“Intelligent 3D Shape Modeling and Image Manipulation“

Olga Sorkine-Hornung

Swiss Federal Institute of Technology, ETH Zuerich

Intelligent 3D Shape Modeling and Image Manipulation

Olga Sorkine-Hornung

Abstract

Computer Graphics and in particular 3D geometric modeling is familiar to everyone from computer-animated films and video games. In fact, digital 3D shape modeling is inherent to our daily lives, even if we don't notice it: practically every manufactured object around us was first designed on a computer, be it a car, a table, a toy or a building. Modern 3D scanning technology enables to capture 3D geometry, much like digital photography or music recording. However, current methods and software for shape modeling and design are tedious and time-consuming to use, they require prolonged training and high professional skills, precluding them from being used by the wider public and causing a significant productivity bottleneck in the industry. In my research, I investigate the underlying mathematics and algorithms, as well as user interfaces that could make the creation of 3D geometry intuitive, accessible and efficient, for professional artists and designers and for the general public alike.

I have developed a novel geometric representation called “differential coordinates”, which extracts and represents local surface details in order to efficiently model and store 3D shapes digitally. The local differential properties turn out to be important for the human perception, and for the function of 3D shape. Based on this representation, together with my collaborators, we have proposed several highly efficient algorithms that enable to easily create, edit and animate digital 3D shapes in real time on commodity hardware, even when the shapes have high resolution and geometric complexity. The discrete differential approach is also beneficial for many image and video processing operations. The current article describes the concept and its applications in the various domains of Computer Graphics.

1. Introduction

My research focuses on two major fields in computer graphics: the processing of digital geometric and visual data, namely 3D shapes, and images and video. One of the core challenges in these fields is to enable higher-level, semantically meaningful processing and modeling of the underlying low-level discrete data. For example, the intuitive manipulation of complex 3D shapes requires an understanding of geometrical features, their relationships, and corresponding high-level editing operations. Likewise, advanced image processing tasks require algorithms that are aware of the content of an image, such as visually important objects or structure. In my past and current research, I have made contributions to these fields, with an emphasis on developing fundamental mathematical concepts, efficient and robust algorithms and effective interactive interfaces, with the ultimate goal to develop intelligent and accessible tools, both for professional experts and ordinary users. This research work is detailed below with an outlook on challenges and directions for future exploration.

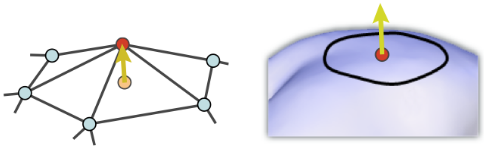
2. Digital geometry processing and geometric modeling

Modeling and manipulation of digital geometric data is essential in many domains, such as engineering design and simulation, product prototyping and manufacturing, surgical planning, prosthesis design, dentistry applications, architecture, geology, computer animated movies and games in the entertainment industry and more. The corresponding digital geometry processing pipeline conceptually consists of three main components:

(i) Shape generation: obtaining geometric data by shape creation or acquisition;
(ii) Processing and optimization: representation, denoising and filtering, parameterization, compression;
(iii) Editing: modification of the geometric data to fit design goals and needs of the end application.

My current and future research agenda contributes to each of the aforementioned components.

Shape representation



$\delta_i = \frac{1}{d_i}\sum_{j\in N(i)}(\mathbf{v}_i - \mathbf{v}_j)$ $\frac{1}{|\gamma|}\int_{\mathbf{v}\in\gamma}(\mathbf{v}_i - \mathbf{v})dl(\mathbf{v})$

Figure 1. The vector of the differential coordinates at a vertex approximates the local shape characteristics of the surface: the normal direction and the mean curvature. Left: discrete differential coordinates. Right: in the limit as the curve length vanishes, this integral converges to the continuous Laplace-Beltrami operator applied to the surface coordinates.

