeletons, which may have different structure, yet corresponding to homeomorphic graphs. Importantly, our approach learns how to retarget without requiring any explicit pairing between the motions in the training set.

We leverage the fact that different homeomorphic skeletons may be reduced to a common *primal skeleton* by a sequence of edge merging operations, which we refer to as *skeletal pooling*. Thus, our main technical contribution is the introduction of novel differentiable convolution, pooling, and unpooling operators. These operators are *skeleton-aware*, meaning that they explicitly account for the skeleton's hierarchical structure and joint adjacency, and together they serve to transform the original motion into a collection of deep temporal features associated with the joints of the primal skeleton. In other words, our operators form the building blocks of a new deep motion processing framework that embeds the motion into a common latent space, shared by a collection of homeomorphic skeletons. Thus, retargeting can be achieved simply by encoding to, and decoding from this latent space.

Our experiments show the effectiveness of our framework for motion retargeting, as well as motion processing in general, compared to existing approaches. Our approach is also quantitatively evaluated on a synthetic dataset that contains pairs of motions applied to different skeletons. To the best of our knowledge, our method is the first to perform retargeting between skeletons with differently sampled kinematic chains, without any paired examples.

CCS Concepts: • **Computing methodologies → Motion processing**; **Neural networks**.

Additional Key Words and Phrases: Neural motion processing, motion retargeting

ACM Reference Format:
Kfir Aberman, Peizhuo Li, Dani Lischinski, Olga Sorkine-Hornung, Daniel Cohen-Or, and Baoquan Chen. 2020. Skeleton-Aware Networks for Deep Motion Retargeting. *ACM Trans. Graph.* 39, 4, Article 62 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392462

*equal contribution
†corresponding author

Authors' addresses: Kfir Aberman, kfiraberman@gmail.com; Peizhuo Li, peizhuo@pku.edu.cn; Dani Lischinski, danix3d@gmail.com; Olga Sorkine-Hornung, sorkine@inf.ethz.ch; Daniel Cohen-Or, cohenor@gmail.com; Baoquan Chen, baoquan@pku.edu.cn.

0730-0301/2020/7-ART62 $15.00
https://doi.org/10.1145/3386569.3392462

1 INTRODUCTION

Capturing the motion of humans is a fundamental task in motion analysis, computer animation, and human-computer interaction. Motion capture (MoCap) systems typically require the performer to wear a set of markers, whose positions are sampled by magnetic or optical sensors, yielding a temporal sequence of 3D skeleton poses. Since different MoCap setups involve different marker configurations and make use of different software, the captured skeletons may differ in their structure and number of joints, in addition to differences in bone lengths and proportions, corresponding to different captured individuals. Thus, *motion retargeting* is necessary, not only for transferring captured motion from one articulated character to another, within the same MoCap setup, but also across different setups. The latter is also essential for using data from multiple

The leopard never changes its spots: realistic pigmentation pattern formation by coupling tissue growth with reaction-diffusion

MARCELO DE GOMENSORO MALHEIROS, FURG, Brazil
HENRIQUE FENSTERSEIFER, UFRGS, Brazil
MARCELO WALTER, UFRGS, Brazil

Previous research in pattern formation using reaction-diffusion mostly focused on static domains, either for computational simplicity or mathematical tractability. In this work, we have explored the expressiveness of combining simple mechanisms as a possible explanation for pigmentation pattern formation, where tissue growth plays a crucial role. Our motivation is not only to realistically reproduce natural patterns but also to get insights into the underlying biological processes. Therefore, we present a novel approach to generate realistic animal skin patterns. First, we describe the approximation of tissue growth by a series of discrete matrix expansion operations. Then, we combine it with an adaptation of Turing's non-linear reaction-diffusion model, which enforces upper and lower bounds to the concentrations of the involved chemical reagents. We also propose the addition of a single-reagent continuous autocatalytic reaction, called reinforcement, to provide a mechanism to maintain an already established pattern during growth. By careful adjustment of the parameters and the sequencing of operations, we closely match the appearance of a few real species. In particular, we reproduce in detail the distinctive features of the leopard skin, also providing a hypothesis for the simultaneous productions of the most common melanin types, eumelanin and pheomelanin.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; **Modeling and simulation**.

Additional Key Words and Phrases: natural phenomena, texturing, pattern formation, reaction-diffusion, Turing model

ACM Reference Format:
Marcelo de Gomensoro Malheiros, Henrique Fensterseifer, and Marcelo Walter. 2020. The leopard never changes its spots: realistic pigmentation pattern formation by coupling tissue growth with reaction-diffusion. *ACM Trans. Graph.* 39, 4, Article 63 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392478

1 INTRODUCTION

Growth is an essential mechanism of life and a continuous process since the inception of all living beings. Although much research has been done on cellular biochemistry, the overall mechanisms responsible for growth have been only partially uncovered by now, and are already very complex [Wolpert et al. 2015].

Reaction-diffusion is a well-known model conceived by Turing [1952], being later used in Computer Graphics by the pioneering works of Turk [1991], Witkin and Kass [1991] and Fowler et al. [1992] for simulating natural phenomena. Although interest in reaction-diffusion faded along the years for the graphics community, it continued to play a major role in theoretical models from Mathematical Biology [Maini et al. 2012; Meinhardt 2009; Murray 2003].

We are interested in the general problem of reproducing the appearance of animal skin patterns. Instead of focusing on procedural methods [Hu et al. 2019], exemplar-based techniques [Raad et al. 2018] or indirectly through machine learning [Zhou et al. 2018], we show that a reduced subset of simulated biological mechanisms is enough to generate accurate textures.

Our main contribution is showing that the adequate approximation of tissue growth combined with reaction-diffusion is the key to the emergence of complex yet realistic patterns. By carefully adjusting model parameters and setting up distinct growth phases, we can match skin pigmentation of mammals and other animals. In particular, we have produced *in silico* the characteristic leopard rosettes, depicted in a 3D rendering in Figure 1.

Fig. 1. Synthetic leopard coat: pigmentation generated by our technique and then rendered in 3D by assigning individual fur colors. The skin below the fur layer has a uniform pink color and it is barely visible.

Authors' addresses: Marcelo de Gomensoro Malheiros, FURG, Computer Science Center, Rio Grande, Brazil, mgmalheiros@gmail.com; Henrique Fensterseifer, UFRGS, Institute of Informatics, Porto Alegre, Brazil, rique.hf@gmail.com; Marcelo Walter, UFRGS, Institute of Informatics, Porto Alegre, Brazil, marcelo.walter@inf.ufrgs.br.

0730-0301/2020/7-ART63 $15.00
https://doi.org/10.1145/3386569.3392478

Unpaired Motion Style Transfer from Video to Animation

KFIR ABERMAN*, AICFVE, Bejing Film Academy & Tel-Aviv University
YIJIA WENG*, CFCS, Peking University & AICFVE, Beijing Film Academy
DANI LISCHINSKI, The Hebrew University of Jerusalem & AICFVE, Beijing Film Academy
DANIEL COHEN-OR, Tel-Aviv University & AICFVE, Beijing Film Academy
BAOQUAN CHEN†, CFCS, Peking University & AICFVE, Beijing Film Academy

Transferring the motion style from one animation clip to another, while preserving the motion content of the latter, has been a long-standing problem in character animation. Most existing data-driven approaches are supervised and rely on paired data, where motions with the same content are performed in different styles. In addition, these approaches are limited to transfer of styles that were seen during training.

In this paper, we present a novel data-driven framework for motion style transfer, which learns from an unpaired collection of motions with style labels, and enables transferring motion styles not observed during training. Furthermore, our framework is able to extract motion styles directly from videos, bypassing 3D reconstruction, and apply them to the 3D input motion.

Our style transfer network encodes motions into two latent codes, for content and for style, each of which plays a different role in the decoding (synthesis) process. While the content code is decoded into the output motion by several temporal convolutional layers, the style code modifies deep features via temporally invariant adaptive instance normalization (AdaIN).

Moreover, while the content code is encoded from 3D joint rotations, we learn a common embedding for style from either 3D or 2D joint positions, enabling style extraction from videos.

Our results are comparable to the state-of-the-art, despite not requiring paired training data, and outperform other methods when transferring previously unseen styles. To our knowledge, we are the first to demonstrate style transfer directly from videos to 3D animations - an ability which enables one to extend the set of style examples far beyond motions captured by MoCap systems.

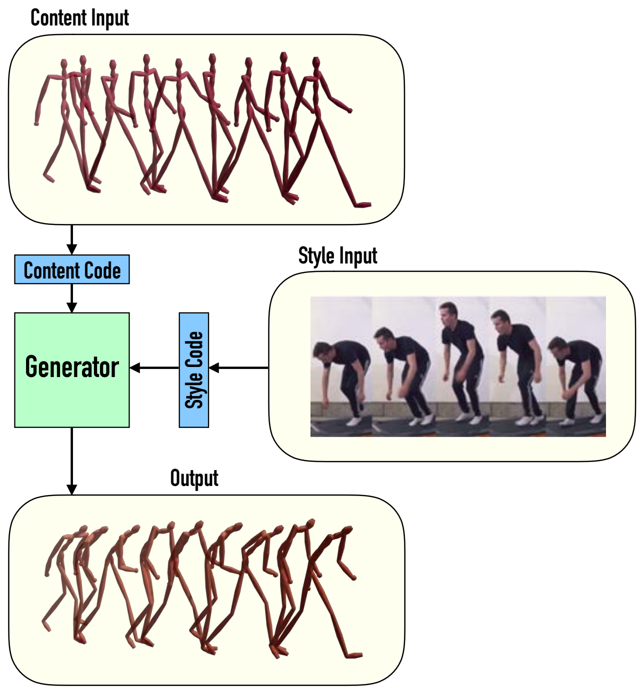


Fig. 1. Style transfer from video to animation. Our network, which is trained with unpaired motion sequences, learns to disentangle content and style. Our trained generator is able to produce a motion sequence that combines the content of a 3D sequence with the style extracted directly from a video.

CCS Concepts: • **Computing methodologies** → **Motion processing**; **Neural networks**.

Additional Key Words and Phrases: motion analysis, style transfer

ACM Reference Format:
Kfir Aberman, Yijia Weng, Dani Lischinski, Daniel Cohen-Or, and Baoquan Chen. 2020. Unpaired Motion Style Transfer from Video to Animation. *ACM Trans. Graph.* 39, 4, Article 64 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392469

1 INTRODUCTION

The style of human motion may be thought of as the collection of motion attributes that convey the mood and the personality of

*equal contribution
†corresponding author

Authors' addresses: Kfir Aberman, kfiraberman@gmail.com; Yijia Weng, halfsummer11@gmail.com; Dani Lischinski, danix3d@gmail.com; Daniel Cohen-Or, cohenor@gmail.com; Baoquan Chen, baoquan@pku.edu.cn.

0730-0301/2020/7-ART64 $15.00
https://doi.org/10.1145/3386569.3392469

Wave Curves: Simulating Lagrangian water waves on dynamically deforming surfaces

TOMAS SKRIVAN, IST Austria, Austria
ANDREAS SODERSTROM, no affiliation, Sweden
JOHN JOHANSSON, Weta Digital, New Zealand
CHRISTOPH SPRENGER, Weta Digital, New Zealand
KEN MUSETH, Weta Digital, New Zealand
CHRIS WOJTAN, IST Austria, Austria

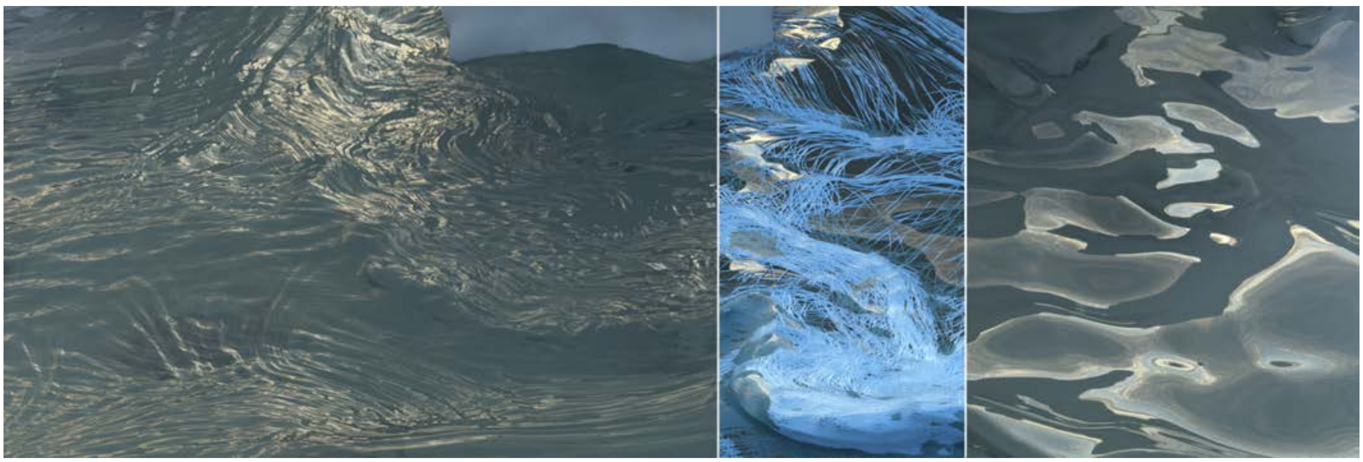


Fig. 1. We introduce an efficient method for adding detailed ripples (left) in the form of curve primitives (middle right) on top of an existing 3D fluid simulation (right). These wave curves evolve according to our extension of linear water wave theory, which naturally models effects like ripples aligned with flow features.

We propose a method to enhance the visual detail of a water surface simulation. Our method works as a post-processing step which takes a simulation as input and increases its apparent resolution by simulating many detailed Lagrangian water waves on top of it. We extend linear water wave theory to work in non-planar domains which deform over time, and we discretize the theory using Lagrangian wave packets attached to spline curves. The method is numerically stable and trivially parallelizable, and it produces high frequency ripples with dispersive wave-like behaviors customized to the underlying fluid simulation.

CCS Concepts: • **Computing methodologies → Physical simulation**; *Simulation by animation.*

Authors' addresses: Tomas Skrivan, IST Austria, Am Campus 1, Klosterneuburg, 3400, Austria, tomas.skrivan@ist.ac.at; Andreas Soderstrom, no affiliation, Sweden, andreas.soderstrom@gmail.com; John Johansson, Weta Digital, New Zealand, jjohansson@wetafx.co.nz; Christoph Sprenger, Weta Digital, New Zealand, christophspr@gmail.com; Ken Museth, Weta Digital, New Zealand, ken.museth@gmail.com; Chris Wojtan, IST Austria, Am Campus 1, Klosterneuburg, 3400, Austria, wojtan@ist.ac.at.

0730-0301/2019/11-ART65
https://doi.org/10.1145/3386569.3392466

Additional Key Words and Phrases: Water surface waves, wave animation, production animation, ripples, wakes, dispersion

ACM Reference Format:
Tomas Skrivan, Andreas Soderstrom, John Johansson, Christoph Sprenger, Ken Museth, and Chris Wojtan. 2019. Wave Curves: Simulating Lagrangian water waves on dynamically deforming surfaces. *ACM Trans. Graph.* 38, 6, Article 65 (November 2019), 12 pages. https://doi.org/10.1145/3386569.3392466

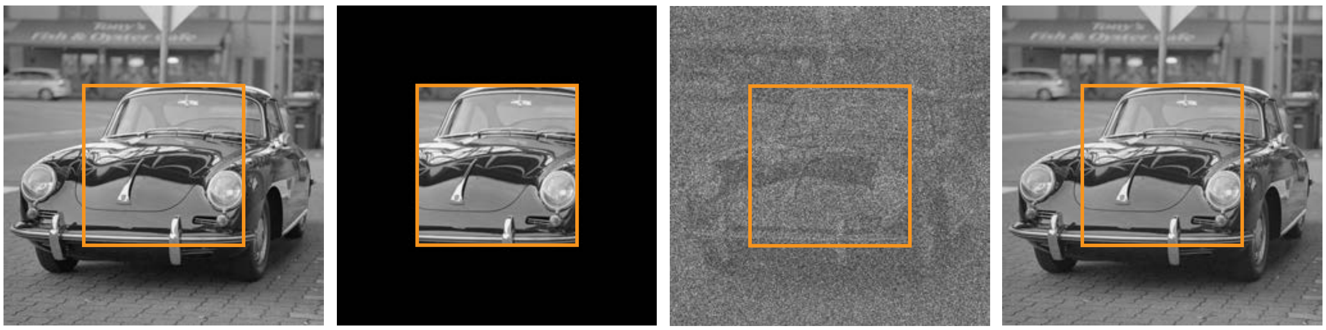
1 INTRODUCTION

Although three-dimensional fluid simulation has led to spectacularly detailed visual effects in the past decade, the expense of three-dimensional fluid simulation strongly limits the amount of detail that can be simulated at the water surface. Several researchers have circumvented this limitation by simulating additional 2D waves directly on top of the 3D water surface. However, these previous approaches have a few significant limitations, like a resolution limit imposed by an underlying grid, mesh, or particle system [Angst et al. 2008; Kim et al. 2013; Mercier et al. 2015; Thürey et al. 2010; Yang et al. 2016; Yu et al. 2012], or a restriction to perfectly flat domains [Canabal et al. 2016; Jeschke et al. 2018; Jeschke and Wojtan 2015, 2017; Yuksel et al. 2007].

In this work, we aim to significantly increase the visible detail on a simulated fluid surface using Lagrangian wave packets, which

High Resolution Étendue Expansion for Holographic Displays

GRACE KUO, Facebook Reality Labs and University of California, Berkeley
LAURA WALLER, University of California, Berkeley
REN NG, University of California, Berkeley
ANDREW MAIMONE, Facebook Reality Labs



Target image — Traditional hologram (simulation) — Expanded étendue One-step backpropagation (simulation) — Expanded étendue Our algorithm (simulation)

Fig. 1. Traditional holographic displays have limited étendue resulting in a tradeoff between field-of-view (FoV) and eyebox size. If the eyebox is held constant, unique imagery cannot be displayed outside of the native FoV (orange box). The addition of a thin scattering mask into the system increases the diffraction angles, and thus the FoV, without sacrificing the eyebox. The scattering mask is taken into account during computation of the hologram through an iterative algorithm that outperforms the one-step backpropagation approach used in prior work. Car source image by Bill Newton (CC BY 2.0).

Holographic displays can create high quality 3D images while maintaining a small form factor suitable for head-mounted virtual and augmented reality systems. However, holographic displays have limited étendue based on the number of pixels in their spatial light modulators, creating a tradeoff between the eyebox size and the field-of-view. Scattering-based étendue expansion, in which coherent light is focused into an image after being scattered by a static mask, is a promising avenue to break this tradeoff. However, to date, this approach has been limited to very sparse content consisting of, for example, only tens of spots.

In this work, we introduce new algorithms to scattering-based étendue expansion that support dense, photorealistic imagery at the native resolution of the spatial light modulator, offering up to a 20 dB improvement in peak signal to noise ratio over baseline methods. We propose spatial and frequency constraints to optimize performance for human perception, and performance is characterized both through simulation and a preliminary benchtop prototype. We further demonstrate the ability to generate content at multiple depths, and we provide a path for the miniaturization of our benchtop prototype into a sunglasses-like form factor.

Authors' addresses: Grace Kuo, Facebook Reality Labs, University of California, Berkeley, gkuo@berkeley.edu; Laura Waller, University of California, Berkeley, waller@berkeley.edu; Ren Ng, University of California, Berkeley, ren@berkeley.edu; Andrew Maimone, Facebook Reality Labs, andrew.maimone@fb.com.

0730-0301/2020/7-ART66
https://doi.org/10.1145/3386569.3392414

CCS Concepts: • **Hardware → Displays and imagers**.

Additional Key Words and Phrases: computer generated holography, computational displays, augmented reality, near-eye displays

ACM Reference Format:
Grace Kuo, Laura Waller, Ren Ng, and Andrew Maimone. 2020. High Resolution Étendue Expansion for Holographic Displays. *ACM Trans. Graph.* 39, 4, Article 66 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392414

1 INTRODUCTION

Computer generated holography allows generation of an arbitrary light distribution from a flat, programmable spatial light modulator (SLM) by controlling the wavefront of a coherent beam of light. This technique is particularly promising for near-eye displays since it enables per-pixel focus control, computational correction of optical aberrations, and simple optical components suitable for miniaturization [Maimone et al. 2017].

However, current holographic displays suffer from a unique challenge: a tradeoff between field-of-view (FoV) and the size of the viewing eyebox, the area in which the eye must be located to see the image. Together, these two quantities describe the *étendue* of the display, a quantity which measures the product of the area and solid angle of emitted light from a surface in an optical system. In conventional, non-holographic displays (e.g., the Oculus Rift S) obtaining large étendue is generally not a challenge and can be provided, for example, by a display panel backlight that has large area and range of emission angles. However, in a holographic display the étendue

Holographic Optics for Thin and Lightweight Virtual Reality

ANDREW MAIMONE, Facebook Reality Labs
JUNREN WANG, Facebook Reality Labs

Fig. 1. *Left:* Photo of full color holographic display in benchtop form factor. *Center:* Prototype VR display in sunglasses-like form factor with display thickness of 8.9 mm. Driving electronics and light sources are external. *Right:* Photo of content displayed on prototype in center image. Car scenes by komba/Shutterstock.

We present a class of display designs combining holographic optics, directional backlighting, laser illumination, and polarization-based optical folding to achieve thin, lightweight, and high performance near-eye displays for virtual reality. Several design alternatives are proposed, compared, and experimentally validated as prototypes. Using only thin, flat films as optical components, we demonstrate VR displays with thicknesses of less than 9 mm, fields of view of over 90° horizontally, and form factors approaching sunglasses. In a benchtop form factor, we also demonstrate a full color display using wavelength-multiplexed holographic lenses that uses laser illumination to provide a large gamut and highly saturated color. We show experimentally that our designs support resolutions expected of modern VR headsets and can scale to human visual acuity limits. Current limitations are identified, and we discuss challenges to obtain full practicality.

CCS Concepts: • **Computing methodologies** → **Virtual reality**.

Additional Key Words and Phrases: near-eye display, holography

ACM Reference Format:
Andrew Maimone and Junren Wang. 2020. Holographic Optics for Thin and Lightweight Virtual Reality. *ACM Trans. Graph.* 39, 4, Article 67 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392416

1 INTRODUCTION

As virtual reality (VR) becomes more ubiquitous, we expect that it will expand beyond entertainment to see broader use in productivity and social interactivity, and these fields will drive VR displays towards more comfortable form factors, higher performance, and improved aesthetics. For example, a VR display used as an immersive computing platform for work would be expected to be used many hours at a time, necessitating a comfortable and lightweight headset. Such a display would also be expected to meet or exceed the performance of conventional displays, and reproduce, for example, small text near the limit of human visual acuity. This use case also brings VR out of the home and in to work and public spaces where socially acceptable sunglasses and eyeglasses form factors prevail.

VR has made good progress in the past few years, and entirely self-contained head-worn systems are now commercially available. However, current headsets still have box-like form factors and provide only a fraction of the resolution of the human eye. Emerging optical design techniques, such as polarization-based optical folding, or "pancake" optics, promise to improve performance while reducing size. However, current implementations rely on curved optics of solid glass or plastic, which has limited designs to goggles-like form factors. In contrast, holographic optical elements can provide arbitrary deflection of light from a flat surface of negligible thickness. However, such elements are difficult to work with due to the need for coherent light sources, wavelength and angle sensitivities, speckle artifacts, and the difficulty of making a full color display.

In this work, we propose combining polarization-based optical folding and holographic optics to gain the performance benefits of both while systematically working through the unique challenges of holography. In particular, we augment these technologies with laser illumination, directional backlighting, and color-multiplexing to achieve the field of view (FOV) and resolution expected of modern VR headsets while reducing thicknesses to ≤ 10 mm to enable sunglasses-like form factors. We demonstrate that our designs scale in resolution to the limits to normal human vision and can exceed the color performance of conventional displays. We propose several design alternatives that are verified in a series of hardware prototypes and discuss challenges to make them fully practical.

1.1 Contributions

We propose designs for near-eye displays that are evaluated across a series of hardware prototypes. Specific contributions include:

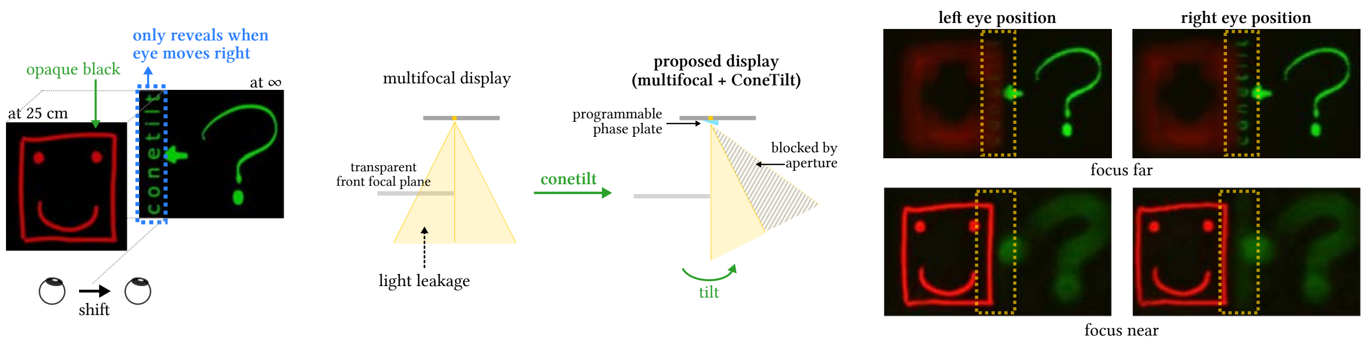
(1) We propose a class of near-eye displays combining laser illumination, directional backlighting, color-multiplexed holographic optics, and polarization-based optical folding that is thinner than previously reported VR displays while achieving the resolution and FOV of a modern VR headset.

Authors' addresses: Andrew Maimone, andrew.maimone@fb.com;
Junren Wang, junren@fb.com, Facebook Reality Labs, Redmond, WA, USA.

0730-0301/2020/7-ART67
https://doi.org/10.1145/3386569.3392416

Towards Occlusion-Aware Multifocal Displays

JEN-HAO RICK CHANG, Carnegie Mellon University, USA
ANAT LEVIN, Technion, Israel
B. V. K. VIJAYA KUMAR, Carnegie Mellon University, USA
ASWIN C. SANKARANARAYANAN, Carnegie Mellon University, USA



(a) Scene (b) Standard multifocal display (c) Proposed ConeTilt display (d) Captured results of the same rendered scene

Fig. 1. Multifocal displays present content at multiple focal planes to satisfy the accommodation cue of the human vision. (a) A scene with content on two focal planes. The text "conetilt" is expected to be occluded when the eye shifts laterally. (b) However, in a standard multifocal display, content on focal planes do not occlude each other and hence content from a farther plane can leak into the area of occluding objects at frontal planes. (c) We propose a ConeTilt multifocal display which provides the ability to tilt the cone of light emerging from each pixel and hence produce the same effect as occlusion. (d) The images captured by our prototype for the scene in (a) under ±0.5 mm lateral translation of a camera. Note how the text next to the boundary is hidden on one viewpoint and revealed on the other, and how defocus cues are faithfully reproduced. This result was produced without gaze tracking or content re-rendering.

The human visual system uses numerous cues for depth perception, including disparity, accommodation, motion parallax and occlusion. It is incumbent upon virtual-reality displays to satisfy these cues to provide an immersive user experience. Multifocal displays, one of the classic approaches to satisfy the accommodation cue, place virtual content at multiple focal planes, each at a different depth. However, the content on focal planes close to the eye do not occlude those farther away; this deteriorates the occlusion cue as well as reduces contrast at depth discontinuities due to leakage of the defocus blur. This paper enables occlusion-aware multifocal displays using a novel ConeTilt operator that provides an additional degree of freedom — tilting the light cone emitted at each pixel of the display panel. We show that, for scenes with relatively simple occlusion configurations, tilting the light cones provides the same effect as physical occlusion. We demonstrate that ConeTilt can be easily implemented by a phase-only spatial light modulator. Using a lab prototype, we show results that demonstrate the presence of occlusion cues and the increased contrast of the display at depth edges.

Authors' addresses: Jen-Hao Rick Chang, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, 15213, USA, rickchang@cmu.edu; Anat Levin, Technion, Haifa, Israel, anat.levin@ee.technion.ac.il; B. V. K. Vijaya Kumar, Carnegie Mellon University, Pittsburgh, USA, kumar@ece.cmu.edu; Aswin C. Sankaranarayanan, Carnegie Mellon University, Pittsburgh, USA, saswin@andrew.cmu.edu.

0730-0301/2020/7-ART68 $15.00
https://doi.org/10.1145/3386569.3392424

CCS Concepts: • **Computing methodologies → Virtual reality.**

Additional Key Words and Phrases: multifocal displays, occlusion, phase modulation, phase spatial light modulator

ACM Reference Format:
Jen-Hao Rick Chang, Anat Levin, B. V. K. Vijaya Kumar, and Aswin C. Sankaranarayanan. 2020. Towards Occlusion-Aware Multifocal Displays. *ACM Trans. Graph.* 39, 4, Article 68 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392424

1 INTRODUCTION

The primary aim of a virtual-reality (VR) display is to present a scene to the eye that is indistinguishable from reality. In the context of depth perception, a VR display has to faithfully reproduce the visual cues pertaining to disparity, accommodation, occlusion, and motion parallax. While there are many types of VR displays, each with differing amounts of fidelity towards satisfying these cues, this paper focuses on multifocal displays with the objective of enhancing the range of perceptual cues that they can satisfy.

In a multifocal display, three-dimensional (3D) content is shown to a user by placing virtual objects on different focal planes, which are optically placed at different depths from the viewer. This has a unique advantage that the display automatically renders the accommodation cues, i.e., supports the focus of our eyes, provided there are a sufficient number of focal planes [Chang et al. 2018; MacKenzie et al. 2010; Rolland et al. 1999; Watt et al. 2012]. In order to display multiple focal planes at different depths, the focal planes are time-multiplexed and content on the planes does not occlude

Attribute2Font: Creating Fonts You Want From Attributes

YIZHI WANG*, Wangxuan Institute of Computer Technology, Peking University, China
YUE GAO*, Wangxuan Institute of Computer Technology, Peking University, China
ZHOUHUI LIAN†, Wangxuan Institute of Computer Technology, Peking University, China

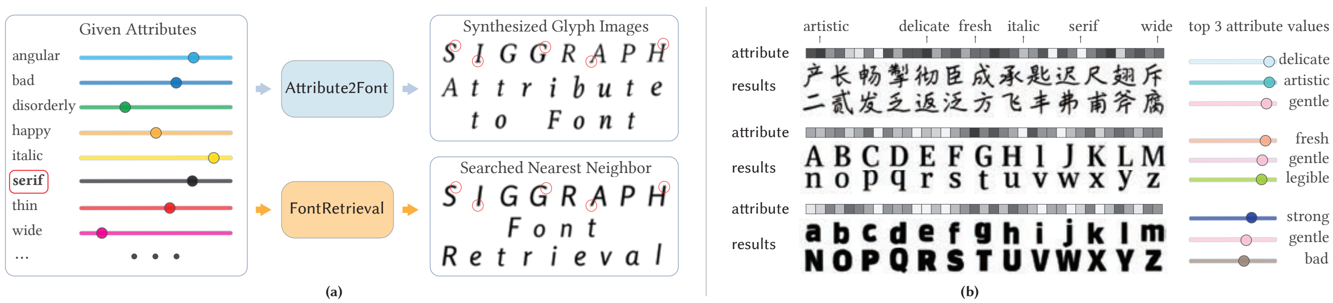


Fig. 1. (a) An overview of our model, given an arbitrary set of predefined font attributes' values, glyph images in the corresponding style can be synthesized. By contrast, the font retrieval methods can only select existing fonts for users, which often cannot satisfy the specific requirements of some users. (b) Examples of synthesized English and Chinese glyph images in different attribute sets. We pre-define 37 different kinds of attributes whose values are shown in the grayscale grids (darker is higher). Please zoom in for better inspection.

Font design is now still considered as an exclusive privilege of professional designers, whose creativity is not possessed by existing software systems. Nevertheless, we also notice that most commercial font products are in fact manually designed by following specific requirements on some attributes of glyphs, such as italic, serif, cursive, width, angularity, etc. Inspired by this fact, we propose a novel model, Attribute2Font, to automatically create fonts by synthesizing visually pleasing glyph images according to user-specified attributes and their corresponding values. To the best of our knowledge, our model is the first one in the literature which is capable of generating glyph images in new font styles, instead of retrieving existing fonts, according to given values of specified font attributes. Specifically, Attribute2Font is trained to perform font style transfer between any two fonts conditioned on their attribute values. After training, our model can generate glyph images in accordance with an arbitrary set of font attribute values. Furthermore, a novel unit named Attribute Attention Module is designed to make those generated glyph images better embody the prominent font attributes. Considering that the annotations of font attribute values are extremely expensive to obtain, a semi-supervised learning scheme is also introduced to exploit a large number of unlabeled fonts. Experimental results demonstrate that our model achieves impressive performance on many tasks, such as creating glyph images in new font styles, editing existing fonts, interpolation among different fonts, etc.

CCS Concepts: • **Computing methodologies** → **Computer vision**; *Image manipulation.*

Additional Key Words and Phrases: Image Synthesis, Font Design, Type Design, Deep Generative Models, Style Transfer

ACM Reference Format:
Yizhi Wang, Yue Gao, and Zhouhui Lian. 2020. Attribute2Font: Creating Fonts You Want From Attributes. *ACM Trans. Graph.* 39, 4, Article 69 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392456

1 INTRODUCTION

Traditional font design workflow sets a high barrier for common users, which requires creativity and expertise in this field. Automatic font designing remains a challenging and ongoing problem in areas of Computer Graphics (CG), Computer Vision (CV) and Artificial Intelligence (AI). In this paper, we aim to handle the task of generating glyph images according to the user-specified font attributes (such as italic, serif, cursive, angularity, etc.) and their values. Our model significantly lowers the barrier and provides various and customizable fonts for common users. We believe that our model can also inspire professional designers and assist them in creating new fonts.

To the best of our knowledge, our work is the first to automatically generate glyph images from the attributes of font styles. Existing works on attribute-controllable image synthesis are unsuitable for this task due to the following three reasons: (1) Existing works aim to generate images such as faces and fashion, whose appearances (such as color and texture) vary with the attributes but shapes generally remain unchanged. By contrast, the shapes of glyphs vary

*Denotes equal contribution
†Corresponding author

Authors' addresses: Yizhi Wang, Wangxuan Institute of Computer Technology, Peking University, China, wangyizhi@pku.edu.cn; Yue Gao, Wangxuan Institute of Computer Technology, Peking University, China, gerry@pku.edu.cn; Zhouhui Lian, Wangxuan Institute of Computer Technology, Peking University, China, lianzhouhui@pku.edu.cn.

0730-0301/2020/7-ART69 $15.00
https://doi.org/10.1145/3386569.3392456

Computational Image Marking on Metals via Laser Induced Heating

SEBASTIAN CUCERCA, Max Planck Institute for Informatics, Germany
PIOTR DIDYK, Università della Svizzera italiana, Switzerland
HANS-PETER SEIDEL, Max Planck Institute for Informatics, Germany
VAHID BABAEI, Max Planck Institute for Informatics, Germany

Fig. 1. Laser marked images on stainless steel using our proposed method (the plates are 13 × 13 cm). Images © Piotr Didyk, Alexandr Ivanov, Sebastian Cucerca.

Laser irradiation induces colors on some industrially important materials, such as stainless steel and titanium. It is however challenging to find marking configurations that create colorful, high-resolution images. The brute-force solution to the *gamut exploration* problem does not scale with the high-dimensional design space of laser marking. Moreover, there exists no color reproduction workflow capable of reproducing color images with laser marking. Here, we propose a measurement-based, data-driven performance space exploration of the color laser marking process. We formulate this exploration as a search for the Pareto optimal solutions to a multi-objective optimization and solve it using an evolutionary algorithm. The explored set of diverse colors is then utilized to mark high-quality, full-color images.

Authors' addresses: Sebastian Cucerca, Max Planck Institute for Informatics, Saarbrücken, Germany, scucerca@mpi-inf.mpg.de; Piotr Didyk, Università della Svizzera italiana, Lugano, Switzerland, piotr.didyk@usi.ch; Hans-Peter Seidel, Max Planck Institute for Informatics, Saarbrücken, Germany, hpseidel@mpi-sb.mpg.de; Vahid Babaei, Max Planck Institute for Informatics, Saarbrücken, Germany, vbabaei@mpi-inf.mpg.de.

0730-0301/2020/7-ART70 $15.00
https://doi.org/10.1145/3386569.3392423

CCS Concepts: • **Applied computing → Computer-aided manufacturing**; • **Computing methodologies → Computer graphics**.

Additional Key Words and Phrases: laser marking, color reproduction, multi-objective optimization, genetic algorithm, computational fabrication

ACM Reference Format:
Sebastian Cucerca, Piotr Didyk, Hans-Peter Seidel, and Vahid Babaei. 2020. Computational Image Marking on Metals via Laser Induced Heating. *ACM Trans. Graph.* 39, 4, Article 70 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392423

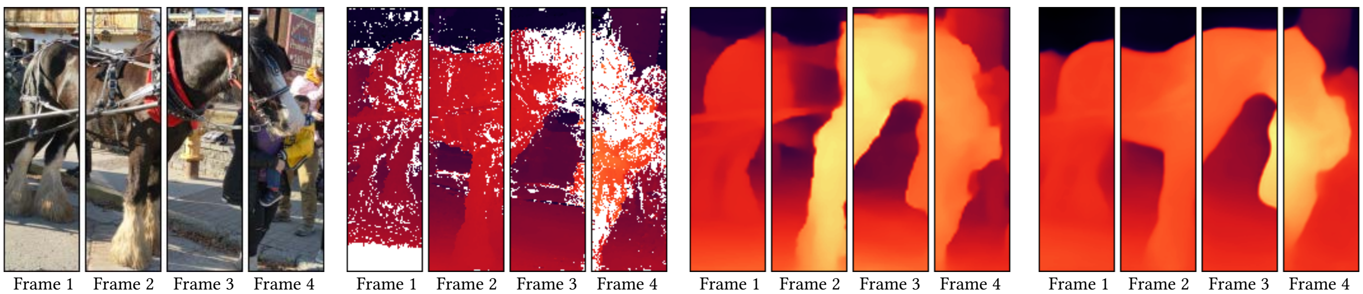
1 INTRODUCTION

Creating visible patterns on surfaces using laser irradiation is a rapidly growing technology with many applications in object identification, customization, and authentication [Liu et al. 2019]. *Laser marking* is an environmentally friendly, low maintenance process with no consumables, dyes, or pigments. While mostly a monochromatic method, some materials exhibit a range of colors when treated with laser, as a result of complex physicochemical phenomena. Among such materials are stainless steel and titanium, some of the most industrially important metals.

Despite the great potential, the industrial adoption of color laser marking is almost non-existent due to its challenging *characterization*. In the absence of such a characterization, the relationship

Consistent Video Depth Estimation

XUAN LUO[*], University of Washington
JIA-BIN HUANG, Virginia Tech
RICHARD SZELISKI, Facebook
KEVIN MATZEN, Facebook
JOHANNES KOPF, Facebook



(a) Input video (b) COLMAP depth (c) Mannequin Challenge depth (d) Our result

Fig. 1. We present a system for estimating temporally coherent and geometrically consistent depth from a casually captured video. Conventional multi-view stereo methods such as COLMAP [Schonberger and Frahm 2016] often produce incomplete depth on moving objects or poorly textured areas. Learning-based methods (e.g., [Li et al. 2019]) predict dense depth for each frame but the video reconstruction is flickering and geometrically inconsistent. Our video depth estimation is fully dense, globally scale-consistent, and capable of handling dynamically moving objects. We evaluate our method on a wide variety of challenging videos and show that our results enable new video special effects.

We present an algorithm for reconstructing dense, geometrically consistent depth for all pixels in a monocular video. We leverage a conventional structure-from-motion reconstruction to establish geometric constraints on pixels in the video. Unlike the ad-hoc priors in classical reconstruction, we use a learning-based prior, i.e., a convolutional neural network trained for single-image depth estimation. At test time, we fine-tune this network to satisfy the geometric constraints of a particular input video, while retaining its ability to synthesize plausible depth details in parts of the video that are less constrained. We show through quantitative validation that our method achieves higher accuracy and a higher degree of geometric consistency than previous monocular reconstruction methods. Visually, our results appear more stable. Our algorithm is able to handle challenging hand-held captured input videos with a moderate degree of dynamic motion. The improved quality of the reconstruction enables several applications, such as scene reconstruction and advanced video-based visual effects.

CCS Concepts: • **Computing methodologies → Reconstruction**; *Computational photography*.

Additional Key Words and Phrases: video, depth estimation

[*]This work was done while Xuan was an intern at Facebook.

Authors' addresses: Xuan Luo, University of Washington; Jia-Bin Huang, Virginia Tech; Richard Szeliski, Facebook; Kevin Matzen, Facebook; Johannes Kopf, Facebook.

0730-0301/2020/7-ART71 $15.00
https://doi.org/10.1145/3386569.3392377

ACM Reference Format:
Xuan Luo, Jia-Bin Huang, Richard Szeliski, Kevin Matzen, and Johannes Kopf. 2020. Consistent Video Depth Estimation. *ACM Trans. Graph.* 39, 4, Article 71 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392377

1 INTRODUCTION

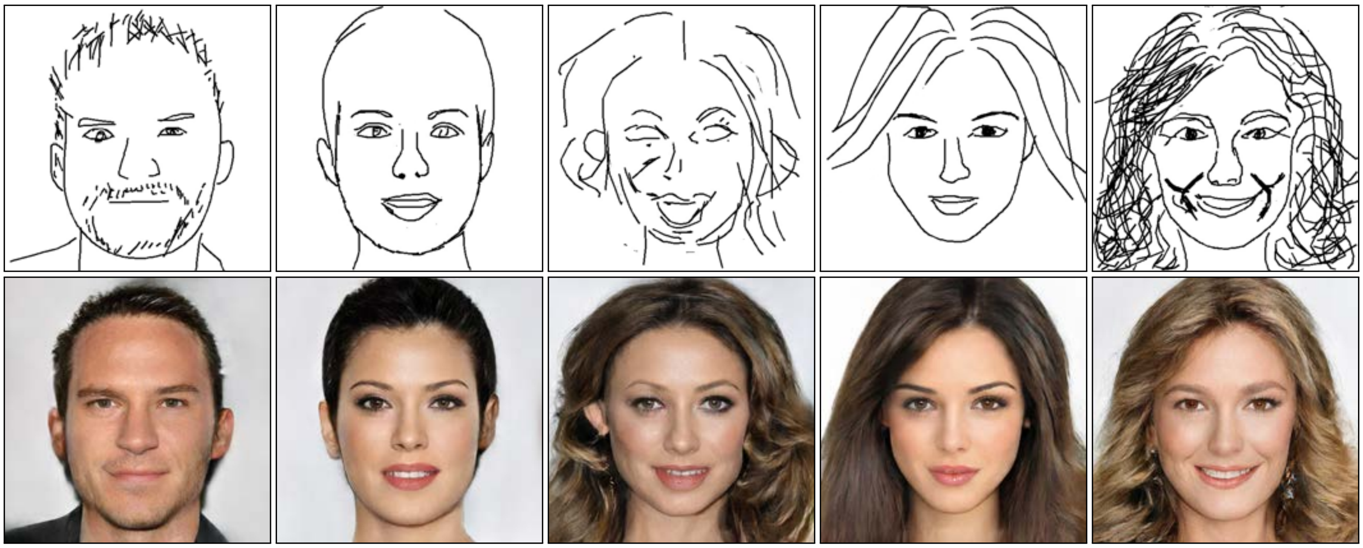
3D scene reconstruction from image sequences has been studied in our community for decades. Until a few years ago, the structure from motion systems for solving this problem were not very robust, and practically only worked "in the lab", with highly calibrated and predictable setups. They also, often, produced only sparse reconstructions, i.e., resolving depth at only a few isolated tracked point features. But in the last decade or so, we have seen good progress towards enabling more *casual* capture and producing *denser* reconstructions, driven by high-quality open-source reconstruction systems and recent advances in learning-based techniques, as discussed in the next section.

Arguably the *easiest* way to capture for 3D reconstruction is using hand-held cell phone video, since these cameras are so readily and widely available, and enable truly spontaneous, impromptu capture, as well as quickly covering large spaces. If we could achieve fully dense and accurate reconstruction from such input it would be immensely useful—however, this turns out to be quite difficult.

Besides the typical problems that any reconstruction system has to deal with, such as poorly textured areas, repetitive patterns, and occlusions, there are several additional challenges with video: higher noise level, shake and motion blur, rolling shutter deformations, small baseline between adjacent frames, and, often, the presence of dynamic objects, such as people. For these reasons, existing methods

DeepFaceDrawing: Deep Generation of Face Images from Sketches

SHU-YU CHEN*, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences
WANCHAO SU*, School of Creative Media, City University of Hong Kong
LIN GAO†, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences
SHIHONG XIA, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences
HONGBO FU, School of Creative Media, City University of Hong Kong



(a) (b) (c) (d) (e)

Fig. 1. Our DeepFaceDrawing system allows users with little training in drawing to produce high-quality face images (Bottom) from rough or even incomplete freehand sketches (Top). Note that our method faithfully respects user intentions in input strokes, which serve more like soft constraints to guide image synthesis.

Recent deep image-to-image translation techniques allow fast generation of face images from freehand sketches. However, existing solutions tend to overfit to sketches, thus requiring professional sketches or even edge maps as input. To address this issue, our key idea is to implicitly model the shape space of plausible face images and synthesize a face image in this space to approximate an input sketch. We take a local-to-global approach. We first learn feature embeddings of key face components, and push corresponding parts of input sketches towards underlying component manifolds defined by the feature vectors of face component samples. We also propose another deep neural network to learn the mapping from the embedded component features to realistic images with multi-channel feature maps as intermediate results to improve the information flow. Our method essentially uses input sketches as soft constraints and is thus able to produce high-quality face images even from rough and/or incomplete sketches. Our tool is easy to use even for non-artists, while still supporting fine-grained control of shape details. Both qualitative and quantitative evaluations show the superior generation ability of our system to existing and alternative solutions. The usability and expressiveness of our system are confirmed by a user study.

CCS Concepts: • **Human-centered computing** → **Graphical user interfaces**; • **Computing methodologies** → *Perception*; *Texturing*; **Image processing**.

Additional Key Words and Phrases: image-to-image translation, feature embedding, sketch-based generation, face synthesis

ACM Reference Format:
Shu-Yu Chen, Wanchao Su, Lin Gao, Shihong Xia, and Hongbo Fu. 2020. DeepFaceDrawing: Deep Generation of Face Images from Sketches. *ACM Trans. Graph.* 39, 4, Article 72 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392386

*Authors contributed equally.
†Corresponding author.

Authors' addresses: Shu-Yu Chen, Lin Gao and Shihong Xia are with the Beijing Key Laboratory of Mobile Computing and Pervasive Device, Institute of Computing Technology, Chinese Academy of Sciences. Wanchao Su and Hongbo Fu are with the School of Creative Media, City University of Hong Kong. Authors' e-mails: chenshuyu@ict.ac.cn, wanchao.su@my.cityu.edu.hk, gaolin@ict.ac.cn, xsh@ict.ac.cn, hongbofu@cityu.edu.hk.

0730-0301/2020/7-ART72 $15.00
https://doi.org/10.1145/3386569.3392386

Interactive Video Stylization Using Few-Shot Patch-Based Training

ONDŘEJ TEXLER, DAVID FUTSCHIK, MICHAL KUČERA, ONDŘEJ JAMRIŠKA, and ŠÁRKA SOCHOROVÁ, Czech Technical University in Prague, Faculty of Electrical Engineering
MENGLEI CHAI and SERGEY TULYAKOV, Snap Inc.
DANIEL SÝKORA, Czech Technical University in Prague, Faculty of Electrical Engineering

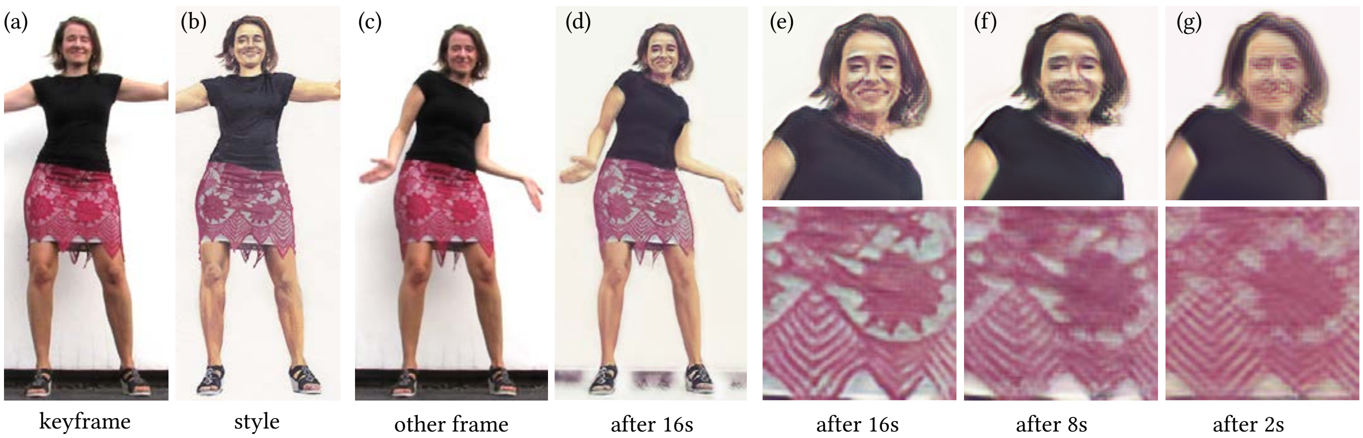


Fig. 1. An example of a sequence stylized using our approach. One frame from the original sequence is selected as a keyframe (a) and an artist stylizes it with acrylic paint (b). We use this single style exemplar as the only data to train a network. After 16 seconds of training, the network can stylize the entire sequence in real-time (c-d) while maintaining the state-of-the-art visual quality and temporal coherence. See the zoom-in views (e-g); even after 2 seconds of training, important structures already start to show up. Video frames (a, c) and style exemplar (b) courtesy of © Zuzana Studená.

In this paper, we present a learning-based method to the keyframe-based video stylization that allows an artist to propagate the style from a few selected keyframes to the rest of the sequence. Its key advantage is that the resulting stylization is semantically meaningful, i.e., specific parts of moving objects are stylized according to the artist's intention. In contrast to previous style transfer techniques, our approach does not require any lengthy pre-training process nor a large training dataset. We demonstrate how to train an appearance translation network from scratch using only a few stylized exemplars while implicitly preserving temporal consistency. This leads to a video stylization framework that supports real-time inference, parallel processing, and random access to an arbitrary output frame. It can also merge the content from multiple keyframes without the need to perform an explicit blending operation. We demonstrate its practical utility in various interactive scenarios, where the user paints over a selected keyframe and sees her style transferred to an existing recorded sequence or a live video stream.

CCS Concepts: • **Computing methodologies → Motion processing**; **Image processing**.

Additional Key Words and Phrases: example-based, apperance translation, style transfer

ACM Reference Format:
Ondřej Texler, David Futschik, Michal Kučera, Ondřej Jamriška, Šárka Sochorová, Menglei Chai, Sergey Tulyakov, and Daniel Sýkora. 2020. Interactive Video Stylization Using Few-Shot Patch-Based Training. *ACM Trans. Graph.* 39, 4, Article 73 (July 2020), 11 pages. https://doi.org/10.1145/3386569.3392453

Authors' addresses: Ondřej Texler, texleond@fel.cvut.cz; David Futschik, futscdav@fel.cvut.cz; Michal Kučera, kucerm22@fel.cvut.cz; Ondřej Jamriška, jamriond@fel.cvut.cz; Šárka Sochorová, sochosar@fel.cvut.cz, Czech Technical University in Prague, Faculty of Electrical Engineering, Karlovo náměstí 13, Praha 2, Czech Republic, 121 35; Menglei Chai, mchai@snap.com; Sergey Tulyakov, stulyakov@snap.com, Snap Inc., 2772 Donald Douglas Loop N, Santa Monica, United States, CA 90405; Daniel Sýkora, sykorad@fel.cvut.cz, Czech Technical University in Prague, Faculty of Electrical Engineering, Karlovo náměstí 13, Praha 2, Czech Republic, 121 35.

0730-0301/2020/7-ART73 $15.00
https://doi.org/10.1145/3386569.3392453

1 INTRODUCTION

Example-based stylization of videos became recently popular thanks to significant advances made in neural techniques [Kotovenko et al. 2019; Ruder et al. 2018; Sanakoyeu et al. 2018]. Those extend the seminal approach of Gatys et al. [2016] into the video domain and improve the quality by adding specific style-aware content losses. Although these techniques can deliver impressive stylization results on various exemplars, they still suffer from the key limitation of being difficult to control. This is due to the fact that they only measure statistical correlations and thus do not guarantee that specific parts of the video will be stylized according to the artist's intention, which is an essential requirement for use in a real production pipeline.

This important aspect is addressed by a concurrent approach—the keyframe-based video stylization [Bénard et al. 2013; Jamriška et al. 2019]. Those techniques employ guided patch-based synthesis [Fišer

Interferometric Transmission Probing with Coded Mutual Intensity

ALANKAR KOTWAL, Carnegie Mellon University
ANAT LEVIN, Technion
IOANNIS GKIOULEKAS, Carnegie Mellon University

We introduce a new interferometric imaging methodology that we term interferometry with coded mutual intensity, which allows selectively imaging photon paths based on attributes such as their length and endpoints. At the core of our methodology is a new technical result that shows that manipulating the spatial coherence properties of the light source used in an interferometric system is equivalent, through a Fourier transform, to implementing light path probing patterns. These patterns can be applied to either the coherent transmission matrix, or the incoherent light transport matrix describing the propagation of light in a scene. We test our theory by building a prototype inspired by the Michelson interferometer, extended to allow for programmable phase and amplitude modulation of the illumination injected in the interferometer. We use our prototype to perform experiments such as visualizing complex fields, capturing direct and global transport components, acquiring light transport matrices, and performing anisotropic descattering, both in steady-state imaging and, by combining our technique with optical coherence tomography, in transient imaging.

