Pattern-Aware Shape Deformation Using Sliding Dockers

Martin Bokeloh* MPI Informatik

Michael Wand* Saarland University and MPI Informatik

Vladlen Koltun† Stanford University

Hans-Peter Seidel* MPI Informatik

Figure 1: *Edit example (left to right): The user places constraints (blue) and manipulates the object by moving constraints. Our deformation model maintains continuous and discrete patterns and adapts the repetition count of discrete patterns, inserting and deleting elements as needed to minimize distortion. Discrete changes are highlighted in orange.*

Abstract

This paper introduces a new structure-aware shape deformation technique. The key idea is to detect continuous and discrete regular patterns and ensure that these patterns are preserved during free-form deformation. We propose a variational deformation model that preserves these structures, and a discrete algorithm that adaptively inserts or removes repeated elements in regular patterns to minimize distortion. As a tool for such structural adaptation, we introduce sliding dockers, which represent repeatable elements that fit together seamlessly for arbitrary repetition counts. We demonstrate the presented approach on a number of complex 3D models from commercial shape libraries.

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

Keywords: shape deformation, shape analysis, symmetry, structural regularity

Links: DL PDF

1 Introduction

Content creation is one of the main bottlenecks in contemporary computer graphics. While sophisticated methods for processing and rendering three-dimensional content are widely available, the creation of detailed custom 3D geometry still requires significant expertise. The issue is not merely of creative ability, but also with the process of directly manipulating detailed 3D models. Such 3D models often feature structural relationships on multiple scales, which need to be manually restored whenever a significant manipulation is performed on the model. Tedious adjustment is often required multiple times in a single modeling session.

Consequently, recent research has begun to investigate *structure-aware* shape editing tools that aim to automate the detailed manipulation required to preserve the structural relationships in a shape as it undergoes manipulation [Kraevoy et al. 2008; Gal et al. 2009; Huang et al. 2009; Wang et al. 2011; Zheng et al. 2011]. Such algorithms analyze the input shape to extract structural features and use the learned structure to assist interactive 3D modeling. They can improve the efficiency of content creation professionals and can assist inexperienced users in adapting existing content to their needs.

In this paper, we present a structure-aware shape editing technique that detects discrete and continuous patterns in the shape and preserves these patterns under free-form deformation. A key distinguishing feature of our approach is that it can change the structure of the object by adding or removing local elements along regular patterns. This structural adaptation is integrated into a global free-form deformation framework that minimizes the overall stretch of the object.

In our approach, the user specifies a small set of constraints and the system computes a new shape that meets these constraints while preserving structural properties of the original model, as shown in Figure 1. As invariants, we extract 1-parameter groups of partial symmetries. In other words, we detect geometry that is replicated in regular patterns. This includes continuous symmetries such as straight lines as well as repeated discrete elements such as windows in a building. We formulate non-local rigidity constraints to maintain these symmetry properties in the output, and allow for adapting the number of discrete repetitions in order to reduce distortions.

In order to add and remove elements along discrete patterns with minimal distortion, we introduce *sliding dockers*. A sliding docker is an element in a local, repeated structure that interfaces with the rest of the model in such way that the structure can be independently replicated with minimal distortion. We develop an algorithm that automatically finds collections of sliding dockers that repeat in one common translational direction and adapts the replication count in a way that minimizes distortions in the overall object.

We evaluate the presented technique on models taken from commercial 3D model libraries and demonstrate that the presented tech-

*e-mail:{mbokeloh,mwand,hpseidel}@mpi-inf.mpg.de
†e-mail:vladlen@stanford.edu

ACM Reference Format
Bokeloh, M., Wand, M., Koltun, V., Seidel, H. 2011. Pattern-Aware Shape Deformation Using Sliding Dockers. *ACM Trans. Graph. 30*, 6, Article 123 (December 2011), 10 pages. DOI = 10.1145/2024156.2024157 http://doi.acm.org/10.1145/2024156.2024157.

Shape Space Exploration of Constrained Meshes

Yong-Liang Yang
KAUST

Yi-Jun Yang
KAUST

Helmut Pottmann
KAUST/TU Vienna

Niloy J. Mitra
KAUST/TU Vienna

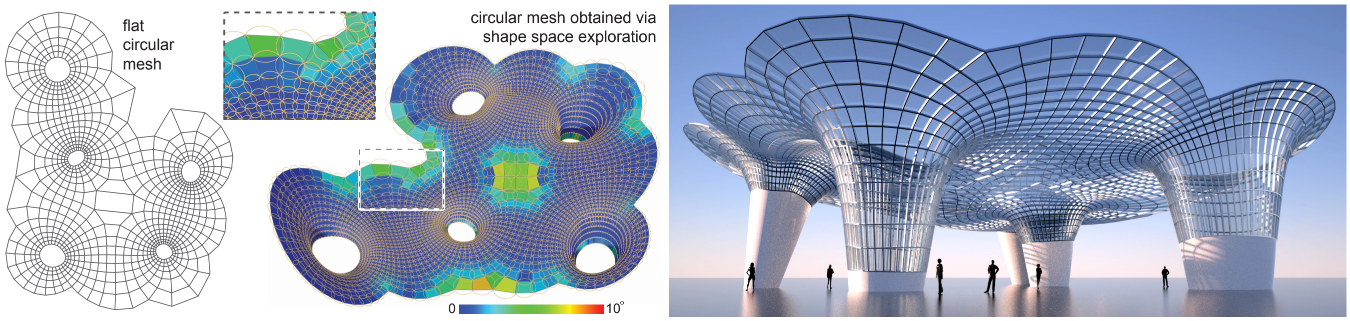


Figure 1: *Starting from a single input mesh along with a set of non-linear constraints, our geometric framework allows local characterization, navigation, and exploration of the corresponding shape space. The figure shows a sample design (right) created using our method, starting from a flat circular mesh (left).*

Abstract

We present a general computational framework to locally characterize any shape space of meshes implicitly prescribed by a collection of non-linear constraints. We computationally access such manifolds, typically of high dimension and co-dimension, through first and second order approximants, namely tangent spaces and quadratically parameterized osculant surfaces. Exploration and navigation of desirable subspaces of the shape space with regard to application specific quality measures are enabled using approximants that are intrinsic to the underlying manifold and directly computable in the parameter space of the osculant surface. We demonstrate our framework on shape spaces of planar quad (PQ) meshes, where each mesh face is constrained to be (nearly) planar, and circular meshes, where each face has a circumcircle. We evaluate our framework for navigation and design exploration on a variety of inputs, while keeping context specific properties such as fairness, proximity to a reference surface, etc.

Keywords: shape space, manifold navigation, design exploration, constrained mesh

Links: DL PDF WEB VIDEO CODE

1 Introduction

In geometry processing, meshes are often specified by a collection of non-linear constraints, typically associated with mesh faces or edges. Exploring and navigating the corresponding shape space, i.e., the possible meshes sharing the same combinatorics as the input mesh while satisfying the constraints, are widely believed to be challenging. Even seemingly simple handle-driven deformations restricted to such shape spaces turn out to be challenging, and remain an active topic of research (see [Botsch et al. 2006; Kilian et al. 2007; Botsch and Sorkine 2008; Gal et al. 2009]).

In this paper, we propose a mathematical framework for the design and manipulation of non-linearly constrained meshes. Our approach is based on the exploration of an appropriate shape space as follows: Geometric models are mapped to points in a high-dimensional space $\mathbb{R}^D$, where the models that satisfy the constraints form a certain manifold $\mathcal{M} \subset \mathbb{R}^D$ (shape space). Modeling proceeds by navigating in the practically useful parts of the manifold $\mathcal{M}$, as prescribed by application specific quality measures. Such a manifold typically has high dimension and co-dimension, making it difficult to directly employ standard differential geometry concepts such as curvatures, especially in an efficient and computationally feasible manner. We locally approximate the manifold using tangent spaces and quadratically parameterized osculant surfaces, and propose how to computationally estimate the local curvature of the manifold to decide between the two representations.

We demonstrate the utility of our framework for two concrete example scenarios: (i) planar quad (PQ) meshes, i.e., meshes with each quad face being planar, and (ii) circular meshes, i.e., meshes with

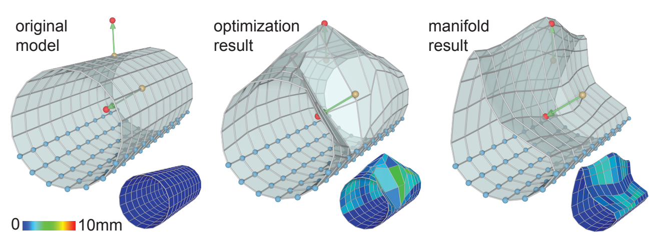


Figure 2: *Decoupling deformation and planarization is undesirable for shape design. Given a PQ mesh (left), the user prescribes a deformation using vertex handles, and the deformed mesh is planarized using an optimization approach [Liu et al. 2006]. The result can be unsatisfactory (middle). In contrast, our PQ mesh manifold exploration characterizes the (non-linearly constrained) design space, allowing direct design (right).*

ACM Reference Format
Yang, Y., Yang, Y., Pottmann, H., Mitra, N. 2011. Shape Space Exploration of Constrained Meshes. *ACM Trans. Graph. 30*, 6, Article 124 (December 2011), 11 pages. DOI = 10.1145/2024156.2024158 http://doi.acm.org/10.1145/2024156.2024158.

Joint Shape Segmentation with Linear Programming

Qixing Huang   Vladlen Koltun   Leonidas Guibas

Stanford University

Abstract

We present an approach to segmenting shapes in a heterogenous shape database. Our approach segments the shapes jointly, utilizing features from multiple shapes to improve the segmentation of each. The approach is entirely unsupervised and is based on an integer quadratic programming formulation of the joint segmentation problem. The program optimizes over possible segmentations of individual shapes as well as over possible correspondences between segments from multiple shapes. The integer quadratic program is solved via a linear programming relaxation, using a block coordinate descent procedure that makes the optimization feasible for large databases. We evaluate the presented approach on the Princeton segmentation benchmark and show that joint shape segmentation significantly outperforms single-shape segmentation techniques.

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

Keywords: shape segmentation, shape correspondence, linear programming

Links: DL PDF

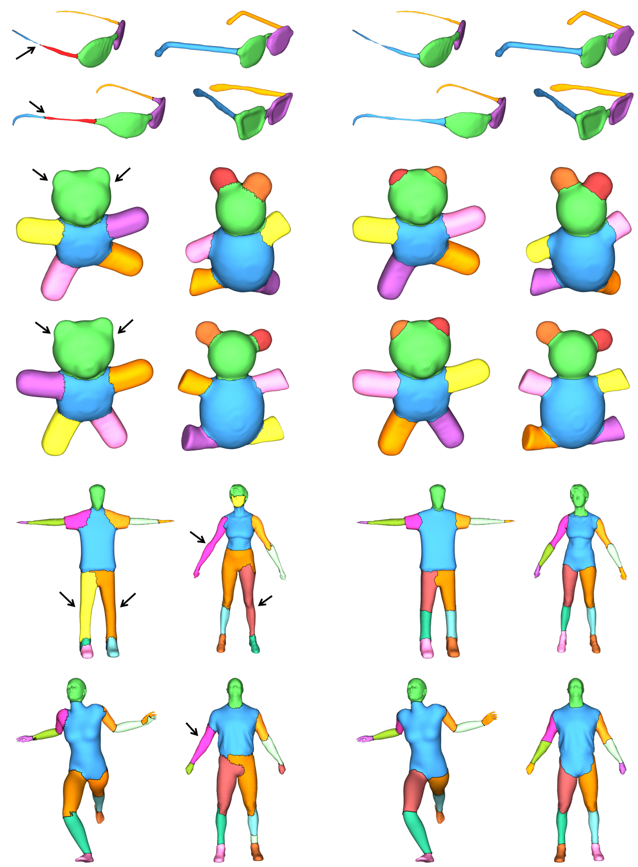


Figure 1: *Comparison of single-shape segmentation (left) and joint shape segmentation (right) on models from the Princeton segmentation benchmark [Chen et al. 2009]. Each segmentation on the left was produced by the top-performing algorithm in the benchmark for that shape. The segmentations on the right were produced by our approach, which jointly optimized segmentations and correspondences across the entire benchmark dataset. The new approach was able to identify meaningful parts despite extraneous geometric cues (top) and low saliency (middle, bottom).*

1 Introduction

Shape segmentation is a fundamental problem in shape analysis. Many shape processing and 3D modeling applications benefit from automatic segmentation of shapes into components that appear natural [Shamir 2008; Chen et al. 2009; Chaudhuri et al. 2011]. Classical shape segmentation techniques analyze the geometric structure of individual shapes in order to detect parts or part boundaries. A variety of geometric features have been investigated, but no single feature or collection of features is known to produce high-quality results for all classes of shapes [Chen et al. 2009]. The underlying difficulty is that a perceptually natural segmentation of a shape is often the result of prior familiarity with other similar shapes and their function. The surface geometry of an individual shape may lack sufficient cues to identify all parts that would be perceived as meaningful to a human observer. Alternatively, a shape may contain strong geometric features within a single perceived part. This can mislead algorithms that consider individual shapes in isolation [Golovinskiy and Funkhouser 2009; Kalogerakis et al. 2010].

In this paper, we present an approach to shape segmentation that jointly analyzes a database of shapes. The approach optimizes segmentations on all shapes together with segment-level correspondences between similar shapes. By considering multiple shapes in concert, our approach is able to identify meaningful parts despite the lack of strong geometric cues on a particular shape. Likewise, the approach is able to identify coherent single parts even when the geometry of the individual shape suggests the presence of multiple segments. This is illustrated in Figure 1. Some eyeglass models (top, left) contain extraneous geometric cues that lead to over-segmentation by approaches that consider single shapes in isolation; our approach correctly detects the presence of a single logical part (the temple) due to counterpart segments in similar shapes in the database. In Figure 1 (middle), the ears of some of the teddy

ACM Reference Format
Huang, Q., Koltun, V., Guibas, L. 2011. Joint Shape Segmentation with Linear Programming. *ACM Trans. Graph. 30*, 6, Article 125 (December 2011), 11 pages. DOI = 10.1145/2024156.2024159 http://doi.acm.org/10.1145/2024156.2024159.

Unsupervised Co-Segmentation of a Set of Shapes via Descriptor-Space Spectral Clustering

Oana Sidi* Oliver van Kaick† Yanir Kleiman* Hao Zhang† Daniel Cohen-Or*
*Tel-Aviv University †Simon Fraser University

Abstract

We introduce an algorithm for unsupervised co-segmentation of a set of shapes so as to reveal the semantic shape parts and establish their correspondence across the set. The input set may exhibit significant shape variability where the shapes do not admit proper spatial alignment and the corresponding parts in any pair of shapes may be geometrically dissimilar. Our algorithm can handle such challenging input sets since, first, we perform co-analysis in a *descriptor space*, where a combination of shape descriptors relates the parts independently of their pose, location, and cardinality. Secondly, we exploit a key enabling feature of the input set, namely, dissimilar parts may be "linked" through third-parties present in the set. The links are derived from the pairwise similarities between the parts' descriptors. To reveal such linkages, which may manifest themselves as anisotropic and non-linear structures in the descriptor space, we perform spectral clustering with the aid of diffusion maps. We show that with our approach, we are able to co-segment sets of shapes that possess significant variability, achieving results that are close to those of a supervised approach.

Keywords: Co-segmentation, shape correspondence, spectral clustering, diffusion maps.

Links: DL PDF

1 Introduction

In recent years, there has been an increasing interest in high-level analysis of 3D shapes. Some methods try to infer high-level knowledge about a given shape from its geometry [Fu et al. 2008; Mitra et al. 2010]. There are also works which utilize semantic knowledge to segment a given shape [Simari et al. 2009; Kalogerakis et al. 2010] or establish a correspondence between a pair of shapes [van Kaick et al. 2011]. The problem of analyzing a *set* of shapes as a whole has received less attention [Golovinskiy and Funkhouser 2009; Xu et al. 2010]. The interesting question about *co-analysis* of a set of shapes is whether more knowledge can be inferred from the set rather than from an individual or pairs of shapes alone. For example, can we better segment the shapes given as a set rather than as individuals? While it seems obvious that a set of shapes contains more knowledge than each individual, it remains a challenge, particularly in the *unsupervised* setting, to extract appropriate knowledge inherent to the set to facilitate fundamental analysis tasks such as segmentation and correspondence.

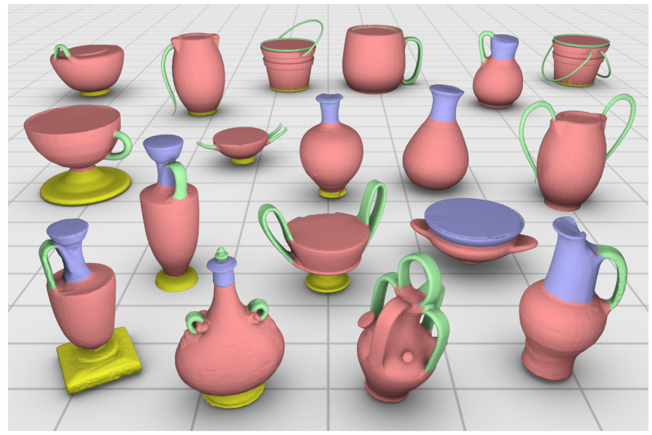


Figure 1: *Unsupervised co-segmentation of a highly varied set of container objects using our algorithm. Corresponding parts differ significantly in their shape, pose, position, and cardinality.*

In this paper, we investigate the problem of unsupervised co-segmentation of a set of shapes so as to reveal the semantic shape parts and establish their correspondence across the set. In our setting, the input shapes belong to a common family, that is, loosely speaking, they share the same functionality and general form. However, their corresponding parts are not necessarily similar; see Figure 1. The co-segmentation is unsupervised in that there is no training set which provides any knowledge to assist the analysis, as opposed to works such as [Kalogerakis et al. 2010] and [van Kaick et al. 2011], where the ability to properly match geometrically dissimilar parts is critically supported by the existence of relevant prior knowledge in the training set. It should also be noted that the analyses performed in the above works do not represent a co-analysis of a *set*. Our work can be seen to complement these knowledge-driven approaches with knowledge extracted from a target set.

The setting and objective of our analysis share similarities with the recent works of Golovinskiy and Funkhouser [2009] and Xu et al. [2010], which both compute an unsupervised co-segmentation of a set of shapes. The method in [Golovinskiy and Funkhouser 2009] pre-aligns the set of shapes and then combines criteria for intra-shape segmentation and inter-shape proximity to cluster mesh faces across the set. Xu et al. [2010] apply a similar co-segmentation scheme with the focus being to remove non-homogeneous part scales from the analysis equation. In our setting, the set consists of shapes with a variety of non-rigid, geometric, and even topological differences; see Figure 1. The dissimilarity between corresponding parts is reaching a point where the power of the set has to be exploited beyond what can be afforded by spatial-domain alignment or clustering. Towards this end, our approach allows correspondence to be inferred indirectly through *third parties* in the set and the analysis is performed in a *descriptor space*; see Figure 2.

Specifically, we treat the unsupervised co-segmentation of a set of shapes as a clustering problem. The clustering is performed in a space of shape descriptors rather than on the spatial coordinates of the shapes themselves. This allows the handling of corresponding

ACM Reference Format
Sidi, O., van Kaick, O., Kleiman, Y., Zhang, H., Cohen-Or, D. 2011. Unsupervised Co-Segmentation of a Set of Shapes via Descriptor-Space Spectral Clustering. *ACM Trans. Graph. 30*, 6, Article 126 (December 2011), 9 pages. DOI = 10.1145/2024156.2024160 http://doi.acm.org/10.1145/2024156.2024160.

Modeling and Generating Moving Trees from Video

Chuan Li† Oliver Deussen‡ Yi-Zhe Song† Phil Willis⋆ Peter Hall†

Media Technology Research Centre, University of Bath, UK† University of Konstanz, Germany‡
Centre for Digital Entertainment, University of Bath, UK⋆

Figure 1: *Using a single video as input (left, background removed using alpha matting), our system outputs a 3D dynamic tree model (middle). Using the model, potentially an infinite number of unique trees with similar appearance and motion can be generated (right).*

Abstract

We present a probabilistic approach for the automatic production of tree models with convincing 3D appearance and motion. The only input is a video of a moving tree that provides us an initial dynamic tree model, which is used to generate new individual trees of the same type. Our approach combines global and local constraints to construct a dynamic 3D tree model from a 2D skeleton. Our modeling takes into account factors such as the shape of branches, the overall shape of the tree, and physically plausible motion. Furthermore, we provide a generative model that creates multiple trees in 3D, given a single example model. This means that users no longer have to make each tree individually, or specify rules to make new trees. Results with different species are presented and compared to both reference input data and state of the art alternatives.

Keywords: tree modeling and animation, generative model.

Links: DL PDF

1 Introduction

Trees are among the Earth's most useful and beautiful products of nature. They have been drawn, painted and modeled for centuries. Contemporary tools make it possible to produce high quality 3D moving models. Typically though, each tree must be individually made by an expert user either by sketching or by providing suitable images — new trees can be grown automatically only if abstract rules are defined. The difficulties of building a tree are magnified when the tree is to move. Overall, tree modeling remains a time consuming process that often relies on expert knowledge.

In this paper we address the tree modeling problem using an approach that is almost entirely automatic. To make a model, the user only has to outline the tree in an initial video frame. The system then creates a full 3D model including motion. This model furthermore can serve as an example to automatically generate new 3D dynamic tree models of the same species. Our approach makes it inexpensive to model and animate a large library of trees for graphics applications. Figure 1 gives an illustrative summary of the process.

Neubert et al. [2007] summarize current tree modeling methods by three categories: rules-based generation, interactive modeling, and image-based production. The first group uses rule-systems such as L-systems [Lindenmayer 1968; Prusinkiewicz and Lindenmayer 1990] or procedural models [Deussen and Lintermann 2005] to generate new trees from an initial state. Talton et al. [2011] present an algorithm for high level controlling grammar-based procedural models and demonstrate the algorithm on tree modeling. Rules tend to be abstract and so are best suited to technical users, yet this is the only current group of methods capable of creating many distinct individual trees. The second group uses interaction to sketch a model in 2D and then create a 3D model from that [Anastacio et al. 2006; Quan et al. 2006; Okabe et al. 2005; Chen et al. 2008]. They provide considerable control to artists skilled enough to create high quality trees. Some methods combine rules and interaction [Lintermann and Deussen 1999; Palubicki et al. 2009]. The third group models trees from image data, with the advantage of increasing realism. Martinez et al. [2004] use a set of registered images to define a model, Neubert et al [2007] allow the construction from two loosely coupled images. Tan et al. [2007; 2008] mixes image input with user interaction to construct trees. Other approaches deal with reconstructing tree models from point clouds [Xu et al. 2007;

ACM Reference Format
Li, C., Deussen, O., Song, Y., Willis, P., Hall, P. 2011. Modeling and Generating Moving Trees from Video. *ACM Trans. Graph. 30*, 6, Article 127 (December 2011), 11 pages. DOI = 10.1145/2024156.2024161 http://doi.acm.org/10.1145/2024156.2024161.

Candid Portrait Selection From Video

Juliet Fiss
University of Washington

Aseem Agarwala
Adobe Systems, Inc.

Brian Curless
University of Washington

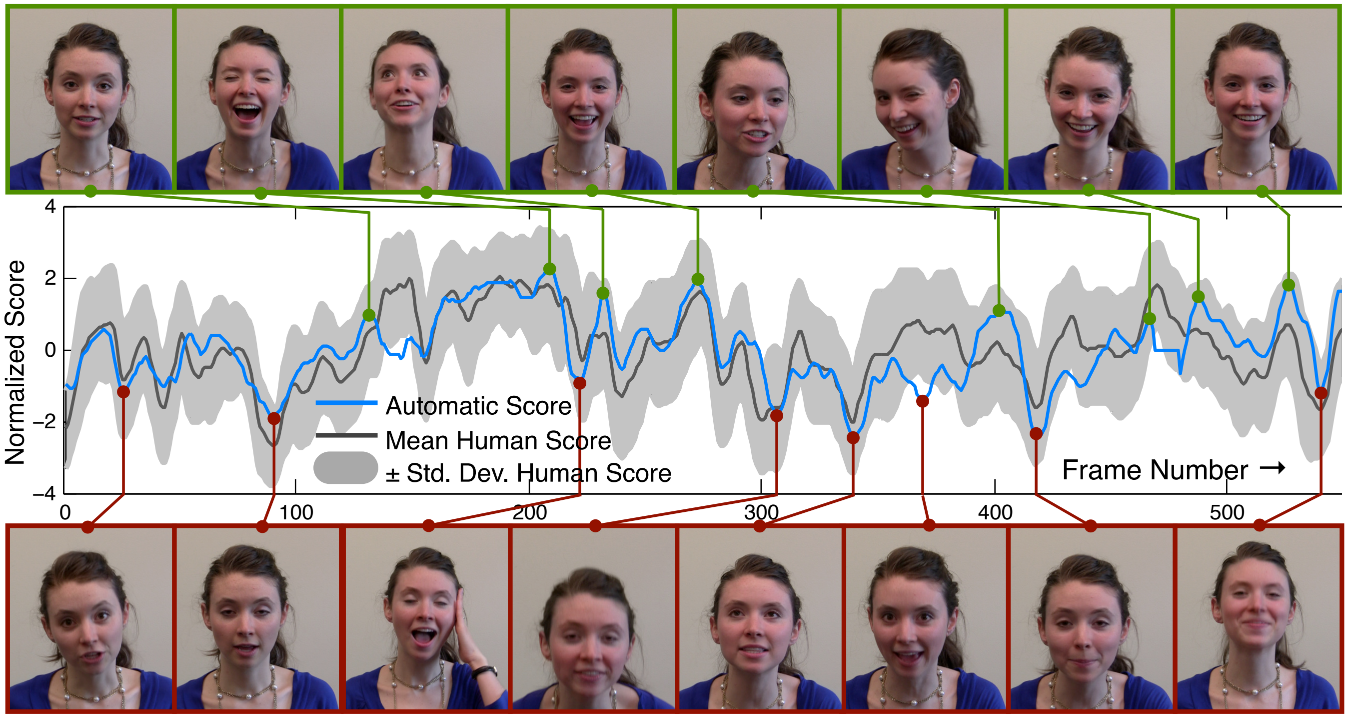


Figure 1: *A plot of the ratings assigned by humans in our psychology study (mean shown in dark gray, per-frame standard deviation shown in light gray), and the ratings assigned by our predictive model (cyan) across the frames of a short video sequence. Both series of ratings have been normalized by their mean and standard deviation. We also show several automatically-selected video frames at peaks (green, top) and valleys (red, bottom) of our predicted rating.*

Abstract

In this paper, we train a computer to select still frames from video that work well as candid portraits. Because of the subjective nature of this task, we conduct a human subjects study to collect ratings of video frames across multiple videos. Then, we compute a number of features and train a model to predict the average rating of a video frame. We evaluate our model with cross-validation, and show that it is better able to select quality still frames than previous techniques, such as simply omitting frames that contain blinking or motion blur, or selecting only smiles. We also evaluate our technique qualitatively on videos that were not part of our validation set, and were taken outdoors and under different lighting conditions.

ACM Reference Format
Fiss, J., Agarwala, A., Curless, B. 2011. Candid Portrait Selection From Video. *ACM Trans. Graph. 30*, 6, Article 128 (December 2011), 8 pages. DOI = 10.1145/2024156.2024162
http://doi.acm.org/10.1145/2024156.2024162.

Links: DL PDF WEB VIDEO DATA

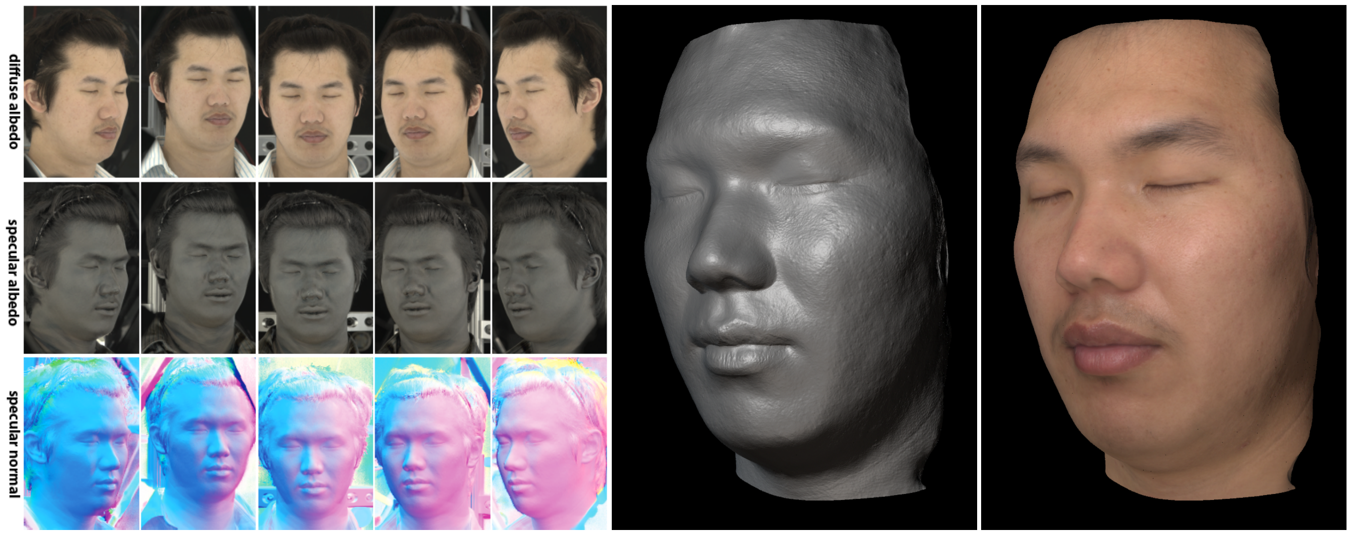
1 Introduction

Cameras sample time very differently than humans do. Still cameras record a frozen moment in time selected by the photographer, and video cameras record a series of moments captured at short, regular intervals. Human vision, on the other hand, is constantly mediated by higher cognitive functions, so that our visual perception of a moment is decidedly different than what a camera captures. This difference is perhaps most glaring when the camera's subject is the human face. It is remarkable how often a photograph or paused video frame of a friend depicts an awkward facial expression that would not be perceived by the human visual system.

While trained photographers are often able to capture the "decisive moment" [Cartier-Bresson 1952], it would be nice if we could simply point our computational cameras at people and automatically acquire only the most desirable photographs. Furthermore, since modern digital cameras can capture very high-resolution video, it would also be nice if we could automatically extract the best moments of people from a captured video. In fact, photographers are starting to record portrait sessions as high-resolution videos; they then select the best frames as a post-process (the cover of the June 2009 issue of Esquire magazine depicting Megan Fox was captured this way [Katz 2009]). In this context, effectively capturing

Multiview Face Capture using Polarized Spherical Gradient Illumination

Abhijeet Ghosh Graham Fyffe Borom Tunwattanapong Jay Busch Xueming Yu Paul Debevec
USC Institute for Creative Technologies



(a) Acquired data (b) Geometry (c) Rendering

Figure 1: *Multiview face capture using polarized spherical gradient illumination. (a) Acquired data from five viewpoints used for stereo reconstruction. (b) Reconstructed geometry. (c) Hybrid normal rendering* [Ma et al. 2007].

Abstract

We present a novel process for acquiring detailed facial geometry with high resolution diffuse and specular photometric information from multiple viewpoints using polarized spherical gradient illumination. Key to our method is a new pair of linearly polarized lighting patterns which enables *multiview* diffuse-specular separation under a given spherical illumination condition from just two photographs. The patterns – one following lines of latitude and one following lines of longitude – allow the use of fixed linear polarizers in front of the cameras, enabling more efficient acquisition of diffuse and specular albedo and normal maps from multiple viewpoints. In a second step, we employ these albedo and normal maps as input to a novel multi-resolution adaptive domain message passing stereo reconstruction algorithm to create high resolution facial geometry. To do this, we formulate the stereo reconstruction from multiple cameras in a commonly parameterized domain for multiview reconstruction. We show competitive results consisting of high-resolution facial geometry with relightable reflectance maps using five DSLR cameras. Our technique scales well for multiview acquisition without requiring specialized camera systems for sensing multiple polarization states.

Keywords: computational illumination, face capture, polarization, message passing.

Links: DL PDF

ACM Reference Format
Ghosh, A., Fyffe, G., Tunwattanapong, B., Busch, J., Yu, X., Debevec, P. 2011. Multiview Face Capture using Polarized Spherical Gradient Illumination. *ACM Trans. Graph. 30*, 6, Article 129 (December 2011), 10 pages. DOI = 10.1145/2024156.2024163 http://doi.acm.org/10.1145/2024156.2024163.

1 Introduction

Digitally reproducing the shape and appearance of real-world subjects is a long-standing goal of computer graphics. In particular, the realistic reproduction of human faces has received increasing attention in recent years. Some of the best techniques use a combination of 3D scanning and photography under different lighting conditions to acquire models of a subject's shape and reflectance. When both of these characteristics are measured, the models can be used to faithfully render how the object would look from any viewpoint, reflecting the light of any environment. An ideal process would accurately model the subject's shape and reflectance with just a few photographs. However, in practice, significant compromises are typically made between the accuracy of the geometry and reflectance model and the amount of data which must be acquired. Ma et al. [2007] introduced polarized spherical gradient illumination for efficiently acquiring diffuse and specular photometric information and employed it in conjunction with structured light scanning to obtain high resolution scans of faces. In addition to the detail in the reconstructed 3D geometry, the photometric data acquired with this technique can be used for realistic rendering in either real-time or offline contexts. However, the technique has significant limitations. Chiefly, Ma et al.'s linear polarization pattern is effective only for the frontal camera viewpoint, forcing the subject to be moved to different positions to scan more than the front of the face. Also, Ma et al.'s lighting patterns require rapidly flipping a polarizer in front of the camera using custom hardware in order to observe both cross- and parallel-polarization states. Finally, Ma et al.'s reliance on structured light for base geometry acquisition adds scanning time and system complexity, while further restricting the process to single-viewpoint scanning.

In order to overcome the viewpoint restriction imposed by active illumination, recent work [Beeler et al. 2010; Bradley et al. 2010] has used advanced multiview stereo (MVS) to derive geometry from several high-resolution cameras under diffuse illumination. While the geometric detail derived by Bradley et al's dynamic system is not at the level of skin mesostructure, [Beeler et al. 2010] infers ad-

Video Face Replacement

Kevin Dale[1] Kalyan Sunkavalli[1] Micah K. Johnson[2] Daniel Vlasic[3] Wojciech Matusik[2,4] Hanspeter Pfister[1]
[1]Harvard University [2]MIT CSAIL [3]Lantos Technologies [4]Disney Research Zurich

(a) Source (b) Target (c) Aligned (d) Three frames of the blended result

Figure 1: *Our method for face replacement requires only single-camera video of the source (a) and target (b) subject, which allows for simple acquisition and reuse of existing footage. We track both performances with a multilinear morphable model then spatially and temporally align the source face to the target footage (c). We then compute an optimal seam for gradient domain compositing that minimizes bleeding and flickering in the final result (d).*

Abstract

We present a method for replacing facial performances in video. Our approach accounts for differences in identity, visual appearance, speech, and timing between source and target videos. Unlike prior work, it does not require substantial manual operation or complex acquisition hardware, only single-camera video. We use a 3D multilinear model to track the facial performance in both videos. Using the corresponding 3D geometry, we warp the source to the target face and retime the source to match the target performance. We then compute an optimal seam through the video volume that maintains temporal consistency in the final composite. We showcase the use of our method on a variety of examples and present the result of a user study that suggests our results are difficult to distinguish from real video footage.

CR Categories: I.4.3 [Image Processing and Computer Vision]: Enhancement—Filtering; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation

Keywords: face replacement, facial animation, video compositing

Links: DL PDF

1 Introduction

Techniques for manipulating and replacing faces in photographs have matured to the point that realistic results can be obtained with minimal user input (e.g., [Agarwala et al. 2004; Bitouk et al. 2008; Sunkavalli et al. 2010]). Face replacement in video, however, poses significant challenges due to the complex facial geometry as well as our perceptual sensitivity to both the static and dynamic elements of faces. As a result, current systems require complex hardware and significant user intervention to achieve a sufficient level of realism (e.g., [Alexander et al. 2009]).

This paper presents a method for face replacement in video that achieves high-quality results using a simple acquisition process. Unlike previous work, our approach assumes inexpensive hardware and requires minimal user intervention. Using a single camera and simple illumination, we capture *source* video that will be inserted into a *target* video (Fig. 1). We track the face in both the source and target videos using a 3D multilinear model. Then we warp the source video in both space and time to align it to the target. Finally, we blend the videos by computing an optimal spatio-temporal seam and a novel mesh-centric gradient domain blending technique.

Our system replaces all or part of the face in the target video with that from the source video. Source and target can have the same person or two different subjects. They can contain similar performances or two very different performances. And either the source or the target can be existing (i.e., uncontrolled) footage, as long as the face poses (i.e., rotation and translation) are approximately the same. This leads to a handful of unique and useful scenarios in film and video editing where video face replacement can be applied.

For example, it is common for multiple takes of the same scene to be shot in close succession during a television or movie shoot. While the timing of performances across takes is very similar, subtle variations in the actor's inflection or expression distinguish one take from the other. Instead of choosing the single best take for the final cut, our system can combine, e.g., the mouth performance from one take and the eyes, brow, and expressions from another to produce a *video montage*.

A related scenario is *dubbing*, where the source and target subject are the same, and the source video depicts an actor in a studio recording a foreign language track for the target footage shot on location. The resulting video face replacement can be far superior to the common approach of replacing the audio track only. In contrast to multi-take video montage, the timing of the dubbing source is completely different and the target face is typically fully replaced,

ACM Reference Format
Dale, K., Sunkavalli, K., Johnson, M., Vlasic, D., Matusik, W., Pfister, H. 2011. Video Face Replacement. *ACM Trans. Graph. 30*, 6, Article 130 (December 2011), 10 pages. DOI = 10.1145/2024156.2024164 http://doi.acm.org/10.1145/2024156.2024164.

A Rendering Framework for Multiscale Views of 3D Models

Wei-Hsien Hsu* Kwan-Liu Ma† Carlos Correa‡
University of California at Davis Lawrence Livermore National Laboratory

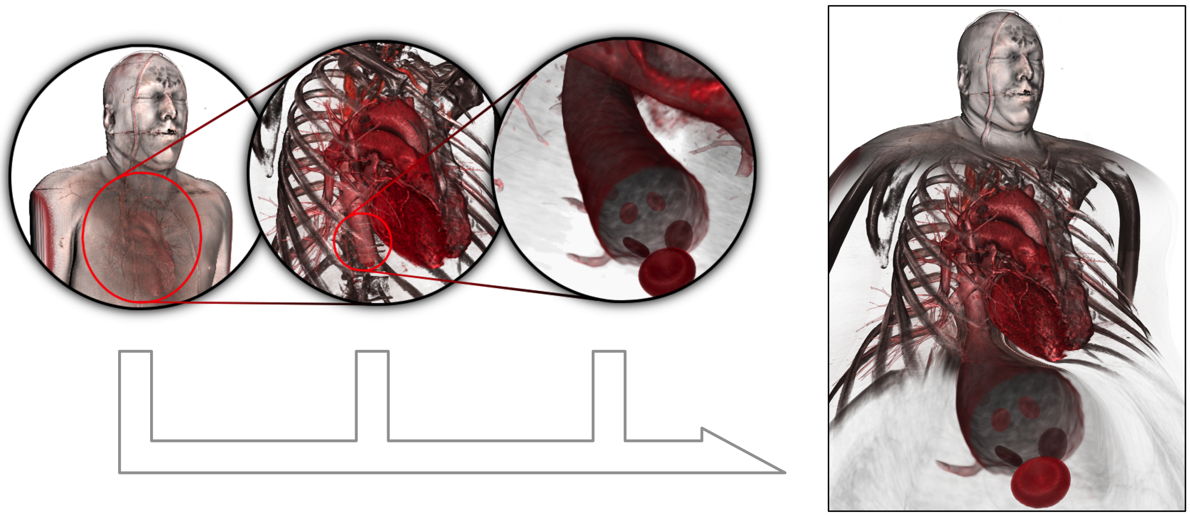


Figure 1: *A continuous multiscale view (right) of a volumetric human body dataset shows three different levels of detail (left three) in a single image. The image on the right is directly rendered with our multiscale framework.*

Abstract

Images that seamlessly combine views at different levels of detail are appealing. However, creating such multiscale images is not a trivial task, and most such illustrations are handcrafted by skilled artists. This paper presents a framework for direct multiscale rendering of geometric and volumetric models. The basis of our approach is a set of non-linearly bent camera rays that smoothly cast through multiple scales. We show that by properly setting up a sequence of conventional pinhole cameras to capture features of interest at different scales, along with image masks specifying the regions of interest for each scale on the projection plane, our rendering framework can generate non-linear sampling rays that smoothly project objects in a scene at multiple levels of detail onto a single image. We address two important issues with non-linear camera projection. First, our streamline-based ray generation algorithm avoids undesired camera ray intersections, which often result in unexpected images. Second, in order to maintain camera ray coherence and preserve aesthetic quality, we create an interpolated 3D field that defines the contribution of each pinhole camera for determining ray orientations. The resulting multiscale camera has three main applications: (1) presenting hierarchical structure in a compact and continuous manner, (2) achieving focus+context visualization, and (3) creating fascinating and artistic images.

