Computational Aesthetics in Graphics, Visualization and Imaging (2005)
L. Neumann, M. Sbert, B. Gooch, W. Purgathofer (Editors)

Computer Generated Stereoscopic Artwork

E. Stavrakis and M. Gelautz

Interactive Media Systems Group, Technical University of Vienna, Austria

Abstract

The focus of this work is to investigate and simulate artistic techniques in stereoscopy that go beyond stereo photography, such as stereoscopic painting. We briefly discuss the topic from a stereo artist's perspective and map some of our observations from traditional stereo techniques to the scientific domain, where we use them to tackle technical tasks involved in the generation of stereo artwork. We describe a framework that we use in our stereoscopic image-based non-photorealistic rendering algorithms, but it can be adopted by other single-view artistic image synthesis techniques in order to generate stereoscopic output.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation I.4.8 [Image Processing and Computer Vision]: Scene Analysis J.5 [Arts Humanities]: Fine Arts

1. Introduction

Pictorial visual artwork is naturally created and presented to the audience on a 2D flat surface. While the use of monocular cues forms a strong toolset for the artist to communicate three-dimensional spatial relationships attempting to overcome the flatness of the medium [Dur02], there is a fundamental discrepancy in this process. The projection of a 3D world onto a 2D surface irreversibly discards critical information from the scene that is otherwise used by our human visual system to better appreciate depth and distances. When interacting with our natural immediate environment, our binocular visual system utilizes two strong depth cues, that of binocular parallax (horizontal eye disparity) and convergence (inwards rotation of the eyes). The two dissimilar retinal images, together with the angle between the two converging lines of eye sight, enable the perceptual mechanisms of our visual system to robustly estimate depth and distances. This mental process of perceiving 3D information by the use of the binocular depth cues is known as stereopsis. By reintroducing the two binocular cues, using two flat stereoscopically correct pictures, an artist may better communicate three dimensional information.

While stereo photography flourished since the discovery of stereoscopy, other visual artforms (i.e. drawing, painting) have fewer examples to present. Nevertheless, similarly to stereo photography, stereo hand-made artwork is composed by two components, one for each of the spectator's eyes. The difference between photography and other 2D visual arts is that the former shifts the technical aspects of stereoscopy onto the photographic device itself, while hand-made stereo techniques require the artist to comprehend and utilize them throughout the creative process.

Early traces of stereo artwork go as far back as the discovery of stereoscopy itself, when stereoscopic drawings were created for idea communication, rather than focusing on aesthetics. The majority of modern visual artists has neglected the use of the binocular cues and disregarded the stereoscopic nature of the human visual system. Only a limited number of artists have experimented with this medium in the form of drawings and paintings, but stereo has not become a widespread technique among artists. The technical details surrounding stereoscopic content generation, as well as the inability of the audience to perceive and appreciate stereo art without the use of viewing devices, have been the main obstacles in its evolution.

While the process of creating stereo artwork remains mostly undocumented to date, recent articles [Bie05,SG05a] show that it is a rather technical activity of which parts could be simulated by using modern computational approaches. We could think of stereoscopic artwork creation as a process where the artist projects his envisioned three-dimensional artwork onto the working 2D surface from two horizontally displaced viewpoints. The two resulting components must be stereoscopically consistent, so that features of the one component can be perceptually mapped onto the second to enable stereoscopic fusion. Apart from the spatial stereoscopic consistency of the depicted objects, their chromaticity should remain consistent across the two components. It is important to point out that the artist may not fully visualise the complete geometry of a scene, be it real or illusionary, but rather work with a pair of projections, ensuring that the projections themselves are stereoscopically correct.

Artists who have created stereo paintings have sometimes used stereoscopic photographs to set a foundation for their pieces (e.g. Salvador Dalí), which allowed them to manage the feature correspondence between their dual compositions. The use of photographs or foundation sketches of stereo artwork usually forms the basis of creating large scale pieces of stereo art (i.e. stereo paintings). These methods assist the artist to preserve the main relationships between objects across the two stereo views. Nevertheless, the cumbersome task of executing the artwork twice by hand, even when using a digital painting studio through a computer interface, retains the time consuming properties of stereo art creation.