CCS Concepts: • **Computing methodologies → Computational photography**.

Additional Key Words and Phrases: spatial coherence, mutual intensity, interferometry, light transport matrix, transmission matrix

ACM Reference Format:
Alankar Kotwal, Anat Levin, and Ioannis Gkioulekas. 2020. Interferometric Transmission Probing with Coded Mutual Intensity. *ACM Trans. Graph.* 39, 4, Article 74 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392384

1 INTRODUCTION

Light propagation is an inherently multi-path phenomenon: When we look at our surroundings, we observe light that has interacted with one or multiple objects, either by reflecting on their surfaces, or by scattering in their interior. Imaging systems typically accumulate contributions from photons traveling along all of these paths, indiscriminately of characteristics such as the paths' origins and lengths. This accumulation process confounds the information that is available in imaging measurements about scene properties of interest, such as the shape and material of objects of interest.

Computational light transport techniques attempt to overcome this confounding effect, by measuring only light that has traveled along specific subsets of all the possible paths in a scene. These subsets can be specified based on characteristics such as the endpoints and length of the paths, or combinations thereof. There is a growing number of systems that can perform this type of *selective probing* of the many light components that make up an image, including systems based on projector-camera combinations, continuous-wave amplitude-modulated sensors, streak cameras, single-photon avalance diodes, as well as interferometry.

Our focus is on the latter type of systems: Interferometric systems operate by simultaneously measuring two light waves that originated at the same light source, and have traveled along different paths in an optical system. When superimposed on an optical sensor, the two waves will produce some measurable *interference*. Depending on the wave-optics properties of the original illumination, only light paths that satisfy certain characteristics will contribute to the interference. Therefore, by controlling the coherence properties of the source the waves originate from, and then measuring their interference after propagation, it is possible to isolate contributions from only specific light paths. This type of interferometric probing has been demonstrated in the past [Gkioulekas et al. 2015], using a system similar to optical coherence tomography [Huang et al. 1991]. However, these previously systems are severely limited in terms of the types of probing they can perform, compared to probing capabilities possible using other imaging technologies.

Our goal on this paper is to significantly expand the probing capabilities that can be realized using interferometry. To this end, we develop a new imaging technique that we term *interferometry with coded mutual intensity*. Our technique is based on a setup similar to the classical Michelson interferometer, augmented with optical components for amplitude and phase modulation. These components enable programmatic control of the spatial coherence properties of the illumination injected in the setup.

We perform a detailed theoretical analysis of our technique, and explan how the underlying wave-optics models relate to the incoherent models of light propagation typically used in computer graphics. Through this analysis, we show that our technique provides several probing capabilities: First, it enables probing the coherent *transmission matrix* of a scene, using arbitrary convolutional probing patterns. Second, it allows probing the incoherent *light transport matrix* of a scene, using probing patterns that are challenging to implement with alternative techniques. Third, it facilitates incorporating these probing capabilities within other interferometric techniques, for example optical coherence tomography.

Our paper begins with background on the Michelson interferometer and the notions of spatial and temporal coherence. We then use this background to relate interferometry to measurements of the transmission matrix characterizing coherent propagation of light. In particular, we explain how, by modulating the spatial coherence properties of the illumination used for interferometry, we can control which elements of the transmission matrix contribute to image

Authors' addresses: Alankar Kotwal, Carnegie Mellon University; Anat Levin, Technion; Ioannis Gkioulekas, Carnegie Mellon University.

0730-0301/2020/7-ART74 $15.00
https://doi.org/10.1145/3386569.3392384

Learning Temporal Coherence via Self-Supervision for GAN-based Video Generation

MENGYU CHU*, YOU XIE*, JONAS MAYER, LAURA LEAL-TAIXÉ, and NILS THUEREY,
Technical University of Munich, Germany

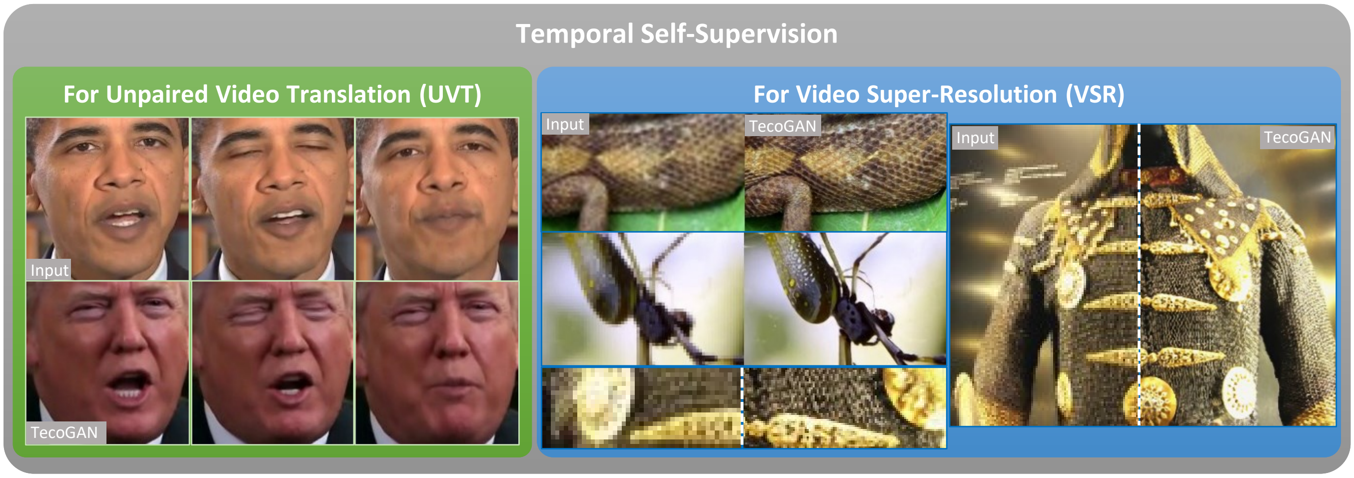


Fig. 1. Using the proposed approach for temporal self-supervision, we achieve realistic results with natural temporal evolution for two inherently different video generation tasks: unpaired video translation (left) and video super-resolution (right). While the resulting sharpness can be evaluated via the still images above, the corresponding videos in our supplemental web-page (Sec. 1 and Sec.2) highlight the high quality of the temporal changes. Obama and Trump video courtesy of the White House (public domain).

Our work explores temporal self-supervision for GAN-based video generation tasks. While adversarial training successfully yields generative models for a variety of areas, temporal relationships in the generated data are much less explored. Natural temporal changes are crucial for sequential generation tasks , e.g. video super-resolution and unpaired video translation. For the former, state-of-the-art methods often favor simpler norm losses such as L^2 over adversarial training. However, their averaging nature easily leads to temporally smooth results with an undesirable lack of spatial detail. For unpaired video translation, existing approaches modify the generator networks to form spatio-temporal cycle consistencies. In contrast, we focus on improving learning objectives and propose a temporally self-supervised algorithm. For both tasks, we show that temporal adversarial learning is key to achieving temporally coherent solutions without sacrificing spatial detail. We also propose a novel Ping-Pong loss to improve the long-term temporal consistency. It effectively prevents recurrent networks from accumulating artifacts temporally without depressing detailed features. Additionally, we propose a first set of metrics to quantitatively evaluate the accuracy as well as the perceptual quality of the temporal evolution. A series of user studies confirm the rankings computed with these metrics. Code, data, models, and results are provided at https://github.com/thunil/TecoGAN.

CCS Concepts: • **Computing methodologies → Neural networks**; *Image processing*.

Additional Key Words and Phrases: Generative adversarial network, temporal cycle-consistency, self-supervision, video super-resolution, unpaired video translation.

ACM Reference Format:
Mengyu Chu, You Xie, Jonas Mayer, Laura Leal-Taixé, and Nils Thuerey. 2020. Learning Temporal Coherence via Self-Supervision for GAN-based Video Generation. *ACM Trans. Graph.* 39, 4, Article 75 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392457

1 INTRODUCTION

Generative adversarial networks (GANs) have been extremely successful at learning complex distributions such as natural images [Isola et al. 2017; Zhu et al. 2017]. However, for sequence generation, directly applying GANs without carefully engineered constraints typically results in strong artifacts over time due to the significant difficulties introduced by the temporal changes. In particular, conditional video generation tasks are very challenging learning problems where generators should not only learn to represent the

*Both authors contributed equally to the paper

Authors' address: Mengyu Chu, mengyu.chu@tum.de; You Xie, you.xie@tum.de; Jonas Mayer, jonas.a.mayer@tum.de; Laura Leal-Taixé, leal.taixe@tum.de; Nils Thuerey, nils.thuerey@tum.de,
Technical University of Munich, Department of Computer Science, Munich, Germany.

0730-0301/2020/7-ART75 $15.00
https://doi.org/10.1145/3386569.3392457

One Shot 3D Photography

JOHANNES KOPF, KEVIN MATZEN, SUHIB ALSISAN, OCEAN QUIGLEY, FRANCIS GE, YANGMING CHONG, JOSH PATTERSON, JAN-MICHAEL FRAHM, SHU WU, MATTHEW YU, PEIZHAO ZHANG, ZIJIAN HE, PETER VAJDA, AYUSH SARAF, and MICHAEL COHEN, Facebook

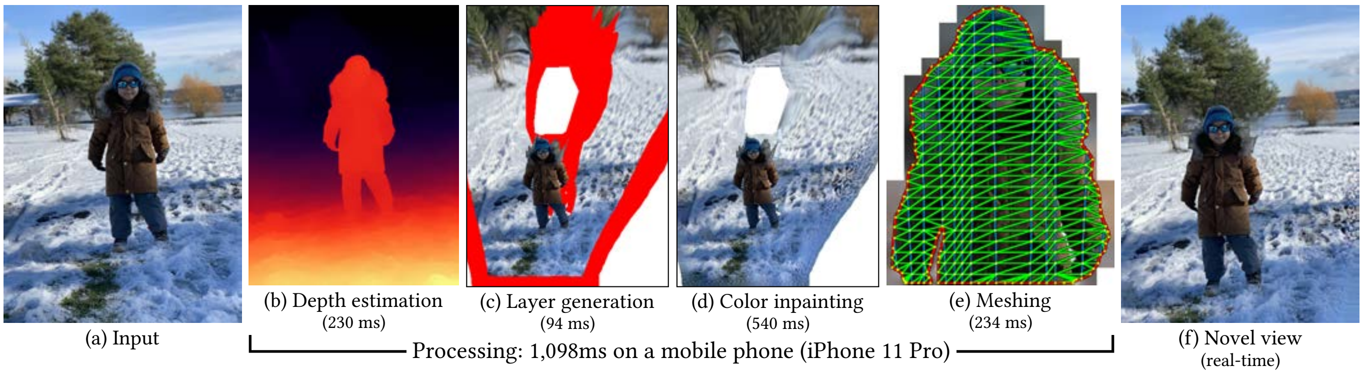


Fig. 1. We present a system for creating *3D photos* from a single mobile phone picture (a). The process involves learning-based algorithms for estimating depth from the 2D input (b) and texture inpainting (d), as well as conventional algorithms for lifting the geometry to 3D and extending it in parallax regions (c), as well as generating a final mesh-based representation (e). All steps are optimized to be fast given the limited compute and memory resources available on a mobile device. The resulting representation (f) can be viewed instantly, generating novel viewpoints at real-time rates.

3D photography is a new medium that allows viewers to more fully experience a captured moment. In this work, we refer to a *3D photo* as one that displays parallax induced by moving the viewpoint (as opposed to a stereo pair with a fixed viewpoint). 3D photos are static in time, like traditional photos, but are displayed with interactive parallax on mobile or desktop screens, as well as on Virtual Reality devices, where viewing it *also* includes stereo. We present an end-to-end system for creating and viewing 3D photos, and the algorithmic and design choices therein. Our 3D photos are captured in a single shot and processed directly on a mobile device. The method starts by estimating depth from the 2D input image using a new monocular depth estimation network that is optimized for mobile devices. It performs competitively to the state-of-the-art, but has lower latency and peak memory consumption and uses an order of magnitude fewer parameters. The resulting depth is lifted to a layered depth image, and new geometry is synthesized in parallax regions. We synthesize color texture and structures in the parallax regions as well, using an inpainting network, also optimized for mobile devices, on the LDI directly. Finally, we convert the result into a mesh-based representation that can be efficiently transmitted and rendered even on low-end devices and over poor network connections. Altogether, the processing takes just a few seconds on a mobile device, and the result can be instantly viewed and shared. We perform extensive quantitative evaluation to validate our system and compare its new components against the current state-of-the-art.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; **Machine learning**.

Additional Key Words and Phrases: 3D Photography, Depth Estimation

ACM Reference Format:
Johannes Kopf, Kevin Matzen, Suhib Alsisan, Ocean Quigley, Francis Ge, Yangming Chong, Josh Patterson, Jan-Michael Frahm, Shu Wu, Matthew Yu, Peizhao Zhang, Zijian He, Peter Vajda, Ayush Saraf, and Michael Cohen. 2020. One Shot 3D Photography. *ACM Trans. Graph.* 39, 4, Article 76 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392420

1 INTRODUCTION

Traditional 2D photography lets us capture the world around us, with a single click, as an instant frozen in time. *3D photography* is a new way to make these captured moments come back alive. We use the term *3D photo* to refer to any representation that can be displayed with parallax induced by viewpoint motion at viewing time (as opposed to a stereo pair, where inter-ocular parallax is baked in at capture time). Although still static in time, 3D photos can be interactively explored. The ability to change the viewpoint is compelling on "flat" mobile or desktop screens, and enables truly life-like experiences in Virtual Reality, by adding stereo viewing to head-motion induced parallax.

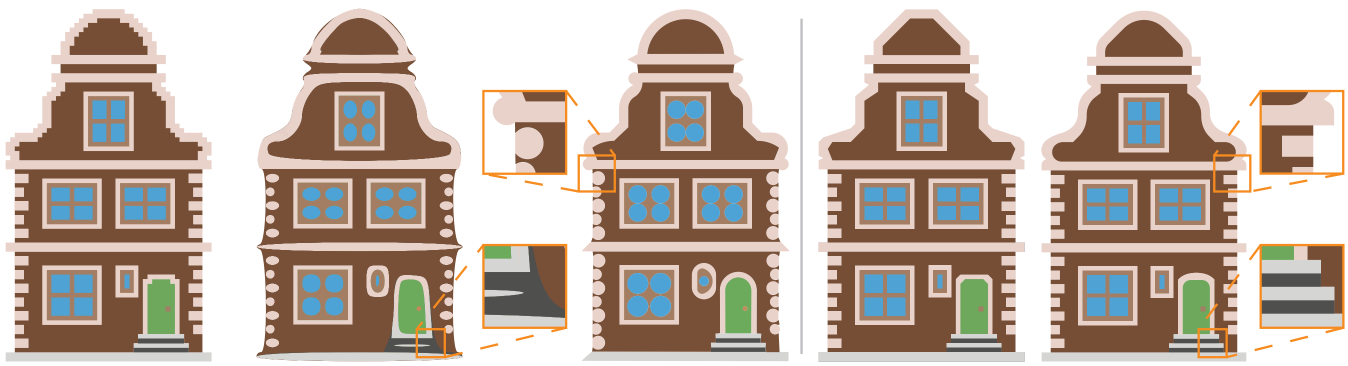
However, creating and displaying 3D photos poses challenges that are not present in 2D or even stereo photography: dense depth is required in addition to color, viewpoint changes reveal previously

Authors' address: Johannes Kopf; Kevin Matzen; Suhib Alsisan; Ocean Quigley; Francis Ge; Yangming Chong; Josh Patterson; Jan-Michael Frahm; Shu Wu; Matthew Yu; Peizhao Zhang; Zijian He; Peter Vajda; Ayush Saraf; Michael CohenFacebook.

0730-0301/2020/7-ART76 $15.00
https://doi.org/10.1145/3386569.3392420

PolyFit: Perception-Aligned Vectorization of Raster Clip-Art via Intermediate Polygonal Fitting

EDOARDO ALBERTO DOMINICI, University of British Columbia
NICO SCHERTLER, University of British Columbia
JONATHAN GRIFFIN, University of British Columbia
SHAYAN HOSHYARI, Adobe Inc.
LEONID SIGAL, University of British Columbia
ALLA SHEFFER, University of British Columbia



(a) Raster Input (b) [Selinger 2003] (c) [Hoshyari et al. 2018] (d) Intermediate Polygon (e) Ours

Fig. 1. Vectorizing raster clip-art inputs (a) using existing methods, here Potrace [Selinger 2003] (b) and [Hoshyari et al. 2018] (c), results in visible artifacts (highlighted in zoomed-in insets). PolyFit outputs (e), created using an intermediate polygonal approximation step (d), are more consistent with viewer expectations than those produced by these alternatives. Please zoom in online to see details. Input image ©IconScout — www.iconscout.com

Raster clip-art images, which consist of distinctly colored regions separated by sharp boundaries typically allow for a clear mental vector interpretation. Converting these images into vector format can facilitate compact lossless storage and enable numerous processing operations. Despite recent progress, existing vectorization methods that target such data frequently produce vectorizations that fail to meet viewer expectations. We present *PolyFit*, a new clip-art vectorization method that produces vectorizations well aligned with human preferences. Since segmentation of such inputs into regions had been addressed successfully, we specifically focus on fitting piecewise smooth vector curves to the raster input region boundaries, a task prior methods are particularly prone to fail on. While perceptual studies suggest the criteria humans are likely to use during mental boundary vectorization, they provide no guidance as to the exact interaction between them; learning these interactions directly is problematic due to the large size of the solution space. To obtain the desired solution, we first approximate the raster region boundaries with coarse intermediate polygons leveraging a combination of perceptual cues with observations from studies of human preferences. We then use these intermediate polygons as auxiliary inputs for computing piecewise smooth vectorizations of raster inputs. We define a finite set of potential polygon to curve primitive maps, and learn the mapping from the polygons to their best fitting primitive configurations from human annotations, arriving at a compact set of local raster and polygon properties whose combinations reliably predict human-expected primitive choices. We use these primitives to obtain a final globally consistent spline vectorization. Extensive comparative user studies show that our method outperforms state-of-the-art approaches on a wide range of data, where our results are preferred three times as often as those of the closest competitor across multiple types of inputs with various resolutions.

Authors' addresses: Edoardo Alberto Dominici, University of British Columbia, dedoardo@cs.ubc.ca; Nico Schertler, University of British Columbia, nschertl@mail.ubc.ca; Jonathan Griffin, University of British Columbia, jcgriff@cs.ubc.ca; Shayan Hoshyari, Adobe Inc., s.hoshyari@gmail.com; Leonid Sigal, University of British Columbia, lsigal@cs.ubc.ca; Alla Sheffer, University of British Columbia, sheffa@cs.ubc.ca.

0730-0301/2020/7-ART77 $15.00
https://doi.org/10.1145/3386569.3392401

CCS Concepts: • **Computing methodologies** → **Image manipulation**.

Additional Key Words and Phrases: clip-art, vectorization

ACM Reference Format:
Edoardo Alberto Dominici, Nico Schertler, Jonathan Griffin, Shayan Hoshyari, Leonid Sigal, and Alla Sheffer. 2020. PolyFit: Perception-Aligned Vectorization of Raster Clip-Art via Intermediate Polygonal Fitting. *ACM Trans. Graph.* 39, 4, Article 77 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392401

Portrait Shadow Manipulation

XUANER (CECILIA) ZHANG, University of California, Berkeley
JONATHAN T. BARRON, Google Research
YUN-TA TSAI, Google Research
ROHIT PANDEY, Google
XIUMING ZHANG, MIT
REN NG, University of California, Berkeley
DAVID E. JACOBS, Google Research

Fig. 1. The results of our portrait enhancement method on real-world portrait photographs. Casual portrait photographs often suffer from undesirable shadows, particularly foreign shadows cast by external objects, and dark facial shadows cast by the face upon itself under harsh illumination. We propose an automated technique for enhancing these poorly-lit portrait photographs by removing unwanted foreign shadows, reducing harsh facial shadows, and adding synthetic fill lights.

Casually-taken portrait photographs often suffer from unflattering lighting and shadowing because of suboptimal conditions in the environment. Aesthetic qualities such as the position and softness of shadows and the lighting ratio between the bright and dark parts of the face are frequently determined by the constraints of the environment rather than by the photographer. Professionals address this issue by adding light shaping tools such as scrims, bounce cards, and flashes. In this paper, we present a computational approach that gives casual photographers some of this control, thereby allowing poorly-lit portraits to be relit post-capture in a realistic and easily-controllable way. Our approach relies on a pair of neural networks—one to remove foreign shadows cast by external objects, and another to soften facial shadows cast by the features of the subject and to add a synthetic fill light to improve the lighting ratio. To train our first network we construct a dataset of real-world portraits wherein synthetic foreign shadows are rendered onto the face, and we show that our network learns to remove those unwanted shadows. To train our second network we use a dataset of Light Stage scans of human subjects to construct input/output pairs of input images harshly lit by a small light source, and variably softened and fill-lit output images of each face. We propose a way to explicitly encode facial symmetry and show that our dataset and training procedure enable the model to generalize to images taken in the wild. Together, these networks enable the realistic and aesthetically pleasing enhancement of shadows and lights in real-world portrait images.[1]

CCS Concepts: • **Computing methodologies → Computational photography**.

Additional Key Words and Phrases: computational-photography

Authors' addresses: Xuaner (Cecilia) Zhang, University of California, Berkeley, cecilia77@berkeley.edu; Jonathan T. Barron, Google Research, barron@google.com; Yun-Ta Tsai, Google Research, yuntatsai@google.com; Rohit Pandey, Google, rohitpandey@google.com; Xiuming Zhang, MIT, xiuming@csail.mit.edu; Ren Ng, University of California, Berkeley, ren@berkeley.edu; David E. Jacobs, Google Research, dejacobs@google.com.

0730-0301/2020/7-ART78 $15.00
https://doi.org/10.1145/3386569.3392390

[1] https://people.eecs.berkeley.edu/~cecilia77/project-pages/portrait

Quanta Burst Photography

SIZHUO MA and SHANTANU GUPTA, University of Wisconsin-Madison, USA
ARIN C. ULKU, CLAUDIO BRUSCHINI, and EDOARDO CHARBON, EPFL, Switzerland
MOHIT GUPTA, University of Wisconsin-Madison, USA

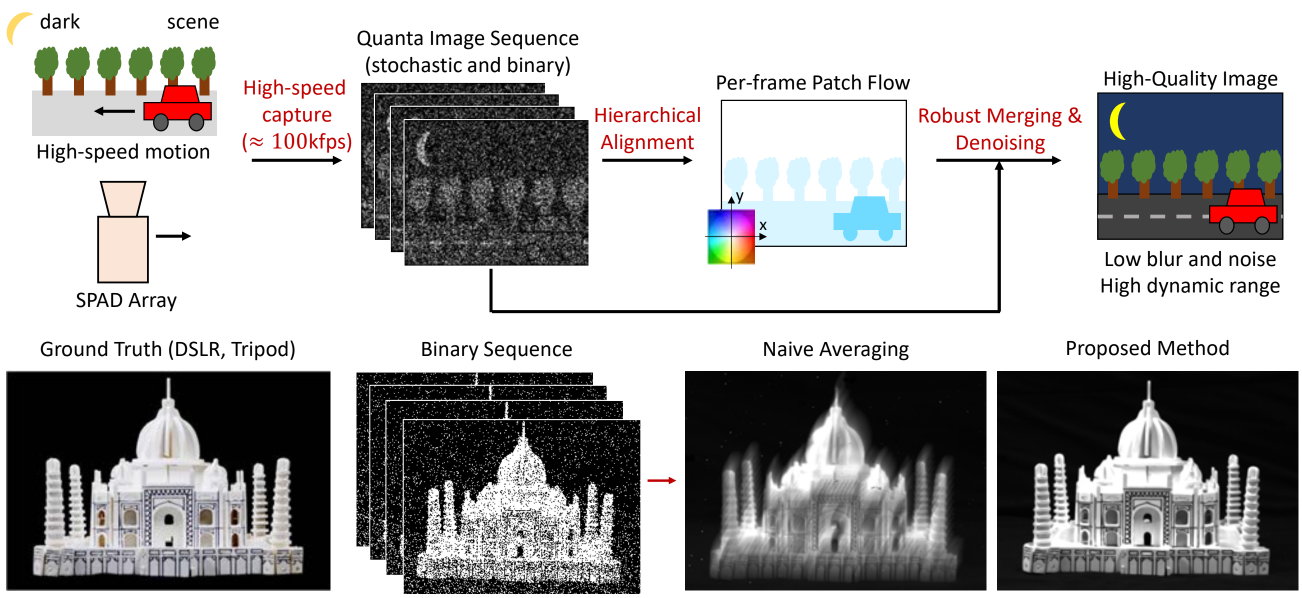


Fig. 1. **Quanta burst photography.** (Top) Single-photon image sensors capture stochastic, binary image sequences at high speeds (∼ 100 kfps). Such high-speed image sequences can be aligned to compensate for scene/camera motion using a spatial-temporal hierarchical alignment algorithm. By merging the aligned sequence robustly, a high-quality image can be reconstructed, with minimal motion blur and noise, and high dynamic range, even in challenging photography conditions. (Bottom, from left to right) An example low-light scene captured by a DSLR camera on a tripod to avoid camera shake; binary image sequence captured by a handheld single-photon camera; image reconstructed by naive averaging of the binary sequence (shown to illustrate the amount of motion during capture); super-resolved image reconstructed using the proposed techniques has low blur and noise. **Zoom in for details.**

Single-photon avalanche diodes (SPADs) are an emerging sensor technology capable of detecting individual incident photons, and capturing their time-of-arrival with high timing precision. While these sensors were limited to single-pixel or low-resolution devices in the past, recently, large (up to 1 MPixel) SPAD arrays have been developed. These single-photon cameras (SPCs) are capable of capturing high-speed sequences of binary single-photon images with no read noise. We present quanta burst photography, a computational photography technique that leverages SPCs as *passive imaging devices* for photography in challenging conditions, including ultra low-light and fast motion. Inspired by recent success of conventional burst photography, we design algorithms that align and merge binary sequences captured by SPCs into intensity images with minimal motion blur and artifacts, high signal-to-noise ratio (SNR), and high dynamic range. We theoretically analyze the SNR and dynamic range of quanta burst photography, and identify the imaging regimes where it provides significant benefits. We demonstrate, via a recently developed SPAD array, that the proposed method is able to generate high-quality images for scenes with challenging lighting, complex geometries, high dynamic range and moving objects. With the ongoing development of SPAD arrays, we envision quanta burst photography finding applications in both consumer and scientific photography.

Authors' addresses: Sizhuo Ma, sizhuoma@cs.wisc.edu; Shantanu Gupta, sgupta226@wisc.edu, University of Wisconsin-Madison, USA; Arin C. Ulku, arin.ulku@epfl.ch; Claudio Bruschini, claudio.bruschini@epfl.ch; Edoardo Charbon, edoardo.charbon@epfl.ch, EPFL, Switzerland; Mohit Gupta, mohitg@cs.wisc.edu, University of Wisconsin-Madison, USA.

0730-0301/2020/7-ART79 $15.00
https://doi.org/10.1145/3386569.3392470

CCS Concepts: • **Computing methodologies** → **Computational photography**.

Additional Key Words and Phrases: Single-photon camera, single-photon avalanche diode, quanta image sensor, burst photography, super-resolution, high dynamic range, high-speed imaging, low-light imaging

ACM Reference Format:
Sizhuo Ma, Shantanu Gupta, Arin C. Ulku, Claudio Bruschini, Edoardo Charbon, and Mohit Gupta. 2020. Quanta Burst Photography. *ACM Trans. Graph.* 39, 4, Article 79 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392470

Single Image HDR Reconstruction Using a CNN with Masked Features and Perceptual Loss

MARCEL SANTANA SANTOS, Centro de Informática, Universidade Federal de Pernambuco
TSANG ING REN, Centro de Informática, Universidade Federal de Pernambuco
NIMA KHADEMI KALANTARI, Texas A&M University

Fig. 1. We propose a novel deep learning system for single image HDR reconstruction by synthesizing visually pleasing details in the saturated areas. We introduce a new feature masking approach that reduces the contribution of the features computed on the saturated areas, to mitigate halo and checkerboard artifacts. To synthesize visually pleasing textures in the saturated regions, we adapt the VGG-based perceptual loss function to the HDR reconstruction application. Furthermore, to effectively train our network on limited HDR training data, we propose to pre-train the network on inpainting task. Our method can reconstruct regions with high luminance, such as the bright highlights of the windows (red inset), and generate visually pleasing textures (green insert). See Figure 7 for comparison against several other approaches. All images have been gamma corrected for display purposes.

Digital cameras can only capture a limited range of real-world scenes' luminance, producing images with saturated pixels. Existing single image high dynamic range (HDR) reconstruction methods attempt to expand the range of luminance, but are not able to hallucinate plausible textures, producing results with artifacts in the saturated areas. In this paper, we present a novel learning-based approach to reconstruct an HDR image by recovering the saturated pixels of an input LDR image in a visually pleasing way. Previous deep learning-based methods apply the same convolutional filters on well-exposed and saturated pixels, creating ambiguity during training and leading to checkerboard and halo artifacts. To overcome this problem, we propose a feature masking mechanism that reduces the contribution of the features from the saturated areas. Moreover, we adapt the VGG-based perceptual loss function to our application to be able to synthesize visually pleasing textures. Since the number of HDR images for training is limited, we propose to train our system in two stages. Specifically, we first train our system on a large number of images for image inpainting task and then fine-tune it on HDR reconstruction. Since most of the HDR examples contain smooth regions that are simple to reconstruct, we propose a sampling strategy to select challenging training patches during the HDR fine-tuning stage. We demonstrate through experimental results that our approach can reconstruct visually pleasing HDR results, better than the current state of the art on a wide range of scenes.

CCS Concepts: • **Computing methodologies → Computational photography**.

Additional Key Words and Phrases: high dynamic range imaging, convolutional neural network, feature masking, perceptual loss

ACM Reference Format:
Marcel Santana Santos, Tsang Ing Ren, and Nima Khademi Kalantari. 2020. Single Image HDR Reconstruction Using a CNN with Masked Features and Perceptual Loss. *ACM Trans. Graph.* 39, 4, Article 80 (July 2020), 10 pages. https://doi.org/10.1145/3386569.3392403

Authors' addresses: Marcel Santana Santos, Centro de Informática, Universidade Federal de Pernambuco, mss8@cin.ufpe.br; Tsang Ing Ren, Centro de Informática, Universidade Federal de Pernambuco, tir@cin.ufpe.br; Nima Khademi Kalantari, Texas A&M University, nimak@tamu.edu.

0730-0301/2020/7-ART80 $15.00
https://doi.org/10.1145/3386569.3392403

1 INTRODUCTION

The illumination of real-world scenes is high dynamic range, but standard digital cameras sensors can only capture a limited range of luminance. Therefore, these cameras typically produce images with

Single-Shot High-Quality Facial Geometry and Skin Appearance Capture

JÉRÉMY RIVIERE, PAULO GOTARDO, and DEREK BRADLEY, DisneyResearch|Studios
ABHIJEET GHOSH, Imperial College London
THABO BEELER, DisneyResearch|Studios

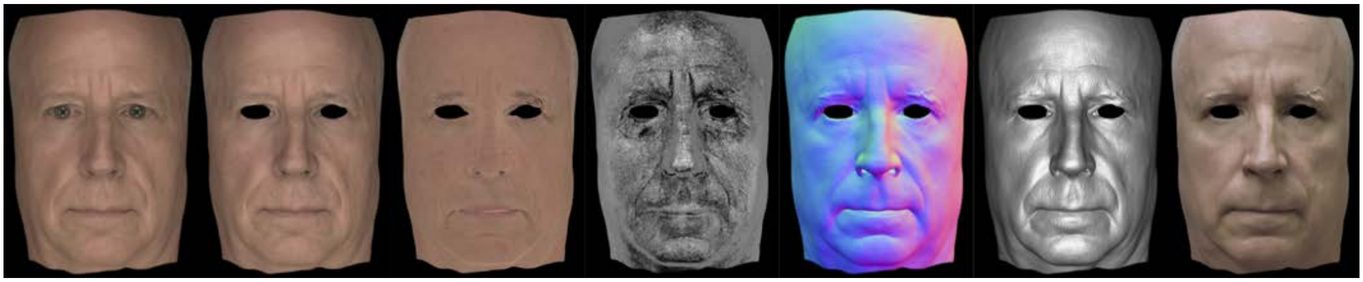


Fig. 1. We present a method to capture complete facial geometry and appearance from a *single* exposure. From left to right: one input image, our matching render, diffuse albedo, specular intensity, normals, high resolution geometry, and a realistic re-render under a different environment map.

We propose a new light-weight face capture system capable of reconstructing both *high-quality geometry and detailed appearance maps* from a *single exposure*. Unlike currently employed appearance acquisition systems, the proposed technology does not require active illumination and hence can readily be integrated with passive photogrammetry solutions. These solutions are in widespread use for 3D scanning humans as they can be assembled from off-the-shelf hardware components, but lack the capability of estimating appearance. This paper proposes a solution to overcome this limitation, by adding appearance capture to photogrammetry systems. The only additional hardware requirement to these solutions is that a subset of the cameras are cross-polarized with respect to the illumination, and the remaining cameras are parallel-polarized. The proposed algorithm leverages the images with the two different polarization states to reconstruct the geometry and to recover appearance properties. We do so by means of an inverse rendering framework, which solves *per texel diffuse albedo, specular intensity, and high-resolution normals, as well as global specular roughness* considering the subsurface scattering nature of skin. We show results for a variety of human subjects of different ages and skin typology, illustrating how the captured fine-detail skin surface and subsurface scattering effects lead to realistic renderings of their digital doubles, also in different illumination conditions.

CCS Concepts: • **Computing methodologies → Reflectance modeling**; **3D imaging**; *Appearance and texture representations.*

Additional Key Words and Phrases: Passive Photogrammetry, Dynamic Face Capture, Appearance Capture, High-Detail Surface, Inverse Rendering

Authors' addresses: Jérémy Riviere, jeremy.riviere@disneyresearch.com; Paulo Gotardo, paulo.gotardo@disneyresearch.com; Derek Bradley, derek.bradley@disneyresearch.com, DisneyResearch|Studios; Abhijeet Ghosh, ghosh@imperial.ac.uk, Imperial College London; Thabo Beeler, thabo.beeler@gmail.com, DisneyResearch|Studios.

0730-0301/2020/7-ART81 $15.00
https://doi.org/10.1145/3386569.3392464

ACM Reference Format:
Jérémy Riviere, Paulo Gotardo, Derek Bradley, Abhijeet Ghosh, and Thabo Beeler. 2020. Single-Shot High-Quality Facial Geometry and Skin Appearance Capture. *ACM Trans. Graph.* 39, 4, Article 81 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392464

1 INTRODUCTION

Digital humans have become omnipresent in today's entertainment landscape, making an appearance in nearly every blockbuster movie and triple-A video game. To create such digital characters it is common practice to 3D scan real humans, digitally capturing their likeness. To accomplish this, passive photogrammetry solutions have become the method of choice for two reasons. Firstly, passive photogrammetry systems can be constructed from off-the-shelf consumer hardware, such as digital cameras and flashes, and are hence much less complex and more cost effective than active technologies, such as structured light scanning or lightstage acquisition. Secondly, a number of software solutions exist, both commercial and open-source, that allow to reconstruct high-quality 3D geometry from the acquired images. This makes 3D shape acquisition readily and widely available.

Standard photogrammetry alone, however, is not sufficient to create photorealistic digital human assets. In addition to 3D shape, high-quality diffuse and specular reflectance properties are also required for realistic rendering. Furthermore, the level of geometric detail provided by photogrammetry has been typically inferior when compared to 3D shapes obtained with more complex setups based on active lighting, such as lightstages [Debevec et al. 2000] and other recent videogrammetry solutions [Gotardo et al. 2018]. Thus, to acquire these appearance properties and fine-detail geometry, studios and digital artists are currently forced to employ costly and complex setups that require expert knowledge to build and operate. As a result, high-quality appearance acquisition is currently only viable for hero assets in high-budget productions.

In this paper, we propose the first light-weight, inexpensive, single-shot acquisition system that can capture both high-quality

XNect: Real-time Multi-Person 3D Motion Capture with a Single RGB Camera

DUSHYANT MEHTA, MPI For Informatics, Saarland Informatics Campus
OLEKSANDR SOTNYCHENKO, MPI For Informatics, Saarland Informatics Campus
FRANZISKA MUELLER, MPI For Informatics, Saarland Informatics Campus
WEIPENG XU, MPI For Informatics, Saarland Informatics Campus
MOHAMED ELGHARIB, MPI For Informatics, Saarland Informatics Campus
PASCAL FUA, École Polytechnique Fédérale de Lausanne
HANS-PETER SEIDEL, MPI For Informatics, Saarland Informatics Campus
HELGE RHODIN, University of British Columbia, École Polytechnique Fédérale de Lausanne
GERARD PONS-MOLL, MPI For Informatics, Saarland Informatics Campus
CHRISTIAN THEOBALT, MPI For Informatics, Saarland Informatics Campus

Fig. 1. Our real-time monocular RGB based 3D motion capture provides temporally coherent estimates of the full 3D pose of multiple people in the scene, handling occlusions and interactions in general scene settings, and localizing subjects relative to the camera. Our design allows the system to handle large groups of people in the scene with the run-time only minimally affected by the number of people in the scene. Our method yields full skeletal pose in terms of joint angles, which can readily be employed for live character animation. Some images courtesy [KNG Music 2019], [Music Express Magazine 2013a]. 3D characters from Mixamo [Adobe 2020].

We present a real-time approach for multi-person 3D motion capture at over 30 fps using a single RGB camera. It operates successfully in generic scenes which may contain occlusions by objects and by other people. Our method operates in subsequent stages. The first stage is a convolutional neural network (CNN) that estimates 2D and 3D pose features along with identity assignments for all visible joints of all individuals. We contribute a new architecture for this CNN, called *SelecSLS Net*, that uses novel selective long and short range skip connections to improve the information flow allowing for a drastically faster network without compromising accuracy. In the second stage, a fully-connected neural network turns the possibly partial (on account of occlusion) 2D pose and 3D pose features for each subject into a complete 3D pose estimate per individual. The third stage applies space-time skeletal model fitting to the predicted 2D and 3D pose per subject to further reconcile the 2D and 3D pose, and enforce temporal coherence. Our method returns the full skeletal pose in joint angles for each subject. This is a further key distinction from previous work that do not produce joint angle results of a coherent skeleton in real time for multi-person scenes. The proposed system runs on consumer hardware at a previously unseen speed of more than 30 fps given 512x320 images as input while achieving state-of-the-art accuracy, which we will demonstrate on a range of challenging real-world scenes.

CCS Concepts: • **Computing methodologies → Motion capture**; *Computer vision*; *Neural networks.*

Additional Key Words and Phrases: human body pose, motion capture, real-time, monocular, RGB

Authors' addresses: Dushyant Mehta, Oleksandr Sotnychenko, Franziska Mueller, Weipeng Xu, Mohamed Elgharib, Hans-Peter Seidel, Gerard Pons-Moll, and Christian Theobalt, MPI For Informatics, Saarland Informatics Campus; Pascal Fua, École Polytechnique Fédérale de Lausanne; Helge Rhodin, University of British Columbia.

0730-0301/2020/7-ART82 $15.00
https://doi.org/10.1145/3386569.3392410

ARAnimator: In-situ Character Animation in Mobile AR with User-defined Motion Gestures

HUI YE*, KIN CHUNG KWAN*, WANCHAO SU, and HONGBO FU†, City University of Hong Kong

Fig. 1. Our *ARAnimator* allows users to move an AR-enabled mobile device to directly control and animate a virtual character situated in a real-world environment. Please refer to the accompanying video for the animation results.

Creating animated virtual AR characters closely interacting with real environments is interesting but difficult. Existing systems adopt video see-through approaches to indirectly control a virtual character in mobile AR, making close interaction with real environments not intuitive. In this work we use an AR-enabled mobile device to directly control the position and motion of a virtual character situated in a real environment. We conduct two guessability studies to elicit user-defined motions of a virtual character interacting with real environments, and a set of user-defined motion gestures describing specific character motions. We found that an SVM-based learning approach achieves reasonably high accuracy for gesture classification from the motion data of a mobile device. We present *ARAnimator*, which allows novice and casual animation users to directly represent a virtual character by an AR-enabled mobile phone and control its animation in AR scenes using motion gestures of the device, followed by animation preview and interactive editing through a video see-through interface. Our experimental results show that with *ARAnimator*, users are able to easily create in-situ character animations closely interacting with different real environments.

CCS Concepts: • **Human-centered computing** → *Gestural input*; *Mixed / augmented reality*; **Mobile devices**.

Additional Key Words and Phrases: Mobile Augmented Reality, Interactive System, Character Animation, User Defined Gestures, Gesture Classification

ACM Reference Format:
Hui Ye, Kin Chung Kwan, Wanchao Su, and Hongbo Fu. 2020. *ARAnimator*: In-situ Character Animation in Mobile AR with User-defined Motion Gestures. *ACM Trans. Graph.* 39, 4, Article 83 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392404

*Both authors contributed equally to the paper.
†Corresponding author.

Authors' address: Hui Ye, huiye4-c@my.cityu.edu.hk; Kin Chung Kwan, kckwan@ieee.org; Wanchao Su, wanchaosu2-c@my.cityu.edu.hk; Hongbo Fu, hongbofu@cityu.edu.hk, City University of Hong Kong, School of Creative Media, 18 Tat Hong Avenue Kowloon Tong, Hong Kong.

0730-0301/2020/7-ART83 $15.00
https://doi.org/10.1145/3386569.3392404

1 INTRODUCTION

Augmented Reality (AR) aims to augment a real world with virtual content, often with proper alignment. With AR technologies, augmenting virtual *static* objects into our real world is relatively easy. However, even with powerful mobile AR platforms recently developed by Apple (i.e., ARKit) and Google (i.e., ARCore), creating animated contents closely interacting with real environments is still challenging. Since virtual objects need to be aligned with real environments at different moments, making the reuse of existing animations difficult. In-situ creation of animation thus becomes more important.

Several tools such as Motion Doodles [Thorne et al. 2004], Spatial Motion Doodles [Garcia et al. 2019] and PuppetPhone [Anderegg et al. 2018] have been designed to allow novice users to create character animations on top of a virtual or real scene. They can be directly applied or extended for in-situ creation of AR animation. However,

HeadBlaster: A Wearable Approach to Simulating Motion Perception using Head-mounted Air Propulsion Jets

SHI-HONG LIU, PAI-CHIEN YEN, YI-HSUAN MAO, YU-HSIN LIN, ERICK CHANDRA, and MIKE Y. CHEN, National Taiwan University

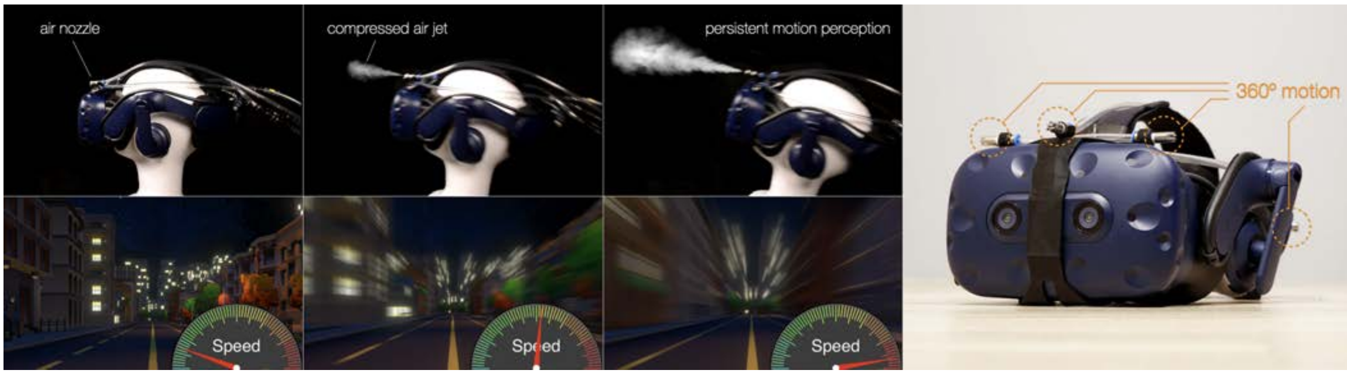


Fig. 1. HeadBlaster: a) applies ungrounded air propulsion force to the head to stimulate the vestibular and proprioception sensory systems to create the perception of persistent self-motion (note: the white smoke is used here only for illustrative purposes; in regular usage, the compressed air is invisible), and b) our system uses 6 air nozzles mounted on VR headsets and combines multiple compressed air jets to generate lateral forces in 360 degrees.

We present HeadBlaster, a novel wearable technology that creates motion perception by applying ungrounded force to the head to stimulate the vestibular and proprioception sensory systems. Compared to motion platforms that tilt the body, HeadBlaster more closely approximates how lateral inertial and centrifugal forces are felt during real motion to provide more persistent motion perception. In addition, because HeadBlaster only actuates the head rather than the entire body, it eliminates the mechanical motion platforms that users must be constrained to, which improves user mobility and enables room-scale VR experiences. We designed a wearable HeadBlaster system with 6 air nozzles integrated into a VR headset, using compressed air jets to provide persistent, lateral propulsion forces. By controlling multiple air jets, it is able to create the perception of lateral acceleration in 360 degrees. We conducted a series of perception and human-factor studies to quantify the head movement, the persistence of perceived acceleration, and the minimal level of detectable forces. We then explored the user experience of HeadBlaster through two VR applications: a custom surfing game, and a commercial driving simulator together with a commercial motion platform. Study results showed that HeadBlaster provided significantly longer perceived duration of acceleration than motion platforms. It also significantly improved realism and immersion, and was preferred by users compared to using VR alone. In addition, it can be used in conjunction with motion platforms to further augment the user experience.

CCS Concepts: • **Human-centered computing → Haptic devices**; *Virtual reality*; *User studies.*

Additional Key Words and Phrases: Motion perception, vestibular system, proprioception

ACM Reference Format:
Shi-Hong Liu, Pai-Chien Yen, Yi-Hsuan Mao, Yu-Hsin Lin, Erick Chandra, and Mike Y. Chen. 2020. HeadBlaster: A Wearable Approach to Simulating Motion Perception using Head-mounted Air Propulsion Jets. *ACM Trans. Graph.* 39, 4, Article 84 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392482

Authors' address: Shi-Hong Liu, r07944009@ntu.edu.tw; Pai-Chien Yen, r07922135@ntu.edu.tw; Yi-Hsuan Mao, r08922005@ntu.edu.tw; Yu-Hsin Lin, r08944013@ntu.edu.tw; Erick Chandra, r06922147@ntu.edu.tw; Mike Y. Chen, mikechen@csie.ntu.edu.tw, National Taiwan University, Computer Science and Information Engineering.

0730-0301/2020/7-ART84 $15.00
https://doi.org/10.1145/3386569.3392482

1 INTRODUCTION

Motion platforms, or motion simulators, create the feelings of being in a real motion environment. They were first invented in 1906, when flight simulators were needed for pilot training as powered aircraft were developed at the beginning of the 20th century [Hancock et al. 2008]. Today, motion platforms are combined with rich visual and sounds to create the sensation of illusory self-motion, called vection, and are popular in arcades, theme parks, and 4D movie theaters. With the recent rise in popularity of VR headsets, interests in consumer motion platforms have been growing rapidly.

Human interprets head and body motion by integrating inputs from our vestibular (semicircular canals and otolith organs), somatosensory (specifically proprioception), and visual systems [Mack et al. 2013]. Visual cues such as displacement and optical flow can create illusion of self-motion called visual vection [Harris et al. 2002;

Human-in-the-Loop Differential Subspace Search in High-Dimensional Latent Space

CHIA-HSING CHIU, National Taiwan University of Science and Technology
YUKI KOYAMA, National Institute of Advanced Industrial Science and Technology (AIST)
YU-CHI LAI, National Taiwan University of Science and Technology
TAKEO IGARASHI, The University of Tokyo
YONGHAO YUE, Aoyama Gakuin University (AGU)

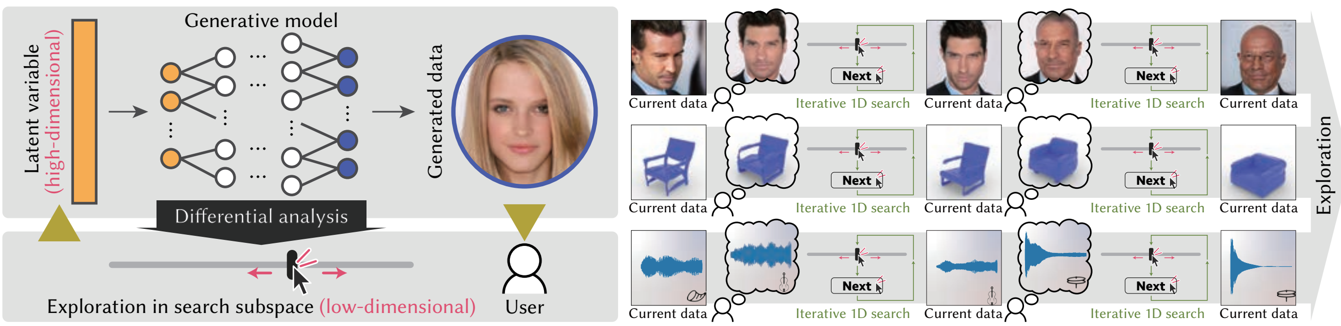


Fig. 1. We present *differential subspace search* for exploring the high-dimensional latent space of a deep generative model, by letting the user perform searches in (low-dimensional) 1D subspaces, where the search directions are provided through differential analysis of the generative model (left). The user iteratively performs searches via a slider interface and updates the subspace by pressing the 'next' button (right). Our method does not rely on domain- or data-specific assumptions and can be applied to exploratory tasks for various generative models for images, sounds, and 3D models.

Generative models based on deep neural networks often have a high-dimensional latent space, ranging sometimes to a few hundred dimensions or even higher, which typically makes them hard for a user to explore directly. We propose *differential subspace search* to allow efficient iterative user exploration in such a space, without relying on domain- or data-specific assumptions. We develop a general framework to extract low-dimensional subspaces based on a local differential analysis of the generative model, such that a small change in such a subspace would provide enough change in the resulting data. We do so by applying singular value decomposition to the Jacobian of the generative model and forming a subspace with the desired dimensionality spanned by a given number of singular vectors stochastically selected on the basis of their singular values, to maintain ergodicity. We use our framework to present 1D subspaces to the user via a 1D slider interface. Starting from an initial location, the user finds a new candidate in the presented 1D subspace, which is in turn updated at the new candidate location. This process is repeated until no further improvement can be made. Numerical simulations show that our method can better optimize synthetic black-box objective functions than the alternatives that we tested. Furthermore, we conducted a user study using complex generative models and the results show that our method enables more efficient exploration of high-dimensional latent spaces than the alternatives.

CCS Concepts: • **Computing methodologies → Computer graphics**; *Search methodologies*.

Additional Key Words and Phrases: Human-in-the-loop optimization, dimensionality reduction, generative models

ACM Reference Format:
Chia-Hsing Chiu, Yuki Koyama, Yu-Chi Lai, Takeo Igarashi, and Yonghao Yue. 2020. Human-in-the-Loop Differential Subspace Search in High-Dimensional Latent Space. *ACM Trans. Graph.* 39, 4, Article 85 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392409

Authors' addresses: C.-H. Chiu, Y.-C. Lai, T4-305-1, No.43, Sec.4, Keelung Rd., Taipei, 106, Taiwan; Y. Koyama, Central 2, 1-1-1 Umezono, Tsukuba, Ibaraki, 305-8568, Japan; T. Igarashi, Science bldg. 7-303, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan; Y. Yue, O-525, Fuchinobe 5-10-1, Chuo-ku, Sagamihara, Kanagawa, 252-5258, Japan.

0730-0301/2020/7-ART85 $15.00
https://doi.org/10.1145/3386569.3392409

1 INTRODUCTION

With the recent developments in deep learning, we have seen a rapid advancement in using machine learning to generate data for multimedia content; there are many impressive generative models with applications ranging from images [Karras et al. 2018; Kingma et al. 2014; Miyato et al. 2018] and sounds [Donahue et al. 2018; Engel et al. 2019] to 3D models [Chen and Zhang 2019; Umetani 2017], to just list a few. A key to this success is the ability to perform efficient and effective training using complex network architectures with more layers and higher-dimensional internal representations (i.e., latent spaces). For instance, IM-GAN for 3D shapes [Chen and Zhang 2019] and SN-GANs for images [Miyato et al. 2018] have 128-dimensional latent spaces, while GANSynth for audio [Engel et al. 2019] has 256. Such richness of network architectures allows

Immersive Light Field Video with a Layered Mesh Representation

MICHAEL BROXTON*, JOHN FLYNN*, RYAN OVERBECK*, DANIEL ERICKSON*, PETER HEDMAN, MATTHEW DUVALL, JASON DOURGARIAN, JAY BUSCH, MATT WHALEN, and PAUL DEBEVEC, Google

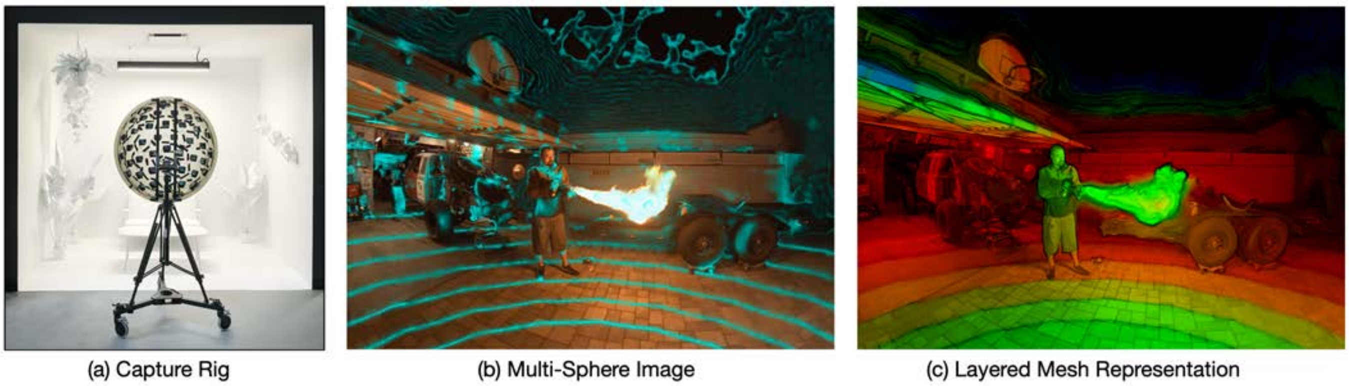


Fig. 1. Our light field capture rig and scene representations. (a) We record immersive light field video using 46 action sports cameras mounted to an acrylic dome. (b) Using deep view synthesis we infer an RGBA multi-sphere image (MSI) from the light field views. Every 10th spherical shell is highlighted. (c) We convert groups of MSI layers into Layered Meshes (each shown as a different color), which are texture atlased and compressed into light field video.

