IMAGE-BASED STEREOSCOPIC STYLIZATION

Efstathios Stavrakis, Michael Bleyer, Danijela Markovic and Margrit Gelautz

Interactive Media Systems Group, Vienna University of Technology
Favoritenstrasse 9-11/188/2, A-1040 Vienna, Austria
{stathis, bleyer, markovic, gelautz}@ims.tuwien.ac.at

ABSTRACT

We present a method to generate stylized stereo imagery that effectively communicates shape and distance of the depicted scene objects. We use computer vision techniques to analyze real stereo image pairs. In particular, a region based stereo matching algorithm with symmetrical treatment of occlusions is used to extract a disparity map and successively the depth information of the scene. The reference image is color segmented for the purpose of color stylization and an algorithm combining intensity image edges and depth discontinuities is applied to depict dominant object contours in the image. We use disparity information to propagate stylized color segments to the second view together with the object outlining contours. The stylized image pairs are consistent across the two views and can be easily fused for stereoscopic viewing. The stereoscopic image fusion provides an extra dimension of depth that is absent on the individual images.

1. INTRODUCTION

Visual artists have used mainly monocular cues (e.g. linear perspective) in order to stimulate their audiences into perceiving inexistent depth information on 2D visual media. However, the human visual system perceives depth information more robustly by using binocular cues (e.g. retinal disparity). Stereoscopic artwork reintroduces these binocular cues by providing two components of the artwork, one for each eye. When fusing the two components into the previously unseen stereo image, depth and distances between objects become apparent and a new world of aesthetics can be perceived. However, creating stereoscopic artwork, whether it is a meticulous painting or a rough cartoon, requires that the artist has extensive technical merit to preserve feature consistency in the two views, as well as additional effort to generate the artwork twice.

We propose that stereo vision is used to tackle the technicalities of the process and we present a method to automatically stylize real stereo image pairs. While many non-photorealistic rendering algorithms produce single view artistic output, most are not tailored to generate stereo. When generating stereoscopic artwork the two components should not be treated individually and the algorithms must account for particular subtleties of stereo artistic content creation, discussed in more detail in [1].

It is important that artistic style corresponds between the two views and hence a relationship must be established between them. In our approach, we use disparity information to achieve this, which we extract by employing a stereo matching algorithm proposed in [2]. Similarly to the algorithm used to stylize single photographs presented in [3], we use color segmentation [4] to transform the colors of the original images into simpler stylized ones. However, because of the stereoscopic nature of our output, to preserve consistency and also increase performance, we stylize only the first view and then propagate style to the second view. While most of the second view can be reconstructed by using disparity information, occluded regions of the second view are consistently synthesized by applying style present in neighboring segments. To further enhance the stylized stereo images we automatically stroke foreground objects, consistently in both views. We do this by combining intensity edge information with depth discontinuities to detect dominant object contours, which we then use as outlines.

2. ALGORITHM

2.1. Computation of the disparity map

In order to obtain a disparity map, we developed a stereo matching algorithm, which we briefly summarize as follows. For an elaborate description the reader is referred to [2]. In a preprocessing step, the algorithm divides the reference (left) image into a set of segments of homogeneous color. It is assumed that disparity inside such segments varies smoothly and depth discontinuities coincide with segment boundaries. We oversegment the image to ensure that these assumptions are met. To describe each segment's disparity, we extract a set of *disparity layers*, which

This work was supported by the Austrian Science Fund (FWF) under project P15663.

are planar surfaces that are likely to occur in the scene. The assignment of segments to layers is done in order to optimize a global cost function. The cost function attempts to minimize pixel dissimilarities and is designed to handle occlusions symmetrically in both views, as well as to motivate smooth disparity solutions. A robust optimization strategy that employs graph cuts is used to find a local optimum of the cost function.