Stereo consistency, viewpoint preservation, time and effort requirements, are technical issues that a visual artist has to undertake when creating stereo artwork. In our work, a mapping of these technical aspects of stereo from the artistic to the scientific domain restricts the scope of using technology into tackling technical problems, rather than attempting to interfere with aesthetic decisions. We used the idea of photographically-assisted artwork to design and implement image-based algorithms that generate stereo artwork and gained insight on how the stereoscopic nature of our process affects the design of those algorithms.

2. Background

Advances in non-photorealistic rendering (NPR) enable computers to simulate natural media and artistic processes. Most of these algorithms produce a single image that appears hand-crafted by processing geometry or images from single or multiple viewpoints. In contrast, our aim is to produce stereoscopically consistent image pairs that are artistically augmented. A key difficulty in producing artistic renderings of real stereo pairs is that many image synthesis algorithms that imitate the artist's hand inject irregularity by using stochastic algorithms and perturbation functions within their processing cycle. If the two images are processed separately, randomness introduces inconsistencies across the two views of a stereo pair, which cause discomfort when viewed stereoscopically. Another challenge in stereoscopic NPR is the preservation of artistic elements on their respective surfaces. Algorithms operating on a single image usually do not have a description of object structure or interpositioning within a depicted scene.

Seitz et al. [SK98] presented a technique for multiple image painting, based on the exploitation of the plenoptic function. Plenoptic image editing allows the manipulation of multiple images simultaneously, but it requires many images of the scene from different viewpoints in order to establish a good set of correspondences between the images. Instead, we use stereo matching to establish a dense per-pixel correspondence between the two views of a single stereo image pair. Stereo matching algorithms [SS02] come in many different types of which the output may be useful within the scope of a particular application. Some of these algorithms handle efficiently untextured regions, or attempt to accurately reconstruct disparity values of occluded regions. These computed disparity maps can be converted to depth maps, but contrary to computer graphics depth-from-raytracing approaches that rely on available 3D models, computer vision depth-from-stereo algorithms produce results that lack the fidelity and reliability of the former. Instead of requiring raytraced-like depth maps in our work, we introduce methods to compensate for the depth-from-stereo inaccuracies according to our requirements.

In NPR a variety of techniques exist that can be used to process individually images of a stereo pair. Usually, these algorithms operate locally on a small pixel neighbourhood and apply simple stroke models, edge detection and colour filtering. They can generate stylized stereo pairs by processing individually the two images, however they lack high level description of scene structure and therefore they cannot separate objects and treat them individually; neither can they simulate artistic media which require surface descriptions and elaborate stroke models.

Available view independent rendering techniques based on 3D scenes, such as [ZISS04], can readily provide stereo artistic output since the artistic primitives used are attached and remain on the 3D model's surface, independent of the location of the viewing point. This, however, becomes particularly challenging in our research, since image-based rendering techniques lack a robust geometrical description of the underlying scene structure.

3. A Stereoscopic NPR Framework

Through our research we have identified problems and limitations that monoscopic NPR algorithms may face when presented with the task of processing stereoscopic inputs. We identify here a set of problems and propose strategies that may be adopted by single view NPR algorithms in order to produce stereoscopic artistic imagery.

Consistency: The problem of stereo consistency is very similar to that of temporal coherence between successive frames. An important difference, however, lies on the fact that artifacts in stereoscopic content, perceived as "ghosting", usually cause eye fatigue, whilst temporal inconsistency has a less severe physical effect to the viewers. Additionally, inputs to image-based algorithms lack the robustness of geometric input information that a 3D scene would equivalently provide. We argue that a relationship should be established between the two stereo views. Establishing a correspondence between projected image points is prone to error and is a subject of research in computer vision referred to as stereo matching.

Depth-from-stereo: The extraction of a disparity map by using stereo matching, and successively depth-from-stereo estimation, may suffer from artifacts produced by stereo matching algorithms, such as inaccurate depth discontinuities' localization or falsely estimated disparities. State of the art object detection and silhouette extraction algorithms from stereo images [GM04] may be used to tackle such problems.

