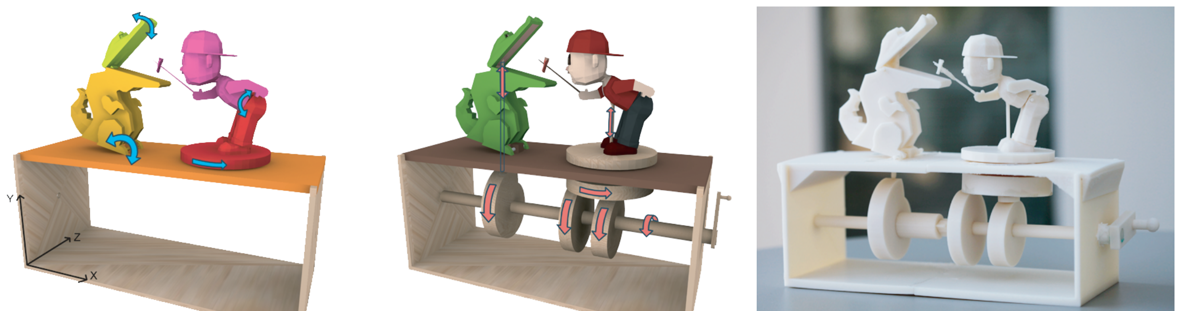
Motion-Guided Mechanical Toy Modeling

Lifeng Zhu* Weiwei Xu†‡ John Snyder § Yang Liu ‡ Guoping Wang* Baining Guo‡
Peking University* Hangzhou Normal University† Microsoft Research Asia‡ Microsoft Research§



(a) (b) (c)

Figure 1: *Example mechanical toy: Crocodile Feeding. (a) Input. The designer specifies the geometry and motion of the toy's features, in this case a boy and a crocodile object, forming two kinematic chains and four color-coded feature components. The feature base is colored orange. (b) Mechanical assembly synthesized by our system to generate the target motion. (c) Fabricated result. Overlayed arrows illustrate the motion, both input for features in (a) and output for the synthesized mechanism in (b), via the rules in [Mitra et al. 2010]. The canonical local coordinate system for the mechanical assembly is shown in (a). Please see the accompanying video for the full animation.*

Abstract

We introduce a new method to synthesize mechanical toys solely from the motion of their features. The designer specifies the geometry and a time-varying rotation and translation of each rigid feature component. Our algorithm automatically generates a mechanism assembly located in a box below the feature base that produces the specified motion. Parts in the assembly are selected from a parameterized set including belt-pulleys, gears, crank-sliders, quick-returns, and various cams (snail, ellipse, and double-ellipse). Positions and parameters for these parts are optimized to generate the specified motion, minimize a simple measure of complexity, and yield a well-distributed layout of parts over the driving axes. Our solution uses a special initialization procedure followed by simulated annealing to efficiently search the complex configuration space for an optimal assembly.

Keywords: forward and inverse kinematics, MCAD, mechanism synthesis, simulated annealing.

Links: DL PDF

*lfzhulf@gmail.com, wgp@pku.edu.cn, Key Lab of Machine Perception(MOE). This work was done when Lifeng Zhu was an intern at Microsoft Research Asia.
†weiwei.xu.g@gmail.com, corresponding author
‡YangLiu@microsoft.com, bainguo@microsoft.com
§johnsny@microsoft.com

ACM Reference Format
Zhu, L., Xu, W., Snyder, J., Liu, Y., Wang, G., Guo, B. 2012. Motion-Guided Mechanical Toy Modeling. *ACM Trans. Graph. 31* 6, Article 127 (November 2012), 10 pages. DOI = 10.1145/2366145.2366146 http://doi.acm.org/10.1145/2366145.2366146.

1 Introduction

Mechanical toy automata produce fascinating motion. Their history dates back to at least 400BC with the flying pigeon of Archytas of Tarentum. Modern versions range from large exhibits such as the Cabaret Mechanical Theater in the United Kingdom and the Swiss Jolly Ball at the Museum of Science and Industry in Chicago to playthings used by millions of children. Mechanisms in these automata drive the motion of the toy's whimsical features, and form an interesting attribute in their own right. Today, mechanical automata represent a popular topic as witnessed by the many books and DVDs covering their design and fabrication (e.g., [Neufeld 2003; Peppe 2005; Frost 2007]).

Mechanical toy design integrates a range of skills from art and craftsmanship to mechanics and mathematics. The designer typically begins by sketching the shapes and motions of the toy's features. He then plans a mechanical assembly to generate this motion. Inevitably, the design is iterated by refining the parameters of its different parts such as cams, gears, cranks, etc., to tune the motion and optimize other relevant factors such as manufacturing cost. The process is difficult and time-consuming, and demands an understanding of the complicated behavior of mechanisms, as well as a balancing of different design desiderata.

We automate design for mechanical toys such as the one in Figure 1. Our larger goal is to integrate kinematic simulation of mechanical assemblies into 3D modeling. Integrated simulation allows a motion-to-form mapping: conversion of user-specified motion into a physical, functioning mechanism. A toy designed in our system can be directly fabricated into a real object using a 3D printer, as shown in Figure 1c. We think custom mechanical design for ordinary users and for recreational and educational purposes will become increasingly important with the advent of inexpensive 3D printing. We also believe our approach can be generalized to other types of high-level mechanism design.

Our algorithm's input is the 3D geometry and motion of the toy's top-level features. These consist of rigid *feature components* connected by joints, to form one or more kinematically-linked chains. Its output is a *mechanical assembly* located below the base of features. The assembly consists of individual *parts*, such as cams,

Recursive Interlocking Puzzles

Peng Song Chi-Wing Fu
Nanyang Technological University

Daniel Cohen-Or
Tel Aviv University

Abstract

Interlocking puzzles are very challenging geometric problems with the fascinating property that once we solve one by putting together the puzzle pieces, the puzzle pieces interlock with one another, preventing the assembly from falling apart. Though interlocking puzzles have been known for hundreds of years, very little is known about the governing mechanics. Thus, designing new interlocking geometries is basically accomplished with extensive manual effort or expensive exhaustive search with computers.

In this paper, we revisit the notion of interlocking in greater depth, and devise a formal method of the interlocking mechanics. From this, we can develop a constructive approach for devising new interlocking geometries that directly guarantees the validity of the interlocking instead of exhaustively testing it. In particular, we focus on an interesting subclass of interlocking puzzles that are recursive in the sense that the assembly of puzzle pieces can remain an interlocking puzzle also after sequential removal of pieces; there is only one specific sequence of assembling, or disassembling, such a puzzle. Our proposed method can allow efficient generation of recursive interlocking geometries of various complexities, and by further realizing it with LEGO bricks, we can enable the hand-built creation of custom puzzle games.

CR Categories: J.6 [Computer Applications]—Computer-aided design; K.8 [Personal Computing]: Games

Keywords: Computer-aided design, interlocking, 3D puzzles

Links: DL PDF

1 Introduction

3D puzzles are generally nontrivial geometric problems that challenge our ingenuity. They have been longstanding stimulating recreational artifacts, where the task is to put together the puzzle pieces to form a meaningful 3D shape. Such an endeavor requires spatial cognitive skills for recognizing and understanding patterns and 3D structures, as observed from different angles. Interestingly enough, while solving a 3D puzzle assembly problem is already an intricate task, the creation of nontrivial puzzles is an even greater challenge.

Among the various families of 3D puzzles, a particularly challenging case is interlocking puzzles, for which one has to identify and follow certain orders to assemble the puzzle pieces into the target shape. The fascinating property of interlocking puzzles is that once they are assembled, the puzzle pieces interlock with one another, preventing the 3D assembly from falling apart. More precisely, all puzzle pieces become immobilized except one single key piece, which is the last puzzle piece inserted to the assembly, and also the first required to be removed in the disassembly.

ACM Reference Format
Song, P., Fu, C., Cohen-Or, D. 2012. Recursive Interlocking Puzzles. *ACM Trans. Graph. 31* 6, Article 128 (November 2012), 10 pages. DOI = 10.1145/2366145.2366147
http://doi.acm.org/10.1145/2366145.2366147.

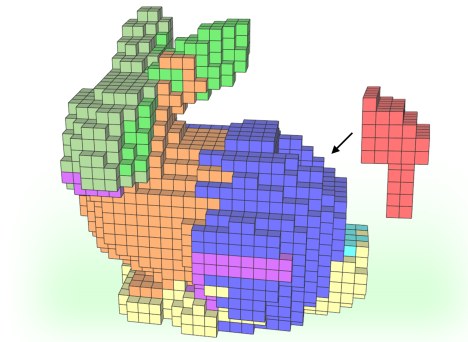


Figure 1: *A* 10*-piece recursive interlocking* BUNNY *puzzle, where the red piece is the first key in the puzzle.*

The geometric properties of interlocking puzzles are highly intriguing, and have been a virtue attracting architecture designers because the structure remains steady and solid while being glue-less, nail-less, and screw-less, and yet can be repeatedly assembled and disassembled. Nevertheless, little is known about the governing mechanics of interlocking puzzles [Cutler 1994; IBM Research 1997; Xin et al. 2011], and we are unaware of any previous research work in various communities, including graphics, computer aided design, and computational geometry. In this paper, we revisit the notion of interlocking in greater depth and develop a computational method for creating interlocking puzzles of varying complexities.

We define an interlocking puzzle as follows:

> An assembly of puzzle pieces (with at least three pieces) is said to be interlocking if there exists only one movable puzzle piece, while all other puzzle pieces, as well as any subset of the puzzle pieces, are immobilized relative to one another.

This definition stresses the difficulty of creating interlocking structures because, geometrically, the immobilization of individual puzzle pieces does not imply the immobilization of subsets of puzzle pieces. Thus, a naive decomposition of a shape into interlocking pieces would require (i) testing the immobilization of all the subsets of pieces, and (ii) testing that the decomposition is not deadlocked. Such tests lead to extremely expensive algorithms. Hence, previous attempts to discover new interlocking puzzles based on exhaustive search [Cutler 1978; Cutler 1994; IBM Research 1997] are extremely expensive. Other attempts, which are created manually, are very limited and still require enormous effort.

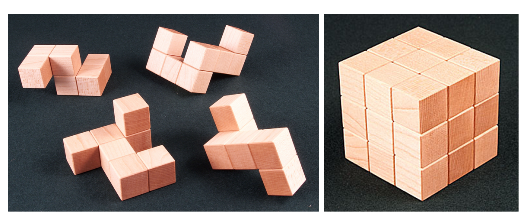


Figure 2: *Coffin's four-piece interlocking cube.*

Chopper: Partitioning Models into 3D-Printable Parts

Linjie Luo[1,2] Ilya Baran[3] Szymon Rusinkiewicz[1] Wojciech Matusik[4]
[1]Princeton University [2]Disney Research Boston [3]Disney Research Zurich [4]MIT CSAIL

Figure 1: *Chopper partitions a given 3D model into parts that are small enough to be 3D-printed and assembled into the original model. Left to right: the input chair model, Chopper's partition (with a printing volume shown as a reference), printed parts, and assembled chair.*

Abstract

3D printing technology is rapidly maturing and becoming ubiquitous. One of the remaining obstacles to wide-scale adoption is that the object to be printed must fit into the working volume of the 3D printer. We propose a framework, called Chopper, to decompose a large 3D object into smaller parts so that each part fits into the printing volume. These parts can then be assembled to form the original object. We formulate a number of desirable criteria for the partition, including assemblability, having few components, unobtrusiveness of the seams, and structural soundness. Chopper optimizes these criteria and generates a partition either automatically or with user guidance. Our prototype outputs the final decomposed parts with customized connectors on the interfaces. We demonstrate the effectiveness of Chopper on a variety of non-trivial real-world objects.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Geometric Algorithms;

Keywords: 3D printing, mesh segmentation and decomposition

Links: DL PDF

1 Introduction

As 3D printing technology matures, becoming cheaper and simpler to use, printing ever-larger objects becomes feasible. At the same time, the maximum size of an object that a 3D printer can produce in one pass (the *printing volume*) is limited by practical considerations. Larger objects must therefore be printed as separate parts and assembled. Existing commercial systems rely on manual partitioning, but this can be tedious if many parts are needed.

Recent work in the field of computer graphics, however, has investigated computation-assisted fabrication approaches in which the user specifies the desired visual or physical properties of an object to be manufactured, and an optimization is used to determine the precise way in which fabrication devices should be driven to produce an object as close as possible to the goal. The optimization may be formulated to consider visual properties such as scattering [Hašan et al. 2010] or physical properties such as deformation [Bickel et al. 2010], but in this paper we focus purely on 3D shape. We propose that this kind of optimization approach may be applied to the problem of decomposition into print volumes, yielding an automated method that is quicker, easier, and in many instances better than can be accomplished by hand.

This work presents a framework, called *Chopper*, for partitioning objects for 3D printing. Because the number of ways in which an object may be partitioned is large, we seek to find a decomposition that optimizes a number of (sometimes conflicting) objectives:

- **Printability:** the parts must fit inside the working volume.
- **Assemblability:** it must be possible to put parts together (without interference) into a finished model.
- **Efficiency:** the partition should avoid small parts and, in general, minimize the number of required sub-volumes.
- **Connector feasibility:** each interface must be large enough to admit connectors, protrusions, or other aids to assembly.
- **Structural soundness:** parts should not have thin slivers, and seams should be away from areas of high mechanical stress.
- **Aesthetics:** seams should be unobtrusive, detracting from appearance as little as possible, and should follow the natural symmetries of the model.

Formalizing these objectives is not easy, since different use cases have different requirements. Moreover, it is easy to propose definitions for the objective functions that lead to sub-optimal partitions or interact in unwanted ways. We propose specific definitions (Sec-

ACM Reference Format
Luo, L., Baran, I., Rusinkiewicz, S., Matusik, W. 2012. Chopper: Partitioning Models into 3D-Printable Parts. *ACM Trans. Graph. 31* 6, Article 129 (November 2012), 9 pages. DOI = 10.1145/2366145.2366148 http://doi.acm.org/10.1145/2366145.2366148.

3D-Printing of Non-Assembly, Articulated Models

Jacques Calì Dan A. Calian Cristina Amati Rebecca Kleinberger Anthony Steed Jan Kautz Tim Weyrich
University College London

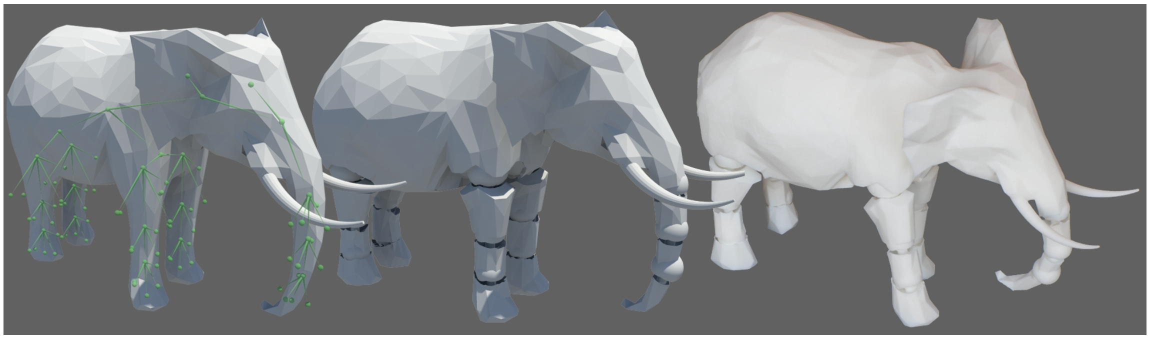


Figure 1: *Starting from a 3D mesh (*left*), our system allows to intuitively add 3D-printable joints (*center*) that, when 3D-printed, create a functional, posable model with joints that exhibit internal friction. The model leaves the printer ready to use; no manual assembly is required.*

Abstract

Additive manufacturing (3D printing) is commonly used to produce physical models for a wide variety of applications, from archaeology to design. While static models are directly supported, it is desirable to also be able to print models with functional articulations, such as a hand with joints and knuckles, without the need for manual assembly of joint components. Apart from having to address limitations inherent to the printing process, this poses a particular challenge for articulated models that should be posable: to allow the model to hold a pose, joints need to exhibit internal friction to withstand gravity, without their parts fusing during 3D printing. This has not been possible with previous printable joint designs. In this paper, we propose a method for converting 3D models into printable, functional, non-assembly models with internal friction. To this end, we have designed an intuitive workflow that takes an appropriately rigged 3D model, automatically fits novel 3D-printable and posable joints, and provides an interface for specifying rotational constraints. We show a number of results for different articulated models, demonstrating the effectiveness of our method.

Keywords: 3D Printing, Articulated Models, Mechanical Joints

Links: DL PDF

1 Introduction

Additive manufacturing (3D printing) is commonly used to produce physical models for a wide variety of applications, from archaeology to industrial prototyping and design. While having been available for more than a decade, the technology recently developed additional momentum. With 3D printers dropping in price and with 3D printing services becoming available to a wider public, many additional use cases arise. For instance, rapid prototyping is now even used to print custom toys and figurines: there are commercial services that fabricate figurines from multiple-choice designs of puppets [MakieLab 2012] and robot toys [Kodama Studios 2012], or even based on models from World of Warcraft and other online games [FigurePrints LLC 2012].

Our work is inspired by—but not limited to—such figurine production. Observing the existing tools for fabrication of articulated models, it becomes apparent that they only allow for the fabrication of *static* models, or employ prefabricated parts to provide a limited degree of articulation of a *manually assembled* model. This evidently does not fully exploit the flexibility of modern 3D printing processes. In this paper, we show how to equip 3D models with *custom articulation*, with minimal user input, leading to designs that can directly be 3D-printed, without the need for manual part assembly.

3D printing of functional models has been researched as long as 3D printers exist. Simple sets of gears have been demonstrated early on, and by now they even attracted a hobbyist community [Thingiverse 2012]. Custom deformable objects have been printed using using a multi-material printer [Bickel et al. 2010]. Closely related to our work, several 3D-printable joints have been designed to be used in a robotic hand [Won et al. 2000; Mavroidis et al. 2001; Laurentis and Mavroidis 2004]; however, these joints require manual assembly and were manually designed for a specific model. Designing a different jointed model would require considerable amount of work. Also, these joints do not exhibit internal friction that would be required for a *posable* model.

In contrast, our work aims for functional, posable models that do not require assembly of any kind, as this would take away some of the advantage of using rapid prototyping techniques. This is challenging, as joints need to provide friction while maintaining certain tolerances in the printing process. Furthermore, we strive to make the final printed models aesthetically pleasing. To this end, we propose an intuitive workflow that takes as input a static 3D model. The user rigs this model for 3D printing using a standard rigging procedure [Autodesk 2012]. We then craft and fit newly developed 3D-printable joints in accordance with the provided rigging while taking into consideration the rotational constraints as specified by

ACM Reference Format
Calì, J., Calian, D., Amati, C., Kleinberger, R., Steed, A., Kautz, J., Weyrich, T. 2012.
3D-Printing of Non-Assembly, Articulated Models. *ACM Trans. Graph. 31* 6, Article 130 (November 2012), 8 pages. DOI = 10.1145/2366145.2366149 http://doi.acm.org/10.1145/2366145.2366149.

Quality Prediction for Image Completion

Johannes Kopf, Microsoft Research
Wolf Kienzle, Microsoft Research
Steven Drucker, Microsoft Research
Sing Bing Kang, Microsoft Research

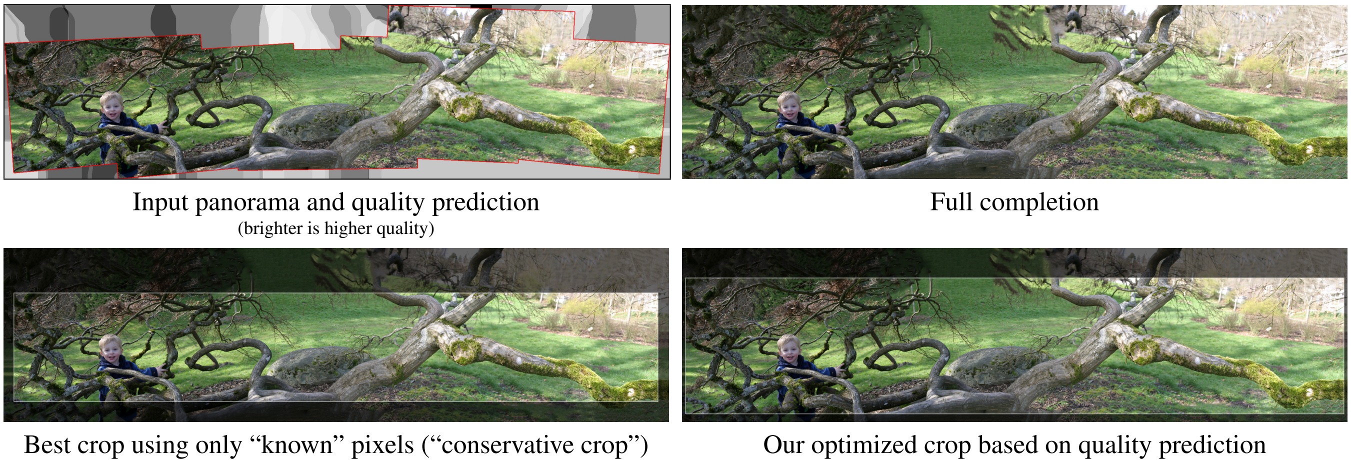


Figure 1: *Our data-driven technique is capable of predicting image completion quality (top left) before the completion is actually computed (top right). Based on our prediction, we compute an optimal crop rectangle that tries to include as many known pixels as possible while avoiding low-quality regions (bottom right). Compared to previous cropping approaches that do not fill in (bottom left) we can usually include a larger amount of the input image in our crop. Our algorithm* only *completes the cropped region, thus saving a significant amount of computation compared to full completion.*

Abstract

We present a data-driven method to predict the quality of an image completion method. Our method is based on the state-of-the-art non-parametric framework of Wexler *et al.* [2007]. It uses automatically derived search space constraints for patch source regions, which lead to improved texture synthesis and semantically more plausible results. These constraints also facilitate performance prediction by allowing us to correlate output quality against features of possible regions used for synthesis. We use our algorithm to first crop and then complete stitched panoramas. Our predictive ability is used to find an optimal crop shape *before* the completion is computed, potentially saving significant amounts of computation. Our optimized crop includes as much of the original panorama as possible while avoiding regions that can be less successfully filled in. Our predictor can also be applied for hole filling in the interior of images. In addition to extensive comparative results, we ran several user studies validating our predictive feature, good relative quality of our results against those of other state-of-the-art algorithms, and our automatic cropping algorithm.

CR Categories: I.4.9 [Image Processing and Computer Vision]: Applications;

Keywords: image completion, quality prediction, cropping

Links: DL PDF WEB

ACM Reference Format
Kopf, J., Kienzle, W., Drucker, S., Kang, S. 2012. Quality Prediction for Image Completion. *ACM Trans. Graph. 31* 6, Article 131 (November 2012), 8 pages. DOI = 10.1145/2366145.2366150 http://doi.acm.org/10.1145/2366145.2366150.

1 Introduction

Image completion or inpainting is a popular image editing tool for object removal and replacement or digital photograph restoration. A variety of completion algorithms have been developed in the scientific community over the past decade, and image completion is now a feature in commercial photo editing software such as Adobe Photoshop. In most previous work, image completion is used to fill holes after unwanted objects are removed. The same algorithms, however, can also be used to extend an image beyond its original boundaries. This is useful for filling beyond the irregular boundaries of a stitched panorama—this application is the focus of this paper.

Casually shot panoramas often have irregular boundaries (e.g., Figure 1). Most users, however, prefer output images with rectangular boundaries. The trivial solution implemented by most stitching software is to crop to the largest box that is fully contained within the panorama. This simple method often removes large parts of the input. The alternative is to apply any existing completion algorithm to fill the missing regions of the panorama bounding box.

Unfortunately, all existing image completion algorithms fail on occasion; the failure typically shows up as either inability to synthesize some textures well or results that are semantically implausible (see Figure 2). In addition, it is difficult to anticipate when and where such algorithms will fail given an arbitrary input image.

In this paper, we use machine learning to predict the quality of image completion. To support this prediction, we design our image completion algorithm to produce high quality results and to allow associations between completed pixels and known pixels to be created. We build on the existing non-parametric optimization framework of Wexler *et al.* [2007] (which is also implemented in the *Content Aware Fill* feature of Adobe Photoshop[1]). Previous work showed that this algorithm performs best if the source locations for patches are constrained to certain areas, e.g., see [Barnes et al. 2009]. We use a heuristic that automatically derives search space

[1] see http://www.adobe.com/technology/projects/content-aware-fill.html

Manifold Preserving Edit Propagation

Xiaowu Chen* Dongqing Zou* Qinping Zhao
State Key Laboratory of Virtual Reality Technology & Systems, Beihang University

Ping Tan
National University of Singapore

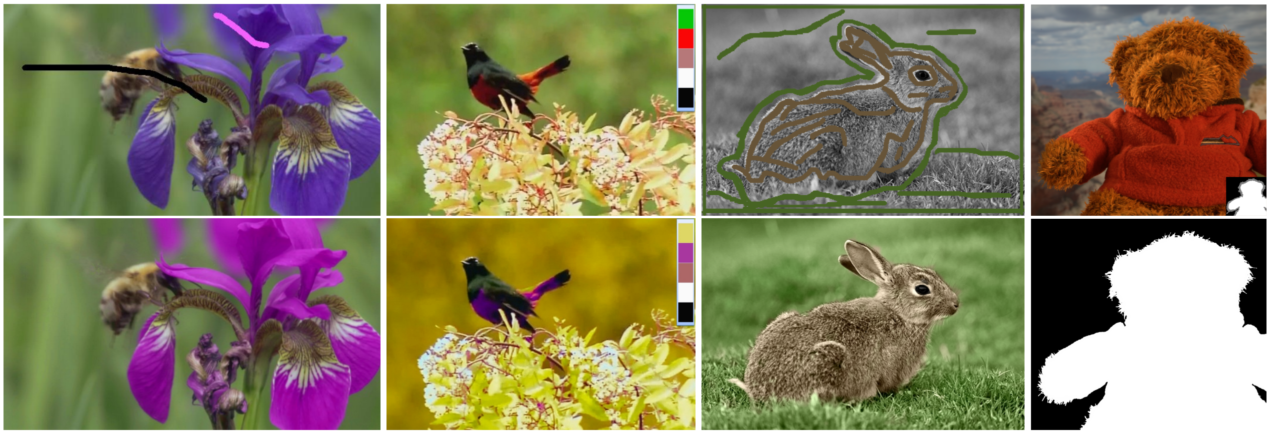


Figure 1: *Our edit propagation algorithm can be applied for various applications including color manipulation and matting. From left to right, these are the results of video recoloring, video color theme editing, image colorization and matting. Shown on the top are the input images (with a few user edits). Below are their corresponding output.*

Abstract

We propose a novel edit propagation algorithm for interactive image and video manipulations. Our approach uses the locally linear embedding (LLE) to represent each pixel as a linear combination of its neighbors in a feature space. While previous methods require similar pixels to have similar results, we seek to maintain the manifold structure formed by all pixels in the feature space. Specifically, we require each pixel to be the same linear combination of its neighbors in the result. Compared with previous methods, our proposed algorithm is more robust to color blending in the input data. Furthermore, since every pixel is only related to a few nearest neighbors, our algorithm easily achieves good runtime efficiency. We demonstrate our manifold preserving edit propagation on various applications.

CR Categories: I.4.9 [Image Processing and Computer Vision]: Applications—;

Keywords: Manifold Preserving, Edit Propagation, Recoloring, Colorization, Matting

Links: DL PDF

*corresponding authors: chen@buaa.edu.cn, zoudq@vrlab.buaa.edu.cn

ACM Reference Format
Chen, X., Zou, D., Zhao, Q., Tan, P. 2012. Manifold Preserving Edit Propagation. *ACM Trans. Graph. 31* 6, Article 132 (November 2012), 7 pages. DOI = 10.1145/2366145.2366151
http://doi.acm.org/10.1145/2366145.2366151.

1 Introduction

Efficiently propagating sparse user edits to a whole image (or video) is an important problem with many applications such as interactive color and tonal editing, colorization, matting or defogging, etc. Typically, this propagation favors nearby similar pixels to be edited in a similar way. It can be achieved by either global optimization [Levin et al. 2004; Lischinski et al. 2006; Pellacini and Lawrence 2007; An and Pellacini 2008; Xu et al. 2009] or smooth function interpolation [Li et al. 2010]. To evaluate the affinity between pixels, previous methods often rely on the Euclidean distance [An and Pellacini 2008; Li et al. 2010] or diffusion distance [Farbman et al. 2010] in some feature space.

We perform edit propagation from a different perspective. Instead of evaluating affinity for far apart pixel pairs, we seek to maintain the manifold structure formed by all pixels in a feature space. Specifically, we follow the Locally Linear Embedding (LLE) [Roweis and Saul 2000] [de Ridder and Duin 2002] to represent each pixel as a linear combination of a few of its nearest neighbors in a feature space. We seek to maintain this linear relationship between every pixel and its neighbors during edit propagation. Essentially, our method tries to preserve the manifold structure formed by all pixels as a whole, instead of looking into individual pixel-pairs.

Real images and videos contain objects with different colors. Pixels at object boundaries have blended color of neighboring objects. Semi-transparency, motion and unfocused blur also cause pixels with color blending. These pixels are dissimilar from those on main image objects, where the user tends to put edit scribbles. As a result, pixels with color blending are often less constrained according to the affinity-based edit propagation, which causes distracting artifacts in results. An example is shown in Figure 2, where the pixel A lies on an image edge and has blended color of the flower and the background. From the 3D color space, we can clearly see that the pixel A is dissimilar to the pixels covered by the user drawn scribbles. Hence, when affinity-based edit propagation methods such as [Xu et al. 2009] are applied to recolor the image, the pixel A will tend to keep its original red hue, which causes a distracting 'halo'

Sparse PDF Maps for Non-Linear Multi-Resolution Image Operations

Markus Hadwiger
KAUST

Ronell Sicat
KAUST

Johanna Beyer
KAUST

Jens Krüger
IVDA, DFKI, Intel VCI

Torsten Möller
Simon Fraser University

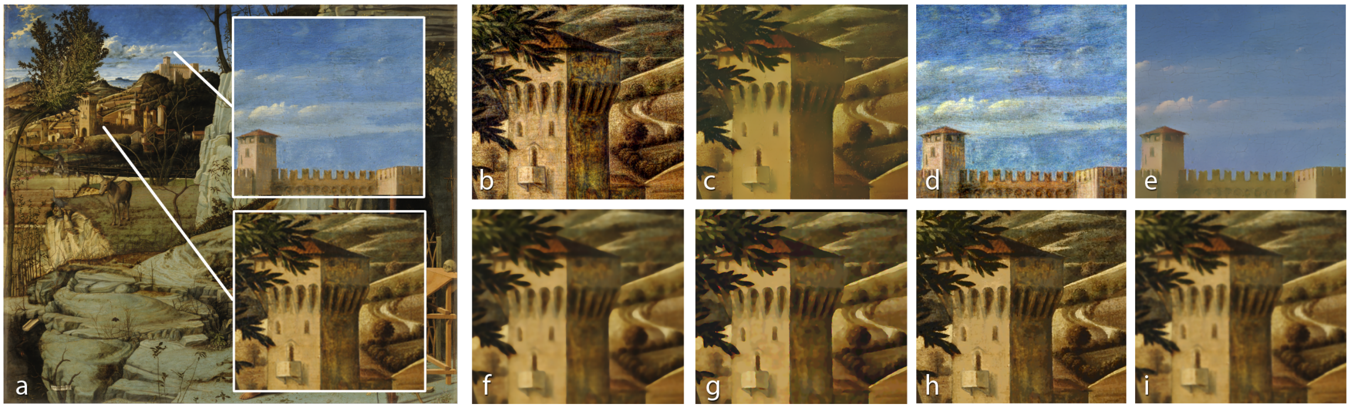


Figure 1: ***sPDF-maps** are a compact multi-resolution image pyramid data structure that sparsely encodes pre-computed pixel neighborhood probability density functions (pdfs) for all pixels in the pyramid. They enable the accurate, anti-aliased evaluation of non-linear image operators directly at any output resolution. A variety of operators can be computed at run time from the same pre-computed data structure in a way that scales to gigapixel images, such as local Laplacian filters for (b,d) detail enhancement or (c,e) smoothing, (f) median filters, (g) dominant mode filters, (h) maximum mode filters, (i) bilateral filters. The original image (a) has resolution* $16,898 \times 14,824$ *(*250 *Mpixels).*

Abstract

We introduce a new type of multi-resolution image pyramid for high-resolution images called *sparse pdf maps* (sPDF-maps). Each pyramid level consists of a sparse encoding of continuous probability density functions (pdfs) of pixel neighborhoods in the original image. The encoded pdfs enable the accurate computation of non-linear image operations directly in any pyramid level with proper pre-filtering for anti-aliasing, without accessing higher or lower resolutions. The sparsity of sPDF-maps makes them feasible for gigapixel images, while enabling direct evaluation of a variety of non-linear operators from the same representation. We illustrate this versatility for anti-aliased color mapping, $\mathcal{O}(n)$ local Laplacian filters, smoothed local histogram filters (e.g., median or mode filters), and bilateral filters.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture;

Keywords: image pyramids, mipmapping, anti-aliasing, multi-resolution filtering, display-aware filtering, smoothed local histogram filtering, bilateral filtering, local Laplacian filtering

Links: DL PDF

ACM Reference Format
Hadwiger, M., Sicat, R., Beyer, J., Krüger, J., Möller, T. 2012. Sparse PDF Maps for Non-Linear Multi-Resolution Image Operations. *ACM Trans. Graph. 31* 6, Article 133 (November 2012), 12 pages. DOI = 10.1145/2366145.2366152 http://doi.acm.org/10.1145/2366145.2366152.

1 Introduction

The recent increase in the resolution of acquired and computed image data has resulted in a need for novel multi-resolution techniques, e.g., representing, processing, and rendering gigapixel images [Kopf et al. 2007; Summa et al. 2010]. One of the most prevalent types of multi-resolution image hierarchies are pyramid representations, such as mipmaps [Williams 1983], or Gaussian pyramids [Burt and Adelson 1983]. Image pyramids store pre-filtered and downsampled versions of the original image, where the pre-filtering is crucial for avoiding aliasing. Because standard pre-filters are linear operators, further linear operators (e.g., further smoothing) can be applied accurately in any pre-computed pyramid level, as they commute with the pre-filter. However, this does not apply to non-linear operators, such as color mapping or edge-preserving filters. In this case, one has to process the original image and re-compute the pyramid. This is impractical for gigapixel images, especially when the goal is the interactive display of processed images at a lower output resolution, which is becoming more important as the gap between the acquired image and display resolutions widens. Therefore, in such a scenario all operators are usually applied directly at the output resolution. Nevertheless, downsampling the image first using a standard pre-filter, e.g., bicubic, followed by a non-linear operation, can introduce false colors. These artifacts due to resampling are a form of aliasing, though not of the same kind as the well-known moiré patterns. A naive solution would be to remove the pre-filter and resort to nearest-neighbor downsampling, but then moiré patterns would appear.

This paper introduces *sparse pdf maps* (sPDF-maps), which are probabilistic image pyramids that sparsely encode probability density functions (pdfs) of local pixel neighborhoods. Our main goals are (1) the accurate evaluation of non-linear operators at any output resolution, and (2) scalability to gigapixel images. The sPDF-map representation is computed in a pre-computation stage. However, at run time different operators with arbitrary parameter settings can be evaluated interactively. sPDF-maps enable the accurate evaluation of non-linear operators directly at any resolution without accessing

DressUp! Outfit Synthesis Through Automatic Optimization

Lap-Fai Yu[1] * Sai-Kit Yeung[2] Demetri Terzopoulos[1] Tony F. Chan[3]
[1]University of California, Los Angeles
[2]Singapore University of Technology and Design [3]Hong Kong University of Science and Technology

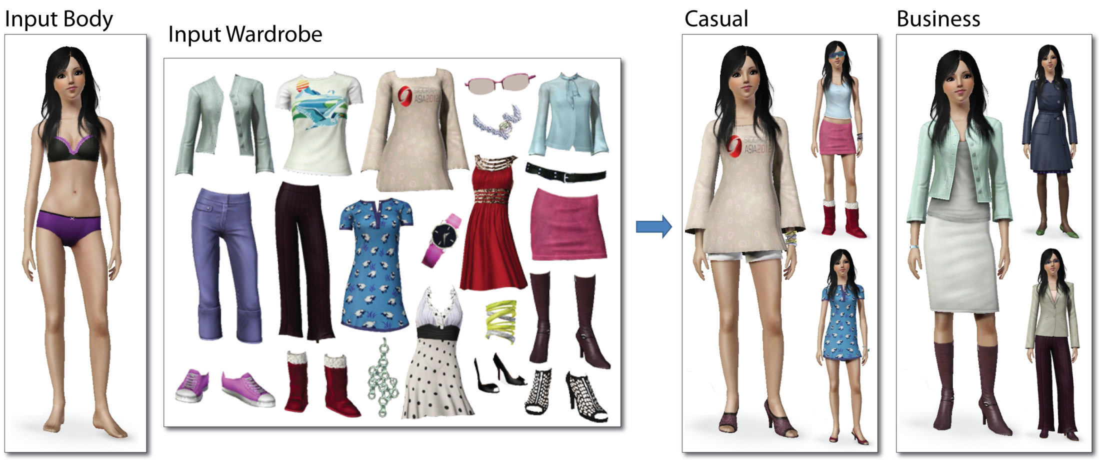


Figure 1: *Outfit optimization with different input dress codes. Left: An input human body with hair color, eye color, and skin color specified, plus a wardrobe of clothing items. Right: Optimized outfits for dress code* Casual *and* Business*.*

Abstract

We present an automatic optimization approach to outfit synthesis. Given the hair color, eye color, and skin color of the input body, plus a wardrobe of clothing items, our outfit synthesis system suggests a set of outfits subject to a particular dress code. We introduce a probabilistic framework for modeling and applying dress codes that exploits a Bayesian network trained on example images of real-world outfits. Suitable outfits are then obtained by optimizing a cost function that guides the selection of clothing items to maximize the color compatibility and dress code suitability. We demonstrate our approach on the four most common dress codes: *Casual*, *Sportswear*, *Business-Casual*, and *Business*. A perceptual study validated on multiple resultant outfits demonstrates the efficacy of our framework.

CR Categories: I.3.7 [Computing Methodologies]: Computer Graphics—Three-Dimensional Graphics and Realism;

Keywords: Clothing combination, perception, variety, functionally realistic, virtual world modeling, fashion, color matching, optimization, procedural modeling

Links: DL PDF

*Part of the work was done when Lap-Fai was visiting SUTD.

ACM Reference Format
Yu, L., Yeung, S., Terzopoulos, D., Chan, T. 2012. DressUp! Outfit Synthesis Through Automatic Optimization. *ACM Trans. Graph. 31* 6, Article 134 (November 2012), 14 pages. DOI = 10.1145/2366145.2366153 http://doi.acm.org/10.1145/2366145.2366153.

1 Introduction

As you awaken each day, there is a simple question that you may need to answer: How should I dress today? Your wardrobe contains various kinds of clothes, such as dress shirts, dress pants, jeans, sweaters, suits, and different types of shoes. What combination of clothing will have you most appropriately dressed for the day's activities, thereby making you most visually appealing? Perhaps you would like multiple suggestions that best coordinate with the new tie that you received from your daughter as a birthday gift. The outfit selection problem also occurs in computer graphics modeling, especially in movie and game production: How should one appropriately dress a large number of human characters with an eye to functionality while avoiding visual awkwardness and repetitiveness? The manual specification of clothing is obviously tedious and it may be prohibitive on a large scale.

We demonstrate that the task at hand, of selecting appropriate subsets of clothing items from a wardrobe, can be addressed formally as a combinatorial optimization problem. A suitable outfit requires jointly combining a variety of clothing items to satisfy functional and certain visual criteria. We do not generally wear a pair of sandals with dress pants to the office, nor do we wear a red dress shirt with a green suit for a business meeting. In addition, to put a wardrobe into full use, we would like to explore as many good solutions as possible, so that we can exhibit sartorial variety. A similar, but much larger-scale problem comes up with regard to online boutique websites, where shoppers can select among many clothing items. Usually it is not difficult for a shopper to locate a desired clothing item; the nontrivial question is how this clothing item should be matched in terms of style and color with other clothing items from the same or different shops or from one's wardrobe at home.

There is no single universal rule that satisfies both the relevant *functional* and *visual* criteria. People generally categorize outfits into *dress codes*, which represent different functionalities. These can

Example-based Synthesis of 3D Object Arrangements

Matthew Fisher* Stanford University
Daniel Ritchie* Stanford University
Manolis Savva* Stanford University
Thomas Funkhouser† Princeton University
Pat Hanrahan* Stanford University

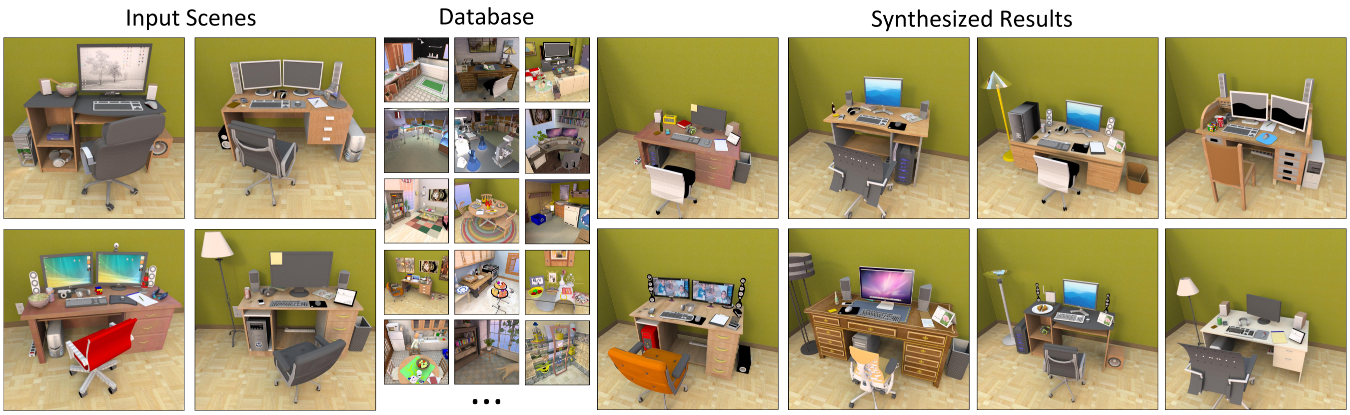


Figure 1: *Example-based scene synthesis. Left: four computer desks modeled by hand and used as input to our algorithm. Right: scenes synthesized using our algorithm. Our algorithm generates plausible results from few examples, incorporating object composition and arrangement information from a database of 3D scenes to increase the variety of its results.*

Abstract

We present a method for synthesizing 3D object arrangements from examples. Given a few user-provided examples, our system can synthesize a diverse set of plausible new scenes by learning from a larger scene database. We rely on three novel contributions. First, we introduce a *probabilistic model for scenes* based on Bayesian networks and Gaussian mixtures that can be trained from a small number of input examples. Second, we develop a clustering algorithm that groups objects occurring in a database of scenes according to their local scene neighborhoods. These *contextual categories* allow the synthesis process to treat a wider variety of objects as interchangeable. Third, we train our probabilistic model on a mix of user-provided examples and relevant scenes retrieved from the database. This *mixed model* learning process can be controlled to introduce additional variety into the synthesized scenes. We evaluate our algorithm through qualitative results and a perceptual study in which participants judged synthesized scenes to be highly plausible, as compared to hand-created scenes.

Keywords: 3D scenes, procedural modeling, automatic layout, probabilistic modeling, data-driven methods

Links: DL PDF WEB DATA

*e-mail: {mdfisher, dritchie, msavva, hanrahan}@stanford.edu
†e-mail: funk@cs.princeton.edu

ACM Reference Format
Fisher, M., Ritchie, D., Savva, M., Funkhouser, T., Hanrahan, P. 2012. Example-based Synthesis of 3D Object Arrangements. *ACM Trans. Graph. 31* 6, Article 135 (November 2012), 11 pages.
DOI = 10.1145/2366145.2366154 http://doi.acm.org/10.1145/2366145.2366154.

1 Introduction

Large-scale games and virtual worlds demand detailed 3D environments. Creating this content requires a lot of time from many artists. In this paper, we focus on scenes: environments composed of arrangements of 3D objects. They include many typical environments in games, animation, and virtual worlds, including most indoor and human-made environments.

A promising solution to this content-creation bottleneck is example-based synthesis: algorithms that can generate new environments similar to a set of input examples. A user provides the system with examples illustrating the desired type of scene, and the system returns a large set of synthesized results. The user then selects the results that she likes best from a ranked list. The user should not have to browse for a long time to find good scenes; as a rule of thumb, at least one out of every three synthesized results should be usable.

To support a tool such as the one described above, there are several criteria an example-based synthesis algorithm should meet. First, it should generate plausible scenes; they should look believable to a casual observer. Second, it should generate a variety of scenes; they should not be copies or small perturbations of the examples. Third, users should only have to provide a few examples, since they are time-consuming to create.

These goals are challenging to meet, and some stand in conflict with one another. Generating a variety of scenes is difficult when the system can only draw data from a few examples. Scenes can contain a large number of different objects, the full range of which can be difficult to specify in a few examples. Just because the user omitted some type of object in the examples, does that mean she does not want it in her scenes? Some objects are connected via precise functional and geometric relationships, while others are more loosely coupled. To generate plausible scenes, an example-based algorithm must infer these relationships from data without additional user guidance.

In this paper, we present an example-based scene synthesis method that meets the above challenges through three main contributions.

An Interactive Approach to Semantic Modeling of Indoor Scenes with an RGBD Camera

Tianjia Shao [*] Weiwei Xu [† §] Kun Zhou [‡] Jingdong Wang [§] Dongping Li[‡] Baining Guo [§]
Tsinghua University [*] Hangzhou Normal University [†] Zhejiang University[‡] Microsoft Research Asia[§]

(a) (b)

Figure 1: *Lab scene. (a) Captured images (Only RGB data are shown). (b) Reconstruction result. Only 6 RGBD images captured by Microsoft Kinect camera are enough for our system to reconstruct its prototype scene with 20 objects of semantic labels (eight monitors are not numbered to avoid possible clutter). The numbers in white indicate the correspondence between objects in the image and their reconstruction results and the overall modeling time of this lab scene is less than 18 minutes.*

Abstract

We present an interactive approach to semantic modeling of indoor scenes with a consumer-level RGBD camera. Using our approach, the user first takes an RGBD image of an indoor scene, which is automatically segmented into a set of regions with semantic labels. If the segmentation is not satisfactory, the user can draw some strokes to guide the algorithm to achieve better results. After the segmentation is finished, the depth data of each semantic region is used to retrieve a matching 3D model from a database. Each model is then transformed according to the image depth to yield the scene. For large scenes where a single image can only cover one part of the scene, the user can take multiple images to construct other parts of the scene. The 3D models built for all images are then transformed and unified into a complete scene. We demonstrate the efficiency and robustness of our approach by modeling several real-world scenes.

