Drawing the Real

Danijela Marković and Margrit Gelautz*
Institute for Software Technology and Interactive Systems, Vienna University of Technology
Favoritenstrasse 9-11/188/2, A-1040 Vienna, Austria

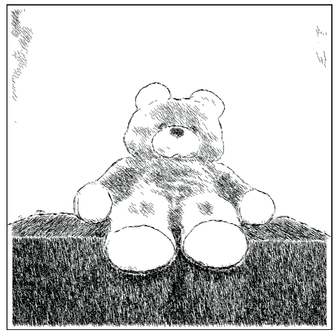


Figure 1: Real scene depicted in drawn-like style using our algorithm.

Abstract

In this paper we present an algorithm to automatically produce artistic drawings from stereo image pairs. The input to the algorithm is a natural scene, along with a user-defined set of parameters that define the tone and stylistic properties of the image to be produced. The stereo analysis yields a depth map that is used to preserve the perspective perception of the stylized image. In the next step, we perform a local image contrast enhancement, which is driven by the color segmentation result of the original image. We then draw each stroke in a direction determined by the stereo-derived disparity layers, with the stroke density and distribution derived from a Poisson disc distribution. To outline important features in the image, we utilize the contour edges provided by an Edge Combination image. The output of the algorithm is a drawn-like form of the original scene, with objects highlighted by their dominant edges.

Keywords: Image-based rendering, drawing, real scene, stereo, Edge Combination.

1 Introduction

As a major fine art technique in itself, drawing is usually defined as a depiction of shapes and forms on a surface by means of lines. Throughout history, drawings were usually created to develop ideas and as preparations for larger or more complex works of art. Over time, people began to consider the drawings, themselves, as art objects. The artwork where the line describes the scene, idea or simply depicts the shapes is no less interesting today. Nowadays computers have become an important and very interesting tool for creating artistic work. Though the computer itself cannot yet produce creative pieces of art, it can artistically depict real or artificial scenes, assist the artist or support the art creation learning process. Consequently, considerable research has been devoted to simulate graphic media. One part of related studies takes an available 3D scene as input, such as [Winkenbach and Salesin 1994; Sousa and Buchanan 1999; Hertzmann and Zorin 2000; Praun et al. 2001]. This geometry-based approach benefits from complete access to the 3D geometry. On the other side there are image-based systems [Salisbury et al. 1997; Hertzmann 1998; Shiraishi and Yamaguchi 2000; Mao et al. 2001; Li and Huang 2003], using available images to produce illustrations. The latter approach has no a-priori comprehension of the scene geometry, and to convey shapes by orienting strokes, it includes either user interaction or image or texture analysis. Recent research has shown the advantage of using combined techniques. The approach in [Secord et al. 2002] is image-oriented and additional 3D information, if available, is used to create a larger variety of styles. In [Xu and Chen 2004] an interactive non-photorealistic rendering system is presented, which stylizes and renders outdoor scenes captured by 3D laser scanning. An earlier work by [Saito and Takahashi 1990] demonstrated the benefit of using depth maps to create a variety of styles to visualize the scene.

In this paper, we exploit the potential of computer vision techniques for image stylization. As opposed to [Secord et al. 2002] and [Xu and Chen 2004], our work is concentrated on recorded stereo image pairs of real scenes. We include additional geometry by computing the disparity image, which is inversely related to the scene depth, to attach to the illustration a feeling of "depth". Furthermore, we use the stereo-derived disparity to compute a so-called Edge Combination image [Markovic et al. 2005] to outline the dominant scene structure.

*e-mail: {markovic, gelautz}@ims.tuwien.ac.at

2 Algorithm

Figure 2 gives a short summary of the involved processing steps. A stereo image pair consisting of a left and right stereo image is processed by a stereo matching algorithm, which delivers a disparity map in the geometry of one of the two input images as output.

