SMPTE Meeting Presentation

Novel Stereoscopic Content Production Tools

Alexander Hornung

Disney Research Zurich, hornung@disneyresearch.com

Aljoscha Smolic

Disney Research Zurich, smolic@disneyresearch.com

Markus Gross

Disney Research Zurich / ETH Zurich, gross@disneyresearch.com

Written for presentation at the

SMPTE International Conference on Stereoscopic 3D for Media and Entertainment

Abstract. *Stereoscopic 3D is on the cusp of becoming a mass consumer product. But despite the recent rise of stereoscopic movies and technological advances in the industry, many challenges regarding 3D content creation and display remain unsolved. A fundamental issue is control and modification of the stereoscopic depth of live-action content after it has been recorded, since changing depth effectively amounts to changing the baseline and vergence of the capture rig.*

With existing post-production tools, the modification of these parameters involves cumbersome and expensive segmentation, dense depth reconstruction, and inpainting. In our paper, we provide an overview on existing technologies for stereoscopic content editing, identify current challenges, and present a number of recent research results which provide novel solutions to problems such as disparity correction and depth authoring, display adaptation, and 2D-to-3D conversion.

Keywords. Stereoscopic 3D content generation, post-production, depth editing,

The authors are solely responsible for the content of this technical presentation. The technical presentation does not necessarily reflect the official position of the Society of Motion Picture and Television Engineers (SMPTE), and its printing and distribution does not constitute an endorsement of views which may be expressed. This technical presentation is subject to a formal peer-review process by the SMPTE Board of Editors, upon completion of the conference. Citation of this work should state that it is a SMPTE meeting paper. EXAMPLE: Author's Last Name, Initials. 2010. Title of Presentation, Meeting name and location.: SMPTE. For information about securing permission to reprint or reproduce a technical presentation, please contact SMPTE at jwelch@smpte.org or 914-761-1100 (3 Barker Ave., White Plains, NY 10601).

Introduction

The current attention of the movie and video industry on stereoscopic 3D content production poses novel challenges, which require a fundamental rethinking and redesign of existing post-production workflows and tools. In contrast to previous technical advances, such as the step from SD to HD, the technical challenges in stereoscopic video do not lie primarily in improved resolution, color fidelity, or transmission bitrates; stereoscopic video introduces *depth* as a completely new visual cue (and challenge) into the process.

Similar to perceiving the real world through our two eyes, our brain can generate the illusion of depth from a stereoscopic image pair consisting of two slightly different viewing positions. However, in order to do so, the properties of such a stereoscopic image pair have to be sufficiently close to what we are used to in the real world. This includes image properties such colors, or geometric distortions (e.g., through lenses). A consequence for post-production is that all processing steps such as color correction, geometric distortion, etc. have to be performed in a synchronized manner for both views. Although technically challenging, these modifications are mostly straightforward extensions of existing technologies for post-production.

In contrast, stereoscopic depth is a property that is *implicitly* contained in the images, and reconstructed by our brain by analyzing a variety of factors, including the displacement of scene elements in the images (disparity), vergence of the eyes, motion parallax, accommodation, or learned knowledge such as relative size and position of objects with respect to each other. The stereographer has certain control over perceived depth by manipulating these parameters, with two of the central technical instruments being the camera baseline (distance) and convergence.

However, it is important to consider that our visual system is subject to several perceptual restrictions. For example, we can only fuse a certain range of disparities into a proper 3D image, and the limitations of existing stereoscopic display technology lead to conflicting depth cues, the so called vergence-accommodation conflict [1]. Violations of such restrictions are the cause of problems like visual fatigue when viewing 3D movies [2]. These considerations render stereoscopic 3D content creation considerably different to 2D movie production, and have led to a complex set of rules and guidelines [3] which should be adhered to during capturing, post-production, and display in order to achieve the desired 3D movie experience. A typical example for these types of guidelines is the so called stereoscopic comfort zone shown in Figure 1, which provides a qualitative illustration of the depth zones around a display where stereo is comfortable to observe, and which zones result in conflicting visual cues.

Now a fundamental problem of stereoscopic content creation is that central depth cues such as image disparity between the two views are fixed in the moment the content is captured; issues such as excessive scene depth cannot be corrected anymore, since changing image disparities effectively amounts to modifying the baseline between the recording cameras. Hence, in order to change perceived depth *after* capturing during post-production, one has to artificially generate virtual views with novel (increased or decreased) camera baselines from the already recorded stereoscopic footage.