Propagation: Since the stereo input pair describes the same scene depicted from two slightly displaced viewpoints, a large amount of information is shared between the two images. Detecting and using these correspondences to propagate style between the two views helps to a) improve the quality of consistency and b) improve the performance of the algorithms.

Occlusions: Since the two images are viewed from horizontally displaced viewpoints not all points have a correspondence from one view to another. These occluded regions, visible only to the one viewpoint, do not have a counterpart in the other view. However, they have to be stylized similarly to their surroundings, within the same view, so that they blend seamlessly with the rest of the artistic image. One approach is to use the neighbouring style to fill in these areas or directly render them.

Randomness: When using stochastic or perturbation functions in artistic stereo image synthesis the two views must remain perceptually consistent. Randomness should be introduced once across a stereo pair, instead of each component individually. Alternatively, the stereo correspondence may be used to propagate style to produce a consistent effect across a given stereo pair.

Paint Spilling: The effect of stylizing elements extending over multiple surfaces that lie at different depths, even though invisible on singe-view techniques, becomes very noticeable through stereo vision. Two separate objects that are interpositioned may be erroneously processed as one if their colour is similar and reversibly a single object with texture variations on its single surface may be treated as two objects. Using the depth discontinuities provided by the depth maps extracted, we can restrict algorithms to process each surface individually and thus prevent the undesired effect of paint spilling.

4. Algorithms

We have used the framework described in the previous section as a set of guidelines to devise and implement image-based stereoscopic non-photorealistic rendering algorithms.

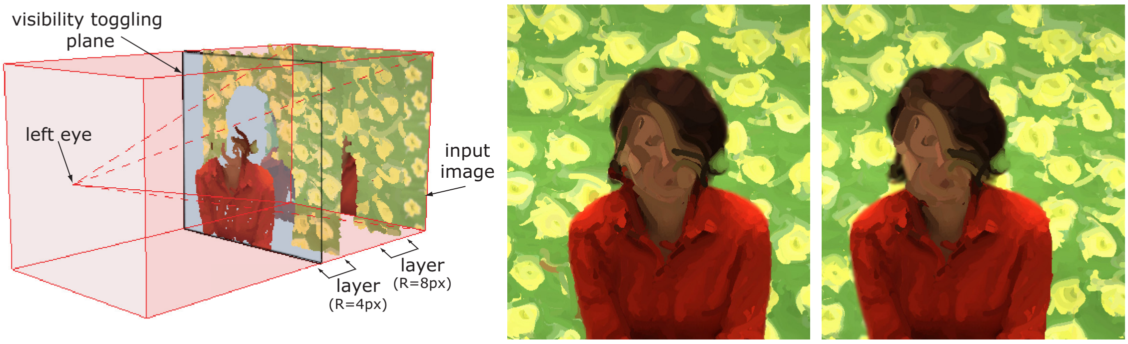
In our initial algorithm [SG04] we have analysed and simulated stereoscopic painting. We based our stereoscopic

Figure 1: *Girl data set. Top row: original stereo pair, bottom row: painted stereo pair.*

painterly algorithm on Hertzmann's [Her98] single-view painting technique, to create a prototype that simulates the artist's hand. A dense disparity map is calculated for a given stereo image pair. The algorithm paints the left image of the input stereo pair to produce an artistic version of the original image. Occluded regions in the second view are detected and specifically treated to compensate for inaccuracies of the stereo matching process. To paint these areas we directly render them in the second view. Finally, to complete the partially painted second view we warp the bitmap describing the painted left view atop the second view, using the disparity map. An example result from this work is shown in Figure 1.

A similar algorithm, using stroke-based warping instead of bitmap warping, has been presented in [SG05b]. Instead of pixel by pixel warping, we project the control points of the brush strokes that compose the left artistic view to the right. This algorithm has the advantage over the initial of having robust corresponding understructure at a brush stroke level. While the first algorithm should refine enough the painting to produce a consistent final layer and would sufficiently tackle scene occlusions, this second algorithm maintains stereo consistency throughout all layers of the progressive refinement steps of the painting process itself. The occlusions are tackled in a similar way, by applying direct rendering in the occluded regions, but these new strokes are warped back to the reference view for consistency. A schematic of this algorithm can be seen in Figure 3.