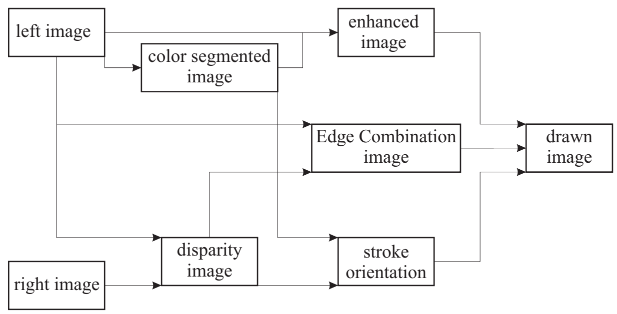


Figure 2: Overview of processing steps.

The disparity map is described by a set of disparity layers, as a result of modeling the 3D scene by a set of planar facets. For further processing, we enhance the reference image by using its color segmentation result. The color segmented image along with the orientation of the disparity layers gives the information to orient strokes. To highlight the dominant structure of the scene we use the Edge Combination algorithm for more accurate recognition of significant contour edges. The various processing steps are described in more detail in subsections 2.1 through 2.4.

2.1 Image Segmentation

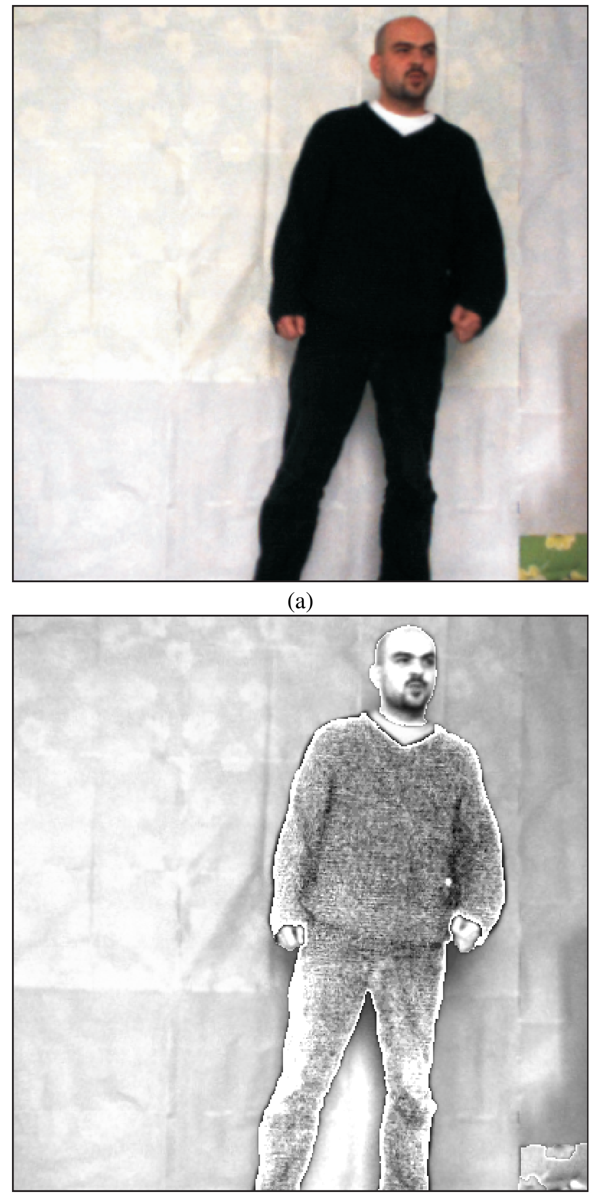
In this step we color segment the reference image employing the Mean Shift Based Image Segmenter (EDISON System) developed by [Comaniciu and Meer 2002]. The image is segmented by applying a high color difference so that each resulting region contains one dominant color. Regions with a size smaller than a user specified threshold are merged to neighboring large regions. The output of the color segmentation step serves to locally adjust the image contrast in order to achieve a more balanced distribution of stroke density across the whole image, as described in the following.

2.2 Stroke Arrangement

Throughout history, artists were using varying tonal values like shadows and illumination to better convey shapes and achieve a "live" appearance of the scene. In this work, we present scenes in drawn-like illustrations derived from a gray scale version of the recorded images. The stroke density varies according to the tone of the image region, from very dense strokes in dark, to almost no strokes in very light areas, depending on the parameters specified by the user. In images recorded from a real scene we often find very dark areas usually with poor contrast. Inevitably, it is necessary to locally enhance the contrast of such images before further processing in order to better convey local shape variations and provide a well-balanced overall impression. The local contrast enhancement is discussed in substep 2.2.1. Substep 2.2.2 explains how the enhanced image is used to calculate the position and local density of the computed strokes.