We present a system for capturing, reconstructing, compressing, and rendering high quality immersive light field video. We accomplish this by leveraging the recently introduced DeepView view interpolation algorithm, replacing its underlying multi-plane image (MPI) scene representation with a collection of spherical shells that are better suited for representing panoramic light field content. We further process this data to reduce the large number of shell layers to a small, fixed number of RGBA+depth layers without significant loss in visual quality. The resulting RGB, alpha, and depth channels in these layers are then compressed using conventional texture atlasing and video compression techniques. The final compressed representation is lightweight and can be rendered on mobile VR/AR platforms or in a web browser. We demonstrate light field video results using data from the 16-camera rig of [Pozo et al. 2019] as well as a new low-cost hemispherical array made from 46 synchronized action sports cameras. From this data we produce 6 degree of freedom volumetric videos with a wide 70 cm viewing baseline, 10 pixels per degree angular resolution, and a wide field of view, at 30 frames per second video frame rates. Advancing over previous work, we show that our system is able to reproduce challenging content such as view-dependent reflections, semi-transparent surfaces, and near-field objects as close as 34 cm to the surface of the camera rig.

*Denotes equal contribution.

Authors' address: Michael Broxton, broxton@google.com; John Flynn, jflynn@google.com; Ryan Overbeck, rover@google.com; Daniel Erickson, danerickson@google.com; Peter Hedman, hedman@google.com; Matthew DuVall, matthewduvall@google.com; Jason Dourgarian, jdourgarian@google.com; Jay Busch, jbusch@google.com; Matt Whalen, mswhalen@google.com; Paul Debevec, debevec@google.com, Google.

0730-0301/2020/7-ART86
https://doi.org/10.1145/3386569.3392485

CCS Concepts: • **Computing methodologies** → *Image compression*; *Neural networks*; **Virtual reality**; **Image-based rendering**; *Computational photography*.

Additional Key Words and Phrases: view synthesis, light fields, image-based rendering, deep learning

ACM Reference Format:
Michael Broxton, John Flynn, Ryan Overbeck, Daniel Erickson, Peter Hedman, Matthew DuVall, Jason Dourgarian, Jay Busch, Matt Whalen, and Paul Debevec. 2020. Immersive Light Field Video with a Layered Mesh Representation. *ACM Trans. Graph.* 39, 4, Article 86 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392485

1 INTRODUCTION

Our visual system is remarkable at perceiving the world around us from a single pair of eyes a few centimeters apart. Part of what makes it so effective is that moving our heads changes our perspective, allowing us to feel a greater sense of depth through motion parallax and a better sense of what the scene is made of by how light plays off of its surfaces. A virtual environment that is able to excite these same senses of motion parallax and view-dependent reflections can be far more immersive and realistic.

Many of today's augmented and virtual reality headsets provide positional tracking, enabling the user's view of the virtual scene to shift perspective properly as they move their head. If the scene is being rendered by a game engine, it is straightforward to feed the positional tracking data into the engine to yield proper motion parallax and view-dependent reflections. But if the scene is a real one recorded by an immersive video camera, changing the perspective is much more complicated. The majority of immersive video cameras deliver panoramic video from a single fixed point of view (e.g. Ricoh Theta Z1, GoPro Max, Insta360 One X) or omnidirectional stereo video fixed to a particular location in space (e.g. Yi Halo, Insta360

MEgATrack: Monochrome Egocentric Articulated Hand-Tracking for Virtual Reality

SHANGCHEN HAN, BEIBEI LIU, RANDI CABEZAS, CHRISTOPHER D. TWIGG, PEIZHAO ZHANG, JEFF PETKAU, TSZ-HO YU, CHUN-JUNG TAI, MUZAFFER AKBAY, ZHENG WANG, ASAF NITZAN, GANG DONG, YUTING YE, LINGLING TAO, CHENGDE WAN, and ROBERT WANG, Facebook Reality Labs

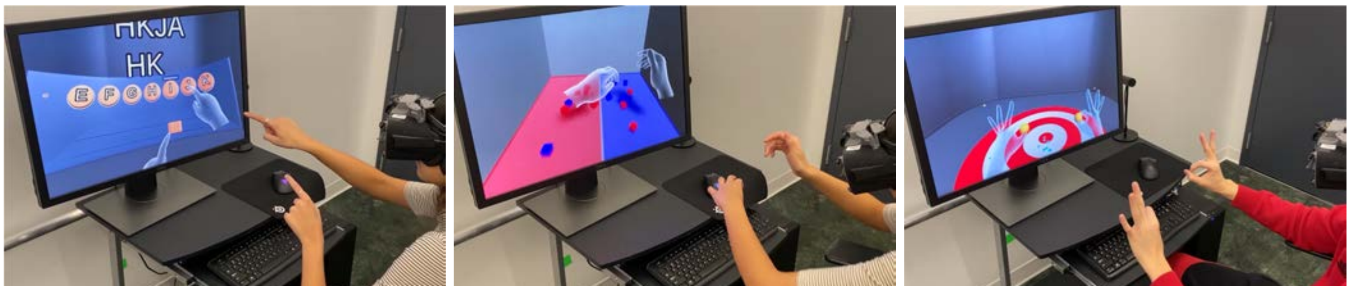


Fig. 1. We present a real-time hand-tracking system using four monochrome cameras mounted on a VR headset. We output the user's skeletal poses and rigged hand model meshes. Here we show some snapshots of users using our system to drive interactive VR experiences.

We present a system for real-time hand-tracking to drive virtual and augmented reality (VR/AR) experiences. Using four fisheye monochrome cameras, our system generates accurate and low-jitter 3D hand motion across a large working volume for a diverse set of users. We achieve this by proposing neural network architectures for detecting hands and estimating hand keypoint locations. Our hand detection network robustly handles a variety of real world environments. The keypoint estimation network leverages tracking history to produce spatially and temporally consistent poses. We design scalable, semi-automated mechanisms to collect a large and diverse set of ground truth data using a combination of manual annotation and automated tracking. Additionally, we introduce a detection-by-tracking method that increases smoothness while reducing the computational cost; the optimized system runs at 60Hz on PC and 30Hz on a mobile processor. Together, these contributions yield a practical system for capturing a user's hands and is the default feature on the Oculus Quest VR headset powering input and social presence.

CCS Concepts: • **Computing methodologies → Computer vision**.

Additional Key Words and Phrases: motion capture, hand tracking, virtual reality

ACM Reference Format:

Shangchen Han, Beibei Liu, Randi Cabezas, Christopher D. Twigg, Peizhao Zhang, Jeff Petkau, Tsz-Ho Yu, Chun-Jung Tai, Muzaffer Akbay, Zheng Wang, Asaf Nitzan, Gang Dong, Yuting Ye, Lingling Tao, Chengde Wan, and Robert Wang. 2020. MEgATrack: Monochrome Egocentric Articulated Hand-Tracking for Virtual Reality. *ACM Trans. Graph.* 39, 4, Article 87 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392452

Authors' address: Shangchen Han, shchhan@oculus.com; Beibei Liu, beibeiliu@fb.com; Randi Cabezas, rcabezas@fb.com; Christopher D. Twigg, cdtwigg@fb.com; Peizhao Zhang, stzpz@fb.com; Jeff Petkau, jpet@fb.com; Tsz-Ho Yu, thyu@fb.com; Chun-Jung Tai, btai@fb.com; Muzaffer Akbay, muzoakbay@fb.com; Zheng Wang, wangz@fb.com; Asaf Nitzan, asafn@fb.com; Gang Dong, gang.dong@fb.com; Yuting Ye, yuting.ye@fb.com; Lingling Tao, linglingt@fb.com; Chengde Wan, vgrasp@fb.com; Robert Wang, rywang@fb.com, Facebook Reality Labs.

0730-0301/2020/7-ART87
https://doi.org/10.1145/3386569.3392452

1 INTRODUCTION

Recent improvements in VR/AR technology have led to the mainstream adoption of commercial headsets such as the Oculus Quest, Microsoft HoloLens and HTC Vive for applications including gaming, virtual training and socializing in virtual worlds. As a new computing platform, VR/AR is still experimenting with various input modalities, including mouse and keyboard, game controllers, 6 degree of freedom (DOF) motion controllers and wearable gloves. Vision-based hand-tracking can potentially provide more convenient and lower-friction input than these peripherals. For instance, users may not need to carry or charge an additional device or put on a wearable. However, to be a truly convenient input modality, hand-tracking must also be robust to environmental and user variations, support a generous working volume and produce responsive and precise (low-jitter) motions for targeting and selection. As VR/AR headsets are increasingly mobile, a hand-tracking input system must also run on a low-compute budget.

Most previous work on hand-tracking has focused on outside-in depth or RGB cameras. A depth camera provides hand geometry in terms of a 2.5D point cloud. However, depth cameras impose extra requirements on hardware design and power usage. In comparison, RGB cameras are easier to integrate and their utility continues to improve as deep learning techniques advance. As a result, predicting hand pose from a single RGB camera, typically with the help of a neural network, has become a popular research topic.

Despite continued progress, several remaining issues have held back RGB-based hand-tracking from being applied in VR/AR. First, predicting 3D hand pose from a single RGB camera is inherently

Sequential Gallery for Interactive Visual Design Optimization

YUKI KOYAMA, National Institute of Advanced Industrial Science and Technology (AIST), Japan
ISSEI SATO, The University of Tokyo, Japan
MASATAKA GOTO, National Institute of Advanced Industrial Science and Technology (AIST), Japan

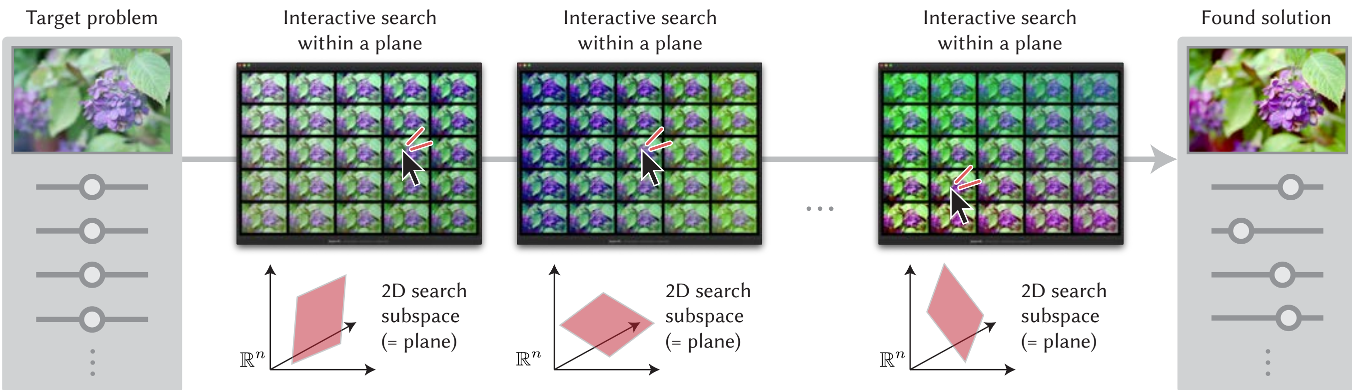


Fig. 1. ***Sequential Gallery*** is an interactive framework for exploring an n-dimensional design space formed by a set of n sliders and then finding an appropriate parameter set from that space. This framework lets the user sequentially select the most preferable option from the options displayed in a grid interface. To enable this framework, we propose a new *Bayesian optimization* method called *sequential plane search*, which decomposes the original high-dimensional search problem into a sequence of two-dimensional search (*i.e.*, plane-search) subtasks.

Visual design tasks often involve tuning many design parameters. For example, color grading of a photograph involves many parameters, some of which non-expert users might be unfamiliar with. We propose a novel user-in-the-loop optimization method that allows users to efficiently find an appropriate parameter set by exploring such a high-dimensional design space through much easier two-dimensional search subtasks. This method, called *sequential plane search*, is based on *Bayesian optimization* to keep necessary queries to users as few as possible. To help users respond to *plane-search* queries, we also propose using a gallery-based interface that provides options in the two-dimensional subspace arranged in an adaptive grid view. We call this interactive framework *Sequential Gallery* since users sequentially select the best option from the options provided by the interface. Our experiment with synthetic functions shows that our sequential plane search can find satisfactory solutions in fewer iterations than baselines. We also conducted a preliminary user study, results of which suggest that novices can effectively complete search tasks with Sequential Gallery in a photo-enhancement scenario.

Authors' addresses: Yuki Koyama, koyama.y@aist.go.jp, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki, Japan, 305-8568; Issei Sato, sato@k.u-tokyo.ac.jp, The University of Tokyo, Hongo 7-3-1, Bunkyo, Tokyo, Japan, 113-8654; Masataka Goto, m.goto@aist.go.jp, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki, Japan, 305-8568.

0730-0301/2020/7-ART88 $15.00
https://doi.org/10.1145/3386569.3392444

CCS Concepts: • **Computing methodologies** → **Computer graphics**; • **Human-centered computing** → *Human computer interaction (HCI)*.

Additional Key Words and Phrases: Visual design exploration, Bayesian optimization, human-in-the-loop optimization.

ACM Reference Format:
Yuki Koyama, Issei Sato, and Masataka Goto. 2020. Sequential Gallery for Interactive Visual Design Optimization. *ACM Trans. Graph.* 39, 4, Article 88 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392444

1 INTRODUCTION

Visual design tasks often involve many parameters that should be carefully adjusted via sliders. The purpose of tweaking these parameters is, for example, to reproduce the desired design in mind or to make the design as aesthetically pleasing as possible. This process is, however, often difficult because the parameters may affect the design in combination and the space of possible parameter configurations is very broad due to the high dimensionality. Moreover, evaluating a certain parameter configuration is also difficult without actually manipulating the slider values and seeing the corresponding visual representation, which thus requires many trials and errors. All this is especially true when users are unfamiliar with the design parameters. For example, photo retouch software has many sliders for color enhancement, including advanced ones such as "shadows (red)" and "highlights (red)" [Adobe 2017a; Instagram, Inc. 2019], which both affect shades of red but in different ways and can produce various effects in combination with other parameters. This complexity requires users to try many slider configurations at the beginning to understand what effects are possible and tweak

Tactile Line Drawings for Improved Shape Understanding in Blind and Visually Impaired Users

ATHINA PANOTOPOULOU, Dartmouth College and Boston University, USA
XIAOTING ZHANG and TAMMY QIU, Boston University, USA
XING-DONG YANG, Dartmouth College, USA
EMILY WHITING, Boston University, USA

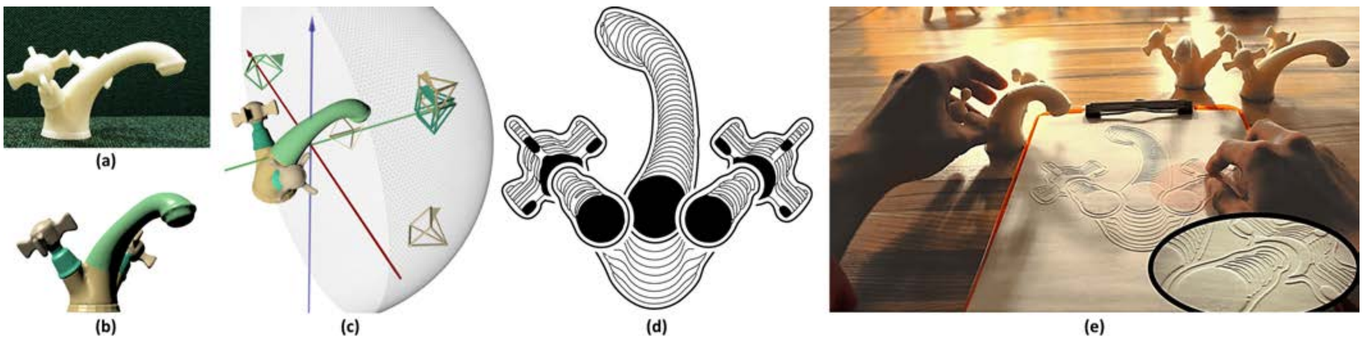


Fig. 1. We present a novel approach for generating tactile illustrations to improve shape understanding in blind individuals. (a) Physical 3D object (3D printed). (b) The input to our pipeline is a pre-partitioned object, colors indicate segmented parts. (c) A local camera is assigned to each part, with a master camera to combine the resulting multi-projection image. (d) Resulting stylized illustration. Cross-sections are used for texturing the interior of each part to communicate surface geometry. (e) We evaluated the technique in a user study with 20 blind participants. Tactile illustrations were fabricated using microcapsule paper.

Members of the blind and visually impaired community rely heavily on tactile illustrations – raised line graphics on paper that are felt by hand – to understand geometric ideas in school textbooks, depict a story in children's books, or conceptualize exhibits in museums. However, these illustrations often fail to achieve their goals, in large part due to the lack of understanding in how 3D shapes can be represented in 2D projections. This paper describes a new technique to design tactile illustrations considering the needs of blind individuals. Successful illustration design of 3D objects presupposes identification and combination of important information in topology and geometry. We propose a twofold approach to improve shape understanding. First, we introduce a part-based multi-projection rendering strategy to display geometric information of 3D shapes, making use of canonical viewpoints and removing reliance on traditional perspective projections. Second, curvature information is extracted from cross sections and embedded as textures in our illustrations.

CCS Concepts: • **Human-centered computing** → *Accessibility systems and tools*; • **Computing methodologies** → *Shape analysis*; **Perception**; **Non-photorealistic rendering**.

Authors' addresses: Athina Panotopoulou, athina.panotopoulou@dartmouth.edu, Dartmouth College and Boston University, USA; Xiaoting Zhang, xiaoting@bu.edu; Tammy Qiu, tqiu@bu.edu; Emily Whiting, whiting@bu.edu, Boston University, USA; Xing-Dong Yang, xing-dong.yang@dartmouth.edu, Dartmouth College, USA.

0730-0301/2020/7-ART89 $15.00
https://doi.org/10.1145/3386569.3392388

Additional Key Words and Phrases: fabrication, design, accessibility, tactile shape perception, tactile images, non-photorealistic rendering

ACM Reference Format:
Athina Panotopoulou, Xiaoting Zhang, Tammy Qiu, Xing-Dong Yang, and Emily Whiting. 2020. Tactile Line Drawings for Improved Shape Understanding in Blind and Visually Impaired Users. *ACM Trans. Graph.* 39, 4, Article 89 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392388

1 INTRODUCTION

Understanding the 3D geometry of everyday objects via 2D media is essential to live, learn, and work. Designers communicate the shape of a product (e.g., a chair) via sketches and renderings. Consumers understand the design by browsing images of the product in a catalogue. Students learn physics, such as why a plane is shaped in a certain way for aerodynamics, from figures in a textbook. Visual perception of complex 3D geometry from a 2D projection is taken for granted by people with healthy vision. However, for people with near or total blindness, understanding 3D geometry of daily objects from current media is extremely challenging.

In this paper we introduce a novel approach to generating *tactile line drawings* that aid 3D shape understanding in users with near or total blindness, enabling the blind community to perceive complex 3D objects from 2D media. Members of the blind community rely heavily on tactile illustrations, defined as raised graphics (e.g. on paper) that are felt by hand. Common uses of tactile illustrations are to make visual information accessible in textbooks, maps [Brock et al. 2015], scientific diagrams [Brown and Hurst 2012], children's literature [Claudet et al. 2008; Stangl et al. 2014; Stangl 2019], or

Tactile Rendering Based on Skin Stress Optimization

MICKEAL VERSCHOOR, DAN CASAS, and MIGUEL A. OTADUY, Universidad Rey Juan Carlos

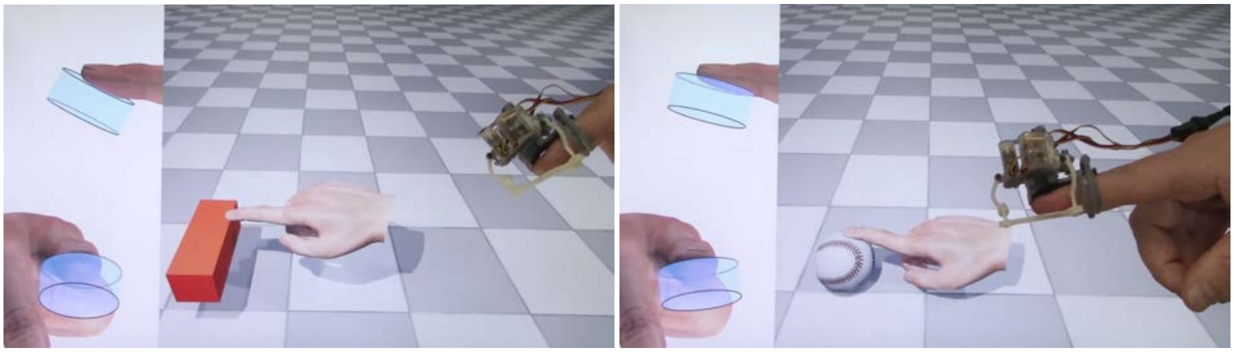


Fig. 1. Our tactile rendering method in action. A virtual hand follows the user and interacts with virtual objects. On each frame, we compute the tactile stimulus (i.e., skin stress) in this simulation, and use it to find the tactile device configuration that produces the best-matching stimulus (see insets). Then, we render this device configuration to the user.

We present a method to render virtual touch, such that the stimulus produced by a tactile device on a user's skin matches the stimulus computed in a virtual environment simulation. To achieve this, we solve the inverse mapping from skin stimulus to device configuration thanks to a novel optimization algorithm. Within this algorithm, we use a device-skin simulation model to estimate rendered stimuli, we account for trajectory-dependent effects efficiently by decoupling the computation of the friction state from the optimization of device configuration, and we accelerate computations using a neural-network approximation of the device-skin model. Altogether, we enable real-time tactile rendering of rich interactions including smooth rolling, but also contact with edges, or frictional stick-slip motion. We validate our algorithm both qualitatively through user experiments, and quantitatively on a BioTac biomimetic finger sensor.

CCS Concepts: • **Computing methodologies** → *Physical simulation*; *Virtual reality*.

Additional Key Words and Phrases: haptics, skin simulation, optimization methods

ACM Reference Format:
Mickeal Verschoor, Dan Casas, and Miguel A. Otaduy. 2020. Tactile Rendering Based on Skin Stress Optimization. *ACM Trans. Graph.* 39, 4, Article 90 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392398

1 INTRODUCTION

Tactile rendering stands for the computer-based generation of virtual touch sensations on the skin. Most works in haptic rendering simulate tool-based interaction using robotic devices, which produce kinesthetic stimulation of muscles, tendons and joints. In contrast, tactile rendering simulates direct touch interaction, by stimulating skin mechanoreceptors directly. Progress on different tactile actuation technologies in the last ten years has opened the door to virtual reality experiences where users touch virtual environments (VEs) directly with their hands [Otaduy et al. 2016].

Most research in tactile haptics has aimed at the design of actuation devices, with little attention to the design of rendering algorithms for VEs. In previous methods, tactile rendering is simplified by approximating VE interaction to some primitive with the same degrees of freedom (DoFs) as the device. However, tactile perception is a very high-dimensional spatiotemporal process, where the stimuli of mechanoreceptors are mapped to high-level percepts in a complex and yet largely unknown way. By limiting the interaction with VEs to simple primitives, previous tactile rendering methods limit the ability to explore the richness of tactile perception. To name some examples, with wearable thimbles, each finger pad is approximated as one 3D point, and the interaction of this point is directly programmed as a force vector or a surface orientation [Minamizawa et al. 2007; Prattichizzo et al. 2013; Solazzi et al. 2011]; active surfaces are programmed to reproduce predefined geometric shapes [Leithinger et al. 2015; Stanley and Okamura 2015]; and air jets or ultrasound haptics are programmed to produce touch sensations at target locations in space [Long et al. 2014; Sodhi et al. 2013].

We follow a different approach to tactile rendering. We simulate a realistic interaction in the VE, using a model of the user's hand and fingers, and thus we compute a rich target stimulus on each simulation frame. Once a target stimulus is computed, we solve an inverse mapping to the tactile device configuration, and we render this configuration to the user. We choose a mechanical characterization of tactile stimuli, using the stress distribution in the skin, and

Authors' address: Mickeal Verschoor, mickeal.verschoor@gmail.com; Dan Casas, dan.casas@urjc.es; Miguel A. Otaduy, miguel.otaduy@urjc.es, Universidad Rey Juan Carlos.

0730-0301/2020/7-ART90 $15.00
https://doi.org/10.1145/3386569.3392398

The Eyes Have It: An Integrated Eye and Face Model for Photorealistic Facial Animation

GABRIEL SCHWARTZ, Facebook Reality Labs
SHIH-EN WEI, Facebook Reality Labs
TE-LI WANG, Facebook Reality Labs
STEPHEN LOMBARDI, Facebook Reality Labs
TOMAS SIMON, Facebook Reality Labs
JASON SARAGIH, Facebook Reality Labs
YASER SHEIKH, Facebook Reality Labs

Fig. 1. **Avatars rendered and driven by our system.** These images highlight the quality of the renders produced by our system. On the left, we show renders of the learned avatars, and on the right, a view of the avatar as driven in our real-time system.

Interacting with people across large distances is important for remote work, interpersonal relationships, and entertainment. While such face-to-face interactions can be achieved using 2D video conferencing or, more recently, virtual reality (VR), telepresence systems currently distort the communication of eye contact and social gaze signals. Although methods have been proposed to redirect gaze in 2D teleconferencing situations to enable eye contact, 2D video conferencing lacks the 3D immersion of real life. To address these problems, we develop a system for face-to-face interaction in VR that focuses on reproducing photorealistic gaze and eye contact. To do this, we create a 3D virtual avatar model that can be animated by cameras mounted on a VR headset to accurately track and reproduce human gaze in VR. Our primary contributions in this work are a jointly-learnable 3D face and eyeball model that better represents gaze direction and upper facial expressions, a method for disentangling the gaze of the left and right eyes from each other and the rest of the face allowing the model to represent entirely unseen combinations of gaze and expression, and a gaze-aware model for precise animation from headset-mounted cameras. Our quantitative experiments show that our method results in higher reconstruction quality, and qualitative results show our method gives a greatly improved sense of presence for VR avatars.

Authors' addresses: Gabriel Schwartz, Facebook Reality Labs, gbschwartz@fb.com; Shih-En Wei, Facebook Reality Labs, swei@fb.com; Te-Li Wang, Facebook Reality Labs, teli@fb.com; Stephen Lombardi, Facebook Reality Labs, stephen.lombardi@fb.com; Tomas Simon, Facebook Reality Labs, tomas.simon@fb.com; Jason Saragih, Facebook Reality Labs, jason.saragih@fb.com; Yaser Sheikh, Facebook Reality Labs, yasers@fb.com.

0730-0301/2020/7-ART91
https://doi.org/10.1145/3386569.3392493

CCS Concepts: • **Computing methodologies → Neural networks**; **Virtual reality**.

Additional Key Words and Phrases: Differentiable Rendering, Eye Modeling

ACM Reference Format:
Gabriel Schwartz, Shih-En Wei, Te-Li Wang, Stephen Lombardi, Tomas Simon, Jason Saragih, and Yaser Sheikh. 2020. The Eyes Have It: An Integrated Eye and Face Model for Photorealistic Facial Animation. *ACM Trans. Graph.* 39, 4, Article 91 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392493

1 INTRODUCTION

Eye contact is a strong and important social signal [Chen 2002], and humans can accurately estimate where someone's eyes are pointing just by looking at them [Cline 1967; Gibson and Pick 1963]. Recent efforts in image-space gaze-correction [Kononenko et al. 2018; Wolf et al. 2010] demonstrate the importance of achieving eye-contact in a telepresence application, but such methods are motivated by the

CNNs on Surfaces using Rotation-Equivariant Features

RUBEN WIERSMA, ELMAR EISEMANN, and KLAUS HILDEBRANDT, Delft University of Technology

This paper is concerned with a fundamental problem in geometric deep learning that arises in the construction of convolutional neural networks on surfaces. Due to curvature, the transport of filter kernels on surfaces results in a rotational ambiguity, which prevents a uniform alignment of these kernels on the surface. We propose a network architecture for surfaces that consists of vector-valued, rotation-equivariant features. The equivariance property makes it possible to locally align features, which were computed in arbitrary coordinate systems, when aggregating features in a convolution layer. The resulting network is agnostic to the choices of coordinate systems for the tangent spaces on the surface. We implement our approach for triangle meshes. Based on circular harmonic functions, we introduce convolution filters for meshes that are rotation-equivariant at the discrete level. We evaluate the resulting networks on shape correspondence and shape classifications tasks and compare their performance to other approaches.

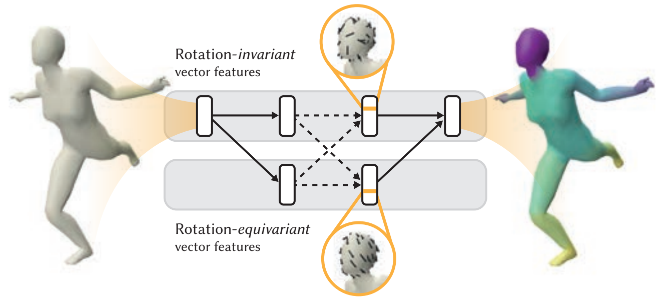


Fig. 1. We propose CNNs on surfaces that operate on vectors and separate rotation-equivariant and rotation-invariant features.

CCS Concepts: • **Computing methodologies → Neural networks**; **Shape analysis**.

Additional Key Words and Phrases: Geometric Deep Learning, CNNs on Surfaces, Surface Networks, Rotation-Equivariance, Circular Harmonic Filters, Shape Classification, Shape Segmentation, Shape Correspondence

ACM Reference Format:
Ruben Wiersma, Elmar Eisemann, and Klaus Hildebrandt. 2020. CNNs on Surfaces using Rotation-Equivariant Features. *ACM Trans. Graph.* 39, 4, Article 92 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392437

1 INTRODUCTION

The success of Deep Learning approaches based on convolutional neural networks (CNNs) in computer vision and image processing has motivated the development of analogous approaches for the analysis, processing, and synthesis of surfaces. Along these lines, approaches have been proposed for problems such as shape recognition [Su et al. 2015], shape matching [Boscaini et al. 2016], shape segmentation [Maron et al. 2017], shape completion [Litany et al. 2018], curvature estimation [Guerrero et al. 2018], and 3D-face synthesis [Ranjan et al. 2018].

In contrast to images, which are described by regular grids in a Euclidean domain, surfaces are curved manifolds and the grids on these surfaces are irregular. In order to still use regular grids, one can work with multiple projections of the surface on planes [Su et al. 2015] or with volumetric grids [Wu et al. 2015].

An alternative to learning on regular grids is generalized deep learning, often referred to as geometric deep learning [Bronstein et al. 2017], which targets irregularly sampled manifolds and graphs. A central element of such geometric CNNs is a generalized convolution operator. For CNNs on images, the convolution layers are built from convolution kernels, which are transported across the image. As a result, the parameters that define one kernel describe the convolution across the whole image, which significantly reduces the number of parameters that need to be learned. This is a motivation for exploring constructions of generalized convolution operators on surfaces based on convolution kernels.

To apply a convolution kernel defined on $\mathbb{R}^2$ to a function at a point on a surface, the Riemannian exponential map is used to locally lift the function from the surface to a function defined on the tangent plane at the point. By identifying the tangent plane with $\mathbb{R}^2$, the convolution of the kernel and the lifted function can be computed. In this way, the convolution kernel can be applied everywhere on the surface. However, a problem arises, since there is a rotational degree of freedom when $\mathbb{R}^2$ is identified with a tangent plane. Moreover, the transport of filters on a surface depends on the chosen path. If a filter is transported along two different ways from one point of a surface to another, the transported filters are rotated against each other. This *rotation ambiguity* problem is fundamental and caused by the curvature of the surface.

The rotation ambiguity problem can be addressed by specifying a coordinate system at each point of the surface, *e.g.* according to the principal curvature directions [Boscaini et al. 2016; Monti et al. 2017; Pan et al. 2018] or the direction of maximum activation [Masci et al. 2015; Sun et al. 2018]. As a consequence, however, the coordinate systems in the local neighborhoods are arranged in different patterns for each point. For a network this means that, when features are aggregated to form the next layer of the network, the features are not only dependent on the sequence of convolution kernels that are applied, but also on the arrangement of coordinate systems in the local neighborhoods. Loosely speaking, the information contained in the features in the neighborhood of a point can be arbitrarily rotated against the coordinate system at the point. One can think of this as in a cubist painting, where the elements that make up

Authors' address: Ruben Wiersma, r.t.wiersma@tudelft.nl; Elmar Eisemann, e.eisemann@tudelft.nl; Klaus Hildebrandt, k.a.hildebrandt@tudelft.nl, Delft University of Technology.

0730-0301/2020/7-ART92 $15.00
https://doi.org/10.1145/3386569.3392437

Code Replicability in Computer Graphics

NICOLAS BONNEEL, Univ Lyon, CNRS, France
DAVID COEURJOLLY, Univ Lyon, CNRS, France
JULIE DIGNE, Univ Lyon, CNRS, France
NICOLAS MELLADO, Univ Toulouse, CNRS, France

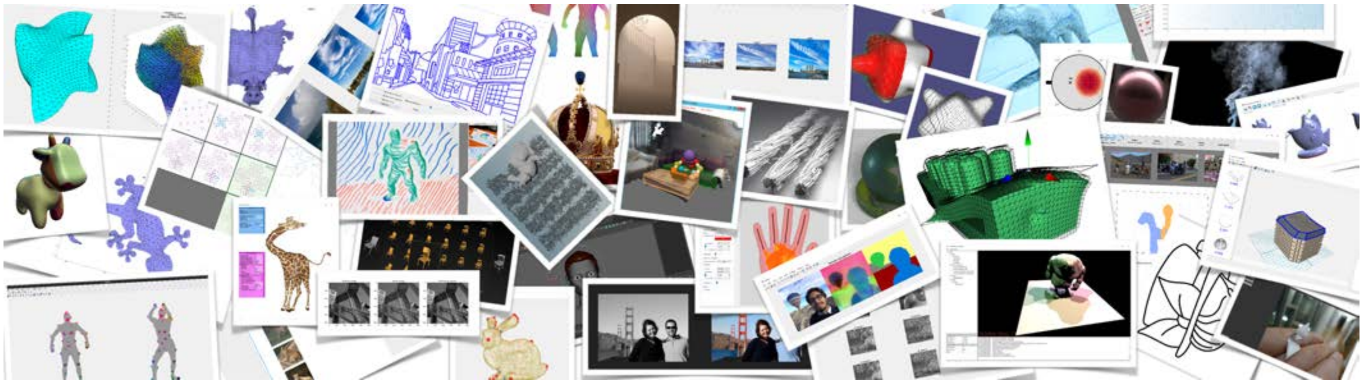


Fig. 1. We ran 151 codes provided by papers published at SIGGRAPH 2014, 2016 and 2018. We analyzed whether these codes could still be run as of 2020 to provide a replicability score, and performed statistical analysis on code sharing. Image credits: Umberto Salvagnin, _Bluenose Girl, Dimitry B., motiqua, Ernest McGray Jr., Yagiz Aksoy, Hillebrand Steve. 3D models by Martin Lubich and Wig42.

Being able to duplicate published research results is an important process of conducting research whether to build upon these findings or to compare with them. This process is called "replicability" when using the original authors' artifacts (e.g., code), or "reproducibility" otherwise (e.g., re-implementing algorithms). Reproducibility and replicability of research results have gained a lot of interest recently with assessment studies being led in various fields, and they are often seen as a trigger for better result diffusion and transparency. In this work, we assess replicability in Computer Graphics, by evaluating whether the code is available and whether it works properly. As a proxy for this field we compiled, ran and analyzed 151 codes out of 374 papers from 2014, 2016 and 2018 SIGGRAPH conferences. This analysis shows a clear increase in the number of papers with available and operational research codes with a dependency on the subfields, and indicates a correlation between code replicability and citation count. We further provide an interactive tool to explore our results and evaluation data.

CCS Concepts: • **Computing methodologies → Computer graphics**; • **Software and its engineering → Open source model**.

Additional Key Words and Phrases: Replicability, reproducibility, open source, code review, siggraph

Authors' addresses: Nicolas Bonneel, nicolas.bonneel@liris.cnrs.fr, Univ Lyon, CNRS, Lyon, France; David Coeurjolly, david.coeurjolly@liris.cnrs.fr, Univ Lyon, CNRS, Lyon, France; Julie Digne, julie.digne@liris.cnrs.fr, Univ Lyon, CNRS, Lyon, France; Nicolas Mellado, nicolas.mellado@irit.fr, Univ Toulouse, CNRS, Toulouse, France.

0730-0301/2020/7-ART93 $15.00
https://doi.org/10.1145/3386569.3392413

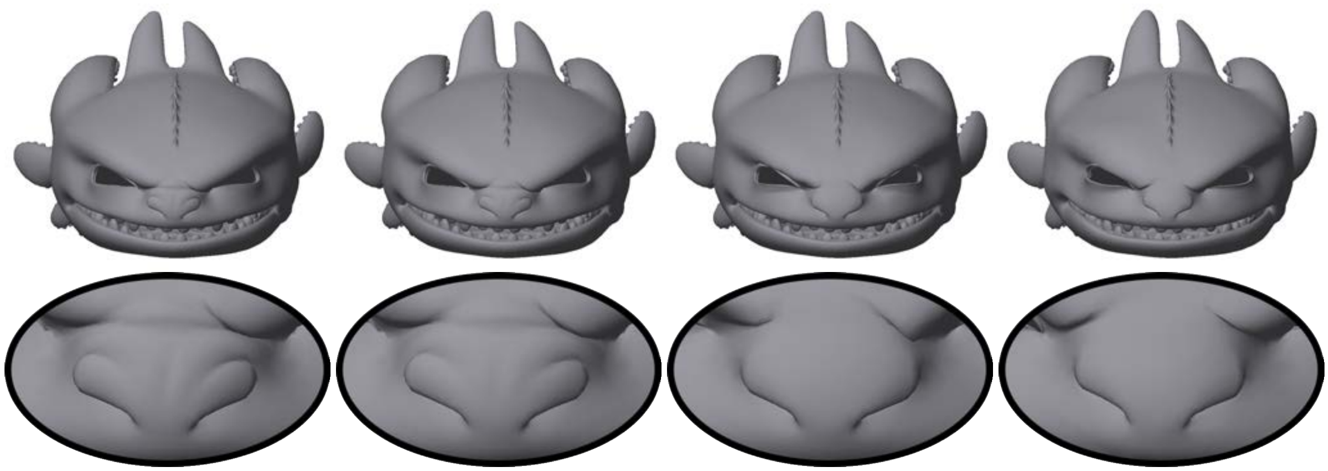
1 INTRODUCTION

The ability to reproduce an experiment and validate its results is a cornerstone of scientific research, a key to our understanding of the world. Scientific advances often provide useful tools, and build upon a vast body of previous work published in the literature. As such, research that cannot be reproduced by peers despite best efforts often has limited value, and thus impact, as it does not benefit to others, cannot be used as a basis for further research, and casts doubt on published results. Reproducibility is also important for comparison purposes since new methods are often seen in the light of results obtained by published competing approaches. Recently serious concerns have emerged in various scientific communities from psychological sciences [Open Science Collaboration et al. 2015] to artificial intelligence [Hutson 2018] over the lack of reproducibility, and one could wonder about the state of computer graphics research in this matter.

In the recent trend of open science and reproducible research, this paper aims at assessing the state of replicability of papers published at ACM Transactions on Graphics as part of SIGGRAPH conferences. Contrary to reproducibility which assesses how results can be obtained by independently reimplementing published papers – an overwhelming task given the hundred papers accepted yearly to this event – replicability ensures the authors' own codes run and produce the published results. While sharing code is not the only available option to guarantee that published results can be duplicated by a practitioner – after all, many contributions can be reimplemented from published equations or algorithm descriptions with more or less effort – it remains an important tool that reduces the time spent in reimplementation, in particular as computer graphics algorithms get more sophisticated.

Fast and Deep Facial Deformations

STEPHEN W. BAILEY, University of California, Berkeley and Unity Technologies
DALTON OMENS, University of California, Berkeley
PAUL DILORENZO, DreamWorks Animation
JAMES F. O'BRIEN, University of California, Berkeley



(a) Ground Truth (b) Refined Approximation (c) Coarse Approximation (d) LBS Approximation

Fig. 1. Side-by-side comparison of facial mesh deformations using our coarse and refined approximations as well as an approximation generated by linear blend skinning. The most noticeable difference, shown on the second row, is observed around the nasal region of the mesh.

Film-quality characters typically display highly complex and expressive facial deformation. The underlying rigs used to animate the deformations of a character's face are often computationally expensive, requiring high-end hardware to deform the mesh at interactive rates. In this paper, we present a method using convolutional neural networks for approximating the mesh deformations of characters' faces. For the models we tested, our approximation runs up to 17 times faster than the original facial rig while still maintaining a high level of fidelity to the original rig. We also propose an extension to the approximation for handling high-frequency deformations such as fine skin wrinkles. While the implementation of the original animation rig depends on an extensive set of proprietary libraries making it difficult to install outside of an in-house development environment, our fast approximation relies on the widely available and easily deployed TensorFlow libraries. In addition to allowing high frame rate evaluation on modest hardware and in a wide range of computing environments, the large speed increase also enables interactive inverse kinematics on the animation rig. We demonstrate our approach and its applicability through interactive character posing and real-time facial performance capture.

CCS Concepts: • **Computing methodologies → Neural networks**; **Shape modeling**; **Animation**.

Additional Key Words and Phrases: character rig, facial animation, mesh deformations, deep learning, function approximation

ACM Reference Format:
Stephen W. Bailey, Dalton Omens, Paul DiLorenzo, and James F. O'Brien. 2020. Fast and Deep Facial Deformations. *ACM Trans. Graph.* 39, 4, Article 94 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392397

Authors' addresses: Stephen W. Bailey, University of California, Berkeley, Unity Technologies; Dalton Omens, University of California, Berkeley; Paul DiLorenzo, DreamWorks Animation; James F. O'Brien, University of California, Berkeley.

0730-0301/2020/7-ART94 $15.00
https://doi.org/10.1145/3386569.3392397

1 INTRODUCTION

Character facial rigs for video games and other real-time applications are often controlled by sets of bones or blendshapes. Although these rigging methods can be computed quickly, they generally sacrifice fine-scale details for speed. Expressing nuanced deformations with these real-time rigs is challenging and often requires additional computational layers added to the underlying rig. Some such additions for increasing the level of detail in the mesh deformation include pose space deformations [Lewis et al. 2000] and wrinkle maps. However, despite these improvements, the level of detail in film-quality facial rigs is noticeably better when compared with real-time rigs. The primary reason why film-quality facial rigs contain more sophisticated mesh deformations is because they are not constrained by real-time requirements.

MichiGAN: Multi-Input-Conditioned Hair Image Generation for Portrait Editing

ZHENTAO TAN, University of Science and Technology of China
MENGLEI CHAI, Snap Inc.
DONGDONG CHEN, Microsoft Cloud AI
JING LIAO, City University of Hong Kong
QI CHU*, University of Science and Technology of China
LU YUAN, Microsoft Cloud AI
SERGEY TULYAKOV, Snap Inc.
NENGHAI YU, University of Science and Technology of China

Fig. 1. Given an input portrait image (a), MichiGAN is capable of enabling multiple input conditions for disentangled hair manipulation. It not only supports editing one attribute individually including appearance (b), structure (c), and shape (d) while keeping the background unchanged, but also allows manipulating multiple attributes (f&g) jointly. The users can also leverage our interactive system to create more customized hair manipulation results (h). (e) is the reconstruction result when all the condition inputs are from (a) itself. Original images courtesy of stephen davis, Eric E Castro, LeafBlue, Pawel Loj, O'Reilly Internal, Lydia Liu, Faculty of Social and Educational Sciences, NWABR, and Zero Emission Resource Organisation.

Despite the recent success of face image generation with GANs, conditional hair editing remains challenging due to the under-explored complexity of its geometry and appearance. In this paper, we present MichiGAN (Multi-Input-Conditioned Hair Image GAN), a novel conditional image generation method for interactive portrait hair manipulation. To provide user control over every major hair visual factor, we explicitly disentangle hair into four orthogonal attributes, including shape, structure, appearance, and background. For each of them, we design a corresponding condition module to represent, process, and convert user inputs, and modulate the image generation pipeline in ways that respect the natures of different visual attributes. All these condition modules are integrated with the backbone generator to form the final end-to-end network, which allows fully-conditioned hair generation from multiple user inputs. Upon it, we also build an interactive portrait hair editing system that enables straightforward manipulation of hair by projecting intuitive and high-level user inputs such as painted masks, guiding strokes, or reference photos to well-defined condition representations. Through extensive experiments and evaluations, we demonstrate the superiority of our method regarding both result quality and user controllability.

*Corresponding author.

Authors' addresses: Zhentao Tan, University of Science and Technology of China, tzt@mail.ustc.edu.cn; Menglei Chai, Snap Inc., mchai@snap.com; Dongdong Chen, Microsoft Cloud AI, cddlyf@gmail.com; Jing Liao, City University of Hong Kong, jingliao@cityu.edu.hk; Qi Chu, University of Science and Technology of China, qchu@ustc.edu.cn; Lu Yuan, Microsoft Cloud AI, luyuan@microsoft.com; Sergey Tulyakov, Snap Inc., stulyakov@snap.com; Nenghai Yu, University of Science and Technology of China, ynh@ustc.edu.cn.

0730-0301/2020/7-ART95 $15.00
https://doi.org/10.1145/3386569.3392488

CCS Concepts: • **Computing methodologies → Image manipulation**; **Neural networks**; *Computer vision representations*; *Graphics systems and interfaces.*

Additional Key Words and Phrases: interactive portrait editing; conditional hair image generation; generative adversarial networks

ACM Reference Format:
Zhentao Tan, Menglei Chai, Dongdong Chen, Jing Liao, Qi Chu, Lu Yuan, Sergey Tulyakov, and Nenghai Yu. 2020. MichiGAN: Multi-Input-Conditioned Hair Image Generation for Portrait Editing. *ACM Trans. Graph.* 39, 4, Article 95 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392488

1 INTRODUCTION

Human hair is so delicate, variable, and expressive that it constantly plays a unique role in depicting the subject in a face image. Given its diversity and flexibility, the urge to manipulate hair in a portrait

NASOQ: Numerically Accurate Sparsity-Oriented QP Solver

KAZEM CHESHMI, University of Toronto, Canada
DANNY M. KAUFMAN, Adobe Research, USA
SHOAIB KAMIL, Adobe Research, USA
MARYAM MEHRI DEHNAVI, University of Toronto, Canada

Quadratic programs (QP), minimizations of quadratic objectives subject to linear inequality and equality constraints, are at the heart of algorithms across scientific domains. Applications include fundamental tasks in geometry processing, simulation, engineering, animation and finance where the accurate, reliable, efficient, and scalable solution of QP problems is critical. However, available QP algorithms generally provide either accuracy or scalability – but not both. Some algorithms reliably solve QP problems to high accuracy but work only for smaller-scale QP problems due to their reliance on dense matrix methods. Alternately, many other QP solvers scale well via sparse, efficient algorithms but cannot reliably deliver solutions at requested accuracies. Towards addressing the need for accurate *and* efficient QP solvers at scale, we develop NASOQ, a new, full-space QP algorithm that provides accurate, efficient, and scalable solutions for QP problems. To enable NASOQ we construct a new row modification method and fast implementation of LDL factorization for indefinite systems. Together they enable efficient updates and accurate solutions of the iteratively modified KKT systems required for accurate QP solves. While QP methods have been previously tested on large synthetic benchmarks, to test and compare NASOQ's suitability for real-world applications we collect here a new benchmark set comprising a wide range of graphics-related QPs across physical simulation, animation, and geometry processing tasks. We combine these problems with numerous pre-existing stress-test QP benchmarks to form, to our knowledge, the largest-scale test set of application-based QP problems currently available. Building off of our base NASOQ solver we then develop and test two NASOQ variants against best, state-of-the-art available QP libraries – both commercial and open-source. Our two NASOQ-based methods each solve respectively 98.8% and 99.5% of problems across a range of requested accuracies from 10^{-3} to 10^{-9} with average speedups ranging from 1.7× to 24.8× over fastest competing methods.

CCS Concepts: • **Mathematics of computing → Quadratic programming**; **Solvers**; **Computations on matrices**; • **Computing methodologies** → *Shape modeling*.

Additional Key Words and Phrases: sparse row modification, indefinite factorization, sparse linear algebra, quadratic programming, numerical optimization, contact simulation, mesh deformation

Authors' addresses: Kazem Cheshmi, kazem@cs.toronto.edu, University of Toronto, Toronto, Ontario, Canada; Danny M. Kaufman, kaufman@adobe.com, Adobe Research, Seattle, Washington, USA; Shoaib Kamil, kamil@adobe.com, Adobe Research, New York, New York, USA; Maryam Mehri Dehnavi, mmehride@cs.toronto.edu, University of Toronto, Toronto, Ontario, Canada.

0730-0301/2020/7-ART96 $15.00
https://doi.org/10.1145/3386569.3392486

ACM Reference Format:
Kazem Cheshmi, Danny M. Kaufman, Shoaib Kamil, and Maryam Mehri Dehnavi. 2020. NASOQ: Numerically Accurate Sparsity-Oriented QP Solver. *ACM Trans. Graph.* 39, 4, Article 96 (July 2020), 17 pages. https://doi.org/10.1145/3386569.3392486

1 INTRODUCTION

Solving a quadratic program (QP) is a core numerical task critical in domains spanning geometry processing [Dvorožňák et al. 2018; Sýkora et al. 2014; Zhu et al. 2018], animation [Jacobson et al. 2011; Righetti and Schaal 2012], physical simulation [Barbic 2007; Erleben 2013], robotics [Pandala et al. 2019], machine learning [Agrawal et al. 2019; Amos and Kolter 2017], engineering, and design [Fesanghary et al. 2008]. Unfortunately, available QP solvers are often neither accurate nor robust enough for many applications [Kaufman et al. 2008; Smith et al. 2012; Yao et al. 2017; Zheng and James 2011; Zhu et al. 2018], necessitating heuristics, approximations and/or multiple failsafe backups to succeed.

A long-standing challenge then has been to provide a single, unified QP solver that is 1) accurate, 2) efficient, and 3) scalable. By accurate we mean that the QP solver converges to all reasonable requested accuracies; by efficient we mean that it converges rapidly in wall-clock time; and by scalable we mean that it efficiently converges across both large- and small-sized QP instances. As we show in Section 6, available QP solver libraries generally succeed for some subsets of QPs, while often failing or becoming impractically slow to achieve success for others. To make matters worse, in many cases, given the algorithms employed, it is not possible to predict in advance when a QP method will succeed or fail per QP problem instance [Zheng and James 2011].

The key challenge for solving a QP is in identifying an *active set* [Fletcher 2013]. An active set is a subset of a QP's linear inequality constraints that are treated as equalities at optimality. All other inequalities can then effectively be safely ignored. If an active set is found, a QP problem instance then reduces to solving a much easier QP subject to just its active constraints set as equalities.

Algorithms for solving large-scale QPs generally treat the entire constraint set as approximately "active" with barrier terms penalizing all constraint violations simultaneously. This allows the application of large-scale, general-purpose sparse linear solvers, but generally comes at the cost of uncertainty in the active set and degraded solution accuracy. On the other hand, to address accuracy, many other QP algorithms employ *active-set methods*. These are a range of methods that iteratively explore and test active-set proposals. Details vary across methods but in all cases each iteration requires solving large numbers of reduced QPs. Each reduced QP is solved subject to a different set of proposed active constraints

Nonlinear Color Triads for Approximation, Learning and Direct Manipulation of Color Distributions

MARIA SHUGRINA, University of Toronto, Vector Institute, and NVIDIA
AMLAN KAR, University of Toronto, Vector Institute, and NVIDIA
SANJA FIDLER, University of Toronto, Vector Institute, and NVIDIA
KARAN SINGH, University of Toronto

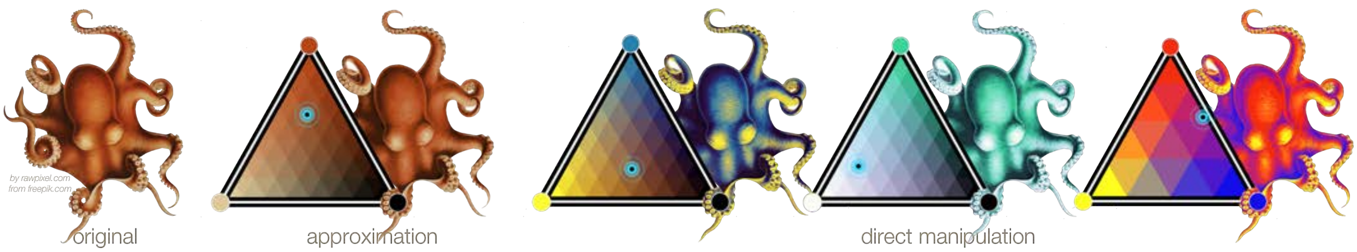


Fig. 1. Color triad is designed to approximate color distributions of shaded objects, enabling direct manipulation of colors in an image and other applications. Original image by rawpixel.com from freepik.com available at https://www.freepik.com/free-vector/octopus-vintage-style_3770655.htm.