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

Keywords: multiscale views, camera model, levels of detail, visualization

Links: DL PDF

*e-mail: whhsu@ucdavis.edu
†e-mail: ma@cs.ucdavis.edu
‡e-mail: correac@llnl.gov

1 Introduction

This project is motivated by an illustration created by artists at the Exploratorium in San Francisco. As shown in Figure 3, this illustration depicts both macro and micro perspectives of the human circulatory system in a continuous landscape across multiple scales. The seamless continuity between scales vividly illustrates how molecules form blood cells, how blood cells are distributed in a blood vessel, how the blood vessel connects to a human heart, and where the heart is located in a human body. The astonishment and fascination evoked by the illustration, along with its high educative value, won it first place in the illustration category of the 2008 U.S. National Science Foundation and Science Magazine Visualization Challenge.

In scientific studies, it is often desirable to illustrate complex physical phenomena, organic structures, and man-made objects. Many of these physical structures are hierarchical in nature. Static multiscale illustrations are frequently used to convey hierarchical structures, such as the anatomy of organ systems and the design of engineered architectures, as shown in Figure 2. Large terrain data, on the other hand, is usually encapsulated in explorable, navigable interactive systems [McCrae et al. 2009; Google 2010]. Animations are also helpful for presenting extremely large datasets

ACM Reference Format
Hsu, W., Ma, K., Correa, C. 2011. A Rendering Framework for Multiscale Views of 3D Models. *ACM Trans. Graph. 30*, 6, Article 131 (December 2011), 9 pages. DOI = 10.1145/2024156.2024165 http://doi.acm.org/10.1145/2024156.2024165.

Mixed-Order Compositing for 3D Paintings

Ilya Baran[1] Johannes Schmid[1,2] Thomas Siegrist[1,2] Markus Gross[1,2] Robert W. Sumner[1]

[1]Disney Research Zurich [2]ETH Zurich

Abstract

We present a method for rendering 3D paintings by compositing brush strokes embedded in space. The challenge in compositing 3D brush strokes is reconciling conflicts between their z-order in 3D and the order in which the strokes were painted, while maintaining temporal and spatial coherence. Our algorithm smoothly transitions between compositing closer strokes over those farther away and compositing strokes painted later over those painted earlier. It is efficient, running in $O(n \log n)$ time, and simple to implement. We demonstrate its effectiveness on a variety of 3D paintings.

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

Keywords: compositing, nonphotorealistic rendering

Links: DL PDF WEB VIDEO

Figure 1: *A 3D painting, consisting of brush strokes embedded in space, composited using our method.* © *Disney Enterprises, Inc.*

1 Introduction

Traditional 2D painting techniques and digital 2D painting systems provide an artist with a lot of expressive freedom in the creation of a painting, allowing a wide range of styles. Translating this freedom into 3D so that paintings can be viewed from multiple angles has been a research challenge for many years. Meier [1996] first proposed generating brush strokes attached to 3D objects. Other systems rely more on the artist. For example, Disney's Deep Canvas [Katanics and Lappas 2003] and, most recently, OverCoat [Schmid et al. 2011], present the artist with a 3D canvas that allows him or her to place stylized paint strokes in space based on 3D proxy geometry. In this paper, we refer to a collection of brush strokes embedded in space as a *3D painting* (Figure 1).

A 3D painting system renders the scene by projecting each brush stroke onto the current view plane and rasterizing it as one or more fragments for every pixel that the stroke overlaps. This rendering method exposes a technical dilemma about the order in which the fragments should be composited. In the 2D painting metaphor, when the artist places a new paint stroke, it obscures all previous paint strokes that it overlaps. Such behavior is achieved by compositing in *stroke order*. From a 3D point of view, however, strokes that are closer to the viewer should obscure those that are farther away, which amounts to compositing in *depth order*. Compositing purely in stroke order negates much of the benefit of 3D painting, as the sense of tangible objects is lost when the view is changed. Compositing purely in depth order, on the other hand, leads to Z-fighting, precluding the artist from painting over existing strokes, and thus ignores an important part of the 2D painting metaphor. This problem dates back to Meier's work [1996] and is illustrated in Figure 2.

Katanics and Lappas [2003] articulated the desire for mixed-order compositing: fragments that, in the artist's mind, belong on the same surface should be composited in stroke order, while those that belong on different surfaces should be composited in depth order. Unfortunately, assigning each fragment to a specific surface is often impossible: the stroke that generated the fragment may span several surfaces, may self-occlude, or may not even conform to a surface at all. The guideline Deep Canvas adopts is therefore to choose the appropriate ordering based on a depth tolerance d: fragments whose depths are within d of each other are assumed to lie on the same surface and composited in stroke order, but fragments that are farther apart are composited in depth order.

In this work, we formalize this idea for the first time. We state the requirements for compositing order and temporal coherence as four properties, which a mixed-order compositing function must satisfy (Section 3.2). We discuss the ways in which existing solutions fail to satisfy one or more of these properties (Section 3.3) and describe the resulting artifacts. We design a function that does satisfy all four properties (Section 3.4) and show how to compute it efficiently (Section 3.5). We present renderings produced by this method and discuss limitations and future extensions (Section 4).

2 Related Work

One of the first uses of paint strokes as rendering primitives was seen in Haeberli's "Paint By Numbers" [1990], which augments

ACM Reference Format
Baran, I., Schmid, J., Siegrist, T., Gross, M., Sumner, R. 2011. Mixed-Order Compositing for 3D Painting. *ACM Trans. Graph. 30*, 6, Article 132 (December 2011), 6 pages. DOI = 10.1145/2024156.2024166 http://doi.acm.org/10.1145/2024156.2024166.

Animated Construction of Line Drawings

Hongbo Fu
City University of Hong Kong

Shizhe Zhou* Ligang Liu
Zhejiang University

Niloy J. Mitra
KAUST/UCL

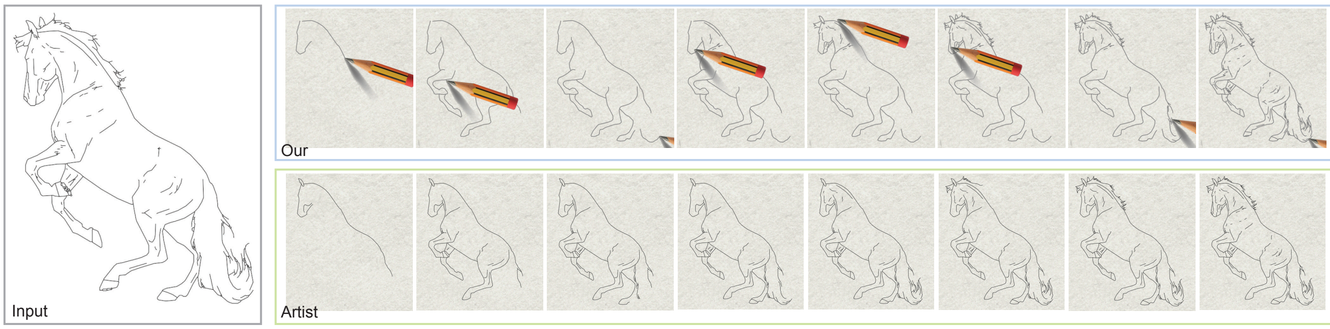


Figure 1: *Guided by sketching principles, we derive a plausible stroke order of an input static line drawing (left) to automatically animate the sketching (right top). A user study shows that the inferred order is comparable to the order used by an artist (right bottom).*

Abstract

Revealing the sketching sequence of a line drawing can be visually intriguing and used for video-based storytelling. Typically this is enabled based on tedious recording of artists' drawing process. We demonstrate that it is often possible to estimate a reasonable drawing order from a static line drawing with clearly defined shape geometry, which looks plausible to a human viewer. We map the key principles of drawing order from drawing cognition to computational procedures in our framework. Our system produces plausible animated constructions of input line drawings, with no or little user intervention. We test our algorithm on a range of input sketches, with varying degree of complexity and structure, and evaluate the results via a user study. We also present applications to gesture drawing synthesis and drawing animation creation especially in the context of video scribing.

Keywords: line art, animation, drawing analysis, video scribing

Links: DL PDF WEB

1 Introduction

Line art is a popular art form, and is widely used for illustrations, caricatures, cartoons, etc. Demonstration videos and commercials (e.g., a series of animations by the RSA) often use animation sequences showing the drawing process of line artworks as a mode of instruction to vividly tell stories. The technique for producing such dynamic line art (with synchronized audio content) is often referred as *video scribing*, which is desirable for building anticipation, directing viewer attention from one object to another, conveying order of action sequences in instructional animations, or simply for maintaining continuity during narrative storytelling. Traditionally, video scribing animations are created by recording the drawing process during the creation of line art images, either by using video cameras or recording functionalities of drawing software. This limits the creation of video scribing to a group of professional drawers.

Our work is motivated by millions of searchable line drawings on the Internet and intends to enable ready reuse of available drawings for video scribing. Given a static line art image, we intend to estimate a drawing order from the image itself, which is visually plausible to a human viewer (see Figure 1). Basic drawing principles [Guptill and Meyer 1997; Willats 1997] naturally demand a solution to the following key problems (see Figure 2): (i) construct a coarse-to-fine hierarchical representation of an input line drawing image, since drawers, both amateurs and professionals, mostly start with a rough sketch and then gradually refine the drawing by introducing additional details, (ii) order a set of drawing strokes, and finally (iii) determine directions for each of the individual strokes.

The order of compilation of a drawing might vary with people or for the same person across time. This implies that our problems typically have multiple plausible solutions. Therefore, we focus on finding one of the *reasonable* solutions that look plausible to a human viewer, instead of searching for the *best* drawing order if any. Although drawing order of line art is subjective and involves personal taste and preferences, drawers indeed follow certain sets of rules due to their stereotypical behavior, as supported by extensive research studies in the cognitive psychology (see [van Sommers 1984; Tversky and Suwa 2009] and references therein).

We propose an effective solution for estimating a reasonable order given a number of 2D lines vectorized from line art images, with most of the curve lines being sharp and cleanly defined. Our work makes the following contributions: (i) introduce the problem of animated construction of line drawings, (ii) summarize geometric guidelines of drawing order from findings in drawing cognition, which are then computationally encoded through analysis of line drawings, and (iii) effectively simulate the drawing process by constructing a coarse-to-fine hierarchy and formulating the ordering of strokes as finding a Hamiltonian path on a graph encoding both the individual properties of lines (e.g., complexity) and their interrelations (e.g., proximity, collinearity, and anchoring).

Since the plausibility of the estimated drawing orders is subjective, we conduct a user study to evaluate the perceived quality of our so-

*Corresponding author: shizhezhou@zju.edu.cn

ACM Reference Format
Fu, H., Zhou, S., Liu, L., Mitra, N. 2011. Animated Construction of Line Drawings. *ACM Trans. Graph. 30*, 6, Article 133 (December 2011), 10 pages. DOI = 10.1145/2024156.2024167 http://doi.acm.org/10.1145/2024156.2024167.

Sketch-based Dynamic Illustration of Fluid Systems

Bo Zhu[1] Michiaki Iwata[3] Ryo Haraguchi[3] Takashi Ashihara[4] Nobuyuki Umetani[2]
Takeo Igarashi[1,2] Kazuo Nakazawa[3]

[1]JST ERATO [2]The University of Tokyo [3]National Cerebral and Cardiovascular Center Research Institute
[4]Shiga University of Medical Science

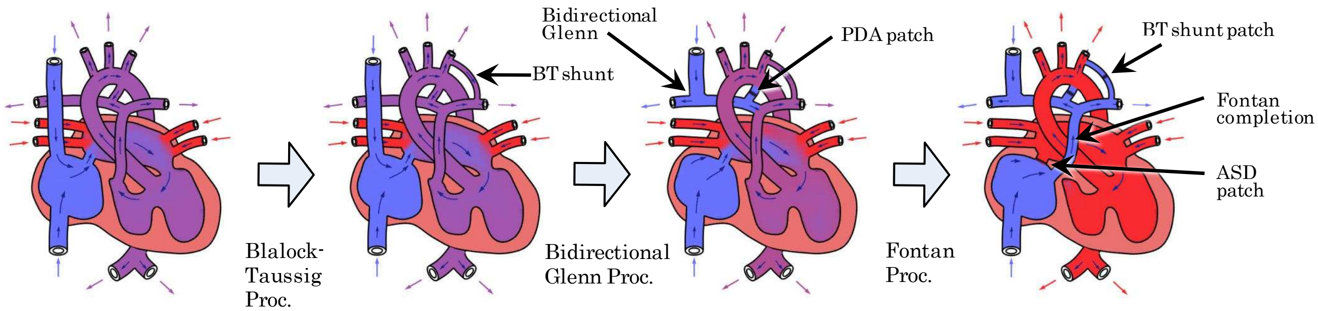


Figure 1: *Illustrations created using our system to describe the surgical repair procedure of tricuspid atresia (TA). The user interactively edits the illustration and the system continuously presents the corresponding blood flow computed by simplified fluid simulation.*

Abstract

This paper presents a lightweight sketching system that enables interactive illustration of complex fluid systems. Users can sketch on a 2.5-dimensional (2.5D) canvas to design the shapes and connections of a fluid circuit. These input sketches are automatically analyzed and abstracted into a hydraulic graph, and a new hybrid fluid model is used in the background to enhance the illustrations. The system provides rich simple operations for users to edit the fluid system incrementally, and the new internal flow patterns can be simulated in real time. Our system is used to illustrate various fluid systems in medicine, biology, and engineering. We asked professional medical doctors to try our system and obtained positive feedback from them.

CR Categories: I.3.8 [Computer Graphics]: Applications—; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques;

Keywords: dynamic illustration, real-time fluid simulation, sketch interface

Links: DL PDF

1 Introduction

Fluid systems are ubiquitous. A typical fluid system includes a fluid to carry the materials to be transported, pipes and regions to distribute the fluid, and external sources and sinks to drive the flow through the system. In the fields of medicine, biology, and engineering, it is important to be able to illustrate how these systems work and how to operate on them dynamically. For example, a doctor may need to illustrate how blood flow patterns inside normal and abnormal hearts differ and how a defective heart can be repaired with a series of surgical operations [Runge and Ohman 2004]; a biologist may need to explain how an animal circulatory system functions and how it transports materials; an engineer may need to explain to a customer how air circulates inside a house after installing new air conditioners. Even though these kinds of illustrations can be enhanced using simple fluid simulations and visualizations, standard simulation systems are too complicated for use in casual and interactive discussions.

This paper presents a sketching system that incorporates a background fluid simulation for illustrating dynamic fluid systems (Figure 1). Our method combines sketching, simulation, and control techniques in one user interface and can produce illustrations of complex fluid systems in real time. Users design the structure of the fluid system using basic sketch operations on a canvas and progressively edit it to show how flow patterns change. The system automatically detects and corrects the structural errors of flow simulation as the user sketches. A fluid simulation runs constantly in the background to enhance flow and material distribution in physically plausible ways.

We developed a hybrid fluid simulation method using multilayered two-dimensional (2D) space for the background simulation. We chose not to use three-dimensional (3D) simulation because this makes the simulation difficult to control and visualize. Two-dimensional illustrations are easier to create and understand, and are therefore more widely used for depicting fluid systems in medical and engineering education. We developed a hybrid method because standard 2D hydrodynamics simulations are too slow for our purposes and cannot handle the multilayered structures typical in these illustrations. Our method combines a hydraulic network model and a multilayered hydrodynamics model to enable efficient flow simulation on different levels. We used a hydraulic graph to represent the entire system on a coarse level and mapped each fluid region onto a multilayered grid to calculate local flow. The flow in

ACM Reference Format
Zhu, B., Iwata, M., Haraguchi, R., Ashihara, T., Umetani, N., Igarashi, T., Nakazawa, K. 2011. Sketch-based Dynamic Illustration of Fluid Systems. *ACM Trans. Graph. 30*, 6, Article 134 (December 2011), 8 pages. DOI = 10.1145/2024156.2024168 http://doi.acm.org/10.1145/2024156.2024168.

Interactive Hybrid Simulation of Large-Scale Traffic

Jason Sewall* David Wilkie† Ming C. Lin‡
Intel Corporation University of North Carolina at Chapel Hill

(a) **(b)**

Figure 1: *(a) Interactive 3D visualization of urban traffic; (b) Augmenting a satellite earth map of a metropolitan region with real-time moving traffic consisting of tens of thousands of vehicles using our method.*

Abstract

We present a novel, real-time algorithm for modeling large-scale, realistic traffic using a hybrid model of both continuum and agent-based methods for traffic simulation. We simulate individual vehicles in regions of interest using state-of-the-art agent-based models of driver behavior, and use a faster continuum model of traffic flow in the remainder of the road network. Our key contributions are efficient techniques for the dynamic coupling of discrete vehicle simulation with the aggregated behavior of continuum techniques for traffic simulation. We demonstrate the flexibility and scalability of our interactive visual simulation technique on extensive road networks using both real-world traffic data and synthetic scenarios. These techniques demonstrate the applicability of hybrid techniques to the efficient simulation of large-scale flows with complex dynamics.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Physically based modeling I.6.8 [Simulation and Modeling]: Types of Simulation—Animation

Keywords: traffic, road networks, hyberbolic models

Links: DL PDF WEB VIDEO

*jason.sewall@intel.com
†wilkie@cs.unc.edu
‡lin@cs.unc.edu

ACM Reference Format
Sewall, J., Wilkie, D., Lin, M. 2011. Interactive Hybrid Simulation of Large-Scale Traffic.
ACM Trans. Graph. 30, 6, Article 135 (December 2011), 11 pages. DOI = 10.1145/2024156.2024169
http://doi.acm.org/10.1145/2024156.2024169.

1 Introduction

Automobile traffic is ubiquitous in the modern world. Traffic simulation techniques for animation, urban planning, and road network design are of increasing interest and importance for analyzing road usage in high-traffic urban environments and for interactive visualization of virtual cityscapes and highway systems.

One of the hallmark applications of 3D graphics is VR flight [Pausch et al. 1992] and driving simulators [Cremer et al. 1997; Donikian et al. 1999; MIT 2011; SUM 2009; Wang et al. 2005] used for training. As today's virtual environments have evolved from the earlier single-user VR systems into online virtual globe systems and open-world games (e.g. Grand Theft Auto), the need to simulate large-scale complex traffic patterns — possibly informed by real-time sensor data — has emerged. These new social, economic, and environmental applications present huge computational demands. Road networks in urban environments can be complex and extensive, and traffic flows on these roads can be enormous, making it a daunting task to model, simulate, and visualize at interactive rates. This paper introduces a hybrid simulation technique that combines the strengths of two broad and disparate classes of traffic simulation to achieve flexible, interactive, and high-fidelity simulation even on very large road networks.

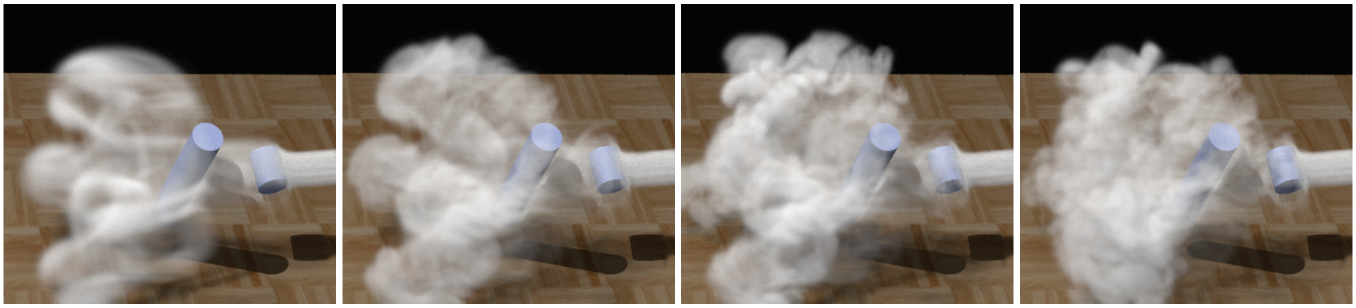
Two classes of simulation techniques are most commonly used in modeling traffic flows. *Agent-based* traffic simulations, also known as *microscopic* methods, determine the motion of each vehicle individually through a series of rules. These rules are easy to vary on a car-to-car basis; such simulation techniques are well-suited to individual vehicles with inhomogeneous governing behaviors. *Continuum*, or *macroscopic*, approaches describe the motion of many vehicles with aggregated behavior; numerical methods are used to solve partial differential equations (PDEs) that model large-scale traffic flows. While agent-based simulation techniques can capture individualistic vehicle behavior, continuum simulations maximize efficiency. Our technique dynamically partitions the simulation domain between these two simulation methodologies to take advantage of their complimentary features.

Pattern-Guided Smoke Animation with Lagrangian Coherent Structure

Zhi Yuan
Kent State University

Fan Chen
Kent State University

Ye Zhao *
Kent State University



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

Figure 1: *Pattern-based fluid animation: high-quality animation is regulated by the flow dynamics pattern from low-resolution simulation. (a) Low-resolution simulation result; (b) High-resolution simulation with pattern regulation; (c) High-resolution simulation with partial pattern regulation on the smoke close to observer; (d) High-resolution simulation without regulation.*

Abstract

Fluid animation practitioners face great challenges from the complexity of flow dynamics and the high cost of numerical simulation. A major hindrance is the uncertainty of fluid behavior after simulation resolution increases and extra turbulent effects are added. In this paper, we propose to regulate fluid animations with predesigned flow patterns. Animators can design their desired fluid behavior with fast, low-cost simulations. Flow patterns are then extracted from the results by the Lagrangian Coherent Structure (LCS) that represents major flow skeleton. Therefore, the final high-quality animation is confined towards the designed behavior by applying the patterns to drive high-resolution and turbulent simulations. The pattern regulation is easily computed and achieves controllable variance in the output. The method makes it easy to design special fluid effects, which increases the usability and scalability of various advanced fluid modeling technologies.

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

Keywords: Flow Pattern, Smoke Animation Design, Fluid Simulation, Lagrangian Coherent Structure, Finite Time Lyapunov Exponent

Links: DL PDF

* {zyuan,fchen,zhao}@cs.kent.edu

ACM Reference Format
Yuan, Z., Chen, F., Zhao, Y. 2011. Pattern-Guided Smoke Animation with Lagrangian Coherent Structure. *ACM Trans. Graph. 30*, 6, Article 136 (December 2011), 8 pages. DOI = 10.1145/2024156.2024170 http://doi.acm.org/10.1145/2024156.2024170.

1 Introduction

In computer graphics, extensive research has been conducted in fluid modeling for a long period, where direct simulation with the application of Computational Fluid Dynamics (CFD) methods has led to substantial success in creating realistic fluid phenomena. However, current fluid modeling technology still imposes great challenge on animators: the nonlinear flow dynamics is laborious to adjust for achieving desired fluid path and shape; and the expensive computational cost hinders their effort to design special effects in an interactive way. Therefore, advanced fluid design tools are direly needed by the animators which, ideally, can provide the functionality as a two-stage process:

- **Design Stage**: Users design fluid behavior with multiple experiments of low-cost simulation. This procedure needs interactive adjustment of initial and boundary conditions, internal structures, and parameters;
- **Output Stage**: Once the expected fluid characteristics (such as major path and shape) are achieved, users can run high-quality simulation to create final animation.

However, unlike image or geometry objects, the transition from experimental design results to final outcome is not straightforward due to the inherent nonlinearity of fluids and numerical dissipation. Flow behavior greatly changes while simulation resolution increases or turbulence enhancement is exploited. The resultant fluid dynamics might not be what an animator prefers and has tested, which will greatly frustrate the animator considering his/her previous design efforts. Although researchers are actively presenting advanced techniques in fluid modeling and simulation, unfortunately, few efforts have been made to support such a two-stage design protocol which can provide great convenience for animators. Lack of such design tools also hamper the usability of the advanced fluid techniques for a wider audience.

In this paper, we propose a novel pattern-based fluid animation approach for advancing fluid modeling in the two-stage animation scenario. Our method focuses on regulating high-quality animation with pre-computed patterns extracted from low-cost simulation results after design experiments. We employ a fast-emerging fluid analysis technique, in particular the Lagrangian Coherent Structure (LCS), to represent flow patterns. LCS defines the *dynamic fluid skeletons* of major flow trends. It provides a geometric in-

Interference-Aware Geometric Modeling

David Harmon[1] Daniele Panozzo[1,2] Olga Sorkine[1,3] Denis Zorin[1]
[1]New York University [2]University of Genova [3]ETH Zurich

Figure 1: *Interference-aware modeling greatly simplifies many complicated modeling tasks. We interactively fit the ogre with a shirt made for a human. We use our ability to fix existing intersections in a mesh and then "shrink-wrap" the shirt on the ogre, ensuring a perfect fit.*

Abstract

While often a requirement for geometric models, there has been little research in resolving the interaction of deforming surfaces during real-time modeling sessions. To address this important topic, we introduce an interference algorithm specifically designed for the domain of geometric modeling. This algorithm is general, easily working within existing modeling paradigms to maintain their important properties. Our algorithm is fast, and is able to maintain interactive rates on complex deforming meshes of over 75K faces, while robustly removing intersections. Lastly, our method is controllable, allowing fine-tuning to meet the specific needs of the user. This includes support for minimum separation between surfaces and control over the relative rigidity of interacting objects.

Links: DL PDF

1 Introduction

Many applications of geometric modeling require constructed shapes to be physically realizable. Shapes created using computer-aided design systems need to be manufactured; if the model is used in a physical simulation—either for special effects, animation, or engineering analysis—its geometric properties should be consistent with those of a real object. One significant impediment to this consistency is intersections within or between modeled objects. These (self-)intersections appear as glaring artifacts, and eliminate the ability to use the final model further down many software pipelines.

Despite its importance, there has been little research on the modeling of surface interference for geometric design. While a number of recent algorithms for collision *detection* are sufficiently fast for interactive applications, collision *response* in the context of geometric modeling has not received much attention. Once interference has been found through detection, the response algorithm modifies positions and trajectories to remove it.

Most existing contact response algorithms are developed for physically based simulations and try to follow the logic of physical laws. As a result, these methods are either too slow (due to strict physical requirements) or cannot be extended for application to general modeling scenarios. Meanwhile, free-form shape design is primarily concerned with surface quality, interactive control, and aesthetics, and typically is not governed by physical equations. The few works on interactive surface deformation that do handle collisions do so as a side effect of their particular modeling paradigm, and are thus limited to those specific tools to model intersection-free surfaces.

In this paper, we present a response algorithm for preventing interference between and within meshed surfaces, formulated in a purely geometric setting. Objects do not have any physical attributes, and their deformation is not necessarily driven by forces. We formulate non-interference constraints on *space-time interference volumes* (STIVs), defined as volumes in space-time traced out by parts of the surface after interpenetration occurs. The advantages of this formulation are two-fold: first, a trajectory-based method can robustly handle problems with thin features, boundaries, and rapid large deformations, without restrictions on the type of geometry. Second, by formulating the non-interference constraint to be zero for each STIV, rather than at the geometric primitive level, the dimension of the numerical problem to be solved is vastly reduced, improving response speed and robustness. Our proposed method has the following features:

- **Independent of deformation model.** Our algorithm is not tied to a specific modeling paradigm; it can resolve intersec-

ACM Reference Format
Harmon, D., Panozzo, D., Sorkine, O., Zorin, D. 2011. Interference-Aware Geometric Modeling. *ACM Trans. Graph. 30*, 6, Article 137 (December 2011), 10 pages. DOI = 10.1145/2024156.2024171 http://doi.acm.org/10.1145/2024156.2024171.

Imperceptible Relaxation of Collision Avoidance Constraints in Virtual Crowds

Richard Kulpa* M2S - Univ. Rennes 2 Golæm S.A.

Anne-Hélène Olivier† INRIA Rennes

Jan Ondřej‡ INRIA Rennes

Julien Pettré§ INRIA Rennes

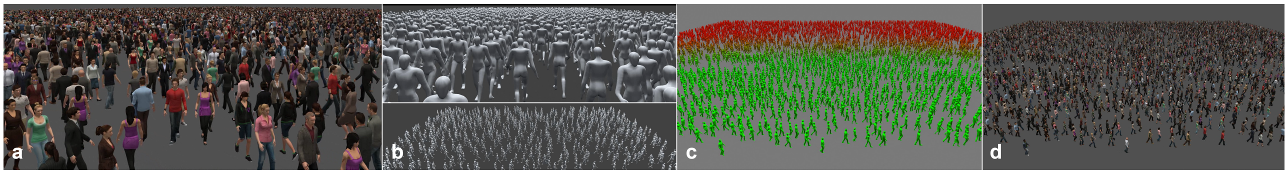


Figure 1: *Hundreds of characters are causing collisions in these crowd scenes. They are however viewed under some conditions that prevent from easily detecting them **(a)**. We performed perceptual studies **(b)** to define a LOD selection function that chooses the characters (in red) for which collisions detection can be avoided **(c)**. Important computation-time savings are obtained from applying the LOD selection function to the design of crowd simulators based on level-of-details methods **(d)**.*

Abstract

The performance of an interactive virtual crowd system for entertainment purposes can be greatly improved by setting a level-of-details (LOD) strategy: in distant areas, collision avoidance can even be stealthy disabled to drastically speed-up simulation and to handle huge crowds. The greatest difficulty is then to select LODs to progressively simplify simulation in an imperceptible but efficient manner. The main objective of this work is to experimentally evaluate spectators' ability to detect the presence of collisions in simulations. Factors related to the conditions of observation and simulation are studied, such as the camera angles, distance to camera, level of interpenetration or crowd density. Our main contribution is to provide a LOD selection function resulting from two perceptual studies allowing crowd system designers to scale a simulation by relaxing the collision avoidance constraint in a least perceptible manner. The relaxation of this constraint is an important source for computational resources savings. Our results reveal several misconceptions in previously used LOD selection functions and suggest yet unexplored variables to be considered. We demonstrate our function efficiency over several evaluation scenarios.

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

Keywords: Crowd simulation, collision avoidance, performance, believability, perception, experimentation

Links: DL PDF

*e-mail: richard.kulpa@univ-rennes2.fr
†e-mail:anne-helene.olivier@inria.fr
‡e-mail:jan.ondrej@gmail.com
§e-mail:julien.pettre@inria.fr

ACM Reference Format
Kulpa, R., Olivier, A., Ondřej, J., Pettré, J. 2011. Imperceptible Relaxation of Collision Avoidance Constraints in Virtual Crowds. *ACM Trans. Graph. 30*, 6, Article 138 (December 2011), 10 pages.
DOI = 10.1145/2024156.2024172 http://doi.acm.org/10.1145/2024156.2024172.

1 Introduction

Interactive virtual crowds require high-performance simulation, animation and rendering techniques to handle numerous characters in real-time. Moreover these characters must be believable in their actions and behaviors. Believability is defined as a trade-off between realism and performance which aims at lowering the complexity of simulation models in favor of performance. The main challenges are to remove the least perceptible details first, to preserve the global visual aspect of results at best and meanwhile, to significantly improve computation times. Level-of-details (LODs) strategies were proposed to combine several realism-performance trade-offs together in a single application. This enables designers to progressively and locally adapt these levels with respect to the visual importance of objects in the displayed scene. A LOD selection function automatically determines the most efficient level to be used for each of the displayed object with respect to several criteria, such as visibility, viewpoint, saliency, etc. Perceptual studies are pertinent answers to the problem of designing LOD selection functions. They were used in previous work on virtual crowds to address some rendering and animation aspects. This paper addresses the problem of scaling crowd simulation models. More specifically, we question the need to solve every single collision within a large moving virtual crowd.

Collision avoidance is a prevailing constraint in the formulation of microscopic crowd simulation models. The absence of interpenetration between bodies ensures that a nominal level of realism is reached when, for example, simulating pedestrian traffic for architectural design. However, in the context of believable crowds, absence of collision is not necessarily the best criterion to guarantee satisfying results and is even sometimes responsible for some strange, but collision-free, maneuvers which are particularly detectable by spectators. In addition, avoidance needs to iteratively check and solve collisions by time-consuming algorithms. Relaxing - in a visually imperceptible manner - the collision avoidance constraint wherever possible simply represents an opportunity for saving important computational resources. Although this artifice has been used in previous work, it was never given an attempt to search for the bounds of collision perception and validate the proposed LOD selection function.

Motivated by these assessments, we propose two successive perceptual studies to inspect the effect of various factors on the visual perception of collisions. The first study focuses on pairwise collision situations out of the context of crowds. Factors related to

A Hybrid Iterative Solver for Robustly Capturing Coulomb Friction in Hair Dynamics

Gilles Daviet Florence Bertails-Descoubes Laurence Boissieux

INRIA, Grenoble, France*

Figure 1: *Comparison of the hair collective behavior between (top) real hair motion sequences and (bottom) our corresponding simulations, based on large assemblies of (up to 2,000) individual fibers with massive self-contacts and Coulomb friction. Our model retains typical emerging effects such as transient coherent motions or stick-slip instabilities. See the accompanying video for the full animations.*

Abstract

Dry friction between hair fibers plays a major role in the collective hair dynamic behavior as it accounts for typical nonsmooth features such as stick-slip instabilities. However, due the challenges posed by the modeling of nonsmooth friction, previous mechanical models for hair either neglect friction or use an approximate smooth friction model, thus losing important visual features. In this paper we present a new generic robust solver for capturing Coulomb friction in large assemblies of tightly packed fibers such as hair. Our method is based on an iterative algorithm where each single contact problem is efficiently and robustly solved by introducing a hybrid strategy that combines a new zero-finding formulation of (exact) Coulomb friction together with an analytical solver as a fail-safe. Our global solver turns out to be very robust and highly scalable as it can handle up to a few thousand densely packed fibers subject to tens of thousands frictional contacts at a reasonable computational cost. It can be conveniently combined to any fiber model with various rest shapes, from smooth to curly. Our results, visually validated against real hair motions, depict typical hair collective effects and greatly enhance the realism of standard hair simulators.

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

Keywords: Hair simulation, hair contacts, Coulomb friction

Links: DL PDF WEB VIDEO

*e-mail: {Gilles.Daviet,Florence.Descoubes,Laurence.Boissieux}@inria.fr

ACM Reference Format
Daviet, G., Bertails-Descoubes, F., Boissieux, L. 2011. A Hybrid Iterative Solver for Robustly Capturing Coulomb Friction in Hair Dynamics. *ACM Trans. Graph. 30*, 6, Article 139 (December 2011), 11 pages. DOI = 10.1145/2024156.2024173 http://doi.acm.org/10.1145/2024156.2024173.

1 Introduction

Hair is an important and distinctive human feature that has an interesting – often visually pleasing – dynamic behavior. Long hair flowing in air or swinging at the pace of a jogger is so common in the real world that one may think hair material arose the curiosity of scientists a long time ago, and that there is no mystery about it anymore. However, unlike fluids or granular materials which have been extensively studied for more than half a century, fibrous materials such as hair have long been neglected, mainly due to the lack of industrial applications at the time. With the recent advances and impacts of Computer Graphics applications in our everyday life, which stimulates the search for realistic models of virtual humans, the modeling and simulation of the hair dynamics has today become an important scientific challenge. Furthermore, this problem now also draws the attention of cosmetologists who intend to better understand and predict the hair mechanical behavior in order to design more adequate and demanding care products.

Modeling hair dynamics nevertheless remains a largely unsolved issue due to the inherent complexity of hair: human hair is composed of 150,000 individual fibers that tightly interact together, leading to a complex collective behavior. Due to the rough surface of the hair fibers, covered with microscopic scales, *dry friction* substantially alters contacts at the fiber level, and consequently greatly influences the hair dynamics at the *macroscopic* level. As illustrated in Figure 2 and in the accompanying video, we have identified three major hair visual features that directly emerge from those nonsmooth frictional contacts occurring at the fiber level:

1. Typical stick-slip instabilities during motion;
2. The spontaneous splitting of hair into multiple untidy wisps and “flyers” during strong motion *vs.* its spontaneous grouping into a few globally coherent locks during gentle motion;
3. The appearance of complex nonsmooth hair patterns that can remain perfectly still at the end of the motion.

We claim that accounting for these phenomena is essential for producing realistic and compelling hair animations.

General Planar Quadrilateral Mesh Design Using Conjugate Direction Field

Yang Liu[*] Weiwei Xu[*] Jun Wang[†] Lifeng Zhu[‡] Baining Guo[*] Falai Chen[†] Guoping Wang[‡]
[*]Microsoft Research Asia [†]University of Science and Technology of China [‡]Peking University

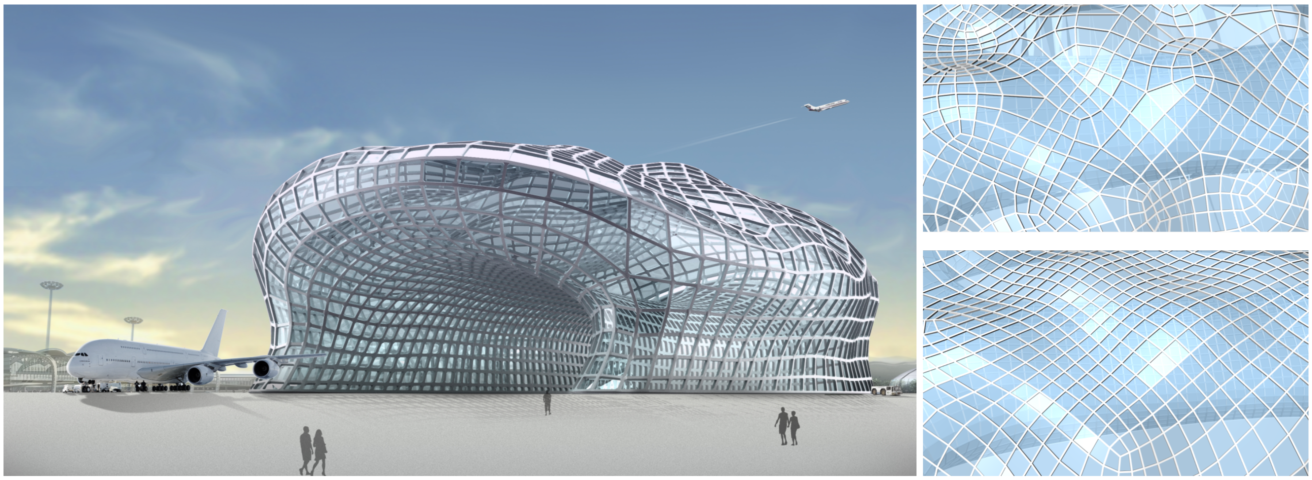


Figure 1: *Left: An airport terminal model with planar quad faces generated by our conjugate direction field method. The maximum value of the planarity measure (the angular difference in degrees between the sum of four internal angles of a quad face and* $360°$*) is* $0.05°$*. Right: A comparison of the planar quad mesh on the roof of this model from the principal curvature network (top) and our method (bottom). Our method allows us to control the layout of the planar quad mesh and reduces the number of singularities (non-four-valence vertices).*

Abstract

We present a novel method to approximate a freeform shape with a planar quadrilateral (PQ) mesh for modeling architectural glass structures. Our method is based on the study of conjugate direction fields (CDF) which allow the presence of $\pm k/4 (k \in \mathbb{Z})$ singularities. Starting with a triangle discretization of a freeform shape, we first compute an as smooth as possible conjugate direction field satisfying the user's directional and angular constraints, then apply mixed-integer quadrangulation and planarization techniques to generate a PQ mesh which approximates the input shape faithfully. We demonstrate that our method is effective and robust on various 3D models.

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

Keywords: planar quadrilateral mesh, conjugate direction field, architectural geometry

Links: DL PDF

ACM Reference Format
Liu, Y., Xu, W., Wang, J., Zhu, L., Guo, B., Chen, F., Wang, G. 2011. General Planar Quadrilateral Mesh Design Using Conjugate Direction Field. *ACM Trans. Graph. 30*, 6, Article 140 (December 2011), 9 pages. DOI = 10.1145/2024156.2024174 http://doi.acm.org/10.1145/2024156.2024174.

1 Introduction

Planar quadrilateral (PQ) meshes are essential in architectural geometry for discretizing a freeform architectural structure with planar quad faces [Glymph et al. 2004; Liu et al. 2006; Pottmann et al. 2007b], and the study of PQ meshes is now an important topic of discrete differential geometry [Pottmann and Wallner 2008; Bobenko and Suris 2008]. Its continuous counterpart, in differential geometry, is the conjugate curve network [Sauer 1970; Liu et al. 2006; Bobenko and Suris 2008], which is defined to be two families of one-parameter curves that cover a smooth surface, and their tangent vectors $\mathbf{v}, \mathbf{w}$ at an arbitrary point $\mathbf{x}$ on a surface are conjugate (*see its formal definition in Section 3*). These two families of tangent directions $\mathbf{v}, \mathbf{w}$ form a general cross field without the requirement of orthogonality, which we call a *conjugate direction field* (CDF) hereafter. The layout of a PQ mesh can be controlled through the design of the CDF.

It has been recognized that an intuitive design tool for smooth CDFs is desirable for architects to control the layout of the PQ mesh [Pottmann et al. 2007a; Eigensatz et al. 2010]. Unfortunately, there is no general solution currently available for CDF design. Existing techniques can handle two special cases. Principal directions, as a typical example of CDFs, have been used in [Liu et al. 2006] to produce PQ meshes. Since the principal directions are unique, there are no degrees of freedom left for the architects. A recent representation-vector based CDF design technique in [Zadravec et al. 2010] is capable of producing a smooth CDF via measuring the smoothness of the representation vector field. However, only singularities with indices of $\pm k/2 (k \in \mathbb{Z})$ can be modeled and it fails in handling $\pm k/4 (k \in \mathbb{Z})$ singularities, such as a surface with convex corners (e.g., a round cube).