On top of the latter method, we built a novel interactive



(a) (b) (c)

Figure 2: *A "sliced" stereo painting with two paint layers. (a) A cross-section of a stereo painting is taken by using the visibility toggling plane. The paint layers are rendered in a coarse-to-fine and back-to-front fashion. (b) Left and (c) right component of the sliced painting as the user views it from an orthographic projection.*

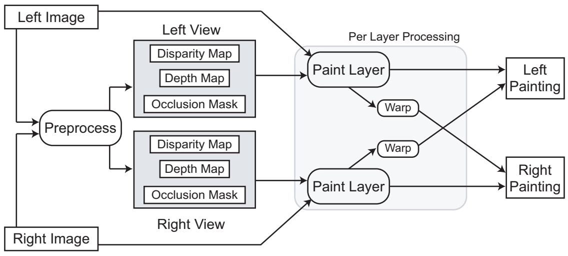


Figure 3: *A schematic of our stereo painting algorithm. Note that the occlusion mask of the right view forces the painting algorithm to operate only in the occluded areas, significantly reducing computations.*

system, also presented in [SG05b], that a potential spectator of the stereo results of this automated process can use to examine and explore the multi-layered stereo paintings. The interface allows a user to slice through the painting by using a stroke visibility toggling plane and examine how the painting has been progressively created, while maintaining the stereoscopic properties of it. The system renders the painting in a back-to-front and coarse-to-fine manner, with attention to handle intersurface paint spilling using depth discontinuities. An example slice of the left view from this system can be seen in Figure 2(a), an equivalent consistent painting exists for the right eye and the stereo painting slices as seen by the user are shown in Figures 2(b) and 2(c). A stereo pair which has been painted with our stereo algorithm is compared with the stereo painting produced by painting individually the two components in Figure 5.

Furthermore, we have recently used purely computer vision techniques to stylize stereo pairs using colour image segmentation [SBMG05]. A stereo matching algorithm is used that provides a layer description of the real scene objects. This information is used to consistently stylize the colour of each layer and transfer style between the stereo components. In this work we also exploit the potential of silhouette edge rendering in stereo images, to depict important objects in the scene. The silhouette edges are partial depth discontinuity edges and are view-dependent features. These silhouettes are frontoparallel to the viewer and can substantially assist the viewer to perceive scene depth by reducing the effort of fusing the stereo images. An example stereo pair is shown in Figure 4.

The results from this method follow the principles initially identified through our experiments with the painterly approach. Even though these algorithms significantly differ, the stereo NPR framework could be successfully applied on them. Therefore it may be possible for other artistic single-view algorithms to tackle the problems in image-based stereoscopic NPR by using this proposed framework.

5. Discussion

Stereo can be perceived fundamentally by forcing each eye into seeing only one of the two stereo images and this has been well assisted by a multitude of stereo viewing devices (e.g. stereoscopes, polarized glasses, etc.). However, it is well known that humans can train their eyes to readily see stereo images by crossing their eyes (cross-eye viewing) or looking at the images in parallel (parallel viewing). The former approach, which involves assisting devices, limits the viewers since all such devices are obtrusive and may be hard to adapt into a gallery or art studio environment. Free-vision remains a skill that is not widely exercised and it is unlikely that it will become popular. A viable solution for stereo artwork presentation and viewing is autostereoscopic displays. Their non-obtrusive design makes them very attractive for a real stereo art presentation scenario.