2.2.1 Image Enhancement

The principal idea of our contrast-enhancement step is to "remove" color from the image. As a natural solution of this problem we use color segmentation to enhance contrast for each segment. The user selects a scaling factor (constant for the whole image), which is applied to each segment: Values greater than the segment's mean value, modified by the scaling factor, turn to white and lower values are stretched between a minimum value and white. An example of this image enhancement step is shown in Figure 3. Some newly revealed details, especially in extremely dark areas of the image (see, for example, the trousers of the recorded person), are obvious in Figure 3(b).



(a)

(b)

Figure 3: (a) Left camera image, (b) enhanced image.

2.2.2 Stroke Centers Generation

To control the density of strokes according to the image intensity values, we first generate random points using the Poisson distribution that we further distribute according to the Poisson disc distribution method with variable disc radii, as explained in [Sato et al. 2004]. The position of the center point of each stroke is specified by calculating the disc radius, which can be considered as a limit that none of the other points can fall inside. The disc radius R of each center point is determined using the intensity value $I_{i,j}$ at that location as:

$$R = R_{df} + 2R_w \left(\tan^{-1} A_I \left(\frac{I_{i,j}}{I_{av}} - 1 \right) / \pi \right), \qquad 0 \le R_w \le R_{df} \quad (1)$$

where I_{av} is the average intensity value and A_I, R_w and R_{df} are user defined constants used to control the density of the strokes. When $I_{i,j}$ decreases (i.e., in darker parts of the image), R decreases to produce more dense strokes and otherwise, when $I_{i,j}$ increases R also increases, which makes the strokes more sparsely distributed. This results in a convenient tuning of the image tone by adjusting the density and distribution of the strokes according to the intensity values, with darker areas represented by a denser stroke pattern than bright regions.

2.3 Stroke Orientation

To depict shapes and to give the image a "depth" impression, a very important factor, the stroke direction should be considered. To produce such effects, each stroke is drawn perpendicularly to the angle θ, which is specified by the orientation of the disparity layers in the disparity image.

In order to obtain a disparity map, we employed the graph-cut based stereo matching algorithm described in [Bleyer and Gelautz 2005]. This method is of particular interest, since it represents disparity by a set of planar layers. The assignment of pixels to those disparity layers can naturally be regarded as disparity segmentation. The disparity maps that we show later in this paper (Figures 4 and 6-9) were computed automatically by our matching algorithm. No additional manual cleaning or noise removal was performed. Further results of our matching algorithm in application to scenes with different structure and complexity can be seen in the corresponding paper and on the Middlebury stereo evaluation website [Scharstein and Szeliski 2002].

In the next step, for each layer, we utilize a square window of a maximum size previously specified by the user, as presented in Figure 4(a). In order to enhance the contrast, the values of the window are rescaled from black to white as in Figure 4(c). To identify the orientation of the projected disparity values, we use image moments as given in [Shiraishi and Yamaguchi 2000].

The image moment of lth degree about the x-axis and mth degree about the y-axis is defined as:

$$M_{lm} = \sum_x \sum_y x^l y^m I(x,y). \qquad (2)$$

M_{lm} is called the image moment of nth degree for $n = l + m$. The image moment of zeroth degree, M_{00}, is simply the sum of all the intensity values of the gray-scale image. The angle θ is calculated as follows:

$$\theta = \frac{1}{2} \tan^{-1} \left(\frac{b}{a-c} \right), \qquad (3)$$

where a, b, and c are defined as:

$$a = \frac{M_{20}}{M_{00}} - x_c^2, \qquad (4)$$

$$b = 2\left(\frac{M_{11}}{M_{00}} - x_c y_c \right), \qquad (5)$$

$$c = \frac{M_{02}}{M_{00}} - y_c^2, \qquad (6)$$

where x_c and y_c are:

$$x_c = \frac{M_{10}}{M_{00}}, \qquad (7)$$

$$y_c = \frac{M_{01}}{M_{00}}. \qquad (8)$$

As an option, the stroke orientation derived from image moments can be combined with additional user-defined preferences. In our application, we decided to replace horizontally and vertically aligned strokes, which we found to produce less appealing effects, by adding/subtracting a value of 45° whenever θ is close to 90° or 0°.

