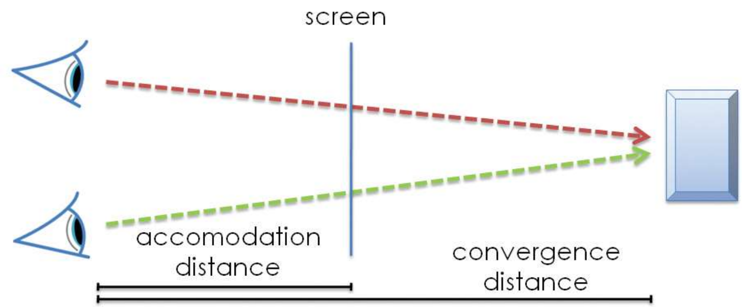


Figure 18.3: Accommodation distance is not equal to convergence distance in a S3D projection on a 2D screen, causing contradicting visual cues.

At the core, using a 2D display at a fixed distance to the viewer cannot produce a natural impression, but is instead an illusion for the human visual system that produces inconsistencies which need to be concealed, if the viewer is not to be overstressed. In particular, the accommodation (focal sharpness) distance is not equal to the convergence distance (Figure 18.3, "convergence-accomodation conflict").

Since the HVS will only tolerate a limited amount of inconsistency between cues, this limits the usable depth around the screen plane. First, the eyes' focus must always be on the screen in order to see sharp images. Now if a viewer converges on an object too far away from the screen, and the maximum sharpness is still on the screen plane, the monocular focus cue and the binocular convergence cue are inconsistent. The brain has to constantly fight against the contradicting input and may get stressed over time.

Content producers therefore carefully plan the use of the z-space, called "stereo budget" (or "depth budget"). Keeping the main action close to the screen at most times ensures the least stressful viewing experience for the audience. This can be characterized as a S3D comfort zone (Figure 18.4). Viewers will receive minimal stress within the green zone, and increasingly toward the red zone. In this manner, creatives today pay a lot of attention to the mindful and expressive depth design of each scene shot.

Additionally, the temporal aspects of depth must be considered. Scene transitions should limit aggressive jumps in depth between salient objects, as the forceful quick realignment of the eyes can be stressful. On the other hand, gentle guidance can stretch the comfort zone, as done, e.g., in a scene of the movie "Hugo"[3] where the main character slowly runs out of the screen while narrative climax, music, and depth-of-field encourage viewers

[3]Hugo movie – http://www.imdb.com/title/tt0970179/

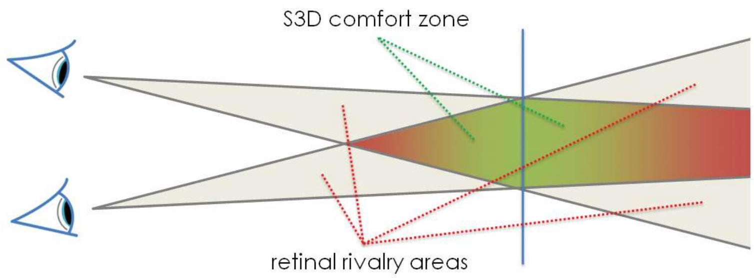


Figure 18.4: Viewer comfort is maximized by staying within the S3D comfort zone.

to fix all attention on the character. This leads to a final depth position that would be considered painful to the eyes when trying to converge on it without any setup.

Cuts should also be not too frequent because adjusting to a new scene requires more time when processing binocular cues [Berends and Erkelens 01]. Therefore, 3D movies typically feature more sweeping shots, contain fewer cuts in action scenes, and keep the depth at the same level during transitions. Applying this style is challenging for trailers, which typically have to include a multitude of cuts. Another factor is the adversarial situation of motion blur in the context of S3D, since depth is perceived more precisely at sharp object boundaries which are absent in the presence of motion blur, encouraging creatives to either increase frame rate or choose a different shooting angle or distance.

Finally, an entire class of post-production issues comes up with stereoscopic footage. The images must be aligned such that vertical disparity is eliminated, since natural scenes never produce it and it is thus perceived as stressful. Even slight inter-camera rotation or lens distortion can introduce this negative effect. Minor differences between camera sensors and lenses remain even after white balancing (more so if a mirror rig is used), and must be fixed via color grading.

18.5 3D Video Representation

Like all media, S3D defines a pipeline ranging from capture to display. When compared to standard 2D video, the additional data implies that 3D is more complex and thus bandwidth-demanding through all stages of processing (Figure 18.1). Even in the simplest case, an additional frame for the second eye has to be produced, transmitted, and displayed. While binocular S3D is currently the only commercial manifestation, other variants appear

Table 18.2: 3D scene representations can be classified as image-, depth-, or geometry-based.

Image-based	Depth-based	Model-based
Lightfield	S3D one + depth	Video textures
Lumigraph	Oriented billboards	3D mesh models
S3D two-frames	Layered depth videos	

either as part of post-production or as prototypes. The variant closest to binocular S3D uses one image plus depth map and renders the scene at the user's display. This confers a fine-tuning advantage to an individual viewer, e.g., interocular baseline adjustment or depth range remapping. Also, this representation is more compressible, an important property when streaming. On the other hand, computational cost for the display system increases, and the required inpainting at object borders and less-than-perfect depth maps may produce errors that the producers cannot remove in advance.

Taking a broader view, many other 3D representations exist which can be classified as a continuum between the two bounds of image-based vs. model-based forms (Table 18.2 and Figure 17.1).

One end consists of purely image-based approaches like dynamic light fields [Levoy and Hanrahan 96, Buehler et al. 01] or multi-view video, which do not use geometry at all and typically require many cameras for dynamic scenes. The other end is represented by approaches that use the full 3D information of a scene, e.g., by reconstructing 3D models from multi-view footage or by creating them manually by modeling and animation. In between are representations that borrow from both sides. Classical two-view S3D is sparsely image-based and borders on depth-based systems, with one-view plus depth[4] on the other side of that border [Fehn 03]. Layered depth videos are still fronto-parallel [Zitnick et al. 04], while oriented billboards [Kilner et al. 06] can already be seen as very coarse 3D video textures [Schödl et al. 00].

Different 3D representations imply different advantages and disadvantages to content producers, as shown in Table 18.3. Once chosen, the type of data is decisive for the design of the entire processing pipeline: It constrains acquisition systems and post-production algorithms, and determines rendering algorithms and view synthesis range.

As a general guideline, reconstruction costs (both computational and in terms of manual post-production refinement) are lowest for image-based systems and highest for model-based ones; conversely, the data storage and transmission demands are highest for images and lowest for models.

[4]known as depth image based rendering (DIBR)

Table 18.3: Different 3D representations imply different reconstruction efforts and amounts of data.

	Image-based	Depth-based	Model-based
Capture	Many cameras Dense sampling	Few cameras	No cameras for manual modeling or many cameras for reconstruction
Reconstruction	No effort	Depth estimation error-prone	Manual modeling labor-intensive Full 3D reconstruction error-prone, high effort for perfect results
Model insertion	Difficult	Simple if depth estimate accurate	Trivial
Coding	High bitrate, depends on number of views	Medium bitrate	Comparably low bitrate, very efficient compression possible
Rendering	Light field rendering or view interpolation, closest to footage	Depth-supported view interpolation, typically close to footage	Classical computer graphics, close to footage when rendering with high resources
Field of view	Narrow, view interpolation only	Narrow, limited extrapolation possible	Full freedom of viewing angle

Content producers must consider both filming and post-production effort, and it seems that two-view S3D movies are currently the most cost-efficient.

18.6 S3D Video Production from Real-World Data

In the 1950s, still in the analog film world, high-quality S3D production was arduous. Analog film from two cameras with limited synchronization caused errors through lens and sensor abberations; analog frames could only barely be color-calibrated by chemical means; and inserting models in post-production was only possible with physical layering, making the inclusion of temporally coherent 3D models very demanding.

Today, stereoscopic capture still suffers from lens/sensor abberations, caused, e.g., by the beam-splitting mirror in two-camera stereo rigs; lens distortion, sensor miscalibration, and inter-camera positioning errors are still present, though to a lesser degree. However, correcting these errors is much easier with digital footage. Typically in the first phase of post-production, artists remove errors with the help of commercial tools, including lens undistortion, color balancing, and vertical disparity removal, before performing depth estimation wherever needed. Because an estimated depth map is often not perfect, manual corrections are necessary and usually represent considerable effort. Alternatively, in 2D-to-3D conversion cases, a depth map may be produced manually from rotoscoping and depth map painting, possibly aided by semi-automatic approaches [Wang et al. 11c].

Depth itself is a dimension that can be edited to enhance the viewing experience as well as the narrative. Artists can use a stereoscopic analyzer like the one contained within TheFoundry Ocula[5] to visualize the depth distribution over a captured shot, and provide corrections if depth parameters get out of the S3D comfort zone (Figure 18.4); in cases where too much disparity would induce viewer discomfort, the cameras can be moved closer through disparity scaling. Capture-time stereoscopic analyzers for use with cameras also exist [Zilly et al. 10].

Another basic correction known as shift convergence adjusts the screen plane, initially determined by inter-camera positioning and their convergence. In post-production, the scene can be shifted forward or backward relative to the screen by shifting the left and right images toward or away from each other [Stelmach et al. 03]. However at the border of the images this leads to disocclusions, which can be circumvented by using bigger sensors capturing more pixels than actually needed on screen, or by cropping or inpainting.

Non-linear depth remapping is another tool-supported technique used to enhance storytelling [Lang et al. 10], e.g., for deep landscapes or to express the relative importance of actors, objects, and motions. For example, in one scene of the movie "Gravity",[6] the main character reaches out of the screen for a crucial handbook which is just floating away. Her arm has been depth-remapped to more than twice the length it could plausibly have, in order to emphasize the criticality of the moment.

A typical production can also feature the insertion of 3D models, be it actors, objects, or environmental effects like fog or fire. The amount

[5]TheFoundry Ocula – http://www.thefoundry.co.uk/products/ocula/
[6]Gravity movie – http://www.imdb.com/title/tt1454468/

of computer-generated graphics with respect to real-world footage is very flexible: Typical cases involve green-screen filmed actors being combined with completely modeled surroundings and effects, or real-world footage being enhanced by virtual objects.

All these depth-based frame editing steps have in common that the depth maps must be of high quality, otherwise depth abberations, damaged textures, and object boundaries must be repaired frame-by-frame in image space in another post-processing step. A typical approach is to fix depth map errors by manual depth map painting, where segmentation from rotoscoping can be re-used as well. This production step is labor-intensive and can generate substantial costs.

Alternatively, the depth can also be estimated and corrected interactively [Ruhl et al. 13], as described in the following. In this approach, initial depth estimation is based on fast cost volume filtering [Rhemann et al. 11], a real-time capable method to estimate depth from a stereo image pair. Given left and right views $I_l(\mathbf{x})$ and $I_r(\mathbf{x})$ with pixel coordinates $\mathbf{x} = (x, y) \in \Omega$ and RGB color values in the range $[0, 1]$, the method aims to attain for each $\mathbf{x}$ an optimal depth $Z_l(\mathbf{x}) \in [d_{\min}, d_{\max}]$, discretized to labels $d \in D = \{d_{\min}, .., d_{\max}\}$ from a set D of depth values.

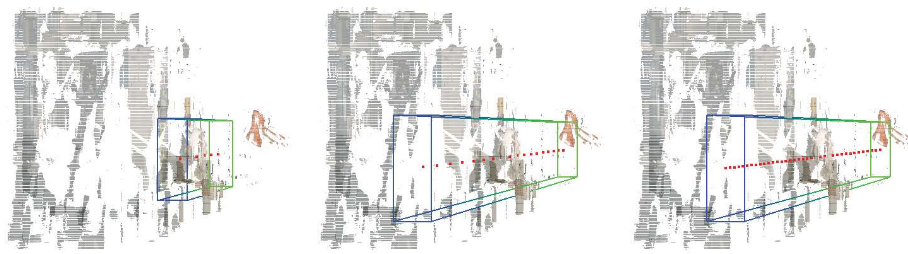
Toward this purpose, a 3-dimensional cost volume $C_l(\mathbf{x}, d)$ for one (e.g., the left) view I_l is constructed. The first two dimensions of C_l are the image size, and the third dimension constitutes the number of depth labels. Each entry within the cost volume is initially a truncated sum of absolute differences (SAD) between the views, using a projection $\pi(\mathbf{x}, d)$ from left to right view based on epipolar geometry from standard calibration [Snavely et al. 06] instead of using disparity.

$$\begin{aligned} C_l(\mathbf{x}, d) = \quad & (1-\alpha)\cdot \quad \min(\tau_1, ||I_l(\mathbf{x})) - I_r(\pi(\mathbf{x}, d))||) \\ + \quad & \alpha\cdot \quad \min(\tau_2, \|\nabla I_l(\mathbf{x})) - \nabla I_r(\pi(\mathbf{x}, d))\|) \end{aligned} \tag{18.1}$$

A value of $\alpha = 0.11$ is used to favor the color term over the gradient term and $\tau_1 = 0.03$ and $\tau_2 = 0.008$ to favor only very exact matches. With the data term set, the next step is to perform a weighted filtering on C_l to arrive at a smoothed cost volume C_l':

$$C_l'(\mathbf{x}, d) = \sum_{\mathbf{x}' \in N_r(\mathbf{x})} W_{\mathbf{x},\mathbf{x}'}(I_l(\mathbf{x}')) \cdot C_l(\mathbf{x}', d) \tag{18.2}$$

The filter weights $W_{\mathbf{x},\mathbf{x}'}$ depend on the guidance image I_l [He et al. 10] similar in spirit to the anisotropic smoothness found in many variational approaches, and are computed on pairs of pixels $(\mathbf{x}, \mathbf{x}')$ in a neighborhood N_r within a filter radius r. Cost filtering is performed on each depth layer, but not between depth layers since there is no guide in the depth direction available.



(a) (b) (c)

Figure 18.5: Cost block within the initial cost volume, visualized as green-blue bounding box; z-layers are shown as red dots in the center. (a) Cost block with a default size in z-direction; (b) the user has increased the z-extent of the cost block; (c) the user has increased the number of z-labels within the cost block.

Runtime is independent of filter radius r (9–24 are sensible radii, depending on image size) when using weighted box filters based on summed area tables, instead of evaluating the weights naively. The OpenCL implementation of [Ruhl et al. 13] uses a tile-based sliding-window variant which works in $O(n)$ on the GPU [Hosni et al. 11a]. Finally, the depth map Z_l is chosen by seeking the depth label with minimal cost per pixel:

$$Z_l(\mathbf{x}) = \arg\min_d C'_l(\mathbf{x}, d) \tag{18.3}$$

Since the results of any stereo algorithm are not perfect, errors cannot be avoided completely particularly on challenging natural scenes which tend to produce artifacts, among them:

Occluded regions. Objects that are occluded differently in the two views can lose significant overlap, preventing unambiguous matching. In a typical stereo configuration, this happens prominently for any object's left and right edges, which are each only visible in one camera. The closer the object is to the camera, the more pronounced the effect becomes. Automated algorithms cannot avoid this error since the information is simply not available; symmetric estimation is at least able to identify occluded regions, but inpainting is still largely heuristic. A human user, on the other hand, is able to provide plausible depth information for those non-visible parts by intuition about the object's shape.

Ill-textured regions. The majority of stereo algorithms for natural scenes (as opposed to controlled lab settings) rely on the color constancy assumption, which may be violated by lighting or camera sensor differences, noise, specular reflectance, translucent objects, caustics, etc. This impedes the matching of objects between both views. Largely uniform or repeating regions in conjunction with different occlusion boundaries in the two views (e.g., columned halls, gratings) are also not solvable with the available information. Again, a human user can assess which objects belong together, and thus distinguish between true and false matches.

The question now is how to integrate human scene understanding in a way that minimizes interaction times. Currently, the most common way is to use image editing tools to select a region via rotoscoping or segmentation, and then use stamp, cloning, and other tools to assign better depth labels.[7] The approach by [Ruhl et al. 13] also starts with a segmentation/mask in 2D image space, but instead of cloning without validation of the resulting depth, a possible range in z-direction is assigned, forming a 3-dimensional "cost block" $K_l \subseteq C_l$ (Figure 18.5). In the first two dimensions, the cost block is a bounding box around the masked or selected pixels and restricts $\mathbf{x}$ to come from $\Omega' \subseteq \Omega$. In the third dimension, the cost block is centered around the median depth of the selection $med(Z_l(\mathbf{x}'))$ with $\mathbf{x}' \in \Omega'$ to expunge outliers (other strategies are also possible) and has some extent that restricts d to come from $D' \subseteq D$. The initial extent in z-direction can either be a fixed parameter or some percentile of $Z_l(\mathbf{x}')$.

In a 3D view of the scene (Figure 18.5), both the current depth estimate and the cost block K_l are visualized. An artist can now shift the cost block along the z-axis until the estimation "snaps" the depth to the most plausible position. With each editing step, $Z_l(\mathbf{x}')$ is locally re-evaluated for all pixels in the mask, providing visual feedback in real-time. The z-extent of the cost block can be widened if objects in the selected area do not fit into it, or narrowed to eliminate superfluous estimates. As a third option, the depth label subset D' can be subdivided to include more depth labels, even to the point where $|D'| > |D|$. This increases the accuracy of z-values but takes longer to compute when the cost block is large.

Using cost blocks does not solve the problem of ill-defined regions in a mathematical sense. Instead, it merely reduces the effect of incorrect cost computation: In a narrowed set of labels d', the cost block $K_l(\mathbf{x}', d')$ merely evaluates to a more plausible depth $Z_l(\mathbf{x}')$, since a search window $I_l(\mathbf{x}')$ has a much lower probability of being matched to a randomly low-cost window $I_r(\pi(\mathbf{x}', d'))$. In the worst case when no support information can be found within filter radius r, the final depth Z_l will be essentially random, but still within the bounds of D'. If implausible, an artist can still narrow down the z-extent of K_l to a thin slice.

In essence, the user interaction cuts away large superfluous blocks from the cost volume, rather than refining the stereo matching itself. Due to

[7] http://www.fxguide.com/featured/art-of-stereo-conversion-2d-to-3d-2012/

Table 18.4: S3D coding standards will facilitate wider adoption of S3D content.

Format/Standard	Properties
Frame compatible coding	Works on existing 2D infrastructure 50% loss of resolution Not a standard
Blu-Ray 3D	Stereo high MVC profile Uses both temporal/inter-view redundancy Backwards compatible to H.264/AVC Standard since 2009
HEVC	Uses both temporal/inter-view redundancy Successor to H.264/MPEG-4 AVC Upcoming standard

the snapping behavior, artists can save considerable effort during depth map correction, and still switch back to depth map painting in case of any remaining failure cases. In the end, a high-quality depth map is available for all further stages of the S3D production and presentation pipeline.

18.7 S3D Delivery

For any type of data delivery, efficient coding and transmission is crucial, most noticeably when streaming individually (and less so when pre-loaded as in a movie theater setting). In order to achieve wide adoption, coding standards are necessary for S3D [Smolic et al. 09], complementing standards such as MPEG or VCEG (H.26x) for 2D movies. A number of coding format standards already exist for some of the 3D representations introduced in Section 18.5, where "representation" determines the format of the data and "coding" refers to the efficient compression of this data. Table 18.4 outlines a selection of current and future formats and standards.

One of the simplest coding formats regarding S3D is frame-compatible coding, whose principle is to embed a S3D stream side-by-side into a 2D stream [Vetro et al. 11]. As input, it takes a two-view stereo video and downsamples each frame by half (either vertically or horizontally). It then combines them back into a single-stream video of the same size as a single-view one, which can then be transmitted and decoded on existing broadcast infrastructure. Because of its simplicity, this approach was used in the first S3D broadcast channels, since only high-level syntax at the receiver is needed to separate the two streams and upsample them again for rendering. However, the loss in resolution is as noticeable as in interleaved 2D video, and is being considered as one of the reasons for the limited success of S3D on home television so far. TV production cost is another factor.

More recent S3D or multi-view video (MVV) approaches encode the full pixel resolution, and exploit spatial (inter-view) as well as temporal (inter-frame) redundancy for efficient compression. Classic 2D video coding makes use of temporal redundancies, i.e., not all pixels will change in neighboring frames, and many movements are regular and can thus be predicted for some number of frames. In S3D video, subsequent frames are not only related temporally but also spatially: Barring occlusion and reflectance effects, objects at the screen plane (same depth) show the same appearance, and objects in front or behind are horizontally shifted [Vetro et al. 11].

Where 2D video uses motion estimation and compensation, S3D allows for additional inter-view disparity estimation and compensation, as e.g., proposed for the 3D-Blu-Ray specification, where the second view typically adds 25-50% bitrate to the first one. Another example is the upcoming HEVC standard, which also features temporal/inter-view prediction [Müller et al. 13].

18.8 Summary

S3D video has come a long way, with a mature production/presentation pipeline and commercially successful products available today. A professional S3D movie ameliorates any adverse side-effects on the human visual system to induce a pleasent viewing experience. While an entire continuum of 3D scene representations exists, efficient hardware, tools, and processes from capture to encoding are most prevalent for two-view S3D videos enhanced by manual modeling. Other approaches are more experimental and do not have cost-effective target hardware at the current time, but are sometimes used for special effects. The perfect theoretical S3D experience featuring full view and focus flexibility is yet to be achieved.

19
Visual Quality Assessment

Holly Rushmeier

19.1 Introduction

The ultimate success of real-world visual computing applications depends on how it is perceived by its users. Integrating an understanding of visual perception into modeling systems and evaluating visual quality of rendered output have always been an essential part of computer graphics research. An example of this from the 1970s is Phong's seminal presentation of a then new reflectance model that used observation of real object highlights as motivation, and a real photograph of a sphere as validation [Phong 75]. Systematic user studies to evaluate and refine methods began to appear in the 1980s, such as Atherton and Caporeal's study of the discretization and shading of curved surfaces [Atherton and Caporael 85].

Visual quality assessment methods build on models from the extensive literature on human visual perception that have been published over more than a century. Perceptual experiments that addressed engineering applications involving human–computer interaction have resulted in useful compendiums of individual models that computer graphics researchers and practitioners can draw on [Boff et al. 88]. For many applications, however, existing human visual models are not yet adequate to address the perceptual questions of computer graphics. There is still a need in many cases to run human subject experiments [Ferwerda 08].

Methods and techniques from the study of human vision are applied both to the development of individual components of visual computing and the development of full systems. Visual assessment determines the accuracy that is needed for geometric reconstruction (Chapter 10) and motion reconstruction (Chapter 11). Visual measures are needed to determine the accuracy of illumination simulations that are used to render the reconstructed geometry and motion (Chapter 6). Alternative assessment techniques are needed for image-based rendering systems (Chapter 17). Image-based rendering systems use natural images to leverage nature's solution to part of the modeling and simulation of a new scene, but may introduce artifacts in the warping and re-sampling processes required to generate new images.

Visual quality assessment for real-world visual computing is problematic because artifacts can arise as the result of individual choices in building a system. To tackle these complex system challenges, traditional models and experiments are gradually being replaced by new tools, new concepts of what aspect of perception is to be measured, the use of crowd sourcing, and the use of new instrumentation.

In this chapter resources are given for perceptual modeling and experimentation that computer graphics researchers can draw on. Examples are given of how visual quality assessment techniques have been applied in illumination simulation, geometric modeling, motion simulation and image-based rendering. Finally, innovative approaches developed in computer graphics for quality assessment are discussed.

19.2 Models from Human Perception

The existing literature in human visual perception is used in two ways in computer graphics. First, computational models of individual aspects of perception are adapted from the biological vision literature to inform algorithms. Second, experimental techniques for studying perception in human subjects are applied to problems that uniquely arise in computer graphics. Human subject experiments are used both to explore new aspects of perception for which models are needed in graphics systems, and to evaluate final rendered results from complex systems.

Computational models used in the design of graphics systems have been drawn from basic models of human vision that were originally developed for printing, photography, and television. Most prominently, graphics has used models of color perception. Radiometry and color have already been discussed in the context of camera systems in Chapter 1. Real-world systems scatter light at wavelengths in the visible spectrum of electromagnetic radiation that ranges from 400 nm to 700 nm. The accuracy needed in spectral values is dictated by human perception of color. Human sensitivity to spectral variation is limited by the response of the eye's cones (at normal daylight conditions) and rods (under low light conditions). Formulae for the conversion of spectral distributions to standard color spaces was developed by researchers in the color and illumination fields in the early twentieth century. Different spectra but with the same coordinates in a standard space (such as the CIE XYZ space) appear identical to humans. In addition to this fundamental insight, there are numerous other models of color appearance that have been and could further be exploited such as color constancy and the change in hue with luminance. Fairchild describes these various effects as well as comprehensive models to characterize color appearance [Fairchild 13].

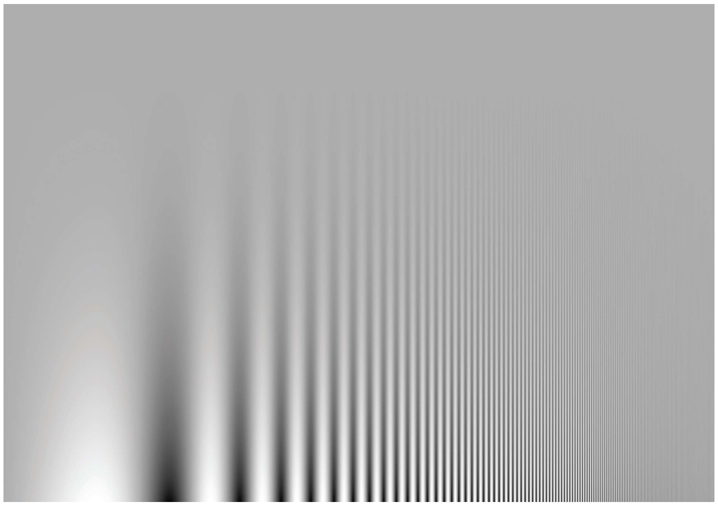


Figure 19.1: The contrast sensitivity function for human vision is illustrated by a spatial variation in intensity, with spatial frequency increasing on the horizontal axis, and contrast decreasing on the vertical axis. The illustration demonstrates that the visibility of a spatial variation is reduced at both very low and very high spatial frequencies.

In addition to color, another fundamental model that applies across many applications is the human contrast sensitivity function (CSF), illustrated in Figure 19.1. Essentially, human sensitivity to contrast is a function of the spatial frequency of the variation, with sensitivity peaking at about 5 cycles per visual degree. This set a lower limit graphics as to how spatially accurate renderings have to be. It is also important for understanding how artifacts that are invisible at one viewing distance become visible at another. For example, for a synthetic texture low spatial variations in intensity are not visible when an image is viewed closeup, but become distracting regular patterns when the image is viewed from a distance where the spatial variations increase in frequency per visual degree. Related to the CSF is the spatial masking effect. The presence of a pattern at a particular spatial frequency can mask other patterns at the same frequency. This is used in various ways to hide artifacts that could be computationally expensive to avoid [Ferwerda et al. 97].

The assessment of overall quality is a long-standing issue in photography and video, and numerous metrics have been developed over the years in the optics, photography, and video communities that combine basic perceptual models. A particularly influential metric in the computer graphics community has been daly's visible difference predictor [Daly 93]. Daly's model combines models of human sensitivity to luminance, contrast sensitivity, and masking to predict the perceivable differences in two images. This model has been used to drive iterative rendering methods [Myszkowski 98] and has served as a model for constructing image-level metrics from low-level vision models.

19.3 Experimental Techniques from Human Perception

Well-established techniques have been developed for running experiments to draw valid conclusions with respect to individual human perception phenomena. An overview of experimental methods and analysis for computer graphics have been summarized by Ferwerda [Ferwerda 08]. Different techniques are needed depending on whether threshold (just noticeable effects) or super-threshold (scaling effects) problems are being studied.

The fundamental techniques still in use today were developed in the nineteenth century. Gustav Fechner (1801–1887) described three ways to approach threshold effects. One, the method of adjustment, asks users to adjust a test display until it matches a standard. A second, the method of limits, displays pairs—a standard and a test. The test display is adjusted to approach the standard, and the user's comparison of the two, in the form of "is the test more or lest *foo* than the standard" (where *foo* is the effect being tested) is recorded. In the third, the method of constant stimuli, random test displays are shown along with the standard and the user's responses are recorded. Each approach has its inherent limitations and requires some different analysis. Super-threshold (scaling) effects can also be measured by various techniques including indirect rating, pair comparison, ranking and category scaling, as well as direct scaling. These methods range from asking directly for a rating of an impression (on, say, a scale of 0 to 100), to making multiple pairwise comparisons and using Louis Thurstone's (1887–1955) "law of comparative judgment" to derive a scale.

In general, before setting up a study, an investigator should either consult an experimental psychologist, or should carefully read the descriptions of many previous studies. The *ACM Transactions on Applied Perception* is an excellent source for descriptions of previous work, since it is particularly oriented to obtaining results that can be applied to computing. An investigator needs to pay attention to the controlled conditions and materials used, conducting the experiment with a protocol and ordering that prevents bias, and appropriate statistical analysis of the results. Controlled conditions require that calibrated display equipment is used. The experimental protocol typically includes randomized ordering of tests for the various subjects to avoid bias.

In the area of statistical analysis of human subject studies, a frequently misunderstood point is the number of subjects required by a study. As a rule of thumb, the study of very low-level effects, such as contrast perception for a sinosoidal display, require a small number of subjects. Higher level effects, such as color preferences, require larger numbers of subjects. The rationale for these numbers is that significant low-level effects are the same across the human population, and if they are found consistently for a small number of observers, they can be expected to be found for the general population. For higher level effects there is no magic number that is "enough" or typical. The adequacy of the number of subjects is indicated by the statistical significance computed for the results.

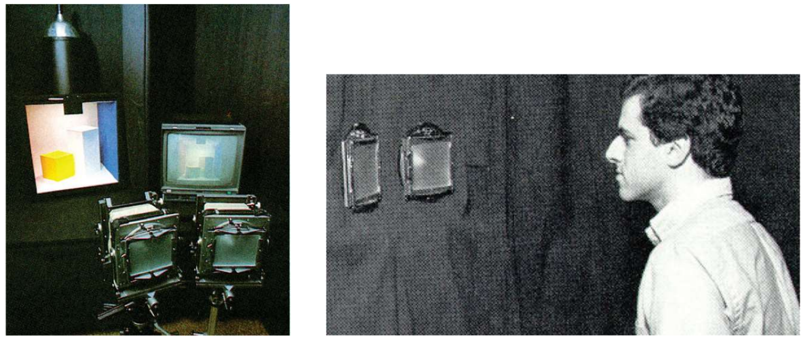
There is synergy between computer graphics and research in human perception. Computer graphics has made possible controlled conditions for investigating various types of perceptual phenomena. One result of this synergy is that a computational tool box has been developed for psychophysical experiments [Brainard 97]. This MATLAB toolbox[1] and lower level code is a long-term stable resource that offers many practical features for the mechanics of setting up an experiment.

Even simple experiments that ask observers to look at pictures and hit a button to respond are human subject experiments. Different procedures for being permitted to do such experiments apply in different countries. In the United States, it is typical that an experimental protocol that documents the measures to ensure subjects' safety and privacy must be submitted to an institutional review board (IRB) before the experiment can be conducted.

19.4 Evaluation in Lighting and Material Modeling

Perceptual methods have been used to extensively evaluate the accuracy of global illumination techniques. An early example is the Cornell box experiment conducted to evaluate radiosity solutions and color calculations [Meyer et al. 86]. The set up for this experiment is shown in Figure 19.2. Rather than establishing a scale or threshold, this work was concerned with whether an observer could reliably discern a simulation from a natural image of a simple scene. The experimental design was the classic "two alternative forced choice." Presented with two images, the observer was forced to make a selection of which was the simulated image. Obvious cues were eliminated by having the observer view both the 3D scene and the image on a display through view cameras. Observers choosing the simulation over the original approximately fifty percent of the time, the same as guessing, indicated the success of the simulation.

[1] ttp://psyctoolbox.org/



(a) (b)

Figure 19.2: The Cornell box experiment was conducted to evaluate global illumination solution techniques. In (a) a real scene and its visual simulation are shown. Human subjects saw the real and simulated scenes through view cameras, as shown in (b). The use of the view cameras made it possible to obscure the obvious cues that indicated which was the real scene. Images from [Meyer et al. 86].

A more extensive dataset for evaluation was provided by Myszkowski et al. in "the Atrium study" [Drago and Myszkowski 01]. Both the Cornell box and Atrium used side-by-side comparisons and were focused on evaluating the ability of particular global illumination techniques in reproducing a natural image. An alternative approach is task-oriented evaluation of real and synthetic scenes introduced by McNamara et al. [McNamara et al. 00]. In a task-oriented evaluation, a user is asked to do the same task in physical and simulated environments. The goal is to find whether the user makes the same choices, whether correct or incorrect, in both the physical and simulated environments.

Local illumination models describe how individual materials scatter light. Perceptual studies to inform local models include a perceptual space for gloss [Pellacini et al. 00] and more recent studies of perceptual material evaluation [Fleming 14]. Pellacini et al. were able to establish a low dimensional gloss space by means of a scaling experiment where observers used sliders to indicate the magnitude of gloss difference between two renderings of a sphere on a checkerboard [Pellacini et al. 00]. Defining and parameterizing a perceptual space for more general scattering properties has proved more difficult because of the high dimensionality of the space and the infeasibility of asking observers to do the required very high number of comparisons. Furthermore, Fleming's more recent work suggests that humans don't recognize materials by estimating the parameters of scattering models. He suggests that human perception works with "statistical

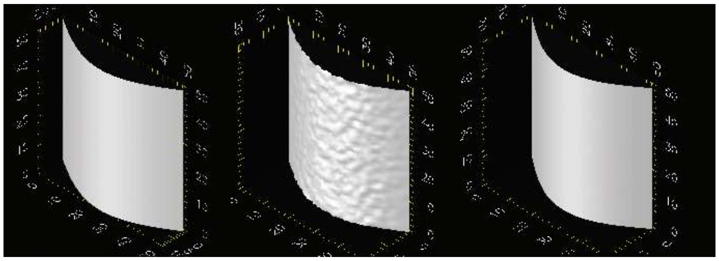


Figure 19.3: Positional accuracy in geometry is not necessarily equivalent to visual accuracy. On the left is a reference surface. The center surface is a vertex-by-vertex more accurate approximation of the reference surface than the surface on the right. However, the distribution of positional errors on the center surface introduce objectionable visual artifacts that are not visible on the right.

appearance models." Understanding what image appearance features are important in judging different classes of materials, and relating modeling parameters to the formation of these features will require much more experimental investigation.

Even if narrowing from the general problem of material appearance to the specific instance of rendering faces, there is still much more experimental work to be done. Fan et al. asked observers to classify each face in a collection of natural and computer generated images whether they were photographs of actual people, or simulations [Fan et al. 12]. In analyzing the results, Fan et al. found variability in the influence of fundamental rendering components such as color, small-scale geometric detail, and gloss. The study demonstrates that there are subtle interacting effects in modeling reflectance, geometry, and illumination that affect the perception of faces that are not yet well understood.

19.5 Geometric Modeling

Following Atherton and Caporeal's model study of geometric representation [Atherton and Caporael 85], studies have been performed to understand the requirements for modeling shape for realistic rendering. As shown in Figure 19.3, higher local positional accuracy doesn't necessarily produce visually more desirable results. Representations that introduce or remove salient, i.e., attention drawing, features are both objectionable.

Real-time rendering typically uses triangle meshes because of their compatibility with the hardware rendering pipeline. The 1990s and early 2000s saw a large volume of work on different techniques for mesh simplification to increase rendering efficiency, and this motivated the need for measuring mesh quality. Innovative work by Watson et al. used "naming time" to study simplification methods [Watson et al. 00]. Instead of comparing side-by-side renderings, model quality was judged by how long it took observers to name the object being represented. "Naming time" is a useful measure in that it does not require the user to compare directly to some "ideal" rendering of a model. In another alternative to side-by-side comparisons, Howlett et al. used gaze tracking to understand the saliency of mesh features, and the effect of simplification of more salient features [Howlett et al. 05]. Subsequently Kim et al. used eye tracking to evaluate various computational models of mesh salience [Kim et al. 10].

Finally, material properties and lighting both affect shape perception [Ferwerda et al. 04], meaning that errors in materials and geometric representations are related.

19.6 Motion

Perception has been used to understand acceptable limits in motion simulation and its presentation. It is well known that humans are sensitive to motion, and that simple animations showing bright dots moving can convey ideas such as a human walking [Hodgins et al. 98]. Since fully realized animation is expensive to compute, previewing a motion sequence in simplified form is desirable. However, in an early paper on human animation and perception Hodgins et al. presented human subject experiments that showed that people are able to detect more subtle variations in motion with full polygonal models [Hodgins et al. 98]. Doing side-by-side comparisons is even more difficult for animations than for static images. Experiments need to be designed so that they don't require too many judgments at a time and don't require judgments that are difficult for the experimenter to convey. In Hodgins et al.'s work the observers were asked to compare the motions of figures rendered with the same geometric models and to assess when they were able to detect subtle differences in motion. In later work in the same area, Hodgins, working with colleagues in neuroscience, found that there is a relationship between the perceived realism or naturalness of character motion and a character's anthropomorphism [Chaminade et al. 07].

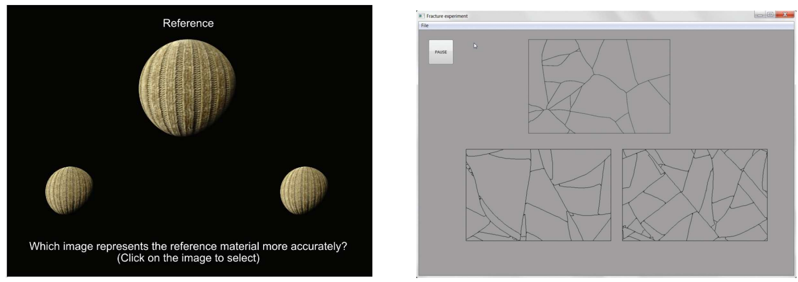
The perception of individual character motion is different from the perception of motion in crowds. From practical industrial experience, realistic large crowd animation does not require every character to be distinct. Instancing is a heavily used technique to reduce computational expense. McDonnell et al. studied the visibility of instanced characters, or clones, in a crowd [McDonnell et al. 08]. The technique used was to present observers with sets of characters that were either appearance clones (i.e., same shape and clothing) or motion clones (walking with the same gait). The time for observers to spot the clones was recorded. It was found that motion clones were much harder to detect than appearance clones.

While humans are very sensitive to nuances in human and character motion, they can also detect errors in the physical motion of inanimate objects. O'Sullivan and Dingliana studied human sensitivity to the accuracy of collisions of simple objects [O'Sullivan and Dingliana 01]. In this work, observers were asked to make judgments about collisions between spheres. For example, they were shown collisions where varying gaps were left between the spheres rather than showing the spheres actually touching during the collisions. The observers were asked whether the spheres had touched during the collision. The result that the observers reported the spheres touching in cases where there was actually a gap suggested a level of detail algorithm for detecting collisions that allowed for error that would not be detected visually [O'Sullivan and Dingliana 01].

19.7 Image-Based Systems

Image-based systems rely on natural images for many component elements to be correct (Chapter 17). However, the source input images frequently need to be warped and resampled to generate new images. Techniques are needed to see if the operations on the images introduce visual artifacts.

An example of the evaluation of an image-based rendering system is the study by Vangorp et al. of perspective distortions in a virtual tour along city streets [Vangorp et al. 13]. The study outlined a model for how virtual tours are created. Multiple images of building facades are captured as a vehicle moves down a street. The images form a panorama which is then displayed for arbitrary views along the route by projecting the images onto planar proxies for each of the buildings. The problem is that if images aren't captured closely enough, the projected images will look obviously incorrect, with corners of buildings that should have faces meeting at 90 degrees making the distortions most obvious. The experimental procedure used had observers indicate their estimate of the angle of corners presented to them by adjusting the relative orientations of two boards on a hinge, rather than asking the observers to give a number. This allowed the observers to report their judgments more directly. The result of the study was a recommendation for how closely images needed to be captured given a particular depth variation for buildings along the street.



(a) (b)

Figure 19.4: Examples of tests for visual equivalence. In (a) filtering techniques for bidirectional textures are tested for the visual equivalence of filtered and downsized images of an object covered by a complex material (image from [Jarabo et al. 14]). In (b) the visual equivalence of crack patterns are tested relative to a reference standard (image from [Glondu et al. 12]).

19.8 Beyond Classic Models and Experimental Techniques

Approaches for modeling and evaluation have in many ways moved past traditional perceptual models and psychophysical experiments. New ideas for visual assessment that have emerged in computer graphics include no-reference metrics, visual equivalence, crowd sourcing, and novel measurement techniques.

In general, the ideal image or animation is not available or is too expensive to compute. An assessment method that doesn't rely on a comparison is desirable. The idea of "no reference" metrics addresses this issue. For assessing the quality of a rendering given reliable geometric, material, and motion representations, a no-reference metric was developed by Herzog et al. [Herzog et al. 12]. The metric is built using machine learning on a database of good and bad renderings. The quality of the rendering is learned as a function of the data that a rendering algorithm has easily available—depth information and diffuse textures.

Taking the full problem of determining the quality of input representation and accuracy needed of the rendering technique, an important issue is that accuracy does not mean pixel-by-pixel equivalence to the "perfect image." It is a rare, and perhaps never occurring case, that images are needed that are pixel-perfect. Instead, what is needed is a picture that represents the same thing, made of the same stuff, in the same environment. Ramanarayanan referred to this type of similarity as "visual equivalence" [Ramanarayanan et al. 07]. For example, from work by Jarabo et al. [Jarabo et al. 14], an object covered by material rendered at two different spatial scales is shown in Figure 19.4(a). The method for filtering the material texture is being tested. It is not expected that the images are pixel-identical because they are at different scales. It is necessary, however, that a successful filtering technique produces an image that appears to be the same object made of the same material—that is, that they are visually equivalent. Another example, from Glondu et al. is shown in Figure 19.4(b) [Glondu et al. 12]. In this example, it is desired to create a material that has been cracked the same as an exemplar. An experiment was conducted to ask observers whether the cracks on the left or right looked like they were from the same process on the same material—that is, whether they were visually equivalent. In this second case, statistics were obtained to characterize when patterns were visually equivalent, and this was used to drive a simulation to create virtual materials that appeared to be equivalent to an example from a natural photograph.

A "weak link" in most perceptual experiments is that observers have to give an active response—verbalizing a response or hitting a button. Recent work has exploited alternative, passive techniques for measuring response. Lindeman et al. used electroencephalograph (EEG) [Lindemann et al. 11] to measure an observer's response to artifacts in image-based rendering. Advances in hardware that have made EEG economically accessible, and in signal processing techniques to deal with the noise in the EEG signal have made this a viable technique. Mustafa et al. demonstrated that certain types of video artifacts can produce consistent EEG responses [Mustafa et al. 12].

Another problematic issue in conducting experiments is recruiting an adequate number of diverse observers. For some cases crowd sourcing is a viable approach [Heer and Bostock 10]. Heer and Bostock found that for some basic perceptual experiments crowd sourcing produced the same results as controlled experiments. They hypothesized that the larger number of subjects that are possible in a crowd-sourced experiment offsets the lack of control in experimental conditions. In a sense, with a large group the diverse conditions are "averaged out" by the large number.

