Stereoscopic Painting with Varying Levels of Detail

Efstathios Stavrakis and Margrit Gelautz

Vienna University of Technology
Interactive Media Systems Group
Favoritenstrasse 9-11/188/2, A-1040 Vienna, Austria

ABSTRACT

We present an algorithm for generating automatically stereoscopic paintings with varying levels of detail. We describe our interactive system built around the algorithm to enable users to adjust the level of detail of the painting. In this context of interactivity we have modified our stereo painting algorithm, presented in previous work, in order to explore the idea of user-driven artistic level-of-detail selection and display. In particular, a stereo painting is composed by two canvases, one for each eye. These canvases contain multiple refining coarse-to-fine layers of brush strokes that compose the final painting. In past research, the underlying coarser layers are obscured and function only as the basis to progressively build the finer painting layers. In contrast, our interactive stereo viewing system enables the user to selectively toggle the visibility of finer strokes to reveal coarser representations of the artwork.

Keywords: Stereo, painting, levels of detail, non-photorealistic rendering

1. INTRODUCTION

Stereoscopic painting traditionally has been explored by visual artists. In the middle of the 20th century, not long after the formal discovery of stereoscopy, a number of painters started to experiment with this technique by painting canvases that could be viewed stereoscopically, either via the use of technical equipment or free viewing. This special technique requires that the artist produces two paintings, simulating each of the retinal images used by our binocular visual system.

The task of creating almost identical pieces of art, even under loosely defined consistency constraints, can be rather tedious. Usually technical means are put into use by artists in order to overcome the bottlenecks of preserving the spatial relations of features in the dually depicted composition. Even though there is scarce information about the technical details of this artistic process, we can categorize stereoscopic artwork in photographically and non-photographically assisted. In the first category, painters (Salvador Dali, Michael Kupka) who explored the possibilities of communicating art via stereoscopy have used real stereo photographs as the basis of their compositions. The non-photographically assisted paintings are composed by abstract shapes and geometric forms (Roger Ferragallo,[1] Heinz Günther Leitner, Oscar Fischinger), where managing feature consistency becomes less difficult than real life subjects, but the artist is able to communicate new spaces, see Fig. 1, rather than reproduce realistic spatial relations.

Independent of the category they belong to, all stereoscopic paintings can be rather immersive and stimulate the brain of the beholder in unconventional ways. As stereo fusion takes place, the two dimensional artwork either protrudes in front of the 2D surface or recedes, captivating the third eye of stereo vision in an exploratory journey of spatial diversity.

Drawing knowledge from stereo painters and observing their methods, as described above, we propose a hybrid algorithm that uses image-based stereo analysis and non-photorealistic rendering that ought to produce stereo paintings. Our technique is more closely related to the photographically assisted stereo paintings, but with the additional possibility of slicing the stereo painting at different distances to explore the underlying paint structure.

Further author information: (Send correspondence to E.S.)
E.S.: E-mail: stathis@ims.tuwien.ac.at, Telephone: +43 (0)1 58 801 18 872
M.G.: E-mail: gelautz@ims.tuwien.ac.at, Telephone: +43 (0)1 58 801 18 849

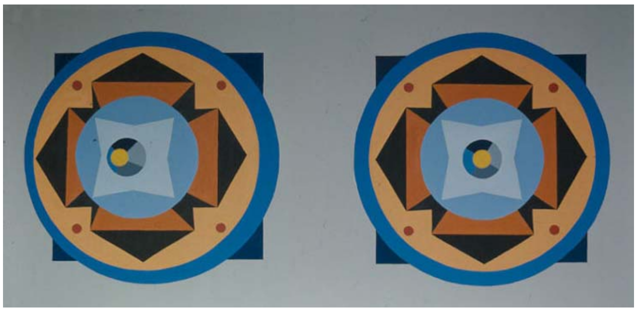


Figure 1. "Apollo Mandala" (1972) acrylic on canvas. Non-photographically assisted stereo painting by Roger Ferragallo. Used by permission.

1.1. Overview

The stereo painting algorithm computes automatically two paintings that are decomposed into depth layers in three dimensions. Our image-based algorithm uses a stereo pair of real images as input. Stereo analysis is performed and a set of auxiliary images is computed. Disparity maps encoding information about the correspondence of scene points between the input stereo images are extracted. For each view an occlusion mask is constructed. We then plan a painting in one view and propagate paint to the second view by using the disparity information. The stereo painting is completed by filling in the occluded areas with paint.