The most basic prerequisite for any kind of geometry processing and editing is the conversion of acquired 3D data samples into a suitable representation. I have explored differential surface representations that are based on polygonal meshes and augment them with information about local geometric detail (see Figure 1). Such representations enable a large variety of detail-preserving shape processing and modeling operations, which are necessary when dealing with complex shapes, such as scans of humans, statues and artefacts, manually designed characters, etc. Differential representations have been applied to shape editing and deformation, optimization, interpolation, compression and approximation, to name a few; the following surveys are now used as standard references for the field [2, 21, 23]. One goal of my current and future research is to continue this work and create *semantically-enabled* shape representations that seamlessly combine both low-level local detail and high-level global features and structure.

Interactive shape creation and editing

The design of intuitive interactive tools to modify the shape of a digital 3D object poses two central challenges:

(i) fast and robust algorithms that implement complex deformation models in real-time;
(ii) effective interfaces that allow versatile manipulation, yet which are simple to use and hide the underlying complexity from the user.

I am interested in both aspects, especially in the context of polygonal meshes.

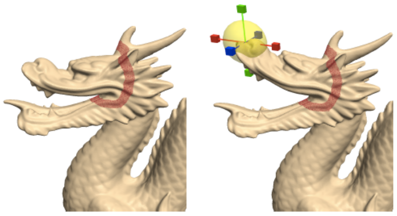


Figure 2. Laplacian mesh editing [26]. The red belt bounds the region of interest; editing is performed by manipulating the handle object represented by the yellow sphere.

Shape deformation. Past research by myself and others concentrated on the "low-level" deformation problem, interpreting intuitive shape modeling behavior as physically plausible deformation similar to real-world materials. I proposed detail-preserving deformation techniques based on differential representations [12,13,26] that employ robust linear variational optimization to compute the deformed surface given user-defined boundary constraints (see Figure 2). The challenge of the linear variational approach is to overcome the linearization errors, which exhibit themselves in local transformations (rotations) of the surface; I have proposed several techniques to deduce plausible fields of local transformations that result in natural surface behavior (see, e.g., Figure 3). Thanks to the use of sparse direct solvers the user can interactively manipulate even highly complex shapes in the order of 105 polygons. This class of algorithms has also been successfully used for animation of articulated shapes [28,32]. The results of this work and follow-up research by others are summarized in our survey [2]. Our work also enabled interactive nonlinear techniques [1,22]; in particular, I have looked at optimization techniques to compute as-rigid-as-possible surface deformations robustly in real-time [22].

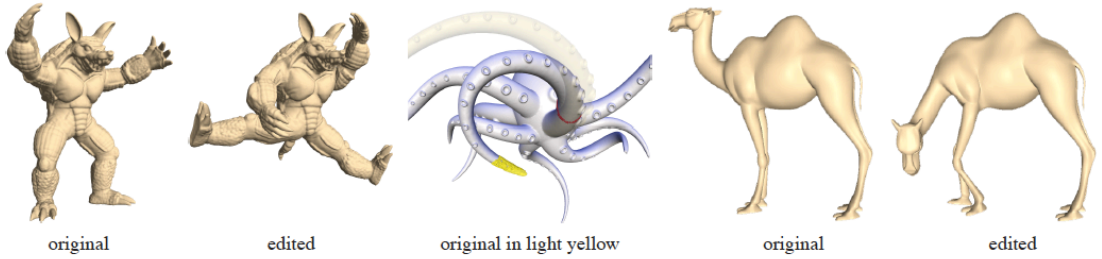


Figure 3. Some editing results using the rotation-invariant representation of [13]. Note the large rotational deformations achieved and the preservation of local surface details.

The low-level deformation problem now largely studied and understood, the next important research challenge is devising *semantically-intuitive* shape modeling algorithms that incorporate high-level knowledge and automatically detect and preserve features and semantic structure during editing operations. This would relieve users from the time consuming and cognitively challenging specification of multiple modeling constraints to manually uphold structure and function of the edited objects, and allow focusing on the specific design goals.

In the FiberMesh system [17], developed in collaboration with the University of Tokyo, shapes can be created from scratch completely by means of feature curves. With the iWires system [8], my colleagues and I showed a first proof-of-concept for semantic editing of man-made shapes, using a simple but powerful set of heuristics for feature representation based on sharp feature curves and their geometric relationships such as symmetry and co-planarity (see Figure 4).