Stereoscopic art, both hand-made and digitally generated,

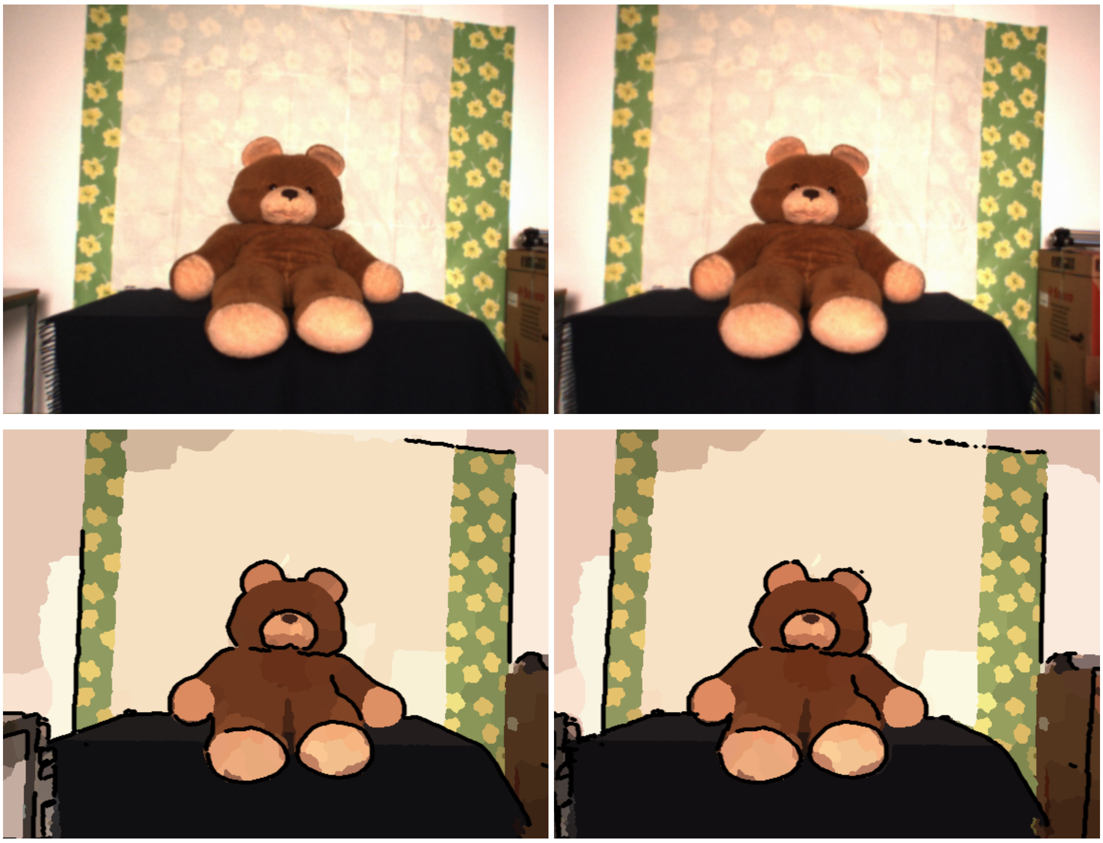


Figure 4: *Teddy data set. Top row: original stereo pair, bottom row: stylized stereo pair.*

is peculiar in that it allows the artwork to detach from the pictorial medium. Relief can be seen on flat surfaces, a property that other 2D pictorial techniques cannot convincingly reproduce. Stereoscopic artwork provides the impression of depth and lively engages the spectator into a journey of spatial diversity. Objects and surfaces detach from the canvas or digital display, and float in 3D space. Combined with monocular depth cues, vastly exploited by both artists and computer visualisation methods, it allows one to experience more naturally a dimension that is not present on any of the two images of a stereo pair.

Computer generated stereoscopic artwork still lacks artistic and creative intention, just like many other artistic rendering techniques. Drawing knowledge from traditional visual artists and mapping technical tasks into the scientific domain, we have provided an approach to the simulation of stereoscopic aesthetics, but there are many open topics that require the attention of the scientific community. Stereoscopic hand-made artwork generation has very scarce documentation and a relatively small number of example works. One open topic is the use of colour tones in stereo artwork. While in our work we assume that the colouration must remain consistent across a stereo pair, there are example painted pieces of which the two stereo components exhibit different chromaticity. Optical blending and colour mixing is possible through stereo vision. Two significantly different components of a stereo pair, when fused, a third view becomes apparent that is coloured and shaded differently than either of the two original components. The perception of this effect is similar to the effect of "ghosting", but the tonal and colour value boundaries within which a pleasing stereo image may be produced need to be identified.

Other challenges remain the automatic generation of stereo images by loosely defined feature correspondences between two views. In our work we use dense disparity data and do a one-to-one mapping between the two views, however artwork is usually perceived on a global level and not through the local structural description of it (i.e. via observation of all brush strokes of a painting). Thus it might be possible to combine object extraction and image understanding algorithms to identify and apply artistic algorithms that operate on a higher level.