19.9 Summary

There is a long history of using perceptual insights and human subject experiments to create realistic images and evaluate them. Past work has built a lot on performing experiments in laboratory-controlled conditions, with an emphasis on traditional approaches such as two-alternatives-forced-choice. Recent work has moved beyond this to systems that predict perceptual impact by validated metrics that encompass a wide range of effects, the use of crowd sourcing, and by using systems to gather user responses that do not require active, conscious, user input. The problem, though, of producing and assessing realistic imagery remains challenging because of the non-linear, convoluted interaction of many perceptual effects simultaneously.

Part V

Applications

20
Facial Capture and Animation in Visual Effects

Darren Cosker, Peter Eisert, and Volker Helzle

20.1 Introduction

In recent years, there has been increasing interest in facial animation research from both academia and the entertainment industry. Visual effects and video game companies both want to deliver new audience experiences—whether that is a hyper-realistic human character [Duncan 09] or a fantasy creature driven by a human performer [Duncan 10]. Having more efficient ways of delivering high-quality animation, as well as increasing the visual realism of performances, has motivated a surge of innovative academic and industrial developments.

Central to many of these developments are key technical advances in computer vision and graphics. Of particular note are advances in multi-view stereo reconstruction, facial tracking and motion capture, dense non-rigid registration of meshes, measurement of skin rendering attributes (e.g., BRDFs for skin and skin subsurface scattering models), and sensing technology.

This chapter builds on concepts already described earlier in this book—such as 3D capture, rigging, and non-rigid registration—and takes a more practical look at how they might typically be applied in visual effects. First, methods and applications for facial static capture and rendering are considered, before dynamic capture is addressed. Finally, a case study is examined called *The Gathering* involving the creation of an animated face from animation to final composite.

20.2 Static Facial Realism and Capture

In today's world, static facial realism has reached a level where humans cannot distinguish 3D facial models from real photographs anymore.

Technology such as the Light-Stage [Ma et al. 07] allow the capture of highly detailed facial surface information. Coupled with sub-surface reflectance data [Donner et al. 08] facial models now display photo-realistic likenesses to real faces. Such technology is now widely used in modern motion pictures, e.g., *Spider Man 2* [Fordham 04] being one recent high-profile use of the Light-Stage. In these circumstances, the high-detail static scans are composited onto either a stunt-actor's body, or a digital double. This is where the actor is to be placed in situations that may not be practical or safe (such as explosions). However, it is still the case that the actor's expression is typically static in these situations, and close examination of such shots reveal a dead-like facial quality. An early high-profile example of facial replacement was in the *Matrix* sequels [Borshukov et al. 05], where passive facial scanning was used to obtain 3D faces with high detail facial texture. An important aspect of the use of facial scans for movies and video games is that faces must be renderable in a wide range of environments so that the face can be convincingly composited into the overall scene. Therefore, the UV map (texture data) is typically diffuse albedo. Skin detail is enhanced through high-resolution normal maps [Ma et al. 07] or geometry [Beeler et al. 10], and rendering is enhanced through sophisticated BRDFs/BSSRDFs modeling material properties [Donner and Jensen 06, Jensen et al. 01]. Acquisition of such reflection properties has advanced widely over recent years, resulting in highly detailed rendering [Donner et al. 08].

An important aspect to the realism of synthetic humans is the realistic rendering of hair, which has made significant progress in the last years. Single hair fibers have been modeled [Marschner et al. 03] as semi-transparent elliptical cylinders. By defining surface reflection as well as scattering inside the hair, the complex lighting characteristics of real hair with its view dependency, highlights, and color changes can be accurately reproduced with moderate rendering complexity [Ren et al. 10]. The possibilities to model several fibers up to a complete hairstyle range from NURBS surfaces via thin shell volumes to strain-based modeling by parameterized clusters, fluid flow, or vector and motion fields [Ward et al. 07]. In order to simulate the complex lighting interaction between strands of hair, Lokovich et al. [Lokovic and Veach 00] propose a deep shadow map which relates visibility to depth for each pixel, yielding realistic but computationally expensive self-shadowing. Approximation algorithms for making the simulation of multiple scattering among hair fibers tractable have been proposed using methods like photon mapping [Moon and Marschner 06b] or spherical harmonics [Moon and Marschner 06a].

Whereas laser scanning technology was initially the most accurate way to derive static facial detail, passive scanning technology using consumer hardware is now popular [Beeler et al. 10, Blumenthal-Barby and Eisert 14]. Multiple consumer-level SLR cameras are used to acquire high-detail images which provide strong features for stereo matching algorithms (Chapter 8), and can result in captures with skin pore (mesoscopic-level) facial details [Beeler et al. 10]. One aspect to consider when using such data is practicality, as the meshes can contain millions of vertices. This is a different approach to those methods currently considered in, for example, movies where low polygonal meshes are used along with high-detail normal maps to display facial meso-structure. There is therefore still a great deal of work to be done on using such technology practically for video games and modern VFX.

Many state-of-the-art stereo and multi-view approaches are local in the sense that they reconstruct the 3D location, and sometimes orientation, of isolated image patches [Furukawa and Ponce 10]. While this strategy is beneficial for parallelization, it requires a post-processing stage to generate a mesh. The reconstruction yields a point cloud with outliers which has to be filtered and meshed with appropriate algorithms (Chapter 10), such as Poisson meshing [Kazhdan et al. 06]. Smoothness priors are often only considered at the meshing stage. Local reconstruction is difficult to combine with efficient interactive tools. As each patch is unaware of its neighbors, the correction of a single mismatched patch by the user will not affect its neighbors, although they are likely to be erroneous as well. Therefore, [Blumenthal-Barby and Eisert 14] follow a similar approach as [Beeler et al. 10] but uses mesh-based deformable image alignment for the reconstruction of high-detail face geometry (including hair) from two or more SLR cameras (Figure 20.1). Instead of iteratively matching small image patches along the epipolar line, an entire view is warped to target views in an uncalibrated framework incorporating a mesh-based deformation model. The additional connectivity information enables the incorporation of surface-dependent smoothness priors and optional user guidance for robust and interactive geometry estimation [Schneider and Eisert 12].

Figure 20.1: Static reconstruction of the head including hair from two images [Blumenthal-Barby and Eisert 14].

Most digital face replacement in movies involves static face replacement, with the actor having little or no movement in facial expression. Although this might be satisfactory for a few frames, as soon as the face moves, or the shot continues for more than a few seconds, this illusion becomes hard to maintain. In the next section, the movement of faces is considered, especially with respect to maintaining an illusion of realism.

20.3 Dynamic Facial Capture and Animation

The holy grail of facial animation research is the portrayal of characters indistinguishable from real humans. This is extremely difficult since humans are experts in detecting the slightest flaws in faces. Even minor defects can break the illusion of realism. In the previous section, static faces are considered where realism has reached a point where it is impossible to distinguish computer graphics from real photographs. However, in order to display a synthetic human that is truly life-like, the movement of the face remains a major challenge.

Arguably, it is easier to convey dynamic realism in the play-back of actual recorded performances than to author a new animation. In order to highlight this, the acquisition of dynamic 3D facial sequences (termed here as 4D for brevity) is first considered.

There are now many commercial companies that market 4D facial capture systems, i.e., those that can obtain 3D mesh data at video recording rate (e.g., Dimensional Imaging,[1] 3DMD[2]). However, the focus here is primarily on academic research in this area. One of the first compelling uses of dynamic facial capture in movies was in the *Matrix* sequels [Borshukov et al. 05]. A passive stereo capture system was constructed where 3D mesh data can be acquired from a face at video rate along with high-resolution texture. The recorded sequences were then composited onto the actors in key action sequences. An extension of this system called *Universal Capture* was later used in several Electronic Arts (EA) promotions and video games (e.g., Tiger Woods Golf [Borshukov et al. 03]). Here, the system was made more robust by adding markers to the actor's face. This could be used to stabilize and track a canonical mesh (i.e., mesh with a known topology) through the captured sequence. Bickel et al. adopt a similar approach with the addition of extra facial paint to appropriately capture wrinkles on the face [Bickel et al. 07].

The use of markers has overcome previous issues related to tracking a face mesh using optical flow. Such methods are notorious to drift, caused

[1] http://www.di3d.com
[2] http://www.3dmd.com

primarily by fast facial changes, for example during speech. Early approaches to avoid the drift in markerless tracking over longer sequences are the incorporation of additional constraints from the silhouette [DeCarlo and Metaxas 00] or the use of an analysis-by-synthesis estimation as in [Eisert 03]. Both methods ensure that estimates are referred to a global reference and avoid error accumulation over time. This approach is also followed by Bradley et al., who propose a multi-view stereo capture system comprising 14 HD cameras mosaiced together [Bradley et al. 10]. This results in a highly detailed set of images upon which to apply optical flow for mesh tracking. Referencing the initial frames of the sequence results in improved mesh stabilization over time. Expanding further on this work, Beeler et al. introduced the concept of anchor frames for stabilizing 4D passive facial capture [Beeler et al. 11]. In this work, neutral frames in the sequences are searched for and then used to essentially reinitialize mesh tracking where possible. This also has the added benefit of offering robustness to certain facial self-occlusions (e.g., as caused by the lips). Although having a lower geometric resolution than previous, passive static capture work [Beeler et al. 10]—which includes approximated skin pore geometry—the extension to 4D including the impressive temporal mesh coherence is a high current benchmark in contemporary facial capture research and development.

One highly successful recent demonstration of the use of 4D capture in industry is from the video game LA Noire.[3] Hundreds of hours of actor footage were recorded in a controlled lighting environment. Key 3D character scenes were then composited with the volumetric facial performances resulting in highly detailed and realistic results. Another high profile use of 3D technology for industrial use was by Alexander et al. [Alexander et al. 10]. The Digital Emily Project was a collaboration between Imagemetrics and USC using Light-Stage technology to capture high detail normal map and surface reflectance properties from an actor's face [Hawkins et al. 07]. A facial blendshape rig was constructed from captured 3D data and then matched to the performance of the actor using proprietary Imagemetrics markerless facial capture technology. Blendshapes are facial poses of different expressions—from stereotypical (happy, sad) to extremely subtle (narrow eyes). The term *rig* is used to describe the complete facial model with all its control parameters. The degrees of freedom of the facial rig are a function of the number and complexity of the blendshapes, and new facial poses are created by combining blendshapes with different weights (Chapter 13). More recently, the Digital Ira Project [von der Pahlen et al. 14] demonstrated how high levels of static and dynamic realism could be animated and rendered in real-time. Thirty high-resolution facial scans were

[3]http://www.rockstargames.com/lanoire

captured using the new Light-Stage X system [Ghosh et al. 11], providing data for a facial rig. Video performance was then captured of an actor and used to animate the facial performance, combined with sophisticated real-time rendering of multiple effects [Jimenez et al. 12].

While the LA Noire production is a high-profile use of 4D capture, it is essentially playback of the recorded data. On the other hand, the Digital Emily and Ira projects demonstrate a degree of performance-driven animation, or retargeting. This type of animation is highly popular in academia and industry, where a performer animates a *puppet* via motion capture or speech (audio only or phonemes). In the case of the both the Digital Emily and Ira projects, the rigs are tracked and animated directly from the actor reference video footage. However, in other cases it is often necessary to retarget between two different rigs one created as a likeness to the actor's face (which is tracked to the input performance) and a second (often a creature or non-human character) animated from output controls of the first rig. In such cases, a rule-based or example-based mapping must be learned between the two rigs. This is a current active area of academic and industry research [Bhat et al. 13]. While it is not the intention of this chapter to give a detailed review of retargeting methods, the excellent course material in [Havaldar et al. 06] encompasses many of the ideas in this area still used today. The aim here is rather to make the distinction between direct playback of captured volumetric animation and the creation of realistic character animation given some reference (e.g., actor performance). However, one important point to make is that even given the best tracking or analysis of a human performer, *automatic* animation of a rig to a level satisfactory for visual effects is still an open problem. Typically, after automatically animating a face in this way, an artist is still required to spend considerable time matching and adding secondary rig movements to the reference performance. In video games, this process can be a hindrance—where hundreds of hours of generated performance may be required for delivery under a short time constraint. In this scenario, a lower level of quality than VFX may therefore be acceptable, as generating VFX quality for current video game productions would add an unrealistic burden on third-party facial animation production or in-house game studios.

The movie *The Curious Case of Benjamin Button* [Duncan 09] contains another successful example of human realistic performance-driven animation and retargeting. MOVA[4] performance capture technology was used to collect 3D scans of Brad Pitt's face and used for blendshape rig construction. Animation was then carried out with the aid of markerless performance mapping from reference footage of the actor. The movie *Avatar* [Duncan 10] also pushed forward the technology of facial

[4]http://www.mova.com

performance capture and retargeting. Although the characters were not human, the movie demonstrated that modern techniques involving motion capture and artistic input could be used for producing large volumes of high-quality performances. The production involved the use of head-mounted cameras, targeted at the actor's face for recording the movements of painted markers. These movements were transferred into a combination of blendshapes per-frame, and the resulting animation used to *block out* an initial animation as a first pass for artists—who later edited and enhanced the performance with the aid of additional video reference (akin to the method previously described earlier in this chapter).

While marker-based motion capture techniques are widely popular e.g., using commercial optical capture systems or painted markers (e.g., Vicon's CARA system[5]), markerless methods provide the potential to capture areas of the face where marker placement is too obtrusive. In addition, it raises the possibility of obtaining a dense capture field for the face, for example based on skin pores. Where the facial rig is based on blendshapes [Havaldar et al. 06], the aim is to optimize a set of weights that approximate the positions of the markers. In marker-less systems, such markers might be located using image-based deformable tracking techniques such as active appearance models [Cootes et al. 98]. Another alternative is to fit the blendshapes to 4D surface data [Weise et al. 09, Valgaerts et al. 12]. This latter method has also been shown to work with consumer 2.5D capture devices such as the Kinect [Weise et al. 11]. However, whether this technology alone can provide the fidelity required for VFX to move beyond optical or marker-based methods remains unclear. It may therefore be sensible in the future to consider a combined approach: markers, high quality RGB, and depth sensors.

In the examples so far, facial dynamics have been captured and replayed, often with considerable artistic manual intervention [Alexander et al. 10]. However, the concept of using such data to author entirely novel performances without reference footage remains a difficult challenge. The success of such methods still largely depends on artistic talent. Advances in interactive facial models, and new methods to create efficient rigs, are promising avenues for improvement. In the last part of this section, some recent advances in blendshape rig construction are briefly considered that could help animators use performance capture data more efficiently and provide better artist tools.

One challenge is how to create effective blendshapes. A standard approach in modern VFX is based on action units (AUs) from the facial action coding system (FACS) [Ekman and Friesen 78], e.g., in *Monsters House* [Havaldar et al. 06] and *Watchmen* [Fordham 09]. Having a FACS

[5]http://www.vicon.com/System/Cara

basis can potentially provide a mapping between different facial rigs. This can be especially useful if one blendshape model is based on actor facial scans and fitted to an actor, and then the weights are transferred onto a puppet model, perhaps of a creature. More recently, Li et al. considered creating blendshape rigs given only a few example expressions and a generic blendshape rig with a wider number of expressions [Li et al. 10]. Such systems can potentially reduce artist time when manually sculpting blendshapes for rigs, and also for reusing existing blendshape models when new rig creating is required. Facial rigs in movies can potentially become very large, with hundreds of blendshapes for *hero rigs*, i.e., rigs required to deliver close-up expressive performances [Fordham 03]. Any technique for increasing efficiency is therefore of high value to industry.

Facial animation bases, or blendshape bases, are also not restricted to artistically sculpted facial expressions or captured 3D scans. Principal component analysis (PCA) also offers a basis for animating faces. However, although this basis is orthogonal—meaning that each expression has a unique solution with respect to the basis—these are often not intuitive enough for artistic animation. In order to address this, Tena et al. [Tena et al. 11] recently proposed a region-based PCA modeling approach that allows more intuitive direct manipulation of local facial regions. Their method also highlights how solving for expression weights locally can provide better approximation of motion capture data. Ultimately however, what an artist will desire of the facial model is a set of controls that are both intuitive and also orthogonal such that altering one expression does not interfere too much with others. To counter this, blendshape rigs become highly complex, with additional shapes (corrective blend shapes) included to counter interference cases. In an ideal world such correctives would not be necessary given the extra work burden they impose, and future work is still required to address this core problem.

Given the discussion of static and dynamic facial capture, the next section considers a case study where facial models based on the likeness of real people were animated and composited onto real footage. This builds on many of the ideas expressed previously in this chapter, and also highlights many of the practical and real-world constraints such a project imposes on animators and technical directors.

20.4 Case Study: The Gathering

The Gathering is a final year short film of Filmakademie Baden-Wuerttemberg.[6] The protagonists in this short film were created digitally

[6]http://www.svendreesbach.com/the-gathering/

based on photo references. The process relied completely on artistic skills since no real reference for 3D scanning or reflectance measurement could be employed. The budget and time constraints of the project demanded to create all assets digitally.

Once the digital models were completed, their facial animation rigs were created by applying the *Adaptable Setup for Performance Driven Facial Animation*[7] [Helzle et al. 04]. This extension for Autodesk Maya[8] enables a rigging artist to apply a generalized library of muscle group movements conforming to the FACS [Ekman and Friesen 78] system to any humanoid geometry. The deformations are driven by a dense data model which includes the non-linear characteristic of facial actions. Compared to static modeling and interpolation of blendshapes, this approach allows for fast and flexible control over the individual facial deformations. The toolset allows complete control in how this data is applied and adjusts to the physiognomies of the geometry. The rigging artist has to manually apply facial landmarks that drive the deformation. The approach has its limitations mainly with respect to the amount and influence of the 69 deformation objects. One way to overcome this limitation was with the use of a limited number of corrective blendshapes, i.e., new blendshapes that trigger other key blendshape combinations to alleviate unwanted or unnatural movements.

Custom extensions to the toolset allow controlling the stickiness of the lips as they part when speaking. This effect is due to moisture on the lips, causing them to open from the inside to the outside as the lips part. Furthermore, the effect of the eyes bulging the upper or lower lids as the gaze changes was realized using a complex constellation of additional deformation rig objects. A fast animation rig allowed quick iterations when animating the sequences and kept the animation artists motivated. All facial animation was realized by rotoscoping the movements recorded from the real actors on set. The head movement could be extracted by rigid body tracking of the markers applied to the actors' heads as shown in the top left of Figure 20.2.

The animated models were rendered using Newteks Lightwave 3D[9] software package. The top right of Figure 20.2 shows the raw rendering which included additional information like motion vectors, reflection and diffuse values embedded inside Open-EXR files, which were provided for compositing. The final compositing was accomplished in The Foundry's Nuke[10] software, integrating the CGI elements into the plates (lower left of Figure 20.2)

[7] http://fat.research.animationsinstitut.de
[8] http://www.autodesk.com/maya
[9] http://www.lightwave3d.com/
[10] http://www.thefoundry.co.uk/nuke/

Figure 20.2: The Gathering: Case study on facial animation. ©Filmakademie Baden-Wuerttemberg, The Gathering, 2011.

before final coloring and additional effects like cigarette smoke were added (lower right of Figure 20.2).

The Gathering shows that it is possible to create convincing digital faces by relying mostly on artistic skills and powerful tools for facial animation. However, it also highlights that the complexities of creating a face and its movements digitally demand a wide set of skills.

20.5 Summary

Capturing real faces with modern technologies like Light-Stage provides 3D face models with high geometric detail and sophisticated material properties that enable face synthesis that is almost indistinguishable from a real picture. While static face models can be used for replacing an actor's face for a few frames, often dynamic face capturing is also desired, which is still a challenging task due to the sensitivity of a human observer for subtle inconsistencies in facial motion. Facial dynamics are usually modeled by a blendshape rig, either from scans, multiview images, or manual work of an artist, while animation data is often derived from marker-based or, more recently, markerless motion capture systems. As shown in the presented case study, convincing digital faces can be animated with such techniques. However, creating realistic facial animations and models still requires significant manual work by artists, leaving room for novel algorithms and toolsets to simplify and automate the process. In terms of research challenges, there is still a large scope for further work in this area. Creating rigs is a time-consuming process, and methods to automate this at production quality are highly desirable. Another core area for future work is in the retargeting of actor faces to new creatures. One challenge in achieving this aim, however, lies perhaps in the contrast between the academic and industrial worlds: academia often has the time to focus on algorithms that could solve this problem while lacking the complex rigs required to test their ideas. Conversely, industry has the expertise to produce such rigs but given practical movie constraints often does not have the time to focus on the algorithms. It is therefore unsurprising that the best advances in this area have been from academic and industrial collaboration.

21
Television and Live Broadcasting

Graham Thomas, Philippe Bekaert, and Robert Dawes

21.1 Introduction

A popular current use of visual computing in live television broadcasting is in the production of graphics and effects for sports analysis. Applications can help the broadcaster illustrate, analyze, and explain sporting events by the generation of images and graphics that can be incorporated in the broadcast, providing visual support to the commentators and pundits. While currently limited to sport, in the future these techniques might spread to other areas of TV broadcasting, to other genres and for use in other aspects of production. Some broadcasters are already experimenting with free-viewpoint, panoramic, and omni-directional video to see how these technologies might change the traditional methods of live TV production.

This chapter looks at some of the challenges involved in developing visual computing tools for use in television coverage and at the techniques that some example systems have developed to tackle them.

21.2 Sports Graphics

Sports directors often refer to "telling a story" to the viewer and are keen to use the best tools available to bring out the key points in a clear, visually interesting, and succinct way. Graphics are a key tool for the TV program maker to help explain sporting events to viewers. The earliest systems, known as telestrators, allowed the pundit or "color commentator" to draw on top of a still image to illustrate some aspect of their analysis. More modern and advanced systems make use of image- or sensor-based systems to extract information about the scene and generate more detailed analysis or visually impressive graphics.

Examples of current state-of-the art graphical analysis systems which involve aspects of visual computing include tools for drawing offside lines and other contextual information into a soccer pitch for tracking the movement of balls in 3D. Other tools provide new views on the action, such as zooming into the image to provide a close-up or overlaying multiple races or heats (such as in downhill skiing) on top of one another so direct comparisons can be made between them. Virtual views of the action can be generated either to provide a new angle on the action that the broadcast camera didn't offer, or to provide a smooth way of transitioning between two different cameras without jumping from one to the other.

These graphics technologies have become so important and integral to their respective sports that some are now used as aids for the judges or referees. For example, the positional information generated for the America's Cup yacht race graphics is used by the event's umpires, and in soccer the Hawk-Eye system[1] is used by the English Premier League for deciding when a ball has crossed the goal line.

21.3 Challenges for Visual Computing in Sports Broadcasting

In some ways visual computing techniques are ideal for use in sports coverage. It is very hard to add extra physical equipment such as sensors into the sporting areas themselves. They might get in the way of the sport being played or affect the performances of the players and athletes. Instead, unobtrusive cameras can be placed around the area of action. Very often the cameras already in place for the television broadcast can be used.

However, there are many significant challenges to the use of visual computing in TV sports coverage that are not found in the laboratory. The environment is out of the control of the system designers; the rules and practices of the sports, such as the type and color of the clothes of the participants, and the size, shape, and fabric of the arenas are normally long established and not necessarily conducive to easy scene analysis. The event can take place over a large area and the locations of the action cannot always be easily predicted. The participants may wear baggy clothing or all wear the same clothing and will often occlude one another. Most sports are played outdoors without controlled lighting or weather conditions and as a result the appearance can change dramatically. Some sporting bodies may be prepared to make changes to their sport if it can be justified by improved television coverage. Perhaps most famously, the fluorescent yellow of tennis balls was chosen to ensure greater visibility on color television.

[1] http://www.hawkeyeinnovations.co.uk/

Sporting events can also be chaotic, with hundreds of people around performing their individual roles. The environment is dynamic, so while a piece of the arena might form a useful calibration point in the morning, it may no longer be there in the afternoon. Similarly, furniture or advertising hoardings may be moved around as the event progresses. Athletics stadiums, for example, will construct or dismantle equipment based on whatever event is happening at that time. This even includes seemingly fixed objects such as the distance markers on the javelin field which consist of pinned down white tape that will disappear when the discus starts from the other end of the field. Some events such as the long jump can even change location when the direction of wind changes.

Broadcasting has historically used interlaced cameras where a series of pairs of fields captured at different times make up each frame. One field contains all the odd lines in the image, the other contains all the even lines. This effectively doubles the temporal resolution of the video, but each field only has half the vertical resolution of the full frame. Combining the two fields into one frame can introduce artifacts because of the time difference between them. The use of interlace is declining but this is still how pictures are transmitted to the home, and if a video feed is made available from a TV camera or recording then it will generally be interlaced. As a result any analysis or graphics systems have to be able to cope with interlace fields rather than the progressive frames as might be expected in the lab.

The use of zoom lenses on broadcast cameras inevitably makes intrinsic calibration more difficult. Lens distortion, image center point, and chromatic aberration can all vary as a function of both zoom and focus settings, and in the absence of data from lens sensors it can be difficult to accurately determine their values. Coverage of sports events also involves a large range of zooms. It is not unusual for a camera covering athletics, for example, to move from a field of 3 or 4 degrees to 45 degrees in a single sequence.

Broadcast cameras have a control known as aperture correction or sometimes "detail," The aperture correction is used to "sharpen" an image and is one element that distinguishes the "TV-look" from "film-look." Effectively, the correction emphasizes high-frequency image components and is therefore a high-boost filter. A high level of aperture correction causes a significant color shift to pixels close to luminance edges. This can affect an area of about 2-3 pixels around contour edges and leads to incorrect segmentation results using color-based segmentation methods. The segmentation can be improved by compensating for the effects of the aperture correction [Grau and Easterbrook 08].

Camera calibration is difficult in uncontrolled scenes which typically have few features with accurately known 3D positions. Although lines on sports pitches are notionally in well-defined positions, pitches are rarely flat (being deliberately domed to improve drainage) and lines are often not painted in exactly the right positions. Techniques such as 3D laser scanning sometimes need to be used to build an accurate model of the environment which can then be used for camera calibration. When calibrating sparsely spaced cameras with a wide baseline, features in the scene can change their appearance significantly between camera viewpoints, making reliable feature matching between cameras difficult. Cameras can often get moved accidentally so calibration may need to be repeated.

Sports broadcasting has its own methods and workflows, into which any graphics system must try and fit. Live sport in particular requires real-time or very fast turnaround results, and the conventions of the coverage may limit the camera choices available. The system also has to be robust enough that the sports director can be confident that it won't break when potentially millions of people are watching.

Unless extra cameras can be installed at a venue, the event will only be available as a video feed from the main broadcast cameras. This means that the operators of any analysis or graphics system will have little or no control over the camera positions and movements. Communication to and from the camera will normally have to pass through the broadcaster's lines of communication with requests for camera movements being passed to the director and then to the camera operator. Camera operators will often cover more than just the game and may start recording the crowd or other interesting events that have no direct relevance to the sport.

A successful system should also offer good value-for-money to the broadcaster compared to other ways of enhancing sports coverage. Sports producers have many ways to improve their coverage and to add significant value. For example, they can add extra cameras around the ground—including special cameras such as infrared heat cameras for cricket analysis, flying cameras suspended from wires attached to the stands, or some of the cameras discussed in Section 21.7 and Chapter 3—microphones on the umpires to hear their decisions and orders or even, and technically rather less interesting, they can just hire a new pundit to provide the analysis.

21.4 Foreground Segmentation

For many visual computing systems it is important to get a high-quality segmentation of the participants from their environment.[2] For example graphics may want to be drawn on the pitch or arena such that they appear

[2]Segmentation is histrionically called keying in the broadcast industry and matting in the movie industry.

to be under the players. If the background is of a reasonably uniform color then some relatively simple image processing can be used. For example, with soccer the players are normally stood on green grass so a color-based segmentation algorithm (known as a chroma-keyer) can be employed such that graphics only appear in regions that are suitably grass-colored. The result is that nothing is drawn on the players themselves but may be drawn on either side of them, giving the impression the graphics are behind him or her.

The color of grass on pitches can vary significantly. This is due to uneven illumination and anisotropic effects in the grass caused by the process of lawn-mowing in alternating directions. After periods of bad weather or late in the season patches of mud can appear to further discolor the pitch. In some applications the fact that the chroma-keyer will not generate a key for discolored areas can actually be an advantage as this adds to the realism that the graphics are painted on the grass, as long as such areas are not large. However, these challenges for the chroma-keyer can give a segmentation that is too noisy to produce the high-quality key needed for applications that require scene reconstruction such as free viewpoint video. Difference keying can produce better results and is also able to segment pitch lines, logos, and other markings of the background. The background model or "background plate" can be created by either taking a picture of the scene without any foreground objects or, if this is not possible, the background plate can be generated by applying a temporal median filter over a sequence to remove moving foreground objects.

For some applications, such as those that draw graphics onto live video, the keying and graphics rendering processes most be performed in real-time.

21.5 Camera Calibration for Sports Events

Early telestrator systems allowed 2D annotations to be drawn on top of a video still. However, if the camera's position, pan, tilt, and zoom are known (together with off-line lens calibration to relate zoom ring position to focal length) then graphics can be overlaid on the image such that they appear in the same orientation as the scene. This is known as registration (Chapter 7). For example, off-side lines can be drawn on the pitch such that they appear to be painted on the soccer field. This camera calibration, combined with knowledge of the geometry of the scene allow measurements such as distance to be taken and displayed. Figures 21.1 and 21.2 illustrates this. American football down lines, for example, must be placed at specific distances along the pitch. Often, some assumptions will be made about the scene to help translate 2D knowledge into 3D. For example, when finding the position of a soccer player it is assumed that he is stood on the

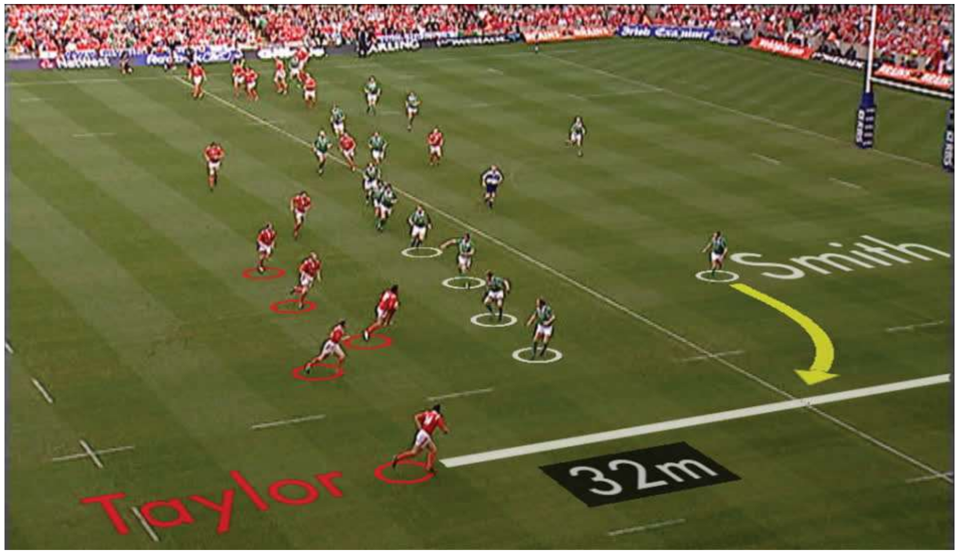


Figure 21.1: Player positions and movements along with distance measurements are added to the scene as if drawn on the pitch.

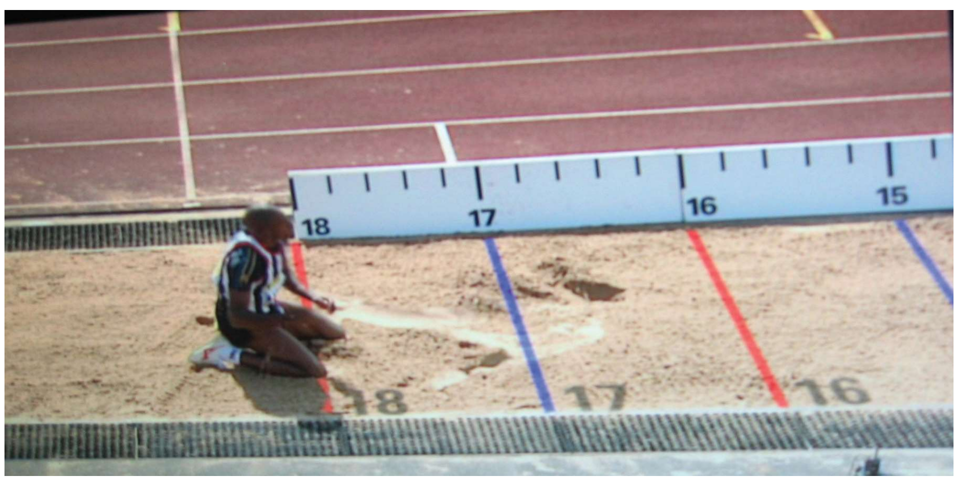


Figure 21.2: Graphics added to the triple jump to aid viewers in judging the distance of a jump. The graphics are drawn on screen as the camera follows the athlete into the pit.

ground so his height off the ground plane will be zero. As a result the other two dimensions can be calculated using the player's 2D position in the image. Multiple calibrated cameras allow for more sophisticated 3D effects, discussed in Section 21.6.

21.5.1 Image-Based Calibration and Moving Cameras

Calibration is most easily obtained from a static camera of known position and orientation. However, it is often desirable to be able to move the camera. Indeed if a broadcast camera is being used it will be moving. As the camera moves, many modern graphics systems such as Piero[3] retain the position of the drawn graphics relative to the real world. This has the effect of making them look even more like they are part of the scene. This is sometimes referred to as tied-to-pitch graphics. Such systems can be applied to many sports, not just those with soccer style pitches, including swimming, running events, and long jump, adding lines to indicate world records or qualifying distances. They can also be used to add logos or sponsorship to the scene either on the pitch or court or to the side as virtual billboards.

To produce this effect the movement of the camera must be tracked so that the calibration information about camera pose can be updated. This was originally achieved using a mechanical sensor on the camera and mount which measured the changing pan, tilt, and zoom. However, there are now several systems, including Piero, that are entirely image-based. This has the benefit that the same system can be used on multiple camera feeds without the need for expensive equipment on each camera or the need to be located near the camera itself. Indeed, such image-based systems can work just as well on recorded material without the need to work live.

21.5.2 Line-Based Tracking

One method of tracking camera pose is to use the lines on the pitch. Sports such as rugby, soccer, and tennis have regular pitch markings with known (or easily measurable) geometry which can be used to calculate the camera's position and pan, tilt, and zoom. The lines on the pitch are in known positions in the real world. If corresponding lines are visible in the camera view the correspondence can be used to compute camera pose. Typically, calibration charts are used (Section 1.6). In the case of sports analysis, however, the use of lines is more practical. With sports coverage it is generally not practical to get access to the camera or to be able to place objects into its view. Instead it is more reliable to use what is known to be visible to the camera in the scene. This has the added benefit that calibration can be performed after the fact from recordings, even if there had been no intention to use the video at the time of filming. A minimum of four lines (which cannot all be parallel) must be visible to solve for the unknowns and fully compute the camera position and pose, fewer are needed if the pose is known. One implementation of this approach is described in [Thomas 07] and forms the basis of the pitch line tracking features within the Piero and

[3]http://www.redbeemedia.com/piero/piero

tOG-Sports[4] graphics systems. The tracking process is split into stages. First, the position of the camera is estimated, then the tracker is started with a view of the pitch from which it must calculate an initial pose. From then on the tracker runs from frame to frame at full video rate.

Most cameras covering events such as soccer generally remain in fixed positions during a match. Indeed, the positions can remain almost unchanged between different matches at the same ground, as the camera mounting points are often rigidly fixed to the stadium structure. It therefore makes sense to use this prior knowledge to compute an accurate camera position which is then used as a constraint during the subsequent tracking process.

Estimating the position of a camera from a set of features in a single image can be a poorly-constrained problem, particularly if focal length also needs to be estimated. Changes to the focal length have a very similar effect on the image to moving the camera along the direction of view. To improve accuracy, multiple images can be used to solve for a common camera position value. The pose for all images is computed simultaneously, and the position is constrained to a common value for all the images. The images should cover a wide range of different pan angles (e.g., in soccer it should cover both goal areas) so that each image's line of position/focal length uncertainty lies in a different direction. In this way the differing views provide complementary information about the pose and solving for all of them together allows for a significant reduction in the ambiguity.

This process can be repeated using images from all the cameras feeds onto which the production team want to add virtual graphics. For soccer, this is likely to include the camera on the center line, the cameras in line with the two 18 yard lines, and possibly the cameras behind the goals. The computed camera positions are then stored for future use.

The frame-to-frame tracking process uses the pose estimate from the previous image and searches a window of the image centered on each predicted line position for points likely to correspond to pitch lines. A straight line is fitted through each set of points, and an iterative minimization process is used to find the values of pan, tilt, and focal length that minimize the distance in the image between the ends of each observed line and the corresponding line in the model when projected into the camera image. The observed and projected lines are highlighted in Figure 21.3. In such a way the lines can be tracked from frame to frame in an efficient and robust manner.

The operators of the graphics software need to be able to start working on generating their analysis clips very quickly after a suitable incident has occurred. For example, if a goal is scored just before half time in a

[4]http://rtsw.co.uk/products/tog-sports/tog-sports-lite/

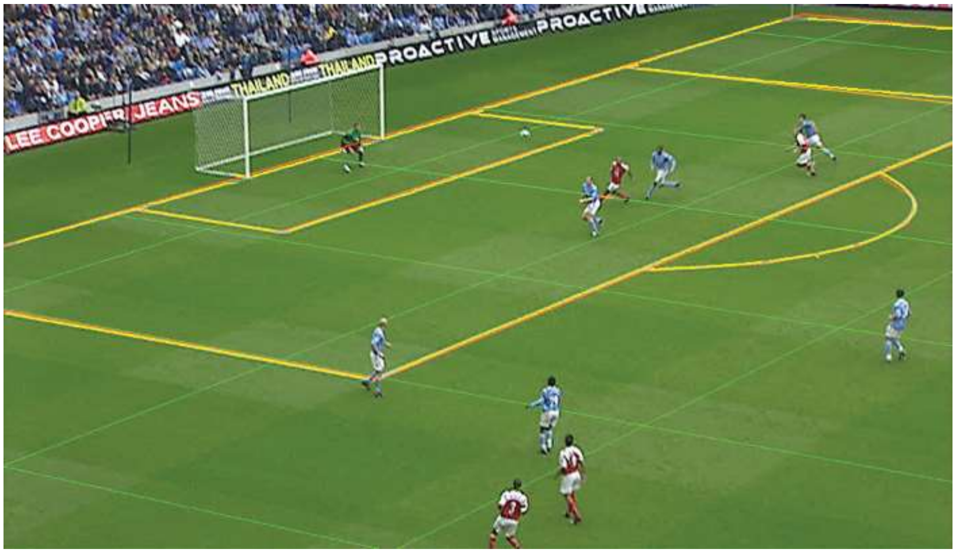


Figure 21.3: The pitch model is projected onto the scene highlighting the lines being tracked. The red lines represent the projected model while the yellow lines are the lines found in the image.

soccer match then a clip to aid analysis and discussion of the goal might be required within around 5 minutes. As such they cannot spend time calibrating the camera with image correspondences every time they want to make a clip. Similarly, if the system is tracking on live footage it needs to get the correct pose very quickly once it starts to receive the video—calibration would again take too long. The system must be able to initialize to a suitable camera pose as quickly as possible when presented with a new scene. Various methods can be used, including classifier based systems that recognize the scene [Dawes et al. 09] or by matching features using the descriptors described in Section 7.2. A very reliable method is to use the camera position saved earlier in the process and then perform an exhaustive search of all possible pan, tilt, and zoom values to find the one that best matches the lines visible in the scene. This, however, would be very slow so the process is sped up by performing the search in Hough space after first generating a Hough image based on lines in the video image [Thomas 07].

21.5.3 Feature-Based Tracking

For sports such as athletics, the camera image will generally show limited numbers of well-defined lines, and those that are visible may be insufficient to allow the camera pose to be computed. For example, lines on a running track are generally all parallel and thus give no indication of the current distance along the track, making pose computation impossible from the lines alone. For events such as long jump, the main area of interest (the sand pit) has no lines in it at all, (see Figure 21.2). Thus for these kinds of events to accurately estimate the camera pose for the insertion of virtual graphics an alternative approach is needed.

One approach is to view the problem as a specific example of SLAM (simultaneous location and mapping [Smith and Cheeseman 86]), in which the pose of the camera and the 3D location of tracked image features are estimated as the camera moves. The system is given an initial pose for the camera and then uses image features to track the pose from frame-to-frame (Chapter 7). There are some special considerations to take into account in the context of sports coverage. Specifically, significant changes in camera focal length can occur. But at the same time, camera position is constrained since it is generally mounted on a fixed point. This is in contrast to most implementations of SLAM which assume a fixed focal length camera but allow full camera movement. A significant degree of motion blur can occur as motion speeds of 40-50 pixels per frame are not uncommon when covering sports events with tightly-zoomed-in cameras. The approach described in [Dawes et al. 09] is designed to meet these requirements. It uses a combination of fixed reference features to prevent long-term drift (whose image texture is always that taken from the first frame in which the feature was seen), and temporary features to allow non-static scene elements (such as faces in the crowd) to play a useful part in the tracking process. The image features are assigned an arbitrary depth, as their depth cannot be determined from a fixed viewpoint. Although it would be possible to treat the whole image as a single texture and determine a homography that maps it to a stored panoramic reference image (since the fixed position of the camera makes the 3D nature of the scene irrelevant), the presence of a large number of moving features (the athletes) makes it advantageous to consider the image as a set of separate features so that outlying features caused by athlete movement can be discarded using a technique such as RANSAC (Section 7.2).

21.6 3D Analysis

As described, the combination of foreground object segmentation and camera calibration can allow the approximate 3D positions of objects to be inferred. This simple approach can also be used to create a crude 3D model of the scene. To place the segmented players into a 3D model of a stadium they can be rendered on flat planes, sometimes know as billboards (Section 17.3), at the estimated locations (Figure 21.4). This allows the generation of virtual views of the game from locations other than those at which real

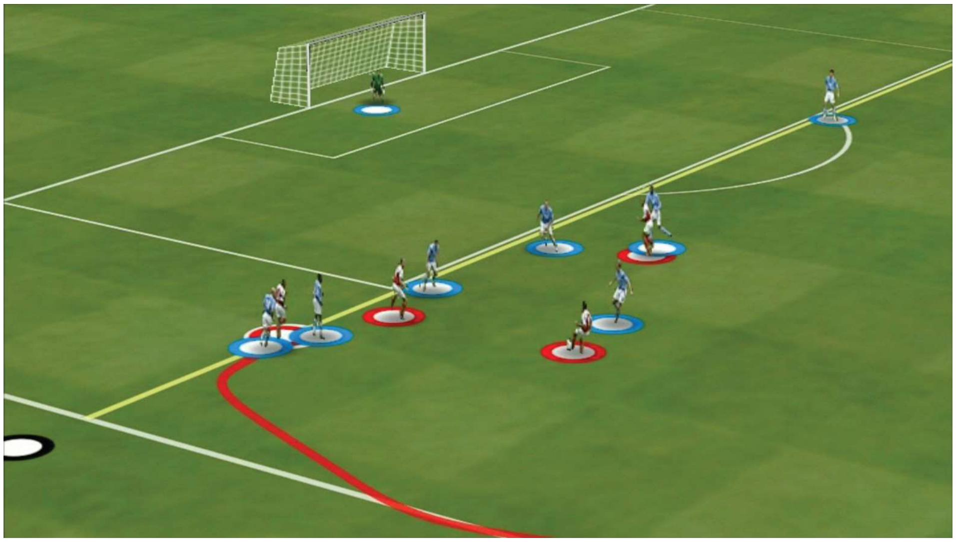


Figure 21.4: Player billboards are placed into a 3D model of the stadium.

cameras are placed. If this process has been conducted on two different cameras, a sequence can be rendered where the virtual camera moves from one real camera to the next. The two sets of flat player textures are blended into one another as the camera transitions. Systems such as Viz Libero[5] make use of this effect.