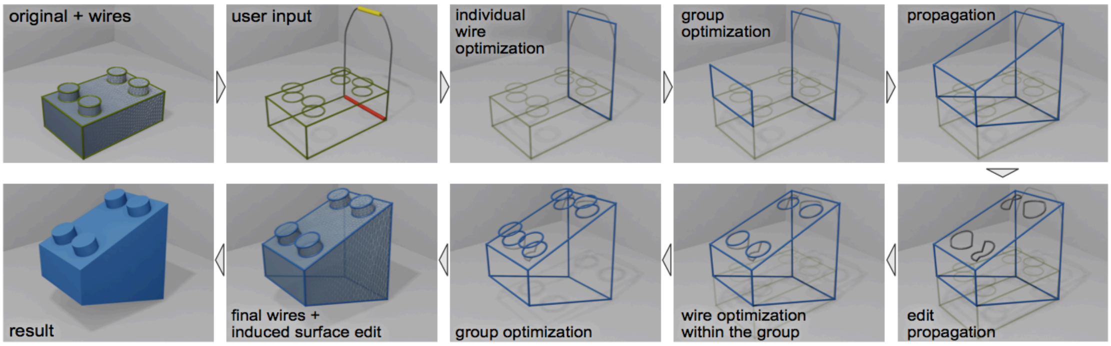


Figure 4. Deforming a Lego model using iWIRES [8]. Given the model geometry, we extract wires (shown in green), learn their individual characteristics and mutual relations, and preserve them when the model undergoes deformation. The user-prescribed edit, defined by the yellow and red handles, is mapped to one of the wires (in gray). We optimize the individual wire (as shown in blue), propagate the edit operation to other wires (gray), and perform a multi-stage constrained optimization to enforce the individual properties and mutual wire relations. The resulting wire positions, shown in blue, determine the final edited shape.

I seek to develop the above direction further into a principled approach that consistently defines and optimizes features and relationships that are meaningful for a given class of shapes. The basic premise is that the important shape features that define its essence can be mathematically described as salient structures of an appropriate function on the shape. Learning the particular characteristic function and the relevant relationships between the features is domain-dependent and can be done by incorporating a priori domain knowledge, training on examples and active learning of the meaningful design space. The required mathematical concepts and algorithms necessitate truly interdisciplinary research, a great opportunity for collaboration and bringing together geometry, scientific computing, machine learning and perceptual science. One of the most important outcomes of this research in the long run would be revolutionizing life cycle management: it will eliminate tedious manual labor required for parametric descriptions of objects, enable re-use of legacy geometric data that lacks construction history and allow working with acquired unstructured geometric data. Moreover, intuitive semantics-driven modeling will make 3D content creation accessible to everyday use, just as 2D paint programs are familiar to most computer users and are even part of basic education in schools.

Interfaces. I have been working on direct surface manipulation and sketching interfaces, where the user can intuitively grab and deform any part of a surface, and control shapes via oversketching of feature curves such as silhouettes, ridges and suggestive contours [8, 17, 18]; sketch-based deformation was also applied to the image domain [6]. See Figure 5 for an example of sketch-based 3D deformation.

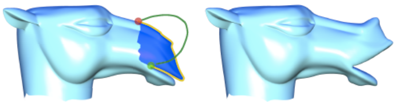


Figure 5. Sketch-based editing of silhouettes [18]. The sketched curve hints at the new desired shape of the silhouette. By weakly weighting the sketch constraints in the reconstruction process, the details of the shape are preserved, and the deformation follows the sketch.

I have also begun to explore the field of expressive modeling in the work on 3D collage creation [9], and would like to research automatic aids for expressiveness deeper. Additionally, in current and future work, I plan to extend the sketch-based modeling metaphor to other domains such as animation and simulation control, and apply novel input metaphors, such as multitouch and affordable pressure-sensitive devices, as well as explore 3D output devices (stereoscopic displays and virtual reality) for both precise and high-level concept modeling tasks.