The paintings are rendered at certain distances from the viewpoint, based on the depth maps that can be inferred from the disparity information. The viewer can use our interactive stereo system to "slice" the painting at different distances to further examine the underlying structure of the multi-layer painting which encodes multiple levels of detail.

1.2. A Practical Framework for Stereo Non-Photorealistic Rendering

Through our technique, as well as experiments with the interactive system, we propose a practical framework of design and implementation considerations that stereoscopic algorithms for artistic content creation should account for. This framework, acting mostly as a set of guidelines rather than a formal ruleset, can further be modified and adapted by other single-view non-photorealistic rendering (NPR) algorithms to generate artistic stereoscopic content.

Feature Correspondence. Features must be consistent across the two views and the epipolar constraint should be reinforced when content is generated procedurally. Painting elements, i.e. brush strokes, that cannot be matched in both views will inhibit stereo fusion and the viewers may experience discomfort. Large non-corresponding areas and deviations in paint color, or style, will produce similar undesired effects. The human visual system is able to tolerate a small percentage of inconsistency, which varies from person to person, but algorithms should strive on providing the best possible correspondence between the two paintings.

Randomness. Stochastic procedures used in generating artistic effects must be carefully designed so that the feature correspondence constraint is met. Non-photorealistic rendering techniques, such as,[2–4] use random numbers to inject irregularity into the artistic process. In image-based stereoscopic NPR, it must be as consistent as possible across a stereo image pair, so that irregularity can be equally modelled within the two views. Perturbation functions applied on stylization parameters must also account for the desired consistency.

Performance. Stereo pairs exhibit correlative information between their images that can be used to optimize the performance of a variety of algorithms[5, 6] (e.g. compression algorithms or rasterization of stereo pairs). Painterly algorithms can take advantage of stereo pair relationships (i.e. stereo disparity) to reduce the computational effort of planning or rendering of the paintings.

Depth from Stereo. Disparity maps acquired by using automated computer vision algorithms, such as stereo matching algorithms,[7] cannot be used to perfectly reconstruct a scene and they often encode artifacts that falsify depth information of scene points. Establishing a feature correspondence by using imperfect depth-from-stereo information may inhibit the artistic process. Artistic algorithms that use such depth maps have to be able to counterbalance the inaccurate geometric descriptions of the scene objects, without degrading the quality of the final artwork. It must be noted that under certain circumstances it may be acceptable to trade off artistic fidelity for consistency, since in stereo viewing inconsistencies have an unpleasant effect to the viewers.

Occlusions. Occluded regions in stereo imaging, which are sets of points in the scene not visible from both viewpoints, must be specially treated. When viewed stereoscopically, occlusions have no correspondence and cannot be fused, however they must be filled in with appropriate content. In an artistic process the style over an occluded region must seamlessly blend with the surrounding regions for perceptual consistency.

Paint Spilling. Paint spilling practically occurs when colour that is naturally expected by the viewer to belong to one object being at a distinct distance from him, appears to be applied on another object, which lies at a different distance. This usually happens because of inaccuracies encoded in the depth map used, e.g. the common effect of block-based stereo matchers that "fatten" the foreground objects in the disparity map. Generally, we call *paint spilling* the result of inadequately extended painting elements across surfaces at different depths. Intersurface paint spilling does not necessarily degrade the quality of a single painting, however in stereo painting it is a more noticeable effect. The viewer may be confused since separation of the depth layers becomes more difficult.

2. RELATED WORK

Image-based stereoscopic painterly rendering as a hybrid computer graphics and computer vision technique has been first introduced by Stavrakis and Gelautz.[8] This technique varies from the current work in both design and implementation. A stereo pair of real images were progressively painted from coarse-to-fine in 2D, using stroke-based rendering,[9] and consistency across the two views was preserved by warping the final painted bitmap to the second view. Further the problem of occlusions was tackled by computing a corrective set of strokes that blend with the overlaying warped bitmap. However, the painting elements correcting the occluded regions on the second view could introduce inconsistencies across the two views, particularly when perturbation functions are used to stylize the colours of the brush strokes. This technique produced a pair of 2D images that could be fused stereoscopically, but the underlying refinement layers where merged. Therefore interactivity with this output is limited in stereoscopically viewing the two painted images and only the separation of the image pair can be adjusted to shift the plane of zero parallax at different distances.