The simple player-modeling approach works well in many situations, but the use of a single camera for creating the models restricts the range of virtual camera movement. The planar nature of the players becomes apparent when the viewing direction changes by more than about 15 degrees from that of the original camera. One solution is to use pre-generated 3D player models, manually selected and positioned to match the view from the camera, as is the case with the Piero "3D Players" feature.[6] It can take a skilled operator several minutes to model such a scene, which is acceptable for a post-match analysis program, but too slow for use in an instant replay. The player models also lack realism. Some tweaking can be made to the flat textures to give them 3D character, such as altering the orientation of some part of the texture, particularly the limbs. However, this has only a limited effect. Very often the transition between two camera views, as described above, will be performed quickly to ensure that any oddities in the flat player models aren't too apparent [Kilner et al. 09] (Chapter 11).

[5]http://www.vizrt.com/products/viz_libero/

[6]http://www.redbeemedia.com/piero/solutions/3d-players

A multi-camera modeling approach provides an alternative. As described in Chapters 2, 12, and 17, multiple cameras arranged around the scene can be used to generate a 3D model. This process can be conducted over a video sequence to add a temporal element to the scene representation. The camera can then be entirely virtual and view the model of the scene from any point. This free viewpoint video enables sport presenters to explore interesting incidents by moving to new viewpoints, like a virtual flight down to pitch level or overhead [Hilton et al. 10, Goorts et al. 13]. This is a powerful tool to visualize spatial relationships between players and their tactics. It also allows for the transition between real cameras without the need to move to an entirely virtual world or try and work round the flat texture problems previously described. However, the few broadcast cameras used might see too little detail of the action or from angles too different to provide the basis for good 3D shape reconstruction. Extra specialist-cameras may be needed in the stadium which adds to the cost and logistical difficulty. Even with multiple cameras, the size of a typical soccer field and the distance of the stadium from the pitch mean that the image quality is not ideal and any small calibration inaccuracies can have a large effect. This can lead to low resolution models.

The regular problem of occlusion becomes more acute while covering team sports where players can easily block a camera's view of other players. The players can also contort themselves into shapes and interact with one another in ways that makes their 3D reconstruction particularly difficult. The crowd watching the games from the stands can also make segmentation particularly difficult from certain views. As a dynamic multi-colored texture they can sometimes prove indistinguishable from the players and ball.

The freeD system, developed by Replay Technologies,[7] avoids some of these issues by concentrating on small areas of known action, such as a piece of gymnastic equipment, the tee off in golf, or the plate in baseball. For recording a baseball sequence around 12 HD cameras were placed around the ground focused and framed on the batter.

Visual effects such as free viewpoint video face the problem of turnaround time. To generate a good quality free viewpoint video may currently take several minutes. However, live sports coverage would ideally like such effects to be available for immediate replay. Even in the case of highlights programming, when there is more time to create sequences, the time a sequence takes to render becomes an issue. Unless the effect is particularly spectacular and insightful, a TV production would rather generate several less complicated sequences than spend that time creating a single complicated one. As processing power improves and better cameras come

[7] http://replay-technologies.com/

onto the market, such effects can be generated far more quickly. The freeD system claims to be able to render out sequences in around 30 seconds.

Free viewpoint video also opens up new possibilities for how viewers would interact with sports coverage. An event could be delivered to the viewer over the web as a free viewpoint video allowing them to view the action from whatever viewpoint they choose [Budd et al. 12].

Hawk-Eye was one of the first commercially available visual computing systems used in sports coverage. Developed for tracking balls in 3D using a multi-camera system, it was first commercially applied in the tracking of cricket balls to aid the analysis featured in TV coverage of international cricket. Since then it has been applied to tennis and, more recently, soccer, among others. The system is deployed with a number of dedicated high-frame-rate, synchronized cameras arranged around and framed on the area of interest, such as the tennis court, goal mouth, cricket wicket, etc. There is normally space around the court or pitch (often on the roofs of the stands) to install the cameras. As static cameras they are easier to calibrate, and because they are not used to generate broadcast footage short shutter times and higher frame rates can be used to create the optimal pictures for processing. The problems of interlaced video can be entirely avoided by using progressive scan. There is a lot of prior knowledge about each sport that the system can use, including the size and appearance of the ball, its motion (once it is hit, its motion can be predicted using the laws of physics), and the area over which it needs to be tracked. The system first identifies possible balls in each camera image by looking for elliptical regions in the expected size range. Candidates for balls are linked with tracks across multiple frames, and plausible tracks are then matched between multiple cameras to generate a trajectory in 3D. This records the path and position of the ball during play and can even be used to predict the likely movement of the ball had it not struck another object or has spin applied. Other multi-camera analysis systems for balls in sporting events include QuesTec's Umpire Information System[8] for baseball and the similar Zone Evaluation System.

21.7 Virtual Broadcast Cameras and Second Screen

Panoramic, omni-directional, and free viewpoint video promise to change the way events are being captured for broadcasting. The common work flow today is to have one camera operator per broadcast camera. The main duty of this operator is to make sure that an event is captured with appropriate framing from the vantage point of the camera position. In high-end live

[8] http://www.questec.com/q2001/prod_uis.htm

productions, the exposure and color settings of the cameras are controlled by shading operators in the outside broadcast van, so a camera operator really can solely focus on framing. This takes particular skills of attention and context awareness in order to predict to a certain extent what is about to happen and how to perform the actual framing.

Different cameras provide raw footage from different vantage points. In soccer, for instance, the camera positions are standardized by FIFA, and assigned numbers so that everyone can refer to a camera position in a non-ambiguous way. But also different types of events are usually captured from established sets of camera positions. Major rock concerts, for instance, are captured by a camera on either side of the stage, a camera middle front stage (shooting in "frogs" perspective), a camera (often on a crane) from within the audience, and sometimes more cameras from back-stage corners.

The raw footage from each of these camera vantage points is presented to a director on an array of video monitors in the outside broadcast van. The director makes live decisions on what stream on the mosaic is dispatched for live broadcasting.

In offline broadcasting, one or more camera operators autonomously choose vantage points, frame an event and control exposure themselves, with the same goal of producing the best possible source footage that can be combined later into the final program. The process of putting together a program after filming has finished is known as post-production. During the editing stage of post-production, footage may be cropped slightly in order to correct framing.

This current work flow is dictated by the fact that camera pixels are expensive, and used to be very expensive. However, pixels are getting cheaper nowadays. As a consequence, it becomes feasible to capture events with a high-resolution camera, at a wide shooting angle, so that framing can be performed by an operator in an outside broadcast van, or in post-production for off-line productions. During the 2014 Super Bowl, for instance, 4K broadcast cameras were used in this way to generate full HD footage. The advent of 8K cameras promises a fourfold zooming capability without loss of quality, offering considerable freedom in reframing events. This comes at its highest benefit in the context of replays, in order to "tell the story" of what has happened at an event, shortly after a particular incident. Post-factum framing no longer requires the prediction skills of a good camera operator today.

Such "panoramic" virtual camera technology is already feasible today at a commercial level by stitching together images of closely spaced, full-HD, cameras. In the context of pitch sports, the stitching can be performed to sub-pixel accuracy today with relatively simple algorithms since the scene being captured is at a large distance compared to the inter-camera spacing

Figure 21.5: 16-times full-HD panoramic camera system for sports broadcasting (left), camera installed in American football stadium (middle), operator reframing footage while programming an instant replay using a joystick controller (right).

Figure 21.6: Omni-directional video capture for broadcast TV (left), director controlling framing and lens effects using a joystick controller in a montage booth (middle), resulting video clip frame with stereographic "virtual lens" (right).

(tens of meters versus centimeters), and depth variations are relatively small (Chapter 3).

Another application scenario of omni-directional cameras is to facilitate video capture on location. A compact omni-directional camera is installed on site and captures whatever happened around the location where it was set up. Actual framing of events takes place in post-production, in a montage or edit booth, at the broadcaster's facility. Creative decisions concerning the field of view or lens effects in order to frame an event are no longer taken in a predictive manner by a camera operator on location, but based on posterior knowledge of what has happened. This work flow may eventually reduce costs, while allowing additional creative freedom. The work flow was tested, among others, in the context of a popular prime time program in "1000 zonnen" by Belgian national broadcaster VRT, with two 1- to 4-minute items produced this way weekly over a period of three months.

In addition to allowing cost reduction and additional creative freedom, panoramic and omni-directional video capture techniques can also be used to provide content for second screen applications that complement a traditional broadcast watched on the main television. A second screen is a viewer's personal device such as a mobile phone or tablet which they may use while watching a program on the shared primary television. A second screen spectator can follow the same event as shown on the primary TV screen but the way the event is shown could be fundamentally different. On the primary screen, a director's cut may be shown, while on the secondary screen, a spectator can be allowed the freedom to "look around" at the event him/herself. Experience from actually bringing this into practice in the context of a popular Belgian talk show, and for sports coverage, showed that the primary screen is by far the best way to show details (close-ups) and steer the spectators' attention (as it always and exclusively is on primary screen), while the surround video content on the second screen helps to convey the ambiance of an event location better than by any other means. In the context of talk shows or political discussions, usually just the person speaking is shown on TV, while surround video on a second screen allows viewers to observe the body language of the others, which may be very informative on occasions.

21.8 Summary

Visual computing systems face various challenges if used in live television broadcasting and in sports coverage. Some systems are already in use, and the approaches they take to tackle the challenges are encouraging. In the next few years, as technology improves there is the potential for visual computing to play an increasingly important role in TV broadcasting. Free viewpoint video systems and wide angle cameras may change how television is produced and even create new ways for the viewer to consume and interact with TV productions.

22
Web-Based Delivery of 3D Mesh Data

Max Limper, Johannes Behr, and Dieter W. Fellner

22.1 Introduction

Visual computing applications are often dealing with 3D mesh data, in order to represent objects that have been captured within the real world. Such data can stem from various 3D scanning methods, which usually output point clouds and then apply some mesh generation method, for example by using the marching cubes algorithm, Delaunay meshing, or the *3D Snakes* method (Chapters 9 and 10). Depending on the size of the data, as well as on the target application, the meshed result might then be processed further, for example through simplification or remeshing methods, before a final, meshed 3D model is ready for presentation.

A trend that can be observed over the last few years is that there is, on the one hand, a growing number of devices with powerful graphics hardware, and on the other hand almost no device any more that is not potentially connected to the Internet, or to a local intranet. As a natural consequence, 3D Web technology serves more and more as a powerful integration platform for visual computing applications: Suddenly, with low-level graphics APIs like WebGL, it has become possible to deliver 3D mesh data to a wide variety of devices, ranging from desktop PCs to smart phones, in a consistent, platform-independent way. Figure 22.1 shows a conceptual overview of this situation.

22.2 3D Meshes vs. Image-Based Representations

There are several possible ways for visually presenting captured real-world scenes or objects to a user. In general, when considering the data transmission format, a distinction can be made according to two basic categories. On the one hand, 3D data may be stored for presentation as polygonal

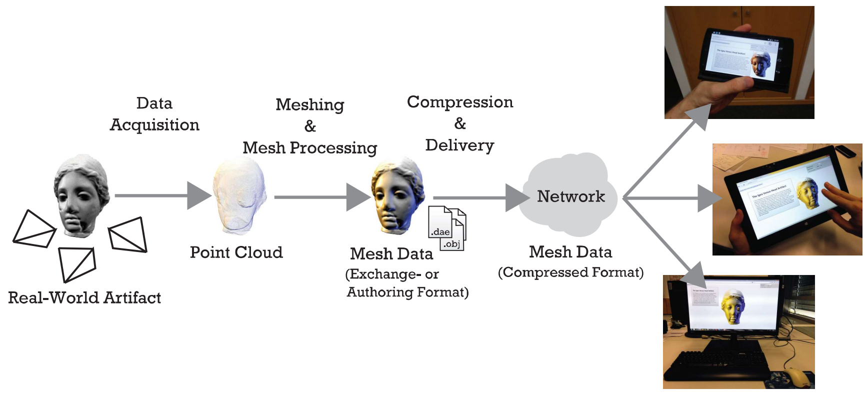


Figure 22.1: Pipeline for a typical visual computing application from the field of cultural heritage. Data is acquired through a 3D scanning method and then meshed. The meshed data may be further processed by simplification algorithms, for instance. It is then compressed for delivery over a network, transmitted, and finally presented on a client device.

mesh. In the most simple case, this mesh is simply a list of triangles, which represent the surface of an object. Appearance details (such as surface normals, displacement or colors) are usually represented at the vertex level, or in the form of texture images. The client application then generates images on-the-fly, using common 3D rendering techniques, as they are also used in computer games, for example. Within this section, such representations are referred to as *mesh-based representations*.

On the other hand, captured data may be directly represented via images. Instead of generating perspective views, along with matching lighting and appearance information, on the client side, all the client has to do is to display the correct image, or to interpolate between a set of given images. Methods that use such representations are usually referred to as image-based rendering (IBR) or video-based rendering (VBR). It is worth noting that there are also other ways of storing captured 3D data for presentation, such as point clouds or depth maps—an overview can be found in Chapter 17. In Figure 17.1, mesh-based 3D representations can be found at the very right edge. Both of the mentioned basic approaches for representing captured data for remote presentation, mesh-based representations as well as IBR/VBR solutions, have specific advantages and disadvantages.

Advantages of IBR and VBR methods. Rendering mesh-based representations in an efficient way requires full access to a dedicated 3D graphics API and hardware. In contrast, VBR or IBR solutions potentially have lower device requirements, and therefore might be a better fit for a maximum portability. If critical data (for instance, protected CAD product data) should be displayed over an unsecure network, security restrictions may also prohibit a delivery of mesh-based representations. With IBR or VBR solutions, deriving such critical data from the displayed images is a much harder (and error-prone) task, therefore these methods might be preferred in such cases. Another great advantage of VBR and IBR solutions is that they are basically independent from the complexity of the model and its appearance properties. Extremely high polygon counts, or the request for a very detailed view of complex material or illumination properties, might therefore also prohibit the use of mesh data, and lead to IBR or VBR solutions instead.

Advantages of mesh-based representations. Depending on the application, the user might not be satisfied with viewing a non-interactive scene, as is usually the case with IBR and VBR solutions. If the user wants, for example, to be able to change object materials, or to move objects or light sources around, IBR or VBR solutions cannot provide this degree of interactivity (Figure 22.4). In such cases, the necessary visual information has to be dynamically generated on the client side, using common 3D graphics techniques and mesh-based representations. If network bandwidth is a critical factor, using mesh-based 3D representations is also the method of choice, especially compared to approaches like light fields, which require the storage of huge amounts of data. As already noted in Chapter 17, this is due to the fact that an optimized, meshed 3D model of a scene, along with a texture atlas, makes very compact scene representations possible. One possible solution to achieve maximum quality along with a maximum degree of interactivity, on almost any client device, is the use of a dedicated server for remote rendering. However, the big disadvantage of this image- or video-based approach is its bad scalability: As soon as multiple clients connect, a dedicated rendering server or process must be maintained for each of them, which is not possible for many kinds of public, large-scale Web applications (such as online shops or exhibitions). Another problem in this context is the need for a connection with minimum latency, in order to be able to provide fluent user interaction. Therefore, mesh-based representations are usually preferred for interactive small and medium-size scenes, which should be accessable for a large number of clients.

Considering the specific advantages and disadvantages of IBR/IBR solutions and mesh-based representations, it strongly depends on the context of the application which representation is best-suited. Using service-oriented architectures (SOA) and RESTful APIs allows to deliver 3D assets along with context-specific application templates, as it is done by Instant3DHub [Jung et al. 12] or XML3DRepo [Doboš et al. 13], for example.

Client devices can then, for instance, automatically receive a specific representation, matching their CPU or GPU capabilities, the available bandwidth, and security-related contraints. This way, the decision between using IBR / VBR methods or using mesh-based representations can even be dynamically performed per client. A server might, for example, decide to share only images instead of real 3D information across unsecure networks, in order to prevent theft of proprietary 3D construction data. Providing an image stream for an interactive 3D experience, on the other hand, is only possible with server-side rendering, which requires a powerful server architecture to scale well, even for a small number of clients. Therefore, the approach of statically providing compact, mesh-based representations is usually the preferred way to deliver 3D assets on the Web, and already applied for a wide variety of use cases.

22.3 Application Scenarios

The field of application for high-performance 3D Web technology is growing every day, including many interesting scenarios from the field of real-world visual computing. One obvious use is the online presentation of scanned products for marketing purposes.

Another use case, which might be not that obvious, is shown in Figure 22.2(right). Here, research data from the CAESAR anthrophometric database has been made accessible as part of an interactive 3D Web application. Such fast, portable, and convenient access to the 3D mesh data allows quick looks on specific poses, without needing to explicitly download the corresponding files and then open them in a specialized viewing application.

One increasingly important class of real-world visual computing applications, using 3D Web technology, can be found in the domain of cultural heritage. Artifacts from this domain are mainly preserved in large archives, mostly belonging to museums all over the world. Usually, such archives contain too many artifacts to fit into a single exhibition, and they are often managed in traditional, paper-based archive infrastructures. The current shift toward digitized archive data also includes the storage and distribution of 3D scans for each artifact (Figure 22.1). This enables people around the world to remotely inspect collections in distant archives via the internet—in an ideal case, such inspection scenarios can be performed using regular Web browsers. An example from a recently initiated Web portal [Martínez et al. 14], using the open-source X3DOM framework for visualization, is depicted in Figure 22.2(left).

As a typical real-world visual computing application from this domain, Schwartz et al. have presented both, a capturing setup as well as a WebGL-

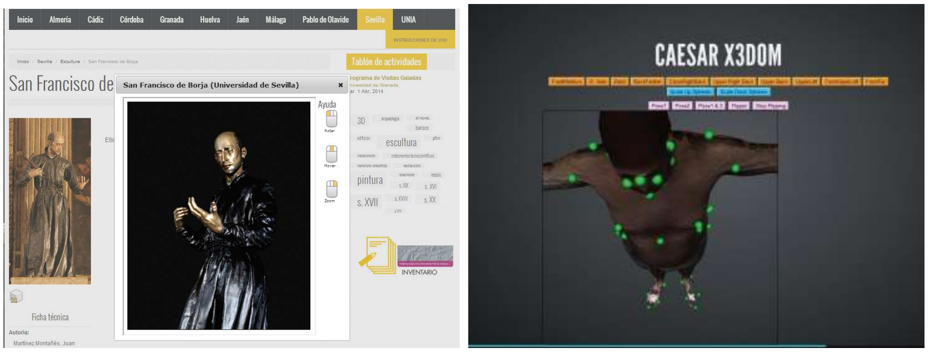


Figure 22.2: 3D Web technology, used for real-world visual computing applications. Left: A scanned cultural heritage object, which is part of an online exhibition. Right: Online inspection of the CAESAR antrophometric database.

based streaming framework, for acquisition and presentation of cultural heritage artifacts. Figure 22.3 gives an impression of their DOME II capturing setup, as well as of the 3D Web application that is used to present the results [Schwartz et al. 13, Schwartz et al. 11a]. Captured reflectance information is represented via bidirectional texture functions (BTF), which are approximated by texture images and progressively transmitted for presentation (Section 22.4). The user is able to interactively modify the lighting conditions inside the viewing application, which allows for a detailed inspection of not only geometry, but also material properties.

In a similar fashion, but using a less sophisticated material representation, the *Radiance Scaling* method enables the user to inspect surface details on an arbitrary 3D mesh [Vergne et al. 10] (Figure 22.4). During

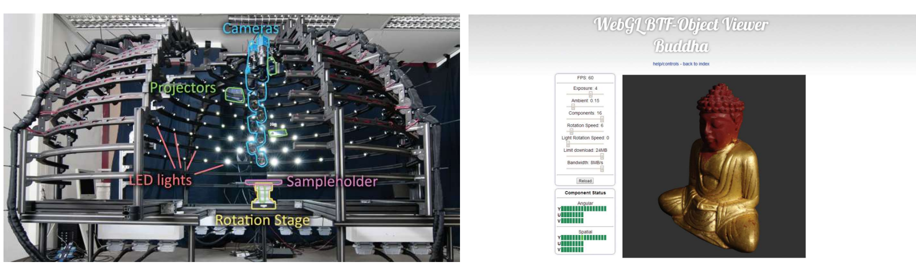


Figure 22.3: First and last stages of a pipeline for online presentation of cultural heritage, as presented by Schwartz et al.: DOME II setup for data acquisition (left), and Web-based BTF streaming and rendering framework (right).

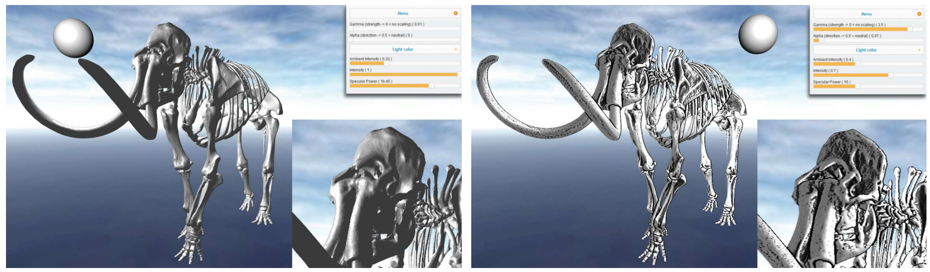


Figure 22.4: WebGL-based viewer, showing a wooly mammoth model that has been digitized and kindly provided by the Smithsonian Institution (see http://3d.si.edu/). The sphere acts as a handle for manipulating the position of the light source. Lighting parameters can be interactively modified to inspect different parts of the model, or to reveal surface details.

the shading stage of the rendering process, surface concavities and convexities are enhanced, depending on some parameters that can be interactively modified by the user. This technique only requires surface normals, and can therefore be used with any capturing setup, even if it only acquires geometry data.

22.4 Compression and Transmission

As soon as the acquired geometry has been reconstructed in the form of a polygonal mesh, additional mesh simplification or resampling steps might take place (Section 10.8). At this stage, mesh data is represented in the form of a topological and geometrical description, which can, for instance, be simply a text-based description in the form of a *Wavefront OBJ* (.obj) or *COLLADA* (.dae) file (Figure 22.1). This representation of the 3D mesh data (potentially enriched by texture data in the form of images) is then transferred to the Web-based delivery pipeline (Figure 22.1). To achieve a convenient user experience, it becomes necessary to apply compression and progressive transmission methods when delivering the data over a network. The conversion from the input format (e.g., from an .obj or .wrl file) to a compressed transmission format is either performed once (for static data) or as soon as needed (for dynamically changing data). The latter approach is, for example, taken by the transcoder component of the Instant3DHub architecture, as proposed by Jung et al. [Jung et al. 12]: As soon as a client requests a page that shows a specific mesh, the server decides whether the current compressed version of the requested mesh inside the server cache is up-to-date. If this is not the case, the outdated version inside the cache is directly replaced by compressing the input file again. Within this step, additional adaption to other parameters like network bandwidth, security constraints, or client device can also take place. A server could, for example, perform additional simplification steps to store representations at different level-of-detail (LOD) inside its cache. Depending on the client device that posts a request, the server might then pick a matching LOD for delivery.

Compressing Mesh Data

A crucial aspect of mobile and Web-based real-world visual computing applications dealing with 3D mesh data is efficient data compression and transmission. In contrast to common desktop 3D applications, the typical user scenario is radically different: users will browse to a specific URL where a 3D showcase should be available, or click a thumbnail in an online catalogue. As soon as the 3D Web application starts to load, users will typically expect *instant* results, as they are used to from other Web pages. This does not necessarily mean that everything should be directly loaded, but that a progressive refinement of the whole page must instantaneously take place.

Regarding compression methods, this means that a very important trade-off has to be made: On one hand, a good compression method could be necessary to deliver the mesh data in acceptable time over the network. On the other hand, a fast decompression is also crucial, since a complex decompression step could take more time than directly sending the uncompressed data would do [Hoppe 98, Pajarola and Rossignac 00, Limper et al. 13b]. This aspect becomes even more crucial for specific application setups, like mobile client devices (with less CPU or GPU power for decoding) in a fast company intranet (using a high bandwidth).

Therefore, only methods that allow a quick and straightforward decompression can be considered as candidate technology for Web-based and mobile 3D graphics applications. For an exhaustive overview over the wide field of research related to mesh compression in general, the survey of Peng et al. might serve as a good starting point [Peng et al. 05].

Lossy compression. The simplest method used to reduce the size of a mesh, and at the same time the most important one, is the use of quantized data. For the most popular quantization in Cartesian space, the idea is fairly simple: if the 3D mesh can be described with its vertex positions lying on a regular grid, without a visible loss of precision, then floating-point coordinates can be safely replaced by integer coordinates. The decoder will compute the original vertex positions from the normalized and quantized representation, using the original size of the mesh's bounding box that has been provided along with the data. Decoding can be performed in parallel on the GPU, inside a vertex shader. This approach has several advantages.

First, it leads to a significant reduction of memory that is needed to store the vertex buffers. Second, and even more important in a browser context, the time spent for CPU-based decoding steps becomes effectively zero, since clients simply need to pass the downloaded data chunks to the GPU, where the decoding is performed during rendering.

It strongly depends on the resolution of the model how many bits of precision are sufficient; typical values vary between 8 and 14 bits per vertex position component [Pajarola and Rossignac 00, Peng et al. 05, Lee et al. 10a].

A simple yet efficient approach to reduce the size of the compressed mesh is the subdivision into several sub-meshes. Such sub-meshes can be obtained from the original mesh with several partitioning methods, and the subdivision does not necessarily relate to manifold surface patches. A simple method is to use a kD-Tree to partition the mesh into regions that contain an approximately equal number of triangles. Another method for subdividing a mesh into multiple sub-meshes is to use a cache-optimization strategy to re-arrange the vertex data, and then cutting slices out of the resulting lists [Chun 12]. A third possibility is to use a hierarchical face clustering approach, where a cost function can be used to produce compact bounding boxes [Lee et al. 10a].

The use of several locally bounded, compact sub-meshes enables the encoder to achieve a very compact representation, since a lower quantization precision can be used to encode the vertex data within each sub-mesh, without having to sacrifice the overall quality of the full mesh after compression (the overhead of additional bounding box data is usually of neglectable size). However, subdividing a surface potentially introduces another problem, which occurs in the form of visible seams between the sub-meshes. To solve this problem, quantization is performed with respect to a common bounding box size, which is, for each spatial dimension, equal to the largest bounding box available among all sub-meshes [Sander and Mitchell 05, Lee et al. 10a].

Besides the vertex positions, other vertex attributes can also be stored in a quantized form for compression purposes. For RGBA colors, it is very common to use an 8 bit range per channel anyway. To store vertex normals in a quantized form, several approaches exist. Jung et al. use spherical coordinates (employing two angles on the sphere instead of 3D Cartesian coordinates) to store normal data for compact transmission, enabling decompression during rendering [Jung et al. 13]. Meyer et al. have provided a study on floating-point normal vectors and several quantized normal vector parameterizations, showing that an octahedron parameterization gives the best results, compared to 3D Cartesian or spherical parameterizations [Meyer et al. 10]. No matter which particular normal storage format is finally picked, it is always the main aim of an interactive 3D Web or mobile application to perform decompression on-the-fly during rendering (usually inside the vertex shader) in order to avoid slow CPU-based decoding.

Aiming at a browser-based inspection of human anatomy, Google conducted experiments that resulted in *WebGL-Loader*, a minimalistic library for compact 3D mesh transmission [Chun 12]. The most interesting aspect of this method is that the UTF-8 file format is a good alternative to binary formats, because it can be parsed very quickly within the browser, while also providing a simple variable-length encoding. The algorithm achieves a comparatively good compression and, combined with the native GZIP implementation of the browser, realizes fast decompression.

It is worth noting that limiting data precision to the minimum amount which is necessary does not only help to reduce transmission time. It has also been shown that it can improve geometric calculations and lighting computations, in terms of execution speed and energy consumption [Hao and Varshney 01, Pool et al. 11]. This aspect becomes especially important in the context of mobile graphics.

Loss-free compression. Besides lossy compression methods, there is a variety of loss-free compression algorithms that potentially fit well with a Web or mobile environment.

A very quick and efficient method is delta-coding and additional entropy-coding, like provided by GZIP compression [Chun 12]. Considering Web-based transmission, GZIP compression is available as a standard encoding method in HTTP, therefore applications can rely on fast, built-in browser implementations.

The open compressed triangle mesh (OpenCTM) format,[1] intended for fast exchange of compressed mesh data, uses LZMA compression, which is potentially superior to GZIP when applied to 3D binary mesh data. However, since there is no Web browser that supports LZMA natively, Web applications have to provide their own LZMA decompression implementation. This in turn might lead to the overall loading time (consisting of download time plus decode time) becoming much longer than for methods that rely on less sophisticated compression algorithms, but provide a faster decoding [Limper et al. 13b].

The Khronos group has recently considered several compression methods as candidate technology for the GL transmission format *glTF*. Among those, an important candidate technology to be included into the glTF specification is the TFAN approach, which has also been considered for ISO standardization within the MPEG-4 standard [Mamou et al. 09]. The idea is to provide good compression rates, support for all kinds of meshes (including non-manifold geometry), along with a low-complexity, fast

[1] http://openctm.sourceforge.net/

decoding method. The encoding algorithm first constructs a set of triangle fans from the input mesh data. The most frequently appearing triangle fan configurations are then used to efficiently encode the mesh connectivity, along with an arithmetic coder [Moffat et al. 95].

Progressive Mesh Data Compression

Since common Web applications demand an instantaneous user experience, the question arises how 3D mesh content can be progressively retrieved. In the ideal case, a compressed representation can be quickly delivered over the network, and at the same time be progressively decompressed without much overhead inside the client application.

A trivial yet efficient method to transmit a mesh in several stages is also provided with mesh partitioning approaches (Section 22.4). However, the term *progressive* is usually used for methods that progressively transmit the data of the full mesh. First approaches were made in this direction in the late 1990s, a long time before Web and mobile graphics APIs and the modern graphics pipeline were available. The pioneering work of Hoppe encodes mesh data in a compact and progressive structure, based on sequential edge collapse and vertex split operators [Hoppe 96]. Mesh data is initially provided in the form of a simplified, coarse mesh, which can be obtained using error-controlled edge collapse operations [Garland and Heckbert 97]. The original high-resolution mesh is then iteratively reconstructed by splitting vertices, which are read from the incoming stream of mesh data. In a later publication, Hoppe provided a short study on when additional compression of vertex data is worth the effort, depending on the decoding speed and bandwidth available [Hoppe 98]. In a similar manner, Pajarola and Rossignac stated that an efficient progressive mesh compression method has to balance between three contradictory constraints [Pajarola and Rossignac 00]. First, it needs to provide fine-grained progressive refinements. Second, high compression ratios should be achieved. Finally, the third important goal is fast decompression.

In the Web and mobile context, fast decompression has a very high priority. But even on desktop machines, sophisticated progressive mesh compression algorithms do not always pay off in terms of decode time compared to download time. Therefore, game developers have tried to port parts of Hoppe's method to the GPU [Svarovsky 99]. The focus of latter work is primarily shifted toward the optimization of rate-distortion performance [Peng et al. 05], which, however, mostly completely ignores the aspect of decode time. Lavoué et al. recently presented a modified progressive mesh compression algorithm which works relatively well on the Web [Lavoué et al. 13] (Figure 22.5). However, it is unable to deal with non-manifold geometry, and decoding steps still have to be performed on

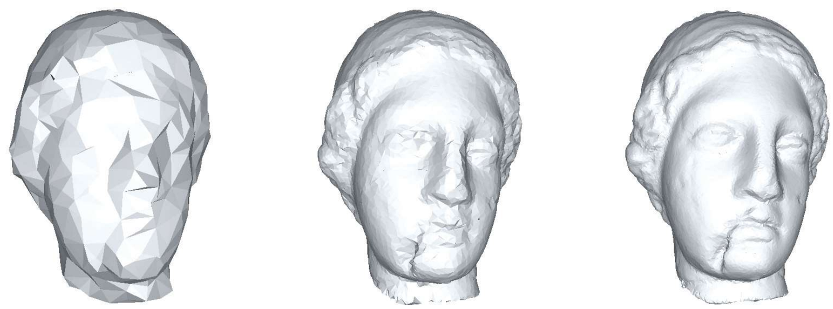


Figure 22.5: Progressive loading, using a progressive mesh. The method produces very good approximations during refinement, while maintainig a good compression rate. On the downside, it is unable to deal with non-manifold geometry, and thereby introduces additional pre-processing steps, and decode time might also become a critical factor.

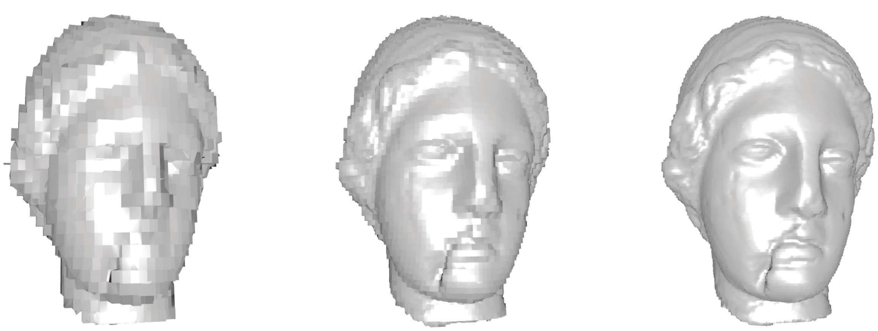


Figure 22.6: Progressive loading, using the POP buffer method. The idea is comparable to progressive PNG images, but transferred to 3D mesh data. It provides fast decoding without any CPU-based steps and is able to handle arbitrary triangle soups, but sacrifices compression performance and quality of the intermediate representations.

the CPU. An alternative, which has been specifically designed for the Web and mobile environment are POP Buffers resembling progressively loaded PNG images [Limper et al. 13a] (Figure 22.6). The idea of the algorithm is very straightforward: Since aggressive quantization of the geometry (for instance, with 5 or 6 bits) results in many triangles being collapsed to a line or even to a single point, it is possible to progressively sort out those triangles when using a coarser quantization resolution. The triangles can then be grouped according to the precision level where they first appear,

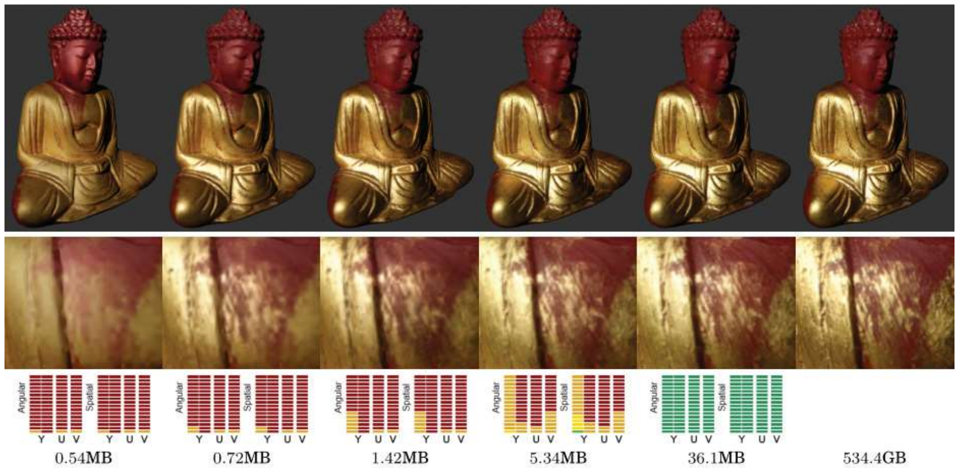


Figure 22.7: Web-based BTF streaming. The BTF data, which is used to shade the textured surface of a 3D mesh, is progressively transmitted as a compressed, decorrelated matrix factorization, with additional wavelet compression for the corresponding images.

which enables the creation of a nested, progressive structure. The method is able to handle arbitrary meshes, and it does not involve any CPU-based decoding steps. However, the quality of intermediate representations is not as good as for other progressive compression methods.

Schwartz et al. have presented a framework for progressively streaming bidirectional texture functions (BTFs), representing the appearance data of digitized artifacts [Schwartz et al. 11a]. The BTF data, which has been acquired from their DOME II capturing setup (Figure 22.3) is compressed using the decorrelated full matrix factorization (DFMF), as proposed by Müller [Müller 09]. Like other compression methods that are used in the Web and mobile context, this method provides a balanced trade-off between compression performance and decode time, and allows for real-time rendering on the client's GPU. The BTF is represented as a product of two functions, one depending on view- and light direction (angular function), and one depending on the spatial position of the sample (spatial function). By performing a singular value decomposition and streaming the resulting components of both functions, this information can be progressively transmitted to the client application, which produces high-quality results after just a few seconds (Figure 22.7). The corresponding images, which are used to store the component data, are compressed for faster transmission, using a wavelet compressor.

22.5 Summary

Remote presentation of real-world data, obtained through different scanning methods, has already become a common use case within many fields of application. Cultural heritage is only one example where the full potential of a visual computing pipeline has not been completely explored yet.

With the current technology trends, like cloud-based applications, the *Industry 4.0* initiative, and the Internet of things, there are also high chances that lightweight, mobile and Web-based 3D graphics applications will continue to gain further importance. Especially, the confluence of mobile vision and mobile graphics, performed on a single, Web-capable device—namely, the smart phone—bears high potential for many new, innovative, real-world visual computing applications. One important field of research in this context remains the search for a commonly accepted standard format for compression and progressive transmission of 3D mesh data.

23
Virtual Production

Volker Helzle, Oliver Grau, and Thomas Knop

23.1 Introduction

The concept of producing digital media content with software operated on affordable workstations has been established during the last two decades. It was mainly influenced by the gain of processing power and is applied in disciplines like visual effects (VFX), 3D character animation, visualization, simulation, or videos games, to name a few. The majority of work, however, is carried out in an offline process which separates the artist from other members of the creative team.

The concept of virtual production foresees to combine key aspects of media production in a real-time, or close to real-time, environment where creative decisions can be taken in direct consultation with other members of the team. Virtual production utilizes existing technology and concepts already established in television (chroma-keying, camera tracking, virtual studio), industrial visualization and design processes (virtual reality and augmented reality) or progress in capturing human motion, for example, in medical or sport applications.

The contemporary situation provides a unique opportunity to cause a major shift in how media is being produced based on the multitude of existing and upcoming hard- and software technologies. Those allow onset content interaction, visualization, and modification, with intuitive methods of controlling creative parameters. This can be achieved by combining modules like accelerated software algorithms for image synthesis, performance capture, sophisticated tracking of cameras and input devices, gesture recognition, virtual and augmented reality display devices, and capturing of additional information such as depth, geometry, or omni-directional video. In this chapter, different aspects of virtual production beginning with the already established virtual studio in television will be looked at. For instance, set extensions and filming on virtual locations offers a unique opportunity compared to the complexities involved in shooting at real locations. Finally, a case study production in which game engine technology and a motion capture system were used for pre-visualization and final render of an advertisement spot will be discussed.

23.2 Virtual Studios

Virtual studio systems were pioneered in broadcast in the 1990s. They were designed to be used in real-time in live broadcast productions. Early systems were using 2D digital video effects hardware systems (DVEs) combined with available studio chroma-keying equipment and mechanical camera sensors to capture pan, tilt, and zoom variations. With the advent of graphic hardware, like the Silicon Graphics Reality Engine™in the mid 1990s, it was possible to render 3D graphics in real-time at video frame rates. An early virtual studio system was developed in the EU-project Mona Lisa [Blonde et al. 96] and led to one of the first commercially available virtual studio system on the market.[1]

A specific problem of virtual studio systems is the tracking of camera movements. Mechanical sensors are only capable of capturing the movement of a TV camera on a pedestal, as pan and tilt[2] and in addition the zoom and focus settings of the camera lens. Early solutions to capture the 3D camera movement (6 DOF pose) were already based on visual tracking of special markers. Examples of this technology are the ORAD pattern-based system or the Vinten-Radamec Free-D system, developed by BBC [Thomas et al. 97]. Recent progress in computer vision in combination with inertial and gyro-sensors led to the development of real-time capable 6 DOF camera tracking without the requirement of special markers in the studio. Moreover, these systems work even outside a studio and are based on natural image features [Chandaria et al. 07] and are now available as commercial products, e.g., from Ncam[3] or SolidTrack.[4]

The take-up of virtual studio systems was initially slow. The hope that virtual techniques would save cost in set design was contrary to very expensive equipment and additional complexity in the preparation and production. With significantly reduced hardware cost virtual studio techniques are today used in many TV productions, in particular in news and factual shows.

Since the techniques developed for virtual studios are real-time capable by design, they were also used in early virtual production approaches in movies. In particular, virtual production systems were used in on-set pre-visualization in a number of movie productions involving real actors

[1] The ELSET™system developed by VAP.

[2] There is usually no roll on a studio camera pedestal.

[3] http://www.ncam-tech.com/

[4] http://www.solid-track.com/

interacting with otherwise invisible virtual actors by giving a real-time preview of the final composited scene [Rosenthal et al. 01]. The ORIGAMI project developed an actor-feedback system using view-dependent projection [Grau et al. 04]. Projector-based feedback was used in TV productions, like the BBC "BAMZOOKi" program for children. Many recent virtual production implementations make use of advances in display technology and are using bright projectors and displays directly in the studio to insert graphical or video content without the need for chroma-keying. This avoids problems of chroma-keying, like color back-spill on presenters and actors and also makes it easier for the presenters on set as they have visual feedback of the inserted content.

Further evolution of virtual studio techniques happened to some extent in techniques developed for use outside the studio space, in particular for sports visualization (Chapter 21).

23.3 Virtual Production for Cinema and TV

Virtual production employs a number of tools throughout the whole pipeline to make a movie or even smaller productions like commercials and drama series. These tools include software and hardware to create virtual assets[5] and/or capture reality with 3D scanning technology. Further, it includes software tools to manage the workflow and to visualize the virtual and virtualized components throughout the production process [Grau 05]. The workflow in movie and (high-end) TV production is divided into phases: The preparation and planning phase is based on the story board and script and implements or acquires all assets (virtual and real). In the on-set phase real assets and actors are captured (filmed) and in the post-production phase final quality renderings are produced and combined with the real camera footage in a process called compositing. The post-production phase is a very expensive step; in some cinema productions the combined cost for visual effects and post-production match the other production cost. It involves highly specialized and skilled jobs, like modeling, animation, virtual lighting, compositing, color grading, and linear editing of the final video/movie. Because of significantly lower budgets in TV, post-production is kept to a minimum and still the tendency in TV is to produce "in real-time."

It became more and more essential to the production team members to experience the sometimes fully or partially virtual scenes and assets. This has led to a drastic change in making moving pictures and resulted in

[5] In media production all items that are created in a creative process, including images, text, 3D content, and video clips are called assets.

higher efficiency, higher quality, and greater variety. It also challenges each team member to adapt to new technologies and methods such as set extensions, character animation, environmental controls, and pre-visualization [Nitsche 08].[6] Cinematography increasingly makes use of virtual data: Its elements and assets are created in the virtual realm all the way from the very first script until the very last frame on the screen.