The applied stereo method incorporates a number of features that particularly favor the generation of stereoscopic stylized images. By using color segmentation, the algorithm is capable of handling regions of low texture, finding precise depth boundaries and assigning occluded areas of the reference image to meaningful disparity values, which are challenging tasks for conventional stereo algorithms. Furthermore, the mapping of segments to layers naturally represents a disparity segmentation. We show the layer map computed for the Venus test set taken from [5] in figure 2(c). Pixels of the same color thereby belong to the same disparity layer. The algorithm was successful in segmenting the five planes that are present in the scene. Visual inspection of the corresponding disparity map presented in figure 2(d) indicates the good quality of the achieved matching results. High grey values represent high disparities in the figure.

2.2. Stylization of the reference view

Stylization of the left image is achieved by using an additional color segmentation step that is different from the one used in the stereo matcher. We employ the algorithm proposed in [4] to identify segments of homogeneous color. The color value of each pixel is then replaced by the mean color of all pixels belonging to the same segment. A post-processing step serves to smooth the edges between neighboring segments. This step changes the color assignments of isolated pixels, i.e. pixels that have a low number of neighbors (in 8-connectivity) being assigned to the same color value.

According to our design principles, we want to avoid generating segments that overlap a depth discontinuity. However, if the whole image is segmented at once, this is very likely to happen. For this reason, we use the information provided in the layer map to divide the left image into a set of regions. A region is built by all pixels of the left image belonging to the same disparity layer. Each of those regions is then stylized individually. In this process, we incorporate the idea that foreground objects are more interesting to the viewer and should therefore contain higher detail than objects of the background. While the parameters of the segmentation algorithm are adjusted to produce a finer segmentation for regions of high disparity, regions of low disparity are stylized using a coarser segmentation. We present the resulting stylization of the left image in figure 2(e).

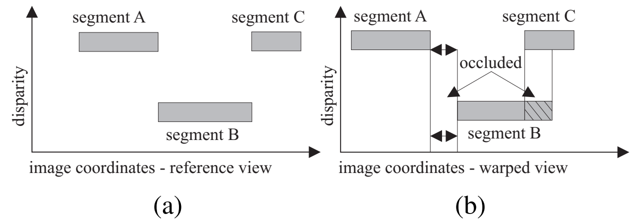

Fig. 1. The warping operation and the occlusion problem. (a) Three segments of the reference view and their disparity values. (b) Segments warped to the second view according to their disparity values.

2.3. Generation of the second view via image warping

Since we want to consistently preserve style across views, we propagate the segments from the left to the right view using the computed disparity map. The right view is thereby obtained by shifting each pixel of the left image horizontally, according to its disparity value. While this procedure is trivial for pixels that are visible in both views, we have to consider the occlusion problem, which is illustrated in figure 1. Some pixels of the warped view receive contributions from more than one pixel of the reference view, which represents an occlusion in the second view. In this case, the pixel of highest disparity is visible, since it is the pixel closest to the viewpoint. Other pixels of the warped image do not receive contribution from any pixel, which corresponds to an occlusion in the reference view. Those regions pose a more complex problem, since no information about their style is available in the reference view. We show this situation in figure 2(f). White pixels mark those areas that are affected by this problem in the figure.

For occluded regions in the warped image we copy the corresponding pixels from the original second view and stylize them by using color segmentation. However, the appearance of the segments that were filled in differs from the appearance of segments that were propagated from the stylized left image. This can, for example, be recognized along the right border of 2(g). To overcome this problem, a segment merging operation is performed in order to copy style from other segments. We compute a dissimilarity score based on the color similarity and the border length between two segments. Segments of identical color values and large border length thereby generate low dissimilarity scores. For each segment of an occluded area, the neighboring segment of lowest dissimilarity score is determined. If that score is smaller than a threshold t_{dis}, the segments are merged. This procedure is then iterated with increasing values for t_{dis} until t_{dis} becomes larger than a user defined threshold. Starting with low values for t_{dis} ensures that very similar segments are merged first and makes the procedure inde-