Keywords: Indoor scene, Depth images, Segmentation, Labeling, Random Regress Forest

Links: DL PDF

[†]email:weiwei.xu.g@gmail.com, corresponding author

[‡]Kun Zhou and Dongping Li are with the State Key Lab of CAD&CG of Zhejiang university.

ACM Reference Format
Shao, T., Xu, W., Zhou, K., Wang, J., Li, D., Guo, B. 2012. An Interactive Approach to Semantic Modeling of Indoor Scenes with an RGBD Camera. *ACM Trans. Graph. 31* 6, Article 136 (November 2012), 11 pages. DOI = 10.1145/2366145.2366155 http://doi.acm.org/10.1145/2366145.2366155.

1 Introduction

With the popularization of commercial RGBD cameras such as Microsoft's Kinect, now everyone can have a low-cost and easy-to-use device to digitalize 3D objects at home [Clark 2011]. It is naturally of great interest to use RGBD cameras to scan and build 3D scenes which closely match the indoor environments where people live or work. Such 3D scenes can be used in applications for many purposes, such as providing an interface for user interaction [Izadi et al. 2011] and rearranging furniture for interior design [Yu et al. 2011; Merrell et al. 2011].

Several techniques have been proposed recently for modeling indoor environments with a depth camera, representing the scene geometry either as point clouds [Du et al. 2011] or signed distance fields defined over a 3D volume grid [Izadi et al. 2011]. While such geometry representations meet application requirements to a certain extent, they do not provide the necessary flexibility required by many other applications due to the lack of semantics. For example, in the context of rearranging furniture, the input scene to state-of-the-art algorithms consists of a set of independent, semantic objects, i.e., each scene object can be manipulated independently and belongs to a category (e.g., sofa, chair, bed, etc.). The algorithms need such semantics to calculate an optimal layout for all objects and transform each object to its target position. As far as we know, none of the existing modeling techniques aim to generate 3D semantic models for indoor environments.

In this paper we present an interactive approach to semantic modeling of indoor scenes with a consumer-level RGBD camera. Using our approach, the user first takes an RGBD image of an indoor scene, which is automatically segmented into a set of regions with semantic labels. If the segmentation is not satisfactory, the user can draw some strokes to guide the algorithm to achieve better results. After the segmentation is finished, the depth data of each semantic region is used to retrieve a matching 3D model from a database. Each model is then transformed according to the image depth to

A *Search-Classify* Approach for Cluttered Indoor Scene Understanding

Liangliang Nan[1] Ke Xie[1] Andrei Sharf[2]

[1] Shenzhen VisuCA Key Lab/SIAT [2] Ben Gurion University

Figure 1: *A raw scan of a highly cluttered indoor scene is given (left). Applying our* search-classify *method, we segment the scene into meaningful objects (middle: chairs (blue) and tables (purple)), followed by a template deform-to-fit reconstruction (right).*

Abstract

We present an algorithm for recognition and reconstruction of scanned 3D indoor scenes. 3D indoor reconstruction is particularly challenging due to object interferences, occlusions and overlapping which yield incomplete yet very complex scene arrangements. Since it is hard to assemble scanned segments into complete models, traditional methods for object recognition and reconstruction would be inefficient. We present a *search-classify* approach which interleaves segmentation and classification in an iterative manner. Using a robust classifier we traverse the scene and gradually propagate classification information. We reinforce classification by a template fitting step which yields a scene reconstruction. We *deform-to-fit* templates to classified objects to resolve classification ambiguities. The resulting reconstruction is an approximation which captures the general scene arrangement. Our results demonstrate successful classification and reconstruction of cluttered indoor scenes, captured in just few minutes.

Keywords: point cloud classification, scene understanding, reconstruction

Links: DL PDF

1 Introduction

Processing of 3D digital environments is an increasingly important research problem, motivated by ambitious applications that aim to build digital copies of cities (e.g., Microsoft Virtual Earth and Google Earth). Advances in laser scanning technology and recent proliferation of GIS services have been driving a strong trend towards processing and modeling of large-scale outdoor scenes based on aerial photography and street-level laser scanners.

Surprisingly, scanned 3D interiors pose a much more difficult problem. In contrast to building exteriors which are relatively piecewise flat, interior scenes are more complicated in their 3D structures. Rooms are densely populated with objects in arbitrary arrangements, for example, chairs pulled underneath tables, semi-open drawers, and etc. Additionally, proper acquisition of interior spaces is challenging in many ways. For example, when scanning densely populated indoor scenes, significant parts remain occluded in all views, yielding a partial representation. The prevalence of thin structures such as doors, walls and table legs poses another significant challenge since their acquisition requires high sampling rate relative to the scale of the scene. Hence, traditional modeling techniques would perform relatively poor on interior scenes, due to typical clutter, missing parts and noise (see Figure 1).

3D scans of large scale environments are relatively new and were made possible due to recent progress in scanning technology. Several algorithms have been proposed for modeling [Sinha et al. 2008; Schnabel et al. 2009; Nan et al. 2010; Livny et al. 2010], and object segmentation in scanned outdoor scenes [Frome et al. 2004b; Anguelov et al. 2005; Golovinskiy et al. 2009]. However, only very recently, researchers have been addressing the complex challenges present in cluttered scanned indoor scenes [Kim et al. 2012; Shao et al. 2012].

We develop a fully automatic algorithm that is capable of understanding and modeling raw scans of cluttered indoor scenes (see Figures 1, 2). In our method, we define a set of shape features that are used for supervised learning of a classifier. We argue that object classification cannot be directly applied to the scene, since object segmentation is unavailable. Moreover, the segmentation of the scene into objects is as challenging as the classification since spatial relationships between points and patches are neither complete nor reliable. For example, it is practically impossible to segment and attribute legs of chairs drawn close together. Thus, classification and segmentation together, constitute a typical *chicken-egg* problem.

Our key idea is to interleave the computations of segmentation and classification of the scene into meaningful parts. We denote this approach *search-classify*, since we search for meaningful segments using a classifier that estimates the probability of a segment to be

ACM Reference Format
Nan, L., Xie, K., Sharf, A. 2012. A *Search-Classify* Approach for Cluttered Indoor Scene Understanding. *ACM Trans. Graph. 31* 6, Article 137 (November 2012), 10 pages. DOI = 10.1145/2366145.2366156 http://doi.acm.org/10.1145/2366145.2366156.

Acquiring 3D Indoor Environments with Variability and Repetition

Young Min Kim
Stanford University

Niloy J. Mitra
Univ. College London / KAUST

Dong-Ming Yan
KAUST

Leonidas Guibas
Stanford University

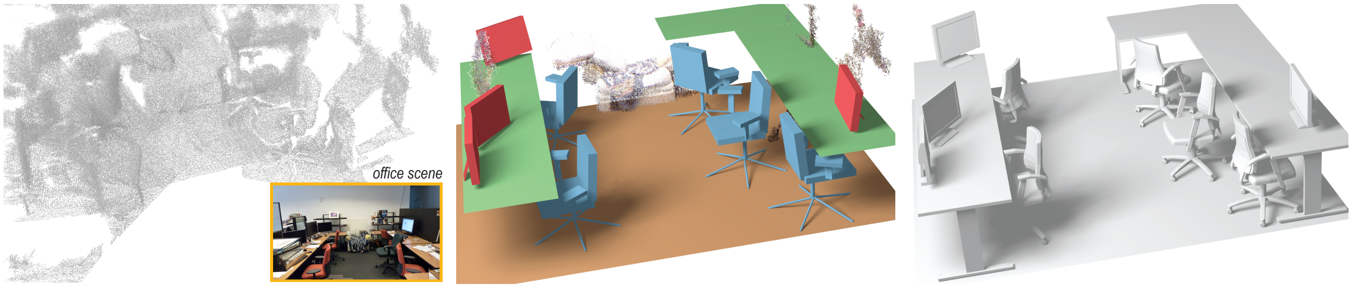


Figure 1: *(Left) Given a single view scan of a 3D environment obtained using a fast range scanner, we perform scene understanding by recognizing repeated objects, while factoring out their modes of variability (middle). The repeating objects have been learned beforehand as low complexity models, along with their joint deformations. We extract the objects despite a poor quality input scan with large missing parts and many outliers. The extracted parameters can then be used to pose 3D models to create a plausible scene reconstruction (right).*

Abstract

Large-scale acquisition of exterior urban environments is by now a well-established technology, supporting many applications in search, navigation, and commerce. The same is, however, not the case for indoor environments, where access is often restricted and the spaces are cluttered. Further, such environments typically contain a high density of repeated objects (e.g., tables, chairs, monitors, etc.) in regular or non-regular arrangements with significant pose variations and articulations. In this paper, we exploit the special structure of indoor environments to accelerate their 3D acquisition and recognition with a low-end handheld scanner. Our approach runs in two phases: (i) a learning phase wherein we acquire 3D models of frequently occurring objects and capture their variability modes from only a few scans, and (ii) a recognition phase wherein from a single scan of a new area, we identify previously seen objects but in different poses and locations at an average recognition time of 200ms/model. We evaluate the robustness and limits of the proposed recognition system using a range of synthetic and real world scans under challenging settings.

Keywords: acquisition, scene understanding, shape analysis, real-time modeling

Links: DL PDF WEB DATA

1 Introduction

Significant advances have been made towards mapping the exteriors of urban environments through large-scale acquisition efforts of Google, Microsoft, Nokia, etc. Capturing 3D indoor environments, however, remains challenging. While sensor-instrumented vehicles can drive down streets to capture exterior spaces, mimicking similar setups for interior acquisition requires customization and manual intervention, and is cumbersome due to unreliable GPS signals, odometry errors, etc. Additionally, unlike building exteriors whose facades are largely flat and have ample clearance for scanning, indoor objects are usually geometrically complex, found in cramped surroundings, and contain significant variations: doors and windows are opened and closed, chairs get moved around, cubicles get rearranged, etc.

The growing popularity of fast, easy-to-use, affordable range cameras (e.g., the Microsoft Kinect) presents new acquisition possibilities. High frame-rate and increased portability, however, come at the cost of resolution and data quality: the scans are at best of modest resolution, often noisy, invariably contain outliers, and suffer from missing parts due to occlusion (see Figure 1). Thus, a traditional single-scan acquisition pipeline is ill-suited: typically, one has to scan the scene multiple times from various viewpoints, semi-automatically align the scans, and finally construct a 3D model, which is often of unsatisfactory quality and provides little understanding of the indoor environment. The process is further complicated when the model deforms between successive acquisitions.

In this paper we focus on interiors of public buildings (e.g. schools, hospitals, hotels, restaurants, airports, train stations) or office buildings. We exploit three observations to make the problem of indoor 3D acquisition tractable: (i) Most such building interiors comprise of basic elements such as walls, doors, windows, furniture (e.g., chairs, tables, lamps, computers, cabinets), which come from a small number of prototypes and repeat many times. (ii) Such building components usually consist of rigid parts of simple geometry, i.e., they have surfaces that are well approximated by planar, cylindrical, conical, spherical proxies. Further, although variability and articulation are dominant (e.g., a chair is moved or rotated, a lamp is folded), such variability is limited and low-dimensional (e.g., translational motion, hinge joint, telescopic joint). (iii) Mutual relationships among the basic objects satisfy strong priors (e.g., a chair stands on the floor, a monitor rests on the table).

We present a simple yet practical pipeline to acquire models of indoor objects such as furniture, together with their variability modes, and discover object repetitions and exploit them to speed up large-

ACM Reference Format
Kim, Y., Mitra, N., Yan, D., Guibas, L. 2012. Acquiring 3D Indoor Environments with Variability and Repetition. *ACM Trans. Graph. 31* 6, Article 138 (November 2012), 11 pages. DOI = 10.1145/2366145.2366157 http://doi.acm.org/10.1145/2366145.2366157.

Structure Extraction from Texture via Relative Total Variation

Li Xu Qiong Yan Yang Xia Jiaya Jia*

Department of Computer Science and Engineering
The Chinese University of Hong Kong

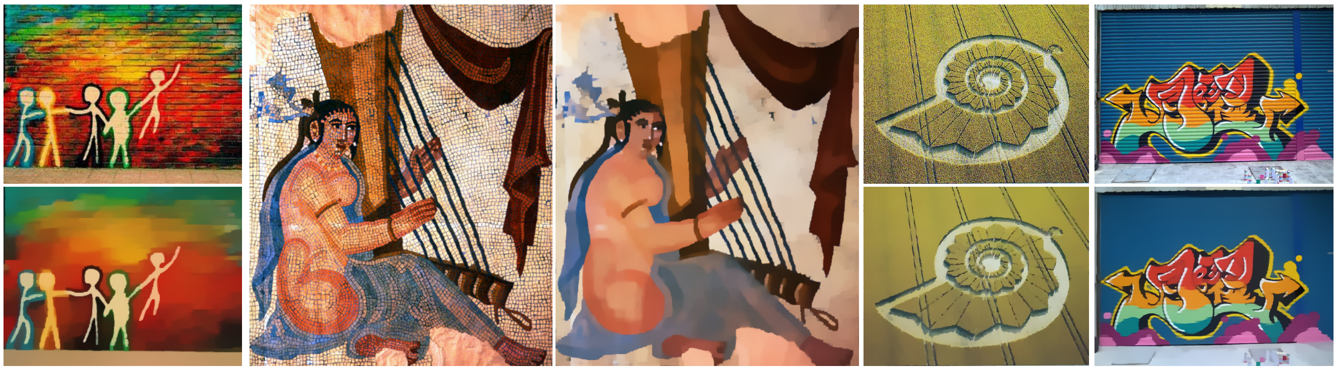


Figure 1: *Meaningful structure extraction from textured surfaces. Examples from left to right are graffiti on brick, marble mosaic (ca. 260 AD), crop circles, and graffiti on gate.*

Abstract

It is ubiquitous that meaningful structures are formed by or appear over textured surfaces. Extracting them under the complication of texture patterns, which could be regular, near-regular, or irregular, is very challenging, but of great practical importance. We propose new inherent variation and relative total variation measures, which capture the essential difference of these two types of visual forms, and develop an efficient optimization system to extract main structures. The new variation measures are validated on millions of sample patches. Our approach finds a number of new applications to manipulate, render, and reuse the immense number of "structure with texture" images and drawings that were traditionally difficult to be edited properly.

CR Categories: I.4.3 [Image Processing and Computer Vision]: Enhancement—Smoothing; G.1.6 [Numerical Analysis]: Optimization—Nonlinear programming

Keywords: texture, structure, smoothing, total variation, relative total variation, inherent variation, prior, regularized optimization

Links: DL PDF WEB CODE

1 Introduction

Many natural scenes and human-created art pieces contain texture. For instance, graffiti and drawings can be commonly seen on brick walls, railroad boxcars, and subways; carpets, sweaters, and other fine crafts contain various geometric patterns. In human history, mosaic has long been be an art form to represent detailed scenes of people and animals, and imitate paintings using stone, glass, ceramic, and other materials. When searching in Google Images, millions of such pictures and drawings can be found quickly.

A few examples from different sources are shown in Figure 1. They share the similarity that semantically meaningful structures are blended with or formed by texture elements. We call them "structure+texture" images. It is particularly interesting that human visual system is fully capable to understand these pictures without needing to remove textures. In psychology [Arnheim 1956], it is also found that "the overall structural features are the primary data of human perception, not the individual details".

Contrary to this almost effortless process, extract structures by a computer is much more challenging. Tedious manual manipulation is needed in all photo editing software that we used. A few approaches [Meyer 2001; Yin et al. 2005; Aujol et al. 2006] employ a total variation image regularizer in optimization. This framework, however, cannot satisfyingly distinguish texture from the main structures because both of them could receive similar penalties during optimization. Recent edge-preserving image editing tools [Farbman et al. 2008; Subr et al. 2009; Farbman et al. 2010; Kass and Solomon 2010; Paris et al. 2011; Xu et al. 2011] do not aim to solve the same problem, and, therefore, are not optimal solutions. More analysis and comparisons will be provided.

We present a simple and yet effective method based on novel local variation measures to accomplish texture removal. We found that with regard to our new *relative total variation*, which will be elaborated later in this paper, texture and main structure exhibit completely different properties, making them surprisingly decomposable. With this finding, we present an optimization framework, in which meaningful content and textural edges are penalized differently. A robust numerical solver is also proposed to decompose the original highly non-convex optimization problem into several linear systems, for which fast and robust solution exists. Note that we do *not* assume specific regularity or symmetry of the texture patterns, and instead allow for a high level of randomness. Non-uniform and anisotropic texture, thus, can be handled in a unified framework.

*e-mail: {xuli, qyan, yxia, leojia}@cse.cuhk.edu.hk

ACM Reference Format
Xu, L., Yan, Q., Xia, Y., Jia, J. 2012. Structure Extraction from Texture via Relative Total Variation. *ACM Trans. Graph. 31* 6, Article 139 (November 2012), 10 pages. DOI = 10.1145/2366145.2366158 http://doi.acm.org/10.1145/2366145.2366158.

Digital Reconstruction of Halftoned Color Comics

Johannes Kopf
Microsoft Research

Dani Lischinski
Hebrew University

Screened Input | Accurate Reconstruction | Painterly Rendering

Figure 1: *We reconstruct an accurate digital representation from a halftoned color print scanned at high resolution. Our result may be viewed at significant magnification, compresses well, and lends itself nicely to painterly re-rendering (shown on the right), as well as advanced editing, such as decomposition into layers for a tour-into-the-picture application (see supplementary video). Input image © Dupuis.*

Abstract

We introduce a method for automated conversion of scanned color comic books and graphical novels into a new high-fidelity rescalable digital representation. Since crisp black line artwork and lettering are the most important structural and stylistic elements in this important genre of color illustrations, our digitization process is geared towards faithful reconstruction of these elements. This is a challenging task, because commercial presses perform halftoning (screening) to approximate continuous tones and colors with overlapping grids of dots. Although a large number of inverse haftoning (descreening) methods exist, they typically blur the intricate black artwork. Our approach is specifically designed to descreen color comics, which typically reproduce color using screened CMY inks, but print the black artwork using non-screened solid black ink. After separating the scanned image into three screening grids, one for each of the CMY process inks, we use non-linear optimization to fit a parametric model describing each grid, and simultaneously recover the non-screened black ink layer, which is then vectorized. The result of this process is a high quality, compact, and rescalable digital representation of the original artwork.

Keywords: comics, descreening, inverse halftoning, vectorization, non-photorealistic rendering

Links: DL PDF WEB

1 Introduction

Emerging from newspaper comic strips in the early 20th century, comic books and graphic novels gradually evolved into a mass medium, and have become a major part of popular culture. Today, these books are purchased, collected, and read by millions of people around the globe. In fact, graphic novels are one of the last varieties of the printed form that are still gaining in popularity [Weiner 2010].

Despite the advent of electronic publishing and the increasing popularity of e-books, facilitated by mobile reading devices such as Amazon's Kindle and Apple's iPad, most classical comic books are still only available in their original printed form. Although some comic book publishers offer a limited amount of classical material as "special digital editions", it is apparent that the creation of these digital editions involved intensive manual intervention. This is not surprising, since as discussed below, simply scanning such books does not do justice to this medium, which often features intricate and detailed line artwork and masterful coloring.

Our goal in this work is to provide a new tool for automated conversion of scanned color comic books, as well as some other types of hand-drawn color illustrations, into a new, high-fidelity rescalable digital representation, suitable for today's high resolution digital reading devices. Black line artwork and lettering are the most important structural and stylistic element of comic art: black lines are used to draw outlines, boundaries of panels and speech bubbles, and as hatching strokes to convey shading and texture. Therefore, both our digitization process and the resulting representation strive to faithfully reconstruct the original line artwork and lettering, as well as to allow significant magnification, enabling full-screen viewing of individual panels on devices such as the 2048×1536 iPad.

The main challenge that we face is rooted in the printing process used to produce color comics. Most printing technologies are unable to produce continuous color or grayscale tones, resorting instead to *halftoning*: approximating continuous tones using various dot patterns. Color comic books, as well as virtually all books, newspapers, and magazines, are printed with AM halftoning, which attempts to simulate changes in tone by varying the size of the dots on a regular grid (*clustered dot screening*). Scanning such prints inevitably captures the halftone screen patterns. Viewing such scans on raster displays results in Moiré patterns due to the interaction between the halftone patterns and the display raster. The effectiveness of various image processing operations is also hindered by the presence of the halftone screen patterns. Thus, it is necessary to invert the halftoning process, an operation also known as *descreening*.

ACM Reference Format
Kopf, J., Lischinski, D. 2012. Digital Reconstruction of Halftoned Color Comics. *ACM Trans. Graph. 31* 6, Article 140 (November 2012), 10 pages. DOI = 10.1145/2366145.2366159
http://doi.acm.org/10.1145/2366145.2366159.

Automatic Stylistic Manga Layout

Ying Cao * Antoni B. Chan Rynson W.H. Lau

Department of Computer Science, City University of Kong Kong

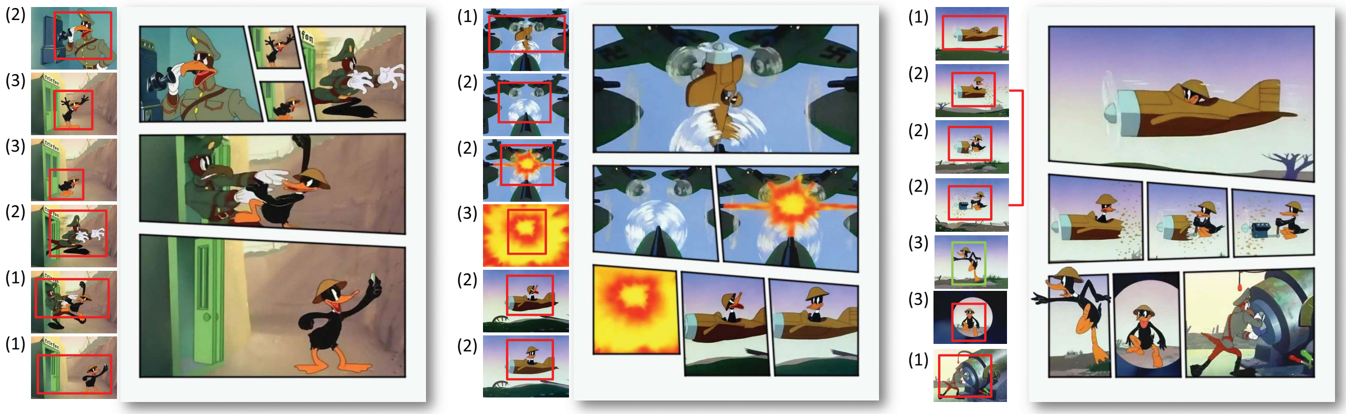


Figure 1: *Three layouts generated by our approach trained on "Fairy Tail". The left side of each example displays the sequence of input artworks ("Daffy: The Commando"(1943) in the public domain), single-panel semantics, including importance-ranking values (within the parenthesis and region of interest (masked by rectangle), as well as optional inter-panel semantics that describe a group of consecutive semantically related panels (grouped by a red line in the rightmost example). The character masked by the green rectangle is chosen for a"fourth wall break" effect. The reading order of each layout is from left to right and then top to bottom.*

Abstract

Manga layout is a core component in manga production, characterized by its unique styles. However, stylistic manga layouts are difficult for novices to produce as it requires hands-on experience and domain knowledge. In this paper, we propose an approach to automatically generate a stylistic manga layout from a set of input artworks with user-specified semantics, thus allowing less-experienced users to create high-quality manga layouts with minimal efforts. We first introduce three parametric style models that encode the unique stylistic aspects of manga layouts, including layout structure, panel importance, and panel shape. Next, we propose a two-stage approach to generate a manga layout: 1) an initial layout is created that best fits the input artworks and layout structure model, according to a generative probabilistic framework; 2) the layout and artwork geometries are jointly refined using an efficient optimization procedure, resulting in a professional-looking manga layout. Through a user study, we demonstrate that our approach enables novice users to easily and quickly produce higher-quality layouts that exhibit realistic manga styles, when compared to a commercially-available manual layout tool.

Keywords: manga, stylistic layout, generative probabilistic model

Links: DL PDF

*yingcao4@student.cityu.edu.hk

ACM Reference Format
Cao, Y., Chan, A., Lau, R. 2012. Automatic Stylistic Manga Layout.
ACM Trans. Graph. 31 6, Article 141 (November 2012), 10 pages. DOI = 10.1145/2366145.2366160
http://doi.acm.org/10.1145/2366145.2366160.

1 Introduction

During the last few decades, manga or Japanese comics have gained in popularity across the world due to its unique styles in terms of screening, ballooning and layout. This has resulted in an increasing involvement of the general population in producing their own manga. However, creating high-quality manga typically requires well-trained skill and unique talent. Although various manga production tools (e.g., Manga Studio [MangaStudio 2011]) have emerged to help novices create manga from scratch, none of these tools provide the insight into how manga layout can be done effectively. Manga layout is at the core of manga production. Effective manga layout is utilized by the artists to help storytelling, guide the reader's attention, and enhance the visual attractiveness of manga pages [Rivkah 2006]. However, manga layout is a complicated task; both in-depth understanding and hands-on experience are required to achieve an effective manga layout [Tsai 2002].

In addition to some basic requirements, such as correct reading order and efficient space utilization, Japanese manga artists typically stylize their layout by introducing some customized features, such as: 1) different layout structures (i.e., spatial arrangement of the panels), rather than a single layout template, which augment the visual richness; 2) variations in panel size, with larger panels for important events and smaller panels for scene or moment transitions, which increase the semantic contrast between the panels and make the storytelling more dynamic; 3) irregular panel shapes, instead of uniform rectangles, which make the contents in the panel more dynamic and engaging. These factors characterize manga layout, distinguishing it from the layout of traditional Western comics, which is more rigid and grid-based (e.g., see Figure 2).

Manga artists rely upon their intuition and experience for layout design, which is hard to completely formulate. Hence, previous layout algorithms, based on either heuristic rules [Taniguchi et al. 1997; Thawonmas and Shuda 2008] or energy optimization [Purvis et al. 2003; Cagan et al. 2002; Geigel and Loui 2003; Merrell et al. 2011; Yu et al. 2011], are not applicable, since there are no clear guidelines that can be incorporated into these frameworks. In contrast

Lazy Selection: A Scribble-based Tool for Smart Shape Elements Selection

Pengfei Xu[1] Hongbo Fu[2] Oscar Kin-Chung Au[2] Chiew-Lan Tai[1]

[1]Hong Kong University of Science and Technology [2]City University of Hong Kong

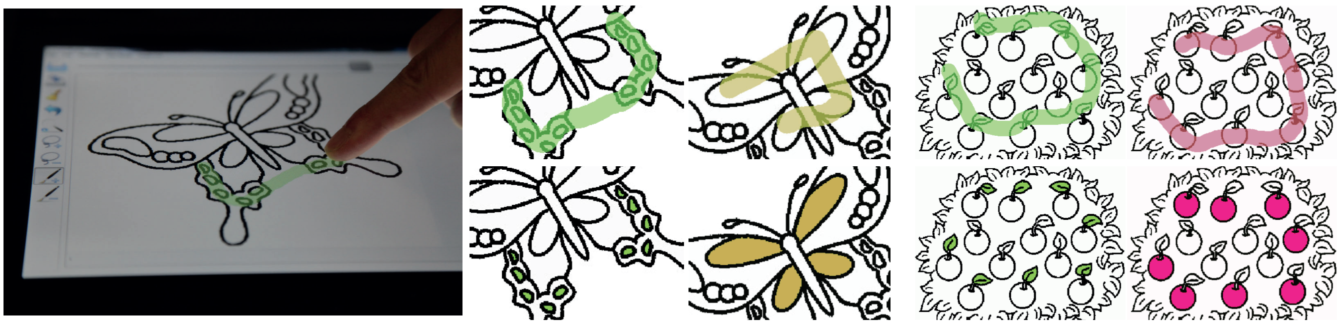


Figure 1: *Our* Lazy Selection *tool allows a user to quickly select desired shape element(s) with a scribble-based user interface. Note that roughly drawn scribbles are sufficient to express the user's intention, making our tool effectively applicable to imprecise input means, e.g., touch input systems.*

Abstract

This paper presents *Lazy Selection*, a scribble-based tool for quick selection of one or more desired shape elements by roughly stroking through the elements. Our algorithm automatically refines the selection and reveals the user's intention. To give the user maximum flexibility but least ambiguity, our technique first extracts selection candidates from the scribble-covered elements by examining the underlying patterns and then ranks them based on their location and shape with respect to the user-sketched scribble. Such a design makes our tool tolerant to imprecise input systems and applicable to touch systems without suffering from the fat finger problem. A preliminary evaluation shows that compared to the standard click and lasso selection tools, which are the most commonly used, our technique provides significant improvements in efficiency and flexibility for many selection scenarios.

Keywords: selection, shape elements, scribble-based UI

Links: DL PDF

1 Introduction

Manually selecting shape elements or objects is one of the most frequently performed tasks in various editing scenarios, e.g., manipulating a vector graphics image or 3D scene. Mainly due to its simplicity, click-based selection has been one of the most commonly used tools, and is adopted in almost every graphics editor (e.g., Adobe Illustrator, Autodesk Maya etc.). However, since it requires every click to be placed exactly within the boundary of an element of interest, such a click-based operation demands intense visual attention and requires accurate means of input. It is expected that its performance would degrade when target elements are small or displayed on screens of small size (e.g., on smartphones). Given the frequent usage of the selection operation, even a slight improvement in its efficiency would save significant user time as a whole.

In this work, we explore a new element selection tool based on a scribble-based user interface (UI). Scribble-based UI has already been successfully used for interactive segmentation of bitmap images and 3D models (Section 2). Compared with click-based UI, we anticipate the scribble-based UI for element selection to provide the following advantages. First, besides the location information of a user-specified scribble, its shape information can be potentially employed to more effectively recognize the user's intent, through shape matching between the scribble and its covered elements (Figure 1). Second, unlike discrete clicks, continuous scribbles naturally support the selection of multiple elements in a single action and more importantly allow the user to quickly select a desired group of elements by analyzing the underlying patterns among elements (Figure 1). Third, such a UI would be more tolerant to errors due to either imprecise input means or casual selection, since the selection decision is made not based on only the input location information alone. This is highly beneficial especially for touch input systems (Figure 1), where accuracy is diminished by the fat finger problem [Au et al. 2012].

However, automatically picking the desired element(s) from the set of elements (partially/fully) covered by a user-specified scribble is not an easy task. For example, it is unclear whether the shape of either individual elements or groups of elements should be used for matching (Figure 2 (a) and (b)). Still, the role of the scribble shape information might vary case by case. In some cases (Figure 2 (c)) there might be no explicit shape relationship between the scribble and the desired element(s). In addition, the grouping information is unknown and has to be derived on the fly, which is a challenging problem on its own. Whether a single group or multiple groups of elements are desired also remains unknown (Figure 2 (d)). In short, there is no obvious decision strategy for revealing the user's intent from the scribble.

To address the above problems, we present a two-phase approach. First, we establish a set of selection candidates, one of which the user might intend to select (Section 3.1). To unify the various selection possibilities involving a single element, a group of elements,

ACM Reference Format
Xu, P., Fu, H., Au, O., Tai, C. 2012. Lazy Selection: A Scribble-based Tool for Smart Shape Elements Selection. *ACM Trans. Graph. 31* 6, Article 142 (November 2012), 9 pages. DOI = 10.1145/2366145.2366161 http://doi.acm.org/10.1145/2366145.2366161.

Material Memex: Automatic Material Suggestions for 3D Objects

Arjun Jain   Thorsten Thormählen   Tobias Ritschel   Hans-Peter Seidel

MPI Informatik

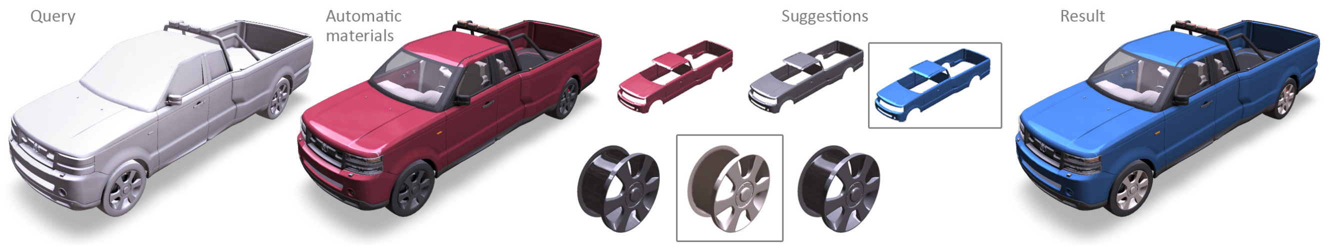


Figure 1: *Given an input query 3D object without materials* (left) *our approach automatically assigns materials* (center) *using database information and suggests alternatives to the user* (insets) *which can be selected interactively to improve the automatic assignment* (right).

Abstract

The material found on 3D objects and their parts in our everyday surroundings is highly correlated with the geometric shape of the parts and their relation to other parts of the same object. This work proposes to model this context-dependent correlation by learning it from a database containing several hundreds of objects and their materials. Given a part-based 3D object without materials, the learned model can be used to fully automatically assign plausible material parameters, including diffuse color, specularity, gloss, and transparency. Further, we propose a user interface that provides material suggestions. This user-interface can be used, for example, to refine the automatic suggestion. Once a refinement has been made, the model incorporates this information, and the automatic assignment is incrementally improved. Results are given for objects with different numbers of parts and with different topological complexity. A user study validates that our method significantly simplifies and accelerates the material assignment task compared to other approaches.

Links: DL PDF WEB VIDEO DATA

1 Introduction

Assigning materials to parts of a 3D object is a difficult and time consuming task that is performed by specially trained color & lighting artists in movie or game productions. The chosen palette of materials strongly influences the overall appearance of the 3D scene and is essential to allow the object to fit into an environment. There is a wide range of different materials, e. g., for a car there is the specular metallic paint work, the diffuse rubber material on the tire, the aluminium of the rim, the fabrics and leather used in the interior, etc. The compositing of materials is also important, as for example all screws on a tire should not just be metallic, but are likely to be of the same metallic material. Furthermore, materials are influenced by their context, e. g., for a part of a car's interior, leather or wood are far more likely materials than for a part of the car's engine.

Despite these observations, current content creation packages assign materials using a tedious manual process, involving the adjustment of rarely intuitive parameters, or by selecting pre-defined materials from a database using a keyword search. Even an experienced artist requires approximately 45 minutes to assign appropriate materials to all 130 parts of the car shown in Fig. 1.

In this work, we propose an approach to computationally model the relation of shape and material by learning it from a database of hundreds of multi-component 3D objects with materials. This model can then be used to automatically assign materials to 3D objects or can be employed in a user-interface to provide a ranked list of the most likely materials (see Fig. 1).

The paper comprises the following contributions:

- A model of the relation between materials and shape as well as context, called the *material memex*,
- Automatic assignment of materials using this model.
- A novel interface to guide a user when assigning materials by providing ranked material suggestions.
- A user study of task performance when using conventional slider or text interfaces compared to our interface.

The rest of the paper is structured as follows. The next section discusses related work. Section 3 introduces the proposed material memex. Section 4 presents two applications of the material memex: automatic material assignment and ranked material suggestion. Results are given in Section 5, and the paper ends with a discussion and conclusion.

2 Related Work

Material Assignment. Exploiting the relations of materials and shapes so far received only limited attention in the computer vision community, and even less interest from computer graphics. For editing and assigning material, different material design interfaces are employed. In most commercial 3D modelling tools it is still common to directly modify the parameters of the analytic reflectance model such as Phong [1975]. Ngan et al. [2006] propose an interface for BRDF selection that displays material variations with several preview images. There are also special solutions [Kautz et al. 2007;

ACM Reference Format
Jain, A., Thormählen, T., Ritschel, T., Seidel, H. 2012. Material Memex: Automatic Material Suggestions for 3D Objects. *ACM Trans. Graph. 31* 6, Article 143 (November 2012), 8 pages.
DOI = 10.1145/2366145.2366162 http://doi.acm.org/10.1145/2366145.2366162.

Interactive Bi-scale Editing of Highly Glossy Materials

Kei Iwasaki* Wakayama University
Yoshinori Dobashi† Hokkaido University
Tomoyuki Nishita‡ The University of Tokyo

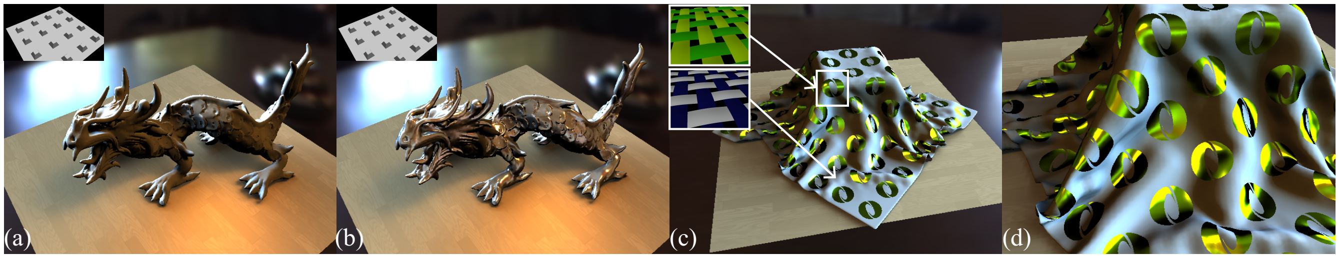


Figure 1: *Interactive (1.1-7.7 fps) bi-scale material design and real-time rendering (87-155 fps) under environment lighting. The small-scale geometries are shown in the insets of top left corner. (a) Rendering result using a small-scale diffuse BRDF. (b) Rendering result using a highly glossy BRDF. (c) Our method can handle spatially-varying bi-scale materials calculated from different small-scale geometries and BRDFs (upper left is used to render the SIGGRAPH logo). (d) is zoomed in image of (c). Two-hue appearance (green and yellow) of specular reflection can be rendered, which is difficult to render by using single-scale BRDFs.*

Abstract

We present a new technique for bi-scale material editing using Spherical Gaussians (SGs). To represent large-scale appearances, an effective BRDF that is the average reflectance of small-scale details is used. The effective BRDF is calculated from the integral of the product of the Bidirectional Visible Normal Distribution (BVNDF) and BRDFs of small-scale geometry. Our method represents the BVNDF with a sum of SGs, which can be calculated on-the-fly, enabling interactive editing of small-scale geometry. By representing small-scale BRDFs with a sum of SGs, effective BRDFs can be calculated analytically by convolving the SGs for BVNDF and BRDF. We propose a new SG representation based on convolution of two SGs, which allows real-time rendering of effective BRDFs under all-frequency environment lighting and real-time editing of small-scale BRDFs. In contrast to the previous method, our method does not require extensive precomputation time and large volume of precomputed data per single BRDF, which makes it possible to implement our method on a GPU, resulting in real-time rendering.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture;

Keywords: material editing, reflectance filtering, normal distribution function, spherical Gaussian

Links: DL PDF

*e-mail:iwasaki@sys.wakayama-u.ac.jp
†e-mail:doba@ime.ist.hokudai.ac.jp
‡e-mail:nis@is.s.u-tokyo.ac.jp

ACM Reference Format
Iwasaki, K., Dobashi, Y., Nishita, T. 2012. Interactive Bi-scale Editing of Highly Glossy Materials. *ACM Trans. Graph. 31* 6, Article 144 (November 2012), 7 pages. DOI = 10.1145/2366145.2366163 http://doi.acm.org/10.1145/2366145.2366163.

1 Introduction

Interactive editing of visual appearances under real-world complex lighting is beneficial in design applications. Recent advances in computer graphics enable us to edit surface materials in real-time [Ben-Artzi et al. 2006; Sun et al. 2007; Cheslack-Postava et al. 2008; Wang et al. 2009]. These methods mainly focus on editing surface materials on a single scale. The visual appearance, however, depends on the scale of the geometry. For example, small-scale bumps on metal surfaces can be seen if the viewer is very close to the surface, while anisotropic reflection of light can be seen as the viewer zooms out from the surface.

In recent years, a physically-based editing method for visual appearance at the bi-scale level (large/small-scale[1]) has been proposed [Wu et al. 2011]. Their method introduced a bidirectional visible normal distribution function (BVNDF), which describes the normal distribution function of small-scale geometry taking into account shadowing and masking effects. The large-scale appearance is described by using the effective BRDF [Han et al. 2007; Wu et al. 2011], which is calculated by integrating the product of BVNDF and the small-scale BRDFs. Their method calculates the effective BRDF by the product of two matrices that store a densely sampled small-scale BRDF and BVNDF. To accelerate the computation time and to compress the matrices, a random projection method is employed [Vempala 2004]. Unfortunately, since their method is not a closed-form solution and hence based on numerical integration, it requires dense sampling of different directions for highly glossy BRDFs, resulting in the requirement for a huge volume of precomputed data and extensive precomputational time for BVNDF and BRDFs.

To address these problems, we propose a closed-form solution to calculate effective BRDFs by using spherical Gaussians (SGs). The key insight of our method is to introduce a unified framework of SG representation for small-scale BRDFs, BVNDFs, and effective BRDFs. Our unified SG framework has many advantages in terms of achieving efficient rendering with effective BRDFs and editing of small-scale geometries and small-scale BRDFs. Our method rep-

[1] In this paper, we follow the terminology of *large/small-scale* [Wu et al. 2011] instead of *micro/mili-scale*. Our method will work as long as the ratio between the large and the small-scale is big, and the small-scale is greater than the wavelength of light.

An Inverse Problem Approach for Automatically Adjusting the Parameters for Rendering Clouds Using Photographs

Yoshinori Dobashi[1,2] Wataru Iwasaki[1] Ayumi Ono[1] Tsuyoshi Yamamoto[1] Yonghao Yue[3] Tomoyuki Nishita[3]
[1]Hokkaido University [2]JST, CREST [3]The University of Tokyo

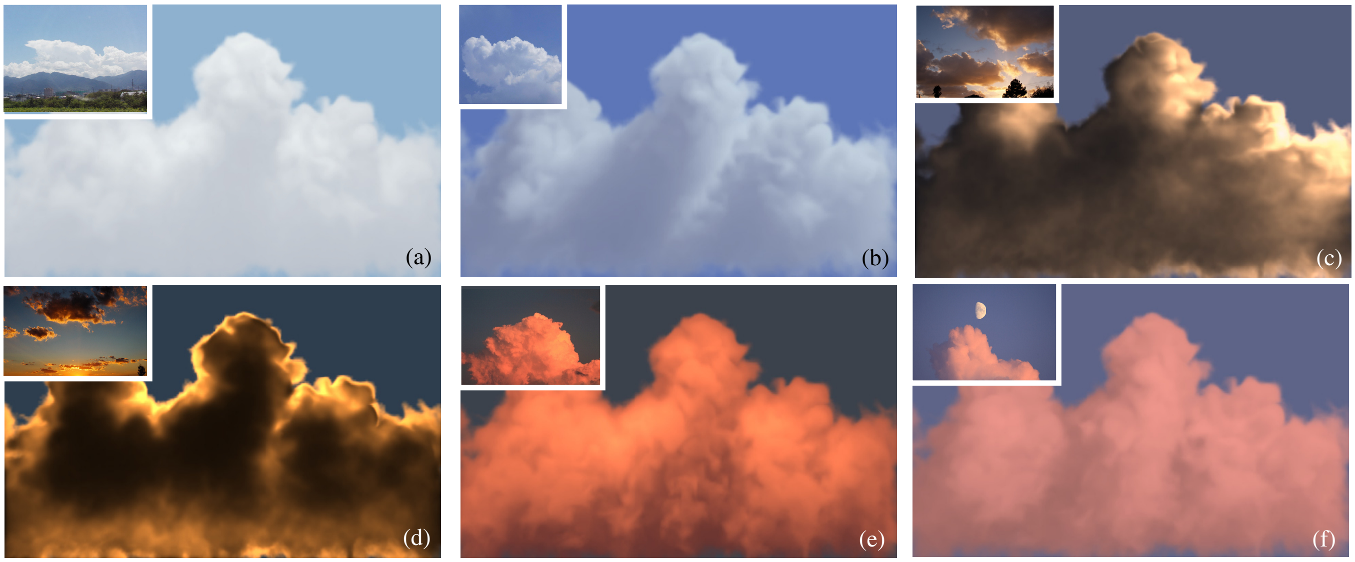


Figure 1: *Examples of our method. The parameters for rendering clouds are estimated from the real photograph shown in the small inset at the top left corner of each image. The synthetic cumulonimbus clouds are rendered using the estimated parameters.*

Abstract

Clouds play an important role in creating realistic images of outdoor scenes. Many methods have therefore been proposed for displaying realistic clouds. However, the realism of the resulting images depends on many parameters used to render them and it is often difficult to adjust those parameters manually. This paper proposes a method for addressing this problem by solving an inverse rendering problem: given a non-uniform synthetic cloud density distribution, the parameters for rendering the synthetic clouds are estimated using photographs of real clouds. The objective function is defined as the difference between the color histograms of the photograph and the synthetic image. Our method searches for the optimal parameters using genetic algorithms. During the search process, we take into account the multiple scattering of light inside the clouds. The search process is accelerated by precomputing a set of intermediate images. After ten to twenty minutes of precomputation, our method estimates the optimal parameters within a minute.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism; I.3.3 [Computer Graphics]: Picture/Image Generation;

Keywords: clouds, rendering parameters, inverse problem

Links: DL PDF

1 Introduction

Clouds are important elements when synthesizing images of outdoor scenes to enhance realism. A volumetric representation is often employed and the intensities of the clouds are computed taking into account the scattering and absorption of light in order to display realistic clouds. However, one of the problems is that the quality of the rendered image depends on many parameters, which need to be adjusted manually by rendering the clouds repeatedly with different parameter settings. This is not an easy task since the relationship between the resulting appearance of the clouds and the parameters is highly nonlinear. The expensive computational cost for the rendering process makes this more difficult. This paper focuses on automatic adjustment of the parameters to address this task.