The main challenge of general CDF design is how to define a correct smoothness measure for a CDF so that singularities of $\pm k/4 (k \in \mathbb{Z})$ are allowed. Since a CDF on two adjacent faces

Connectivity Editing for Quadrilateral Meshes

Chi-Han Peng*
Arizona State University

Eugene Zhang†
Oregon State University

Yoshihiro Kobayashi‡
Arizona State University

Peter Wonka§
Arizona State University / KAUST

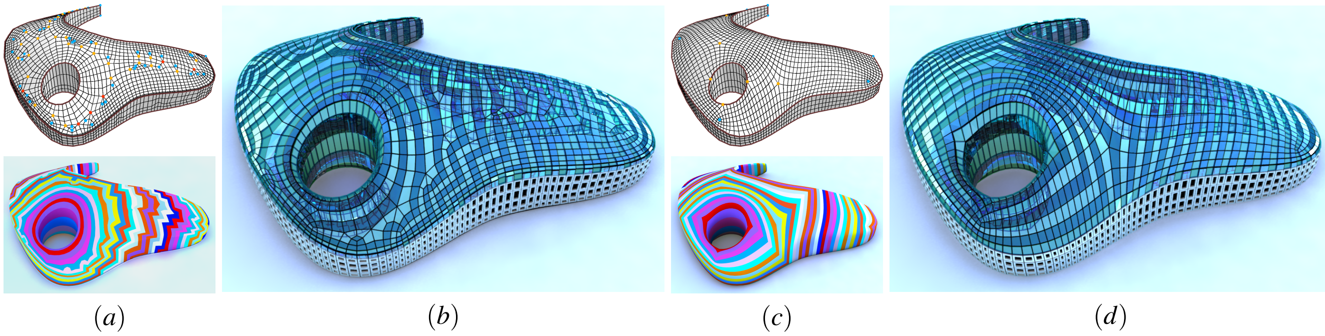


Figure 1: *We introduce connectivity editing operations to control irregular vertices in quadrilateral meshes. This can lead to improved results in the design of a glass structure: (a) top: the original mesh with irregular vertices as colored dots, (a) bottom: a stripe pattern applied to the mesh, (b) a rendering of the design as glass construction. In (c) and (d) we show the edited mesh. The glass panels on the roof are generated from the edges in the meshes.*

Abstract

We propose new connectivity editing operations for quadrilateral meshes with the unique ability to explicitly control the location, orientation, type, and number of the irregular vertices (valence not equal to four) in the mesh while preserving sharp edges. We provide theoretical analysis on what editing operations are possible and impossible and introduce three *fundamental* operations to move and re-orient a pair of irregular vertices. We argue that our editing operations are fundamental, because they only change the quad mesh in the smallest possible region and involve the fewest irregular vertices (i.e., two). The irregular vertex movement operations are supplemented by operations for the splitting, merging, canceling, and aligning of irregular vertices. We explain how the proposed high-level operations are realized through graph-level editing operations such as quad collapses, edge flips, and edge splits. The utility of these mesh editing operations are demonstrated by improving the connectivity of quad meshes generated from state-of-art quadrangulation techniques.

Keywords: quadrilateral mesh editing, irregular vertex editing, mesh optimization, mesh-based design, topology, geometry processing

Links: DL PDF WEB

*e-mail:pchihan@asu.edu
†e-mail:zhange@eecs.oregonstate.edu
‡e-mail:ykobaya@asu.edu
§e-mail:peter.wonka@asu.edu

ACM Reference Format
Peng, C., Zhang, E., Kobayashi, Y., Wonka, P. 2011. Connectivity Editing for Quadrilateral Meshes. *ACM Trans. Graph. 30*, 6, Article 141 (December 2011), 12 pages. DOI = 10.1145/2024156.2024175 http://doi.acm.org/10.1145/2024156.2024175.

1 Introduction

Quadrilateral and hexahedral meshes are popular choices in simulation and shape modeling due to the natural tensor product property that they possess. Quadrilateral meshes can also facilitate architectural modeling as well as texture and geometry synthesis. Important aspects of a quadrilateral mesh include the location, orientation, type, and number of irregular vertices. While there has been some work in quad mesh connectivity editing [Daniels et al. 2008; Bommes et al. 2011], achieving irregular vertex control is challenging and many questions about what editing operations are possible and impossible still need to be answered.

In this paper, we propose three operations that move an irregular vertex pair (two valence 3, two valence 5, or one valence 3 and one valence 5) over the mesh. To show that these are fundamental operations for quad mesh editing, we will establish the following properties:

- These editing operations impact the smallest possible region on the mesh and are therefore as local as possible (in a convex region).
- A region containing only one irregular vertex cannot be edited.
- A region containing two irregular vertices can be edited by changing the location of the irregular vertices within the region. However, they cannot be canceled. Some irregular vertex pairs can be merged while others cannot, depending on their graph distance in the initial configuration.
- A region with three irregular vertices can be edited by canceling or merging the irregular vertices.
- Our three movement operations can perform all possible edits within a (convex) region that contains two irregular vertices.

While the three vertex pair movement operations are at the core of this paper, we also introduce several supplementary operations to control the type and number of irregular vertices: splitting, merging, cancellation, and alignment. All of these operations can be realized through three graph-level editing operations: quad collapses,

Simple Quad Domains for Field Aligned Mesh Parametrization

Marco Tarini[1,2] Enrico Puppo[3] Daniele Panozzo[3] Nico Pietroni[1] Paolo Cignoni[1]

[1]ISTI-CNR, Italy [2]Università dell'Insubria, Italy [3]Università di Genova, Italy

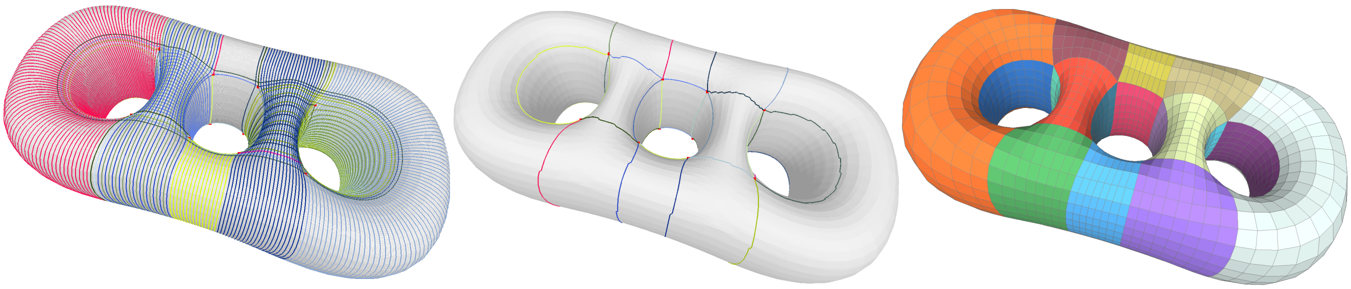


Figure 1: *(Left) An input mesh of quads induces a cross field with an entangled graph of separatrices defining almost eight thousand domains; (center) the graph is disentangled with small distortion from the input field to obtain just twenty parametrization domains; (right) parametrization is smoothed to make it conformal; an example of remeshing from the parametrization.*

Abstract

We present a method for the global parametrization of meshes that preserves alignment to a cross field in input while obtaining a parametric domain made of few coarse axis-aligned rectangular patches, which form an abstract base complex without T-junctions. The method is based on the topological simplification of the cross field in input, followed by global smoothing.

CR Categories: I.3.5 [Computer graphics]: Computational geometry and object modeling—Curve, surface, solid and object repres.

Keywords: quad mesh, mesh parametrization

Links: DL PDF WEB

1 Introduction

Finding a high-quality parameterization $f : D \to M$ for a given 3D polygonal mesh M is a prerequisite in a number of applications, such as quad-based semiregular remeshing, texture mapping, compression, fitting high order surfaces, physical simulations, tangent space geometry processing, and even tasks outside computer graphics, like physical modeling with metal sheets.

The definition of quality for f depends on the application, but usually encompasses criteria like injectivity, isometry (implying angle preservation and area preservation), smoothness (continuity of gradient vectors), and alignment of gradient vectors with geometric features of M. This problem has been addressed with a wide arsenal of tools [Hormann et al. 2008] and good automatic results are becoming increasingly common.

A recent trend is to first define a cross field $\mathcal{C}$ over M, and then to find f such that its gradient vectors match $\mathcal{C}$ as much as possible. Interestingly, each of the criteria above can be redefined in terms of desired properties of $\mathcal{C}$. Thus the task is shifted from the definition of a good parameterization to the definition of a good cross field $\mathcal{C}$ for a given M. High quality parametrization can be obtained following this approach [Ray et al. 2006; Bommes et al. 2009]. It is now appearent that the definition of a good cross field $\mathcal{C}$ implies, among other things, the good placing of a few irregular points (a.k.a. cone singularities). Irregular points tend to be needed, for example, in places where M exhibits high Gaussian curvature.

In this paper, we focus on an important *additional* criterion for the quality of f, which we refer to as the *simplicity* of domain D (see below for an informal definition). Simplicity determines how much a parametrization f will be effectively useful in most applications, just as much as the other criteria listed above. As we will show, a cross field $\mathcal{C}$ designed to satisfy all the above conventional criteria, but not simplicity, will usually fail producing an acceptably simple domain. Still, it is often the case that a slightly modified cross field $\mathcal{C}'$ exists, which is able to generate a parametrization f with a dramatically simpler domain, while preserving to a large extent the other qualities of $\mathcal{C}$. This work presents a way to obtain the cross field $\mathcal{C}'$, given $\mathcal{C}$.

1.1 Objective: Domain Simplicity

For topologies of M other than the disk, the domain D must necessarily include discontinuities (a.k.a. cuts, or seams): two infinitesimally close points m_0 and m_1 of M lying of different sides of the cut may be mapped by f^{-1} to arbitrarily distant positions d_0 and d_1 of D. The values d_0 and d_1 are often constrained to be reciprocally associated with a "transition function" associated to that cut.

Simplicity of domain D is a concept encapsulating: how many discontinuities are needed (the fewer, the simpler); how simple the discontinuity lines are in D (e.g., straight axis-aligned lines are simpler than curved or jagged lines); and also how constrained and straightforward the transition functions are (the more constrained, the simpler). For example, a domain D consisting of a single flat unit square, with no seams, would exhibit the maximal possible domain simplicity (possible only for disk-like M). On the other

ACM Reference Format
Tarini, M., Puppo, E., Panozzo, D., Pietroni, N., Cignoni, P. 2011. Simple Quad Domains for Field Aligned Mesh Parametrization. *ACM Trans. Graph. 30*, 6, Article 142 (December 2011), 11 pages.
DOI = 10.1145/2024156.2024176 http://doi.acm.org/10.1145/2024156.2024176.

Boundary Aligned Smooth 3D Cross-Frame Field

Jin Huang[†] Yiying Tong[‡] Hongyu Wei[†] Hujun Bao[†]
[†] State Key Lab of CAD&CG, Zhejiang University [‡] Michigan State University

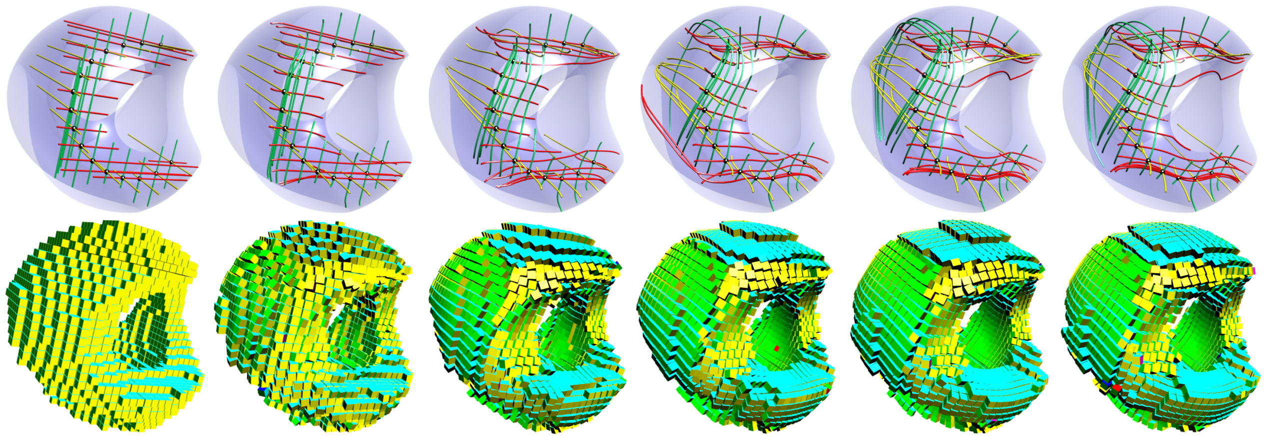


Figure 1: *Snapshots of the optimization procedure to construct a boundary aligned 3D cross-frame field. The top row shows the internal streamlines. The next row contains another visualization with cubes spread by a parameterization along the current cross-frame field and rotated by the current local frame* $R(\Phi)$*. The corresponding number of iteration is shown at the bottom.*

Abstract

In this paper, we present a method for constructing a 3D *cross-frame field*, a 3D extension of the 2D cross-frame field as applied to surfaces in applications such as quadrangulation and texture synthesis. In contrast to the surface cross-frame field (equivalent to a 4-Way Rotational-Symmetry vector field), symmetry for 3D cross-frame fields cannot be formulated by simple one-parameter 2D rotations in the tangent planes. To address this critical issue, we represent the 3D frames by spherical harmonics, in a manner invariant to combinations of rotations around any axis by multiples of $\pi/2$. With such a representation, we can formulate an efficient smoothness measure of the cross-frame field. Through minimization of this measure under certain boundary conditions, we can construct a smooth 3D cross-frame field that is aligned with the surface normal at the boundary. We visualize the resulting cross-frame field through restrictions to the boundary surface, streamline tracing in the volume, and singularities. We also demonstrate the application of the 3D cross-frame field to producing hexahedron-dominant meshes for given volumes, and discuss its potential in high-quality hexahedralization, much as its 2D counterpart has shown in quadrangulation.

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

Keywords: hexahedral, spherical harmonics, N-RoSy frame field

Links: DL PDF

ACM Reference Format
Huang, J., Tong, Y., Wei, H., Bao, H. 2011. Boundary Aligned Smooth 3D Cross-Frame Field.
ACM Trans. Graph. 30, 6, Article 143 (December 2011), 8 pages. DOI = 10.1145/2024156.2024177
http://doi.acm.org/10.1145/2024156.2024177.

1 Introduction

Many recent quadrangulation methods start by constructing a smooth field of orientations determined up to a rotation of π or $\pi/2$. Substantial progress has been made towards the generation of quadrilateral meshes with controlled element sizes and edge directions by optimizing such fields. However, many applications require discretization of 3D volumes rather than just their boundary surfaces. Applications such as simulated elasticity of 3D volumetric objects, computational electromagnetics, and computational fluid dynamics require Finite Element, Finite Volume, or Finite Difference methods on a discretized domain. These methods benefit from a high-quality hexahedral mesh, since hexhedral meshes offer several numerical advantages over tetrahedral meshes due to their tensor product nature. They are also desirable for applications such as geometric design and B-spline fitting, and amenable to applications such as 3D texture atlases. In addition, hexahedral meshes often capture the symmetries of 3D objects and domains better than tetrahedral meshes, thus making the model more intuitive to designers or animators. However, the automatic generation of a hexahedral mesh for a given curved 2D boundary with feature alignment, sizing, and regularity control remains far more challenging than automatic tetrahedralization.

In this paper, we focus on frame field construction, a critical step towards automatic generation of hexahedral meshes that meet the aforementioned requirements. We propose constructing a 3D local *frame field* akin to the cross field in curved 2D surfaces. We use a spherical harmonics-based optimization to obtain a smooth frame field, which can be seen as three vector fields that are continuous up to a simultaneous 3D rotation of all three fields composed of $\pi/2$ rotations around the orthonormal axes. We show as an example that a 3D parameterization guided by the frame field can be constructed and used to generate the final boundary- and feature-aligned mesh with well-shaped hexahedral elements.

Material Matting

Daniel Lepage Jason Lawrence
University of Virginia

Abstract

Despite the widespread use of measured real-world materials, intuitive tools for editing measured reflectance datasets are still lacking. We present a solution inspired by natural image matting and texture synthesis to the *material matting* problem, which allows separating a measured spatially-varying material into simpler foreground and background component materials and a corresponding opacity map. We approach this problem in the context of Bayesian statistics and introduce a new prior on materials that favors those with highly self-similar stochastic structure. We describe a prototype system that iteratively performs these separations based on small sets of user scribbles and demonstrate multiple separations and edits.

Keywords: Appearance Models, Spatially-Varying BRDFs, Material Separation, Texture Synthesis, Matting

Links: DL PDF

1 Introduction

Modern computer graphics systems rely on accurate models of the appearance of real-world materials. Due to the difficulty of constructing convincing appearance models by hand, they are often measured from physical samples. However, artists still desire the ability to edit and manipulate appearance models derived from measured data. For example, the material at the top of Figure 1, measured from the cover of an old book, has three components: scuff marks, red ink, and an underlying golden paper. An artist may wish to make the paper shinier, change the color of the ink, or alter the spatial pattern of the scuffs. This paper describes a new separation method that can assist in these tasks.

The appearance of a uniform opaque material is fully characterized by a four-dimensional function, $f_r(\omega_i, \omega_o)$, called the Bidirectional Reflectance Distribution Function, or BRDF. The BRDF is equal to the fraction of light reflected in the direction ω_o due to light striking the surface from the direction ω_i [Nicodemus et al. 1977]. Materials whose reflectance properties vary over their surface can be described by assigning a BRDF to every point on the surface, resulting in a six-dimensional function $S_r(\vec{x}, \omega_i, \omega_o)$, called the Spatially-Varying BRDF or SVBRDF.

While many methods exist for editing BRDFs [Dorsey et al. 2007], or SVBRDFs in which every point on the surface is expressed as a linear combination of a small number of basis BRDFs [Lawrence et al. 2006; Goldman et al. 2005], many real-world SVBRDFs are more naturally thought of as *mixtures of spatially-varying components*. The book in Figure 1, for example, is composed of several different materials that each exhibit some degree of spatial variation, such as the different shades of red within the red stripes, or the varying color of the golden paper.

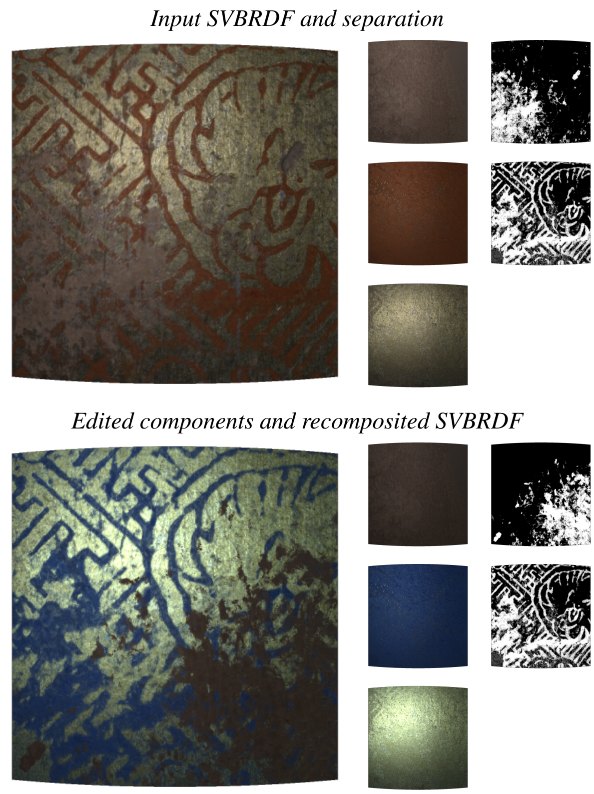


Figure 1: *A separation and edit performed with our system.* ***Top:*** *The input to our system (left) is an SVBRDF, in this case an old book with two materials on its cover that has been scuffed in several places. A user decomposes this image into a set of three component SVBRDFs (center) – the scuffed marks, the red ink, and the underlying gold paper – with corresponding per-pixel opacity maps for each layer (right).* ***Bottom:*** *The user then edits these components by moving and darkening the scuffs, changing the color of the ink and both increasing the shininess and changing the color of the paper. Recompositing the modified layers yields a realistic alteration of the original SVBRDF.*

We present a method for separating an SVBRDF into multiple simpler layers with corresponding opacities, so that a user can manipulate their appearance (specularity, color, contrast, etc.) and spatial distribution independently. Inspired by natural image matting, we associate an *opacity map* α with an SVBRDF layer F and define *composition* of F *over* another SVBRDF layer B as

$$F \oplus^{\alpha} B \equiv \alpha F + (1-\alpha)B. \tag{1}$$

Thus, the *material matting problem* is to decompose a composite SVBRDF V into a set of N layers L_i each with per-pixel opacity encoded in an opacity map α_i such that

$$V = L_1 \oplus^{\alpha_1} \left(L_2 \oplus^{\alpha_2} \left(\cdots \left(L_{N-1} \oplus^{\alpha_{N-1}} L_N\right) \cdots\right)\right). \tag{2}$$

ACM Reference Format
Lepage, D., Lawrence, J. 2011. Material Matting. *ACM Trans. Graph. 30*, 6, Article 144 (December 2011), 10 pages. DOI = 10.1145/2024156.2024178 http://doi.acm.org/10.1145/2024156.2024178.

Physically-Based Interactive Bi-Scale Material Design

Hongzhi Wu Julie Dorsey Holly Rushmeier
Computer Graphics Group, Yale University

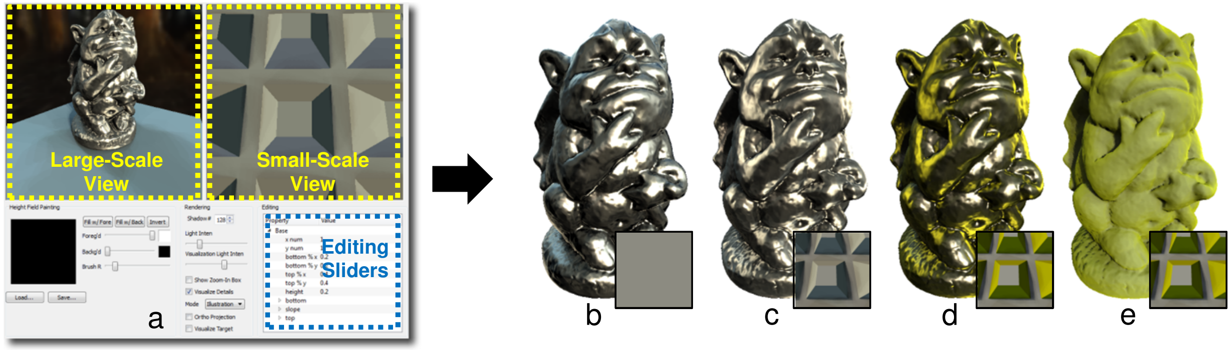


Figure 1: *Bi-scale material design: large-scale appearance changes produced by editing of small-scale geometry (b & c), color (d) and BRDFs (e), using our interactive system (a), which quickly updates the appearance of the large-scale object after small-scale edits. The small-scale details (rendered with a Lambertian BRDF for a better visualization) are shown in the bottom right corner of (b-e). In (d), the color of the small-scale side faces is adjusted to yellow. In (e), the small-scale material is changed from a measured silver-metallic-paint2 BRDF to a Lambertian model.*

Abstract

We present the first physically-based interactive system to facilitate the appearance design at *different scales* consistently, through manipulations of both small-scale geometry and materials. The core of our system is a novel reflectance filtering algorithm, which rapidly computes the large-scale appearance from small-scale details, by exploiting the low-rank structures of the Bidirectional Visible Normal Distribution Function and pre-rotated BRDFs in the matrix formulation of our rendering problem. Our algorithm is three orders of magnitude faster than a ground-truth method. We demonstrate various editing results of different small-scale geometry with analytical and measured BRDFs. In addition, we show the applications of our system to physical realization of appearance, as well as modeling of real-world materials using very sparse measurements.

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

Keywords: bi-scale, material editing, reflectance filtering, low-rank matrix

Links: DL PDF WEB VIDEO

1 Introduction

The appearance of materials can vary considerably when viewed at different scales. For example, individual grains of sand or fabric threads that are visible on close view merge into a material described by a single reflectance function when viewed from a distance. Physically speaking, the large-scale appearance is uniquely determined by averaging the look of small-scale details [Bruneton and Neyret 2011]. Therefore, it would be desirable to build an editing system for interactive appearance design at *different scales*, by manipulating small-scale structures. This could be useful in applications like building exterior design, where the user edits the looks of a building at different view distances.

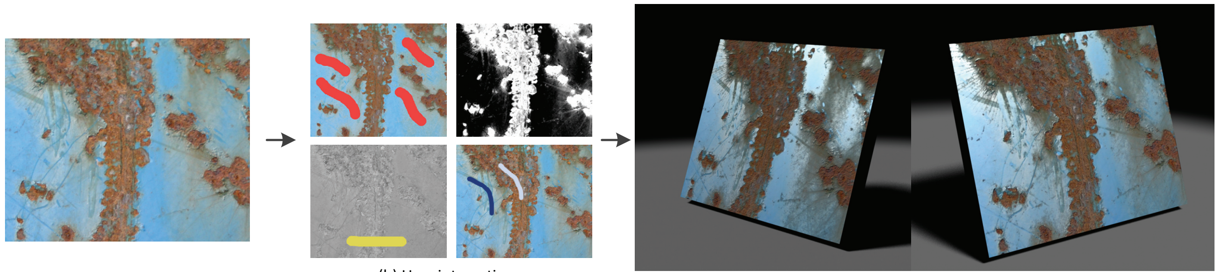
Existing interactive material editing systems (e.g. [Ben-Artzi et al. 2006; Pellacini and Lawrence 2007]) focus on adjusting material appearance only at a *single scale*. On the other hand, previous work [Westin et al. 1992; Gondek et al. 1994], which computes realistic large-scale appearance by simulating light interactions in small-scale details, is too slow to provide interactive feedback. Although converting small-scale structures to large-scale appearance is essentially performing reflectance filtering, related techniques [Bruneton and Neyret 2011] are not suitable for our purpose, due to the lack of support for general geometry and materials [Han et al. 2007], or costly computational overhead [Wu et al. 2009].

This paper presents, to our knowledge, the first physically-based interactive bi-scale material editing system, which manipulates small-scale geometry and Bidirectional Reflectance Distribution Functions (BRDFs), to facilitate appearance design at two different scales consistently. The user can freely change both small-scale geometry and materials, then our system quickly computes the large-scale appearance to provide interactive visual feedback. As illustrated in Fig. 1, various small-scale edits can have dramatic effects on appearance. We achieve an acceleration rate of over 5000:1, when compared with a ground-truth method similar to [Westin et al. 1992], implemented on modern hardware. The key to the performance of our system is a novel reflectance filtering algorithm, which efficiently processes the *Bidirectional Visible Normal Distribution Function* (BVNDF) and pre-rotated BRDFs, derived from the changing small-scale details. We observe and exploit the low-rank structures in both quantities to accelerate the large-scale appearance computation, using Singular Value Decomposition (SVD) combined with the random projection method [Vempala 2004].

ACM Reference Format
Wu, H., Dorsey, J., Rushmeier, H. 2011. Physically-Based Interactive Bi-Scale Material Design. *ACM Trans. Graph. 30*, 6, Article 145 (December 2011), 10 pages. DOI = 10.1145/2024156.2024179
http://doi.acm.org/10.1145/2024156.2024179.

AppGen: Interactive Material Modeling from a Single Image

Yue Dong[*,†] Xin Tong[†] Fabio Pellacini[‡,§] Baining Guo[*,†]
[*]Tsinghua University [†]Microsoft Research Asia [‡]Dartmouth College [§]Sapienza University of Rome



(a) Input image (b) User interaction (c) Rendering result

Figure 1: *Given a single image (a), our system models the spatially varying reflectance properties and normals with a few strokes specified by the user (b). The resulting material can be rendered under different lighting and viewing conditions (c).*

Abstract

We present *AppGen*, an interactive system for modeling materials from a single image. Given a texture image of a nearly planar surface lit with directional lighting, our system models the detailed spatially-varying reflectance properties (diffuse, specular and roughness) and surface normal variations with minimal user interaction. We ask users to indicate global shading and reflectance information by roughly marking the image with a few user strokes, while our system assigns reflectance properties and normals to each pixel. We first interactively decompose the input image into the product of a diffuse albedo map and a shading map. A two-scale normal reconstruction algorithm is then introduced to recover the normal variations from the shading map and preserve the geometric features at different scales. We finally assign the specular parameters to each pixel guided by user strokes and the diffuse albedo. Our system generates convincing results within minutes of interaction and works well for a variety of material types that exhibit different reflectance and normal variations, including natural surfaces and man-made ones.

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

Keywords: appearance modeling, user interaction

Links: DL PDF

1 Introduction

Modeling Realistic Materials. The use of realistic materials is necessary when rendering high-quality images. Measured materials provide the highest quality datasets, but are cumbersome to use in practice due to complex acquisition setups, lengthy measurement times and the size of the generated data [Weyrich et al. 2009]. Today, the vast majority of applications use materials painstakingly modeled by artists. Typically, artists start from a single texture image, and use that to generate spatially-varying diffuse, specular and roughness coefficients of an analytic reflectance model together with a bump map to enrich the surface details. For many materials, this process takes hours to perform, involving the use of image manipulation programs (e.g. Photoshop), inverse shading tools (e.g. CrazyBump), and shading networks in 3D software (e.g. Maya). Not only is this process cumbersome, but it often does not lead to the highest quality materials since no robust method can be used to easily derive detailed reflectance and normal maps from the image. Fig. 3 shows two example materials generated by an experienced artist in roughly one hour each, using the standard toolset that includes Photoshop and CrazyBump.

AppGen. In this paper, we present *AppGen*, an interactive system for modeling material from a single image. We focus on modeling spatially-varying reflectance (i.e. diffuse, specular and roughness parameters) and normal variations from a texture image that is captured from a nearly planar surface lit by directional lighting. Such images are easily found in texture collections since they are widely used by artists when manually modeling materials. Our goal is not to determine the exact reflectance and normals from such single images, which is a well-known ill-posed problem. Instead, we are interested in significantly speeding up the workflow of artists when modeling such materials. Our key idea is that we can keep user interaction minimal by asking the user to specify shading or reflectance information on a few pixels with sparse strokes, while our algorithm efficiently infers the reflectance and normal details for all pixels in the image. Fig. 1 shows one example of a material modeled using our system with just a few user strokes. Note the highly detailed, realistic look of the output material. Our experienced artist was able to regenerate the example materials in Fig. 3 in a few minutes using *AppGen*.

Our system consists of four steps, illustrated in Fig. 2. (1) First, we remove the highlight and shadow pixels in the input image and fill them by image inpainting. After that, we are left with an image of only the diffuse contribution. (2) We present an algorithm for interactively separating the texture image into the product of shading and diffuse albedo. We assume that in each local region pixels with the same chroma value belong to the same material and have the same albedo intensity, while groups of pixels with different chroma values share the same global geometry and thus have the same average shading. Based on these two assumptions, we formu-

ACM Reference Format
Dong, Y., Tong, X., Pellacini, F., Guo, B. 2011. AppGen: Interactive Material Modeling from a Single Image. *ACM Trans. Graph. 30*, 6, Article 146 (December 2011), 10 pages. DOI = 10.1145/2024156.2024180 http://doi.acm.org/10.1145/2024156.2024180.

AppWarp: Retargeting Measured Materials by Appearance-Space Warping

Xiaobo An* Xin Tong† Jonathan D. Denning* Fabio Pellacini*‡
*Dartmouth College †Microsoft Research Asia ‡Sapienza University of Rome

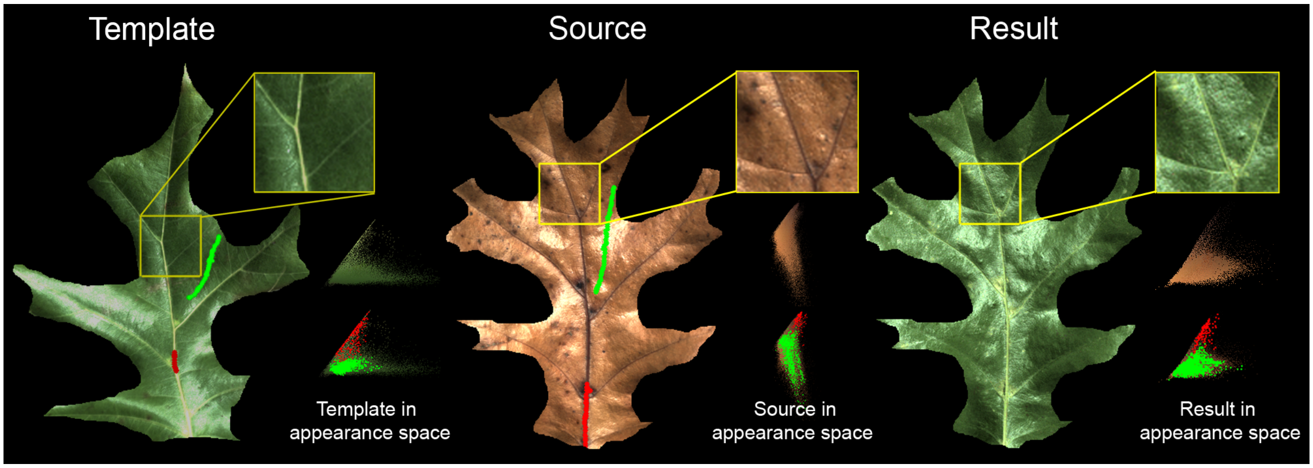


Figure 1: *A material retargeting result, generated by our* AppWarp *algorithm, where a* source *measured material (brown leaf) is edited by applying the reflectance functions found on a* template *measured material (green leaf). To guide the retargeting process, the user marks pairs of strokes to specify corresponding regions on source and template. In this case, the user marks the stem (red strokes) and body (green strokes) on the two leaves. From these correspondences,* AppWarp *computes a retargeted material that maintains the spatial patterns of the source while exhibiting reflectance behaviors of the template. Below the zoom-in, the reflectance functions are shown projected into two dimensions with corresponding user's strokes overlaid. Data from [Lawrence 2009].*

Abstract

We propose a method for retargeting measured materials, where a source measured material is edited by applying the reflectance functions of a template measured dataset. The resulting dataset is a material that maintains the spatial patterns of the source dataset, while exhibiting the reflectance behaviors of the template. Compared to editing materials by subsequent selections and modifications, retargeting shortens the time required to achieve a desired look by directly using template data, just as color transfer does for editing images. With our method, users have to just mark corresponding regions of source and template with rough strokes, with no need for further input.

This paper introduces *AppWarp*, an algorithm that achieves retargeting as a user-constrained, appearance-space warping operation, that executes in tens of seconds. Our algorithm is independent of the measured material representation and supports retargeting of analytic and tabulated BRDFs as well as BSSRDFs. In addition, our method makes no assumption of the data distribution in appearance-space nor on the underlying correspondence between source and target. These characteristics make *AppWarp* the first general formulation for appearance retargeting. We validate our method on several types of materials, including leaves, metals, waxes, woods and greeting cards. Furthermore, we demonstrate how retargeting can be used to enhance diffuse texture with high quality reflectance.

ACM Reference Format
An, X., Tong, X., Denning, J., Pellacini, F. 2011. AppWarp: Retargeting Measured Materials by Appearance-Space Warping. *ACM Trans. Graph. 30*, 6, Article 147 (December 2011), 9 pages.
DOI = 10.1145/2024156.2024181 http://doi.acm.org/10.1145/2024156.2024181.

Links: DL PDF WEB

1 Introduction

Editing Measured Materials. In the past years, the use of measured materials in Computer Graphics has grown since these datasets capture the nuances of real-world surface appearance. For many applications, editing these datasets is desired to permit artistic control. The editing process is typically comprised of performing soft selections on the data and applying edits to each selected region. For example, one might want to increase the roughness of the body of a leaf without changing its stem. Prior work has focused on simplifying selection [Pellacini and Lawrence 2007; An and Pellacini 2008], but does not address the issue of finding the proper editing parameters. For anything but the simplest cases, the latter remains remarkably cumbersome since editing parameters are not intuitively related to appearance changes. This implies that significant trial-and-error is needed after selection, e.g., minutes for simplest cases as reported in [Kerr and Pellacini 2010]. This is further complicated by the fact that different material datasets use different representations: from analytic (e.g. Cook-Torrance [1982] BRDFs) to sampled tables (e.g., [Ashikhmin et al. 2000]). These different representations require entirely different editing operations and potentially hundreds of parameters for the sampled representations, making the process even harder.

Retargeting Measured Materials. In this paper, we focus on editing spatially-varying measured materials that can be described by spatially-varying BRDFs (bidirectional reflectance distribution function) for opaque surfaces or heterogenous BSSRDFs (bidirectional subsurface reflectance distribution function) for optically-thick translucent materials. Inspired by color transfer methods, we propose to edit a *source* measured material by applying the reflectance functions defined in a *template* dataset. For example, as seen in Fig. 1, one might want to apply the reflectance functions

Fusion of Depth Maps with Multiple Scales

Simon Fuhrmann
TU Darmstadt

Michael Goesele
TU Darmstadt

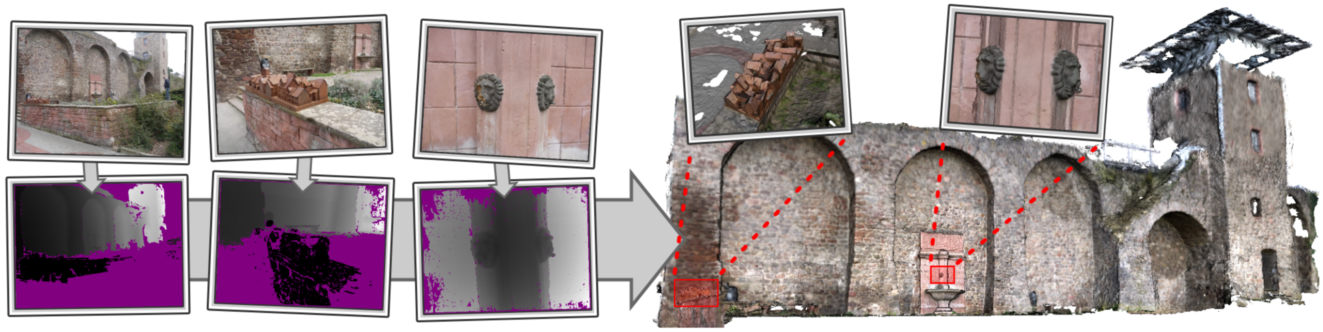


Figure 1: *Reconstruction pipeline. Input photographs (top left) depicting objects at different levels of detail. Multi-view stereo yields depth maps (bottom left), which inherit these multi-scale properties. Our system is able to fuse such depth maps and produce an adaptive mesh (right) with coarse regions as well as fine scale details (insets).*

Abstract

Multi-view stereo systems can produce depth maps with large variations in viewing parameters, yielding vastly different sampling rates of the observed surface. We present a new method for surface reconstruction by integrating a set of registered depth maps with dramatically varying sampling rate. The method is based on the construction of a hierarchical signed distance field represented in an incomplete primal octree by incrementally adding triangulated depth maps. Due to the adaptive data structure, our algorithm is able to handle depth maps with varying scale and to consistently represent coarse, low-resolution regions as well as small details contained in high-resolution depth maps. A final surface mesh is extracted from the distance field by construction of a tetrahedral complex from the scattered signed distance values and applying the Marching Tetrahedra algorithm on the partition. The output is an adaptive triangle mesh that seamlessly connects coarse and highly detailed regions while avoiding filling areas without suitable input data.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Surface Reconstruction I.3.6 [Computer Graphics]: Methodology and Techniques—Graphics data structures and data types

Keywords: multi-scale depth map fusion, multi-view stereo depth maps, depth map integration, hierarchical signed distance field, surface reconstruction, marching tetrahedra

Links: DL PDF WEB VIDEO

ACM Reference Format
Fuhrmann, S., Goesele, M. 2011. Fusion of Depth Maps with Multiple Scales. *ACM Trans. Graph. 30*, 6, Article 148 (December 2011), 8 pages. DOI = 10.1145/2024156.2024182
http://doi.acm.org/10.1145/2024156.2024182.

1 Introduction

Surface reconstruction is an important problem with huge practical applications and a long history in computer graphics. The goal is to build high quality 3D surface representations from captured real-word data. Important applications include the preservation of cultural heritage, model reverse engineering, and prototyping in the multi-media industry. Typical inputs to surface reconstruction algorithms are either unorganized points or more structured data such as depth maps. In this work we will focus on the latter kind of data, which is produced by range scanners and some multi-view stereo algorithms. To fully capture an object of interest, multiple overlapping depth maps are necessary, each covering parts of the object surface. In a general acquisition framework, these depth maps need to be aligned into a common coordinate system and fused into a single, non-redundant surface representation. This process is called the *integration* or *fusion* of depth maps.

One source of depth maps are multi-view stereo (MVS) systems, which recently attained renewed interest [Seitz et al. 2006]. These algorithms reconstruct the scene geometry from photographs of the scene by regaining the 3D information lost during capture. Current structure-from-motion systems [Snavely et al. 2008] are able to recover the camera parameters of thousands of photographs under very uncontrolled conditions. This enables modern MVS algorithms to make use of the massive amount of Internet imagery for geometry reconstruction [Goesele et al. 2007; Agarwal et al. 2009; Frahm et al. 2010].