Bartesaghi et al. presented a technique for generating non-photorealistic renderings using multiple images.[10] They generate pen-and-ink illustrations by reconstructing surface normals and principal directions to guide a texture synthesis process that produces the stylized output. As the authors point out, the surface normals' recovery can be rather unreliable especially when using depth-from-stereo. Similarly, our technique does not require a complete geometrical representation of the real scene.

Raskar et al.[11] have developed a novel hardware assisted method to automatically create comprehensible illustrations of real images by using depth discontinuities to enhance legibility. They use multiple images taken by a multi-flash camera where cast shadows are used to detect depth discontinuities in the scene. These depth edges together with texture of the scene objects are stylized to convey and better communicate the important information of even complex scenes.

A limited number of other techniques[12,13] have proposed the use of depth information to restrict stylization overruns of salient image features. However, they use highly accurate depth information computed by raytracing of 3D scenes and therefore are not exposed to the problems present in the depth maps estimated by using depth-from-stereo.

Apart from the stereoscopic painterly algorithm,[8] all other techniques described aim to generate monoscopic content and do not provide stereo output, which requires consistent stylization on both stereo views. We also observe that techniques using depth-from-stereo do not attempt to extract a geometric description of the scene objects and then process them for an artistic effect in 3D, but rather use either reliable features (e.g. depth discontinuities) related to the extracted depth information or provide alternative mechanisms in the stylization process that compensate for those inaccuracies.

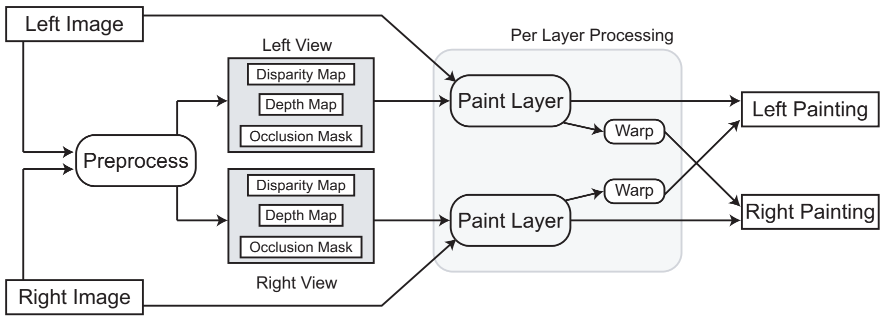


Figure 2. A schematic of our stereo painting algorithm.

3. STEREOSCOPIC PAINTING

Our stereo painting algorithm, shown in Fig. 2, uses still images acquired from spatially displaced cameras, which simulate the viewpoints from which an artist would depict his envisioned artwork. In a preprocessing step, explained in Sect. 3.1, the system processes the two input images to create a stereo image pair and also compute a set of auxiliary data that are required by the core stereo painting algorithm. The stereo painting algorithm, described in Sect. 3.2, generates interactive stereoscopic paintings that encapsulate multiple levels of detail. Finally in Sect. 3.3, we give a description of the system's functionality that enables the user to interactively examine the generated stereo paintings.

3.1. Input Processing and Auxiliary Data

Images acquired from the two horizontally displaced cameras cannot be readily used in stereo image analysis. Unless a stereo image acquisition device is used, a pair of cameras usually needs to be calibrated. In our case, we correct the two images for lens distortions of the cameras and bring them into epipolar geometry[14] by using a common calibration procedure proposed by Zhang.[15]

We perform automated stereo image analysis to calculate a dense disparity map in the geometry of the left image. We have used both a dynamic programming based stereo matching algorithm introduced by Birchfield and Tomasi,[16] see Fig. 8(a), but also a graph-based technique proposed by Bleyer and Gelautz,[17] see Fig. 8(b). We have considered using two greatly different stereo matching algorithms to evaluate and understand how our stereo painting technique works together with different quality disparity maps. Assuming that perfect disparity information is available when using depth-from-stereo is not a realistic scenario in most practical cases.