We present nonlinear color triads, an extension of color gradients able to approximate a variety of natural color distributions that have no standard interactive representation. We derive a method to fit this compact parametric representation to existing images and show its power for tasks such as image editing and compression. Our color triad formulation can also be included in standard deep learning architectures, facilitating further research.

CCS Concepts: • **Computing methodologies** → **Image manipulation**.

Additional Key Words and Phrases: color palettes, recoloring, neural networks, deep learning, interactive techniques.

ACM Reference Format:
Maria Shugrina, Amlan Kar, Sanja Fidler, and Karan Singh. 2020. Nonlinear Color Triads for Approximation, Learning and Direct Manipulation of Color Distributions. *ACM Trans. Graph.* 39, 4, Article 97 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392461

1 INTRODUCTION

Existing color theme representations are either restricted to a few color swatches or are fully unconstrained. In this work, we show the power of a color representation that strikes a careful balance between expressiveness and structure. We propose non-linear color triads, which can both approximate a wide array of color distributions and naturally lend themselves to many applications.

Authors' addresses: Maria Shugrina, University of Toronto, Vector Institute, NVIDIA, shumash@cs.toronto.edu; Amlan Kar, University of Toronto, Vector Institute, NVIDIA, amlan@cs.toronto.edu; Sanja Fidler, University of Toronto, Vector Institute, NVIDIA, fidler@cs.toronto.edu; Karan Singh, University of Toronto, karan@dgp.toronto.edu.

0730-0301/2020/7-ART97 $15.00
https://doi.org/10.1145/3386569.3392461

The simplicity of our representation is inspired by discrete color palettes, typically comprised of 5-10 independent colors. Despite their limited representative power, simplicity and the ease of construction have made discrete palettes a popular tool for artists and researchers alike, and they have found use in recoloring, color theme extraction and machine learning. While discrete palettes may be suitable for some domains, they are a poor fit for larger, continuous color distributions found in most art and design. Palettes modeling more complex color distributions [Meier et al. 2004; Nguyen et al. 2015; Shugrina et al. 2017, 2019] focus on supporting the creation of novel artwork and have not been applied to representing, analyzing and editing existing images, perhaps due to their freeform unstructured nature. Color triads combine versatility and representational power with simplicity, enabling new applications.

A color triad is simply a triangular patch of RGB space governed by five parameters: three colors defining the linear interpolation space, and the fourth parameter setting a constrained amount of nonlinearity. This non-linearity is critical for modeling an array of color distributions and blending behaviors, while constraints prevent degeneracies during user editing. An additional fifth parameter defines the level of discretization, making it possible to model both coarsely and densely sampled distributions. We explain the intuition behind this simple interactive representation, and demonstrate a number of useful applications. Our contributions are:

- analysis and intuition for color distribution modeling (§3)
- nonlinear color triad formulation (§4)
- algorithm for fitting color triads to images (§5)
- application of color triads to interactive editing of images (§6)
- demonstration of color triad formulation incorporated into a fully differentiable deep learning architecture (§7)

We also sketch out potential applications of our model to image compression and paint pigment modeling in the Supplemental Material. We evaluate our model quantitatively to show its representative power and qualitatively with a user study (§8).

RoboCut: Hot-wire Cutting with Robot-controlled Flexible Rods

SIMON DUENSER, ETH Zurich
ROI PORANNE, ETH Zurich and University of Haifa
BERNHARD THOMASZEWSKI, Université de Montréal and ETH Zurich
STELIAN COROS, ETH Zurich

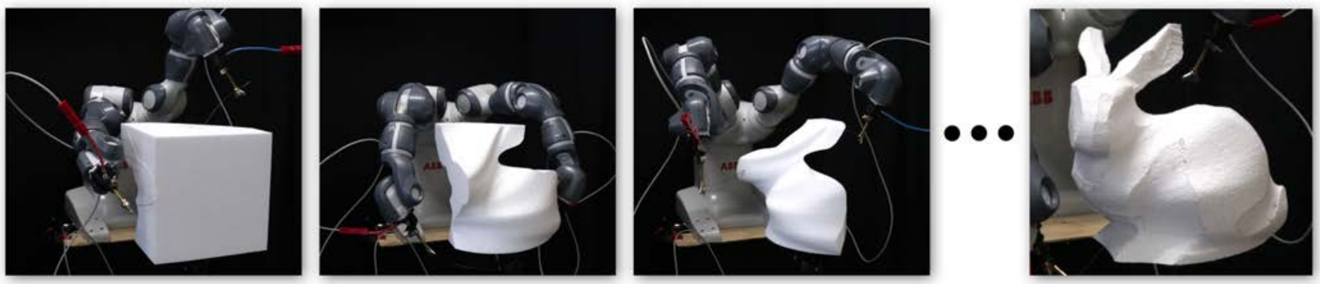


Fig. 1. Our method enables robotically-controlled hot wire cutting of complex shapes through tight integration of physical simulation, surface approximation, and path planning. Our algorithm anticipates and controls the deformations of the cutting rod to maximize the efficiency of every sweep, allowing this bunny shape to emerge after only 2 cuts, and be completed after 10.

Hot-wire cutting is a subtractive fabrication technique used to carve foam and similar materials. Conventional machines rely on straight wires and are thus limited to creating piecewise ruled surfaces. In this work, we propose a method that exploits a dual-arm robot setup to actively control the shape of a flexible, heated rod as it cuts through the material. While this setting offers great freedom of shape, using it effectively requires concurrent reasoning about three tightly coupled sub-problems: 1) modeling the way in which the shape of the rod and the surface it sweeps are governed by the robot's motions; 2) approximating a target shape through a sequence of surfaces swept by the equilibrium shape of an elastic rod; and 3) generating collision-free motion trajectories that lead the robot to create desired sweeps with the deformable tool. We present a computational framework for robotic hot wire cutting that addresses all three sub-problems in a unified manner. We evaluate our approach on a set of simulated results and physical artefacts generated with our robotic fabrication system.

CCS Concepts: • **Computer graphics → Computational geometry and object modeling**; *Physically based modeling*;

Additional Key Words and Phrases: Computer graphics, robotics, fabrication, sensitivity analysis

ACM Reference Format:
Simon Duenser, Roi Poranne, Bernhard Thomaszewski, and Stelian Coros. 2020. RoboCut: Hot-wire Cutting with Robot-controlled Flexible Rods. *ACM Trans. Graph.* 39, 4, Article 98 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392465

Authors' addresses: Simon Duenser, ETH Zurich, simon.duenser@inf.ethz.ch; Roi Poranne, ETH Zurich , University of Haifa; Bernhard Thomaszewski, Université de Montréal , ETH Zurich; Stelian Coros, ETH Zurich.

1 INTRODUCTION

The ability to use tools is a hallmark of intelligence and it has profoundly shaped the evolution of human culture. For example, carving—the process of sculpting away material from a workpiece in order to reveal a desired artefact—has been used to create art forms and functional objects since several thousand years ago. Carving demands skillful manipulation of various tools and intuitive understanding of the physical interactions between the tools and the objects they are applied to. Fueled by technological advances, the craft of carving has grown into CNC milling and cutting processes that now enjoy widespread use. With the continued evolution of carving techniques as a long-term goal, in this paper we present a computational framework for robotic hot-wire cutting.

A hot-wire cutter is a tool used to carve polystyrene foam and other materials that melt or vaporize when subjected to a source of intense heat. The tool consists of a thin metal wire that is typically held under tension using a bow. When attached to an electrical power source, the wire heats up past the melting temperature of the material. The cutter can therefore create thin cuts through the workpiece without the need to engage in any physical contact.

Hot wire cutters can be easily mounted on robotic platforms, and together with milling tools, they are routinely employed in machine shops. As they are equipped with a *straight* wire, each cut generated by a typical hot wire cutter is a *ruled surface*. Generating efficient toolpaths for hot-wire cutting is thus a matter of approximating a desired shape using piecewise ruled surfaces—a topic of intense ongoing research. Nevertheless, finding a high quality approximation is only part of the challenge; another crucial problem in generating *feasible* toolpath trajectories is to ensure that the work-space of the fabrication machine and collision avoidance constraints are taken into account—a point we will return to shortly.

Sliced Optimal Transport Sampling

LOIS PAULIN, Univ Lyon, CNRS
NICOLAS BONNEEL, Univ Lyon, CNRS
DAVID COEURJOLLY, Univ Lyon, CNRS
JEAN-CLAUDE IEHL, Univ Lyon, CNRS
ANTOINE WEBANCK, Univ Lyon, CNRS
MATHIEU DESBRUN, ShanghaiTech/Caltech
VICTOR OSTROMOUKHOV, Univ Lyon, CNRS

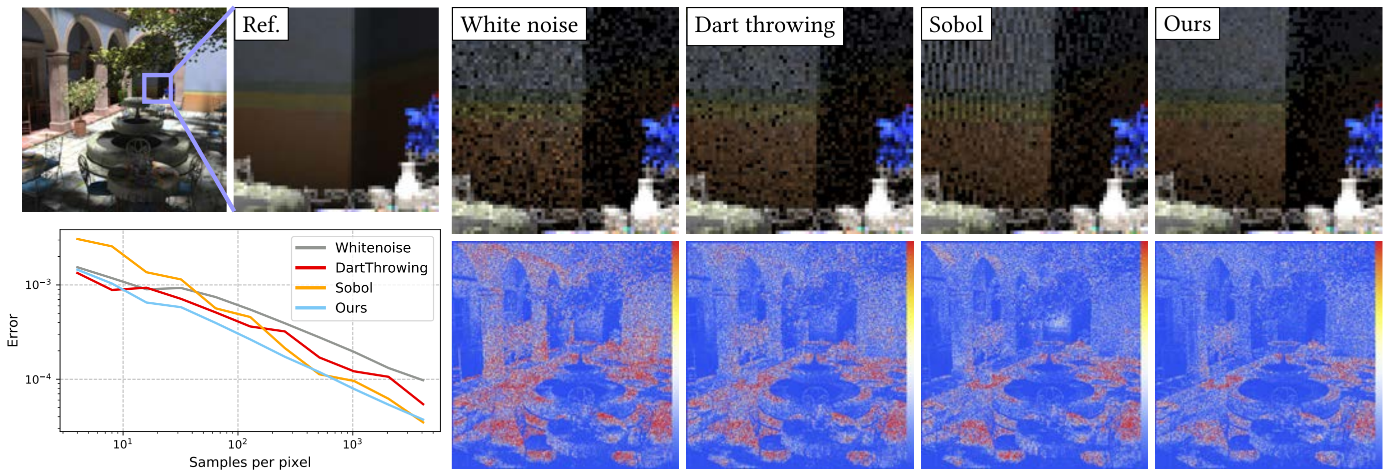


Fig. 1. Sliced Optimal Transport Sampling. Global illumination of a scene (top left, San Miguel) requires integrating radiance over a high-dimensional space of light paths. The projective variant of our sliced optimal transport (SOT) sampling technique, leveraging the particular nature of integral evaluation in rendering and further combined with a micro-Cranley-Patterson rotation per pixel, outperforms standard Monte Carlo and Quasi-Monte Carlo techniques, exhibiting less noise and no structured artifact (top right, 32spp) while offering a better spatial distribution of error (bottom right, errors from blue (small) to red (large)). Moreover, our projective SOT sampling produces better convergence of the mean absolute error for the central 7×7 zone of the highlighted reference window as a function of the number of samples per pixel (from 4spp to 4096spp, bottom-left graph) in the case of indirect lighting with one bounce.

In this paper, we introduce a numerical technique to generate sample distributions in arbitrary dimension for improved accuracy of Monte Carlo integration. We point out that optimal transport offers theoretical bounds on Monte Carlo integration error, and that the recently-introduced numerical framework of sliced optimal transport (SOT) allows us to formulate a novel and efficient approach to generating well-distributed high-dimensional pointsets. The resulting sliced optimal transport sampling, solely involving repeated 1D solves, is particularly simple and efficient for the common case of a uniform density over a d-dimensional ball. We also construct a volume-preserving map from a d-ball to a d-cube (generalizing the Shirley-Chiu mapping to arbitrary dimensions) to offer fast SOT sampling over d-cubes. We provide ample numerical evidence of the improvement in Monte Carlo integration accuracy that SOT sampling brings compared to existing QMC techniques, and derive a projective variant for rendering which rivals, and at times outperforms, current sampling strategies using low-discrepancy sequences or optimized samples.

CCS Concepts: • **Computing methodologies** → **Computer graphics**.

Additional Key Words and Phrases: multidimensional sampling, optimal transport, Radon transform, Monte Carlo integration, point distributions

ACM Reference Format:
Lois Paulin, Nicolas Bonneel, David Coeurjolly, Jean-Claude Iehl, Antoine Webanck, Mathieu Desbrun, and Victor Ostromoukhov. 2020. Sliced Optimal Transport Sampling. *ACM Trans. Graph.* 39, 4, Article 99 (July 2020), 17 pages. https://doi.org/10.1145/3386569.3392395

Authors' addresses: Lois Paulin, Nicolas Boneel, David Coeurjolly, Jean-Claude Iehl, Antoine Webanck, Victor Ostromoukhov: CNRS/LIRIS, Lyon, France; Mathieu Desbrun: California Institute of Technology, Pasadena, CA, USA, on sabbatical at SIST in ShanghaiTech University, Shanghai, China.

0730-0301/2020/7-ART99 $15.00
https://doi.org/10.1145/3386569.3392395

1 INTRODUCTION

The need to evaluate integrals of high-dimensional signals arises in a number of applications such as finance or machine learning. It is particularly crucial in global illumination where the radiance through a pixel must be integrated across the multidimensional space of possible light transport paths. Monte Carlo integration, which approximates an integral through averaging the values of

Unsupervised K-modal Styled Content Generation

OMRY SENDIK, Tel Aviv University, Israel
DANI LISCHINSKI, The Hebrew University of Jerusalem, Israel
DANIEL COHEN-OR, Tel Aviv University, Israel

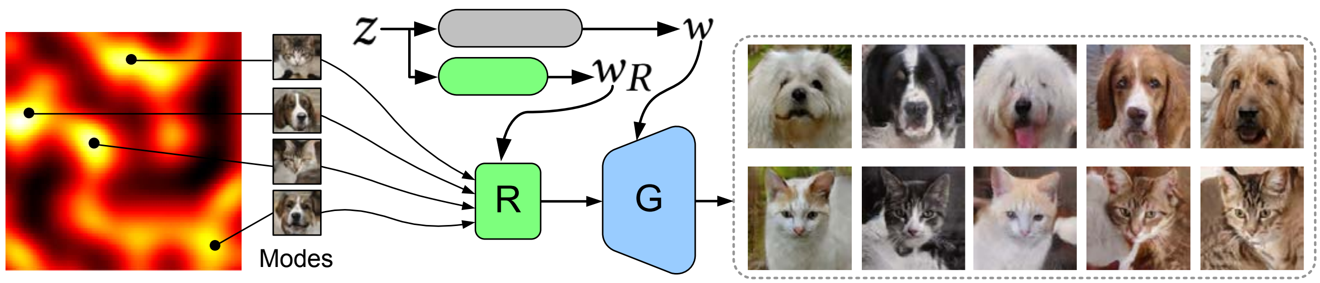


Fig. 1. Our unsupervised K-modal GAN (uMM-GAN with $K = 4$) trained on a dataset of images of cats and dogs, with various visual attributes shared by both kinds of images. The complex distribution of the training data is approximated by the generator by sampling a mixture of $K = 4$ modes, which are learned in an unsupervised manner. In this case, two of the modes (visualized on the left) correspond to cats, and two to dogs. Thus, at test time it is possible to control the species in the generated samples. Furthermore, the mode and the style in our generator are disentangled, enabling changing one while preserving the other, as demonstrated by the generated samples on the right. In each column the mode is switched, while keeping the style parameters fixed, yielding similar fur colors and patterns.

The emergence of deep generative models has recently enabled the automatic generation of massive amounts of graphical content, both in 2D and in 3D. Generative Adversarial Networks (GANs) and style control mechanisms, such as Adaptive Instance Normalization (AdaIN), have proved particularly effective in this context, culminating in the state-of-the-art StyleGAN architecture. While such models are able to learn diverse distributions, provided a sufficiently large training set, they are not well-suited for scenarios where the distribution of the training data exhibits a multi-modal behavior. In such cases, reshaping a uniform or normal distribution over the latent space into a complex multi-modal distribution in the data domain is challenging, and the generator might fail to sample the target distribution well. Furthermore, existing unsupervised generative models are not able to control the mode of the generated samples independently of the other visual attributes, despite the fact that they are typically disentangled in the training data.

In this paper, we introduce uMM-GAN, a novel architecture designed to better model multi-modal distributions, in an unsupervised fashion. Building upon the StyleGAN architecture, our network learns multiple modes, in a *completely unsupervised manner*, and combines them using a set of learned weights. We demonstrate that this approach is capable of effectively approximating a complex distribution as a superposition of multiple simple ones. We further show that uMM-GAN effectively disentangles between modes and style, thereby providing an independent degree of control over the generated content.

Authors' addresses: Omry Sendik, Tel Aviv University, Israel, omrysendik@gmail.com; Dani Lischinski, The Hebrew University of Jerusalem, Israel; Daniel Cohen-Or, Tel Aviv University, Israel.

0730-0301/2020/7-ART100 $15.00
https://doi.org/10.1145/3386569.3392454

CCS Concepts: • **Computing methodologies** → **Neural networks**; **Image manipulation**.

Additional Key Words and Phrases: Generative Adversarial Networks, StyleGAN, multi-modal distributions

ACM Reference Format:
Omry Sendik, Dani Lischinski, and Daniel Cohen-Or. 2020. Unsupervised K-modal Styled Content Generation. *ACM Trans. Graph.* 39, 4, Article 100 (July 2020), 10 pages. https://doi.org/10.1145/3386569.3392454

1 INTRODUCTION

Content generation has been a major bottleneck since the dawn of computer graphics. Recently, the emergence of generative models based on deep neural networks, finally carries a promise for being able to automatically generate massive amounts of diverse content. Although the visual quality of deep generative models could not initially rise up to the high visual fidelity bar of the field, it has been improving rapidly. Some of the most promising approaches, in terms of visual fidelity are Generative Adversarial Networks (GANs) [Goodfellow et al. 2014], which learn to generate samples whose distribution closely resembles that of the training data.

While GANs are able to generate a large amount of varied data, provided a sufficiently large training set, they are not explicitly designed for scenarios where the distribution of training data exhibits a multi-modal behavior. Consider, for example, a dataset consisting of several different species of animals, or several different kinds of cars. In such cases, reshaping a simple distribution over the latent space into a complex multi-modal one is challenging, and the learned distribution might fail to approximate that of the training

A Low-Parametric Rhombic Microstructure Family for Irregular Lattices

DAVI COLLI TOZONI, New York University
JÉRÉMIE DUMAS, Adobe Research, nTopology
ZHONGSHI JIANG, New York University
JULIAN PANETTA, EPFL
DANIELE PANOZZO, New York University
DENIS ZORIN, New York University

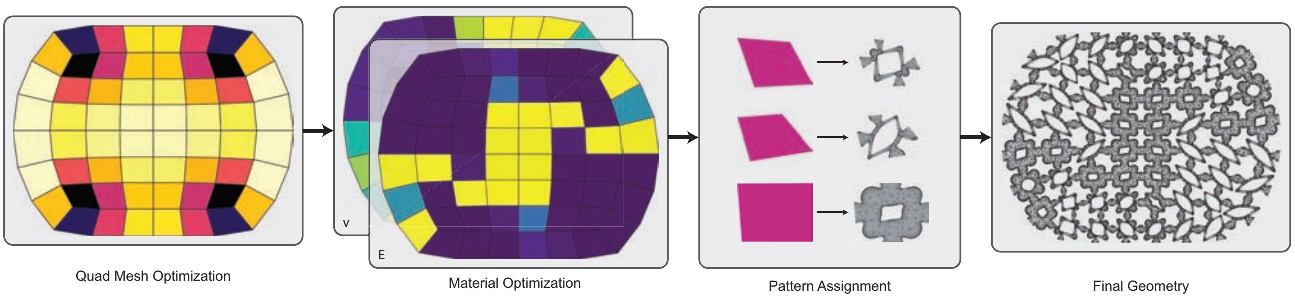


Fig. 1. The cells of a quadrilateral mesh are optimized to become quasi-rhombic; then material properties (variable Young's moduli and Poisson's ratio) are assigned to the cells. The assigned material properties are used to evaluate the geometric parameters of a tileable microstructure, encoded with a smooth spline map.

New fabrication technologies have significantly decreased the cost of fabrication of shapes with highly complex geometric structure. One important application of complex fine-scale geometric structures is to create variable effective elastic material properties in shapes manufactured from a single material. Modification of material properties has a variety of uses, from aerospace applications to soft robotics and prosthetic devices. Due to its scalability and effectiveness, an increasingly common approach to creating spatially varying materials is to partition a shape into cells and use a parametric family of small-scale geometric structures with known effective properties to fill the cells.

We propose a new approach to solving this problem for extruded, planar microstructures. Differently from existing methods for two-scale optimization based on regular grids with square periodic cells, which cannot conform to an arbitrary boundary, we introduce cell decompositions consisting of (nearly) rhombic cells. These meshes have far greater flexibility than those with square cells in terms of approximating arbitrary shapes, and, at the same time, have a number of properties simplifying small-scale structure construction. Our main contributions include a new family of 2D cell geometry structures, explicitly parameterized by their effective Young's moduli E, Poisson's ratios ν, and rhombic angle α with the geometry parameters expressed directly as smooth spline functions of E, ν, and α. This family leads to smooth transitions between the tiles and can handle a broad range of rhombic cell shapes. We introduce a complete material design pipeline based on this microstructure family, composed of an algorithm to generate rhombic tessellation from quadrilateral meshes and an algorithm to synthesize the microstructure geometry. We fabricated a number of models and experimentally demonstrated how our method, in combination with material optimization, can be used to achieve the desired deformation behavior.

CCS Concepts: • **Computing methodologies → Mesh models**.

Additional Key Words and Phrases: Additive Fabrication, Microstructures, Deformable Objects, Homogenization, Shape Optimization

ACM Reference Format:
Davi Colli Tozoni, Jérémie Dumas, Zhongshi Jiang, Julian Panetta, Daniele Panozzo, and Denis Zorin. 2020. A Low-Parametric Rhombic Microstructure Family for Irregular Lattices. *ACM Trans. Graph.* 39, 4, Article 101 (July 2020), 20 pages. https://doi.org/10.1145/3386569.3392451

Authors' addresses: Davi Colli Tozoni, New York University, davi.tozoni@nyu.edu; Jérémie Dumas, Adobe Research, nTopology, jeremie.dumas@ens-lyon.org; Zhongshi Jiang, New York University, jiangzs@nyu.edu; Julian Panetta, EPFL, julian.panetta@epfl.ch; Daniele Panozzo, New York University, panozzo@nyu.edu; Denis Zorin, New York University, dzorin@cs.nyu.edu.

0730-0301/2020/7-ART101 $15.00
https://doi.org/10.1145/3386569.3392451

1 INTRODUCTION

Advances in fabrication of highly complex geometry using additive manufacturing and other technologies resulted in new opportunities for shape design. In particular, *small-scale*, topologically and geometrically complex structures make it possible to achieve variable effective material properties using a single material for fabrication, including material properties not easily obtained by other means (such as negative Poisson's ratio) or material properties needed for precise control of deformation behavior of a shape. Realizing the potential of microstructures requires automatic generation and high-level control of the geometry, absent from commonly used geometric modeling tools.

Automatic Structure Synthesis for 3D Woven Relief

RUNDONG WU, Cornell University, USA
CLAIRE HARVEY, T.E.A.M. Inc., USA
JOY XIAOJI ZHANG, Cornell University, USA
SEAN KROSZNER, T.E.A.M. Inc., USA
BROOKS HAGAN, Rhode Island School of Design, USA
STEVE MARSCHNER, Cornell University, USA

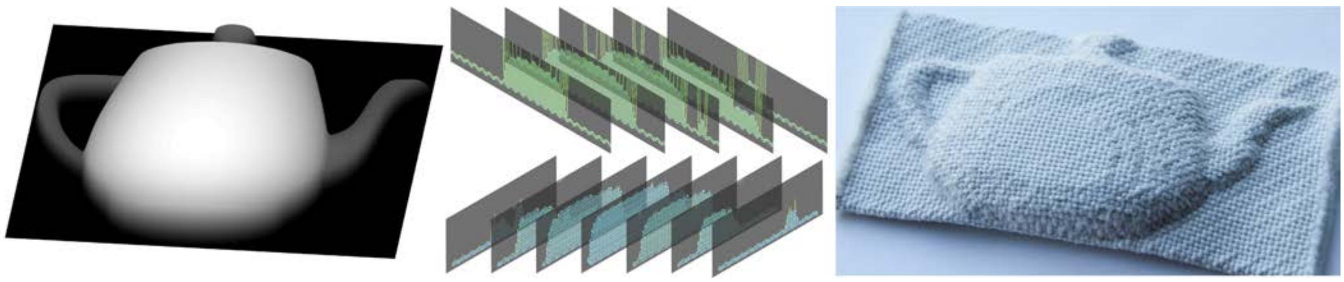


Fig. 1. Given a height field (left), our method automatically generates a 3D weavable structure (middle, showing slices of cutting planes intersecting with the model) including both warp (green) and weft (blue) paths, which can be fabricated on a Jacquard loom to approximate the input shape (right).

3D weaving is a manufacturing technique that creates multilayer textiles with substantial thickness. Currently, the primary use for these materials is in regularly structured carbon-polymer or glass-polymer composites, but in principle a wide range of complex shapes can be achieved, providing the opportunity to customize the fiber structure for individual parts and also making 3D weaving appealing in many soft-goods applications. The primary obstacle to broader use is the need to design intricate weave structures, involving tens to hundreds of thousands of yarn crossings, which are different for every shape to be produced. The goal of this research is to make 3D weaving as readily usable as CNC machining or 3D printing, by providing an algorithm to convert an arbitrary 3D solid model into machine instructions to weave the corresponding shape. We propose a method to generate 3D weaving patterns for height fields by slicing the shape along intersecting arrays of parallel planes and then computing the paths for all the warp and weft yarns, which travel in these planes. We demonstrate the method by generating weave structures for different shapes and fabricating a number of examples in polyester yarn using a Jacquard loom.

CCS Concepts: • **Computing methodologies → Shape modeling**;

Additional Key Words and Phrases: Fabrication, 3D Weaving

Authors' addresses: Rundong Wu, Cornell University, Ithaca, USA; Claire Harvey, T.E.A.M. Inc. Woonsocket, USA; Joy Xiaoji Zhang, Cornell University, Ithaca, USA; Sean Kroszner, T.E.A.M. Inc. Woonsocket, USA; Brooks Hagan, Rhode Island School of Design, Providence, USA; Steve Marschner, Cornell University, Ithaca, USA.

0730-0301/2020/7-ART102 $15.00
https://doi.org/10.1145/3386569.3392449

ACM Reference Format:
Rundong Wu, Claire Harvey, Joy Xiaoji Zhang, Sean Kroszner, Brooks Hagan, and Steve Marschner. 2020. Automatic Structure Synthesis for 3D Woven Relief. *ACM Trans. Graph.* 39, 4, Article 102 (July 2020), 10 pages. https://doi.org/10.1145/3386569.3392449

1 INTRODUCTION

The traditional process of weaving is used to make textiles for a broad range of applications. The fabrics are typically 2D—they are thin in height compared to their width and length, and are therefore used as sheet material. With the advent of fiber-reinforced composites requiring thick volumes to be filled with fibers for strength, the process of 3D weaving has emerged. Using industrial looms with some modifications to the way yarns are supplied, 3D weaving produces fabrics up to dozens of layers thick, which are comparable to their other dimensions. Using 3D fabrics in composites comes with advantages in strength and durability compared to a stack of separate layers, which is weak in the thickness direction and prone to delamination.

For reasonably flat shapes, 3D fabrics can be made in large uniform panels and bent to fit the shape of a part. For high-performance applications with complex geometry, however, the best solution is to use a "preform," a fabric that is woven specially for the part, with a shape close to the desired final shape so that it easily deforms to fit and provides a uniform and well-aligned distribution of fibers throughout. The typical practice is to use preforms only for the most demanding applications, and the required weave structures are designed manually, with some low-level support from software tools, in a process requiring hours of design time and multiple weaving trials for every new part.

Designing 3D woven fabrics is complicated because the structures in which the yarns interlace need to be consistent and uniform in

Bézier Guarding: Precise Higher-Order Meshing of Curved 2D Domains

MANISH MANDAD and MARCEL CAMPEN, Osnabrück University, Germany

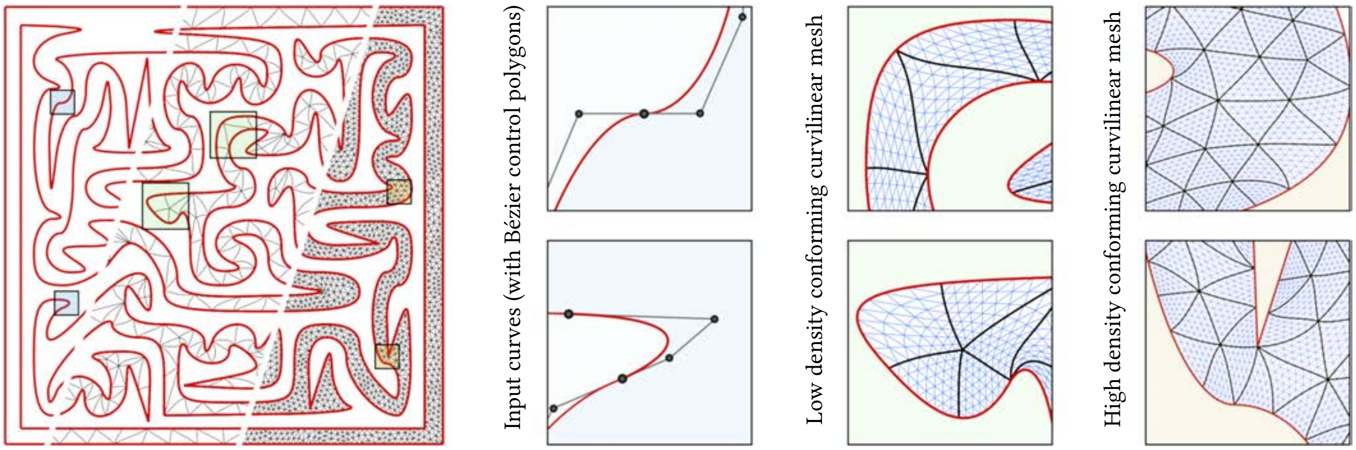


Fig. 1. Example planar domain with piecewise polynomial boundary curve (red). Our method constructs higher-order (here: cubic) curved triangle meshes (of controllable density) conforming to the prescribed curved boundary. The triangular elements (black edges) are provably regular, i.e., have injective geometric maps, and provably conform precisely to the domain's boundary. The blue blow-ups show the input curve's Bézier control polygons, while the green and orange blow-ups visualize the triangles' regular geometric maps using blue parametric iso-curves between their black and red edges.

We present a mesh generation algorithm for the curvilinear triangulation of planar domains with piecewise polynomial boundary. The resulting mesh consists of regular, injective higher-order triangular elements and precisely conforms with the domain's curved boundary. No smoothness requirements are imposed on the boundary. Prescribed piecewise polynomial curves in the interior, like material interfaces or feature curves, can be taken into account for precise interpolation by the resulting mesh's edges as well. In its core, the algorithm is based on a novel explicit construction of guaranteed injective Bézier triangles with certain edge curves and edge parametrizations prescribed. Due to the use of only rational arithmetic, the algorithm can optionally be performed using exact number types in practice, so as to provide robustness guarantees.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; **Mesh models**; **Mesh geometry models**; *Shape modeling*; • **Applied computing** → *Computer-aided design*; • **Mathematics of computing** → **Mesh generation**.

Additional Key Words and Phrases: Bézier triangle, Bézier simplex, curvilinear mesh, isogeometric analysis

Author's addresses: Manish Mandad, Marcel Campen, Institute for Computer Science, Osnabrück University, Germany.

0730-0301/2020/7-ART103 $15.00
https://doi.org/10.1145/3386569.3392372

ACM Reference Format:
Manish Mandad and Marcel Campen. 2020. Bézier Guarding: Precise Higher-Order Meshing of Curved 2D Domains. *ACM Trans. Graph.* 39, 4, Article 103 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392372

1 INTRODUCTION

The meshing of given domains using conforming triangular elements is a cornerstone in graphics, geometry processing, numerical simulation, and other fields. While often *linear* elements, triangles with straight edges, are used, the potentially significant benefits of higher-order polynomial elements with curved edges have been discussed and demonstrated [Babuška and Guo 1996; Hu et al. 2019; Oden 1994; Wang et al. 2013; Zlámal 1973]. The importance of accurate curved boundary conformance was studied as well, e.g. [Bassi and Rebay 1997; Ciarlet and Raviart 1972b; Luo et al. 2001].

We describe a method, based on a construction we call *Bézier guarding*, to generate higher-order polynomial triangle meshes with curved edges for arbitrary 2D domains with piecewise polynomial boundary curves, interface curves, and constraint curves—collectively referred to as *domain curves* in the following. We do not impose any smoothness requirements on these curves. The algorithm is general in that it supports arbitrary polynomial order.

To the best of our knowledge this method is the first to offer both of the following output properties in combination:

(1) the meshes precisely *conform* to (instead of approximate) the domain curves;
(2) the elements come with strictly *injective* and *polynomial* geometric maps.

The latter point guarantees that each curved element is the image of a straight-edge reference triangle under some injective polynomial map—and that this map is known explicitly. In other words,

C-Space Tunnel Discovery for Puzzle Path Planning

XINYA ZHANG, The University of Texas at Austin
ROBERT BELFER, McGill University
PAUL G. KRY, McGill University
ETIENNE VOUGA, The University of Texas at Austin

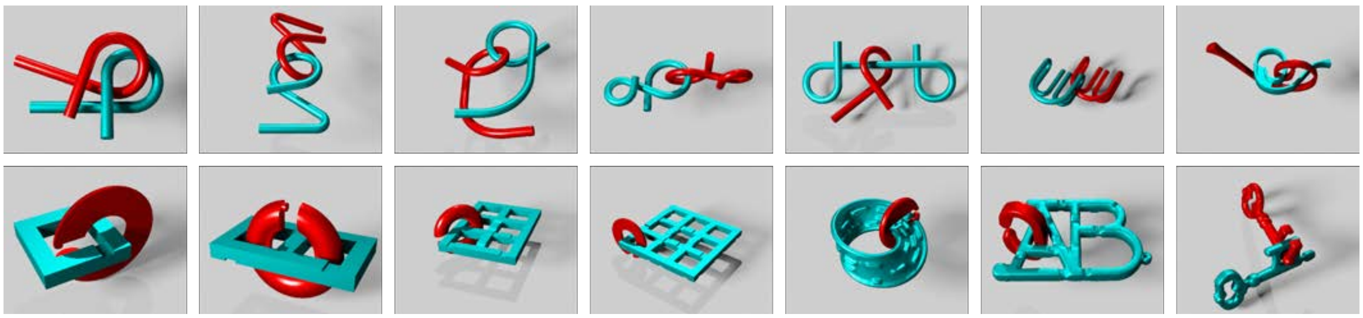


Fig. 1. Puzzles we solve using our approach, roughly ordered by difficulty from left to right. *Top row*: alpha, alpha-z, alpha-j, alpha-g, double-alpha, claw, and enigma. *Bottom row*: duet (with 4 different grid sizes), Mobius, ABC, and Key. All but the alpha variants are Hanayama puzzles.

Rigid body disentanglement puzzles are challenging for both humans and motion planning algorithms because their solutions involve tricky twisting and sliding moves that correspond to navigating through narrow tunnels in the puzzle's configuration space (C-space). We propose a tunnel-discovery and planning strategy for solving these puzzles. First, we locate important features on the pieces using geometric heuristics and machine learning, and then match pairs of these features to discover collision free states in the puzzle's C-space that lie within the narrow tunnels. Second, we propose a Rapidly-exploring Dense Tree (RDT) motion planner variant that builds tunnel escape roadmaps and then connects these roadmaps into a solution path connecting start and goal states. We evaluate our approach on a variety of challenging disentanglement puzzles and provide extensive baseline comparisons with other motion planning techniques.

CCS Concepts: • **Computing methodologies → Motion path planning**; *Neural networks.*

Additional Key Words and Phrases: sampling strategies

ACM Reference Format:
Xinya Zhang, Robert Belfer, Paul G. Kry, and Etienne Vouga. 2020. C-Space Tunnel Discovery for Puzzle Path Planning. *ACM Trans. Graph.* 39, 4, Article 104 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392468

Authors' addresses: Xinya Zhang, xinyazhang@utexas.edu, The University of Texas at Austin; Robert Belfer, belfer2470@gmail.com, McGill University; Paul G. Kry, kry@cs.mcgill.ca, McGill University; Etienne Vouga, evouga@cs.utexas.edu, The University of Texas at Austin.

0730-0301/2020/7-ART104 $15.00
https://doi.org/10.1145/3386569.3392468

1 INTRODUCTION

The *Piano Mover's Problem* asks whether one can move a piano between rooms through a sequence of rigid-body motions. This problem has inspired a great deal of work in computational geometry and robotics because it has a vast number of practical applications, from collision-free navigation in automated warehouses, to path planning in pharmaceutical drug design. Rigid disentanglement puzzles are an interesting variant of the problem because they are specifically designed to be difficult to take apart: they are notoriously difficult for both humans and computers to solve. One of the easiest puzzle in Figure 1, the *alpha puzzle* at top-left, requires a counter-intuitive twisting motion, and is frequently used as the most difficult benchmark when testing path-planning algorithms [Amato et al. 1998b; Kuffner 2004; Zhang et al. 2008; Zhang and Manocha 2008]. The reason for the difficulty can be understood by considering the geometry of the admissible configuration space.

Tunnels and Bubbles. Consider the 2D disentanglement puzzle shown in Figure 2, where a red brick can only escape the lower chamber through a thin gap. The configuration space (*C-space*) is $SE(2)$, that is, rigid translation and rotation of the red brick in the plane. The admissible region of C-space, C_{free}, is all of the collision-free points in $SE(2)$, which we visualize as a volume in $\mathbb{R}^3$ at right in Figure 2. The red brick has a large range of motion within both lower and upper chambers. We call these large and open regions of C_{free} *bubbles*. Two very thin *tunnels* connect the bubbles, and correspond to translations of the brick as it slides through the gap at one of two possible vertical orientations. Any solution path for this puzzle must find and navigate through one of these tunnels, while not getting caught in dead-ends or complex geometric features of C_{free} (the creased boundary on the left side of the right tower in Figure 2 corresponds to configurations where the brick jams into the gap at non-vertical orientations). This tunnel-bubble geometry

Computational Design of Skintight Clothing

JUAN MONTES, Concordia University
BERNHARD THOMASZEWSKI, Université de Montréal & ETH Zürich
SUDHIR MUDUR, Concordia University
TIBERIU POPA, Concordia University

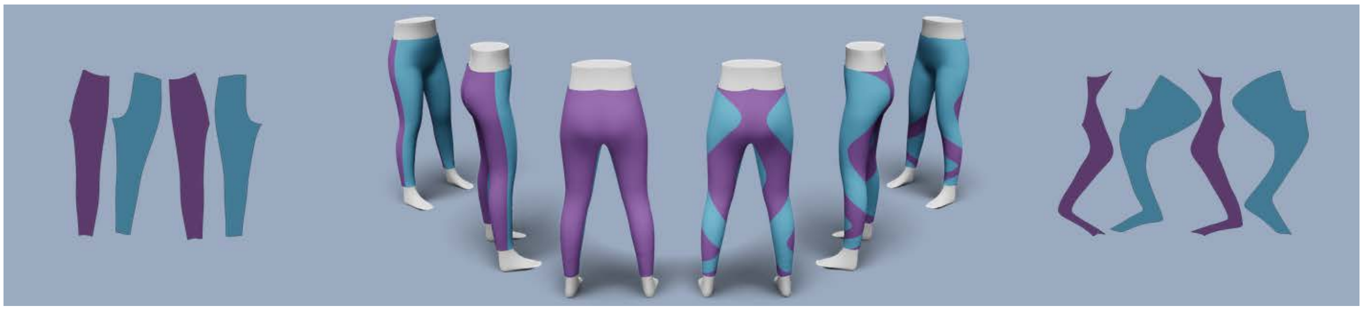


Fig. 1. Our method automatically generates patterns for skintight clothing, considering design objectives related to shape, comfort, and function. In this example, a set of initial patterns (*left*) is optimized such as to reduce the traction forces acting on the seams, yielding complex patterns (*right*) that lead to aesthetically pleasing results *(middle, right)*.

We propose an optimization-driven approach for automated, physics-based pattern design for tight-fitting clothing. Designing such clothing poses particular challenges since large nonlinear deformations, tight contact between cloth and body, and body deformations have to be accounted for. To address these challenges, we develop a computational model based on an embedding of the two-dimensional cloth mesh in the surface of the three-dimensional body mesh. Our *Lagrangian-on-Lagrangian* approach eliminates contact handling while coupling cloth and body. Building on this model, we develop a physics-driven optimization method based on sensitivity analysis that automatically computes optimal patterns according to design objectives encoding body shape, pressure distribution, seam traction, and other criteria. We demonstrate our approach by generating personalized patterns for various body shapes and a diverse set of garments with complex pattern layouts.

CCS Concepts: • **Computing methodologies → Physical simulation**; • **Applied computing → Computer-aided manufacturing**.

Additional Key Words and Phrases: computational design, pattern optimization, physically-based modelling, seam design

Authors' addresses: Juan Montes, Concordia University, juans.montes@gmail.com; Bernhard Thomaszewski, Université de Montréal & ETH Zürich, bernhard@iro.umontreal.ca; Sudhir Mudur, Concordia University, mudur@cs.concordia.ca; Tiberiu Popa, Concordia University, tiberiu.popa@concordia.ca.

0730-0301/2020/7-ART105 $15.00
https://doi.org/10.1145/3386569.3392477

ACM Reference Format:
Juan Montes, Bernhard Thomaszewski, Sudhir Mudur, and Tiberiu Popa. 2020. Computational Design of Skintight Clothing. *ACM Trans. Graph.* 39, 4, Article 105 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392477

1 INTRODUCTION

Whether as casual clothing, functional sportswear, or medical compression garments—skintight clothing has many applications, and *fit* is of central importance to all of them. The fit of a garment is determined by its design which, from a technical perspective, consists of two components: (1) a layout that determines the number of patterns and how they connect to each other and (2) the shape of the individual patterns. When fitting a design to a given body shape, the layout is typically kept fix, whereas the pattern shapes are adjusted in order to accommodate different body shapes and sizes. This task of *pattern grading* is a challenging problem, since the designer has to simultaneously consider multiple criteria that relate to the state of the garment once worn.

Although shape is largely determined by the underlying body, there is often substantial room for shape control within the limits of comfort and physics. The shape and location of the seams on the body is another design consideration, important for both aesthetic and functional goals. Apart from these visual criteria, there are several objectives relating to the deformations induced in clothing and body. For example, excessive tensile deformations will affect comfort and may cause fabric and seams to deteriorate prematurely. Compressions, on the other hand, induce wrinkles that are typically perceived as design flaws in tight-fitting clothing. Designing pattern shapes that strike an ideal balance between these criteria requires time and expertise, both of which are important cost factors.

In this work, we present an automated, optimization-driven fitting approach for skintight clothing. As the technical core of our method,

Cut-enhanced PolyCube-Maps for Feature-aware All-Hex Meshing

HAO-XIANG GUO*, Tsinghua University and Microsoft Research Asia
XIAOHAN LIU* and DONG-MING YAN, NLPR, Institute of Automation, Chinese Academy of Sciences and School of Artificial Intelligence, University of Chinese Academy of Sciences
YANG LIU†, Microsoft Research Asia

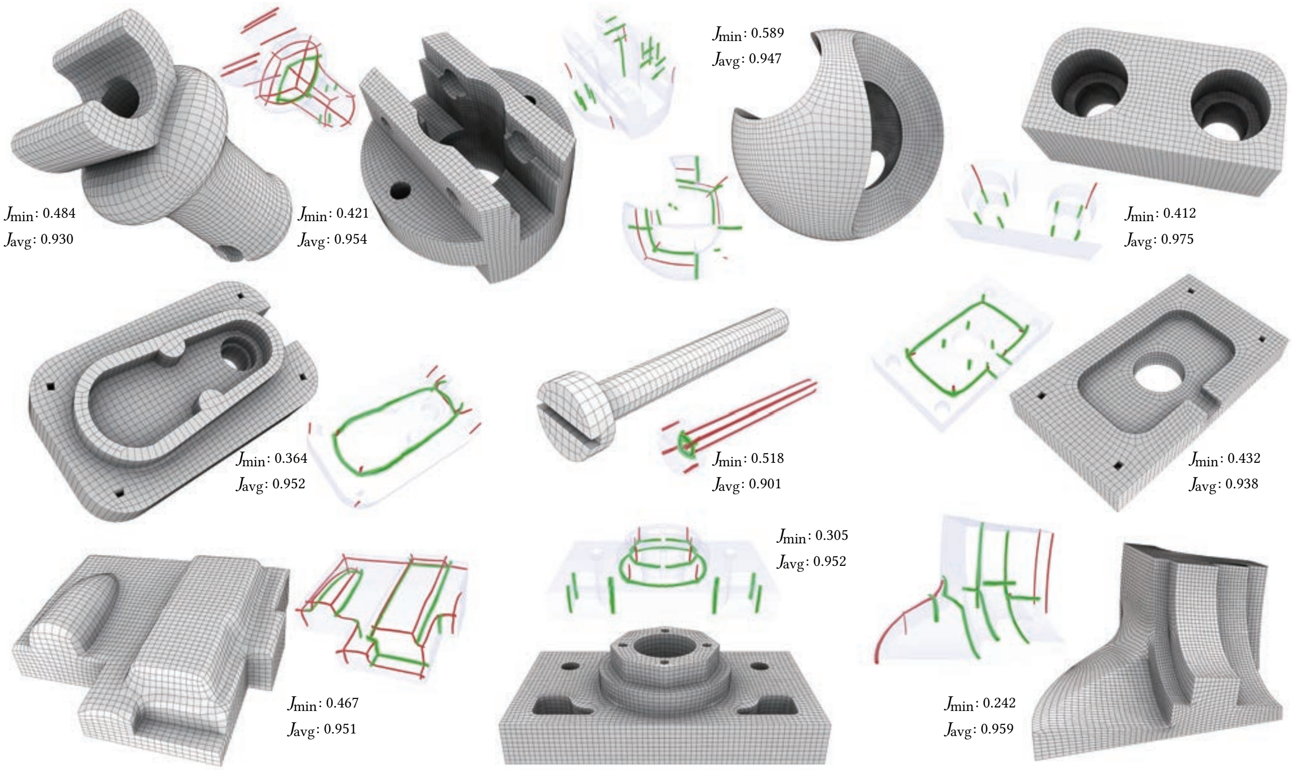


Fig. 1. High-quality hexahedral meshes generated by our CE-PolyCubeMap algorithm. Their interior singular graphs are illustrated in the insets. $J_{\min}$ and J_{avg} are the minimal and average scaled Jacobian of the hexahedral mesh.

Volumetric PolyCube-Map-based methods offer automatic ways to construct all-hexahedral meshes for closed 3D polyhedral domains, but their meshing quality is limited by the lack of interior singularities and feature alignment. In the presented work, we propose *cut-enhanced PolyCube-Maps*, to introduce essential interior singularities and preserve most input features. Our main idea is simple and intuitive: by inserting proper parameterization seams into the initial PolyCube-Map via novel PolyCube cutting operations, the mapping distortion can be reduced significantly.

The cut-enhanced PolyCube-Map computation includes feature-aware PolyCube-Map construction and cut-enhanced PolyCube deformation. The former aims to preserve input feature edges during the initial PolyCube-Map construction. The latter introduces seams into the volumetric PolyCube shape by cutting it through selective PolyCube edges and deforms the modified PolyCube under the seamless constraints to compute a low-distortion PolyCube-Map. The hexahedral mesh induced by the final PolyCube-Map can be further enhanced by our mesh improvement algorithm.

We validate the efficacy of our method on a collection of more than one hundred CAD models and demonstrate its advantages over other automatic all-hex meshing methods and padding strategies. The limitations of cut-enhanced PolyCube-Maps are also discussed thoroughly.

CCS Concepts: • **Computing methodologies → Volumetric models**.

Additional Key Words and Phrases: hexahedral meshing, PolyCube-Map, cut-enhanced, feature-aware

*Both authors were intern students at Microsoft Research Asia.
†Corresponding author

Authors' addresses: Hao-Xiang Guo, ghx17@mails.tsinghua.edu.cn, Tsinghua University, Microsoft Research Asia; Xiaohan Liu, liuxiaohan2017@ia.ac.cn; Dong-Ming Yan, yandongming@gmail.com, NLPR, Institute of Automation, Chinese Academy of Sciences, School of Artificial Intelligence, University of Chinese Academy of Sciences; Yang Liu, yangliu@microsoft.com, Microsoft Research Asia.

0730-0301/2020/7-ART106 $15.00
https://doi.org/10.1145/3386569.3392378

Data-driven Extraction and Composition of Secondary Dynamics in Facial Performance Capture

GASPARD ZOSS, DisneyResearch|Studios and ETH Zurich
EFTYCHIOS SIFAKIS, University of Wisconsin-Madison and DisneyResearch|Studios
MARKUS GROSS, DisneyResearch|Studios and ETH Zurich
THABO BEELER, DisneyResearch|Studios
DEREK BRADLEY, DisneyResearch|Studios

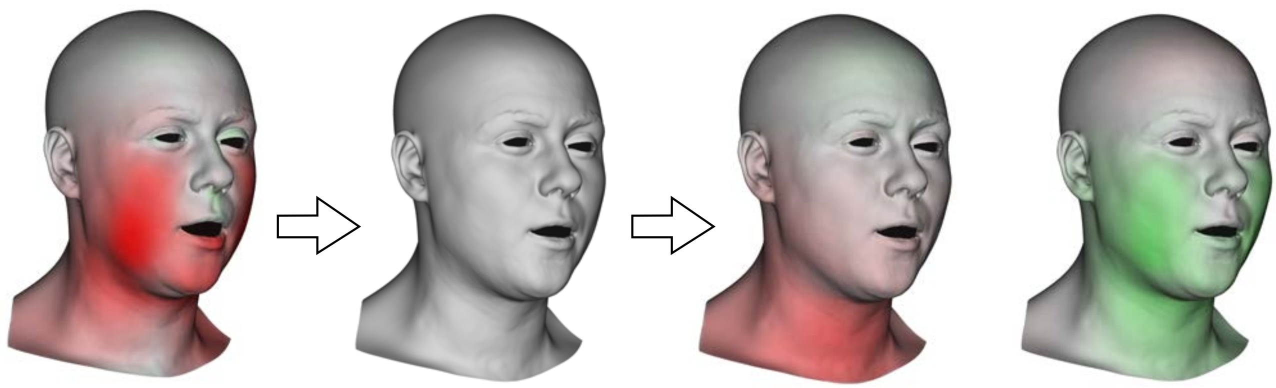


Fig. 1. We present a data-driven method to remove secondary dynamic effects from performance capture data, such as jiggling skin due to root skull motion, and a complementary method to synthesize dynamics under different root motion. Here, the color (red/green) represents the signed distance from the dynamics-free performance (gray).

Performance capture of expressive subjects, particularly facial performances acquired with high spatial resolution, will inevitably incorporate some fraction of motion that is due to inertial effects and dynamic overshoot due to ballistic motion. This is true in most natural capture environments where the actor is able to move freely during their performance, rather than being tethered to a fixed position. Normally these secondary dynamic effects are unwanted, as the captured facial performance is often retargeted to different head motion, and sometimes to completely different characters, and in both cases the captured dynamic effects should be removed and new secondary effects should be added. This paper advances the hypothesis that for a highly constrained elastic medium such as the human face, these secondary inertial effects are predominantly due to the motion of the underlying bony structures (cranium and mandible). Our work aims to compute and characterize the difference between the captured dynamic facial performance, and a speculative quasistatic variant of the same motion should the inertial effects have been absent. This is used to either subtract parasitic secondary dynamics that resulted from unintentional motion during capture, or compose such effects on top of a quasistatic performance to simulate a new dynamic motion of the actor's body and skull, either artist-prescribed or acquired via motion capture. We propose a data-driven technique that comprises complementary removal and synthesis networks for secondary dynamics in facial performance capture. We show how such a system can be effectively trained from a collection of acquired dynamic deformations under varying expressions where the actor induces rigid head motion from walking and running, as well as forced oscillatory body motion in a controlled setting by external actuators.

CCS Concepts: • **Computing methodologies → Motion processing**; *Motion capture.*

Additional Key Words and Phrases: Secondary Dynamics Prediction, Soft-Tissue Motion, Facial Performance Capture, Data-Driven Animation

ACM Reference Format:
Gaspard Zoss, Eftychios Sifakis, Markus Gross, Thabo Beeler, and Derek Bradley. 2020. Data-driven Extraction and Composition of Secondary Dynamics in Facial Performance Capture. *ACM Trans. Graph.* 39, 4, Article 107 (July 2020), 10 pages. https://doi.org/10.1145/3386569.3392463

Authors' addresses: Gaspard Zoss, DisneyResearch|Studios, ETH Zurich, gaspard.zoss@disneyresearch.com; Eftychios Sifakis, University of Wisconsin-Madison, DisneyResearch|Studios, sifakis@cs.wisc.edu; Markus Gross, DisneyResearch|Studios, ETH Zurich, gross@disneyresearch.com; Thabo Beeler, DisneyResearch|Studios, thabo.beeler@gmail.com; Derek Bradley, DisneyResearch|Studios, derek.bradley@disneyresearch.com.