We desire to construct surface representations from the depth maps delivered by these acquisition systems, which is still an unsolved problem and difficult for various reasons. In particular, the photographs may be at different resolutions and show large variations in viewing parameters. The resulting depth maps inherit these properties and imply vastly different sampling rates of the surface. As in almost all acquisition processes, individual depth map samples are not ideal point samples. Instead, they represent the surface at a particular scale depending on viewing distance, focal length and image resolution. The extent of individual pixels when projected into 3D space can therefore dramatically vary in size. We call this the *pixel footprint*. The issue of scale and pixel footprints is crucial and requires particular care when mixing samples at different

Global Parametrization of Range Image Sets

Nico Pietroni[1] Marco Tarini[1,2] Olga Sorkine[3,4] Denis Zorin[4]
[1]ISTI, CNR, Italy [2]University of Insubria, Italy [3]ETH Zurich [4]New York University

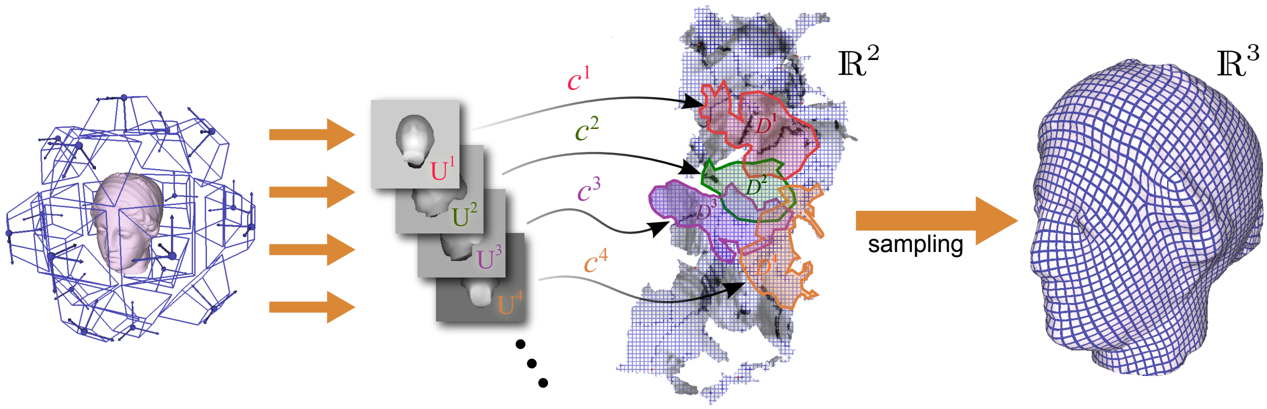


Figure 1: *A range image set and its parametrization. A set of range images $\mathcal{U}^i$ is captured by direct measurement of real world objects or by means of digital rendering of virtual objects. Each range scan $\mathcal{U}^i$ is mapped over a domain $\mathcal{D}^i$ by a parametrization function c^i, in a globally consistent way. Among other uses, a semi-regular quad remeshing of the original object can be obtained by regularly sampling $\cup\mathcal{D}^i$.*

Abstract

We present a method to globally parameterize a surface represented by height maps over a set of planes (range images). In contrast to other parametrization techniques, we do not start with a manifold mesh. The parametrization we compute defines a manifold structure, it is seamless and globally smooth, can be aligned to geometric features and shows good quality in terms of angle and area preservation, comparable to current parametrization techniques for meshes. Computing such global seamless parametrization makes it possible to perform quad remeshing, texture mapping and texture synthesis and many other types of geometry processing operations. Our approach is based on a formulation of the Poisson equation on a manifold structure defined for the surface by the range images. Construction of such global parametrization requires only a way to project surface data onto a set of planes, and can be applied directly to implicit surfaces, nonmanifold surfaces, very large meshes, and collections of range scans. We demonstrate application of our technique to all these geometry types.

CR Categories: I.3.5 [Computational Geometry and Object Modeling]: Geometric algorithms, languages, and systems

Links: DL PDF WEB

Keywords: geometry processing, parametrization, range scans

ACM Reference Format
Pietroni, N., Tarini, M., Sorkine, O., Zorin, D. 2011. Global Parametrization of Range Image Sets. *ACM Trans. Graph. 30*, 6, Article 149 (December 2011), 10 pages. DOI = 10.1145/2024156.2024183 http://doi.acm.org/10.1145/2024156.2024183.

1 Introduction

A high-quality global parametrization greatly simplifies many operations on surfaces. Recent techniques made substantial progress in improving the quality and robustness of global parametrization. At the same time, the work on parametrization, with few notable exceptions, focuses on manifold meshes, rather than on other forms of geometric data. In this paper, we describe how global parametrization techniques based on solving the Poisson equation (or another PDE) on the surface can be extended to a surface represented by a set of projections to planes. In some cases (e.g., range scanning) raw surface data is directly given in this format. In many other cases, it can be easily computed from a given arbitrary geometry representation: for example, if a geometry description can be rendered with depth values, it can serve as the input to our algorithm.

Range image sets occupy an intermediate place between point clouds or triangle soups, and manifold meshes. On one hand, they exhibit a regular connectivity and implicitly define a global manifold structure for the object, with transition maps determined by reprojection. On the other hand, each point on the surface may be represented by multiple positions inside different range images, and the connectivities of different range images, while highly regular, are inconsistent with each other.

Our method directly recovers a global parametrization from a range image set, entirely avoiding the need to construct a consistent manifold mesh first; this parametrization itself can be used to create high quality regular meshes. This considerably reduces the complexity of the meshing pipeline, replacing a more difficult step of manifold mesh reconstruction by simple and robust projections, followed directly by parametrization and quadrangulation.

Our method is based on a novel discretization of the *seamless* global parametrization equations and constraints on a collection of overlapping triangles, in contrast to conventional discretization on a single mesh. Our parametrization is globally consistent (images of a point in each range image are assigned the same parametric coordinates), seamless and globally smooth. It has comparable area and angle preservation quality to similar approaches for meshes, and can be aligned to geometric features.

Image-Based Bidirectional Scene Reprojection

Lei Yang[1] Yu-Chiu Tse[1] Pedro V. Sander[1] Jason Lawrence[2] Diego Nehab[3,4] Hugues Hoppe[3] Clara L. Wilkins[5]

[1]Hong Kong UST [2]University of Virginia [3]Microsoft Research [4]IMPA [5]Wesleyan University

Abstract

We introduce a method for increasing the framerate of real-time rendering applications. Whereas many existing temporal upsampling strategies only reuse information from previous frames, our bidirectional technique reconstructs intermediate frames from a pair of consecutive rendered frames. This significantly improves the accuracy and efficiency of data reuse since very few pixels are simultaneously occluded in both frames. We present two versions of this basic algorithm. The first is appropriate for fill-bound scenes as it limits the number of expensive shading calculations, but involves rasterization of scene geometry at each intermediate frame. The second version, our more significant contribution, reduces both shading and geometry computations by performing reprojection using only image-based buffers. It warps and combines the adjacent rendered frames using an efficient iterative search on their stored scene depth and flow. Bidirectional reprojection introduces a small amount of lag. We perform a user study to investigate this lag, and find that its effect is minor. We demonstrate substantial performance improvements (3–4×) for a variety of applications, including vertex-bound and fill-bound scenes, multi-pass effects, and motion blur.

Keywords: real-time rendering, temporal upsampling

Links: DL PDF WEB VIDEO

1 Introduction

Reprojection is a general approach for improving real-time rendering by reusing expensive pixel shading from nearby frames [Scherzer et al. 2011]. It has proven beneficial in popular games. For instance, reverse reprojection [e.g. Nehab et al. 2007; Scherzer et al. 2007] is used in *Gears of War II* to accelerate low-frequency lighting effects, and in *Crysis 2* to antialias distant geometry.

Such data reuse techniques can be broadly categorized based on three separate algorithmic choices:

- *Temporal direction:* Whether shading data is propagated forward, backward, or in both directions in animation time;
- *Data access:* Whether pixels are "pushed" (scattered) onto the current frame, or "pulled" (gathered) from other frames;
- *Correspondence domain:* Whether the motion data (e.g., velocity vectors) used to reproject the samples is defined over the source image domain or over the rendered target.

As reviewed in Section 2, different techniques follow different strategies for each of the choices above (see also Table 1).

In this paper, we present reprojection techniques for temporally upsampling rendered content by inserting interpolated frames between

Table 1: *Strategic algorithmic choices in reprojection techniques.*

	Temporal direction	*Data access*	*Correspondence domain*
e.g., Badt [1988]	Forward	Scatter	Source
e.g., Nehab et al. [2007]	Forward	Gather	Target
e.g., Didyk et al. [2010a]	Forward	Hybrid	Hybrid
Our scene-assisted scheme	**Both**	Gather	Target
Our image-based scheme	**Both**	Gather	**Source**

pairs of rendered frames. Borrowing terminology from video compression, we refer to rendered frames as *intra-* or simply *I-frames*, and to interpolated frames as *bidirectionally predicted-* or *B-frames*.

Our approach offers two major contributions: (1) *bidirectional reprojection* which combines samples both forward and backward in time, and (2) *image-based reprojection* which establishes reprojection correspondences based on velocity fields stored in the I-frames.

Temporal direction A fundamental limitation of existing reverse reprojection techniques [e.g. Nehab et al. 2007] is that they incur a drop in performance whenever there are disoccluded regions in the scene—elements visible in the current frame that were not visible in the preceding frame. This is because such regions must be reshaded from scratch. Since the number of disoccluded pixels varies over time, framerates may fluctuate undesirably. In addition, the entire scene geometry must be processed in order to reshade, incurring significant overhead in complex scenes.

Our *bidirectional* reprojection temporally upsamples rendered content by reusing data from *both* backward and forward temporal directions. This provides two clear benefits:

- *Smooth shading interpolation:* The vast majority of pixels in a B-frame are also visible in *both* I-frames. This lets us fetch shading information from both directions and create an interpolated signal that greatly attenuates the popping artifacts associated with one-sided reconstruction (i.e., sample-and-hold extrapolation). This is particularly important for fast changing shading signals (e.g., dynamic shadows and glossy lighting).
- *Higher, more stable framerate:* Disoccluded regions are extremely rare since they must be occluded in both I-frames. Thus with bidirectional reprojection we can avoid reshading and achieve higher and steadier framerates.

One downside of bidirectional reprojection is that it introduces a lag in the resulting image sequence. This lag is not present in forward-only reprojection schemes. We present a careful analysis of this lag, showing that it is small (less than one I-frame). Moreover, results of a user study we conducted allow us to conclude that it is beneficial to use bidirectional reprojection in a real-time gaming scenario.

Data access Some reprojection techniques use a forward-mapping strategy to scatter shading samples from prior frames into the new frame. However, scatter is difficult to realize efficiently in prevalent graphics systems. Accurate filtering of irregularly scattered samples is also challenging. Like recent reverse reprojection techniques, we use a gather strategy, which simply involves texture lookups into a previously rendered image. The gather operations thus benefit from the accurate, efficient texture-sampling hardware.

Correspondence domain Prior gather-based reprojection techniques require that geometry be rasterized in the target frame to establish a correspondence with source frame pixels. We develop an

ACM Reference Format
Yang, L., Tse, Y., Sander, P., Lawrence, J., Nehab, D., Hoppe, H., Wilkins, C. 2011. Image-Based Bidirectional Scene Reprojection. *ACM Trans. Graph. 30*, 6, Article 150 (December 2011), 10 pages. DOI = 10.1145/2024156.2024184 http://doi.acm.org/10.1145/2024156.2024184.

A Shading Reuse Method for Efficient Micropolygon Ray Tracing

Qiming Hou Kun Zhou
State Key Lab of CAD&CG, Zhejiang University*

Abstract

We present a shading reuse method for micropolygon ray tracing. Unlike previous shading reuse methods that require an explicit object-to-image space mapping for shading density estimation or shading accuracy, our method performs shading density control and actual shading reuse in different spaces with uncorrelated criterions. Specifically, we generate the shading points by shooting a user-controlled number of shading rays from the image space, while the evaluated shading values are assigned to antialiasing samples through object-space nearest neighbor searches. Shading samples are generated in separate layers corresponding to first bounce ray paths to reduce spurious reuse from very different ray paths. This method eliminates the necessity of an explicit object-to-image space mapping, enabling the elegant handling of ray tracing effects such as reflection and refraction. The overhead of our shading reuse operations is minimized by a highly parallel implementation on the GPU. Compared to the state-of-the-art micropolygon ray tracing algorithm, our method is able to reduce the required shading evaluations by an order of magnitude and achieve significant performance gains.

Keywords: micropolygon, GPU, Reyes, ray tracing

Links: DL PDF

Figure 1: *An animated scene rendered with motion blur and reflection. This scene contains 1.56M primitives and is rendered with 2 reflective bounces at* 1920×1080 *resolution and* 11×11 *supersampling. The total render time is 289 seconds. On average only 3.48 shading evaluations are performed for each pixel and are reused for other samples.*

1 Introduction

Shading is typically the performance bottleneck in cinematic-quality rendering, which is often based on the Reyes architecture and uses micropolygons to represent high order surfaces or highly detailed objects [Cook et al. 1987]. In order to reduce shading costs, state-of-the-art micropolygon renderers (e.g., Pixar's RenderMan) perform shading computation on micropolygon vertices, and reuse the shading values to evaluate the color of each visibility sample (or antialiasing sample) and composite the final image. Such a shading reuse strategy enables a shading rate significantly lower than the visibility sampling rate, which is vital for efficient high-quality rendering where extremely high supersampling of visibility is necessary, especially when rendering defocus and motion blur.

Existing shading reuse methods for micropolygon rendering are primarily designed for rasterization based pipelines. Ray tracing effects such as reflection and refraction are typically considered as a part of shading in such methods. Consequently, all reflected/refracted samples have to be shaded, incurring significant overhead. As ray tracing achieves more significance in modern high-quality rendering [Parker et al. 2010], this may become a major obstacle in future applications.

In this paper, we introduce a simple but effective method to reuse shading evaluations for efficient micropolygon ray tracing. Compared to the state-of-the-art micropolygon ray tracing algorithm, our method is able to reduce the required shading evaluations by an order of magnitude and achieve significant performance gains.

1.1 Related Work

Extensive research has been done on micropolygon rendering and ray tracing.

Researchers have explored efficient parallel implementations of micropolygon rendering on GPUs [Wexler et al. 2005; Patney and Owens 2008; Zhou et al. 2009; Hou et al. 2010]. In particular, Hou et al. [2010] introduced a GPU-based micropolygon ray tracing algorithm. They demonstrated that for high-quality defocus and motion blur ray tracing can greatly outperform rasterization methods. In their method, ray tracing is only used for visibility sampling and shading is still performed on micropolygon vertices. Another branch of research also seeks to accelerate micropolygon rendering using GPUs [Fisher et al. 2009; Fatahalian et al. 2009; Fatahalian et al. 2010; Ragan-Kelley et al. 2011; Burns et al. 2010]. The key difference between our work and theirs is that while they propose new GPU architecture designs that support real-time micropolygon rasterization, we aim to accelerate high-quality, off-line ray tracing using software approaches on current GPU hardware.

A majority of micropolygon rendering algorithms are designed to reuse the expensive shading computations across multiple visibility samples, assuming that shading is continuous and does not vary significantly across adjacent visibility samples. Existing shading reuse methods can be categorized into object-space methods [Cook et al. 1987; Burns et al. 2010] and image-space methods [Ragan-Kelley et al. 2011]. In [Cook et al. 1987], shading is performed on micropolygon vertices and reused within the same micropolygon. Burns et al. [2010] define a shading grid based on a priori shading density estimation, while shading values are evaluated lazily after visibility

*Email: hqm03ster@gmail.com, kunzhou@acm.org

ACM Reference Format
Hou, Q., Zhou, K. 2011. A Shading Reuse Method for Efficient Micropolygon Ray Tracing.
ACM Trans. Graph. 30, 6, Article 151 (December 2011), 7 pages. DOI = 10.1145/2024156.2024185
http://doi.acm.org/10.1145/2024156.2024185.

Genetic Programming for Shader Simplification

Pitchaya Sitthi-amorn Nicholas Modly Westley Weimer Jason Lawrence
University of Virginia

Abstract

We present a framework based on Genetic Programming (GP) for automatically simplifying procedural shaders. Our approach computes a series of increasingly simplified shaders that expose the inherent trade-off between speed and accuracy. Compared to existing automatic methods for pixel shader simplification [Olano et al. 2003; Pellacini 2005], our approach considers a wider space of code transformations and produces faster and more faithful results. We further demonstrate how our cost function can be rapidly evaluated using graphics hardware, which allows tens of thousands of shader variants to be considered during the optimization process. Our approach is also applicable to multi-pass shaders and perceptual-based error metrics.

Keywords: procedural texturing, pixel shaders, code simplification, genetic programming

Links: DL PDF WEB VIDEO

1 Introduction

The complexity of procedural shaders [Cook 1984; Perlin 1985] has continued to grow alongside the steady increase in performance and programmability of graphics hardware. Modern interactive rendering systems often contain hundreds of pixel shaders, each of which may perform thousands of arithmetic operations and texture fetches to generate a single frame.

Although this rise in complexity has brought considerable improvements to the realism of interactive 3D content, there is a growing need for automated tools to *optimize* procedural shaders to meet a computational budget or set of hardware constraints. For example, the popular virtual world Second Life allows users to supply custom models and textures, but not custom shaders: the performance of a potentially-complex custom shader cannot be guaranteed on all client hardware. Automatic optimization algorithms could adapt such shaders to multiple platform capabilities while retaining the intent of the original.

As with traditional computer programs, pixel shaders can be executed faster through a variety of semantics-preserving transformations like dead code elimination or constant folding [Muchnick 1997]. Unlike traditional programs, however, shaders also admit *lossy* optimizations [Olano et al. 2003]. A user will likely tolerate a shader that is incorrect in a minority of cases or that deviates from its ideal value by a small percentage in exchange for a significant performance boost.

One common way to achieve this type of optimization is through code simplification. The methods proposed by Olano et al. [2003] and Pellacini [2005] automatically generate a sequence of progressively simplified versions of an input pixel shader. These may be used in place of the original to improve performance at an acceptable reduction in detail. However, these methods have a number of important disadvantages. The system proposed by Olano et al. [2003] only considers code transformations that replace a texture with its average color. This overlooks many possible opportunities involving source-level modifications. The system proposed by Pellacini [2005] does consider source-level simplifications. However, that approach is limited to a small number of code transformations and uses a brute-force optimization strategy that can easily miss profitable areas of the objective function. Furthermore, both approaches were demonstrated only on shaders that require a single rendering pass. Modern shaders often involve multiple interdependent passes. Finally, in both of these systems, only a single error metric is used to evaluate the fidelity of a modified shader. It would be desirable for a simplification algorithm to support a range of error metrics, including those that are designed to predict perceptual differences [Wang et al. 2004].

Intuitively, we hypothesize that a shader contains the seeds of its own optimization: relevant functions such as **`sin`** and **`cos`**, operations such as `*` or `+`, or constants such as **`1.0`** are already present in the shader source code. We propose to produce optimized shader variants by copying, reordering and deleting the statements and expressions already available. We also hypothesize that the landscape of possible shader variants is sufficiently complex that a simple hill-climbing search will not suffice to avoid being trapped in local optima. We thus present a novel framework for simplifying shaders that is based on Genetic Programming (GP). GP is a computational method inspired by biological evolution which evolves computer programs tailored to a particular task [Koza 1992]. GP maintains a population of program variants, each of which is evaluated for suitability using a task-specific fitness function. High-fitness variants are selected for continued evolution. Computational analogs of biological crossover and mutation help to locate global optima by combining partial solutions and variations of the high-fitness programs; the process repeats until a fit program is located.

Our approach is related to recent software engineering work that applies GP to automatic program repair [Weimer et al. 2009, 2010]. Our approach is novel not only in the domain considered (continuous shader output vs. discrete software test cases) but also in the techniques used: to take advantage of the special structure of shader software, we apply new mutation operations, new approaches to select a diverse population of variants, new handling for multiple competing objectives, and optimizations to rapidly approximate the performance of a shader variant.

Our approach offers a number of benefits over existing methods for shader simplification. In particular, it explores optimizations beyond just texture lookups (cf. Olano et al. [2003]), is not limited to an *a priori* set of simplifying transformations (cf. Pellacini [2005]), does not require the user to specify continuous domains for shader input parameters (cf. Pellacini [2005]), is demonstrably applicable to multi-pass shaders and perceptual-based error metrics, and outperforms previous work in a direct comparison.

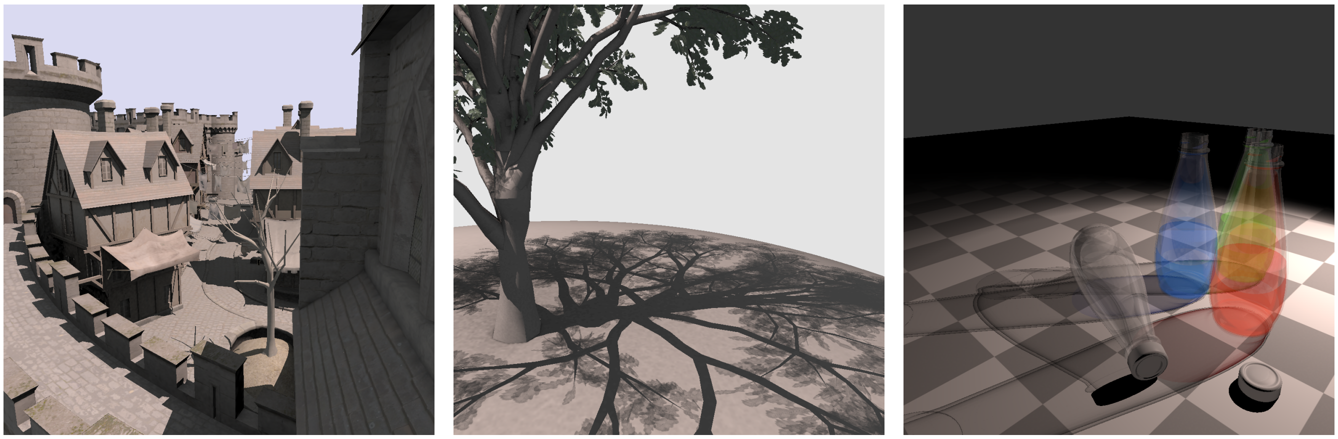
ACM Reference Format
Sitthi-amorn, P., Modly, N., Weimer, W., Lawrence, J. 2011. Genetic Programming for Shader Simplification. *ACM Trans. Graph. 30*, 6, Article 152 (December 2011), 11 pages. DOI = 10.1145/2024156.2024186 http://doi.acm.org/10.1145/2024156.2024186.

An Efficient Alias-free Shadow Algorithm for Opaque and Transparent Objects using per-triangle Shadow Volumes

Erik Sintorn*
Chalmers University Of Technology

Ola Olsson†
Chalmers University Of Technology

Ulf Assarsson‡
Chalmers University Of Technology



(a) (b) (c)

Figure 1: *Images rendered with the novel shadow algorithm. All images rendered in 1024x1024, time taken to generate shadow buffers in parenthesis. (a) Pixel accurate hard shadows in a game scene (7.29ms, 60k triangles). (b) Alpha-textured shadow casters (13ms, 35k triangles). (c) Colored transparent shadows. Image rendered using depth peeling of 8 layers (75.66ms, 5-19 ms per layer, 60k triangles).*

Abstract

This paper presents a novel method for generating pixel-accurate shadows from point light-sources in real-time. The new method is able to quickly cull pixels that are not in shadow and to trivially accept large chunks of pixels thanks mainly to using the whole triangle shadow volume as a primitive, instead of rendering the shadow quads independently as in the classic Shadow-Volume algorithm. Our CUDA implementation outperforms z-fail consistently and surpasses z-pass at high resolutions, although these latter two are hardware accelerated, while inheriting none of the robustness issues associated with these methods. Another, perhaps even more important property of our algorithm, is that it requires no pre-processing or identification of silhouette edges and so robustly and efficiently handles arbitrary triangle soups. In terms of view sample test and set operations performed, we show that our algorithm can be an order of magnitude more efficient than z-pass when rendering a game-scene at multi-sampled HD resolutions. We go on to show that the algorithm can be trivially modified to support textured, semi-transparent and colored semi-transparent shadow-casters and that it can be combined with either depth-peeling or stochastic transparency to also support transparent shadow receivers. Compared to recent alias-free shadow-map algorithms, our method has a very small memory footprint, does not suffer from load-balancing issues, and handles omni-directional lights without modification. It is easily incorporated into any deferred rendering pipeline and combines many of the strengths of shadow maps and shadow volumes.

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

Keywords: shadows, alias-free, real time, transparency

*e-mail: erik.sintorn@chalmers.se
†e-mail:ola.olsson@chalmers.se
‡e-mail:uffe@chalmers.se

1 Introduction

Generating accurate shadows from point light-sources for each pixel sample remains a challenging problem for real-time applications. Despite generations of research, we have yet to see a pixel-accurate shadow-algorithm for point lights that requires no pre-processing, works on any arbitrary set of triangles and that runs at stable real-time frame rates for typical game-scenes on consumer level hardware. Traditional shadow mapping [Williams 1978] techniques generate shadows from a discretized image representation of the scene and so alias when queried for light visibility in screen space. Real-time techniques based on irregular rasterization [Sintorn et al. 2008] tend to generate unbalanced workloads that fit current GPUs poorly and consequently, frame rates are often very unstable. Real-time ray tracing algorithms rely heavily on geometry pre-processing to generate efficient acceleration structures. Finally, robust implementations of the Shadow-Volume algorithm require pre-processing the mesh to find edge connectivity, work poorly or not at all for polygon soups without connectivity, and have frame rates that are all but stable as the view of a complex scene changes. Nevertheless, the idea of directly rasterizing the volumes that represent shadows onto the view samples remains compelling. In the

ACM Reference Format
Sintorn, E., Olsson, O., Assarsson, U. 2011. An Efficient Alias-free Shadow Algorithm for Opaque and Transparent Objects using per-triangle Shadow Volumes. *ACM Trans. Graph. 30*, 6, Article 153 (December 2011), 10 pages. DOI = 10.1145/2024156.2024187 http://doi.acm.org/10.1145/2024156.2024187.

Data-driven Visual Similarity for Cross-domain Image Matching

Abhinav Shrivastava
Carnegie Mellon University

Tomasz Malisiewicz
MIT

Abhinav Gupta
Carnegie Mellon University

Alexei A. Efros
Carnegie Mellon University

Figure 1: *In this paper, we are interested in defining visual similarity between images across different domains, such as photos taken in different seasons, paintings, sketches, etc. What makes this challenging is that the visual content is only similar on the higher scene level, but quite dissimilar on the pixel level. Here we present an approach that works well across different visual domains.*

Abstract

The goal of this work is to find *visually similar* images even if they appear quite different at the raw pixel level. This task is particularly important for matching images across visual domains, such as photos taken over different seasons or lighting conditions, paintings, hand-drawn sketches, etc. We propose a surprisingly simple method that estimates the relative importance of different features in a query image based on the notion of "data-driven uniqueness". We employ standard tools from discriminative object detection in a novel way, yielding a generic approach that does not depend on a particular image representation or a specific visual domain. Our approach shows good performance on a number of difficult cross-domain visual tasks e.g., matching paintings or sketches to real photographs. The method also allows us to demonstrate novel applications such as *Internet re-photography*, and `painting2gps`. While at present the technique is too computationally intensive to be practical for interactive image retrieval, we hope that some of the ideas will eventually become applicable to that domain as well.

CR Categories: I.2.10 [Artificial Intelligence]: Vision and Scene Understanding—Learning; I.4.10 [Image Processing and Computer Vision]: Image Representation—Statistical;

Keywords: image matching, visual similarity, saliency, image retrieval, paintings, sketches, re-photography, visual memex

Links: DL PDF WEB

1 Introduction

Powered by the availability of Internet-scale image and video collections coupled with greater processing speeds, the last decade has witnessed the rise of data-driven approaches in computer graphics and computational photography. Unlike traditional methods, which employ parametric models to capture visual phenomena, the data-driven approaches use visual data directly, without an explicit intermediate representation. These approaches have shown promising results on a wide range of challenging computer graphics problems, including super-resolution and de-noising [Freeman et al. 2002; Buades et al. 2005; HaCohen et al. 2010], texture and video synthesis [Efros and Freeman 2001; Schodl et al. 2000], image analogies [Hertzmann et al. 2001], automatic colorization [Torralba et al. 2008], scene and video completion [Wexler et al. ; Hays and Efros 2007; Whyte et al. 2009], photo restoration [Dale et al. 2009], assembling photo-realistic virtual spaces [Kaneva et al. 2010; Chen et al. 2009], and even making CG imagery more realistic [Johnson et al. 2010], to give but a few examples.

The central element common to all the above approaches is searching a large dataset to find visually similar matches to a given query – be it an image patch, a full image, or a spatio-temporal block. However, defining a good visual similarity metric to use for matching can often be surprisingly difficult. Granted, in many situations where the data is reasonably homogeneous (e.g., different patches within the same texture image [Efros and Freeman 2001], or different frames within the same video [Schodl et al. 2000]), a simple pixel-wise sum-of-squared-differences (L2) matching works quite well. But what about the cases when the visual content is only similar on the higher scene level, but quite dissimilar on the pixel level? For instance, methods that use scene matching e.g., [Hays and Efros 2007; Dale et al. 2009] often need to match images across different illuminations, different seasons, different cameras, etc. Likewise, retexturing an image in the style of a painting [Hertzmann et al. 2001; Efros and Freeman 2001] requires making visual correspondence between two very different domains – photos and paintings. Cross-domain matching is even more critical for applications such as Sketch2Photo [Chen et al. 2009] and CG2Real [Johnson et al. 2010], which aim to bring domains as different as sketches and CG renderings into correspondence with natural photographs. In all of these cases, pixel-wise matching fares quite poorly, because small perceptual differences can result in arbitrarily large pixel-wise differences. What is needed is a visual metric that can capture the important visual structures that make two images appear similar, yet show robustness to small, unimportant visual details. This is precisely what makes this problem so difficult – the visual similarity algorithm somehow needs to know which visual structures are important for a human observer and which are not.

Currently, the way researchers address this problem is by using var-

ACM Reference Format
Shrivastava, A., Malisiewicz, T., Gupta, A., Efros, A. 2011. Data-driven Visual Similarity for Cross-domain Image Matching. *ACM Trans. Graph. 30*, 6, Article 154 (December 2011), 9 pages.
DOI = 10.1145/2024156.2024188 http://doi.acm.org/10.1145/2024156.2024188.

Arcimboldo-like Collage Using Internet Images

Hua Huang* Lei Zhang Hong-Chao Zhang
Xi'an Jiaotong University, China

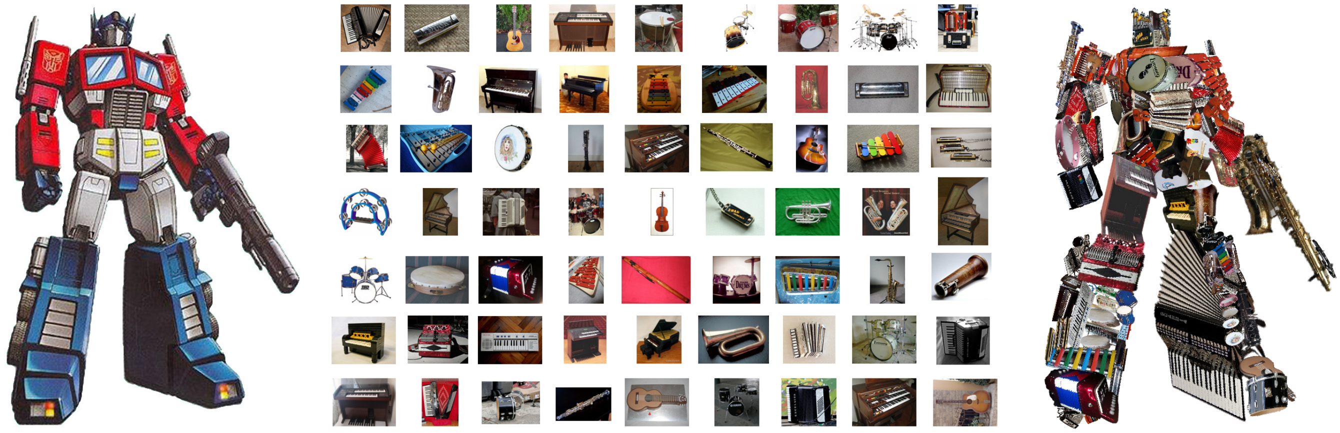


Figure 1: *From an input images (left, Transformers ©2011 Hasbro), our algorithm creates the fascinating Arcimboldo-like collage (right) by using a collection of Internet images (middle). The Internet images are retrieved with the key words "musical instrument".*

Abstract

Collage is a composite artwork made from assemblage of different material forms. In this work, we present a novel approach for creating a fantastic collage artform, namely Arcimboldo-like collage, which represents an input image with multiple thematically-related cutouts from the filtered Internet images. Due to the massive data of Internet images, competent image cutouts can almost always be discovered to match the segmented components of the input image. The selected cutouts are purposefully arranged such that as a whole assembly, they can represent the input image with disguise in both shape and color; but separately, individual cutout is still recognizable as its own being. Experimental results and user study show that our algorithm can effectively produce the entertaining Arcimboldo-like collages.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithm; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques; I.4.6 [Image Processing and Computer Vision]: Segmentation—Pixel classification

Keywords: Arcimboldo collage, internet image, segmentation, layer ordering

Links: DL PDF

*e-mail: huanghua@mail.xjtu.edu.cn

ACM Reference Format
Huang, H., Zhang, L., Zhang, H. 2011. Arcimboldo-like Collage Using Internet Images.
ACM Trans. Graph. 30, 6, Article 155 (December 2011), 7 pages. DOI = 10.1145/2024156.2024189
http://doi.acm.org/10.1145/2024156.2024189.

1 Introduction

"Our body is overcoat of the soul inside" — Bhagwad Gita. It has long been discussed by psychologists and artists about essence and its physical representation. The essence can still be recognizable even with marked changes on its appearance, while conveys more information in a compelling way (see Figure 1). This principle is artistically realized in the painting artworks by Giuseppe Arcimboldo, an Italian painter in the Renaissance and honored as the vanguard of *Surrealism*. In his paintings, portrait heads are represented by a variety of symbolic elements like fruits, books (see Figure 2). Although this sort of "hodgepodge" images looks rather eccentric, it provides a fascinating form to display the content in an extravagant yet lucid style [Maiorino 1983], which is also fashionable nowadays. This paper focuses on the creation of such collages with the Arcimboldo-like style.

From their origin dating back to the year 200 BC in China, collages have gained growing popularity as artistic and collective expression of photo assemblage [Wikipedia 2011]. However, manually creating collage is labor intensive and time consuming, which needs delicate cutouts as materials and reciprocity on their assembly. Thus, many approaches work on automating collage construction [Rother et al. 2006; Wang et al. 2006; Goferman et al. 2010]. These approaches often produce commendable collages, but confine themselves in a regular canvas, which disable the Arcimboldo-like effect as shown in Figure 2. In [Gal et al. 2007], 3D models instead of image cutouts are assembled to approximate a 3D shape. Although exhibiting the Arcimboldo-like style, such collage needs a 3D model as the template, which prohibits stylization of various 2D images without 3D representation.

The visual mechanism by which we recognize the Arcimboldo-like collage is the so-called *Apophenia*, which suggests the experienced meaningfulness coming from specific connections of coherent representation [Maiorino 1983]. Hence, there are two challenges to create effective Arcimboldo-like collages: selecting competent image cutouts and assembling them consistently for resemblance. The Internet provides large easy-to-access image database

Semantic Colorization with Internet Images

Alex Yong-Sang Chia[1] Shaojie Zhuo[2] Raj Kumar Gupta[3] Yu-Wing Tai[4]
Siu-Yeung Cho[3] Ping Tan[2] Stephen Lin[5]
[1]Institute for Infocomm Research [2]National University of Singapore [3]Nanyang Technological University
[4]Korea Advanced Institute of Science and Technology [5]Microsoft Research Asia

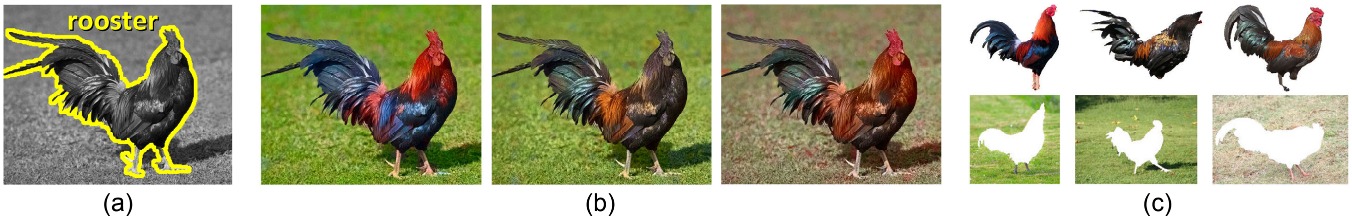


Figure 1: *Our system takes a grayscale photo with labeled and segmented foreground objects (a) as input and generates a set of colorization results (b) using reference image regions automatically searched from the internet and filtered to obtain the most suitable examples. Some of these examples are shown in (c).*

Abstract

Colorization of a grayscale photograph often requires considerable effort from the user, either by placing numerous color scribbles over the image to initialize a color propagation algorithm, or by looking for a suitable reference image from which color information can be transferred. Even with this user supplied data, colorized images may appear unnatural as a result of limited user skill or inaccurate transfer of colors. To address these problems, we propose a colorization system that leverages the rich image content on the internet. As input, the user needs only to provide a semantic text label and segmentation cues for major foreground objects in the scene. With this information, images are downloaded from photo sharing websites and filtered to obtain suitable reference images that are reliable for color transfer to the given grayscale photo. Different image colorizations are generated from the various reference images, and a graphical user interface is provided to easily select the desired result. Our experiments and user study demonstrate the greater effectiveness of this system in comparison to previous techniques.

Links: DL PDF

1 Introduction

Image colorization can bring a grayscale photo to life, but often demands extensive user interaction. In techniques such as [Levin et al. 2004; Huang et al. 2005], a user typically needs to specify many color scribbles on the image to achieve a desirable result. Moreover, it can be difficult for a novice user to provide these color scribbles in a consistent and perceptually coherent manner. Other methods take a different approach by using a color image of a similar scene as a reference, and transferring its colors to the grayscale input image [Reinhard et al. 2001; Welsh et al. 2002]. This requires less skill from the user, but a suitable reference image may take much effort to find. In addition, inaccuracies in color transfer can lead to results that appear unnatural.

To colorize grayscale photos with less manual labor, we present a system that takes advantage of the tremendous amount of image data available on the internet. The internet is almost certain to contain images suitable for colorizing a given grayscale input, but finding those images in a sea of photos is a challenging task, especially since search engines often return images with incompatible content. We address this problem with a novel image filtering method that analyzes spatial distributions of local and regional image features to identify candidate reference regions most compatible with the grayscale target. The user needs only to provide semantic labels and segmentation cues for major foreground objects in the grayscale image, which is more intuitive than previous scribble based user interaction. For each foreground object, a multitude of images is downloaded from the internet using the semantic label as a search term, and our system filters them down to a small number of best matches. To minimize the amount of user input, our system does not require the user to label and segment background regions. Rather, it exploits correlations between the foregrounds and backgrounds of scenes by re-using the images downloaded for the foreground objects, which likely contain some backgrounds that can serve as a reference for background colorization.

From the filtered reference images, the system transfers colors to the corresponding foreground objects and background with a graph-based optimization based on local properties at a super-pixel resolution. Since the filtering method seeks reference objects with spatial distributions of features most consistent with the target object, color transfer becomes more reliable, as corresponding locations between the reference and target can be identified more accurately. Various colorization results are generated from the set of reference images, and the user is provided an intuitive interface to rapidly explore the results and select the most preferred colorization.

2 Related Work

Colorization methods can be roughly divided into those based on user drawn scribbles and those that utilize example images. Scribble based methods propagate the colors from an initial set of user drawn strokes to the whole image. For example, Levin et al. [2004]

ACM Reference Format
Chia, A., Zhuo, S., Gupta, R., Tai, Y., Cho, S., Tan, P., Lin, S. 2011. Semantic Colorization with Internet Images. *ACM Trans. Graph. 30*, 6, Article 156 (December 2011), 7 pages. DOI = 10.1145/2024156.2024190 http://doi.acm.org/10.1145/2024156.2024190.

Rendering Synthetic Objects into Legacy Photographs

Kevin Karsch   Varsha Hedau   David Forsyth   Derek Hoiem

University of Illinois at Urbana-Champaign
{karsch1,vhedau2,daf,dhoiem}@uiuc.edu

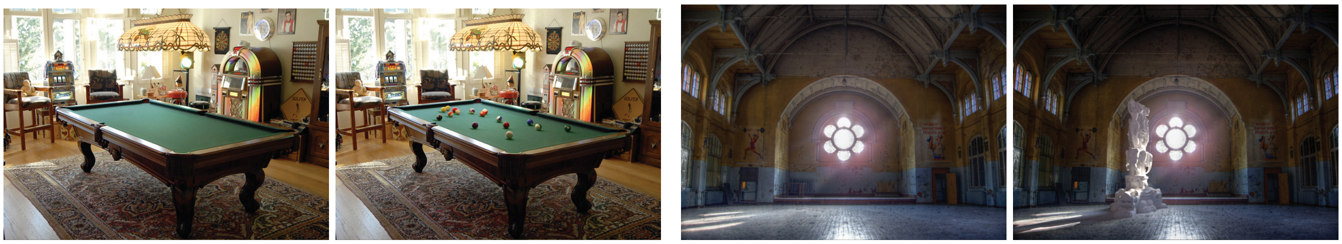


Figure 1: *With only a small amount of user interaction, our system allows objects to be inserted into legacy images so that perspective, occlusion, and lighting of inserted objects adhere to the physical properties of the scene. Our method works with only a single LDR photograph, and no access to the scene is required.*

Abstract

We propose a method to realistically insert synthetic objects into existing photographs without requiring access to the scene or any additional scene measurements. With a single image and a small amount of annotation, our method creates a physical model of the scene that is suitable for realistically rendering synthetic objects with diffuse, specular, and even glowing materials while accounting for lighting interactions between the objects and the scene. We demonstrate in a user study that synthetic images produced by our method are confusable with real scenes, even for people who believe they are good at telling the difference. Further, our study shows that our method is competitive with other insertion methods while requiring less scene information. We also collected new illumination and reflectance datasets; renderings produced by our system compare well to ground truth. Our system has applications in the movie and gaming industry, as well as home decorating and user content creation, among others.