Recently, many real-time methods have been proposed for editing the parameters used in rendering images [Harris and Lastra 2001; Bouthors et al. 2008; Zhou et al. 2008]. These methods are fast so-

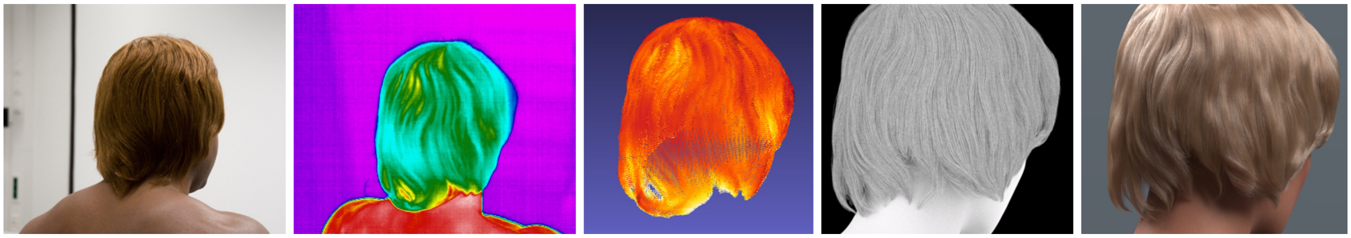
ACM Reference Format
Dobashi, Y., Iwasaki, W., Ono, A., Yamamoto, T., Yue, Y., Nishita, T. 2012. An Inverse Problem Approach for Automatically Adjusting the Parameters for Rendering Clouds Using Photographs. *ACM Trans. Graph. 31* 6, Article 145 (November 2012), 10 pages. DOI = 10.1145/2366145.2366164 http://doi.acm.org/10.1145/2366145.2366164.

Lighting Hair From The Inside: A Thermal Approach To Hair Reconstruction

Tomas Lay Herrera*
University of Bonn

Arno Zinke†
TeamUp Technologies, Bonn

Andreas Weber‡
University of Bonn



Abstract

Generating plausible hairstyles is a very challenging problem. Despite recent efforts no definite solution was presented so far. Many of the current limitations are related to the optical complexity of hair. In this paper we present a technique for hair reconstruction based on thermal imaging. By using this technique several issues of conventional image-based techniques, such as shadowing and anisotropy in reflectance, can be avoided. Moreover, hair-skin segmentation becomes a trivial problem, and no special care about lighting has to be taken, as the hair is "lit from inside" with the head as light source. The capture process is fast and requires a single hand-held device only. The potential of the proposed method is demonstrated by several challenging examples.

Keywords: hair modeling, image-based modeling, 3-D reconstruction, thermal imaging

Links: DL PDF

1 Introduction

Hair is one of the key features affecting human appearance. Because of its visual importance, hair modeling is also a critical step in the creation of digital avatars.

Generating a believable hairstyle, however, is a very challenging task as even slightest errors may lead to visually unconvincing results. On the other hand geometrical complexity is a serious issue. A typical human hairstyle consists of 100000–150000 filaments whose shape is a result of physical and chemical effects including complex fiber-fiber interactions and cosmetic products. Therefore, most hairstyle modeling techniques presented to date rely on manual work, requiring special skills from the user to obtain satisfactory results [Kim and Neumann 2000; Kim and Neumann 2002; Ward et al. 2003; Yuksel et al. 2009].

To reduce the degree of manual interaction some approaches reproduce hairstyles with the help of physical models describing the behavior of the hair strands [Hadap and Magnenat-Thalmann 2000; Bertails et al. 2005; Choe and Ko 2005; Sobottka et al. 2006].

Other works attempt to reproduce existing hairstyles taking an image-based approach. Ideally, this would the solve the modeling problem, as hair is captured from real samples. However, all methods presented so far require either a very complex setup, with controlled lighting, multiple cameras [Paris et al. 2004; Paris et al. 2008] or substantial manual pre-processing, e.g. for hair-skin segmentation to obtain acceptable results [Wei et al. 2005]. Moreover, they are likely to give unreliable results in cases where global optical effects –e.g. self-shadowing or strong multiple scattering– cannot be correlated with local shape. Another limitation is related to the fact that classical image-based approaches can only image the surface of a hair volume but give no insights about the volumetric structure inside.

Following previous image-based work we present a novel, fast and more robust approach for reconstructing a given hairstyle from video streams. The key idea is to use far infrared imaging to overcome most of the issues related to the visual range. With the human head as "light source" hair-skin segmentation becomes a trivial task and a dedicated photo-consistency-based approach can be used for capturing a full hairstyles under arbitrary lighting conditions using a single hand-held thermal video camera.

To improve accuracy and computational efficiency we extend existing state-of-the-art along several lines. Taking advantage of the local relationship between temperature and distance to skin, photo-consistency is accelerated by one order of magnitude enabling high resolution reconstruction of a "hairstyle boundary". Moreover, multi-scale orientation analysis leads to a more robust estimation of local hair orientation, a key component for synthesis of the hair model. Finally, we propose a novel two step approach for synthesizing individual hair strands where more reliable data is considered first and smoothness and curvature constraints help to recover continuous strands, even in case of curly hair.

1.1 Related Work

According to Ward et al. [2007] hair modeling techniques can be classified with respect to required input as either geometry-based, physically-based or image-based.

*e-mail: tomas@cs.uni-bonn.de
†e-mail: a.zinke@getteamup.com
‡e-mail: weber@cs.uni-bonn.de

ACM Reference Format
Herrera, T., Zinke, A., Weber, A. 2012. Lighting Hair From The Inside: A Thermal Approach To Hair Reconstruction. *ACM Trans. Graph. 31* 6, Article 146 (November 2012), 9 pages. DOI = 10.1145/2366145.2366165 http://doi.acm.org/10.1145/2366145.2366165.

New Measurements Reveal Weaknesses of Image Quality Metrics in Evaluating Graphics Artifacts

Martin Čadík* Robert Herzog* Rafał Mantiuk◇ Karol Myszkowski* Hans-Peter Seidel*
*MPI Informatik Saarbrücken, Germany ◇Bangor University, United Kingdom

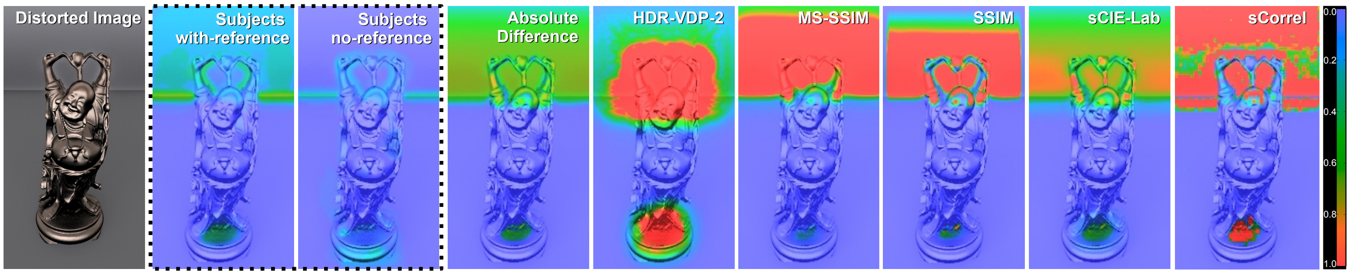


Figure 1: *State-of-the-art image quality metrics often fail in the prediction of the human-perceived distortions in complex images. Here, we show the predicted detection probabilities (color-coded) for gradient-based tone mapping artifacts [Fattal et al. 2002] in a synthetic image.*

Abstract

Reliable detection of global illumination and rendering artifacts in the form of localized distortion maps is important for many graphics applications. Although many quality metrics have been developed for this task, they are often tuned for compression/transmission artifacts and have not been evaluated in the context of synthetic CG-images. In this work, we run two experiments where observers use a brush-painting interface to directly mark image regions with noticeable/objectionable distortions in the presence/absence of a high-quality reference image, respectively. The collected data shows a relatively high correlation between the with-reference and no-reference observer markings. Also, our demanding per-pixel image-quality datasets reveal weaknesses of both simple (PSNR, MSE, sCIE-Lab) and advanced (SSIM, MS-SSIM, HDR-VDP-2) quality metrics. The most problematic are excessive sensitivity to brightness and contrast changes, the calibration for near visibility-threshold distortions, lack of discrimination between plausible/implausible illumination, and poor spatial localization of distortions for multi-scale metrics. We believe that our datasets have further potential in improving existing quality metrics, but also in analyzing the saliency of rendering distortions, and investigating visual equivalence given our with- and no-reference data.

CR Categories: I.3.0 [Computer Graphics]: General;

Keywords: Image quality metrics (IQM), perceptual experiments, global illumination, noticeable and objectionable distortions

Links: DL PDF WEB DATA

*e-mail: mcadik@mpi-inf.mpg.de, the complete dataset is available at: http://www.mpii.de/resources/hdr/iqm-evaluation/

ACM Reference Format
Cadik, M., Herzog, R., Mantiuk, R., Myszkowski, K., Seidel, H. 2012. New Measurements Reveal Weaknesses of Image Quality Metrics in Evaluating Graphics Artifacts. *ACM Trans. Graph. 31* 6, Article 147 (November 2012), 10 pages. DOI = 10.1145/2366145.2366166 http://doi.acm.org/10.1145/2366145.2366166.

1 Introduction

Rendering techniques, in particular global illumination, are prone to image artifacts, which might arise due to specific scene configurations, imbalanced scene complexity that might lead to a locally varying convergence-rate of the solution, and numerous simplifications in the rendering algorithms themselves. With the proliferation of 3D rendering services, where the user may often arbitrarily interact with the content, the role of automatic rendering-quality control gains in importance. Even in well-established industries such as gaming a massive approach to automatic quality testing is desirable. In practice, objective image quality metrics (IQM) that are successful in lossy image compression and transmission applications [Wang and Bovik 2006] are predominantly used in graphics, including advanced attempts of their adaptation to actively steer rendering [Rushmeier et al. 1995; Bolin and Meyer 1998; Ramasubramanian et al. 1999]. Such objective IQM are trained to predict a single value of mean opinion score (MOS) for image blockiness, noise, blur, or ringing distortions. However, their performance for other distortion types as well as their spatial localization within an image has not been systematically validated so far.

The goal of this work is to generate a new rendering-oriented dataset with localized distortion maps and use it for the evaluation of existing IQM. For this purpose we prepare a set of images with distortions that are typical for popular global illumination and rendering techniques as well as the corresponding distortion-free reference images. Table 1 presents a summary of our stimuli. In two separate experiments (Sec. 3) we ask the observers to locally mark *noticeable* and *objectionable* distortions where the reference image is either shown or hidden, respectively. We demonstrate that the observers can reliably perform both tasks, yielding high coefficients of agreement (Sec. 4.1). In general, our results show a high correlation between the observer marking for the *with-reference* and *no-reference* datasets, but we also indicate the most common sources of discrepancies in such marking (Sec. 4.2).

We use the with-reference dataset to evaluate the performance of state-of-the-art *full-reference* (FR) IQM in detecting and localizing rendering distortions (Sec. 5). We show that even advanced IQM fail for some common computer graphics artifacts (e.g., Fig. 1). Our data shows that in general no IQM performs better than any other, even including the simple absolute difference (AD), which is equivalent to the peak signal-to-noise ratio (PSNR) or mean-

Large-scale Fluid Simulation using Velocity-Vorticity Domain Decomposition

Abhinav Golas*[1], Rahul Narain†[2], Jason Sewall‡[3], Pavel Krajcevski§[1], Pradeep Dubey¶[3], and Ming Lin‖[1]

[1]University of North Carolina at Chapel Hill
[2]University of California, Berkeley
[3]Intel Corporation

(a) (b) (c)

Figure 1: *Examples of fluids simulated with our technique: (a) a city block hit by a tsunami (vortex domain in yellow) (b) seagulls flying through smoke (c) smoke flow around a sphere. We achieve up to three orders of magnitude of performance over standard grid-only techniques.*

Abstract

Simulating fluids in large-scale scenes with appreciable quality using state-of-the-art methods can lead to high memory and compute requirements. Since memory requirements are proportional to the product of domain dimensions, simulation performance is limited by memory access, as solvers for elliptic problems are not compute-bound on modern systems. This is a significant concern for large-scale scenes. To reduce the memory footprint and memory/compute ratio, vortex singularity bases can be used. Though they form a compact bases for incompressible vector fields, robust and efficient modeling of nonrigid obstacles and free-surfaces can be challenging with these methods.

We propose a hybrid domain decomposition approach that couples Eulerian velocity-based simulations with vortex singularity simulations. Our formulation reduces memory footprint by using smaller Eulerian domains with compact vortex bases, thereby improving the memory/compute ratio, and simulation performance by more than 1000x for single phase flows as well as significant improvements for free-surface scenes. Coupling these two heterogeneous methods also affords flexibility in using the most appropriate method for modeling different scene features, as well as allowing robust interaction of vortex methods with free-surfaces and nonrigid obstacles.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation; I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Physically based modeling;

Keywords: Computational fluid dynamics, vortex methods, stable fluids, Navier-Stokes equations

*golas@cs.unc.edu
†narain@eecs.berkeley.edu
‡jason.sewall@intel.com
§pavel@cs.unc.edu
¶pradeep.dubey@intel.com
‖lin@cs.unc.edu

Links: DL PDF WEB

1 Introduction

State-of-the-art methods for fluid simulation, including velocity-based Eulerian methods and smoothed particle hydrodynamics, model the entire *spatial* extent of the fluid. Discretization of this space is often chosen to be able to sample sufficiently fine details under the restriction of limited computational resources. As a result, scenes with large spatial scales can only be simulated to coarse detail on PCs, relying on procedural methods to infuse detail. The simple computational kernels of these methods are largely *memory-bandwidth bound*, since domains of interest cannot reside in caches of current generation CPUs, and computational complexity cannot mask the cost of memory accesses. An alternate approach to modeling fluids is to model fluid *detail*, represented by the *vorticity* of the fluid, *i.e.* the curl of the velocity field. For incompressible flows, vorticity can be compactly represented by Lagrangian singularity elements. They are thus free of numerical dissipation, which can be a significant issue with Eulerian methods, and do not need to explicitly model the pressure of the fluid. Though this leads to computational savings for scenes with unbounded fluid, robust and efficient modeling of obstacles or free-surfaces with two-way coupling using vorticity methods is challenging. Vortex singularity elements also serve as intuitive models for visual fluid detail, e.g. a smoke ring can be modeled as a vortex curve or filament.

These aspects make *detail* modeling of fluids with vortex singularity

ACM Reference Format
Golas, A., Narain, R., Sewall, J., Krajcevski, P., Dubey, P., Lin, M. 2012. Large-Scale Fluid Simulation using Velocity-Vorticity Domain Decomposition. *ACM Trans. Graph. 31* 6, Article 148 (November 2012), 9 pages. DOI = 10.1145/2366145.2366167 http://doi.acm.org/10.1145/2366145.2366167.

Staggered Meshless Solid-Fluid Coupling

Xiaowei He *†‡, Ning Liu‡, Guoping Wang‡, Fengjun Zhang†, Sheng Li‡, Songdong Shao§, and Hongan Wang†
†Institute of Software, Chinese Academy of Sciences ‡Peking University §University of Bradford

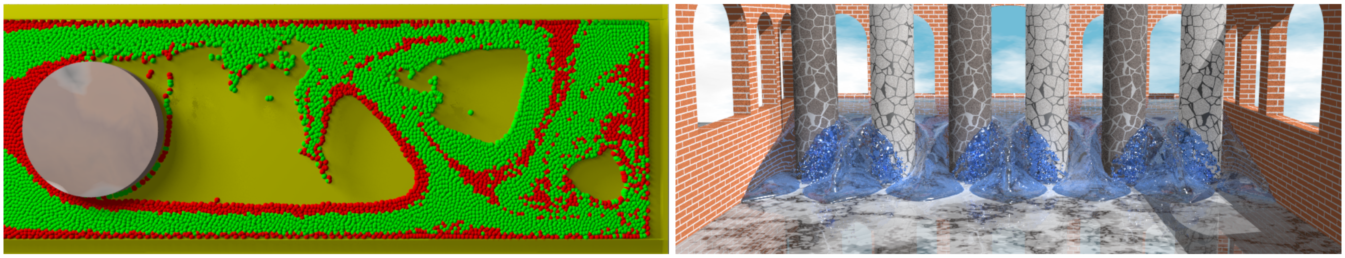


Figure 1: *Left: fluid flowing around a cylinder with no-slip boundary condition in two dimensions. Right: dam break of fluid flowing around rotating cylinders with no-slip boundary condition in three dimensions.*

Abstract

Simulating solid-fluid coupling with the classical meshless methods is an difficult issue due to the lack of the Kronecker delta property of the shape functions when enforcing the essential boundary conditions. In this work, we present a novel staggered meshless method to overcome this problem. We create a set of staggered particles from the original particles in each time step by mapping the mass and momentum onto these staggered particles, aiming to stagger the velocity field from the pressure field. Based on this arrangement, an new approximate projection method is proposed to enforce divergence-free on the fluid velocity with compatible boundary conditions. In the simulations, the method handles the fluid and solid in a unified meshless manner and generalizes the formulations for computing the viscous and pressure forces. To enhance the robustness of the algorithm, we further propose a new framework to handle the degeneration case in the solid-fluid coupling, which guarantees stability of the simulation. The proposed method offers the benefit that various slip boundary conditions can be easily implemented. Besides, explicit collision handling for the fluid and solid is avoided. The method is easy to implement and can be extended from the standard SPH algorithm in a straightforward manner. The paper also illustrates both one-way and two-way couplings of the fluids and rigid bodies using several test cases in two and three dimensions.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: staggered SPH, physically-based animation, solid-fluid coupling

Links: DL PDF

*clouddon.he@gmail.com

ACM Reference Format
He, X., Liu, N., Wang, G., Zhang, F., Li, S., Shao, S., Wang, H. 2012. Staggered Meshless Solid-Fluid Coupling. *ACM Trans. Graph. 31* 6, Article 149 (November 2012), 12 pages. DOI = 10.1145/2366145.2366168 http://doi.acm.org/10.1145/2366145.2366168.

1 Introduction

The solid-fluid interactions are common in everyday life, such as when pouring beer into a glass, dropping a stone into the water, etc. In the documented meshless methods, much attention has been paid to the fluid animations, such as the works [Adams et al. 2007; Solenthaler and Pajarola 2009; Solenthaler and Gross 2011] based on the Lagrangian *Smoothed Particle Hydrodynamics* (SPH) method. However, under this purely Lagrangian framework, only few works have dealt with the coupling problem since there are still several problems in the SPH which have not been fully addressed [Liu and Liu 2010]. For example, due to the lack of the Kronecker delta property of the shape functions, the free-slip and no-slip boundary conditions cannot be imposed as easily as in a mesh-based method. The spurious zero-energy modes are also well known for the null-space issues, which is caused by the fact that the field variables and their derivatives are calculated at the same positions. As a result, high-frequency oscillations in the variable field may persist or even grow up that will lead to the simulation failure.

Referring to the grid-based methods, a novel "MAC-grid"as shown in Figure 2 (left) [Harlow et al. 1965] is commonly used to overcome the problem of the zero-energy modes. Inspired by this brilliant work, [Vignjevic et al. 2000; Randles and Libersky 2000] extended the SPH to a stress-point formulation which was composed of the stress particles and velocity particles. With this arrangement, the zero-energy modes can be trivially avoided as the colocational nature of the standard SPH is removed by adding the stress particles. The boundary treatment is also simple, since it is possible to use either type of the particles to enforce the boundary conditions. However, this method has several disadvantages that make it inappropriate to apply directly to the solid-fluid interactions. One major problem comes from the irregular distributions of the particles in the simulation. Both compression and expansion of the particles can severely degrade the solution, resulting in the simulation failure. Besides, the interleaved pattern of the stress particles and velocity particles cannot be well preserved if complex fluid interactions with the solid occur. In addition, the problems, such as extra computing cost of finding the neighbors for stress particles and the numerical dissipations, also limit the application of this method.

In this paper, a new arrangement of different particles is devised as shown in Figure 2 (right). To ensure the stability, we enforce the density constraint on the carrier particles, which represent the discretization of the problem domain, and create the staggered particles at each time step at the positions between each pair of the

Automated Constraint Placement to Maintain Pile Shape

Shu-Wei Hsu
Texas A&M University

John Keyser
Texas A&M University

Figure 1: *To create a pile of a given shape, a user specifies a desired shape and starting configuration of objects. At right, a pile starting in a desired configuration will collapse under simulation. The middle shows the result of adding constraints judiciously to maintain the pile shape. At right is the desired control mesh (in white), with the objects that have had constraints applied colored in cyan.*

Abstract

We present a simulation control to support art-directable stacking designs by automatically adding constraints to stabilize the stacking structure. We begin by adapting equilibrium analysis in a local scheme to find "stable" objects of the stacking structure. Next, for stabilizing the structure, we pick suitable objects from those passing the equilibrium analysis and then restrict their DOFs by managing the insertion of constraints on them. The method is suitable for controlling stacking behavior of large scale. Results show that our control method can be used in varied ways for creating plausible animation. In addition, the method can be easily implemented as a plug-in into existing simulation solvers without changing the fundamental operations of the solvers.

CR Categories: I.3.5 [Computer Graphics]: Computational Geoemtry and Object Modeling—Physically Based Modeling I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: rigid bodies, simulation control, equilibrium analysis

Links: DL PDF

1 Introduction

Modeling massive, disordered piles of objects, e.g. stone piles or food piles is used in the creation of digital content such as video games and films. To create appealing results and to satisfy specific needs of an artist, a goal-directable method is desired.

However, this goal-directable problem is very challenging because simulations are usually formulated as initial configuration problems. Given an initial configuration of a simulation scene, the simulator will evolve the initial configuration step by step. There is no guarantee that the results of the simulation will still be close to the design that a user provided. In addition, directly tweaking the simulation by adjusting the physics parameters is inefficient, and often ineffective, since we do not know the mapping between the parameters and our goal.

While procedural synthesis approaches are able to synthesize piles explicitly, the resulting piles cannot be directly used in a dynamic simulation since the synthesis process mainly considers the geometry properties, not the dynamic simulation properties. If users want to use this resulting pile in a dynamic simulation further, e.g. stack more objects on it or destroy it with other objects, it is very possible the shape will fall as soon as dynamic simulations begin.

The main result of this paper is a method for stabilizing the stacking process of simulations, helping preserve the design of object stacking. We demonstrate that our method is able to guide the stacking simulation to desired pile shapes. Plausibility of the simulation is also maintained. During simulation, our method eliminates the degrees of freedom (DOFs) of certain stacking objects which are considered "stable" locally. It helps preserve the volumes of the desired shapes. Our method includes two stages. The goal of the first stage is to find object candidates that satisfy static equilibrium conditions locally. In the second stage, we evaluate whether the DOFs of candidates can be removed to stabilize the structure without decreasing plausibility. Then we group eligible candidates to remove some DOFs temporarily, stabilizing the stacking structure.

2 Related Works

Contact and friction Modeling contact and friction is essential to rigid body simulation. Baraff[1989; 1994], Hahn[1988], and Mirtich and Canny[1995] presented some fundamental work. For an overview of the concepts and technologies developed in this field, we refer the reader to a recent state-of-the-art report on rigid body simulation in computer graphics[Bender et al. 2012].

Contact and friction stabilization Modeling contact and friction accurately helps objects rest stably in simulations. Solving for contact and friction separately, which most solvers do, leads to stability issues. Kaufman et al.[2008] propose a coupled model for solving contact and friction together and demonstrate that the method is able to simulate structures that need a lot of friction for support(e.g. a card tower as in Figure 6).

Temporarily stopping some objects is another way to stabilize structures. Mattikalli et al.[1994] proposed a method for stably assembling objects under gravity. The method finds critical positions

ACM Reference Format
Hsu, S., Keyser, J. 2012. Automated Constraint Placement to Maintain Pile Shape.
ACM Trans. Graph. 31 6, Article 150 (November 2012), 6 pages. DOI = 10.1145/2366145.2366169
http://doi.acm.org/10.1145/2366145.2366169.

Speculative Parallel Asynchronous Contact Mechanics

Samantha Ainsley, Columbia University
Etienne Vouga, Columbia University
Eitan Grinspun, Columbia University
Rasmus Tamstorf, Walt Disney Animation Studios

Abstract

We extend the Asynchronous Contact Mechanics algorithm [Harmon et al. 2009] and improve its performance by two orders of magnitude, using only optimizations that do not compromise ACM's three guarantees of safety, progress, and correctness. The key to this speedup is replacing ACM's timid, forward-looking mechanism for detecting collisions—locating and rescheduling separating plane kinetic data structures—with an optimistic speculative method inspired by Mirtich's rigid body Time Warp algorithm [2000]. Time warp allows us to perform collision detection over a window of time containing many of ACM's asynchronous trajectory changes; in this way we cull away large intervals as being collision free. Moreover, by replacing force processing intermingled with KDS rescheduling by windows of pure processing followed by collision detection, we transform an algorithm that is very difficult to parallelize into one that is embarrassingly parallel.

CR Categories: I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: contact, collision, simulation, parallelization

Links: DL PDF

1 Introduction

The design of physical simulation algorithms can require difficult decisions weighing slower, more principled approaches against faster shortcuts. In the short term, it can be tempting or necessary to prioritize speed. However, in the words of Sutter and Alexandrescu [2004], "it is far, far easier to make a correct program fast than it is to make a fast program correct." In this work we adopt the longer term perspective that as computers become ever faster, methods designed from the ground up to guarantee correctness will prove to be the longest-lasting.

Moreover, while available processing power continues to increase exponentially, this increase is no longer in the raw speed of CPU cores, but in the number of cores available per die [Borkar and Chien 2011]. Algorithms that cannot be parallelized are less likely to survive in the long term.

We seek algorithms to simulate thin flexible materials subject to complex collisions and contact geometries. *Asynchronous contact mechanics* (ACM) [Harmon et al. 2009] addresses this goal by focusing on three built-in guarantees: 1) the simulation is *safe* and provably stops all interpenetrations; 2) it conserves momentum and energy, *physical invariants*, for physical systems with the appropriate symmetries; and 3) for well-posed problems, ACM is guaranteed to make *progress*, in the sense of terminating in finite time. There is, however, a wide gulf between "finite time" and "fast," and the earlier paper was presented as a foundation for systems-style research into correct and *fast* simulations.

We describe a new system for ACM that dramatically decreases the amount of time spent on bookkeeping and collision detection, by an order of magnitude. While the original ACM was difficult to parallelize, we employ *speculation* to expose easy parallelization, leading to another order of magnitude speedup.

Our implementation yields speedups of more than two orders of magnitudes on a 12-core work station, enabling practical computational cost for simulations with complex contact geometries. The algorithm retains ACM's three aforementioned guarantees, and therefore serves as one step toward realizing Sutter and Alexandrescu's statement in the context of physical simulation.

1.1 Overview

Summary of ACM The ACM algorithm upon which our method is based is described in detail by Harmon et al. [2009] and in follow-up work [Harmon 2010; Vouga et al. 2011; Harmon et al. 2011a]. We only briefly summarize the salient features of the method here.

ACM guarantees safety and correctness of the simulation by activating nested *penalty layers* of decreasing thickness $\eta_1 > \eta_2 > \ldots$ and increasing stiffness in anticipation of collisions. Whenever two objects approach each other with distance less than η_1, ACM adds a penalty force to the simulation that opposes the collision. If the objects continue approaching and reach a distance η_2 from each other, a second penalty force is added that's stiffer than the first, and so on; the total force applied to the objects grows unbounded as the distance between them decreases, so that it is guaranteed to stop the collision no matter how hard the impact. Each penalty force is stepped *asynchronously* [Lew et al. 2003] instead of in lockstep, allowing each force to be integrated at its own stable time step. An *event priority queue* keyed by time maintains which force is to be processed next.

To detect when to activate a new penalty force, ACM uses *kinetic data structures* (KDSs) [Guibas 1998]. For each pair of primitives in the simulation, a separating slab of thickness η_1 is found that separates the primitives, guaranteeing that they are farther than η_1 apart. This slab certifies that no collisions can occur between them for some time interval into the future. A *KDS event* is placed on the event queue at this expiration time. Whenever the velocity of either primitive changes, this time must be recomputed, or *rescheduled.* When the KDS event is popped from the queue, either a new separating slab is found and its event pushed onto the queue, or the first penalty layer is activated for the two primitives under the conservative assumption that a collision is imminent. The above process is then repeated to detect when to activate the next deepest penalty layer at distance η_2.

An alternative to KDSs The above approach to activating penalty layers guarantees that collisions cannot be missed. Unfortunately, this guarantee carries a heavy cost: in typical simulations

ACM Reference Format
Ainsley, S., Vouga, E., Grinspun, E., Tamstorf, R. 2012. Speculative Parallel Asynchronous Contact Mechanics. *ACM Trans. Graph. 31* 6, Article 151 (November 2012), 8 pages. DOI = 10.1145/2366145.2366170 http://doi.acm.org/10.1145/2366145.2366170.

Adaptive Anisotropic Remeshing for Cloth Simulation

Rahul Narain Armin Samii James F. O'Brien

University of California, Berkeley

Figure 1: *A motion-captured character performs a jumping kick. His clothing is dynamically remeshed to capture detail such as wrinkles, while having larger elements in smooth areas. Here and elsewhere in the paper, large elements are shown in blue, small equilateral elements in red, and anisotropic elements in yellow.*

Abstract

We present a technique for cloth simulation that dynamically refines and coarsens triangle meshes so that they automatically conform to the geometric and dynamic detail of the simulated cloth. Our technique produces anisotropic meshes that adapt to surface curvature and velocity gradients, allowing efficient modeling of wrinkles and waves. By anticipating buckling and wrinkle formation, our technique preserves fine-scale dynamic behavior. Our algorithm for adaptive anisotropic remeshing is simple to implement, takes up only a small fraction of the total simulation time, and provides substantial computational speedup without compromising the fidelity of the simulation. We also introduce a novel technique for strain limiting by posing it as a nonlinear optimization problem. This formulation works for arbitrary non-uniform and anisotropic meshes, and converges more rapidly than existing solvers based on Jacobi or Gauss-Seidel iterations.

Keywords: Cloth simulation, dynamic remeshing, anisotropic remeshing, strain limiting, augmented Lagrangian method.

Links: DL PDF VIDEO WEB

1 Introduction

Cloth in real life simultaneously exhibits both highly detailed wrinkles and folds, and flat or smoothly curving regions. As the cloth moves, the fine wrinkles appear and disappear at different locations. Wrinkles may also migrate across the surface or travel in a wave-like fashion. With a fixed simulation mesh, adequately resolving these details requires a large number of extremely small elements over the entire mesh. The resulting computation and memory costs for uniformly high-resolution simulations can be burdensome, and it is instead desirable to focus simulation resolution in regions that exhibit complex shape and motion, while representing flat regions with coarser elements.

In this paper, we present a technique to dynamically refine and coarsen a finite element mesh used for cloth simulation. Our scheme creates anisotropic elements that follow the curvature of wrinkles and creases in the cloth. It preserves dynamic behaviors by maintaining resolution conforming to high velocity field gradients. It also anticipates the buckling of the material by refining the mesh where it is beginning to become compressed.

The resulting simulation method efficiently produces results that are visually equivalent to those produced with more costly uniformly high-resolution meshes. By adaptively refining and coarsening the mesh, elements are concentrated in detail regions. The anisotropic nature of our remeshing algorithm means that refining near clusters of long parallel wrinkles, which occur commonly in cloth, will require roughly linear instead of quadratic growth in the number of elements. Anisotropic remeshing also tends to align mesh edges with wrinkles so that visually they appear smoother. These features can be seen in Figures 1 and 2. Our adaptive, anisotropic remeshing procedure is computationally inexpensive and simple to implement.

We also introduce a fast technique for performing strain limiting on large meshes. Existing methods based on nonlinear Jacobi or Gauss-Seidel iterations converge slowly for large meshes. By casting strain limiting as a constrained optimization problem, we obtain faster convergence using the augmented Lagrangian method along with nonlinear conjugate gradients.

2 Related Work

Cloth simulation. Cloth simulation has been a major topic of computer graphics research for over two decades. Many successful techniques have been proposed for modeling the dynamics of cloth and for handling collisions [Baraff and Witkin 1998; Bridson et al. 2002; Choi and Ko 2002; Bridson et al. 2003]. Summaries

Contact email: {narain, samii, job}@eecs.berkeley.edu

ACM Reference Format
Narain, R., Samii, A., O'Brien, J. 2012. Adaptive Anisotropic Remeshing for Cloth Simulation. *ACM Trans. Graph. 31* 6, Article 152 (November 2012), 10 pages. DOI = 10.1145/2366145.2366171 http://doi.acm.org/10.1145/2366145.2366171.

Motion Graphs++:
a Compact Generative Model for Semantic Motion Analysis and Synthesis

Jianyuan Min
Texas A&M University

Jinxiang Chai*
Texas A&M University

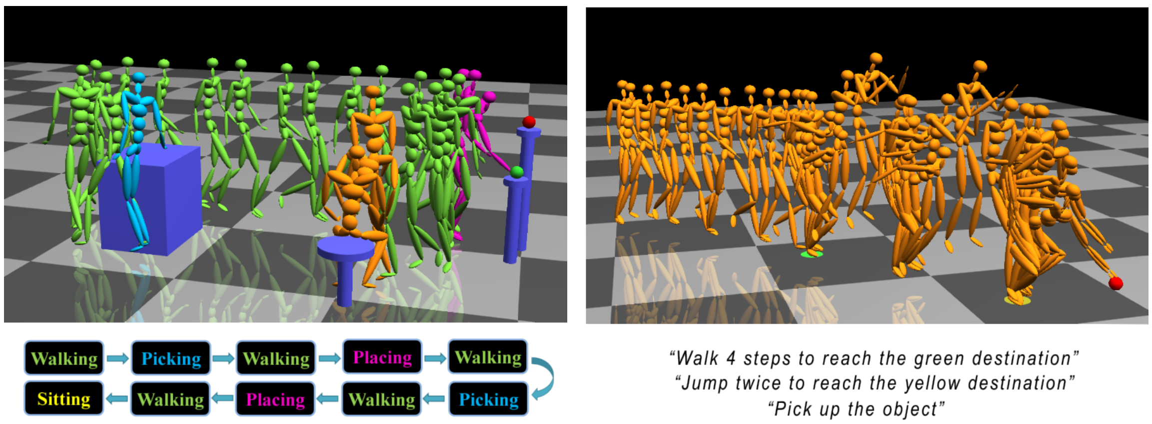


Figure 1: *Semantic motion analysis (left) and synthesis (right) with our generative statistical model.*

Abstract

This paper introduces a new generative statistical model that allows for human motion analysis and synthesis at both semantic and kinematic levels. Our key idea is to decouple complex variations of human movements into *finite structural variations* and *continuous style variations* and encode them with a concatenation of morphable functional models. This allows us to model not only a rich repertoire of behaviors but also an infinite number of style variations within the same action. Our models are appealing for motion analysis and synthesis because they are *highly structured*, *contact aware*, and *semantic embedding*. We have constructed a compact generative motion model from a huge and heterogeneous motion database (about two hours mocap data and more than 15 different actions). We have demonstrated the power and effectiveness of our models by exploring a wide variety of applications, ranging from automatic motion segmentation, recognition, and annotation, and online/offline motion synthesis at both kinematics and behavior levels to semantic motion editing. We show the superiority of our model by comparing it with alternative methods.

Keywords: character animation, generative motion models, semantic motion analysis and synthesis, motion planning

Links: DL PDF

*e-mail: jianyuan|jchai@cs.tamu.edu

ACM Reference Format
Min, J., Chai, J. 2012. Motion Graphs++: a Compact Generative Model for Semantic Motion Analysis and Synthesis. *ACM Trans. Graph. 31* 6, Article 153 (November 2012), 12 pages.
DOI = 10.1145/2366145.2366172 http://doi.acm.org/10.1145/2366145.2366172.

1 Introduction

This paper focuses on constructing a generative motion model to create a rich repertoire of behaviors for virtual humans. Thus far, one of the most successful solutions to this problem is to build *generative statistical models* from prerecorded motion data. Generative statistical models are appealing for motion analysis and synthesis because they are compact, they have strong generalization ability to create motions that are not in prerecorded motion data, and they can generate an infinite number of motion variations with a small number of hidden variables. Despite the progress made over the last decade, creating appropriate generative statistical models for human motion synthesis and control remains challenging for a number of key reasons.

Scalability. A responsive lifelike human character must possess a rich repertoire of activities and display a wide range of styles within the same action. This inevitably requires generative statistical models to scale up well to huge and heterogeneous motion datasets. However, current generative statistical models (e.g., [Chai and Hodgins 2007; Lau et al. 2009; Min et al. 2009]) are often behavior-specific and focus on modeling detailed style variations within the same action. They have not demonstrated that they can model a rich repertoire of behaviors and a wide range of style variations at the same time.

Semantic embedding. Current approaches often fail to encode semantic information into generative models. For example, they cannot identify when and where to pick up an object and cannot count how many walking steps are in output animation. This prohibits the user from creating and controlling an animation at the semantic level (e.g., "walk five steps and pick up the object"). However, for many applications, users are more interested in which actions to perform than how to animate the character at the kinematics level.

Contact awareness. Current generative models are not contact aware because they encode little or no information about environmental contact information such as "left foot on ground." As a result, generated motions often violate environmental contact con-

Terrain Runner: Control, Parameterization, Composition, and Planning for Highly Dynamic Motions

Libin Liu* KangKang Yin† Michiel van de Panne‡ Baining Guo§

* Tsinghua University † National University of Singapore ‡ University of British Columbia § Microsoft Research Asia

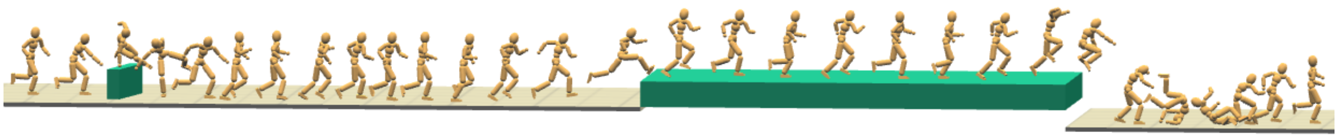


Figure 1: *An animated character runs, vaults, jumps, and drop-rolls across a parkour terrain during a real-time physics-based simulation. Given a single motion capture clip of each of these four skills as input, our method uses an offline process to develop robust control policies for parameterized versions of these skills, as well as robust transition motions.*

Abstract

In this paper we learn the skills required by real-time physics-based avatars to perform parkour-style fast terrain crossing using a mix of running, jumping, speed-vaulting, and drop-rolling. We begin with a single motion capture example of each skill and then learn reduced-order linear feedback control laws that provide robust execution of the motions during forward dynamic simulation. We then parameterize each skill with respect to the environment, such as the height of obstacles, or with respect to the task parameters, such as running speed and direction. We employ a continuation process to achieve the required parameterization of the motions and their affine feedback laws. The continuation method uses a predictor-corrector method based on radial basis functions. Lastly, we build control laws specific to the sequential composition of different skills, so that the simulated character can robustly transition to obstacle clearing maneuvers from running whenever obstacles are encountered. The learned transition skills work in tandem with a simple online step-based planning algorithm, and together they robustly guide the character to achieve a state that is well-suited for the chosen obstacle-clearing motion.

Keywords: physics-based animation, motion control, parkour

Links: DL PDF

1 Introduction

The physics-based animation of human-like characters has seen many significant advances, especially for locomotion tasks. However, the abilities and overall agility of the proposed control methods still fall far short of many human motions, such as that exhibited by a freerunner performing parkour skills [Edwardes 2009]. Parkour is a physical discipline which focuses on efficient movement around obstacles, and is considered as a sport, an art, or sometimes even a philosophy. The control, agility, expressivity, and versatility that human athletes demonstrate in highly dynamic motions such as parkour makes them mesmerizing to watch.

Reproducing the abilities of a high-level parkour athlete can be considered a "grand challenge" problem for character animation for several reasons. First, reference motions are difficult to capture. Capturing highly dynamic motions such as those used in parkour requires large space, special equipment, and trained athletes, in contrast to easy-to-capture motions such as walking. We thus desire a method that requires only a sparse set of motion capture examples. Second, designing robust motion controllers usually requires domain knowledge and human insight. However, the required insights are hard to obtain for unfamiliar and highly dynamic motions. A largely-automated method that is broadly applicable to a variety of motions is thus desirable. Third, the sequential composition of highly dynamic skills is difficult to develop using existing approaches for physics-based control. The end state resulting from one controller can easily fall outside the basin of attraction of the next controller, and the physical barriers in terrain-crossing tasks can appear at any distance with arbitrary height.

We present the first attempt at achieving the real-time physics-based simulation of an integrated subset of parkour skills, including running, jumping, speed vaulting, and drop-rolling. While this set of skills falls far short of a complete repertoire of parkour skills, we believe that it provides an important proof-of-concept that integrated sets of parameterized skills can be designed for highly dynamic motions. We present three primary contributions:

- Parameterization: Given a single motion-capture example clip for each skill, we develop closed-loop feedback controllers for parameterized versions of the motion. This entails learning *parameterized* low-dimensional feedback strategies and developing a multidimensional continuation method.
- Sequential Composition: We introduce a structured composition scheme for transitioning between running and the obstacle clearing maneuvers. The composition scheme is coupled to both the offline parameterization and the online planning. In this way the running steps used to approach an obstacle can be adjusted in the best way to achieve more realistic and

*e-mail:llb05@mails.tsinghua.edu.cn
†e-mail:kkyin@comp.nus.edu.sg
‡e-mail:van@cs.ubc.ca
§e-mail:bainguo@microsoft.com

ACM Reference Format
Liu, L., Yin, K., van de Panne, M., Guo, B. 2012. Terrain Runner: Control, Parametrization, Composition, and Planning for Highly Dynamic Motions. *ACM Trans. Graph. 31* 6, Article 154 (November 2012), 9 pages. DOI = 10.1145/2366145.2366173 http://doi.acm.org/10.1145/2366145.2366173.

Falling and Landing Motion Control for Character Animation

Sehoon Ha Yuting Ye C. Karen Liu
Georgia Institute of Technology

Figure 1: *A simulated character lands on the roof of a car, leaps forward, dive-rolls on the sidewalk, and gets back on its feet, all in one continuous motion.*

Abstract

We introduce a new method to generate agile and natural human landing motions in real-time via physical simulation without using any mocap or pre-scripted sequences. We develop a general controller that allows the character to fall from a wide range of heights and initial speeds, continuously roll on the ground, and get back on its feet, without inducing large stress on joints at any moment. The character's motion is generated through a forward simulator and a control algorithm that consists of an airborne phase and a landing phase. During the airborne phase, the character optimizes its moment of inertia to meet the ideal relation between the landing velocity and the angle of attack, under the laws of conservation of momentum. The landing phase can be divided into three stages: impact, rolling, and getting-up. To reduce joint stress at landing, the character leverages contact forces to control linear momentum and angular momentum, resulting in a rolling motion which distributes impact over multiple body parts. We demonstrate that our control algorithm can be applied to a variety of initial conditions with different falling heights, orientations, and linear and angular velocities. Simulated results show that our algorithm can effectively create realistic action sequences comparable to real world footage of experienced freerunners.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: Character Animation, Physics-based Animation, Optimal Control

Links: DL PDF

ACM Reference Format
Ha, S., Ye, Y., Liu, C. 2012. Falling and Landing Motion Control for Character Animation.
ACM Trans. Graph. 31 6, Article 155 (November 2012), 9 pages. DOI = 10.1145/2366145.2366174
http://doi.acm.org/10.1145/2366145.2366174.

1 Introduction

One of the great challenges in computer animation is to physically simulate a virtual character performing highly dynamic motion with agility and grace. A wide variety of athletic movements, such as acrobatics or freerunning (parkour), involve frequent transitions between airborne and ground-contact phases. How to land properly to break a fall is therefore a fundamental skill athletes must acquire. A successful landing should minimize the risk of injury and disruption of momentum because the quality of performance largely depends on the athlete's ability to safely absorb the shock at landing, while maintaining readiness for the next action. To achieve a successful landing, the athlete must plan coordinated movements in the air, control contacting body parts at landing, and execute fluid follow-through motion. The basic building blocks of these motor skills can be widely used in other sports that involve controlled falling and rolling, such as diving, gymnastics, judo, or wrestling.

We introduce a new method to generate agile and natural human falling and landing motions in real-time via physical simulation without using motion capture data or pre-scripted animation (Figure 1). We develop a general controller that allows the character to fall from a wide range of heights and initial speeds, continuously roll on the ground, and get back on its feet, without inducing large stress on joints at any moment. Previous controllers for acrobat-like motions either precisely define the sequence of actions and contact states in a state-machine structure, or directly track a specific motion capture sequence. Both cases fall short of creating a generic controller capable of handling a wide variety of initial conditions, overcoming drastic perturbations in runtime, and exploiting unpredictable contacts.

Our method is inspired by three landing principles informally developed in freerunning community. First, reaching the ground with flexible arms or legs provides cushion time to dissipate energy over a longer time window rather than absorbing it instantly at impact. It also protects the important and fragile body parts, such as the head, the pelvis, and the tailbone. Second, it is advisable to distribute the landing impact over multiple body parts to reduce stress on any particular joint. Third, it is crucial to utilize the friction force generated by landing impact to steer the forward direction and control the angular momentum for rolling, a technique referred to as "blocking" in the freerunning community. These three principles outline the most commonly employed landing strategy in practice: landing with feet or hands as the first point of contact, gradually lowering the center of mass (COM) to absorb vertical impact, and

Synthesis of Concurrent Object Manipulation Tasks

Yunfei Bai* Kristin Siu† C. Karen Liu‡
Georgia Institute of Technology

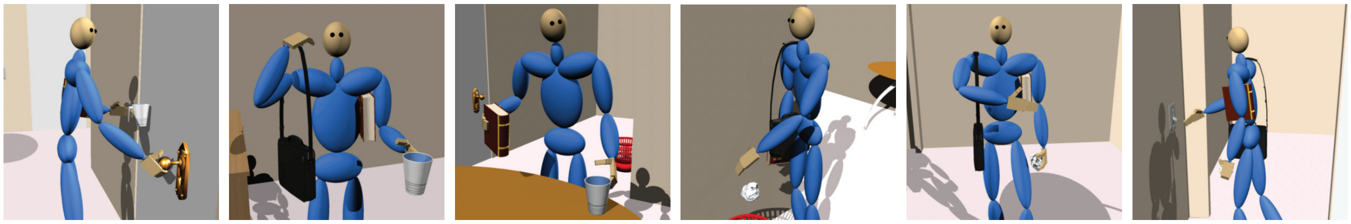


Figure 1: *A simulated character manipulating multiple objects concurrently in different scenarios.*

Abstract

We introduce a physics-based method to synthesize concurrent object manipulation using a variety of manipulation strategies provided by different body parts, such as grasping objects with the hands, carrying objects on the shoulders, or pushing objects with the elbows or the torso. We design dynamic controllers to physically simulate upper-body manipulation and integrate it with procedurally generated locomotion and hand grasping motion. The output of the algorithm is a continuous animation of the character manipulating multiple objects and environment features concurrently at various locations in a constrained environment. To capture how humans deftly exploit different properties of body parts and objects for multitasking, we need to solve challenging planning and execution problems. We introduce a graph structure, a *manipulation graph*, to describe how each object can be manipulated using different strategies. The problem of manipulation planning can then be transformed to a standard graph traversal. To achieve the manipulation plan, our control algorithm optimally schedules and executes multiple tasks based on the dynamic space of the tasks and the state of the character. We introduce a "task consistency" metric to measure the physical feasibility of multitasking. Furthermore, we exploit the redundancy of control space to improve the character's ability to multitask. As a result, the character will try its best to achieve the current tasks while adjusting its motion continuously to improve the multitasking consistency for future tasks.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: human simulation; physically based animation; motion planning; optimization; physically based animation

*e-mail: ybai30@mail.gatech.edu
†e-mail:kasiu@gatech.edu
‡e-mail:karenliu@cc.gatech.edu

Links: DL PDF

1 Introduction

Performing multiple object manipulation tasks concurrently is an essential human activity in everyday environments. A mundane morning routine before going to work can involve numerous consecutive and concurrent tasks: picking up the briefcase on the floor, opening the refrigerator to fetch a lunch box, using the elbow to close the refrigerator door, tucking the lunch box under the arm so the hand can search for keys in the pocket, and pushing the front door open by leaning on it. This sequence of tasks, which humans perform effortlessly, requires sophisticated planning and dynamic motion control, which have not been broadly explored in physics-based computer animation or robotics. Unlike existing robots, humans can employ a variety of *manipulation strategies* to interact with objects, such as using their hands, shoulders, elbows, torso, or even their head. Consequently, synthesizing full-body manipulation requires not only simulating physically realistic joint motion, but also capturing how humans deftly exploit different properties of body parts and objects for multitasking.