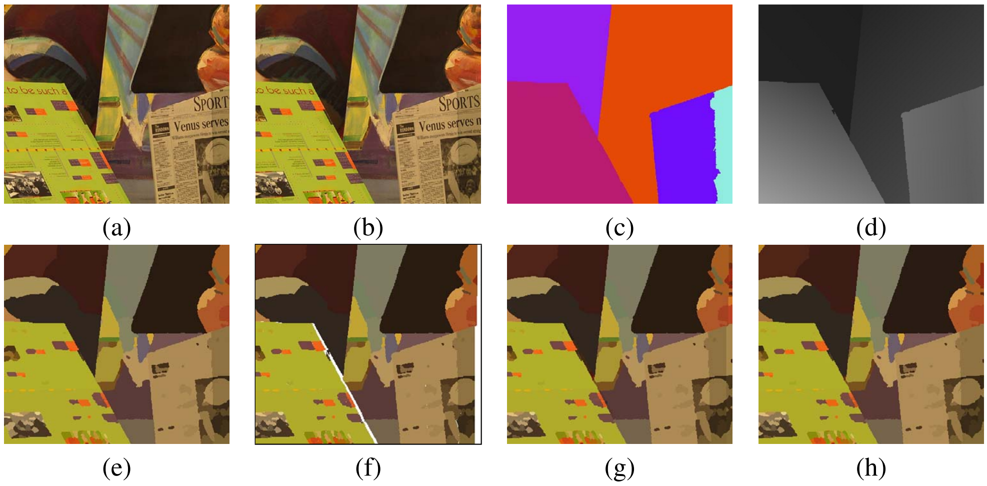

Fig. 2. Stereoscopic stylization of the Venus data set. (a) Left image. (b) Right image. (c) Layer map of the left image. (d) Disparity map for the left image. (e) Stylized left image. (f) Warped second view. (g) Warped second view with occlusions filled in. (h) Warped second view after segment merging.

pendent of the order of applied merge operations.

2.4. Outlining

Line is an important factor in image comprehension and in the context of stereoscopy it can assist the viewer to seperate the various depth layers. Following the principles of consistency in stereo imaging we have devised a method to partially stroke dominant contours of objects in the scene. These dominant contours usually represent parts of silhouettes and creases of objects. We construct two edge images, one by detecting the depth discontinuities using the disparity map in the geometry of the reference view, and a second by applying edge detection to the intensity image. Depth discontinuities may be displaced or have gaps, in relation to their intensity edge counterparts, because of disparity artifacts. We combine the information of the two edge images into another outline image in which those artifacts are greatly suppressed.

We label all edges in both edge images and apply per pixel similarity measurements for each of them. The similarity metrics we use are edge orientation and position. Lines in the depth discontinuity image whose similarity to corresponding intensity edges is above a user-defined threshold are marked as dominant edges. This algorithm produces a set of outlines that may be disconnected, but still belong to the same outline. We use the initial labels and similarity metrics to identify and connect endpoints of disconnected outlines. Similarly to the propagation of style from left to right, we assign those edge pixels to foreground objects and warp them onto the right view, using the disparity map. A more detailed description of these steps are part of the edge combination algorithm described in [6]. Figure 3 shows the final stylized output with the stereo outlining algorithm applied to the Venus data set.

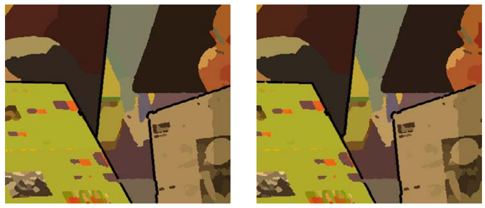

Fig. 3. Black outlining of dominant edges in the Venus set. Left and right components.

3. EXPERIMENTAL RESULTS

We have tested our algorithm with real image datasets. Apart from the Venus dataset of [5], we have recorded the Teddy dataset, shown in figure 4. Figures 4(a) and 4(b) show the original left and right components of the stereo image pair. This scene has a mixture of textured and untextured surfaces, which are split into layers, figure 4(c), and the respective disparity map extracted by our stereo matching algorithm is shown in figure 4(d). The outlining algorithm uses the intensity edge image, figure 4(e), in combination with the depth discontinuities, as shown in 4(f), to detect dominant object contours, as presented in 4(g). In figures 4(h) and 4(i), the output of the color stylization step is shown and the final result after stroking in them the dominant edge outlines can be seen in 4(j) and 4(k).