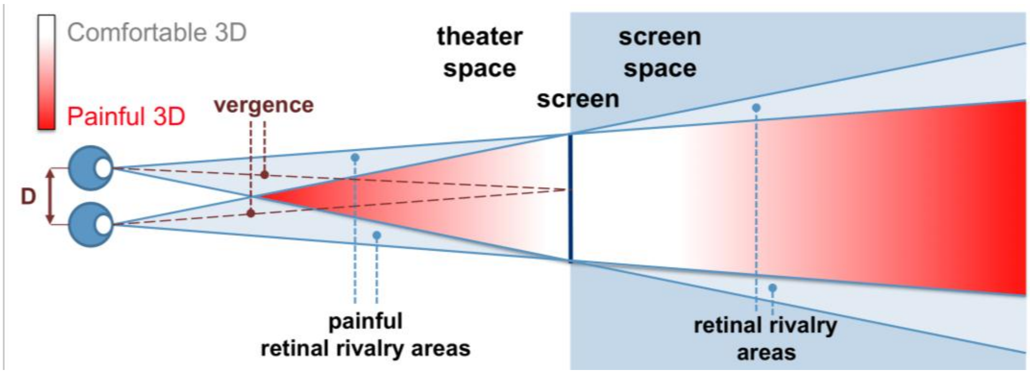

Figure 1: Illustration of the stereoscopic comfort zone.

This novel view synthesis problem has been investigated in computer graphics and vision research for many years, and the basic approach is well known; the core task is to reconstruct an explicit depth map (or disparity map) from the implicit depth information contained in the stereo footage, so that for each scene point its respective distance to the cameras is available. Given this explicit depth information it is then possible to compute novel views of a scene. However, due to the ill-posed, difficult nature of this problem, to date there is no sufficiently robust, fully automatic algorithm available to solve this problem for arbitrary scenes with cinema quality. Consequently, post-production of stereoscopic 3D movies requires cumbersome and expensive manual editing such as segmentation and rotoscoping, depth map generation and correction, or inpainting of occluded scene content.

In the following sections we first describe the classical image-based rendering pipeline using dense depth maps, and discuss some of the major challenges in reconstructing such depth maps from stereoscopic footage. Then we describe advanced user interfaces, which enable an efficient semi-automatic correction of dense depth maps, or allow for the generation of stereoscopic 3D video from a single 2D video source. Finally, we provide an overview on a recently developed approach for stereoscopic image editing, which is based on image *warping* with sparse depth information instead of dense depth, and which provides a powerful new alternative to existing techniques. We conclude with a discussion and outlook on remaining research challenges.

The Classical Depth Image-based Rendering Pipeline

The classical approach for novel view synthesis relies on a video plus depth representation as illustrated in Figure 2. For each pixel in the color image there is a corresponding depth value describing it's depth in the scene. In this example depth is quantized with 8 bit on a logarithmic scale between a minimum and maximum distance Z_{near} and Z_{far}. Brighter areas are closer to the camera; darker areas are further in the background.

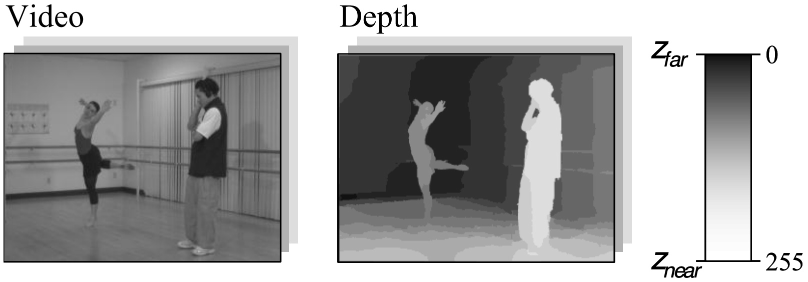

Figure 2: Video and associated per pixel depth data [4]. Images courtesy of Zitnick et al.