We use this disparity map to infer depth information of the scene. For a given disparity value of a scene point the z value of that point from the viewpoint is inversely proportional to the disparity value multiplied with a scale factor ($Depth = s \cdot \frac{1}{Disparity}$). Warping the left depth map by using the disparity information we construct a partial depth map in the geometry of the right view. Occluded regions and artifacts from disparity information subsampling are encoded into a binary occlusion mask for each view separately. We apply a morphological operation of dilation to this mask to assist the algorithm compensate for inaccurate disparity information. The depth values in these occluded regions are filled in with reliable depth values of neighboring pixels. At the end of this procedure we have for each view the following:
(a) an original RGB image, after calibration and distortions correction,
(b) disparity values that map each point to another point on the other view,
(c) depth values that provide an estimate of the distance of the scene point from the viewpoint,
(d) an occlusion mask encoding the positions of those points in the original image that have no correspondence to the other view.

A self-recorded dataset showing the auxiliary images along with the painted images can be seen in in Fig. 9.

3.2. Image-Based Stereoscopic Painterly Rendering Algorithm

The stereo image synthesis algorithm uses the data described in the previous section to produce the stereoscopic painting. We use the stroke-based painting algorithm introduced by Hertzmann,[4] which we modified to prevent the effect of paint spilling and further tackle the problem of occluded regions, while preserving consistency across the two views.

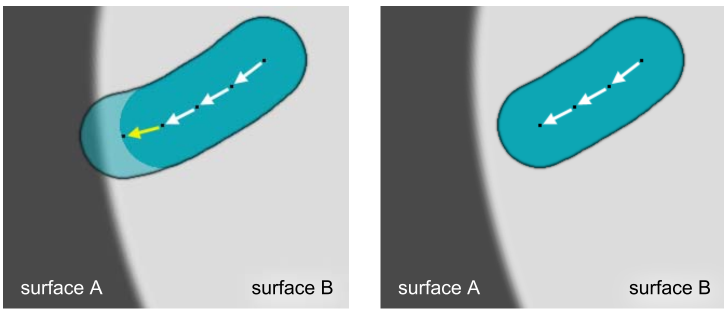
Given a set of radii, R_i, where $2 \leq R_i \leq R_{max}$, the painting algorithm* calculates for each R_i a collection of strokes that compose a single paint layer. The individual brush strokes are piecewise B-Spline curves that follow the contours of the input image that will be painted. As each brush stroke is planned by adding control points, we check whether the candidate control point P is near a depth discontinuity. For a brush stroke of radius R_i, we calculate in a rectangular neighborhood M of size $(2 * R_i + 1) \times (2 * R_i + 1)$ centered at P, the depth difference between P and every neighboring pixel P_n, $n \in M$. If any of the depth differences is above a user-defined depth difference threshold T_D, the control point is discarded and the stroke is terminated. This prevents paint spilling, see Fig. 3, since all strokes are restricted to surfaces with small depth variations.

A paint layer planned for the left view requires a corresponding and stereoscopically consistent paint layer to be generated for the right view. Using the previously computed disparity data, we warp the control points of each brush stroke from the left to the right view. We use the notion of an adaptive stroke to preserve consistency between the two views, but also to tackle the crucial problem of occlusions. During the warping operation of a brush stroke's control points, we detect whether the control point belongs to an occluded region of the source image. If the control point is within an occluded region, it is evident that there is no corresponding point that this control point should be mapped to on the destination view, hence the control point is discarded and the stroke warping is terminated. This way long strokes in the left view, spanning from unoccluded to occluded regions, will be split and only the part of the stroke that has correspondence may be warped.

After completing the propagation of paint, the right view will not be completely painted, since occlusions occur in the left view also. In practise these areas on the right view where paint was not possible to be propagated from the left view. These areas need to be painted in a successive step. We plan a painting in the right view of which strokes may only be initiated from occluded points, using the occlusion mask. The same procedure of painting as with the left view is used. The result of this process will be a small set of strokes emanating from occluded regions to the rest of the painting. Even though the occluded area is filled in with paint, inconsistencies arise between the two views, since these new strokes span outside occluded areas, which are visible on the other view. To tackle possible inconsistencies we apply the warping operation in reverse, from right to left, to propagate any unoccluded portions of those strokes to the left view.