We introduce a physics-based technique to synthesize human activities involving concurrent full-body manipulation of multiple objects. We view full-body manipulation as three interrelated layers of motor control: locomotion, upper-body manipulation, and detailed hand manipulation. This paper focuses on the second layer – we design dynamic controllers to physically simulate upper-body manipulation and integrate it with procedurally generated locomotion and hand manipulation. The main algorithm must overcome major challenges in both planning and execution.

Planning a valid sequence of manipulation strategies for a character is difficult because humans have abundant choices for manipulating an object. To circumvent this issue, our key insight is that, instead

ACM Reference Format
Bai, Y., Siu, K., Liu, C. 2012. Synthesis of Concurrent Object Manipulation Tasks.
ACM Trans. Graph. 31 6, Article 156 (November 2012), 9 pages. DOI = 10.1145/2366145.2366175
http://doi.acm.org/10.1145/2366145.2366175.

Sculpting by Numbers

Alec Rivers
MIT CSAIL

Andrew Adams
MIT CSAIL

Frédo Durand
MIT CSAIL

(a) Target 3D model (b) Guidance projected onto material (c) Sculpted physical replica

Figure 1: *We assist users in creating physical objects that match digital 3D models. Given a target 3D model* **(a)***, we project different forms of guidance onto a work in progress* **(b)** *that indicate how it must be deformed to match the target model. As the user follows this guidance, the physical object's shape approaches that of the target* **(c)***. With our system, unskilled users are able to produce accurate physical replicas of complex 3D models. Here, we recreate the Stanford bunny model (courtesy of the Stanford Computer Graphics Laboratory) out of polymer clay.*

Abstract

We propose a method that allows an unskilled user to create an accurate physical replica of a digital 3D model. We use a projector/camera pair to scan a work in progress, and project multiple forms of guidance onto the object itself that indicate which areas need more material, which need less, and where any ridges, valleys or depth discontinuities are. The user adjusts the model using the guidance and iterates, making the shape of the physical object approach that of the target 3D model over time. We show how this approach can be used to create a duplicate of an existing object, by scanning the object and using that scan as the target shape. The user is free to make the reproduction at a different scale and out of different materials: we turn a toy car into cake. We extend the technique to support replicating a sequence of models to create stop-motion video. We demonstrate an end-to-end system in which real-world performance capture data is retargeted to claymation. Our approach allows users to easily and accurately create complex shapes, and naturally supports a large range of materials and model sizes.

Keywords: personal digital fabrication, spatially augmented reality, sculpting

Links: DL PDF

ACM Reference Format
Rivers, A., Adams, A., Durand, F. 2012. Sculpting by Numbers. *ACM Trans. Graph. 31* 6, Article 157 (November 2012), 7 pages. DOI = 10.1145/2366145.2366176 http://doi.acm.org/10.1145/2366145.2366176.

1 Introduction

Most people find it challenging to sculpt, carve or manually form a precise shape. We argue that this is usually not because they lack manual dexterity – the average person is able to perform very precise manipulations – but rather because they lack precise 3D information, and cannot figure out what needs to be done to modify a work in progress in order to reach a goal shape. An analogy can be made to the task of reproducing a 2D painting: when given outlines that need only be filled in, as in a child's coloring book or a paint-by-numbers kit, even an unskilled user can accurately reproduce a complex painting; the challenge lies not in placing paint on the canvas but in knowing where to place it. Motivated by this observation, we present Sculpting by Numbers, a method to provide analogous guidance for the creation of 3D objects, which assists a user in making an object that precisely matches the shape of a target 3D model.

We employ a spatially-augmented reality approach (see e.g. Raskar et al. [1998] or Bimber and Raskar [2005] for an overview of spatially-augmented reality), in which visual feedback illustrates the discrepancy between a work in progress and a target 3D shape. This approach was first proposed by Skeels and Rehg [2007]. In this approach, a projector-camera pair is used to scan the object being created using structured light. The scanned shape is compared with the target 3D model, and the projector then annotates the object with colors that indicate how the object ought to be changed to match the target. The user follows this guidance to adjust the object and rescans when necessary, bringing the object closer to the target shape over time.

Our proposed method provides guidance that illustrates depth disparities between the current work and the target, similar to the approach of Skeels and Rehg [2007], as well as an additional form of guidance, which we call edge guidance, which aids in reproducing high-frequency surface details. We demonstrate an application to reproducing an existing physical object at a different scale or

Stackabilization

Honghua Li[1,2] Ibraheem Alhashim[1] Hao Zhang[1] Ariel Shamir[3] Daniel Cohen-Or[4]
[1]Simon Fraser University * [2]National University of Defense Technology [3]Interdisciplinary Center [4]Tel Aviv University

Abstract

We introduce the geometric problem of *stackabilization*: how to geometrically modify a 3D object so that it is more amenable to stacking. Given a 3D object and a stacking direction, we define a measure of stackability, which is derived from the gap between the lower and upper envelopes of the object in a stacking configuration along the stacking direction. The main challenge in stackabilization lies in the desire to modify the object's geometry only subtly so that the intended functionality and aesthetic appearance of the original object are not significantly affected. We present an automatic algorithm to deform a 3D object to meet a target stackability score using energy minimization. The optimized energy accounts for both the scales of the deformation parameters as well as the preservation of pre-existing geometric and structural properties in the object, e.g., symmetry, as a means of maintaining its functionality. We also present an intelligent editing tool that assists a modeler when modifying a given 3D object to improve its stackability. Finally, we explore a few fun variations of the stackabilization problem.

Keywords: stackability, stackabilization, shape optimization, structure-preserving deformation

Links: DL PDF WEB

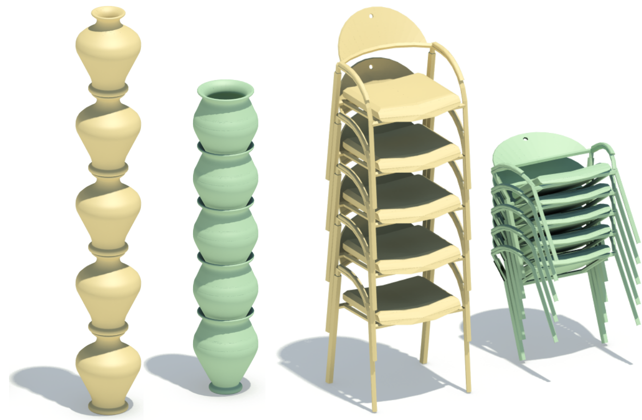


Figure 1: *Stackabilization: improving the stackability of 3D objects (yellow denotes input and green denotes output) by analyzing and optimizing geometry. The object geometry is only altered subtly so as to maintain the intended functionality and aesthetic appearance of the original object.*

1 Introduction

Stacking objects on top of each other is a common task performed by humans. Objects are often stacked to make or save space, for instance, while shipping or storing them. In product design and engineering, efforts have been invested to allow compact stacking without compromising the product's intended functionality. One of the most celebrated examples of stackable objects are chairs, where many space-saving and aesthetic designs of stackable chairs have been realized [Fiell and Fiell 2000].

An intriguing geometric question about stacking is: what makes some 3D objects more amenable to stacking than others. It is clear that concavity plays an important role for simple shapes such as those of bowls or cups. However, to precisely define and improve the *stackability* of a 3D shape turns into an interesting and challenging geometry problem, especially for shapes with more complex structures, such as tables and chairs.

In this paper, we are interested first in a geometric characterization of stackability for 3D shapes and second, in the development of algorithms to make a given 3D object more stackable. We use a new term for such an operation — *stackabilization*. The main challenge in stackabilization lies in the desire to only modify the shape's geometry and structure in subtle ways so as to maximize its stackability, while not adversely affecting its shape, functionality and aesthetic appearance. Understanding stackability and developing tools for stackabilization will assist designers in creating 3D models that can be packed together more compactly. Figure 1 shows the before-and-after of stackabilization of two shapes and illustrates the saving in space achieved.

The need for a specific tool, automatic or semi-automatic, and algorithms to assist a designer or modeler in solving stackabilization is emphasized by the global and unpredictable nature of the problem. As many examples in this paper illustrate, a small modification to the shape can have a great effect on the stackability of objects, while large modifications can have little to no effect at all. The reason is that modifying the shape on one side affects the contact with the opposite side. This problem is amplified when the shapes become more complex with certain parts cluttered, occluded, or inaccessible. Hence, the key problem for designers, even when manually modifying the shape, is to recognize which modifications are better and where to apply them in order to have the largest effect on stackability and the smallest effect on the design. This is exactly the strength of our automatic stackability analysis and the corresponding stackabilization algorithm.

Stackability. Geometrically, stacking an object on top of a copy of itself along a stacking direction can be achieved by translating one object copy along the stacking direction, just until the the two copies have no overlap, i.e., they are just touching each other. We call such a configuration the *stacking configuration* and the surface regions where the two copies maintain contact the *contact regions*. The extent of the minimal stacking translation can be found by looking at the maximum gap between the *upper* and *lower envelopes* of the object, and the contact regions are the regions inside these envelopes where this maximum gap is achieved; see Figure 2.

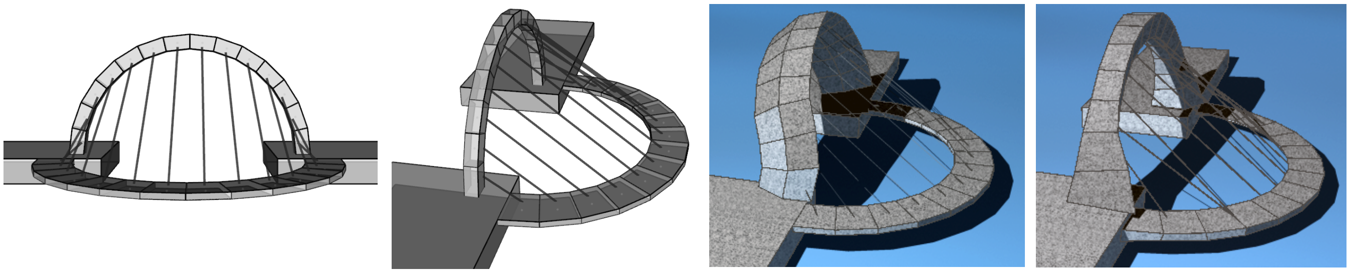
*e-mail: {honghual, iaa7, haoz}@sfu.ca

ACM Reference Format
Li, H., Alhashim, I., Zhang, H., Shamir, A., Cohen-Or, D. 2012. Stackabilization.
ACM Trans. Graph. 31 6, Article 158 (November 2012), 9 pages. DOI = 10.1145/2366145.2366177
http://doi.acm.org/10.1145/2366145.2366177.

Structural Optimization of 3D Masonry Buildings

Emily Whiting Hijung Shin Robert Wang John Ochsendorf Frédo Durand

Massachusetts Institute of Technology



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

Figure 1: *We present a method to compute the gradient for the stability of a structure composed of rigid blocks, and demonstrate how we enable the optimization of stable structures. For example: (a) A cable bridge structure originally infeasible. (b) Side view of the input model. (c) Output feasible model. The horizontal arched walkway and cables are fixed, only the vertical arch is optimized. (d) An alternative feasible output. The horizontal arch and cable joints are free to deform with the constraint that top faces (walking surface) remain horizontal.*

Abstract

In the design of buildings, structural analysis is traditionally performed after the aesthetic design has been determined and has little influence on the overall form. In contrast, this paper presents an approach to guide the form towards a shape that is more structurally sound. Our work is centered on the study of how variations of the geometry might improve structural stability. We define a new measure of structural soundness for masonry buildings as well as cables, and derive its closed-form derivative with respect to the displacement of all the vertices describing the geometry. We start with a gradient descent tool which displaces each vertex along the gradient. We then introduce displacement operators, imposing constraints such as the preservation of orientation or thickness; or setting additional objectives such as volume minimization.

Keywords: Statics, structural stability, architecture, optimization

Links: DL PDF

1 Introduction

While computer graphics and computer-aided-design (CAD) have dramatically broadened the range of shapes available for architectural design, structural considerations have often been ignored. Structural analysis of a building is usually performed after the aesthetic design has been determined and has little influence on the overall form. An architect designs the shape, which is passed to structural engineers to make the building stable through the use of appropriate material and reinforcement. Existing structural analysis software, such as finite element analysis, is a powerful method for analyzing a given structure, but does not directly suggest ways to improve the geometry in order to reduce internal forces and required material. In contrast, we seek to propose modifications to the geometry that enhance structural soundness. We focus on masonry structures because their stability is the direct result of their geometry, but we argue that the central principle of sound masonry design – minimization of non-axial forces – extends to other materials.

The input to our method is a building geometry described as a set of blocks specified by their vertex coordinates. The central component of our approach is the notion of a structural gradient, which expresses, for each vertex, the displacement direction that maximally improves structural soundness. The gradients can be used in a steepest-descent manner. Alternatively, constraints can be introduced to modify the gradient direction, such as preservation of horizontal and vertical directions, or constant thickness of blocks. Objectives can also be added such as volume minimization to reduce material usage. We explore a number of gradient modifications and show that they enable variations in structurally sound models.

We base our notion of structural soundness on static analysis [Livesley 1978; Livesley 1992] and focus on masonry materials, comprising stone and brick structures. Masonry is the dominant material for traditional architecture and is also used in modern architecture, especially in developing countries. In contrast to contemporary steel or reinforced concrete, traditional masonry relies on forms which are inherently stable, because the material resists only axial compressive forces [Allen and Zalewski 2009]. Though we focus on the case of masonry, our approach can be used to minimize non-axial forces in general. Even with materials that resist tension, such as reinforced concrete or steel, a good structural form with reduced non-axial force requires less material, leading to cheaper, more environmentally-friendly, and robust buildings. In addition, we extend our approach to enable the treatment of cables as tension-only elements, using the same principles of static analysis and resistance to axial forces.

The heart of our approach is to compute the gradient of a stability metric with respect to geometry modification. First, we show that previous expressions of masonry instability [Whiting et al. 2009] do

ACM Reference Format
Whiting, E., Shin, H., Wang, R., Ochsendorf, J., Durand, F. 2012. Structural Optimization of 3D Masonry Buildings. *ACM Trans. Graph. 31* 6, Article 159 (November 2012), 11 pages.
DOI = 10.1145/2366145.2366178 http://doi.acm.org/10.1145/2366145.2366178.

Depth-Presorted Triangle Lists

Ge Chen[1] Pedro V. Sander[1] Diego Nehab[2] Lei Yang[1,3] Liang Hu[4]
[1]Hong Kong UST [2]IMPA [3]Bosch Research North America [4]Google Inc.

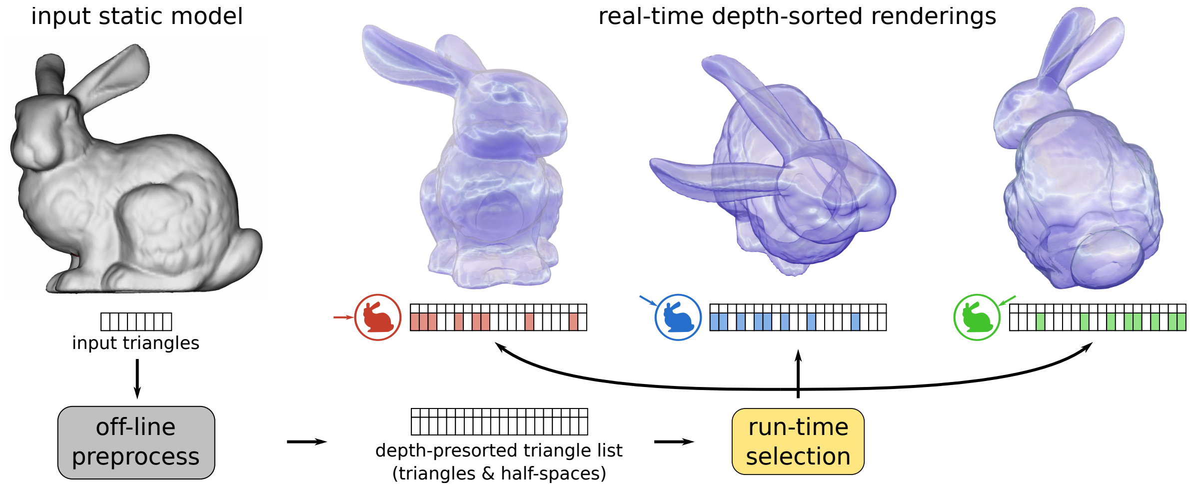


Figure 1: *Our off-line preprocessing algorithm produces a depth-presorted triangle list from a static 3D input model. The list contains copies of the model that are depth-sorted relative to* any *viewpoint outside its bounding volume. Each triangle comes annotated with a half-space test that enables a fast run-time algorithm to select exactly the triangles needed to render a complete depth-sorted model for a given viewpoint.*

Abstract

We present a novel approach for real-time rendering of static 3D models front-to-back or back-to-front relative to any viewpoint outside its bounding volume. The approach renders depth-sorted triangles using a single draw-call. At run-time, we replace the traditional *sorting* strategy of existing algorithms with a faster triangle *selection* strategy. The selection process operates on an extended sequence of triangles annotated by test planes, created by our off-line preprocessing stage. Based on these test planes, a simple run-time procedure uses the given viewpoint to select a subsequence of triangles for rasterization. Selected subsequences are statically presorted by depth and contain each input triangle exactly once. Our method runs on legacy hardware and renders depth-sorted static models significantly faster than previous approaches. We conclude demonstrating the real-time rendering of order-independent transparency effects.

Links: DL PDF

1 Introduction

In real-time rendering applications that employ the Z-buffer for visibility determination [Catmull 1974], there are still many scenarios in which depth-sorting is necessary or desirable. The most common is order-independent transparency or translucency. Since the compositing operation is not commutative [Porter and Duff 1984], blending must happen in depth-sorted order.

As we discuss in section 2, a large number of techniques have been proposed for performing real-time depth sorting. In this work, we present a technique that possesses a unique combination of desirable properties. It can be implemented with the standard graphics pipeline, requires a single rendering pass, uses a fixed amount of memory, produces exact results, is very simple to integrate with existing rendering engines, and is extremely efficient.

These advantages come with certain limitations. Whereas most previous depth-sorting algorithms work seamlessly with deformable geometry at run-time, our method assumes static geometry viewed from outside the model's bounding volume. Finally, our preprocessing stage can take hours to complete when run on larger models, and the resulting data-structure consumes more memory than the input.

We target performance-critical applications that must render a number of moderately complex static objects with translucency effects, such as computer games. In this scenario, which we demonstrate in the results section, the relative order between objects is determined by the CPU, and our method ensures correct triangle ordering within each object. During game development, instant feedback can be provided to artists using earlier, less efficient methods. At the end of the release cycle, required models can be preprocessed and the engine set up to take advantage of the simplicity of our run-time component, and of the large performance gains that ensue. Since only the transparent components of objects with translucency effects must be preprocessed, the increase in run-time memory is not a significant limitation either.

Our key insight is that the space of different triangle orders that result from depth-sorting a triangle model under each possible viewpoint constitutes a tiny fraction of all triangle permutations. Moreover, this "space of depth-sorted orders" is extremely redundant in the sense that, with few modifications, the same order is valid for large portions of the viewpoint space.

ACM Reference Format
Chen, G., Sander, P., Nehab, D., Yang, L., Hu, L. 2012. Depth-Presorted Triangle Lists.
ACM Trans. Graph. 31 6, Article 160 (November 2012), 9 pages. DOI = 10.1145/2366145.2366179
http://doi.acm.org/10.1145/2366145.2366179.

Softshell: Dynamic Scheduling on GPUs

Markus Steinberger* Bernhard Kainz* Bernhard Kerbl Stefan Hauswiesner* Michael Kenzel Dieter Schmalstieg*
Graz University of Technology, Austria

Abstract

In this paper we present Softshell, a novel execution model for devices composed of multiple processing cores operating in a single instruction, multiple data fashion, such as graphics processing units (GPUs). The Softshell model is intuitive and more flexible than the kernel-based adaption of the stream processing model, which is currently the dominant model for general purpose GPU computation. Using the Softshell model, algorithms with a relatively low local degree of parallelism can execute efficiently on massively parallel architectures. Softshell has the following distinct advantages: **(1)** work can be dynamically issued directly on the device, eliminating the need for synchronization with an external source, *i.e.*, the CPU; **(2)** its three-tier dynamic scheduler supports arbitrary scheduling strategies, including dynamic priorities and real-time scheduling; and **(3)** the user can influence, pause, and cancel work already submitted for parallel execution. The Softshell processing model thus brings capabilities to GPU architectures that were previously only known from operating-system designs and reserved for CPU programming. As a proof of our claims, we present a publicly available implementation of the Softshell processing model realized on top of CUDA. The benchmarks of this implementation demonstrate that our processing model is easy to use and also performs substantially better than the state-of-the-art kernel-based processing model for problems that have been difficult to parallelize in the past.

CR Categories: I.3.6 [Computer Graphics]: Methodology and Techniques—Languages; I.3.1 [Computer Graphics]: Hardware Architecture—Parallel processing;

Keywords: priority scheduling, GPU, priority work queue, real-time scheduling, persistent threads, dynamic parallelism

Links: DL PDF WEB CODE

1 Introduction

Over the last decade, parallel computing has become increasingly available to a wide audience, largely due to graphics processing units (GPUs), which have evolved into massively-parallel general-purpose co-processors. GPU hardware is evolving rapidly, eliminating the drawbacks of previous designs with the introduction of a multi-level cache or stack [NVIDIA 2009]. On the software side, new programming languages such as CUDA have advanced, but the processing model itself has mostly remained untouched since the beginning of general-purpose GPU programming. This model inherently has several limitations:

*e-mail: {steinberger, kainz, hauswiesner, schmalstieg}@icg.tugraz.at

ACM Reference Format
Steinberger, M., Kainz, B., Kerbl, B., Hauswiesner, S., Kenzel, M., Schmalstieg, D. 2012.
Softshell: Dynamic Scheduling on GPUs. *ACM Trans. Graph. 31* 6, Article 161 (November 2012), 11 pages.
DOI = 10.1145/2366145.2366180 http://doi.acm.org/10.1145/2366145.2366180.

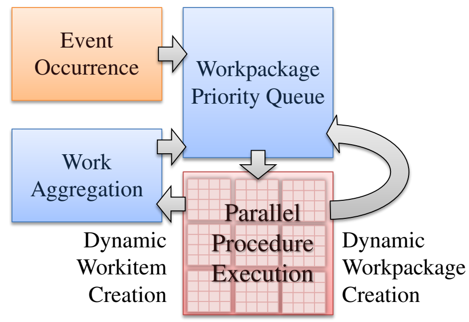


Figure 1: *The Softshell processing model: Due to an event, new work becomes available for execution on a GPU. Work items are grouped in workpackages, waiting for their parallel execution in a priority queue. When a SIMD processing unit becomes available, it draws the highest priority workpackage from the queue and runs the procedure associated with it. During execution, new work can dynamically be created on the GPU, either in form of entire workpackages or as single work items to be automatically aggregated. In this way, algorithms with various degrees of parallelism can efficiently be mapped to SIMD architectures.*

The first limitation is that sufficient data parallelism is required in every stage of an algorithm. These stages are captured by *kernel* function calls. A fixed number of threads are launched for a single kernel and start executing the same code. There is no way to dynamically adjust the parallelism during a single kernel launch. While from a hardware perspective it would be sufficient if a high number of coherently executing thread groups are available, good performance using *kernel* functions can only be achieved, if kernels are started for thousands of threads [NVIDIA 2011; Khronos 2008]. This rigid requirement often impairs the straight forward mapping of common algorithms for GPU execution. Such a problem is, for example, traversing a tree and executing an operation for every node. One way of parallelizing these classes of algorithms is to subsequently launch a kernel for each tree level. For non-trivial applications, the local tree depth may strongly vary, resulting in underutilization if there are not enough tree nodes available. Additionally, determining the number of nodes per level and mapping them to threads requires synchronization after each level and inefficient parallel reduction methods. The easier and more efficient solution would be to launch new threads for every child node dynamically. Many other problems in graphics show a similar dependency between control flow and parallelism.

The second limitation is due to the fact that kernel launches are entirely controlled by the CPU. Therefore, GPU/CPU synchronization between kernel launches is necessary, if decisions concerning the next kernel launch must be made. The overhead of passing control back and forth between the GPU and CPU can be omitted if the execution of new tasks could be initiated on the GPU itself. Additionally, there is currently no way to interrupt or terminate the execution of a running kernel. Interrupting events can only be considered after a kernel has finished. However, if a user changes input parameters – for example in an interactive visualization – the current kernel launch may become obsolete and any further computations become useless.

High-Quality Curve Rendering using Line Sampled Visibility

Rasmus Barringer
Lund University

Carl Johan Gribel
Lund University

Tomas Akenine-Möller
Lund University and Intel Corporation

Figure 1: *Our novel thin curve rendering algorithm used on a test production model to compute accurate visibility. The model has 32,000 unique hair strands, which consists of over one million Bézier curves with varying thickness. As can be seen, our algorithm works at all different scales, from cases where there are hundreds of hair strands per pixel to zooming in on the hair strands. All images were rendered at* 1024×1024 *pixels with our GPU implementation. The leftmost image took 109 ms to render, while the close-up on the face took 468 ms. The rightmost image showcases our ability to handle thick curves. Hair model courtesy of Weta Digital.*

Abstract

Computing accurate visibility for thin primitives, such as hair strands, fur, grass, at all scales remains difficult or expensive. To that end, we present an efficient visibility algorithm based on spatial line sampling, and a novel intersection algorithm between line sample planes and Bézier splines with varying thickness. Our algorithm produces accurate visibility both when the projected width of the curve is a tiny fraction of a pixel, and when the projected width is tens of pixels. In addition, we present a rapid resolve procedure that computes final visibility. Using an optimized implementation running on graphics processors, we can render tens of thousands long hair strands with noise-free visibility at near-interactive rates.

CR Categories: I.3.3 [Picture/Image Generation]: Antialiasing; I.3.7 [Three-Dimensional Graphics and Realism]: Color, shading, shadowing, and texture;

Keywords: analytical visibility, anti-aliasing, curve rendering

Links: DL PDF

1 Introduction

High quality rendering of thin, curved primitives, e.g., hair, fibers, fur, and grass, is an important ingredient in today's computer generated imagery. This is particularly true for offline rendering for feature films, but also increasingly so for real-time rendering in games. One approach is to model such thin curves as ribbons with varying width, e.g., using RenderMan's `riCurves` primitive, and then sample visibility using point sampling. A similar modeling and rendering technique was used by Marschner et al. [2003] when developing an accurate appearance model for hair. Another common approach is to rasterize lines with alpha blending to simulate line widths smaller than one pixel [Leblanc et al. 1991; Sintorn and Assarsson 2008]. A third approach is to model and render hair with volumetric textures using ray marching [Kajiya and Kay 1989].

While shading for some types of thin primitives, in particular hair [Marschner et al. 2003; Moon et al. 2008; Zinke et al. 2008; Zinke 2008; Hery and Ramamoorthi 2012], is well understood, computing accurate visibility rapidly for a large number of curves remains a challenge. A major problem is that when point sampling is used, noise is inevitable unless a very large number of samples per pixel is used. This is especially true at a macro-scale, when the viewer is relatively far away from the curves, and the projected width of the curve is only, say, 10% or less, of the pixel width. In such cases, hundreds of samples per pixel may be needed for accurate visibility. For comparison, the diameter of a hair strand is about 0.1 mm [Hadap et al. 2007]. Another problem is that the ribbon model breaks down at the microscale, i.e., when a curve's width project to relatively large number of pixels. In those cases, a hair strand, for example, does not appear as a cylinder as expected.

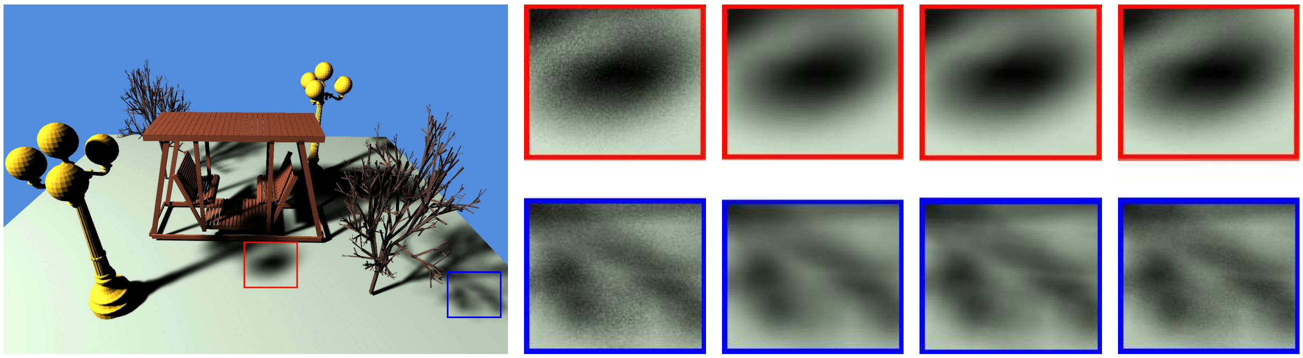
As a solution to this challenge, we present a visibility engine based on line sampling [Jones and Perry 2000] in the spatial domain. Our curves are modeled as Bézier splines with varying thickness. We develop a novel intersection algorithm between such curves and line samples and present a new interval resolve procedure. As can be seen in Figure 1, our approach renders practically noise-free images at large spectrum of scales. For rapid rendering, we have also implemented our visibility engine on a graphics processor running in parallel.

ACM Reference Format
Barringer, R., Gribel, C., Akenine-Möller, T. 2012. High-Quality Curve Rendering using Line Sampled Visibility. *ACM Trans. Graph. 31* 6, Article 162 (November 2012), 10 pages. DOI = 10.1145/2366145.2366181 http://doi.acm.org/10.1145/2366145.2366181.

Axis-Aligned Filtering for Interactive Sampled Soft Shadows

Soham Uday Mehta Brandon Wang Ravi Ramamoorthi

University of California, Berkeley *



(a) *our method, 37 spp (samples per pixel)* **(b)** *equal time, unfiltered, 37 spp* **(c)** *our method, 37 spp* **(d)** *ground truth 4000 spp* **(e)** *equal error, unfiltered, 153 spp*

Figure 1: *(a) Soft shadows from a planar area light are computed accurately by raytracing at 2.3 fps, with adaptive sampling and filtering using our method. The scene has 300K vertices and complex shadows. Our method is simple, requires no precomputation and directly plugs into NVIDIA's OptiX or other real-time raytracer. (b) Soft shadows without filtering, equal time; note the considerable noise. (c) Our method compares well to ground truth (d). (e) Equal error without filtering still has some noise making it visually less acceptable. The scene is based on one first used in [Overbeck et al. 2006]. Readers are encouraged to zoom into the PDF in all figures to see the noise and shadow details.*

Abstract

We develop a simple and efficient method for soft shadows from planar area light sources, based on explicit occlusion calculation by raytracing, followed by adaptive image-space filtering. Since the method is based on Monte Carlo sampling, it is accurate. Since the filtering is in image-space, it adds minimal overhead and can be performed at real-time frame rates. We obtain interactive speeds, using the Optix GPU raytracing framework. Our technical approach derives from recent work on frequency analysis and sheared pixel-light filtering for offline soft shadows. While sample counts can be reduced dramatically, the sheared filtering step is slow, adding minutes of overhead. We develop the theoretical analysis to instead consider *axis-aligned* filtering, deriving the sampling rates and filter sizes. We also show how the filter size can be reduced as the number of samples increases, ensuring a consistent result that converges to ground truth as in standard Monte Carlo rendering.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Shadow Algorithms

Keywords: Fourier, filtering, shadows, raytracing

*(sohamum,ravir)@cs.berkeley.edu, brandonwang@berkeley.edu

ACM Reference Format
Mehta, S., Wang, B., Ramamoorthi, R. 2012. Axis-Aligned Filtering for Interactive Sampled Soft Shadows. *ACM Trans. Graph. 31* 6, Article 163 (November 2012), 10 pages. DOI = 10.1145/2366145.2366182 http://doi.acm.org/10.1145/2366145.2366182.

Links: DL PDF VIDEO CODE

1 Introduction

This paper focuses on accurate and efficient rendering of soft shadows from planar area lights. Soft shadows are a key effect in photo-realistic rendering, but are expensive because every location on the area light must be considered and sampled. While real-time techniques [Hasenfratz et al. 2003; Johnson et al. 2009] have achieved impressive results, they rely on a variety of approximations, making no guarantee of convergence to ground truth while retaining some artifacts. On the other hand, Monte Carlo shadow-ray tracing is the preferred offline rendering method since it is physically-based, accurate and artifacts (noise that goes away with more samples) are well understood. Monte Carlo rendering can now be GPU-accelerated, and ready-to-use raytracers are available; we use NVIDIA's Optix [Parker et al. 2010]. However, the number of samples per pixel for soft shadows remains too large for interactive use.

We build on [Egan et al. 2011b] to significantly reduce the number of Monte Carlo samples needed, while still keeping the benefits of accurate raytraced occlusion. [Egan et al. 2011b] developed a sheared filter in the 4D pixel-light space, that combines samples from many different pixels, at different light locations. However, the filtering step introduces considerable overhead of minutes, and the technique is offline. In this paper, we use simpler *axis-aligned* (rather than sheared) filters. (In this context, axis-aligned or sheared refers to the shadow light field in the pixel-light domain, rather than the 2D image—we also use separable 1D filters along the image axes as a practical optimization, but this is less critical.)

While the number of samples per pixel is somewhat increased in our axis-aligned method as opposed to sheared filtering, post-processing reduces to a simple adaptive 2D image-space filter, rather than needing to search over the irregular sheared filter for samples in the full 4D shadow light field. Our method is easily integrated with existing Monte Carlo renderers, reducing the samples

Foveated 3D Graphics

Brian Guenter Mark Finch Steven Drucker Desney Tan John Snyder
Microsoft Research

Abstract

We exploit the falloff of acuity in the visual periphery to accelerate graphics computation by a factor of 5-6 on a desktop HD display (1920×1080). Our method tracks the user's gaze point and renders three image layers around it at progressively higher angular size but lower sampling rate. The three layers are then magnified to display resolution and smoothly composited. We develop a general and efficient antialiasing algorithm easily retrofitted into existing graphics code to minimize "twinkling" artifacts in the lower-resolution layers. A standard psychophysical model for acuity falloff assumes that minimum detectable angular size increases linearly as a function of eccentricity. Given the slope characterizing this falloff, we automatically compute layer sizes and sampling rates. The result looks like a full-resolution image but reduces the number of pixels shaded by a factor of 10-15.

We performed a user study to validate these results. It identifies two levels of foveation quality: a more conservative one in which users reported foveated rendering quality as equivalent to or better than non-foveated when directly shown both, and a more aggressive one in which users were unable to correctly label as increasing or decreasing a short quality progression relative to a high-quality foveated reference. Based on this user study, we obtain a slope value for the model of 1.32-1.65 arc minutes per degree of eccentricity. This allows us to predict two future advantages of foveated rendering: (1) bigger savings with larger, sharper displays than exist currently (e.g. 100 times speedup at a field of view of 70° and resolution matching foveal acuity), and (2) a roughly linear (rather than quadratic or worse) increase in rendering cost with increasing display field of view, for planar displays at a constant sharpness.

Keywords: antialiasing, eccentricity, minimum angle of resolution (MAR), multiresolution gaze-contingent display (MGCD).

Links: DL PDF WEB VIDEO

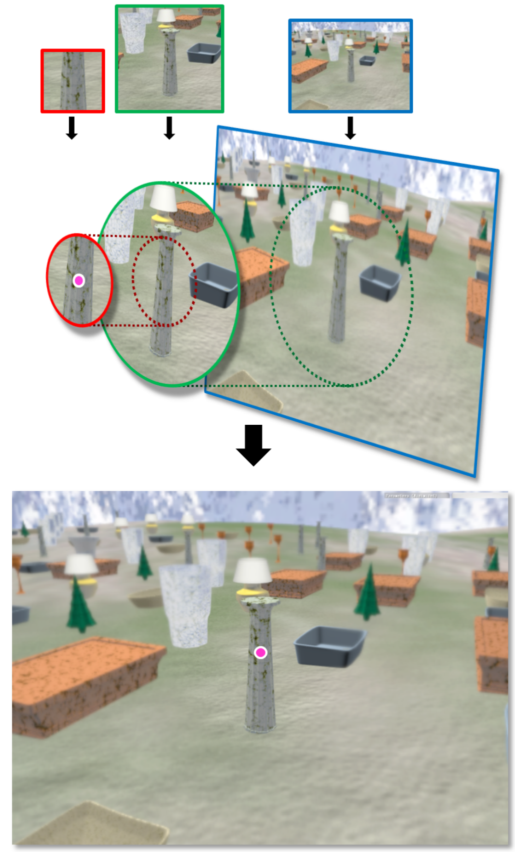


Figure 1: *Foveated rendering. We render three eccentricity layers (red border = inner layer, green = middle layer, blue = outer layer) around the tracked gaze point (pink dot), shown at their correct relative sizes in the top row. These are interpolated to native display resolution and smoothly composited to yield the final image at the bottom. Foveated rendering greatly reduces the number of pixels shaded and overall graphics computation.*

1 Introduction

We see 135° vertically and 160° horizontally, but sense fine detail only within a 5° central circle. This tiny portion of the visual field projects to the retinal region called the fovea, tightly packed with color cone receptors.[1] The angular distance away from the central gaze direction is called *eccentricity*. Acuity falls off rapidly as eccentricity increases due to reduced receptor and ganglion density in the retina, reduced optical nerve "bandwidth", and reduced "processing" devoted to the periphery in the visual cortex. A commonly-used psychophysical model, first discovered by Aubert and Foerster in 1857, asserts that the minimum discernible angular size (the reciprocal of visual acuity) increases roughly linearly with eccentricity. This model accurately predicts performance on many low-level vision tasks [Strasburger et al. 2011]. We use the term *foveation* as a shorthand for the decrease in acuity with eccentricity in the human visual system.

Current CG practice ignores user gaze and renders a high-resolution image over the whole display. This is tremendously wasteful of power and computing resources. The 5° foveal region fills a mere 0.8% of the solid angle of a 60° display. By tracking eye gaze and adapting image resolution and geometric level of detail (LOD) to eccentricity, we can omit unperceived detail and draw far fewer pixels and triangles.

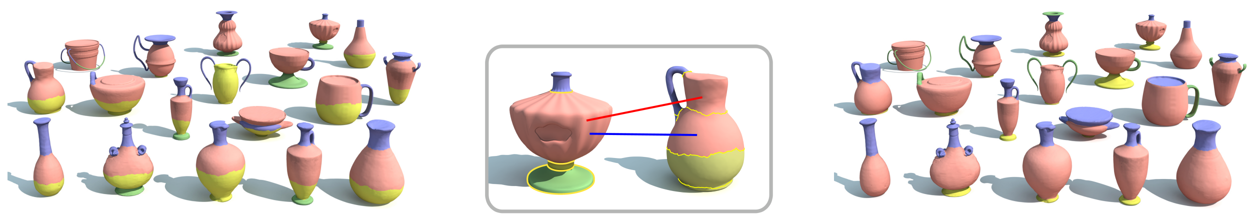
Our system exploits foveation on existing graphics hardware by rendering three nested and overlapping render targets or *eccentricity layers* centered around the current gaze point. Refer to Figure 1. These layers are denoted the *inner/foveal layer*, *middle layer*,

[1] A smaller region of 1° diameter, called the foveola, is often considered the site of foveal vision.

ACM Reference Format
Guenter, B., Finch, M., Drucker, S., Tan, D., Snyder, J. 2012. Foveated 3D Graphics.
ACM Trans. Graph. 31 6, Article 164 (November 2012), 10 pages. DOI = 10.1145/2366145.2366183
http://doi.acm.org/10.1145/2366145.2366183.

Active Co-Analysis of a Set of Shapes

Yunhai Wang* Shmulik Asafi† Oliver van Kaick‡ Hao Zhang‡ Daniel Cohen-Or† Baoquan Chen*
*Shenzhen VisuCA Key Lab/SIAT †Tel-Aviv University ‡Simon Fraser University



(a) (b) (c)

Figure 1: *Overview of our active co-analysis: (a) We start with an initial unsupervised co-segmentation of the input set. (b) During active learning, the system automatically suggests constraints which would refine results and the user interactively adds constraints as appropriate. In this example, the user adds a cannot-link constraint (in red) and a must-link constraint (in blue) between segments. (c) The constraints are propagated to the set and the co-segmentation is refined. The process from (b) to (c) is repeated until the desired result is obtained.*

Abstract

Unsupervised co-analysis of a set of shapes is a difficult problem since the geometry of the shapes alone cannot always fully describe the semantics of the shape parts. In this paper, we propose a semi-supervised learning method where the user actively assists in the co-analysis by iteratively providing inputs that progressively constrain the system. We introduce a novel constrained clustering method based on a spring system which embeds elements to better respect their inter-distances in feature space together with the user-given set of constraints. We also present an active learning method that suggests to the user where his input is likely to be the most effective in refining the results. We show that each single pair of constraints affects many relations across the set. Thus, the method requires only a sparse set of constraints to quickly converge toward a consistent and error-free semantic labeling of the set.

Keywords: semi-supervised learning, active learning

Links: DL PDF WEB VIDEO DATA

1 Introduction

Recently, there is an increasing interest in the co-analysis of sets of shapes, since current works have shown that more information can be extracted by simultaneously analyzing a set, rather than analyzing each shape individually [Golovinskiy and Funkhouser 2009; Xu et al. 2010; Sidi et al. 2011]. The main task in co-analysis is to simultaneously segment all the shapes in the set in a consistent manner, which is of great utility for modeling and texturing [Kalogerakis et al. 2010; Xu et al. 2011]. That is, besides partitioning the shapes into segments, we also obtain a labeling of the segments across the set, where the parts with the same label serve the same semantic purpose, albeit possibly being geometrically dissimilar.

Previous attempts to co-analyze a set of shapes can be classified into supervised and unsupervised. In the supervised setting [Kalogerakis et al. 2010; van Kaick et al. 2011], a training set with enough pre-analyzed shapes is assumed to be given. The training set is then used to probabilistically label a set of unknown shapes. Although supervised methods are not strictly speaking a co-analysis (since the shapes are not simultaneously analyzed), the result of the labeling is a consistent segmentation of the set. The unsupervised setting is more challenging, since no prior information is given and the entire knowledge must be extracted from the input set [Golovinskiy and Funkhouser 2009; Xu et al. 2010; Sidi et al. 2011]. In general, supervised methods have superior performance, but their performance hinges upon the relevance and quality of the training set.

In this paper, we consider the use of semi-supervised learning (SSL) for co-analysis. Semi-supervised methods [Zhu 2005] can be viewed as supervised methods with a rather small training set, but which also consider the latent information in the entire set. Alternatively, SSL methods can also be viewed as unsupervised methods assisted by rather minimal input coming out of the set (which we call *external input*). Viewed as augmented unsupervised methods, SSL methods outperform unsupervised methods at the minimal cost of requiring some educated external input. At the same time, an effective SSL method should outperform supervised methods in cases where the training set is too small, or poorly suited to the input shapes. Note that no method, be it supervised or unsupervised, can guarantee a perfect co-analysis of a set, since the geometry alone cannot always fully convey the semantics of parts. In particular, no descriptors can capture all possible geometric variations of a part.

Our contribution is the introduction of an SSL technique where the user interactively provides the external input as a means to iteratively correct and improve the result of the co-analysis (Figure 1). The external input consists of a sparse set of pairs of segments that the user marks as *must-link* or *cannot-link* constraints. Our SSL system is based on two main components. Firstly, we introduce a novel constrained clustering method based on a spring system which embeds elements to better respect both their inter-distances in a feature space and a given set of constraints. In our interactive setting, the user marks a few pairs of segments at each refinement

ACM Reference Format
Wang, Y., Asafi, S., van Kaick, O., Zhang, H., Cohen-Or, D., Chen, B. 2012. Active Co-Analysis of a Set of Shapes. *ACM Trans. Graph. 31* 6, Article 165 (November 2012), 10 pages.
DOI = 10.1145/2366145.2366184 http://doi.acm.org/10.1145/2366145.2366184.