Computing challenges. While mesh modeling is already quite commonplace in the entertainment industry, strict requirements on efficiency, robustness and precision have hindered its establishment in engineering applications. To address these issues I am working on robust numerical solutions to variational or PDE-based geometric optimization. Together with my colleagues at INRIA I have explored out-of-core freeform modeling that allows editing huge data sets that do not fit in main memory [3]. Furthermore, my group at the Courant Institute is currently addressing the problem of generating surfaces with high-order smoothness for which the traditional FEM solutions are impractical. An additional untapped challenge for practical modeling is collision detection and handling. Based on my previous research [11] I plan to handle this problem by exploiting the high coherence and temporal predictability that is inherent to the interactive modeling setting.

Mesh processing and optimization

I have been working on several fundamental aspects of digital geometry processing, with an emphasis on efficient geometric optimization algorithms.

Denoising, filtering and shape optimization. Denoising and smoothing of measured geometric data is essential for accurate geometry analysis and processing. Our work on Laplacian mesh optimization [16] provides a feature-aware, non-iterative smoothing method that optimizes both the shape and the mesh quality while observing the curvature distribution of the surface. In ongoing work I am developing filtering methods that explicitly preserve high-level features, such as the topological complex of a scalar function on the surface. As a first step in this direction, my group at the Courant Institute developed a height-field smoothing algorithm that allows explicitly prescribing the desired Morse-Smale complex [33]. To achieve higher-order continuity of the result and maintain the prescribed structure, we perform constrained (nonlinear) bi-Laplacian optimization. A long term research challenge is further exploration of topological and semantic feature definitions and the required surface optimization frameworks, with clear links between surface representation, filtering, and editing.

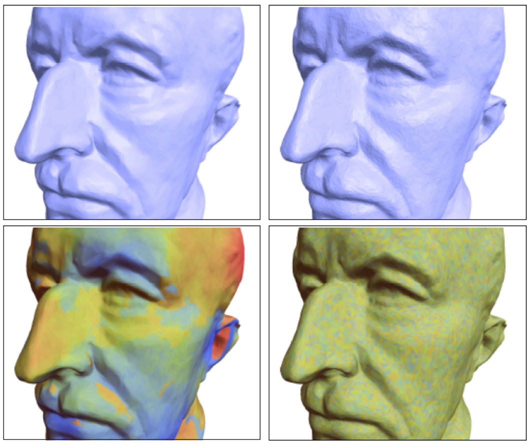


Figure 6. Quantization of the differential coordinates to 5 bits/coordinate (left) introduces low-frequency errors, whereas Cartesian quantization to 11 bits/coordinate (right) introduces noticeable errors.

Shape approximation and geometry compression. Acquired discrete geometric models are often complex and require a vast amount of storage space. I have developed a compression method for highly-detailed data, called high-pass quantization [27]. The method produces visually imperceptible low-frequency geometric errors, in contrast to the high-frequency errors of standard quantization, and therefore allows significantly better detail preservation (see Figure 6). The error bounds of this type of compression were analyzed in [4] and a practical algorithm to uphold the bounds was developed. Feature analysis of the shape can lead to even more

Original model, 49864 vertices

100 basis vectors, e=0.6
0.6KB (0.09 bits/vertex)

500 basis vectors, e=0.21
2.8KB (0.46 bits/vertex)

4000 basis vectors, e=0.07
22.2KB (3.65 bits/vertex)

9000 basis vectors, e=0.04
50.1KB (8.23 bits/vertex)

Figure 7. Reconstruction of the Feline model using an increasing number of geometry-aware basis vectors. The sizes of the encoded geometry files are displayed below the models. The letter e *denotes the* L^2 *error value, given in units of* 10^{-4}.

dramatically compact representation, as I have shown in my work on shape approximation using "geometry-aware" basis vectors [25] (see Figure 7). It has also been employed for example-space compression in computer animation [32]. In ongoing and future research, I am interested in exploring the above approaches in the context of compactly representing time-varying phenomena, and further development of structure- and semantics-aware compact shape representations.