0730-0301/2020/7-ART107 $15.00
https://doi.org/10.1145/3386569.3392463

Deep Geometric Texture Synthesis

AMIR HERTZ*, Tel Aviv University
RANA HANOCKA*, Tel Aviv University
RAJA GIRYES, Tel Aviv University
DANIEL COHEN-OR, Tel Aviv University

Recently, deep generative adversarial networks for image generation have advanced rapidly; yet, only a small amount of research has focused on generative models for irregular structures, particularly meshes. Nonetheless, mesh generation and synthesis remains a fundamental topic in computer graphics. In this work, we propose a novel framework for synthesizing geometric textures. It learns geometric texture statistics from local neighborhoods (*i.e.,* local triangular patches) of a single reference 3D model. It learns deep features on the faces of the input triangulation, which is used to subdivide and generate offsets across multiple scales, without parameterization of the reference or target mesh. Our network displaces mesh vertices in any direction (*i.e.,* in the normal *and* tangential direction), enabling synthesis of geometric textures, which cannot be expressed by a simple 2D displacement map. Learning and synthesizing on local geometric patches enables a genus-oblivious framework, facilitating texture transfer between shapes of different genus.

CCS Concepts: • **Computing methodologies → Neural networks**; **Shape analysis**.

Additional Key Words and Phrases: Geometric Deep Learning, Surface Reconstruction, Shape Analysis

ACM Reference Format:
Amir Hertz, Rana Hanocka, Raja Giryes, and Daniel Cohen-Or. 2020. Deep Geometric Texture Synthesis. *ACM Trans. Graph.* 39, 4, Article 108 (July 2020), 11 pages. https://doi.org/10.1145/3386569.3392471

1 INTRODUCTION

In recent years, neural networks for geometry processing have emerged rapidly and changed the way we approach geometric problems. Yet, common 3D modeling representations are irregular and unordered, which challenges the straightforward adaptation from image-based techniques. Recent advances enable applying convolutional neural networks (CNNs) on irregular structures, like point clouds and meshes [Li et al. 2018a; Hanocka et al. 2019]. So far, these CNN-based methods have demonstrated promising success for discriminative tasks like classification and segmentation. On the other hand, only a small amount of research has focused on generative models for irregular structures, particularly meshes [Gao et al. 2019].

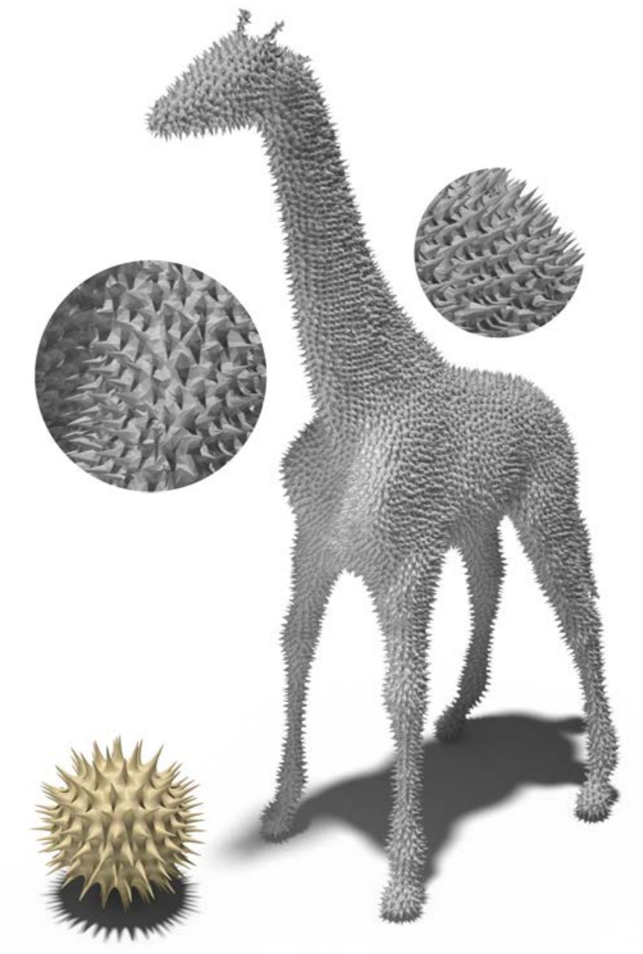


Fig. 1. Learning local geometric textures from a reference mesh (gold) and transferring it to a target mesh (giraffe).

In this work, we take a step forward in developing generative models for meshes. We present a deep neural network that learns the geometric texture of a single 3D reference mesh, and can transfer its texture to any arbitrary target mesh. Our generative framework uses a CNN to learn to model the unknown distribution of geometric textures directly from an input triangular mesh. Our network learns local neighborhoods (*i.e.,* local triangular patches) from a reference model, which is used to subdivide and generate offsets over the target mesh to match the local statistics of the reference model. For example, see Figure 1, where the geometric spikes of the reference

*Joint first authors.

Authors' addresses: Amir Hertz, Tel Aviv University; Rana Hanocka, Tel Aviv University; Raja Giryes, Tel Aviv University; Daniel Cohen-Or, Tel Aviv University.

0730-0301/2020/7-ART108 $15.00
https://doi.org/10.1145/3386569.3392471

Developability of Heightfields via Rank Minimization

SILVIA SELLÁN, University of Toronto
NOAM AIGERMAN, Adobe Research
ALEC JACOBSON, University of Toronto and Adobe Research

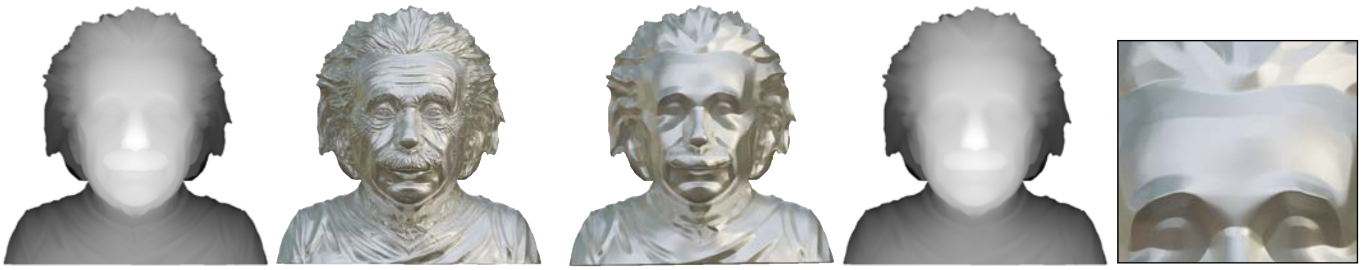


Fig. 1. Our method, inspired by compressed sensing and the problem of rank minimization, takes as input a heightfield describing a surface (left) and outputs a heightfield describing a piecewise developable surface (right) which approximates the input. 3D model by Lloyd Chidgzey under CC BY-SA 3.0.

This work concerns the computation and approximation of developable surfaces — surfaces that are locally isometric to the two-dimensional plane. These surfaces are heavily studied in differential geometry, and are also of great interest to fabrication, architecture and fashion. We focus specifically on developability of heightfields. Our main observation is that developability can be cast as a rank constraint, which can then be plugged into theoretically-grounded rank-minimization techniques from the field of compressed sensing. This leads to a convex semidefinite optimization problem, which receives an input heightfield and recovers a similar heightfield which is developable. Due to the sparsifying nature of compressed sensing, the recovered surface is piecewise developable, with creases emerging between connected developable pieces. The convex program includes one user-specified parameter, balancing adherence to the original surface with developability and number of patches. We moreover show, that in contrast to previous techniques, our discretization does not introduce a bias and the same results are achieved across resolutions and orientations, and with no limit on the number of creases and patches. We solve this convex semidefinite optimization problem efficiently, by devising a tailor-made ADMM solver which leverages matrix-projection observations unique to our problem. We employ our method on a plethora of experiments, from denoising 3D scans of developable geometry such as documents and buildings, through approximating general heightfields with developable ones, and up to interpolating sparse annotations with a developable heightfield.

Authors' addresses: Silvia Sellán, University of Toronto, 40 St. George Street, Toronto, Ontario, M5S2E4, sgsellan@cs.toronto.edu; Noam Aigerman, Adobe Research, 601 Townsend St., San Francisco, California, 94103, aigerman@adobe.com; Alec Jacobson, University of Toronto and Adobe Research, 40 St. George Street, Toronto, Ontario, M5S2E4, jacobson@cs.toronto.edu.

0730-0301/2020/7-ART109 $15.00
https://doi.org/10.1145/3386569.3392419

CCS Concepts: • **Computing methodologies** → *Image processing*; *Shape analysis*; *Mesh geometry models*; • **Mathematics of computing** → *Convex optimization.*

Additional Key Words and Phrases: geometry processing, compressed sensing, rank minimization, developable surface, heightfield

ACM Reference Format:
Silvia Sellán, Noam Aigerman, and Alec Jacobson. 2020. Developability of Heightfields via Rank Minimization . *ACM Trans. Graph.* 39, 4, Article 109 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392419

1 INTRODUCTION

Developability determines an important subclass of surfaces in three-dimensions. A (piecewise) developable surface is one that can be constructed by folding, creasing, bending or welding planar surfaces without stretching. Piecewise developable surfaces are all around us: paper pages of a book, mechanical objects manufactured with a 5-axis CNC-mill, the wooden-plank hulls of boats, and the steel and glass panels of modern architecture.

While manufacturing techniques for developable surfaces enjoy a long history and ubiquitous use, computational methods for developable surfaces have been notoriously elusive. Mapping the curvature criteria of developability to common discrete surface representations can be tricky: for example, a triangle mesh is trivially piecewise developable; meanwhile, a quad mesh is in general non-planar. There has been a recent surge of advances building new discrete notions of developability for these and other common surface representations. Many if not most works focus on defining developability for a single smooth patch without crease or weld curves, while others require a small number of explicitly provided curves. These methods focus on forward simulation of bending planar patches into a design or surfacing provided boundary curves. Relatively few works consider the inverse problem: which piecewise developable surface best explains an input observation (see Fig. 1).

Discrete Differential Operators on Polygonal Meshes

FERNANDO DE GOES, Pixar Animation Studios
ANDREW BUTTS, Pixar Animation Studios
MATHIEU DESBRUN, ShanghaiTech/Caltech

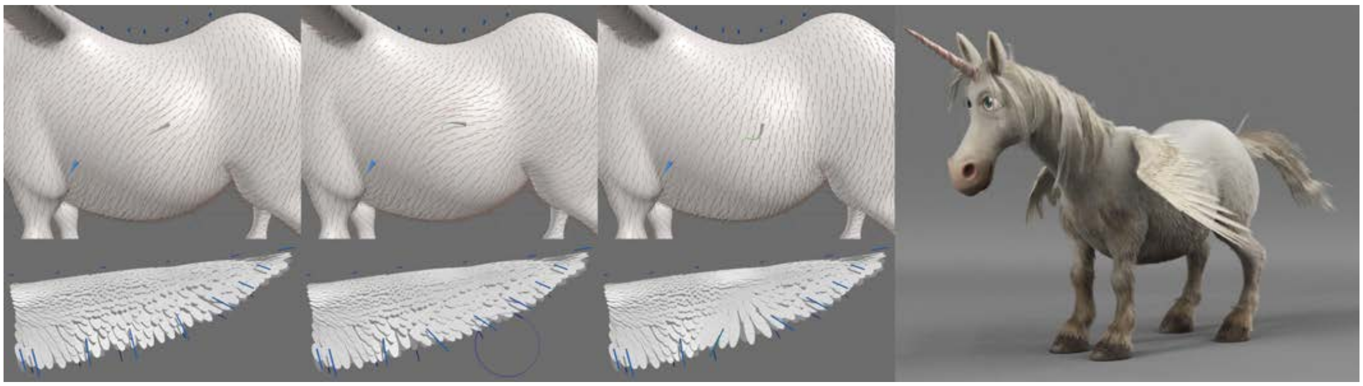


Fig. 1. **Geometry processing on polygonal meshes:** Our approach enriches the available numerical toolbox for polygonal meshes by offering discrete differential operators (gradient, Laplacian, covariant derivative, shape operator, etc) acting either on discrete forms [Desbrun et al. 2008] or directional fields [Vaxman et al. 2017]. Our operators allow for the seamless extension of existing geometry processing algorithms to meshes with arbitrary 3D polygons, including non-planar and non-convex faces. Here, an example showing grooming on a quad mesh where fur and feathers are designed interactively by solving scalar and vector Poisson equations. Offset blue curves represent handles to guide the tangent direction and local shape of strands of fur and feathers.

Geometry processing of surface meshes relies heavily on the discretization of differential operators such as gradient, Laplacian, and covariant derivative. While a variety of discrete operators over triangulated meshes have been developed and used for decades, a similar construction over polygonal meshes remains far less explored despite the prevalence of non-simplicial surfaces in geometric design and engineering applications. This paper introduces a principled construction of discrete differential operators on surface meshes formed by (possibly non-flat and non-convex) polygonal faces. Our approach is based on a novel mimetic discretization of the gradient operator that is linear-precise on arbitrary polygons. Equipped with this discrete gradient, we draw upon ideas from the Virtual Element Method in order to derive a series of discrete operators commonly used in graphics that are now valid over polygonal surfaces. We demonstrate the accuracy and robustness of our resulting operators through various numerical examples, before incorporating them into existing geometry processing algorithms.

CCS Concepts: • **Mathematics of computing → Discretization**; • **Computing methodologies → Mesh models**.

Additional Key Words and Phrases: Discrete differential operators, geometry processing, polygonal meshes.

Authors' addresses: Fernando de Goes, Andrew Butts, Pixar Animation Studios, Emeryville, CA, USA; Mathieu Desbrun, California Institute of Technology, Pasadena, CA, USA, on sabbatical at SIST in ShanghaiTech University, Shanghai, China.

0730-0301/2020/7-ART110
https://doi.org/10.1145/3386569.3392389

ACM Reference Format:
Fernando de Goes, Andrew Butts, and Mathieu Desbrun. 2020. Discrete Differential Operators on Polygonal Meshes. *ACM Trans. Graph.* 39, 4, Article 110 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392389

1 INTRODUCTION

Discrete differential operators play a central role in geometry processing, allowing for the analysis and manipulation of 3D surfaces. The vast majority of discrete differential operators assume a triangulated surface, as its simplicial and piecewise-linear nature renders the derivation and error analysis of operators particularly simple. Yet, modeling and engineering applications often rely on polygonal meshes to better capture geometric features and ease both artistic design and fabrication. Deriving discrete differential operators for polygonal meshes by first triangulating each polygonal face is ill-advised since it introduces unnecessary computational bias due to the dependence of the results on the choice of triangulation. While recent works have offered extensions of the Laplacian operator on polygonal meshes [Alexa and Wardetzky 2011; Sharp et al. 2019; Bunge et al. 2020], there remains a long list of basic discrete differential operators used in geometry processing for which variants that are valid on non-simplicial meshes do not exist.

In this paper, we introduce a systematic construction of discrete differential operators on surface meshes made out of (possibly non-planar and non-convex) polygonal faces. At its core is a new discretization of the gradient operator that is linear-precise for arbitrary polygons. From this gradient, we bootstrap the derivation of various first- and second-order discrete operators, including the covariant derivative and the shape operator, by mimicking key structural

Efficient Bijective Parameterizations

JIAN-PING SU, University of Science and Technology of China, China
CHUNYANG YE, University of Science and Technology of China, China
LIGANG LIU, University of Science and Technology of China, China
XIAO-MING FU*, University of Science and Technology of China, China

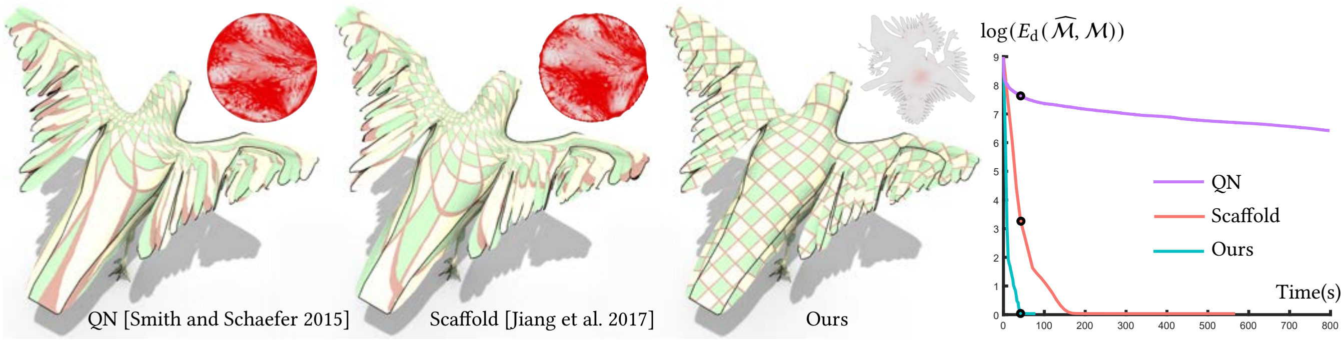


Fig. 1. We present an efficient method for generating bijective parameterizations which is much faster than existing methods. The left three figures respectively show the snapshots of the bijective parameterization results on a disk topology mesh (1445322 triangles) generated by the three methods at similar times: 39.974 seconds for QN (quasi-Newton), 41.905 seconds for Scaffold, 39.971 seconds for ours, marked as the black circles on the rightmost curves. All methods start from the same initialization [Tutte 1963]. The Scaffold method takes 159.177 seconds and the QN method takes at least five hours to obtain our result. The color of the triangles in the parameterizations encodes the symmetric Dirichlet distortion metric (with large values in red and optimal values in white).

We propose a novel method to efficiently compute bijective parameterizations with low distortion on disk topology meshes. Our method relies on a second-order solver. To design an efficient solver, we develop two key techniques. First, we propose a coarse shell to substantially reduce the number of collision constraints that are used to guarantee overlap-free boundaries. During the optimization process, the shell ensures the Hessian matrix with a fixed nonzero structure and a low density, thereby significantly accelerating the optimization. The second is a triangle inequality-based barrier function that effectively ensures non-intersecting boundaries. Our barrier function is C^∞ inside the locally supported region and its convex second-order approximation is able to be analytically obtained. Compared to state-of-the-art methods for optimizing bijective parameterizations, our method exhibits better scalability and is about six times faster. The performance of our bijective parameterization algorithm is comparable to state-of-the-art methods of locally flip-free parameterizations. A large number of experimental results have shown the capability and feasibility of our method.

CCS Concepts: • **Computing methodologies → Shape modeling**.

*The corresponding author

Authors' addresses: Jian-Ping Su, University of Science and Technology of China, China, SJPing@mail.ustc.edu.cn; Chunyang Ye, University of Science and Technology of China, China, yechyang@mail.ustc.edu.cn; Ligang Liu, University of Science and Technology of China, China, lgliu@ustc.edu.cn; Xiao-Ming Fu, University of Science and Technology of China, China, fuxm@ustc.edu.cn.

0730-0301/2020/7-ART111 $15.00
https://doi.org/10.1145/3386569.3392435

Additional Key Words and Phrases: bijective parameterizations, coarse shell meshes, triangle inequality-based barriers

ACM Reference Format:
Jian-Ping Su, Chunyang Ye, Ligang Liu, and Xiao-Ming Fu. 2020. Efficient Bijective Parameterizations. *ACM Trans. Graph.* 39, 4, Article 111 (July 2020), 8 pages. https://doi.org/10.1145/3386569.3392435

1 INTRODUCTION

Computing low-distortion parameterizations is a fundamental problem in computer graphics applications, such as texture mapping, remeshing, and shape analysis [Floater and Hormann 2005; Hormann et al. 2007]. A number of methods have been proposed to compute *locally flip-free* parameterizations [Liu et al. 2018; Rabinovich et al. 2017; Shtengel et al. 2017; Smith et al. 2019; Zhu et al. 2018]. Generating bijective parameterizations has started to attract attention in recent years [Jiang et al. 2017; Smith and Schaefer 2015].

In addition to being locally flip-free, a bijection should also be *globally overlap-free*, i.e., no intersecting boundary edges. The general method to optimize bijective parameterizations starts from a bijective initialization and then minimizes the distortion metric while ensuring bijection during the optimization process. However, efficiently optimizing bijective parameterizations with low isometric distortion is still a challenging problem. The reasons are two-fold. First, preventing overlaps for boundary edges leads to non-linear collision constraints. Second, the number of potential collisions is quadratic in the number of boundary edges. To achieve high efficiency, a possible solution is to develop a fast second-order solver. In general, there are three requirements for designing such a solver.

First, since it takes much more time to solve a sparse linear system with an updated nonzero structure than with a fixed nonzero

ENIGMA: Evolutionary Non-Isometric Geometry MAtching

MICHAL EDELSTEIN, Technion - Israel Institute of Technology, Israel
DANIELLE EZUZ, Technion - Israel Institute of Technology, Israel
MIRELA BEN-CHEN, Technion - Israel Institute of Technology, Israel

In this paper we propose a *fully automatic* method for shape correspondence that is widely applicable, and especially effective for non isometric shapes and shapes of different topology. We observe that fully-automatic shape correspondence can be decomposed as a hybrid discrete/continuous optimization problem, and we find the best *sparse* landmark correspondence, whose sparse-to-dense extension minimizes a *local metric distortion*. To tackle the *combinatorial* task of landmark correspondence we use an evolutionary *genetic algorithm*, where the local distortion of the sparse-to-dense extension is used as the objective function. We design novel *geometrically* guided genetic operators, which, when combined with our objective, are highly effective for non isometric shape matching. Our method outperforms state of the art methods for automatic shape correspondence both quantitatively and qualitatively on challenging datasets.

CCS Concepts: • **Computing methodologies** -> **Shape analysis**.

Additional Key Words and Phrases: Shape Analysis, Shape Correspondence

ACM Reference Format:
Michal Edelstein, Danielle Ezuz, and Mirela Ben-Chen. 2020. ENIGMA: Evolutionary Non-Isometric Geometry MAtching. *ACM Trans. Graph.* 39, 4, Article 112 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392447

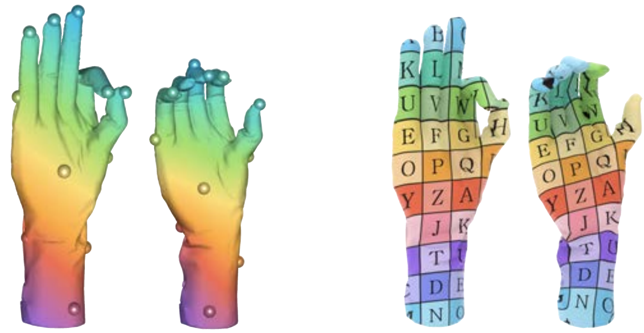


Fig. 1. A map between shapes of different genus obtained by our approach. (left) Output landmark correspondence and functional map, visualized using color transfer. (right) Final pointwise map visualized using texture transfer.

1 INTRODUCTION

Shape correspondence is a fundamental task in shape analysis: given two shapes, the goal is to compute a semantic correspondence between points on them. Shape correspondence is required when two shapes are analyzed jointly, which is common in many applications such as texture and deformation transfer [Sumner and Popović 2004], statistical shape analysis [Munsell et al. 2008] and shape classification [Ezuz et al. 2017], to mention just a few examples.

The difficulty of the shape matching problem depends on the class of *deformations* that can be applied to one shape to align it with the second. For example, if only *rigid* transformations are allowed it is easier to find a correspondence than if non-rigid deformations are also possible, since the number of degrees of freedom is small and the space of allowed transformations is easy to parameterize. Similarly, if only *isometric* deformations are allowed, the matching is easier than if non-isometry is possible, since then there is a clear criteria of the quality of the map, namely the preservation of geodesic distances. The hardest case is when the two shapes belong to the same semantic class, but are not necessarily isometric. In this case, the correspondence algorithm should achieve two goals: (1) put in correspondence semantically meaningful points on both shapes, and (2) reduce the *local* metric distortion.

Hence, the non-isometric shape correspondence problem is often considered as a *two step* process. First, the *global semantics* of the matching is given by a sparse set of *corresponding landmarks* of salient points on both shapes. If this set is informative enough, then the full shapes can be matched by extending the landmark correspondence to a full map from the source to the target in a *consistent* and *smooth* way. The first problem is *combinatorial*, requiring the computation of a permutation of a subset of the landmarks, whereas the second problem is *continuous*, requiring the definition and computation of local differential properties of the map. Whereas the second problem has been tackled by multiple methods [Aigerman and Lipman 2016; Ezuz et al. 2019a,b; Mandad et al. 2017] which yield excellent results for non-isometric shapes, methods that address the sparse landmark correspondence problem [Dym et al. 2017; Kezurer et al. 2015; Maron et al. 2016; Sahillioğlu 2018] have so far been limited either to the nearly isometric case, or to a very small set of landmarks.

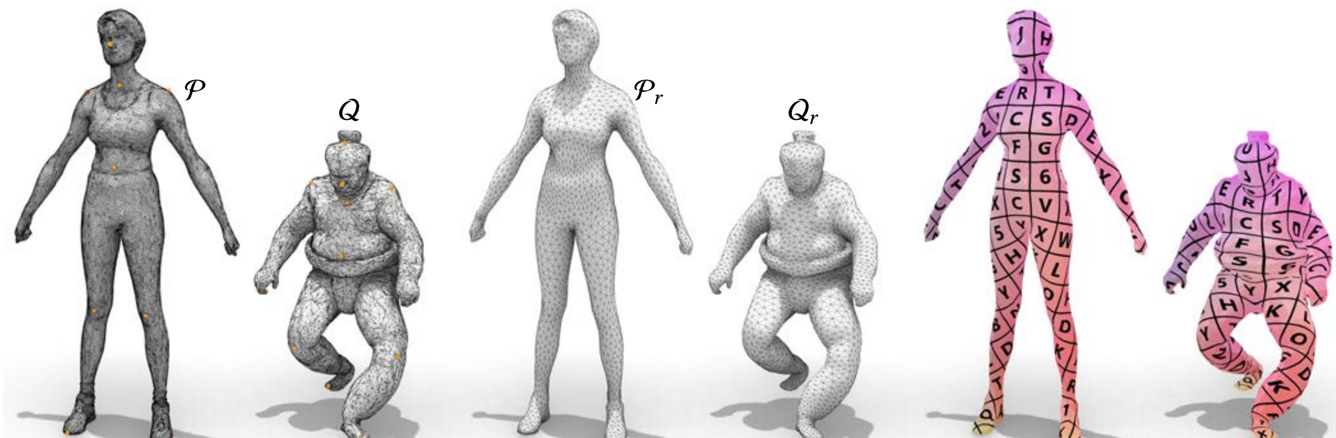
We propose to leverage the efficient algorithms for solving the second problem to generate a framework for solving the first. Specifically, we suggest a combinatorial optimization for matching a sparse set of landmarks, such that the best obtainable local distortion of the corresponding sparse-to-dense extension is minimized. As the optimization tool, we propose to use a *genetic algorithm*, as these have been used for combinatorial optimization for a few decades [Holland 1992], and are quite general in the type of objectives they can optimize. Despite their success in other fields, though, to the best of our knowledge, their use in shape analysis has been limited so far to isometric matching [Sahillioğlu 2018].

Authors' addresses: Michal Edelstein, Technion - Israel Institute of Technology, Israel, smichale@cs.technion.ac.il; Danielle Ezuz, Technion - Israel Institute of Technology, Israel, danielle.ezuz@gmail.com; Mirela Ben-Chen, Technion - Israel Institute of Technology, Israel, mirela@cs.technion.ac.il.

0730-0301/2020/7-ART112
https://doi.org/10.1145/3386569.3392447

Error-bounded Compatible Remeshing

YANG YANG, University of Science and Technology of China, China
WEN-XIANG ZHANG, University of Science and Technology of China, China
YUAN LIU, Hefei University of Technology, China
LIGANG LIU, University of Science and Technology of China, China
XIAO-MING FU*, University of Science and Technology of China, China



(a) Input **(b)** Compatible meshes **(c)** Mapping

Fig. 1. Error-bounded compatible remeshing. (a) Two input surface meshes with a small set of corresponding landmarks (orange points), one of which is dense with 44026 vertices, and the other is uneven with 16000 vertices. (b) The fairly regular output compatible meshes, each of which contains 3405 vertices and satisfies the error-bounded constraint. Namely, $d_H(\mathcal{P}_r, \mathcal{P}) \leq \epsilon_p$ and $d_H(\mathcal{Q}_r, \mathcal{Q}) \leq \epsilon_q$. Here, $d_H(\cdot, \cdot)$ denotes the Hausdorff distance between two meshes, and $\epsilon_p = 0.3\% d_{bb}^p$ and $\epsilon_q = 0.3\% d_{bb}^q$, where d_{bb}^p and d_{bb}^q are the diagonal lengths of the bounding boxes of $\mathcal{P}$ and $\mathcal{Q}$, respectively. (c) Textures and colors are used to indicate correspondences between models.

We present a novel method to construct compatible surface meshes with bounded approximation errors. Given two oriented and topologically equivalent surfaces and a sparse set of corresponding landmarks, our method contains two steps: (1) generate compatible meshes with bounded approximation errors and (2) reduce mesh complexity while ensuring that approximation errors are always bounded. Central to the first step is a parameterization-based remeshing technique, which is capable of isotropically remeshing the input surfaces to be compatible and error-bounded. By iteratively performing a novel edge-based compatible remeshing and increasing the compatible target edge lengths, the second step effectively reduces mesh complexity while explicitly maintaining compatibility, regularity, and bounding approximation errors. Tests on various pairs of complex models demonstrate the efficacy and practicability of our method for constructing high-quality compatible meshes with bounded approximation errors.

CCS Concepts: • **Computing methodologies → Shape modeling**.

Additional Key Words and Phrases: compatible remeshing, bounded approximation errors, bounded Hausdorff distances, low mesh complexity

ACM Reference Format:
Yang Yang, Wen-Xiang Zhang, Yuan Liu, Ligang Liu, and Xiao-Ming Fu. 2020. Error-bounded Compatible Remeshing. *ACM Trans. Graph.* 39, 4, Article 113 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392434

*The corresponding author

Authors' addresses: Yang Yang, University of Science and Technology of China, China, yang1993@mail.ustc.edu.cn; Wen-Xiang Zhang, University of Science and Technology of China, China, zwx111@mail.ustc.edu.cn; Yuan Liu, Hefei University of Technology, China, liuyzz@hfut.edu.cn; Ligang Liu, University of Science and Technology of China, China, lgliu@ustc.edu.cn; Xiao-Ming Fu, University of Science and Technology of China, China, fuxm@ustc.edu.cn.

0730-0301/2020/7-ART113 $15.00
https://doi.org/10.1145/3386569.3392434

1 INTRODUCTION

Compatible meshes possess the same connectivity structures. The task of compatible remeshing is fundamental in many geometric processing and scientific computing applications, such as shape morphing, attribute transfers, and synchronized model editing.

Given two 3D meshes with a sparse set of corresponding landmarks, high-quality compatible meshes usually satisfy the following properties: (1) the two output meshes have identical connectivity structures; (2) the piecewise linear mapping between two output meshes has low distortion; (3) the two output meshes are fairly regular; (4) the approximation error between each input mesh and the corresponding output mesh is bounded; and (5) the complexity of the output meshes is low. Satisfying the fourth property, which

Exact and Efficient Polyhedral Envelope Containment Check

BOLUN WANG, Beihang University, China and New York University, USA
TESEO SCHNEIDER, New York University, USA
YIXIN HU, New York University, USA
MARCO ATTENE, Italian National Research Council, Italy
DANIELE PANOZZO, New York University, USA

We introduce a new technique to check containment of a triangle within an envelope built around a given triangle mesh. While existing methods conservatively check containment within a Euclidean envelope, our approach makes use of a non-Euclidean envelope where containment can be checked both *exactly* and *efficiently*. Exactness is crucial to address major robustness issues in existing geometry processing algorithms, which we demonstrate by integrating our technique in two surface triangle remeshing algorithms and a volumetric tetrahedral meshing algorithm. We provide a quantitative comparison of our method and alternative algorithms, showing that our solution, in addition to being exact, is also more efficient. Indeed, while containment within large envelopes can be checked in a comparable time, we show that our algorithm outperforms alternative methods when the envelope becomes thin.

CCS Concepts: • **Mathematics of computing** → **Mesh generation**.

Additional Key Words and Phrases: Robust Geometric Computation, Geometric Predicates, Shape Proximity

ACM Reference Format:
Bolun Wang, Teseo Schneider, Yixin Hu, Marco Attene, and Daniele Panozzo. 2020. Exact and Efficient Polyhedral Envelope Containment Check. *ACM Trans. Graph.* 39, 4, Article 114 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392426

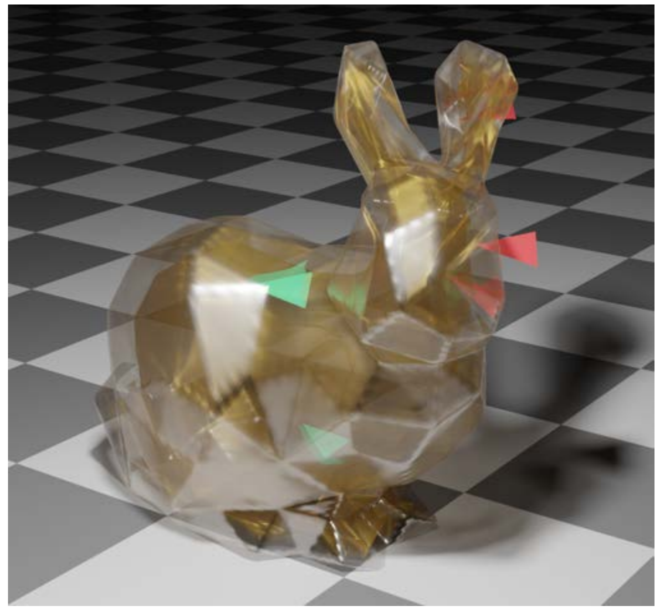


Fig. 1. Our method *exactly* detects if a triangle is inside (green) or outside (red) of an envelope (glass shell) of a bunny model (bronze).

1 INTRODUCTION

The computation of distances between surfaces is a basic building block in geometry processing. In particular, the computation of the Hausdorff distance between an individual triangle $\mathcal{T}$ and a triangle mesh $\mathcal{M}$ is often used by meshing and remeshing algorithms (e.g., [Cheng et al. 2019; Hu et al. 2020, 2018]) to ensure geometric preservation up to a small distance ϵ. This distance allows algorithms to smooth out small details, fill small gaps, remove noise, and perform other operations to generate a high quality mesh, while at the same time bounding the geometrical approximation error. This bound is used, for example, in graphics applications to ensure sub pixels accuracy, or in finite element analysis to bound the error on the solution.

The Euclidean ϵ-envelope is the space of all points whose L^2 distance from a reference surface is less than ϵ (Figure 2). While checking if a point is contained within the envelope is a simple task, checking if an edge or triangle is contained within the envelope is a challenging problem, despite its apparent simplicity.

Many existing algorithms in the literature perform this operation *inexactly* (e.g. by sampling the triangles), whereas just a few can be implemented exactly. A major limitation of inexact checks is that the running time (and memory usage) depends on ϵ: a thinner envelope will require more computations (e.g., more sampling points, larger number of refinements) to compensate for inaccuracy. This fact makes inexact checks impracticable (in terms of both memory and running time) for thin envelopes (Figure 23).

Additionally, while an inexact check is sufficient for certain applications, we discovered that it is problematic when used for remeshing. Remeshing algorithms use the envelope check during local operations, preventing any operation that will move the tracked surface outside of the envelope. Thus, these algorithms are based on a strong invariant as they assume that all the triangles remain

Authors' addresses: Bolun Wang, Beihang University, China, New York University, USA, wangbolun@buaa.edu.cn; Teseo Schneider, New York University, USA, teseo.schneider@nyu.edu; Yixin Hu, New York University, USA, yixin.hu@nyu.edu; Marco Attene, Italian National Research Council, Italy, marco.attene@ge.imati.cnr.it; Daniele Panozzo, New York University, USA, panozzo@nyu.edu.

0730-0301/2020/7-ART114 $15.00
https://doi.org/10.1145/3386569.3392426

Exertion-Aware Path Generation

WANWAN LI*, George Mason University
BIAO XIE*, University of Massachusetts Boston
YONGQI ZHANG, George Mason University
WALTER MEISS, University of Massachusetts Boston
HAIKUN HUANG, University of Massachusetts Boston
LAP-FAI YU, George Mason University

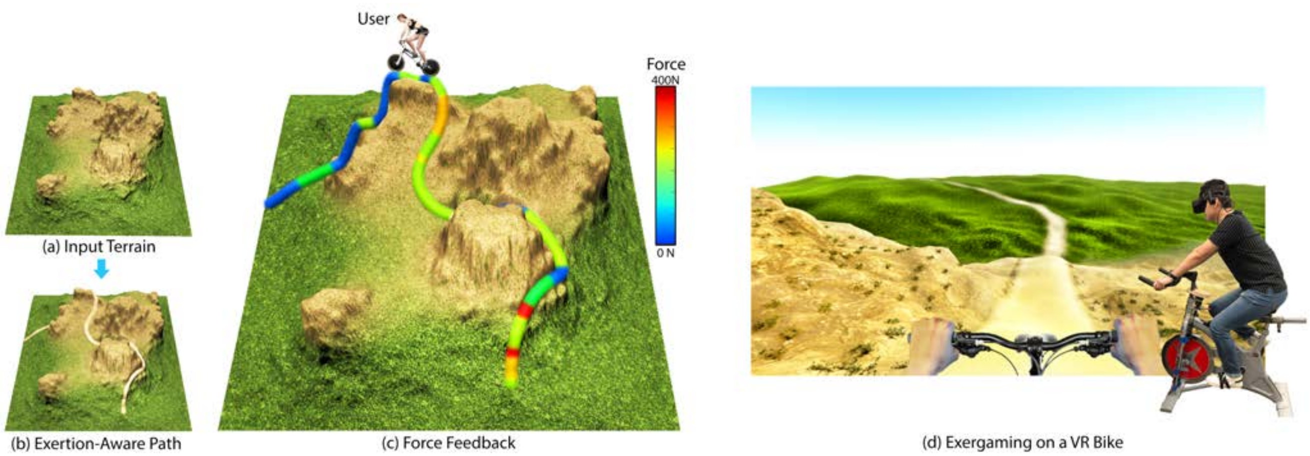


Fig. 1. Given (a) a terrain as input, our approach automatically generates (b) an exertion-aware path that satisfies the user-specified exertion goals such as the the total work and perceived level of path difficulty. (c) Forces that a user experiences along the path. (d) The user bikes along the generated path in virtual reality via a VR headset and our custom-built exercise bike whose feedback force changes with the generated path, giving a highly immersive experience.

We propose a novel approach for generating paths with desired exertion properties, which can be used for delivering highly realistic and immersive virtual reality applications that help users achieve exertion goals. Given a terrain as input, our optimization-based approach automatically generates feasible paths on the terrain which users can bike to perform body training in virtual reality. The approach considers exertion properties such as the total work and the perceived level of path difficulty in generating the paths. To verify our approach, we applied it to generate paths on a variety of terrains with different exertion targets and constraints. To conduct our user studies, we built an exercise bike whose force feedback was controlled by the elevation angle of the generated path over the terrain. Our user study results showed that users found exercising with our generated paths in virtual reality more enjoyable compared to traditional exercising approaches. Their energy expenditure in biking the generated paths also matched with the specified targets, validating the efficacy of our approach.

*Equal Contributors

Authors' addresses: Wanwan Li and Lap-Fai Yu, Computer Science Department, George Mason University, 4400 University Dr, Fairfax, VA 22030, US.

0730-0301/2020/7-ART115 $15.00
https://doi.org/10.1145/3386569.3392393

CCS Concepts: • **Computing methodologies → Graphics systems and interfaces**; **Virtual reality**;

Additional Key Words and Phrases: procedural modeling, level design, path generation, haptics

ACM Reference Format:
Wanwan Li, Biao Xie, Yongqi Zhang, Walter Meiss, Haikun Huang, and Lap-Fai Yu. 2020. Exertion-Aware Path Generation. *ACM Trans. Graph.* 39, 4, Article 115 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392393

1 INTRODUCTION

"The ultimate display would, of course, be a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal."

— Ivan Sutherland, 1965

Ivan Sutherland, inventor of the first virtual reality (VR) head-mounted display, envisioned that the ultimate display would deliver virtual contents that are highly realistic in terms of visuals and haptics. The tremendous efforts of computer graphics researchers have led to the invention of generative models and procedural modeling algorithms capable of synthesizing visually stunning virtual contents such as streets [Chen et al. 2008], roads [Beneš et al. 2014; Galin et al. 2010; Nishida et al. 2016], terrains [Cordonnier et al. 2018; Guérin et al. 2017], cities [Parish and Müller 2001; Vanegas

Fabrication-in-the-Loop Co-Optimization of Surfaces and Styli for Drawing Haptics

MICHAL PIOVARČI, Università della Svizzera italiana, Switzerland
DANNY M. KAUFMAN, Adobe Research
DAVID I. W. LEVIN, University of Toronto, Canada
PIOTR DIDYK, Università della Svizzera italiana, Switzerland

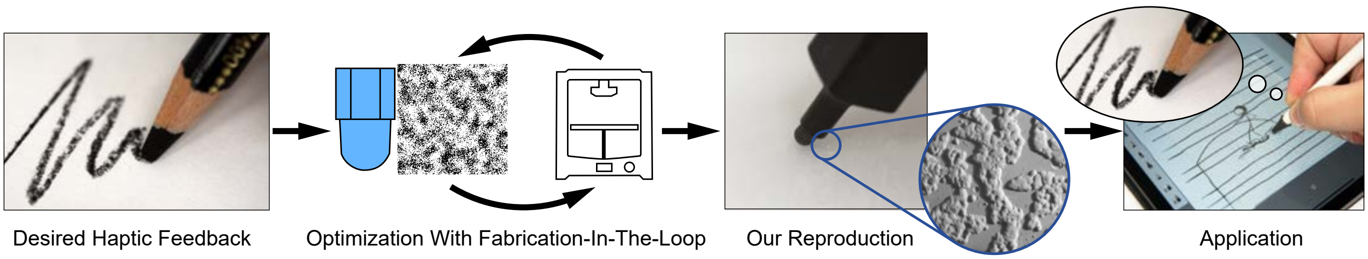


Fig. 1. We propose a data-driven method for mimicking haptic feedback of drawing tools. Our method uses fabrication-in-the-loop design enabled by our data-driven surrogate model which automatically handles exploration-exploitation trade-offs and minimizes the amount of printed samples. The final stylus-surface combinations are manufacturable on commonly available hardware and can be directly integrated into current digital drawing solutions.

Digital drawing tools are now standard in art and design workflows. These tools offer comfort, portability, and precision as well as native integration with digital-art workflows, software, and tools. At the same time, artists continue to work with long-standing, traditional drawing tools. One feature of traditional tools, well-appreciated by many artists and lacking in digital tools, is the specific and diverse range of haptic responses provided by them. Haptic feedback in traditional drawing tools provides unique, per-tool responses that help determine the precision and character of individual strokes. In this work, we address the problem of fabricating digital drawing tools that closely match the haptic feedback of their traditional counterparts. This requires the formulation and solution of a complex, co-optimization of both digital styli and the drawing surfaces they move upon. Here, a potentially direct formulation of this optimization with numerical simulation-in-the-loop is not yet viable. As in many complex design tasks, state-of-the-art methods do not currently offer predictive modeling at rates and scales that can account for the numerous, coupled, physical behaviors governing the haptics of styli and surfaces, nor for the limitations and uncertainties inherent in their fabrication processes. To address these challenges, we propose fabrication-in-the-loop optimization. Critical to making this strategy practical we construct our objective via a Gaussian Process that does not require computing derivatives with respect to design parameters. Our Gaussian Process surrogate model then provides both function estimates and confidence intervals that guide the efficient sampling of our design space. In turn, this sampling critically reduces the numbers of fabricated examples during exploration and automatically handles exploration-exploitation trade-offs. We apply our method to fabricate drawing tools that provide a wide range of haptic feedback, and demonstrate that they are often hard for users to distinguish from their traditional drawing-tool analogs.

CCS Concepts: • **Computing methodologies → Graphics input devices**; **Perception**; • **Human-centered computing → Haptic devices**.

Additional Key Words and Phrases: perception, fabrication, drawing tools, haptics, fabrication-in-the-loop, co-optimization

ACM Reference Format:
Michal Piovarči, Danny M. Kaufman, David I. W. Levin, and Piotr Didyk. 2020. Fabrication-in-the-Loop Co-Optimization of Surfaces and Styli for Drawing Haptics. *ACM Trans. Graph.* 39, 4, Article 116 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392467

Authors' addresses: Michal Piovarči, Università della Svizzera italiana, Via Giuseppe Buffi 13, Lugano, 6900, Switzerland, michal.piovarci@usi.ch; Danny M. Kaufman, Adobe Research, Seattle; David I. W. Levin, University of Toronto, Toronto, Canada; Piotr Didyk, Università della Svizzera italiana, Lugano, Switzerland.

0730-0301/2020/7-ART116 $15.00
https://doi.org/10.1145/3386569.3392467

1 INTRODUCTION

Drawing and writing are among the oldest methods of recording and communicating information. While the culture and technology have evolved, traditional drawing and writing instruments, such as charcoal, crayons, pencil, ballpoint or fountain pens, remain unchanged and are still widely used. They are made of different materials and often used with a different drawing substrate producing a drastically different style and serving a different application. Each of them also has specific haptic feedback easily recognizable by professional and casual users. The relation between the haptic feedback and the tool is not only appreciated by artists but also used to their advantage [Annett et al. 2014; Danna and Velay 2015]. The specific haptic feedback determines the degree to which the artist can control the tool. While in some cases, less precise control is desired to achieve less regular strokes; in other cases, excellent control is critical. The

Fast Tetrahedral Meshing in the Wild

YIXIN HU, New York University, USA
TESEO SCHNEIDER, New York University, USA
BOLUN WANG, Beihang University, China and New York University, USA
DENIS ZORIN, New York University, USA
DANIELE PANOZZO, New York University, USA

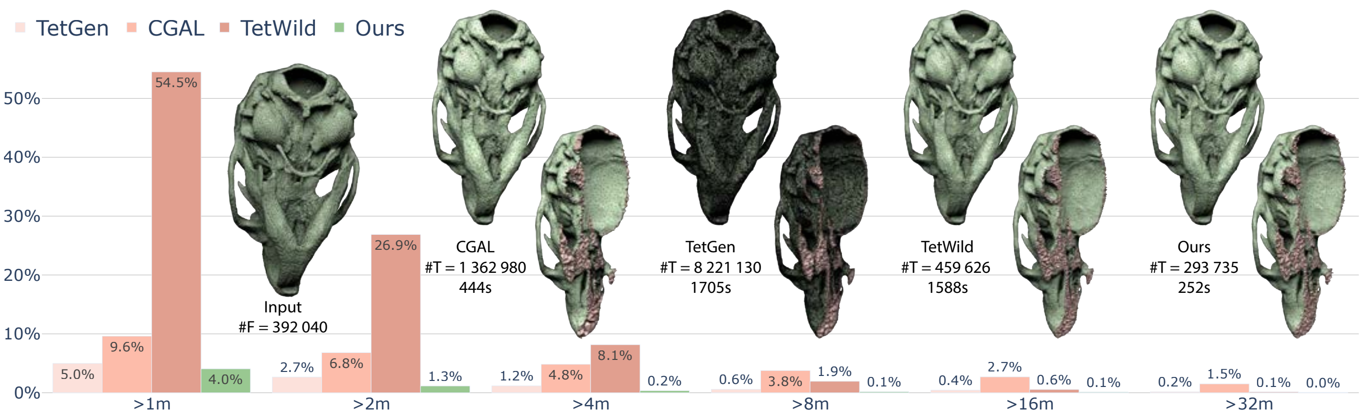


Fig. 1. The bar charts show the percentage of models requiring more than the indicated time for the different approaches over 4 540 inputs (the subset of Thingi10k where all 4 compared algorithms succeed). Our algorithm successfully meshes 98.7% of the input models in less than 2 minutes, and processes all models within 32 minutes. The comparison has been done using the experimental setup of TetWild [Hu et al. 2018] and selecting a similar target resolution for all methods. The CGAL surface approximation parameter has been selected to be comparable to the envelope size used for TetWild and for our method. The images above the plot show a mouse skull model (from micro-CT) tetrahedralized with fTetWild (right) compared with other popular tetrahedral meshing algorithms.

We propose a new tetrahedral meshing method, fTetWild, to convert triangle soups into high-quality tetrahedral meshes. Our method builds on the TetWild algorithm, replacing the rational triangle insertion with a new incremental approach to construct and optimize the output mesh, interleaving triangle insertion and mesh optimization. Our approach makes it possible to maintain a valid floating-point tetrahedral mesh at all algorithmic stages, eliminating the need for costly constructions with rational numbers used by TetWild, while maintaining full robustness and similar output quality. This allows us to improve on TetWild in two ways. First, our algorithm is significantly faster, with running time comparable to less robust Delaunay-based tetrahedralization algorithms. Second, our algorithm is guaranteed to produce a valid tetrahedral mesh with floating-point vertex coordinates, while TetWild produces a valid mesh with rational coordinates which is not guaranteed to be valid after floating-point conversion. As a trade-off, our algorithm no longer guarantees that all input triangles are present in the output mesh, but in practice, as confirmed by our tests on the Thingi10k dataset, the algorithm always succeeds in inserting all input triangles.

CCS Concepts: • **Mathematics of computing → Mesh generation**.

Additional Key Words and Phrases: Mesh Generation, Tetrahedral Meshing, Robust Geometry Processing

ACM Reference Format:
Yixin Hu, Teseo Schneider, Bolun Wang, Denis Zorin, and Daniele Panozzo. 2020. Fast Tetrahedral Meshing in the Wild. *ACM Trans. Graph.* 39, 4, Article 117 (July 2020), 18 pages. https://doi.org/10.1145/3386569.3392385

Authors' addresses: Yixin Hu, New York University, USA, yixin.hu@nyu.edu; Teseo Schneider, New York University, USA, teseo.schneider@nyu.edu; Bolun Wang, Beihang University, China, New York University, USA, wangbolun@buaa.edu.cn; Denis Zorin, New York University, USA, dzorin@cs.nyu.edu; Daniele Panozzo, New York University, USA, panozzo@nyu.edu.

0730-0301/2020/7-ART117 $15.00
https://doi.org/10.1145/3386569.3392385

1 INTRODUCTION

Tetrahedral meshes are commonly used in graphics and engineering applications. Tetrahedral meshing algorithms usually take a 3D surface triangle mesh as input and output a volumetric tetrahedral mesh filling the volume bounded by the input mesh. Traditional tetrahedral meshing algorithms have strong assumptions on the input, requiring it to be a closed manifold, free of self-intersections and numerical unstably close elements, and so on. However, those assumptions often do not hold on imperfect 3D geometric data in the wild.

The recently proposed Tetrahedral Meshing in the Wild (TetWild) [Hu et al. 2018] algorithm makes it possible to reliably tetrahedralize

Graph2Plan: Learning Floorplan Generation from Layout Graphs

RUIZHEN HU, Shenzhen University
ZEYU HUANG, Shenzhen University
YUHAN TANG, Shenzhen University
OLIVER VAN KAICK, Carleton University
HAO ZHANG, Simon Fraser University
HUI HUANG*, Shenzhen University

Fig. 1. Our deep neural network Graph2Plan is a learning framework for automated floorplan generation from layout graphs. The trained network can generate floorplans based on an input building boundary only (a-b), like in previous works. In addition, we allow users to add a variety of constraints such as room counts (c), room connectivity (d), and other layout graph edits. Multiple generated floorplans which fulfill the input constraints are shown.

We introduce a learning framework for automated floorplan generation which combines generative modeling using deep neural networks and user-in-the-loop designs to enable human users to provide sparse design constraints. Such constraints are represented by a *layout graph*. The core component of our learning framework is a deep neural network, Graph2Plan, which converts a layout graph, along with a building boundary, into a floorplan that fulfills both the layout and boundary constraints. Given an input building boundary, we allow a user to specify room counts and other layout constraints, which are used to retrieve a set of floorplans, with their associated layout graphs, from a database. For each retrieved layout graph, along with the input boundary, Graph2Plan first generates a corresponding raster floorplan image, and then a refined set of boxes representing the rooms. Graph2Plan is trained on RPLAN, a large-scale dataset consisting of 80K annotated floorplans. The network is mainly based on convolutional processing over both the layout graph, via a graph neural network (GNN), and the input building boundary, as well as the raster floorplan images, via conventional image convolution. We demonstrate the quality and versatility of our floorplan generation framework in terms of its ability to cater to different user inputs. We conduct both qualitative and quantitative evaluations, ablation studies, and comparisons with state-of-the-art approaches.

CCS Concepts: • **Computing methodologies → Shape modeling**; *Neural networks*.

Additional Key Words and Phrases: floorplan generation, layout graph, deep generative modeling

ACM Reference Format:
Ruizhen Hu, Zeyu Huang, Yuhan Tang, Oliver van Kaick, Hao Zhang, and Hui Huang. 2020. Graph2Plan: Learning Floorplan Generation from Layout Graphs. *ACM Trans. Graph.* 39, 4, Article 118 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392391

1 INTRODUCTION

One of the hottest recent trends in the field of designs, in particular, architectural designs, is the adoption of AI and machine learning techniques. The volume and efficiency afforded by automated and AI-enabled generative models are expected to complement and enrich the architects' workflow, providing a profound and long-lasting impact on the design process. As one of the most fundamental elements of architecture, floor and building plans have drawn recent

*Corresponding author: Hui Huang (hhzhiyan@gmail.com)

Authors' addresses: Ruizhen Hu, College of Computer Science & Software Engineering, Shenzhen University, ruizhen.hu@gmail.com; Zeyu Huang, Shenzhen University; Yuhan Tang, Shenzhen University; Oliver van Kaick, Carleton University; Hao Zhang, Simon Fraser University; Hui Huang, College of Computer Science & Software Engineering, Shenzhen University.