CR Categories: I.2.10 [Computing Methodologies]: Artificial Intelligence—Vision and Scene Understanding; I.3.6 [Computing Methodologies]: Computer Graphics—Methodology and Techniques

Keywords: image-based rendering, computational photography, light estimation, photo editing

Links: DL PDF WEB

ACM Reference Format
Karsch, K., Hedau, V., Forsyth, D., Hoiem, D. 2011. Rendering Synthetic Objects into Legacy Photographs. *ACM Trans. Graph. 30*, 6, Article 157 (December 2011), 12 pages. DOI = 10.1145/2024156.2024191 http://doi.acm.org/10.1145/2024156.2024191.

1 Introduction

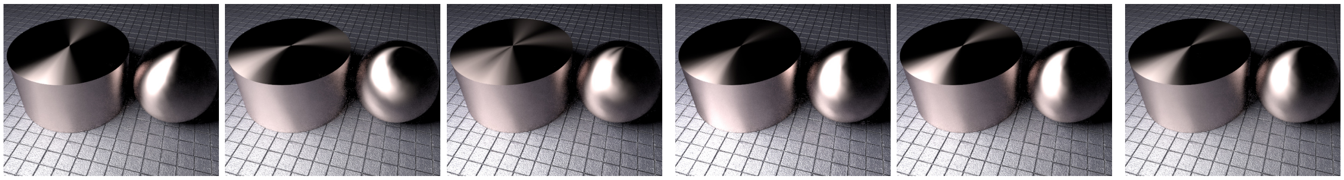
Many applications require a user to insert 3D meshed characters, props, or other synthetic objects into images and videos. Currently, to insert objects into the scene, some scene geometry must be manually created, and lighting models may be produced by photographing mirrored light probes placed in the scene, taking multiple photographs of the scene, or even modeling the sources manually. Either way, the process is painstaking and requires expertise.

We propose a method to realistically insert synthetic objects into existing photographs without requiring access to the scene, special equipment, multiple photographs, time lapses, or any other aids. Our approach, outlined in Figure 2, is to take advantage of small amounts of annotation to recover a simplistic model of geometry and the position, shape, and intensity of light sources. First, we automatically estimate a rough geometric model of the scene, and ask the user to specify (through image space annotations) any additional geometry that synthetic objects should interact with. Next, the user annotates light sources and light shafts (strongly directed light) in the image. Our system automatically generates a physical model of the scene using these annotations. The models created by our method are suitable for realistically rendering synthetic objects with diffuse, specular, and even glowing materials while accounting for lighting interactions between the objects and the scene.

In addition to our overall system, our primary technical contribution is a semiautomatic algorithm for estimating a physical lighting model from a single image. Our method can generate a full lighting model that is demonstrated to be physically meaningful through a ground truth evaluation. We also introduce a novel image decomposition algorithm that uses geometry to improve lightness estimates, and we show in another evaluation to be state-of-the-art for single image reflectance estimation. We demonstrate with a user study that the results of our method are confusable with real scenes, even for people who believe they are good at telling the difference. Our study also shows that our method is competitive with other insertion methods while requiring less scene information. This method has become possible from advances in recent literature. In the past few years, we have learned a great deal about extracting high level information from indoor scenes [Hedau et al. 2009; Lee et al. 2009; Lee et al. 2010], and that detecting shadows in images is relatively straightforward [Guo et al. 2011]. Grosse *et al.* [2009] have also shown that simple lightness assumptions lead to powerful surface estimation algorithms; Retinex remains among the best methods.

Displacement Interpolation Using Lagrangian Mass Transport

Nicolas Bonneel[1,3] Michiel van de Panne[1] Sylvain Paris[2] Wolfgang Heidrich[1]
[1] University of British Columbia [2]Adobe Systems, Inc. [3]ALICE/INRIA Nancy



(a) BRDF A (anisotropic angle = $0°$) (b) BRDF B (anisotropic angle = $45°$) (c) linear interpolation (unrealistic double highlight; error= 40.5%) (d) displacement interpolation in Euclidean space (error= 36.6%) (e) ground truth (single highlight at $22.5°$; error= 0%) (f) our displacement interpolation on a spherical space (error= 24.3%)

Figure 1: *Examples of interpolation between two anisotropic sampled BRDFs (a,b). Naive linear interpolation generates an unrealistic double highlight (c). Working with a flattened BRDF representation in a Euclidean space produces large conspicuous errors (d). Since we used a parametric BRDF model, we can compute the ground-truth in-between BRDF by interpolating the parameters (e). Our approach has no knowledge of the parametric BRDF representation but is nonetheless able to produce a similar output (f).*

Abstract

Interpolation between pairs of values, typically vectors, is a fundamental operation in many computer graphics applications. In some cases simple linear interpolation yields meaningful results without requiring domain knowledge. However, interpolation between pairs of distributions or pairs of functions often demands more care because features may exhibit translational motion between exemplars. This property is not captured by linear interpolation. This paper develops the use of *displacement interpolation* for this class of problem, which provides a generic method for interpolating between distributions or functions based on advection instead of blending. The functions can be non-uniformly sampled, high-dimensional, and defined on non-Euclidean manifolds, e.g., spheres and tori. Our method decomposes distributions or functions into sums of radial basis functions (RBFs). We solve a mass transport problem to pair the RBFs and apply partial transport to obtain the interpolated function. We describe practical methods for computing the RBF decomposition and solving the transport problem. We demonstrate the interpolation approach on synthetic examples, BRDFs, color distributions, environment maps, stipple patterns, and value functions.

CR Categories: I.3.7 [Computing Methodologies]: Computer Graphics—Three-Dimensional Graphics and Realism, G.1.1 [Mathematics of Computing]: Numerical Analysis—Interpolation

Keywords: displacement interpolation, mass transport

Links: DL PDF WEB VIDEO

ACM Reference Format
Bonneel, N., van de Panne, M., Paris, S., Heidrich, W. 2011. Displacement Interpolation Using Lagrangian Mass Transport. *ACM Trans. Graph. 30*, 6, Article 158 (December 2011), 11 pages.
DOI = 10.1145/2024156.2024192 http://doi.acm.org/10.1145/2024156.2024192.

1 Introduction

Interpolation is a fundamental operation in many mathematical models and algorithms. Linear interpolation between a pair of values is particularly simple. However, linear interpolation does not always achieve meaningful results. In particular, this is often the case for probability distributions. Consider two Gaussian distributions, $\mathcal{N}_A(\mu_A, \sigma_A)$ and $\mathcal{N}_B(\mu_B, \sigma_B)$, that model height in two different populations of humans, group A with mostly short people and group B with mostly tall people. There are two plausible answers as to what an intermediate interpolated distribution might look like, as shown in Figure 2. A first answer is to assume an equal mix of people from both groups, which therefore results in the bimodal height distribution shown in Fig.2 (b). This *mixture model* interpretation is the answer that arises from linear interpolation of the two initial distributions at every point in their domain. An alternative answer can be obtained by interpolating the underlying parameters of the distributions, yielding the distribution $\mathcal{N}_C(\mu_C, \sigma_C)$, where $\mu_C = 0.5\mu_A + 0.5\mu_B$ and $\sigma_C^2 = 0.5\sigma_A^2 + 0.5\sigma_B^2$. Intuitively, this interpretation says that interpolating between the height distribution of a group of tall people and that of a group of short people should result in a distribution of medium-height people.

The need for something other than the mixture model interpretation is in fact rather common. For example, the mean of two bidirectional reflectance distribution functions (BRDFs), which each have a single large reflectance lobe but each pointing in a different direction is, subjectively speaking, better defined as having a single large reflectance lobe in a halfway direction, rather than two half-sized lobes in the original directions. The difference between these two interpretations is shown in Figure 1 for a BRDF example where we also know the underlying parameterization.

The *displacement interpolation* methods developed in this paper provide an efficient way to achieve the type of *advection-based* or *motion-based* interpolation that characterizes the above examples and a wide range of applications. Importantly, it does this in a *generic way* without requiring access to a domain-specific underlying parameterization. Displacement interpolation casts interpolation as a mass transport problem, wherein each unit of 'mass' of one distribution needs to move to the second distribution. The goal is to find a minimal total cost solution, where the cost for each unit

Adaptive Sampling and Reconstruction using Greedy Error Minimization

Fabrice Rousselle*
University of Bern

Claude Knaus †
University of Bern

Matthias Zwicker‡
University of Bern

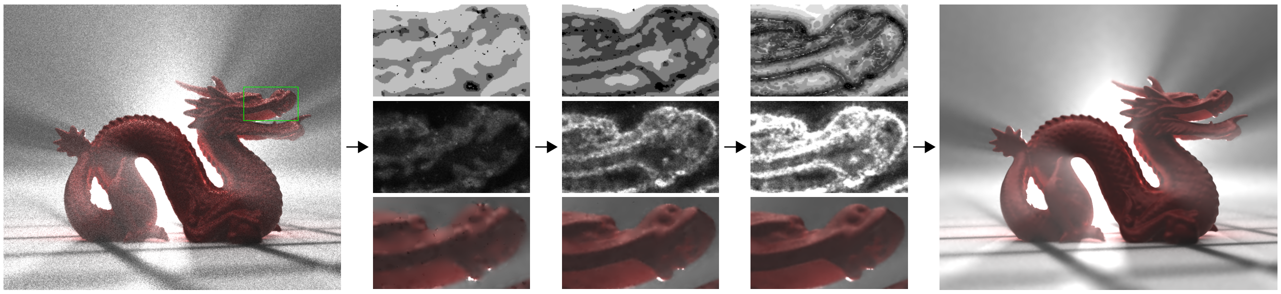


Figure 1: *We minimize* MSE *in Monte Carlo rendering by adaptive sampling and reconstruction in image space. We iterate over two steps: given current samples, optimize over a set of filters at each pixel to minimize* MSE*; then, given a filter at each pixel, distribute more samples to further reduce* MSE*. Left: initialization with* 4 *samples per pixel. Insets: each column is one iteration (top to bottom): filter selection (smooth filters shaded white, sharp ones black), sample density map, reconstruction. Right: result at an average of* 32 *samples per pixel. This image features single scattering participating media, indirect illumination using photon mapping, depth of field, and area lighting.*

Abstract

We introduce a novel approach for image space adaptive sampling and reconstruction in Monte Carlo rendering. We greedily minimize relative mean squared error (MSE) by iterating over two steps. First, given a current sample distribution, we optimize over a discrete set of filters at each pixel and select the filter that minimizes the pixel error. Next, given the current filter selection, we distribute additional samples to further reduce MSE. The success of our approach hinges on a robust technique to select suitable per pixel filters. We develop a novel filter selection procedure that robustly solves this problem even with noisy input data. We evaluate our approach using effects such as motion blur, depth of field, interreflections, etc. We provide a comparison to a state-of-the-art algorithm based on wavelet shrinkage and show that we achieve significant improvements in numerical error and visual image quality. Our approach is simple to implement, requires a single user parameter, and is compatible with standard Monte Carlo rendering.

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
†e-mail:knaus@iam.unibe.ch
‡e-mail:zwicker@iam.unibe.ch

ACM Reference Format
Rousselle, F., Knaus, C., Zwicker, M. 2011. Adaptive Sampling and Reconstruction using Greedy Error Minimization. *ACM Trans. Graph. 30*, 6, Article 159 (December 2011), 11 pages.
DOI = 10.1145/2024156.2024193 http://doi.acm.org/10.1145/2024156.2024193.

1 Introduction

Monte Carlo techniques compute pixel colors by (quasi-)randomly sampling an integration domain that covers all light paths transporting light from a source to the camera. The integration domain may include effects such as depth of field, motion blur, and light paths with multiple interreflections. Unless one computes an excessive number of samples, this often leads to high pixel variance and the typical noise artifacts in Monte Carlo rendering. There are two main strategies to address this. The first is to distribute samples in an optimal fashion, with respect to the problem at hand. The second is to smooth out noise by applying suitable filters. Both strategies can be applied in the high dimensional space of light paths or in the image plane. We focus on strategies that operate in the image plane.

We formulate the problem as follows: given a certain budget of Monte Carlo samples, obtain an image that minimizes the relative mean squared error (MSE) by distributing samples in a suitable fashion in the image plane and by filtering the image with appropriate filters. We can interpret this as an optimization problem over the space of sample distributions and image filters. Our core idea is to make the problem tractable by restricting the space of filters to a discrete set of predetermined filters per pixel. Each pixel may have a different set of filters, but the set is predefined and not itself part of the optimization. We use a simple greedy strategy to obtain an approximate solution to the MSE minimization problem. Starting from an initial set of samples, we iterate over two steps. First, for each pixel we select the filter from its discrete set that minimizes the pixel MSE given the current samples. Second, given the currently chosen pixel filters, we distribute a new batch of samples that try to further reduce MSE as much as possible. This process is repeated until a termination criterion is fulfilled.

To minimize pixel MSE we express it as the sum of the squared bias, i.e., expected error, and variance. We define the set of filters at each pixel such that it provides a trade-off between reducing bias and increasing variance. Then we attempt to minimize pixel MSE by selecting the filter that offers an optimal compromise. The main challenge in practice is that we only have access to noisy data to estimate bias and variance. Therefore, an important component of our algorithm is a robust method to solve this filter selection problem.

T&I Engine: Traversal and Intersection Engine for Hardware Accelerated Ray Tracing

Jae-Ho Nah* Jeong-Soo Park* Chanmin Park† Jin-Woo Kim* Yun-Hye Jung* Woo-Chan Park‡ Tack-Don Han*
*Yonsei University, Korea †Samsung Electronics, Korea ‡Sejong University, Korea

Abstract

Ray tracing naturally supports high-quality global illumination effects, but it is computationally costly. Traversal and intersection operations dominate the computation of ray tracing. To accelerate these two operations, we propose a hardware architecture integrating three novel approaches. First, we present an ordered depth-first layout and a traversal architecture using this layout to reduce the required memory bandwidth. Second, we propose a three-phase ray-triangle intersection architecture that takes advantage of early exit. Third, we propose a latency hiding architecture defined as the ray accumulation unit. Cycle-accurate simulation results indicate our architecture can achieve interactive distributed ray tracing.

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

Keywords: ray tracing, ray tracing hardware, global illumination

Links: DL PDF

1 Introduction

Ray tracing [Whitted 1980; Cook et al. 1984] is the most commonly-used algorithm for photorealistic rendering. Ray tracing generates a more realistic image than does rasterization, but it requires tremendous computational power for traversal and ray-primitive intersections. For this reason, it has been used for offline rendering for most of the last decade.

For real-time ray tracing, many approaches utilizing CPUs, GPUs, or custom hardware have recently been studied. These approaches do not yet provide sufficient performance for processing 1G rays/s for real-time distributed ray tracing [Govindaraju et al. 2008].

Most performance bottlenecks in ray tracing are in traversal and intersection tests [Benthin 2006]. Traversal is the process of searching an acceleration structure (AS), such as a kd-tree or bounding volume hierarchy (BVH), to find a small subset of the primitives for testing by the ray. A ray-primitive intersection test determines the visibility of primitives found during the traversal.

We believe a dedicated hardware unit for traversal and the intersection test is a suitable solution for real-time distributed ray tracing. In this paper, we present a custom hardware architecture, called T&I (traversal and intersection) engine. This architecture can be integrated with existing programmable shaders, as with raster operations pipelines (ROPs) or texture mapping units. Also, it comprises three novel approaches that are applicable to the traversal and intersection test processes.

First, an ordered depth-first layout (ODFL) and its traversal architecture are presented. The ODFL is the enhancement of an eight-byte kd-tree node layout [Pharr and Humphreys 2010]. It arranges the child node, which has a larger surface area than its sibling, adjacent to its parent to improve parent-child locality. We apply this layout to our traversal architecture to effectively reduce the miss rate of the traversal cache. The ODFL also can be easily applied to other CPU or GPU ray tracers. This concept was previously announced in the extended abstract [Nah et al. 2010].

Second, we propose a three-phase intersection test unit, which divides the intersection test stage into three phases. Phase 1 is the ray-plane test, Phase 2 is the barycentric coordinate test, and Phase 3 is the final hit point calculation. This configuration reduces the need for further computation and memory requests for missed triangles that are identified in either Phases 1 or 2. Phases 1 and 2 are performed in a common module because they use roughly the same arithmetic operations.

Third, a ray accumulation unit is proposed for hiding memory latency. This unit manages memory requests and accumulates rays that induce a cache miss. While the waiting missed block is fetched, other rays can perform their operations. When the missed block is fetched, the accumulated rays are flushed to the pipeline.

We verify the performance of our architecture with a cycle-accurate simulator and evaluate resource requirements and performance. We also perform a simulation with three types of rays that have different coherence. The proposed architecture achieves 44-1188 Mrays/s ray tracing performance at 500 MHz on 65 nm process.

The remainder of this paper is structured as follows. Section 2 describes related work. Section 3 gives an overview of the proposed architecture. In Sections 4 to 6, we cover the details of our three approaches (a traversal unit with the ODFL, a three-phase intersection test unit, and a ray accumulation unit). In Section 7, we describe the experimental results of the proposed architecture simulation. Finally, we conclude the paper in Section 8.

2 Related Work

2.1 Dedicated ray tracing hardware

SaarCOR [Schmittler et al. 2004] is a ray tracing pipeline that consists of a ray generation/shading unit, a 4-wide SIMD traversal unit, a list unit, a transformation unit, and an intersection test unit. Woop et al. [2005] presented the programmable RPU architecture, which performs ray generation, shading, and intersection tests with programmable shaders. For dynamic scenes, D-RPU [Woop et al. 2006a; Woop 2007] has a node update unit [Woop et al. 2006b] unlike RPU. RTE [Davidovic et al. 2011] is an optimized version of D-RPU that uses tail recursive shaders with treelets.

SaarCOR, RPU, D-RPU, and RTE are based on packet tracing [Wald et al. 2001], but packet tracing may result in low SIMD efficiency with incoherent rays as described in [Gribble and Ramani

ACM Reference Format
Nah, J., Park, J., Park, C., Kim, J., Jung, Y., Park, W., Han, T. 2011. T & I Engine: Traversal and Intersection Engine for Hardware Accelerated Ray Tracing. *ACM Trans. Graph. 30*, 6, Article 160 (December 2011), 10 pages. DOI = 10.1145/2024156.2024194 http://doi.acm.org/10.1145/2024156.2024194.

Coherent Parallel Hashing

Ismael García[1,2] Sylvain Lefebvre[2] Samuel Hornus[2] Anass Lasram[2]

[1]GGG / Universitat de Girona [2]ALICE / INRIA

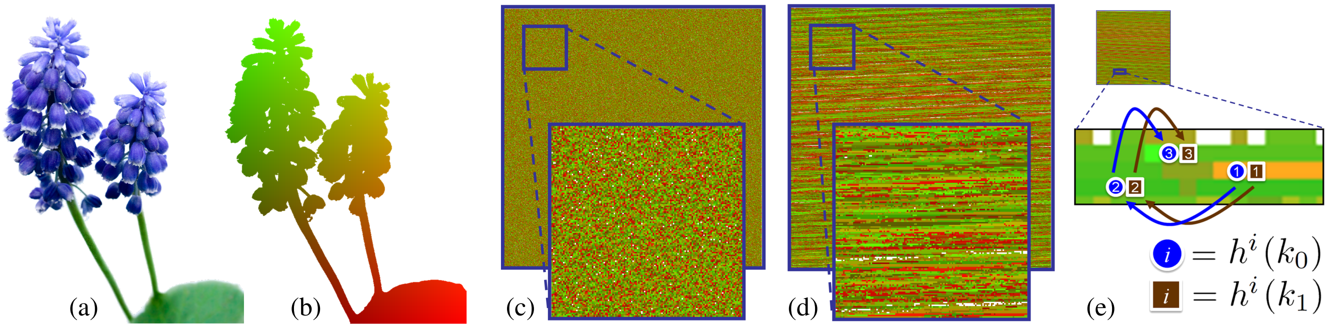


Figure 1: (a) *The flower image is* 3820×3820 *image (*14.5 *million pixels) and contains* 3.7 *million non–white pixels. The coordinates of these pixels are shown as colors in* (b)*. We store the image in a hash table under a* 0.99 *load factor: the hash table contains only* 3.73 *million entries. These are used as keys for hashing.* (c) *The table obtained with a typical randomizing hash function: Keys are randomly spread and all coherence is lost.* (d) *Our spatially coherent hash table, built in parallel on the GPU. The table is built in* 15 *ms on a GeForce GTX 480, and the image is reconstructed from the hash in* 3.5 *ms. The visible structures are due to preserved coherence. This translates to faster access as neighboring threads perform similar operations and access nearby memory.* (e) *Neighboring keys are kept together during probing, thereby improving the coherence of memory accesses of neighboring threads.*

Abstract

Recent spatial hashing schemes hash millions of keys in parallel, compacting sparse spatial data in small hash tables while still allowing for fast access from the GPU. Unfortunately, available schemes suffer from two drawbacks: Multiple runs of the construction process are often required before success, and the random nature of the hash functions decreases access performance.

We introduce a new parallel hashing scheme which reaches high load factor with a very low failure rate. In addition our scheme has the unique advantage to exploit coherence in the data and the access patterns for faster performance. Compared to existing approaches, it exhibits much greater locality of memory accesses and consistent execution paths within groups of threads. This is especially well suited to Computer Graphics applications, where spatial coherence is common. In absence of coherence our scheme performs similarly to previous methods, but does not suffer from construction failures.

Our scheme is based on the Robin Hood scheme modified to quickly abort queries of keys that are not in the table, and to preserve coherence. We demonstrate our scheme on a variety of data sets. We analyze construction and access performance, as well as cache and threads behavior.

ACM Reference Format
García, I., Lefebvre, S., Hornus, S., Lasram, A. 2011. Coherent Parallel Hashing.
ACM Trans. Graph. 30, 6, Article 161 (December 2011), 8 pages. DOI = 10.1145/2024156.2024195
http://doi.acm.org/10.1145/2024156.2024195.

CR Categories: I.3.6 [Computer Graphics]: Methodology and Techniques—Graphics data structures and data types;

Keywords: spatial, parallel, coherent, hashing, sparse data

Links: DL PDF

1 Introduction

Sparse spatial data is very common in Computer Graphics and finding the good tradeoff between access performance and efficient storage is an ongoing challenge. Spatial hashing has proven useful in these situations, enabling data to be tightly packed while still allowing fast random access. It has been successfully applied for texturing, rendering, collision detection and animation.

Using a spatial hash, data is stored in a single array—the *hash table*—addressed through a *hash function*. The hash function computes the data location in the hash table from the query coordinates, or *keys*. There have been several developments lately, improving query and construction times, and in particular enabling fast parallel construction on GPUs.

The first spatial hashing schemes focused essentially on reaching good load factors while having a constant time and simple access to the data from the GPU. Lefebvre and Hoppe [2006] proposed a static hash construction enabling access to the data with as little as two memory accesses and one addition. However, to achieve this result the hash has to be *perfect*: All keys corresponding to defined data should map to different locations in the hash table. In other words, there are no *collisions*. Building such a constrained hash function requires an off–line construction process, limiting this approach to static cases.

Alcantara *et al.* [2009; 2011] propose less constrained hashing schemes that achieve fast, parallel construction on the GPU. These schemes may produce collisions. However, querying a key never requires more than four independent memory accesses. The par-

Artist Friendly Facial Animation Retargeting

Yeongho Seol* Jaewoo Seo† Paul Hyunjin Kim‡ KAIST

J.P. Lewis§ Weta Digital

Junyong Noh¶ KAIST

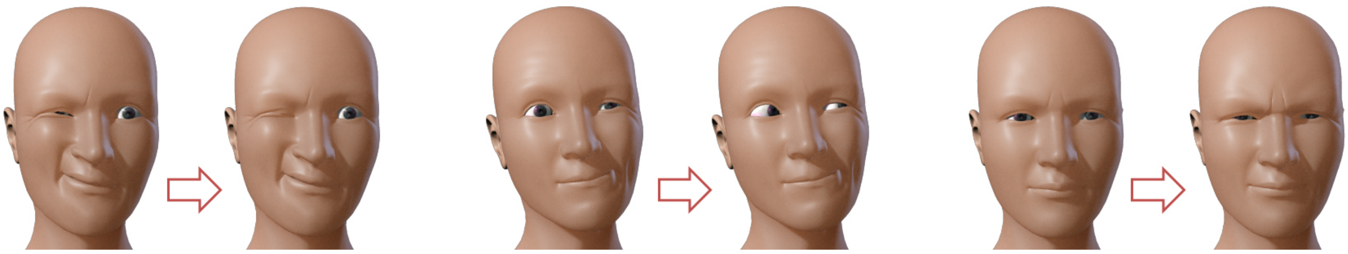


Figure 1: *Examples of the initially retargeted facial animation (left of each pair) and final animation (right) after an animator's editing is applied. By design, our retargeting method seamlessly integrates with the animator's editing process.*

Abstract

This paper presents a novel facial animation retargeting system that is carefully designed to support the animator's workflow. Observation and analysis of the animators' often preferred process of key-frame animation with blendshape models informed our research. Our retargeting system generates a similar set of blendshape weights to those that would have been produced by an animator. This is achieved by rearranging the group of blendshapes into several sequential retargeting groups and solving using a matching pursuit-like scheme inspired by a traditional key-framing approach. Meanwhile, animators typically spend a tremendous amount of time simplifying the dense weight graphs created by the retargeting. Our graph simplification technique effectively produces editable weight graphs while preserving the visual characteristics of the original retargeting. Finally, we automatically create GUI controllers to help artists perform key-framing and editing very efficiently. The set of proposed techniques greatly reduce the time and effort required by animators to achieve high quality retargeted facial animations.

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

Keywords: Animation, Face, Retargeting, Animator, Editing

Links: DL PDF

*e-mail: seolyeongho@kaist.ac.kr
†e-mail: goongsang@kaist.ac.kr
‡e-mail: salbur@kaist.ac.kr
§e-mail: noisebrain@gmail.com
¶e-mail: junyongnoh@kaist.ac.kr

ACM Reference Format
Seol, Y., Seo, J., Kim, P., Lewis, J., Noh, J. 2011. Artist Friendly Facial Animation Retargeting. *ACM Trans. Graph. 30*, 6, Article 162 (December 2011), 10 pages. DOI = 10.1145/2024156.2024196 http://doi.acm.org/10.1145/2024156.2024196.

1 Introduction

Creating facial animation that is both realistic and emotionally compelling continues to be a challenge. Increasingly, the generation of realistic facial animations starts with retargeting results from motion capture. This approach has been used for recent high-quality character animations such as the ones in the films *King Kong*, *The Curious Case of Benjamin Button*, and *Avatar*. However, these are extremely labor intensive efforts. Although motion capture and retargeting are sometimes naively assumed to be automated processes, as a ballpark figure current movie projects can require up to a day of manual editing for each second of finished retargeted animation.

Most current approaches do not produce facial animation retargeting results that are directly usable for high quality applications. Even in a relatively simple case of identical source and target faces, the animator's participation is inevitably needed to achieve a high quality animation as evidenced by *The Digital Emily Project* [Alexander et al. 2009]. Havaldar [2006] described the reasons why manual editing is always necessary:

(a) The combination of artistically designed blendshapes cannot perfectly match the actor's motion.

(b) The proportions of CG face model and an actor's face can be significantly different.

(c) Motion capture marker placements differ from day to day.

(d) The desired performance is not what the actor performed. Either the required expression is not present in the motion capture data or it needs to be exaggerated.

Animators frequently comment that, even if the retargeting results are "90%" correct, fixing the remaining "10%" requires almost as much work as starting from scratch [Luamanuvae 2011]. Considering the successful research outcomes reported by previous methods [Chuang and Bregler 2005; Weise et al. 2009; Li et al. 2010], the amount of manual editing in current practice seems hard to justify. To exactly identify the source of the difficulties in editing, we performed a simple experiment in which two skilled animators (each with five years of experience) edit a given retargeting result produced by a retargeting method based on Non-Negative Least Squares (NNLS) similar to Joshi et al. [2003] and Chuang and Bregler [2005]. Then, they were asked to produce the same animation

Controlling Physics-Based Characters Using Soft Contacts

Sumit Jain C. Karen Liu

Georgia Institute of Technology*

Abstract

In this paper, we investigate the impact of the deformable bodies on the control algorithms for physically simulated characters. We hypothesize that ignoring the effect of deformable bodies at the site of contact negatively affects the control algorithms, leading to less robust and unnatural character motions. To verify the hypothesis, we introduce a compact representation for an articulated character with deformable soft tissue and develop a practical system to simulate two-way coupling between rigid and deformable bodies in a robust and efficient manner. We then apply a few simple and widely used control algorithms, such as pose-space tracking control, Cartesian-space tracking control, and a biped controller (SIMBICON), to simulate a variety of behaviors for both full-body locomotion and hand manipulation. We conduct a series of experiments to compare our results with the motion generated by these algorithms on a character comprising only rigid bodies. The evaluation shows that the character with soft contact can withstand larger perturbations in a more noisy environment, as well as produce more realistic motion.

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

Keywords: character animation, articulated rigid body, deformable body, linear complementarity problem

Links: DL PDF

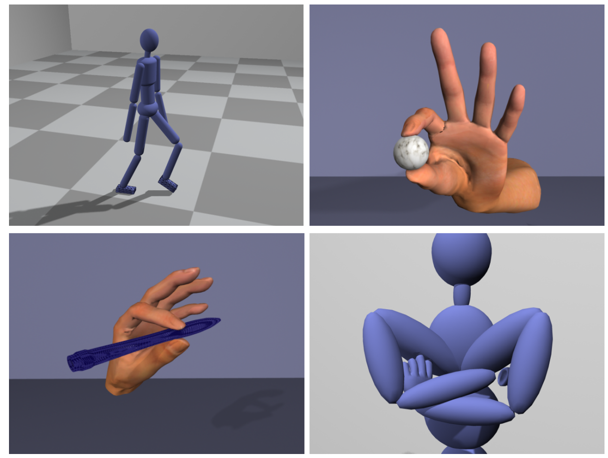


Figure 1: *Various controllers for character animation can be improved by simulating soft tissue deformation at the site of contact.*

1 Introduction

One of the fundamental simplifications that researchers in physics-based human motion synthesis make, is that motion is the product of an articulated rigid body system with actuated joints representing bones and active skeletal muscles. On the surface this abstraction does capture the most fundamental aspects of the human musculoskeletal system. Utilizing this assumption, researchers have developed several control algorithms that can synthesize movement for various tasks like balance and walking. Although these controllers work well in their specific problem domain, they still cannot achieve the same level of agility the human body displays.

In this paper we revisit the fundamental assumption that an articulated rigid body system, by itself, captures the fundamental properties that enable human-like motion. We focus on one aspect of the motion that is not captured by this simplified model: the contact with the environment primarily occurs through the soft tissue. This factor comes into play in any situation where there is a collision between the character and the environment. Collisions between rigid bodies usually result in sporadic contact points and highly discontinuous pressure distribution. Although a character consist of only rigid bodies is ideal for efficiently simulating human movement, we postulate that the simplified rigid contact model inadvertently increases the difficulty in controller design and results in unrealistic motion.

The primary contribution of this paper is to demonstrate that simple control strategies coupled with the simulation of soft tissue deformation at the site of contact can achieve very robust and realistic motion. We develop a practical system that allows us to simulate two-way coupling between rigid and deformable bodies in a robust and efficient manner. We then apply a few simple and widely used control algorithms, such as pose-space tracking control, Cartesian-space tracking control, and SIMBICON, to simulate a variety of both full-body locomotion and hand manipulation. The resulting motions are compared with the motion generated by these algorithms on a character comprising only of rigid bodies. These simple controllers demonstrate that the character with soft contacts can withstand larger perturbations in a more noisy environment, without the need of designing more sophisticated control algorithms.

Simulating deformable bodies can be achieved in a few different ways and the design choice often has to balance the required accuracy and performance. We hypothesize that the accuracy offered by sophisticated but expensive methods, such as Finite Element Method (FEM) is unnecessary for our application for two reasons. First, unlike most previous work that simulates deformation of complex volumetric meshes for aesthetic purpose, the primary goal of our work is to produce deformation for more physically correct contacts. Second, average human body deforms marginally due to the support of bones. In particular, the deformation due to contacts is typically small and localized. We take advantage of these properties to design a simple and accurate model that only computes the surface of deformable bodies, rather than the entire volume.

To this end, we introduce a new representation for human skele-

*email:{sumit,karenliu}@cc.gatech.edu

ACM Reference Format
Jain, S., Liu, C. 2011. Controlling Physics-Based Characters Using Soft Contacts. *ACM Trans. Graph. 30*, 6, Article 163 (December 2011), 10 pages. DOI = 10.1145/2024156.2024197 http://doi.acm.org/10.1145/2024156.2024197.

Compression and Direct Manipulation of Complex Blendshape Models

Jaewoo Seo* KAIST
Geoffrey Irving† Weta Digital
J. P. Lewis‡ Weta Digital
Junyong Noh§ KAIST

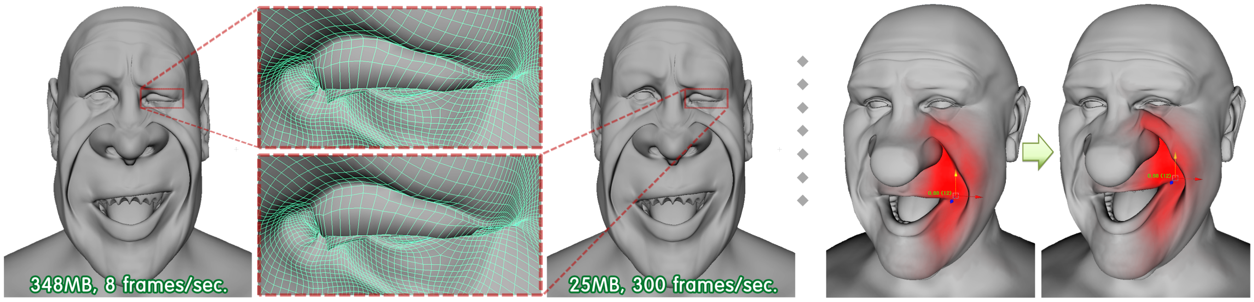


Figure 1: *Left: This facial model has 348 MB of uncompressed blendshape data, and runs at 8 frames per second on 8 CPUs. Our compression method reduces the storage to 25.4 MB and achieves 300 frames per second with a GPU implementation. No artifacts are visible. Right: The user can interactively manipulate the blendshape puppet by dragging any vertex on the model. The deforming region is colored red for visual feedback.*

Abstract

We present a method to compress complex blendshape models and thereby enable interactive, hardware-accelerated animation of these models. Facial blendshape models in production are typically large in terms of both the resolution of the model and the number of target shapes. They are represented by a single huge blendshape matrix, whose size presents a storage burden and prevents real-time processing. To address this problem, we present a new matrix compression scheme based on a hierarchically semi-separable (HSS) representation with matrix block reordering. The compressed data are also suitable for parallel processing. An efficient GPU implementation provides very fast feedback of the resulting animation. Compared with the original data, our technique leads to a huge improvement in both storage and processing efficiency without incurring any visual artifacts. As an application, we introduce an extended version of the direct manipulation method to control a large number of facial blendshapes efficiently and intuitively.

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

Keywords: blendshape, compression, direct manipulation

Links: DL PDF

*e-mail: goongsang@kaist.ac.kr
†e-mail: irving@naml.us
‡e-mail: zilla@computer.org
§e-mail: junyongnoh@kaist.ac.kr

ACM Reference Format
Seo, J., Irving, G., Lewis, J., Noh, J. 2011. Compression and Direct Manipulation of Complex Blendshape Models. *ACM Trans. Graph. 30*, 6, Article 164 (December 2011), 10 pages.
DOI = 10.1145/2024156.2024198 http://doi.acm.org/10.1145/2024156.2024198.

1 Introduction

Blendshapes are a predominant technique for the creation of realistic and semi-realistic facial animation. The strength of the blendshape method lies in its simple and interpretable parameterization: the artist defines an ‘expression space’ with shapes such as raise-inner-brow-right, and then models each shape exactly as desired. Blendshape deformers are available in most commercial graphics packages such as Autodesk Maya and 3ds Max, and virtual characters created with this technique abound in recent popular films.

The blendshape formulation represents the face as a linear combination of a set of predefined shapes (i.e. morph targets),

$$o = \hat{B}w$$

where o is a vector containing the resulting vertex positions and w is a vector storing the blending weights. $\hat{B}$, the *blendshape matrix*, has the vertex positions of target shapes at its columns. In practice, the delta blendshape formulation is often utilized. In this form the targets B are offsets from a neutral shape vector n:

$$o = n + Bw \tag{1}$$

Typically, each target shape deforms only some part of the face, so this approach exposes sparsity in the matrix B. Our blendshape model is based on this formula.

One downside of the blendshape method is that it requires a large number of target shapes to produce high quality facial animation. For example, the *Gollum* character in *the Lord of the Rings* movies had over 900 target shapes [Raitt 2004]. Contemporary film-resolution models, such as the examples in this paper, generally require more than 1000 target shapes. With a high-resolution facial mesh, the result is a large blendshape matrix B with at least tens of thousands of rows and hundreds of columns, requiring hundreds of megabytes of memory. Although the processing involves only a matrix-vector multiplication, this size is too large to achieve real-time performance on current hardware. For example, after incorporating a blendshape model similar to that in Figure 1 into a realistic “rig” with additional deformers (e.g. for the neck), a coarse body model, textures, another character, and an environment, the resulting scene plays at less than 1 frame per second.

Stretchable and Twistable Bones for Skeletal Shape Deformation

Alec Jacobson Olga Sorkine
New York University & ETH Zurich

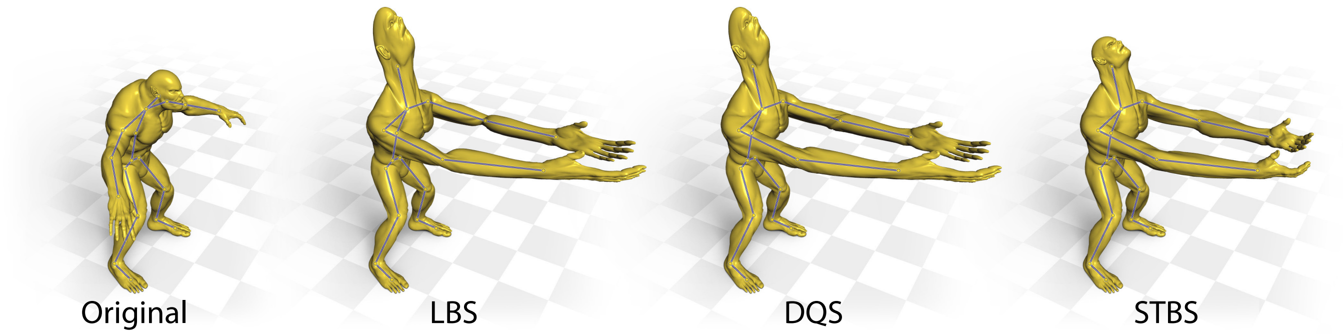


Figure 1: *Left to right: the* Beast *model is rigged to a skeleton in its rest pose. The neck is stretched and the arms are twisted and stretched using linear blend skinning. LBS relies solely on per-bone scalar weight functions, resulting in the explosion of the head and hands. The candy-wrapper artifact of LBS is also noticeable at the elbows. The dual quaternion skinning (DQS) solution [Kavan et al. 2008] correctly blends rotations, avoiding the candy-wrapper artifact, but reliance on bone weights alone unnaturally concentrates the twisting near the elbows. DQS also does not alleviate the stretching artifacts. Our solution, stretchable, twistable bones skinning (STBS), uses an extra set of weights per bone, allowing stretching without explosions and smooth twisting along the entire length of each arm.*

Abstract

Skeleton-based linear blend skinning (LBS) remains the most popular method for real-time character deformation and animation. The key to its success is its simple implementation and fast execution. However, in addition to the well-studied elbow-collapse and candy-wrapper artifacts, the space of deformations possible with LBS is inherently limited. In particular, blending with only a scalar weight function per bone prohibits properly handling stretching, where bones change length, and twisting, where the shape rotates along the length of the bone. We present a simple modification of the LBS formulation that enables stretching and twisting without changing the existing skeleton rig or bone weights. Our method needs only an extra scalar weight function per bone, which can be painted manually or computed automatically. The resulting formulation significantly enriches the space of possible deformations while only increasing storage and computation costs by constant factors.

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

Keywords: shape deformation, articulated character animation, linear blend skinning

Links: DL PDF WEB VIDEO

ACM Reference Format
Jacobson, A., Sorkine, O. 2011. Stretchable and Twistable Bones for Skeletal Shape Deformation.
ACM Trans. Graph. 30, 6, Article 165 (December 2011), 7 pages. DOI = 10.1145/2024156.2024199
http://doi.acm.org/10.1145/2024156.2024199.

1 Introduction

Skinning and skeletal deformation remain standard for character animation because the associated deformation metaphor is directly intuitive for many situations: most characters are creatures or humans who ought to behave as if a skeleton was moving underneath their skin. The motion capture pipeline, for example, explicitly relies on this metaphor to build a subspace representation of human motion [Anguelov et al. 2005]. At the cost of performance, some applications demand physical accuracy, ensuring preservation of volume or simulating muscles [Teran et al. 2005]. Other applications, such as video games, crowd simulation and interactive animation editing, cannot afford to compromise real-time performance, so they trade accuracy for speed and often adopt the simplest and most efficient implementation of skeletal deformation.

The long-standing standard real-time skeletal deformation method is linear blend skinning (LBS), also known as skeletal subspace deformation or enveloping [Magnenat-Thalmann et al. 1988; Lewis et al. 2000]. In a typical workflow, a trained rigging artist manually constructs and fits a skeleton of rigid bones within the target shape. The skeleton is bound to the shape by assigning a set of correspondence weights for each bone, a process which can be tedious and labor-intensive. To deform the shape, animators assign transformations to each skeleton bone, either directly or with the assistance of an inverse kinematics engine or motion capture data. These transformations are propagated to the shape by blending them linearly as matrix operations according to the bone weights.