Currently, there are barely any available systems to connect the different stages of a virtual production, and if so, they are proprietary and not accessible to other companies and parties. The vast variety in technical development (new cameras, new codecs, programming languages, interfaces, sensors, data transfers, etc.) renders the integration of different systems a very complicated task. The different approaches in producing moving pictures challenge at the same time vendors of technology to adapt their products for these needs, and demand from users new and more knowledge.

Using virtual production provides a direct and interactive visual feedback for everyone on the set, making decisions easier and more reliable. For instance, seeing the actual surrounding gives the director the chance to adjust camera framing and moves while shooting. Further, it gives actors visual feedback of virtual objects. Although actors are trained to perform without visual reference, it is hard to impossible to achieve a natural appearing interaction between real and virtual actors without on-set visualization, leading, e.g., to not matching eye-lines.

Besides these creative benefits, virtual production makes the film crew more independent of physical facts. For example, it is not necessary to wait until the right lighting situation outdoors arises, leaving the crew only a small time slot to shoot. The background as well as the fine-tuning of the lighting can be done later in a post-production step. This gives the film team much more time to retry a specific shot or to adjust it if needed. Furthermore, weather conditions or spatial limitations are almost no issue. It might be very costly to block a whole street or bring the whole team to the desired location in the Arctic. Using virtual production, the location in question can be filmed separately or created by artist on the PC, while the shooting of the scene takes place in the backyard or a film studio.

Apart from the direct interaction on set and the independence from time and space, virtual production can be used to decrease the time and cost that is needed for post-production. The actual amount of time that can be saved depends on the type and complexity of a specific production. It is possible to record the virtual background as well as the real foreground and compose both in real-time. This recorded material can be handed to the final cut, without being touched by any post-production. With TV and

[6] Also called previz.

movie productions getting visually more and more complex, this is rather an exception today.

Still, virtual production makes the post-production process easier in many ways. Recording a composed version of the shot beside the foreground gives the video editor, directors, producers, and 2D artists a solid and reliable impression of what the actual scene should look like. Using this reference can decrease the time that would be needed for an iterative process between director and artist until all scene elements are in their proper places. Misunderstandings are less likely to happen as the reference shot visually explains the scene.

Virtual production is a tool for the interaction between everyone involved in the production process. It is ideally designed for all different departments necessary to produce a film, TV show, commercial, live show, etc., and to connect those departments in an early stage and all through the whole process until the finished product.

Virtual production enables a very close relationship between clients and service providers, with the aim of ensuring that any questions and issues can be dealt with at a very early stage of production. The exchange of the (virtual) production data enables an efficient workflow between the departments. This can only be achieved with software and hardware systems that allow the exchange between the components involved in the production. Tasks can be split up after the idea or vision has been discussed and it is possible to imagine the result with the aid of pre-visualization.

Set Extensions and Virtual Backdrops in Commercials and Film

Filming at virtual locations can dramatically reduce costs compared to shooting on a real location. Critical production scenarios which allow filming only at short frames, like dawn or dusk, can be used continuously for an arbitrary duration in a virtual environment. Logistics of blocking streets for a production involving a dialogue in a driving car are no longer required if the scene can be convincingly produced in a virtual production scenario.

Moreover, this technique also allows situating film sets in locations that are inaccessible, or completely based on fictitious content. This applies for studio productions as well as for productions at outside locations where the real foreground is merged with virtual environment at medium to far distance. An example is shown in Figure 23.1 showcasing the real location which was shot at a former airfield in East Germany. The lower image displays the final scenery in which the location was extended to be situated in Guantanamo, Cuba.

For extending the scenery, digital assets and matte painting were created and combined with the filmed material. To fit the virtual elements into the filmed scene, it was necessary to match the set, the camera, and the

Figure 23.1: Set extension before and after. ©Filmakademie Baden-Wuerttemberg, 5 Jahre Leben, 2013.

lighting conditions. Additionally, the virtual camera needed to be as close as possible to the real one, including movement, lens effects, color, and even imperfections and artifacts caused by the real camera. The first step was to perform an exact match move to get information about the camera's movement and the lens geometry. Lighting and the surface definition were typically realized during, or directly after, the 3D asset creation process using digital content creation applications like Autodesk Maya or Maxon Cinema 3D. All 2D extensions, matte paintings, and cardboard styled assets were painted in tools like Adobe Photoshop or similar. These elements could also be animated or mixed with filmed footage. In the final step the generated assets were combined, visually matched and color grated in a compositing tool like the Foundry's Nuke or Adobe After Effects. For the given example the effective working time was about 20 man-days: 12 for the asset generation, 5 for the match move, 3 days compositing, and 1 day for revision requests.

A complete technical solution that enables virtual set production scenarios with realistic backdrops has been developed by Stargate Studios.[7] To meet the growing demand for low-cost alternatives to location shooting, Stargate Studios has developed the "next generation" stock footage library which specializes in virtual environments and customizable stock "locations." Branded by Stargate Studios in 2002 as the VIRTUAL BACKLOT (VB), the VB Library allows actors to literally walk into pre-shot footage—as if they were actually on location—without ever having to leave town.

The Virtual Backlot uses various techniques and methods to provide a range of options for creative production staff. The major components on producing in studio or outdoors are 3D rendering, camera tracking, and color keying (green and blue screen), all in real-time. Essential to the whole workflow is pre-production where, for example, the locations defined in the script are pre-recorded only with a small VFX team, instead of moving the whole production crew of up to 100 people to a designated destination. The recordings are called plates or background plates because they are used to replace the green screen.

Already a proven success in broadcast television, the VB "immersive environments" Library is beyond anything available in today's stock footage market. The environments vary from panoramic photographed and video sequences up to full 3D sceneries. Hundreds of hours of immersive footage have been acquired from multiple locations and successfully used by more than 165 feature, television, and commercial productions like the PanAm Demo[8] shown in Figure 23.2. As in 2014, 12 major television series are using VB on a regular basis. The cost model is so compelling that television networks are actually adopting creative content to match what is available in the VB library.

The most challenging and therefore also expensive piece of technology is the real-time component because its development is both time- and knowledge-consuming, and the market is still small. To register the movements of a film camera in real-time, only a few different technologies are available that allow the use in- and outdoors and are simple enough to be set up in as short as 1-3 hours on a stage or outdoors.

One of the most recent systems used in VB is the Lightcraft Previzion[9] system using Intersense technology. Here fiducial markers with circular barcodes are recorded by a small camera. The signal is processed in real-time to identify where each of the up to 300 circular barcodes is placed on the set and allow the real-time 3D render engine to define the current position and render a matching video stream according to the practical camera signal with only a short delay of 7 frames (@25 fps in 1920x1080).

[7] http://www.stargatestudios.net/
[8] http://vimeo.com/25483317
[9] http://www.lightcrafttech.com

Figure 23.2: PanAm Demo using Stargate VB, top to bottom: camera green screen material, on set previz (pre-composite), final composite after post-production. ©Stargate Studios.

A weaknesses of the Intersense tracking system is that its tracking camera has only a preset resolution, operational frame rate and light sensitivity. This means the fiducial markers have to be lit properly, have to be 100% fix (no movement is allowed). In some cases the limitation of only 300 fiducials allows for only a limited space to be tracked. The mounting of both fiducials and the Intersense camera on the production camera can be complicated.

Another extra procedure added to the production is the requirement to calibrate the lenses of the production camera before green screen shooting. This requires extra effort, and in the case of the Intersense system, a complex rig using pattern recognition on a large board of roughly the size of a mid-size desktop. To measure a prime lens (for example 35 mm Ultra Prime) takes from 30 to 60 minutes. To measure a zoom lens takes about 6-7 hours per lens. It is common to use, on average, film shooting 10 to 20 different lenses. Every single individual lens has to be calibrated because none is alike.

23.4 Real-Time Rendering with Game Engines

Game engines include reliable tools for real-time visualization addressing some of the needs of virtual production scenarios [Nitsche 08]. One solution currently in development is the Cinebox™system from the game developer studio Crytek™. The Cinebox™system is based on the Cryengine™[10] and provides extended functionality to meet filmmaking requirements. The software was used in the 3-minute branded-entertainment virtual production graduation short-film Dark Matter[11] [Y. Sahin and Backhaus 14].

The Dark Matter production combined digital video footage in 3k (ARRIRAW 16:9, 2880 x 1620 pixels) with real-time renderings from the game engine for live, on-set visualization, and the final render for the post-production. The game engine was the main and only tool for lighting and rendering of the short film. The workflow allowed having an editor on-set working with the produced media and preparing new versions of the short-film by the end of each production day, which could then be reviewed by the entire team.

One challenging issue in the workflow was to synchronize the game engine output, the live camera imagery, and the motion-capture data in a frame accurate way. This was accomplished by manually delaying the data streams using a custom software solution. Nevertheless, the game engine signal was not delivered at a fixed delay, which in minor cases resulted in asynchronous image output. A possible solution for this problem would be the integration of the live camera image directly in the game engine, thus reducing the complexity of syncing two sources to just one. Another aspect was camera tracking accuracy. Given that the motion capture systems precision was in the sub-millimeter range, the position definition of the real

[10] http://www.crytek.com/cryengine
[11] http://youtu.be/XlDL8hHVdDw

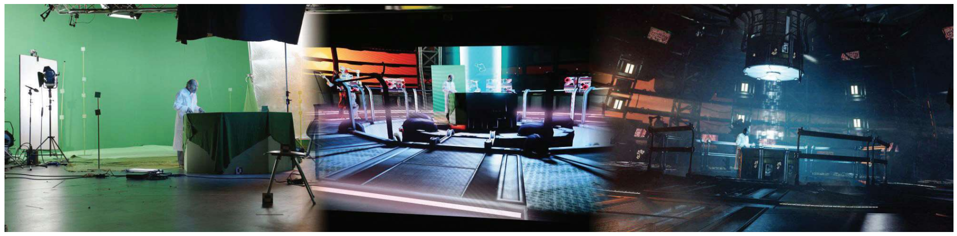


Figure 23.3: The Dark Matter production combined real camera footage with computer-generated imagery from game engine to provide interactive high-quality previsualization directly on a film set. ©Filmakademie Baden-Wuerttemberg, Dark Matter, 2014.

cameras film-back and the lenses nodal point was not sufficient. This led to a mismatching overlay of the real-footage and the CG elements. In post-production, the accuracy was still sufficient for most of the takes, and only 4 of the total of 33 sequences involving dynamic camera movement needed additional match move handling.

The alignment of the virtual with the real scene elements turned out to be another important issue. Lens and camera parameters were encoded as metadata [ALE 14] located in non-picture regions of the individual image frames, known as vertical ancillary data (VANC), to match the virtual lenses with the real ones. Real-time, high-quality chroma-keying and extraction of the actors, and practical assets have been other major obstacles to overcome in rendering the final images directly on set. In traditional VFX productions this is established as a classical post-production process, whereas in this production a dedicated hardware chroma-keying device was used that combined the game engine output with the keyed camera signal.

The Dark Matter short-film showcased an exemplary production pipeline for generating interactive, high-quality previsualization directly on the film set. Lighting parameters of the real set could be adjusted in the virtual scenery on the fly. The director of photography was able to preview the composed video signal directly in the camera. The visual consistency between on-set and final image was extremely close due to the continuous use of the same game engine technology throughout the entire production process. Exchanging and integrating assets in the post process was also possible in real-time. This resulted in a much more efficient post-production phase requiring only half of the time compared to a traditional workflow. Figure 23.3 shows the live camera, the real-time output of the game engine combined with the camera feed, and the final image.

23.5 Summary

Virtual production processes promise a new, more collaborative and interactive form of filmmaking. Advances in visual computing, in particular in sensing real-world parameters of cameras, assets and actors on one hand and progress in rendering and visualization on the other side show currently a major impact on the production process. With more visual computing tools to come it is expected that the process will be using more and more virtual techniques.

The state-of-the-art today is in most cases a rather complex custom development of all required technological aspects. Further development toward open standards will help to overcome some of these obstacles and allow for a more seamless exchange of data between individual hard- and software solutions. There are only a few turn-key solutions which can be picked up effortlessly. Game engines provide flexible, real-time tools, and their rendering quality and capabilities are constantly improving. Many film productions, however, demand physically correct image synthesis which in most cases relies on ray tracing. The future will show if these two can be combined in a meaningful way to support virtual production processes.

Preparing for virtual productions will also require a lot of rethinking of how assets and creative decisions will be made. The new creative freedom to make visual decisions directly on-set might also mix up established production pipelines (where most decisions are taken in post-production). Virtual production tools are also highly capable of democratizing the filmmaking process to individuals, independent filmmakers, tight budget, and academic productions.

Bibliography

[Adabala et al. 03] Neeharika Adabala, Nadia Magnenat-Thalmann, and Guangzheng Fei. "Visualization of Woven Cloth." In *Proc. of the 14th Eurographics Workshop on Rendering*, pp. 178–185. Eurographics Association, 2003.

[Adam et al. 09] Michael Adam, Christoph Jung, Stefan Roth, and Guido Brunnett. "Real-Time Stereo-Image Stitching Using GPU-Based Belief Propagation." In *Proc. Vision, Modeling and Visualization (VMV)*. Braunschweig, Germany, 2009.

[Addison et al. 95] Alonzo C. Addison, Douglas MacLeod, Gerald Margolis, Beit Hashoah, Michael Naimark, and Hans-Peter Schwarz. "Museums without Walls (panel session): New Media for New Museums." In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '95*, pp. 480–481. New York, NY, USA: ACM, 1995.

[Adobe Systems Inc. 13] Adobe Systems Inc. "After Effects." 2013. Http://www.adobe.com/products/aftereffects.html.

[Agarwal and Triggs 06] Ankur Agarwal and Bill Triggs. "Recovering 3D Human Pose from Monocular Images." *IEEE Transactions on PAMI* 28:1 (2006), 44–58.

[Agarwal et al. 10] Sameer Agarwal, Noah Snavely, Steven M. Seitz, and Richard Szeliski. "Bundle Adjustment in the Large." In *Computer Vision–ECCV 2010*, pp. 29–42. Springer, 2010.

[Agarwal et al. 11] Sameer Agarwal, Yasutaka Furukawa, Noah Snavely, Ian Simon, Brian Curless, Steven M. Seitz, and Richard Szeliski. "Building Rome in a Day." *Comm. ACM*. 2011.

[Aggarwal and Ahuja 04] Manoj Aggarwal and Narendra Ahuja. "Split Aperture Imaging for High Dynamic Range." *International Journal of Computer Vision* 58:1 (2004), 7–17.

[Agisoft 12] Agisoft. "PhotoScan." http://www.agisoft.ru/, 2012.

[Agrawal et al. 06] Amit K. Agrawal, Ramesh Raskar, and Rama Chellappa. "What Is the Range of Surface Reconstructions from a Gradient Field?" In *ECCV*, pp. 578–591, 2006.

[Ahmed et al. 01] Amr Ahmed, Farzin Mokhtarian, and Adrian Hilton. "Parametric Motion Blending through Wavelet Analysis." In *EUROGRAPHICS. Proceedings of Short Presentations*, pp. 347–353, 2001.

[Alahi et al. 12] Alexandre Alahi, Raphael Ortiz, and Pierre Vandergheynst. "Freak: Fast Retina Keypoint." In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pp. 510–517. IEEE, 2012.

[ALE 14] "Metadata in the ALEXA—Digital Workflow Solutions." Technical report, Arnold & Richter Cine Technik GmbH & Co. Betriebs KG, 2014. Available online (http://www.arri.com/camera/alexa/downloads).

[Alexa and Adamson 09] Marc Alexa and Anders Adamson. "Interpolatory Point Set Surfaces—Convexity and Hermite Data." *ACM Transactions on Graphics* 28:2 (2009), 20:1–20:10.

[Alexa et al. 01] Marc Alexa, Johannes Behr, Daniel Cohen-Or, Shachar Fleishman, David Levin, and Claudio T. Silva. "Point Set Surfaces." In *Proceedings of the IEEE Conference on Visualization '01*, pp. 21–28, 2001.

[Alexa et al. 04] M. Alexa, S. Rusinkiewicz, and Anders Adamson. "On Normals and Projection Operators for Surfaces Defined by Point Sets." In *Proceedings of the Eurographics Symposium on Point-Based Graphics 2004*, pp. 149–155, 2004.

[Alexander et al. 10] Oleg Alexander, Mike Rogers, William Lambeth, Jen-Yuan Chiang, Wan-Chun Ma, Chuan-Chang Wang, and Paul Debevec. "The Digital Emily Project: Achieving a Photorealistic Digital Actor." *IEEE Computer Graphics and Applications.*

[Allen et al. 02] B. Allen, B. Curless, and Z. Popović. "Articulated Body Deformation from Range Scan Data." *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 21:3 (2002), 612–619.

[Allen et al. 03] B. Allen, B. Curless, and Z. Popović. "The Space of Human Body Shapes: Reconstruction and Parameterization from Range Scans." *ACM Transactions on Graphics* 22:3 (2003), 587–594. Proceedings of SIGGRAPH.

[Allen et al. 06] B. Allen, B. Curless, Z. Popović, and A. Hertzmann. "Learning a Correlated Model of Identity and Pose-Dependent Body Shape Variation for Real-Time Synthesis." In *Symposium on Computer Animation*, pp. 147–156, 2006.

[Alvarez et al. 00] Luis Alvarez, Joachim Weickert, and Javier Sánchez. "Reliable Estimation of Dense Optical Flow Fields with Large Displacements." *International Journal of Computer Vision* 39:1 (2000), 41–56.

[Alvarez et al. 07] Luis Alvarez, Rachid Deriche, Théo Papadopoulo, and Javier Sánchez. "Symmetrical Dense Optical Flow Estimation with Occlusions Detection." *International Journal of Computer Vision* 75:3 (2007), 371–385.

[Amenta et al. 01] Nina Amenta, Sunghee Choi, and Krishna Kolluri. "The Power Crust." In *Proceedings of the 6th ACM Symposium on Solid Modeling*, pp. 249–260, 2001.

[Anandan 89] P. Anandan. "A Computational Framework and an Algorithm for the Measurement of Visual Motion." *International Journal of Computer Vision* 2:3 (1989), 283–310.

[Anderson et al. 11] Robert Anderson, Björn Stenger, and Roberto Cipolla. "Color Photometric Stereo for Multicolored Surfaces." In *ICCV*, 2011.

[Anguelov et al. 04a] D. Anguelov, D. Koller, H. C. Pang, P. Srinivasan, and S. Thrun. "Recovering Articulated Object Models from 3D Range Data." In *Proceedings of AUAI*, pp. 18–26, 2004.

[Anguelov et al. 04b] Dragomir Anguelov, Praveen Srinivasan, Hoi-Cheung Pang, Daphne Koller, Sebastian Thrun, and James Davis. "The Correlated Correspondence Algorithm for Unsupervised Registration of Nonrigid Surfaces." In *NIPS*, 2004.

[Anguelov et al. 05] Dragomir Anguelov, Praveen Srinivasan, Daphne Koller, Sebastian Thrun, Jim Rodgers, and James Davis. "SCAPE: Shape Completion and Animation of People." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 24:3 (2005), 408–416.

[Anisimov et al. 77] Sergei I. Anisimov, V. A. Benderski, and G. Farkas. "Nonlinear Photoelectric Emission from Metals Induced by a Laser Radiation." *Soviet Physics Uspekhi* 20:6 (1977), 467. Available online (http://stacks.iop.org/0038-5670/20/i=6/a=R01).

[Arikan and Forsyth 02] O. Arikan and D.A. Forsyth. "Synthesizing Constrained Motions from Examples." In *Proc. ACM SIGGRAPH*, 2002.

[Arvo 86] James Arvo. "Backward Ray Tracing." In *Developments in Ray Tracing, ACM SIGGRAPH Course Notes*, pp. 259–263, 1986.

[Ashikmin et al. 00] Michael Ashikmin, Simon Premože, and Peter Shirley. "A Microfacet-Based BRDF Generator." In *Proc. of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 65–74. ACM Press/Addison-Wesley Publishing Co., 2000.

[Atherton and Caporael 85] Peter R. Atherton and Linnda R. Caporael. "A Subjective Judgment Study of Polygon Based Curved Surface Imagery." *SIGCHI Bull.* 16:4 (1985), 27–34. Available online (http://doi.acm.org/10.1145/1165385.317462).

[Aujay et al. 07] G. Aujay, F. Hétroy, F. Lazarus, and C. Depraz. "Harmonic Skeleton for Realistic Character Animation." In *Proceedings of the ACM Symposium on Computer Animation (SCA)*, pp. 151–160, 2007.

[Autodesk 12] Autodesk. "123D Catch." http://www.123dapp.com/catch, 2012.

[Azouz et al. 06] Z. Ben Azouz, C. Shu, and A. Mantel. "Automatic Locating of Anthropometric Landmarks on 3D Human Models." In *Symposium on 3D Data Processing, Visualization, and Transmission*, pp. 750–757, 2006.

[Baak et al. 13] Andreas Baak, Meinard Müller, Gaurav Bharaj, Hans-Peter Seidel, and Christian Theobalt. "A Data-Driven Approach for Real-Time Full Body Pose Reconstruction from a Depth Camera." In *Consumer Depth Cameras for Computer Vision*, pp. 71–98. Springer, 2013.

[Badler et al. 87] N. I. Badler, K. H. Manoochehri, and Baraff D. "Multi-Dimensional Input Techniques and Articulated Figure Positioning by Multiple Constraints." In *Proc. of SI3D*, pp. 151–169, 1987.

[Baker et al. 11] Simon Baker, Daniel Scharstein, J. P. Lewis, Stefan Roth, Michael J. Black, and Richard Szeliski. "A Database and Evaluation Methodology for Optical Flow." *Int. J. Comput. Vision* 92:1 (2011), 1–31.

[Balan and Black 08] A. Balan and M. Black. "The Naked Truth: Estimating Body Shape under Clothing." In *European Conference on Computer Vision*, pp. 15–29, 2008.

[Balan et al. 07] A. Balan, L. Sigal, M. Black, J. Davis, and H. Haussecker. "Detailed Human Shape and Pose from Images." In *Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2007.

[Ballan et al. 10] Luca Ballan, Gabriel J. Brostow, Jens Puwein, and Marc Pollefeys. "Unstructured Video-Based Rendering: Interactive Exploration of Casually Captured Videos." *ACM Trans. on Graphics (Proc. SIGGRAPH)* 29:3 (2010), 87:1–87:11.

[Ballan et al. 12] Luca Ballan, Aparna Taneja, Jürgen Gall, Luc Van Gool, and Marc Pollefeys. "Motion Capture of Hands in Action Using Discriminative Salient Points." In *Computer Vision–ECCV 2012*, pp. 640–653. Springer, 2012.

[Baraff and Witkin 98] David Baraff and Andrew Witkin. "Large Steps in Cloth Simulation." In *Proc. 25th Annual Conf. on Computer Graphics and Interactive Techniques, SIGGRAPH '98*, pp. 43–54. New York, NY, USA: ACM, 1998.

[Baran and Popović 07] Ilya Baran and Jovan Popović. "Automatic Rigging and Animation of 3D Characters." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 26:3 (2007), 72.

[Baran et al. 09] I. Baran, D Vlasic, E. Grinspun, and J. Popovic. "Semantic Deformation Transfer." *ACM Transactions on Graphics (Proc. ACM SIGGRAPH 2009)* 28:3.

[Barron et al. 94] John L. Barron, David J. Fleet, and Steven S. Beauchemin. "Performance of Optical Flow Techniques." *Int. J. Comput. Vision* 12:1 (1994), 43–77.

[Bartczak and Koch 09] Bogumil Bartczak and Reinhard Koch. "Dense Depth Maps from Low Resolution Time-of-Flight Depth and High Resolution Color Views." In *ISVC (2)*, pp. 228–239, 2009.

[Baumgart 74] Bruce Guenther Baumgart. "Geometric modeling for computer vision." Ph.D. thesis, Stanford, CA, USA, 1974. AAI7506806.

[Bay et al. 06] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. "Surf: Speeded Up Robust Features." In *Computer Vision–ECCV 2006*, pp. 404–417. Springer, 2006.

[Beardsley et al. 97] Paul A. Beardsley, Andrew Zisserman, and David W. Murray. "Sequential Updating of Projective and Affine Structure from Motion." *International Journal of Computer Vision* 23:3 (1997), 235–259.

[Beder and Steffen 06] Christian Beder and Richard Steffen. "Determining an Initial Image Pair for Fixing the Scale of a 3D Reconstruction from an Image Sequence." In *Pattern Recognition, 4174*, edited by K. Franke, K.-R. Müller, B. Nickolay, and R. Schäfer, pp. 657–666. Springer, 2006.

[Beder et al. 07] Christian Beder, Bogumil Bartczak, and Reinhard Koch. "A Combined Approach for Estimating Patchlets from PMD Depth Images and Stereo Intensity Images." In *DAGM-Symposium*, pp. 11–20, 2007.

[Beeler et al. 10] Thabo Beeler, Bernd Bickel, Paul Beardsley, Bob Sumner, and Markus Gross. "High-Quality Single-Shot Capture of Facial Geometry." *ACM Transactions on Graphics (TOG)* 29:4 (2010), 40.

[Beeler et al. 11] Thabo Beeler, Fabian Hahn, Derek Bradley, Bernd Bickel, Paul Beardsley, Craig Gotsman, Robert W. Sumner, and Markus Gross. "High-Quality Passive Facial Performance Capture Using Anchor Frames." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 30 (2011), 75:1–75:10.

[Beier and Neely 92] Thaddeus Beier and Shawn Neely. "Feature-Based Image Metamorphosis." *Computer Graphics (Proc. of SIGGRAPH'93)* 26:2 (1992), 35–42.

[Belbachir et al. 12] Ahmed Nabil Belbachir, Manfred Mayerhofer, Daniel Matolin, and Joseph Colineau. "360SCAN: High-Speed Rotating Line Sensor for Real-Time 360 Panoramic Vision." In *Proc. 6th Int. Conf on Distributed Smart Cameras (ICDSC)*, pp. 1–6. Hongkong, 2012.

[Bentley 75] Jon Louis Bentley. "Multidimensional Binary Search Trees Used for Associative Searching." *Communication of the ACM* 18:9 (1975), 509–517.

[Berends and Erkelens 01] Ellen M Berends and Casper J Erkelens. "Adaptation to Disparity but Not to Perceived Depth." *Vision Research* 41:7 (2001), 883–892.

[Berent and Dragotti 06] J. Berent and P.L. Dragotti. "Segmentation of Epipolar-Plane Image Volumes with Occlusion and Disocclusion Competition." In *IEEE 8th Workshop on Multimedia Signal Processing*, pp. 182–185, 2006.

[Berger et al. 11] Kai Berger, Kai Ruhl, Christian Brümmer, Yannic Schröder, Alexander Scholz, and Marcus Magnor. "Markerless Motion Capture Using Multiple Color-Depth Sensors." In *Vision, Modeling, and Visualization (VMV)*, pp. 317–324, 2011.

[Berger et al. 14] M. Berger, A. Tagliasacchi, L. M. Seversky, P. Alliez, J. A. Levine, A. Sharf, and C. T. Silva. "State of the Art in Surface Reconstruction from Point Clouds." In *EUROGRAPHICS State-of-the-Art Report*, 2014.

[Besl and McKay 92] P. Besl and N. McKay. "A Method for Registration of 3-D Shapes." *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)* 14:2 (1992), 239–258.

[Bhat et al. 03] Kiran. S. Bhat, Christopher. D. Twigg, Jessica. K. Hodgins, Pradeep. K. Khosla, Zoran. Popović, and Steven. M. Seitz. "Estimating Cloth Simulation Parameters from Video." In *Proc. ACM SIGGRAPH/Eurographics Symp. on Computer Animation*, pp. 37–51. Eurographics Association, 2003.

[Bhat et al. 13] Kiran Bhat, Rony Goldenthal, Yuting Ye, Ronald Mallet, and Michael Koperwas. "High Fidelity Facial Animation Capture and Retargeting with Contours." *ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp. 7–14.

[Bickel et al. 07] Bernd Bickel, Mario Botsch, Roland Angst, Wojciech Matusik, Miguel Otaduy, Hanspeter Pfister, and Markus Gross. "Multi-Scale Capture of Facial Geometry and Motion." *ACM Transactions on Graphics* 26:3 (2007).

[Birchfield and Tomasi 98] Stan Birchfield and Carlo Tomasi. "A Pixel Dissimilarity Measure That Is Insensitive to Image Sampling." *IEEE Trans. Pattern Anal. Mach. Intell.* 20:4 (1998), 401–406.

[Bishop and Favaro 12] T. Bishop and P. Favaro. "The Light Field Camera: Extended Depth of Field, Aliasing, and Superresolution." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34:5 (2012), 972–986.

[Blache et al. 14] Ludovic Blache, Céline Loscos, Olivier Nocent, and Laurent Lucas. "3D Volume Matching for Mesh Animation of Moving Actors." In *Eurographics Workshop on 3D Object Retrieval (3DOR)*, pp. 69–76. Eurographics, 2014.

[Black and Anandan 96] Michael J. Black and P. Anandan. "The Robust Estimation of Multiple Motions: Parametric and Piecewise-smooth Flow Fields." *Comput. Vis. Image Underst.* 63:1 (1996), 75–104.

[Bleyer and Breiteneder 13] Michael Bleyer and Christian Breiteneder. "Stereo Matching—State-of-the-Art and Research Challenges." In *Advanced Topics in Computer Vision*, pp. 143–179, 2013.

[Bleyer and Gelautz 07] Michael Bleyer and Margrit Gelautz. "Graph-cut-based Stereo Matching Using Image Segmentation with Symmetrical Treatment of Occlusions." *Image Commun.* 22:2 (2007), 127–143.

[Bleyer and Gelautz 08] Michael Bleyer and Margrit Gelautz. "Simple but Effective Tree Structures for Dynamic Programming-Based Stereo Matching." In *Conference on Computer Vision Theory and Applications*, pp. 415–422, 2008.

[Bleyer et al. 11a] Michael Bleyer, Christoph Rhemann, and Carsten Rother. "PatchMatch Stereo—Stereo Matching with Slanted Support Windows." In *British Machine Vision Conference*, pp. 14.1–14.11, 2011.

[Bleyer et al. 11b] Michael Bleyer, Carsten Rother, Pushmeet Kohli, Daniel Scharstein, and Sudipta Sinha. "Object Stereo—Joint Stereo Matching an Object Segmentation." In *Computer Vision and Pattern Recognition*, 2011.

[Bleyer et al. 12] Michael Bleyer, Christoph Rhemann, and Carsten Rother. "Extracting 3D Scene Consistent Object Proposals and Depth from Stereo Images." In *European Conference on Computer Vision*, 2012.

[Blinn and Newell 76] James F. Blinn and Martin E. Newell. "Texture and Reflection in Computer Generated Images." *Commun. ACM* 19:10 (1976), 542–547.

[Blonde et al. 96] Laurent Blonde, Wolfgang Niem, Yakup Paker, Matthias Buck, Ricardo Galli, Wolfgang Schmidt, and Graham Thomas. "A Virtual Studio for Live Broadcasting: The Mona Lisa Project." *IEEE Multimedia* 3:2 (1996), 18–29.

[Blumenthal-Barby and Eisert 14] David Blumenthal-Barby and Peter Eisert. "High-Resolution Depth for Binocular Image-Based Modelling." *Computers & Graphics* 39 (2014), 89–100.

[Boff et al. 88] K.R. Boff and J.E. Lincoln. *Engineering Data Compendium–Human Perception and Performance, Vol. I, Sensory Process and Perception.* New York: John Wiley & Sons, 1988.

[Bolles et al. 87] R.C. Bolles, H.H. Baker, and D.H. Marimont. "Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion." *International Journal of Computer Vision* 1:1 (1987), 7–55.

[Bonnard et al. 13] J. Bonnard, G. Valette, C. Loscos, and J.-M. Nourrit. "3D HDR Acquisition and Restitution." In *3D Video: From Capture to Diffusion.* John Wiley & Sons, 2013.

[Borshukov et al. 03] George Borshukov, Jefferson Montgomery, and John Hable. *Playable Universal Capture, GPU Gems 3.* Addison-Wesley, 2003.

[Borshukov et al. 05] George Borshukov, Dan Piponi, Oystein Larsen, John Peter Lewis, and Christina Tempelaar-Lietz. "Universal Capture—Image-Based Facial Animation for "The Matrix Reloaded."" In *SIGGRAPH Course Notes*, 2005.

[Botsch and Sorkine 08] M. Botsch and O. Sorkine. "On Linear Variational Surface Deformation Methods." *IEEE Transactions on Visualization and Computer Graphics* 4:1 (2008), 213–230.

[Botsch et al. 10] Mario Botsch, Leif Kobbelt, Mark Pauly, Pierre Alliez, and Bruno Levy. *Polygon Mesh Processing.* A. K. Peters/CRC Press, 2010.

[Bougnoux 98] Sylvain Bougnoux. "From Projective to Euclidean Space under Any Practical Situation, a Criticism of Self-Calibration." In *Computer Vision, 1998. Sixth International Conference on*, pp. 790–796. IEEE, 1998.

[Bouguet 08] Jean-Yves Bouguet. "Camera Calibration Toolbox for MATLAB®.", 2008. Available online (http://www.vision.caltech.edu/bouguetj/calib_doc/.).

[Bouman et al. 13] Katherine L. Bouman, Bei Xiao, Peter Battaglia, and William T. Freeman. "Estimating the Material Properties of Fabric from Video." In *Proc. Int. Conf. on Computer Vision*, pp. 1984–1991, 2013.

[Boykov et al. 01] Yuri Boykov, Olga Veksler, and Ramin Zabih. "Fast Approximate Energy Minimization via Graph Cuts." *IEEE Trans. Pattern Anal. Mach. Intell.* 23:11 (2001), 1222–1239.

[Bradley et al. 08] Derek Bradley, Tiberiu Popa, Alla Sheffer, Wolfgang Heidrich, and Tamy Boubekeur. "Markerless Garment Capture." *ACM Trans. Graph.* 27:3 (2008), 99:1–99:9.

[Bradley et al. 10] Derek Bradley, Wolfgang Heidrich, Tiberiu Popa, and Alla Sheffer. "High Resolution Passive Facial Performance Capture." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 29:4 (2010), 41:1–41:10.

[Bradski 00] Gary Bradski. "The OpenCV Library." *Doctor Dobbs Journal* 25:11 (2000), 120–126.

[Brainard 97] David H Brainard. "The Psychophysics Toolbox." *Spatial Vision* 10:4 (1997), 433–436.

[Breen et al. 94] David E. Breen, Donald H. House, and Michael J. Wozny. "Predicting the Drape of Woven Cloth Using Interacting Particles." In *Proc. of the 21st Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '94*, pp. 365–372. New York, NY, USA: ACM, 1994.

[Bregler et al. 04] C. Bregler, J. Malik, and K. Pullen. "Twist Based Acquisition and Tracking of Animal and Human Kinematics." *IJCV* 56:3 (2004), 179–194.

[Bridson et al. 03] Robert Bridson, Sebastian Marino, and Ronald Fedkiw. "Simulation of Clothing with Folds and Wrinkles." In *Proc. of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, SCA '03*, pp. 28–36. Eurographics Association, 2003.

[Bronstein et al. 07] Alex M. Bronstein, Michael M. Bronstein, and Ron Kimmel. "Calculus of Non-Rigid Surfaces for Geometry and Texture Manipulation." *IEEE Transactions on VCG* 13(5) (2007), 902–913.

[Brown and Rusinkiewicz 07] Benedict J. Brown and Szymon Rusinkiewicz. "Global Non-rigid Alignment of 3-D Scans." *ACM Transactions on Graphics* 26:3.

[Brown 66] D. C. Brown. "Decentering Distortion of Lenses." *Photogrammetric Engineering* 32:3 (1966), 444–462.

[Brox and Malik 11] Thomas Brox and Jitendra Malik. "Large Displacement Optical Flow: Descriptor Matching in Variational Motion Estimation." *IEEE Trans. Pattern Anal. Mach. Intell.* 33:3 (2011), 500–513.

[Brox et al. 04] Thomas Brox, Andrés Bruhn, Nils Papenberg, and Joachim Weickert. "High Accuracy Optical Flow Estimation based on a Theory for Warping." In *European Conference on Computer Vision (ECCV)*, Lecture Notes in Computer Science, 3024, pp. 25–36. Springer, 2004.

[Buades et al. 05] Antoni Buades, Bartomeu Coll, and J-M. Morel. "A Non-Local Algorithm for Image Denoising." In *IEEE CVPR*, pp. 60–65, 2005.

[Budd et al. 12] Chris Budd, Oliver Grau, and Peter Schübel. "Web Delivery of Free-Viewpoint Video of Sport Events."

[Budd et al. 13] Chris Budd, Peng Huang, Martin Klaudiny, and Adrian Hilton. "Global Non-rigid Alignment of Surface Sequences." *International Journal of Computer Vision* 102:1-3 (2013), 256–270.

[Buehler et al. 01] Chris Buehler, Michael Bosse, Leonard McMillan, Steven Gortler, and Michael Cohen. "Unstructured Lumigraph Rendering." In *ACM SIGGRAPH*, pp. 425–432, 2001.

[Butler et al. 12] Alex Butler, Shahram Izadi, Otmar Hilliges, David Molyneaux, Steve Hodges, and David Kim. "Shake'n'Sense: Reducing Structured Light Interference when Multiple Depth Cameras Overlap." In *Proceedings of Human Factors in Computing Systems (ACM CHI)*. New York, NY, USA: ACM, 2012.

[Cagniart et al. 10] Cedric Cagniart, Edmond Boyer, and Slobodan Ilic. "Probabilistic Deformable Surface Tracking from Multiple Videos." In *ECCV*, pp. 326–339, 2010.

[Calonder et al. 10] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. "Brief: Binary Robust Independent Elementary Features." In *Computer Vision–ECCV 2010*, pp. 778–792. Springer, 2010.

[Carranza et al. 03] J. Carranza, C. Theobalt, M. Magnor, and H.-.P. Seidel. "Free-Viewpoint Video of Human Actors." *ACM Transactions on Graphics (Proc. ACM SIGGRAPH)* 22:3 (2003), 569–577.

[Casas et al. 12a] Dan Casas, Margara Tejera, Jean-Yves Guillemaut, and Adrian Hilton. "4D Parametric Motion Graphs for Interactive Animation." In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pp. 103–110. ACM, 2012.

[Casas et al. 12b] Dan Casas, Margara Tejera, Jean-Yves Guillemaut, and Adrian Hilton. "Parametric Animation of Performance-Captured Mesh Sequences." *Journal of Visualization and Computer Animation* 23:2 (2012), 101–111.

[Casas et al. 13] Dan Casas, Margara Tejera, Jean-Yves Guillemaut, and Adrian Hilton. "Interactive Animation of 4D Performance Capture." *IEEE Transactions on Visualization and Computer Graphics (TVCG)* 19:5 (2013), 762–773.

[Casas et al. 14] Dan Casas, Marco Volino, John Collomosse, and Adrian Hilton. "4D Video Textures for Interactive Character Appearance." *Computer Graphics Forum (Proc. EUROGRAPHICS 2014)* 33:2.

[Chai et al. 00] Jin-Xiang Chai, Xin Tong, Shing-Chow Chan, and Heung-Yeung Shum. "Plenoptic Sampling." In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 307–318. ACM Press/Addison-Wesley Publishing Co., 2000.

[Chaminade et al. 07] Thierry Chaminade, Jessica Hodgins, and Mitsuo Kawato. "Anthropomorphism Influences Perception of Computer-Animated Characters Actions." *Social Cognitive and Affective Neuroscience* 2:3 (2007), 206–216.

[Chan et al. 08] Derek Chan, Hylke Buisman, Christian Theobalt, and Sebastian Thrun. "A Noise-Aware Filter for Real-Time Depth Upsampling." In *ECCV Workshop on Multi-Camera and Multi-Modal Sensor Fusion Algorithms and Applications*, edited by Andrea Cavallaro and Hamid Aghajan, pp. 1–12. Marseille, France, 2008.

[Chandaria et al. 07] Jigna Chandaria, Graham A. Thomas, and Didier Stricker. *J. Real-Time Image Processing* 2:2-3 (2007), 69–79.

[Chatterjee et al. 11] Priyam Chatterjee, Neel Joshi, Sing Bing Kang, and Yasuyuki Matsushita. "Noise Suppression in Low-Light Images through Joint Denoising and Demosaicing." In *IEEE CVPR*, pp. 321–328, 2011.

[Chaurasia et al. 11] Gaurav Chaurasia, Olga Sorkine, and George Drettakis. "Silhouette-Aware Warping for Image-Based Rendering." *Computer Graphics Forum* 30:4 (2011), 1223–1232.

[Chazal and Lieuter 05] F. Chazal and A. Lieuter. "The λ Medial Axis." *Graphical Models* 67:4 (2005), 304–331.

[Chen and Williams 93] Shenchang Eric Chen and Lance Williams. "View Interpolation for Image Synthesis." In *Proc. of ACM SIGGRAPH'93*, pp. 279–288. New York: ACM Press/ACM SIGGRAPH, 1993.

[Chen et al. 03] Chia-Yen Chen, Reinhard Klette, and Chi-Fa Chen. "Shape from Photometric Stereo and Contours." In *CAIP*, pp. 377–384, 2003.

[Chen et al. 13a] J. Chen, D. Bautembach, and S. Izadi. "Scalable Real-time Volumetric Surface Reconstruction." *ACM Trans. Graph. (Proc. SIGGRAPH)* 32:4 (2013), 113:1–113:16.

[Chen et al. 13b] Y. Chen, Z. Liu, and Z. Zhang. "Tensor-Based Human Body Modeling." In *Conference on Computer Vision and Pattern Recognition*, 2013.

[Chen 95] Shenchang Eric Chen. "QuickTime VR: An Image-Based Approach to Virtual Environment Navigation." In *Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pp. 29–38, 1995.

[Choi and Ko 02] Kwang-Jin Choi and Hyeong-Seok Ko. "Stable but Responsive Cloth." *ACM Trans. Graph.* 21:3 (2002), 604–611.

[Choi and Ko 05] Kwang-Jin Choi and Hyeong-Seok Ko. "Research Problems in Clothing Simulation." *Comput. Aided Des.* 37:6 (2005), 585–592.

[Chum et al. 07] Ondrej Chum, James Philbin, Josef Sivic, Michael Isard, and Andrew Zisserman. "Total Recall: Automatic Query Expansion with a Generative Feature Model for Object Retrieval." In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pp. 1–8. IEEE, 2007.

[Chun 12] Won Chun. "WebGL Models: End-to-End." In *OpenGL Insights*, edited by Patrick Cozzi and Christophe Riccio, pp. 431–454. CRC Press, 2012.

[Collins 96] Robert T. Collins. "A Space-Sweep Approach to True Multi-Image Matching." In *Computer Vision and Pattern Recognition*, pp. 358–363. IEEE, 1996.

[Cootes et al. 98] Tim Cootes, Gareth J. Edwards, and Christopher J.Taylor. "Active Appearance Models." In *Proc. European Conference on Computer Vision ECCV98*, pp. 484–498, 1998.

[Cordier and Magnenat-Thalmann 05] Frederic Cordier and Nadia Magnenat-Thalmann. "A Data-Driven Approach for Real-Time Clothes Simulation." *Comput. Graph. Forum* 24 (2005), 173–183.

[Criminisi et al. 03] A. Criminisi, P. Perez, and K. Toyama. "Object Removal by Exemplar-Based Inpainting." In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, 2, 2, pp. 721–728. IEEE, 2003.

[Criminisi et al. 05] A. Criminisi, S.B. Kang, R. Swaminathan, R. Szeliski, and P. Anandan. "Extracting Layers and Analyzing their Specular Properties using Epipolar-Plane-Image Analysis." *Computer Vision and Image Understanding* 97:1 (2005), 51–85.