Such a 3D data representation combining pixel data with 3D geometry information allows rendering of virtual views nearby the available color image. This process known as depth image-based rendering (DIBR) is illustrated in Figure 3. In this example, two original camera views with associated depth data are available. In DIBR, each pixel is projected into the 3D space using the depth map and camera calibration information. Then, each pixel can be projected back onto an arbitrary virtual camera plane, creating a novel virtual view [7]. However, disocclusion artifacts increase with distance of the virtual view from the available view, limiting the virtual view navigation range. For intermediate view synthesis in between two not too distant original cameras disocclusion artifacts are limited. Remaining holes in the rendered images have to be filled using inpainting algorithms [10]. Specific handling of depth discontinuities helps to reduce artifacts along object borders [4], [5], [6]. Different combinations of such novel views can then be used to modify the image disparity and, hence, the perceived depth.

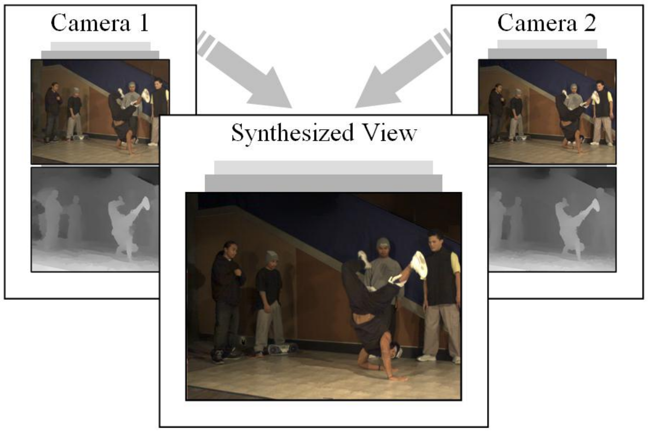

Figure 3: Intermediate view synthesis from multiple video plus depth data. Images courtesy of Zitnick et al.

A general problem of DIBR approaches is that they rely on availability of accurate, dense depth data, i.e., a depth value per image pixel. Depth generation is estimation by nature. The true information is in general not accessible. Robustness of the estimation depends on many theoretical and practical factors. There is always a residual error probability which may affect the quality of the finally rendered output views.

Depth estimation is an important core research area of computer vision and numerous algorithms for this purpose have been proposed [26]-[29] ([29] provides a collective repository of papers, test data and quality evaluation tools). Although these algorithms may overall differ drastically, they rely in the core on establishment of dense correspondences between images related to projection of the same 3D scene points to different cameras (dense depth or disparity maps). Inherent problems arise from occlusions and ambiguities, which is why reliable correspondence estimation cannot be guaranteed. Practical problems include noise, varying illumination conditions, and other types of inaccuracies of signals and camera setups. Therefore, reliable, accurate, robust and automatic depth estimation is still a difficult and unresolved task, and a very active and important research area.

Some newer approaches combine color cameras with specific depth sensors that directly capture scene depth [8], [9]. Typical time-of-flight (ToF) depth sensors are of relatively low resolution, e.g. 204x204 pixel. The generated depth maps have to be registered with the color images, which is not an easy task since both cameras are inherently located at different positions. Further, ToF sensors are very sensitive to noise and temperature, and the depth range they capture is quite limited. Also non-linear distortions of the measured depth values are a problem. A promising approach is to combine ToF depth maps with high-resolution depth maps computed by classical stereo algorithms to get the best of both worlds [9].

Other 3D geometry reconstruction algorithms can be applied in certain application scenarios as alternative to the aforementioned shape-from-stereo. One class is shape-from-silhouette or visual hull reconstruction [11]. Such algorithms typically use a multi-camera setup. An object of interest is segmented in each of the camera views. The 3D volume of the object of interest or more precisely its convex hull can be reconstructed using volume carving [12]. Numerous variants and improvements of shape-from-silhouette have been described over the years, e.g. [13], [14]. Other types of 3D structure recovery include shape-from-focus and -defocus [15], [16], shape-from-shading [17], [18], and structure-from-motion [19]-[22]. Some methods extract geometrical scene properties like vanishing points and lines to get a 3D reconstruction [23]-[25].