Co-Abstraction of Shape Collections

Mehmet Ersin Yumer   Levent Burak Kara
Carnegie Mellon University

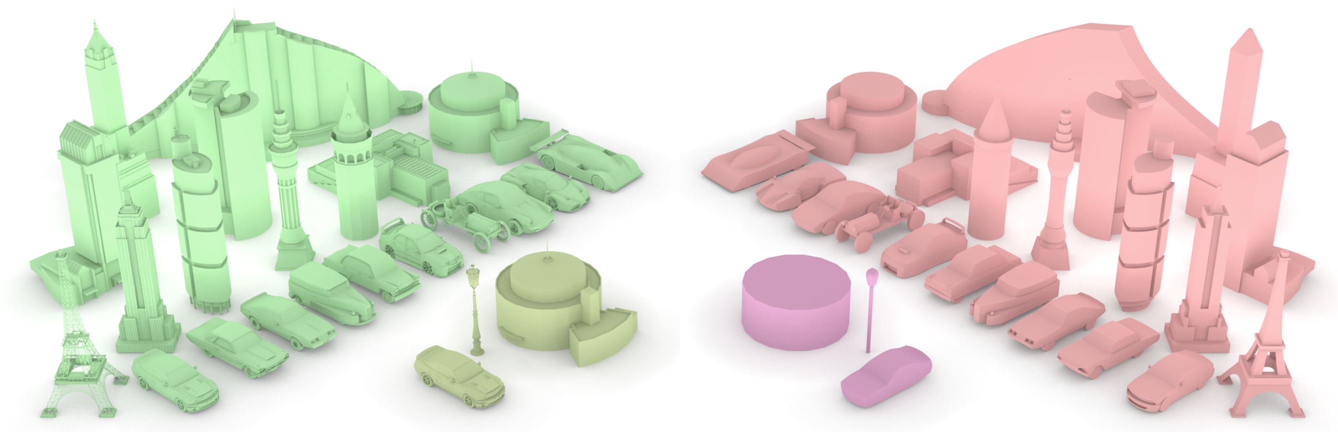


Figure 1: *Given a shape collection, our system creates a mutually abstracted set. We compute a spectrum of abstractions for each model. The most appropriate abstraction level is then determined in relation to other abstractions in the collection. When a collection includes geometrically similar models such as the cars (**green set**), the resulting abstractions (**red set**) preserve details unique to each model. When the original models are dissimilar (**yellow set**), the abstraction of the same models can become much simpler (**magenta set**). (The original models used throughout this paper are the copyright of the respective owners from Google 3D Warehouse, Archive-3D and Artist-3D.)*

Abstract

We present a co-abstraction method that takes as input a collection of 3D objects, and produces a mutually consistent and individually identity-preserving abstraction of each object. In general, an abstraction is a simpler version of a shape that preserves its main characteristics. We hypothesize, however, that there is no single abstraction of an object. Instead, there is a variety of possible abstractions, and an admissible one can only be chosen conjointly with other objects' abstractions. To this end, we introduce a new approach that hierarchically generates a spectrum of abstractions for each model in a shape collection. Given the spectra, we compute the appropriate abstraction level for each model such that shape simplification and inter-set consistency are collectively maximized, while individual shape identities are preserved.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Applications;

Keywords: co-abstraction, mutual abstraction, shape simplification, shape character extraction.

Links: DL PDF

ACM Reference Format
Yumer, M., Kara, L. 2012. Co-Abstraction of Shape Collections. *ACM Trans. Graph. 31* 6, Article 166 (November 2012), 11 pages. DOI = 10.1145/2366145.2366185 http://doi.acm.org/10.1145/2366145.2366185.

1 Introduction

> "Our experience is not of the real world itself but of an internal representation, a miniature virtual-reality replica of the world." *–Lehar, 2003.*

An object, independent of its true physical form, may manifest itself differently to different observers [Lehar 2003]. As the subject of many philosophical debates for centuries [Shorey 1901], indirect realism has influenced researchers in cognitive sciences and human perception, who have long studied the relation between perception and experience [Pylyshyn 2007; Foster 2000]. Inspired by the strong coupling between the two, we introduce co-abstraction. Instead of searching for a single abstraction of a 3D object in isolation, we hypothesize that an abstraction is meaningful only in the context of other objects. In particular, every object can have a multitude of possible abstract forms, but the most suitable abstraction can be identified by conjointly studying the abstractions of all the objects that make up a collection of models.

This view is based on the previous studies which argue that human cognition system tends to encode only a few characteristic features of a given object, rather than storing every little detail associated with it [Noë et al. 2000]. The amount of information preserved in this way is determined by the observer's knowledge of other objects. Only minimal amount of information is preserved sufficient enough to help the observer discriminate the objects in his or her knowledge base. To this end, we define an abstraction as a geometrically simplified object which preserves the features unique to its shape. From this perspective, our vision is similar to Mehra *et al.* [2009], where an abstraction is sought in the form of characteristic curves generated from the 3D geometry. However, in the quest for the "right abstraction," we believe it is further necessary to generate a spectrum of abstractions, while maintaining the ability to quantify and measure the difference in the perceived level of detail along the spectrum.

An Optimization Approach for Extracting and Encoding Consistent Maps in a Shape Collection

Qi-Xing Huang* Guo-Xin Zhang+ Lin Gao+ Shi-Min Hu+ Adrian Butscher* Leonidas Guibas*

*Stanford University +Tsinghua University

Abstract

We introduce a novel approach for computing high quality point-to-point maps among a collection of related shapes. The proposed approach takes as input a sparse set of imperfect initial maps between pairs of shapes and builds a compact data structure which implicitly encodes an improved set of maps between all pairs of shapes. These maps align well with point correspondences selected from initial maps; they map neighboring points to neighboring points; and they provide cycle-consistency, so that map compositions along cycles approximate the identity map.

The proposed approach is motivated by the fact that a complete set of maps between all pairs of shapes that admits nearly perfect cycle-consistency are highly redundant and can be represented by compositions of maps through a single base shape. In general, multiple base shapes are needed to adequately cover a diverse collection. Our algorithm sequentially extracts such a small collection of base shapes and creates correspondences from each of these base shapes to all other shapes. These correspondences are found by global optimization on candidate correspondences obtained by diffusing initial maps. These are then used to create a compact graphical data structure from which globally optimal cycle-consistent maps can be extracted using simple graph algorithms.

Experimental results on benchmark datasets show that the proposed approach yields significantly better results than state-of-the-art data-driven shape matching methods.

CR Categories: I.3.5 [Computing Methodologies]: Computer Graphics—Computational Geometry and Object Modeling;

Keywords: Data-driven methods, shape matching, loop closure, geodesic consistency, heat diffusion, quadratic programming

Links: DL PDF WEB

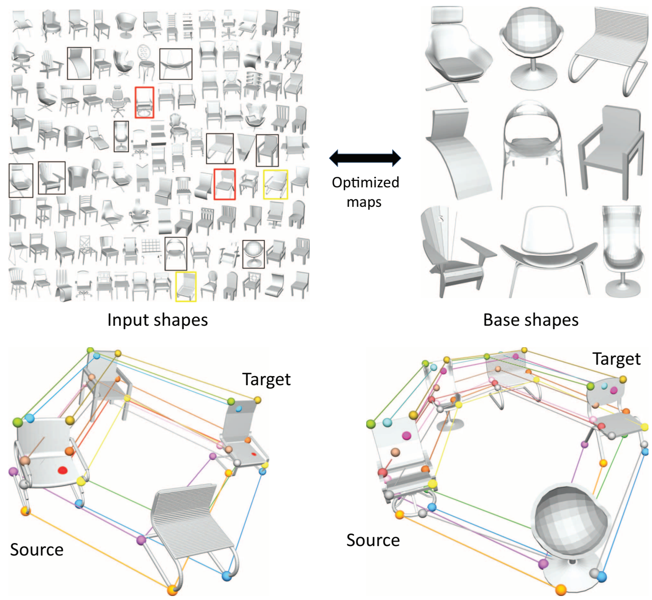


Figure 1: Given a set of input shapes, the proposed approach simultaneously selects a representative subset (the base shapes) and optimizes a set of point-to-point correspondences between each of the base shapes and the entire collection. The final pairwise maps are derived by following shortest paths in the graph specified by these correspondences. (Top) Input shapes and the base shapes selected. (Bottom) Induced maps between example pairs of shapes. For clarity, we only show a subset of correspondences. We see a diversity of paths between the source and target shapes, passing through different intermediate base shapes.

1 Introduction

With an increasing amount of data describing 3D shapes becoming available, research focus is shifting from processing a single shape to simultaneously processing a collection of shapes, aiming at combining information from multiple sources to improve the processing of each individual shape. In this direction, researchers have studied data-driven shape analysis [Golovinskiy and Funkhouser 2009; Kalogerakis et al. 2010; Huang et al. 2011; Nguyen et al. 2011; Sidi et al. 2011; van Kaick et al. 2011; Kim et al. 2012] and data-driven shape modeling [Chaudhuri et al. 2011; Fisher et al. 2011].

In this paper, we study the shape analysis problem of finding point-to-point maps between shapes belonging to a collection of loosely related shapes. High quality point-to-point maps are crucial to a variety of applications, including information transfer across shapes [Sumner and Popović 2004], shape modeling by assembly [Funkhouser et al. 2004] and detecting the shared structure among a shape collection [James and Twigg 2005]. So far most existing approaches have focused on matching pairs of shapes in isolation. These methods typically find mappings that optimally preserve some invariant property across shapes. One common example is the preservation of geodesic distances amongst near-isometrically deformed shapes, and there are several excellent contributions in this direction [Lipman and Funkhouser 2009; Kim et al. 2011; Ovsjanikov et al. 2012]. These methods work well if the invariant property is at least approximately satisfied but tend to fail when the differences between the input shapes are large. In this case, the search procedure may be trapped in local minima, or the global minimum may be far from the semantic ground truth.

In contrast, considering a collection of shapes together provides additional regularization constraints that help to detect and mitigate these issues. To explain this, we introduce the concept of a

ACM Reference Format
Huang, Q., Zhang, G., Gao, L., Hu, S., Butscher, A., Guibas, L. 2012. An Optimization Approach for Extracting and Encoding Consistent Maps in a Shape Collection. *ACM Trans. Graph. 31* 6, Article 167 (November 2012), 11 pages. DOI = 10.1145/2366145.2366186 http://doi.acm.org/10.1145/2366145.2366186.

Inverse Design of Urban Procedural Models

Carlos A. Vanegas
Purdue University
U.C. Berkeley

Ignacio Garcia-Dorado
Purdue University

Daniel G. Aliaga
Purdue University

Bedrich Benes
Purdue University

Paul Waddell
U.C. Berkeley

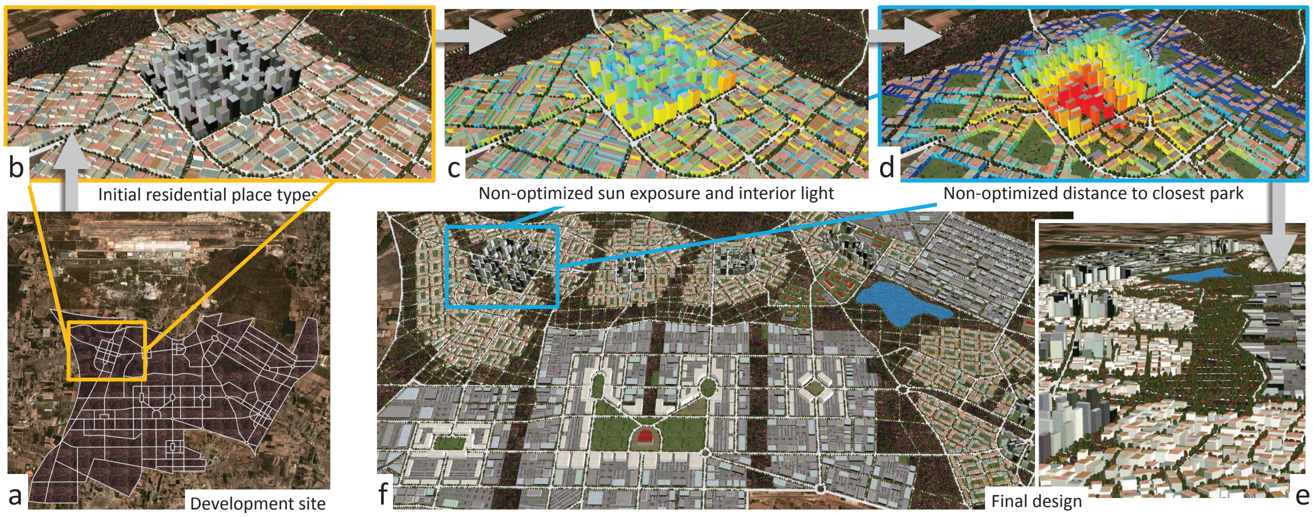


Figure 1: ***Example Urban Design Process.*** *The user interactively controls a 3D urban model by altering the parameters of an underlying procedural model (forward modeling) or by changing the values of arbitrary target indicator functions (inverse modeling). a) Starting with a development site, b) the user selects an initial urban layout (using templates or "place types"). The layout now has parcel egress but has c) initial undesired values of building sun-exposure and natural interior light, and d) unwanted average distance from buildings to closest parks. The model alterations needed to obtain user-specified local or global values for these indicator values are computed "inversely" resulting in a final design, which is e) seen up close or f) from a distance. In contrast, accomplishing such an output with traditional forward modeling would require either specifically writing a procedural model with the desired parameters or very careful parameter tuning by an expert to obtain the intended results.*

Abstract

We propose a framework that enables adding intuitive high level control to an existing urban procedural model. In particular, we provide a mechanism to interactively edit urban models, a task which is important to stakeholders in gaming, urban planning, mapping, and navigation services. Procedural modeling allows a quick creation of large complex 3D models, but controlling the output is a well-known open problem. Thus, while forward procedural modeling has thrived, in this paper we add to the arsenal an inverse modeling tool. Users, unaware of the rules of the underlying urban procedural model, can alternatively specify arbitrary target indicators to control the modeling process. The system itself will discover how to alter the parameters of the urban procedural model so as to produce the desired 3D output. We label this process inverse design.

CR Categories: I.3 [Computer Graphics] I.3.3 [Computer Graphics]: Picture/Image Generation I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling I.3.6 [Computer Graphics]: Methodology and Techniques

Keywords: interactive, procedural modeling, inverse procedural modeling, urban models

Links: DL PDF

1 Introduction

Urban procedural modeling is becoming increasingly popular in computer graphics and urban planning applications. A key basis for the popularity of city-scale urban procedural modeling is that once the procedural model is defined, it encapsulates the complex interdependencies within realistic urban spaces [Batty 2007] and enables users, who need not be aware of the internal details of the procedural model, to quickly create large complex 3D city models (e.g., [Parish and Müller 2001; Honda et al. 2004; Weber et al. 2009; Vanegas et al. 2009]). Effectively, the detail amplification inherently provided by procedural modeling is exploited: a small set of succinct *input rules* and *input parameters* can yield very complex and coherent outputs. However, the succinctness of urban pro-

ACM Reference Format
Vanegas, C., Garcia-Dorado, I., Aliaga, D., Benes, B., Waddell, P. 2012. Inverse Design of Urban Procedural Models. *ACM Trans. Graph. 31* 6, Article 168 (November 2012), 11 pages. DOI = 10.1145/2366145.2366187 http://doi.acm.org/10.1145/2366145.2366187.

Capturing and Animating the Morphogenesis of Polygonal Tree Models

Sören Pirk[1*] Till Niese[1*] Oliver Deussen[1*] Boris Neubert[2*]

[1]University of Konstanz, Germany
[2]École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

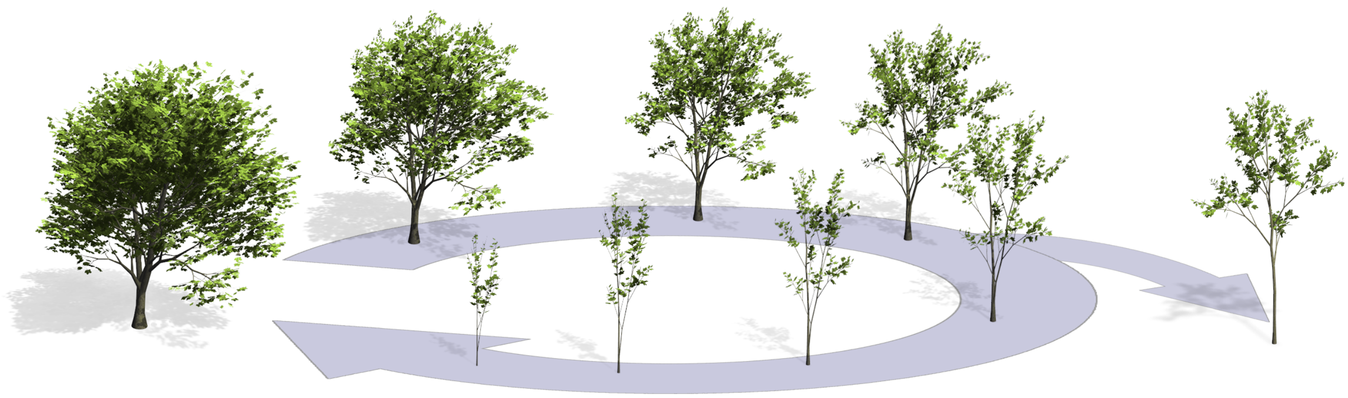


Figure 1: *The static tree model on the left is converted into a developmental model (middle part) that encompasses the ability to create arbitrary intermediate stages between a very young model and the given geometry. We define a "growth space" that allows the user to edit the model in an enhanced way. A corresponding model is shown on the right.*

Abstract

Given a static tree model we present a method to compute developmental stages that approximate the tree's natural growth. The tree model is analyzed and a graph-based description its skeleton is determined. Based on structural similarity, branches are added where pruning has been applied or branches have died off over time. Botanic growth models and allometric rules enable us to produce convincing animations from a young tree that converge to the given model. Furthermore, the user can explore all intermediate stages. By selectively applying the process to parts of the tree even complex models can be edited easily. This form of reverse engineering enables users to create rich natural scenes from a small number of static tree models.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction Techniques 1.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation I.6.8 [Simulation and Modeling]: Types of Simulation—Visual

Keywords: Generative Tree Modeling, Interactive Procedural Modeling, Plant Growth, Animation, Visual Models of Trees

Links: DL PDF WEB

*email: boris.neubert@gmail.com, {soeren.pirk, till.niese, od}@uni.kn

ACM Reference Format
Pirk, S., Niese, T., Deussen, O., Neubert, B. 2012. Capturing and Animating the Morphogenesis of Polygonal Tree Models. *ACM Trans. Graph. 31* 6, Article 169 (November 2012), 11 pages.
DOI = 10.1145/2366145.2366188 http://doi.acm.org/10.1145/2366145.2366188.

1 Introduction

Modeling complex natural scenes and objects was actively investigated for many years; trees and plants belong to almost every animated movie or computer game. Various methods have been proposed that successfully ease the modeling process while still creating realistic models. Today, thousands of such tree models can be found on the internet, large libraries exist that encompass all sorts of plants. While their visual fidelity has steadily improved over time, most of these models are still stored in the form of static geometries that cannot be altered easily, if needed. Specialized programs have to be used and sets of quite specific parameters have to be known to create variations or even animations –a task typically much too tedious for most content creators.

In this paper we present an automatic method for analyzing and re-modeling such tree models. The method allows developmental stages to be generated from a single input and supports animating growth between these states. In addition, our method offers new forms of editing through selectively applying growth development to parts of the plants. This dramatically reduces the tedium of editing such complex models and opens up what we call a "growth space": a space for developmental edits that influences the model in a quite natural way.

Most polygonal tree models (generated by L-Systems, Laser-scanning, Xfrog or other systems) can serve as the input to our system. Using a mesh contraction algorithm we reduce them to a graph structure that represents the tree skeleton. The allometry – the geometric relations within the tree– and the branching statistics can be obtained from this data structure.

During tree growth the lower branches of a tree typically die off or are pruned by humans, so we have to artificially add back these branches to our developmental model. This is done by analyzing the main branching structure and by using structural similarity –the fact that trees typically repeat branching structures along their main axes. Furthermore, we adapt the geometric relations during growth since branches change their thickness and bend as the tree develops.

Analysis and Synthesis of Point Distributions based on Pair Correlation

A. Cengiz Öztireli*
ETH Zürich

Markus Gross†
ETH Zürich

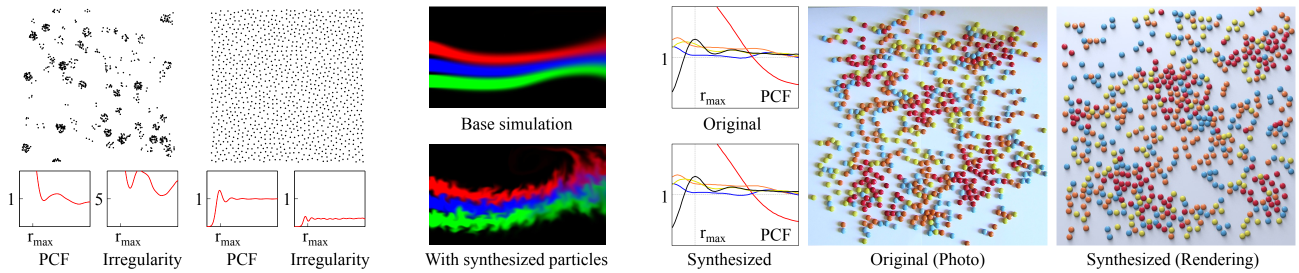


Figure 1: *Left: Our analysis and synthesis methods offer a unified treatment of point distributions, middle: to understand and recreate complex structures such as turbulence, right: by generating distributions of possibly multiple classes of objects with pair correlation functions (PCF) precisely matching those of provided input examples, such as the distributions of these candies of four colors.*

Abstract

Analyzing and synthesizing point distributions are of central importance for a wide range of problems in computer graphics. Existing synthesis algorithms can only generate white or blue-noise distributions with characteristics dictated by the underlying processes used, and analysis tools have not been focused on exploring relations among distributions. We propose a unified analysis and general synthesis algorithms for point distributions. We employ the pair correlation function as the basis of our methods and design synthesis algorithms that can generate distributions with given target characteristics, possibly extracted from an example point set, and introduce a unified characterization of distributions by mapping them to a space implied by pair correlations. The algorithms accept example and output point sets of different sizes and dimensions, are applicable to multi-class distributions and non-Euclidean domains, simple to implement and run in $O(n)$ time. We illustrate applications of our method to real world distributions.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—;

Keywords: Sampling, distributions, point processes, multi-class, blue-noise, Poisson-disk sampling, anti-aliasing

Links: DL PDF WEB

*e-mail: cengizo@inf.ethz.ch
†e-mail: grossm@inf.ethz.ch

ACM Reference Format
Öztireli, A., Gross, M. 2012. Analysis and Synthesis of Point Distributions based on Pair Correlation. *ACM Trans. Graph. 31* 6, Article 170 (November 2012), 10 pages. DOI = 10.1145/2366145.2366189 http://doi.acm.org/10.1145/2366145.2366189.

1 Introduction

Observing and recreating complex structures and phenomena of the real world is one of the main goals of computer graphics. The structures often exhibit repetitive patterns and symmetries and can be explained well in a distributional sense. Analysis of a specific phenomena is followed by fitting appropriate models and developing specialized algorithms for synthesis. Modeling complex object distributions [Deussen et al. 1998; Lagae and Dutré 2006; Wei 2010; Ma et al. 2011], turbulence synthesis for fluids [Selle et al. 2005a; Pfaff et al. 2009], crowd simulations [Narain et al. 2009] and many more applications can be seen as instances of this process. Hence, a very important open problem is learning and synthesizing general distributions.

Although point distributions arise in many contexts, analysis tools and synthesis algorithms in graphics have been mostly focused on blue-noise distributions, where points are distributed randomly with a minimum distance between each pair. It is well-known that using such a distribution for sampling leads to high quality anti-aliasing. They are also used in many contexts such as image sampling [Cook 1986], geometry processing and synthesis [Alliez et al. 2002; Ma et al. 2011], object placement [Deussen et al. 1998; Lagae and Dutré 2006; Wei 2010], or procedural noise generation [Lewis 1989]. This has led to various synthesis algorithms that generate distributions with different regularity, density, and randomness. However, the characteristics of the generated distributions are solely determined by the underlying point processes used and limited to blue-noise sampling. In contrast, nature is full of distributions with complex characteristics and many applications in graphics require analyzing and synthesizing such datasets.

Point distributions have also been extensively studied in physics and spatial statistics [Torquato 2002; Illian et al. 2008]. The emphasis in these fields is put on analyzing general distributions and fitting models to understand natural processes. Statistics that depend on correlations of locations and marks of points are used to analyze a diversity of distributions and have been proven to be powerful and discriminative.

In this paper, we introduce methods for analysis and synthesis of general multi-class point distributions based on the statistical measure *pair correlation function* (PCF). To explore the nature of this measure, we introduce an analysis based on the interpretation of it as a mean in a high dimensional vector space that we call the *pair*

Blue Noise through Optimal Transport

Fernando de Goes, Caltech
Katherine Breeden, Stanford
Victor Ostromoukhov, Lyon 1 U./CNRS
Mathieu Desbrun, Caltech

Abstract

We present a fast, scalable algorithm to generate high-quality blue noise point distributions of arbitrary density functions. At its core is a novel formulation of the recently-introduced concept of capacity-constrained Voronoi tessellation as an optimal transport problem. This insight leads to a continuous formulation able to enforce the capacity constraints exactly, unlike previous work. We exploit the variational nature of this formulation to design an efficient optimization technique of point distributions via constrained minimization in the space of power diagrams. Our mathematical, algorithmic, and practical contributions lead to high-quality blue noise point sets with improved spectral and spatial properties.

CR Categories: I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Sampling.

Keywords: Blue noise, power diagram, capacity constraints.

Links: DL PDF WEB CODE

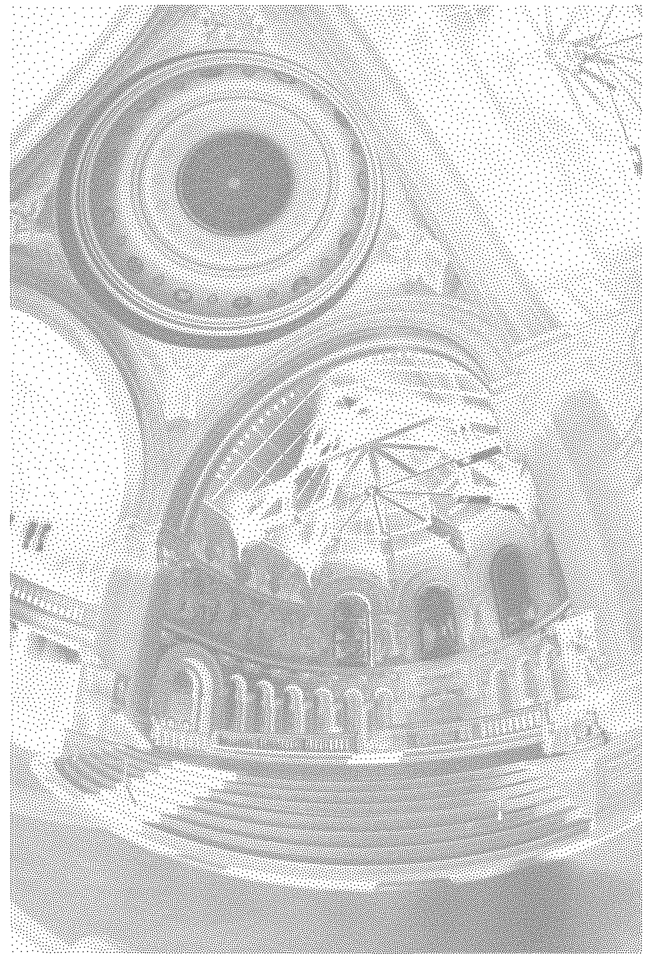


Figure 1: ***Memorial.*** *Our variational approach allows sampling of arbitrary functions (e.g., a high-dynamic range image courtesy of P. Debevec), producing high-quality, detail-capturing blue noise point distributions without spurious regular patterns (100K points, 498 s).*

1 Introduction

Coined by Ulichney [1987], the term *blue noise* refers to an even, isotropic, yet unstructured distribution of points. Blue noise was first recognized as crucial in dithering of images since it captures the intensity of an image through its local point density, without introducing artificial structures of its own. It rapidly became prevalent in various scientific fields, especially in computer graphics, where its isotropic properties lead to high-quality sampling of multidimensional signals, and its absence of structure prevents aliasing. It has even been argued that its visual efficacy (used to some extent in stippling and pointillism) is linked to the presence of a blue-noise arrangement of photoreceptors in the retina [Yellott 1983].

1.1 Previous Work

Over the years, a variety of research efforts targeting both the characteristics and the generation of blue noise distributions have been conducted in graphics. Arguably the oldest approach to algorithmically generate point distributions with a good balance between density control and spatial irregularity is through error diffusion [Floyd and Steinberg 1976; Ulichney 1987], which is particularly well adapted to low-level hardware implementation in printers. Concurrently, a keen interest in uniform, regularity-free distributions appeared in computer rendering in the context of anti-aliasing [Crow 1977]. Cook [1986] proposed the first dart-throwing algorithm to create Poisson disk distributions, for which no two points are closer together than a certain threshold. Considerable efforts followed to modify and improve this original algorithm [Mitchell 1987; McCool and Fiume 1992; Jones 2006; Bridson 2007; Gamito and Maddock 2009]. Today's best Poisson disc algorithms are very efficient and versatile [Dunbar and Humphreys 2006; Ebeida et al. 2011], even running on GPUs [Wei 2008; Bowers et al. 2010; Xiang et al. 2011]. Fast generation of irregular low-discrepancy sequences have also been proposed [Niederreiter 1992; Lemieux 2009]; however, these methods based on the radical-inverse function rarely generate high-quality blue noise.

In an effort to allow fast blue noise generation, the idea of using patterns computed offline was raised in [Dippé and Wold 1985]. To remove potential aliasing artifacts due to repeated patterns, Cohen et al. [2003] recommended the use of non-periodic Wang tiles, which subsequently led to improved hierarchical sampling [Kopf et al. 2006] and a series of other tile-based alternatives [Ostromoukhov et al. 2004; Lagae and Dutré 2006; Ostromoukhov 2007]. However, all precalculated structures used in this family of approaches rely on the offline generation of high-quality blue noise.

ACM Reference Format
de Goes, F., Breeden, K., Ostromoukhov, V., Desbrun, M. 2012. Blue Noise through Optimal Transport. *ACM Trans. Graph. 31* 6, Article 171 (November 2012), 11 pages. DOI = 10.1145/2366145.2366190 http://doi.acm.org/10.1145/2366145.2366190.

GPU-accelerated Path Rendering

Mark J. Kilgard Jeff Bolz
NVIDIA Corporation*

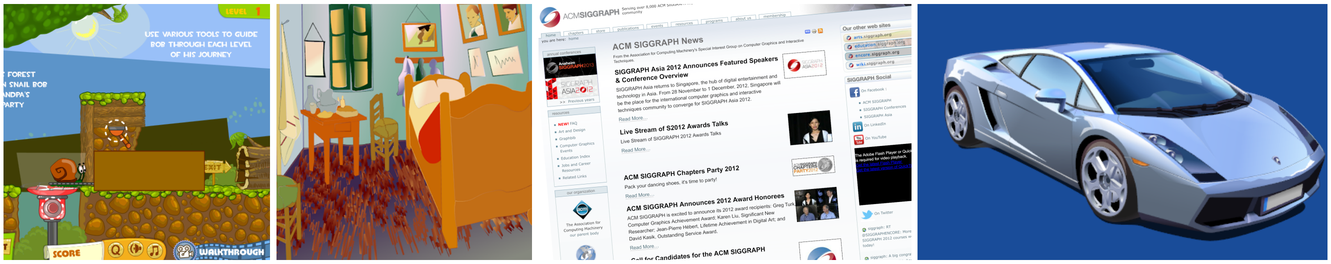


Figure 1: *GPU-accelerated scenes rendered at super-real-time rates with our system: Snail Bob Flash-based game (5ms) by permission of Andrey Kovalishin and Maxim Yurchenko, Van Gogh SVG scene with gradients (5.26ms) by permission of Enrique Meza C, complete (shown clipped) SIGGRAPH web page (4.8ms), and SVG scene with path clipping (1.9ms) by permission of Michael Grosberg, all rendered on a GeForce 560M laptop.*

Abstract

For thirty years, resolution-independent 2D standards (e.g. PostScript, SVG) have depended on CPU-based algorithms for the filling and stroking of paths. Advances in graphics hardware have largely ignored accelerating resolution-independent 2D graphics rendered from paths.

We introduce a two-step "Stencil, then Cover" (StC) programming interface. Our GPU-based approach builds upon existing techniques for curve rendering using the stencil buffer, but we *explicitly* decouple in our programming interface the *stencil step* to determine a path's filled or stroked coverage from the subsequent *cover step* to rasterize conservative geometry intended to test and reset the coverage determinations of the first step while shading color samples within the path. Our goals are completeness, correctness, quality, and performance—yet we go further to unify path rendering with OpenGL's established 3D and shading pipeline. We have built and productized our approach to accelerate path rendering as an OpenGL extension.

CR Categories: I.3.2 [Computer Graphics]: Graphics Systems—Stand-alone systems;

Keywords: path rendering, vector graphics, OpenGL, stencil buffer

Links: DL PDF

1 Introduction

*e-mail: {mjk,jbolz}@nvidia.com

ACM Reference Format
Kilgard, M., Bolz, J. 2012. GPU-accelerated Path Rendering. *ACM Trans. Graph. 31* 6, Article 172 (November 2012), 10 pages. DOI = 10.1145/2366145.2366191 http://doi.acm.org/10.1145/2366145.2366191.

When people surf the web, read PDF documents, interact with smart phones or tablets, use productivity software, play casual games, or in everyday life encounter the full variety of visual output created digitally (advertising, books, logos, signage, etc.), they are experiencing resolution-independent 2D graphics.

While 3D graphics dominates graphics research, we observe that most visual interactions between humans and computers involve 2D graphics. Sometimes this type of computer graphics is called *vector graphics*, but we prefer the term *path rendering* because the latter term emphasizes the path as the unifying primitive for this approach to rendering.

1.1 Terminology of Path Rendering

A *path* is a sequence of trajectories and contours. In this context, a *trajectory* is a connected sequence of *path commands*. Path commands include line segments, Bézier curve segments, and partial elliptical arcs. Each path command has an associated set of numeric parameters known as *path coordinates*. When a pair of path coordinates defines a 2D (x, y) location, this pair is a *control point*. Intuitively a trajectory corresponds to pressing a pen's tip down on paper, dragging it to draw on the paper, and eventually lifting the pen.

A *contour* is a trajectory with the same start and end point; in other words, a closed trajectory. These contours and trajectories may be convex, self-intersecting, nested in other contours, or may intersect other trajectories/contours in the path. There is generally no bound on the number of path segments or trajectories/contours in a path. For a rendering "primitive," paths can be quite complex.

Paths are rendered by either *filling* or *stroking* the path. Conceptually, path filling corresponds to determining what points (framebuffer sample locations) are logically "inside" the path. Stroking is roughly the region swept out by a fixed-width pen—centered on the trajectory—that travels along the trajectory orthogonal to the trajectory's tangent direction.

1.2 History, Standards, Motivation, and Contributions

Seminal work by Warnock and Wyatt [1982] introduced a coherent model for path rendering. Since that time, many standards and programming interfaces have incorporated path rendering constructs

A Vectorial Solver for Free-form Vector Gradients

Simon Boyé Pascal Barla Gaël Guennebaud
Inria - U. Bordeaux - IOGS - CNRS

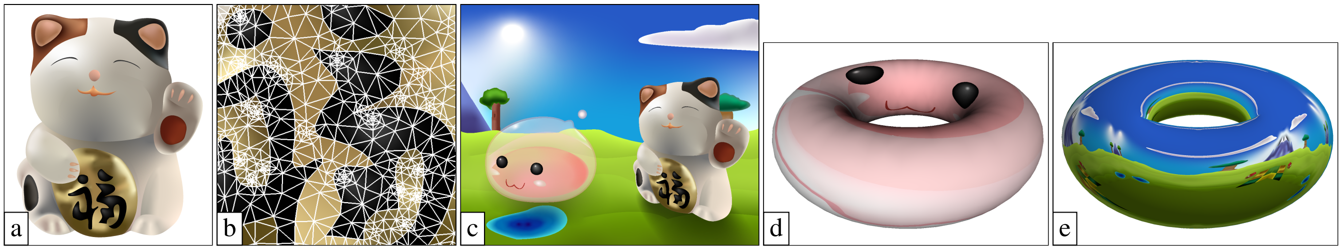


Figure 1: ***Free-form vector gradients:*** *Our vectorial solver permits to create complex vector gradients (a). Our method is based on a triangular representation (b) that is output-insensitive and thus works with arbitrary high image resolutions. Our solver does not need to be updated for a variety of operations: (c) instancing, layering and deformation; (d) texture mapping; and even (e) environment mapping.*

Abstract

The creation of free-form vector drawings has been greatly improved in recent years with techniques based on (bi)-harmonic interpolation. Such methods offer the best trade-off between sparsity (keeping the number of control points small) and expressivity (achieving complex shapes and gradients). In this paper, we introduce a *vectorial* solver for the computation of free-form vector gradients. Based on Finite Element Methods (FEM), its key feature is to output a low-level vector representation suitable for very fast GPU accelerated rasterization and close-form evaluation. This intermediate representation is hidden from the user: it is dynamically updated using FEM during drawing when control points are edited. Since it is output-insensitive, our approach enables novel possibilities for (bi)-harmonic vector drawings such as instancing, layering, deformation, texture and environment mapping. Finally, in this paper we also generalize and extend the set of drawing possibilities. In particular, we show how to locally control vector gradients.

Keywords: Vector graphics, diffusion curves, finite elements

Links: DL PDF VIDEO

1 Introduction

Vector graphics techniques have been tremendously improved in recent years, both for image vectorization and free-form drawing purposes. Vectorization methods usually require a dense set of control points to accurately represent all the details of an input image. With drawing tools, this is quite the contrary: a sparse set of vector primitives with rich expressivity is preferred, since it provides for more direct editing and creation. In this paper, we are primarily interested in free-form vector drawing.

One of the first vector primitives that made tractable the creation of complex color gradients is the Gradient Mesh [Sun et al. 2007] (GM). It lets artists specify color gradients by assigning color values and derivatives to the vertices of a 2D control mesh. With increasing mesh resolution, this technique permits to represent very complex color gradients. On the down side, the denser the mesh, the more difficult it becomes to edit it (Section 2.1).

Recent work has shown that complex color gradients can advantageously be produced through the harmonic or bi-harmonic interpolation of a few spline curves and sample points defining colors [Orzan et al. 2008; Finch et al. 2011]. Such Diffusion Curves (DC) based representations have also been successfully used for the design of 3D images, height fields or normal maps (Section 2.2).

Although DC-based representations provide an appealing alternative to GMs, computing such images involves a global differential equation which is rather challenging when seeking for real-time performance, high-quality, and robustness. Current methods strive to *rasterize* DCs directly in the regular pixel grid of the destination image using finite differences. As explained in more details in Section 2.3, such an approach is subject to aliasing when multiple curves pass through a single pixel, and zooming-in can only be approximated via expensive multiple evaluation passes thus breaking the infinite resolution property of the vector representation. Moreover, the expensive differential equation has to be fully solved not only when the curves are edited, but also for simple operations such as translation, rotation and scaling of the entire drawing. This is a major practical limitation, since it makes DCs not tractable for applications that require instancing, deformation or 3D mapping.

In this paper we address all these limitations by introducing a novel *vectorial* solver for DC images, i.e., a solver that outputs a vectorial representation. The key idea is to dynamically convert a DC image into a high-order mesh-based representation that is automatically adapted to the complexity of the input. To this end, we propose a robust solution that builds on constrained Delaunay triangulations and Finite Element Methods (FEM). Contrary to mesh-based approaches, this intermediate representation is hidden from the user, and updated only when DCs are edited; hence the approach takes the best of both worlds. A direct benefit is that a variety of common operations do not require any update of the solution, as illustrated in Figure 1a-c. Moreover, since the solution is vectorial, it may be accessed randomly, permitting novel uses of free-form vector graphics such as texture or environment mapping (Figure 1d-e).

In addition, we introduce the notion of asymmetric transmission that permits to progressively and precisely limit the influence of some primitives, thus enabling the mix of *global* and *local* diffusion. We present this extension in Section 3, together with a unified taxonomy of curve constraints. Our vectorial solver and intermediate representation are detailed in Section 4, and we show in Section 5 a variety of applications leveraging their benefits. Finally, we discuss limitations as well as further extensions in Section 6.

ACM Reference Format
Boyé, S., Barla, P., Guennebaud, G. 2012. A Vectorial Solver for Free-form Vector Gradients.
ACM Trans. Graph. 31 6, Article 173 (November 2012), 9 pages. DOI = 10.1145/2366145.2366192
http://doi.acm.org/10.1145/2366145.2366192.

Gaze Correction for Home Video Conferencing

Claudia Kuster[1] Tiberiu Popa[1] Jean-Charles Bazin[1] Craig Gotsman[1,2] Markus Gross[1]
[1]ETH Zurich [2]Technion - Israel Institute of Technology

Figure 1: *Top: frames recorded during video-conferencing. Note the lack of eye contact because of the disparity between the locations of the participant and the camera. Bottom: real-time gaze correction with the proposed algorithm.*

Abstract

Effective communication using current video conferencing systems is severely hindered by the lack of eye contact caused by the disparity between the locations of the subject and the camera. While this problem has been partially solved for high-end expensive video conferencing systems, it has not been convincingly solved for consumer-level setups. We present a gaze correction approach based on a single Kinect sensor that preserves both the integrity and expressiveness of the face as well as the fidelity of the scene as a whole, producing nearly artifact-free imagery. Our method is suitable for mainstream home video conferencing: it uses inexpensive consumer hardware, achieves real-time performance and requires just a simple and short setup. Our approach is based on the observation that for our application it is sufficient to synthesize only the corrected face. Thus we render a gaze-corrected 3D model of the scene and, with the aid of a face tracker, transfer the gaze-corrected facial portion in a seamless manner onto the original image.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Radiosity;

Keywords: video conferencing, depth camera, gaze correction

Links: DL PDF

ACM Reference Format
Kuster, C., Popa, T., Bazin, J., Gotsman, C., Gross, M. 2012. Gaze Correction for Home Video Conferencing. *ACM Trans. Graph. 31* 6, Article 174 (November 2012), 6 pages. DOI = 10.1145/2366145.2366193 http://doi.acm.org/10.1145/2366145.2366193.

1 Introduction

It has been firmly established [Argyle and Cook 1976; Chen 2002; Macrae et al. 2002] that mutual gaze awareness (i.e. eye contact) is a critical aspect of human communication, both in person or over an electronic link such as a video conferencing system [Grayson and Monk 2003; Mukawa et al. 2005; Monk and Gale 2002]. Thus, in order to realistically imitate real-world communication patterns in virtual communication, it is critical that the eye contact is preserved. Unfortunately, conventional hardware setups for consumer video conferencing inherently prevent this. During a session we tend to look at the face of the person talking, rendered in a window within the display, and not at the camera, typically located at the top or bottom of the screen. Therefore it is not possible to make eye contact. People who use consumer video conferencing systems, such as Skype, experience this problem frequently. They constantly have the illusion that their conversation partner is looking somewhere above or below them. The lack of eye contact makes communication awkward and unnatural. This problem has been around since the dawn of video conferencing [Stokes 1969] and has not yet been convincingly addressed for consumer-level systems.

While full gaze awareness is a complex psychological phenomenon [Chen 2002; Argyle and Cook 1976], mutual gaze or eye contact has a simple geometric description: the subjects making eye contact must be in the center of their mutual line of sight [Monk and Gale 2002]. Using this simplified model, the gaze problem can be cast as a novel view synthesis problem: render the scene from a virtual camera placed along the line of sight [Chen 2002]. One way to do this is through the use of custom-made hardware setups that change the position of the camera using a system of mirrors [Okada et al. 1994; Ishii and Kobayashi 1992]. These setups are usually too expensive for a consumer-level system.

The alternative is to use software algorithms to synthesize an image from a novel viewpoint different from that of the real camera. Systems that can convincingly do novel view synthesis typically consist

Discontinuity-Aware Video Object Cutout

Fan Zhong† Xueying Qin†* Qunsheng Peng‡ Xiangxu Meng†

†*School of Computer Science and Technology, Shandong University*
†*Shandong Provincial Key Laboratory of Software Engineering*
‡*State Key Lab. of CAD&CG, Zhejiang University*

Abstract

Existing video object cutout systems can only deal with limited cases. They usually require detailed user interactions to segment real-life videos, which often suffer from both inseparable statistics (similar appearance between foreground and background) and temporal discontinuities (e.g. large movements, newly-exposed regions following disocclusion or topology change).

In this paper, we present an efficient video cutout system to meet this challenge. A novel directional classifier is proposed to handle temporal discontinuities robustly, and then multiple classifiers are incorporated to cover a variety of cases. The outputs of these classifiers are integrated via another classifier, which is learnt from real examples. The foreground matte is solved by a coherent matting procedure, and remaining errors can be removed easily by additive spatio-temporal local editing. Experiments demonstrate that our system performs more robustly and more intelligently than existing systems in dealing with various input types, thus saving a lot of user labor and time.

CR Categories: I.4.6 [Computer Graphics]: Image Processing and Computer Vision—Segmentation - Pixel classification;

Keywords: video segmentation, object cutout, pixel classification

Links: DL PDF WEB VIDEO

1 Introduction

Extraction of dynamic video object is normally a labor-intensive and time-consuming task. Although great successes have been achieved over the past decade [Chuang et al. 2002; Agarwala et al. 2004; Li et al. 2005; Bai et al. 2009], existing methods for video object cutout are not yet efficient enough for real-life videos, and numerous user interactions are required to remove the errors caused by inseparable statistics and temporal discontinuities. Inseparable statistics are notoriously problematic in the fields of image and video segmentation. Temporal discontinuities can be caused by sudden occlusion/disocclusion, topology changes and fast motion, as demonstrated in Fig. 1. Discontinuous regions are difficult to segment correctly due to the lack of reliable temporal constraints.

Early video cutout methods based on global classifiers [Li et al. 2005; Wang et al. 2005] have proved to be sensitive to inseparable statistics. To overcome this problem, Bai et al. [2009] proposes local classifiers. The resulting system is known as Video Snapcut, and has been successfully incorporated into Adobe After Effect CS5 as the Rotobrush tool. However, as shown in Fig. 1, when compared with global classifiers, local classifiers are more sensitive to temporal discontinuities because of their limited coverage. In fact, the problem is not restricted to the local classifier, all methods relying on local temporal continuity suffer from the same problem, including the 3D graph-cut [Li et al. 2005; Tong et al. 2011] and the 3D extension [Tang et al. 2011] of the color line model [Levin et al. 2008] in video cube.

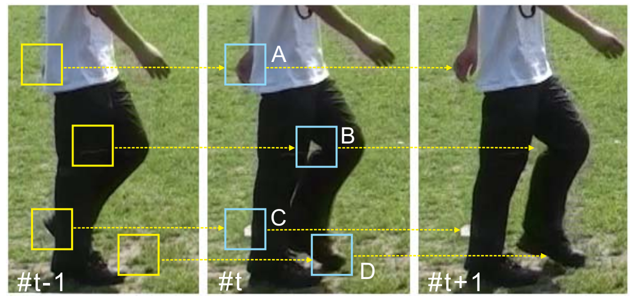


Figure 1: *Temporal discontinuities in video. Local classifiers cannot correctly segment the newly-exposed regions in frame t by sampling only in the corresponding windows in frame $t-1$, and the errors then would be magnified in frame $t+1$.*

Some efforts have been made to address the problem of temporal discontinuity. Bai et al. [2010] proposes a multi-size classifier with a size which is adaptively adjusted based on the local registration error. However, in newly-exposed regions such as regions A and B in Fig. 1, the proper window size is obviously independent of local registration errors, and is thus difficult to determine. In [Zhong et al. 2010], local and global classifiers are combined in order to deal with occlusion. However, because global classifier is sensitive to inseparable statistics, this method cannot achieve robust segmentation for videos of complex scenes.