Like temporal coherence on real video sequences remains an open and active topic of research, accurate stereo correspondence requires similar attention. The effect of inconsistencies on stereo imaging causes severe discomfort to the viewer. Qualitative measures and psychophysical studies may assist artistic algorithm designers to understand and tune their methods to produce pleasing images.

6. Conclusion

We have started by discussing a special form of traditional visual arts, that of stereoscopic hand-made art. By observing the artist we have identified key areas where computer algorithms can assist artists to expand and experiment with the medium. Through our experience with the designed and implemented artistic stereo algorithms we have proposed a framework that other image-based algorithms may adopt in order to provide stereoscopically consistent artistic imagery. It is our main intention to further extend our research to make the task of creating stereoscopic artwork on a computer more user friendly by taking up the repetitive and cumbersome tasks, while letting the user take the aesthetic decisions for the stereoscopic composition. Non-photorealistic rendering has been mostly focused on simulating widespread artistic techniques and media. Simulations are one area were computer graphics software and hardware excel over other alternatives. We believe that equally well computer systems can enable stereo artists to further explore new media by introducing novel interfaces that can transparently assist them with technical tasks.

7. Acknowledgements

This work was supported by the Austrian Science Fund (FWF) under project P15663. We would like to thank fine artist Roger Ferragallo for sharing with us his experience on stereoscopic painting.

References

[Bie05] BIEGON G.: Stereoscopic synergy: Twin-relief sculpture and painting. *Leonardo 38*, 2 (2005), 92–100.

[Dur02] DURAND F.: Limitations of the medium and pictorial techniques. *Perceptual and Artistic Principles for Effective Computer Depiction, Course Notes for ACM SIGGRAPH 2002* (July 2002), pp. 27–45.

[GM04] GELAUTZ M., MARKOVIC D.: Recognition of object contours from stereo images: an edge combination approach. In *Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission* (2004), pp. 774–780.

[Her98] HERTZMANN A.: Painterly rendering with curved brush strokes of multiple sizes. In *Proc. SIGGRAPH '98* (1998), ACM Press, pp. 453–460.

[SBMG05] STAVRAKIS E., BLEYER M., MARKOVIC D., GELAUTZ M.: Image-based stereoscopic stylization. In *IEEE Internation Conference on Image Processing (to appear)* (September 2005).

[SG04] STAVRAKIS E., GELAUTZ M.: Image-based stereoscopic painterly rendering. In *Rendering Techniques 2004 (Proceedings of Eurographics Symposium on Rendering)* (Norrköping, Sweden, June 2004), pp. 53–60.