Parameterization. Parameterization is needed in a wide range of geometry processing tasks, such as texture mapping, filtering, resampling and remeshing, as well as for illumination and simulation computations. I proposed a simple and very efficient solution to simultaneously segment and parameterize triangle meshes, which provably bounds the introduced metric distortion [24]. Currently I am interested in meshless parameterization techniques that deal directly with the raw scanning data (range images). This would allow accurate and efficient processing of real-world geometric data, such as the huge volumes of geometric and photometric data acquired in urban scanning. I have also worked on fast ray space parameterization for from-region visibility computations [11], leveraging high coherence of spatial arrangement in geometric data and parallel computation on the GPU. I see great potential in exploring parallel visibility computation in the context of collision detection for physical simulation and geometric modeling, as previously mentioned.

3. Visual media editing and retargeting

Similar to the field of geometry processing, research on content-aware, structure preserving manipulation of visual media such as images or video has gained considerable importance in computer vision and graphics research in recent years. Due to the growing variety in capture devices and displays and the availability of huge amounts of image and video data, media content has to be adapted to the characteristics of different output devices. For instance, legacy data that was captured with limited sensor capability and intended for a particular type of display device needs to be adapted to newer display technology. A typical example is viewing narrow (4:3) low-dynamic range footage on high dynamic range wide screen displays. To this end, images and videos are analyzed, segmented, and semantic information is extracted to assist higher-level manipulation algorithms. In my research, I am interested in content-aware editing of visual data for authoring and retargeting purposes.

Aspect ratio and resolution

Changing the aspect ratio of images and video by simple scaling or cropping does not provide satisfactory results since only the geometric constraints of the output display are considered. In [7], together with R. Gal, I proposed a feature-aware image warping method that deforms an image into an arbitrary prescribed 2D shape while preserving the shape of manually marked important features, at the expense of placing the distortion in unimportant areas. This was the first content-aware warping method, and several image retargeting works followed, including discrete seam carving approaches and additional continuous optimization approaches (see our survey in [20]). I have explored image retargeting further in [30] together with my student and colleagues from NCKU, where we proposed an automatic warping approach that relies on an image saliency map and allows uniform scaling of prominent content while hiding distortion in background areas (see Figure 8). The real challenge in aspect ratio retargeting lies in video information, since temporal coherence is of outmost importance there. As my team showed in [29], simple retargeting of each video frame independently, or previously employed simple temporal coherence constraints do not suffice. We were the first to propose a principled approach to motion-aware retargeting, taking camera and dynamic object motion into account. Future research would study motion retargeting further, as well as relevant hardware-amenable optimization techniques, e.g., replacing the costly pixel-based variational optimization with coordinate-based warping [19,31].

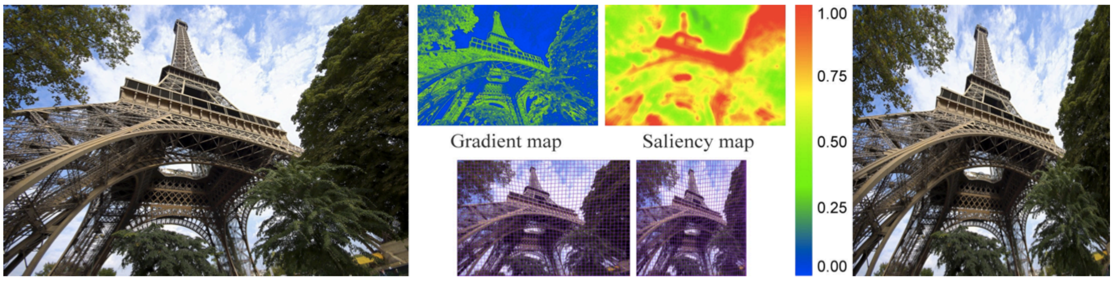


Figure 8. Image retargeting [30]. We partition the original image (left) into a grid mesh and deform it to fit the new desired dimensions (right), such that the quad faces covering important image regions are optimized to scale uniformly while regions with homogeneous content are allowed to be distorted. The scaling and stretching of the image content is guided by a significance map which combines the gradient and the saliency maps.