Our Approach: We present a video cutout system which is capable of handling all common cases well; our system produces much fewer errors than existing systems, and therefore can greatly reduce the number of user interactions. We achieve this mainly by a combination of three approaches:

• *A medium-scale classifier to robustly handle temporal discontinuity.* We propose the unbiased directional classifiers (UDCs). Unlike local classifiers in previous methods, whose support windows are always square or circular [Bai et al. 2009; Zhong et al. 2010; Bai et al. 2010], UDCs reside in a set of long-narrow and directional support windows, which enables them to explore long-distance region similarities while maintaining relatively small coverage, making them robust against both inseparable statistics and temporal discontinuities. By integrating the classifiers in different directions, large image movements in any direction can be captured, and spatial context information can also be explored more effectively.

*Email: qxy@sdu.edu.cn

ACM Reference Format
Zhong, F., Qin, X., Peng, Q., Meng, X. 2012. Discontinuity-Aware Video Object Cutout.
ACM Trans. Graph. 31 6, Article 175 (November 2012), 10 pages. DOI = 10.1145/2366145.2366194
http://doi.acm.org/10.1145/2366145.2366194.

Transfusive Image Manipulation

Kaan Yücer[1,2] Alec Jacobson[1] Alexander Hornung[2] Olga Sorkine[1]
[1]ETH Zurich [2]Disney Research, Zurich

Figure 1: *A trained artist makes detailed edits to a single source image (left) and our method transfers the edits to the 8 target images (right).*

Abstract

We present a method for consistent automatic transfer of edits applied to one image to many other images of the same object or scene. By introducing novel, content-adaptive weight functions we enhance the non-rigid alignment framework of Lucas-Kanade to robustly handle changes of view point, illumination and non-rigid deformations of the subjects. Our weight functions are content-aware and possess high-order smoothness, enabling to define high-quality image warping with a low number of parameters using spatially-varying weighted combinations of affine deformations. Optimizing the warp parameters leads to subpixel-accurate alignment while maintaining computation efficiency. Our method allows users to perform precise, localized edits such as simultaneous painting on multiple images in real-time, relieving them from tedious and repetitive manual reapplication to each individual image.

Keywords: non-rigid alignment, Lucas-Kanade, content-aware warping, image edit transfer

Links: DL PDF WEB VIDEO

1 Introduction

The process of editing photographs is nearly as old as photography itself. Digital techniques in recent years have greatly expanded the spectrum of possibilities and improved the quality of these edits. Types of editing operations range from global tone adjustments, to color histograms (e.g., [Cohen-Or et al. 2006]) to localized pixel adjustments achieved by highly trained artists using specialized user-interfaces and software (e.g., [Photoshop 2012]). With the increasing availability of large digital photo collections, we currently witness a growing demand to process entire sets of images of similar scenes, taken from different viewpoints, exhibiting varying illumination, dynamic changes such as different facial expressions, and so on [Hays and Efros 2007; Hasinoff et al. 2010; HaCohen et al. 2011]. As pointed out by Hasinoff and colleagues [2010], the manual effort of applying the same localized edit to a multitude of photographs of the subject is often too great, causing users to simply discard some images from a collection.

Many recent works have reduced the user effort required for making image adjustments by using image content to intuitively propagate sparse user edits to the entire image (e.g., [Levin et al. 2004; Lischinski et al. 2006]). This is especially successful for the type of edits which demand less detailed or precise direction by the user, and therefore may be casually transferred to similar images or neighboring images in a video sequence [Li et al. 2010]. Such edits rely on a layer of indirection to disguise imperfect matching or correspondences. For example, in image colorization by sparse scribbles, the chrominance channels produced by [Levin et al. 2004] may contain discontinuities or may be matched incorrectly, but the final result is still acceptable when composed beneath the original (and the most perceptually salient) lightness channel. Recent works improve upon the practicality of such methods, e.g., by supporting more complex macros for photo manipulation that can be applied to larger collections of images [Berthouzoz et al. 2011], but effectively edit propagation is still supported only on a global scale rather than at the pixel level.

In contrast, edits such as local deformations or hand-painted pixel adjustments like the ones shown in Figure 1 may require tedious hours of a trained artist. Because of this cost, it would be advantageous to propagate such detailed edits to similar images of the same subjects. The nature of these edits requires accurate, semantically-meaningful sub-pixel matching between the relevant parts of the images. Simple matching based on color and/or spatial proximity, as used, e.g., in [Levin et al. 2004; Li et al. 2010], proves insufficient for this task. General purpose matching techniques such as optical

ACM Reference Format
Yücer, K., Jacobson, A., Hornung, A., Sorkine, O. 2012. Transfusive Image Manipulation.
ACM Trans. Graph. 31 6, Article 176 (November 2012), 9 pages. DOI = 10.1145/2366145.2366195
http://doi.acm.org/10.1145/2366145.2366195.

All-Hex Meshing using Singularity-Restricted Field

Yufei Li† Yang Liu* Weiwei Xu* Wenping Wang† Baining Guo*
*Microsoft Research Asia †The University of Hong Kong

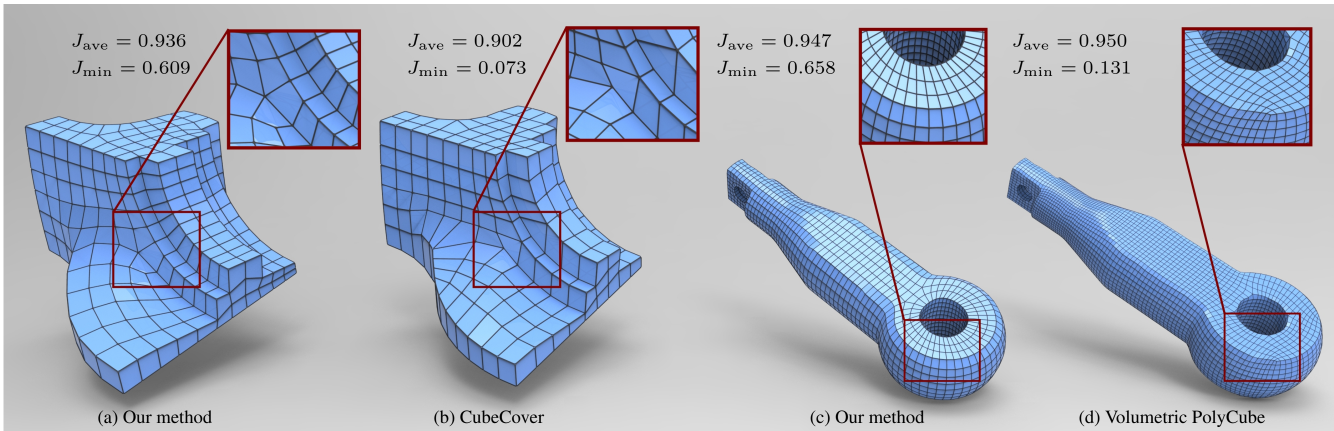


Figure 1: *High quality all-hex meshes generated by our method. Comparisons with CubeCover [Nieser et al. 2011] and volumetric PolyCube [Gregson et al. 2011] demonstrate that the hex meshes by our method are superior in mesh quality (the minimal scaled Jacobian of hexes is shown in the figure, bigger is better) and singularity placement (see the zoom-in views for comparison).*

Abstract

Decomposing a volume into high-quality hexahedral cells is a challenging task in geometric modeling and computational geometry. Inspired by the use of cross field in quad meshing and the CubeCover approach in hex meshing, we present a complete all-hex meshing framework based on *singularity-restricted field* that is essential to induce a valid all-hex structure. Given a volume represented by a tetrahedral mesh, we first compute a boundary-aligned 3D frame field inside it, then convert the frame field to be singularity-restricted by our effective topological operations. In our all-hex meshing framework, we apply the CubeCover method to achieve the volume parametrization. For reducing degenerate elements appearing in the volume parametrization, we also propose novel tetrahedral split operations to preprocess singularity-restricted frame fields. Experimental results show that our algorithm generates high-quality all-hex meshes from a variety of 3D volumes robustly and efficiently.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Geometric algorithms, languages, and systems; I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Curve, surface, solid, and object representations

Keywords: all-hex meshing, singularity-restricted field, 3D frame field.

Links: DL PDF

ACM Reference Format
Li, Y., Liu, Y., Xu, W., Wang, W., Guo, B. 2012. All-Hex Meshing using Singularity-Restricted Field. *ACM Trans. Graph. 31* 6, Article 177 (November 2012), 11 pages. DOI = 10.1145/2366145.2366196 http://doi.acm.org/10.1145/2366145.2366196.

1 Introduction

All-hex meshes possess nice numerical properties, such as a reduced number of elements and high approximation accuracy in physical simulation and mechanical engineering [Shimada 2006; Shepherd and Johnson 2008], but it is still a challenging task to automatically decompose a general 3D volume with complex boundary into high-quality hex elements [Owen 1998]. It is generally accepted that a high-quality all-hex mesh should have three properties: (1) Boundary conforming. The generated all-hex mesh should conform to the boundary surface of the 3D volume. (2) Feature alignment. Feature edges in the input 3D volume should be aligned. (3) Most important, low distortion. A single inverted hex can ruin finite element simulation. To achieve this point, singularities, the interior edges with other than four incident hexes in an all-hex mesh, should be introduced and embedded in the interior of the volume to reduce distortion.

Recently several methods have been proposed for the generation of high-quality all-hex meshes. Octree-based method [Maréchal 2009] and volumetric PolyCube [Gregson et al. 2011] embed axis-aligned boxes into the 3D volume to produce all-hex meshes. However, since these two methods do not generate interior singularities, large distortion might be introduced in the resulting all-hex meshes. Another approach, CubeCover [Nieser et al. 2011], generates all-hex meshes with the guidance of a *valid* 3D frame field via global parametrization. It first constructs a 3D frame field with the assistance of manually designed *meta-mesh*, computes a volumetric parametrization guided by this field, and then extracts an all-hex mesh accordingly. However, for a 3D volume with complex boundary, it requires a lot of manual efforts for a skilled user to design a suitable meta-mesh that can characterize the complex geometry. How to generate a valid frame field automatically is still an open problem. Recently Huang *et al.* [2011] design boundary-aligned 3D frame fields for guiding hex meshing but generate only hex-dominant meshes, not all-hex meshes.

In this paper we present a method for generating a high-quality all-hex mesh from a given tetrahedral mesh. A key ingredient of our work is a novel frame field called *singularity-restricted field*

Design-Driven Quadrangulation of Closed 3D Curves

Mikhail Bessmeltsev[1] Caoyu Wang[1] Alla Sheffer[1] Karan Singh[2]
[1]University of British Columbia [2]University of Toronto

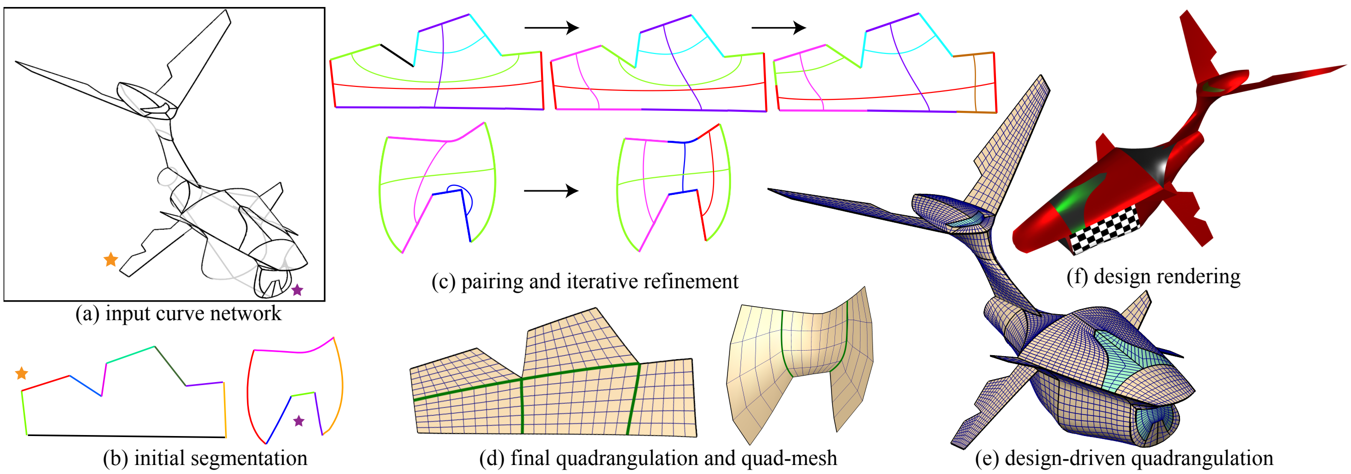


Figure 1: *Steps to quadrangulating a design network of closed 3D curves (a) : Closed curves are independently segmented (b) and iteratively paired and refined to capture dominant flow-lines as well as overall flow-line quality (c); final quadrangulation in green and dense quad-mesh (d); quadrangulations are aligned across adjacent cycles to generate a single densely sampled mesh (e), suitable for design rendering and downstream applications (f).*

Abstract

We propose a novel, design-driven, approach to quadrangulation of closed 3D curves created by sketch-based or other curve modeling systems. Unlike the multitude of approaches for quad-remeshing of existing surfaces, we rely solely on the input curves to both conceive and construct the quad-mesh of an artist imagined surface bounded by them. We observe that viewers complete the intended shape by envisioning a dense network of smooth, gradually changing, *flow-lines* that interpolates the input curves. Components of the network bridge pairs of input curve segments with similar orientation and shape. Our algorithm mimics this behavior. It first segments the input closed curves into pairs of *matching* segments, defining dominant flow line sequences across the surface. It then interpolates the input curves by a network of quadrilateral cycles whose iso-lines define the desired flow line network. We proceed to interpolate these networks with all-quad meshes that convey designer intent. We evaluate our results by showing convincing quadrangulations of complex and diverse curve networks with concave, non-planar cycles, and validate our approach by comparing our results to artist generated interpolating meshes.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—;

Keywords: curve-based modeling, flow-lines, quad-meshing

Links: DL PDF

1 Introduction

Curves are a quintessential primitive of visual communication dating back millenia to early cave drawings. In current 3D design practice, curves are universally used to depict both form and function, conveying the essence of a shape [Bordegoni and Rizzi 2011]. Sparse networks of closed 3D curves are indeed the foundation of shape in both, traditional CAD modeling [Farin and Hansford 1999] and increasingly popular sketch-based modeling interfaces [Bae et al. 2008; Nealen et al. 2007; Schmidt et al. 2009]. Unsurprisingly, recent research affirms that such 3D curve networks do effectively convey complex 3D shape [Mehra et al. 2009; de Goes et al. 2011; McCrae et al. 2011] (Figure 1 (a)). We aim to recover and compactly represent this conveyed shape (Figure 1 (f)), for designer-drawn curve networks, such as those generated by Abbasinejad et al. [2011] from sketched 3D curves [Bae et al. 2008].

While arbitrary 3D curve cycles have highly ambiguous interpolating surfaces (Figure 2 (top)), designer created curve cycles, even when highly complex, typically convey a uniquely imagined surface (Figure 2 (bottom)). These curves are designed to serve as a visual proxy of the 3D object, with the expectation that every element of surface detail is explicitly captured by the network [Gahan 2010]. To this end, design texts repeatedly emphasize the significance of

ACM Reference Format
Bessmeltsev, M., Wang, C., Sheffer, A., Singh, K. 2012. Design-Driven Quadrangulation of Closed 3D Curves. *ACM Trans. Graph. 31* 6, Article 178 (November 2012), 11 pages. DOI = 10.1145/2366145.2366197 http://doi.acm.org/10.1145/2366145.2366197.

Field-Guided Registration for Feature-Conforming Shape Composition

Hui Huang* Minglun Gong† Daniel Cohen-Or‡ Yaobin Ouyang* Fuwen Tan* Hao Zhang§
*Shenzhen VisuCA Key Lab / SIAT †Memorial Univ. of Newfoundland ‡Tel-Aviv Univ. §Simon Fraser Univ.

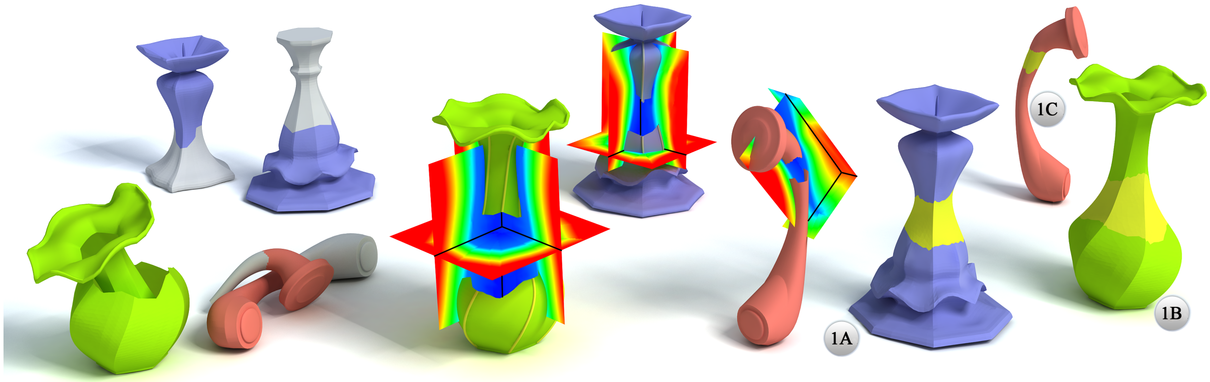


Figure 1: *Composing parts, possibly with sharp features and non-overlapping boundaries, presents challenges to both part alignment and blending. Our field-guided approach (see middle for a visualization of the fields) leads to alignment of parts away from each other and feature-conforming surface blending. The bridging surfaces generated (colored yellow on the right) are piecewise smooth.*

Abstract

We present an automatic shape composition method to fuse two shape parts which may not overlap and possibly contain sharp features, a scenario often encountered when modeling man-made objects. At the core of our method is a novel *field-guided* approach to automatically align two input parts in a *feature-conforming* manner. The key to our field-guided shape registration is a *natural continuation* of one part into the ambient field as a means to introduce an overlap with the distant part, which then allows a surface-to-field registration. The ambient vector field we compute is feature-conforming; it characterizes a piecewise smooth field which respects and naturally extrapolates the surface features. Once the two parts are aligned, gap filling is carried out by spline interpolation between matching feature curves followed by piecewise smooth least-squares surface reconstruction. We apply our algorithm to obtain feature-conforming shape composition on a variety of models and demonstrate generality of the method with results on parts with or without overlap and with or without salient features.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—[Curve, surface, solid, and object representations]

Keywords: field-guided registration, feature-conforming, shape composition, interpolation, extrapolation, piecewise smooth

Links: DL PDF WEB

ACM Reference Format
Huang, H., Gong, M., Cohen-Or, D., Ouyang, Y., Tan, F., Zhang, H. 2012. Field-Guided Registration for Feature-Conforming Shape Composition. *ACM Trans. Graph. 31* 6, Article 179 (November 2012), 11 pages. DOI = 10.1145/2366145.2366198 http://doi.acm.org/10.1145/2366145.2366198.

1 Introduction

Model creation in 3D is one of the most central problems in computer graphics. With the rapid growth of online shape repositories, advanced approaches to shape creation have opted to exploit the large number of available models. Perhaps the most widely applied modeling paradigm is to create or modify shapes by means of composing parts belonging to existing shapes [Funkhouser et al. 2004]. Often, the composing parts need to be fused together in a conforming, seamless, and coherent manner so that the resulting shapes can be readily used for subsequent processing. Despite much work on shape composition with a focus on high-level tasks [Funkhouser et al. 2004; Kreavoy et al. 2007; Chaudhuri et al. 2011; Kalogerakis et al. 2012; Xu et al. 2012], the low-level part composition problem has received considerably less attention.

The process of composing two parts together consists of two subproblems: (i) *alignment*: the two parts should be properly aligned spatially to allow a coherent composition; and (ii) *blending*: the surface(s) connecting the two parts should conform gracefully to the part geometry. These two problems are inter-related as clearly a good alignment simplifies the blending, and a powerful blending compensates for an imperfect alignment.

Most works on shape composition require a user to bring the parts into the right positions and pose [Yu et al. 2004; Sorkine et al. 2004; Kreavoy et al. 2007; Schmidt and Singh 2010]; this can be a delicate task when the parts contain sharp features to be aligned. Methods developed for automatic part alignment have appeared mostly in the shape registration literature [van Kaick et al. 2011], not shape composition. The major difference between the two is that registration almost always assumes and relies on the two parts to sufficiently overlap [Rusinkiewicz and Levoy 2001; Gelfand et al. 2005; Aiger et al. 2008], which is typically not the case for shape composition since the parts originate from different models.

Existing approaches to both part alignment and blending have all been designed to handle smooth geometry [Yu et al. 2004; Sorkine et al. 2004; Kreavoy et al. 2007; Sharf et al. 2006; Lin et al. 2008; Schmidt and Singh 2010]. In the presence of sharp features, which are ubiquitous in man-made models, the composition problem be-

Structure Recovery by Part Assembly

Chao-Hui Shen[1] Hongbo Fu[2] Kang Chen[1] Shi-Min Hu[1]
[1]TNList, Tsinghua University, Beijing [2]City University of Hong Kong

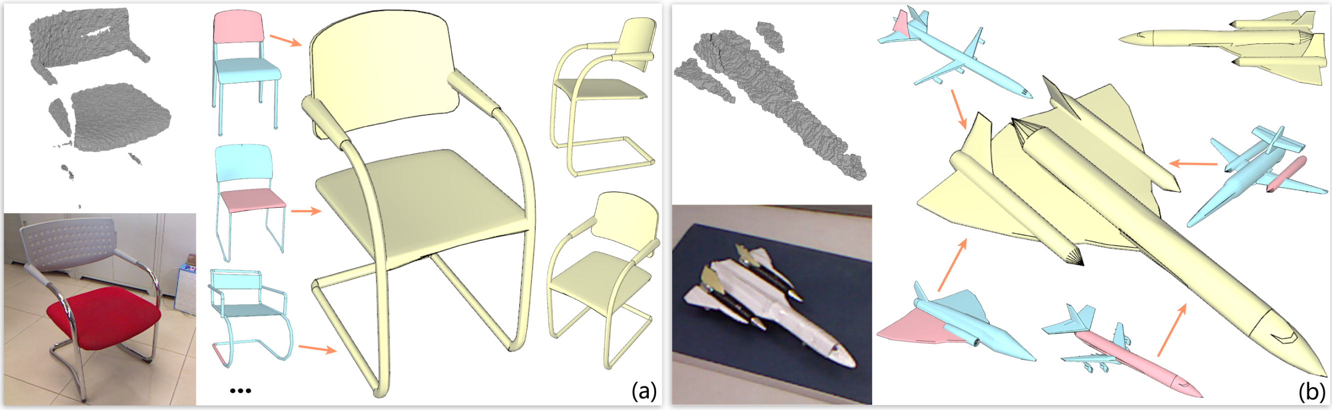


Figure 1: *Given single-view scans by the Kinect system, containing highly noisy and incomplete 3D scans (upper left) and corresponding RGB images (lower left), our approach is able to faithfully recover their underlying structures (yellow) by assembling suitable parts (red) in the repository models (blue).*

Abstract

This paper presents a technique that allows quick conversion of acquired low-quality data from consumer-level scanning devices to high-quality 3D models with labeled semantic parts and meanwhile their assembly reasonably close to the underlying geometry. This is achieved by a novel structure recovery approach that is essentially local to global and bottom up, enabling the creation of new structures by assembling existing labeled parts with respect to the acquired data. We demonstrate that using only a small-scale shape repository, our part assembly approach is able to faithfully recover a variety of high-level structures from only a single-view scan of man-made objects acquired by the Kinect system, containing a highly noisy, incomplete 3D point cloud and a corresponding RGB image.

Links: DL PDF WEB

1 Introduction

3D scanning devices provide a quick way to acquire 3D models of real-world objects or environment, which benefit a variety of applications. However, the acquired models, typically represented as unorganized point clouds, are often corrupted with noise and outlier. Worse, large regions or even entire parts might remain missing (see an example in Figure 1), possibly due to occlusions, grazing angle views, or scanner-unfriendly lighting/materials (e.g., highly reflective materials). These problems further deteriorate for consumer-level scanning devices like the Kinect system of Microsoft, which provide an economical solution to 3D capturing but at the cost of low-quality acquisition of geometry and appearance.

It is challenging to faithfully recover the underlying geometry or structure from such highly incomplete and noisy scan data. Most of the existing works (e.g., [Sharf et al. 2004; Shalom et al. 2010]) focus on geometry completion or reconstruction, and tackle inputs with small deficiencies or simple missing geometry only. Still, it is unclear how to effectively recover the underlying structure even if the geometry gets completed. The template-based approaches [Pauly et al. 2005; Kraevoy and Sheffer 2005] have great potential in completing larger, more complex holes. It is possible to transfer the structural information from the templates to the scan data. However, the existing approaches largely operate in a global-to-local manner, and thus heavily rely on the availability of template models that are globally similar to the underlying object. Although there exist a few online shape repositories like Google 3D Warehouse, the available models are still far from capturing real-world objects exhibiting complex structures, causing the main bottleneck for the existing template-based approaches.

The recent advance in mesh segmentation greatly simplifies the segmentation and labeling of parts in a set of 3D models [Kalogerakis et al. 2010; Huang et al. 2011; Sidi et al. 2011]. The recent works demonstrate how to significantly enlarge the existing database of 3D models via shape synthesis by part composition [Kalogerakis et al. 2012; Jain et al. 2012; Xu et al. 2012]. However, in practice this would result in a 3D model database that grows exponentially, making both the storage and the retrieval challenging to manage. We show that it is unnecessary to explicitly prepare such larger database by part composition and it is possible to retrieve and assemble suitable parts on the fly for structure recovery.

We propose a part assembly approach for structure recovery from a highly incomplete, noisy 3D scan of a man-made object together with the corresponding RGB image acquired by the Kinect system (Figure 1). Our approach is based on the key fact that many classes of man-made objects (e.g., chairs, bicycles etc.) lie in a low-dimensional shape space defined with respect to the relative sizes and positions of shape parts [Ovsjanikov et al. 2011]. This allows us

ACM Reference Format
Shen, C., Fu, H., Chen, K., Hu, S. 2012. Structure Recovery by Part Assembly.
ACM Trans. Graph. 31 6, Article 180 (November 2012), 11 pages. DOI = 10.1145/2366145.2366199
http://doi.acm.org/10.1145/2366145.2366199.

Multi-Scale Partial Intrinsic Symmetry Detection

Kai Xu[*†] Hao Zhang[‡] Wei Jiang[*] Ramsay Dyer[§] Zhiquan Cheng[*] Ligang Liu[⁎] Baoquan Chen[†]
[*]National University of Defense Technology [†]Shenzhen VisuCA Key Lab/SIAT
[‡]Simon Fraser University [§]INRIA, GEOMETRICA [⁎]University of Science and Technology of China

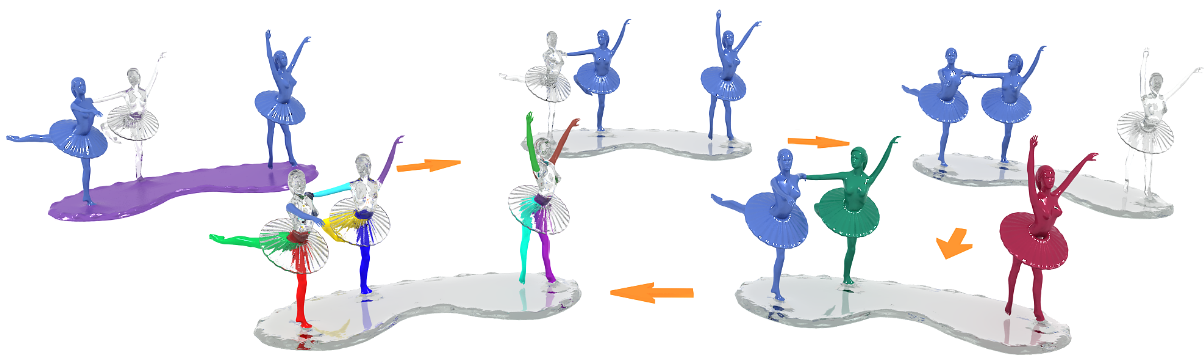


Figure 1: *Multi-scale partial intrinsic symmetry detection: five symmetry scales (large to small) are detected. Each symmetric region is shown in uniform color. Note the detection of inter- and intra-object symmetries, as well as cylindrical symmetry of the limbs.*

Abstract

We present an algorithm for *multi-scale* partial intrinsic symmetry detection over 2D and 3D shapes, where the scale of a symmetric region is defined by intrinsic distances between symmetric points over the region. To identify prominent symmetric regions which overlap and vary in form and scale, we decouple scale extraction and symmetry extraction by performing two levels of clustering. First, significant symmetry scales are identified by clustering sample point *pairs* from an input shape. Since different point pairs can share a common point, shape regions covered by points in different scale clusters can overlap. We introduce the *symmetry scale matrix* (SSM), where each entry estimates the likelihood two point pairs belong to symmetries at the same scale. The *pair-to-pair* symmetry affinity is computed based on a pair signature which encodes scales. We perform spectral clustering using the SSM to obtain the scale clusters. Then for all points belonging to the same scale cluster, we perform the second-level spectral clustering, based on a novel *point-to-point* symmetry affinity measure, to extract partial symmetries at that scale. We demonstrate our algorithm on complex shapes possessing rich symmetries at multiple scales.

Links: DL PDF WEB DATA

1 Introduction

Symmetry is ubiquitous in nature and in manufactured artifacts. The study of shape symmetry has attracted much attention in computer graphics lately [Mitra et al. 2012]. Most existing works focus on detecting extrinsic symmetries and a popular approach is transformation space voting [Mitra et al. 2006]. In the intrinsic case, most efforts are devoted to global symmetries [Ovsjanikov et al. 2008; Lipman et al. 2010]. It is generally recognized that partial intrinsic symmetry detection is a challenging problem since it has to deal with larger search spaces for both the symmetric regions (compared to global symmetry analysis) and the symmetry-revealing transforms (compared to extrinsic symmetry detection).

A partial intrinsic symmetry over a shape is a subregion with associated self-homeomorphisms that preserve all pairwise intrinsic distances [Mitra et al. 2012]. In this paper, we address the problem of partial intrinsic symmetry detection over 2D and 3D shapes. More importantly, our goal is to detect such symmetries at *multiple scales*, where we define the scale of a symmetric region based on the intrinsic distances between symmetric points over the region.

Complex shapes often exhibit multiple symmetries that overlap and vary in form and scale (see Figure 1). Multi-scale analysis enables the construction of high-level, coarse-to-fine representations for such shapes [Wang et al. 2011] to improve shape understanding and facilitate solutions to such problems as shape correspondence, editing, and synthesis. However, multi-scale symmetry analysis poses additional challenges. The problem is at first complicated by a new search dimension, the scale dimension. Also, existing approaches to intrinsic symmetry detection including those based on region growing [Xu et al. 2009], partial matching [Raviv et al. 2010], and symmetry correspondence [Lipman et al. 2010], cannot extract symmetries that physically overlap.

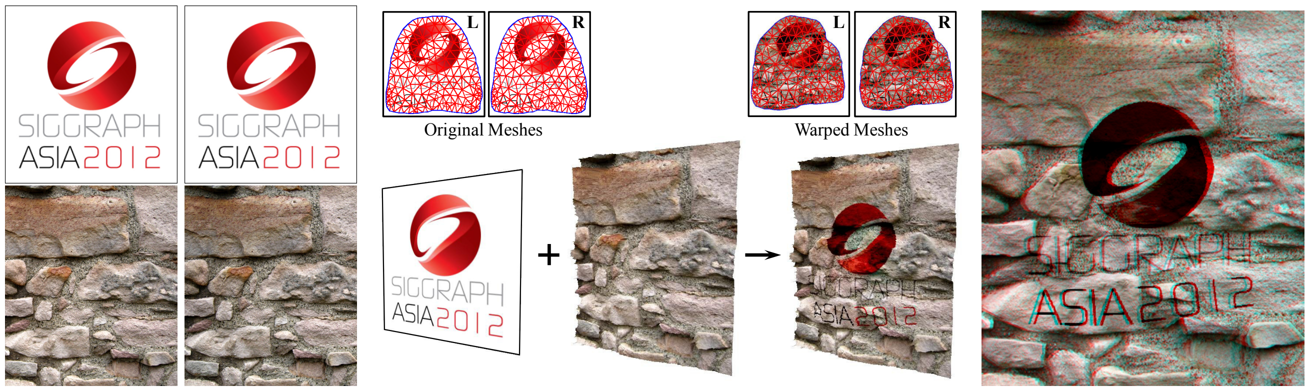
Symmetry detection inherently involves a grouping of shape elements deemed to be symmetric to each other. A clustering approach often facilitates the detection of prominent groups. Lipman et al. [2010] cluster sample points taken from an input shape with respect to a symmetry correspondence matrix (SCM). In the intrinsic setting, each SCM entry measures how symmetric two points are, based on the geometric similarity (a necessity for symmetry) between the local neighborhoods of the points. Xu et al. [2009] let point pairs vote for their intrinsic reflection symmetry axes and perform symmetry grouping via region growing. However, neither approach considered scale information.

ACM Reference Format
Xu, K., Zhang, H., Jiang, W., Dyer, R., Cheng, Z., Liu, L., Chen, B. 2012. Multi-Scale Partial Intrinsic Symmetry Detection. *ACM Trans. Graph. 31* 6, Article 181 (November 2012), 11 pages.
DOI = 10.1145/2366145.2366200 http://doi.acm.org/10.1145/2366145.2366200.

Perspective-Aware Warping for Seamless Stereoscopic Image Cloning

Sheng-Jie Luo[1] I-Chao Shen[2] Bing-Yu Chen[1] Wen-Huang Cheng[2] Yung-Yu Chuang[1]

[1]National Taiwan University [2]Academia Sinica



(a) Input images (b) Perspective-aware warping with disparity adaptation (c) Result anaglyph

Figure 1: *We present a novel technique for seamlessly cloning content from one stereoscopic image pair to another. Given a synthetic 3D SIGGRAPH Asia 2012 logo as the source image pair and a target stereoscopic image pair of a bumpy wall (a), we use perspective-aware warping to adjust the structure of the logo and paste it on to the bumpy wall (c). The perceived depth and projection of the pasted logo (b) are adjusted locally and adaptively to fit onto the bumpy surface. (Note that the resultant left and right images are included in the supplemental materials. It is recommended to watch them with stereoscopic displays for better visual effects.)*

Abstract

This paper presents a novel technique for seamless stereoscopic image cloning, which performs both shape adjustment and color blending such that the stereoscopic composite is seamless in both the perceived depth and color appearance. The core of the proposed method is an iterative disparity adaptation process which alternates between two steps: disparity estimation, which re-estimates the disparities in the gradient domain so that the disparities are continuous across the boundary of the cloned region; and perspective-aware warping, which locally re-adjusts the shape and size of the cloned region according to the estimated disparities. This process guarantees not only depth continuity across the boundary but also models local perspective projection in accordance with the disparities, leading to more natural stereoscopic composites. The proposed method allows for easy cloning of objects with intricate silhouettes and vague boundaries because it does not require precise segmentation of the objects. Several challenging cases are demonstrated to show that our method generates more compelling results compared to methods with only global shape adjustment.

Keywords: Seamless cloning, stereoscopic images, image gradients, disparity gradients, Poisson equation, disparity adaptation, perspective-aware warping.

Links: DL PDF WEB VIDEO

ACM Reference Format
Luo, S., Shen, I., Chen, B., Cheng, W., Chuang, Y. 2012. Perspective-Aware Warping for Seamless Stereoscopic Image Cloning. *ACM Trans. Graph. 31* 6, Article 182 (November 2012), 8 pages.
DOI = 10.1145/2366145.2366201 http://doi.acm.org/10.1145/2366145.2366201.

1 Introduction

The success of stereoscopic 3D movies has ignited the rapid development of 3D cameras and displays for consumers. Once users are able to capture and display stereoscopic 3D media easily, the next requirement will be the ability to manipulate these 3D media similar to the way 2D media are manipulated. However, it can be challenging to directly apply 2D media editing tools to stereoscopic 3D media, because the additional information (i.e., depth) in stereoscopic 3D images introduces additional constraints in maintaining a comfortable and enjoyable 3D viewing experience. Naive extensions of existing 2D image editing methods usually fail as they do not take these constraints into account.

This paper focuses on stereoscopic image cloning, that is, selecting a region from a source image and pasting it into a target image. Although successful 2D image cloning methods have been proposed [Pérez et al. 2003; Jia et al. 2006; Farbman et al. 2009; Yang et al. 2009], stereoscopic 3D image cloning has its own challenges: (1) we must adjust disparity values within the cloned region for depth continuity; (2) we must alter the projected shape of the cloned region according to the disparity change to model perspective effects such as foreshortening; (3) we must maintain the coordination between the left and right views for comfortable 3D viewing. Lo *et al.* [2010] proposed a cut-and-paste system for stereoscopic images which uses a segmentation technique to accurately select the object that users intend to clone. Although their system successfully meets some of these challenges, it still has two shortcomings. First, it is difficult to accurately segment out objects with complex silhouettes and objects without obvious boundaries between them and the background. Second, their system models the cloned objects as stereo billboards, which is more effective for approximating objects standing on a ground plane and for flat objects without much depth variation.

To address the above challenges and limitations, we propose a novel technique for seamless stereoscopic image cloning which performs both shape adjustment and color blending such that the stereoscopic

Enabling Warping on Stereoscopic Images

Yuzhen Niu Wu-Chi Feng Feng Liu
Portland State University

(a) Input (b) Basic rotation result and disparities (c) Our result

Figure 1: *Stereoscopic image warping. Applying the same warping to the left and right image of a stereoscopic image often introduces vertical disparities, as shown in (b). Our method enables existing image warping algorithms on stereoscopic images. (c) shows that our warping result is free from vertical disparities. In this paper, we show stereoscopic images using the red-cyan anaglyph representation.*

Abstract

Warping is one of the basic image processing techniques. Directly applying existing monocular image warping techniques to stereoscopic images is problematic as it often introduces vertical disparities and damages the original disparity distribution. In this paper, we show that these problems can be solved by appropriately warping both the disparity map and the two images of a stereoscopic image. We accordingly develop a technique for extending existing image warping algorithms to stereoscopic images. This technique divides stereoscopic image warping into three steps. Our method first applies the user-specified warping to one of the two images. Our method then computes the target disparity map according to the user specified warping. The target disparity map is optimized to preserve the perceived 3D shape of image content after image warping. Our method finally warps the other image using a spatially-varying warping method guided by the target disparity map. Our experiments show that our technique enables existing warping methods to be effectively applied to stereoscopic images, ranging from parametric global warping to non-parametric spatially-varying warping.

CR Categories: I.4.m [Image Processing and Computer Vision]: Miscellaneous;

Keywords: Stereoscopic image warping, disparity mapping

Links: DL PDF

ACM Reference Format
Niu, Y., Feng, W., Liu, F. 2012. Enabling Warping on Stereoscopic Images. *ACM Trans. Graph. 31* 6, Article 183 (November 2012), 7 pages. DOI = 10.1145/2366145.2366202 http://doi.acm.org/10.1145/2366145.2366202.

1 Introduction

These years we have been observing a tremendous resurgence of interest in stereoscopic 3D. A variety of stereoscopic displays and cameras are available. This brings in the demand for tools for authoring and processing stereoscopic content. However, extending existing tools to stereoscopic content is often non-trivial as stereoscopic content has an extra dimension of disparity that needs to be correctly taken care of to deliver a pleasant viewing experience.

This paper focuses on stereoscopic image warping. Warping is one of the basic image processing techniques and a wide range of image warping methods have been developed [Wolberg 1990]. Applying the same warping to the left and right image of a stereoscopic image can transform them consistently; however this straightforward solution is often problematic. Figure 1 shows a simple 2D image transformation: rotation. If we rotate the left and right image with the same rotation matrix, we introduce vertical disparities, which often bring "3D fatigue" to viewers [Mendiburu 2009]. Moreover, the original horizontal disparity distribution is changed.

This paper presents a technique that extends existing image warping algorithms to stereoscopic images. Our idea is to warp one of the two images of a stereoscopic image using the user-specified warping and warp the other to both follow the user-specified warping and meet the disparity requirement. Our technique consists of three steps. We first apply the user-specified warping to one of the two images. Without loss of generality, we always warp the left image using the user-specified warping. We then compute the target disparity map according to the user-specified warping. We consider that a good target disparity map should be consistent with the warping applied to the input image and maintain the perceived roundness of the image objects. For example, if an object is stretched in the image space, it should also be stretched in depth. If the disparity remains the same, the warped image may suffer from the *cardboarding* artifacts where the perceived object becomes flattened [Mendiburu 2009]. Based on this observation, we develop an automatic disparity mapping technique that scales the local disparity range according to how this region is warped. Thus, the target disparity map is optimized to preserve the perceived 3D shape of image content after warping.

A Luminance-Contrast-Aware Disparity Model and Applications

Piotr Didyk[1,2] Tobias Ritschel[1] Elmar Eisemann[3] Karol Myszkowski[1] Hans-Peter Seidel[1] Wojciech Matusik[2]

[1]MPI Informatik [2]CSAIL MIT [3]Delft University of Technology / Télécom ParisTech

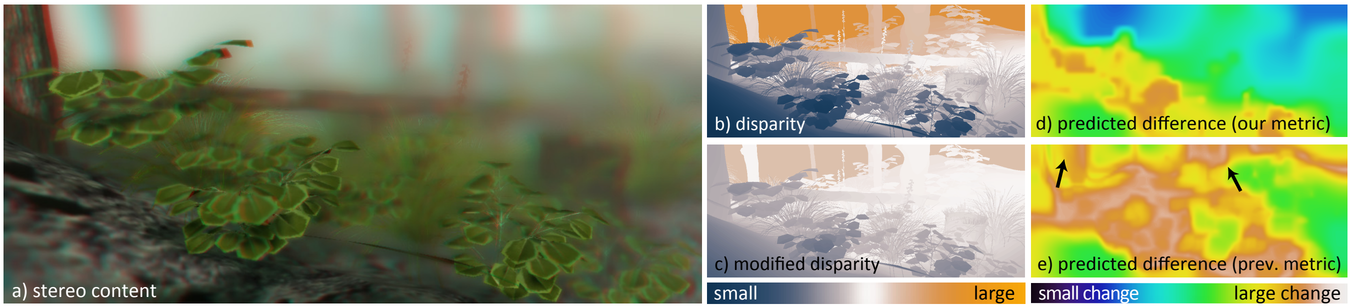


Figure 1: *When stereo content* (a; b) *is manipulated* (c)*, we quantify the perceived change considering luminance, and disparity* (d)*, whereas previous work leads to wrong predictions* (e) *e. g., for low-texture areas, fog, or depth-of-field* (arrows)*.Please note that all images in the paper, except for disparity and response maps are presented in anaglyph colors.*

Abstract

Binocular disparity is one of the most important depth cues used by the human visual system. Recently developed stereo-perception models allow us to successfully manipulate disparity in order to improve viewing comfort, depth discrimination as well as stereo content compression and display. Nonetheless, all existing models neglect the substantial influence of luminance on stereo perception. Our work is the first to account for the interplay of luminance contrast (magnitude/frequency) and disparity and our model predicts the human response to complex stereo-luminance images. Besides improving existing disparity-model applications (e. g., difference metrics or compression), our approach offers new possibilities, such as joint luminance contrast and disparity manipulation or the optimization of auto-stereoscopic content. We validate our results in a user study, which also reveals the advantage of considering luminance contrast and its significant impact on disparity manipulation techniques.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image generation—display algorithms,viewing algorithms;

Keywords: perception; stereoscopy
Links: DL PDF WEB VIDEO DATA

1 Introduction

The human visual system (HVS) combines information coming from many different cues [Howard and Rogers 2002] to determine spatial layout. Binocular disparity, due to differences of the projected retinal positions of the same object in both eyes, is one of the strongest cues, in particular for short ranges (up to 30 meters) [Cutting and Vishton 1995]. Current 3D display technology allows us to make use of binocular disparity, but, in order to ensure viewing comfort, disparity should be limited to a so-called comfort zone [Rushton et al. 1994; Lambooij et al. 2009; Shibata et al. 2011]. Smaller screens often imply a smaller disparity range, auto-stereoscopic displays only have a reduced depth of field, and artistic manipulations can enhance certain features [Ware et al. 1998; Jones et al. 2001; Lang et al. 2010; Didyk et al. 2011]. Whenever such modifications are applied, it is important to analyze the impact. Furthermore, such a prediction also leads to a better control of the changes. However, so far, no existing perception model considers the influence of RGB image content on depth perception. Intuitively, a certain magnitude of luminance contrast is required to make disparity visible, while stereopsis is likely to be weaker for low-contrast and blurry patterns. In this work, we show that luminance contrast (magnitude/frequency) does have a significant impact on depth perception and should be taken into account for a more faithful computational model. One key challenge of a combined luminance contrast and disparity model is the growing dimensionality, which we limit to 4D by considering: spatial frequency and magnitude of disparity, as well as spatial frequency and magnitude of luminance contrast. We ignore image brightness, pixel color and saturation, which seem to have a lower impact on depth perception (Sec. 7). Our model improves the performance of existing applications [Lang et al. 2010; Didyk et al. 2011] such as disparity retargeting, difference metrics, compression, and even enables previously-impossible applications. Precisely, we make the following contributions:

- A disparity-perception model accounting for image content;
- Measurements of perceived disparity changes for stimuli with different luminance and disparity patterns;
- New methods to automatically retarget disparity and to manipulate luminance contrast to improve depth perception;
- A user study to validate and illustrate the advantages of our method.

ACM Reference Format
Didyk, P., Ritschel, T., Eisemann, E., Myszkowski, K., Seidel, H., Matusik, W. 2012. A Luminance-Contrast-Aware Disparity Model and Applications. *ACM Trans. Graph. 31* 6, Article 184 (November 2012), 10 pages. DOI = 10.1145/2366145.2366203 http://doi.acm.org/10.1145/2366145.2366203.