However, all these methods only work well under specific conditions and for specific scenes. General robustness and accuracy cannot be guaranteed, in particular for stereoscopic content captured with only two cameras. User-assisted content generation is an option for specific application scenarios to improve performance, where a human operator selects a method and parameters and corrects errors until the result is satisfying. But these types of approaches generally involve a considerable amount of expensive manual interaction. Finally, there exist a couple of techniques for automatic, real-time conversion of 2D to 3D content. Although these techniques may provide acceptable results in some cases, they may also totally fail and do not provide a general solution to the problem. All these issues render the use of classical DIBR difficult in the context of high quality cinema or broadcast applications.

Semi-Automatic Generation and Correction of Dense Depth

A core problem for the above depth-based rendering pipeline is a sufficiently robust and accurate automatic computation of the dense depth maps. Since this process is still challenging and error prone, research has been invested into building sophisticated semi-automatic interfaces, which assist the user to correct wrong depth estimates, or to generate depth maps manually “from scratch”, e.g., for 2D to 3D video conversion.

One of the simplest ways to create or modify stereoscopic depth is to perform the depth assignment on a per-object basis, i.e., assign a constant depth value to each object in a scene. For example, given a 2D video that should be converted to stereoscopic 3D, a typical workflow would consist of first rotoscoping and segmenting [30] the individual scene elements and objects in the source video. Then, each object gets a depth value assigned, which effectively corresponds to a left- or right-shift of the object (the disparity or displacement) in the target video. Although this approach provides a basic impression of depth, each particular object does not feature any internal depth variation, so that the resulting scene looks as if it was composed of many flat billboards (cardboard effect).

Assigning more complex geometric proxies instead of planes (e.g., cylinders, spheres, etc.) can improve the situation. But although this can provide an improved, more natural sensation of depth per object, the difference to the true depth can obviously be arbitrarily large and distracting for the viewer, so that additional manual editing and correction is necessary. Overall, the major disadvantage of this approach is that it mostly consists of manual editing without any higher-level support by the computer.

As an example, more sophisticated semi-automatic proxy generation and augmentation of video with depth has been described in [31]. The user provides a sketch of the dominant surface parts of an object, which are then tracked by the system throughout the video sequence. From the tracked surfaces it is then easily possible to reconstruct a 3D model, which captures the geometry of the object much more faithfully than the purely manual techniques described above. In the paper this approach is used to create textured 3D models, but the same technique would be applicable to creating and correcting stereoscopic depth maps. A variety of similar approaches are referenced in [31]. However, although these types of techniques are most suitable for man-made, static structures such as architecture, machines, etc, they are less suitable for more general, organic and dynamic scenes involving as characters.

A system that targets more directly at stereoscopic depth map creation and editing has been presented in [32]. In this system, the user provides a number of scribbles in a sparse set of video frames. Then system then combines this sparse information with local visual saliency, object motion, and machine learning techniques in order to assign depth to the complete image. In contrast to the previous approach the system can handle dynamic scenes, but it is restricted to rather small depth variation and it is difficult to reproduce fine depth details and features. The main application presented in the paper is 2D to 3D conversion.

These types of interfaces provide a first step into semi-automatic tools for dense depth correction and editing in post-production, but many research challenges remain.

Image Warping Technologies

Despite the recent advances of algorithms for dense depth reconstruction, improved user interfaces, and novel sensor technologies, controlling stereoscopic depth in post-production based on dense depth maps and inpainting of occluded regions remains a difficult challenge and cannot be solved yet in a fully automatic manner.

An alternative family of approaches for novel view synthesis employs *image-based warping* [33][34]; instead of generating novel views by rendering the scene from a dense 3D reconstruction, warping-based methods deform or morph image content directly in image space. The underlying idea is to define a sparse set of image features between one or more source images, and the desired position of these features in a target view. Such features can be simple pixel correspondences, or higher-level features such as lines, shapes, etc. The sparse feature positions in the target view are then used as deformation constraints to compute a dense nonlinear warping field, which maps the input pixels from the source image(s) to the target view. The goal of this warping is to fulfill the feature position constraints (i.e., move an object from one position to another), and at the same time minimize the visual distortion in the remaining image regions.

A simple example that utilizes per-pixel warping [35] to create a stereoscopic image pair from a single 2D input image is shown in Figure 4. The middle image shows the statue from the original left image moved slightly to the right, while the background sky remains fixed. The necessary image deformation is hidden in a narrow band around the statue (visualized in the right image). When viewed cross-eyed, the left and middle image can be fused as a stereoscopic image pair and provide a basic impression of depth. Note that this image pair has been generated without the necessity for accurate boundary segmentation or inpainting of occluded image regions.