[Cryer et al. 95] James Edwin Cryer, Ping-Sing Tsai, and Mubarak Shah. "Integration of Shape from Shading and Stereo." *Pattern Recognition Society* 28:7 (1995), 1033–1043.

[Cui et al. 10] Yan Cui, Sebastian Schuon, Derek Chan, Sebastian Thrun, and Christian Theobalt. "3D Shape Scanning with a Time-of-Flight Camera." In *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 1173–1180. IEEE, 2010.

[Curless and Levoy 96] B. Curless and M. Levoy. "A Volumetric Method for Building Complex Models from Range Images." In *Proc. Comp. Graph. & Interact. Techn.*, pp. 303–312, 1996.

[Dabov et al. 07] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering." *IEEE Transactions on Image Processing* 16:8 (2007), 2080–2095.

[Dachsbacher et al. 13] Carsten Dachsbacher, Jaroslav Křivánek, Milos Hasan, Adam Arbree, Bruce Walter, and Jan Novák. "Scalable Realistic Rendering with Many-Light Methods." *Computer Graphics Forum* DOI: 10.1111/cgf.12256.

[Daly 93] Scott Daly. *Digital Images and Human Vision.* edited by Andrew B. Watson, The Visible Differences Predictor: An Algorithm for the Assessment of Image Fidelity, pp. 179–206. Cambridge, MA, USA: MIT Press, 1993.

[Danette Allen et al. 01] B. Danette Allen, Gary Bishop, and Greg Welch. "Tracking: Beyond 15 Minutes of Thought." Technical Report Course 11, SIGGRAPH 2001, 2001.

[Dansereau et al. 13] D.G. Dansereau, O. Pizarro, and S.B. Williams. "Calibration and Rectification for Lenselet-Based Plenoptic Cameras." In *Proc. International Conference on Computer Vision and Pattern Recognition*, 2013.

[Davis et al. 12] Abe Davis, Marc Levoy, and Frdo Durand. "Unstructured Light Fields." *Comput. Graph. Forum* 31:2 (2012), 305–314.

[Dawes et al. 09] Robert Dawes, Jigna Chandaria, and Graham Thomas. "Image-Based Camera Tracking for Athletics." In *Broadband Multimedia Systems and Broadcasting, 2009. BMSB'09. IEEE International Symposium on*, pp. 1–6. IEEE, 2009.

[de Aguiar and Ukita 14] E. de Aguiar and N. Ukita. "Representing Mesh-Based Character Animations." *Computers and Graphics* 38:0 (2014), 10–17.

[de Aguiar et al. 07] Edilson de Aguiar, Christian Theobalt, Carsten Stoll, and Hans-Peter Seidel. "Marker-Less Deformable Mesh Tracking for Human Shape and Motion Capture." In *IEEE CVPR*, 2007.

[de Aguiar et al. 08a] E. de Aguiar, C. Theobalt, S. Thrun, and H.-P. Seidel. "Automatic Conversion of Mesh Animations into Skeleton-Based Animations." *Computer Graphics Forum* 27:2.

[de Aguiar et al. 08b] E.de de Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H.-P. Seidel, and S. Thrun. "Performance Capture from Sparse Multi-View Video." *ACM Transactions on Graphics (Proc. SIGGRAPH 2008)* 27:3 (2008), 1–10.

[de Aguiar et al. 10] E. de Aguiar, L. Sigal, A. Treuille, and J. K. Hodgins. "Stable Spaces for Real-Time Clothing." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 9:4 (2010), Art.106.

[Debevec and Malik 97] Paul E. Debevec and Jitendra Malik. "Recovering High Dynamic Range Radiance Maps from Photographs." In *Proceedings of ACM Siggraph '97 (Computer Graphics)*, pp. 369–378, 1997.

[Debevec et al. 96] Paul E. Debevec, Camillo J. Taylor, and Jitendra Malik. "Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach." In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '96*, pp. 11–20. New York, NY, USA: ACM, 1996.

[Debevec et al. 98] Paul Debevec, Yizhou Yu, and George Borshukov. "Efficient View-Dependent Image-Based Rendering with Projective Texture-Mapping." In *Rendering Techniques 98*, *Eurographics*, edited by George Drettakis and Nelson Max, pp. 105–116. Springer Vienna, 1998.

[Debevec et al. 00] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. "Acquiring the Reflectance Field of a Human Face." In *SIGGRAPH '00*, pp. 145–156, 2000.

[Debevec 98] Paul Debevec. "Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-Based Graphics with Global Illumination and High Dynamic Range Photography." In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '98*, pp. 189–198. New York, NY, USA: ACM, 1998.

[DeCarlo and Metaxas 00] Douglas DeCarlo and Dimitris Metaxas. "Optical Flow Constraints on Deformable Models with Applications to Face Tracking." *International Journal of Computer Vision* 38:2 (2000), 99–127.

[Decaudin et al. 06] Philippe Decaudin, Dan Julius, Jamie Wither, Laurence Boissieux, Alla Sheffer, and Marie-Paule Cani. "Virtual Garments: A Fully Geometric Approach for Clothing Design." *Computer Graphics Forum* 25:3.

[Dellaert et al. 00] Frank Dellaert, Steven M. Seitz, Charles E. Thorpe, and Sebastian Thrun. "Structure from Motion without Correspondence." In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, 2, 2, pp. 557–564. IEEE, 2000.

[DeMenthon and Davis 95] D. DeMenthon and L.S. Davis. "Model-Based Object Pose in 25 Lines of Code." *International Journal of Computer Vision* :15 (1995), 123–141.

[Dempster et al. 77] A. P. Dempster, N. M. Laird, and D. B. Rubin. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society, Series B* 39:1 (1977), 1–38.

[Deng et al. 05] Yi Deng, Qiong Yang, Xueyin Lin, and Xiaoou Tang. "A Symmetric Patch-Based Correspondence Model for Occlusion Handling." In *International Conference on Computer Vision*, pp. 1316–1322, 2005.

[Deutscher et al. 00] J. Deutscher, A. Blake, and I. Reid. "Articulated Body Motion Capture by Annealed Particle Filtering." In *CVPR*, pp. 1144–1149, 2000.

[Devernay and Beardsley 10] Frédéric Devernay and Paul Beardsley. "Stereoscopic Cinema." In *Image and Geometry Processing for 3-D Cinematography*, Geometry and Computing, 5, edited by R. Ronfard and G. Taubin, 5, pp. 11–51. Springer Berlin Heidelberg, 2010.

[Devernay and Faugeras 01] F. Devernay and O. Faugeras. "Straight Lines Have to Be Straight: Automatic Calibration and Removal of Distortion from Scenes of Structured Environments." *Machine Vision and Applications* 13:1 (2001), 14–24.

[Doboš et al. 13] Jozef Doboš, Kristian Sons, Dmitri Rubinstein, Philipp Slusallek, and Anthony Steed. "XML3DRepo: A REST API for Version Controlled 3D Assets on the Web." In *Proceedings of the 18th International Conference on 3D Web Technology, Web3D '13*. ACM, 2013. To be published.

[Dodgson 02] Neil A. Dodgson. "Analysis of the Viewing Zone of Multi-View Autostereoscopic Displays." In *Stereoscopic Displays and Applications XIII*, pp. 254–265. Proc SPIE 4660, 2002.

[Donner and Jensen 06] Craig Donner and Henrik Wann Jensen. "A Spectral BSSRDF for Shading Human Skin." In *Eurographics Symposium on Rendering*, 2006.

[Donner et al. 08] Craig Donner, Tim Weyrich, Eugene d'Eon, Ravi Ramamoorthi, and Szymon Rusinkiewicz. "A Layered, Heterogeneous Reflectance Model for Acquiring and Rendering Human Skin." *ACM Transactions on Graphics (Proc. SIGGRAPH Asia 2008)* 27:5 (2008), 1–12.

[Dorrington et al. 11] A. Dorrington, J. Godbaz, M. Cree, A. Payne, and L. Streeter. "Separating True Range Measurements from Multi-Path and Scattering Interference in Commercial Range Cameras." In *Proc. IS&T/SPIE Electronic Imaging*, pp. 786404–786404, 2011.

[Drago and Myszkowski 01] Frédéric Drago and Karol Myszkowski. "Validation Proposal for Global Illumination and Rendering Techniques." *Computers & Graphics* 25:3 (2001), 511–518.

[Drettakis et al. 97] George Drettakis, Luc Robert, and Sylvain Bougnoux. "Interactive Common Illumination for Computer Augmented Reality." In *Rendering Techniques*, pp. 45–56, 1997.

[Dryden and Mardia 02] I. Dryden and K. Mardia. *Statistical Shape Analysis.* Wiley, 2002.

[Dubois 01] Eric Dubois. "A Projection Method to Generate Anaglyph Stereo Images." In *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, pp. 1661–1664. IEEE Computer Society Press, 2001.

[Duda et al. 01] R. Duda, P. Hart, and D. Stork. *Pattern Classification, Second Edition.* John Wiley & Sons, Inc., 2001.

[Duncan 09] Jody Duncan. "The Unusual Birth of Benjamin Button." In *Cinefex 116*, 2009.

[Duncan 10] Jody Duncan. "The Seduction of Reality." In *Cinefex 120*, 2010.

[Duveau et al. 12] Estelle Duveau, Simon Courtemanche, Lionel Reveret, and Edmond Boyer. "Cage-based Motion Recovery using Manifold Learning." In *IEEE 3DIMPVT*, pp. 206–213, 2012.

[DŹmura 91] Michel DŹmura. "Shading Ambiguity: Reflectance and Illumination." *Computational Models of Visual Processing*, pp. 187–207.

[Eberhardt et al. 96] Bernhard Eberhardt, Andreas Weber, and Wolfgang Strasser. "A Fast, Flexible, Particle-System Model for Cloth Draping." *IEEE Comput. Graph. Appl.* 16:5 (1996), 52–59.

[Eden et al. 06] Ashley Eden, Matthew Uyttendaele, and Richard Szeliski. "Seamless Image Stitching of Scenes with Large Motions and Exposure Differences." In *Proc. Int. Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 2498–2505. New York, USA, 2006.

[Efros and Freeman 01] Alexei A. Efros and William T. Freeman. "Image Quilting for Texture Synthesis and Transfer." In *Proceedings of the 28th*

Annual Conference on Computer Graphics and Interactive Techniques, pp. 341–346. ACM, 2001.

[Einarsson et al. 06] Per Einarsson, Charles-Felix Chabert, Andrew Jones, Wan-Chun Ma, Bruce Lamond, Tim Hawkins, Mark Bolas, Sebastian Sylwan, and Paul Debevec. “Relighting Human Locomotion with Flowed Reflectance Fields.” In *Rendering Techniques 2006: 17th Eurographics Workshop on Rendering*, pp. 183–194, 2006.

[Eisemann et al. 07] Martin Eisemann, Anita Sellent, and Marcus Magnor. “Filtered Blending: A New, Minimal Reconstruction Filter for Ghosting-Free Projective Texturing with Multiple Images.” In *Vision, Modeling and Visualization (VMV)*, pp. 119–126, 2007.

[Eisemann et al. 08] Martin Eisemann, Bert De Decker, Marcus Magnor, Philippe Bekaert, Edilson de Aguiar, Naveed Ahmed, Christian Theobalt, and Anita Sellent. “Floating Textures.” *Computer Graphics Forum (Proc. Eurographics)* 27:2 (2008), 409–418.

[Eisert 03] Peter Eisert. “MPEG-4 Facial Animation in Video Analysis and Synthesis.” *International Journal of Imaging Systems and Technology* 13:5 (2003), 245–256.

[Ekman and Friesen 78] Paul Ekman and Wallace Friesen. *Facial Action Coding System: A Technique for the Measurement of Facial Movement.* Palo Alto: Consulting Psychologists Press, 1978.

[Elhayek et al. 12] A. Elhayek, Carsten Stoll, Nils Hasler, Kwang In Kim, H. Seidel, and Christian Theobalt. “Spatio-Temporal Motion Tracking with Unsynchronized Cameras.” In *Proc. CVPR*, pp. 1870–1877. IEEE, 2012.

[English and Bridson 08] Elliot English and Robert Bridson. “Animating Developable Surfaces Using Nonconforming Elements.” *ACM Trans. Graph.* 27:3 (2008), 66:1–66:5.

[Etzmuss et al. 03] Olaf Etzmuss, Michael Keckeisen, and Wolfgang Strasser. “A Fast Finite Element Solution for Cloth Modelling.” In *Proc. of the 11th Pacific Conference on Computer Graphics and Applications, PG ’03*, pp. 244–251. Washington, DC, USA: IEEE Computer Society, 2003.

[Fairchild 13] Mark D. Fairchild. *Color Appearance Models.* John Wiley & Sons, 2013.

[Falie and Buzuloiu 08] Dragos Falie and Vasile Buzuloiu. “Distance Errors Correction for the Time of Flight (ToF) Cameras.” In *Proc. European Conf. on Circuits and Systems for Communications*, pp. 193–196, 2008.

[Fan et al. 12] Shaojing Fan, Tian-Tsong Ng, Jonathan S. Herberg, Bryan L. Koenig, and Shiqing Xin. “Real or Fake?: Human Judgments about Photographs and Computer-Generated Images of Faces.” In *SIGGRAPH Asia 2012 Technical Briefs, SA ’12*, pp. 17:1–17:4. New York, NY, USA: ACM, 2012.

[Fechteler et al. 07] P. Fechteler, P. Eisert, and J. Rurainsky. “Fast and High Resolution 3D Face Scanning.” In *ICIP07*, 3, 3, pp. 81–84, 2007.

[Fehn et al. 07] Christoph Fehn, Christoph Weissig, Ingo Feldmann, Markus Müller, Peter Eisert, Peter Kauff, and Hans Bloß. "Creation of High-Resolution Video Panoramas for Sport Events." *International Journal of Semantic Computing (IJSC)* 1:2 (2007), 493–505.

[Fehn 03] Christoph Fehn. "A 3D-TV Approach Using Depth-Image-Based Rendering (DIBR)." In *Proc. of VIIP*, 3, 3, 2003.

[Fehn 04] Christoph Fehn. "Depth-Image-Based Rendering (DIBR), Compression, and Transmission for a New Approach on 3D-TV." *Proc. SPIE* 5291:93 (2004), 93–104.

[Fender and Julesz 67] Derek Fender and Bela Julesz. "Extension of Panums Fusional Area in Binocularly Stabilized Vision." *JOSA* 57:6 (1967), 819–826.

[Feng et al. 10] Wei-Wen Feng, Yizhou Yu, and Byung-Uck Kim. "A Deformation Transformer for Real-Time Cloth Animation." *ACM Trans. Graph.* 29:4 (2010), 108:1–108:9.

[Ferwerda et al. 97] James A. Ferwerda, Peter Shirley, Sumanta N. Pattanaik, and Donald P. Greenberg. "A Model of Visual Masking for Computer Graphics." In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '97*, pp. 143–152. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1997.

[Ferwerda et al. 04] James A. Ferwerda, Stephen H. Westin, Randall C. Smith, and Richard Pawlicki. "Effects of Rendering on Shape Perception in Automobile Design." In *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization, APGV '04*, pp. 107–114. New York, NY, USA: ACM, 2004.

[Ferwerda 08] James A. Ferwerda. "Psychophysics 101: How to Run Perception Experiments in Computer Graphics." In *ACM SIGGRAPH 2008 Classes, SIGGRAPH '08*, pp. 87:1–87:60. New York, NY, USA: ACM, 2008.

[Fischler and Bolles 81] Martin A. Fischler and Robert C. Bolles. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography." *Communications of the ACM* 24:6 (1981), 381–395.

[Fleishman et al. 05] Shachar Fleishman, Daniel Cohen-Or, and Cláudio T. Silva. "Robust Moving Least-Squares Fitting with Sharp Features." *ACM Transactions on Graphics* 24:3 (2005), 544–552.

[Fleming 14] Roland W. Fleming. "Visual Perception of Materials and Their Properties." *Vision Research* 94 (2014), 62–75.

[Foote et al. 04] Jonathan Foote, Qiong Liu, Don Kimber, Patrick Chiu, and Frank Zhao. "Reach-Through-the-Screen: A New Metaphor for Remote Collaboration." In *Proc. 5th Pacific Rim Conference on Multimedia (PCM)*, pp. 73–80. Tokyo, Japan, 2004.

[Fordham 03] Joe Fordham. "Middle Earth Strikes Back." In *Cinefex 92*, 2003.

[Fordham 04] Joe Fordham. "Armed and Dangerous." In *Cinefex 99*, 2004.

[Fordham 09] Joe Fordham. "The Manhattan Project." In *Cinefex 117*, 2009.

[Fournier et al. 93] Alain Fournier, Atjeng S. Gunawan, and Chris Romanzin. "Common Illumination between Real and Computer Generated Scenes." In *Proceedings of Graphics Interface '93*, pp. 254–262, 1993.

[Frahm et al. 10] Jan-Michael Frahm, Pierre Fite-Georgel, David Gallup, Tim Johnson, Rahul Raguram, Changchang Wu, Yi-Hung Jen, Enrique Dunn, Brian Clipp, Svetlana Lazebnik, and Marc Pollefeys. "Building Rome on a Cloudless Day." In *Proceedings of the 11th European Conference on Computer Vision: Part IV, ECCV'10*, pp. 368–381, 2010.

[Franco and Boyer 09] Jean-Sébastien Franco and Edmond Boyer. "Efficient Polyhedral Modeling from Silhouettes." *IEEE Transactions on PAMI* 31:3 (2009), 414–427.

[Fuchs and May 08] Stefan Fuchs and Stefan May. "Calibration and Registration for Precise Surface Reconstruction with Time-Of-Flight Cameras." *Int. J. on Intell. Systems Techn. and App., Issue on Dynamic 3D Imaging* 5:3/4 (2008), 278–284.

[Fuchs et al. 05] Martin Fuchs, Volker Blanz, and Hans-Peter Seidel. "Bayesian Relighting." In *Eurographics Symposium on Rendering (EGSR)*, pp. 157–164, 2005.

[Fuchs et al. 07] Martin Fuchs, Volker Blanz, Hendrik P.A. Lensch, and Hans-Peter Seidel. "Adaptive Sampling of Reflectance Fields." *ACM Transactions on Graphics* 26:2 (2007), 10.

[Furukawa and Ponce 08] Yasutaka Furukawa and Jean Ponce. "Dense 3D Motion Capture from Synchronized Video Streams." In *CVPR*, pp. 193–211. IEEE Computer Society, 2008.

[Furukawa and Ponce 10] Yasutaka Furukawa and Jean Ponce. "Accurate, Dense, and Robust Multiview Stereopsis." *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 32:8 (2010), 1362–1376.

[Fusiello et al. 97] Aandrea Fusiello, Vito Roberto, and Emanuele Trucco. "Efficient Stereo with Multiple Windowing." In *Conference on Computer Vision and Pattern Recognition*, pp. 858–863, 1997.

[Fusiello 07] A. Fusiello. "Specifying Virtual Cameras in Uncalibrated View Synthesis." *Circuits and Systems for Video Technology, IEEE Transactions on* 17:5 (2007), 604 –611.

[Gagvani and Silver 99] N. Gagvani and D. Silver. "Parameter Controlled Volume Thinning." *Graphical Models and Image Processing* 61:3 (1999), 149–164.

[Gal et al. 10] R. Gal, Y. Wexler, E. Ofek, H. Hoppe, and D. Cohen-Or. "Seamless Montage for Texturing Models." In *Computer Graphics Forum*, 29, pp. 479–486. Wiley Online Library, 2010.

[Gall et al. 09] Juergen Gall, Carsten Stoll, Edilson De Aguiar, Christian Theobalt, Bodo Rosenhahn, and Hans-Peter Seidel. "Motion Capture Using Joint Skeleton Tracking and Surface Estimation." In *IEEE Computer Vision and Patttern Recognition (CVPR)*, pp. 1746–1753, 2009.

[Gall et al. 10] Juergen Gall, Bodo Rosenhahn, Thomas Brox, and Hans-Peter Seidel. "Optimization and Filtering for Human Motion Capture." *International Journal of Computer Vision* 87:1-2 (2010), 75–92.

[Gallup et al. 07] David Gallup, Jan-Michael Frahm, Philippos Mordohai, Qingxiong Yang, and Marc Pollefeys. "Real-Time Plane-Sweeping Stereo with Multiple Sweeping Directions." In *Conference on Computer Vision and Pattern Recognition*, 2007.

[Ganapathi et al. 10] Varun Ganapathi, Christian Plagemann, Daphne Koller, and Sebastian Thrun. "Real-Time Motion Capture Using a Single Time-of-Flight Camera." In *CVPR*, pp. 755–762, 2010.

[Garg et al. 06] Gaurav Garg, Eino-Ville Talvala, Marc Levoy, and Hendrik P. A. Lensch. "Symmetric Photography: Exploiting Data-Sparseness in Reflectance Fields." In *Rendering Techniques*, pp. 251–262, 2006.

[Gargallo and Sturm 05] Pau Gargallo and Peter Sturm. "Bayesian 3D Modeling from Images Using Multiple Depth Maps." In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) Volume 02*, pp. 885–891, 2005.

[Garland and Heckbert 97] M. Garland and P.S. Heckbert. "Surface Simplification Using Quadric Error Metrics." In *ACM SIGGRAPH*, pp. 209–216, 1997.

[Gehrig et al. 12] Stefan K. Gehrig, Hernán Badino, and Uwe Franke. "Improving Sub-Pixel Accuracy for Long Range Stereo." *Computer Vision and Image Understanding* 116:1 (2012), 16–24.

[Geiger 12] Andreas Geiger. "Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite." In *Computer Vision and Pattern Recognition*, pp. 3354–3361, 2012.

[Georgiev et al. 11] T. Georgiev, A. Lumsdaine, and G. Chunev. "Using Focused Plenoptic Cameras for Rich Image Capture." *CGA* 31:1 (2011), 62–73.

[Georgiev et al. 12] Iliyan Georgiev, Jaroslav Krivanek, Tomas Davidovic, and Philipp Slusallek. "Light Transport Simulation with Vertex Connection and Merging." *ACM Transactions on Graphics (Proc. SIGGRAPH Asia* 31:6 (2012), 192:1–192:10.

[Germann et al. 10] Marcel Germann, Alexander Hornung, Richard Keiser, Remo Ziegler, Stephan Würmlin, and Markus Gross. "Articulated Billboards for Video-Based Rendering." *Comput. Graphics Forum (Proc. Eurographics)* 29:2 (2010), 585–594.

[Germann et al. 12] Marcel Germann, Tiberiu Popa, Richard Keiser, Remo Ziegler, and Markus Gross. "Novel-View Synthesis of Outdoor Sport

Events Using an Adaptive View-Dependent Geometry." *Comput. Graphics Forum (Proc. Eurographics)* 31:2 (2012), 325–333.

[Germann 12] Marcel Germann. "Video-Based Rendering Techniques." Ph.D. thesis, Zürich, 2012. Diss., Eidgenssische Technische Hochschule ETH Zürich, Nr. 20290, 2012.

[Gershun 36] A. Gershun. "The Light Field." *J. Math. and Physics* 18 (1936), 51–151.

[Ghosh et al. 10] Abhijeet Ghosh, Wolfgang Heidrich, Shruthi Achutha, and Matthew O'Toole. "A Basis Illumination Approach to BRDF Measurement." *International Journal on Computer Vision* 90:2 (2010), 183–197.

[Ghosh et al. 11] Abhijeet Ghosh, Graham Fyffee, Borom Tunwattanapong, Jay Busch, Xueming Yu, and Paul Debevec. "Multiview Face Capture Using Polarized Spherical Gradient Illumination." In *ACM Transactions on Graphics (Proc. SIGGRAPH Asia 2011)*, 2011.

[Giachetti et al. 98] Andrea Giachetti, Marco Campani, and Vincent Torre. "The Use of Optical Flow for Road Navigation." *IEEE T. Robotics and Automation* 14:1 (1998), 34–48.

[Gibson and Mirtich 97] Sarah F. F. Gibson and Brian Mirtich. "A Survey of Deformable Modeling in Computer Graphics, Technical Report TR-97-19, MERL." Technical report, Cambridge, MA, 1997.

[Giesen et al. 09] J. Giesen, B. Miklos, M. Pauly, and C. Wormser. "The Scale Axis Transform." In *Proceedings of the Annual Symposium on Computational Geometry*, pp. 106–115, 2009.

[Giraudot et al. 13] Simon Giraudot, David Cohen-Steiner, and Pierre Alliez. "Noise-Adaptive Shape Reconstruction from Raw Point Sets." *Computer Graphics Forum* 32:5 (2013), 229–238.

[Glondu et al. 12] Loeiz Glondu, Lien Muguercia, Maud Marchal, Carles Bosch, Holly Rushmeier, Georges Dumont, and George Drettakis. "Example-Based Fractured Appearance." *Computer Graphics Forum* 31 (2012), 1547–1556. Available online (http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8659.2012.03151.x/full).

[Goesele et al. 00] Michael Goesele, Wolfgang Heidrich, Hendrik P. A. Lensch, and Hans-Peter Seidel. "Building a Photo Studio for Measurement Purposes." In *VMV*, pp. 231–238, 2000.

[Goesele et al. 03] Michael Goesele, Xavier Granier, Wolfgang Heidrich, and Hans-Peter Seidel. "Accurate Light Source Acquisition and Rendering." *ACM Trans. Graph.* 22:3 (2003), 621–630.

[Goesele et al. 04] Michael Goesele, Hendrik P. A. Lensch, Jochen Lang, Christian Fuchs, and Hans-Peter Seidel. "DISCO: Acquisition of Translucent Objects." *ACM Trans. Graph.* 23:3 (2004), 835–844.

[Goesele et al. 06] Michael Goesele, Brian Curless, and Steven M. Seitz. "Multi-View Stereo Revisited." In *Computer Vision and Pattern Recognition*, CVPR '06, 2, CVPR '06, 2, pp. 2402–2409, 2006.

[Goesele et al. 07] Michael Goesele, Noah Snavely, Brian Curless, Hugues Hoppe, and Steven M. Seitz. "Multi-View Stereo for Community Photo Collections." In *International Conference on Computer Vision*, pp. 1–8, 2007.

[Goesele et al. 10] Michael Goesele, Jens Ackermann, Simon Fuhrmann, Carsten Haubold, Ronny Klowsky, and T.U. Darmstadt. "Ambient Point Clouds for View Interpolation." *ACM Trans. Graph.* 29 (2010), 95:1–95:6.

[Goldlücke and Magnor 03] B. Goldlücke and M. Magnor. "Real-time Microfacet Billboarding for Free-Viewpoint Video Rendering." In *IEEE International Conference on Image Processing (ICIP)*, 3, pp. 713–716. IEEE, 2003.

[Goldlücke and Magnor 04] B. Goldlücke and M. Magnor. "Space-time Isosurface Evolution for Temporally Coherent 3D Reconstruction." In *IEEE Computer Vision and Pattern Recognition (CVPR)*, 1, pp. 350–355, 2004.

[Goldlücke et al. 02] Bastian Goldlücke, Marcus Magnor, and Bennett Wilburn. "Hardware-Accelerated Dynamic Light Field Rendering." In *Vision, Modeling, and Visualization (VMV)*, pp. 455–462, 2002.

[Goldlücke and Magnor 04] Bastian Goldlücke and Marcus Magnor. "Weighted Minimal Hypersurfaces and Their Applications in Computer Vision." In *European Conference on Computer Vision (ECCV)*, pp. 366–378, 2004.

[Goldlücke and Wanner 13] B. Goldlücke and S. Wanner. "The Variational Structure of Disparity and Regularization of 4D Light Fields." In *Proc. International Conference on Computer Vision and Pattern Recognition*, 2013.

[google 12] google. "Google Sketchup." http://sketchup.google.com, 2012.

[Goorts et al. 13] Patrik Goorts, Cosmin Ancuti, Maarten Dumont, Sammy Rogmans, and Philippe Bekaert. "Real-Time Video-Based View Interpolation of Soccer Events Using Depth-Selective Plane Sweeping."

[Gortler et al. 96] S.J. Gortler, R. Grzeszczuk, R. Szeliski, and M.F. Cohen. "The Lumigraph." In *Proc. SIGGRAPH*, pp. 43–54, 1996.

[Granados et al. 10] Miguel Granados, Boris Ajdin, Michael Wand, Christian Theobalt, H.-P. Seidel, and H. Lensch. "Optimal HDR Reconstruction with Linear Digital Cameras." In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 215–222. IEEE, 2010.

[Grau and Easterbrook 08] Oliver Grau and Jim Easterbrook. "Effects of Camera Aperture Correction on Keying of Broadcast Video."

[Grau et al. 04] Oliver Grau, Tim Pullen, and Graham A. Thomas. "A Combined Studio Production System for 3-D Capturing of Live Action and Immersive Actor Feedback." *IEEE Transactions on Circuits and Systems for Video Technology* 14:3 (2004), 370–380.

[Grau 05] Oliver Grau. "A 3D Production Pipeline for Special Effects in TV and Film." In *Mirage 2005, Computer Vision/Computer Graphics Collaboration Techniques and Applications*, 2005.

[Grosch et al. 07] Thorsten Grosch, Tobias Eble, and Stefan Mueller. "Consistent Interactive Augmentation of Live Camera Images with Correct Near-Field Illumination." In *Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology*, pp. 125–132, 2007.

[Grosch 05] Thorsten Grosch. "Differential Photon Mapping—Consistent Augmentation of Photographs with Correction of all Light Paths." In *EG Short Presentations*, pp. 53–56, 2005.

[Grossberg and Nayar 04] Michael D. Grossberg and Shree K. Nayar. "Modeling the Space of Camera Response Functions." *IEEE Trans. on Pattern Analysis Machine Intelligence* 26:10 (2004), 1272–1282.

[Grossman and Balakrishnan 06] Tovi Grossman and Ravin Balakrishnan. "An Evaluation of Depth Perception on Volumetric Displays." In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 193–200. ACM, 2006.

[Guan et al. 09] P. Guan, A. Weiss, A. Bălan, and M. Black. "Estimating Human Shape and Pose from a Single Image." In *International Conference on Computer Vision*, 2009.

[Guan et al. 12] Peng Guan, Loretta Reiss, David Hirshberg, Alex Weiss, and Micheal J. Black. "DRAPE: DRessing Any PErson." *ACM Trans. Graph.* 31:3 (2012), 35:1–35:10.

[Gudmundsson et al. 08] Sigurjón Árni Gudmundsson, Henrik Aanæs, and Rasmus Larsen. "Fusion of Stereo Vision and Time-of-Flight Imaging for Improved 3D Estimation." *IJISTA* 5:3/4 (2008), 425–433.

[Guennebaud and Gross 07] Gaël Guennebaud and Markus Gross. "Algebraic Point Set Surfaces." *ACM Transactions on Graphics* 26:3.

[Guillemaut et al. 10] J.-Y. Guillemaut, M. Sarim, and A. Hilton. "Stereoscopic Content Production of Complex Dynamic Scenes Using a Wide-Baseline Monoscopic Camera Set-Up." In *Proc. International Conference on Image Processing (ICIP 2010), Special Session on Image Processing for Stereo Digital Cinema Production*, pp. 9–12, 2010.

[Guillemot et al. 12] Thiery Guillemot, Andrés Almansa, and Tamy Boubekeur. "Non-Local Point Set Surfaces." In *Proceedings of 3DIMPVT 2012*, pp. 61–70, 2012.

[Guomundsson and Sveinsson 11] SA Guomundsson and Johannes R Sveinsson. "TOF-CCD Image Fusion Using Complex Wavelets." In *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1557–1560, 2011.

[Guskov et al. 03] Igor Guskov, Sergey Klibanov, and Benjamin Bryant. "Trackable Surfaces." In *Proc. Eurographics Symp. on Computer Animation (SCA)*, pp. 251–257, 2003.

[Gvili et al. 03] Ronen Gvili, Amir Kaplan, Eyal Ofek, and Giora Yahav. "Depth Keying." In *Proc SPIE, Video-Based Image Techniques and Emerging Work*, 5006, 5006, 2003. DOI: 10.1117/12.474052.

[Habbecke and Kobbelt 06] Martin Habbecke and Leif Kobbelt. "Iterative Multi-View Plane Fitting." In *Vision, Modelling and Visualization*, pp. 73–80, 2006.

[Habbecke and Kobbelt 07] Martin Habbecke and Leif Kobbelt. "A Surface-Growing Approach to Multi-View Stereo Reconstruction." In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE, Conference on*, pp. 1–8, 2007.

[Haber et al. 09] Tom Haber, Christian Fuchs, Philippe Bekaert, Hans-Peter Seidel, Michael Goesele, and Hendrik P. A. Lensch. "Relighting Objects from Image Collections." In *Computer Vision and Pattern Recognition*, pp. 627–634, 2009.

[Hachisuka et al. 08] Toshiya Hachisuka, Shinji Ogaki, and Henrik Wann Jensen. "Progressive Photon Mapping." *ACM Transactions on Graphics* 27:5 (2008), 130:1–130:8.

[Hachisuka et al. 12] Toshiya Hachisuka, Jacopo Pantaleoni, and Henrik Wann Jensen. "A Path Space Extension for Robust Light Transport Simulation." *ACM Transactions on Graphics (Proc. SIGGRAPH Asia)* 31:6 (2012), 191:1–191:10.

[Hackbusch 99] Wolfgang Hackbusch. "A Sparse Matrix Arithmetic Based on H-matrices. Part I: Introduction to H-matrices." *Computing* 62:2 (1999), 89–108. Available online (http://dx.doi.org/10.1007/s006070050015).

[Hadap et al. 99] Sunil Hadap, Endre Bangerter, Pascal Volino, and Nadia Magnenat-Thakmann. "Animating Wrinkles on Cloth." In *Proc. of IEEE Conf. on Visualization*, pp. 175–182, 1999.

[Hahne and Alexa 09] Uwe Hahne and Marc Alexa. "Depth Imaging by Combining Time-of-Flight and On-Demand Stereo." In *Dyn3D*, pp. 70–83, 2009.

[Hall-Holt and Rusinkiewicz 01] Olaf Hall-Holt and Szymon Rusinkiewicz. "Stripe Boundary Codes for Real-Time Structured-light Range Scanning of Moving Objects." In *Proceedings of the Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001*, 2, 2, pp. 359–366, 2001.

[Hanika and Dachsbacher 14] Johannes Hanika and Carsten Dachsbacher. "Efficient Monte Carlo Rendering with Realistic Lenses." *Computer Graphics Forum (Proc. of Eurographics)* 33:2, 2014.

[Hao and Varshney 01] Xuejun Hao and Amitabh Varshney. "Variable-Precision Rendering." In *Proc. I3D*, pp. 149–158, 2001.

[Hara et al. 08] Kenji Hara, Ko Nishino, and Katsushi Ikeuchi. "Mixture of Spherical Distributions for Single-View Relighting." *IEEE Trans. Pattern Anal. Mach. Intell.* 30:1 (2008), 25–35.

[Haralick et al. 94] Robert M. Haralick, Chung-Nan Lee, Karsten Ottenberg, and Michael Nölle. "Review and Analysis of Solutions of the Three Point Perspective Pose Estimation Problem." *Int. J. Comput. Vision* 13:3 (1994), 331–356.

[Harker and O'Leary 08] Matthew Harker and Paul O'Leary. "Least Squares Surface Reconstruction from Measured Gradient Fields." In *CVPR*, 2008.

[Harris and Stephens 88] Chris Harris and Mike Stephens. "A Combined Corner and Edge Detector." In *Alvey Vision Conference*, 15, p. 50. Manchester, UK, 1988.

[Hartley and Kang 05] R. Hartley and S. B. Kang. "Parameter Free Radial Distortion Correction with Center of Distortion Estimation." Technical Report TR=2005-42, Microsoft Research, 2005.

[Hartley and Sturm 97] Richard Hartley and Peter Sturm. "Triangulation." *Computer Vision and Image Understanding* 68:2 (1997), 146–157.

[Hartley and Zisserman 03] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*, Second edition. Cambridge University Press, 2003.

[Hartley 93] Richard Hartley. "Extraction of Focal Lengths from the Fundamental Matrix." Technical Report, G.E. CRD, Schenectady, NY.

[Hartley 97] Richard Hartley. "In Defense of the Eight-Point Algorithm." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 19:6 (1997), 580–593.

[Hasinoff et al. 10] Samuel W. Hasinoff, Fredo Durand, and William T. Freeman. "Noise-Optimal Capture for High Dynamic Range Photography." *2013 IEEE Conference on Computer Vision and Pattern Recognition* (2010), 553–560.

[Hasler et al. 06] Nils Hasler, Mark Asbach, Bodo Rosenhahn, Jens-Rainer Ohm, and Hans-Peter Seidel. "Physically-Based Tracking of Cloth." In *Proc. of the Int. Workshop on Vision, Modeling, and Visualization, VMV*, pp. 49–56, 2006.

[Hasler et al. 09a] N. Hasler, B. Rosenhahn, T. Thormählen, M. Wand, J. Gall, and H.-P. Seidel. "Markerless Motion Capture with Unsynchronized Moving Cameras." In *Computer Vision and Pattern Recognition (CVPR)*, pp. 224–231, 2009.

[Hasler et al. 09b] N. Hasler, C. Stoll, B. Rosenhahn, T. Thormählen, and H.-P. Seidel. "Estimating Body Shape of Dressed Humans." *Computers and Graphics* 33:3 (2009), 211–216. Proceedings of Shape Modeling International.

[Hasler et al. 09c] N. Hasler, C. Stoll, M. Sunkel, B. Rosenhahn, and H.-P. Seidel. "A Statistical Model of Human Pose and Body Shape." *Computer Graphics Forum* 2:28 (2009), 337–346. Proceedings of Eurographics.

[Hasler et al. 10a] N. Hasler, H. Ackermann, B. Rosenhahn, T. Thormählen, and H.-P. Seidel. "Multilinear Pose and Body Shape Estimation of Dressed Subjects from Image Sets." In *Conference on Computer Vision and Pattern Recognition*, pp. 1823–1830, 2010.

[Hasler et al. 10b] N. Hasler, T. Thormahlen, B. Rosenhahn, and H.-P. Seidel. "Learning Skeletons for Shape and Pose." In *I3D10: Proc. of Symp. on Interactive 3D Graphics and Games*, pp. 23–30, 2010.

[Hauswiesner et al. 11] Stefan Hauswiesner, Matthias Straka, and Gerhard Reitmayr. "Image-Based Clothes Transfer." In *Proc. Int. Symp.on Mixed and Augmented Reality (ISMAR)*, pp. 169–172. IEEE Computer Society, 2011.

[Hauswiesner et al. 13] Stefan Hauswiesner, Matthias Straka, and Gerhard Reitmayr. "Virtual Try-On through Image-Based Rendering." *IEEE Trans. on Visualization and Computer Graphics* 19:9 (2013), 1552–1565.

[Havaldar et al. 06] Parag Havaldar, Fred Pighin, and John Peter Lewis. "Performance Driven Facial Animation." In *SIGGRAPH Course Notes 2006*, 2006.

[Hawkins et al. 07] Tim Hawkins, Jonathan Cohen, Chris Tchou, and Paul Debevec. "Light Stage 2.0." In *SIGGRAPH 2001 Sketch*, p. 229, 2007.

[He et al. 10] Kaiming He, Jian Sun, and Xiaoou Tang. "Guided Image Filtering." In *Computer Vision–ECCV 2010*, pp. 1–14. Springer, 2010.

[Heck and Gleicher 07] R. Heck and M. Gleicher. "Parametric Motion Graphs." In *ACM Symposium on Interactive 3D Graphics and Graphics*, 2007.

[Heer and Bostock 10] Jeffrey Heer and Michael Bostock. "Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 203–212. ACM, 2010.

[Heide et al. 13] Felix Heide, Matthias B. Hullin, James Gregson, and Wolfgang Heidrich. "Low-Budget Transient Imaging Using Photonic Mixer Devices." *ACM Trans. Graph.* 32:4 (2013), 45:1–45:10.

[Heidrich et al. 99] Wolfgang Heidrich, Hartmut Schirmacher, Hendrik Kück, and Hans-Peter Seidel. "A Warping-Based Refinement of Lumigraphs." In *Proc. WSCG*, 99, 99, pp. 102–109, 1999.

[Heikkila and Silven 97] J. Heikkila and O. Silven. "A Four-step Camera Calibration Procedure with Implicit Image Correction." In *IEEE CVPR*, 1997.

[Heinly et al. 12] Jared Heinly, Enrique Dunn, and Jan-Michael Frahm. "Comparative Evaluation of Binary Features." In *Computer Vision–ECCV 2012*, pp. 759–773. Springer, 2012.

[Helten et al. 13] T. Helten, A. Baak, G. Bharai, M. Müller, H.-P. Seidel, and C. Theobalt. "Personalization and Evaluation of a Real-Time Depth-Based Full Body Scanner." In *3D Vision*, 2013.

[Helzle et al. 04] Volker Helzle, Christoph Biehn, Thomas Schlömer, and Florian Linner. "Adaptable Setup for Performance Driven Facial Animation." *ACM Transactions on Graphics (Proc. SIGGRAPH 2004).*

[Herbst et al. 09] Evan Herbst, Steve Seitz, and Simon Baker. "Occlusion Reasoning for Temporal Interpolation Using Optical Flow." Technical report, Microsoft Research Technical Report, MSR-TR-2009-2014, 2009. No. MSR-TR-2009-2014.

[Herrera C. et al. 12] D. Herrera C., J. Kannala, and J. Heikkilä. "Joint Depth and Color Camera Calibration with Distortion Correction." *IEEE Trans. Pattern Anal. Mach. Intell.* 34:10 (2012), 2058–2064.

[Herzog et al. 12] Robert Herzog, Martin Cadik, Tunc O. Aydin, Kwang In Kim, Karol Myszkowski, and Hans-P. Seidel. "NoRM: No-Reference Image Quality Metric for Realistic Image Synthesis." *Comp. Graph. Forum* 31:2pt4 (2012), 545–554.

[Hildebrandt et al. 12] K. Hildebrandt, C. Schulz, C. von Tycowicz, and K. Polthier. "Interactive Space-Time Control of Deformable Objects." *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 31:4.

[Hill 53] A. J. Hill. "A Mathematical and Experimental Foundation for Stereoscopic Photography." In *Journal of SMPTE*, 61, pp. 461–486, 1953.

[Hilsmann et al. 10] Anna Hilsmann, David C. Schneider, and Peter Eisert. "Realistic Cloth Augmentation in Single View Video under Occlusions." *Computers & Graphics* 34:5 (2010), 567–574.

[Hilsmann et al. 11] Anna Hilsmann, David C. Schneider, and Peter Eisert. "Warp-Based Near-Regular Texture Analysis for Image-Based Texture Overlay." In *Proc. Int. Workshop on Vision, Modeling, and Visualization (VMV)*, pp. 73–80. Eurographics Association, 2011.

[Hilsmann et al. 13] Anna Hilsmann, Philipp Fechteler, and Peter Eisert. "Pose Space Image Based Rendering." *Comput. Graph. Forum (Proc. Eurographics)* 32:2 (2013), 265–274.

[Hilton et al. 10] Adrian Hilton, Jean-Yves Guillemaut, Joe Kilner, Oliver Grau, and Graham Thomas. "Free-Viewpoint Video for TV Sport Production." In *Image and Geometry Processing for 3-D Cinematography*, pp. 77–106. Springer, 2010.

[Hilton et al. 11] A. Hilton, J.-Y. Guillemaut, J. Kilner, O. Grau, and G. Thomas. "3D-TV Production from Conventional Cameras for Sports Broadcast." *IEEE Transactions on Broadcasting* 57:2 (2011), 462–476.