0730-0301/2020/7-ART118 $15.00
https://doi.org/10.1145/3386569.3392391

Inter-Surface Maps via Constant-Curvature Metrics

PATRICK SCHMIDT, RWTH Aachen University
MARCEL CAMPEN, Osnabrück University
JANIS BORN, RWTH Aachen University
LEIF KOBBELT, RWTH Aachen University

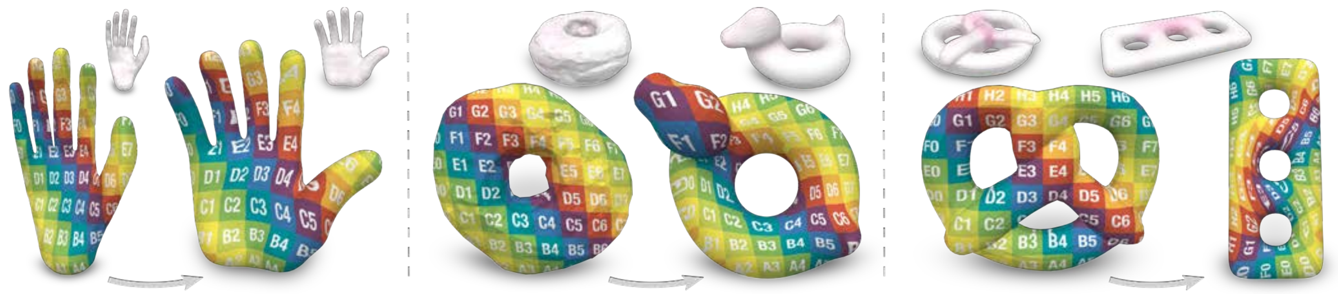


Fig. 1. Visualization of inter-surface maps for pairs of surfaces of varying genus, optimized for low distortion while guaranteeing bijectivity. We represent and optimize such maps flexibly and compactly via discrete constant-curvature metrics of spherical (genus 0), flat (genus 1), or hyperbolic (genus 2+) type.

We propose a novel approach to represent maps between two discrete surfaces of the same genus and to minimize intrinsic mapping distortion. Our maps are well-defined at every surface point and are guaranteed to be continuous bijections (surface homeomorphisms). As a key feature of our approach, only the images of vertices need to be represented explicitly, since the images of all other points (on edges or in faces) are properly defined implicitly. This definition is via unique geodesics in metrics of constant Gaussian curvature. Our method is built upon the fact that such metrics exist on surfaces of arbitrary topology, without the need for any cuts or cones (as asserted by the uniformization theorem). Depending on the surfaces' genus, these metrics exhibit one of the three classical geometries: Euclidean, spherical or hyperbolic. Our formulation handles constructions in all three geometries in a unified way. In addition, by considering not only the vertex images but also the discrete metric as degrees of freedom, our formulation enables us to simultaneously optimize the images of these vertices and images of all other points.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; **Mesh models**; **Mesh geometry models**; *Shape modeling*.

Additional Key Words and Phrases: cross-parametrization, surface parametrization, mesh overlay, bijection, discrete homeomorphism

ACM Reference Format:
Patrick Schmidt, Marcel Campen, Janis Born, and Leif Kobbelt. 2020. Inter-Surface Maps via Constant-Curvature Metrics. *ACM Trans. Graph.* 39, 4, Article 119 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392399

Author's addresses: Patrick Schmidt, Janis Born, Leif Kobbelt, Visual Computing Institute, RWTH Aachen University, Germany; Marcel Campen, Institute of Computer Science, Osnabrück University, Germany.

0730-0301/2020/7-ART119 $15.00
https://doi.org/10.1145/3386569.3392399

1 INTRODUCTION

Maps between surfaces have a variety of uses in Computer Graphics and Geometry Processing. Classical applications include the transfer of various types of information between surfaces, such as textures, geometric detail, deformations, or tessellations. The parametrization or registration of exemplars over a common base model is another application scenario. Such inter-surface maps are furthermore of increasing importance for advanced shape processing tasks, in the context of co-processing of shape collections, or the analysis of frame sequences of time-varying or animated shapes.

In these various fields, inter-surface maps are used as fundamental building blocks of complex methods. Being able to reliably compute, optimize, and provide such maps is therefore of significant practical interest. Properties of maps that commonly are relevant in such applications are bijectivity, continuity, and low distortion.

We present a novel approach to represent inter-surface maps with guaranteed bijectivity and continuity (i.e., surface homeomorphisms) and a method to optimize such maps for low distortion in a direct manner. Our approach is general in that it supports discrete surfaces (triangle meshes) of arbitrary genus.

1.1 Piecewise Maps

Piecewise-linear maps are a common choice for mapping from a discrete surface (triangle mesh) to the plane, e.g. for surface parametrization purposes. These maps are easily represented by the images of the vertices alone, i.e. a finite list of 2D point coordinates. The rest of the map is defined implicitly via barycentric interpolation: the image of any non-vertex surface point $\mathbf{p} = \alpha\mathbf{a} + \beta\mathbf{b} + (1-\alpha-\beta)\mathbf{c}$, lying in a triangle with vertices $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, is defined as

$$f(\mathbf{p}) = \alpha f(\mathbf{a}) + \beta f(\mathbf{b}) + (1-\alpha-\beta) f(\mathbf{c}).$$

The fact that the map can be represented in this compact manner, and that important quantities like the map's Jacobian are linear in

Lifting Simplices to Find Injectivity

XINGYI DU, Washington University in St. Louis, USA
NOAM AIGERMAN and QINGNAN ZHOU, Adobe Research, USA
SHAHAR Z. KOVALSKY, Duke University, USA
YAJIE YAN, Facebook, USA
DANNY M. KAUFMAN, Adobe Research, USA
TAO JU, Washington University in St. Louis, USA

Mapping a source mesh into a target domain while preserving local injectivity is an important but highly non-trivial task. Existing methods either require an already-injective starting configuration, which is often not available, or rely on sophisticated solving schemes. We propose a novel energy form, called Total Lifted Content (TLC), that is equipped with theoretical properties desirable for injectivity optimization. By lifting the simplices of the mesh into a higher dimension and measuring their contents (2D area or 3D volume) there, TLC is smooth over the entire embedding space and its global minima are always injective. The energy is simple to minimize using standard gradient-based solvers. Our method achieved 100% success rate on an extensive benchmark of embedding problems for triangular and tetrahedral meshes, on which existing methods only have varied success.

CCS Concepts: • **Computing methodologies → Mesh models**.

Additional Key Words and Phrases: Parameterization, injective embedding

ACM Reference Format:
Xingyi Du, Noam Aigerman, Qingnan Zhou, Shahar Z. Kovalsky, Yajie Yan, Danny M. Kaufman, and Tao Ju. 2020. Lifting Simplices to Find Injectivity. *ACM Trans. Graph.* 39, 4, Article 120 (July 2020), 17 pages. https://doi.org/10.1145/3386569.3392484

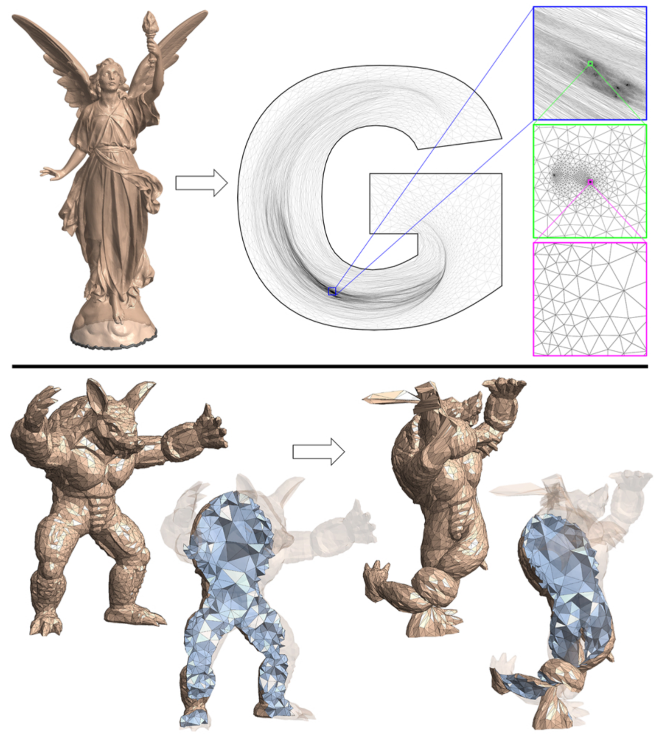


Fig. 1. Injectively mapping a complex surface mesh (Lucy, 48K vertices) to a non-convex boundary (letter "G", with zoom-ins), at the top, and mapping a tetrahedral mesh (Armadillo, 6K vertices) to a highly deformed target surface, at the bottom, as a result of minimizing our novel energy. These two examples are part of our new benchmark data set.

1 INTRODUCTION

Computing constrained mappings between domains is a fundamental task, performed across a wide range of geometric and physical applications ranging from parameterization and UV-mapping, to deformation modeling and the simulation of elastica. In all of these applications, it is in most cases critical to generate a one-to-one, *injective* mapping. This ensures that the inverse map exists and that the correspondence between domains is well-defined. Injectivity is critical for various applications, such as painting textures in UV space, co-analyzing shapes based on correspondences, obtaining good-looking deformations, and generating physically correct simulations of materials, to name just a few.

Most of the time, (local) injectivity is formulated computationally as preservation of all mesh simplices' orientation, *i.e.*, no triangle or tetrahedron is flipped. Unfortunately, the injectivity constraint is not only highly non-convex, but also an *open* set, making any optimization involving it non-trivial. As a result, many mapping and deformation algorithms focus on *preserving* triangle's orientation while improving the map's quality, i.e., by minimizing distortion measures [Rabinovich et al. 2017; Zhu et al. 2018] that also act as a *barrier* that pushes them away from degenerating triangles on the closure of the locally-injective set. This in turn entails that they require a *feasible* embedding - one that is locally injective and satisfies all given constraints - as initialization to begin the minimization process. Indeed, if the input is non-injective, most distortion metrics that act as a barrier to prevent triangles from inverting also fight against *un-inverting* the initially-inverted triangles.

Authors' addresses: Xingyi Du, Washington University in St. Louis, USA, du.xingyi@wustl.edu; Noam Aigerman; Qingnan Zhou, Adobe Research, USA; Shahar Z. Kovalsky, Duke University, USA; Yajie Yan, Facebook, USA; Danny M. Kaufman, Adobe Research, USA; Tao Ju, Washington University in St. Louis, USA.

0730-0301/2020/7-ART120 $15.00
https://doi.org/10.1145/3386569.3392484

LoopyCuts: Practical Feature-Preserving Block Decomposition for Strongly Hex-Dominant Meshing

MARCO LIVESU*, CNR IMATI, Italy
NICO PIETRONI*, University of Technology Sydney, Australia
ENRICO PUPPO, University of Genoa, Italy
ALLA SHEFFER, University of British Columbia, Canada
PAOLO CIGNONI, CNR ISTI, Italy

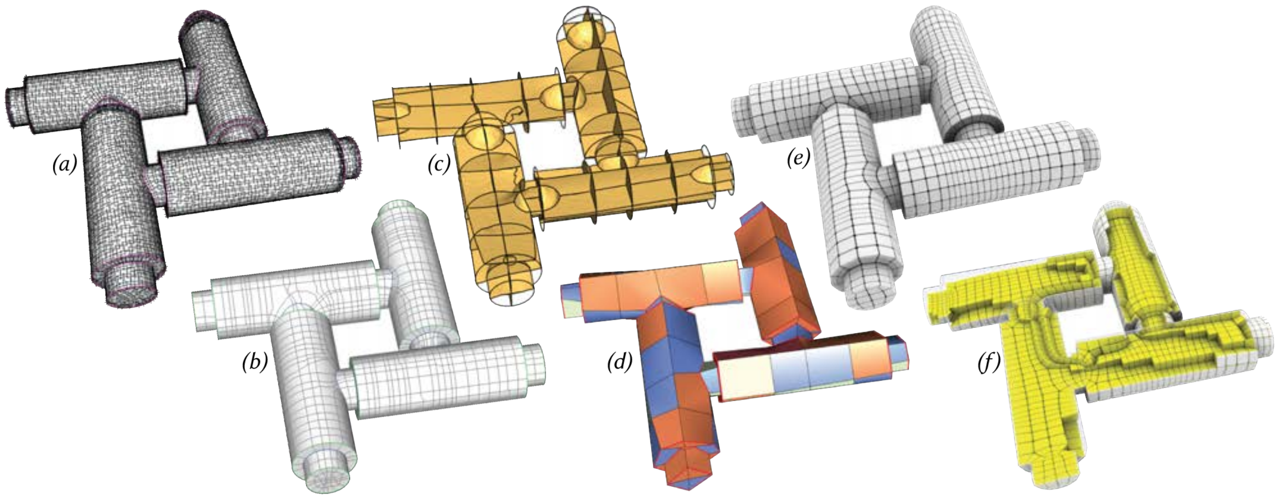


Fig. 1. Given a surface mesh and a curvature and feature aligned cross-field (a) *LoopyCuts* generates a sequence of field-aware cutting loops (b), and uses these loops to generate solid cuts through the object (c), decomposing the model into a metamesh consisting of hex (green), prism (blue) and other (orange) simple blocks (d). It converts the metamesh into a hex-mesh via midpoint refinement. The output hex-mesh (e,f) is well-shaped and well-aligned with the input field.

We present a new fully automatic block-decomposition algorithm for feature-preserving, *strongly* hex-dominant meshing, that yields results with a drastically larger percentage of hex elements than prior art. Our method is guided by a surface field that conforms to both surface curvature and feature lines, and exploits an ordered set of cutting loops that evenly cover the input surface, defining an arrangement of loops suitable for hex-element generation. We decompose the solid into coarse blocks by iteratively cutting it with surfaces bounded by these loops. The vast majority of the obtained blocks can be turned into hexahedral cells via simple midpoint subdivision. Our method produces pure hexahedral meshes in approximately 80% of the cases, and hex-dominant meshes with less than 2% non-hexahedral cells in the remaining cases. We demonstrate the robustness of our method on 70+ models, including CAD objects with features of various complexity, organic and synthetic shapes, and provide extensive comparisons to prior art, demonstrating its superiority.

CCS Concepts: • **Computing methodologies → Mesh models**; **Mesh geometry models**; **Volumetric models**; *Shape analysis*.

Additional Key Words and Phrases: mesh generation, volumetric meshing, hex-meshing, hex-dominant meshing, cross field

ACM Reference Format:
Marco Livesu, Nico Pietroni, Enrico Puppo, Alla Sheffer, and Paolo Cignoni. 2020. *LoopyCuts*: Practical Feature-Preserving Block Decomposition for Strongly Hex-Dominant Meshing. *ACM Trans. Graph.* 39, 4, Article 121 (July 2020), 17 pages. https://doi.org/10.1145/3386569.3392472

*joint first authors

Authors' addresses: Marco Livesu, CNR IMATI, Genoa, Italy, marco.livesu@gmail.com; Nico Pietroni, University of Technology Sydney, Sydney, Australia; Enrico Puppo, University of Genoa, Genoa, Italy; Alla Sheffer, University of British Columbia, Vancouver, Canada; Paolo Cignoni, CNR ISTI, Pisa, Italy.

0730-0301/2020/7-ART121 $15.00
https://doi.org/10.1145/3386569.3392472

1 INTRODUCTION

Hexahedral and hex-dominant volumetric meshing of 3D shapes is a well investigated, yet still open, research topic. At their core, hexahedral meshing algorithms balance fidelity to the input surface geometry against element quality. They seek to generate meshes with well shaped, or box-like, elements whose outer surface closely aligns with that of the input model. To achieve high surface fidelity

MGCN: Descriptor Learning using Multiscale GCNs

YIQUN WANG, NLPR, Institute of Automation, CAS; School of AI, University of Chinese Academy of Sciences; KAUST
JING REN, KAUST
DONG-MING YAN*, NLPR, Institute of Automation, CAS; School of AI, University of Chinese Academy of Sciences
JIANWEI GUO, NLPR, Institute of Automation, CAS; School of AI, University of Chinese Academy of Sciences
XIAOPENG ZHANG, NLPR, Institute of Automation, CAS; School of AI, University of Chinese Academy of Sciences
PETER WONKA, KAUST

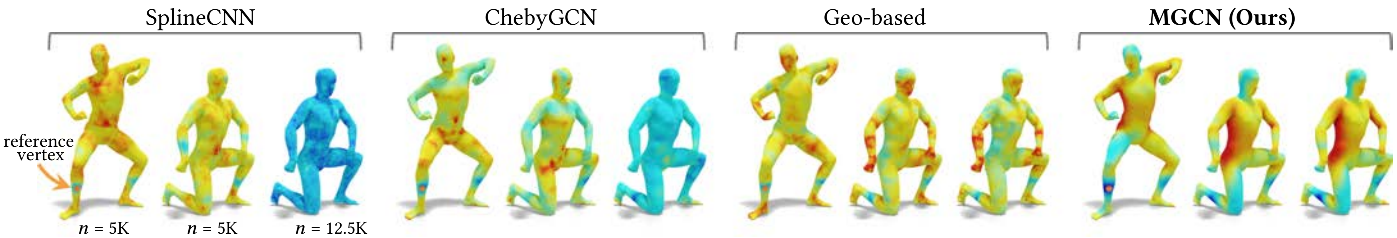


Fig. 1. Visualization of dissimilarity maps. For a vertex highlighted by the orange arrow, we take its learned desriptor and visualize the difference to other vertex descriptors on the same shape, another shape with the same 5K resolution, and the other shape in a different 12.5K resolution. We compare four different learned descriptors, from left to right: SplineCNN, ChebyGCN, Geodesic-based method [Wang et al. 2019a], and MGCN. All networks are trained on 5K and tested on 5K and 12.5K resolution. We can see that our network MGCN is most consistent between different resolutions.

We propose a novel framework for computing descriptors for characterizing points on three-dimensional surfaces. First, we present a new non-learned feature that uses graph wavelets to decompose the Dirichlet energy on a surface. We call this new feature *Wavelet Energy Decomposition Signature* (WEDS). Second, we propose a new *Multiscale Graph Convolutional Network* (MGCN) to transform a non-learned feature to a more discriminative descriptor. Our results show that the new descriptor WEDS is more discriminative than the current state-of-the-art non-learned descriptors and that the combination of WEDS and MGCN is better than the state-of-the-art learned descriptors. An important design criterion for our descriptor is the robustness to different surface discretizations including triangulations with varying numbers of vertices. Our results demonstrate that previous graph convolutional networks significantly overfit to a particular resolution or even a particular triangulation, but MGCN generalizes well to different surface discretizations. In addition, MGCN is compatible with previous descriptors and it can also be used to improve the performance of other descriptors, such as the heat kernel signature, the wave kernel signature, or the local point signature.

*Corresponding author.

Authors' addresses: Yiqun Wang, NLPR, Institute of Automation, CAS; School of AI, University of Chinese Academy of Sciences; KAUST, yiqun.wang@nlpr.ia.ac.cn; Jing Ren, KAUST, jing.ren@kaust.edu.sa; Dong-Ming Yan, NLPR, Institute of Automation, CAS; School of AI, University of Chinese Academy of Sciences, yandongming@gmail.com; Jianwei Guo, NLPR, Institute of Automation, CAS; School of AI, University of Chinese Academy of Sciences, jianwei.guo@nlpr.ia.ac.cn; Xiaopeng Zhang, NLPR, Institute of Automation, CAS; School of AI, University of Chinese Academy of Sciences, xiaopeng.zhang@ia.ac.cn; Peter Wonka, KAUST, pwonka@gmail.com.

0730-0301/2020/7-ART122 $15.00
https://doi.org/10.1145/3386569.3392443

CCS Concepts: • **Computing methodologies** → **Shape analysis**.

Additional Key Words and Phrases: Multiscale, Energy Decomposition, Wavelet Convolution, Shape Matching

ACM Reference Format:
Yiqun Wang, Jing Ren, Dong-Ming Yan, Jianwei Guo, Xiaopeng Zhang, and Peter Wonka. 2020. MGCN: Descriptor Learning using Multiscale GCNs. *ACM Trans. Graph.* 39, 4, Article 122 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392443

1 INTRODUCTION

Designing descriptors for surface points is a fundamental problem in geometry processing as descriptors are a building block for many applications, such as shape matching, registration, segmentation, and retrieval.

A good descriptor should satisfy two criteria: (1) The descriptor should be *discriminative* to map similar surface points to similar values and dissimilar surface points to dissimilar values. The definition of similarity depends on the application. In our setting, we consider the very popular requirement that descriptors should be invariant to rigid and near-isometric deformations of the surface. (2) The descriptor should be *robust* to different discretizations of the surface, *e.g.*, meshes of different resolution and triangulation. If the descriptor discriminates surface points based on the discretization, we also say it overfits or lacks generalization.

Generally, we can distiguish two types of descriptor computation: supervised and non-learned. Examples of non-learned descriptors are the *Wave Kernel Signature* (WKS) and the *Heat Kernal Signature* (HKS). While these descriptors are robust to different surface discretization, there is a lot of room for improvement making them more discriminative. This can been done successfully using neural networks to compute supervised descriptors. A very promising type of network architecture are graph convolutional networks, such as

Monte Carlo Geometry Processing: A Grid-Free Approach to PDE-Based Methods on Volumetric Domains

ROHAN SAWHNEY and KEENAN CRANE, Carnegie Mellon University

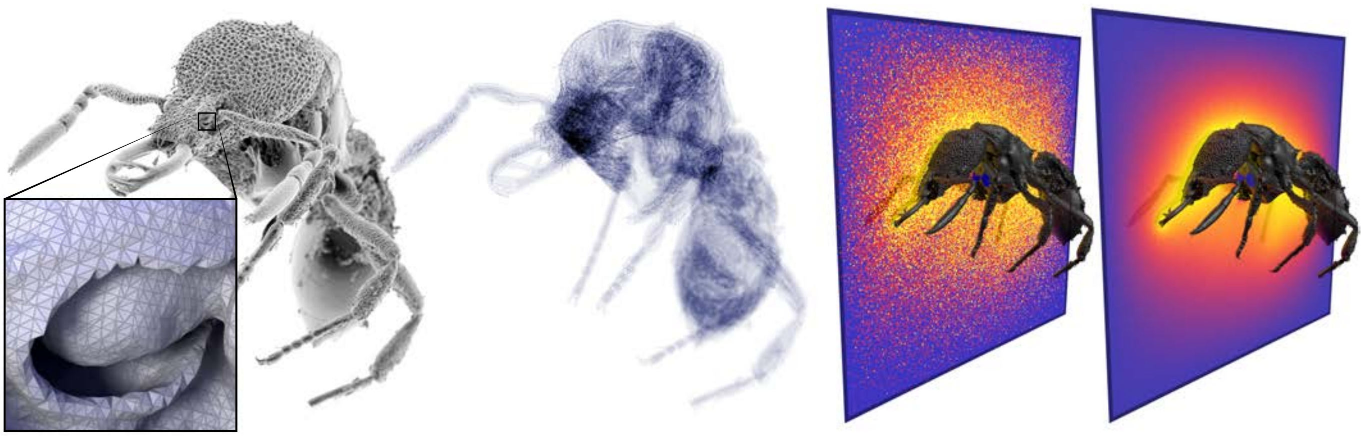


Fig. 1. Real-world geometry has not only rich surface detail *(left)* but also intricate internal structure *(center)*. On such domains, FEM-based geometric algorithms struggle to mesh, setup, and solve PDEs—in this case taking more than 14 hours and 30GB of memory just for a basic Poisson equation. Our Monte Carlo solver uses about 1GB of memory and takes less than a minute to provide a preview *(center right)* that can then be progressively refined *(far right)*. *[Boundary mesh of Fijian strumigenys FJ13 used courtesy of the Economo Lab at OIST.]*

This paper explores how core problems in PDE-based geometry processing can be efficiently and reliably solved via grid-free Monte Carlo methods. Modern geometric algorithms often need to solve Poisson-like equations on geometrically intricate domains. Conventional methods most often mesh the domain, which is both challenging and expensive for geometry with fine details or imperfections (holes, self-intersections, *etc.*). In contrast, grid-free Monte Carlo methods avoid mesh generation entirely, and instead just evaluate closest point queries. They hence do not discretize space, time, nor even function spaces, and provide the exact solution (in expectation) even on extremely challenging models. More broadly, they share many benefits with Monte Carlo methods from photorealistic rendering: excellent scaling, trivial parallel implementation, view-dependent evaluation, and the ability to work with any kind of geometry (including implicit or procedural descriptions). We develop a complete "black box" solver that encompasses integration, variance reduction, and visualization, and explore how it can be used for various geometry processing tasks. In particular, we consider several fundamental linear elliptic PDEs with constant coefficients on solid regions of $\mathbb{R}^n$. Overall we find that Monte Carlo methods significantly broaden the horizons of geometry processing, since they easily handle problems of size and complexity that are essentially hopeless for conventional methods.

CCS Concepts: • **Computing methodologies → Shape analysis**.

Additional Key Words and Phrases: numerical methods, stochastic solvers

Authors' address: Rohan Sawhney; Keenan Crane, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, 15213.

0730-0301/2020/7-ART123
https://doi.org/10.1145/3386569.3392374

ACM Reference Format:
Rohan Sawhney and Keenan Crane. 2020. Monte Carlo Geometry Processing: A Grid-Free Approach to PDE-Based Methods on Volumetric Domains. *ACM Trans. Graph.* 39, 4, Article 123 (July 2020), 18 pages. https://doi.org/10.1145/3386569.3392374

1 INTRODUCTION

The complexity of geometric models has increased dramatically in recent years, but is still far from matching the complexity found in nature—consider, for instance, detailed microstructures that give rise to physical or biological behavior (Fig. 1). PDE-based methods provide powerful tools for processing and analyzing such data, but have not yet reached a point where algorithms "just work": even basic tasks still entail careful preprocessing or parameter tuning, and robust algorithms can exhibit poor scaling in time or memory. Monte Carlo methods provide new opportunities for geometry processing, making a sharp break with traditional finite element methods (FEM). In particular, by avoiding the daunting challenge of *mesh generation* they offer a framework that is highly scalable, parallelizable, and numerically robust, and significantly expands the kind of geometry that can be used in PDE-based algorithms.

Photorealistic rendering experienced an analogous development around the 1990s: finite element radiosity [Goral et al. 1984] gave way to Monte Carlo integration of the light transport equation [Kajiya 1986], for reasons that are nicely summarized by Wann Jensen [2001, Chapter 1]. Although this shift was motivated in part by a desire for more complex illumination, it has also made it possible to work with scenes of extreme *geometric* complexity—modern renderers handle trillions of effective polygons [Georgiev et al. 2018] and, in stark contrast to FEM, yield high-quality results even for

Neural Subdivision

HSUEH-TI DEREK LIU, University of Toronto
VLADIMIR G. KIM, Adobe Research
SIDDHARTHA CHAUDHURI, Adobe Research, IIT Bombay
NOAM AIGERMAN, Adobe Research
ALEC JACOBSON, University of Toronto, Adobe Research

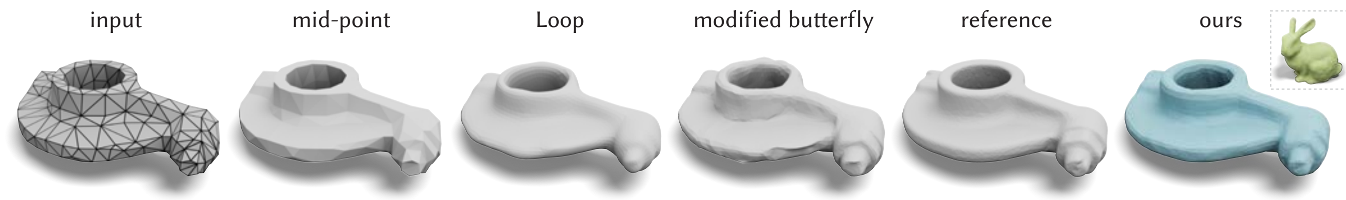


Fig. 1. Our neural subdivision framework performs geometry-aware subdivision, reconstructing the reference rocker arm that we decimated to obtain the coarse input with high accuracy, even though it was only trained on one single model - the Stanford bunny. Neural subdivision does not suffer from the inherent limitations of classic subdivisions, such as volume shrinkage and over-smoothing (Loop [1987]), or amplification of tessellation artifacts (modified butterfly [Zorin et al. 1996]). Throughout this paper, we use green to denote the training shape, and blue for the neural subdivision output.

This paper introduces *Neural Subdivision*, a novel framework for data-driven coarse-to-fine geometry modeling. During inference, our method takes a coarse triangle mesh as input and recursively subdivides it to a finer geometry by applying the fixed topological updates of Loop Subdivision, but predicting vertex positions using a neural network conditioned on the local geometry of a patch. This approach enables us to learn complex non-linear subdivision schemes, beyond simple linear averaging used in classical techniques. One of our key contributions is a novel self-supervised training setup that only requires a set of high-resolution meshes for learning network weights. For any training shape, we stochastically generate diverse low-resolution discretizations of coarse counterparts, while maintaining a bijective mapping that prescribes the exact target position of every new vertex during the subdivision process. This leads to a very efficient and accurate loss function for conditional mesh generation, and enables us to train a method that generalizes across discretizations and favors preserving the manifold structure of the output. During training we optimize for the same set of network weights across all local mesh patches, thus providing an architecture that is not constrained to a specific input mesh, fixed genus, or category. Our network encodes patch geometry in a local frame in a rotation- and translation-invariant manner. Jointly, these design choices enable our method to generalize well, and we demonstrate that even when trained on a single high-resolution mesh our method generates reasonable subdivisions for novel shapes.

Authors' addresses: Hsueh-Ti Derek Liu, University of Toronto, hsuehtil@cs.toronto.edu; Vladimir G. Kim, Adobe Research, vokim@adobe.com; Siddhartha Chaudhuri, Adobe Research, IIT Bombay, sidch@cse.iitb.ac.in; Noam Aigerman, Adobe Research, aigerman@adobe.com; Alec Jacobson, University of Toronto, Adobe Research, jacobson@cs.toronto.edu.

0730-0301/2020/7-ART124 $15.00
https://doi.org/10.1145/3386569.3392418

CCS Concepts: • **Computing methodologies → Shape modeling**; **Machine learning**.

Additional Key Words and Phrases: geometry processing, shape modeling, subdivision surfaces, machine learning

ACM Reference Format:
Hsueh-Ti Derek Liu, Vladimir G. Kim, Siddhartha Chaudhuri, Noam Aigerman, and Alec Jacobson. 2020. Neural Subdivision. *ACM Trans. Graph.* 39, 4, Article 124 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392418

1 INTRODUCTION

Subdivision surfaces are defined by deterministic, recursive upsampling of a discrete surface mesh. Classic methods work by performing two steps: each input mesh element is *divided* into many elements (e.g., one triangle becomes three) by splitting edges and adding vertices. The positions of the mesh vertices are then smoothed by taking a weighted average of their neighbors' positions according to a weighting scheme based purely on the local mesh connectivity. Subdivision surfaces are well studied and have rich theory connecting their limit surfaces (applying an infinite number of subdivide-and-smooth iterations) to traditional splines. They are a standard paradigm in surface modeling tools, allowing modelers to sculpt shapes in a coarse-to-fine manner. A modeler may start with a very coarse cage, adjust vertex positions, then subdivide once, adjust the finer mesh vertices, and repeat this process until satisfied.

While existing subdivision methods are well-suited for this sort of interactive modeling, they fall short when used to automatically upsample a low resolution asset. Without a user's guidance, classic methods will overly smooth the entire shape (see Fig. 1). Popular methods based on simple linear averaging do not identify details to maintain or accentuate during upsampling. They make no use of the geometric context of a local patch of a surface. Furthermore, classic methods based on fixed one-size-fits-all weighting rules are

On Elastic Geodesic Grids and Their Planar to Spatial Deployment

STEFAN PILLWEIN, TU Wien
KURT LEIMER, TU Wien
MICHAEL BIRSAK, KAUST
PRZEMYSLAW MUSIALSKI, NJIT and TU Wien

We propose a novel type of planar–to–spatial deployable structures that we call elastic geodesic grids. Our approach aims at the approximation of freeform surfaces with spatial grids of bent lamellas which can be deployed from a planar configuration using a simple kinematic mechanism. Such elastic structures are easy–to–fabricate and easy–to–deploy and approximate shapes which combine physics and aesthetics. We propose a solution based on networks of geodesic curves on target surfaces and we introduce a set of conditions and assumptions which can be closely met in practice. Our formulation allows for a purely geometric approach which avoids the necessity of numerical shape optimization by building on top of theoretical insights from differential geometry. We propose a solution for the design, computation, and physical simulation of elastic geodesic grids, and present several fabricated small-scale examples with varying complexity. Moreover, we provide an empirical proof of our method by comparing the results to laser-scans of the fabricated models. Our method is intended as a form-finding tool for elastic gridshells in architecture and other creative disciplines and should give the designer an easy-to-handle way for the exploration of such structures.

CCS Concepts: • **Computing methodologies → Shape modeling**; Optimization algorithms.

Additional Key Words and Phrases: geometric modeling, fabrication, elastic deformation, physical simulation, architectural geometry, elastic gridshells, active bending

ACM Reference Format:
Stefan Pillwein, Kurt Leimer, Michael Birsak, and Przemyslaw Musialski. 2020. On Elastic Geodesic Grids and Their Planar to Spatial Deployment. *ACM Trans. Graph.* 39, 4, Article 125 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392490

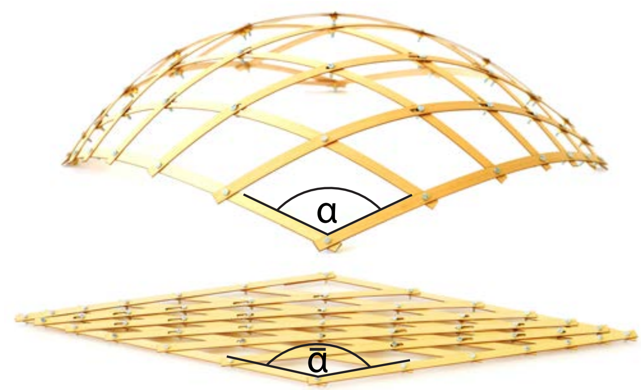


Fig. 1. A deployed elastic geodesic gridshell (top) and its planar lattice in the rest state (bottom) fabricated of wooden lamellas. The deployment of the whole kinematic system is based on changing angle $\overline{\alpha}$, such that $\overline{\alpha} \rightarrow \alpha$.

1 INTRODUCTION

Design and construction of structures composed of curved elastic elements has a long history in the field of architecture. Alongside their aesthetical aspects imposed by nature, they have a lot of functional advantages: they are compact, lightweight and easy to build; nonetheless practicable, durable, and of high structural performance.

They have been utilized for a long time dating back to ancient vernacular architecture for formal as well as for performance reasons, however, the possibilities of their form-finding in the past were limited [Lienhard et al. 2013].

Authors' addresses: Stefan Pillwein, stefan.pillwein@tuwien.ac.at, TU Wien; Kurt Leimer, kurt.leimer@tuwien.ac.at, TU Wien; Michael Birsak, michael.birsak@kaust.edu.sa, KAUST; Przemyslaw Musialski, przem@njit.edu, NJIT and TU Wien.

0730-0301/2020/7-ART125 $15.00
https://doi.org/10.1145/3386569.3392490

Fortunately, the currently available computational capabilities and advances in computer science open up avenues for direct modeling of complex shapes composed of elastically bending members. This goes beyond traditional architectural design and allows to aim at many general purpose products composed of such elements. The range of potential objects encompasses gridshells, formwork, paneling, various types of furniture, sun and rain protectors, pavilions and similar small-scale buildings, home decoration and accessories, like vases, bowls, or lamps, etc., and finally, also elements of future's functional digital fabrics that can be utilized in engineering as well as in fashion.

This vision leads directly to the objective of this paper: a designer provides a target surface and a computational method finds a planar grid of flat lamellas, that—when deployed—approximates the surface well. Figure 1 shows a planar and a deployed grid of wooden strips, where a surface with the curved lamellas being tangential to it can be imagined. The joints between the lamellas allow for rotation and partially also for sliding. As the lamellas connecting opposite edges of the planar boundary quadrilateral are not parallel to each other, the grid is rigid in the plane. Given the flexibility of wooden lamellas with regard to bending and twisting, the grid is not rigid in space. By adjusting only one degree of freedom, for example the angle $\overline{\alpha} \rightarrow \alpha$ at one corner, the planar kinematic configuration *elastically bends* continuously into a spatial gridshell which approximates the desired surface. The deployment process is governed by the rules of physics, seeing the lamellas as thin elastic minimal energy beams, allowed to bend as well as to rotate and slide at their intersections.

Our goal is to find a suitable planar setup of the lamellas that can be deformed into a spatial grid, fitting the target surface as closely as possible. To achieve this goal, we propose a solution based on

Point2Mesh: A Self-Prior for Deformable Meshes

RANA HANOCKA, Tel Aviv University
GAL METZER, Tel Aviv University
RAJA GIRYES, Tel Aviv University
DANIEL COHEN-OR, Tel Aviv University

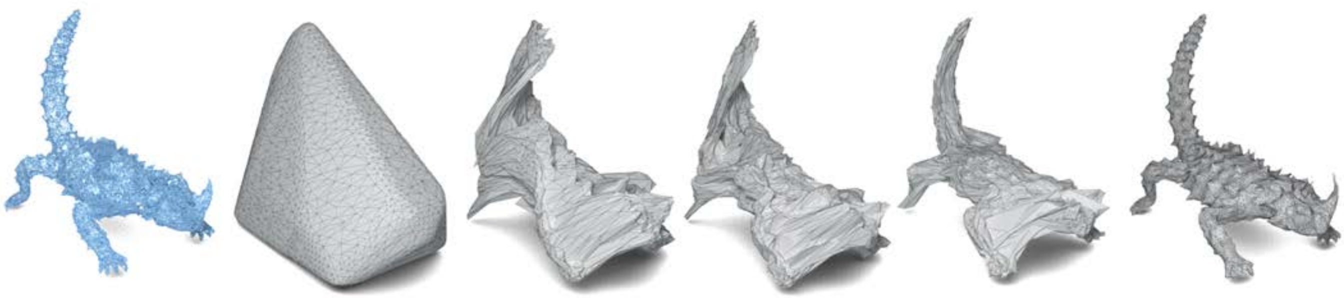


Fig. 1. Starting with an input point cloud (left) and a deformable mesh, we iteratively shrink-wrap the input, leading to a watertight reconstruction.

In this paper, we introduce *Point2Mesh*, a technique for reconstructing a surface mesh from an input point cloud. Instead of explicitly specifying a prior that encodes the expected shape properties, the prior is defined automatically using the input point cloud, which we refer to as a *self-prior*. The self-prior encapsulates reoccurring geometric repetitions from a single shape within the weights of a deep neural network. We optimize the network weights to deform an initial mesh to *shrink-wrap* a single input point cloud. This explicitly considers the entire reconstructed shape, since shared local kernels are calculated to fit the overall object. The convolutional kernels are optimized globally across the entire shape, which inherently encourages local-scale geometric self-similarity across the shape surface. We show that shrink-wrapping a point cloud with a self-prior converges to a desirable solution; compared to a prescribed smoothness prior, which often becomes trapped in undesirable local minima. While the performance of traditional reconstruction approaches degrades in non-ideal conditions that are often present in real world scanning, *i.e.,* unoriented normals, noise and missing (low density) parts, Point2Mesh is robust to non-ideal conditions. We demonstrate the performance of Point2Mesh on a large variety of shapes with varying complexity.

CCS Concepts: • **Computing methodologies → Neural networks**; **Shape analysis**.

Additional Key Words and Phrases: Geometric Deep Learning, Surface Reconstruction, Shape Analysis

Authors' addresses: Rana Hanocka, Tel Aviv University; Gal Metzer, Tel Aviv University; Raja Giryes, Tel Aviv University; Daniel Cohen-Or, Tel Aviv University.

0730-0301/2020/7-ART126 $15.00
https://doi.org/10.1145/3386569.3392415

ACM Reference Format:
Rana Hanocka, Gal Metzer, Raja Giryes, and Daniel Cohen-Or. 2020. Point2Mesh: A Self-Prior for Deformable Meshes. *ACM Trans. Graph.* 39, 4, Article 126 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392415

1 INTRODUCTION

Reconstructing a mesh from a point cloud is a long-standing problem in computer graphics. In recent decades, various approaches have been developed to reconstruct shapes for an array of applications [Berger et al. 2017]. The reconstruction problem is ill-posed, making it necessary to define a prior which incorporates the expected properties of the reconstructed mesh. Traditionally, priors are manually designed to encourage general properties, like piece-wise smoothness or local uniformity.

The recent emergence of deep neural networks carries new promise to bypass manually specified priors. Studies have shown that filtering data with a collection of convolution and pooling layers results in a salient feature representation, even with randomly initialized weights [Saxe et al. 2011; Gaier and Ha 2019]. Convolutions exploit local spatial correlations, which are shared and aggregated across the entire data; while pooling reduces the dimensionality of the learned representation. Since shapes, like natural images, are not random, they have a distinct distribution which fosters the powerful intrinsic properties of CNNs to garner self-similarities.

In this paper, we introduce *Point2Mesh*, a method for reconstructing meshes from point clouds, where the prior is defined automatically by a convolutional neural network (CNN). Instead of explicitly specifying a prior, it is *learned* automatically from a single input point cloud, without relying on any training data or pre-training, in other words, a *self-prior*. In contrast, supervised learning paradigms demand large amounts of input (point cloud) and ground-truth (surface) training pairs (*i.e., data-driven priors*), which often entails modeling the acquisition process. An appealing aspect of Point2Mesh is

Principal Symmetric Meshes

DAVIDE PELLIS*, TU Wien
HUI WANG*, DUT
MARTIN KILIAN, TU Wien
FLORIAN RIST, KAUST, TU Wien
HELMUT POTTMANN, KAUST, TU Wien
CHRISTIAN MÜLLER, TU Wien

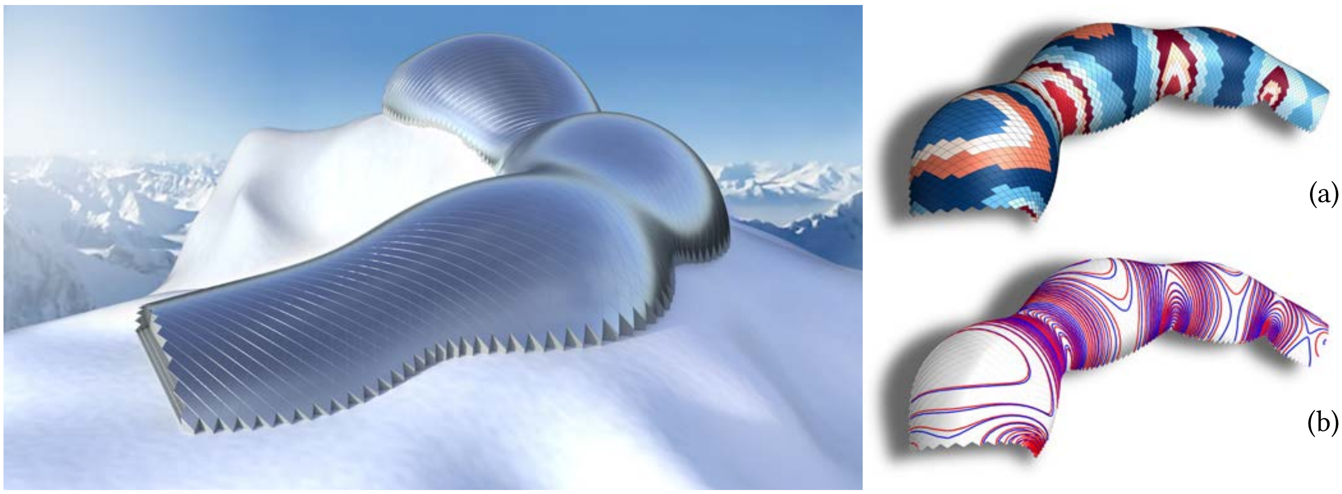


Fig. 1. Architectural design of a modern mountain leisure and tourism center. The geometric shape of the hull is a discrete principal symmetric parametrization (an S-net) of a Weingarten surface of type $a\kappa_1 + b\kappa_2 = c$ (see Sec. 3.5). The big advantage lies in the fabrication of its doubly curved panels by a substantial reduction of necessary molds, as the same mold can be used along curves of constant curvature illustrated by isolines (b). Panels with identical molds are clustered together and highlighted by the same color (a). The network of this S-net assumes constant intersection angles which contributes to its aesthetic.

The isolines of principal symmetric surface parametrizations run symmetrically to the principal directions. We describe two discrete versions of these special nets/quad meshes which are dual to each other and show their usefulness for various applications in the context of fabrication and architectural design. Our discretization of a principal symmetric mesh comes naturally with a family of spheres, the so-called Meusnier and Mannheim spheres. In our representation of principal symmetric meshes, we have direct control over the radii of theses spheres and the intersection angles of the parameter lines. This facilitates tasks such as generating Weingarten surfaces including constant mean curvature surfaces and minimal surfaces. We illustrate the potential of Weingarten surfaces for paneling doubly curved freeform facades by significantly reducing the number of necessary molds. Moreover, we have direct access to curvature adaptive tool paths for cylindrical CNC milling with circular edges as well as flank milling with rotational cones. Furthermore, the construction of curved support structures from congruent circular strips is easily managed by constant sphere radii. The underlying families of spheres are in a natural way discrete curvature spheres in analogy to smooth Möbius and Laguerre geometry which further leads to a novel discrete curvature theory for principal symmetric meshes.

CCS Concepts: • **Computing methodologies → Shape modeling**; *Optimization algorithms.*

Additional Key Words and Phrases: discrete differential geometry, architectural geometry, computational fabrication, paneling, normal curvature, sphere geometries, curvature adaptive milling

ACM Reference Format:
Davide Pellis, Hui Wang, Martin Kilian, Florian Rist, Helmut Pottmann, and Christian Müller. 2020. Principal Symmetric Meshes. *ACM Trans. Graph.* 39, 4, Article 127 (July 2020), 17 pages. https://doi.org/10.1145/3386569.3392446

1 INTRODUCTION

Meshes are omnipresent in geometric computing: as a geometry representation, as a basis for simulations, and for shape optimization.

*Joint first authors

Authors' addresses: Davide Pellis, TU Wien; Hui Wang, DUT; Martin Kilian, TU Wien; Florian Rist, KAUST, TU Wien; Helmut Pottmann, KAUST, TU Wien; Christian Müller, TU Wien.

0730-0301/2020/7-ART127 $15.00
https://doi.org/10.1145/3386569.3392446

Quad-mesh based isometric mappings and developable surfaces

CAIGUI JIANG, KAUST
CHENG WANG, KAUST
FLORIAN RIST, KAUST / TU Wien
JOHANNES WALLNER, TU Graz
HELMUT POTTMANN, KAUST

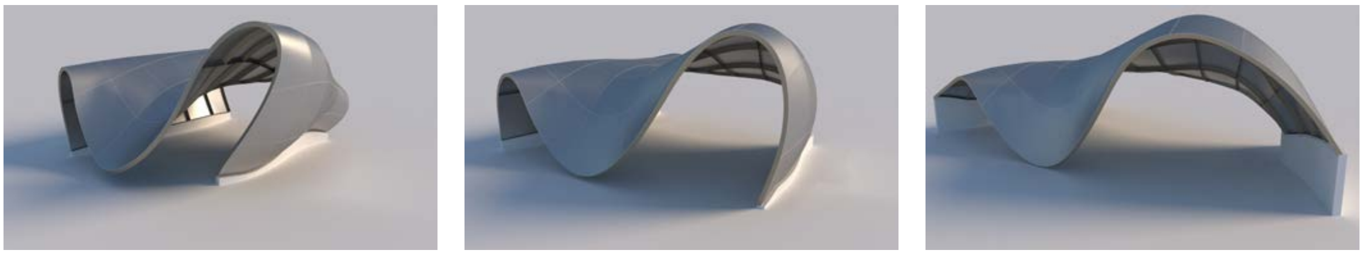


Fig. 1. *Design with developables in freeform architecture.* We propose a simple approach to isometric mappings between quad meshes, which immediately leads to a way to model developable surfaces. For this example we started from a segmentation into panels of part of a freeform concrete shell by Zaha Hadid Architects, built in 2014 in Baku. We approximated it by a quad mesh which is piecewise discrete-developable in the sense of our framework. Afterwards this mesh undergoes an isometric deformation which respects the panelization. Our approach to discrete-developable surfaces is very flexible in the sense that the edges of meshes do not have to be aligned with rulings or otherwise special curves on developables, and can be aligned with boundaries and features instead.

We discretize isometric mappings between surfaces as correspondences between checkerboard patterns derived from quad meshes. This method captures the degrees of freedom inherent in smooth isometries and enables a natural definition of discrete developable surfaces. This definition, which is remarkably simple, leads to a class of discrete developables which is much more flexible in applications than previous concepts of discrete developables. In this paper, we employ optimization to efficiently compute isometric mappings, conformal mappings and isometric bending of surfaces. We perform geometric modeling of developables, including cutting, gluing and folding. The discrete mappings presented here have applications in both theory and practice: We propose a theory of curvatures derived from a discrete Gauss map as well as a construction of watertight CAD models consisting of developable spline surfaces.

CCS Concepts: • **Computing methodologies → Shape modeling**; *Optimization algorithms.*

Additional Key Words and Phrases: discrete differential geometry, computer-aided design, computational fabrication, shape optimization, discrete isometry, developable surface, developable spline surface

Authors' addresses: Caigui Jiang, KAUST; Cheng Wang, KAUST; Florian Rist, KAUST / TU Wien; Johannes Wallner, TU Graz; Helmut Pottmann, KAUST.

0730-0301/2020/7-ART128 $15.00
https://doi.org/10.1145/3386569.3392430

ACM Reference Format:
Caigui Jiang, Cheng Wang, Florian Rist, Johannes Wallner, and Helmut Pottmann. 2020. Quad-mesh based isometric mappings and developable surfaces. *ACM Trans. Graph.* 39, 4, Article 128 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392430

1 INTRODUCTION

The geometric modeling of developable surfaces is a topic attracting attention since many years. One reason for that is the great practical importance of developables, which represent shapes made by bending flat pieces of inextensible sheet material into space. Materials which fall into this category include paper and sufficiently thin plates [Audoly and Pomeau 2009]. New algorithms and computational representations of developables continue to emerge. This steady progress is a sign that the problem of modeling developables still has no complete and satisfactory solution. The mathematical theory of developables is far from simple, which probably accounts for a part of the computational difficulties which occur. The approach to developables presented in this paper is via a systematic theory of isometric mappings. It is based on correspondences between quad meshes, with no specific restrictions on the meshes themselves.

1.1 Overview and Contributions

We propose to manage discrete surfaces and their mappings not directly, via properties of the vertices, edges and faces, but via properties of a checkerboard pattern inscribed in the original mesh. That pattern is created by inserting midpoints of edges – the edge midpoints belonging to a face will always form a parallelogram, regardless of the shape of the original face.

TilinGNN: Learning to Tile with Self-Supervised Graph Neural Network

HAO XU*, KA-HEI HUI*, and CHI-WING FU, The Chinese University of Hong Kong
HAO ZHANG, Simon Fraser University

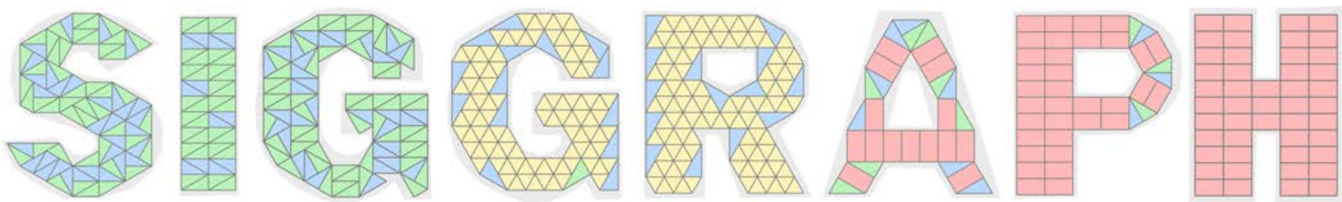


Fig. 1. Our self-supervised neural network, TilinGNN, produces tiling results in time roughly *linear* to the number of candidate tile locations, significantly outperforming traditional combinatorial search methods. The average runtime of our network for tiling a character is only 25.71s. The character shapes to be tiled are shown in grey and different types of tiles are displayed using different colors (note that mirror reflections count as different tile types).

We introduce the first *neural optimization* framework to solve a classical instance of the *tiling* problem. Namely, we seek a *non-periodic* tiling of an arbitrary 2D shape using *one or more* types of tiles—the tiles *maximally* fill the shape's interior without overlaps or holes. To start, we reformulate tiling as a graph problem by modeling candidate tile locations in the target shape as graph nodes and connectivity between tile locations as edges. Further, we build a *graph convolutional neural network*, coined TilinGNN, to progressively propagate and aggregate features over graph edges and predict tile placements. TilinGNN is trained by maximizing the tiling coverage on target shapes, while avoiding overlaps and holes between the tiles. Importantly, our network is *self-supervised*, as we articulate these criteria as loss terms defined on the network outputs, without the need of ground-truth tiling solutions. After training, the runtime of TilinGNN is roughly *linear* to the number of candidate tile locations, significantly outperforming traditional combinatorial search. We conducted various experiments on a variety of shapes to showcase the speed and versatility of TilinGNN. We also present comparisons to alternative methods and manual solutions, robustness analysis, and ablation studies to demonstrate the quality of our approach.

CCS Concepts: • **Computing methodologies** → **Shape modeling**; *Neural networks.*

Additional Key Words and Phrases: Tiling, neural combinatorial optimization, graph neural network

ACM Reference Format:
Hao Xu, Ka-Hei Hui, Chi-Wing Fu, and Hao Zhang. 2020. TilinGNN: Learning to Tile with Self-Supervised Graph Neural Network. *ACM Trans. Graph.* 39, 4, Article 129 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392380

*Both authors contributed equally to the paper.

Authors' addresses: Hao Xu; Ka-Hei Hui; Chi-Wing Fu, The Chinese University of Hong Kong; Hao Zhang, Simon Fraser University.

0730-0301/2020/7-ART129 $15.00
https://doi.org/10.1145/3386569.3392380

1 INTRODUCTION

Many geometry processing tasks in computer graphics require solving discrete and combinatorial optimization problems, e.g., decomposition, packing, set cover, and assignment. Conventional approaches resort to approximation algorithms with guaranteed bounds or heuristic schemes exhibiting favorable average performance. With the rapid adoption of machine learning techniques in all facets of visual computing, an intriguing question is whether difficult combinatorial problems involving geometric primitives can be effectively and efficiently solved using a *machine learning* approach.