Figure 4: Image warping applied to the problem of 2D to 3D conversion.

Algorithms for image warping have shown to be practical and powerful solutions to a variety of problems, which have been difficult to solve using view interpolation techniques based on dense depth. Examples include content-aware rescaling of images and video between different aspect ratios [35], undistorting and optimizing image content [36], modification of perspective [37], and stabilization of camera paths [38].

In recent work, similar principles have been applied to stereoscopic depth editing [39]. This work discusses the general problem of modifying the depth impression in stereoscopic video and proposes a general framework for nonlinear disparity mapping. The framework considers the previously mentioned perceptual and technical restrictions when generating and displaying 3D video, and proposes a solution that warps a stereoscopic image pair in order to modify perceived scene depth instead of applying the classical depth-based approaches for view synthesis. The basic principle is illustrated in Figure 5. First, the stereoscopic video is analyzed, and a set of sparse stereo correspondences is computed automatically for each pair of input frames. In Figure 5 the leftmost image shows a left view of a stereoscopic image pair and the detected sparse stereo correspondences. Red lines depict scene points with positive image disparity (background), while blue lines represent the negative disparity of the statue in the foreground.

Perceived depth can now be controlled by changing the disparity values of the stereo correspondences (intuitively this means changing the “length” of the lines), and using the resulting positions of the correspondences as input constraints to a method for image warping. In order to control, which parts of the input frames may be deformed, one additionally computes visual importance maps (middle-left image). During the warping, only the dark regions are deformed, so that the visible distortion is minimized. The difference between an original and a warped video frame is visualized in the two images on the right. The deformation (i.e., change of disparity) can be seen on the overlayed grid lines, e.g., around the statues’ heads. Without the grid, the deformation is imperceptible, while the stereoscopic depth is effectively changed.

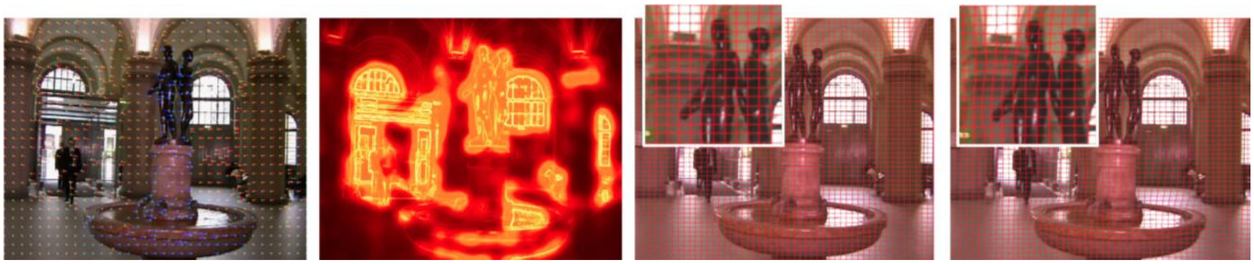

Figure 5: Stereoscopic depth editing by image warping.

Note that the output images computed with warping-based approaches are not *geometrically accurate* since image content is deformed by the warping function. However, the deformations necessary for the above types of problems are generally so small that our visual system easily compensates for slight visual inconsistencies. For many practically relevant problems, the resulting images are *visually plausible*, without noticeable artifacts.

Moreover, the previously discussed depth-based methods are only able to generate geometrically accurate images under the assumption of perfectly accurate dense depth maps and inpainting. These assumptions are unrealistic in practice, such that depth-based approaches are effectively confronted with similar restrictions as the stereoscopic warping.

The major advantage of warping-based techniques is their conceptual simplicity and that they can mostly be computed fully automatically and in real-time. They do not have requirements such as accurate camera calibration, dense depth estimates, inpainting, or other challenging processing steps known from depth-based view synthesis. Therefore, this new class of warping-based methods to video post-production and stereoscopic depth editing could develop into a promising alternative to existing post-production technologies, and open up a variety of applications, which are difficult to realize with depth-based methods, e.g., real-time display adaptation, online correction of stereoscopic framing problems, etc.

Conclusion

In this paper we provided an overview of currently existing tools for stereoscopic depth editing, ranging from the classical technologies based on dense depth and inpainting to recently developed approaches, which employ image-warping and sparse stereo correspondences.