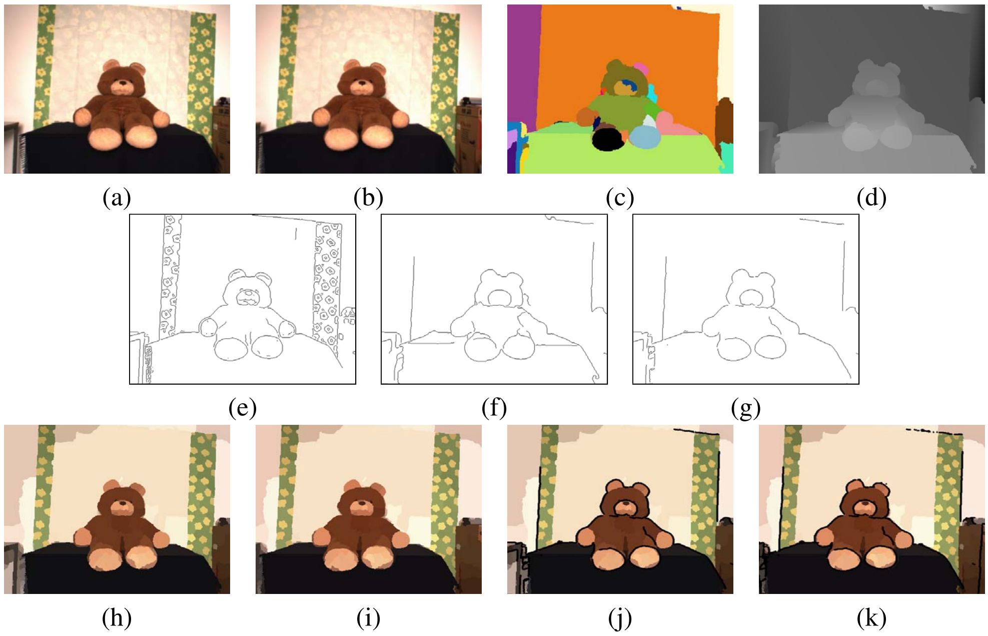

Fig. 4. Self-recorded Teddy data set. (a) Left image. (b) Right image. (c) Layer map (left view). (d) Disparity map (left view). (e) Intensity edge image (left view). (f) Depth discontinuities (left view). (g) Dominant contours (left view). (h) Stylized left image. (i) Stylized right image. (j) Stylized left image with dominant outlines. (k) Stylized right image with dominant outlines.

Our experimental results show that there is great potential in generating stereoscopic artwork by using purely computer vision algorithms. We found that the idea of propagating information from one view to the other, further tackling occlusions, and stroking outlines around objects generates stylized images that are comfortable and pleasing when viewed stereoscopically.

4. CONCLUSIONS

We have proposed a multiview image-based method that generates stereoscopic stylized images. Our approach enabled us to couple our stereo matching algorithm with the color stylization and object outlining processes to provide consistent stereoscopic output with a hand crafted appearance. We have taken advantage of the stereo matcher's layered design to treat occlusions, as well as preserve style within and across the two views. Finally, we have presented how outlines can further enhance the stereo experience by assisting the viewer in separating distinct scene objects, without producing image clutter.

5. REFERENCES

[1] E. Stavrakis and M. Gelautz, "Image-based stereoscopic painterly rendering," in *Proceedings of Eurographics Symposium on Rendering*, 2004, pp. 53–60.

[2] M. Bleyer and M. Gelautz, "Graph-based surface reconstruction from stereo pairs using image segmentation," in *Proceedings of SPIE*, 2005, vol. 5665.

[3] D. DeCarlo and A. Santella, "Stylization and abstraction of photographs," in *SIGGRAPH '02*, 2002, pp. 769–776.

[4] C. Christoudias, B. Georgescu, and P. Meer, "Synergism in low-level vision," in *International Conference on Pattern Recognition*, 2002, vol. 4, pp. 150–155.

[5] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1/2/3, pp. 7–42, 2002. http://www.middlebury.edu/stereo/ accessed April 28, 2005.

[6] M. Gelautz and D. Markovic, "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.