In this paper, we explore a learning-based approach to solve a combinatorial geometric optimization problem, *tiling*, which has drawn interests from the computer graphics community in different contexts [Kaplan 2009], e.g., sampling [Kopf et al. 2006; Ostromoukhov 2007], texture generation [Cohen et al. 2003], architectural construction [Eigensatz et al. 2010; Fu et al. 2010; Singh and Schaefer 2010], and puzzle design [Duncan et al. 2017]. In general, tiling refers to the partition of a domain into regions, the tiles, of one or more types. So far, most works on tiling have stayed in the 2D domain; see Figure 2 for some typical examples and applications.

As a first attempt, we focus on *non-periodic* tiling[1] of an arbitrary 2D shape using *one or more* types of tiles. Specifically, we seek a tiling that *maximally* fills the shape's *interior* without overlaps, holes, or tiles exceeding the shape boundary. Even such an elementary

[1]Non-periodicity means that no finite shifts of a tiling can reproduce it. A better known special case of such tilings are *aperiodic* tilings, e.g., Penrose tilings. Aperiodicity has the additional requirement that the tiling cannot contain arbitrarily large periodic patches, which is not necessarily respected by our method (see Figure 1).

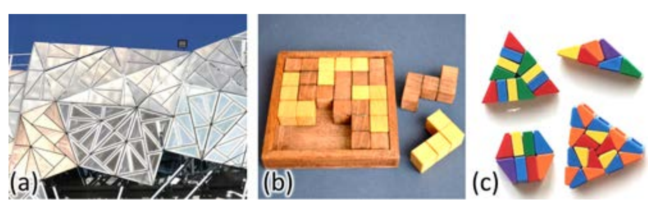


Fig. 2. Applications of tilings: (a) Federation Square in Melbourne; (b) a puzzle called jags and hooks designed by Erhan Cubukcuoglu; and (c) cheese slope mosaic using LEGO bricks from Katie Walker's Flickr page.

Uncertainty Quantification for Multi-Scan Registration

XIANGRU HUANG, The University of Texas at Austin, USA
ZHENXIAO LIANG, The University of Texas at Austin, USA
QIXING HUANG, The University of Texas at Austin, USA

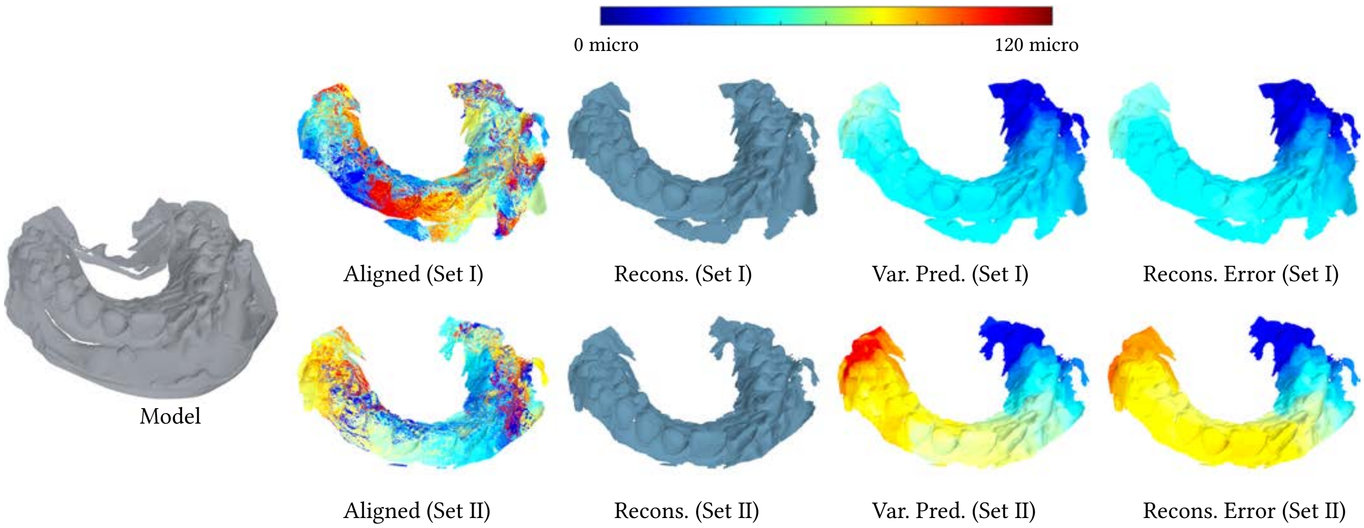


Fig. 1. Two reconstructions from two sets of scans of the same underlying teeth object. Both reconstructions nicely reconstruct local features, yet the bottom reconstruction possesses larger global drifts than the top reconstruction. Such global drifts are hard to detect manually. Our approach computes quantified uncertainties that accurately reveal such global drifts. The predictions are consistent with ground-truth reconstruction errors. From left to right: The ground-truth model obtained by an Artec Micro scanner, the two sets of scans obtained by an Artec Space Spider scanner, the corresponding reconstructions, the predicted uncertainties of the reconstructions, and the ground-truth reconstruction errors.

A fundamental problem in scan-based 3D reconstruction is to align the depth scans under different camera poses into the same coordinate system. While there are abundant algorithms on aligning depth scans, few methods have focused on assessing the quality of a solution. This quality checking problem is vital, as we need to determine whether the current scans are sufficient or not and where to install additional scans to improve the reconstruction. On the other hand, this problem is fundamentally challenging because the underlying ground-truth is generally unavailable, and it is challenging to predict alignment errors such as global drifts manually. In this paper, we introduce a local uncertainty framework for geometric alignment algorithms. Our approach enjoys several appealing properties, such as it does not require re-sampling the input, no need for the underlying ground-truth, informative, and high computational efficiency. We apply this framework to two multi-scan alignment formulations, one minimizes geometric distances between pairs of scans, and another simultaneously aligns the input scans with a deforming model. The output of our approach can be seamlessly integrated with view selection, enabling uncertainty-aware view planning. Experimental results and user studies justify the effectiveness of our approach on both synthetic and real datasets.

CCS Concepts: • **Computing methodologies → Point-based models**; **Shape analysis**.

Additional Key Words and Phrases: uncertainty quantification, multi-scan registration, approximation error, view planning

ACM Reference Format:
Xiangru Huang, Zhenxiao Liang, and Qixing Huang. 2020. Uncertainty Quantification for Multi-Scan Registration. *ACM Trans. Graph.* 39, 4, Article 130 (July 2020), 24 pages. https://doi.org/10.1145/3386569.3392402

1 INTRODUCTION

Reconstructing 3D models from depth scans is a fundamental problem in geometry processing and beyond. A standard 3D reconstruction pipeline typically combines a geometry registration phase and a geometry reconstruction phase (c.f. [Huber 2002]). The goal of geometry registration is to bring depth scans obtained from different camera poses into the same coordinate system. The accuracy of this phase largely determines the quality of the reconstructed 3D model. Existing multi-scan registration techniques have predominantly

Authors' addresses: Xiangru Huang, The University of Texas at Austin, 2317 Speedway, Austin, TX, 78712, USA; Zhenxiao Liang, The University of Texas at Austin, 2317 Speedway, Austin, TX, 78712, USA; Qixing Huang, The University of Texas at Austin, 2317 Speedway, Austin, TX, 78712, USA.

0730-0301/2020/7-ART130 $15.00
https://doi.org/10.1145/3386569.3392402

Variable-width contouring for Additive Manufacturing

SAMUEL HORNUS, Université de Lorraine, CNRS, Inria, LORIA
TIM KUIPERS, Ultimaker and Department of Sustainable Design Engineering, Delft University of Technology
OLIVIER DEVILLERS and MONIQUE TEILLAUD, Université de Lorraine, CNRS, Inria, LORIA
JONÀS MARTÍNEZ, Université de Lorraine, CNRS, Inria, LORIA
MARC GLISSE, Inria
SYLVAIN LAZARD and SYLVAIN LEFEBVRE, Université de Lorraine, CNRS, Inria, LORIA

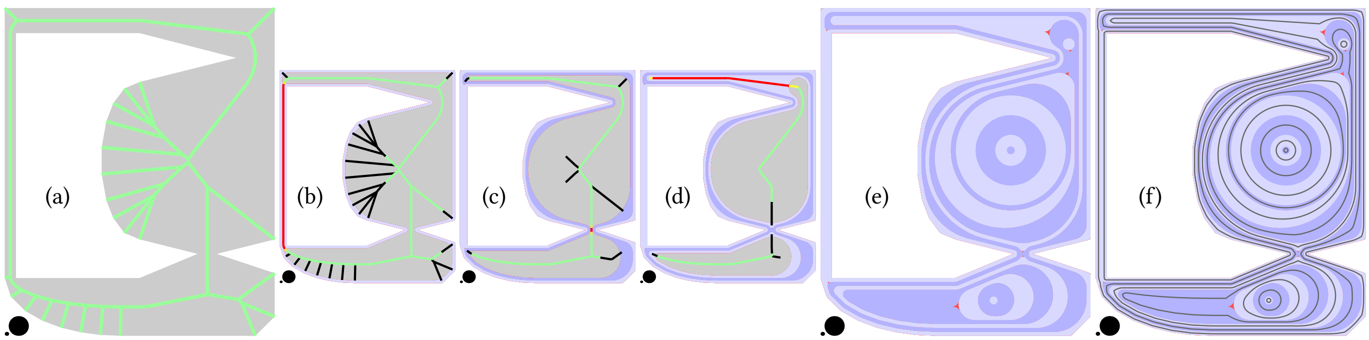


Fig. 1. In this example, we fill the input shape (a) with beads of width varying between 0.3 mm and 1.5 mm, as visualized as the diameter of the black disks. We trim the medial axis and label parts variously (colors) in order to model the (b) first, (c) second, (d) third bead, etc. All the beads are visible in (e) and the print trajectories in (f). In (e,f), the eight small triangular red areas are gaps not covered by any bead.

In most layered additive manufacturing processes, a tool solidifies or deposits material while following pre-planned trajectories to form solid *beads*. Many interesting problems arise in this context, among which one concerns the planning of trajectories for filling a planar shape as densely as possible. This is the problem we tackle in the present paper. Recent works have shown that allowing the bead width to vary along the trajectories helps increase the filling density. We present a novel technique that, given a deposition width range, constructs a set of closed beads whose width varies within the prescribed range and fill the input shape. The technique outperforms the state of the art in important metrics: filling density (while still guaranteeing the absence of bead overlap) and trajectories smoothness. We give a detailed geometric description of our algorithm, explore its behavior on example inputs and provide a statistical comparison with the state of the art. We show that it is possible to obtain high quality fabricated layers on commodity FDM printers.

CCS Concepts: • **Computing methodologies → Shape modeling**; • **Applied computing → Computer-aided design**.

Authors' addresses: Samuel Hornus, Université de Lorraine, CNRS, Inria, LORIA, F-54000, Nancy, France; Tim Kuipers, Ultimaker, Utrecht , Department of Sustainable Design Engineering, Delft University of Technology, The Netherlands; Olivier Devillers; Monique Teillaud, Université de Lorraine, CNRS, Inria, LORIA, F-54000, Nancy, France; Jonàs Martínez, Université de Lorraine, CNRS, Inria, LORIA, F-54000, Nancy, France; Marc Glisse, Inria, Palaiseau, France; Sylvain Lazard; Sylvain Lefebvre, Université de Lorraine, CNRS, Inria, LORIA, F-54000, Nancy, France.

0730-0301/2020/7-ART131
https://doi.org/10.1145/3386569.3392448

Additional Key Words and Phrases: 3D printing, dense infill, medial axis

ACM Reference Format:
Samuel Hornus, Tim Kuipers, Olivier Devillers, Monique Teillaud, Jonàs Martínez, Marc Glisse, Sylvain Lazard, and Sylvain Lefebvre. 2020. Variable-width contouring for Additive Manufacturing. *ACM Trans. Graph.* 39, 4, Article 131 (July 2020), 17 pages. https://doi.org/10.1145/3386569.3392448

1 INTRODUCTION

In additive manufacturing (AM), *process planning* refers to the sequence of operations required to transform an input 3D model into a set of instructions for a target machine. These instructions define the strategies the machine will execute to solidify material and form a solid object. Our work is developed in the context of those strategies that rest on the ability to solidify material along planar trajectories or toolpaths. The solidified material along a trajectory is called a *bead*. On many technologies – such as selective laser sintering (SLS), stereolithography (SLA) and fused filament fabrication (FFF) – the instructions are vector trajectories enriched with information such as energy levels or material flow. When following the trajectories, the machine produces solidified beads that join together to form a planar solid layer. The layers are then sequentially solidified on top of each other to obtain a final solid object.

Process planning raises many interesting geometric problems, with direct practical implications. Amongst these, the question of how to solidify the area of each object layer has received much attention (we review prior works in Section 2).

In this context, we focus on the specific problem of densely covering a layer. A critical aspect of such *dense infills* is to avoid situations

Vid2Curve: Simultaneous Camera Motion Estimation and Thin Structure Reconstruction from an RGB Video

PENG WANG, The University of Hong Kong
LINGJIE LIU, Max Planck Institute for Informatics
NENGLUN CHEN, The University of Hong Kong
HUNG-KUO CHU, National Tsing Hua University
CHRISTIAN THEOBALT, Max Planck Institute for Informatics
WENPING WANG, The University of Hong Kong

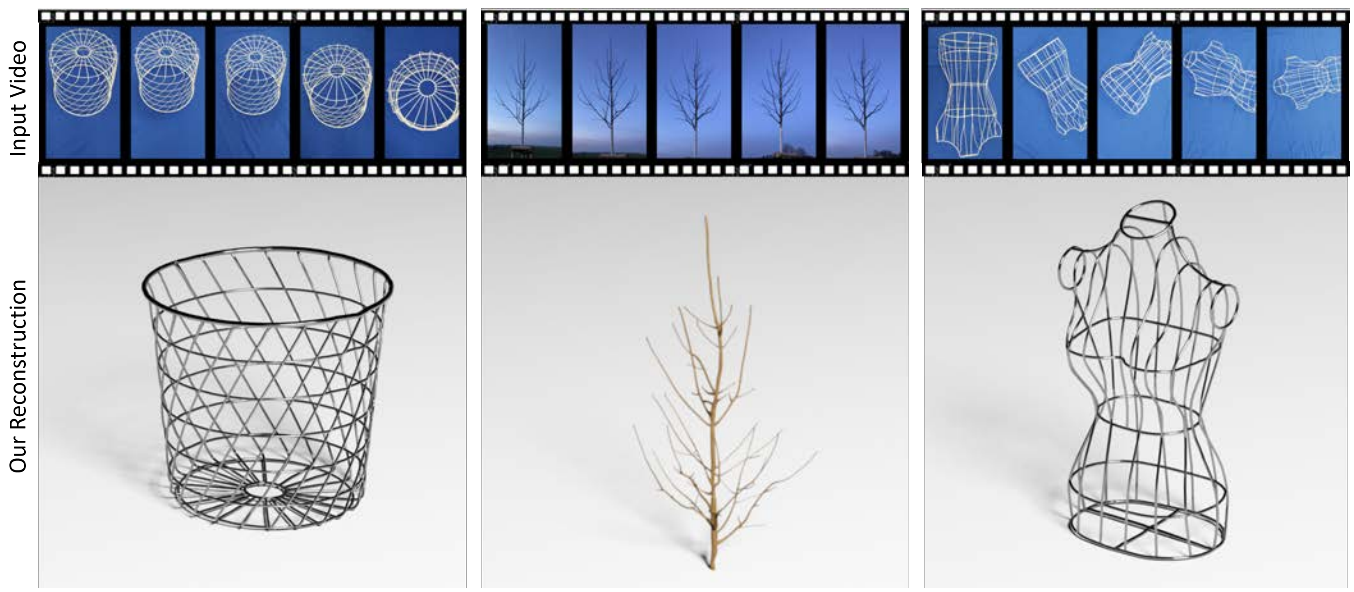


Fig. 1. This figure shows three thin structured objects reconstructed using our method, together with selected frames of the input videos. Using an RGB video as input, our method performs curve-based camera pose estimation and reconstructs complex 3D thin structures in better quality than other existing methods. See the comparisons with the other methods on the bucket model and the hanger model in Section 5.5.

Thin structures, such as wire-frame sculptures, fences, cables, power lines, and tree branches, are common in the real world. It is extremely challenging to acquire their 3D digital models using traditional image-based or depth-based reconstruction methods, because thin structures often lack distinct point features and have severe self-occlusion. We propose the first approach that simultaneously estimates camera motion and reconstructs the geometry of complex 3D thin structures in high quality from a color video captured by a handheld camera. Specifically, we present a new curve-based approach to estimate accurate camera poses by establishing correspondences between featureless thin objects in the foreground in consecutive video frames, without requiring visual texture in the background scene to lock on. Enabled by this effective curve-based camera pose estimation strategy, we develop an iterative optimization method with tailored measures on geometry, topology as well as self-occlusion handling for reconstructing 3D thin structures. Extensive validations on a variety of thin structures show that our method achieves accurate camera pose estimation and faithful reconstruction of 3D thin structures with complex shape and topology at a level that has not been attained by other existing reconstruction methods.

CCS Concepts: • **Computing methodologies** → **Parametric curve and surface models**.

Additional Key Words and Phrases: curve reconstruction, delicate structure, image-based reconstruction

ACM Reference Format:
Peng Wang, Lingjie Liu, Nenglun Chen, Hung-Kuo Chu, Christian Theobalt, and Wenping Wang. 2020. Vid2Curve: Simultaneous Camera Motion Estimation and Thin Structure Reconstruction from an RGB Video . *ACM*

Authors' addresses: Peng Wang, The University of Hong Kong; Lingjie Liu, Max Planck Institute for Informatics; Nenglun Chen, The University of Hong Kong; Hung-Kuo Chu, National Tsing Hua University; Christian Theobalt, Max Planck Institute for Informatics; Wenping Wang, The University of Hong Kong.

0730-0301/2020/7-ART132 $15.00
https://doi.org/10.1145/3386569.3392476

A perceptual model of motion quality for rendering with adaptive refresh-rate and resolution

GYORGY DENES, University of Cambridge
AKSHAY JINDAL, University of Cambridge
ALIAKSEI MIKHAILIUK, University of Cambridge
RAFAŁ K. MANTIUK, University of Cambridge

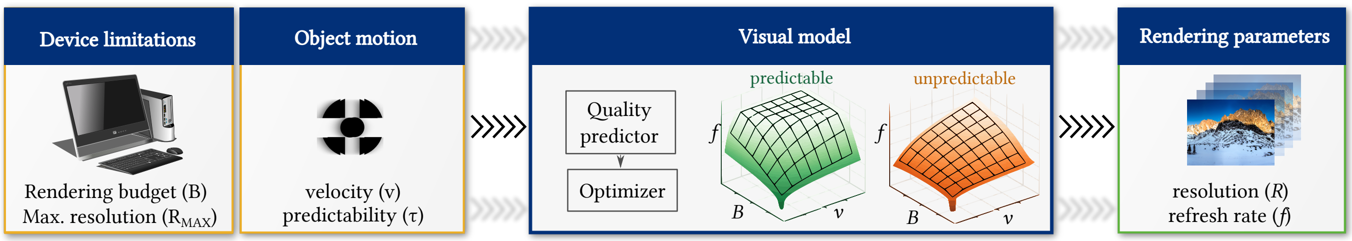


Fig. 1. Our proposed perceptual model of motion quality takes object motion, refresh rate and device limitations (such as the rendering budget and the maximum screen resolution) to predict the perceived quality. This model can be then used to find the combination of resolution and refresh rate that produces the highest animation quality under the given conditions. Surface plots visualize model predictions.

Limited GPU performance budgets and transmission bandwidths mean that real-time rendering often has to compromise on the spatial resolution or temporal resolution (refresh rate). A common practice is to keep either the resolution or the refresh rate constant and dynamically control the other variable. But this strategy is non-optimal when the velocity of displayed content varies. To find the best trade-off between the spatial resolution and refresh rate, we propose a perceptual visual model that predicts the quality of motion given an object velocity and predictability of motion. The model considers two motion artifacts to establish an overall quality score: non-smooth (juddery) motion, and blur. Blur is modeled as a combined effect of eye motion, finite refresh rate and display resolution. To fit the free parameters of the proposed visual model, we measured eye movement for predictable and unpredictable motion, and conducted psychophysical experiments to measure the quality of motion from 50 Hz to 165 Hz. We demonstrate the utility of the model with our on-the-fly motion-adaptive rendering algorithm that adjusts the refresh rate of a G-Sync-capable monitor based on a given rendering budget and observed object motion. Our psychophysical validation experiments demonstrate that the proposed algorithm performs better than constant-refresh-rate solutions, showing that motion-adaptive rendering is an attractive technique for driving variable-refresh-rate displays.

CCS Concepts: • **Computing methodologies → Perception**; *Rendering*.

Authors' addresses: Gyorgy Denes, University of Cambridge, gyorgy.denes@cl.cam.ac.uk; Akshay Jindal, University of Cambridge; Aliaksei Mikhailiuk, University of Cambridge; Rafał K. Mantiuk, Department of Computer Science and Technology, University of Cambridge, rafal.mantiuk@cl.cam.ac.uk.

0730-0301/2020/7-ART133 $15.00
https://doi.org/10.1145/3386569.3392411

Additional Key Words and Phrases: motion quality, adaptive refresh rate

ACM Reference Format:
Gyorgy Denes, Akshay Jindal, Aliaksei Mikhailiuk, and Rafał K. Mantiuk. 2020. A perceptual model of motion quality for rendering with adaptive refresh-rate and resolution. *ACM Trans. Graph.* 39, 4, Article 133 (July 2020), 17 pages. https://doi.org/10.1145/3386569.3392411

1 INTRODUCTION

Modern displays can offer both high spatial resolution (up to 8K) and high refresh rates (above 100 Hz). Such high spatio-temporal resolution is needed to reach the perceptual limits of the visual system and to deliver high-fidelity content. This is particularly important for VR/AR headsets, which still offer resolutions far below the perceptual limits. However, a major obstacle is the limited computational power and bandwidth of modern GPUs: only the most powerful GPUs can render 4K content at 100 Hz or more.

Rendering for modern display technologies often requires a trade-off between spatial and temporal resolution. For example, as VR/AR headsets require constant and sustained refresh rates, the quality control mechanism in rendering engines needs to dynamically adjusts the rendering resolution to fit within the rendering budget. Another approach, employed on G-/Free-Sync capable displays, is to render at a fixed resolution but vary the refresh rate according to the available rendering budget. However, depending on camera and content motion, keeping either the spatial *or* temporal resolution constant may not produce the best visual quality. A better approach is to manipulate the refresh rate and resolution simultaneously, *i.e.* to dynamically adjust the trade-off based on the content of the animation. For example, when a scene is static, the application should maximize spatial resolution, but when movement is fast, the application should optimize for higher refresh rates which result in better perceived quality. Such a mechanism can be introduced

Analytic Spherical Harmonic Gradients for Real-Time Rendering with Many Polygonal Area Lights

LIFAN WU, University of California, San Diego
GUANGYAN CAI, University of California, San Diego
SHUANG ZHAO, University of California, Irvine
RAVI RAMAMOORTHI, University of California, San Diego

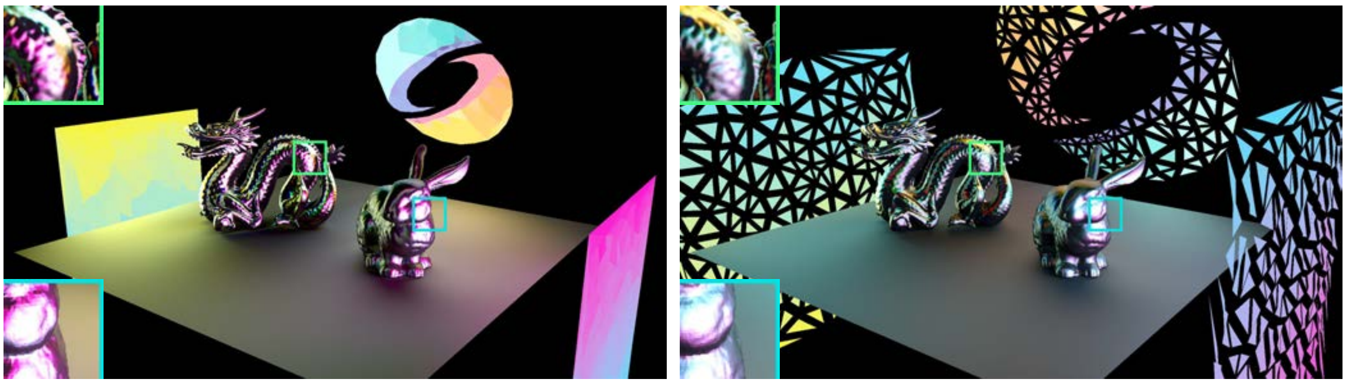


Fig. 1. We present an analytic formula for spherical harmonic (SH) gradients from uniform polygonal area lights, and show how this new theoretical result enables scaling Precomputed Radiance Transfer (PRT) to hundreds of area lights. We first compute the lighting SH coefficients and gradients on a sparse 3D grid. To evaluate SH coefficients for any intermediate point (vertex), we exploit SH gradients and use accurate Hermite interpolation. Here we render a glossy scene with 713 polygonal (triangular) lights and 1.3M polygons at 36fps. Each light transforms independently (in terms of color, location, orientation etc.), enabling the appearance of textured lights or more complex patterns, and causing significant changes in glossy highlights (compare left and right images).

Recent work has developed analytic formulae for spherical harmonic (SH) coefficients from uniform polygonal lights, enabling near-field area lights to be included in Precomputed Radiance Transfer (PRT) systems, and in offline rendering. However, the method is inefficient since coefficients need to be recomputed at each vertex or shading point, for each light, even though the SH coefficients vary smoothly in space. The complexity scales linearly with the number of lights, making many-light rendering difficult. In this paper, we develop a *novel analytic formula for the spatial gradients of the spherical harmonic coefficients for uniform polygonal area lights.* The result is a significant generalization, involving the Reynolds transport theorem to reduce the problem to a boundary integral for which we derive a new analytic formula, showing how to reduce a key term to an earlier recurrence for SH coefficients. The implementation requires only minor additions to existing code for SH coefficients. The results also hold implications for recent efforts on differentiable rendering. We show that SH gradients enable very sparse spatial sampling, followed by accurate Hermite interpolation. This enables *scaling PRT to hundreds of area lights* with minimal overhead and real-time frame rates. Moreover, the SH gradient formula is a new mathematical result that potentially enables many other graphics applications.

CCS Concepts: • **Computing methodologies → Rendering**.

Additional Key Words and Phrases: analytic gradients, spherical harmonics, area lighting, differentiable rendering

ACM Reference Format:
Lifan Wu, Guangyan Cai, Shuang Zhao, and Ravi Ramamoorthi. 2020. Analytic Spherical Harmonic Gradients for Real-Time Rendering with Many Polygonal Area Lights. *ACM Trans. Graph.* 39, 4, Article 134 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392373

1 INTRODUCTION

In this paper, we address a fundamental mathematical question, deriving an analytic formula for the spatial gradients of spherical harmonic (SH) coefficients from a uniform polygonal area light. While both area lights and spherical harmonics are widely used in rendering, to our knowledge, there has been no previous work on finding analytic SH gradients for them. We believe the result has implications for many problems in rendering and beyond.

Our immediate practical motivation is for real-time rendering with precomputed radiance transfer (PRT) [Sloan et al. 2002]. PRT and SH lighting enable dynamic low-frequency environments with realistic highlights and real-time shading, including soft shadows. Hence, they are widely used in real-time applications like games and even in offline rendering [Pantaleoni et al. 2010]. However, the

Authors' addresses: Lifan Wu, University of California, San Diego, liw086@eng.ucsd.edu; Guangyan Cai, University of California, San Diego, g5cai@ucsd.edu; Shuang Zhao, University of California, Irvine, shz@ics.uci.edu; Ravi Ramamoorthi, University of California, San Diego, ravir@cs.ucsd.edu.

0730-0301/2020/7-ART134 $15.00
https://doi.org/10.1145/3386569.3392373

Compositional Neural Scene Representations for Shading Inference

JONATHAN GRANSKOG, ETH Zurich & NVIDIA
FABRICE ROUSSELLE, NVIDIA
MARIOS PAPAS, DisneyResearch|Studios
JAN NOVÁK, NVIDIA

We present a technique for adaptively partitioning neural scene representations. Our method disentangles lighting, material, and geometric information yielding a scene representation that preserves the orthogonality of these components, improves interpretability of the model, and allows compositing new scenes by mixing components of existing ones. The proposed adaptive partitioning respects the uneven entropy of individual components and permits compressing the scene representation to lower its memory footprint and potentially reduce the evaluation cost of the model. Furthermore, the partitioned representation enables an in-depth analysis of existing image generators. We compare the flow of information through individual partitions, and by contrasting it to the impact of additional inputs (G-buffer), we are able to identify the roots of undesired visual artifacts, and propose one possible solution to remedy the poor performance. We also demonstrate the benefits of complementing traditional forward renderers by neural representations and synthesis, e.g. to infer expensive shading effects, and show how these could improve production rendering in the future if developed further.

CCS Concepts: • **Computing methodologies → Rendering**; **Neural networks**.

Additional Key Words and Phrases: rendering, neural networks, neural scene representations, disentanglement, attribution

ACM Reference Format:
Jonathan Granskog, Fabrice Rousselle, Marios Papas, and Jan Novák. 2020. Compositional Neural Scene Representations for Shading Inference. *ACM Trans. Graph.* 39, 4, Article 135 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392475

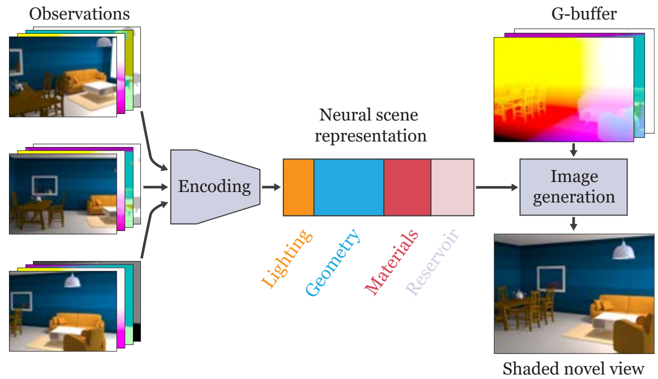


Fig. 1. We present an approach for extracting adaptively partitioned, compressible neural scene representations of 3D scenes that disentangle lighting, material, and geometry information. The compositionality of the representation aids interpretability and analysis of the inner workings of the model, which are key for remedying visual artifacts and combining neural approaches with traditional forward renderers.

1 INTRODUCTION

Data-driven realistic image synthesis has recently achieved a number of notable breakthroughs, such as rendering realistic human faces [Karras et al. 2018], or high-quality relighting of photographs [Philip et al. 2019]. Remarkable achievements have been demonstrated also in data-driven simulation of light transport, where neural networks predict various radiative quantities [Hermosilla et al. 2019; Kallweit et al. 2017; Nalbach et al. 2017; Ren et al. 2013] or improve their unbiased estimation [Müller et al. 2019; Zheng and Zwicker 2019]. Common to all these is the utilization of neural networks to perform the task in its entirety, all at once. The black-box nature, however, hinders interpretability, generalization, and makes further development less intuitive.

An alternative approach to performing the synthesis at once is to introduce an intermediate neural scene representation [Eslami et al. 2018; Kulkarni et al. 2015; Sitzmann et al. 2019], by breaking the rendering task into: (i) extracting a learned scene representation, and (ii) using it to render an image. The intermediate (latent) scene representation allows enforcing certain behaviors upon the model, e.g. ensuring consistency of images rendered from different views of the scene. It also presents an opportunity for increasing robustness, improving generalization, and accelerating training by injecting physically-based constraints to regularize the model. While the scene complexity and rendering quality of these approaches may appear limited at present, it is foreseeable that these will improve in the future, especially if the neural renderer merely *augments* classical rendering pipelines. The key advantage of the neural approach is the end-to-end training, which allows the neural representation to carry information that is complementary to classical inputs and tailored to the task at hand. We carry out our investigations in one such scenario, where a classical rasterizer determines directly visible objects and the neural renderer infers their appearance.

We present two distinct contributions in this article. First, we extend the works of Eslami et al. [2018] and Sitzmann et al. [2019] with mechanisms to disentangle material, lighting, and geometric content of the scene. We do not prescribe a specific encoding between images and the latent scene representation (this shall be extracted from data), but we introduce additional constraints to *adaptively partition* the latent scene representation. The adaptive partitioning

Authors' addresses: Jonathan Granskog, ETH Zurich & NVIDIA, jgranskog@nvidia.com; Fabrice Rousselle, NVIDIA, frousselle@nvidia.com; Marios Papas, DisneyResearch|Studios, marios.papas@disneyresearch.com; Jan Novák, NVIDIA, jnovak@nvidia.com.

0730-0301/2020/7-ART135 $15.00
https://doi.org/10.1145/3386569.3392475

Continuous Multiple Importance Sampling

REX WEST, The University of Tokyo, Japan
ILIYAN GEORGIEV, Autodesk, United Kingdom
ADRIEN GRUSON, McGill University, Canada
TOSHIYA HACHISUKA, The University of Tokyo, Japan

Fig. 1. Three applications of our framework to light transport simulation. We reformulate each application as a problem of combining a continuum of sampling techniques and leverage our continuous MIS (CMIS) formulation to derive an efficient weighting scheme. Based on this scheme, our practical stochastic MIS (SMIS) estimator outperforms existing state-of-the-art methods. For each image we report error in SMAPE units (see Section 4.4).

Multiple importance sampling (MIS) is a provably good way to combine a finite set of sampling techniques to reduce variance in Monte Carlo integral estimation. However, there exist integration problems for which a continuum of sampling techniques is available. To handle such cases we establish a continuous MIS (CMIS) formulation as a generalization of MIS to uncountably infinite sets of techniques. Our formulation is equipped with a base estimator that is coupled with a provably optimal balance heuristic and a practical stochastic MIS (SMIS) estimator that makes CMIS accessible to a broad range of problems. To illustrate the effectiveness and utility of our framework, we apply it to three different light transport applications, showing improved performance over the prior state-of-the-art techniques.

CCS Concepts: • **Computing methodologies → Rendering**; **Ray tracing**.

Additional Key Words and Phrases: multiple importance sampling, light transport, spectral rendering, path reuse, volume rendering

Authors' addresses: Rex West, The University of Tokyo, Japan; Iliyan Georgiev, Autodesk, United Kingdom; Adrien Gruson, McGill University, Canada; Toshiya Hachisuka, The University of Tokyo, Japan.

0730-0301/2020/7-ART136 $15.00
https://doi.org/10.1145/3386569.3392436

ACM Reference Format:
Rex West, Iliyan Georgiev, Adrien Gruson, and Toshiya Hachisuka. 2020. Continuous Multiple Importance Sampling. *ACM Trans. Graph.* 39, 4, Article 136 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392436

1 INTRODUCTION

Multiple importance sampling [Veach and Guibas 1995] (MIS) provides a framework for combining a set of sampling techniques in Monte Carlo integration. This combination is done by weighting the contribution of each sample produced by each sampling technique according to some heuristic.

MIS can be directly applied to problems where the set of sampling techniques is countable. However, there are certain problems where *an uncountably infinite number* (i.e., a *continuum*) of techniques arises naturally. A generalization of MIS is needed for these problems as the classical formulation does not consider such cases.

We formally establish this continuous generalization of MIS, which we call *continuous MIS* (CMIS). Based on our formulation, we devise a CMIS estimator that combines a continuum of sampling techniques using a provably optimal balance heuristic. Since this estimator is not always practical, we propose an approximation to it—our *stochastic MIS* (SMIS) estimator—which is unbiased and extends classical MIS to stochastic technique selection.

To demonstrate the utility of our framework, we apply it to three applications in light transport simulation shown in Fig. 1. In path space filtering [Keller et al. 2014], CMIS allows us to reformulate

Converting stroked primitives to filled primitives

DIEGO NEHAB, IMPA, Brazil

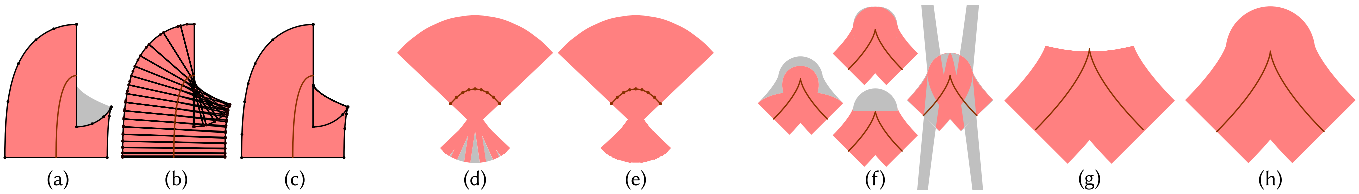


Fig. 1. We analyzed 22 distinct third-party strokers. (a) 14 of them confuse the stroke-to-fill conversion problem with the curve-offsetting problem, and produce incorrect results near high-curvature regions. (b) The remaining 8 do treat high-curvature regions, but offer no accuracy guarantees, and often output too many segments. (c) Our stroker correctly handles these regions using evolutes. (d) All but 2 strokers ignore inner joins between segments, leaving unexpected gaps. (e) Our stroker detects when such joins are visible (even between curved segments) and outputs them. (f) All but 2 strokers fail spectacularly in the vicinity of intra-segment cusps and "almost cusps". (g) Following the standards to the letter produces discontinuous results at cusps. (h) Our stroker robustly detects cusps and "almost cusps" to produce continuous, intuitive results. — input path — output path ■ filled region ■ missing or extraneous region

Vector graphics formats offer support for both filled and stroked primitives. Filled primitives paint all points in the region bounded by a set of outlines. Stroked primitives paint all points covered by a line drawn over the outlines. Editors allow users to convert stroked primitives to the outlines of equivalent filled primitives for further editing. Likewise, renderers typically convert stroked primitives to equivalent filled primitives prior to rendering. This conversion problem is deceivingly difficult to solve. Surprisingly, it has received little to no attention in the literature. Existing implementations output too many segments, do not satisfy accuracy requirements, or fail under a variety of conditions, often spectacularly. In this paper, we present a solution to the stroke-to-fill conversion problem that addresses these issues. One of our key insights is to take into account the evolutes of input outlines, in addition to their offsets, in regions of high curvature. Furthermore, our approach strives to maintain continuity between the input and the set of painted points. Our implementation is available in open source.

CCS Concepts: • **Computing methodologies** → *Graphics file formats*; *Parametric curve and surface models*; *Rasterization.*

Additional Key Words and Phrases: stroke, vector graphics, offset curves

ACM Reference Format:
Diego Nehab. 2020. Converting stroked primitives to filled primitives. *ACM Trans. Graph.* 39, 4, Article 137 (July 2020), 17 pages. https://doi.org/10.1145/3386569.3392392

1 INTRODUCTION

Vector graphics is the standard way of describing scalable visual information, such as pages of text, illustrations, maps, etc. Well-known vector graphics file formats include PS [1999], PDF [2006], SVG [2011], OpenXPS [2009], and CGM [1999]. All these formats closely follow the framework put forth in the seminal work by Warnock and Wyatt [1982]. They define an illustration as a series of *shapes* to be *painted* in order, each one on top of the previous. Each shape-paint combination forms a *layer*. The process mimics silk-screen printing, with shapes working as stencils that restrict the area of the illustration affected by paint. Paints can be constant, vary spatially in gradients, or even consist of arbitrary textures. Shapes can be *filled* or *stroked*. Both types of shapes are defined by *paths* (i.e., piecewise-polynomial outlines). Points are *inside* a filled shape based on their *winding number* relative to the path. (Both *odd* and *non-zero* tests are widely supported.) Stroked shapes, in contrast, consist of the set of points covered when a line of given *width* is drawn over the path.

We must often find a set of outlines defining a shape that, when filled, paints the same set of points as those painted by a given stroked shape. The focus of this paper is this *stroke-to-fill conversion problem* (or simply *stroking*). The fact that filled and stroked shapes define interior points in entirely different ways is inconvenient for vector graphics renderers. Rather than using two distinct algorithms at rendering time, virtually all rendering engines convert stroked shapes to equivalent filled shapes prior to rendering. (The only exception we found is NVpr [Kilgard and Bolz 2012].) A solution to the stroke-to-fill problem is also useful outside the context of rendering: Vector graphics editors allow users to convert stroked primitives to their outlines to enable further editing (e.g., Inkscape's "Stroke to Path" or Adobe Illustrator's "Outline Stroke").

Surprisingly, in this problem's nearly 40 years of history, it has rarely, if ever, been mentioned in the literature. As a consequence, developers of vector graphics rendering engines and editors keep "reinventing the wheel". The approaches followed by Anti-Grain Geometry, the Cairo Graphics Library, Microsoft's Direct2D, Apple's Quartz, Ghostscript, MuPDF, MPVG [Ganacim et al. 2014], Java 2D, the livarot library (used by Inkscape), Qt 5, the OpenVG Reference Implementation, the Skia Graphics Library, and Adobe Illustrator are all different *and* inconsistent with each other.

These implementations can be broadly categorized according to two axes: *flat* vs. *curve-based*, and *local* vs. *global*. *Flat* algorithms operate on piecewise-linear approximations of the input path (i.e., they *flatten* the input). Although this makes the problem easier to solve, an undesirably large number of output segments are needed

Author's address: Diego Nehab, IMPA, Rio de Janeiro, Brazil.

0730-0301/2020/7-ART137 $15.00
https://doi.org/10.1145/3386569.3392392

GS-PAT: High-Speed Multi-Point Sound-Fields for Phased Arrays of Transducers

DIEGO MARTINEZ PLASENCIA, University of Sussex, UK
RYUJI HIRAYAMA, University of Sussex, UK
ROBERTO MONTANO-MURILLO, University of Sussex, UK
SRIRAM SUBRAMANIAN, University of Sussex, UK

Phased Arrays of Transducers (PATs) allow accurate control of ultrasound fields, with applications in haptics, levitation (i.e. displays) and parametric audio. However, algorithms for multi-point levitation or tactile feedback are usually limited to computing solutions in the order of hundreds of sound-fields per second, preventing the use of multiple high-speed points, a feature that can broaden the scope of applications of PATs. We present *GS-PAT*, a GPU multi-point phase retrieval algorithm, capable of computing 17K solutions per second for up to 32 simultaneous points in a mid-end consumer grade GPU (NVidia GTX 1660). We describe the algorithm and compare it to state of the art multi-point algorithms used for ultrasound haptics and levitation, showing similar quality of the generated sound-fields, and much higher computation rates. We then illustrate how the shift in paradigm enabled by *GS-PAT* (i.e. real-time control of several high-speed points) opens new applications for PAT technologies, such as in volumetric fully coloured displays, multi-point spatio-temporal tactile feedback, parametric audio and simultaneous combinations of these modalities.

CCS Concepts: • **Computing methodologies → Physical simulation**.

Additional Key Words and Phrases: Multi-point phase optimization; Parametric sound; Ultrasound Levitation; Mid-air haptics; Phased Arrays of Transducers

ACM Reference Format:
Diego Martinez Plasencia, Ryuji Hirayama, Roberto Montano-Murillo, and Sriram Subramanian. 2020. GS-PAT: High-Speed Multi-Point Sound-Fields for Phased Arrays of Transducers. *ACM Trans. Graph.* 39, 4, Article 138 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392492

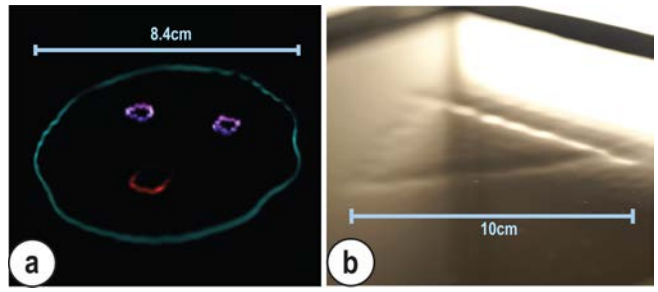


Fig. 1. Example applications enabled by GS-PAT. (A) Volumetric POV visual content using 6 high-speed scanning particles. (B) Multi-point tactile feedback using three high-speed moving tactile points.

1 INTRODUCTION

Phased Arrays of Transducers (PATs) provide accurate control of the phase and amplitude of dense arrays of transducers (e.g. 16x16 transducers), and have demonstrated capabilities to present visual, auditive and tactile content.

Visual content using PATs typically relied on sparse sets of levitated particles, initially constrained to specific locations or arrangements [Ochiai et al. 2014; Omirou et al. 2015]. Algorithmic advances allowed unconstrained 3D positioning of single particles first [Marzo et al. 2015] and then extended this to several particles [Marzo and Drinkwater 2019]. Recently, the use of a single levitated high-speed particle [Hirayama et al. 2019] has allowed the creation of 3D displays using the Persistence of Vision (POV) effect. For audio, PATs have allowed the steering of single [Ochiai et al. 2017; Olszewski et al. 2005] and then multiple [Shi et al. 2015] audible beams in different directions. Haptics also demonstrated mid-air feedback, first at a single tactile point in 3D [Hoshi et al. 2010], then at several tactile points [Carter et al. 2013; Inoue et al. 2015; Long et al. 2014] and finally using a single high-speed tactile point [Frier et al. 2019].

From an algorithmic point of view, all these PAT approaches have followed a similar evolution, with single-point approaches demonstrating an initial potential [Carter et al. 2013; Marzo et al. 2015; Olszewski et al. 2005]; multi-point extending it [Long et al. 2014; Marzo and Drinkwater 2019; Shi et al. 2015]; and single-point high-speed algorithms demonstrating further applications, such as better tactile feedback [Frier et al. 2019], or even simultaneous delivery of visual, tactile and auditive content [Hirayama et al. 2019].

The next natural step in this evolution lies in exploiting multi-point sound-fields computed at high rates (e.g. such as >10KHz, as used for single-particle POV displays [Hirayama et al. 2019]; or 17KHz, used for single-point haptics [Frier et al. 2019]). However, this step remains hindered by the performance of existing multi-point algorithms, typically limited to computing rates of hundreds of hertz. This paper fills this gap, and demonstrates how this shift in paradigm from a single high-speed point to multiple high-speed points enables new applications of PATs for all 3 modalities.

We present *GS-PAT*, a high performance approximation to the Gerchberg-Saxton (GS) algorithm [Gerchberg 1972] specifically tailored to phase and amplitude PATs. *GS-PAT* reduces the algorithmic complexity of GS, as to deal with more points or PAT setups with

Authors' addresses: Diego Martinez Plasencia, University of Sussex, UK, dm372@sussex.ac.uk; Ryuji Hirayama, University of Sussex, UK, r.hirayama@sussex.ac.uk; Roberto Montano-Murillo, University of Sussex, UK, r.montano-murillo@sussex.ac.uk; Sriram Subramanian, University of Sussex, UK, sriram@sussex.ac.uk.

0730-0301/2020/7-ART138 $15.00
https://doi.org/10.1145/3386569.3392492

Image-Based Acquisition and Modeling of Polarimetric Reflectance

SEUNG-HWAN BAEK, KAIST
TIZIAN ZELTNER, EPFL
HYUN JIN KU and INSEUNG HWANG, KAIST
XIN TONG, Microsoft Research Asia
WENZEL JAKOB, EPFL
MIN H. KIM, KAIST

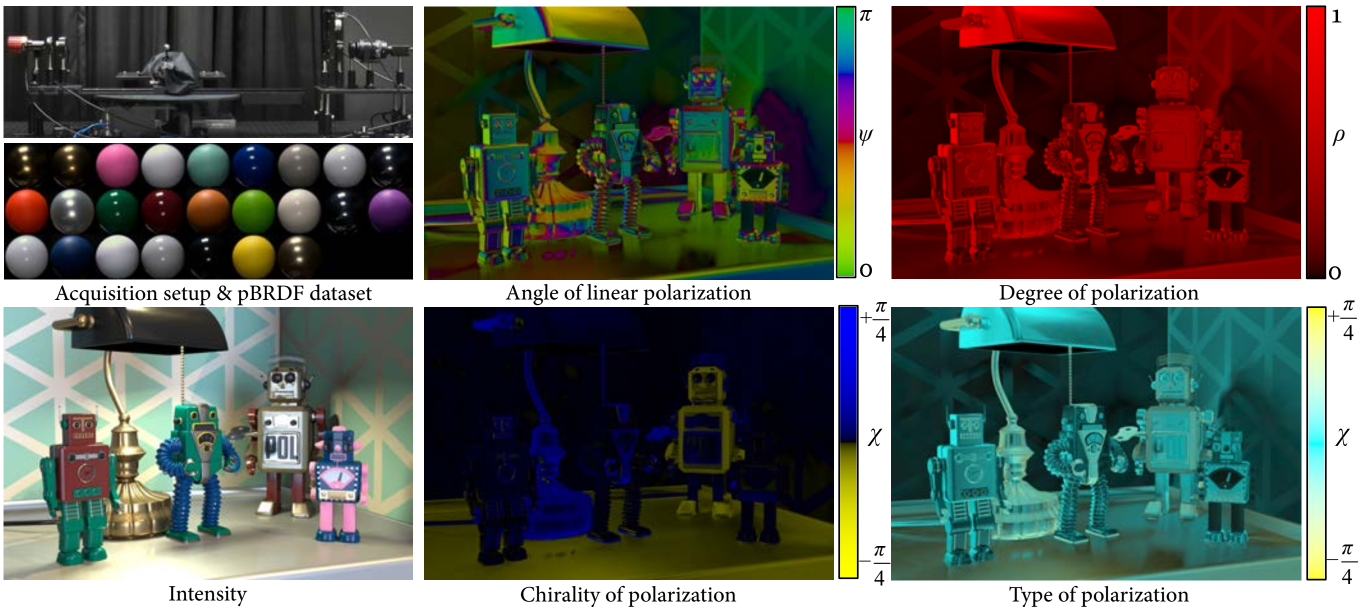


Fig. 1. We have acquired the first polarimetric BRDF dataset of real-world materials that provides coverage of arbitrary viewing and lighting configurations using our novel acquisition setup (shown on the left top). Acquired pBRDFs can be used in physically based simulations that correctly account for the change in polarization state during multiple scattering. We visualize the polarimetric information using the degree, azimuth, chirality, and type of polarization at the wavelength of 550 nm, following Wilkie and Weidlich [2010].

Realistic modeling of the bidirectional reflectance distribution function (BRDF) of scene objects is a vital prerequisite for any type of physically based rendering. In the last decades, the availability of databases containing real-world material measurements has fueled considerable innovation in the development of such models. However, previous work in this area was mainly focused on increasing the visual realism of images, and hence ignored the effect of scattering on the polarization state of light, which is normally imperceptible to the human eye. Existing databases thus only capture scattered flux, or polarimetric BRDF datasets are too directionally sparse (e.g., in-plane) to be usable for simulation.

While subtle to human observers, polarization is easily perceived by any optical sensor (e.g., using polarizing filters), providing a wealth of additional information about shape and material properties of the object under observation. Given the increasing application of rendering in the solution of inverse problems via analysis-by-synthesis and differentiation, the ability to realistically model polarized radiative transport is thus highly desirable.

Polarization depends on the wavelength of the spectrum, and thus we provide the first polarimetric BRDF (pBRDF) dataset that captures the polarimetric properties of real-world materials over the full angular domain, and at multiple wavelengths. Acquisition of such reflectance data is challenging due to the extremely large space of angular, spectral, and polarimetric configurations that must be observed, and we propose a scheme combining image-based acquisition with spectroscopic ellipsometry to perform measurements in a realistic amount of time. This process yields raw Mueller matrices, which we subsequently transform into Rusinkiewicz-parameterized pBRDFs that can be used for rendering.

Our dataset provides 25 isotropic pBRDFs spanning a wide range of appearances: diffuse/specular, metallic/dielectric, rough/smooth, and different

Authors' addresses: Seung-Hwan Baek, KAIST, School of Computing, Daejeon, South Korea, 34141; Tizian Zeltner, EPFL, , Hyun Jin Ku; Inseung Hwang, KAIST, School of Computing, Daejeon, South Korea, 34141; Xin Tong, Microsoft Research Asia, Wenzel Jakob, EPFL, , Min H. Kim, KAIST, School of Computing, Daejeon, South Korea, 34141, minhkim@kaist.ac.kr.

0730-0301/2020/7-ART139
https://doi.org/10.1145/3386569.3392387

Langevin Monte Carlo Rendering with Gradient-based Adaptation

FUJUN LUAN, Cornell University
SHUANG ZHAO, University of California, Irvine
KAVITA BALA, Cornell University
IOANNIS GKIOULEKAS, Carnegie Mellon University

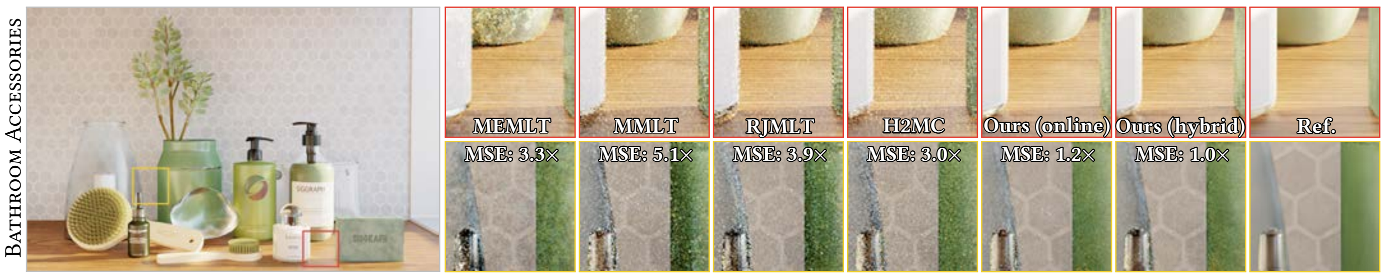


Fig. 1. Equal-time (20 minutes) comparison between MEMLT, MMLT, RJMLT, H2MC and two variants of our methods. The scene presents complex glossy and specular interreflections with difficult visibilities. The reference (Ref.) is rendered by BDPT in roughly a day. MEMLT suffers from correlated noise in the glass bottle region due to insufficient local exploration. MMLT and RJMLT both have severe noise because they sometimes get trapped in regions with hard-to-find features using Kelemen-style isotropic mutations. H2MC is more efficient with the help of anisotropic Gaussian mutations, but the computational overhead from Hessian computations and dense matrix operations results in a low sample budget and insufficient exploration. Our Langevin Monte Carlo methods efficiently address these challenges by exploiting first-order gradient information to robustly balance the tradeoff between adaptation and cost.