Dynamic range and color

Together with my colleagues from the University of Zaragoza and MPI I have explored tone mapping of legacy data for high dynamic range displays, in particular under varying exposure conditions [14,15]. By means of psychophysical experiments we have evaluated the performance of state of the art reverse tone mapping techniques and discovered previously unknown limitations, for which we proposed a new, perceptually validated tone mapping method. Furthermore, I worked on color harmonization of images [5] – the first technique of its kind to optimize the colors of an image according to classical models of harmony [10]. An interesting next step is computing tone mapping that exploits simultaneous contrast rather than just luminance contrast.

In my ongoing and future research I am interested in conducting further perceptual evaluation of retargeting methods, in particular image and video resizing, with the goal to find out what phenomena and artifacts affect the human perception of retargeting and to devise a perceptually valid quality metric and algorithms that optimize it. I believe that perceptual understanding of nonlinear image and video editing is absolutely necessary not only in the retargeting scope but generally for evaluation of image processing and computational photography research.

Semantic processing of geometric and image/video data clearly share high-level concepts and methodologies, offering great potential for cross-fertilization between the two fields which I am excited to continue exploring in my future research.

Acknowledgement

I am deeply honored to receive the ETH Latsis Prize and I would like to thank the Latsis Foundation for its generous support of science in Switzerland and internationally. I am also infinitely grateful to all my collaborators and students for the years of challenging, fruitful and exciting work together.

References

[1] Mario Botsch, Mark Pauly, Markus Gross, and Leif Kobbelt. PriMo: Coupled prisms for intuitive surface modeling. In Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, pages 11–20, 2006.

[2] Mario Botsch and Olga Sorkine. On linear variational surface deformation methods. IEEE Transactions on Visualization and Computer Graphics, 14(1):213–230, 2008.

[3] Tamy Boubekeur, Olga Sorkine, and Christophe Schlick. SIMOD: Making freeform deformation size-insensitive. In Proceedings of IEEE/Eurographics Symposium on Point-Based Graphics, pages 47–56, 2007.

[4] Doron Chen, Daniel Cohen-Or, Olga Sorkine, and Sivan Toledo. Algebraic analysis of high-pass quantization. ACM Transactions on Graphics, 24(4):1259–1282, 2005.

[5] Daniel Cohen-Or, Olga Sorkine, Ran Gal, Tommer Leyvand, and Ying-Qing Xu. Color harmonization. ACM Transac- tions on Graphics (Proceedings of ACM SIGGRAPH), 25(3):624–630, 2006.

[6] Mathias Eitz, Olga Sorkine, and Marc Alexa. Sketch based image deformation. In Proceedings of Vision, Modeling and Visualization (VMV), pages 135–142, 2007.

[7] Ran Gal, Olga Sorkine, and Daniel Cohen-Or. Feature-aware texturing. In Proceedings of Eurographics Symposium on Rendering, pages 297–303, 2006.

[8] Ran Gal, Olga Sorkine, Niloy Mitra, and Daniel Cohen-Or. iWires: An analyze-and-edit approach to shape manipula- tion. ACM Transactions on Graphics (proceedings of ACM SIGGRAPH), 28(3), 2009.

[9] Ran Gal, Olga Sorkine, Tiberiu Popa, Alla Sheffer, and Daniel Cohen-Or. 3D collage: Expressive non-realistic modeling. In NPAR: Proceedings of the 5th International Symposium on Non-Photorealistic Animation and Rendering, pages 7–14. ACM Press, 2007.

[10] Johannes Itten. Kunst der Farbe. Subjektives Erleben und objektives Erkennen als Wege zur Kunst. Urania Verlag, 1970.

[11] Tommer Leyvand, Olga Sorkine, and Daniel Cohen-Or. Ray space factorization for from-region visibility. ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH), 22(3):595–604, 2003.

[12] Yaron Lipman, Olga Sorkine, Daniel Cohen-Or, David Levin, Christian Roessl, and Hans-Peter Seidel. Differential coordinates for interactive mesh editing. In Proceedings of Shape Modeling International, pages 181–190. IEEE Computer Society Press, 2004.

[13] Yaron Lipman, Olga Sorkine, David Levin, and Daniel Cohen-Or. Linear rotation-invariant coordinates for meshes. ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH), 24(3):479–487, 2005.