[Hirschmüller et al. 02] Heiko Hirschmüller, Peter R. Innocent, and Jon M. Garibaldi. "Real-Time Correlation-Based Stereo Vision with Reduced Border Errors." *International Journal of Computer Vision* 47:1-3 (2002), 229–246.

[Hirschmüller 05] Heiko Hirschmüller. "Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information." In *Computer Vision and Pattern Recognition*, 2, pp. 807–814, 2005.

[Hodgins et al. 98] Jessica K Hodgins, James F O'Brien, and Jack Tumblin. "Perception of Human Motion with Different Geometric Models." *Visualization and Computer Graphics, IEEE Transactions on* 4:4 (1998), 307–316.

[Högg et al. 13] T. Högg, D. Lefloch, and A. Kolb. "Real-Time Motion Artifact Compensation for PMD-ToF Images." In *Proc. Workshop Imaging New Modalities, German Conference of Pattern Recognition (GCPR)*, LNCS, 8200, LNCS, 8200, pp. 273–288. Springer, 2013.

[Hong and Chen 04] Li Hong and George Chen. "Segment-Based Stereo Matching Using Graph Cuts." In *Conference on Computer Vision and Pattern Recognition*, pp. 74–81, 2004.

[Hong et al. 10] Q. Y. Hong, S. I. Park, and J. K. Hodgins. "A Data-Driven Segmentation for the Shoulder Complex." *Computer Graphics Forum.* 29:2 (2010), 537–544.

[Hoppe et al. 92] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. "Surface Reconstruction from Unorganized Points." In *Proceedings of ACM SIGGRAPH 1992*, pp. 71–78, 1992.

[Hoppe 96] Hugues Hoppe. "Progressive Meshes." In *Proc. SIGGRAPH*, pp. 99–108, 1996.

[Hoppe 98] Hugues Hoppe. "Efficient Implementation of Progressive Meshes." *Computers & Graphics* 22 (1998), 27–36.

[Horaud et al. 09] Radu Horaud, Matti Niskanen, Guillaume Dewaele, and Edmond Boyer. "Human Motion Tracking by Registering an Articulated Surface to 3-D Points and Normals." *IEEE Transactions on PAMI* 31:1 (2009), 158–163.

[Horn and Brooks 86] Berthold K. P. Horn and Michael J. Brooks. "The Variational Approach to Shape from Shading." *Computer Vision, Graphics, and Image Processing* 33:2 (1986), 174–208.

[Horn and Schunck 81] Berthold K. P. Horn and Brian G. Schunck. "Determining Optical Flow." *Artificial Intelligence* 17 (1981), 185–203.

[Hornung and Kobbelt 06] Alexander Hornung and Leif Kobbelt. "Robust Reconstruction of Watertight 3D Models from Non-uniformly Sampled Point Clouds without Normal Information." In *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, pp. 41–50, 2006.

[Hornung and Kobbelt 09] Alexander Hornung and Leif Kobbelt. "Interactive Pixel-Accurate Free Viewpoint Rendering from Images with Silhouette Aware Sampling." *Computer Graphics Forum* 28:8 (2009), 2090–2103.

[Hosni et al. 09] Asmaa Hosni, Michael Bleyer, Margrit Gelautz, and Christoph Rhemann. "Local Stereo Matching Using Geodesic Support Weights." In *International Conference on Image Processing*, pp. 2069–2072, 2009.

[Hosni et al. 11a] Asmaa Hosni, Michael Bleyer, Christoph Rhemann, Margrit Gelautz, and Carsten Rother. "Real-Time Local Stereo Matching Using Guided Image Filtering." In *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, pp. 1–6. IEEE, 2011.

[Hosni et al. 11b] Asmaa Hosni, Christoph Rhemann, Michael Bleyer, Carsten Rother, and Margrit Gelautz. "Fast Cost-Volume Filtering for Visual Correspondence and Beyond." In *Computer Vision and Pattern Recognition*, pp. 3017–3024, 2011.

[Hosni et al. 13] Asmaa Hosni, Michael Bleyer, and Margrit Gelautz. "Secrets of Adaptive Support Weight Techniques for Local Stereo Matching." *Comput. Vis. Image Underst.* 117:6 (2013), 620–632.

[Hostica et al. 06] B. Hostica, P. Seitz, and A. Simoni. *Encyclopedia of Sensors*, 7, Chapter Optical Time-of-Flight Sensors for Solid-State 3D-vision, pp. 259–289. American Scientific Pub, 2006.

[House and Breen 00] Donald H. House and David E. Breen, editors. *Cloth Modeling and Animation.* Natick, MA, USA: A. K. Peters, Ltd., 2000.

[Howlett et al. 05] Sarah Howlett, John Hamill, and Carol O'Sullivan. "Predicting and Evaluating Saliency for Simplified Polygonal Models." *ACM Trans. Appl. Percept.* 2:3 (2005), 286–308.

[Hua et al. 07] Hong Hua, Narendra Ahuja, and Chunyu Gao. "Design Analysis of a High-Resolution Panoramic Camera Using Conventional Imagers and a Mirror Pyramid." *IEEE Trans. on Pattern Analysis and Machine Intelligence* 29:2 (2007), 356–945.

[Huang et al. 09] P. Huang, A. Hilton, and J. Starck. "Human Motion Synthesis from 3D Video." In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1478–1485, 2009.

[Huang et al. 10] Peng Huang, Adrian Hilton, and Jonathan Starck. "Shape Similarity for 3D Video Sequences of People." *International Journal of Computer Vision* 89:2-3 (2010), 362–381.

[Huang et al. 13] Chun-Hao Huang, Edmond Boyer, and Slobodan Ilic. "Robust Human Body Shape and Pose Tracking." In *3DV*, 2013.

[Hullin et al. 11] Matthias B. Hullin, Elmar Eisemann, Hans-Peter Seidel, and Sungkil Lee. "Physically-Based Real-Time Lens Flare Rendering." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 30:4 (2011), 108:1–108:9.

[Hypr3D Development Team 12] Hypr3D Development Team. "hypr3D." http://www.agisoft.ru/, 2012.

[Iddan and Yahav 01] G. J. Iddan and G. Yahav. "3D Imaging in the Studio." In *Proc. of SPIE*, 4298, 4298, pp. 48–56, 2001.

[Ihrke et al. 08] I. Ihrke, T. Stich, H. Gottschlich, M. Magnor, and H. Seidel. "Fast Incident Light Field Acquisition and Rendering." *Journal of WSCG*, pp. 177–184.

[Ihrke et al. 10] Ivo Ihrke, Kiriakos N Kutulakos, Hendrik Lensch, Marcus Magnor, and Wolfgang Heidrich. "Transparent and Specular Object Reconstruction." *Computer Graphics Forum (CGF)* 29:8 (2010), 2400–2426.

[Imre and Hilton 12] Evren Imre and Adrian Hilton. "Through-the-Lens Synchronisation for Heterogeneous Camera Networks." In *Proceedings of the British Machine Vision Conference*, pp. 97.1–97.11. BMVA Press, 2012.

[Imre et al. 12] Evren Imre, Jean-Yves Guillemaut, and Adrian Hilton. "Through-the-Lens Multi-camera Synchronisation and Frame-Drop Detection for 3D Reconstruction." In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on*, pp. 395–402, 2012.

[Ince and Konrad 08] Serdar Ince and Janusz Konrad. "Occlusion-Aware View Interpolation." *EURASIP Journal on Image and Video Processing* (2008), 1–15.

[Irani and Peleg 91] Michael Irani and Shmuel Peleg. "Improving Resolution by Image Registration." *CVGIP: Graph. Models Image Process.* 53:3 (1991), 231–239.

[Irawan and Marschner 12] Piti Irawan and Steve Marschner. "Specular Reflection from Woven Cloth." *ACM Transactions on Graphics (TOG)* 31:1 (2012), 11.

[Irawan 08] Piti Irawan. "Appearance of Woven Cloth." Ph.D. thesis, Cornell University, 2008.

[Ismael et al. 14] Muhannad Ismael, Stéphanie Prévost, Céline Loscos, and Yannick Rémion. "Materiality Maps: A Novel Scene-Based Framework for Direct Multi-View Stereovision Reconstruction." In *IEEE International Conference on Image Processing*, p. to appear. IEEE, 2014.

[Ives 03] F. Ives. "Parallax Stereogram and Process of Making Same." 1903.

[Izadi et al. 11] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon. "KinectFusion: Real-Time 3D Reconstruction and Interaction Using a Moving Depth Camera." In *Proc. ACM Symp. User Interface Softw. & Tech.*, pp. 559–568, 2011.

[Jacobson et al. 12] A. Jacobson, I. Baran, L. Kavan, J. Popović, and O. Sorkine. "Fast Automatic Skinning Transformations." *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 31:4.

[Jain et al. 10] A. Jain, T. Thormählen, H.-P. Seidel, and C. Theobalt. "MovieReshape: Tracking and Reshaping of Humans in Videos." *ACM Transactions on Graphics* 29 (2010), 148:1–10. Proceedings of SIGGRAPH Asia.

[Jakob and Marschner 12] Wenzel Jakob and Steve Marschner. "Manifold Exploration: A Markov Chain Monte Carlo Technique for Rendering Scenes with Difficult Specular Transport." *ACM Transactions on Graphics (Proc. of SIGGRAPH)* 31:4 (2012), 58:1–58:13.

[James and Twigg 05] D. James and C. Twigg. "Skinning Mesh Animations." *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 24:3, 2005.

[Jarabo et al. 14] Adrian Jarabo, Hongzhi Wu, Julie Dorsey, Holly Rushmeier, and Diego Gutierrez. "Effects of Approximate Filtering on the Appearance of Bidirectional Texture Functions." *IEEE Transactions on Visualization and Computer Graphics.*

[Jensen et al. 01] Henrik Wann Jensen, Stephen R. Marschner, Marc Levoy, and Pat Hanrahan. "A Practical Model for Subsurface Light Transport." *ACM Transactions on Graphics (Proc. SIGGRAPH 2001).*

[Jensen 96] Henrik Wann Jensen. "Global Illumination Using Photon Maps." In *Proc. Eurographics Workshop on Rendering*, pp. 21–30, 1996.

[Jian and Vemuri 11] Bing Jian and Baba C. Vemuri. "Robust Point Set Registration Using Gaussian Mixture Models." *IEEE Transactions on PAMI* 33:8 (2011), 1633–1645.

[Jimenez et al. 12] Jorge Jimenez, Adrian Jarabo, and Diego Gutierrez. "Separable Subsurface Scattering and Photorealistic Eyes Rendering." In *SIGGRAPH Course Notes 2012*, 2012.

[Johannsen et al. 13] O. Johannsen, C. Heinze, B. Goldlücke, and C. Perwass. "On the Calibration of Focused Plenoptic Cameras." In *GCPR Workshop on Imaging New Modalities*, 2013.

[Johnson and Hebert 97] A. Johnson and M. Hebert. "Recognizing Objects by Matching Oriented Points." In *Conference on Computer Vision and Pattern Recognition*, pp. 684–692, 1997.

[Johnson 02] T. Johnson. "Methods for Characterizing Colour Scanners and Digital Cameras." In *Colour Engineering: Achieving Device Independent Colour*, pp. 165–178. John Wiley & Sons Inc., 2002.

[Jones et al. 01] Graham Jones, Delman Lee, Nicolas Holliman, and David Ezra. "Controlling Perceived Depth in Stereoscopic Images." 4297, pp. 42–53, 2001.

[Ju and Kang 09] Myung-Ho Ju and Hang-Bong Kang. "Constant Time Stereo Matching." *International Machine Vision and Image Processing Conference* 0 (2009), 13–17.

[Ju et al. 02] Tao Ju, Frank Losasso, Scott Schaefer, and Joe Warren. "Dual Contouring of Hermite Data." *ACM Transactions on Graphics* 21:3 (2002), 339–346.

[Jung et al. 12] Yvonne Jung, Johannes Behr, Timm Drevensek, and Sebastian Wagner. "Declarative 3D Approaches for Distributed Web-Based Scientific Visualization Services." In *Dec3D*, 869, edited by Johannes Behr, Donald P. Brutzman, Ivan Herman, Jacek Jankowski, and Kristian Sons, 869, 2012.

[Jung et al. 13] Yvonne Jung, Max Limper, Pasquale Herzig, Karsten Schwenk, and Johannes Behr. "Fast and Efficient Vertex Data Representations for the Web." In *IVAPP*, pp. 77–86, 2013.

[Kahlmann et al. 07] T. Kahlmann, F. Remondino, and S. Guillaume. "Range Imaging Technology: New Developments and Applications for People Identification and Tracking." In *Proc. of Videometrics IX - SPIE-IS&T Electronic Imaging*, 6491, 6491, 2007. DOI: 10.1117/12.702512.

[Kajiya 86] James T. Kajiya. "The Rendering Equation." *Computer Graphics (Proc. of SIGGRAPH)*, pp. 143–150.

[Kaldor et al. 08] Jonathan M. Kaldor, Doug L. James, and Steve Marschner. "Simulating Knitted Cloth at the Yarn Level." *ACM Trans. Graph.* 27:3 (2008), 65.

[Kaldor et al. 10] Jonathan M. Kaldor, Doug L. James, and Steve Marschner. "Efficient Yarn-Based Cloth with Adaptive Contact Linearization." *ACM Trans. Graph.* 29:4 (2010), 105.

[Kán 12] Hannes Kán, Peter and Kaufmann. "Physically-Based Depth of Field in Augmented Reality." In *Proceedings of EUROGRAPHICS 2012*, pp. 89–92, 2012.

[Kanatani et al. 08] Kenichi Kanatani, Yasuyuki Sugaya, and Hirotaka Niitsuma. "Triangulation from Two Views Revisited: Hartley-Sturm vs. Optimal Correction." *In Practice* 4 (2008), 5.

[Kang and Szeliski 04] Sing Bing Kang and Richard Szeliski. "Extracting View-Dependent Depth Maps from a Collection of Images." *International Journal of Computer Vision* 58 (2004), 139–163.

[Kang et al. 04] Sing Bing Kang, Richard Szeliski, and Matthew Uyttendaele. "Seamless Stitching Using Multi-Perspective Plane Sweep." Technical Report MSR- TR-2004-48, June 2004.

[Kaplanyan and Dachsbacher 13a] Anton S. Kaplanyan and Carsten Dachsbacher. "Adaptive Progressive Photon Mapping." *ACM Transactions on Graphics* 32:2, 2013.

[Kaplanyan and Dachsbacher 13b] Anton S. Kaplanyan and Carsten Dachsbacher. "Path Space Regularization for Holistic and Robust Light Transport." *Computer Graphics Forum (Proc. of Eurographics)* 32:2, 2013.

[Kaplanyan et al. 14] Anton S. Kaplanyan, Johannes Hanika, and Carsten Dachsbacher. "The Natural-Constraint Representation of the Path Space for Efficient Light Transport Simulation." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 33:4, 2014.

[Karsch et al. 11] Kevin Karsch, Varsha Hedau, David Forsyth, and Derek Hoiem. "Rendering Synthetic Objects into Legacy Photographs." *ACM Trans. Graph.* 30:6 (2011), 157:1–157:12.

[Kastrinaki et al. 03] V. Kastrinaki, Michael E. Zervakis, and Kostas Kalaitzakis. "A Survey of Video Processing Techniques for Traffic Applications." *Image Vision Comput.* 21:4 (2003), 359–381.

[Katz and Tal 03] S. Katz and A. Tal. "Hierarchical Mesh Decomposition Using Fuzzy Clustering and Cuts." *ACM Transactions on Graphics* 22:3 (2003), 954–961.

[Kavan et al. 10] L. Kavan, P-P. Sloan, and C. O'Sullivan. "Fast and Efficient Skinning of Animated Meshes." *Computer Graphics Forum* 29:2 (2010), 327–336.

[Kavan et al. 11] Ladislav Kavan, Dan Gerszewski, Adam W. Bargteil, and Peter-Pike Sloan. "Physics-Inspired Upsampling for Cloth Simulation in Games." *ACM Trans. Graph.* 30:4 (2011), 93:1–93:10.

[Kazhdan et al. 06] Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe. "Poisson Surface Reconstruction." In *Proceedings of the Fourth Eurographics Symposium on Geometry Processing, SGP '06*, pp. 61–70. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2006.

[Keller et al. 13] M. Keller, D. Lefloch, M. Lambers, S. Izadi, T. Weyrich, and A. Kolb. "Real-Time 3D Reconstruction in Dynamic Scenes Using Point-Based Fusion." In *Proc. Conf. 3D Vision (3DV)*, 2013. DOI:10.1109/3DV.2013.9.

[Khan et al. 06] E. A. Khan, A. O. Akyz, and E. Reinhard. "Ghost Removal in High Dynamic Range Images." In *IEEE International Conference on Image Processing*, pp. 2005–2008, 2006.

[Kholgade et al. 14] Natasha Kholgade, Tomas Simon, Alexei Efros, and Yaser Sheikh. "3D Object Manipulation in a Single Photograph Using Stock 3D Models." *ACM Transactions on Computer Graphics* 33:4, 2014.

[Khoshelham and Elberink 12] Kourosh Khoshelham and Sander Oude Elberink. "Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications." *Sensors* 12:2 (2012), 1437–1454.

[Kilner et al. 06] J.J. Kilner, J.R. Starck, and A. Hilton. "A Comparative Study of Free Viewpoint Video Techniques for Sports Events." In *IET European Conference on Visual Media Production*, pp. 87–96. IET, 2006.

[Kilner et al. 09] J. Kilner, J.-Y. Guillemaut, and A. Hilton. "3D Action Matching with Key-Pose Detection." In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pp. 1–8. IEEE, 2009.

[Kim and Vendrovsky 08] Tae-Yong Kim and Eugene Vendrovsky. "DrivenShape: A Data-driven Approach for Shape Deformation." In *Proc. of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, SCA '08*, pp. 49–55. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2008.

[Kim et al. 07] Seon Joo Kim, Jan-Michael Frahm, and Marc Pollefeys. "Joint Feature Tracking and Radiometric Calibration from Auto-Exposure

Video." In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pp. 1–8. IEEE, 2007.

[Kim et al. 09] Y. M. Kim, Christian Theobalt, J. Diebel, J. Kosecka, B. Micusik, and S. Thrun. "Multi-View Image and ToF Sensor Fusion for Dense 3D Reconstruction." In *IEEE Workshop on 3-D Digital Imaging and Modeling (3DIM)*, edited by Adrian Hilton, Takeshi Masuda, and Chang Shu, pp. 1542–1549. Kyoto, Japan: IEEE, 2009.

[Kim et al. 10] Youngmin Kim, Amitabh Varshney, David W. Jacobs, and François Guimbretière. "Mesh Saliency and Human Eye Fixations." *ACM Trans. Appl. Percept.* 7:2 (2010), 12:1–12:13. Available online (http://doi.acm.org/10.1145/1670671.1670676).

[Kim et al. 12] Min H. Kim, Todd Alan Harvey, David S. Kittle, Holly Rushmeier, Julie Dorsey, Richard O. Prum, and David J. Brady. "3D Imaging Spectroscopy for Measuring Hyperspectral Patterns on Solid Objects." *ACM Trans. Graph.* 31:4 (2012), 38:1–38:11. Available online (http://doi.acm.org/10.1145/2185520.2185534).

[Kim et al. 13] C. Kim, H. Zimmer, Y. Pritch, A. Sorkine-Hornung, and M. Gross. "Scene Reconstruction from High Spatio-Angular Resolution Light Fields." In *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 32, 32, pp. 73:1–73:12, 2013.

[Kircher and Garland 06] S. Kircher and M. Garland. "Editing Arbitrarily Deforming Surface Animations." *ACM Transactions on Graphics (Proceedings of SIGGRAPH)*, pp. 1098–1107, 2006.

[Klaudiny et al. 12] Martin Klaudiny, Chris Budd, and Adrian Hilton. "Towards Optimal Non-Rigid Surface Tracking." In *ECCV*, pp. 743–756, 2012.

[Klose et al. 11] Felix Klose, Kai Ruhl, Christian Lipski, and Marcus Magnor. "Flowlab—An Interactive Tool for Editing Dense Image Correspondences." In *European Conference on Visual Media Production (CVMP)*, pp. 59–66, 2011.

[Kobbelt et al. 01] Leif P. Kobbelt, Mario Botsch, Ulrich Schwanecke, and Hans-Peter Seidel. "Feature Sensitive Surface Extraction from Volume Data." In *Proceedings of ACM SIGGRAPH 2001*, pp. 57–66, 2001.

[Kolb et al. 10] A. Kolb, E. Barth, R. Koch, and R. Larsen. "Time-of-Flight Cameras in Computer Graphics." *J. Computer Graphics Forum* 29:1 (2010), 141–159.

[Kolmogorov and Zabih 01] Vladimir Kolmogorov and Ramin Zabih. "Computing Visual Correspondence with Occlusions Using Graph Cuts." In *International Conference on Computer Vision*, 2, pp. 508–515 vol. 2, 2001.

[Kolmogorov and Zabih 02] Vladimir Kolmogorov and Ramin Zabih. "Multi-Camera Scene Reconstruction via Graph Cuts." In *Proceedings of the 7th European Conference on Computer Vision-Part III*, pp. 82–96, 2002.

[Kovar et al. 02] L. Kovar, M. Gleicher, and F. Pighin. "Motion Graphs." In *Proc. ACM SIGGRAPH*, pp. 473–482, 2002.

[Kriegel et al. 09] Hans-Peter Kriegel, Peer Kröger, Erich Schubert, and Arthur Zimek. "LoOP: Local Outlier Probabilities." In *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, pp. 1649–1652, 2009.

[Krishnan and Nayar 08] Gurunandan Krishnan and Shree Nayar. "Cata-Fisheye Camera for Panoramic Imaging." In *Proc. IEEE Workshop on Application of Computer Vision (WACV)*, 2008.

[Křivánek et al. 13] Jaroslav Křivánek, Iliyan Georgiev, Anton Kaplanyan, and Juan Canada. "Recent Advances in Light Transport Simulation: Theory and Practice." In *ACM SIGGRAPH Courses*, 2013.

[Kry et al. 02] P. G. Kry, D. L. James, and D. K. Pai. "EigenSkin: Real-Time Large Deformation Character Skinning in Hardware." In *Proceedings of Symposium on Computer Animation (SCA)*, pp. 153–160, 2002.

[Kuhnert and Stommel 06] Klaus-Dieter Kuhnert and Martin Stommel. "Fusion of Stereo-Camera and PMD-Camera Data for Real-Time Suited Precise 3D Environment Reconstruction." In *IROS*, pp. 4780–4785, 2006.

[Kunitomo et al. 10] Shoji Kunitomo, Shinsuke Nakamura, and Shigeo Morishima. "Optimization of Cloth Simulation Parameters by Considering Static and Dynamic Features." In *ACM SIGGRAPH 2010 Posters, SIGGRAPH '10*, pp. 15:1–15:1. New York, NY, USA: ACM, 2010.

[Kuster et al. 11] C. Kuster, T. Popa, C. Zach, C. Gotsman, and M. Gross. "FreeCam: A Hybrid Camera System for Interactive Free-Viewpoint Video." In *Proceedings of Vision, Modeling, and Visualization (VMV)*, pp. 17–24, 2011.

[Kutulakos and Seitz 00] Kiriakos N. Kutulakos and Steven M. Seitz. "A Theory of Shape by Space Carving." *Int. J. Comput. Vision* 38:3 (2000), 199–218.

[Kwatra et al. 03] Vivek Kwatra, Arno Schödl, Irfan Essa, Greg Turk, and Aaron Bobick. "Graphcut Textures: Image and Video Synthesis Using Graph Cuts." In *ACM Transactions on Graphics (ToG)*, 22, 22, pp. 277–286. ACM, 2003.

[Lafortune and Willems 93] Eric P. Lafortune and Yves D. Willems. "Bi-Directional Path Tracing." In *Compugraphics '93*, pp. 145–153, 1993.

[Lalonde et al. 09] Jean-Francois Lalonde, Alexei A. Efros, and Srinivasa G. Narasimhan. "Estimating Natural Illumination from a Single Outdoor Image." In *IEEE International Conference on Computer Vision*, 2009.

[Lamond et al. 09] Bruce Lamond, Pieter Peers, Abhijeet Ghosh, and Paul Debevec. "Image-based Separation of Diffuse and Specular Reflections Using Environmental Structured Illumination." In *IEEE International Conference on Computational Photography*, 2009.

[Lang et al. 10] Manuel Lang, Alexander Hornung, Oliver Wang, Steven Poulakos, Aljoscha Smolic, and Markus Gross. "Nonlinear Disparity Mapping for Stereoscopic 3D." *ACM Transactions on Graphics (TOG)* 29:4 (2010), 75.

[Lanman et al. 08] D. Lanman, R. Raskar, A. Agrawal, and G. Taubin. "Shield Fields: Modeling and Capturing 3D Occluders." *Proc. SIGGRAPH Asia* 27:5 (2008), 1–10.

[Lantz 07] Ed Lantz. "A Survey of Large-scale Immersive Displays." In *Proceedings of the 2007 Workshop on Emerging Displays Technologies: Images and Beyond: The Future of Displays and Interacton*, p. 1. ACM, 2007.

[Laurentini 94] Aldo Laurentini. "The Visual Hull Concept for Silhouette-Based Image Understanding." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16:2 (1994), 150–162.

[Lavoué et al. 13] Guillaume Lavoué, Laurent Chevalier, and Florent Dupont. "Streaming Compressed 3D Data on the Web Using JavaScript and WebGL." In *Proceedings of the 18th International Conference on 3D Web Technology, Web3D '13*, pp. 19–27, 2013.

[Lee et al. 98] S. Lee, G. Wolberg, and S. Shin. "Polymorph: Morphing among Multiple Images." *IEEE Computer Graphics and Applications*, pp. 58–71.

[Lee et al. 02] J. Lee, J. Chai, P.S.A. Reitsma, J.K. Hodgins, and N.S. Pollard. "Interactive Control of Avatars Animated with Human Motion Data." In *Proc. ACM SIGGRAPH 2002*, pp. 491–500, 2002.

[Lee et al. 09] S.-H. Lee, E. Sifakis, and D. Terzopoulos. "Comprehensive Biomechanical Modeling and Simulation of the Upper Body." *ACM Transactions on Graphics* 28:4 (2009), 99:1–99:17.

[Lee et al. 10a] Jongseok Lee, Sungyul Choe, and Seungyong Lee. "Mesh Geometry Compression for Mobile Graphics." In *Proc. CCNC*, pp. 301–305, 2010.

[Lee et al. 10b] Sungkil Lee, Elmar Eisemann, and Hans-Peter Seidel. "Real-Time Lens Blur Effects and Focus Control." *ACM Transaction on Graphics (Proc. SIGGRAPH Asia)* 29:4.

[Lefloch et al. 13] D. Lefloch, T. Hoegg, and A. Kolb. "Real-Time Motion Artifacts Compensation of ToF Sensors Data on GPU." In *Proc. SPIE Defense, Security—Three-Dimensional Imaging, Visualization, and Display*, pp. 87380U–87380U–7, 2013.

[Lempitsky and Ivanov 07] Victor Lempitsky and Denis Ivanov. "Seamless Mosaicing of Image-Based Texture Maps." In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pp. 1–6. IEEE, 2007.

[Lengyel 98] J. Lengyel. "The Convergence of Graphics and Vision." *Computer* 31:7 (1998), 46–53.

[Lensch et al. 03] Hendrik P. A. Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel. "Image-Based Reconstruction of Spatial Appearance and Geometric Detail." *ACM Trans. Graph.* 22:2 (2003), 234–257.

[Leutenegger et al. 11] Stefan Leutenegger, Margarita Chli, and Roland Yves Siegwart. "BRISK: Binary Robust Invariant Scalable Keypoints." In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 2548–2555. IEEE, 2011.

[Levin 98] David Levin. "The Approximation Power of Moving Least-Squares." *Mathematics of Computation* 67 (1998), 1517–1531.

[Levin 03] David Levin. "Mesh-Independent Surface Interpolation." *Geometric Modeling for Scientific Visualization* 3 (2003), 37–49.

[Levoy and Hanrahan 96] Marc Levoy and Pat Hanrahan. "Light Field Rendering." In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, pp. 31–42. New York: ACM, 1996.

[Levoy et al. 00] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, and J. Ginsberg. "The Digital Michelangelo Project: 3D Scanning of Large Statues." In *Proc. Conf. Computer Graphics and Interactive Techniques (SIGGRAPH)*, pp. 131–144, 2000.

[Levoy 06] M. Levoy. "Light Fields and Computational Imaging." *Computer* 39:8 (2006), 46–55.

[Lewis and Anjyo 10] J. P. Lewis and K. Anjyo. "Direct Manipulation Blendshapes." In *Proceeedings of IEEE CGAA*, 2010.

[Lewis et al. 00] J. P. Lewis, Matt Cordner, and Nickson Fong. "Pose Space Deformation: A Unified Approach to Shape Interpolation and Skeleton-Driven Deformation." In *ACM SIGGRAPH*, pp. 165–172, 2000.

[Lhuillier and Quan 05] Maxime Lhuillier and Long Quan. "A Quasi-Dense Approach to Surface Reconstruction from Uncalibrated Images." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 27:3 (2005), 418–433.

[Li et al. 03] Ming Li, Marcus A. Magnor, and Hans-Peter Seidel. "Hardware-Accelerated Visual Hull Reconstruction and Rendering." In *Graphics Interface (GI)*, pp. 65–71, 2003.

[Li et al. 08] Xin Li, Bahadir Gunturk, and Lei Zhang. "Image Demosaicing: A Systematic Survey." In *Electronic Imaging 2008*, pp. 68221J–68221J. International Society for Optics and Photonics, 2008.

[Li et al. 10] Hao Li, Thibaut Weise, and Mark Pauly. "Example-Based Facial Rigging." *ACM Transactions on Graphics (Proc. SIGGRAPH 2010)* 29:4.

[Li et al. 12a] Hao Li, Linjie Luo, Daniel Vlasic, Pieter Peers, Jovan Popović, Mark Pauly, and Szymon Rusinkiewicz. "Temporally Coherent Completion of Dynamic Shapes." *ACM Trans. Graph.* 31:1 (2012), 2:1–2:11.

[Li et al. 12b] K. Li, Q. Dai, W. Xu, and J. Yang. "Temporal-Dense Dynamic 3-D Reconstruction with Low Frame Rate Cameras." *Selected Topics in Signal Processing, IEEE Journal of* 6:5 (2012), 447–459.

[Li et al. 13] Guannan Li, Chenglei Wu, Carsten Stoll, Yebin Liu, Kiran Varanasi, Qionghai Dai, and Christian Theobalt. "Capturing Relightable Human Performances under General Uncontrolled Illumination." In *Computer Graphics Forum*, 32, 32, pp. 275–284, 2013.

[Liang et al. 08] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, and H. Chen. "Programmable Aperture Photography: Multiplexed Light Field Acquisition." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 27:3 (2008), 1–10.

[Lien et al. 06] J. M. Lien, J. Keyser, and N. M. Amato. "Simultaneous Shape Decomposition and Skeletonization." In *Proceedings of SPM*, pp. 219–228, 2006.

[Lim et al. 05] SukHwan Lim, Jon G. Apostolopoulos, and Abbas El Gamal. "Optical Flow Estimation Using Temporally Oversampled Video." *IEEE Transactions on Image Processing* 14:8 (2005), 1074–1087.

[Limper et al. 13a] Max Limper, Yvonne Jung, Johannes Behr, and Marc Alexa. "The POP Buffer: Rapid Progressive Clustering by Geometry Quantization." *Computer Graphics Forum* 32:7 (2013), 197–206.

[Limper et al. 13b] Max Limper, Stefan Wagner, Christian Stein, Yvonne Jung, and André Stork. "Fast Delivery of 3D Web Content: A Case Study." In *Proceedings of the 18th International Conference on 3D Web Technology, Web3D '13*, pp. 11–17, 2013.

[Lin and Tomasi 04] Micahel H. Lin and Carlo Tomasi. "Surfaces with Occlusions from Layered Stereo." *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (2004), 1073–1078.

[Lindemann et al. 11] Lea Lindemann, Stephan Wenger, and Marcus Magnor. "Evaluation of Video Artifact Perception Using Event-Related Potentials." In *ACM Symposium on Applied Perception in Graphics and Visualization (APGV)*, pp. 53–58, 2011.

[Lindner and Kolb 06] M. Lindner and A. Kolb. "Lateral and Depth Calibration of PMD-Distance Sensors." In *Proc. Int. Symp. on Visual Computing, LNCS*, pp. 524–533. Springer, 2006.

[Lindner and Kolb 07] M. Lindner and A. Kolb. "Calibration of the Intensity-Related Distance Error of the PMD ToF-Camera." In *Proc. SPIE, Intelligent Robots and Computer Vision*, 6764, pp. 6764–35, 2007.

[Lindner and Kolb 09] M. Lindner and A. Kolb. "Compensation of Motion Artifacts for Time-of-Flight Cameras." In *Proc. Dynamic 3D Imaging*, LNCS, 5742, pp. 16–27. Springer, 2009.

[Lindner et al. 08] M. Lindner, M. Lambers, and A. Kolb. "Sub-Pixel Data Fusion and Edge-Enhanced Distance Refinement for 2D/3D Images." *Int. J. on Intell. Systems and Techn. and App., Special Issue on Dynamic 3D Imaging* 5:3/4 (2008), 344–354.

[Lindner et al. 10] M. Lindner, I. Schiller, A. Kolb, and R. Koch. "Time-of-Flight Sensor Calibration for Accurate Range Sensing." *Computer Vision and Image Understanding* 114:12 (2010), 1318–1328.

[Lindstrom 10] Peter Lindstrom. "Triangulation Made Easy." In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pp. 1554–1561. IEEE, 2010.

[Linz et al. 10a] Christian Linz, Christian Lipski, and Marcus Magnor. "Multi-Image Interpolation Based on Graph-Cuts and Symmetric Optic Flow." In *Vision, Modeling, and Visualization (VMV)*, pp. 115–122, 2010.

[Linz et al. 10b] Christian Linz, Christian Lipski, Lorenz Rogge, Christian Theobalt, and Marcus Magnor. "Space-Time Visual Effects as a Post-Production Process." In *ACM International Workshop on 3D Video Processing (3DVP)*, pp. 1–6, 2010.

[Linz 11] Christian Linz. "Correspondence Estimation and Image Interpolation for Photo-Realistic Rendering." Ph.D. thesis, Braunschweig, 2011.

[Lippmann 08] G. Lippmann. "Épreuves réversibles donnant la sensation du relief." *J. Phys. Theor. Appl.* 7:1 (1908), 821–825.

[Lipski et al. 10a] Christian Lipski, Christian Linz, Kai Berger, Anita Sellent, and Marcus Magnor. "Virtual Video Camera: Image-Based Viewpoint Navigation through Space and Time." *Computer Graphics Forum* 29:8 (2010), 2555–2568.

[Lipski et al. 10b] Christian Lipski, Christian Linz, Thomas Neumann, Markus Wacker, and Marcus Magnor. "High Resolution Image Correspondences for Video Post-Production." In *European Conference on Visual Media Production (CVMP)*, pp. 33–39, 2010.

[Lipski et al. 12] Christian Lipski, Christian Linz, Thomas Neumann, Markus Wacker, and Marcus Magnor. "High Resolution Image Correspondences for Video Post-Production." *Journal of Virtual Reality and Broadcasting (JVRB)* 9:8 (2012), 1–12.

[Lipski et al. 14] Christian Lipski, Felix Klose, and Marcus Magnor. "Correspondence and Depth-Image Based Rendering: A Hybrid Approach for Free-Viewpoint Video." *IEEE Trans. Circuits and Systems for Video Technology (CSVT)* 24:6 (2014), 942–951.

[Lipski 12] Christian Lipski. "Making of 3D-Gauss, Vision, Modeling, and Visualization Workshop 2009." http://vmv09.tu-bs.de/index.php?page=gauss, 2012.

[Liu et al. 03] P. C. Liu, F. C. Wu, W. C. Ma, R. H Liang, and M. Ouhyoung. "Automatic Animation Skeleton Using Repulsive Force Field." In *Proceedings of the 11th Pacific Conference on Computer Graphics and Applications*, pp. 409–413, 2003.

[Liu et al. 08] Ce Liu, Jenny Yuen, Antonio Torralba, Josef Sivic, and William T. Freeman. "SIFT Flow: Dense Correspondence Across Different Scenes." In *European Conference on Computer Vision*, pp. 28–42, 2008.

[Liu et al. 10a] Kai Liu, Yongchang Wang, Daniel L. Lau, Qi Hao, and Laurence G. Hassebrook. "Dual-Frequency Pattern Scheme for High-Speed 3-D Shape Measurement." *Optics Express* 18:5 (2010), 5229–5244.

[Liu et al. 10b] Yebin Liu, Qionghai Dai, and Wenli Xu. "A Point-Cloud-Based Multiview Stereo Algorithm for Free-Viewpoint Video." *IEEE Transactions on Visualization and Computer Graphics* 16:3 (2010), 407–418.

[Liu et al. 11] Yebin Liu, Carsten Stoll, Juergen Gall, Hans-Peter Seidel, and Christian Theobalt. "Markerless Motion Capture of Interacting Characters Using Multi-View Image Segmentation." In *CVPR*, pp. 1249–1256, 2011.

[Liu et al. 13] Yebin Liu, Juergen Gall, Carsten Stoll, Qionghai Dai, Hans-Peter Seidel, and Christian Theobalt. "Markerless Motion Capture of Multiple Characters Using Multiview Image Segmentation." *IEEE Transactions on Pattern Analysis Machine Intelligence (PAMI)* 35:11 (2013), 2720–2735.

[Lokovic and Veach 00] Tom Lokovic and Eric Veach. "Deep Shadow Maps." *ACM Transactions on Graphics (Proc. SIGGRAPH 2000)*, pp. 385–392.

[Longere et al. 02] P. Longere, Xuemei Zhang, P.B. Delahunt, and D.H. Brainard. "Perceptual Assessment of Demosaicing Algorithm Performance." *Proceedings of the IEEE* 90:1 (2002), 123–132.

[Lorensen and Cline 87] William E. Lorensen and Harvey E. Cline. "Marching Cubes: A High Resolution 3D Surface Construction Algorithm." In *Proceedings of ACM SIGGRAPH 1987*, pp. 163–169, 1987.

[Loscos and Jacobs 10] C. Loscos and K. Jacobs. "High-Dynamic Range Imaging for Dynamic Scenes." In *Computational Photography, Methods and Applications*, p. 259–281. CRC Press, 2010.

[Loscos et al. 99] Cline Loscos, Marie-Claude Frasson, George Drettakis, Bruce Walter, Xavier Granier, and Pierre Poulin. "Interactive Virtual Relighting and Remodeling of Real Scenes." In *Rendering Techniques*, pp. 329–340, 1999.

[Lowe 04] David G. Lowe. "Distinctive Image Features from Scale-Invariant Keypoints." *International Journal of Computer Vision* 60:2 (2004), 91–110.

[Lucas and Kanade 81] Bruce D. Lucas and Takeo Kanade. "An Iterative Image Registration Technique with an Application to Stereo Vision." In *International Joint Conference on Artificial Intelligence*, 2, pp. 674–679, 1981.

[Lucas et al. 13] Laurent Lucas, Philippe Souchet, Muhannad Ismael, Olivier Nocent, Cédric Niquin, Céline Loscos, Ludovic Blache, Stéphanie Prévost, and Yannick Remion. "RECOVER3D: A Hybrid Multi-View System for 4D Reconstruction of Moving Actors." In *4th International Conference and Exhibition on 3D Body Scanning Technologies*, pp. 219–230. Hometrica Consulting, 2013.

[Luenberger 73] D. G. Luenberger. "Introduction to Linear and Non-Linear Programming." Reading MA: Addison-Wesley, 1973.

[Lumsdaine and Georgiev 09] Andrew Lumsdaine and Todor Georgiev. "The Focused Plenoptic Camera." In *Proc. IEEE International Conference on Computational Photography*, pp. 1–8, 2009.

[Lytro, Inc. 12] Lytro, Inc. "The Lytro Camera." https://www.lytro.com/, 2012.

[Ma et al. 07] Wan-Chun Ma, Tim Hawkins, Pieter Peers, Charles-Felix Chabert, Malte Weiss, and Paul Debevec. "Rapid Acquisition of Specular and Diffuse Normal Maps from Polarized Spherical Gradient Illumination." In *Eurographics Symposium on Rendering*, pp. 183–194, 2007.

[Magnenat-Thalmann and Volino 05] Nadia Magnenat-Thalmann and Pascal Volino. "From Early Draping to Haute Couture Models: 20 Years of Research." *The Visual Computer* 21:8-10 (2005), 506–519.

[Magnenat-Thalmann et al. 88] N. Magnenat-Thalmann, R. Laperrière, and D. Thalmann. "Joint-Dependent Local Deformations for Hand Animation and Object Grasping." In *Proceedings on Graphics Interface '88*, pp. 26–33. Canadian Information Processing Society, 1988.

[Magnenat-Thalmann et al. 04] Nadia Magnenat-Thalmann, Frederic Cordier, Michael Keckeisen, Stefan Kimmerle, Reinhard Klein, and Jan Meseth. "Simulation of Clothes for Real-Time Applications." In *Eurographics 2004, Tutorials 1: Simulation of Clothes for Real-time Applications*. Eurographics Association, 2004.

[Magnor 05] M. Magnor. *Video-Based Rendering.* A. K. Peters, 2005.

[Malvar et al. 04] Henrique S. Malvar, Li-wei He, and Ross Cutler. "High-Quality Linear Interpolation for Demosaicing of Bayer-Patterned Color Images." In *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 3, pp. 485–488, 2004.

[Mamou et al. 09] Khaled Mamou, Titus Zaharia, and Françoise Prêteux. "TFAN: A Low Complexity 3D Mesh Compression Algorithm." *Comput. Animat. Virtual Worlds* 20:3 (2009), 343–354.

[Manakov et al. 13] Alkhazur Manakov, John F. Restrepo, Oliver Klehm, Ramon Hegedüs, Elmar Eisemann, Hans-Peter Seidel, and Ivo Ihrke. "A Reconfigurable Camera Add-on for High Dynamic Range, Multi-Spectral, Polarization, and Light-Field Imaging." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 32:4 (2013), 47:1–47:14.

[Mann and Picard 95] Steve Mann and Rosalind W. Picard. "Being Undigital with Digital Cameras: Extending Dynamic Range by Combining Differently Exposed Pictures." In *IST 46th Annual Conference*, pp. 422–428, 1995.

[Marr and Poggio 76] D. Marr and T. Poggio. "Cooperative Computation of Stereo Disparity." *Science* 194 (1976), 283–287.

[Marschner et al. 03] Stephen Marschner, Henrik Wann Jensen, Mike Cammarano, Steve Worley, and Pat Hanrahan. "Light Scattering from Human Hair Fibers." *ACM Transactions on Graphics (Proc. SIGGRAPH 2003).*

[Martin and Crowley 95] Jerome Martin and James L. Crowley. "Comparison of Correlation Techniques." In *Intelligent Autonomous Systems*, pp. 86–93, 1995.

[Martínez et al. 14] Leonardo Sánchez-Mesa Martínez, Francisco Javier Melero Rus, and Jorge Revelles Moreno. "Portal Virtual de Patrimonio de las Universidades Andaluzas." www.http://patrimonio3d.ugr.es, February 28, 2014.

[Marwah et al. 13] K. Marwah, G. Wetzstein, Y. Bando, and R. Raskar. "Compressive Light Field Photography Using Overcomplete Dictionaries and Optimized Projections." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 32:4 (2013), 46:1–46:11.