Linearly blending matrix transformations with scalar weight functions has a number of limitations. Many improvements of LBS focus on the problems arising from linearly blending rotations as matrices, which results in shape collapses near joints. Multi-weight enveloping (MWE) [Wang and Phillips 2002; Merry et al. 2006] and Dual Quaternions [Kavan et al. 2008] have been proposed as alternative rotation blending methods. However, a different set of limitations arises from the fact that using only a single scalar weight function per bone limits the space of possible deformations. We show that neither LBS nor its improvements properly handle stretching, where bones change length, nor twisting, where the skin

Freeform Vector Graphics with Controlled Thin-Plate Splines

Mark Finch John Snyder Hugues Hoppe
Microsoft Research

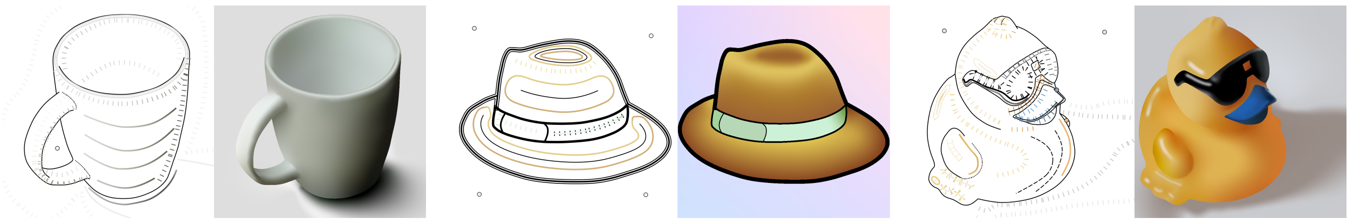


Figure 1: *We build on thin-plate splines to enrich vector graphics with a variety of powerful and intuitive controls.*

Abstract

Recent work defines vector graphics using diffusion between colored curves. We explore higher-order fairing to enable more natural interpolation and greater expressive control. Specifically, we build on thin-plate splines which provide smoothness everywhere except at user-specified tears and creases (discontinuities in value and derivative respectively). Our system lets a user sketch discontinuity curves without fixing their colors, and sprinkle color constraints at sparse interior points to obtain smooth interpolation subject to the outlines. We refine the representation with novel contour and slope curves, which anisotropically constrain interpolation derivatives. Compound curves further increase editing power by expanding a single curve into multiple offsets of various basic types (value, tear, crease, slope, and contour). The vector constraints are discretized over an image grid, and satisfied using a hierarchical solver. We demonstrate interactive authoring on a desktop CPU.

Keywords: bilaplacian/biharmonic PDE, slope/contour curves

Links: DL PDF WEB VIDEO DATA CODE

1 Introduction

Traditional vector graphics fills each closed shape independently with a simple color function. Recent work applies more global and powerful Laplacian interpolation between diffusion curves with colors on each side [Orzan et al. 2008; Jeschke et al. 2009].

A Laplacian solution yields a *membrane* function which is "as-constant-as-possible". Its low-order smoothness objective has drawbacks as illustrated in Figure 2. The solution is smooth only away from constrained points. Value constraints yield tent-like responses at isolated points and form creases along curves. The Laplacian objective is also incompatible with derivative constraints, because it already seeks zero first-derivatives everywhere in all directions. Only a higher-order notion of smoothness supports sparse constraints on directional derivatives.

ACM Reference Format
Finch, M., Snyder, J., Hoppe, H. 2011. Freeform Vector Graphics with Controlled Thin-Plate Splines. *ACM Trans. Graph. 30*, 6, Article 166 (December 2011), 10 pages. DOI = 10.1145/2024156.2024200 http://doi.acm.org/10.1145/2024156.2024200.

Our approach builds on *thin-plate splines* (TPS) [Courant and Hilbert 1953], which define a higher-order interpolating function that is "as-harmonic-as-possible". This smoothness objective overcomes previous limitations (Figure 2). Thin-plate splines have been applied in several areas including geometric modeling [e.g. Welch and Witkin 1992; Botsch and Kobbelt 2004; Sorkine and Cohen-Or 2004; Botsch and Sorkine 2008], computer vision [Terzopoulos 1983], and machine learning [Bookstein 1989]. They have also been adapted to allow discontinuity control with explicit tears and creases [Terzopoulos 1988]. We extend these controls and demonstrate their usefulness in vector graphics authoring.

In the simplest case, an artist sketches some outlines (tears) without fixing their colors, and specifies color constraints at a few interior points or curves to obtain a smooth color wash within the outlines. This ink-and-paint ordering of tasks is similar to hand drawing. The result is then refined by adding creases, contour curves, slope curves, and critical points. These features increase editing power by anisotropically constraining interpolation derivatives (e.g. along or across the curves, or in both directions).

In addition to the basic curves, we introduce *compound* curves, with user-assigned widths, for more complex effects. These internally yield several offset curves, of possibly different types. For instance, a value-slope curve juxtaposes the two basic types to create a smooth ridge-like feature. A wide contour uses two offset contours to form a constant-colored strip without fixing its color. Other combinations produce a variety of interesting and useful results.

We demonstrate a prototype system based on these ideas. Like other variational approaches such as diffusion curves, our system is easy to use and supports "freeform" input based on a general network of curves, which we augment with points. Smoother interpolation and more flexible constraints enhance naturalness and editing power and produce rich results from a compact input (Figure 1).

Our contributions include:

- Extension of the diffusion curves framework to benefit from higher-order interpolation and general discontinuity control.
- A discretized least-squares kernel for accurate modeling of crease curves.
- Contour and slope curves that constrain derivatives anisotropically for intuitive control.
- A variety of compound curve types for added expressiveness.
- Discontinuity-aware upsampling for improved accuracy in a multiresolution setting.

Multiscale Vector Volumes

Lvdi Wang
Tsinghua University

Yizhou Yu
University of Illinois at Urbana-Champaign
The University of Hong Kong

Kun Zhou
Zhejiang University

Baining Guo
Microsoft Research Asia
Tsinghua University

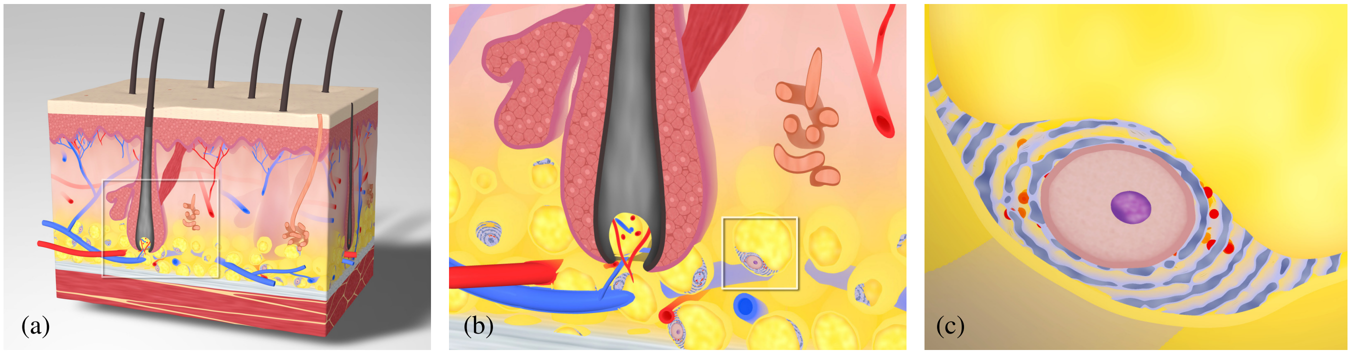


Figure 1: *A multiscale vector volume depicting a portion of human skin viewed at different scales. (b) Zoomed view of the selected area in (a). (c) Zoomed view of the selected area in (b). This model was created from scratch with our representation and consumes 6.8 MB storage.*

Abstract

We introduce *multiscale vector volumes*, a compact vector representation for volumetric objects with complex internal structures spanning a wide range of scales. With our representation, an object is decomposed into components and each component is modeled as an *SDF tree*, a novel data structure that uses multiple signed distance functions (SDFs) to further decompose the volumetric component into regions. Multiple signed distance functions collectively can represent non-manifold surfaces and deliver a powerful vector representation for complex volumetric features. We use *multiscale embedding* to combine object components at different scales into one complex volumetric object. As a result, regions with dramatically different scales and complexities can co-exist in an object. To facilitate volumetric object authoring and editing, we have also developed a scripting language and a GUI prototype. With the help of a recursively defined spatial indexing structure, our vector representation supports fast random access, and arbitrary cross sections of complex volumetric objects can be visualized in real time.

Keywords: volumetric modeling, multiscale representations

Links: DL PDF WEB

1 Introduction

Most natural organisms and materials, such as the human body, fruits and sedimentary rocks, have rich and complex volumetric structures, patterns and color variations. Constructing high-quality digital models for such natural organisms and materials is of vital importance because they exist everywhere, and literally everything we see is either a living being, a natural material or something made from these two. Since volumetric models allow us to visualize the internal structure of an object, they are valuable graphical contents that can be used in biomedical research, scientific visualization as well as educational and training activities.

Nevertheless, constructing high-quality digital models of natural organisms and materials with complex volumetric properties is an extremely challenging task. *First*, how can we compactly represent volumetric structures and patterns spanning a wide range of scales? Taking the human body as an example, it has a skeleton, organs and tissues at the macroscopic scale, cellular structures and neuronal fibers at an intermediate scale as well as proteins and genes at the molecular scale. *Second*, how can we represent high-frequency features in a resolution-independent way? Physical and appearance properties (e.g. color) change abruptly across different materials or volumetric components. Such discontinuities typically form thin surface sheets, which may join or split following an irregular pattern, giving rise to a non-manifold structure (Figure 2). To prevent visual artifacts when zooming into these high-frequency structures, a resolution-independent vector representation is desired. A volumetric object representation not only needs to depict complex multiscale structures, but also should be easy to use, which means it should be easy to create novel objects and easy to view existing objects in this representation. Thus, a volumetric object representation should have the following desired properties:

- *Expressiveness*: It should be able to represent volumetric objects with spatial structures spanning a wide range of scales and including complex non-manifold features. High-frequency features should remain sharp during magnification.
- *Ease of editing*: It should be easy and intuitive to create novel objects and edit existing ones using this representation.
- *Random access*: To be able to provide a timely response to user interactions, fast visualization is required, which further demands efficient random access to the volumetric content.
- *Compactness*: Given the limited memory on current graphics cards, the representation should be as compact as possible.

ACM Reference Format
Wang, L., Yu, Y., Zhou, K., Guo, B. 2011. Multiscale Vector Volumes. *ACM Trans. Graph. 30*, 6, Article 167 (December 2011), 8 pages. DOI = 10.1145/2024156.2024201 http://doi.acm.org/10.1145/2024156.2024201.

Slices: A Shape-proxy Based on Planar Sections

James McCrae   Karan Singh
University of Toronto

Niloy J. Mitra
KAUST/UCL

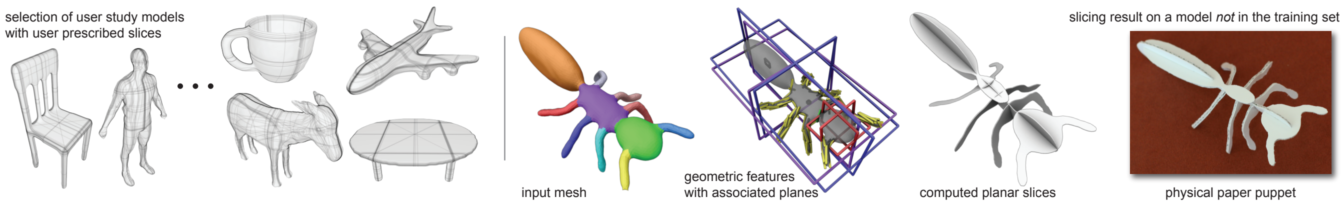


Figure 1: *Based on a user study, where participants abstracted meshes of common objects using a small collection of planar slices, we develop an automatic algorithm to create planar slice based abstractions of (untrained) models. Starting from a set of planar slices, approximating the object's geometric features, the algorithm picks a subset of planes based on relative feature importance learned from the user study.*

Abstract

Minimalist object representations or *shape-proxies* that spark and inspire human perception of shape remain an incompletely understood, yet powerful aspect of visual communication. We explore the use of planar sections, i.e., the contours of intersection of planes with a 3D object, for creating shape abstractions, motivated by their popularity in art and engineering. We first perform a user study to show that humans do define consistent and similar planar section proxies for common objects. Interestingly, we observe a strong correlation between user-defined planes and geometric features of objects. Further we show that the problem of finding the minimum set of planes that capture a set of 3D geometric shape features is both NP-hard and not always the proxy a user would pick. Guided by the principles inferred from our user study, we present an algorithm that progressively selects planes to maximize feature coverage, which in turn influence the selection of subsequent planes. The algorithmic framework easily incorporates various shape features, while their relative importance values are computed and validated from the user study data. We use our algorithm to compute planar slices for various objects, validate their utility towards object abstraction using a second user study, and conclude showing the potential applications of the extracted planar slice shape proxies.

Keywords: abstraction, shape proxy, shape perception

Links: DL PDF WEB DATA CODE

1 Introduction

Over the last few decades, great strides have been made in the area of acquisition and modeling of 3D geometry. The underlying shape representation is typically a collection of polygons, popular for their generality, rendering efficiency, and amenability to geometry processing algorithms. Unfortunately, such a representation can be expensive and, in itself, neither conveys the essence of the depicted object nor aids in our understanding of the represented shape.

Artists and sculptors have long explored minimalist shape proxies to highlight defining aspects of familiar objects (see Figure 2). As humans, we effortlessly perceive the underlying shapes even from such sectional representations, which greatly differ from their surface representations. In fact, the sparse nature of these representations allows us to see otherwise occluded details and the absence of unnecessary detail makes such artforms attractive, fascinating, and sometimes mysterious. Evidence suggests that symbolic abstractions [Edwards 2002] dominate our mental model of objects. Thus across cultures, we both recognize shape proxies quickly and tend to communicate objects by drawing them as symbolic abstractions.

Planar section proxies are also motivated by medical and engineering visualization where section planes are used to illustrate the interior details of complex shapes (see Figure 2). These planar-sections often pass through anatomic landmarks or engineering features such as channels or bosses, reaffirming our use of high-level shape features like segments, symmetries, ridges and valleys to define planar sections in Section 4.

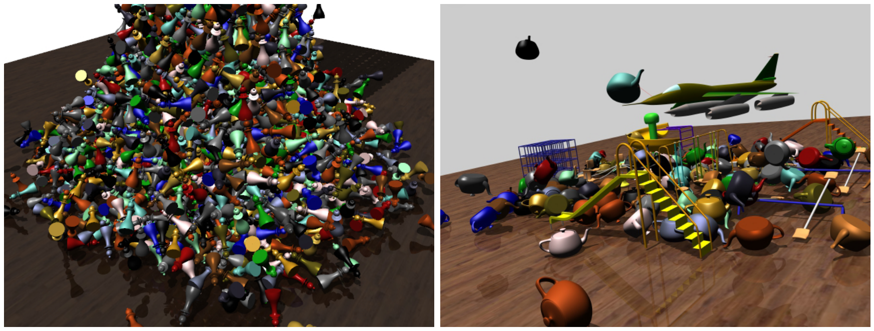
In geometry processing, model simplification [Cignoni et al. 1997] and variational shape approximation [Cohen-Steiner et al. 2004] remain the dominant modes of shape abstraction. Recently, there has been work on shape abstraction using curve networks [Mehra et al. 2009; de Goes et al. 2011]. These methods and their variants evaluate the quality of an approximation by the geometric deviation of the proxy from the original shape. In contrast, we aim to explore shape abstractions that are based on human perception of form. Given a shape S, our goal is to produce a proxy $\mathscr{S}$ using planar sections of the shape, such that perceptually S and $\mathscr{S}$ are comparable, while representationally $|\mathscr{S}| \ll |S|$. Note that $\mathscr{S}$ are S likely to be quite different from a purely geometric or topological standpoint. The fundamental difficulty in proposing or evaluating

Figure 2: *Planar sections in art (©Alexander Calder), sculpture (wood puppet), design (section chair, ©Hebert Franco) and scientific visualization.*

ACM Reference Format
McCrae, J., Singh, K., Mitra, N. 2011. Slices: A Shape-proxy based on Planar Sections.
ACM Trans. Graph. 30, 6, Article 168 (December 2011), 11 pages. DOI = 10.1145/2024156.2024202
http://doi.acm.org/10.1145/2024156.2024202.

Coons BVH for Freeform Geometric Models

Yong-Joon Kim[1] Young-Taek Oh[1] Seung-Hyun Yoon[2] Myung-Soo Kim[1]* Gershon Elber[3]
[1]Seoul National University [2]Dongguk University [3]Technion



(a) (b)

Figure 1: *(a) 10000 freeform geometric models (chess pieces) falling into a pile, where the same models share a common BVH of Coons patches approximating the given freeform NURBS surfaces within an error bound* 10^{-5}*, where the unit length is taken as the largest side length of the minimum bounding box of the model, and (b) the minimum distance computation between a flying B58 model and a complex dynamic scene with many Utah teapots falling to the playground.*

Abstract

We present a compact representation for the bounding volume hierarchy (BVH) of freeform NURBS surfaces using Coons patches. Following the Coons construction, each subpatch can be bounded very efficiently using the bilinear surface determined by the four corners. The BVH of freeform surfaces is represented as a hierarchy of Coons patch approximation until the difference is reduced to within a given error bound. Each leaf node contains a single Coons patch, where a detailed BVH for the patch can be represented very compactly using two lists (containing curve approximation errors) of length proportional only to the height of the BVH. We demonstrate the effectiveness of our compact BVH representation using several experimental results from real-time applications in collision detection and minimum distance computation for freeform models.

Keywords: Coons patch, freeform surface, bilinear surface, NURBS, bounding volume hierarchy (BVH), tetrahedron, offset, collision detection, minimum distance computation

Links: DL PDF

1 Introduction

Hierarchical spatial data structures play an essential role in the design of efficient geometric algorithms for three-dimensional objects [Samet 2006]. Real-time algorithms for polygonal meshes employ various different types of BVHs that are built in a preprocessing stage of the geometric computation [Akenine-Möller et al. 2008]. The BVH for a polygonal mesh usually requires a much larger memory space compared to the original model itself [Yoon and Manocha 2006]. Thus it is an important subject of research to develop compact representations for BVH structures.

Freeform geometric models are more compact than polygonal meshes. The BVH structure of freeform geometry can be generated by recursively subdividing the freeform surfaces [Johnson and Cohen 1998]. Nevertheless, it is unclear, in general, where to stop the recursive subdivision and how to proceed with the geometric computation when we reach the leaf level. In this paper, we address these two important issues and propose a compact BVH construction scheme for freeform geometry that is based on the special structure of the Coons patch.

The Coons patch is one of the earliest freeform representation schemes in CAGD and was developed in the early 1960's [Coons 1964]. (For an introduction to Coons patches, see Chapter 14 of [Cohen et al. 2001] and Chapter 22 of [Farin 2002].) Compared with other freeform surfaces, such as B-spline or Bézier surfaces, Coons patches are seldom used in contemporary freeform modeling applications. Nevertheless, there are many useful properties of Coons patches that we employ in this work for the acceleration of geometric algorithms for freeform shapes. The most important property, for our purpose, is that Coons patches are uniquely determined by their boundary curves. As a direct consequence, Coons patches can be subdivided very efficiently by evaluating points only on their boundary curves.

The hierarchy of recursive Coons approximations generates a BVH, where the leaf nodes contain the Coons patches approximating the original freeform surface to within a given error bound. The interior nodes of the BVH correspond to freeform surface patches that are recursively subdivided. While it may seem there is nothing dramatically different from conventional BVH approaches, the significant difference is in the size of the surface patches that are stored in the leaf nodes, and hence the size of the entire BVH. Because the Coons patches approximate the freeform surfaces very

*Corresponding author; e-mail: mskim@snu.ac.kr

ACM Reference Format
Kim, Y., Oh, Y., Yoon, S., Kim, M., Elber, G. 2011. Coons BVH for Freeform Geometric Models. *ACM Trans. Graph. 30*, 6, Article 169 (December 2011), 8 pages. DOI = 10.1145/2024156.2024203 http://doi.acm.org/10.1145/2024156.2024203.

Practical Spectral Characterization of Trichromatic Cameras

Martin Rump* Universität Bonn
Arno Zinke† GfaR mbH
Reinhard Klein‡ Universität Bonn

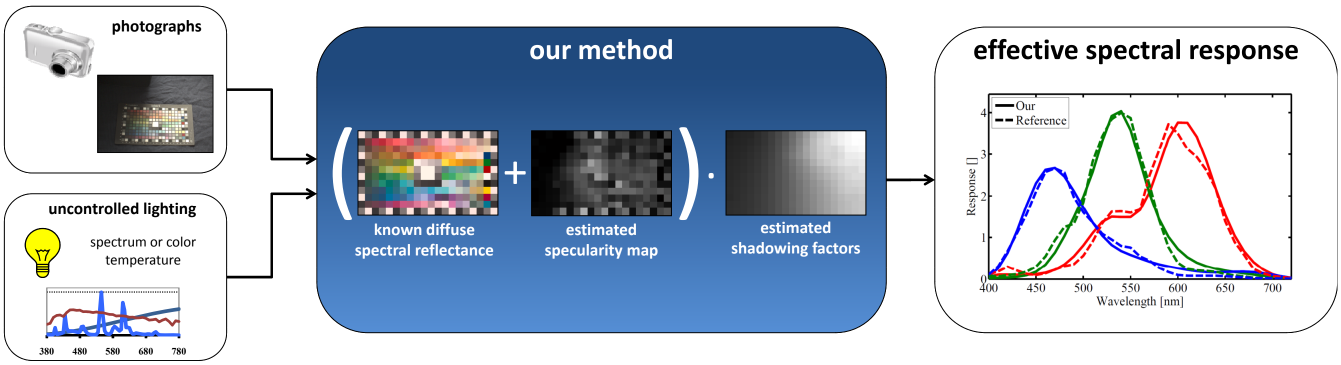


Figure 1: *Illustrating the workflow of our approach. Our method is capable of reconstructing the effective spectral response of trichromatic cameras from photographs of a color chart captured under uncontrolled lighting conditions. The only assumption is that the spectral power distribution of the dominant illuminant is given. Alternatively, in many relevant cases even knowing a color temperature of the illumination is sufficient. By introducing a novel imaging model accounting for specularity and spatially varying illumination, our approach is robust with respect to out-of-lab capturing conditions. While greatly simplifying spectral calibration of trichromatic cameras we achieve a quality comparable to very costly reference measurements for practical applications such as color processing or imaging system design.*

Abstract

Simple and effective geometric and radiometric calibration of camera devices has enabled the use of consumer digital cameras for HDR photography, for image based measurement and similar applications requiring a deeper understanding about the camera characteristics. However, to date no such practical methods for estimating the spectral response of cameras are available. Existing approaches require costly hardware and controlled acquisition conditions limiting their applicability. Consequently, even though being highly desirable for color correction and color processing purposes as well as for designing image-based measurement or photographic setups, the spectral response of a camera is rarely considered. Our objective is to close this gap. In this work a practical approach for multispectral characterization of trichromatic cameras is presented. Taking photographs of a color chart and measuring the average lighting using a spectrophotometer the effective spectral response of a camera can be estimated for a wide range of out-of-lab environments. By comprehensive cross validation experiments we prove that the new method performs well compared to costly reference measurements. Moreover, we show that our technique can also be used to generate ICC profiles with higher accuracy and less constrained capturing conditions compared to state-of-the-art ICC profilers.

CR Categories: I.3.6 [Computer Graphics]: Methodology and Techniques—Device independence; I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Camera calibration

Keywords: spectral, camera, color filter, calibration

*e-mail: rump@cs.uni-bonn.de
†e-mail: a.zinke@gfar.de
‡e-mail: rk@cs.uni-bonn.de

Links: DL PDF

1 Introduction

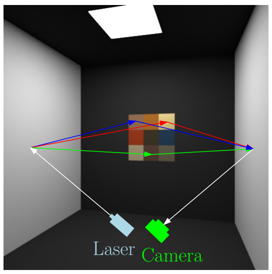
Calibration of camera devices is of fundamental importance for different applications such as image based measurement and color accurate photography. The calibration process may include several steps: geometric calibration to estimate parameters determining image projection, radiometric calibration to compensate for nonlinearities in energy response and spectral calibration that determines the spectral response of the red, green and blue color channels. To make a method usable in practice and to make it accessible to a wide range of users simplicity is a key factor. For example, the success of HDR imaging is not thinkable without practical solutions for recovering the energy response of a camera. To date no such practical methods for estimating the spectral response of cameras are available. However, the spectral response is highly desirable for color correction and color processing purposes like white balancing [Cherdhirunkorn et al. 2006] as well as for designing image-based measurement or photographic setups [Farrell et al. 2008]. Another field of application is spectral imaging where techniques utilizing trichromatic cameras with known spectral response have been proposed ([Imai and Berns 1999], [Rump and Klein 2010]). Unfortunately, existing approaches for measuring the spectral response require costly hardware such as either tunable filters or light sources (e.g. monochromators) or they need strictly controlled acquisition

ACM Reference Format
Rump, M., Zinke, A., Klein, R. 2011. Practical Spectral Characterization of Trichromatic Cameras. *ACM Trans. Graph. 30*, 6, Article 170 (December 2011), 9 pages. DOI = 10.1145/2024156.2024204 http://doi.acm.org/10.1145/2024156.2024204.

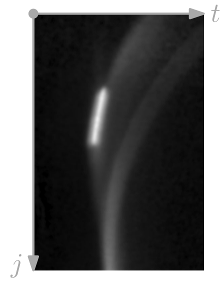
Single View Reflectance Capture using Multiplexed Scattering and Time-of-flight Imaging

Nikhil Naik [1] Shuang Zhao [2] Andreas Velten [1] Ramesh Raskar [1] Kavita Bala [2]

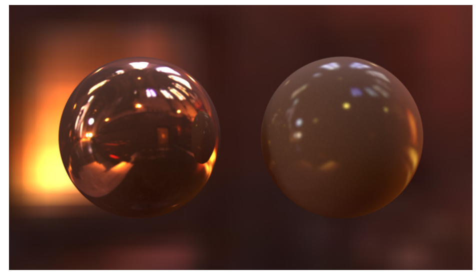
[1]MIT Media Lab [2]Cornell University



(a)



(b)



(c)

Figure 1: *By analyzing indirectly scattered light, our approach can recover reflectance from a single viewpoint of a camera and a laser projector. (a) To recover reflectance of patches on the back wall, we successively illuminate points on the left wall to indirectly create a range of incoming directions. The camera observes points on the right wall which in turn are illuminated by a range of outgoing directions on the patches. (b) Using space-time images ("streak images") from a time-of flight camera, we recover the reflectances of multiple patches simultaneously even though multiple pairs of patch and outgoing-direction contribute to the same point on the right wall. The figure shows a real streak image captured with two patches on the back wall. (c) We render two spheres with the simultaneously recovered parametric reflectance models of copper (left) and jasper (right) in simulation.*

Abstract

This paper introduces the concept of time-of-flight reflectance estimation, and demonstrates a new technique that allows a camera to rapidly acquire reflectance properties of objects from a single viewpoint, over relatively long distances and without encircling equipment. We measure material properties by indirectly illuminating an object by a laser source, and observing its reflected light indirectly using a time-of-flight camera. The configuration collectively acquires dense angular, but low spatial sampling, within a limited solid angle range - all from a single viewpoint. Our ultra-fast imaging approach captures space-time "streak images" that can separate out different bounces of light based on path length. Entanglements arise in the streak images mixing signals from multiple paths if they have the same total path length. We show how reflectances can be recovered by solving for a linear system of equations and assuming parametric material models; fitting to lower dimensional reflectance models enables us to disentangle measurements.

We demonstrate proof-of-concept results of parametric reflectance models for homogeneous and discretized heterogeneous patches, both using simulation and experimental hardware. As compared to lengthy or highly calibrated BRDF acquisition techniques, we demonstrate a device that can rapidly, on the order of seconds, capture meaningful reflectance information. We expect hardware advances to improve the portability and speed of this device.

ACM Reference Format
Naik, N., Zhao, S., Velten, A., Raskar, R., Bala, K. 2011. Single View Reflectance Capture using Multiplexed Scattering and Time-of-flight Imaging. *ACM Trans. Graph. 30*, 6, Article 171 (December 2011), 10 pages. DOI = 10.1145/2024156.2024205 http://doi.acm.org/10.1145/2024156.2024205.

Keywords: computational photography, multipath light transport, reflectance acquisition, global illumination, time of flight

Links: DL PDF

1 Introduction

Acquiring material properties of real-world materials has a long and rich history in computer graphics; existing techniques directly image the sample being measured to acquire different properties including tabulated reflectance functions, spatially varying reflectances, and parametric models (see [Weyrich et al. 2009] for a survey of state-of-the-art techniques.). These reflectance functions, are necessary for relighting, material editing, and rendering, as well as for matching and material identification.

In this paper, we present a new acquisition approach to reflectance measurement. Our approach is unique in two ways: we exploit ultra-fast time-of-flight (ToF) imaging to achieve rapid acquisition of materials; and we use indirect observation to acquire many samples simultaneously, and in fact, even permit around-the-corner measurement of reflectance properties. The key insight of this research is to exploit ultra-fast imaging to measure individual light transport paths, based on the distance traveled at the speed of light. This capability uniquely lets us separately measure the direct (0-bounce), 1-bounce, 2-bounce, and more, light paths; in comparison, traditional approaches use controlled laboratory settings to minimize the impact of multi-bounce light transport, or must explicitly separate direct and indirect lighting from all bounces.

We make the following contributions:

a) We present a new technique for reflectance acquisition by separating light multiplexed along different transport paths. Our approach uses indirect viewing with 3-bounce scattering coupled with time-of-flight imaging to capture reflectances. Our proof-of-concept system demonstrates first steps towards rapid material acquisition.

Estimating Dual-scale Properties of Glossy Surfaces from Step-edge Lighting

Chun-Po Wang* Cornell University

Noah Snavely† Cornell University

Steve Marschner‡ Cornell University

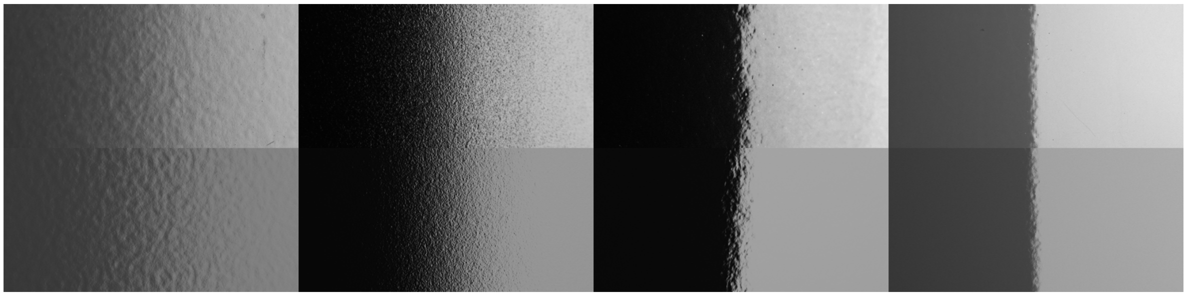


Figure 1: *Example results of capturing four different surfaces under step-edge illumination. From left to right: painted bookend, metal bookend, shiny black folder, and whiteboard. The upper row shows real photos, and the bottom row shows images synthesized from the parameters acquired by our method.*

Abstract

This paper introduces a rapid appearance capture method suited for a variety of common indoor surfaces, in which a single photograph of the reflection of a step edge is used to estimate both a BRDF and a statistical model for visible surface geometry, or mesostructure. It is applicable to surfaces with statistically stationary variation in surface height, even when these variations are large enough to produce visible texture in the image. Results are shown from a prototype system using a separate camera and LCD, demonstrating good visual matches for a range of man-made indoor materials.

Keywords: appearance capture, reflectance, rendering

Links: DL PDF WEB DATA

1 Introduction

Acquiring the appearance of real-world surfaces is important when real scenes need to be modeled faithfully by computer graphics. However, even the near-homogeneous manufactured surfaces—paints, metals, plastics—that are ubiquitous in man-made environments have evaded easy measurement. With the assumption of homogeneity, surfaces can be described by a single reflectance function that can be estimated from images [Yu et al. 1999] or rapidly acquired by handheld devices [Dong et al. 2010; X-Rite 2011], but the resulting surfaces always look too featureless, because surfaces with no visible texture are rare. Good, detailed appearance can be achieved by measuring high-resolution parameter maps for particular samples [Gardner et al. 2003; Ghosh et al. 2009; Ghosh et al. 2010], but these methods are considerably less convenient—too much work to add subtle texture to a basically homogeneous surface.

In this paper we explore a middle path, suited for many common indoor surfaces, in which a rapid single-image measurement results in a statistical description of both visible and microscopic surface roughness that is sufficient to yield renderings of a surface that qualitatively *match* its appearance. Figure 2 shows some examples of surfaces that appear in a typical indoor scene. Note that these surfaces can be described not only in terms of reflectance—the diffuseness of the wall, the gloss evident in the reflection off the cabinet—but also by visible "bumps" with characteristic frequency content. Both of these phenomena are important to the appearance of surfaces (particularly in high-resolution imagery), and our approach explicitly handles both. However, our goal is *not* to measure exact surface properties—full BRDFs, accurate normal maps—but instead to capture enough information about the statistics of a surface to achieve a qualitative appearance match. To this end, we propose a dual-level model of surface appearance, with one level modeling microscopic surface geometry (described by the BRDF), and another modeling visible surface bumps (the mesostructure often represented in a normal or bump map). To estimate the parameters of our model, we propose an appearance capture system that uses a single image of the reflection of *step-edge* illumination from a planar sample of a surface to estimate both the BRDF and the statistics of meso-scale geometry (Figure 1). While we present a prototype capture system using a separate camera and LCD, our system could potentially be implemented on a consumer device with a display and a front-facing camera, such as an Apple iPad, and eventually could use natural illumination (which often includes step edges). Our appearance model and capture system are designed as a whole with the goal of enabling very simple, robust acquisition, while still handling many interesting real-world surfaces. We begin by briefly describing these two components.

Dual-level appearance model. Our reflectance model includes both diffuse and specular components; to represent specular reflection, we use a Cook-Torrance model (a microfacet-based model). The BRDF is then modeled by the slope distribution of the microfacets with a single parameter, the slope variance. Similarly, we observe that the visible bumps (the meso-scale structure) can be seen

*e-mail:cpwang@cs.cornell.edu
†e-mail:snavely@cs.cornell.edu
‡e-mail:srm@cs.cornell.edu

ACM Reference Format
Wang, C., Snavely, N., Marschner, S. 2011. Estimating Dual-scale Properties of Glossy Surfaces from Step-edge Lighting. *ACM Trans. Graph. 30*, 6, Article 172 (December 2011), 11 pages. DOI = 10.1145/2024156.2024206 http://doi.acm.org/10.1145/2024156.2024206.

Interactive Hair Rendering and Appearance Editing under Environment Lighting

Kun Xu[1] Li-Qian Ma[1] Bo Ren[1] Rui Wang[2] Shi-Min Hu[1]

[1] TNList, Department of Computer Science and Technology, Tsinghua University, Beijing
[2] Department of Computer Science, University of Massachusetts, Amherst

Figure 1: *Our algorithm achieves interactive hair rendering and appearance editing under environment lighting, including both single and multiple scattering effects. In this example, the user directly paints onto the hair to edit the spatially-varying scattering parameters. This results in a dynamic simulation of hair coloring. From left to right, the hair is dyed with a progressively more vivid color. The environment map is represented by 40 SRBF lights, and our algorithm runs at 8.3 fps on an NVIDIA GTX 580.*

Abstract

We present an interactive algorithm for hair rendering and appearance editing under complex environment lighting represented as spherical radial basis functions (SRBFs). Our main contribution is to derive a compact 1D circular Gaussian representation that can accurately model the hair scattering function introduced by [Marschner et al. 2003]. The primary benefit of this representation is that it enables us to evaluate, at run-time, closed-form integrals of the scattering function with each SRBF light, resulting in efficient computation of both single and multiple scatterings. In contrast to previous work, our algorithm computes the rendering integrals entirely on the fly and does not depend on expensive precomputation. Thus we allow the user to dynamically change the hair scattering parameters, which can vary spatially. Analyses show that our 1D circular Gaussian representation is both accurate and concise. In addition, our algorithm incorporates the eccentricity of the hair. We implement our algorithm on the GPU, achieving interactive hair rendering and simultaneous appearance editing under complex environment maps for the first time.

Keywords: Hair rendering, environment lighting, appearance editing, SRBFs, circular Gaussian, single and multiple scattering, GPU.

Links: DL PDF WEB

1 Introduction

ACM Reference Format
Xu, K., Ma, L., Ren, B., Wang, R., Hu, S. 2011. Interactive Hair Rendering and Appearance Editing under Environment Lighting. *ACM Trans. Graph. 30*, 6, Article 173 (December 2011), 10 pages.
DOI = 10.1145/2024156.2024207 http://doi.acm.org/10.1145/2024156.2024207.

In hair rendering, it is often desirable to support the dynamic changing of hair's scattering properties. This enables artists and designers to edit hair appearance at will, and receive realistic rendering feedbacks at interactive speed. Existing methods already support such capability under simple lighting, such as a few point or directional lights [Zinke et al. 2008; Yuksel and Keyser 2008; Shinya et al. 2010; Sadeghi et al. 2010]. However, it remains a challenge to render and simultaneously edit hair appearance under complex lighting such as environment maps [Debevec and Malik 1997]. Such lighting is important to convey the rich look of the hair in natural illumination conditions.

The main difficulty with environment lighting is the large number of directional lights that must be considered. An effective solution is to approximate the environment map as a set of spherical radial basis functions (SRBFs), yielding a low-dimensional representation. This approach has been studied in the context of precomputed light transport and BRDF aprpoximation [Tsai and Shih 2006; Green et al. 2007; Wang et al. 2009]. For hair rendering, Ren et al. [2010] proposed to integrate the hair scattering function with each SRBF light to produce realistic rendering effects. Their method incorporates both single and multiple scatterings. Unfortunately, as they precompute the integrals of the scattering function with sampled SRBF lights into 4D tables, their method requires fixing hair scattering parameters at precomputation time, disabling online editing.

In this paper, our goal is to enable realistic rendering and simultaneous editing of hair appearance under complex environment lighting. Similar to previous work, we represent an environment map using SRBF lights. Our main contribution is to derive a compact 1D circular Gaussian representation that can accurately model the hair scattering function introduced by [Marschner et al. 2003]. Analyses show that our representation is both concise and accurate. Exploiting the properties of Gaussian functions, the primary benefit of this representation is that it enables the run-time evaluation of closed-form integrals of the scattering function with SRBF lights. This results in efficient computation of both single and multiple scatterings, without the need of expensive precomputation. By using this approach, our algorithm evaluates the rendering integrals entirely on the fly, allowing the user to dynamically change hair scattering parameters at will. Our approach can be seen as an accurate model of the scattering function that is particularly suitable for integration

Image Smoothing via L_0 Gradient Minimization

Li Xu* Cewu Lu* Yi Xu Jiaya Jia

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

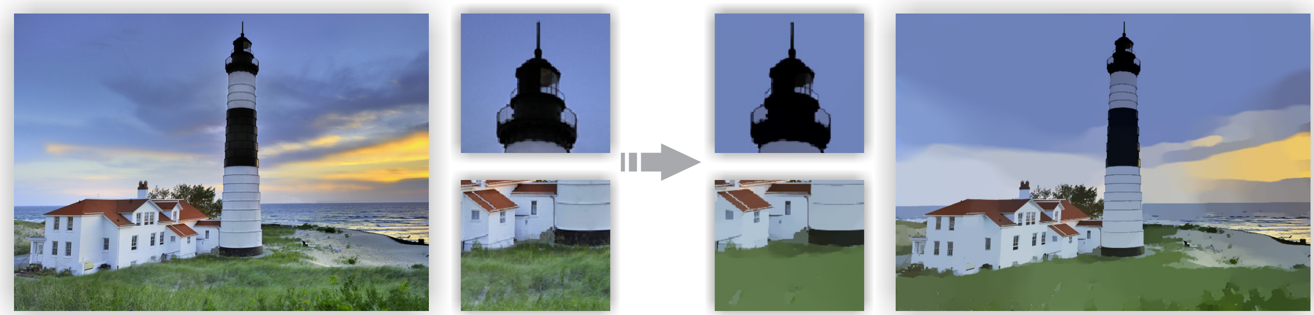


Figure 1: *L_0 smoothing accomplished by global small-magnitude gradient removal. Our method suppresses low-amplitude details. Meanwhile it globally retains and sharpens salient edges. Even the high-contrast thin edges on the tower are preserved.*

Abstract

We present a new image editing method, particularly effective for sharpening major edges by increasing the steepness of transition while eliminating a manageable degree of low-amplitude structures. The seemingly contradictive effect is achieved in an optimization framework making use of L_0 gradient minimization, which can globally control how many non-zero gradients are resulted in to approximate prominent structure in a sparsity-control manner. Unlike other edge-preserving smoothing approaches, our method does not depend on local features, but instead globally locates important edges. It, as a fundamental tool, finds many applications and is particularly beneficial to edge extraction, clip-art JPEG artifact removal, and non-photorealistic effect generation.

Keywords: image smoothing, L_0 sparsity, sharpening, filtering

Links: DL PDF WEB VIDEO CODE

1 Introduction

Photos comprise rich and well-structured visual information. In human visual perception, edges are effective and expressive stimulation, vital for neural interpretation to make the best sense of the scene. In manipulating and understanding pictures, high-level inference with regard to salient structures was intensively attended to. Research following this line embodies generality and usefulness in a wide range of applications, including image recognition, segmentation, object classification, and many other photo editing and non-photorealistic rendering tasks.