Brush strokes are given a unique z value, which is naively chosen to be the depth value at the first control point of the stroke. We do not attempt to follow surface orientation, since we found that surface reconstruction

*For a thorough description of the painting algorithm please consult Hertzmann's article.[4]



(a) (b)

Figure 3. Paint Spilling. (a) The next control point of a stroke inadequately extends over surfaces A and B at different depths. (b) We discard the candidate control point and terminate the stroke to prevent paint spilling.

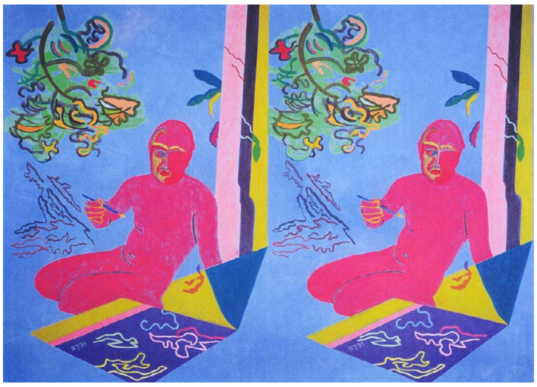


Figure 4. "Self Portrait with Brush" (1983) by Heinz Günther Leitner. Used by permission.

is not sufficiently accurate. The problem, contrary to monoscopic methods that use surface orientation, is that when these paint elements are positioned in false orientations the stereo fused result, even though it is consistent, may appear to have unnaturally oriented objects that confuse the user. We chose to draw the painting strokes normal to the line of sight at a distance z instead, hence a certain amount of cardboarding† must be expected. However, we observed that a certain amount of cardboarding is present in artists' works also, see Fig. 4, and in general does not significantly affect the perception of depth in the final stereo painting.

The above procedure is applied for each paint layer to generate a multi-layer stereo painting. Finally the paint layers are rendered on a coarse-to-fine manner, so that a coarse layer is refined progressively with the rendering of each additional layer. The painting elements of those layers are sorted and rendered according to their z value from far to near. The level of detail of the painting is defined by the amount of brush stroke sizes chosen. Coarse paint elements that vastly abstract the input images are further refined by the painting algorithm and the refining layers are positioned closer to the viewer. The coarser strokes remain obscured and further away from the viewer, at the most part, serving mostly as guiding elements for the generation of the finer strokes.

Once the two views are completed, the initial images used to guide the painting's planning should be positioned behind all layers in order to fill in missing pixels. We also found that users found it adequate to be able to see how the strokes relate to the original images.

The final stereo painting appears as a pair of stereo images, that are viewed from an orthographic projection. However, due to its progressively refined layers it contains precalculated multiple levels of detail, that are positioned at known distances from the viewer. You can see two finished stereo paintings, "Boy" and "Girl", in Fig. 6 and Fig. 7 respectively.

3.3. Interacting with Stereoscopic Paintings

To interact with the stereo paintings generated with this algorithm we established two modes of interactivity and implemented a supportive user interface. Our interactive system provides a page-flipping hardware accelerated stereo rendering of the two canvases. The user is assisted into fusing the stereo images by using active shutter stereo glasses. The plane of zero parallax in the fused stereo image coincides with the surface of the digital display. At this plane features between the two stereo images coincide. Any points that have negative parallax will be perceived to occupy space between the viewer and the display surface, while points with positive parallax will appear to be behind the display.