[Masselus et al. 02] Vincent Masselus, Philip Dutré, and Frederik Anrys. "The Free-Form Light Stage." In *Proceedings of the 13th Eurographics Workshop on Rendering, EGRW '02*, pp. 247–256. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2002.

[Masselus et al. 03] Vincent Masselus, Pieter Peers, Philip Dutré, and Yves D. Willems. "Relighting with 4D Incident Light Fields." *ACM Transactions on Graphics (Proc. SIGGRAPH 2003)* 22:3 (2003), 613–620.

[Matas et al. 04] Jiri Matas, Ondrej Chum, Martin Urban, and Tomás Pajdla. "Robust Wide-Baseline Stereo from Maximally Stable Extremal Regions." *Image and Vision Computing* 22:10 (2004), 761–767.

[Matsushita et al. 06] Yasuyuki Matsushita, Eyal Ofek, Weina Ge, Xiaoou Tang, and Heung-Yeung Shum. "Full-Frame Video Stabilization with Motion Inpainting." *IEEE Trans. Pattern Anal. Mach. Intell.* 28:7 (2006), 1150–1163.

[Matsuyama et al. 04] T. Matsuyama, X. Wu, T. Takai, and S. Nobuhara. "Real-Time 3D Shape Reconstruction, Dynamic 3D Mesh Deformation, and High Fidelity Visualization for 3D Video." *Computer Vision and Image Understanding* 96:3 (2004), 393–434.

[Matusik et al. 00] Wojciech Matusik, Chris Buehler, Ramesh Raskar, Steven J. Gortler, and Leonard McMillan. "Image-Based Visual Hulls." In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '00*, pp. 369–374. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 2000.

[Matusik et al. 02] Wojciech Matusik, Hanspeter Pfister, Addy Ngan, Paul Beardsley, Remo Ziegler, and Leonard McMillan. "Image-Based 3D Photography Using Opacity Hulls." *ACM Trans. Graph.* 21:3 (2002), 427–437.

[Matusik et al. 04] Wojciech Matusik, Matthew Loper, and Hanspeter Pfister. "Progressively-Refined Reflectance Functions from Natural Illumination." In *Proceedings of the Fifteenth Eurographics Conference on Rendering Techniques, EGSR'04*, pp. 299–308. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2004. Available online (http://dx.doi.org/10.2312/EGWR/EGSR04/299-308).

[McDonnell et al. 08] Rachel McDonnell, Michéal Larkin, Simon Dobbyn, Steven Collins, and Carol O'Sullivan. "Clone Attack! Perception of Crowd Variety." In *ACM Transactions on Graphics (TOG)*, 27, p. 26. ACM, 2008.

[McGlone 13] J. Chris McGlone, editor. *Manual of Photogrammetry*. Sixth Edition, American Society for Photogrammetry and Remote Sensing, 2013.

[McMillan and Bishop 95] Leonard McMillan and Gary Bishop. "Plenoptic Modeling: An Image-Based Rendering System." In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '95*, pp. 39–46. New York, NY, USA: ACM, 1995.

[McNamara et al. 00] Ann McNamara, Alan Chalmers, Tom Troscianko, and Iain Gilchrist. "Comparing Real & Synthetic Scenes Using Human Judgements of Lightness." In *Proceedings of the Eurographics Workshop on Rendering Techniques 2000*, pp. 207–218. Springer-Verlag, 2000.

[Meyer et al. 86] Gary W. Meyer, Holly E. Rushmeier, Michael F. Cohen, Donald P. Greenberg, and Kenneth E. Torrance. "An Experimental Evaluation of Computer Graphics Imagery." *ACM Trans. Graph.* 5:1 (1986), 30–50.

[Meyer et al. 02] M. Meyer, M. Desbrun, P. Schroder, and A. H. Barr. "Discrete Differential-Geometry Operators for Triangulated 2-Manifolds." In *Proc. VisMath*, pp. 35–57, 2002.

[Meyer et al. 10] Quirin Meyer, Jochen Sümuth, Gerd Sussner, Marc Stamminger, and Günther Greiner. "On Floating-Point Normal Vectors." *Computer Graphics Forum* 29:4 (2010), 1405–1409.

[Micusik and Kosecka 10] Branislav Micusik and Jana Kosecka. "Multi-View Superpixel Stereo in Urban Environments." *International Journal of Computer Vision* 89:1 (2010), 106–119.

[Miguel et al. 12] Eder Miguel, Derek Bradley, Bernhard Thomaszewski, Bernd Bickel, Wojciech Matusik, Miguel A. Otaduy, and Steve Marschner. "Data-Driven Estimation of Cloth Simulation Models." *Comp. Graph. Forum* 31:2pt2 (2012), 519–528.

[Miguel et al. 13] Eder Miguel, Rasmus Tamstorf, Derek Bradley, Sara C. Schvartzman, Bernhard Thomaszewski, Bernd Bickel, Wojciech Matusik, Steve Marschner, and Miguel A. Otaduy. "Modeling and Estimation of Internal Friction in Cloth." *ACM Trans. Graph.* 32:6 (2013), 212:1–212:10.

[Mikolajczyk and Schmid 05] Krystian Mikolajczyk and Cordelia Schmid. "A Performance Evaluation of Local Descriptors." *TPAMI* 27:10.

[Mileva et al. 07] Yana Mileva, André Bruhn, and Joachim Weickert. "Illumination-Robust Variational Optical Flow with Photometric Invariants." In *DAGM Conference on Pattern Recognition*, pp. 152–162, 2007.

[Mitsunaga and Nayar 99] Tomoo Mitsunaga and Shree K. Nayar. "Radiometric Self Calibration." In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 374–380, 1999.

[Moeslund et al. 06] Thomas B. Moeslund, Adrian Hilton, and Volker Krüger. "A Survey of Advances in Vision-Based Human Motion Capture and Analysis." *Computer Vision and Image Understanding* 104:2 (2006), 90–126.

[Moffat et al. 95] A. Moffat, R. Neal, and Ian H. Witten. "Arithmetic Coding Revisited." In *Data Compression Conference, 1995. DCC '95. Proceedings*, pp. 202–211, 1995.

[Mohan et al. 08] Ankit Mohan, Ramesh Raskar, and Jack Tumblin. "Agile Spectrum Imaging: Programmable Wavelength Modulation for Cameras and Projectors." *Comput. Graph. Forum* 27:2 (2008), 709–717.

[Molina-Tanco and Hilton 00] Luis Molina-Tanco and Adrian Hilton. "Realistic Synthesis of Novel Human Movements from a Database of Motion Capture Examples." In *Workshop on Human Motion (HUMO)*, pp. 137–142, 2000.

[Moon and Marschner 06a] Jonathan T. Moon and Stephen R. Marschner. "Efficient Multiple Scattering in Hair Using Spherical Harmonics." *ACM Transactions on Graphics* 27:3.

[Moon and Marschner 06b] Jonathan T. Moon and Stephen R. Marschner. "Simulating Multiple Scattering in Hair Using a Photon Mapping Approach." *ACM Transactions on Graphics* 25:3.

[Moreno-Noguer et al. 07] Francesc Moreno-Noguer, Peter N. Belhumeur, and Shree K. Nayar. "Active Refocusing of Images and Videos." In *ACM SIGGRAPH 2007 Papers, SIGGRAPH '07*, pp. 67:1–67:9. New York, NY, USA: ACM, 2007.

[Mostafa et al. 99] Mostafa G.-H. Mostafa, Sameh M. Yamany, and Aly A. Farag. "Integrating Shape from Shading and Range Data Using Neural Networks." In *CVPR*, pp. 2015–2020, 1999.

[Mühlmann et al. 02] Karsten Mühlmann, Dennis Maier, Jürgen Hesser, and Reinhard Männer. "Calculating Dense Disparity Maps from Color Stereo Images, an Efficient Implementation." *International Journal of Computer Vision* 47:1-3 (2002), 79–88.

[Müller and Chentanez 10] Matthias Müller and Nuttapong Chentanez. "Wrinkle Meshes." In *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, SCA '10*, pp. 85–92. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2010.

[Müller et al. 04] Gero Müller, Jan Meseth, Mirko Sattler, Ralf Sarlette, and Reinhard Klein. "Acquisition, Synthesis and Rendering of Bidirectional Texture Functions." In *Eurographics 2004, State of the Art Reports*, edited by Christophe Schlick and Werner Purgathofer, pp. 69–94. INRIA and Eurographics Association, 2004.

[Müller et al. 08] K. Müller, A. Smolic, K. Dix, P. Kauff, and T. Wiegand. "Reliability-Based Generation and View Synthesis in Layered Depth Video." In *Multimedia Signal Processing, 2008 IEEE 10th Workshop on*, pp. 34–39. IEEE, 2008.

[Müller et al. 13] K. Müller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, F.H. Rhee, G. Tech, M. Winken, and T. Wiegand. "3D High-Efficiency Video Coding for Multi-View Video and Depth Data." *Image Processing, IEEE Transactions on* 22:9 (2013), 3366–3378.

[Müller 09] Gero Müller. "Data-Driven Methods for Compression and Editing of Spatially Varying Appearance." Ph.D. thesis, University of Bonn, 2009.

[Mundermann et al. 07] Lars Mundermann, Stefano Corazza, and Thomas P. Andriacchi. "Accurately Measuring Human Movement Using Articulated ICP with Soft-joint Constraints and a Repository of Articulated Models." In *IEEE CVPR*, 2007.

[Murray et al. 94] Richard M. Murray, Zexiang Li, S. Shankar Sastry, and S. Shankara Sastry. *A Mathematical Introduction to Robotic Manipulation.* CRC Press, 1994.

[Mustafa et al. 12] Maryam Mustafa, Lea Lindemann, and Marcus Magnor. "EEG Analysis of Implicit Human Visual Perception." In *SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pp. 513–516. ACM, 2012.

[Myronenko and Song 10] Andreiy Myronenko and Xubo Song. "Point Set Registration: Coherent Point Drift." *IEEE Transactions on PAMI* 32:12 (2010), 2262–2275.

[Myszkowski 98] Karol Myszkowski. "The Visible Differences Predictor: Applications to Global Illumination Problems." In *Rendering Techniques 98*, *Eurographics*, edited by George Drettakis and Nelson Max, pp. 223–236. Springer Vienna, 1998. Available online (http://dx.doi.org/10.1007/978-3-7091-6453-2_21).

[Nagel and Enkelmann 86] Hans-Helmut Nagel and Wilfried Enkelmann. "An Investigation of Smoothness Constraints for the Estimation of Displacement Vector Fields from Image Sequences." *IEEE Trans. Pattern Anal. Mach. Intell.* 8:5 (1986), 565–593.

[Nair et al. 12] Rahul Nair, Frank Lenzen, Stephan Meister, Henrik Schäfer, Christoph S. Garbe, and Daniel Kondermann. "High Accuracy TOF and Stereo Sensor Fusion at Interactive Rates." In *ECCV Workshops (2)*, pp. 1–11, 2012.

[Narain et al. 12] Rahul Narain, Armin Samii, and James F O'Brien. "Adaptive Anisotropic Remeshing for Cloth Simulation." *ACM Transactions on Graphics (TOG)* 31:6 (2012), 152.

[Narayanan et al. 98] P.J. Narayanan, Peter W. Rander, and Takeo Kanade. "Constructing Virtual Worlds Using Dense Stereo." In *Proceedings of the Sixth International Conference on Computer Vision*, pp. 3–11, 1998.

[Nayar and Mitsunaga 00] S.K. Nayar and T. Mitsunaga. "High Dynamic Range Imaging: Spatially Varying Pixel Exposures." In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1, 1, pp. 472–479, 2000.

[Nayar and Nakagawa 94] S.K. Nayar and Y. Nakagawa. “Shape from Focus.” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16:8 (1994), 824–831.

[Nayar et al. 06] Shree K. Nayar, Gurunandan Krishnan, Michael D. Grossberg, and Ramesh Raskar. “Fast Separation of Direct and Global Components of a Scene Using High Frequency Illumination.” *ACM Transactions on Graphics (Proc. SIGGRAPH 2006)* 25:3 (2006), 935–944.

[Nayar 97] Shree Nayar. “Catadioptric Omnidirectional Camera.” In *Proc. CVPR*. San Juan, 1997.

[Nealen et al. 05] Andrew Nealen, Matthias Müller, Richard Keiser, Eddy Boxermann, and Mark Carlson. “Physically Based Deformable Models in Computer Graphics.” In *Eurographics 2005. STAR—State of the Art Reports*, pp. 71–94, 2005.

[Nehab et al. 05] Diego Nehab, Szymon Rusinkiewicz, James Davis, and Ravi Ramamoorthi. “Efficiently Combining Positions and Normals for Precise 3D Geometry.” *SIGGRAPH* 24:3, 2005.

[Neumann et al. 13a] T. Neumann, K. Varanasi, N. Hasler, M. Wacker, M. Magnor, and C. Theobalt. “Capture and Statistical Modeling of Arm-Muscle Deformations.” *Computer Graphics Forum* 32:2pt3 (2013), 285–294.

[Neumann et al. 13b] T. Neumann, K. Varanasi, S. Wenger, M. Wacker, M. Magnor, and C. Theobalt. “Sparse Localized Deformation Components.” *ACM Transactions on Graphics (Proc. SIGGRAPH Asia)* 32:6 (2013), Art.179.

[Newcombe et al. 11] R. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. Davison, P. Kohli, J. Shotton, S. Hodges, and A. Fitzgibbon. “KinectFusion: Real-Time Dense Surface Mapping and Tracking.” In *Proc. IEEE Int. Symp. Mixed and Augm. Reality*, pp. 127–136, 2011.

[Ng et al. 02] A. Y. Ng, M. Jordan, and Y. Weiss. “On Spectral Clustering: Analysis and an Algorithm.” In *Proc. NIPS*, 2002.

[Ng 06] R. Ng. “Digital Light Field Photography.” PhD thesis, Stanford University, 2006. Note: thesis led to commercial light field camera, see also www.lytro.com.

[Nguyen et al. 12] H.M. Nguyen, B. Wünsche, P. Delmas, and C. Lutteroth. “3D Models from the Black Box: Investigating the Current State of Image-Based Modeling.” *Journal of WSCG* 20:1 (2012), 1–10.

[Nießner et al. 13] M. Nießner, M. Zollhöfer, S. Izadi, and M. Stamminger. “Real-Time 3D Reconstruction at Scale Using Voxel Hashing.” *ACM Transactions on Graphics (TOG)* 32:6 (2013), 169.

[Niquin et al. 10] Cédric Niquin, Stéphanie Prévost, and Yannick Rémion. “An Occlusion Approach with Consistency Constraint for Multiscopic Depth Extraction.” *Int. J. Digital Multimedia Broadcasting* 2010 (2010), 1–8.

[Nistér and Stewenius 06] David Nistér and Henrik Stewenius. "Scalable Recognition with a Vocabulary Tree." In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, 2, pp. 2161–2168. IEEE, 2006.

[Nistér 04] David Nistér. "An Efficient Solution to the Five-Point Relative Pose Problem." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26:6 (2004), 756–770.

[Nitsche 08] Michael Nitsche. "Experiments in the Use of Game Technology for Pre-Visualization." In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, pp. 160–165. ACM, 2008.

[Oggier et al. 05] T. Oggier, B. Büttgen, F. Lustenberger, G. Becker, B. Rüegg, and A. Hodac. "SwissRanger SR3000 and First Experiences Based on Miniaturized 3D-ToF Cameras." In *Proc. of the First Range Imaging Research Day at ETH Zurich*, 2005.

[Oikonomidis et al. 11a] Iason Oikonomidis, Nikolaos Kyriazis, and Antonis A Argyros. "Efficient Model-Based 3D Tracking of Hand Articulations Using Kinect." In *BMVC*, pp. 1–11, 2011.

[Oikonomidis et al. 11b] Iasonas Oikonomidis, Nikolaos Kyriazis, and Antonis A Argyros. "Full dof Tracking of a Hand Interacting with an Object by Modeling Occlusions and Physical Constraints." In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 2088–2095. IEEE, 2011.

[Okutomi and Kanade 93] Masatoshi Okutomi and Takeo Kanade. "A Multiple-Baseline Stereo." *IEEE Trans. Pattern Anal. Mach. Intell.* 15:4 (1993), 353–363.

[O'Sullivan and Dingliana 01] Carol O'Sullivan and John Dingliana. "Collisions and Perception." *ACM Transactions on Graphics (TOG)* 20:3 (2001), 151–168.

[O'Toole and Kutulakos 10] Matthew O'Toole and Kiriakos N. Kutulakos. "Optical Computing for Fast Light Transport Analysis." *ACM Trans. Graph.* 29:6 (2010), 164:1–164:12.

[O'Toole et al. 12] Matthew O'Toole, Ramesh Raskar, and Kiriakos N. Kutulakos. "Primal-Dual Coding to Probe Light Transport." *ACM Trans. Graph.* 31:4 (2012), 39:1–39:11. Available online (http://doi.acm.org/10.1145/2185520.2185535).

[Ozawa et al. 12] Tomohiro Ozawa, Kris M. Kitani, and Hideki Koike. "Human-Centric Panoramic Imaging Stitching." In *Proceedings of the 3rd Augmented Human International Conference*, p. 20. ACM, 2012.

[Pajarola and Rossignac 00] Renato B. Pajarola and Jarek Rossignac. "SQUEEZE: Fast and Progressive Decompression of Triangle Meshes." In *Proc. CGI*, pp. 173–182, 2000.

[Park and Hodgins 06] S. I. Park and J. K. Hodgins. "Capturing and Animating Skin Deformation in Human Motion." *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 25:3.

[Park and Hodgins 08] S. I. Park and J. K. Hodgins. "Data-Driven Modeling of Skin and Muscle Deformation." *ACM Transactions on Graphics*, 27:3(2008), 96:1–96:6.

[Parker et al. 98] S. Parker, P. Shirley, Y. Livnat, Ch. Hansen, and P. Sloan. "Interactive Ray Tracing for Isosurface Rendering." In *Proc. Visualization*, pp. 233–238, 1998.

[Peers and Dutré 03] Pieter Peers and Philip Dutré. "Wavelet Environment Matting." In *Rendering Techniques 2003 (Proc. Eurographics Symposium on Rendering)*, pp. 157–166, 2003. Available online (http://www.eg.org/EG/DL/WS/EGWR/EGWR03/157-166.pdf).

[Peers and Dutré 05] Pieter Peers and Philip Dutré. "Inferring Reflectance Functions from Wavelet Noise." In *Rendering Techniques*, pp. 173–182, 2005.

[Peers et al. 09] Pieter Peers, Dhruv K. Mahajan, Bruce Lamond, Abhijeet Ghosh, Wojciech Matusik, Ravi Ramamoorthi, and Paul Debevec. "Compressive Light Transport Sensing." *ACM Trans. Graph.* 28:1 (2009), 3:1–3:18. Available online (http://doi.acm.org/10.1145/1477926.1477929).

[Peinsipp-Byma et al. 09] E. Peinsipp-Byma, N. Rehfeld, and R. Eck. "Evaluation of Stereoscopic 3D Displays for Image Analysis Tasks." In *Stereoscopic Displays and Applications XX*, edited by Andrew J. Woods, Nicolas S. Holliman, and John O. Merritt, p. 72370L. SPIE, 2009.

[Peleg et al. 99] Shmuel Peleg, Yael Pritch, and Moshe Ben-Ezra. "Stereo Panorama with a Single Camera." In *Proc. CVPR*. Fort Collins, USA, 1999.

[Peleg et al. 00] Shmuel Peleg, Yael Pritch, and Moshe Ben-Ezra. "Cameras for Stereo Panoramic Imaging." In *Proc. CVPR*, pp. 208–214. Hilton Head Island, 2000.

[Peleg et al. 01] Shmuel Peleg, Moshe Ben-Ezra, and Yael Pritch. "Omnistereo: Panoramic Stereo Imaging." *IEEE Trans. Pattern Analysis and Machine Intelligence* 23:3 (2001), 279–290.

[Pellacini et al. 00] Fabio Pellacini, James A. Ferwerda, and Donald P. Greenberg. "Toward a Psychophysically-Based Light Reflection Model for Image Synthesis." In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '00*, pp. 55–64. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 2000.

[Peng et al. 05] Jingliang Peng, Chang-Su Kim, and C. C. Jay Kuo. "Technologies for 3D Mesh Compression: A Survey." *J. Vis. Commun. Image Represent.*, pp. 688–733.

[Perez and Luke 09] N. Perez and J. Luke. "Simultaneous Estimation of Super-Resolved Depth and All-in-Focus Images from a Plenoptic Camera." In *3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video*, pp. 1–4. IEEE, 2009.

[Perlin et al. 00] K. Perlin, S. Paxia, and J. S. Kollin. “An Autostereoscopic Display.” In *ACM SIGGRAPH 2000 Conference Proceedings*, 33, 33, pp. 319–326, 2000.

[Perwass and Wietzke 12] C. Perwass and L. Wietzke. “Single Lens 3D-Camera with Extended Depth-of-Field.” *SPIE Electronic Imaging*, pp. 22–26.

[Petit et al. 11] Benjamin Petit, Antoine Letouzey, Edmond Boyer, and Jean-Sébastien Franco. “Surface Flow from Visual Cues.” In *VMV 2011—Vision, Modeling and Visualization Workshop*, pp. 1–8. Berlin, Germany, 2011.

[Pharr and Humphreys 10] Matt Pharr and Greg Humphreys. *Physically Based Rendering: From Theory To Implementation*, Second edition. Morgan Kaufmann Publishers Inc., 2010.

[Phong 75] Bui Tuong Phong. “Illumination for Computer Generated Pictures.” *Commun. ACM* 18:6 (1975), 311–317.

[Pollefeys et al. 04] Marc Pollefeys, Luc Van Gool, Maarten Vergauwen, Frank Verbiest, Kurt Cornelis, Jan Tops, and Reinhard Koch. “Visual Modeling with a Hand-Held Camera.” *International Journal of Computer Vision* 59:3 (2004), 207–232.

[Pool et al. 11] Jeff Pool, Anselmo Lastra, and Montek Singh. “Precision Selection for Energy-Efficient Pixel Shaders.” In *Proceedings of the ACM SIGGRAPH Symposium on High Performance Graphics, HPG ’11*, pp. 159–168. New York, NY, USA: ACM, 2011. Available online (http://doi.acm.org/10.1145/2018323.2018349).

[Popa et al. 09] Tiberiu Popa, Quan Zhou, Derek Bradley, Vladislav Kraevoy, Hongbo Fu, Alla Sheffer, and Wolfgang Heidrich. “Wrinkling Captured Garments Using Space-Time Data-Driven Deformation.” *Comput. Graph. Forum* 28:2 (2009), 427–435.

[Potmesil 87] Michael Potmesil. “Generating Octree Models of 3D Objects from Their Silhouettes in a Sequence of Images.” *Computer Vision, Graphics, and Image Processing* 40:1 (1987), 1–29.

[Press et al. 07] William Press, Saul Tuekolski, William Vetterling, and Brian Flannery. *Numerical Recipes: The Art of Scientific Computing, 3rd edition.* Cambridge University Press, 2007.

[Prévost et al. 13] Stéphanie Prévost, Cédric Niquin, Sylvie Chambon, and Guillaume Gales. “Multi- and Stereoscopic Matching, Depth and Disparity.” In *In 3D Video: From Capture to Diffusion*, edited by Yannick Rémion Laurent Lucas and Cline Loscos, pp. 137–154. Wiley-ISTE, 2013.

[Prévoteau et al. 10] Jessica Prévoteau, Sylvia Chalençon-Piotin, Didier Debons, Laurent Lucas, and Yannick Remion. “Multi-View Shooting Geometry for Multiscopic Rendering with Controlled Distortion.” *International Journal of Digital Multimedia Broadcasting (IJDMB), special issue Advances in 3DTV: Theory and Practice* 2010 (2010), 1–11.

[Pritchard and Heidrich 03] David Pritchard and Wolfgang Heidrich. "Cloth Motion Capture." *Comput. Graph. Forum* 22:3 (2003), 263–272.

[Provot 95] Xavier Provot. "Deformation Constraints in a Mass-Spring Model to Describe Rigid Cloth Behavior." In *In Graphics Interface*, pp. 147–154, 1995.

[Pujades et al. 14] S. Pujades, B. Goldlücke, and F. Devernay. "Bayesian View Synthesis and Image-Based Rendering Principles." In *Proc. International Conference on Computer Vision and Pattern Recognition*, 2014.

[Qi and Cooperstock 08] Zhi Qi and Jeremy R. Cooperstock. "Depth-Based Image Mosaicing for Both Static and Dynamic Scenes." In *Proc. Int. Conf. on Pattern Recognition (ICPR)*, pp. 1–4. Tampa, USA, 2008.

[Quan and Lan 99] Long Quan and Zhongdan Lan. "Linear N-Point Camera Pose Estimation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21:7.

[Raguram et al. 11] Rahul Raguram, Changchang Wu, Jan-Michael Frahm, and Svetlana Lazebnik. "Modeling and Recognition of Landmark Image Collections Using Iconic Scene Graphs." *International Journal of Computer Vision* 95:3 (2011), 213–239.

[Ramamoorthi and Hanrahan 01a] Ravi Ramamoorthi and Pat Hanrahan. "An Efficient Representation for Irradiance Environment Maps." In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 497–500. ACM, 2001.

[Ramamoorthi and Hanrahan 01b] Ravi Ramamoorthi and Pat Hanrahan. "On the Relationship between Radiance and Irradiance: Determining the Illumination from Images of a Convex Lambertian Object." *Journal of the Optical Society of America* 18:10 (2001), 2448–2459.

[Ramamoorthi and Hanrahan 01c] Ravi Ramamoorthi and Pat Hanrahan. "A Signal-Processing Framework for Inverse Rendering." In *Proceedings of SIGGRAPH 2001, Computer Graphics Proceedings, Annual Conference Series*, edited by Eugene Fiume, pp. 117–128. ACM Press / ACM SIGGRAPH, 2001. ISBN 1-58113-292-1.

[Ramanarayanan et al. 07] Ganesh Ramanarayanan, James Ferwerda, Bruce Walter, and Kavita Bala. "Visual Equivalence: Towards a New Standard for Image Fidelity." *ACM Trans. Graph.* 26:3. Available online (http://doi.acm.org/10.1145/1276377.1276472).

[Raposo et al. 13] Carolina Raposo, Joao Pedro Barreto, and Urbano Nunes. "Fast and Accurate Calibration of a Kinect Sensor." In *Proceedings of the 2013 International Conference on 3D Vision, 3DV '13*, pp. 342–349. Washington, DC, USA: IEEE Computer Society, 2013.

[Read and Bohr 14] Jenny C.A. Read and Iwo Bohr. "User Experience while Viewing Stereoscopic 3D Television." *Ergonomics*: ahead-of-print (2014), 1–14.

[Reibel et al. 03] Y. Reibel, M. Jung, M. Bouhifd, B. Cunin, and C. Draman. "CCD or CMOS Camera Noise Characterisation." *The European Physical Journal Applied Physics* 21 (2003), 75–80.

[Reinhard et al. 08] Erik Reinhard, Erum Arif Khan, Ahmet Oguz Akyz, and Garrett M Johnson. *Color Imaging: Fundamentals and Applications.* A.K. Peters, Ltd., 2008.

[Reinhard et al. 10] Erik Reinhard, Greg Ward, Summant Pattanaik, Paul Debevec, Wolfgang Heidrich, and Karol Myszkowski. *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*, Second edition. The Morgan Kaufmann series in Computer Graphics, Burlington, MA: Elsevier (Morgan Kaufmann), 2010.

[Ren et al. 10] Zhong Ren, Kun Zhou, Tengfei Li, Wei Hua, and Baining Guo. "Interactive Hair Rendering under Environment Lighting." *ACM Transactions on Graphics (Proc. SIGGRAPH 2010).*

[Rhemann et al. 11] Christoph Rhemann, Asmaa Hosni, Michael Bleyer, Carsten Rother, and Margrit Gelautz. "Fast Cost-Volume Filtering for Visual Correspondence and beyond." In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 3017–3024. IEEE, 2011.

[Richardt et al. 10] Christian Richardt, Douglas Orr, Ian Davies, Antonio Criminisi, and Neil A. Dodgson. "Real-Time Spatiotemporal Stereo Matching Using the Dual-Cross-Bilateral Grid." In *Computer Vision Conference on Computer Vision: Part III*, pp. 510–523, 2010.

[Richardt et al. 13] Christian Richardt, Yael Pritch, Henning Zimmer, and Alexander Sorkine-Hornung. "Megastereo: Constructing High-Resolution Stereo Panoramas." In *Proc. CVPR 2013*, pp. 1256–1263. Portland, USA, 2013.

[Ritschel et al. 12] Tobias Ritschel, Carsten Dachsbacher, Thorsten Grosch, and Jan Kautz. "The State of the Art in Interactive Global Illumination." *Computer Graphics Forum* 31:1 (2012), 160–188.

[Robinette et al. 99] K. Robinette, H. Daanen, and E. Paquet. "The CAESAR Project: A 3-D Surface Anthropometry Survey." In *Conference on 3D Digital Imaging and Modeling*, pp. 180–186, 1999.

[Rodriguez-Ramos et al. 11] J.M. Rodriguez-Ramos, J.G. Marichal-Hernandez, J.P. Luke, J. Trujillo-Sevilla, M. Puga, M. Lopez, J.J. Fernandez-Valdivia, C. Dominguez-Conde, J.C. Sanluis, F. Rosa, V. Guadalupe, H. Quintero, C. Militello, L.F. Rodriguez-Ramos, R. Lopez, I. Montilla, and B. Femenia. "New Developments at CAFADIS Plenoptic Camera." In *Information Optics (WIO), 2011 10th Euro-American Workshop on*, pp. 1 –3, 2011.

[Rogge et al. 14] Lorenz Rogge, Felix Klose, Michael Stengel, Martin Eisemann, and Marcus Magnor. "Garment Replacement in Monocular Video Sequences." *ACM Transactions on Graphics (TOG).* To appear.

[Rogmans et al. 09] Sammy Rogmans, Jiangbo Lu, Philippe Bekaert, and Gauthier Lafruit. "Real-Time Stereo-Based View Synthesis Algorithms: A Unified Framework and Evaluation on Commodity GPUs." *Image Communication* 24:1-2 (2009), 49–64.

[Rohmer et al. 10] Damien Rohmer, Tiberiu Popa, Marie-Paule Cani, Stefanie Hahmann, and Alla Sheffer. "Animation Wrinkling: Augmenting Coarse Cloth Simulations with Realistic-Looking Wrinkles." *ACM Trans. Graph.* 29:6 (2010), 157:1–157:8.

[Rose et al. 98] C. Rose, M. Cohen, and B. Bodenheimer. "Verbs and Adverbs: Multidimensional Motion Interpolation." *IEEE Computer Graphics and Applications* 18:5 (1998), 32–40.

[Rosenthal et al. 01] S. Rosenthal, D. Griffin, and M Sanders. "Real-Time Compter Graphics for On-set Visualization: "A.I." and "The Mummy Returns."" In *Siggraph 2001, Sketches and Applications*, 2001.

[Roth and Vona 12] Henry Roth and Marsette Vona. "Moving Volume KinectFusion." In *Proc. BMVC*, pp. 1–11, 2012.

[Rublee et al. 11] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. "ORB: An Efficient Alternative to SIFT or SURF." In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 2564–2571. IEEE, 2011.

[Rucci 08] Michele Rucci. "Fixational Eye Movements, Natural Image Statistics, and Fine Spatial Vision." *Network: Computation in Neural Systems* 19:4 (2008), 253–285.

[Ruhl et al. 12a] Kai Ruhl, Benjamin Hell, Felix Klose, Christian Lipski, Sören Petersen, and Marcus Magnor. "Improving Dense Image Correspondence Estimation with Interactive User Guidance." In *ACM Multimedia*, pp. 1129–1132, 2012.

[Ruhl et al. 12b] Kai Ruhl, Felix Klose, Christian Lipski, and Marcus A. Magnor. "Integrating Approximate Depth Data into Dense Image Correspondence Estimation." In *European Conference on Visual Media Production (CVMP)*, pp. 26–31, 2012.

[Ruhl et al. 13] Kai Ruhl, Martin Eisemann, and Marcus Magnor. "Cost Volume-Based Interactive Depth Editing in Stereo Post-Processing." In *European Conference on Visual Media Production (CVMP)*, pp. 1–6, 2013.

[Rupprecht et al. 13] C. Rupprecht, O. Pauly, C. Theobalt, and S. Ilic. "3D Semantic Parameterization for Human Shape Modeling: Application to 3D Animation." In *3D Vision*, 2013.

[Rusinkiewicz and Levoy 01] Szymon Rusinkiewicz and Marc Levoy. "Efficient Variants of the ICP Algorithm." In *Third International Conference on 3D Digital Imaging and Modeling (3DIM)*, pp. 145–152, 2001.

[Rusinkiewicz et al. 02] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy. "Real-Time 3D Model Acquisition." *ACM Trans. Graph* 21:3 (2002), 438–446.

[Ryan White 07] David Forsyth Ryan White, Keenan Crane. "Data Driven Cloth Animation." SIGGRAPH Technical Sketch, 2007.

[Sabov and Krüger 10] A. Sabov and J. Krüger. "Identification and Correction of Flying Pixels in Range Camera Data." In *Proc. Spring Conf. on Computer Graphics*, pp. 135–142, 2010.

[Sadeghi et al. 13] Iman Sadeghi, Oleg Bisker, Joachim De Deken, and Henrik Wann Jensen. "A Practical Microcylinder Appearance Model for Cloth Rendering." *ACM Transactions on Graphics (TOG)* 32:2 (2013), 14.

[Sand and Teller 06] Peter Sand and Seth J. Teller. "Particle Video: Long-Range Motion Estimation Using Point Trajectories." In *Computer Vision and Pattern Recognition*, pp. 2195–2202, 2006.

[Sander and Mitchell 05] Pedro V. Sander and Jason L. Mitchell. "Progressive Buffers: View-Dependent Geometry and Texture LOD Rendering." In *Proc. SGP*, pp. 129–138, 2005.

[Sato et al. 99] Imari Sato, Yoichi Sato, and Katsushi Ikeuchi. "Acquiring a Radiance Distribution to Superimpose Virtual Objects onto a Real Scene." *IEEE Transactions on Visualization and Computer Graphics* 5:1 (1999), 1–12.

[Sattler et al. 03] Mirko Sattler, Ralf Sarlette, and Reinhard Klein. "Efficient and Realistic Visualization of Cloth." In *Eurographics Symposium on Rendering 2003*, 2003.

[Schaefer and Yuksel 07] S. Schaefer and C. Yuksel. "Example-Based Skeleton Extraction." In *Proceedings of the Symposium on Geometry Processing (SGP)*, pp. 153–162, 2007.

[Scharstein and Szeliski 02] Daniel Scharstein and Richard Szeliski. "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms." *International Journal of Computer Vision* 47:1/2/3 (2002), 7–42.

[Scharstein and Szeliski 03] Daniel Scharstein and Richard Szeliski. "High-Accuracy Stereo Depth Maps Using Structured Light." In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 195–202, 2003.

[Schmeing and Jiang 11] Michael Schmeing and Xiaoyi Jiang. "Time-Consistency of Disocclusion Filling Algorithms in Depth Image-Based Rendering." In *3DTV Conference: The True Vision—Capture, Transmission and Display of 3D Video (3DTV-CON), 2011*, pp. 1–4, 2011.

[Schmidt and Jahne 11] M. Schmidt and B. Jahne. "Efficient and Robust Reduction of Motion Artifacts for 3D Time-of-Flight Cameras." In *Proc. Int. Conf. 3D Imaging (IC3D)*, pp. 1–8, 2011.

[Schneider and Eisert 12] David C. Schneider and Peter Eisert. "On User-Interaction in 3D Reconstruction." In *Eurographics (poster)*. Cagliari, Italy, 2012.

[Schödl et al. 00] Arno Schödl, Richard Szeliski, David H Salesin, and Irfan Essa. "Video Textures." In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 489–498. ACM Press/Addison-Wesley Publishing Co., 2000.

[Schoemake 94] Ken Schoemake. "Euler Angle Conversion." In *Graphics Gems, IV*, pp. 222–229. Academic Press, 1994.

[Scholz and Magnor 06] Volker Scholz and Marcus Magnor. "Texture Replacement of Garments in Monocular Video Sequences." In *Proc. Eurographics Symposium on Rendering (EGSR)*, pp. 305–312, 2006.

[Scholz et al. 05] Volker Scholz, Timo Stich, Michael Keckeisen, Markus Wacker, and Marcus Magnor. "Garment Motion Capture Using Color-Coded Patterns." *Computer Graphics Forum* 24:3 (2005), 439–448.

[Schreer et al. 12] Oliver Schreer, Peter Kauff, Peter Eisert, Christian Weissig, and Jean-Claude Rosenthal. "Geometrical Design Concept for Panoramic 3D Video Acquisition." In *Proc. 20th European Signal Processing Conference (EUSICPO 2012)*, pp. 2757–2761. Bucharest, Romania, 2012.

[Schreer et al. 13] Oliver Schreer, Ingo Feldmann, Christian Weissig, Peter Kauff, and Ralf Schäfer. "Ultrahigh-Resolution Panoramic Imaging for Format-Agnostic Video Production." *Proceedings of the IEEE* 101:1 (2013), 99–114.

[Schröder et al. 11] Kai Schröder, Reinhard Klein, and Arno Zinke. "A Volumetric Approach to Predictive Rendering of Fabrics." In *Computer Graphics Forum*, 30, 30, pp. 1277–1286. Wiley Online Library, 2011.

[Schröder et al. 12] Kai Schröder, Shuang Zhao, and Arno Zinke. "Recent Advances in Physically-Based Appearance Modeling of Cloth." In *SIGGRAPH Asia 2012 Courses, SA '12*, pp. 12:1–12:52, 2012.

[Schuon et al. 09] Sebastian Schuon, Christian Theobalt, James Davis, and Sebastian Thrun. "Lidarboost: Depth Superresolution for ToF 3D Shape Scanning." In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 343–350. IEEE, 2009.

[Schwartz et al. 10] Christopher Schwartz, Ruwen Schnabel, Patrick Degener, and Reinhard Klein. "PhotoPath: Single Image Path Depictions from Multiple Photographs." *Journal of WSCG* 18:1-3.

[Schwartz et al. 11a] Christopher Schwartz, Roland Ruiters, Michael Weinmann, and Reinhard Klein. "WebGL-Based Streaming and Presentation Framework for Bidirectional Texture Functions." In *The 12th International Symposium on Virtual Reality, Archeology and Cultural Heritage VAST 2011*, pp. 113–120. Eurographics Association, Eurographics Association, 2011.

[Schwartz et al. 11b] Christopher Schwartz, Michael Weinmann, Roland Ruiters, and Reinhard Klein. "Integrated High-Quality Acquisition of Geometry and Appearance for Cultural Heritage." In *The 12th International Symposium on Virtual Reality, Archeology and Cultural*

Heritage VAST 2011, pp. 25–32. Eurographics Association, Eurographics Association, 2011.

[Schwartz et al. 13] Christopher Schwartz, Ralf Sarlette, Michael Weinmann, and Reinhard Klein. "DOME II: A Parallelized BTF Acquisition System." In *Eurographics Workshop on Material Appearance Modeling: Issues and Acquisition*, edited by Holly Rushmeier and Reinhard Klein, pp. 25–31, 2013.

[Segal et al. 92] Marc Segal, Carl Korobkin, Rolf van Widenfelt, Jim Foran, and Paul Haeberli. "Fast Shadows and Lighting Effects Using Texture Mapping." *SIGGRAPH Comput. Graph.* 26:2 (1992), 249–252.

[Seitz and Dyer 96] Steven M. Seitz and Charles R. Dyer. "View Morphing." In *Proc. of ACM SIGGRAPH'96*, pp. 21–30. New York: ACM Press/ACM SIGGRAPH, 1996.

[Seitz and Dyer 99] Steven M. Seitz and Charles R. Dyer. "Photorealistic Scene Reconstruction by Voxel Coloring." *Int. J. Comput. Vision* 35:2 (1999), 151–173.

[Seitz et al. 05] Steven M. Seitz, Yasuyuki Matsushita, and Kiriakos N. Kutulakos. "A Theory of Inverse Light Transport." In *Proceedings of IEEE International Conference on Computer Vision*, pp. 1440–1447, 2005.

[Seitz et al. 06] Steven M. Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms." In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, 1, pp. 519–528, 2006.

[Sellent et al. 11] Anita Sellent, Martin Eisemann, Bastian Goldlücke, Daniel Cremers, and Marcus Magnor. "Motion Field Estimation from Alternate Exposure Images." *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* 33:8 (2011), 1577–1589.

[Sellent et al. 12] Anita Sellent, Kai Ruhl, and Marcus Magnor. "A Loop-Consistency Measure for Dense Correspondences in Multi-View Video." *Journal of Image and Vision Computing* 30:9 (2012), 641–654.

[Sen and Darabi 09] Pradeep Sen and Soheil Darabi. "Compressive Dual Photography." *Computer Graphics Forum* 28:2 (2009), 609–618.

[Sen et al. 05] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. "Dual Photography." 24 (2005), 745755. Available online (http://dl.acm.org/citation.cfm?id=1073257).

[Shade et al. 98] J. Shade, S. Gortler, L. He, and R. Szeliski. "Layered Depth Images." In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 231–242. ACM, 1998.

[Shaheen et al. 09] Mohammed Shaheen, Jürgen Gall, Robert Strzodka, Luc J. Van Gool, and Hans-Peter Seidel. "A Comparison of 3D Model-Based Tracking Approaches for Human Motion Capture in Uncontrolled Environments." In *WACV*, 2009.

[Sharf et al. 07] A. Sharf, T. Lewiner, A. Shamir, and L. Kobbelt. "On-the-Fly Skeleton Computation for 3D Shapes." *Computer Graphics Forum (Proceedings of EUROGRAPHICS)*, pp. 323–328.

[Shi and Tomasi 94] Jianbo Shi and Carlo Tomasi. "Good Features to Track." In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94, 1994 IEEE Computer Society Conference on*, pp. 593–600. IEEE, 1994.

[Shotton et al. 11] Jamie Shotton, Andrew W. Fitzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, and Andrew Blake. "Real-Time Human Pose Recognition in Parts from Single Depth Images." In *IEEE Computer Vision and Pattern Recognition (CVPR)*, pp. 1297–1304, 2011.

[Shum and He 99] Heung-Yeung Shum and Li-Wei He. "Rendering with Concentric Mosaics." In *SIGGRAPH '99*, pp. 299–306. Los Angeles, USA, 1999.

[Shum and Kang 00] Heung-Yeung Shum and Sing Bing Kang. "A Review of Image-Based Rendering Techniques." *IEEE/SPIE Visual Communications and Image Processing (VCIP)* 213.

[Sigal et al. 12] Leonid Sigal, Michael Isard, Horst Haussecker, and Michael J. Black. "Loose-Limbed People: Estimating 3D Human Pose and Motion Using Non-Parametric Belief Propagation." *IJCV* 98:1 (2012), 15–48.

[Sinha et al. 11] Sudipta N. Sinha, Jan-Michael Frahm, Marc Pollefeys, and Yakup Genc. "Feature Tracking and Matching in Video Using Programmable Graphics Hardware." *Machine Vision and Applications* 22:1 (2011), 207–217.

[Sinha et al. 12] Sudipta N. Sinha, Drew Steedly, and Richard Szeliski. "A Multi-Stage Linear Approach to Structure from Motion." In *Trends and Topics in Computer Vision*, pp. 267–281. Springer, 2012.