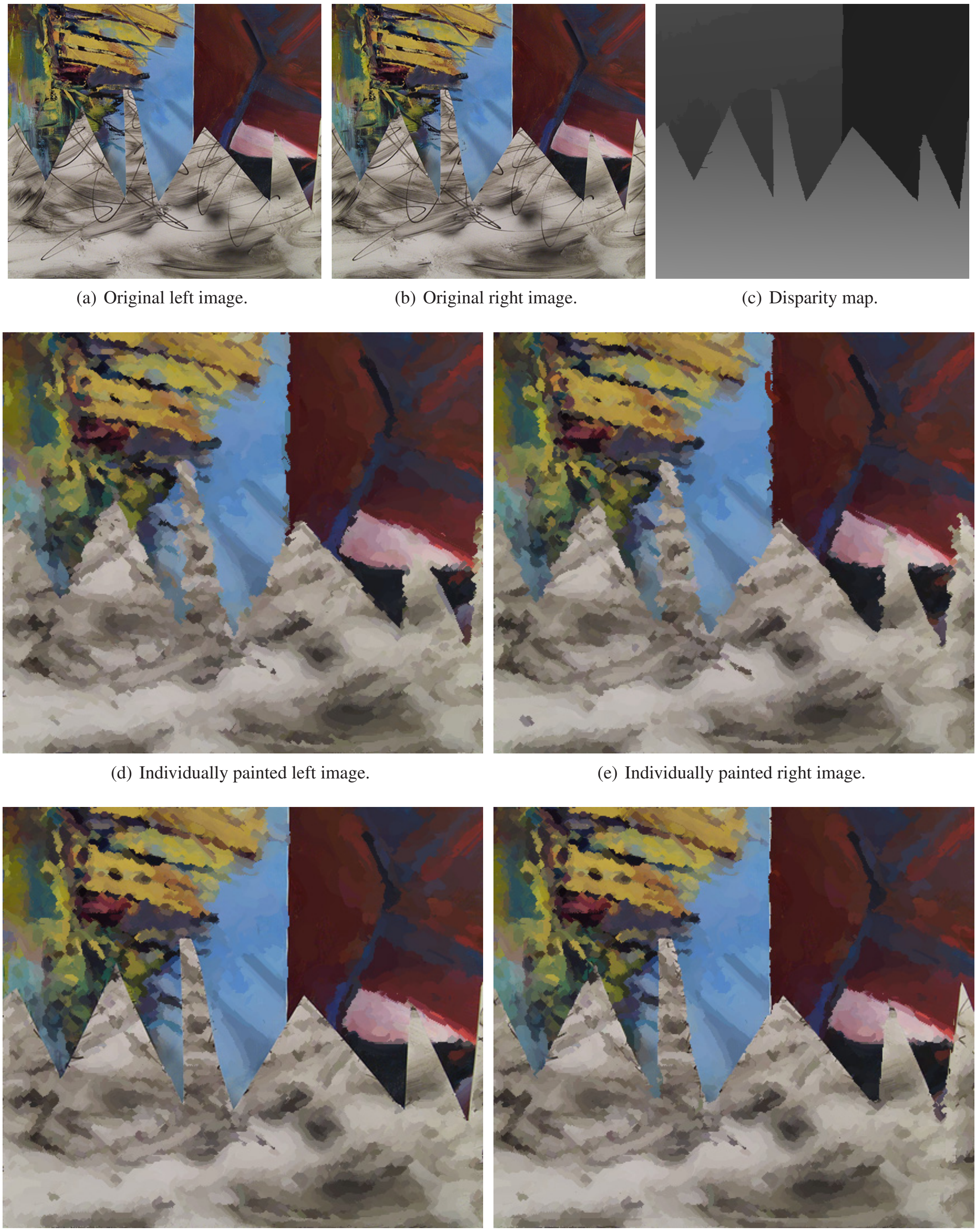
[SG05a] STAVRAKIS E., GELAUTZ M.: Stereo painting: Pleasing the third eye. *Journal of 3D Imaging*, 168 (Spring 2005), 20–23.

[SG05b] STAVRAKIS E., GELAUTZ M.: Stereoscopic painting with varying levels of detail. In *Proceedings of SPIE - Stereoscopic Displays and Virtual Reality Systems XII* (San Jose, California, USA, 2005), vol. 5564.

[SK98] SEITZ S. M., KUTULAKOS K. N.: Plenoptic image editing. In *Proceedings of the Sixth International Conference on Computer Vision* (1998), IEEE Computer Society, pp. 17–24.

[SS02] SCHARSTEIN D., SZELISKI R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision 22*, 1/2/3 (April-June 2002), 7–47.

[ZISS04] ZANDER J., ISENBERG T., SCHLECHTWEG S., STROTHOTTE T.: High quality hatching. *Computer Graphics Forum 23*, 3 (September 2004).



(a) Original left image. (b) Original right image. (c) Disparity map.

(d) Individually painted left image. (e) Individually painted right image.

(f) Stereoscopically painted left image. (g) Stereoscopically painted right image.

Figure 5: *Sawtooth data set (Courtesy of [SS02]). Notice in the individually painted images, 5(d) and 5(e), how paint spills between different surfaces, that are clearly indicated by the disparity map 5(c). The images when viewed stereoscopically also exhibit inconsistencies in the distribution of paint within the body of those surfaces, which causes eyestrain. In contrast, the results from our method [SG05b], shown in 5(f) and 5(g), are consistently painted. In addition, our algorithm efficiently tackles the problem of paint spilling to preserve intact the edges along depth discontinuities.*