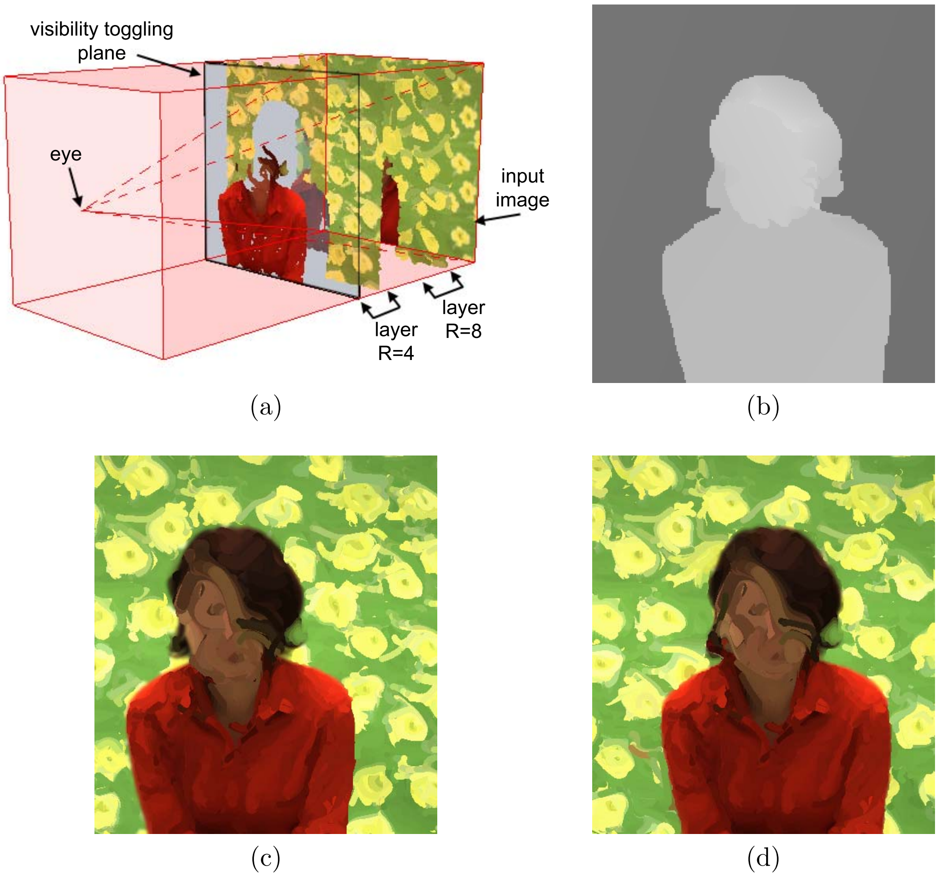
The user can adjust the separation of the stereo pair to set different points at the plane of zero parallax, effectively bringing features of the stereo image out of the display or sending them behind it. The separation adjustment is done by allowing the user to interact with the disparity map. When the user points at a particular

†Cardboarding is the effect of smooth surfaces to be perceived as flat cutouts in the stereo fused image.

location on it, using common input devices, the system automatically brings the plane of zero parallax at this point by calculating the appropriate separation for the two views.

Apart from adjusting the depth impression of the stereo painting, we enable the users to toggle the visibility of painting elements at different distances. The amount of brush strokes composing each layer usually exceeds a human manageable number and it is impractical to expect the user to hide individual elements in order to see the coarse detail obscured by them. Therefore, we bind a stroke visibility-toggling plane to the finest paint layer. The user can then interactively move this plane, which is normal to the line of sight also, to toggle the visibility of the front refining brush strokes effectively revealing the different levels of detail of the painting that are present behind the initial orthographic projection. An example of the user interacting with the painting is shown in Fig. 5.

The two interaction methods provided, allow a user to affect the look and feel of the painting. Another advantage of our interactive system in comparison to other systems and real hand-painted canvases, stereoscopic or not, is that the paintings can be decomposed or explored by the spectator, further enhancing the experience of viewing the artwork.



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

Figure 5. A "sliced" stereo painting with two paint layers. (a) A perspective view of the right component of the stereo painting. The user looks from the eye position. (b) Disparity map, in the geometry of the left view, used to generate the painting. (c) Right component of the sliced painting as the user views it from the orthographic projection. (d) Left component.

4. CONCLUSION

We presented a system for viewing and interacting with stereoscopic paintings generated automatically by a combination of stereo image analysis and non-photorealistic rendering techniques. Our system performs well, but there are inherent the problems of stereo analysis. Even though we do not attempt to reconstruct the underlying geometry of the scene the process may fail to achieve good results from very inaccurate disparity information. Since a complete geometric description is not available, depending on the amount of different depth layers encoded in the disparity map, the stereo painting may have limited depth layers and objects may appear to be flattened. This effect is perceived as cardboarding, but is better accepted by the users than inconsistencies or unnaturally oriented surfaces.