All these technologies are still in a state where important research challenges remain to be solved in order to build effective tools for stereoscopic post-production. In classical depth image-based rendering, the core problems are still sufficiently reliable computation of dense depth maps for arbitrary scenes, accurate segmentation, and automatic inpainting of occluded regions. The manual effort that often has to be invested in order to address the deficiencies of these technologies can be alleviated by advanced semi-automatic user interfaces, which assist the user in correcting or even creating dense depth. Image-warping techniques described in the last section are conceptually much simpler and avoid many of the above problems since they only require sparse feature correspondences and work directly in image-space. However, since this technology has only recently been developed, additional perceptual studies would be helpful in order to determine whether the introduced image-distortions and geometrical inaccuracies have effects on stereo perception.

Although many of these recent techniques are still research prototypes and require technology transfer to actual post-production, they extend the currently existing spectrum of tools considerably and provide an interesting alternative for the industry to long established, but still not sufficiently matured technology for stereoscopic post-production.

References

[1] I. P. Howard and B. J. Rogers, *Seeing in Depth*, Oxford University Press, New York, USA, 2002

[2] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks: “Vergence-accommodation conflicts hinder visual performance and cause visual fatigue”, Journal of Vision 8, 3(3), 2008, 1-30

[3] B. Mendiburu, “3D Movie Making – Stereoscopic Digital Cinema from Script to Screen.” Elsevier, ISBN: 978-0-240-81137-6, 2008.

[4] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, “High-Quality Video View Interpolation Using a Layered Representation”, *ACM SIGGRAPH and ACM Trans. on Graphics*, Los Angeles, CA, USA, August 2004.

[5] K. Mueller, A. Smolic, K. Dix, P. Merkle, P. Kauff, and T. Wiegand, “View Synthesis for Advanced 3D Video Systems”, *EURASIP Journal on Image and Video Processing*, 2008, doi:10.1155/2008/438148.

[6] A. Hornung, and L. Kobbelt, "Interactive Pixel-Accurate Free Viewpoint Rendering from Images with Silhouette Aware Sampling”, Computer Graphics Forum, 2009, vol. 28, 8, pp. 2090 – 2103.

[7] C. Fehn, P. Kauff, M. Op de Beeck, F. Ernst, W. Ijsselsteijn, M. Pollefeys, L. Vangool, E. Ofek, and I. Sexton, “An Evolutionary and Optimised Approach on 3D-TV”, Proc. of IBC 2002, Int. Broadcast Convention, Amsterdam, Netherlands, Sept. 2002.

[8] http://www.3d4you.eu/index.php.

[9] B. Bartczak and R. Koch, “Dense Depth Maps from Low Resolution Time-of-Flight Depth and High Resolution Color Views”, Proc. ISVC 09, Las Vegas, Nevada, USA, December 1-3, 2009.

[10] M. Bertalmío, G. Sapiro, V. Caselles, and C. Ballester, "Image Inpainting", in Proc. of ACM SIGGRAPH, New Orleans, USA, July 2000.

[11] A. Laurentini, "The Visual Hull Concept for Silhouette-Based Image Understanding", in IEEE Transactions on Pattern Analysis and Machine Intelligence 16, no. 2, pp. 150–162, Feb. 1994.

[12] W. Matusik, C. Buehler, R. Raskar, S. J. Gortler, and L. McMillan, “Image-based Visual Hulls”, in SIGGRAPH’00 Proceedings, pages 369–374, July 2000.

[13] C. Liang, and K. K. Wong, "3D reconstruction using silhouettes from unordered viewpoints", Image and Vision Computing 28 (4), pp. 579-589, 2010.

[14] E. de Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H. P. Seidel, and S. Thrun, “Performance capture from sparse multi-view video”, in SIGGRAPH '08: ACM SIGGRAPH 2008 papers (2008), pp. 1-10.

[15] Y.-F. Liu, "A Unified Approach to Image Focus and Defocus Analysis", PhD. Thesis, State University of New York at Stony Brook, New York, 1998

[16] T.-C. Wei, "Three Dimensional Machine Vision Using Image Defocus", PhD. Thesis, State University of New York at Stony Brook, New York, 1994.