Both authors contributed equally to this work.

ACM Reference Format
Xu, L., Lu, C., Xu, Y., Jia, J. 2011. Image Smoothing via L_0 Gradient Minimization. *ACM Trans. Graph. 30*, 6, Article 174 (December 2011), 11 pages. DOI = 10.1145/2024156.2024208
http://doi.acm.org/10.1145/2024156.2024208.

(a) Abstraction (b) Pencil Sketch Rendering

Figure 2: *Our L_0 smoothing results avail non-photorealistic effect generation.*

We in this paper present a new editing tool, greatly helpful for characterizing and enhancing fundamental image constituents, i.e., salient edges, and in the meantime for diminishing insignificant details. Our method relates in spirit to edge-preserving smoothing [Tomasi and Manduchi 1998; Durand and Dorsey 2002; Paris and Durand 2006; Farbman et al. 2008; Subr et al. 2009; Kass and Solomon 2010] that aims to retain primary color change, and yet differs from them in essence in focus and in mechanism. Our objective is to *globally* maintain and possibly enhance the most prominent set of edges by increasing steepness of transition while not affecting the overall acutance. It enables faithful principal-structure representation.

Algorithmically, we propose a sparse gradient counting scheme in an optimization framework. The main contribution is a new strategy to confine the discrete *number* of intensity changes among neighboring pixels, which links mathematically to the L_0 norm for information sparsity pursuit. This idea also leads to an unconventional global optimization procedure involving a discrete metric, whose solution enables diversified edge manipulation according to saliency. The qualitative effect of our method is to thin salient edges, which makes them easier to be detected and more visually distinct. Different from color quantization and segmentation, our enhanced edges are generally in line with the original ones. Even small-resolution objects and thin edges can be faithfully maintained if they are structurally conspicuous, as shown in Fig. 1.

The framework is general and finds several applications. We apply it to compression-artifact degraded clip-art recovery. High quality results can be obtained in our extensive experiments. Our method

Convolution Pyramids

Zeev Farbman
The Hebrew University

Raanan Fattal
The Hebrew University

Dani Lischinski
The Hebrew University

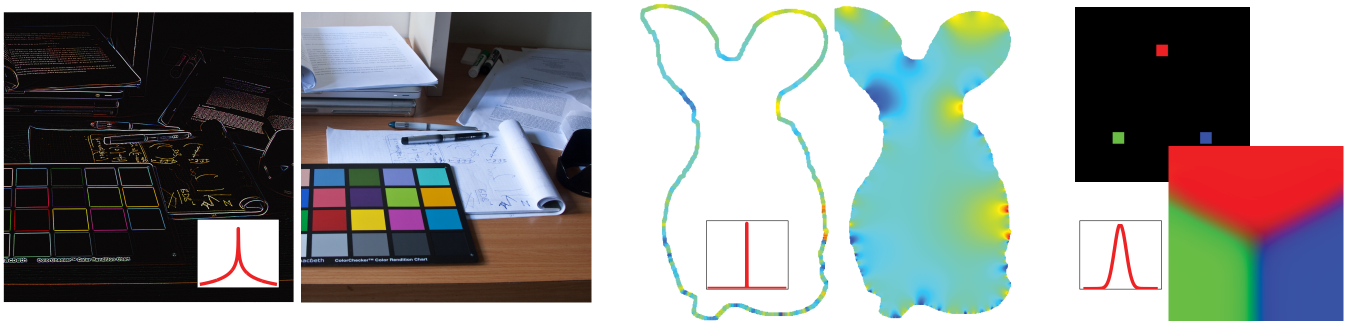


Figure 1: *Three examples of applications that benefit from our fast convolution approximation. Left: gradient field integration; Middle: membrane interpolation; Right: scattered data interpolation. The insets show the shapes of the corresponding kernels.*

Abstract

We present a novel approach for rapid numerical approximation of convolutions with filters of large support. Our approach consists of a multiscale scheme, fashioned after the wavelet transform, which computes the approximation in linear time. Given a specific large target filter to approximate, we first use numerical optimization to design a set of small kernels, which are then used to perform the analysis and synthesis steps of our multiscale transform. Once the optimization has been done, the resulting transform can be applied to any signal in linear time. We demonstrate that our method is well suited for tasks such as gradient field integration, seamless image cloning, and scattered data interpolation, outperforming existing state-of-the-art methods.

Keywords: convolution, Green's functions, Poisson equation, seamless cloning, scattered data interpolation, Shepard's method

Links: DL PDF WEB

1 Introduction

Many tasks in computer graphics and image processing involve applying large linear translation-invariant (LTI) filters to images. Common examples include low- and high-pass filtering of images, and measuring the image response to various filter banks. Some less obvious tasks that can also be accomplished using large LTI filters are demonstrated in Figure 1: reconstructing images by integrating their gradient field [Fattal et al. 2002], fitting a smooth membrane to interpolate a set of boundary values [Pérez et al. 2003; Agarwala 2007], and scattered data interpolation [Lewis et al. 2010].

ACM Reference Format
Farbman, Z., Fattal, R., Lischinski, D. 2011. Convolution Pyramids. *ACM Trans. Graph. 30*, 6, Article 175 (December 2011), 8 pages. DOI = 10.1145/2024156.2024209 http://doi.acm.org/10.1145/2024156.2024209.

While convolution is the most straightforward way of applying an LTI filter to an image, it comes with a high computational cost: $O(n^2)$ operations are required to convolve an n-pixel image with a kernel of comparable size. The Fast Fourier Transform offers a more efficient, $O(n \log n)$ alternative for periodic domains [Brigham 1988]. Other fast approaches have been proposed for certain special cases. For example, Burt [1981] describes a multiscale approach, which can *approximate* a convolution with a large Gaussian kernel in $O(n)$ time at hierarchically subsampled locations. We review this and several other related approaches in the next section.

In this work, we generalize these ideas, and describe a novel multiscale framework that is not limited to approximating a specific kernel, but can be tuned to reproduce the effect of a number of useful large LTI filters, while operating in $O(n)$ time. Specifically, we demonstrate the applicability of our framework to convolutions with the Green's functions that span the solutions of the Poisson equation, inverse distance weighting kernels for membrane interpolation, and wide-support Gaussian kernels for scattered data interpolation. These applications are demonstrated in Figure 1.

Our method consists of a multiscale scheme, resembling the Laplacian Pyramid, as well as certain wavelet transforms. However, unlike these more general purpose transforms, our approach is to custom-tailor the transform to directly approximate the effect of a given LTI operator. In other words, while previous multiscale constructions are typically used to transform the problem into a space where it can be better solved, in our approach the transform itself directly yields the desired solution.

Specifically, we repeatedly perform convolutions with three small, fixed-width kernels, while downsampling and upsampling the image, so as to operate on all of its scales. The weights of each of these kernels are numerically optimized such that the *overall* action of the transform best approximates a convolution operation with some *target filter*. The optimization only needs to be done once for each target filter, and then the resulting multiscale transform may be applied to any input signal in $O(n)$ time. The motivation behind this design was to avoid dealing with the analytical challenges that arise from the non-idealness of small finite filters, on the one hand, while attempting to make the most out of the linear computational budget, on the other.

Our scheme's ability to closely approximate convolutions with the

GPU-Efficient Recursive Filtering and Summed-Area Tables

Diego Nehab[1] André Maximo[1] Rodolfo S. Lima[2] Hugues Hoppe[3]
[1]IMPA [2]Digitok [3]Microsoft Research

Abstract

Image processing operations like blurring, inverse convolution, and summed-area tables are often computed efficiently as a sequence of 1D recursive filters. While much research has explored parallel recursive filtering, prior techniques do not optimize across the entire filter sequence. Typically, a separate filter (or often a causal-anticausal filter pair) is required in each dimension. Computing these filter passes independently results in significant traffic to global memory, creating a bottleneck in GPU systems. We present a new algorithmic framework for parallel evaluation. It partitions the image into 2D blocks, with a small band of additional data buffered along each block perimeter. We show that these perimeter bands are sufficient to accumulate the effects of the successive filters. A remarkable result is that the image data is read only twice and written just once, independent of image size, and thus total memory bandwidth is reduced even compared to the traditional serial algorithm. We demonstrate significant speedups in GPU computation.

Links: DL PDF WEB VIDEO DATA CODE

1 Introduction

Linear filtering (i.e. convolution) is commonly used to blur, sharpen, or downsample images. A direct implementation evaluating a filter of support d on an $h \times w$-image has cost $O(hwd)$. For filters with a wide impulse response, the Fast Fourier Transform reduces the cost to $O(hw \log hw)$, regardless of filter support. Often, similar results can be obtained with a *recursive filter*, in which the computation reuses prior outputs, e.g. $y_i = x_i - \frac{1}{2} y_{i-1}$. Such feedback allows for an infinite impulse response (IIR), i.e. an effectively large filter support, at reduced cost $O(hwr)$, where the number r of recursive feedbacks (a.k.a. the filter *order*) is small relative to d. Recursive filters are a key computational tool in several applications:

- **Low-pass filtering.** Filters like Gaussian kernels are well approximated by a pair of low-order causal and anticausal recursive filters [e.g. Deriche 1992; van Vliet et al. 1998].
- **Inverse convolution.** If an image X is the result of convolving an image V with a compactly supported filter F, i.e. $X = V * F$, the original image can be recovered as $V = X * F^{-1}$. Although the inverse filter F^{-1} generally has infinite support, it can be expressed *exactly* as a sequence of low-order recursive filters.
- **Summed-area tables.** Such tables store the sum of all pixel values above and to the left of each pixel [Crow 1984]. They have many uses in graphics and vision. On a GPU, summed-area tables are typically computed with prefix sums over all columns then all rows of the image [Hensley et al. 2005]. Crucially, a prefix sum is a special case of a 1D first-order recursive filter.

These applications all have in common the fact that they invoke a *sequence* of recursive filters. First, a 2D operation is decomposed into separate 1D filters. Second, except for the case of summed-area tables, one usually desires a centered and well-shaped impulse response function, and this requires the combination of a causal and anticausal filter pair in each dimension.

As an example, consider the problem of finding the coefficient image V such that reconstruction with the bicubic B-spline interpolates the pixel values of a given image X. (Such a preprocess is required in techniques for high-quality image filtering [Blu et al. 1999, 2001].) The coefficients and image are related by the convolution $X = V * \frac{1}{6}[1\ 4\ 1] * \frac{1}{6}[1\ 4\ 1]^T$. The inverse convolution can be decomposed into four successive 1D recursive filters:

$$
\begin{aligned}
y_{i,j} &= 6x_{i,j} + \alpha y_{i-1,j}, && \text{(causal in } i\text{)}\\
z_{i,j} &= -\alpha y_{i,j} + \alpha z_{i+1,j}, && \text{(anticausal in } i\text{)}\\
u_{i,j} &= 6z_{i,j} + \alpha u_{i,j-1}, && \text{(causal in } j\text{)}\\
v_{i,j} &= -\alpha u_{i,j} + \alpha v_{i,j+1}, \text{ with } \alpha = \sqrt{3} - 2. && \text{(anticausal in } j\text{)}
\end{aligned}
$$

Although the parallelization of such recursive filters is challenging due to the many feedback data dependencies, efficient algorithms have been developed as reviewed in section 2.

However, an important element not considered in previous work is optimization over the entire sequence of recursive filters. If each filter is applied independently as is common, the entire image must be read from global memory, and then written back each time. Due to the low computational intensity of recursive filters, this memory traffic becomes a clear bottleneck. Because the data access pattern is different in successive filters (down, up, right, left), the different computation passes cannot be directly fused.

Contribution In this paper, we present a new algorithmic framework to reduce memory bandwidth by *overlapping* computation over the full sequence of recursive filters, across both multiple dimensions and causal-anticausal pairs. The resulting algorithm scales efficiently to images of arbitrary size.

Our approach partitions the image into 2D blocks of size $b \times b$. This blocking serves several purposes. Inter-block parallelism provides sufficient independent tasks to hide memory latency. The blocks fit in on-chip memory for fast intra-block computations. And, the square blocks enables efficient propagation of data in both dimensions.

We associate a set of narrow rectangular buffers with each block's perimeter. These buffers' widths are the same as the filter order r. Intuitively, the goal of these buffers is to measure the filter response of each block in isolation (i.e. given zero boundary conditions), as well as to efficiently propagate the boundary conditions across the blocks given these measured responses.

The surprising result is that we are able to accumulate, with a single block-parallel pass over the input image, all the necessary responses to the filter sequence (down, up, right, left), so as to efficiently initialize boundary conditions for a second block-parallel pass that produces the final solution. Thus, the source image is read only twice, and written just once. A careful presentation requires algebraic derivations, and is detailed in sections 4–5.

Our framework also benefits the computation of summed-area tables (section 6). Even though in that setting there are no anticausal filters, the reduction in bandwidth due solely to the overlapped processing of the two dimensions still provides performance gains.

ACM Reference Format
Nehab, D., Maximo, A., Lima, R., Hoppe, H. 2011. GPU-Efficient Recursive Filtering and Summed-Area Tables. *ACM Trans. Graph. 30*, 6, Article 176 (December 2011), 11 pages. DOI = 10.1145/2024156.2024210 http://doi.acm.org/10.1145/2024156.2024210.

Multigrid and Multilevel Preconditioners for Computational Photography

Dilip Krishnan *
Department of Computer Science
New York University

Richard Szeliski †
Interactive Visual Media Group
Microsoft Research

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

Figure 1: *Colorization using a variety of multilevel techniques: (a) input gray image with color strokes overlaid; (b) "gold" (final) solution; (c) after one iteration of adaptive basis preconditioners (ABF) with partial sparsification; (d) after one iteration of regular coarsening algebraic multigrid (AMG) with Jacobi smoothing and (e) with four-color smoothing. The computational cost for the ABF approach is less than half of the AMG variants and the error with respect to the gold solution is lower (zoom in to see the differences). Four-color smoothing provides faster convergence (lower error) than Jacobi.*

Abstract

This paper unifies multigrid and multilevel (hierarchical) preconditioners, two widely-used approaches for solving computational photography and other computer graphics simulation problems. It provides detailed experimental comparisons of these techniques and their variants, including an analysis of relative computational costs and how these impact practical algorithm performance. We derive both theoretical convergence rates based on the condition numbers of the systems and their preconditioners, and empirical convergence rates drawn from real-world problems. We also develop new techniques for sparsifying higher connectivity problems, and compare our techniques to existing and newly developed variants such as algebraic and combinatorial multigrid. Our experimental results demonstrate that, except for highly irregular problems, adaptive hierarchical basis function preconditioners generally outperform alternative multigrid techniques, especially when computational complexity is taken into account.

Keywords: Computational photography, Poisson blending, colorization, multilevel techniques, fast PDE solution, parallel algorithms

Links: DL PDF

*dilip@cs.nyu.edu
†szeliski@microsoft.com

ACM Reference Format
Krishnan, D., Szeliski, R. 2011. Multigrid and Multilevel Preconditioners for Computational Photography. *ACM Trans. Graph. 30*, 6, Article 177 (December 2011), 9 pages. DOI = 10.1145/2024156.2024211 http://doi.acm.org/10.1145/2024156.2024211.

1 Introduction

Multigrid and multilevel preconditioning techniques have long been widely used in computer graphics and computational photography as a means of accelerating the solution of large gridded optimization problems such as geometric modeling [Gortler and Cohen 1995], high-dynamic range tone mapping [Fattal et al. 2002], Poisson and gradient-domain blending [Pérez et al. 2003; Levin et al. 2004b; Agarwala et al. 2004], colorization [Levin et al. 2004a] (Fig. 1), and natural image matting [Levin et al. 2008]. They have also found widespread application in the solution of computer vision problems such as surface interpolation, stereo matching, optical flow, and shape from shading [Terzopoulos 1986; Szeliski 1990; Pentland 1994], as well as large-scale finite element and finite difference modeling [Briggs et al. 2000; Trottenberg et al. 2000].

While the locally adaptive hierarchical basis function technique developed by Szeliski [Szeliski 2006] showed impressive speedups over earlier non-adaptive basis functions [Szeliski 1990], it was never adequately compared to state-of-the art multigrid techniques such as algebraic multigrid [Briggs et al. 2000; Trottenberg et al. 2000] or to newer techniques such as combinatorial multigrid [Koutis et al. 2009]. Furthermore, the original technique was restricted to problems defined on four neighbor ($\mathcal{N}_4$) grids.

In this paper, we generalize the sparsification method introduced in [Szeliski 2006] to handle a larger class of grid topologies, and show how multi-level preconditioners can be enhanced with smoothing to create hybrid algorithms that accrue the advantages of both adaptive basis preconditioning and multigrid relaxation. We also provide a detailed study of the convergence properties of all of these algorithms using both condition number analysis and empirical observations of convergence rates on real-world problems in computer graphics and computational photography.

Our experimental results demonstrate that locally adaptive hierarchical basis functions (ABF) [Szeliski 2006] combined with the extensions proposed in this paper generally outperform algebraic and combinatorial multigrid techniques, especially once computational complexity is taken into account. However, for highly irregular and inhomogeneous problems, techniques that use adaptive coarsening strategies (Section 3), which our approach does not currently use,

Modular Radiance Transfer

Bradford J. Loos[1,2] Lakulish Antani[1,3] Kenny Mitchell[1] Derek Nowrouzezahrai[4] Wojciech Jarosz[4] Peter-Pike Sloan[1]
[1]Disney Interactive Studios [2]University of Utah [3]UNC Chapel Hill [4]Disney Research Zurich

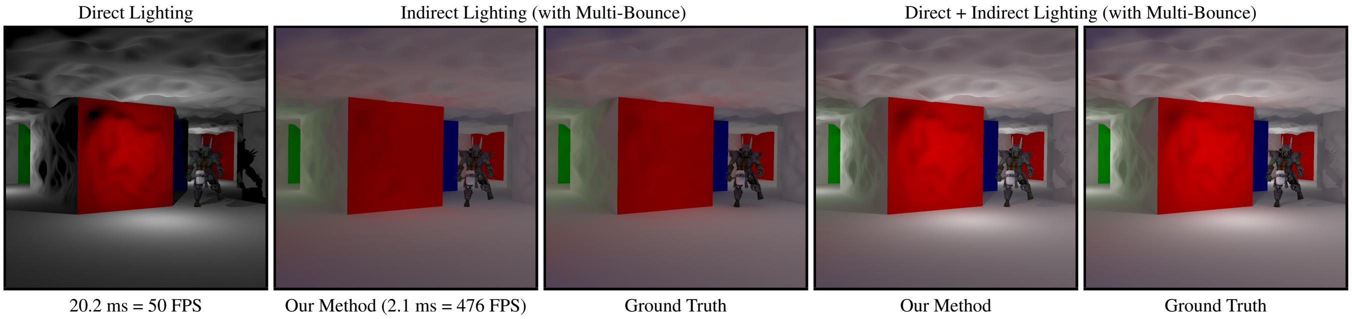


Figure 1: *Indirect light computed in reduced subspaces for a cave with 19 blocks and 4 lights. We derive low-dimensional transport operators, on simple proxy shapes, that are warped and combined at run-time, at* > 475 *FPS on high-end GPUs and* > 45 *FPS on mobile platforms, and can model indirect light at surfaces (with detailed normal variation) and within volumes of large-scale scene geometry.* © *2011 The Authors.*

Abstract

Many rendering algorithms willingly sacrifice accuracy, favoring plausible shading with high-performance. Modular Radiance Transfer (MRT) models coarse-scale, distant indirect lighting effects in scene geometry that scales from high-end GPUs to low-end mobile platforms. MRT eliminates scene-dependent precomputation by storing compact transport on simple shapes, akin to bounce cards used in film production. These shapes' modular transport can be instanced, warped and connected on-the-fly to yield approximate light transport in large scenes. We introduce a prior on incident lighting distributions and perform all computations in low-dimensional subspaces. An *implicit lighting environment* induced from the low-rank approximations is in turn used to model secondary effects, such as volumetric transport variation, higher-order irradiance, and transport through lightfields. MRT is a new approach to precomputed lighting that uses a novel low-dimensional subspace simulation of light transport to uniquely balance the need for high-performance and portable solutions, low memory usage, and fast authoring iteration.

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

Keywords: Global Illumination, GPU, Interactive

Links: DL PDF

1 Introduction

Indirect illumination increases the realism of computer generated images. The ambient term is a simple inexpensive approximation that does not respond to dynamic lighting. Accurate real-time techniques [Keller 1997] have difficulty scaling to complex scenes and often have significant performance requirements, particularly on modern console and mobile platforms. Techniques that approximate different elements of indirect lighting have been extremely successful in interactive graphics applications. Precomputation techniques used in video games [Chen 2008; Larsson and Halen 2009] tend to assume static scenes and lighting, but suffer from long authoring iteration times and memory requirements. Ambient Occlusion (AO) [Zhukov et al. 1998] captures only salient shading effects. Variants of Precomputed Radiance Transfer (PRT) [Sloan et al. 2002] generate soft lighting results. These techniques are favorable compared to more accurate techniques due to their lower storage and computation costs and the pleasing nature of their approximation. Modular Radiance Transfer (MRT) targets **coarse-scale, distant** indirect lighting in scene geometry, responds plausibly and smoothly to dynamic lighting, has extremely high-performance, and allows fast author iteration.

Our shapes are motivated by bounce cards used in live-action films. These planes only approximate indirect light from geometry in the real-world but offer a high level of control to produce the desired lighting. In digital film production, non-shadow casting lights are commonly used to allow artists to quickly iterate and achieve a desired look. The ease-of-use and controllability of these approximations outweighs their physically incorrect nature.

Our approach is very efficient, uses very little data and a quick, one-time, *scene-independent* precomputation step. It also allows real-time computation of approximate indirect light, and is designed with rapid iteration of light design in mind. We precompute light transport operators (LTOs) for a handful of simple canonical "shapes", then interactively warp and combine these shapes, along with their LTOs, to more complex geometry. These shape proxies are used to model direct-to-indirect transport which is then applied as a light map to the actual scene geometry. The flow of indirect light between proxies is modeled with lightfields, and all computations are performed on very low-dimensional subspaces. MRT results in plausible, dynamic global-illumination effects, rendered at high frame rates with low memory overhead. We design special LTOs for secondary transport effects such as light volumes for dynamic characters and higher-order irradiance for normal mapping. We illustrate our solutions ability to scale from high-end to mobile platforms and, like PRT, to provide smooth results which respond to light change.

ACM Reference Format
Loos, B., Antani, L., Mitchell, K., Nowrouzezahrai, D., Jarosz, W., Sloan, P. 2011. Modular Radiance Transfer. *ACM Trans. Graph. 30*, 6, Article 178 (December 2011), 10 pages. DOI = 10.1145/2024156.2024212 http://doi.acm.org/10.1145/2024156.2024212.

LightSlice: Matrix Slice Sampling for the Many-Lights Problem

Jiawei Ou*
Dartmouth College

Fabio Pellacini†
Dartmouth College

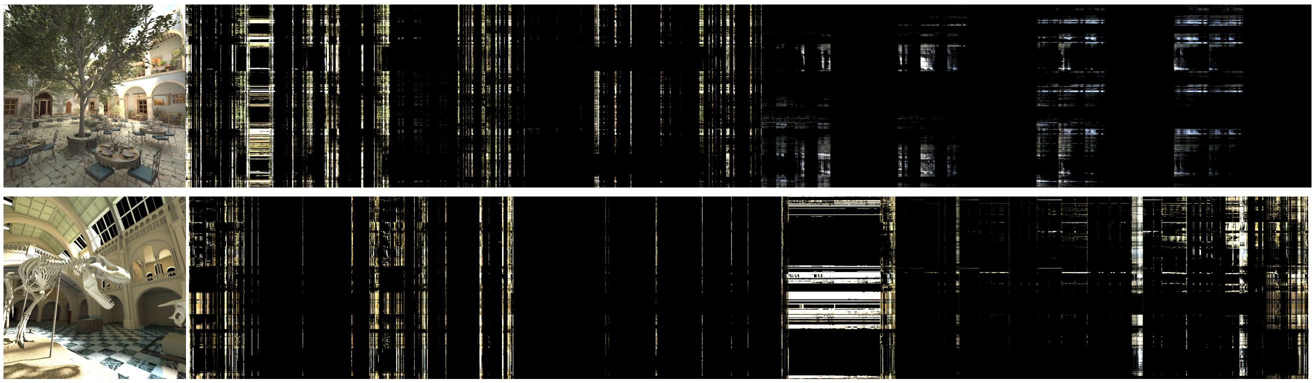


Figure 1: *The light transport matrices of two complex scenes (subsampled from the original). Note the existence of repeating patterns and large areas of near black in the matrices. Our algorithm,* LightSlice*, effectively exploits these typical structures of light transport matrices, and efficiently solves the many-lights problem by seeking locally optimized light clustering for each slice of the light trasnport matrix.*

Abstract

Recent work has shown that complex lighting effects can be well approximated by gathering the contribution of hundreds of thousands of virtual point lights (VPLs). This final gathering step is known as the many-lights problem. Due to the large number of VPLs, computing all the VPLs' contribution is not feasible. This paper presents *LightSlice*, an algorithm that efficiently solves the many-lights problem for large environments with complex lighting. As in prior work, we derive our algorithm from a matrix formulation of the many-lights problem, where the contribution of each VPL corresponds to a column, and computing the final image amounts to computing the sum of all matrix columns. We make the observation that if we cluster similar surface samples together, the slice of the matrix corresponding to these surface samples has significantly lower rank than the original matrix. We exploit this observation by deriving a two-step algorithm where we first globally cluster all lights, to capture the global structure of the matrix, and then locally refine these clusters to determine the most important lights for each slice. We then reconstruct a final image from only these locally-important lights. Compared to prior work, our algorithm has the advantage of discovering and exploiting the global as well as local matrix structure, giving us a speedup of between three and six times compared to state-of-the-art algorithms.

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

Keywords: many light, global illumination, matrix sampling

Links: DL PDF

*e-mail: ouj@cs.dartmouth.edu
†e-mail: fabio@cs.dartmouth.edu

ACM Reference Format
Ou, J., Pellacini, F. 2011. LightSlice: Matrix Slice Sampling for the Many-Lights Problem.
ACM Trans. Graph. 30, 6, Article 179 (December 2011), 8 pages. DOI = 10.1145/2024156.2024213
http://doi.acm.org/10.1145/2024156.2024213.

1 Introduction

Realistic Rendering as Many-Lights. Fast computation of global illumination in large scenes with a complex lighting configuration is still a challenging problem in computer graphics. Many methods have been proposed to compute fast global illumination solutions, e.g. bidirectional path tracing and photon mapping. (see [Pharr and Humphreys 2010] for a recent review). In this paper, we focus on computing images using a variant of instant global illumination [Keller 1997], where direct and indirect illumination are approximated by converting the original light sources into a large number of virtual point lights (VPLs) distributed across the entire scene. In this model, computing a global illumination solution is equivalent to computing an image lit solely by a large number of point light sources, i.e. the many-lights problem. Prior work in offline, high-fidelity rendering [Hašan et al. 2007; Walter et al. 2005] has shown that for scenes with diffuse and low gloss materials, hundreds or thousands of VPLs effectively approximate complex direct and indirect illumination effects, while having the advantage of treating both equally within the same algorithm framework. VPLs have also been used in real-time applications, where they handle a smaller number of lights at the price of the accuracy of approximation [Ritschel et al. 2009]. VPLs have also found much use in feature film production [Christensen 2008]. In this paper, we focus on high-fidelity rendering in complex environments rather than interactive applications.

Matrix Intepretation of Many-Lights. It is useful to consider an alternative interpretation of the many-lights problem as a matrix sampling problem. Let us arrange all pixels of an image as a long column vector. We can then arrange all columns corresponding to each VPL as a large unknown matrix. Computing the final image amounts to computing the sum of each row in the matrix. Figure 1 shows an example of such matrix. While many-lights algorithms

Practical Filtering for Efficient Ray-Traced Directional Occlusion

Kevin Egan *
Columbia University

Frédo Durand
MIT CSAIL

Ravi Ramamoorthi
University of California, Berkeley

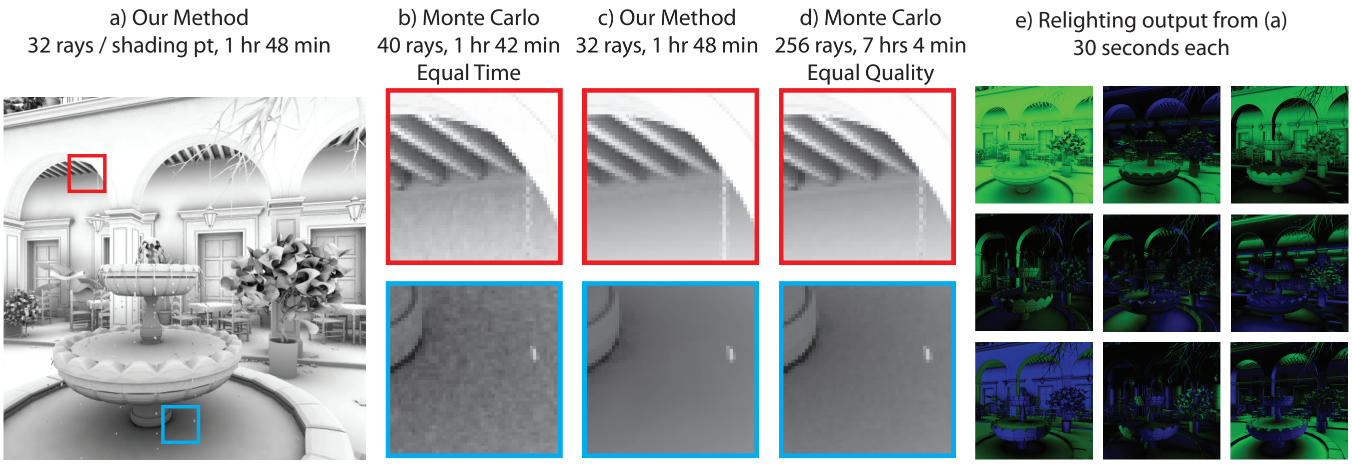


Figure 1: *(a) A visualization of ambient occlusion produced by our method. This scene used 32 samples per shading point, 13 rays in the sparse sampling pass (41 min) and 19 rays in the second pass in areas with contact shadows (1 hr 7 min). Total running time for both passes was 1 hr 48 min. (b) Closeups of Monte Carlo using equal time (40 samples, 1 hr 42 min), noise can be seen. (c) Closeups of our method. (d) Closeups of Monte Carlo with equal quality (256 samples, 7 hrs 4 min). (e) At little extra cost our method can also compute spherical harmonic occlusion for low frequency lighting. While computing (a) our method also outputs directional occlusion information for 9 spherical harmonic coefficients (green is positive, blue is negative).*

Abstract

Ambient occlusion and directional (spherical harmonic) occlusion have become a staple of production rendering because they capture many visually important qualities of global illumination while being reusable across multiple artistic lighting iterations. However, ray-traced solutions for hemispherical occlusion require many rays per shading point (typically 256-1024) due to the full hemispherical angular domain. Moreover, each ray can be expensive in scenes with moderate to high geometric complexity. However, many nearby rays sample similar areas, and the final occlusion result is often low frequency. We give a frequency analysis of shadow light fields using distant illumination with a general BRDF and normal mapping, allowing us to share ray information even among complex receivers. We also present a new rotationally-invariant filter that easily handles samples spread over a large angular domain. Our method can deliver 4x speed up for scenes that are computationally bound by ray tracing costs.

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

Keywords: ambient occlusion, relighting, sampling, frequency analysis

*e-mail:ktegan@cs.columbia.edu

Links: DL PDF

1 Introduction

Modern production rendering algorithms often compute low frequency hemispherical occlusion, where the surrounding environment is approximated to either be a solid white dome (ambient occlusion), or a series of low frequency spherical harmonics. Two different bodies of work related to ambient occlusion were given scientific Academy Awards in 2010 [AcademyAwards 2010], and the movie Avatar used ray-traced ambient and spherical harmonic occlusion for lighting and final rendering [Pantaleoni et al. 2010]. While fully sampling the surrounding illumination at each receiver is the completely accurate way to compute global illumination, these approximations of distant lighting work well in practice. Another advantage is that the ambient occlusion and spherical harmonic calculations are independent of the final lighting environment and can be reused throughout the lighting process.

Ray-traced occlusion is often very expensive to compute due to the large number of incoherent ray casts. The authors of the PantaRay system state that they typically shoot 512 or 1024 rays per shading

ACM Reference Format
Egan, K., Durand, F., Ramamoorthi, R. 2011. Practical Filtering for Efficient Ray-Traced Directional Occlusion. *ACM Trans. Graph. 30*, 6, Article 180 (December 2011), 10 pages.
DOI = 10.1145/2024156.2024214 http://doi.acm.org/10.1145/2024156.2024214.

Progressive Photon Beams

Wojciech Jarosz[1] Derek Nowrouzezahrai[1] Robert Thomas[1] Peter-Pike Sloan[2] Matthias Zwicker[3]

[1]Disney Research Zürich [2]Disney Interactive Studios [3]University of Bern

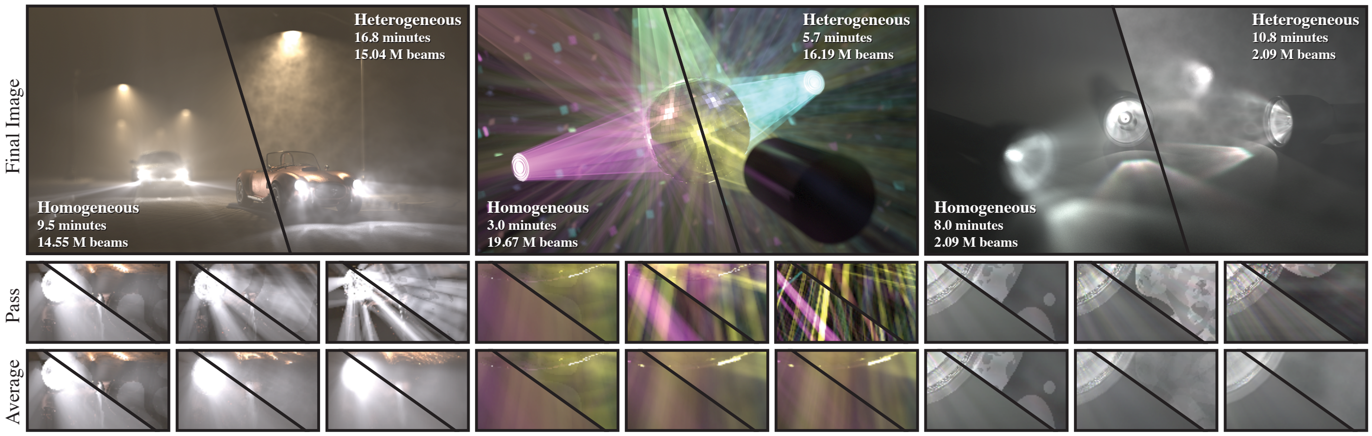


Figure 1: *The* Cars, Disco, *and* FlashLights *scenes rendered using progressive photon beams for both homogeneous and heterogeneous media. We generate a sequence of independent render passes (middle) where we progressively reduce the photon beam radii. This slightly increases variance in each pass, but we prove that the average of all passes (bottom) converges to the ground truth.*

Abstract

We present progressive photon beams, a new algorithm for rendering complex lighting in participating media. Our technique is efficient, robust to complex light paths, and handles heterogeneous media and anisotropic scattering while provably converging to the correct solution using a bounded memory footprint. We achieve this by extending the recent photon beams variant of volumetric photon mapping. We show how to formulate a progressive radiance estimate using photon beams, providing the convergence guarantees and bounded memory usage of progressive photon mapping. Progressive photon beams can robustly handle situations that are difficult for most other algorithms, such as scenes containing participating media and specular interfaces, with realistic light sources completely enclosed by refractive and reflective materials. Our technique handles heterogeneous media and also trivially supports stochastic effects such as depth-of-field and glossy materials. Finally, we show how progressive photon beams can be implemented efficiently on the GPU as a splatting operation, making it applicable to interactive and real-time applications. These features make our technique scalable, providing the same physically-based algorithm for interactive feedback and reference-quality, unbiased solutions.

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

Keywords: participating media, global illumination, photon mapping, photon beams, density estimation

Links: DL PDF Web Video

ACM Reference Format
Jarosz, W., Nowrouzezahrai, D., Thomas, R., Sloan, P., Zwicker, M. 2011. Progressive Photon Beams. *ACM Trans. Graph. 30*, 6, Article 181 (December 2011), 11 pages. DOI = 10.1145/2024156.2024215 http://doi.acm.org/10.1145/2024156.2024215.

1 Introduction

The scattering of light creates stunning visual complexity, in particular with interactions in participating media such as clouds, fog, and even air. Rendering this complex light transport requires solving the radiative transport equation [Chandrasekar 1960] combined with the rendering equation [Kajiya 1986] as a boundary condition.

The gold standard for rendering is arguably computing unbiased, noise-free images. Unfortunately, the only options to achieve this are variants of brute force path tracing [Kajiya 1986; Lafortune and Willems 1993; Veach and Guibas 1994; Lafortune and Willems 1996] and Metropolis light transport [Veach and Guibas 1997; Pauly et al. 2000], which are slow to converge to noise-free images despite recent advances [Raab et al. 2008; Yue et al. 2010]. This becomes particularly problematic when the scene contains so-called SDS (specular-diffuse-specular) or SMS (specular-media-specular) subpaths, which are actually quite common in physical scenes (e.g. illumination due to a light source inside a glass fixture). Unfortunately, path tracing methods cannot robustly handle these situations, especially in the presence of small light sources. Methods based on volumetric photon mapping [Jensen and Christensen 1998; Jarosz et al. 2008] do not suffer from these problems. They can robustly handle S(D|M)S subpaths, and generally produce less noise. However, these methods suffer from bias, which can be eliminated *in theory* by using infinitely many photons, but *in practice* this is not feasible since it requires unlimited memory.

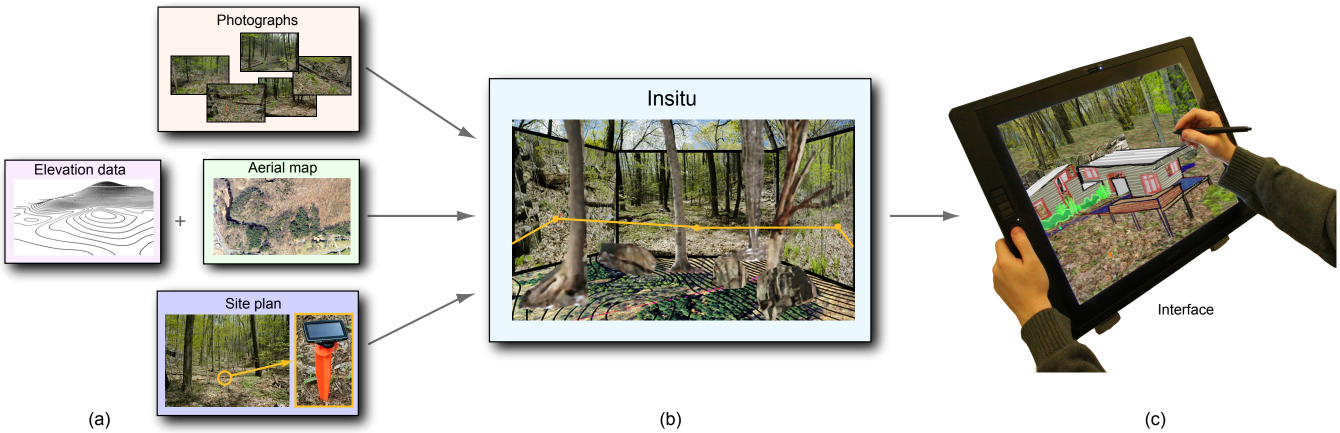
We combine the benefits of these two classes of algorithms. Our approach converges to the gold standard much faster than path tracing, is robust to S(D|M)S subpaths, and has a bounded memory footprint. In addition, we show how the algorithm can be accelerated on the GPU. This allows for interactive lighting design in the presence of complex light sources and participating media. We also obtain reference-quality results at real-time rates for simple scenes containing complex volumetric light interactions. Our algorithm gracefully handles a wide spectrum of fidelity settings, ranging from real-time frame rates to reference quality solutions. This makes it possible to produce interactive previews with the same technique used for a high-quality final render — providing visual consistency, an essential property for interactive lighting design tools.

Our approach draws upon two recent advances in rendering: pho-

Insitu: Sketching Architectural Designs in Context

Patrick Paczkowski† Min H. Kim† Yann Morvan§ Julie Dorsey† Holly Rushmeier† Carol O'Sullivan§

Yale University† Trinity College Dublin§



(a) (b) (c)

Figure 1: *An overview of our approach to presenting context in a lightweight conceptual design system. (a) Input includes site photos, aerial maps, elevation data, and GPS coordinates, from which linked pop-ups and terrain are produced through our novel positioning method, as shown in (b). (c) These representations are used to guide the development of sketches reconciled with the site.*

Abstract

Architecture is design in spatial context. The only current methods for representing context involve designing in a heavyweight computer-aided design system, using a full model of existing buildings and landscape, or sketching on a panoramic photo. The former is too cumbersome; the latter is too restrictive in viewpoint and in the handling of occlusions and topography. We introduce a novel approach to presenting context such that it is an integral component in a lightweight conceptual design system. We represent sites through a fusion of data available from different sources. We derive a site model from geographic elevation data, on-site point-to-point distance measurements, and images of the site. To acquire and process the data, we use publicly available data sources, multidimensional scaling techniques and refinements of recent bundle adjustment techniques. We offer a suite of interactive tools to acquire, process, and combine the data into a lightweight stroke and image-billboard representation. We create multiple and linked *pop-ups* derived from images, forming a lightweight representation of a three-dimensional environment. We implemented our techniques in a stroke-based conceptual design system we call *Insitu*. We developed our work through continuous interaction with professional designers. We present designs created with our new techniques integrated in a conceptual design system.