Correcting for Optical Aberrations using Multilayer Displays

Fu-Chung Huang[1] Douglas Lanman[2,4] Brian A. Barsky[1,3] Ramesh Raskar[2]

[1]Computer Science Division, UC Berkeley [2]MIT Media Lab [3]School of Optometry, UC Berkeley [4]NVIDIA Research

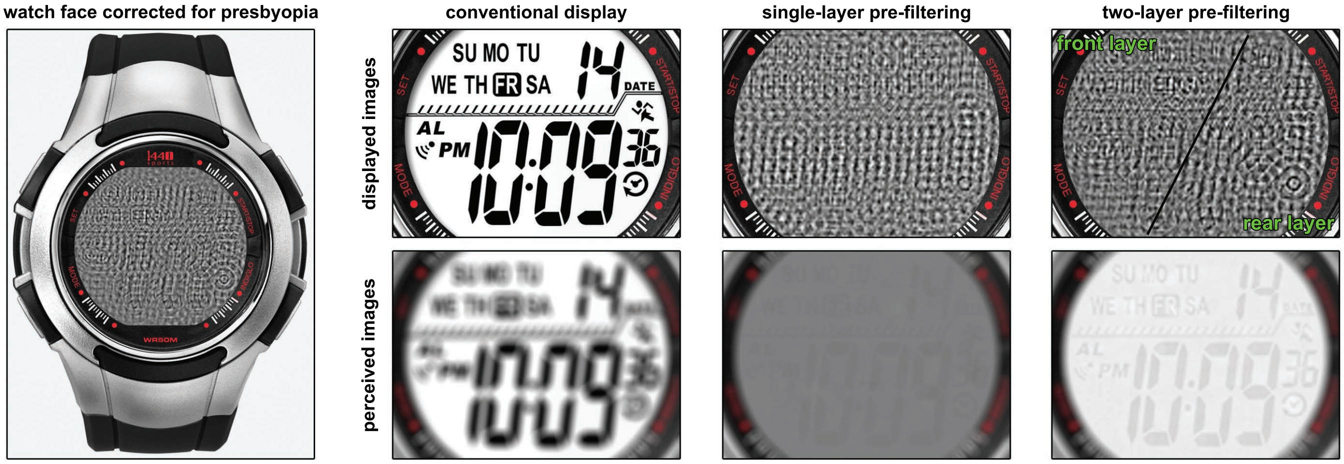


Figure 1: *Correcting presbyopia using multilayer displays. A presbyopic individual observes a watch at a distance of* 45 *cm. The watch appears out of focus due to the limited range of accommodation. To read the watch, corrective eyewear (e.g., bifocals) must be worn with a* +2.50 *diopter spherical lens. (Left) As a substitute for eyewear, the watch can be modified to use a multilayer display containing two semi-transparent, light-emitting panels. The images displayed on these layers are pre-filtered such that the watch face appears in focus when viewed by the defocused eye. (Right) From left to right along the bottom row: the perceived image using a conventional display (e.g., an unmodified LCD), using prior single-layer pre-filtering methods, and using the proposed multilayer pre-filtering method. Corresponding images of the watch face are shown along the top row. Two-layer pre-filtering, while increasing the watch thickness by* 6 *mm, enhances contrast and eliminates ringing artifacts, as compared to prior single-layer pre-filtering methods. (Watch image* © *Timex Group USA, Inc.)*

Abstract

Optical aberrations of the human eye are currently corrected using eyeglasses, contact lenses, or surgery. We describe a fourth option: modifying the composition of displayed content such that the perceived image appears in focus, after passing through an eye with known optical defects. Prior approaches synthesize pre-filtered images by deconvolving the content by the point spread function of the aberrated eye. Such methods have not led to practical applications, due to severely reduced contrast and ringing artifacts. We address these limitations by introducing multilayer pre-filtering, implemented using stacks of semi-transparent, light-emitting layers. By optimizing the layer positions and the partition of spatial frequencies between layers, contrast is improved and ringing artifacts are eliminated. We assess design constraints for multilayer displays; autostereoscopic light field displays are identified as a preferred, thin form factor architecture, allowing synthetic layers to be displaced in response to viewer movement and refractive errors. We assess the benefits of multilayer pre-filtering versus prior light field pre-distortion methods, showing pre-filtering works within the constraints of current display resolutions. We conclude by analyzing benefits and limitations using a prototype multilayer LCD.

ACM Reference Format
Huang, F., Lanman, D., Barsky, B., Raskar, R. 2012. Correcting for Optical Aberrations using Multilayer Displays. *ACM Trans. Graph. 31* 6, Article 185 (November 2012), 12 pages. DOI = 10.1145/2366145.2366204 http://doi.acm.org/10.1145/2366145.2366204.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithms; I.4.4 [Image Processing and Computer Vision]: Restoration—Inverse filtering; K.4.2 [Computers and Society]: Social Issues—Assistive technologies for persons with disabilities

Keywords: optical aberrations, deconvolution, pre-filtering, pre-compensation, multilayer displays, tailored displays, light fields, assistive devices

Links: DL PDF WEB VIDEO DATA

1 Introduction

Recent studies indicate an increasing prevalence of refractive errors: Vitale et al. [2009] found the incidence of myopia increased from 25.0% to 41.6% in the United States between 1971-1972 and 1999-2004. Today, individuals requiring correction have three options: eyeglasses, contact lenses, or refractive surgery. Oridinary eyeglasses can only correct lower-order aberrations (i.e., defocus and astigmatism). Higher-order aberrations, such as those induced by disorders including keratoconus or pellucid marginal degeneration, can be difficult to correct and are currently addressed using contact lenses or surgery. We describe a fourth option: modifying the composition of displayed imagery, as well as the underlying display hardware, to correct optical aberrations without eyewear or invasive surgery.

Our multilayer technique for the correction of optical aberrations using pre-filtering builds upon the work of Huang and Barsky [2011] and addresses visual artifacts occurring with earlier pre-filtering approaches proposed by Alonso and Barreto [2003]

The Magic Lens: Refractive Steganography

Marios Papas[1,2] Thomas Houit[2] Derek Nowrouzezahrai[1,3] Markus Gross[1,2] Wojciech Jarosz[1]
[1]Disney Research Zürich [2]ETH Zürich [3]Université de Montréal

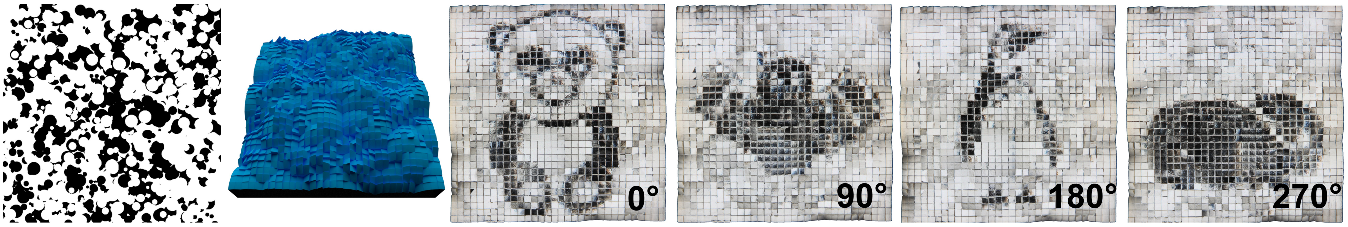


Figure 1: *We automatically design and manufacture magic lenses to warp source images into specified target images. Here we photograph a source image (far left) viewed through a manufactured lens with 32×32 facets (left), resulting in four images depending on the lens' orientation atop the source.*

Abstract

We present an automatic approach to design and manufacture passive display devices based on optical hidden image decoding. Motivated by classical steganography techniques we construct *Magic Lenses*, composed of refractive lenslet arrays, to reveal hidden images when placed over potentially unstructured printed or displayed source images. We determine the refractive geometry of these surfaces by formulating and efficiently solving an inverse light transport problem, taking into account additional constraints imposed by the physical manufacturing processes. We fabricate several variants on the basic magic lens idea including using a single source image to encode several hidden images which are only revealed when the lens is placed at prescribed orientations on the source image or viewed from different angles. We also present an important special case, the *universal lens*, that forms an injection mapping from the lens surface to the source image grid, allowing it to be used with arbitrary source images. We use this type of lens to generate hidden animation sequences. We validate our simulation results with many real-world manufactured magic lenses, and experiment with two separate manufacturing processes.

CR Categories: I.3.7 [Image and Video Processing]: Novel Display Technologies—Multi-View and 3D

Keywords: steganography, image morphing, lens fabrication

Links: DL PDF WEB

1 Introduction

Steganographic techniques, from simple hidden message decoders to invisible inks and complex watermarking schemes, have led to active areas of research and have been applied in a wide variety of fields. Searching for and finding structure in unexpected places is also a fun and insightful process. Some common day examples of this expedition include the pursuits of a child armed with only a magnifying glass and their imagination, to a family huddled around a table, completing a jigsaw puzzle.

We leverage and incite this sense of wonder, encountered when inanimate objects suddenly convey a unexpected message or reveal surprising behavior, by combining the ideas of steganography, hands-on physical user manipulation, and structure from unstructured patterns. We design and construct several different types of *Magic Lenses*, using a custom computational procedure, capable of warping both structured and unstructured image sequences into unexpected target images. Our magic lenses are composed of lenslets that, when placed atop an image/video and viewed from prescribed locations, warp the image through refraction to form the desired images specified during lens generation.

We pose secret image encoding as an inverse light transport problem and present a fully-automatic approach for designing and manufacturing various types of magic lenses (see Figure 1). We experiment with various use-cases, for example enabling multiple target images to be warped from a single source image depending on the viewing angle between the user and the lens, or depending on the relative rotation or alignment of the lens and the source (see Section 6 for more results). In addition, while we experiment with two manufacturing processes to generate physical prototypes of hand-sized magic lenses, nothing about our technique precludes more exotic use-cases such as those depicted in Figure 2: e.g., replacing architectural fixtures with large-scale magic mirrors, revealing hidden messages for interactive and exploratory museum exhibitions, sending secret messages that can only be viewed with a user's magic lens, or embedding thin, flexible magic lenses in paper currency as an anti-counterfeiting and validation measure.

We are motivated by recent work on computationally embedding images into physical material properties, classic steganographic techniques such as the Cardan grille, as well as "magic decoder rings" which reveal secret messages already present in the source image using masking or subtractive transmission. In contrast, our lenses use optical refraction (or reflection), and we require little relation between the input and output images as long as the original image contains all the colors of the target image. Furthermore, our approach can be passive, removing the need to carefully design or modify the source image (carrier signal) to encode the secret image. We also present an important special case of a magic lens called a *universal lens* (Section 5) that completely removes the dependence

ACM Reference Format
Papas, M., Houit, T., Nowrouzezahrai, D., Gross, M., Jarosz, W. 2012.
The Magic Lens: Refractive Steganography. *ACM Trans. Graph. 31* 6, Article 186 (November 2012), 10 pages. DOI = 10.1145/2366145.2366205 http://doi.acm.org/10.1145/2366145.2366205.

Lightweight Binocular Facial Performance Capture under Uncontrolled Lighting

Levi Valgaerts[1] * Chenglei Wu[1,2] † Andrés Bruhn[3] ‡ Hans-Peter Seidel[1] § Christian Theobalt[1] ¶
[1]MPI for Informatics [2]Intel Visual Computing Institute [3]University of Stuttgart

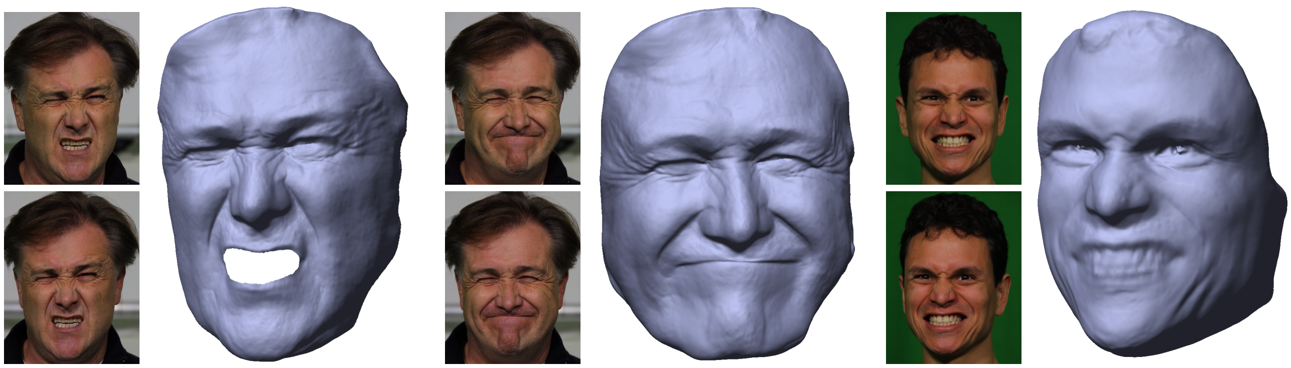


Figure 1: *Three results of our facial performance capture method for two indoor sequences with fast and expressive motion.*

Abstract

Recent progress in passive facial performance capture has shown impressively detailed results on highly articulated motion. However, most methods rely on complex multi-camera set-ups, controlled lighting or fiducial markers. This prevents them from being used in general environments, outdoor scenes, during live action on a film set, or by freelance animators and everyday users who want to capture their digital selves. In this paper, we therefore propose a lightweight passive facial performance capture approach that is able to reconstruct high-quality dynamic facial geometry from only a single pair of stereo cameras. Our method succeeds under uncontrolled and time-varying lighting, and also in outdoor scenes. Our approach builds upon and extends recent image-based scene flow computation, lighting estimation and shading-based refinement algorithms. It integrates them into a pipeline that is specifically tailored towards facial performance reconstruction from challenging binocular footage under uncontrolled lighting. In an experimental evaluation, the strong capabilities of our method become explicit: We achieve detailed and spatio-temporally coherent results for expressive facial motion in both indoor and outdoor scenes – even from low quality input images recorded with a hand-held consumer stereo camera. We believe that our approach is the first to capture facial performances of such high quality from a single stereo rig and we demonstrate that it brings facial performance capture out of the studio, into the wild, and within the reach of everybody.

CR Categories: I.3.7 [COMPUTER GRAPHICS]: Three-Dimensional Graphics and Realism; I.4.1 [IMAGE PROCESSING]: Digitization and Image Capture—Scanning; I.4.8 [IMAGE PROCESSING]: Scene Analysis;

Keywords: Facial Performance Capture, Scene Flow, Shading-based Refinement, Uncontrolled Lighting

*e-mail: valgaerts@mpi-inf.mpg.de
†e-mail: chenglei@mpi-inf.mpg.de
‡e-mail: andres.bruhn@vis.uni-stuttgart.de
§e-mail: hpseidel@mpi-inf.mpg.de
¶e-mail: theobalt@mpi-inf.mpg.de

Links: DL PDF WEB VIDEO DATA

1 Introduction

Two essential features of a realistic virtual actor are a convincingly rendered face and a convincingly animated facial performance. If virtual facial detail is not believably modeled and lit, and if facial motion and expression does not exhibit authentic high spatial and temporal detail, it will not be perceived as realistic. To meet these high quality demands, the research community has developed a variety of facial performance capture techniques that aim to reconstruct very detailed dynamic facial geometry, motion and possibly appearance from sensor measurements of real subjects.

On the one hand, there are active optical systems that use markers, active illumination or invisible paint to capture facial performance [Bickel et al. 2007; Zhang et al. 2004; Furukawa and Ponce 2009]. However, such reconstructions often lack detail and appearance capture is difficult or impossible. On the other hand, passive approaches use multiple cameras and vision-based reconstruction techniques to capture facial performance, e.g. [Bradley et al. 2010]. Reconstructions are of high quality, but pore-level detail is often missing. Moreover, accumulating drift makes it hard to capture very expressive motion. Active lighting methods can bring out pore-level shape detail, but the price to be paid is a complex controlled light and camera set-up [Vogiatzis and Hernández 2011; Wilson et al. 2010]. In other words, to capture facial performance with high-quality spatial and temporal detail, current state-of-the-art techniques require a large number of cameras in a controlled indoor environment, possibly actively controlled illumination, and in many cases some form of active interference with the scene.

ACM Reference Format
Valgaerts, L., Wu, C., Bruhn, A., Seidel, H., Theobalt, C. 2012. Lightweight Binocular Facial Performance Capture under Uncontrolled Lighting. *ACM Trans. Graph. 31* 6, Article 187 (November 2012), 11 pages. DOI = 10.1145/2366145.2366206 http://doi.acm.org/10.1145/2366145.2366206.

Accurate Realtime Full-body Motion Capture Using a Single Depth Camera

Xiaolin Wei Peizhao Zhang Jinxiang Chai*
Texas A&M University

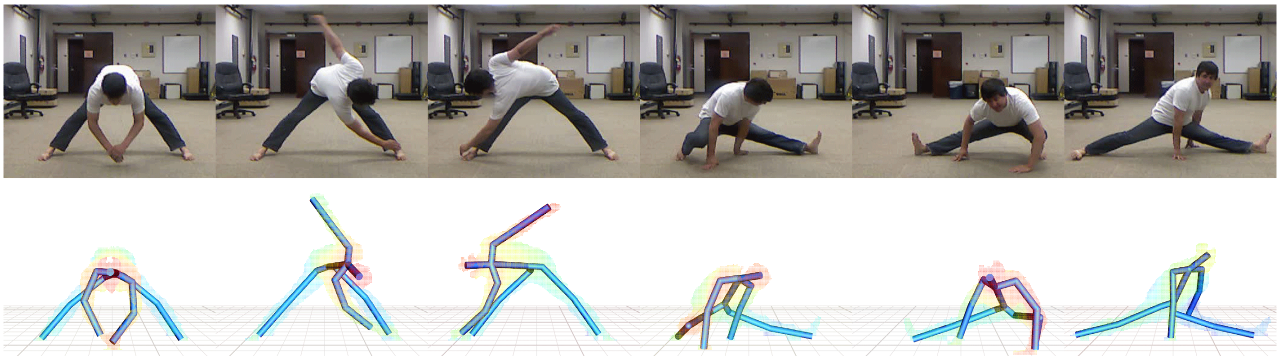


Figure 1: *Our system automatically and accurately reconstructs 3D skeletal poses in real time using monocular depth data obtained from a single camera. (top) reference image data; (bottom) the reconstructed poses overlaying depth data.*

Abstract

We present a fast, automatic method for accurately capturing full-body motion data using a single depth camera. At the core of our system lies a realtime registration process that accurately reconstructs 3D human poses from single monocular depth images, even in the case of significant occlusions. The idea is to formulate the registration problem in a *Maximum A Posteriori* (MAP) framework and iteratively register a 3D articulated human body model with monocular depth cues via linear system solvers. We integrate depth data, silhouette information, full-body geometry, temporal pose priors, and occlusion reasoning into a unified MAP estimation framework. Our 3D tracking process, however, requires manual initialization and recovery from failures. We address this challenge by combining 3D tracking with 3D pose detection. This combination not only automates the whole process but also significantly improves the robustness and accuracy of the system. Our whole algorithm is highly parallel and is therefore easily implemented on a GPU. We demonstrate the power of our approach by capturing a wide range of human movements in real time and achieve state-of-the-art accuracy in our comparison against alternative systems such as *Kinect* [2012].

Keywords: motion capture, performance-based animation, human motion tracking, 3D pose detection, vision-based motion modeling

Links: DL PDF

*e-mail: xwei|stapz|jchai@cs.tamu.edu

ACM Reference Format
Wei, X., Zhang, P., Chai, J. 2012. Accurate Realtime Full-body Motion Capture Using a Single Depth Camera. *ACM Trans. Graph. 31* 6, Article 188 (November 2012), 12 pages. DOI = 10.1145/2366145.2366207 http://doi.acm.org/10.1145/2366145.2366207.

1 Introduction

The ability to accurately reconstruct 3D human poses in real time would allow interactive control of human characters in games, virtual environments, and teleconferences. Such a system would also facilitate natural user interaction with computers, robots, and machines. A major milestone in realtime full-body pose reconstruction is achieved by the recent launch of Microsoft *Kinect* [2012], which automatically infers 3D joint positions from a single depth image. The *Kinect* system is appealing because it is robust, superfast, low-cost, and fully automatic. In addition, it is non-intrusive and easy to set up because the system requires no markers, no motion sensors, and no special suits.

Kinect, despite its high popularity and wide applications, does not provide an accurate reconstruction for complex full-body movements. In particular, the system often produces poor results when significant occlusions occur. This significantly limits the application of *Kinect* because depth data from a single camera frequently contains significant occlusions. The primary contribution of this paper is a robust full-body motion capture system that accurately reconstructs 3D poses even in the case of significant occlusions (see Figure 1). Our system advances the state of the art because it shares the same advantages of *Kinect* but significantly improves the accuracy of the capturing system.

At the core of our system lies a realtime full-body motion tracking process that accurately reconstructs 3D skeletal poses using monocular depth images obtained from a single camera. The idea is to formulate the tracking problem in a *Maximum A Posteriori* (MAP) framework and iteratively register a 3D articulated human model with depth data via linear system solvers. The system is accurate because we integrate depth data, silhouette information, full-body geometry, and temporal pose priors into a unified framework. In addition, we incorporate occlusion reasoning into MAP estimation in order to handle significant occlusions caused by a single camera.

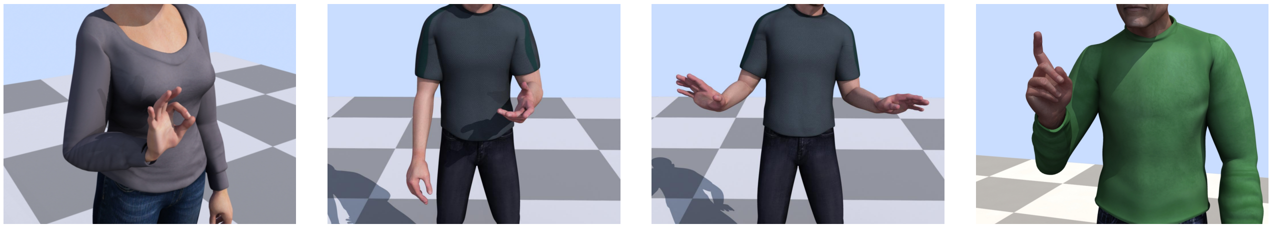
Our 3D pose tracking process, however, requires manual initialization and might be prone to local minima because it initializes current poses with previously reconstructed poses. We address this challenge by complementing tracking with 3D pose detection. 3D pose tracking and detection are complementary to each other. At

Data-driven Finger Motion Synthesis for Gesturing Characters

Sophie Jörg*
Carnegie Mellon University
Clemson University

Jessica Hodgins†
Carnegie Mellon University
Disney Research, Pittsburgh

Alla Safonova‡
Disney Research, Pittsburgh



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

Figure 1: *Animations with synthesized finger motions: (a) ok gesture, (b)-(c) extracts from a conversation, (d) attention gesture.*

Abstract

Capturing the body movements of actors to create animations for movies, games, and VR applications has become standard practice, but finger motions are usually added manually as a tedious post-processing step. In this paper, we present a surprisingly simple method to automate this step for gesturing and conversing characters. In a controlled environment, we carefully captured and post-processed finger and body motions from multiple actors. To augment the body motions of virtual characters with plausible and detailed finger movements, our method selects finger motion segments from the resulting database taking into account the similarity of the arm motions and the smoothness of consecutive finger motions. We investigate which parts of the arm motion best discriminate gestures with leave-one-out cross-validation and use the result as a metric to select appropriate finger motions. Our approach provides good results for a number of examples with different gesture types and is validated in a perceptual experiment.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: finger motions, human animation, gesture synthesis, hand motions, virtual characters

Links: DL PDF

*e-mail: sjoerg@clemson.edu
†e-mail: jkh@cs.cmu.edu
‡e-mail: alla.safonova5@gmail.com

ACM Reference Format
Jörg, S., Hodgins, J., Safonova, A. 2012. Data-driven Finger Motion Synthesis for Gesturing Characters. *ACM Trans. Graph. 31* 6, Article 189 (November 2012), 7 pages. DOI = 10.1145/2366145.2366208 http://doi.acm.org/10.1145/2366145.2366208.

1 Introduction

Hand and finger motions are omnipresent in our daily life. We use our hands to interact with objects and with each other. When communicating, we use them to punctuate our speech, point to an object of interest, signal our opinion, and convey emotions. Communication – and therefore hand and finger motions – also plays an increasingly important role in digital applications. Examples of such applications include virtual worlds, such as Second Life, digital games, such as L.A. Noire, or any type of teaching, training, or advice application that uses embodied conversational agents. If we want to create believable and compelling virtual characters, we therefore need to generate convincing hand motions.

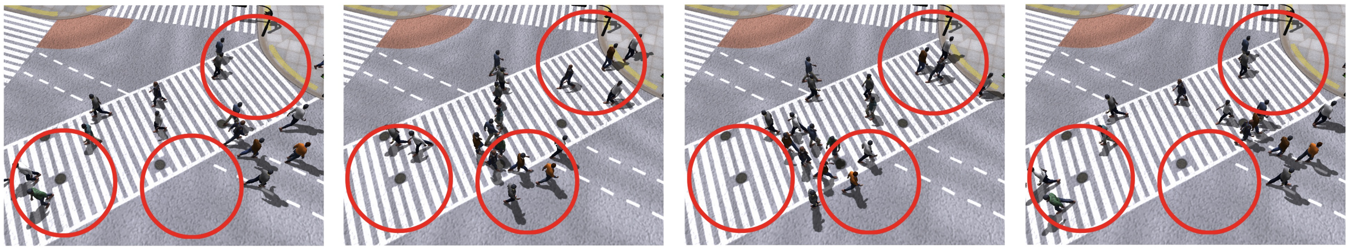
Motion capture has become a widely used technique for animating realistic and human-like characters for movies or games. However, the elaborate motions of the fingers with their high number of degrees of freedom and small size are still not easy to capture. As a consequence, hands are rarely captured in their full complexity. In general, two or three markers on fingertips (e.g., on the thumb, index, and pinky) and a few on the back of the hand are used as a reference [Kitagawa and Windsor 2008] and the fingers are keyframed manually, which is a tedious, labor-intensive process.

In this paper, we propose a method to automatically add plausible finger motions to body motions (Figure 1). Our algorithm uses a previously recorded database of body and finger motions to generate finger movements that match an input body motion (Figure 2). We locate suitable finger motions in the database using a metric based on the similarity of the arm movements and the smoothness of the reconstructed finger motions. Our key observation is that plausible finger motions can be inferred from the wrist motion. Our approach preserves the naturalness and subtlety of motion captured movements, without requiring time-consuming manual post-processing. As our method is intended as a post-processing step, it does not impact the motion capture session, saving valuable production time. We apply our algorithm to gesturing and conversing characters. It is not intended for object manipulations, such as grasping, where the fingers must be precisely positioned. We demonstrate our approach with several example databases, consisting of gestures, casual conversations, debates, and giving directions.

Our main contribution is the development of a data-driven method to automatically add realistic and detailed finger motions to the body motions of conversational characters. We explore the validity of several parameters as predictors for the similarity of finger

A Statistical Similarity Measure for Aggregate Crowd Dynamics

Stephen J. Guy* UNC-Chapel Hill
Jur van den Berg University of Utah
Wenxi Liu, Rynson Lau City University of Hong Kong
Ming C. Lin, Dinesh Manocha UNC-Chapel Hill



(a) Real-world Data (b) Entropy Metric: 4.7 (c) Entropy Metric: 3.8 (d) Entropy Metric: 2.7

Figure 1: *A comparison between a rendering of real-world crowd data (a), and stills from three different simulation algorithms applied to the same scenario (b-d). Our entropy metric is used to measure the similarity of simulation algorithm to real-world data. A small value of the metric, as in (d), indicates a better match to the data. Differences between the simulations are highlighted with circles.*

Abstract

We present an information-theoretic method to measure the similarity between a given set of observed, real-world data and visual simulation technique for aggregate crowd motions of a complex system consisting of many individual agents. This metric uses a two-step process to quantify a simulator's ability to reproduce the collective behaviors of the whole system, as observed in the recorded real-world data. First, Bayesian inference is used to estimate the simulation states which best correspond to the observed data, then a maximum likelihood estimator is used to approximate the prediction errors. This process is iterated using the EM-algorithm to produce a robust, statistical estimate of the magnitude of the prediction error as measured by its entropy (smaller is better). This metric serves as a simulator-to-data similarity measurement. We evaluated the metric in terms of robustness to sensor noise, consistency across different datasets and simulation methods, and correlation to perceptual metrics.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation; I.6.4 [Simulation and Modeling]: Model Validation and Analysis; I.2.10 [Artificial Intelligence]: Vision and Scene Understanding—Perceptual Reasoning.

Keywords: crowd simulation, validation, data-driven simulations

Links: DL PDF WEB VIDEO

*Email: {sjguy,lin,dm}@cs.unc.edu. The first author is currently an Assistant Professor at the University of Minnesota. This work was done while the second and the third authors were at the University of North Carolina (UNC-Chapel Hill).

ACM Reference Format
Guy, S., van den Berg, J., Liu, W., Lau, R., Lin, M., Manocha, D. 2012. A Statistical Similarity Measure for Aggregate Crowd Dynamics. *ACM Trans. Graph. 31* 6, Article 190 (November 2012), 11 pages.
DOI = 10.1145/2366145.2366209 http://doi.acm.org/10.1145/2366145.2366209.

1 Introduction

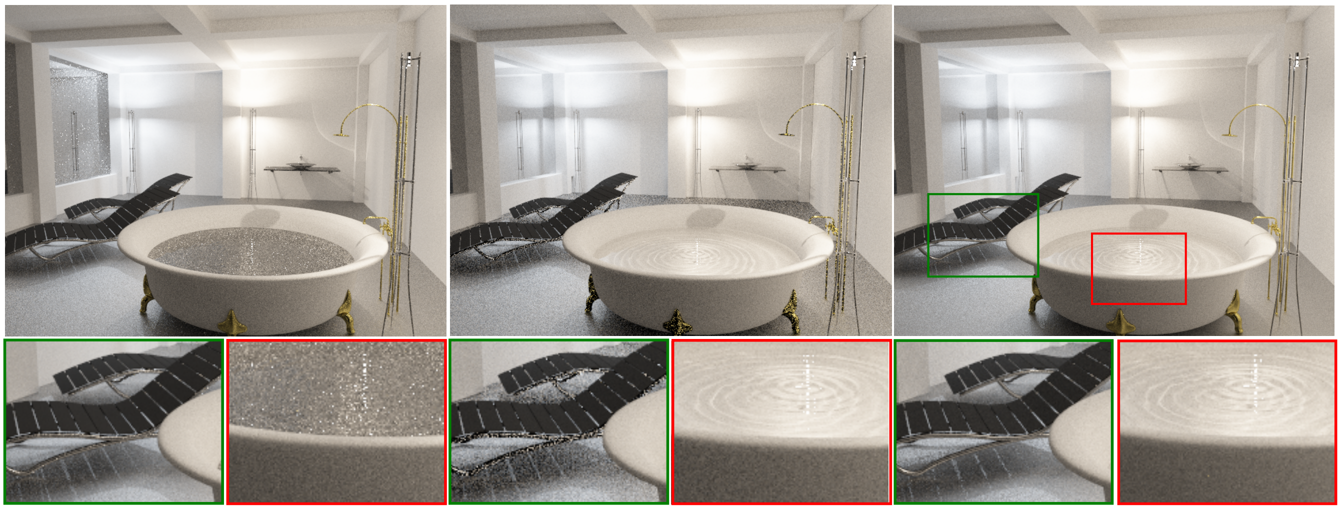
Visual simulation of aggregates systems, including human crowds, animal herds, and insect swarms is a growing area of interest in computer graphics, with applications in diverse areas such as social sciences, swarm intelligence, and city planning. For applications in entertainment, providing artists and animators with high-level control while maintaining visual plausibility of motion is often sufficient. However, for many other training and planning applications, such as virtual reality based training, fire-safety planning, and crowd control and management, it is often critical to model accurate motion, in addition to producing a compelling visual rendering. In this context, we define a measure of a simulator's accuracy based on the similarity of the motion from the simulator to the motion captured in real-world observations. While some previous work has studied the visual plausibility of simulation techniques, we present a new metric for quantifying the similarity between a set of real-world observations and any algorithm designed to simulate the aggregate crowd dynamics captured in the data.

Evaluating the correctness or predictability of the results from a crowd simulation method presents several interesting challenges, many of which arise from the inherent nature of a crowd as a *complex system*. Complex systems are systems composed of several components or elements that interact to exhibit emergent patterns that cannot be easily predicted from the properties of the individual components alone [Schadschneider et al. 2011; Gallagher and Appenzeller 1999]. Because of issues inherent in these systems, such as uncertainty and non-determinism, the study of complex systems generally must draw on techniques from the fields of statistics, information theory, and non-linear dynamics. We likewise draw on inspiration from these fields, in proposing a new method to compare aggregate simulation methods with real-world data that accounts for these challenges.

Real-world data of crowds is becoming increasingly common, driven in part by recent improvement in sensor technology, such as LiDAR and GPS; the proliferation of high-resolution cameras; and advances in computer vision and motion tracking. However, several aspects of a crowd and its aggregate motions make it difficult to directly compare such data against any simulation results. For example, given two very similar initial states, a small crowd can reach two very different configurations after just a few seconds, because the effects of small changes in states can quickly compound

A Path Space Extension for Robust Light Transport Simulation

Toshiya Hachisuka[1,3] Jacopo Pantaleoni[2] Henrik Wann Jensen[3]
[1]Aarhus University [2]NVIDIA Research [3]UC San Diego



Monte Carlo Path Integration (RMSE: 0.07836) Photon Density Estimation (RMSE: 0.04638) Unified Framework (RMSE: 0.01246)

Figure 1: *Equal-time comparison of rendered images of a bathroom scene with realistic lighting fixtures. This scene includes both glossy reflections and complex caustics due to lighting fixtures which are common in interior design. Existing light transport simulation methods including Monte Carlo path integration and photon density estimation cannot efficiently render scenes with such lighting phenomena. Our new framework for light transport simulation automatically combines Monte Carlo path integration and photon density estimation by extending the sampling space of light transport paths, and produces a significantly more accurate solution in the same rendering time.*

Abstract

We present a new sampling space for light transport paths that makes it possible to describe Monte Carlo path integration and photon density estimation in the same framework. A key contribution of our paper is the introduction of vertex perturbations, which extends the space of paths with loosely coupled connections. The new framework enables the computation of path probabilities in the same space under the same measure, which allows us to use multiple importance sampling to combine Monte Carlo path integration and photon density estimation. The resulting algorithm, *unified path sampling*, can robustly render complex combinations and glossy surfaces and caustics that are problematic for existing light transport simulation methods.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Raytracing;

Keywords: global illumination, multiple importance sampling

Links: DL PDF

ACM Reference Format
Hachisuka, T., Pantaleoni, J., Jensen, H. 2012. A Path Space Extension for Robust Light Transport Simulation. *ACM Trans. Graph. 31* 6, Article 191 (November 2012), 10 pages. DOI = 10.1145/2366145.2366210 http://doi.acm.org/10.1145/2366145.2366210.

1 Introduction

Efficiently simulating light transport under general scene configurations is a difficult task. Currently, the most successful approaches are based on Monte Carlo path integration and photon density estimation. Both approaches solve the rendering equation introduced by Kajiya [1986]. Unfortunately, neither Monte Carlo path integration nor photon density estimation can simulate *all* types of light transport efficiently. For example, Monte Carlo path integration has problems with specular-diffuse-specular (SDS) transport, while photon density estimation techniques can suffer in the presence of highly glossy materials.

Recently, Hachisuka and Jensen [2009] introduced stochastic progressive photon mapping. They combined photon density estimation and distributed ray tracing [Cook et al. 1984] to solve the rendering equation in general scenes. This algorithm can handle both SDS paths and glossy reflections efficiently, however, the algorithm uses a manual classification of surface materials into either “glossy” or “diffuse”. This binary classification can lead to inefficient sampling in some scenes such as glossy surfaces lit by a directional light source.

A key challenge in combining density estimation and Monte Carlo path integration is the lack of a shared theoretical foundation for both approaches. It is well known that both approaches solve the rendering equation. However, while Monte Carlo path integration has a solid theoretical foundation with the path integral formulation [Veach 1998], photon density estimation has not been formulated under the same foundation. This difference in their theoretical formulations limits the ability to analyze both methods in the same framework.

In this paper, we propose a novel path sampling framework for light transport. The key contribution is a new extension of the path space

Light Transport Simulation with Vertex Connection and Merging

Iliyan Georgiev
Saarland University
Intel VCI, Saarbrücken

Jaroslav Křivánek
Charles University in Prague
Faculty of Mathematics and Physics

Tomáš Davidovič
Saarland University
Intel VCI, Saarbrücken

Philipp Slusallek
Saarland University
Intel VCI & DFKI, Saarbrücken

Figure 1: *A comparison of our new progressive vertex connection and merging (VCM) algorithm against bidirectional path tracing (BPT) and stochastic progressive photon mapping (PPM) after 30 minutes of rendering. BPT fails to reproduce the reflected caustics produced by the vase, while PPM has difficulties in handling the illumination coming from the room seen in the mirror. Our new VCM algorithm automatically computes a good mixture of sampling techniques from BPT and PPM to robustly capture the entire illumination in the scene. The rightmost column shows the relative contributions of the BPT and PPM techniques to the VCM image in false color.*

Abstract

Developing robust light transport simulation algorithms that are capable of dealing with arbitrary input scenes remains an elusive challenge. Although efficient global illumination algorithms exist, an acceptable approximation error in a reasonable amount of time is usually only achieved for specific types of input scenes. To address this problem, we present a reformulation of photon mapping as a bidirectional path sampling technique for Monte Carlo light transport simulation. The benefit of our new formulation is twofold. First, it makes it possible, for the first time, to explain in a formal manner the relative efficiency of photon mapping and bidirectional path tracing, which have so far been considered conceptually incompatible solutions to the light transport problem. Second, it allows for a seamless integration of the two methods into a more robust combined rendering algorithm via multiple importance sampling. A progressive version of this algorithm is consistent and efficiently handles a wide variety of lighting conditions, ranging from direct illumination, diffuse and glossy inter-reflections, to specular-diffuse-specular light transport. Our analysis shows that this algorithm inherits the high asymptotic performance from bidirectional path tracing for most light path types, while benefiting from the efficiency of photon mapping for specular-diffuse-specular lighting effects.

ACM Reference Format
Georgiev, I., Krivánek, J., Davidovic, T., Slusallek, P. 2012. Bidirectional Light Transport with Vertex Connection and Merging. *ACM Trans. Graph. 31* 6, Article 192 (November 2012), 10 pages.
DOI = 10.1145/2366145.2366211 http://doi.acm.org/10.1145/2366145.2366211.

CR Categories: I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Realism—Raytracing

Keywords: light transport, global illumination, importance sampling, bidirectional path tracing, photon mapping, density estimation

Links: DL PDF WEB

1 Introduction

Light transport simulation is a central problem in photo-realistic image synthesis. It has been an active area of research for decades due to its utility in many applications, including architectural visualization, industrial design, as well as the entertainment industry. In the past years, considerable advances have been made with respect to the efficiency of light transport algorithms, but the improvements usually come with some sort of bias: often, some types of light interactions are disregarded, or handled inefficiently. Such approximations are sometimes acceptable, but often lead to a severe loss of image fidelity (see e.g. [Křivánek et al. 2010]). Developing truly robust light transport algorithms that can efficiently and accurately render a wide variety of scenes remains an important challenge that we address in this paper.

Bidirectional path tracing (BPT) [Lafortune and Willems 1993; Veach and Guibas 1994] is among the most versatile light transport algorithms. The true key to its robustness is the provably good combination of various path sampling techniques using multiple importance sampling (MIS) [Veach and Guibas 1995]. It has been, however, widely acknowledged that BPT is not efficient for transport paths with specular-diffuse-specular (SDS) configurations, where the notion of 'specular' also includes sharp glossy interactions. This is indeed an important practical limitation, because such paths occur in all scenes containing specular objects and their image contribution is especially important in some very common cases such as an object enclosed in glass, an interior of a car or a building, etc. The reason for this problem is that the path sampling techniques in BPT usually

Practical Hessian-Based Error Control for Irradiance Caching

Jorge Schwarzhaupt* UC San Diego
Henrik Wann Jensen† UC San Diego
Wojciech Jarosz‡ Disney Research Zürich

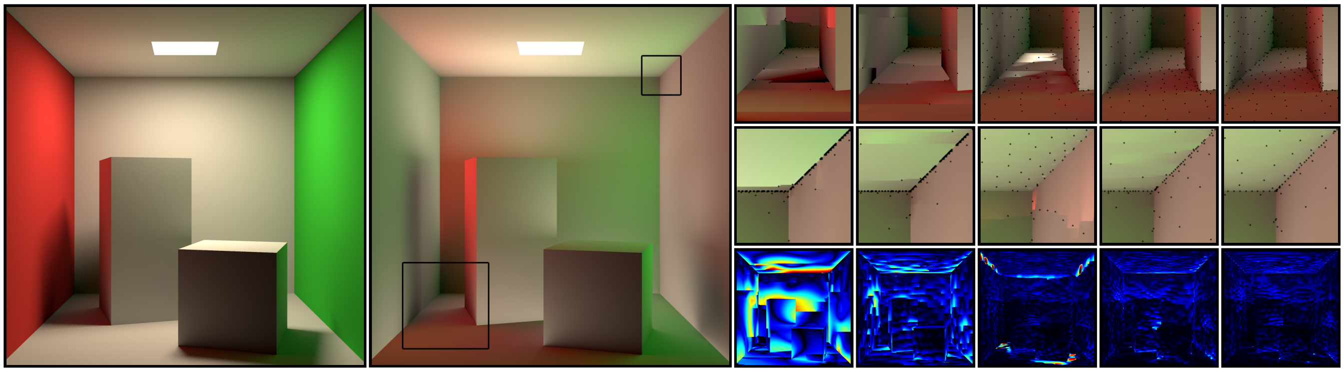


Figure 1: *Our new Occlusion Hessian significantly outperforms both the Pure and the Bounded Split-Sphere (clamped to the gradient and 150px max spacing) for irradiance caching. It also performs significantly better than the recently published occlusion-unaware Hessian error metrics by Jarosz et al. [2012].*

Abstract

This paper introduces a new error metric for irradiance caching that significantly outperforms the classic Split-Sphere heuristic. Our new error metric builds on recent work using second order gradients (Hessians) as a principled error bound for the irradiance. We add occlusion information to the Hessian computation, which greatly improves the accuracy of the Hessian in complex scenes, and this makes it possible for the first time to use a radiometric error metric for irradiance caching. We enhance the metric making it based on the relative error in the irradiance as well as robust in the presence of black occluders. The resulting error metric is efficient to compute, numerically robust, supports elliptical error bounds and arbitrary hemispherical sample distributions, and unlike the Split-Sphere heuristic it is not necessary to arbitrarily clamp the computed error thresholds. Our results demonstrate that the new error metric outperforms existing error metrics based on the Split-Sphere model and occlusion-unaware Hessians.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Raytracing

Keywords: Global illumination, Irradiance caching, Monte Carlo ray tracing

Links: DL PDF WEB

*jschwarz@cs.ucsd.edu
†henrik@cs.ucsd.edu
‡wjarosz@disneyresearch.com

ACM Reference Format
Schwarzhaupt, J., Jensen, H., Jarosz, W. 2012. Practical Hessian-Based Error Control for Irradiance Caching. *ACM Trans. Graph. 31* 6, Article 193 (November 2012), 10 pages. DOI = 10.1145/2366145.2366212 http://doi.acm.org/10.1145/2366145.2366212.

1 Introduction

Computing indirect illumination in complex scenes both accurately and efficiently has been a long-standing challenge in computer graphics. Since its introduction by Ward et al. over twenty years ago, irradiance caching [Ward et al. 1988; Ward and Heckbert 1992] has become a very popular choice for accelerating the computation of diffuse indirect illumination. Irradiance caching exploits the fact that indirect illumination varies slowly across diffuse surfaces by computing it accurately only at a sparse set of locations and interpolating (or extrapolating) between the computed and cached irradiance values whenever possible.

The decision of whether to interpolate or not is central to the quality and efficiency of the algorithm. Ward et al. proposed to determine the cache point radii using the "Split-Sphere" heuristic which approximates an upper-bound on the rate of change of irradiance at the cache points. Unfortunately there are many common failure cases where this basic approach produces unacceptable results. This has led to several modifications and additions to the original metric to more robustly deal with such failure cases, by e.g. clamping to minimum/maximum radii, constraining the radius based on the gradient, enforcing the triangle inequality between cache records, and many more [Křivánek and Gautron 2009]. Though these additions can lead to a more robust approach, they are inherently trying to augment a sub-optimal heuristic and, in the process, introducing many more parameters, which makes it more difficult to control the algorithm.

Instead of modifying the original heuristic, Jarosz et al. [2012] recently proposed an alternative. While analyzing global illumination in 2D, they derived second order derivatives of irradiance for scenes with no occlusions, and showed how this could potentially be applied to obtain an improved error metric for irradiance caching. The core idea was to use a second-order Taylor expansion as a principled error term for the gradient extrapolation used in irradiance caching. By bounding the allowed error, they showed that the cache point radii could be derived from an irradiance Hessian instead of the Split-Sphere heuristic, without relying excessively on clamping and other corrections. Though they demonstrated that this idea shows promise, their preliminary investigation fell short of a full practical

SURE-based Optimization for Adaptive Sampling and Reconstruction

Tzu-Mao Li Yu-Ting Wu Yung-Yu Chuang

National Taiwan University

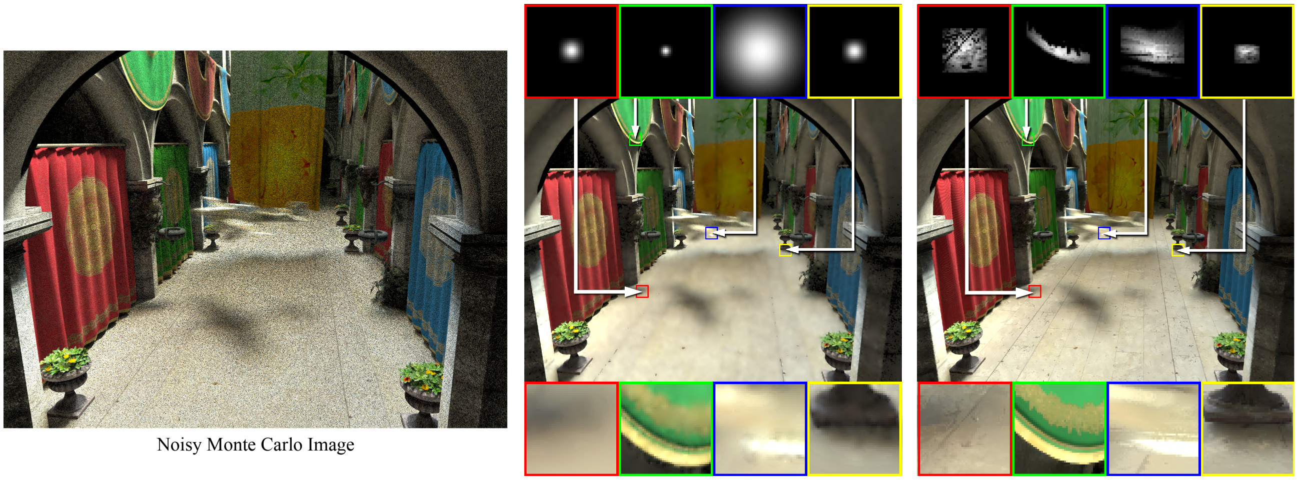


Figure 1: *Comparisons between greedy error minimization (GEM) [Rousselle et al. 2011] and our SURE-based filtering. With SURE, we are able to use kernels (cross bilateral filters in this case) that are more effective than GEM's isotropic Gassians. Thus, our approach better adapts to anisotropic features (such as the motion blur pattern due to the motion of the airplane) and preserves scene details (such as the textures on the floor and curtains). The kernels of both methods are visualized for comparison.*

Abstract

We apply Stein's Unbiased Risk Estimator (SURE) to adaptive sampling and reconstruction to reduce noise in Monte Carlo rendering. SURE is a general unbiased estimator for mean squared error (MSE) in statistics. With SURE, we are able to estimate error for an arbitrary reconstruction kernel, enabling us to use more effective kernels rather than being restricted to the symmetric ones used in previous work. It also allows us to allocate more samples to areas with higher estimated MSE. Adaptive sampling and reconstruction can therefore be processed within an optimization framework. We also propose an efficient and memory-friendly approach to reduce the impact of noisy geometry features where there is depth of field or motion blur. Experiments show that our method produces images with less noise and crisper details than previous methods.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—RayTracing.