[17] R. Zhang, P.-S. Tsai, J. E. Cryer, and M. Shah, "Shape from Shading: A Survey", in IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (1999), no. 8, pp. 690–706.

[18] J.-D. Durou, M. Falcone, and M. Sagona, "Numerical Methods for Shape-from- Shading: A New Survey with Benchmarks", in Computer Vision and Image Understanding 109, no. 1, pp. 22–43, Jan. 2008.

[19] P. Beardsley, P. H. S. Torr, and A. Zisserman, "3D Model Acquisition from Extended Image Sequences", in Proc. of the European Conf. on Computer Vision (ECCV), pp. 683-695, 1996.

[20] T. Jebara, A. Azarbayejani, and A. Pentland, “3D structure from 2D motion”, IEEE Signal Processing Magazine, May 1999, Vol. 16. No. 3, p. 66-84.

[21] R.-I. Hartley, and A. Zisserman, "Multiple View Geometry in Computer Vision", Cambridge University Press, 2. Edition, 2004.

[22] M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch, "Visual modeling with a hand-held camera", International Journal of Computer Vision 59(3), 207-232, 2004.

[23] D. Hoiem, A. A. Efros, and M. Hebert, "Automatic Photo Pop-up", in ACM Transactions on Graphics 24, no. 3, pp. 577–584, 2005.

[24] S. Lee, D. Feng, and Gooch, "Perception-Based Construction of 3D Models from Line Drawings", in Symposium on Interactive 3D Graphics and Games, 2008.

[25] P. Parodi, and G. Piccioli, "3D Shape Reconstruction by Using Vanishing Points", in IEEE Transactions on Pattern Analysis and Machine Intelligence 18, no. 2, pp. 211–217, Feb. 1996.

[26] N. Atzpadin, P. Kauff, and O. Schreer, "Stereo Analysis by Hybrid Recursive Matching for Real-Time Immersive Video Conferencing", in IEEE Transactions on Circuits and Systems for Video Technology, no. 14, pp. 321-334, 2004.

[27] D. Scharstein, and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms", in International Journal of Computer Vision, no. 47, pp. 7-42, 2002.

[28] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, „A comparison and evaluation of multi-view stereo reconstruction algorithms”, in Proc. IEEE Conf. Comp. Vision and Pattern Recognition, pages 519–528, 2006.

[29] http://vision.middlebury.edu/stereo/

[30] X. Bai, J. Wang, D. Simons, G. Sapiro: “Video SnapCut: robust video object cutout using localized classifiers”, ACM Trans. Graph. 28(3), 2009

[31] A. van den Hengel, A. R. Dick, T. Thormählen, B. Ward, Philip H. S. Torr: VideoTrace: rapid interactive scene modelling from video. ACM Trans. Graph. 26(3): 86, 2007

[32] M. Guttmann, L. Wolf and D. Cohen-Or: “Semi-Automatic Stereo Extraction From Video Footage”, IEEE International Conference on Computer Vision (ICCV), 2009

[33] S. Seitz and C. Dyer: “View morphing”, SIGGRAPH 96, 21-30, 1996

[34] D. Mahajan, F.-C. Huang, W. Matusik, R. Ramamoorthi, and P. N. Belhumeur: “Moving gradients: a path-based method for plausible image interpolation.”, ACM Trans. Graph. 23, 3, 814-824, 2009

[35] P. Krähenbühl, M. Lang, A. Hornung, and M. Gross: “A system for retargeting of streaming video”, ACM Trans. Graph. 28, 5, 2009

[36] R. Carroll, M. Agrawala, and A. Agrawala: “ Optimizing content-preserving projections for wide-angle images”, ACM Trans. Graph. 28, 3, 2009

[37] R. Carroll, M. Agrawala, and A. Agrawala: “Image Warps for Artistic Perspective Manipulation”, ACM Trans. Graph. (Proc. SIGGRAPH 2010), vol. 29, 3, 2010, to appear

[38] F. Liu, M. Gleicher, H. Jin, and A. Agrawala: “Content-preserving warps for 3D video stabilization”, ACM Trans. Graph. 28, 3, 2009

[39] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross: “Nonlinear Disparity Mapping for Stereoscopic 3D”, ACM Trans. Graph. (Proc. SIGGRAPH 2010), vol. 29, 3, 2010, to appear