Keywords: Design, Context, Sketching, Human Factors

Links: DL PDF WEB VIDEO

ACM Reference Format
Paczkowski, P., Kim, M., Morvan, Y., Dorsey, J., Rushmeier, H., O'Sullivan, C. 2011. InSitu: Sketching Architectural Designs in Context. *ACM Trans. Graph. 30*, 6, Article 182 (December 2011), 10 pages. DOI = 10.1145/2024156.2024216 http://doi.acm.org/10.1145/2024156.2024216.

1 Introduction

Computer-aided design (CAD) systems have been extraordinarily successful in design, especially in architecture. Recently there has been considerable interest in 3D modeling systems for early phases in the design of structures. Sketching programs that allow users to rough out three-dimensional definitions from simple strokes and gestures are widely used by architects. Yet even as computers are ubiquitous in the design of the built environment, the existing array of computational aids does not offer assistance in the early conceptual design of structures relative to existing natural and manmade environments – a central concern of architectural design.

The design of architectural structures in situ poses substantive and unique challenges. The spatial fit between architecture and its context is a key concern, as architecture includes exterior and interior space [McHarg 1991] and is never designed in a vacuum. There are three basic relationships between architecture and its surround: *contrast*, *merger*, and *reciprocity*. *Contrast* juxtaposes architecture with the natural context – for example, the relationship of New York's Central Park to the surrounding fabric. *Merger* is the opposite of contrast: a building is designed to appear as a harmonious integral part of the surround. Much of Frank Lloyd Wright's work, including his famous Fallingwater House, shown in Fig. 2, aspired to this condition. Here, the building and context are conceived as one, making it impossible to design independent of the context; such a condition cannot be represented effectively by current CAD systems. *Reciprocity* represents a hybrid condition, in which a building and its surround reflect one another, and enter into a sort of spatial dialog. The considerations and devices used to achieve these various relationships are numerous, and include such factors as massing, geometry, view(s), and scale, to name a few.

There are two current approaches available for representing the surrounding visual context: 3D models and panoramic images. Full 3D models of sites allow a designer to envision designs in multiple views and relative to the context. However, such models are rarely, if ever, used in practice because full-3D models, particularly of landscapes, are difficult to acquire, and the representation is too unwieldy to support conceptual design. Design sketching over pho-

Structure-Preserving Retargeting of Irregular 3D Architecture

Jinjie Lin[1] Daniel Cohen-Or[2] Hao Zhang[3] Cheng Liang[1] Andrei Sharf[4,1] Oliver Deussen[5] Baoquan Chen[1]
[1]SIAT [2]Tel-Aviv University [3]Simon Fraser University [4]Ben-Gurion University [5]University of Konstanz

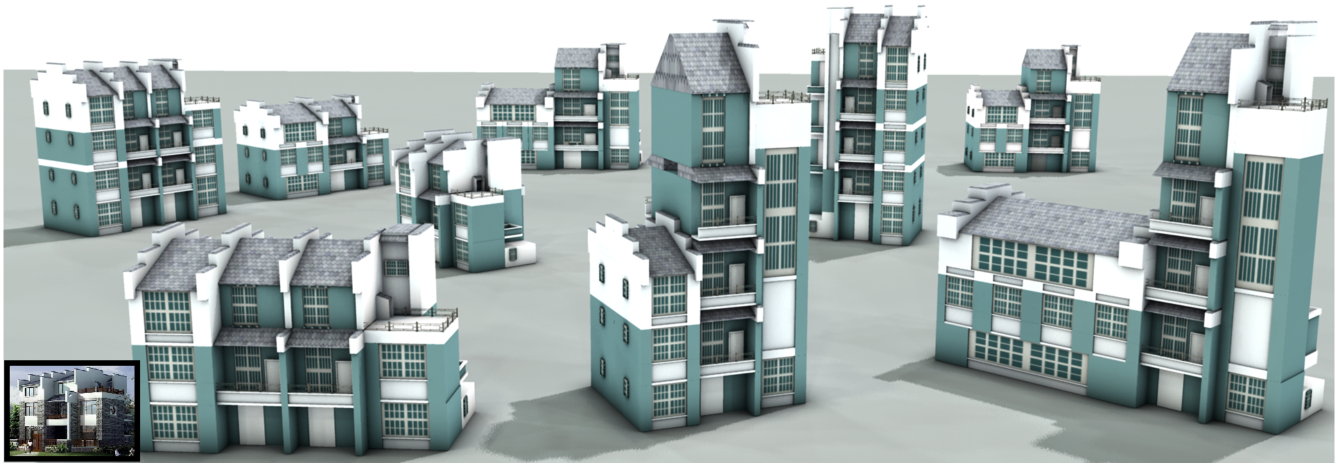


Figure 1: *Retargeting of a real-world irregular 3D architectural model (photo at bottom-left corner) while preserving its structural style.*

Abstract

We present an algorithm for interactive *structure-preserving* retargeting of *irregular* 3D architecture models, offering the modeler an easy-to-use tool to quickly generate a variety of 3D models that resemble an input piece in its structural style. Working on a more global and structural level of the input, our technique allows and even encourages replication of its structural elements, while taking into account their semantics and expected geometric interrelations such as alignments and adjacency. The algorithm performs automatic replication and scaling of these elements while preserving their structures. Instead of formulating and solving a complex constrained optimization, we decompose the input model into a set of sequences, each of which is a 1D structure that is relatively straightforward to retarget. As the sequences are retargeted in turn, they progressively constrain the retargeting of the remaining sequences. We demonstrate interactivity and variability of results from our retargeting algorithm using many examples modeled after real-world architectures exhibiting various forms of irregularity.

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

Keywords: Model retargeting, irregular 3D architecture

Links: DL PDF

1 Introduction

Generating a scene with a variety of models that share a common characteristic or style is a challenging task. In the absence of a concrete description of the target style, a common approach is to synthesize by examples, where some notion of style in the example is preserved in the generated models. A straightforward means of style-preserving synthesis generates models consisting of contents taken from the example while preserving their general relation. A primary example is example-based texture synthesis [Wei et al. 2009] where the imitated styles are mainly of a local nature. Recently, such techniques have been extended to preserve more global structures in images [Risser et al. 2010; Wu et al. 2010].

Works on inverse procedural modeling can be regarded as abstracting the structure present in the examples into a set of rules and then synthesizing novel models based on the rules [Aliaga et al. 2007; Stava et al. 2010; Bokeloh et al. 2010]. For effective rule extraction, dominant presence of regularities in the examples is essential. In a rather loose sense, works on image or model retargeting [Shamir and Sorkine 2009] also represent a form of style-preserving synthesis, where the style is signified by the salient features of the source model. This direction has led to recent work on example-bases synthesis of facade images [Lefebvre et al. 2010].

In this paper, we are interested in retargeting of 3D architectural geometry which preserves the *structural style* of an input piece. Our focus is on the handling of complex structures, especially *irregularities* in the example architectural piece. Specifically, the input model does not exhibit dominant regular grids; it is typically composed of regular substructures that vary in form and are arranged in an irregular manner in 3D space. Our retargeting technique offers a modeler an easy-to-use interactive tool to quickly generate a variety of 3D models possessing a common structural style; see Figure 1. The large collection of models obtained provides a means for

ACM Reference Format
Lin, J., Cohen-Or, D., Zhang, H., Liang, C., Sharf, A., Deussen, O., Chen, B. 2011. Structure-Preserving Retargeting of Irregular 3D Architecture. *ACM Trans. Graph. 30*, 6, Article 183 (December 2011), 10 pages. DOI = 10.1145/2024156.2024217 http://doi.acm.org/10.1145/2024156.2024217.

Adaptive Partitioning of Urban Facades

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

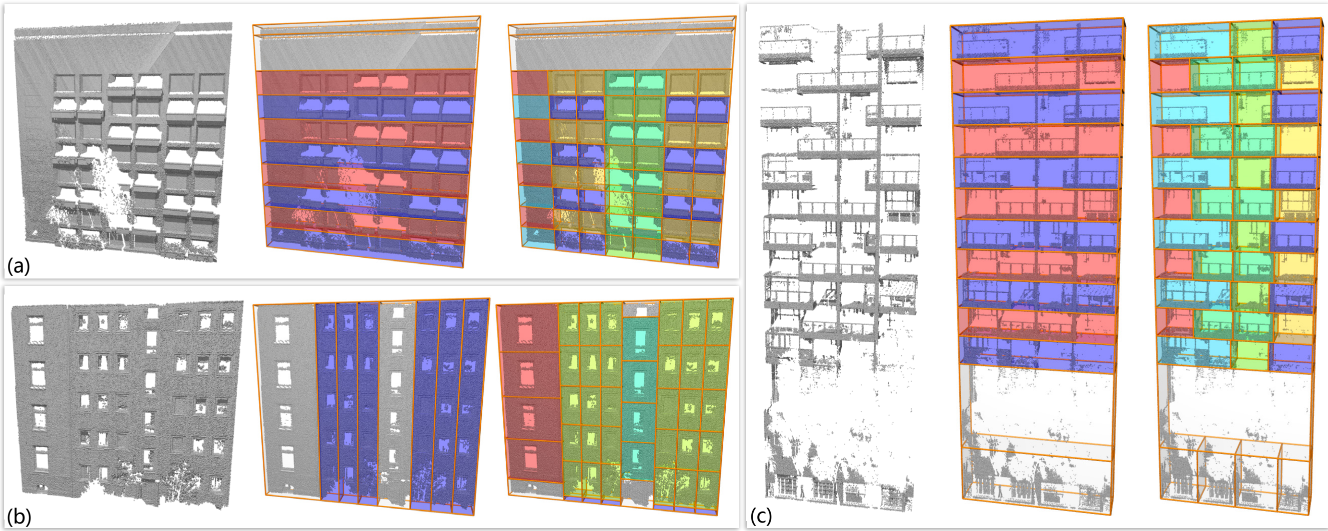


Figure 1: *Given input unorganized point clouds of 3D urban facades, a recursive adaptive partitioning is automatically performed to build a hierarchy of building blocks upon them (from left to right). The splitting direction, number and location of splitting planes are all adaptively determined in each step. Repetitive patterns are indicated by different colors.*

Abstract

Automatically discovering high-level facade structures in unorganized 3D point clouds of urban scenes is crucial for applications like digitalization of real cities. However, this problem is challenging due to poor-quality input data, contaminated with severe missing areas, noise and outliers. This work introduces the concept of *adaptive partitioning* to automatically derive a flexible and hierarchical representation of 3D urban facades. Our key observation is that urban facades are largely governed by concatenated and/or interlaced grids. Hence, unlike previous automatic facade analysis works which are typically restricted to globally rectilinear grids, we propose to automatically partition the facade in an adaptive manner, in which the splitting direction, the number and location of splitting planes are all adaptively determined. Such an adaptive partition operation is performed recursively to generate a hierarchical representation of the facade. We show that the concept of adaptive partitioning is also applicable to flexible and robust analysis of image facades. We evaluate our method on a dozen of LiDAR scans of various complexity and styles, and the image facades from the eTRIMS database and the Ecole Centrale Paris database. A series of applications that benefit from our approach are also demonstrated.

Links: DL PDF WEB

ACM Reference Format
Shen, C., Huang, S., Fu, H., Hu, S. 2011. Adaptive Partitioning of Urban Facades.
ACM Trans. Graph. 30, 6, Article 184 (December 2011), 9 pages. DOI = 10.1145/2024156.2024218
http://doi.acm.org/10.1145/2024156.2024218.

1 Introduction

With the recent advances in LiDAR scanning devices, the acquisition of 3D point clouds from urban buildings is getting more efficient and more convenient. However, the captured point cloud often suffers from severe missing data, noise and outliers, making the reconstruction of architectural models with faithful geometry and topology from such data rather challenging. The key to this problem is to explore and utilize the characteristics of urban scenes as prior knowledge, especially repetition of building elements in facades. To obtain such architectural characteristics, the state-of-the-art works [Zheng et al. 2010; Nan et al. 2010] rely on user assistance, which would be labor intensive for applications like digitalization of real cities.

There exist automatic solutions for discovering facade structures in single- or multi-view facade images [Müller et al. 2007; Xiao et al. 2009; Musialski et al. 2010]. A common assumption made in those works is that facades are inherently governed by *global rectilinear structures*. That is, a facade can be split into building blocks (e.g., windows) by a single rectilinear grid. Although there indeed exist many real facades satisfying such assumption, it is not general enough to handle many other patterns like asymmetric patterns (Figure 1), which are also ubiquitous in urban scenes.

This work aims for a more flexible representation of high-level facade structures, which is based on the key observation: the high-level structure of a facade is largely governed by either a rectilinear grid or a mixture of rectilinear grids by concatenation and/or interlacing. For examples, Figure 2 shows facades with concatenated grids and interlaced grids. Therefore, how to identify different grids of repetitive elements and their relations is our core problem.

To address the problems, the concept of *adaptive partitioning* of urban facades is introduced in this paper (Figure 1). Explicitly searching for repetitive patterns over the entire facade is challenging given poor data quality and complex hidden structures. The underlying

Conjoining Gestalt Rules for Abstraction of Architectural Drawings

Liangliang Nan[1] Andrei Sharf[2] Ke Xie[1] Tien-Tsin Wong[3] Oliver Deussen[4] Daniel Cohen-Or[5] Baoquan Chen[1]

[1] SIAT [2] Ben Gurion Univ. [3] CUHK [4] Konstanz Univ. [5] Tel Aviv Univ.

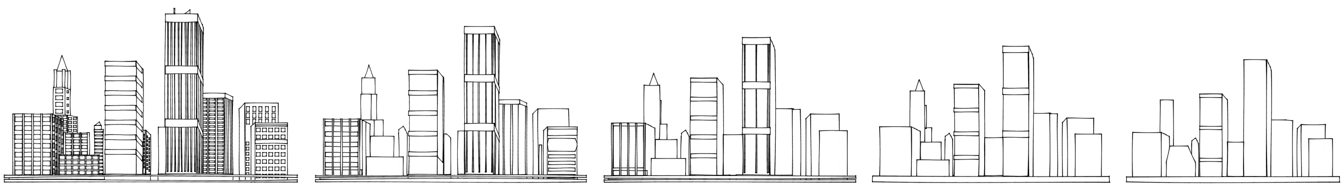


Figure 1: *Simplification of a complex cityscape line-drawing obtained using our Gestalt-based abstraction.*

Abstract

We present a method for structural summarization and abstraction of complex spatial arrangements found in architectural drawings. The method is based on the well-known Gestalt rules, which summarize how forms, patterns, and semantics are perceived by humans from bits and pieces of geometric information. Although defining a computational model for each rule alone has been extensively studied, modeling a conjoint of Gestalt rules remains a challenge. In this work, we develop a computational framework which models Gestalt rules and more importantly, their complex interactions. We apply *conjoining rules* to line drawings, to detect groups of objects and repetitions that conform to Gestalt principles. We summarize and abstract such groups in ways that maintain structural semantics by displaying only a reduced number of repeated elements, or by replacing them with simpler shapes. We show an application of our method to line drawings of architectural models of various styles, and the potential of extending the technique to other computer-generated illustrations, and three-dimensional models.

CR Categories: I.3.5 [Computer Graphics]: Geometric Modeling and Level of Detail—Geometric algorithms;

Keywords: simplification, abstraction, Gestalt

Links: DL PDF

1 Introduction

Artistic imagery, architectural renderings, cartography and games often exploit abstraction to clarify, exaggerate, simplify or emphasize the visual content. Abstraction is a strategy for communicating information effectively. It allows artists to highlight specific visual information and thereby direct the viewer to important aspects of the structure and organization of the scene. In this paper, we present an approach to the abstraction of 2D shapes, in particular those of architectural models. Our approach to abstracting shape directly aims to clarify shape and preserve meaningful structures using Gestalt principles.

The well-known Gestalt principles by Wertheimer [1923], reflect strategies of the human visual system to group objects into forms and create internal representations for them. Whenever groups of visual element have one or several characteristics in common, they get grouped and form a new larger visual object - a *gestalt*. Psychologists have tried to simulate and model these principles, by finding computational means to predict what human perceive as gestalts in images.

The notion of Gestalt is very well-known and widely used in various fields. In particular, it explains the tendency of the human visual recognition to form whole shapes and forms just from bits and pieces of geometric information. Naturally, Gestalt principles have been used in computer vision, primarily in context with object recognition and scene understanding. In computer graphics, Gestalt principles have been applied to a variety of applications, like scene completion [Drori et al. 2003], image and scene abstraction [Wang et al. 2004; Mehra et al. 2009], stroke synthesis [Barla et al. 2006; Ijiri et al. 2008] and emerging images generation [Mitra et al. 2009]. In general, these works rely on discrete Gestalt principles, but none addresses the complex interactions emerging from the multitude of Gestalt principles operating simultaneously.

A difficult problem while dealing with gestalts is the conjoined effect of two or more Gestalt principles operating at the same time on the same site. Modeling gestalts in such cases is especially challenging due to the complexity and ambiguity of the scene. Recently, attempts to discover how grouping principles interact were made in psychology and computer vision [Desolneux et al. 2002; Feldman 2003; Cao et al. 2007; Kubovy and van den Berg 2008]. These works model limited gestalt interactions, by finding computational means which are physiologically plausible. Kubovy and van den Berg [2008] explore the quantification of perceptual groupings formed conjointly by two grouping principles: similarity and proximity. Nevertheless, providing general computational means for modeling the interaction of multiple Gestalt principles remains a difficult challenge.

In this paper, we take a first step in developing a computational model for conjoining Gestalt rules. We model a subset of Gestalt rules and their mutual interaction for abstracting architectural line drawings. We choose to focus on architectural drawings since typically their visual elements are of rather low complexity and their spatial arrangement is strongly biased to the main axes (due to engineering considerations). Hence, architectural drawings consist of prevalent similarities, proximities and regularities among their el-

ACM Reference Format
Nan, L., Sharf, A., Xie, K., Wong, T., Deussen, O., Cohen-Or, D., Chen, B. 2011. Conjoining Gestalt Rules for Abstraction of Architectural Drawings. *ACM Trans. Graph. 30*, 6, Article 185 (December 2011), 10 pages. DOI = 10.1145/2024156.2024219 http://doi.acm.org/10.1145/2024156.2024219.

Polarization Fields: Dynamic Light Field Display using Multi-Layer LCDs

Douglas Lanman[1] Gordon Wetzstein[2] Matthew Hirsch[1] Wolfgang Heidrich[2] Ramesh Raskar[1]
[1]MIT Media Lab [2]University of British Columbia

Figure 1: *Dynamic light field display using polarization field synthesis with multi-layered LCDs. (Left) We construct an optically-efficient polarization field display by covering a stack of liquid crystal panels with crossed linear polarizers. Each layer functions as a polarization rotator, rather than as a conventional optical attenuator. (Right, Top) A target light field. (Right, Bottom) Light fields are displayed, at interactive refresh rates, by tomographically solving for the optimal rotations to be applied at each layer. (Middle) A pair of simulated views is compared to corresponding photographs of the prototype on the left and right, respectively. Inset regions denote the relative position with respect to the display layers, shown as black lines, demonstrating objects can extend beyond the display surface.*

Abstract

We introduce polarization field displays as an optically-efficient design for dynamic light field display using multi-layered LCDs. Such displays consist of a stacked set of liquid crystal panels with a single pair of crossed linear polarizers. Each layer is modeled as a spatially-controllable polarization rotator, as opposed to a conventional spatial light modulator that directly attenuates light. Color display is achieved using field sequential color illumination with monochromatic LCDs, mitigating severe attenuation and moiré occurring with layered color filter arrays. We demonstrate such displays can be controlled, at interactive refresh rates, by adopting the SART algorithm to tomographically solve for the optimal spatially-varying polarization state rotations applied by each layer. We validate our design by constructing a prototype using modified off-the-shelf panels. We demonstrate interactive display using a GPU-based SART implementation supporting both polarization-based and attenuation-based architectures. Experiments characterize the accuracy of our image formation model, verifying polarization field displays achieve increased brightness, higher resolution, and extended depth of field, as compared to existing automultiscopic display methods for dual-layer and multi-layer LCDs.

Keywords: computational displays, light fields, automultiscopic 3D displays, tomography, multi-layer LCDs

Links: DL PDF WEB VIDEO DATA CODE

ACM Reference Format
Lanman, D., Wetzstein, G., Hirsch, M., Heidrich, W., Raskar, R. 2011. Polarization Fields: Dynamic Light Field Display using Multi-Layer LCDs. *ACM Trans. Graph. 30*, 6, Article 186 (December 2011), 9 pages. DOI = 10.1145/2024156.2024220 http://doi.acm.org/10.1145/2024156.2024220.

1 Introduction

The emergence of consumer, glasses-based stereoscopic displays has renewed interest in glasses-free automultiscopic alternatives. Manufacturers are beginning to offer such displays, primarily using two technologies: parallax barriers [Ives 1903] and integral imaging [Lippmann 1908]. These approaches have well-documented limitations compared to stereoscopic displays: decreased resolution, potentially reduced brightness, and narrow depths of field (objects separated from the display appear blurred). Alternatives are being pursued, spanning volumetric to holographic systems; yet, particularly for mobile devices, a display is required that leverages existing or emerging spatial light modulation technologies compatible with thin form factors and having minimal power consumption.

We are inspired by systems that address these issues using well-developed LCD technology. Jacobs et al. [2003] demonstrate dual-layer LCDs can be operated as parallax barriers, allowing full-resolution 2D content and 3D modes with reduced resolution and brightness. Lanman et al. [2010] increase the optical efficiency of dual-layer LCDs with content-adaptive parallax barriers, although at the cost of increased computation. Several researchers have considered automultiscopic multi-layer LCDs. Loukianitsa and Putilin [2002; 2006] evaluate three-layer designs. More recently, Gotoda [2010] and Wetzstein et al. [2011] propose tomographically-optimized multi-layer LCDs. Yet, these works share a common architecture: LCDs are stacked such that each layer implements a spatial light modulator that *attenuates* light.

This paper introduces optically-efficient architectures and computationally-efficient algorithms for automultiscopic display using multi-layered LCDs. In contrast to prior work, we operate these layered architectures as *polarization field displays*: constructed by covering multiple liquid crystal panels with a single pair of crossed linear polarizers. Each layer, if properly designed, functions as a *polarization rotator*, rather than as a light attenuator. We propose an efficient tomographic solver, specially-suited to this design, that enables interactive applications. Through polarization field displays we endeavor to leverage existing and emerging LCD technology for practical dynamic light field display.

Computing and Fabricating Multilayer Models

Michael Holroyd
University of Virginia
Disney Research Zürich

Ilya Baran
Disney Research Zürich

Jason Lawrence
University of Virginia

Wojciech Matusik
MIT CSAIL
Disney Research Zürich

Abstract

We present a method for automatically converting a digital 3D model into a *multilayer model*: a parallel stack of high-resolution 2D images embedded within a semi-transparent medium. Multilayer models can be produced quickly and cheaply and provide a strong sense of an object's 3D shape and texture over a wide range of viewing directions. Our method is designed to minimize visible cracks and other artifacts that can arise when projecting an input model onto a small number of parallel planes, and avoid layer transitions that cut the model along important surface features. We demonstrate multilayer models fabricated with glass and acrylic tiles using commercially available printers.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Display algorithms; I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Curve, surface, solid, and object representation

Keywords: multilayer models, fabrication, volumetric displays

Links: DL PDF WEB VIDEO

Figure 1: *A real object next to a multilayer model fabricated in acrylic using our proposed algorithm.*

1 Introduction

We describe a method for converting digital 3D models into multilayer models. Our work is inspired by artists like Carol Cohen[1] and Dustin Yellin[2] who reproduce three-dimensional forms by painting on glass or acrylic sheets and then stacking them together (Figure 2). A multilayer model thus consists of a small number of high-resolution 2D images stacked along the 3rd dimension, and creates a natural 3D effect by displaying parts of the object at the appropriate depth over a range of viewing directions.

There is no substitute for the experience of holding and examining a physical 3D object in your hands. Technologies capable of manufacturing 3D objects have seen significant advances in recent years, most notably 3D printing [Dimitrov et al. 2006] and multi-axis milling. However, despite these advances, producing a 3D prototype in full color remains expensive and time-consuming. Furthermore, objects with thin features and disconnected parts (e.g., the tree in Figure 14) cannot be printed at all using existing techniques. In contrast, a multilayer model (as in Figure 1) can be fabricated in minutes and for a fraction of the cost and provides many of the benefits of a physical replica.

The process of converting a 3D surface into a multilayer model requires a number of algorithmic advances in order to produce high-quality output. First, we observe that a naïve projection of the input surface onto multiple parallel planes creates artifacts caused by visible seams or cracks and salient features being split between different layers. We propose a novel algorithm that warps each layer based on the way it is occluded by the layers above it in order to avoid these seams while simultaneously seeking cuts along textureless regions. Second, the shadows that each layers casts onto those below it can undermine the intended 3D effect. We describe a fast method for computing a correction factor that compensates for these shadows. Finally, one also needs to consider the contrast loss caused by light absorption in the embedding medium (e.g., glass or acrylic). We propose a simple measurement process to estimate parameters of an analytic absorption model. Based on this model, we restrict the colors printed on each layer to a reduced color gamut that can be achieved throughout the volume.

We show multiple examples of multilayer models produced with our prototype system. Additionally, we include a comparison of our method to simple alternatives that illustrate the importance of properly handling cracks, seams, self shadowing, and volumetric attenuation. We believe multilayer models can serve as useful rapid prototypes, teaching aids, art, and personalized memorabilia.

[1] carolcohen.com
[2] dustinyellin.com

ACM Reference Format
Holroyd, M., Baran, I., Lawrence, J., Matusik, W. 2011. Computing and Fabricating Multilayer Models. *ACM Trans. Graph. 30*, 6, Article 187 (December 2011), 8 pages. DOI = 10.1145/2024156.2024221 http://doi.acm.org/10.1145/2024156.2024221.

2 Related Work

Volumetric Glass Models A source of inspiration for this work are artists who use stacks of painted glass or acrylic sheets to create three-dimensional forms (Figure 2). This stunning art makes it apparent that they have overcome similar technical challenges to the ones we address in this work. A related computer-driven technique

C1x6: A Stereoscopic Six-User Display for Co-located Collaboration in Shared Virtual Environments

Alexander Kulik*,1 André Kunert*,1 Stephan Beck*,1 Roman Reichel*,1 Roland Blach†,2
Armin Zink‡,3 Bernd Froehlich*,1
1 Bauhaus-Universität Weimar 2 Fraunhofer IAO 3 digital IMAGE

Abstract

Stereoscopic multi-user systems provide multiple users with individual views of a virtual environment. We developed a new projection-based stereoscopic display for six users, which employs six customized DLP projectors for fast time-sequential image display in combination with polarization. Our intelligent high-speed shutter glasses can be programmed from the application to adapt to the situation. For instance, it does this by staying open if users do not look at the projection screen or switch to a VIP high brightness mode if less than six users use the system. Each user is tracked and can move freely in front of the display while perceiving perspectively correct views of the virtual environment.

Navigating a group of six users through a virtual world leads to situations in which the group will not fit through spatial constrictions. Our augmented group navigation techniques ameliorate this situation by fading out obstacles or by slightly redirecting individual users along a collision-free path. While redirection goes mostly unnoticed, both techniques temporarily give up the notion of a consistent shared space. Our user study confirms that users generally prefer this trade-off over naïve approaches.

CR Categories: B4.2 [INPUT/OUTPUT AND DATA COMMUNICATIONS]: Input/Output Devices—Image Display; H5.1 [INFORMATION INTERFACES AND PRESENTATION]: Multimedia Information Systems—Artificial, augmented, and virtual realities

Keywords: Virtual Reality, Display Technology, Collaboration

Links: DL PDF WEB

1 Introduction

3D television sets and 3D cinemas display only a single stereoscopic image stream, and thus there is only a single location from which a person observes a perspectively correct view of the displayed scenes. All of the other spectators perceive the 3D scene more or less distorted, which inhibits precise spatial perception of the displayed geometry. While this may not matter much in movie theaters, these distortions significantly hamper the acceptance of 3D technology in many other application areas. In order to compensate for this, each user must be provided with individual stereoscopic image pairs, which are rendered for the exact position of the user in front of a display. While the computing power to generate multiple views of interactive content is available, the display technology for presenting large individual stereoscopic images for multiple users is still lacking

We developed the C1x6, a new projection-based stereoscopic display for six users (Figure 1). Our system consists of six customized DLP projectors, each of which projects six fast time-sequential images in one of the primary colors. By differently polarizing the light output of the first set of three single color projectors (red, green, blue) than those of the second set, we are able to project twelve separable full-color images onto a projection screen. Our intelligent high-speed shutter glasses can be fully controlled from the application level. This feature is used to keep the glasses open if users look away from the screen or for supporting a VIP high brightness mode if less than six users are involved. We developed the software and hardware infrastructure to generate, warp and feed the stereoscopic images for the six tracked users into the projectors.

Multi-user displays enable co-located collaborative work in shared virtual environments. For collaborative design reviews we developed the Spheron, an input device which makes interactions transparent to other co-located users. However, when navigating a group of users through a virtual building, many situations arise in which there is not enough space to place the users in the virtual world in the same way as they are positioned relative to each other in the real world. This problem did not exist in common projection-based virtual reality systems, where all the observers share the same perspective as the head-tracked navigator. Therefore, we present several approaches to facilitate group navigation in such situations by avoiding collisions of group members with surrounding objects such as walls and other obstacles.

The main contributions of our work fall into three areas:

- A synchronized DLP projector array that is capable of displaying twelve high resolution (1920x1200) full color image streams at 360Hz—60Hz per user. Left and right eye images are simultaneously projected and separated by polarization.
- Application-level programmable shutter glasses consisting of double-cell liquid crystals that enable intelligent shutter control and provide fast switching speeds as well as high-contrast.
- Augmented group navigation techniques that avoid collisions when traveling through narrow spaces. Our user study reveals that these techniques are preferred over naïve approaches.

Besides these central contributions, significant amounts of engineering are necessary to build and run such a complex system, including a custom digital video-multiplexing hardware, synchronization of all the components and real-time color convergence through image warping. Measurements show that our six-user system achieves almost the same brightness per user as a stereoscopic single-user display based on the same type of projector would.

*e-mail:<first name>.<last name>@uni-weimar.de
†email:roland.blach@iao.fraunhofer.de
‡email:zink@digital-image.de

ACM Reference Format
Kulik, A., Kunert, A., Beck, S., Reichel, R., Blach, R., Zink, A., Froehlich, B. 2011.
C1x6: A Stereoscopic Six-User Display for Co-located Collaboration in Shared Virtual Environments.
ACM Trans. Graph. 30, 6, Article 188 (December 2011), 11 pages. DOI = 10.1145/2024156.2024222
http://doi.acm.org/10.1145/2024156.2024222.

OSCAM - Optimized Stereoscopic Camera Control for Interactive 3D

Thomas Oskam [1,2] Alexander Hornung [2] Huw Bowles [3,4] Kenny Mitchell [3,4] Markus Gross [1,2]

[1] ETH Zurich [2] Disney Research Zurich [3] Black Rock Studio [4] Disney Interactive Studios

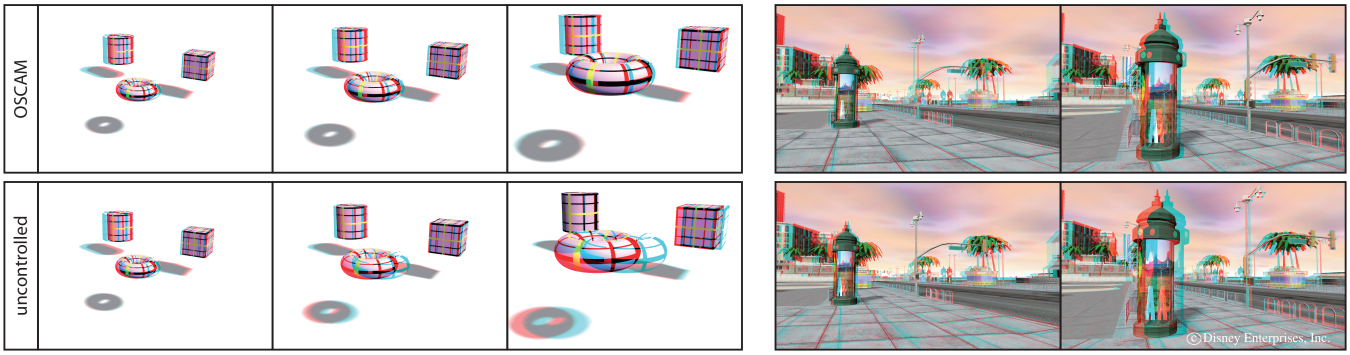


Figure 1: *Two stereoscopic shots of the camera moving towards objects. Our method keeps a constant target depth range when moving close to the objects. Uncontrolled stereoscopy, in contrast, can cause large disparities and destroy stereoscopic perception.*

Abstract

This paper presents a controller for camera convergence and interaxial separation that specifically addresses challenges in *interactive* stereoscopic applications like games. In such applications, unpredictable viewer- or object-motion often compromises stereopsis due to excessive binocular disparities. We derive constraints on the camera separation and convergence that enable our controller to automatically adapt to any given viewing situation and 3D scene, providing an exact mapping of the virtual content into a comfortable depth range around the display. Moreover, we introduce an interpolation function that linearizes the transformation of stereoscopic depth over time, minimizing nonlinear visual distortions. We describe how to implement the complete control mechanism on the GPU to achieve running times below 0.2ms for full HD. This provides a practical solution even for demanding real-time applications. Results of a user study show a significant increase of stereoscopic comfort, without compromising perceived realism. Our controller enables 'fail-safe' stereopsis, provides intuitive control to accommodate to personal preferences, and allows to properly display stereoscopic content on differently sized output devices.

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

Keywords: stereoscopic 3D, disparity control, real-time graphics, games, interactive 3D

Links: DL PDF

ACM Reference Format
Oskam, T., Hornung, A., Bowles, H., Mitchell, K., Gross, M. 2011. OSCAM - Optimized Stereoscopic Camera Control for Interactive 3D. *ACM Trans. Graph. 30*, 6, Article 189 (December 2011), 8 pages. DOI = 10.1145/2024156.2024223 http://doi.acm.org/10.1145/2024156.2024223.

1 Introduction

Stereoscopic content creation, processing, and display has become a pivotal element in movies and entertainment, yet the industry is still confronted with various difficult challenges. Recent research has made substantial progress in some of these areas [Lang et al. 2010; Koppal et al. 2011; Didyk et al. 2011; Heinzle et al. 2011]. Most of these works focus on the classical production pipeline, where the consumer views ready-made content that has been optimized in (post-) production to ensure a comfortable stereoscopic experience. See Tekalp et al. [2011] for an overview.

In interactive applications that create stereoscopic output in real-time, one faces a number of fundamentally different challenges [Gateau and Neuman 2010]. For example, in a first-person game where the player is in control of the view, a simple collision with a wall or another object will result in excessive disparities that cause visual fatigue or destroy stereopsis (see Figure 1). In order to guarantee proper stereoscopy, one needs a controller that adjusts the range of disparities to the viewer's preferences. An example for such a controller is the work of Lang et al. [2010] which, however, has been designed for *post-capture* disparity range adaptation using complex image-domain warping techniques. In a game environment where the stereoscopic output is *created and displayed* in real-time, it is advisable to optimize the stereoscopic rendering parameters, i.e., camera convergence and interaxial separation, and to avoid computationally expensive solutions.

The problem can be formulated as one of controlling perceived depth. We use the term 'perceived depth' in the geometrical sense, where the distances reconstructed by the viewer are dominated by the observed screen disparities. Even though there are other important cues such as vertical size, focus that influence perceived depth [Backus et al. 1999; Watt et al. 2005], the work of Held and Banks [2008] showed that the geometrical approach is a valid approximation. The range of perceived depth around the screen that can be viewed comfortably is generally referred to as the *comfort zone*, and is defined as the range of positive and negative disparities that can be comfortably watched by each individual viewer [Smolic et al. 2011; Shibata et al. 2011]. Therefore, we are looking for an exact mapping of a specific range of distances in the scene into this depth volume around the screen. In the course of this article, we will refer to this volume as the *target depth range*. While we concentrate on the control of the mapping between the virtual and real space there exists prior work on how to derive a comfortable target depth range [Woods et al. 1993; Shibata et al. 2011]

Multi-Perspective Stereoscopy from Light Fields

Changil Kim[1,2] Alexander Hornung[2] Simon Heinzle[2] Wojciech Matusik[2,3] Markus Gross[1,2]

[1]ETH Zurich [2]Disney Research Zurich [3]MIT CSAIL

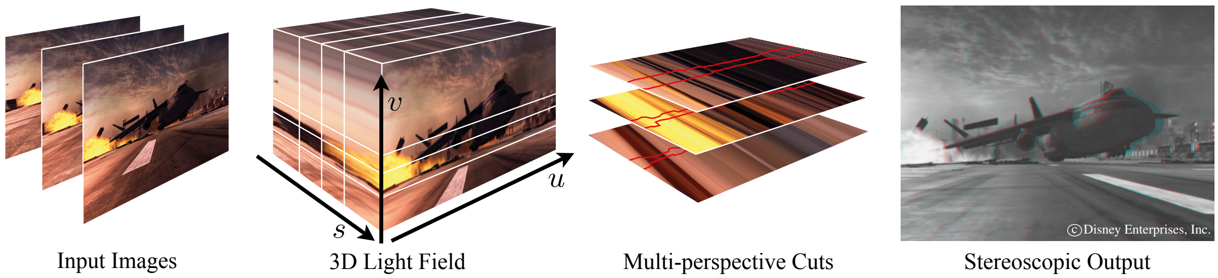


Figure 1: *We propose a framework for flexible stereoscopic disparity manipulation and content post-production. Our method computes multi-perspective stereoscopic output images from a 3D light field that satisfy arbitrary prescribed disparity constraints. We achieve this by computing piecewise continuous cuts (shown in red) through the light field that enable per-pixel disparity control. In this particular example we employed gradient domain processing to emphasize the depth of the airplane while suppressing disparities in the rest of the scene.*

Abstract

This paper addresses stereoscopic view generation from a light field. We present a framework that allows for the generation of stereoscopic image pairs with per-pixel control over disparity, based on multi-perspective imaging from light fields. The proposed framework is novel and useful for stereoscopic image processing and post-production. The stereoscopic images are computed as piecewise continuous cuts through a light field, minimizing an energy reflecting prescribed parameters such as depth budget, maximum disparity gradient, desired stereoscopic baseline, and so on. As demonstrated in our results, this technique can be used for efficient and flexible stereoscopic post-processing, such as reducing excessive disparity while preserving perceived depth, or retargeting of already captured scenes to various view settings. Moreover, we generalize our method to multiple cuts, which is highly useful for content creation in the context of multi-view autostereoscopic displays. We present several results on computer-generated content as well as live-action content.

Keywords: stereoscopy, light field, multi-perspective imaging, autostereoscopic display, post-production

Links: DL PDF

1 Introduction

Three-dimensional stereoscopic television, movies, and games have been gaining more and more popularity both within the entertainment industry and among consumers. An ever increasing amount of content is being created, distribution channels including live-broadcast are being developed, and stereoscopic monitors and TV sets are being sold in all major electronic stores. With novel generations of autostereoscopic and multi-view autostereoscopic displays even glasses-free solutions become available to the consumer.

However, the task of creating convincing yet perceptually pleasing stereoscopic content remains difficult. This is mainly because post-processing tools for stereo are still underdeveloped, and one often has to resort to traditional monoscopic tools and workflows, which are generally ill-suited for stereo-specific issues [Mendiburu 2009]. This situation creates an opportunity to rethink the whole post-processing pipeline for stereoscopic content creation and editing. In the past the computer graphics community has greatly contributed to the development of novel tools for image and video processing. One particular example in the context of this work is the recent progress on light field capture and processing, which enables *post-acquisition* content modification such as depth-of-field, focus, or viewpoint changes. A variety of prototypes for light field acquisition have been developed [Adelson and Wang 1992; Yang et al. 2002; Ng et al. 2005; Wilburn et al. 2005; Georgiev et al. 2006; Veeraraghavan et al. 2007] such that we can expect plenoptic cameras to become available in the near future. However, the concept of post-acquisition control and editing is missing in stereoscopic post-processing.

The main cue responsible for stereoscopic scene perception is binocular parallax (or binocular disparity) and therefore tools for its manipulation are extremely important. One of the most common methods for controlling the amount of binocular parallax is based on setting the baseline, or the inter-axial distance, of two cameras prior to acquisition. However, the range of admissible baselines is quite limited since most scenes exhibit more disparity than humans can tolerate when viewing the content on a stereoscopic display. Reducing baseline decreases the amount of binocular disparity; but it also causes scene elements to be overly flat. The second, more sophisticated approach to disparity control requires remapping image disparities (or remapping the depth of scene elements), and then re-synthesizing new images. This approach has considerable disadvantages as well; for content captured with stereoscopic camera rigs, it typically requires accurate disparity computation and hole filling of scene elements that become visible in the re-synthesized views. For computer-generated images, changing the depth of the underlying scene elements is generally not an option, because changing the 3D geometry compromises the scene composition, lighting calculations, visual effects, etc [Neuman 2010].

In this paper we propose a novel concept for stereoscopic post-production to resolve these issues. The main contribution is a

ACM Reference Format
Kim, C., Hornung, A., Heinzle, S., Matusik, W., Gross, M. 2011. Multi-Perspective Stereoscopy from Light Fields. *ACM Trans. Graph. 30*, 6, Article 190 (December 2011), 10 pages. DOI = 10.1145/2024156.2024224 http://doi.acm.org/10.1145/2024156.2024224.