We introduce a suite of Langevin Monte Carlo algorithms for efficient photorealistic rendering of scenes with complex light transport effects, such as caustics, interreflections, and occlusions. Our algorithms operate in primary sample space, and use the Metropolis-adjusted Langevin algorithm (MALA) to generate new samples. Drawing inspiration from state-of-the-art stochastic gradient descent procedures, we combine MALA with adaptive preconditioning and momentum schemes that re-use previously-computed first-order gradients, either in an online or in a cache-driven fashion. This combination allows MALA to adapt to the local geometry of the primary sample space, without the computational overhead associated with previous Hessian-based adaptation algorithms. We use the theory of controlled Markov chain Monte Carlo to ensure that these combinations remain ergodic, and are therefore suitable for unbiased Monte Carlo rendering. Through extensive experiments, we show that our algorithms, MALA with online and cache-driven adaptation, can successfully handle complex light transport in a large variety of scenes, leading to improved performance (on average more than 3× variance reduction at equal time, and 7× for motion blur) compared to state-of-the-art Markov chain Monte Carlo rendering algorithms.

CCS Concepts: • **Computing methodologies** → **Rendering**.

Additional Key Words and Phrases: global illumination, photorealistic rendering, Langevin Monte Carlo

Authors' addresses: Fujun Luan, Cornell University, fujun@cs.cornell.edu; Shuang Zhao, University of California, Irvine, shz@ics.uci.edu; Kavita Bala, Cornell University, kb@cs.cornell.edu; Ioannis Gkioulekas, Carnegie Mellon University, igkioule@andrew.cmu.edu.

0730-0301/2020/7-ART140 $15.00
https://doi.org/10.1145/3386569.3392382

ACM Reference Format:
Fujun Luan, Shuang Zhao, Kavita Bala, and Ioannis Gkioulekas. 2020. Langevin Monte Carlo Rendering with Gradient-based Adaptation. *ACM Trans. Graph.* 39, 4, Article 140 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392382

1 INTRODUCTION

The development of general-purpose and efficient global illumination algorithms is one of the foundational problems in computer graphics. Physics-based Monte Carlo rendering algorithms [Dutre et al. 2006; Pharr et al. 2016] can accurately simulate complicated light transport effects such as caustics, strong interreflections, subsurface scattering, and motion blur. They achieve this by aggregating intensity contributions from a large number of randomly generated *light paths*, representing the different ways in which light can propagate through the scene that is being simulated. As the complexity of the underlying transport effects increases, retaining efficiency requires using Markov chain Monte Carlo (MCMC) techniques [Veach and Guibas 1997]: unlike traditional Monte Carlo algorithms, which generate independent paths through local importance sampling, MCMC algorithms use Markov chains to create statistically-correlated paths. This allows them to efficiently perform both a global exploration of the space of possible paths, searching for clusters of paths with high contributions to the image, and a local exploration of such newly discovered clusters.

The success of MCMC algorithms depends critically on the design of proposal distributions for producing new path samples given previously sampled ones. Whereas so-called *path space* MCMC rendering algorithms directly modify the path within the scene, *primary sample space* algorithms [Kelemen et al. 2002] modify the random numbers provided as input to a black-box path tracing algorithm. Operating in the mathematically-tractable space of real random

Massively Parallel Rendering of Complex Closed-Form Implicit Surfaces

MATTHEW J. KEETER, Independent researcher

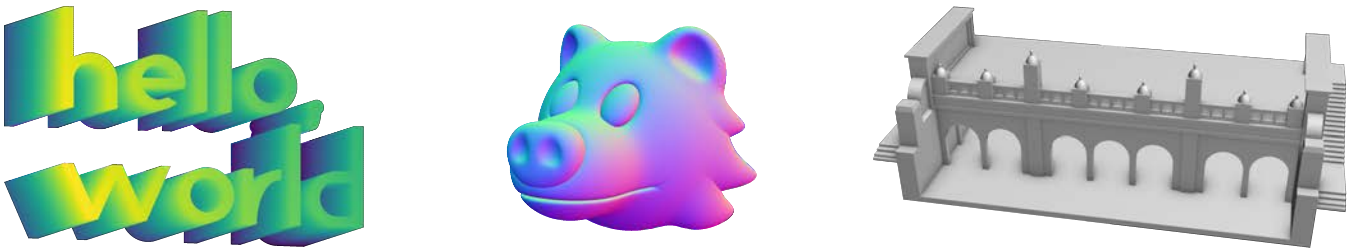


Fig. 1. An assortment of implicit surfaces rendered using our technique. Left: an extruded text string, rotated and rendered as a heightmap. Center: a bear head sculpted using smooth blending operations, with normals found by automatic differentiation. Right: a complex architectural model rendered with screen-space ambient occlusion and perspective. All models are rendered directly from their mathematical representations, without triangulation or raytracing.

We present a new method for directly rendering complex closed-form implicit surfaces on modern GPUs, taking advantage of their massive parallelism. Our model representation is unambiguously solid, can be sampled at arbitrary resolution, and supports both constructive solid geometry (CSG) and more unusual modeling operations (e.g. smooth blending of shapes). The rendering strategy scales to large-scale models with thousands of arithmetic operations in their underlying mathematical expressions. Our method only requires C^0 continuity, allowing for warping and blending operations which break Lipshitz continuity.

To render a model, its underlying expression is evaluated in a shallow hierarchy of spatial regions, using a high branching factor for efficient parallelization. Interval arithmetic is used to both skip empty regions and construct reduced versions of the expression. The latter is the optimization that makes our algorithm practical: in one benchmark, expression complexity decreases by two orders of magnitude between the original and reduced expressions. Similar algorithms exist in the literature, but tend to be deeply recursive with heterogeneous workloads in each branch, which makes them GPU-unfriendly; our evaluation and expression reduction both run efficiently as massively parallel algorithms, entirely on the GPU.

The resulting system renders complex implicit surfaces in high resolution and at interactive speeds. We examine how performance scales with computing power, presenting performance results on hardware ranging from older laptops to modern data-center GPUs, and showing significant improvements at each stage.

CCS Concepts: • **Computing methodologies → Rasterization**; **Volumetric models**.

Additional Key Words and Phrases: implicit surface, signed distance field, freps, octrees, rasterization, gpu, cuda

Author's address: Matthew J. Keeter, Independent researcher, matt.j.keeter@gmail.com.

0730-0301/2020/7-ART141 $15.00
https://doi.org/10.1145/3386569.3392429

ACM Reference Format:
Matthew J. Keeter. 2020. Massively Parallel Rendering of Complex Closed-Form Implicit Surfaces. *ACM Trans. Graph.* 39, 4, Article 141 (July 2020), 10 pages. https://doi.org/10.1145/3386569.3392429

1 INTRODUCTION

Implicit surfaces and functional representations are a powerful way to represent solid models [Bloomenthal and Wyvill 1997; Gomes et al. 2009]. Compared to boundary representations (e.g. triangle meshes or NURBS surfaces), they offer unambiguous inside-outside checking, easy constructive solid geometry (CSG) operations, and arbitrary resolution. In recent years, functional representations (f-reps) have been used as the kernel of both commercial [Courter 2019] and open-source [Keeter 2019] CAD packages. They are a fundamental building block in the demoscene community [Burger et al. 2002; Quilez 2008], used as a representation for generative art [Moen 2019], and even as the underlying technology for a recent PlayStation 4 game [Evans 2015].

Unlike boundary representations, implicit surfaces cannot easily be rendered in their native forms. This paper presents a new method for rendering the family of implicit surfaces represented by arbitrary closed-form arithmetic expressions, i.e., representing a sphere as

$$f(x,y,z) < 0 \text{ where } f(x,y,z) = \sqrt{x^2 + y^2 + z^2} - 1$$

This representation is particularly flexible and can be treated as an "assembly language for shapes" which is targeted by higher-level tools. The space of higher-level tools spans the gamut from advanced solid modeling packages [Allen 2019] to user-friendly content generation tools [Keeter 2015].

Our rendering strategy runs in both 2D and 3D, making efficient use of modern GPU hardware and APIs. Unlike previous work, it scales to complex expressions, maintaining interactive framerates while rendering models built from hundreds or thousands of arithmetic operations. It requires no continuity higher than C^0, which allows for extremely flexible modeling and unusual spatial transformations. Finally, it scales well with GPU power; as GPU performance

Neural Supersampling for Real-time Rendering

LEI XIAO, SALAH NOURI, MATT CHAPMAN, ALEXANDER FIX, DOUGLAS LANMAN, and ANTON KAPLANYAN, Facebook Reality Labs

Fig. 1. Results of our real-time, learned 4×4 supersampling are shown for four sample scenes. From top to bottom: the rendered low-resolution color input, our reconstruction, and the rendered reference images. Our supersampling method takes the color, depth, and motion vectors of multiple low-resolution frames, and produces high-fidelity reconstructions by reducing aliasing and recovering scene details.

Due to higher resolutions and refresh rates, as well as more photorealistic effects, real-time rendering has become increasingly challenging for video games and emerging virtual reality headsets. To meet this demand, modern graphics hardware and game engines often reduce the computational cost by rendering at a lower resolution and then upsampling to the native resolution. Following the recent advances in image and video superresolution in computer vision, we propose a machine learning approach that is specifically tailored for high-quality upsampling of rendered content in real-time applications. The main insight of our work is that in rendered content, the image pixels are point-sampled, but precise temporal dynamics are available. Our method combines this specific information that is typically available in modern renderers (i.e., depth and dense motion vectors) with a novel temporal network design that takes into account such specifics and is aimed at maximizing video quality while delivering real-time performance. By training on a large synthetic dataset rendered from multiple 3D scenes with recorded camera motion, we demonstrate high fidelity and temporally stable results in real-time, even in the highly challenging 4×4 upsampling scenario, significantly outperforming existing superresolution and temporal antialiasing work.

CCS Concepts: • **Computing methodologies → Machine learning**; **Rendering**.

Additional Key Words and Phrases: deep learning, rendering, upsampling, superresolution, virtual reality

ACM Reference Format:
Lei Xiao, Salah Nouri, Matt Chapman, Alexander Fix, Douglas Lanman, and Anton Kaplanyan. 2020. Neural Supersampling for Real-time Rendering. *ACM Trans. Graph.* 39, 4, Article 142 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392376

1 INTRODUCTION

Real-time rendering for modern desktop, mobile, and virtual reality applications is challenging due to increasing display resolutions and demands for photorealistic visual quality. For example, virtual reality (VR) headsets such as the Valve Index require rendering 2880×1600 pixels at 90-144Hz and recent gaming monitors support 3840×2160 resolution at 144Hz, which, together with the recent advances in physically based shading and real-time ray tracing, sets a high demand on computational power for high-quality rendering.

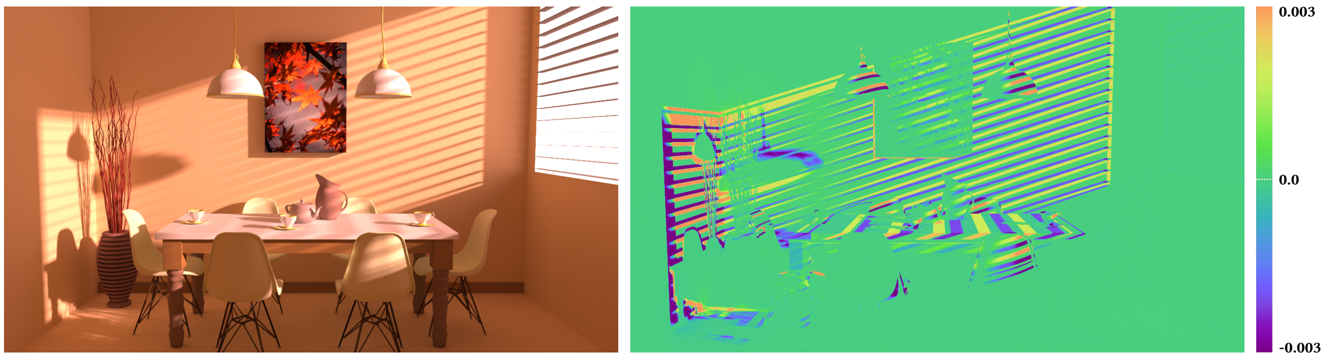
A multitude of techniques have been introduced to address this problem in recent years. Oculus Quest applies fixed foveated rendering, for which peripheral regions are rendered at low resolution. Kaplanyan et al. [2019] employ gaze-contingent foveated reconstruction by rendering non-uniform sparse pixel samples followed by

Authors' address: Lei Xiao, lei.xiao@fb.com; Salah Nouri, snouri@fb.com; Matt Chapman, mchapman@fb.com; Alexander Fix, alexander.fix@fb.com; Douglas Lanman, douglas.lanman@fb.com; Anton Kaplanyan, kaplanyan@fb.com, Facebook Reality Labs.

0730-0301/2020/7-ART142
https://doi.org/10.1145/3386569.3392376

Path-Space Differentiable Rendering

CHENG ZHANG, University of California, Irvine
BAILEY MILLER, Carnegie Mellon University
KAI YAN, University of California, Irvine
IOANNIS GKIOULEKAS, Carnegie Mellon University
SHUANG ZHAO, University of California, Irvine



Original | **Derivative** with respect to sun location

Fig. 1. We introduce **path-space differentiable rendering**, a new theoretical framework to estimate derivatives of radiometric measurements with respect to arbitrary scene parameters (e.g., material properties and object geometries). By directly differentiating full path integrals, we derive the *differential path integral* framework, enabling the design of new unbiased Monte Carlo methods capable of efficiently estimating derivatives in virtual scenes with complex geometry and light transport effects. This example shows a dinning room scene lit by the sun from outside the window. On the right, we show the corresponding derivative image with respect to the vertical location of the sun. (Please use Adobe Acrobat to view the teaser images to see them animated.)

Physics-based differentiable rendering, the estimation of derivatives of radiometric measures with respect to arbitrary scene parameters, has a diverse array of applications from solving analysis-by-synthesis problems to training machine learning pipelines incorporating forward rendering processes. Unfortunately, general-purpose differentiable rendering remains challenging due to the lack of efficient estimators as well as the need to identify and handle complex discontinuities such as visibility boundaries.

In this paper, we show how path integrals can be differentiated with respect to arbitrary differentiable changes of a scene. We provide a detailed theoretical analysis of this process and establish new differentiable rendering formulations based on the resulting differential path integrals. Our path-space differentiable rendering formulation allows the design of new Monte Carlo estimators that offer significantly better efficiency than state-of-the-art methods in handling complex geometric discontinuities and light transport phenomena such as caustics.

We validate our method by comparing our derivative estimates to those generated using the finite-difference method. To demonstrate the effectiveness of our technique, we compare inverse-rendering performance with a few state-of-the-art differentiable rendering methods.

Authors' addresses: Cheng Zhang, University of California, Irvine, chengz20@uci.edu; Bailey Miller, Carnegie Mellon University, baileymark.miller@gmail.com; Kai Yan, University of California, Irvine, kyan8@uci.edu; Ioannis Gkioulekas, Carnegie Mellon University, igkioule@andrew.cmu.edu; Shuang Zhao, University of California, Irvine, shz@ics.uci.edu.

0730-0301/2020/7-ART143 $15.00
https://doi.org/10.1145/3386569.3392383

CCS Concepts: • **Computing methodologies** → **Rendering**.

Additional Key Words and Phrases: Differentiable rendering, path integral, Monte Carlo rendering

ACM Reference Format:
Cheng Zhang, Bailey Miller, Kai Yan, Ioannis Gkioulekas, and Shuang Zhao. 2020. Path-Space Differentiable Rendering. *ACM Trans. Graph.* 39, 4, Article 143 (July 2020), 19 pages. https://doi.org/10.1145/3386569.3392383

1 INTRODUCTION

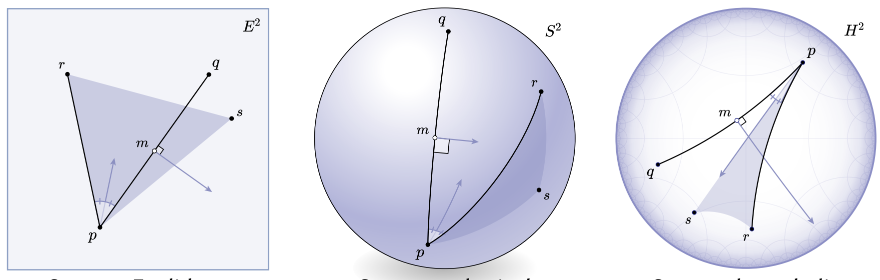
Physics-based light transport simulation, a core research topic in computer graphics since the field's inception, focus on numerically estimating radiometric sensor responses in fully specified virtual scenes. Previous research efforts have led to mature *forward rendering* algorithms that can efficiently and accurately simulate light transport in virtual environments with high complexities.

Differentiable rendering computes the derivatives of radiometric measurements with respect to differential changes of such environments. These techniques can enable, for example: (i) *gradient-based optimization* when solving *inverse-rendering* problems; and (ii) efficient integration of physics-based light transport simulation in *machine learning* and *probabilistic inference* pipelines.

Penrose: From Mathematical Notation to Beautiful Diagrams

KATHERINE YE, Carnegie Mellon University
WODE NI, Carnegie Mellon University
MAX KRIEGER, Carnegie Mellon University
DOR MA'AYAN, Technion and Carnegie Mellon University
JENNA WISE, Carnegie Mellon University
JONATHAN ALDRICH, Carnegie Mellon University
JOSHUA SUNSHINE, Carnegie Mellon University
KEENAN CRANE, Carnegie Mellon University

```
Point p, q, r, s
Segment a := {p, q}
Segment b := {p, r}
Point m := Midpoint(a)
Angle theta := ∠(q, p, r)
Triangle t := {p, r, s}
Ray w := Bisector(theta)
Ray h := PerpendicularBisector(a)
```



Style — Euclidean   Style — spherical   Style — hyperbolic

Fig. 1. Penrose is a framework for specifying how mathematical statements should be interpreted as visual diagrams. A clean separation between abstract mathematical objects and their visual representation provides new capabilities beyond existing code- or GUI-based tools. Here, for instance, the same set of statements *(left)* is given three different visual interpretations *(right)*, via Euclidean, spherical, and hyperbolic geometry. (Further samples are shown in Fig. 29.)

We introduce a system called Penrose for creating mathematical diagrams. Its basic functionality is to translate abstract statements written in familiar math-like notation into one or more possible visual representations. Rather than rely on a fixed library of visualization tools, the visual representation is user-defined in a constraint-based specification language; diagrams are then generated automatically via constrained numerical optimization. The system is user-extensible to many domains of mathematics, and is fast enough for iterative design exploration. In contrast to tools that specify diagrams via direct manipulation or low-level graphics programming, Penrose enables rapid creation and exploration of diagrams that faithfully preserve the underlying mathematical meaning. We demonstrate the effectiveness and generality of the system by showing how it can be used to illustrate a diverse set of concepts from mathematics and computer graphics.

CCS Concepts: • **Human-centered computing → Visualization toolkits**; • **Software and its engineering → Domain specific languages**.

Additional Key Words and Phrases: mathematical diagrams

Authors' addresses: Katherine Ye, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, 15213; Wode Ni, Carnegie Mellon University; Max Krieger, Carnegie Mellon University; Dor Ma'ayan, Technion and Carnegie Mellon University, Haifa, Israel; Jenna Wise, Carnegie Mellon University; Jonathan Aldrich, Carnegie Mellon University; Joshua Sunshine, Carnegie Mellon University; Keenan Crane, Carnegie Mellon University.

0730-0301/2020/7-ART144
https://doi.org/10.1145/3386569.3392375

ACM Reference Format:
Katherine Ye, Wode Ni, Max Krieger, Dor Ma'ayan, Jenna Wise, Jonathan Aldrich, Joshua Sunshine, and Keenan Crane. 2020. Penrose: From Mathematical Notation to Beautiful Diagrams. *ACM Trans. Graph.* 39, 4, Article 144 (July 2020), 16 pages. https://doi.org/10.1145/3386569.3392375

1 INTRODUCTION

"Mathematicians usually have fewer and poorer figures in their papers and books than in their heads."
—William Thurston

Effective communication of mathematical ideas is a major challenge for students, educators, and researchers. Though modern mathematics has a strong cultural bias toward formal language [Mashaal 2006], visualization and illustration undoubtedly have an equally profound impact on mathematical progress [Thurston 1998]. Yet the ability to translate abstract concepts into concrete illustrations is often limited to the select few individuals who have both a deep understanding of mathematics and an intimate knowledge of graphical tools. As a result, diagrams are rather scarce in mathematical writing—for instance, recent mathematical papers from *arXiv* have on average only one figure every ten pages. A central goal of this work is to lower the barrier to turning mathematical ideas into effective, high-quality visual diagrams. In the same way that TEX and LATEX have democratized mathematical writing by algorithmically codifying best practices of professional typesetters [Beeton and Palais 2016], Penrose aims to codify best practices of mathematical illustrators into a format that is reusable and broadly accessible.

Polar Stroking: New Theory and Methods for Stroking Paths

MARK J. KILGARD, NVIDIA

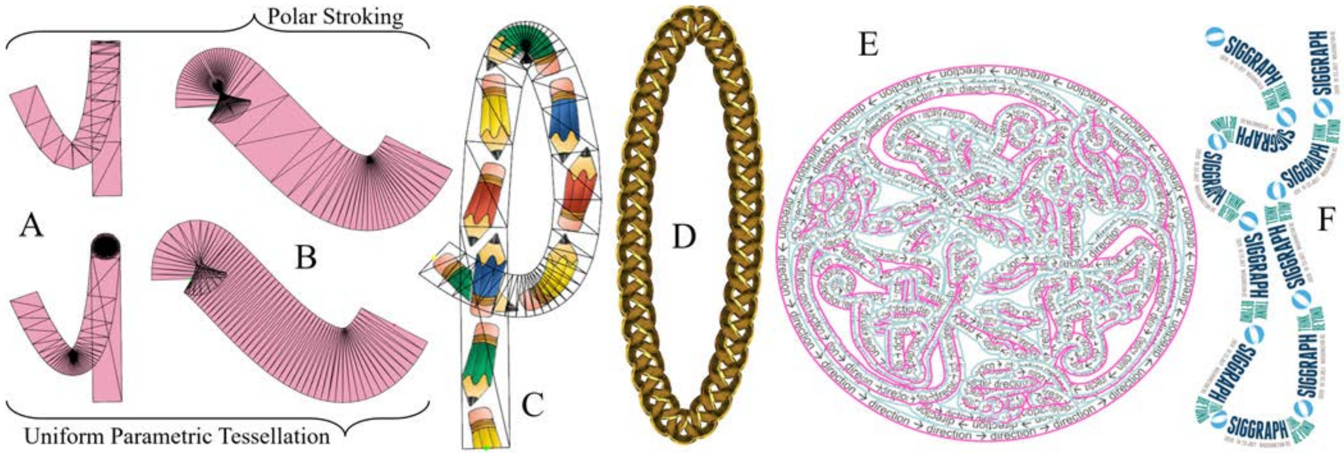


Fig. 1. Polar stroking samples: **A** cubic Bézier segment with a cusp rendered properly with polar stroking while uniform parametric tessellation has no cusp, both using 134 triangles; **B** polar stroking improves the facet angles distribution compared to uniform tessellation, both using 126 triangles; **C** arc length texturing; **D** ellipse drawn as just 2 conic segments, one external; **E** complex cubic Bézier path (5,031 path commands, 29,058 scalar path coordinates) with cumulative arc length texturing; **F** centripetal Catmull-Rom spline.

Stroking and filling are the two basic rendering operations on paths in vector graphics. The theory of filling a path is well-understood in terms of contour integrals and winding numbers, but when path rendering standards specify stroking, they resort to the analogy of painting pixels with a brush that traces the outline of the path. This means important standards such as PDF, SVG, and PostScript lack a rigorous way to say what samples are inside or outside a stroked path. Our work fills this gap with a principled theory of stroking.

Guided by our theory, we develop a novel *polar stroking* method to render stroked paths robustly with an intuitive way to bound the tessellation error without needing recursion. Because polar stroking guarantees small uniform steps in tangent angle, it provides an efficient way to accumulate arc length along a path for texturing or dashing. While this paper focuses on developing the theory of our polar stroking method, we have successfully implemented our methods on modern programmable GPUs.

CCS Concepts: • **Computing methodologies → Rasterization**.

Additional Key Words and Phrases: path rendering, vector graphics, stroking, offset curves

Author's address: Mark J. Kilgard, NVIDIA, mjk@nvidia.com.

0730-0301/2020/7-ART145 $15.00
https://doi.org/10.1145/3386569.3392458

ACM Reference Format:
Mark J. Kilgard. 2020. Polar Stroking: New Theory and Methods for Stroking Paths. *ACM Trans. Graph.* 39, 4, Article 145 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392458

1 INTRODUCTION

Vector graphics standards such as PDF [Adobe Systems 2008], SVG [SVG Working Group 2011], PostScript [Adobe Systems 1985], HTML5 Canvas [Whatwg.org 2011], PCL [Hewlett-Packard 1992], and XPS [ECMA International 2009] support two basic rendering operations on paths: stroking and filling.

The intuition of stroking a path is like a child drawing in a coloring book by "tracing over the lines" and treating each path as the outline to trace. Filling a path is like "coloring inside the lines."

The stroking operation on paths—mandated and specified by all the listed standards above—lacks a mathematically grounded theory to define what stroking means. To remedy this situation, we aim to provide a principled theory for stroking and show our theory motivates robust, useful, and GPU-amendable methods for stroking.

1.1 A Quick Theory of Path Filling

We first review the theory of path filling to show filling indeed has a principled theory—in contrast to path stroking.

When a path is filled, pixels "inside" the path get shaded and composited. At first glance, path filling sounds simple, but a path can be arbitrarily complex. It can be empty, concave (perhaps extremely so), intersect itself, contain multiple closed regions (some which wind clockwise while others the reverse), contain curved sections as well as straight ones, and may be degenerate in various ways (exhibiting cusps or closed regions with no interior). So a computer's decision

Radiative Backpropagation: An Adjoint Method for Lightning-Fast Differentiable Rendering

MERLIN NIMIER-DAVID, École Polytechnique Fédérale de Lausanne (EPFL)
SÉBASTIEN SPEIERER, École Polytechnique Fédérale de Lausanne (EPFL)
BENOÎT RUIZ, École Polytechnique Fédérale de Lausanne (EPFL)
WENZEL JAKOB, École Polytechnique Fédérale de Lausanne (EPFL)

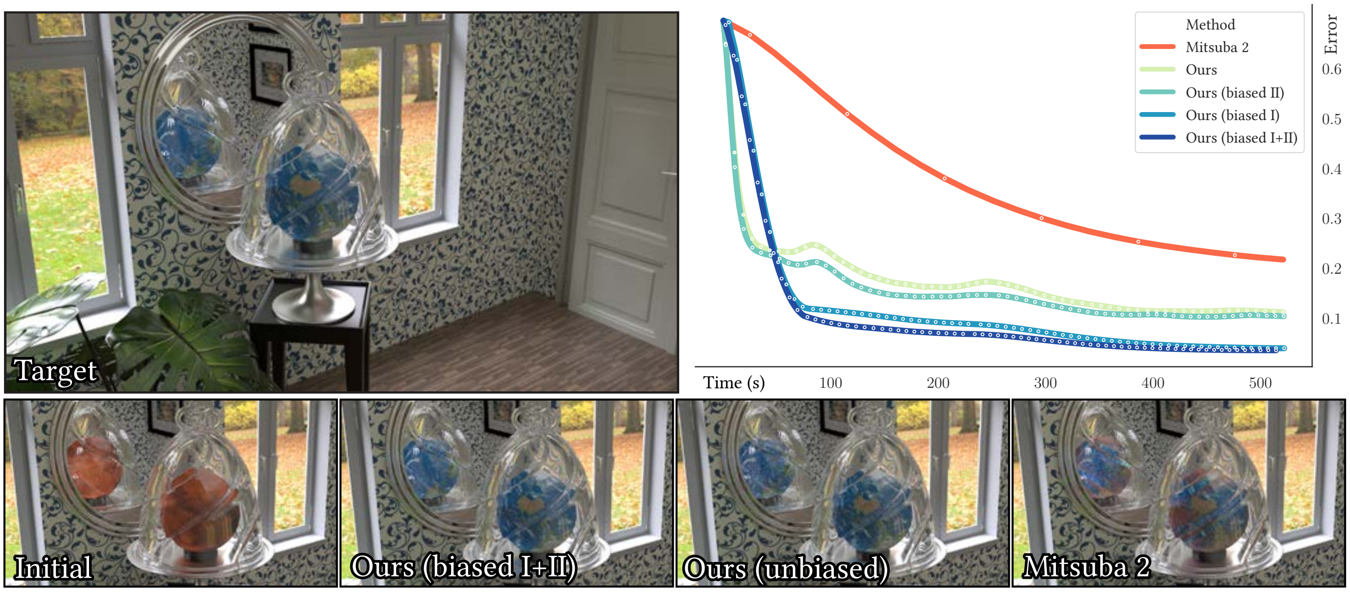


Fig. 1. Globe: Our method is able to reconstruct the texture of a globe seen through a bell jar in this interior scene with complex materials and interreflection. Starting from a different initialization (Mars), it attempts to match a reference rendering by differentiating scene parameters with respect to L_2 image distance. The plot on the right shows convergence over time for prior work [Nimier-David et al. 2019] and multiple variants of radiative backpropagation. Our method removes the severe overheads of differentiation compared to ordinary rendering, and we demonstrate speedups of up to ~1000× compared to prior work.

Physically based differentiable rendering has recently evolved into a powerful tool for solving inverse problems involving light. Methods in this area perform a differentiable simulation of the physical process of light transport and scattering to estimate partial derivatives relating scene parameters to pixels in the rendered image. Together with gradient-based optimization, such algorithms have interesting applications in diverse disciplines, e.g., to improve the reconstruction of 3D scenes, while accounting for interreflection and transparency, or to design meta-materials with specified optical properties.

The most versatile differentiable rendering algorithms rely on reverse-mode differentiation to compute all requested derivatives at once, enabling optimization of scene descriptions with millions of free parameters. However, a severe limitation of the reverse-mode approach is that it requires a detailed transcript of the computation that is subsequently replayed to back-propagate derivatives to the scene parameters. The transcript of typical renderings is extremely large, exceeding the available system memory by many orders of magnitude, hence current methods are limited to simple scenes rendered at low resolutions and sample counts.

We introduce *radiative backpropagation*, a fundamentally different approach to differentiable rendering that does not require a transcript, greatly improving its scalability and efficiency. Our main insight is that reverse-mode propagation through a rendering algorithm can be interpreted as the solution of a continuous transport problem involving the partial derivative of radiance with respect to the optimization objective. This quantity is "emitted" by sensors, "scattered" by the scene, and eventually "received" by objects with differentiable parameters. Differentiable rendering then decomposes into two separate primal and adjoint simulation steps that scale to complex scenes rendered at high resolutions. We also investigated biased variants of this algorithm and find that they considerably improve both runtime and convergence speed. We showcase an efficient GPU implementation of radiative backpropagation and compare its performance and the quality of its gradients to prior work.

Authors' addresses: Merlin Nimier-David, École Polytechnique Fédérale de Lausanne (EPFL), merlin.nimier-david@epfl.ch; Sébastien Speierer, École Polytechnique Fédérale de Lausanne (EPFL), sebastien.speierer@epfl.ch; Benoît Ruiz, École Polytechnique Fédérale de Lausanne (EPFL), benoit.ruiz@epfl.ch; Wenzel Jakob, École Polytechnique Fédérale de Lausanne (EPFL), wenzel.jakob@epfl.ch.

0730-0301/2020/7-ART146 $15.00
https://doi.org/10.1145/3386569.3392406

Robust Fitting of Parallax-Aware Mixtures for Path Guiding

LUKAS RUPPERT*, University of Tübingen, Germany
SEBASTIAN HERHOLZ*, University of Tübingen, Germany and Intel Corporation
HENDRIK P. A. LENSCH, University of Tübingen, Germany

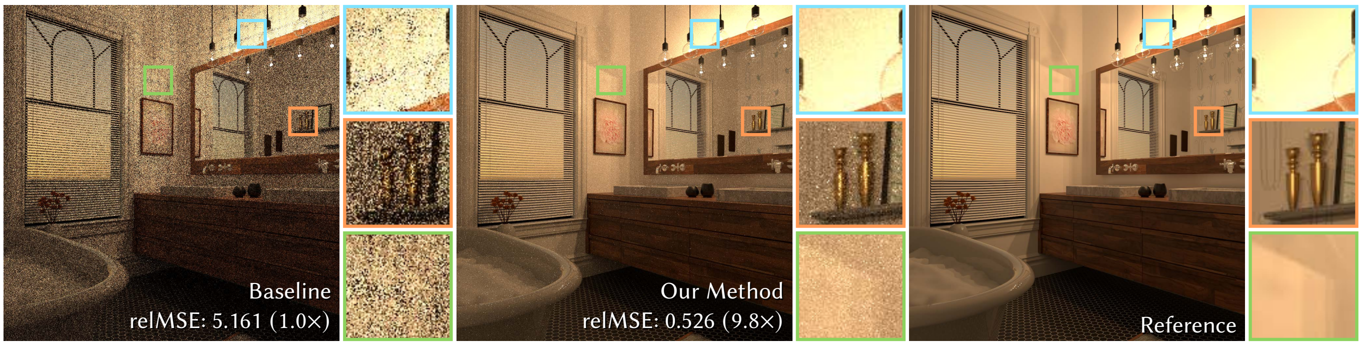


Fig. 1. Bathroom scene rendered using our guided path tracer implementation with next-event estimation in 10 minutes each. The "baseline" (left) corresponds to a configuration similar to previous guiding approaches. "Our method" (middle) features significantly improved direction sampling in local path guiding by introducing a robust fitting procedure for parallax-aware mixtures, achieving a speedup of 9.8 in this scene.

Effective local light transport guiding demands for high quality guiding information, i.e., a precise representation of the directional incident radiance distribution at every point inside the scene. We introduce a parallax-aware distribution model based on parametric mixtures. By parallax-aware warping of the distribution, the local approximation of the 5D radiance field remains valid and precise across large spatial regions, even for close-by contributors. Our robust optimization scheme fits parametric mixtures to radiance samples collected in previous rendering passes. Robustness is achieved by splitting and merging of components refining the mixture. These splitting and merging decisions minimize and bound the expected variance of the local radiance estimator. In addition, we extend the fitting scheme to a robust, iterative update method, which allows for incremental training of our model using smaller sample batches. This results in more frequent training updates and, at the same time, significantly reduces the required sample memory footprint. The parametric representation of our model allows for the application of advanced importance sampling methods such as radiance-based, cosine-aware, and even product importance sampling. Our method further smoothly integrates next-event estimation (NEE) into path guiding, avoiding importance sampling of contributions better covered by NEE. The proposed robust fitting and update scheme, in combination with the parallax-aware representation, results in faster learning and lower variance compared to state-of-the-art path guiding approaches.

*Both authors contributed equally to the paper.
Authors' addresses: Lukas Ruppert, University of Tübingen, Tübingen, Germany, lukas.ruppert@uni-tuebingen.de; Sebastian Herholz, University of Tübingen, Tübingen, Germany, Intel Corporation, sebastian.herholz@gmail.com; Hendrik P. A. Lensch, University of Tübingen, Tübingen, Germany, hendrik.lensch@uni-tuebingen.de.

0730-0301/2020/7-ART147 $15.00
https://doi.org/10.1145/3386569.3392421

CCS Concepts: • **Computing methodologies** → **Ray tracing**.

Additional Key Words and Phrases: Global Illumination, Ray Tracing, Sampling and Reconstruction, Stochastic Sampling

ACM Reference Format:
Lukas Ruppert, Sebastian Herholz, and Hendrik P. A. Lensch. 2020. Robust Fitting of Parallax-Aware Mixtures for Path Guiding. *ACM Trans. Graph.* 39, 4, Article 147 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392421

1 INTRODUCTION

Path-based rendering algorithms have become the industry standard for solving light transport simulations [Fascione et al. 2017, 2018, 2019]. Especially uni-directional path tracing [Kajiya 1986] is now the favorable choice, due to its simplicity to implement and its extensibility to incorporate different sampling strategies or estimators (e.g., next-event estimation). The basic path tracing algorithm achieves a significant gain in quality and efficiency with proper importance sampling, which for global illumination needs to consider both the distribution of the scattering function as well as an approximation of the true radiance distribution. Building on initial work of Jensen [1995] and Lafortune and Willems [1995], recent path guiding approaches, e.g. [Vorba et al. 2014; Müller et al. 2017], use representations of the 5D-radiance distribution that are either learned form a pre-processing step or online during rendering (a.k.a., forward-learning). Approaches based on forward-learning are preferred in practice, as their ability to provide previews more quickly can improve the turnaround times of artists. With local path guiding, even complex light transport effects such as caustics and multiple diffuse bounces can be rendered reliably using simple forward path tracing and, thus, it is now used in production rendering [Vorba et al. 2019].

As an imprecise guiding function generally increases variance rather than reducing it, both the chosen representation and a robust

Spatiotemporal reservoir resampling for real-time ray tracing with dynamic direct lighting

BENEDIKT BITTERLI, Dartmouth College
CHRIS WYMAN, NVIDIA
MATT PHARR, NVIDIA
PETER SHIRLEY, NVIDIA
AARON LEFOHN, NVIDIA
WOJCIECH JAROSZ, Dartmouth College

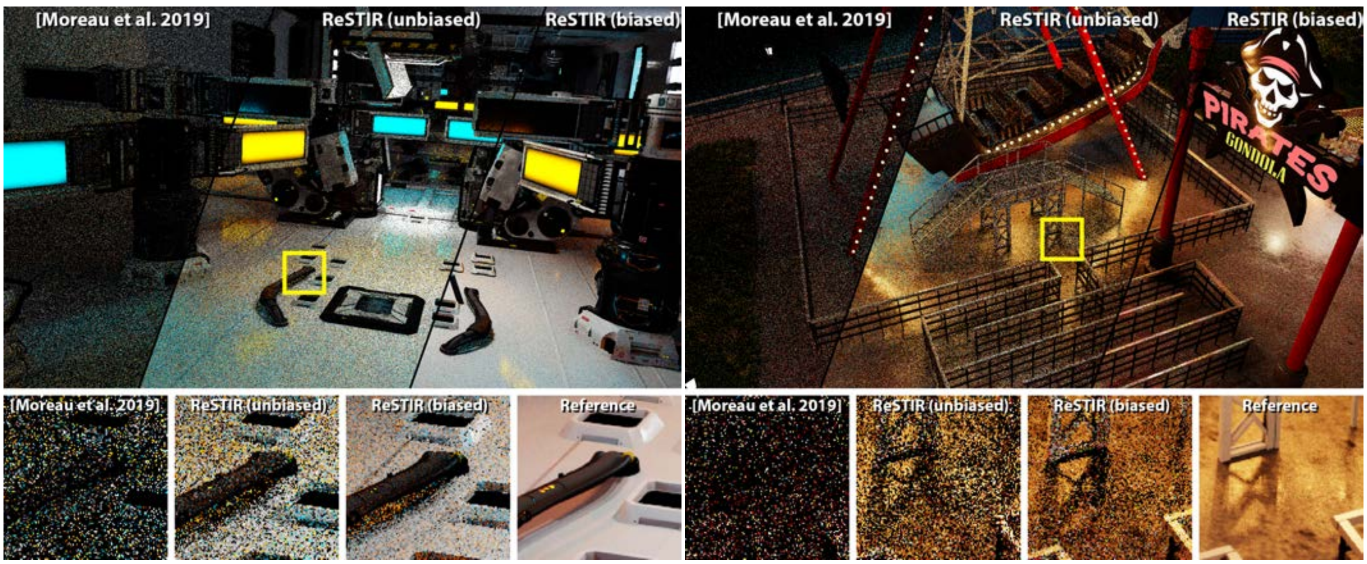


Fig. 1. Two complex scenes ray traced with direct lighting from many dynamic lights. (Left) A still from the Zero Day video [Winkelmann 2015] with 11,000 dynamic emissive triangles. (Right) A view of one ride in an Amusement Park scene containing 3.4 million dynamic emissive triangles. Both images show three methods running in equal time on a modern GPU, from left to right: Moreau et al. [2019]'s efficient light-sampling BVH, our new unbiased estimator, and our new biased estimator. The Zero Day image is rendered in 15 ms and Amusement Park in 50 ms, both at 1920 × 1080 resolution. Zero Day ©beeple, Pirate Ship ©sema edis

Efficiently rendering direct lighting from millions of dynamic light sources using Monte Carlo integration remains a challenging problem, even for off-line rendering systems. We introduce a new algorithm—ReSTIR—that renders such lighting interactively, at high quality, and without needing to maintain complex data structures. We repeatedly resample a set of candidate light samples and apply further spatial and temporal resampling to leverage information from relevant nearby samples. We derive an unbiased Monte Carlo estimator for this approach, and show that it achieves equal-error 6×-60× faster than state-of-the-art methods. A biased estimator reduces noise further and is 35×-65× faster, at the cost of some energy loss. We implemented our approach on the GPU, rendering complex scenes containing up to 3.4 million dynamic, emissive triangles in under 50 ms per frame while tracing at most 8 rays per pixel.

Authors' addresses: Benedikt Bitterli, Dartmouth College, benedikt.m.bitterli.gr@dartmouth.edu; Chris Wyman, NVIDIA, chris.wyman@acm.org; Matt Pharr, NVIDIA, matt.pharr@gmail.com; Peter Shirley, NVIDIA, ptrshrl@gmail.com; Aaron Lefohn, NVIDIA, 2788 San Tomas Expressway, Santa Clara, CA, 95051, alefohn@nvidia.com; Wojciech Jarosz, Dartmouth College, Department of Computer Science, 9 Maynard St. Hanover, NH, 03755, wojciech.k.jarosz@dartmouth.edu.

0730-0301/2020/7-ART148 $15.00
https://doi.org/10.1145/3386569.3392481

CCS Concepts: • **Computing methodologies → Ray tracing**.

Additional Key Words and Phrases: Photorealistic rendering, resampled importance sampling, real-time rendering, reservoir sampling

ACM Reference Format:
Benedikt Bitterli, Chris Wyman, Matt Pharr, Peter Shirley, Aaron Lefohn, and Wojciech Jarosz. 2020. Spatiotemporal reservoir resampling for real-time ray tracing with dynamic direct lighting. *ACM Trans. Graph.* 39, 4, Article 148 (July 2020), 17 pages. https://doi.org/10.1145/3386569.3392481

Specular Manifold Sampling for Rendering High-Frequency Caustics and Glints

TIZIAN ZELTNER, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
ILIYAN GEORGIEV, Autodesk, United Kingdom
WENZEL JAKOB, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Fig. 1. Rendering of a shop window featuring a combination of challenging-to-sample light transport paths with specular-diffuse-specular ("SDS") interreflection: the two golden normal-mapped pedestals are illuminated by spot lights and project intricate caustic patterns following a single reflection from the metallic surface, while the transparent center pedestal generates caustics via double refraction. The glinty appearance of the shoes arises due to specular microgeometry encoded in a high-frequency normal map. This image was rendered by an ordinary unidirectional path tracer using our new specular manifold sampling strategy. The remaining noise is due to indirect lighting by caustics, which is not explicitly sampled by our technique. The background image is "Hexactinellae" from *Art Forms in Nature* by Ernst Haeckel.

Scattering from specular surfaces produces complex optical effects that are frequently encountered in realistic scenes: intricate caustics due to focused reflection, multiple refraction, and high-frequency glints from specular microstructure. Yet, despite their importance and considerable research to this end, sampling of light paths that cause these effects remains a formidable challenge.

In this article, we propose a surprisingly simple and general sampling strategy for specular light paths including the above examples, unifying the previously disjoint areas of caustic and glint rendering into a single framework. Given two path vertices, our algorithm stochastically finds a specular subpath connecting the endpoints. In contrast to prior work, our method supports high-frequency normal- or displacement-mapped geometry, samples specular-diffuse-specular ("SDS") paths, and is compatible with standard Monte Carlo methods including unidirectional path tracing. Both unbiased and biased variants of our approach can be constructed, the latter often significantly reducing variance, which may be appealing in applied settings (e.g. visual effects). We demonstrate our method on a range of challenging scenes and evaluate it against state-of-the-art methods for rendering caustics and glints.

CCS Concepts: • **Computing methodologies → Rendering**.

Additional Key Words and Phrases: Specular light pahts, SDS paths, Caustics, Glints

ACM Reference Format:
Tizian Zeltner, Iliyan Georgiev, and Wenzel Jakob. 2020. Specular Manifold Sampling for Rendering High-Frequency Caustics and Glints. *ACM Trans. Graph.* 39, 4, Article 149 (July 2020), 15 pages. https://doi.org/10.1145/3386569.3392408

Authors' addresses: Tizian Zeltner, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, tizian.zeltner@epfl.ch; Iliyan Georgiev, Autodesk, United Kingdom, iliyan.georgiev@autodesk.com; Wenzel Jakob, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, wenzel.jakob@epfl.ch.

0730-0301/2020/7-ART149 $15.00
https://doi.org/10.1145/3386569.3392408

The design and evolution of the UberBake light baking system

DARIO SEYB*, Dartmouth College
PETER-PIKE SLOAN*, Activision Publishing
ARI SILVENNOINEN, Activision Publishing
MICHAŁ IWANICKI, Activision Publishing
WOJCIECH JAROSZ, Dartmouth College

Fig. 1. Our system allows for *player-driven* lighting changes at run-time. Above we show a scene where a door is opened during gameplay. The image on the left shows the final lighting produced by our system as seen in the game. In the middle, we show the scene without the methods described here (top). Our system enables us to efficiently precompute the associated lighting change (bottom). This functionality is built on top of a dynamic light set system which allows for levels with hundreds of lights who's contribution to global illumination can be controlled individually at run-time (right). ©Activision Publishing, Inc.

We describe the design and evolution of UberBake, a global illumination system developed by Activision, which supports limited lighting changes in response to certain player interactions. Instead of relying on a fully dynamic solution, we use a traditional static light baking pipeline and extend it with a small set of features that allow us to dynamically update the precomputed lighting at run-time with minimal performance and memory overhead. This means that our system works on the complete set of target hardware, ranging from high-end PCs to previous generation gaming consoles, allowing the use of lighting changes for gameplay purposes. In particular, we show how to efficiently precompute lighting changes due to individual lights being enabled and disabled and doors opening and closing. Finally, we provide a detailed performance evaluation of our system using a set of production levels and discuss how to extend its dynamic capabilities in the future.

*Both authors contributed equally to this research.

Authors' addresses: Dario Seyb, dario.r.seyb.gr@dartmouth.edu, Dartmouth College; Peter-Pike Sloan, ppsloan@activision.com, Activision Publishing; Ari Silvennoinen, ari.silvennoinen@activision.com, Activision Publishing; Michał Iwanicki, michal.iwanicki@activision.com, Activision Publishing; Wojciech Jarosz, wjarosz@dartmouth.edu, Dartmouth College.

0730-0301/2020/7-ART150 $15.00
https://doi.org/10.1145/3386569.3392394

CCS Concepts: • **Computing methodologies → Ray tracing**; *Graphics systems and interfaces.*

Additional Key Words and Phrases: global illumination, baked lighting, real time systems

ACM Reference Format:
Dario Seyb, Peter-Pike Sloan, Ari Silvennoinen, Michał Iwanicki, and Wojciech Jarosz. 2020. The design and evolution of the UberBake light baking system. *ACM Trans. Graph.* 39, 4, Article 150 (July 2020), 13 pages. https://doi.org/10.1145/3386569.3392394

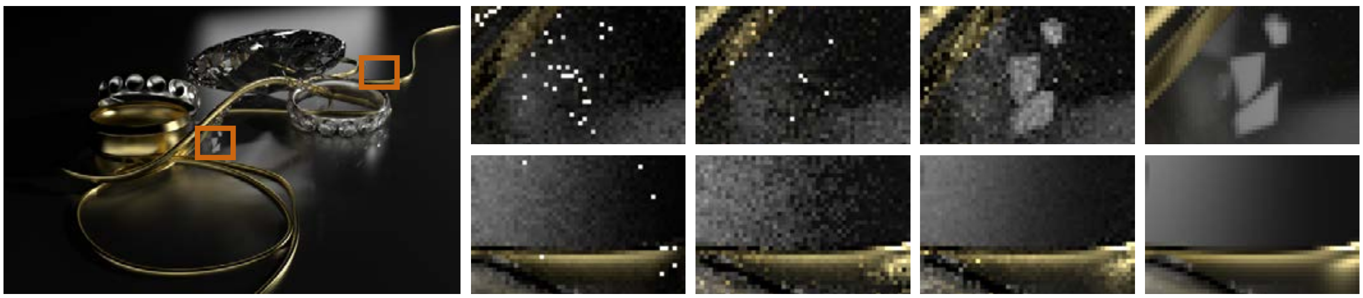
1 INTRODUCTION

AAA games today produce images at real-time frame rates (usually 30 or 60 frames per second) that can rival the realism and complexity of offline rendered movies from just a few years ago. This leaves just 16–30 ms to simulate the virtual environment, react to player input, and produce images showing a wide range of complex light-transport phenomena. This last goal can be especially challenging, as players enjoy games on a variety of hardware platforms and comparable quality needs to be achieved on all of them, including ones less powerful than the state of the art, such as mobile devices or previous generation consoles.

One of the difficulties of the rendering process is computing *global illumination*—the component of the lighting that arrives at each point not directly from a light source, but after some number of bounces off other surfaces in the scene. Given the limited time budget, most modern game engines rely on some form of precomputation or *baking*. Parts of the lighting are computed offline, stored in

Variance-Aware Path Guiding

ALEXANDER RATH, DFKI and Saarland University, Germany
PASCAL GRITTMANN, Saarland University, Germany
SEBASTIAN HERHOLZ, University of Tübingen, Germany and Intel Corporation
PETR VÉVODA, Charles University, Prague and Chaos Czech a.s., Czech Republic
PHILIPP SLUSALLEK, DFKI and Saarland University, Germany
JAROSLAV KŘIVÁNEK, Charles University, Prague and Chaos Czech a.s., Czech Republic



NECKLACE | (a) Path tracer 89.9 | (b) Radiance-based 0.4 (baseline) | (c) Our target density **0.16 (2.6x)** | (d) Reference *relMSE*

Fig. 1. The guiding approach of Müller et al. [2017] benefits greatly from our target densities, e.g., for caustics on glossy surfaces, as shown here. Our method consists of a trivial modification applicable to a variety of path guiding algorithms without additional parameters or computational overhead.

Path guiding is a promising tool to improve the performance of path tracing algorithms. However, not much research has investigated what target densities a guiding method should strive to learn for optimal performance. Instead, most previous work pursues the zero-variance goal: The local decisions are guided under the assumption that all other decisions along the random walk will be sampled perfectly. In practice, however, many decisions are poorly guided, or not guided at all. Furthermore, learned distributions are often marginalized, e.g., by neglecting the BSDF. We present a generic procedure to derive theoretically optimal target densities for local path guiding. These densities account for variance in nested estimators, and marginalize provably well over, e.g., the BSDF. We apply our theory in two state-of-the-art rendering applications: a path guiding solution for unidirectional path tracing [Müller et al. 2017] and a guiding method for light source selection for the many lights problem [Vévoda et al. 2018]. In both cases, we observe significant improvements, especially on glossy surfaces. The implementations for both applications consist of trivial modifications to the original code base, without introducing any additional overhead.

Authors' addresses: Alexander Rath, alexander.rath@dfki.de, DFKI, Saarland University, Saarbrücken, Germany; Pascal Grittmann, Saarland University, Saarbrücken, Germany, grittmann@cg.uni-saarland.de; Sebastian Herholz, University of Tübingen, Tübingen, Germany, Intel Corporation, sebastian.herholz@gmail.com; Petr Vévoda, Charles University, Prague, Chaos Czech a.s., Czech Republic, petrvevoda@seznam.cz; Philipp Slusallek, DFKI, Saarland University, Saarbrücken, Germany, philipp.slusallek@dfki.de; Jaroslav Křivánek, Charles University, Prague, Chaos Czech a.s., Czech Republic.

0730-0301/2020/7-ART151 $15.00
https://doi.org/10.1145/3386569.3392441

CCS Concepts: • **Computing methodologies → Ray tracing**.

Additional Key Words and Phrases: global illumination, importance sampling, path guiding

ACM Reference Format:
Alexander Rath, Pascal Grittmann, Sebastian Herholz, Petr Vévoda, Philipp Slusallek, and Jaroslav Křivánek. 2020. Variance-Aware Path Guiding. *ACM Trans. Graph.* 39, 4, Article 151 (July 2020), 12 pages. https://doi.org/10.1145/3386569.3392441

1 INTRODUCTION

The majority of rendering systems today rely on unidirectional path tracers [Burley et al. 2018; Fascione et al. 2018; Georgiev et al. 2018; Keller et al. 2015]. The simplicity, flexibility, and extensibility of the algorithm is what makes it so appealing. The performance, however, depends heavily on the employed importance sampling strategy. Ideally, paths should be sampled proportionally to their pixel contribution. Unfortunately, computing that ideal distribution is a harder problem than rendering the image, because it would require knowledge of the full light transport in the scene. Hence, many implementations construct paths by locally sampling from coarse approximations, like BSDF importance sampling.

Path guiding methods learn better importance sampling densities, either locally or for full paths, based on information gathered from previous rendering iterations [Vorba et al. 2019]. The learned densities are then used to importance sample paths in future iterations.

Learning the optimal sampling density for a complete path is often infeasible, due to the high dimensionality [Müller et al. 2018; Zheng and Zwicker 2019]. Alternatively, it is theoretically possible to construct an optimal path with only local decisions. To achieve