Keywords: Sampling, reconstruction, ray tracing, cross bilateral filter, Stein's unbiased risk estimator (SURE).

Links: DL PDF WEB

ACM Reference Format
Li, T., Wu, Y., Chuang, Y. 2012. SURE-based Optimization for Adaptive Sampling and Reconstruction. *ACM Trans. Graph. 31* 6, Article 194 (November 2012), 9 pages. DOI = 10.1145/2366145.2366213 http://doi.acm.org/10.1145/2366145.2366213.

1 Introduction

Monte Carlo (MC) integration is a common technique for rendering images with distributed effects such as antialiasing, depth of field, motion blur, and global illumination. It simulates a variety of sophisticated light transport paths in a unified manner; it estimates pixel values by using stochastic point samples in the integral domain. Despite its generality and simplicity, however, the MC approach converges slowly. A complex scene with multiple distributed effects usually requires several thousand expensive samples per pixel to produce a noise-free image.

Adaptive sampling and reconstruction (or filtering, used interchangeably in the paper) are two effective techniques for reducing noise. Given a fixed budget of samples, adaptive sampling determines the optimal sample distribution by concentrating more samples on difficult regions. To decide which pixels are worth more effort, we require a robust criterion for measuring errors. Accurate estimation of errors is challenging in our application because the ground truth is not available. Reconstruction algorithms, in contrast, properly construct smooth results from the discrete samples at hand. One key issue that reconstruction must resolve is how to select the filters for each pixel, as the optimal reconstruction kernels are usually spatially-varying and anisotropic. Recently, approaches have been developed to address the challenge of spatially-varying filters [Chen et al. 2011; Rousselle et al. 2011], producing better results than those that use a single filter across the whole image. However, these methods are limited to symmetric filters and do not work well for scenes with anisotropic features such as high-frequency textures on the floor and curtains in Figure 1.

We here propose an adaptive sampling and reconstruction algorithm to improve the efficiency of Monte Carlo ray tracing. The core idea is to adopt Stein's Unbiased Risk Estimator (SURE) [Stein 1981], a general unbiased estimator for mean squared error (MSE)

Adaptive Rendering with Non-Local Means Filtering

Fabrice Rousselle* University of Bern
Claude Knaus University of Bern
Matthias Zwicker University of Bern

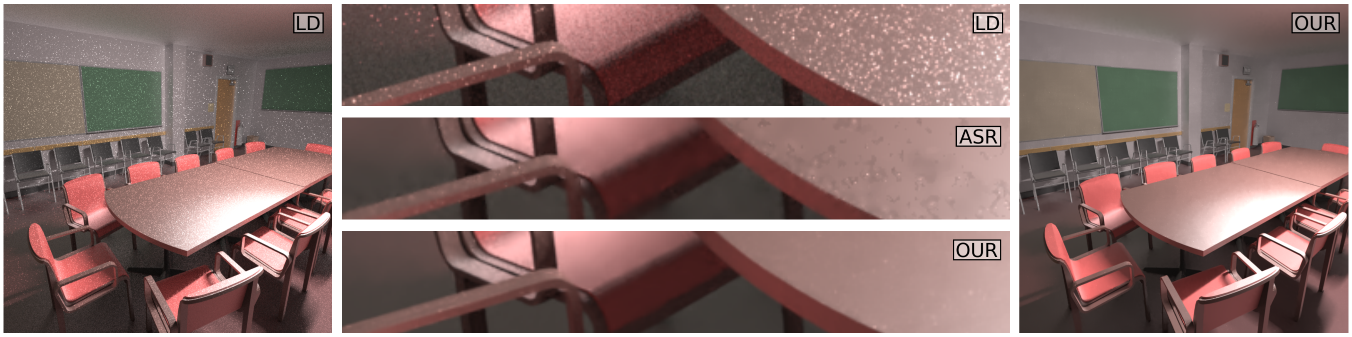


Figure 1: *Our adaptive rendering method handles arbitrary light transport effects, employs a state of the art denoising filter, and incurs little computational overhead. The "conference" scene includes indirect illumination and moderately glossy surfaces. We compare path tracing using low-discrepancy sampling (far left) to our approach (far right). The close-ups in the middle show low-discrepancy sampling (LD), adaptive sampling and reconstruction (ASR) by Rousselle et al. [2011], and our approach (OUR), all at equal rendering time.*

Abstract

We propose a novel approach for image space adaptive sampling and filtering in Monte Carlo rendering. We use an iterative scheme composed of three steps. First, we adaptively distribute samples in the image plane. Second, we denoise the image using a non-linear filter. Third, we estimate the residual per-pixel error of the filtered rendering, and the error estimate guides the sample distribution in the next iteration. The effectiveness of our approach hinges on the use of a state of the art image denoising technique, which we extend to an adaptive rendering framework. A key idea is to split the Monte Carlo samples into two buffers. This improves denoising performance and facilitates variance and error estimation. Our method relies only on the Monte Carlo samples, allowing us to handle arbitrary light transport and lens effects. In addition, it is robust to high noise levels and complex image content. We compare our approach to a state of the art adaptive rendering technique based on adaptive bandwidth selection and demonstrate substantial improvements in terms of both numerical error and visual quality. Our framework is easy to implement on top of standard Monte Carlo renderers and it incurs little computational overhead.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Raytracing;

Keywords: adaptive sampling and reconstruction

Links: DL PDF

*e-mail: roussell@iam.unibe.ch

ACM Reference Format
Rousselle, F., Knaus, C., Zwicker, M. 2012. Adaptive Rendering with Non-Local Means Filtering. *ACM Trans. Graph. 31* 6, Article 195 (November 2012), 11 pages. DOI = 10.1145/2366145.2366214 http://doi.acm.org/10.1145/2366145.2366214.

1 Introduction

In this paper we present an algorithm to reduce numerical error and visual artifacts in Monte Carlo rendering. Our approach relies on image space adaptive sampling and filtering, which is attractive because it makes few assumptions about the specific light transport or lens effects being rendered. Recent techniques that follow this strategy [Overbeck et al. 2009; Rousselle et al. 2011] have been shown to effectively improve the visual quality over standard sampling and filtering for a wide range of effects that are prone to noise artifacts, such as indirect illumination, soft shadows, participating media, depth of field, or motion blur. We observe, however, that previous work in adaptive rendering is based on image denoising techniques that are not competitive with the state of the art in image processing. For example, Overbeck et al.'s approach [Overbeck et al. 2009] uses straightforward wavelet shrinkage, and Rousselle et al.'s algorithm [Rousselle et al. 2011] is based on adaptive bandwidth selection with isotropic Gaussian filters. Both techniques leave ample room for improvement, as we show in this paper.

The main contribution of our approach is to extend an image denoising filter that is competitive with the state of the art in image processing and employ it in an adaptive rendering framework. We show that our approach is more robust under severe noise than previous techniques, significantly reducing numerical error and visual artifacts. In addition, our approach is more effective at removing noise even in complex image regions while minimizing the smoothing of image features. Finally, our technique is compatible with efficient low discrepancy sampling while previous techniques assumed random sampling, which affords additional improvements in output quality. Our technique shares the advantages of other image based approaches. It can deal with arbitrary light transport and lens effects, and it only requires the Monte Carlo samples as its input such that it is straightforward to implement it on top of existing renderers.

Our framework builds on an iterative strategy consisting of three components in each iteration step. First, we distribute a given budget of Monte Carlo samples over the image. We sample the image uniformly in the initial iteration step, while consecutive iterations employ adaptive sampling. Second, we filter the image to reduce

Elasticity-Inspired Deformers for Character Articulation

Ladislav Kavan* Olga Sorkine

ETH Zurich

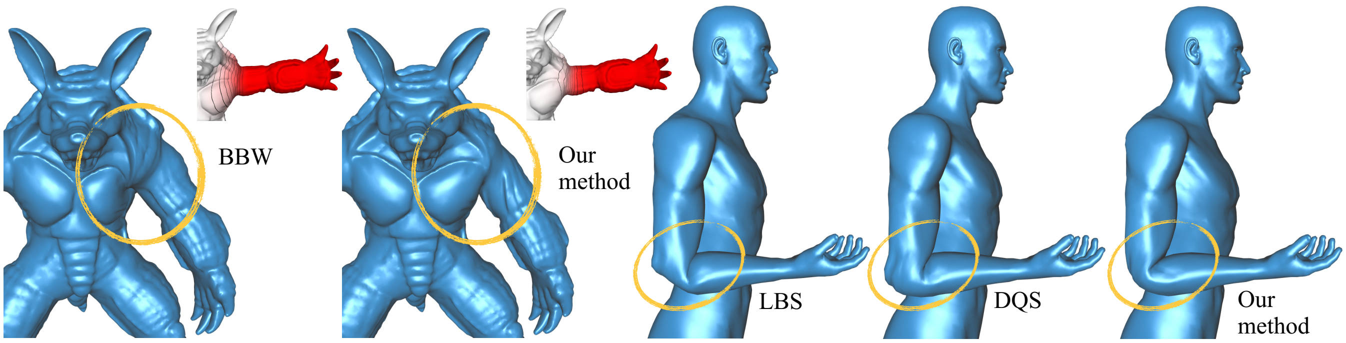


Figure 1: *Previous automatic weighting schemes, such as bounded biharmonic weights (BBW) are deformation-agnostic, whereas our method optimizes weights to minimize shape distortion. Combined with our joint-based deformers, we achieve higher-quality results than both linear blend (LBS) and dual quaternion skinning (DQS). Our technique requires no additional user input and its speed is competitive to linear skinning.* © *2012 The Authors*

Abstract

Current approaches to skeletally-controlled character articulation range from real-time, closed-form skinning methods to offline, physically-based simulation. In this paper, we seek a closed-form skinning method that approximates nonlinear elastic deformations well while remaining very fast. Our contribution is two-fold: (1) we optimize skinning weights for the standard linear and dual quaternion skinning techniques so that the resulting deformations minimize an elastic energy function. We observe that this is not sufficient to match the visual quality of the original elastic deformations and therefore, we develop (2) a new skinning method based on the concept of *joint-based deformers*. We propose a specific deformer which is visually similar to nonlinear variational deformation methods. Our final algorithm is fully automatic and requires little or no input from the user other than a rest-pose mesh and a skeleton. The runtime complexity requires minimal memory and computational overheads compared to linear blend skinning, while producing higher quality deformations than both linear and dual quaternion skinning.

CR Categories: I.3.7 [Computer Graphics]: Three Dimensional Graphics and Realism—Animation

Keywords: Skeletal shape deformation, skinning, elasticity.

Links: DL PDF WEB VIDEO

*e-mail: ladislav.kavan@gmail.com

ACM Reference Format
Kavan, L., Sorkine, O. 2012. Elasticity-Inspired Deformers for Character Articulation.
ACM Trans. Graph. 31 6, Article 196 (November 2012), 8 pages. DOI = 10.1145/2366145.2366215
http://doi.acm.org/10.1145/2366145.2366215.

1 Introduction

Real-time approaches for character articulation, known as skinning or skeletal subspace deformation, are necessary for interactive animation interfaces and applications such as computer games and crowd simulation. Direct geometric methods, e.g., linear or dual quaternion skinning deliver the required speed, but often at the cost of compromising quality. Physically-based models produce high-quality, realistic deformations at the expense of much more complex computation. In this paper, we study the question of how to design a real-time, direct skinning method that delivers visually similar results to offline elastic simulation, albeit without collisions.

This problem is also motivated by requests from professional rigging artists who enjoy the improvements offered by dual quaternion skinning, but dislike the joint-bulging artifacts (Fig. 1), that require them to apply manual fix-ups. Physically-based methods deliver high-quality deformations automatically, but even the latest highly optimized algorithms [McAdams et al. 2011] are not fast enough for interactive posing. In this paper, we consider a common physically-based model – an elastic solid with rigid bones embedded inside. An input skeletal pose induces deformations through minimization of a nonlinear elastic energy, subject to rigid bone constraints. While being only a gross simplification of real anatomy, this model yields intuitive, high-quality deformations applicable to both realistic and stylized characters.

Our initial idea to improve the results of closed-form skinning techniques was to find weights that minimize nonlinear elastic energy over a range of skeletal poses. While achieving better results than deformation-agnostic weights, such as bounded biharmonic weights [Jacobson et al. 2011], we found that the results still contained objectionable visual artifacts when compared to elastic simulation. Indeed, rigging artists also experimentally confirmed that the space of traditional skinning deformations is not rich enough to approximate the nonlinear elastic behavior well. We identify the culprit to be the isotropy of linear and dual quaternion blending operators, causing these methods to treat all directions in material-space as equal. This contrasts physics, which, even with homogeneous elasticity, results in quite different deformations along the direction of the bone (which remains rigid) and in the orthogonal directions (where squash and stretch occur).

Simulation of Complex Nonlinear Elastic Bodies using Lattice Deformers

Taylor Patterson
University of Wisconsin-Madison

Nathan Mitchell
University of Wisconsin-Madison

Eftychios Sifakis
University of Wisconsin-Madison

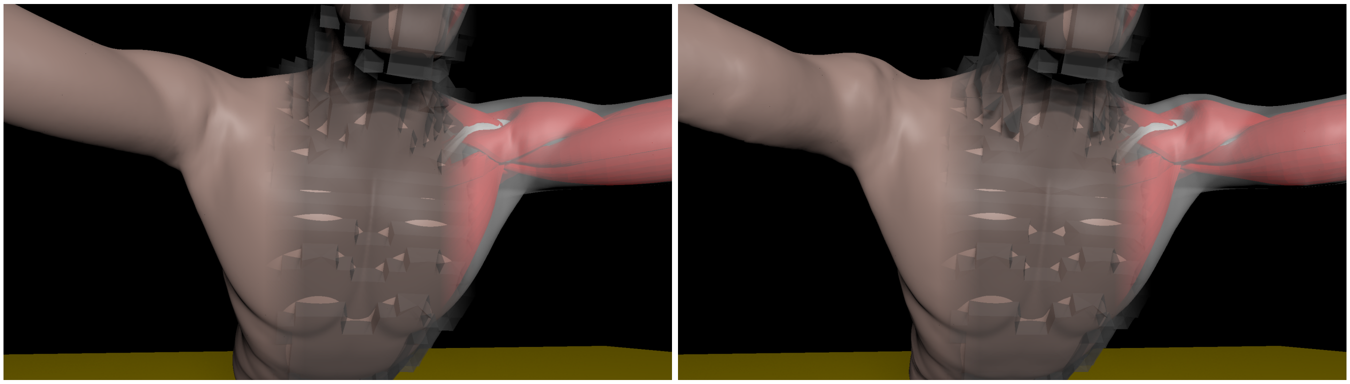


Figure 1: *Anatomic simulation with skin, deformer lattice and embedded muscles shown. Left: Muscles inactive, Right: Muscles fully flexed.*

Abstract

Lattice deformers are a popular option for modeling the behavior of elastic bodies as they avoid the need for conforming mesh generation, and their regular structure offers significant opportunities for performance optimizations. Our work expands the scope of current lattice-based elastic deformers, adding support for a number of important simulation features. We accommodate complex nonlinear, optionally anisotropic materials while using an economical one-point quadrature scheme. Our formulation fully accommodates near-incompressibility by enforcing accurate nonlinear constraints, supports implicit integration for large time steps, and is not susceptible to locking or poor conditioning of the discrete equations. Additionally, we increase the accuracy of our solver by employing a novel high-order quadrature scheme on lattice cells overlapping with the model boundary, which are treated at sub-cell precision. Finally, we detail how this accurate boundary treatment can be implemented at a minimal computational premium over the cost of a voxel-accurate discretization. We demonstrate our method in the simulation of complex musculoskeletal human models.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation

Keywords: nonlinear elasticity, incompressibility, cut-cell methods

Links: DL PDF

1 Introduction

Simulation of elastic deformable models is ubiquitous in computer graphics and remains a vibrant area of research. Algorithmic techniques for deformable body simulation, pioneered by Terzopoulos et al [1987] have attained a significant level of maturity, leading to broad adoption in visual effects, games, virtual environments and biomechanics applications. However, numerous theoretical and technical challenges remain. Research efforts often emphasize improved computational performance for cost-conscious interactive applications. Simulation of complex materials and concerns about accuracy and fidelity, especially in biomechanics applications, place additional strain on simulation techniques. Finally, ease of use and deployment in production environments is an important trait that scholarly research work needs to be sensitive to. Our paper proposes a grid-based simulation technique with a number of original components that enhance performance and parallelism, natively accommodate complex materials (including skin, flesh and muscles) while offering the simple and familiar front-end of a lattice deformer for easy integration into an animation pipeline.

Lattice-based volumetric deformers are popular components in both physics-based and procedural animation techniques. In the case of physics-based simulation, one of their key advantages is that they avoid having to construct a simulation-ready conforming volume mesh, which is a delicate preprocessing task often requiring supervision and fine-tuning. Another crucial benefit is that the regularity of such data structures enables aggressive performance optimizations as vividly demonstrated by shape matching techniques [Rivers and James 2007]. Cartesian lattices have also been leveraged to accelerate performance in physics-based approaches, albeit predominantly for simple models such as linear or corotated elasticity [Müller et al. 2004; Georgii and Westermann 2008; McAdams et al. 2011]. Prior graphics work, however, has not demonstrated such aggressive performance gains from lattice-based discretizations when highly nonlinear, anisotropic or incompressible materials are involved. In part, this is attributed to the fact that simulation of complex materials commands an increased level of attention to issues of robust convergence and reliable treatment of incompressibility. Mature solutions to these concerns have predominantly been demonstrated in the context of specific discretizations (e.g. explicit tetrahedral meshes) where regularity of data structures, compactness of memory footprint and parallelization/vectorization potential were not inherently emphasized. Furthermore, as applications requiring the use of complex materials are also likely to emphasize geometric accuracy, they often opt for conforming mesh discretizations due to their superior performance in capturing intricate boundary features, even if their computational cost is higher.

ACM Reference Format
Patterson, T., Mitchell, N., Sifakis, E. 2012. Simulation of complex nonlinear elastic bodies using lattice deformers. *ACM Trans. Graph. 31* 6, Article 197 (November 2012), 10 pages.
DOI = 10.1145/2366145.2366216 http://doi.acm.org/10.1145/2366145.2366216.

RigMesh: Automatic Rigging for Part-Based Shape Modeling and Deformation

Péter Borosán[1]* Ming Jin[1]† Doug DeCarlo[1] Yotam Gingold[1,2] Andrew Nealen[1]
[1]Rutgers University [2]Columbia University

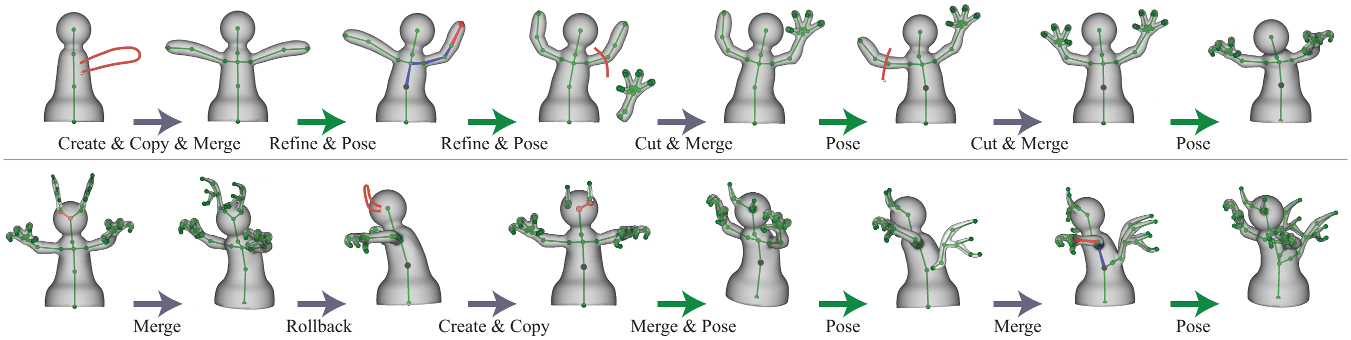


Figure 1: *An example RigMesh modeling sequence: the artist experiments with different ideas for a shape by modeling (blue arrows) and posing (green arrows) using our skeletal rig. In the traditional pipeline, modeling and rigging are separate procedures, so the model must be re-rigged every time the shape changes (i.e. after each modeling operation). RigMesh unifies modeling and rigging; the model is always rigged, and the artist can pose freely, allowing for iterative modeling, deformation, and key-frame animation with real-time response.*

Abstract

The creation of a 3D model is only the first stage of the 3D character animation pipeline. Once a model has been created, and before it can be animated, it must be *rigged*. Manual rigging is laborious, and automatic rigging approaches are far from real-time and do not allow for incremental updates. This is a hindrance in the real world, where the shape of a model is often revised after rigging has been performed. In this paper, we introduce algorithms and a user-interface for sketch-based 3D modeling that unify the modeling and rigging stages of the 3D character animation pipeline. Our algorithms create a rig for each sketched part in real-time, and update the rig as parts are merged or cut. As a result, users can freely pose and animate their shapes and characters while rapidly iterating on the base shape. The rigs are compatible with the state-of-the-art character animation pipeline; they consist of a low-dimensional skeleton along with skin weights identifying the surface with bones of the skeleton.

Keywords: rigging, skeletonization, skinning, animation, sketch-based modeling

Links: DL PDF

*e-mail: pborosan@cs.rutgers.edu
†e-mail: mjin@cs.rutgers.edu
Péter Borosán and Ming Jin contributed equally to this work and are placed in alphabetical order.

ACM Reference Format
Borosán, P., Jin, M., DeCarlo, D., Gingold, Y., Nealen, A. 2012. RigMesh: Automatic Rigging for Part-Based Shape Modeling and Deformation. *ACM Trans. Graph. 31* 6, Article 198 (November 2012), 9 pages. DOI = 10.1145/2366145.2366217 http://doi.acm.org/10.1145/2366145.2366217.

1 Introduction

The task of creating ready-to-animate 3D models is fundamentally hard; designers and artists spend years becoming proficient at commercial, state-of-the-art tools such as Maya and 3ds Max [Autodesk 2012c; Autodesk 2012a]. In an effort to make freeform 3D modeling more accessible to novices and to enable rapid iterative prototyping, sketch-based tools such as Teddy [Igarashi et al. 1999] and a large body of follow-up work [Tai et al. 2004; Cherlin et al. 2005; Schmidt et al. 2005; Alexe et al. 2005; Karpenko and Hughes 2006; Nealen et al. 2007; Bernhardt et al. 2008; Sugihara et al. 2008; Pihuit et al. 2010; Cordier et al. 2011] introduced techniques to create plausible 3D models from 2D freeform strokes. (See [Olsen et al. 2009] for a recent survey.) While these tools greatly simplify shape modeling, modeling is only the first stage of the 3D animation pipeline. Once a model has been created, and before it can be animated, it must be *rigged*. In the state-of-the-art character animation pipeline, a rig takes the form of a *skeleton*, a cycle-free graph whose nodes are called joints and whose edges are called bones, and *skin weights* identifying the surface of the model with the bones of the skeleton [Magnenat-Thalmann et al. 1988; Lewis et al. 2000]. Rigging can be performed manually, by designing the skeleton and then laboriously painting the surface with skin weights for each bone. Various computational methods exist for automatic skeleton extraction [Sharf et al. 2007; Cornea et al. 2007; Au et al. 2008; Pan et al. 2009] and skinning [Lewis et al. 2000; Bloomenthal 2002; Kry et al. 2002; Mohr and Gleicher 2003; Weber et al. 2007; Baran and Popović 2007; Wang et al. 2007; Yang and Wünsche 2010; Miller et al. 2010].

In this paper, we take advantage of these recent advances in the literature, and further simplify the process of creating ready-to-animate 3D models by combining shape creation, modification and posing into a single coherent framework. However, this cannot be achieved trivially by integrating existing tools. Sketch-based modeling tools like Teddy [1999] and FiberMesh [2007] do not explicitly maintain a skeleton. Automatically adding the skeleton (e.g. [Au et al. 2008]) and rigging (e.g. [Baran and Popović 2007]) as a post-process typically does not allow for real-time interactions. Furthermore, in our discussions with professional anima-

Smooth Skinning Decomposition with Rigid Bones

Binh Huy Le* Zhigang Deng†
University of Houston

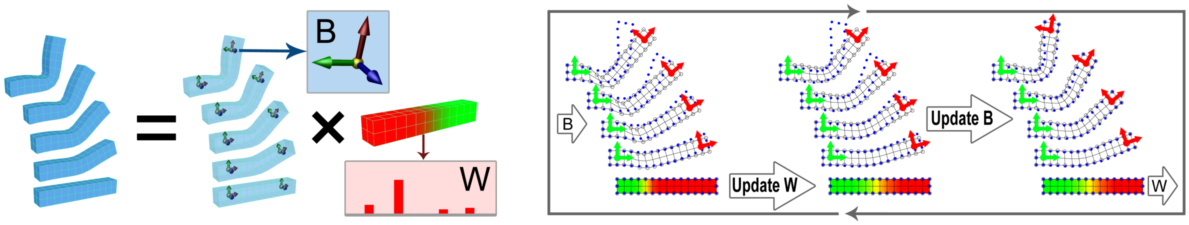


Figure 1: *A set of example poses are decomposed into rigid bone transformations B and a sparse, convex bone-vertex weight map W (left hand side) by our block coordinate descent algorithm (right hand side). During the process, the example poses (indicated as blue dots) can be reconstructed more accurately by alternatively updating W and B while the other is kept fixed.*

Abstract

This paper introduces the Smooth Skinning Decomposition with Rigid Bones (SSDR), an automated algorithm to extract the linear blend skinning (LBS) from a set of example poses. The SSDR model can effectively approximate the skin deformation of nearly articulated models as well as highly deformable models by a low number of rigid bones and a sparse, convex bone-vertex weight map. Formulated as a constrained optimization problem where the least squared error of the reconstructed vertices by LBS is minimized, the SSDR model can be solved by a block coordinate descent-based algorithm to iteratively update the weight map and the bone transformations. By employing the sparseness and convex constraints on the weight map, the SSDR model can be used for traditional skinning decomposition tasks such as animation compression and hardware-accelerated rendering. Moreover, by imposing the orthogonal constraints on the bone rotation matrices (rigid bones), the SSDR model can also be applied in motion editing, skeleton extraction, and collision detection tasks. Through qualitative and quantitative evaluations, we show the SSDR model can measurably outperform the state-of-the-art skinning decomposition schemes in terms of accuracy and applicability.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Geometric algorithms; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: Skinning Decomposition, Linear Blend Skinning, Geometric Deformation, Block Coordinate Descent

Links: DL PDF WEB VIDEO

*e-mail: bhle2@cs.uh.edu
†e-mail:zdeng@cs.uh.edu

ACM Reference Format
Le, B., Deng, Z. 2012. Smooth Skinning Decomposition with Rigid Bones. *ACM Trans. Graph. 31* 6, Article 199 (November 2012), 10 pages. DOI = 10.1145/2366145.2366218 http://doi.acm.org/10.1145/2366145.2366218.

1 Introduction

Numerous research efforts have been focused on efficiently skinning mesh animations. Among many proposed techniques, *linear blend skinning* (LBS) is widely known to be the most popular skinning computational model due to its effectiveness, simplicity, and efficiency [Lewis et al. 2000; Gain and Bechmann 2008]. It has many different names over the years including *skeleton subspace deformation*, *enveloping*, *vertex blending*, *smooth skinning* (Autodesk Maya), *bones skinning* (Autodesk 3D Studio Max), or *linear blend skinning* (the open-source Blender).

In the LBS model, skin deformation is driven by a set of bones. Every vertex is associated with the bones via a bone-vertex weight map which quantifies the influence of each bone to the vertices. The skin is deformed by transforming each vertex through a weighted combination of bone transformations from the rest pose. Assuming w_{ij} is the influence of *j*-th bone to the *i*-th vertex, p_i is the position of the *i*-th vertex at the rest pose, $|B|$ is the number of bones, and R_j^t and T_j^t are the rotation matrix and translation vector of the *j*-th bone at the *t*-th configuration, respectively, then the deformed *i*-th vertex, v_i^t, can be computed as follows:

$$v_i^t = \sum_{j=1}^{|B|} w_{ij}(R_j^t p_i + T_j^t) \quad (1)$$

Depending on specific applications, the above LBS model may impose certain constraints. The weight map w_{ij} is normally required to be convex, i.e., $w_{ij} \geq 0$ and $\sum_{j=1}^{|B|} w_{ij} = 1$. The first *non-negativity* constraint makes the transformation blending additive. The second *affinity* constraint normalizes the influences/weights to prevent over-fitting and deformation artifacts. The two constraints are critical for certain applications such as animation editing. In addition, the *sparseness* constraint on the weight map, which limits the number of non-zero bone weights per vertex, may be applied to take advantage of graphic hardware capabilities. Many applications also require R_j^t matrix to be orthogonal, e.g. animation editing, collision detection, and skeleton extraction (please refer to Section 6 for more details). The *orthogonal* constraint avoids any shearing or scaling effect on the bone transformations, thus put the transformations into the rigid group. For this reason, the bone transformation with orthogonal rotation matrix is also called the "rigid bone".

In this work, we propose a novel example-based method (called

User-guided White Balance for Mixed Lighting Conditions

Ivaylo Boyadzhiev
Cornell University

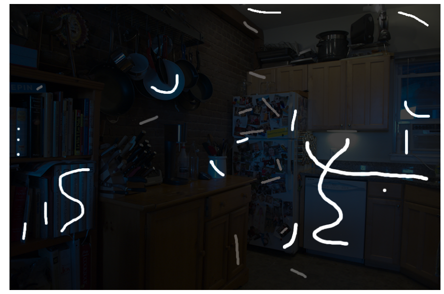
Kavita Bala
Cornell University

Sylvain Paris
Adobe

Frédo Durand
MIT CSAIL

(a) input with mixed lighting (daylight + neon under the cabinets + low-energy bulbs on ceiling) exhibits unsightly color casts everywhere

(b) naive single-light white balance makes the ceiling white, but other color casts remain



(c) user indicates regions that are neutral (white strokes) or correct after the single-light white balance (grey strokes)

(d) the image is improved, but color variations can still be observed, e.g., on the wooden cabinet

(e) user adds marks to specify uniform color, e.g., the cabinet and the wall

(f) our final output with no color casts

Figure 1: *In this photo, the ambient lighting, the cabinet light, and the ceiling lights all have different colors, which produces unpleasant color casts (a). In such situations, the single-light white balance tool provided in all photo editing software only improves a portion of the image, but the result is not satisfying (b). We address this issue by letting users make annotations on the photo. First, they mark objects of neutral color (i.e., white or gray), and regions that look fine after the standard white balance (c). This improves the result, but undesirable color variations are still visible, e.g., on the cabinetry and on the wall (d). Users can indicate that these elements should have a constant color (e), which yields a result free of color cast (f).*

Abstract

Proper white balance is essential in photographs to eliminate color casts due to illumination. The single-light case is hard to solve automatically but relatively easy for humans. Unfortunately, many scenes contain multiple light sources such as an indoor scene with a window, or when a flash is used in a tungsten-lit room. The light color can then vary on a per-pixel basis and the problem becomes challenging at best, even with advanced image editing tools.

We propose a solution to the ill-posed mixed light white balance problem, based on user guidance. Users scribble on a few regions that should have the same color, indicate one or more regions of neutral color , and select regions where the current color looks correct. We first expand the provided scribble groups to more regions using pixel similarity and a robust voting scheme. We formulate the spatially varying white balance problem as a sparse data interpolation problem in which the user scribbles and their extensions form constraints. We demonstrate that our approach can produce satisfying results on a variety of scenes with intuitive scribbles and without any knowledge about the lights.

Keywords: white balance, mixed lighting

1 Introduction

White balance correction is a critical photography step, where the goal is ”to compensate for different colour temperatures of scene illuminants” [Jacobson 2000]. For example, tungsten lights cause images to have a yellowish cast. Proper white balance compensates for this color cast and yields photos where objects have their natural colors, as if taken under a neutral light [Hedgecoe 2009]. When all the lights have the same color, this problem is easy to solve for

ACM Reference Format
Boyadzhiev, I., Bala, K., Paris, S., Durand, F. 2012. User-guided White Balance for Mixed Lighting Conditions. *ACM Trans. Graph. 31* 6, Article 200 (November 2012), 10 pages. DOI = 10.1145/2366145.2366219 http://doi.acm.org/10.1145/2366145.2366219.

Calibrated Image Appearance Reproduction

Erik Reinhard[1]* Tania Pouli[1] Timo Kunkel[2] Ben Long[3] Anders Ballestad[4] Gerwin Damberg[4]
[1]MPI for Informatics [2]Dolby Laboratories [3]University of Bristol [4]MTT Innovation

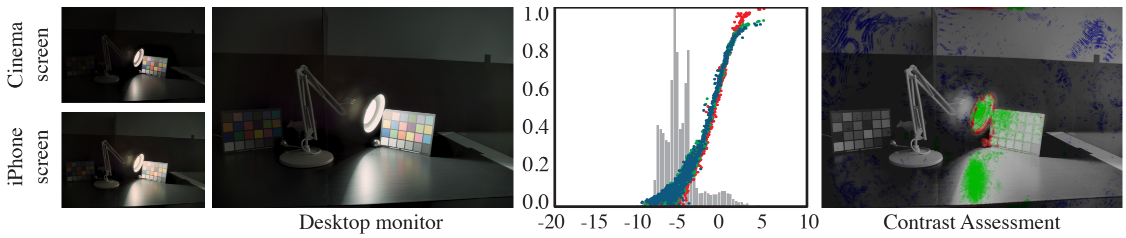


Figure 1: *Our appearance reproduction model faithfully reproduces color, taking into account different viewing environments and display types as shown here for different combinations of viewing environment and display. The plot shows the image histogram (grey) as well as the input/output mapping (red, green and blue) for this particular image. The right panel shows that contrast is accurately reproduced for most pixels (shown grey), as determined by the dynamic range independent image quality metric [Aydin et al. 2008].*

Abstract

Managing the appearance of images across different display environments is a difficult problem, exacerbated by the proliferation of high dynamic range imaging technologies. Tone reproduction is often limited to luminance adjustment and is rarely calibrated against psychophysical data, while color appearance modeling addresses color reproduction in a calibrated manner, albeit over a limited luminance range. Only a few image appearance models bridge the gap, borrowing ideas from both areas. Our take on scene reproduction reduces computational complexity with respect to the state-of-the-art, and adds a spatially varying model of lightness perception. The predictive capabilities of the model are validated against all psychophysical data known to us, and visual comparisons show accurate and robust reproduction for challenging high dynamic range scenes.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Display Algorithms;

Keywords: Color Appearance, High Dynamic Range Imaging, Tonemapping, Lightness Perception

Links: DL PDF WEB VIDEO CODE

1 Introduction

Traditional imaging pipelines are designed around the abilities of conventional capture and display devices, and therefore do not handle dynamic range beyond what can be represented with a single byte per pixel per color channel. Light in the world around us cannot be well represented by such highly quantized values. This has led to the development of a collection of capture, processing and display technologies that are collectively termed high dynamic range imaging (HDRI) [Reinhard et al. 2010; McCann and Rizzi 2012]. In particular capture and display hardware technologies are rapidly maturing [Tocci et al. 2011; Seetzen et al. 2004], opening up new opportunities in entertainment and broadcasting, but bringing new demands on encoding, storage and especially on display.

For instance, dynamic range reduction [Myszkowski et al. 2008; Reinhard et al. 2010] address luminance mismatches between image and display, and may take display capabilities into account [Mantiuk et al. 2008]. Many of these techniques offer sophisticated mechanisms and models to handle the extensive luminance range of HDR, often inspired by aspects of the visual system. More often than not though, tone reproduction operators treat color as a separate modality, and with much less precision. Usually, post-processing steps are applied to reduce the saturation of colors [Schlick 1994; Mantiuk et al. 2009]. Although this may lead to satisfying results in many cases, they are not sufficient for accurately modelling the appearance of colors under different conditions. For instance, a color timer preparing a movie for the cinema will need to take into account the specific viewing conditions likely to occur (dark room, large display, varying viewing distances). On the other hand, when preparing the same movie for DVD, to ensure that the viewer will have the same experience as the cinema goer, different viewing conditions are considered.

Color appearance models (CAMs) predict the appearance of a given color under different conditions (which are specified as inputs to the algorithm) [Fairchild 2005]. As opposed to tonemapping, most CAMs are designed with a focus on color and less so on dynamic range, making them less appropriate for dealing with HDR data as they offer little compression. Although CAMs offer high predictive power for single patches of color, spatial relations in an image can greatly affect the appearance of colors, requiring spatially varying image appearance models instead [Kuang et al. 2007].

A difference between tonemapping and color appearance solutions is the treatment of the viewing environment. In the former, display capabilities are rarely taken into account, albeit with one notable exception [Mantiuk et al. 2008]. In CAMs on the other hand, the room illumination is taken into account in a symmetric manner to the original scene environment. Mixed adaptation to both room and display is not considered in any model that we are aware of.

*Contact e-mail: reinhard@mpi-inf.mpg.de

ACM Reference Format
Reinhard, E., Pouli, T., Kunkel, T., Long, B., Ballestad, A., Damberg, G. 2012. Calibrated Image Appearance Reproduction. *ACM Trans. Graph. 31* 6, Article 201 (November 2012), 11 pages.
DOI = 10.1145/2366145.2366220 http://doi.acm.org/10.1145/2366145.2366220.

Coherent Intrinsic Images from Photo Collections

Pierre-Yves Laffont[1] Adrien Bousseau[1] Sylvain Paris[2] Frédo Durand[3] George Drettakis[1]

[1] REVES / INRIA Sophia Antipolis [2] Adobe [3] MIT CSAIL

Input photo-collection Coherent reflectance Individual illuminations

Figure 1: *Our method leverages the heterogeneity of photo collections to automatically decompose photographs of a scene into reflectance and illumination layers. The extracted reflectance layers are coherent across all views, while the illumination captures the shading and shadow variations proper to each picture. Here we show the decomposition of three photos in the collection.*

Abstract

An intrinsic image is a decomposition of a photo into an illumination layer and a reflectance layer, which enables powerful editing such as the alteration of an object's material independently of its illumination. However, decomposing a single photo is highly under-constrained and existing methods require user assistance or handle only simple scenes. In this paper, we compute intrinsic decompositions using several images of the same scene under different viewpoints and lighting conditions. We use multi-view stereo to automatically reconstruct 3D points and normals from which we derive relationships between reflectance values at different locations, across multiple views and consequently different lighting conditions. We use robust estimation to reliably identify reflectance ratios between pairs of points. From these, we infer constraints for our optimization and enforce a coherent solution across multiple views and illuminations. Our results demonstrate that this constrained optimization yields high-quality and coherent intrinsic decompositions of complex scenes. We illustrate how these decompositions can be used for image-based illumination transfer and transitions between views with consistent lighting.

Keywords: intrinsic images, photo collections

Links: DL PDF

1 Introduction

Image collections aggregate many images of a scene from a variety of viewpoints and are often captured under different illuminations. The variation of illumination in a collection has often been seen as a nuisance that is distracting during navigation or, at best an interesting source of visual diversity. Inspired by existing work on time-lapse sequences [Weiss 2001; Matsushita et al. 2004], we consider these variations as a rich source of information to compute *intrinsic images*, i.e., to decompose photos into the product of an illumination layer by a reflectance layer [Barrow and Tenenbaum 1978]. This decomposition is an ill-posed problem since an infinity of reflectance and illumination configurations can produce the same image, and so far automatic techniques are limited to simple objects [Grosse et al. 2009], while real-world scenes require user assistance [Bousseau et al. 2009], detailed geometry [Troccoli and Allen 2008; Haber et al. 2009], or varying illumination with a fixed or restricted viewpoint [Weiss 2001; Liu et al. 2008].

In this paper, we exploit the rich information provided by multiple viewpoints and illuminations in an image collection to process complex scenes without user assistance, nor precise and complete geometry. Furthermore, we enforce that the decomposition be *coherent*, which means that the reflectance of a scene point should be the same in all images.

The observation of a point under different unknown illuminations does not help directly with the fundamental ambiguity of intrinsic images. Any triplet R, G, B is a possible reflectance solution for which the illumination of the point in each image is its pixel value divided by R, G, B. We overcome this difficulty by processing pairs of points. We consider the ratio of radiance between two points, which is equal to the ratio of reflectance if the points share the same illumination. A contribution of this paper is to identify pairs of points that are likely to have similar illumination across most conditions. For this, we leverage sparse 3D information from multi-view stereo as well as a simple statistical criterion on the distribution of the observed ratios. These ratios give us a set of equations relating the reflectance of pairs of sparse scene points, and consequently of sparse pixels where the scene points project in the input images. To infer the reflectance and illumination for all the pixels, we build on image-guided propagation [Levin et al. 2008; Bousseau et al. 2009]. We augment it with a term to force the estimated reflectance of a given 3D point to be the same in all the images in which it is visible. This yields a large sparse linear system, which we solve in an interleaved manner. By enforcing coherence in the reflectance layer we obtain a common "reflectance space" for all input views, while we extract the color variations proper to each image in the illumination layer.

Our automatic estimation of coherent intrinsic image decompositions from photo collections relies on the following contributions:

- A method to robustly identify reliable reflectance constraints between pairs of pixels, based on multi-view stereo and a statistical criterion.

ACM Reference Format
Laffont, P., Bousseau, A., Paris, S., Durand, F., Drettakis, G. 2012. Coherent Intrinsic Images from Photo Collections. *ACM Trans. Graph. 31* 6, Article 202 (November 2012), 11 pages.
DOI = 10.1145/2366145.2366221 http://doi.acm.org/10.1145/2366145.2366221.

Robust Patch-Based HDR Reconstruction of Dynamic Scenes

Pradeep Sen[1] Nima Khademi Kalantari[1] Maziar Yaesoubi[2] Soheil Darabi[2]
Dan B Goldman[3] Eli Shechtman[3]

[1]University of California, Santa Barbara [2]UNM Advanced Graphics Lab [3]Adobe

Input LDR sources Reconstructed LDR images Final tonemapped HDR result

Figure 1: *Our algorithm takes as input a sequential set of bracketed exposures of a dynamic scene (not pre-aligned) and outputs a high-quality HDR result along with a reconstructed set of aligned images at each exposure. On the left are three of seven low-dynamic range (LDR) input sources taken with a standard, hand-held digital camera that have both subject and camera motion. These are followed by the outputs of our algorithm at each exposure, aligned to the reference image which is not shown. Not only does our algorithm properly align the images despite the complex motion, but it also maintains the subtle lighting detail in each exposure (e.g., highlights on the hat, shading on the shirt, detail in the Christmas tree) that will contribute information to the final radiance map. On the right is our tonemapped HDR result. For a set like this with 7 input images, our algorithm takes less than 3 minutes to compute the final result at* 1350×900 *resolution.*

Abstract

High dynamic range (HDR) imaging from a set of sequential exposures is an easy way to capture high-quality images of static scenes, but suffers from artifacts for scenes with significant motion. In this paper, we propose a new approach to HDR reconstruction that draws information from all the exposures but is more robust to camera/scene motion than previous techniques. Our algorithm is based on a novel patch-based energy-minimization formulation that integrates alignment and reconstruction in a joint optimization through an equation we call the HDR image synthesis equation. This allows us to produce an HDR result that is aligned to one of the exposures yet contains information from all of them. We present results that show considerable improvement over previous approaches.

CR Categories: I.4.1 [Computing Methodologies]: Image Processing and Computer Vision—Digitization and Image Capture

Keywords: High-dynamic range imaging, image alignment

Links: DL PDF WEB

ACM Reference Format
Sen, P., Kalantari, N., Yaesoubi, M., Darabi, S., Goldman, D., Shechtman, E. 2012. Robust patch-based HDR reconstruction of dynamic scenes. *ACM Trans. Graph. 31* 6, Article 203 (November 2012), 11 pages. DOI = 10.1145/2366145.2366222 http://doi.acm.org/10.1145/2366145.2366222.

1 Introduction

High-dynamic range (HDR) imaging has the potential to transform the world of photography. Unlike traditional low-dynamic range (LDR) images that measure only a small range of the total illumination of a scene, HDR images capture a much wider range and therefore more closely resemble what photographers see with their own eyes. However, despite their tremendous potential, existing approaches for high-quality HDR imaging have serious limitations. For example, specialized camera hardware has been proposed to capture HDR content directly (e.g., [Nayar and Mitsunaga 2000; Tocci et al. 2011]), but these devices are typically expensive and are currently unavailable to the general public.

To make high-quality HDR imaging widespread, we must focus on approaches that use standard digital cameras. The most common approach is to take sequential LDR images at different exposure levels (known as bracketed exposures) and then merge them into an HDR image [Mann and Picard 1995; Debevec and Malik 1997]. Although this technique can produce spectacular results (see, e.g., [Ratcliff 2012]), the original approaches work only for static scenes because they typically assume a constant radiance at each pixel over all exposures. When the scene has moving content (or the camera is hand-held), this method produces ghost-like artifacts from even small misalignments between exposures. This is a serious limitation, since real-world scenes often have moving objects and real-world cameras are not often mounted on tripods.

The problem of removing motion artifacts for sequential HDR imaging has been the subject of extensive research and has led to two major kinds of approaches. The first kind assume that the images are mostly static and that only small parts of the scene have