We have proposed a set of guidelines that monoscopic NPR techniques may adopt in order to produce stereo artwork and have shown how one of those algorithms may be modified to use it. Further experimentation is required with other algorithms using different stylizing elements, not necessarily brush strokes. Combining stereo matching algorithms with NPR techniques can be more robust in estimating and preserving consistency rather than post-processing the output of stereo analysis (e.g. per pixel disparity information).

Interactivity with a stereo painting can be enhanced by using eye tracking data. Using a combination of gaze-directed adjustment of the plane of zero parallax and haptic devices to slice the painting may enhance the user's experience.

ACKNOWLEDGMENTS

This work was supported by the Austrian Science Fund (FWF) under project P15663. We would like to thank fine artist Roger Ferragallo for granting us permission to use his stereo painting 'Apollo Mandala' (1972), as well as VESCON Systemtechnik GmbH (Gleisdorf, Austria) for making available the stereo painting 'Self Portrait with Brush' (1983) by Heinz Günther Leitner. We also wish to thank Michael Bleyer for providing us with an implementation of his stereo matching algorithm.

REFERENCES

1. R. Ferragallo, "On stereoscopic painting," *Leonardo* **7**, pp. 97–104, Spring 1974.
2. M. P. Salisbury, *Image-Based Pen-and-Ink Illustration.* PhD thesis, Department of Computer Science and Engineering, University of Washington, 1997.
3. P. Litwinowicz, "Processing images and video for an impressionist effect," in *Proc. SIGGRAPH '97*, **31**, pp. 407–414, 1997.
4. A. Hertzmann, "Painterly rendering with curved brush strokes of multiple sizes," in *Proc. SIGGRAPH '98*, pp. 453–460, ACM Press, 1998.
5. H. S. Sawhney, Y. Guo, K. Hanna, R. Kumar, S. Adkins, and S. Zhou, "Hybrid stereo camera: an IBR approach for synthesis of very high resolution stereoscopic image sequences," in *Proc. SIGGRAPH '01*, pp. 451–460, ACM Press, 2001.
6. S. J. Adelson, J. B. Bentley, I. S. Chong, L. F. Hodges, and J. Winograd, "Simultaneous generation of stereoscopic views," *cgforum* **10**(1), pp. 3–10, 1991.
7. D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision* **22**, pp. 7–47, April-June 2002.
8. E. Stavrakis and M. Gelautz, "Image-based stereoscopic painterly rendering," in *Eurographics Symposium on Rendering*, (Norrköping, Sweden), 2004.
9. A. Hertzmann, "A survey of stroke-based rendering," *IEEE Computer Graphics & Applications, Special Issue on Non-Photorealistic Rendering* **23**, pp. pp. 70–81, July/August 2003.
10. A. Bartesaghi, G. Sapiro, T. Malzbender, and D. Gelb, "Non-photorealistic rendering from multiple images," in *IEEE International Conference on Image Processing (ICIP04)*, pp. 2403–2406, (Singapore), October 24-27 2004.
11. R. Raskar, K. han Tan, R. Feris, J. Yu, and M. Turk, "Non-photorealistic camera: Depth edge detection and stylized rendering using multi-flash imaging," *ACM Transactions on Graphics* **23**(3), pp. 679–688, 2004.
12. B. Gooch, G. Coombe, and P. Shirley, "Artistic vision: Painterly rendering using computer vision techniques," in *Proc. NPAR '02*, pp. 83–90, ACM Press, 2002.

13. M. Shiraishi, "Adaptive parameter control for image moment-based painterly rendering," in *Proc. 9th International Conference on Geometry and Graphics*, 2000.
14. O. Faugeras, *Three-Dimensional Computer Vision: a Geometric Viewpoint*, The MIT Press, 1993.
15. Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(11), pp. 1330–1334, 2000.
16. S. Birchfield and C. Tomasi, "Depth discontinuities by pixel-to-pixel stereo," *International Journal of Computer Vision* **35**, pp. 269–293, December 1999.
17. M. Bleyer and M. Gelautz, "Graph-based surface reconstruction from stereo pairs using image segmentation," in *Proceedings of SPIE (to appear)*, **5665**, 2005.

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

Figure 6. "Boy" dataset. (a) Right input image, (b) Left input image, (c) Right painted image, (d) Left painted image.

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

Figure 7. "Girl" dataset. (a) Right input image, (b) Left input image, (c) Right painted image, (d) Left painted image.