[SizeGermany 07] SizeGermany, 2007. Available online (http://www.sizegermany.de).

[Slama 80] C. C. Slama. *Manual of Photogrammetry.* Americal Society of Photogrammetry, 1980.

[Smisek et al. 11] J. Smisek, M. Jancosek, and T. Pajdla. "3D with Kinect." In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pp. 1154–1160, 2011.

[Smith and Cheeseman 86] Randall C. Smith and Peter Cheeseman. "On the Representation and Estimation of Spatial Uncertainly." *Int. J. Rob. Res.* 5:4 (1986), 56–68.

[Smolic et al. 09] A. Smolic, K. Mueller, P. Merkle, P. Kauff, and T. Wiegand. "An Overview of Available and Emerging 3D Video Formats and Depth Enhanced Stereo as Efficient Generic Solution." In *Picture Coding Symposium, 2009. PCS 2009*, pp. 1–4, 2009.

[Snavely et al. 06] N. Snavely, S.M. Seitz, and R. Szeliski. "Photo Tourism: Exploring Photo Collections in 3D." In *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 25, 25, pp. 835–846, 2006.

[Snavely et al. 08a] Noah Snavely, Rahul Garg, Steven M. Seitz, and Richard Szeliski. "Finding Paths through the World's Photos." *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2008)* 27:3 (2008), 11–21.

[Snavely et al. 08b] Noah Snavely, Steven M Seitz, and Richard Szeliski. "Modeling the World from Internet Photo Collections." *International Journal of Computer Vision* 80:2 (2008), 189–210.

[Snavely et al. 08c] Noah Snavely, Steven M Seitz, and Richard Szeliski. "Skeletal Graphs for Efficient Structure from Motion." In *CVPR*, 1, p. 2, 2008.

[Snavely 12] Noah Snavely. "Bundler: Structure from Motion (SfM) for Unordered Image Collections." http://phototour.cs.washington.edu/bundler/, 2012.

[Son et al. 07] Jung-Young Son, Yuri N. Gruts, Kae-Dal Kwack, Kyung-Hun Cha, and Sung-Kyu Kim. "Stereoscopic Image Distortion in Radial Camera and Projector Configurations." In *J. Opt. Soc. Am. A*, 24, pp. 643–650. OSA, 2007.

[Sorkine et al. 04] Olga Sorkine, Daniel Cohen-Or, Yaron Lipman, Marc Alexa, Christian Rssl, and Hans-Peter Seidel. "Laplacian Surface Editing." In *ACM Symposium on Geometry Processing*, pp. 175–184, 2004.

[Starck and Hilton 05] Jonathan Starck and Adrian Hilton. "Spherical Matching for Temporal Correspondence of Non-Rigid Surfaces." In *IEEE ICCV*, 2005.

[Starck and Hilton 07a] Jonathan Starck and Adrian Hilton. "Correspondence Labelling for Wide-Time Free-Form Surface Matching." In *IEEE ICCV*, 2007.

[Starck and Hilton 07b] Jonathan Starck and Adrian Hilton. "Surface Capture for Performance-Based Animation." *IEEE Computer Graphics and Application* 27 (2007), 21–31.

[Steinbrücker et al. 09] Frank Steinbrücker, Thomas Pock, and Daniel Cremers. "Large Displacement Optical Flow Computation without Warping." In *International Conference on Computer Vision, ICCV*, pp. 1609–1614, 2009.

[Steitz and Pannekamp 05] A. Steitz and J. Pannekamp. "Systematic Investigation of Properties of PMD-Sensors." In *Proc. 1st Range Imaging Research Day*, pp. 59–69. ETH Zurich, 2005.

[Stelmach et al. 03] Lew B Stelmach, Wa James Tam, Filippo Speranza, Ronald Renaud, and Taali Martin. "Improving the Visual Comfort of Stereoscopic Images." In *Electronic Imaging 2003*, pp. 269–282. International Society for Optics and Photonics, 2003.

[Stewart et al. 03] J. Stewart, J. Yu, S. J. Gortler, and L. McMillan. "A New Reconstruction Filter for Undersampled Light Fields." In *Proceedings of the 14th Eurographics Workshop on Rendering, EGRW '03*, pp. 150–156. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2003.

[Stich et al. 08] Timo Stich, Christian Linz, Georgia Albuquerque, and Marcus Magnor. "View and Time Interpolation in Image Space." *Computer Graphics Forum* 27:7 (2008), 1781–1787.

[Stich et al. 11] Timo Stich, Christian Linz, Christian Wallraven, Douglas Cunningham, and Marcus Magnor. "Perception-Motivated Interpolation of Image Sequences." *ACM Transactions on Applied Perception (TAP)* 8:2 (2011), 11:1–11:25.

[Stolfi 91] J. Stolfi. *Oriented Projective Geometry: A Framework for Geometric Computation.* Boston: Academic Press, 1991.

[Stoll et al. 10] Carsten Stoll, Juergen Gall, Edilson de Aguiar, Sebastian Thrun, and Christian Theobalt. "Video-Based Reconstruction of Animatable Human Characters." *ACM Transactions on Graphics (Proc. SIGGRAPH ASIA 2011)* 29:6 (2010), 139:1–139:10.

[Straka et al. 12] Matthias Straka, Stefan Hauswiesner, Matthias Rüther, and Horst Bischof. "Simultaneous Shape and Pose Adaption of Articulated Models Using Linear Optimization." In *ECCV*, pp. 724–737, 2012.

[Stumpfel et al. 04] Jessi Stumpfel, Chris Tchou, Andrew Jones, Tim Hawkins, Andreas Wenger, and Paul Debevec. "Direct HDR Capture of the Sun and Sky." In *Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa, AFRIGRAPH '04*, pp. 145–149. New York, NY, USA: ACM, 2004.

[Sturm and Triggs 96] Peter Sturm and Bill Triggs. "A Factorization-Based Algorithm for Multi-Image Projective Structure and Motion." In *4th European Conference on Computer Vision*, Cambridge, England, pp. 709–720, 1996.

[Sturm et al. 11] Peter Sturm, Srikumar Ramalingam, Jean-Philippe Tardif, and Joao Barreto. "Camera Models and Fundamental Concepts Used in Geometric Computer Vision." *Foundations and Trends in Computer Graphics and Vision* 6:1-2.

[Sturm 01] Peter Sturm. "On Focal Length Calibration from Two Views." In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 2, pp. II–145. IEEE, 2001.

[Sumner and Popović 04] R. W. Sumner and J. Popović. "Deformation Transfer for Triangle Meshes." *ACM Transactions on Graphics (Proceedings of SIGGRAPH)* 23:3 (2004), 399–405.

[Sumner et al. 07] R. W. Sumner, J. Schmid, and M. Pauly. "Embedded Deformation for Shape Manipulation." *ACM Transactions on Graphics (Proceedings of SIGGRAPH)*, 2007.

[Sun and Abidi 01] Y. Sun and M. Abidi. "Surface Matching by 3D Point's Fingerprint." In *International Conference on Computer Vision*, 2, pp. 263–269, 2001.

[Sun et al. 03] Jian Sun, Nan-Ning Zheng, and Heung-Yeung Shum. "Stereo Matching Using Belief Propagation." *IEEE Trans. Pattern Anal. Mach. Intell.* 25:7 (2003), 787–800.

[Sun et al. 05] Jian Sun, Yin Li, Sing Bing Kang, and Heung-Yeung Shum. "Symmetric Stereo Matching for Occlusion Handling." In *Computer Vision and Pattern Recognition*, pp. 399–406, 2005.

[Sun et al. 10] Deqing Sun, Stefan Roth, and Michael J. Black. "Secrets of Optical Flow estimation and Their Principles." In *Computer Vision and Pattern Recognition*, pp. 2432–2439, 2010.

[Svarovsky 99] Jan Svarovsky. "Extreme Detail Graphics." In *Proc. Game Developers Conference*, pp. 889–904, 1999.

[Svoboda et al. 05] Tomáš Svoboda, Daniel Martinec, and Tomáš Pajdla. "A Convenient Multi-Camera Self-Calibration for Virtual Environments." *PRESENCE: Teleoperators and Virtual Environments* 14:4 (2005), 407–422.

[Szeliski and Golland 99] Richard Szeliski and Polina Golland. "Stereo Matching with Transparency and Matting." *Int. J. Comput. Vision* 32:1 (1999), 45–61.

[Szeliski et al. 08] Richard Szeliski, Ramin Zabih, Daniel Scharstein, Olga Veksler, Vladimir Kolmogorov, Aseem Agarwala, Marshall Tappen, and Carsten Rother. "A Comparative Study of Energy Minimization Methods for Markov Random Fields with Smoothness-Based Priors." *IEEE Trans. Pattern Anal. Mach. Intell.* 30:6 (2008), 1068–1080.

[Szeliski 96] Richard Szeliski. "Video-Mosaics for Virtual Environments." *IEEE Computer Graphics and Applications* 16:2 (1996), 22–30.

[Szeliski 99] Richard Szeliski. "A Multi-View Approach to Motion and Stereo." In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*, 1, p. 163 Vol. 1, 1999.

[Szeliski 06] Richard Szeliski. "Image Alignment and Stitching: A Tutorial." *Foundations and Trends in Computer Graphics and Vision* 2:1 (2006), 1–104.

[Szeliski 11] Richard Szeliski. *Computer Vision: Algorithms and Applications.* Springer, 2011.

[Tagliasacchi et al. 09] A. Tagliasacchi, H. Zhang, and D. Cohen-Or. "Curve Skeleton Extraction from Incomplete Point Cloud." *ACM Transactions on Graphics (Proceedings of SIGGRAPH).*

[Taguchi et al. 08] Yuichi Taguchi, Bennett Wilburn, and C. Lawrence Zitnick. "Stereo Reconstruction with Mixed Pixels Using Adaptive Over-Segmentation." In *Computer Vision and Pattern Recognition*, 2008.

[Tam et al. 13] G. Tam, Z.-Q. Cheng, Y.-K. Lai, F. Langbein, Y. Liu, D. Marshall, R. Martin, X.-F. Sun, and P. Rosin. "Registration of 3D Point Clouds and Meshes: A Survey from Rigid to Non-Rigid." *Transactions on Visualization and Computer Graphics* 19:7 (2013), 1199–1217.

[Tan et al. 04] Kar-Han Tan, Hong Hua, and Narendra Ahuja. "Multiview Panoramic Cameras Using Mirror Pyramids." *IEEE Trans. on Pattern Analysis and Machine Intelligence* 26:7 (2004), 941–945.

[Tauber et al. 07] Z. Tauber, Ze-Nian Li, and M.S. Drew. "Review and Preview: Disocclusion by Inpainting for Image-Based Rendering." *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 37:4 (2007), 527–540.

[Taylor 03] Camillo J. Taylor. "Surface Reconstruction from Feature Based Stereo." In *Proceedings of the Ninth IEEE International Conference on Computer Vision—Volume 2*, pp. 184–192, 2003.

[Teichman et al. 13] Alex Teichman, Stephen Miller, and Sebastian Thrun. "Unsupervised Intrinsic Calibration of Depth Sensors via SLAM." In *Proceedings of Robotics: Science and Systems*. Berlin, Germany, 2013.

[Teichmann and Teller 98] M. Teichmann and S. Teller. "Assisted Articulation of Closed Polygonal Models." *Computer Animation and Simulation*, pp. 87–102.

[Tejera et al. 13] M. Tejera, D. Casas, and A. Hilton. "Animation Control of Surface Motion Capture." *IEEE Transactions on Cybernetics* 43:6 (2013), 1532–1545.

[Tena et al. 11] Jose Rafael Tena, Fernando De la Torre, and Iain Matthews. "Interactive Region-Based Linear 3D Face Models." *ACM Transactions on Graphics (Proc. SIGGRAPH 2011)* 30:4.

[Terzopoulos et al. 87] Demetri Terzopoulos, John Platt, Alan Barr, and Kurt Fleischer. "Elastically Deformable Models." In *Proc. of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '87*, pp. 205–214. New York, NY, USA: ACM, 1987.

[The Foundry 13] The Foundry. "Nuke." 2013. Http://www.thefoundry.co.uk/.

[Theobalt et al. 07] Christian Theobalt, Naveed Ahmed, Hendrik Lensch, Marcus Magnor, and H-P Seidel. "Seeing People in Different Light-Joint Shape, Motion, and Reflectance Capture." *IEEE Transactions on Visualization and Computer Graphics (TVCG)* 13:4 (2007), 663–674.

[Thomas et al. 97] Graham A Thomas, J Jin, T Niblett, and C Urquhart. "A Versatile Camera Position Measurement System for Virtual Reality TV Production."

[Thomas 07] Graham Thomas. "Real-Time Camera Tracking Using Sports Pitch Markings." *Journal of Real-Time Image Processing* 2:2-3 (2007), 117–132.

[Thomaszewski et al. 09] Bernhard Thomaszewski, Simon Pabst, and Wolfgang Straßer. "Continuum-Based Strain Limiting." *Comput. Graph. Forum* 28:2 (2009), 569–576.

[Tocci et al. 11] M. D. Tocci, C. Kiser, N. Tocci, and P. Sen. "A Versatile HDR Video Production System." *ACM Transactions on Graphics (TOG) (Proc. SIGGRAPH)* 30:4.

[Tompkin et al. 13] James Tompkin, Fabrizio Pece, Rajvi Shah, Shahram Izadi, Jan Kautz, and Christian Theobalt. "Video Collections in Panoramic Contexts." In *Proc. of Symposium on User Interface Software and Technology*, pp. 1256–1263. Portland, USA, 2013.

[Triggs et al. 00] Bill Triggs, Philip F. McLauchlan, Richard Hartley, and Andrew W. Fitzgibbon. "Bundle Adjustment Modern Synthesis." In *Vision Algorithms: Theory and Practice*, pp. 298–372. Springer, 2000.

[Tsai 87] R. Y. Tsai. "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses." *IEEE Journal of Robotics and Automation*, pp. 323–344.

[Tunwattanapong et al. 13] Borom Tunwattanapong, Graham Fyffe, Paul Graham, Jay Busch, Xueming Yu, Abhijeet Ghosh, and Paul Debevec. "Acquiring Reflectance and Shape from Continuous Spherical Harmonic Illumination." In *SIGGRAPH*. Anaheim, CA, 2013.

[Tuytelaers and Mikolajczyk 07] Tinne Tuytelaers and Krystian Mikolajczyk. "Local Invariant Feature Detectors: A Survey." *Foundations and Trends in Computer Graphics and Vision* 3:3 (2007), 177–280.

[Urey et al. 11] Hakan Urey, Kishore V Chellappan, Erdem Erden, and Phil Surman. "State of the Art in Stereoscopic and Autostereoscopic Displays." *Proceedings of the IEEE* 99:4 (2011), 540–555.

[Vaish and Adams 12] Vaibhav Vaish and Andrew Adams. "The (New) Stanford Light Field Archive." http://lightfield.stanford.edu/, 2012.

[Vaish et al. 04] V. Vaish, B. Wilburn, N. Joshi, and M. Levoy. "Using Plane + Parallax for Calibrating Dense Camera Arrays." In *Proc. International Conference on Computer Vision and Pattern Recognition*, 2004.

[Valgaerts et al. 12] Levi Valgaerts, Chenglei Wu, Andres Bruhn, Hans-Peter Seidel, and Christian Theobalt. "Lightweight Binocular Facial Performance Capture under Uncontrolled Lighting." *ACM Transactions on Graphics (Proc. SIGGRAPH ASIA 2012)*.

[van den Hengel et al. 07] Anton van den Hengel, Anthony Dick, Thorsten Thormählen, Ben Ward, and Philip H. S. Torr. "VideoTrace: Rapid Interactive Scene Modelling from Video." *ACM Trans. Graph.* 26:3 (2007), 86:1–86:5.

[van Kaick et al. 11] O. van Kaick, H. Zhang, G. Hamarneh, and D. Cohen-Or. "A Survey on Shape Correspondence." *Computer Graphics Forum* 3:6 (2011), 1681–1707.

[Vangorp et al. 11] Peter Vangorp, Gaurav Chaurasia, Pierre-Yves Laffont, Roland Fleming, and George Drettakis. "Perception of Visual Artifacts in Image-Based Rendering of Façades." *Computer Graphics Forum (Proceedings of the Eurographics Symposium on Rendering)* 30:4 (2011), 1241–1250.

[Vangorp et al. 13] Peter Vangorp, Christian Richardt, Emily A. Cooper, Gaurav Chaurasia, Martin S. Banks, and George Drettakis. "Perception of Perspective Distortions in Image-Based Rendering." *ACM Trans. Graph.* 32:4 (2013), 58:1–58:12.

[Varanasi et al. 08] Kiran Varanasi, Andrei Zaharescu, Edmond Boyer, and Radu Horaud. "Temporal Surface Tracking Using Mesh Evolution." In *ECCV*, pp. 30–43, 2008.

[Veach and Guibas 94] Eric Veach and Leonidas Guibas. "Bidirectional Estimators for Light Transport." In *Proc. Eurographics Workshop on Rendering*, pp. 147–162, 1994.

[Veach and Guibas 97] Eric Veach and Leonidas J. Guibas. "Metropolis Light Transport." In *SIGGRAPH '97*, pp. 65–76, 1997.

[Veach 97] Eric Veach. "Robust Monte Carlo Methods for Light Transport Simulation." PhD thesis, Stanford University, 1997.

[Vedaldi and Fulkerson 08] Andrea Vedaldi and Brian Fulkerson. "VLFeat: An Open and Portable Library of Computer Vision Algorithms." http://www.vlfeat.org/, 2008.

[Vedula et al. 05] S. Vedula, S. Baker, and T. Kanade. "Image Based Spatio-Temporal Modeling and View Interpolation of Dynamic Events." *ACM Trans. on Graphics* 24:2 (2005), 240–261.

[Veksler 02] Olga Veksler. "Stereo Correspondence with Compact Windows via Minimum Ratio Cycle." *IEEE Trans. Pattern Anal. Mach. Intell.* 24:12 (2002), 1654–1660.

[Veksler 05] Olga Veksler. "Stereo Correspondence by Dynamic Programming on a Tree." In *Computer Vision and Pattern Recognition*, 2, pp. 384–390, 2005.

[Velten et al. 13] Andreas Velten, Di Wu, Adrian Jarabo, Belen Masia, Christopher Barsi, Chinmaya Joshi, Everett Lawson, Moungi Bawendi, Diego Gutierrez, and Ramesh Raskar. "Femto-Photography: Capturing and Visualizing the Propagation of Light." *ACM Trans. Graph.* 32:4 (2013), 44:1–44:8.

[Vergne et al. 10] Romain Vergne, Romain Pacanowski, Pascal Barla, Xavier Granier, and Christophe Schlick. "Radiance Scaling for Versatile Surface Enhancement." In *Proceedings of the 2010 ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pp. 143–150, 2010.

[Vetro et al. 11] A. Vetro, A.M. Tourapis, K. Muller, and Tao Chen. "3D-TV Content Storage and Transmission." *Broadcasting, IEEE Transactions on* 57:2 (2011), 384–394.

[Vlasic et al. 05] D. Vlasic, M. Brand, H. Pfister, and J. Popović. "Face Transfer with Multi-Linear Models." *ACM Transactions on Graphics (Proceedings SIGGRAPH)*, 2005.

[Vlasic et al. 08] D. Vlasic, I. Baran, W. Matusik, and J. Popović. "Articulated Mesh Animation from Multi-View Silhouettes." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 27:3 (2008), 97:1–97:9.

[Vogiatzis et al. 06] George Vogiatzis, Carlos Hernández, and Roberto Cipolla. "Reconstruction in the Round Using Photometric Normals and Silhouettes." In *CVPR*, pp. 1847–1854, 2006.

[Volino et al. 09] Pascal Volino, Nadia Magnenat-Thalmann, and Francois Faure. "A Simple Approach to Nonlinear Tensile Stiffness for Accurate Cloth Simulation." *ACM Trans. Graph.* 28:4 (2009), 105:1–105:16.

[von der Pahlen et al. 14] Javier von der Pahlen, Jorge Jimenez, Etienne Danvoye, Paul Debevec, Graham Fyffe, and Hao Li. "Digital Ira and Beyond: Creating Photoreal Real-Time Digital Characters." In *SIGGRAPH Course Notes 2014*, 2014.

[Wade 00] L. Wade. "Automated Generation of Control Skeletons for Use in Animation." PhD thesis, Ohio State University, 2000.

[Wang et al. 01] Lifeng Wang, Sing Bing Kang, Richard Szeliski, and Heung-Yeung Shum. "Optimal Texture Map Reconstruction from Multiple Views." In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 1, pp. I–347. IEEE, 2001.

[Wang et al. 07] Huamin Wang, Mingxuan Sun, and Ruigang Yang. "Space-Time Light Field Rendering." *IEEE Transactions on Visualization and Computer Graphics* 13:4 (2007), 697–710.

[Wang et al. 08] Jiaping Wang, Shuang Zhao, Xin Tong, John Snyder, and Baining Guo. "Modeling Anisotropic Surface Reflectance with Example-Based Microfacet Synthesis." In *ACM Transactions on Graphics (TOG)*, 27, 27, p. 41. ACM, 2008.

[Wang et al. 09] Jiaping Wang, Yue Dong, Xin Tong, Zhouchen Lin, and Baining Guo. "Kernel Nyström Method for Light Transport." *ACM Transactions on Graphics (TOG)* 28:3 (2009), 29.

[Wang et al. 10] Huamin Wang, Florian Hecht, Ravi Ramamoorthi, and James O'Brien. "Example-Based Wrinkle Synthesis for Clothing Animation." *ACM Trans. Graph.* 29:4 (2010), 107:1–107:8.

[Wang et al. 11a] Huamin Wang, James F. O'Brien, and Ravi Ramamoorthi. "Data-Driven Elastic Models for Cloth: Modeling and Measurement." *ACM Trans. Graph.* 30:4 (2011), 71:1–71:12.

[Wang et al. 11b] Lili Wang, Kees Teunissen, Yan Tu, Li Chen, Panpan Zhang, Tingting Zhang, and Ingrid Heynderickx. "Crosstalk Evaluation in Stereoscopic Displays." *Display Technology, Journal of* 7:4 (2011), 208–214.

[Wang et al. 11c] O. Wang, M. Lang, M. Frei, A. Hornung, A. Smolic, and M. Gross. "StereoBrush: Interactive 2D to 3D Conversion Using Discontinuous Warps." In *Proceedings of the Eighth Eurographics Symposium on Sketch-Based Interfaces and Modeling, SBIM '11*, pp. 47–54. New York, NY, USA: ACM, 2011.

[Wang et al. 13a] J. Wang, K. Xu, L. Liu, J. Cao, S. Liu, Z. Yu, and X. D. Gu. "Consolidation of Low-Quality Point Clouds from Outdoor Scenes." *Computer Graphics Forum* 32:5 (2013), 207–216.

[Wang et al. 13b] Yangang Wang, Jianyuan Min, Jianjie Zhang, Yebin Liu, Feng Xu, Qionghai Dai, and Jinxiang Chai. "Video-Based Hand Manipulation Capture through Composite Motion Control." *ACM Transactions on Graphics (TOG)* 32:4 (2013), 43.

[Wanner and Goldlücke 12] S. Wanner and B. Goldlücke. "Spatial and Angular Variational Super-Resolution of 4D Light Fields." In *Proc. European Conference on Computer Vision*, 2012.

[Wanner and Goldlücke 13] S. Wanner and B. Goldlücke. "Reconstructing Reflective and Transparent Surfaces from Epipolar Plane Images." In *German Conference on Pattern Recognition (GCPR)*, 2013.

[Wanner and Goldlücke 14] S. Wanner and B. Goldlücke. "Variational Light Field Analysis for Disparity Estimation and Super-Resolution." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36:3 (2014), 606–619.

[Wanner et al. 11] S. Wanner, J. Fehr, and B. Jähne. "Generating EPI Representations of 4D Light Fields with a Single Lens Focused Plenoptic Camera." *Advances in Visual Computing*, pp. 90–101.

[Ward et al. 07] Kelly Ward, Florence Bertails, Tae-Yong Kim, Stephen R. Marschner, Marie-Paule Cani, and Ming Lin. "A Survey on Hair Modelling: Styling, Simulation and Rendering." *IEEE Transactions on Visualization and Computer Graphics TVCG* 13:2.

[Waschbüsch et al. 07] M. Waschbüsch, S. Würmlin, and M. Gross. "3D Video Billboard Clouds." In *Computer Graphics Forum*, 26, 26, pp. 561–569. Wiley Online Library, 2007.

[Watson et al. 00] Benjamin Watson, Alinda Friedman, and Aaron McGaffey. "Using Naming Time to Evaluate Quality Predictors for Model Simplification." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '00*, pp. 113–120. New York, NY, USA: ACM, 2000.

[Weber et al. 07] O. Weber, O. Sorkine, Y. Lipman, and C. Gotsman. "Context-Aware Skeletal Shape Deformation." *Computer Graphics Forum (Proceedings of Eurographics)* 26:3 (2007), 267–273.

[Weber et al. 09] O. Weber, M. Ben-Chan, and C. Gotsman. "Complex Barycentric Coordinates with Applications to Planar Shape Deformations." *Computer Graphics Forum (Proceedings EUROGRAPHICS)* 28:2.

[Wei et al. 12] Xiaolin Wei, Peizhao Zhang, and Jinxiang Chai. "Accurate Realtime Full-Body Motion Capture Using a Single Depth Camera." *ACM Transactions on Graphics (TOG)* 31:6 (2012), 188.

[Weise et al. 09] Thibaut Weise, Hao Li, Luc Van Gool, and Mark Pauly. "Face/Off: Live Facial Puppetry." In *ACM/Eurographics Symposium on Computer Animation*, 2009.

[Weise et al. 11] Thibaut Weise, Sofien Bouaziz, Hao Li, and Mark Pauly. "Realtime Performance Based Facial Animation." *ACM Transactions on Graphics (Proc. SIGGRAPH 2011)* 30:4.

[Weiss et al. 11] A. Weiss, D. Hirshberg, and M. Black. "Home 3D Body Scans from Noisy Image and Range Data." In *International Conference on Computer Vision*, pp. 1951–1958, 2011.

[Weissig et al. 12] Christian Weissig, Oliver Schreer, Peter Eisert, and Peter Kauff. "The Ultimate Immersive Experience: Panoramic 3D Video Acquisition." In *Proc. 18th Int. Conf. on MultiMedia Modelling*, pp. 671–681. Klagenfurt, Austria, 2012.

[Wendland 95] H. Wendland. "Piecewise Polynomial, Positive Definite and Compactly Supported Radial Functions of Minimal Degree." *Adv. Comput. Math.* 4:4 (1995), 389–396.

[Wenger et al. 05] Andreas Wenger, Andrew Gardner, Chris Tchou, Jonas Unger, Tim Hawkins, and Paul Debevec. "Performance Relighting and Reflectance Transformation with Time-Multiplexed Illumination." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 24:3 (2005), 756–764.

[Werlberger et al. 09] Manuel Werlberger, Werner Trobin, Thomas Pock, Andreas Wedel, Daniel Cremers, and Horst Bischof. "Anisotropic Huber-L1 Optical Flow." In *British Machine Vision Conference*, pp. 108.1–108.11, 2009.

[Wetzstein et al. 07] Gordon Wetzstein, Oliver Bimber, et al. "Radiometric Compensation through Inverse Light Transport." In *Pacific Conference on Computer Graphics and Applications*, pp. 391–399, 2007.

[Wetzstein et al. 11] G. Wetzstein, I. Ihrke, D Lanman, and W. Heidrich. "State of the Art in Computational Plenoptic Imaging." In *STAR Proceedings of Eurographics*, 2011.

[Weyrich et al. 08] Tim Weyrich, Jason Lawrence, Hendrik Lensch, Szymon Rusinkiewicz, and Todd Zickler. "Principles of Appearance Acquisition and Representation." *Foundations and Trends in Computer Graphics and Vision* 4:2 (2008), 75–191.

[White and Forsyth 06] Ryan White and David A. Forsyth. "Retexturing Single Views Using Texture and Shading." In *Proc. Europ. Conf. on Computer Vision (ECCV)*, pp. 70–81. Springer, 2006.

[White et al. 07] Ryan White, Keenan Crane, and David A. Forsyth. "Capturing and Animating Occluded Cloth." *ACM Trans. Graph.* 26:3.

[Wietzke 12] Lennart Wietzke. "Raytrix 3D Lightfield Camera Technology." http://www.raytrix.de/, 2012.

[Wilburn et al. 05] Bennett Wilburn, Neel Joshi, Vaibhav Vaish, Eino-Ville Talvala, Emilio Antunez, Adam Barth, Andrew Adams, Mark Horowitz, and Marc Levoy. "High Performance Imaging Using Large Camera Arrays." *ACM Transactions on Graphics* 24 (2005), 765–776.

[Wilson and Snavely 13] Kyle Wilson and Noah Snavely. "Network Principles for SfM: Disambiguating Repeated Structures with Local Context." In *International Conference on Computer Vision*. IEEE, 2013.

[Witkin and Popovic 95] A. Witkin and Z. Popovic. "Motion Warping." In *Proceedings of ACM SIGGRAPH*, pp. 105–108, 1995.

[Wood et al. 00] Daniel N. Wood, Daniel I. Azuma, Ken Aldinger, Brian Curless, Tom Duchamp, David H. Salesin, and Werner Stuetzle. "Surface Light Fields for 3D Photography." In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 287–296. ACM Press/Addison-Wesley Publishing Co., 2000.

[Woodford et al. 09] Oliver Woodford, Philip Torr, Ian Reid, and Andrew Fitzgibbon. "Global Stereo Reconstruction under Second-Order Smoothness Priors." *IEEE Trans. Pattern Anal. Mach. Intell.* 31:12 (2009), 2115–2128.

[Woodham 80] R. J. Woodham. "Photometric Method for Determining Surface Orientation from Multiple Images." In *Optical Engineering*, 19(1), 19(1), pp. 139–144, 1980.

[Wu et al. 11] Changchang Wu, Sameer Agarwal, Brian Curless, and Steven M. Seitz. "Multicore Bundle Adjustment." In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 3057–3064. IEEE, 2011.

[Wu et al. 12a] Chenglei Wu, Kiran Varanasi, and Christian Theobalt. "Full Body Performance Capture under Uncontrolled and Varying Illumination: A Shading-Based Approach." In *Computer Vision–ECCV 2012*, pp. 757–770. Springer, 2012.

[Wu et al. 12b] Di Wu, M. O'Toole, A. Velten, A. Agrawal, and R. Raskar. "Decomposing Global Light Transport Using Time of Flight Imaging." In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pp. 366–373, 2012.

[Wu et al. 13] Chenglei Wu, Carsten Stoll, Levi Valgaerts, and Christian Theobalt. "On-Set Performance Capture of Multiple Actors with a Stereo Camera." *ACM Trans. Graph. (Proc. of SIGGRAPH Asia)* 32:6.

[Wu 11] Changchang Wu. "VisualSFM: A Visual Structure from Motion System." http://ccwu.me/vsfm/, 2011.

[Wuhrer and Shu 13] S. Wuhrer and C. Shu. "Estimating 3D Human Shapes from Measurements." *Machine Vision and Applications* 24:6 (2013), 1133–1147.

[Wuhrer et al. 10] S. Wuhrer, Z. Ben Azouz, and C. Shu. "Semi-Automatic Prediction of Landmarks on Human Models in Varying Poses." In *Canadian Conference on Computer and Robot Vision*, pp. 136–142, 2010.

[Wuhrer et al. 11] S. Wuhrer, C. Shu, and P. Xi. "Landmark-Free Posture Invariant Human Shape Correspondence." *The Visual Computer* 27:9 (2011), 843–852.

[Wuhrer et al. 12] S. Wuhrer, C. Shu, and P. Xi. "Posture-Invariant Statistical Shape Analysis Using Laplace Operator." *Computers & Graphics* 36:5 (2012), 410–416. *Proceedings of Shape Modeling International.*

[Xiao et al. 06] Jiangjian Xiao, Hui Cheng, Harpreet S. Sawhney, Cen Rao, and Michael A. Isnardi. "Bilateral Filtering-Based Optical Flow Estimation with Occlusion Detection." In *European Conference on Computer Vision*, 3951, 3951, pp. 211–224, 2006.

[Xu et al. 98] Z. Xu, R. Schwarte, H. Heinol, B. Buxbaum, and T. Ringbeck. "Smart Pixel—Photonic Mixer Device (PMD)." In *Proc. Int. Conf. on Mechatron. & Machine Vision*, pp. 259–264, 1998.

[Xu et al. 11] F. Xu, Y. Liu, C. Stoll, J. Tompkin, G. Bharaj, Q. Dai, H.-P. Seidel, J. Kautz, and C. Theobalt. "Video-Based Characters—Creating New Human Performances from a Multi-View Video Database." *Proc. ACM SIGGRAPH*, 2011.

[Y. Sahin and Backhaus 14] S. Spielmann Y. Sahin and M. Backhaus. "Dark Matter—A Tale of Virtual Production." *In ACM SIGGRAPH 2014 Talks (SIGGRAPH '14)*, ACM. New York, NY, USA, Article 25, 1 page DOI=10.1145/2614106,2614195 http://doi.acm.org/1145/2614106.2614195.

[Yahav et al. 07] G. Yahav, G. J. Iddan, and D. Mandelbaum. "3D Imaging Camera for Gaming Application." In *Digest of Technical Papers of Int. Conf. on Consumer Electronics*, 2007. DOI: 10.1109/ICCE.2007.341537.

[Yamanoue 06] H. Yamanoue. "The Differences between Toed-in Camera Configurations and Parallel Camera Configurations in Shooting Stereoscopic Images." *Multimedia and Expo, IEEE International Conference on* 0 (2006), 1701–1704.

[Yamauchi et al. 05] H. Yamauchi, S. Gumhold, R. Zayer, and Seidel H.-P. "Mesh Segmentation Driven by Gaussian Curvature." *The Visual Computer* 21 (2005), 649–658.

[Yamazaki et al. 02] Shuntaro Yamazaki, Ryusuke Sagawa, Hiroshi Kawasaki, Katsushi Ikeuchi, and Masao Sakauchi. "Microfacet Billboarding." In *Proceedings of the 13th Eurographics Workshop on Rendering, EGRW '02*, pp. 169–180. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2002.

[Yamazoe et al. 12] H. Yamazoe, H. Habe, I. Mitsugami, and Y. Yagi. "Easy Depth Sensor Calibration." In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pp. 465–468, 2012.

[Yang and Pollefeys 03] Ruigang Yang and Marc Pollefeys. "Multi-Resolution Real-Time Stereo on Commodity Graphics Hardware." In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, 1, pp. I–211. IEEE, 2003.

[Yang et al. 93] Yibing Yang, Alan Yuille, and Jie Lu. "Local, Global, and Multilevel Stereo Matching." In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR '93*, pp. 274–279. IEEE, 1993.

[Yang et al. 06a] Qingxiong Yang, Liang Wang, Ruigang Yang, Henrik Stewénius, and David Nistér. "Stereo Matching with Color-Weighted Correlation, Hierarchical Belief Propagation and Occlusion Handling." In *Computer Vision and Pattern Recognition*, pp. 2347–2354, 2006.

[Yang et al. 06b] Qingxiong Yang, Liang Wang, Ruigang Yang, Shengnan Wang, Miao Liao, and David Nistér. "Real-Time Global Stereo Matching Using Hierarchical Belief Propagation." In *British Machine Vision Conference*, 2006.

[Yang et al. 06c] R. Yang, D. Guinnip, and L. Wang. "View-Dependent Textured Splatting." *The Visual Computer* 22:7 (2006), 456–467.

[Yang et al. 10] Qingxiong Yang, Kar-Han Tan, W. Bruce Culbertson, and John G. Apostolopoulos. "Fusion of Active and Passive Sensors for Fast 3D Capture." In *MMSP*, pp. 69–74, 2010.

[Yaraş et al. 10] Fahri Yaraş, Hoonjong Kang, and Levent Onural. "State of the Art in Holographic Displays: A Survey." *Journal of Display Technology* 6:10 (2010), 443–454.

[Ye et al. 11] Mao Ye, Xianwang Wang, Ruigang Yang, Liu Ren, and Marc Pollefeys. "Accurate 3D Pose Estimation from a Single Depth Image." In *Proceedings of the 2011 International Conference on Computer Vision, ICCV '11*, pp. 731–738. Washington, DC, USA: IEEE Computer Society, 2011.

[Ye et al. 12] Genzhi Ye, Yebin Liu, Nils Hasler, Xiangyang Ji, Qionghai Dai, and Christian Theobalt. "Performance Capture of Interacting Characters with Handheld Kinects." In *ECCV*, pp. 828–841, 2012.

[Yoon and Kweon 05] Kuk-Jin Yoon and In-So Kweon. "Locally Adaptive Support-Weight Approach for Visual Correspondence Search." In *Computer Vision and Pattern Recognition*, pp. 924–931, 2005.

[Yuksel et al. 12] Cem Yuksel, Jonathan M. Kaldor, Doug L. James, and Steve Marschner. "Stitch Meshes for Modeling Knitted Clothing with Yarn-Level Detail." *ACM Trans. Graph.* 31:4 (2012), 37:1–37:12.

[Zabih and Woodfill 94] Ramin Zabih and John Woodfill. "Non-Parametric Local Transforms for Computing Visual Correspondence." In *Proceedings of the Third European Conference on Computer Vision (Vol. II), ECCV '94*, pp. 151–158. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1994.

[Zach et al. 07] C. Zach, T. Pock, and H. Bischof. "A Duality Based Approach for Realtime TV-L 1 Optical Flow." In *Proc. German Conference on Pattern Recognition (DAGM)*, pp. 214–223. Springer, 2007.

[Zach et al. 08] Christopher Zach, David Gallup, Jan-Michael Frahm, and Marc Niethammer. "Fast Global Labeling for Real-Time Stereo Using Multiple Plane Sweeps." In *Vision, Modeling and Visualization*, pp. 243–252, 2008.

[Zalevsky et al. 07] Zeev Zalevsky, Alexander Shput, Aviad Maizels, and Javier Garcia. "Method and System for Object Reconstruction." 2007.

[Zhang and Huang 04] Song Zhang and Peisen Huang. "High-Resolution, Real-Time 3D Shape Acquisition." In *Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04)*, 3, 3, pp. 28–36. Washington, DC, USA: IEEE Computer Society, 2004.

[Zhang and Zhang 11] Cha Zhang and Zhengyou Zhang. "Calibration between Depth and Color Sensors for Commodity Depth Cameras." In *Proceedings of the 2011 IEEE International Conference on Multimedia and Expo, ICME '11*, pp. 1–6. Washington, DC, USA: IEEE Computer Society, 2011.

[Zhang et al. 02] Li Zhang, Brian Curless, and Steven M. Seitz. "Rapid Shape Acquisition Using Color Structured Light and Multi-Pass Dynamic Programming." In *The 1st IEEE International Symposium on 3D Data Processing, Visualization, and Transmission*, pp. 24–36, 2002.

[Zhang et al. 06] Qingshan Zhang, Zicheng Liu, Baining Guo, and Harry Shum. "Geometry-Driven Photorealistic Expression Synthesis." *IEEE Trans. Vis. Comput. Graphics* 12:1 (2006), 48–60.

[Zhang et al. 08] Yilei Zhang, Minglun Gong, and Yee-Hong Yang. "Local Stereo Matching with 3D Adaptive Cost Aggregation for Slanted Surface Modeling and Sub-Pixel Accuracy." In *International Conference on Pattern Recognition*, pp. 1–4, 2008.

[Zhang et al. 10a] Ke Zhang, Gauthier Lafruit, Rudy Lauwereins, and Luc J. Van Gool. "Joint Integral Histograms and Its Application in Stereo Matching." In *International Conference on Image Processing*, pp. 817–820, 2010.

[Zhang et al. 10b] Song Zhang, Daniel Van Der Weide, and James Oliver. "Superfast Phase-Shifting Method for 3-D Shape Measurement." *Optics Express* 18:9 (2010), 9684–9689.

[Zhang et al. 11] Lei Zhang, Xiaolin Wu, Antoni Buades, and Xin Li. "Color Demosaicking by Local Directional Interpolation and Nonlocal Adaptive Thresholding." *Journal of Electronic Imaging* 20:2.

[Zhang et al. 12] Qing Zhang, Mao Ye, Ruigang Yang, Yasuyuki Matsushita, Bennett Wilburn, and Huimin Yu. "Edge-Preserving Photometric Stereo via Depth Fusion." In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), CVPR '12*, pp. 2472–2479, 2012.

[Zhang 00] Z. Zhang. "A Flexible New Technique for Camera Calibration." *IEEE Transactions on Patterns Analysis and Machine Intelligence* 22:11 (2000), 1330–1334.

[Zhao et al. 11] Shuang Zhao, Wenzel Jakob, Steve Marschner, and Kavita Bala. "Building Volumetric Appearance Models of Fabric Using Micro CT Imaging." In *ACM Transactions on Graphics (TOG)*, 30, 30, p. 44. ACM, 2011.

[Zheng et al. 09] K.C. Zheng, A. Colburn, A. Agarwala, M. Agrawala, D. Salesin, B. Curless, and M.F. Cohen. "Parallax Photography: Creating 3D Cinematic Effects from Stills." In *Proceedings of Graphics Interface 2009*, pp. 111–118. Canadian Information Processing Society, 2009.

[Zhou et al. 10] S. Zhou, H. Fu, L. Liu, D. Cohen-Or, and X. Han. "Parametric Reshaping of Human Bodies in Images." *ACM Transactions on Graphics* 29 (2010), 126:1–10. Proceedings of SIGGRAPH.

[Zhou et al. 12] Zhenglong Zhou, Bo Shu, Shaojie Zhuo, Xiaoming Deng, Ping Tan, and Stephen Lin. "Image-Based Clothes Animation for Virtual Fitting." In *ACM SIGGRAPH Asia 2012 Technical Briefs*, pp. 33:1–33:4. ACM, 2012.

[Zhu et al. 11] Jiejie Zhu, Liang Wang, Ruigang Yang, James E. Davis, and Zhigeng Pan. "Reliability Fusion of Time-of-Flight Depth and Stereo Geometry for High Quality Depth Maps." *IEEE Trans. Pattern Anal. Mach. Intell.* 33:7 (2011), 1400–1414.

[Zilly et al. 10] Frederik Zilly, M Muller, Peter Eisert, and Peter Kauff. "The Stereoscopic Analyzer—An Image-Based Assistance Tool for Stereo Shooting and 3D Production." In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pp. 4029–4032. IEEE, 2010.

[Zimmer et al. 11] Henning Zimmer, Andrés Bruhn, and Joachim Weickert. "Optic Flow in Harmony." *Int. J. Comput. Vision* 93:3 (2011), 368–388.

[Zitnick et al. 04] C Lawrence Zitnick, Sing Bing Kang, Matthew Uyttendaele, Simon Winder, and Richard Szeliski. "High-Quality Video View Interpolation Using a Layered Representation." *ACM Transactions on Graphics (Proc. SIGGRAPH)* 23:3 (2004), 600–608.

[Zongker et al. 99] Douglas E. Zongker, Dawn M. Werner, Brian Curless, and David H. Salesin. "Environment Matting and Compositing." In *SIGGRAPH '99: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 205–214. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1999.

[Zurdo et al. 13] Javier S. Zurdo, Juan P. Brito, and Miguel A. Otaduy. "Animating Wrinkles by Example on Non-Skinned Cloth." *IEEE Trans. on Visualization and Computer Graphics* 19:1 (2013), 149–158.

Index