[14] Miguel Martin, Roland Fleming, Olga Sorkine, and Diego Gutierrez. Understanding exposure for reverse tone mapping. In Proceedings of CEIG, pages 189–198. EG Spanish Chapter, 2008.

[15] Belen Masia, Sandra Agustin, Roland Fleming, Olga Sorkine, and Diego Gutierrez. Evaluation of reverse tone mapping through varying exposure conditions. ACM Transactions on Graphics (proceedings of ACM SIGGRAPH Asia), 28(5), 2009.

[16] Andrew Nealen, Takeo Igarashi, Olga Sorkine, and Marc Alexa. Laplacian mesh optimization. In Proceedings of ACM GRAPHITE, pages 381–389, 2006.

[17] Andrew Nealen, Takeo Igarashi, Olga Sorkine, and Marc Alexa. FiberMesh: Designing freeform surfaces with 3D curves. ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH), 26(3):article no. 41, 2007.

[18] Andrew Nealen, Olga Sorkine, Marc Alexa, and Daniel Cohen-Or. A sketch-based interface for detail-preserving mesh editing. ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH), 24(3):1142–1147, 2005.

[19] Scott Schaefer, Travis McPhail, and Joe Warren. Image deformation using moving least squares. ACM Trans. Graph., 25(3):533–540, 2006.

[20] Ariel Shamir and Olga Sorkine. Visual media retargeting. In ACM SIGGRAPH Asia Courses, 2009.

[21] Olga Sorkine. Differential representations for mesh processing. Computer Graphics Forum, 25(4):789–807, 2006.

[22] Olga Sorkine and Marc Alexa. As-rigid-as-possible surface modeling. In Proceedings of Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, pages 109–116, 2007.

[23] Olga Sorkine and Mario Botsch. Tutorial: Interactive shape modeling and deformation. In Eurographics, 2009. [24] Olga Sorkine, Daniel Cohen-Or, Rony Goldenthal, and Dani Lischinski. Bounded-distortion piecewise mesh parameterization. In Proceedings of IEEE Visualization, pages 355–362. IEEE Computer Society, 2002.

[25] Olga Sorkine, Daniel Cohen-Or, Dror Irony, and Sivan Toledo. Geometry-aware bases for shape approximation. IEEE Transactions on Visualization and Computer Graphics, 11(2):171–180, 2005.

[26] Olga Sorkine, Daniel Cohen-Or, Yaron Lipman, Marc Alexa, Christian Roessl, and Hans-Peter Seidel. Laplacian surface editing. In Proceedings of the Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, pages 179–188. ACM Press, 2004.

[27] Olga Sorkine, Daniel Cohen-Or, and Sivan Toledo. High-pass quantization for mesh encoding. In Proceedings of the Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, pages 42–51. Eurographics Association, 2003.

[28] Robert Y. Wang, Kari Pulli, and Jovan Popovic. Real-time enveloping with rotational regression. ACM Trans. Graph., 26(3):73, 2007.

[29] Yu-Shuen Wang, Hongbo Fu, Olga Sorkine, Tong-Yee Lee, and Hans-Peter Seidel. Motion-aware temporal coherence for video resizing. ACM Transactions on Graphics (proceedings of ACM SIGGRAPH Asia), 28(5), 2009.

[30] Yu-Shuen Wang, Chiew-Lan Tai, Olga Sorkine, and Tong-Yee Lee. Optimized scale-and-stretch for image resizing. ACM Transactions on Graphics (proceedings of ACM SIGGRAPH Asia), 27(5):article no. 118, 2008.

[31] Ofir Weber, Mirela Ben-Chen, and Craig Gotsman. Complex barycentric coordinates with applications to planar shape deformation. Computer Graphics Forum (Proceedings of Eurographics), 28(2), 2009.

[32] Ofir Weber, Olga Sorkine, Yaron Lipman, and Craig Gotsman. Context-aware skeletal shape deformation. Computer Graphics Forum (Proceedings of Eurographics), 26(3):265–273, 2007.

[33] Tino Weinkauf, Yotam Gingold, and Olga Sorkine. Topology-based smoothing of 2D scalar fields with C^1-continuity. Computer Graphics Forum (proceedings of EuroVis 2010), 29(3):1221–1230, 2010.