(a) (b)

Figure 8. Disparity images on the geometry of the left view. (a) Disparity image for "Boy" dataset, (b) Disparity image for "Girl" dataset.

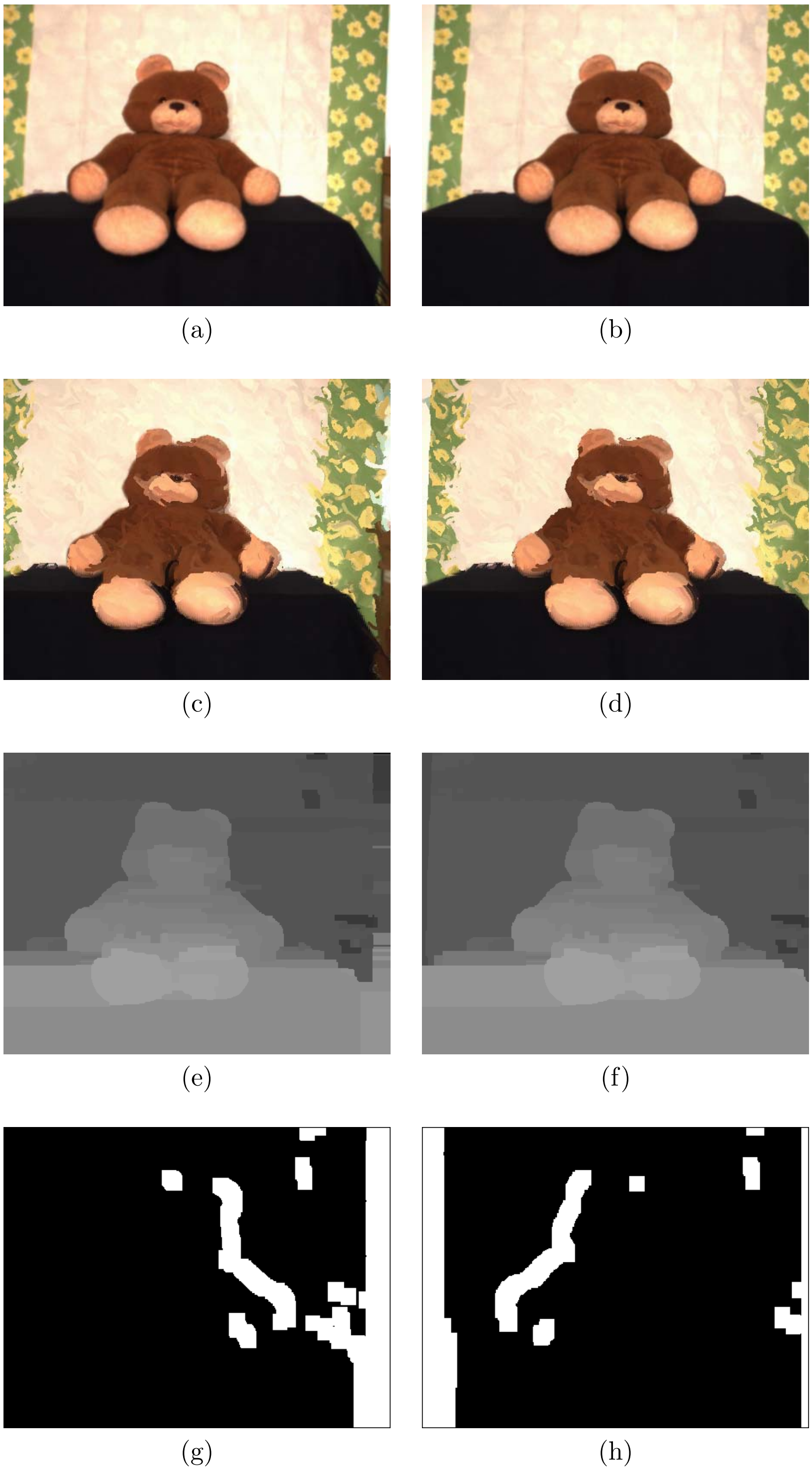


Figure 9. "Teddy" complete dataset with auxiliary images. (a) Right input image. (b) Left left image (c) Right painted image. (d) Left painted image. (e) Right disparity image. (f) Left disparity image. (g) Right occlusion mask. (h) Left occlusion mask.