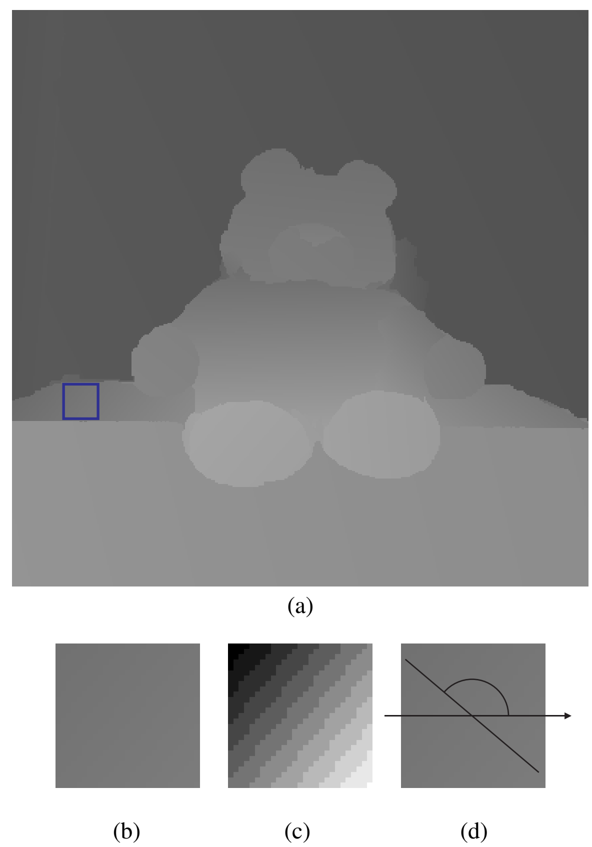


Figure 4: Process of disparity layer orientation estimation: (a) disparity image, (b) square window of the disparity layer, (c) rescaled values of the window in (b) from black to white, (d) estimated angle θ of the disparity layer.

2.4 Outlining

A very important element often used by artists are outlines or silhouettes, which define the edges of objects, describe their shape and indicate their volume. The approach we use to emphasize the dominant structure of the scene is based on our previous work [Markovic et al. 2005], in which we suggest the computation of a so-called

Edge Combination image. The idea of the Edge Combination algorithm that we devised is to combine the edges from the disparity edge image, which suggests the location of scene object discontinuities, with the higher positional accuracy of the edges from the original image for correct recognition of significant contours.

For strokes resulting from the Edge Combination algorithm, we derive the orientation using image moments applied to the color segmented image. This distinguishes the contour strokes from the rest of the image, whereby the stroke orientation was computed from the enhanced intensity image. As a result, outlines are emphasized in the drawn image.

3 Results

The obtained results of our algorithm applied to images of real scenes are illustrated in Figures 6 through 9. The size of the images is 400 x 400 pixels for all four test scenes in Figures 6 through 9. Figures (a) and (b) show a pair of stereo video frames in epipolar geometry. Their resulting disparity image is given in Figure (c), and Figure (e) shows the labeled disparity layers. Figure (d) shows the Edge Combination image, which combines the edges derived from the original image (a) and the disparity image (c), to outline the dominant structure of the scene. The result of the image color segmentation is given in Figure (f). Finally, Figure (g) shows the final drawing result: The real scene is illustrated using strokes oriented according to the disparity layers with the strokes' density reflecting local brightness variations encountered in the contrast-enhanced original image. In addition, dominant edges in the image (e.g., object contours) are highlighted to produce an artistic hand-drawn effect.

To illustrate the advantage of using a disparity map, Figure 5 presents the same scene as Figure 6, but generated without disparity information. One can recognize that the front and the top side of the Teddy's desk cannot be separated from each other in Figure 5, due to the uniform stroke orientation caused by the homogeneous appearance of the whole desk in the original image from Figure 6(a). This results in an obvious absence of the impression of "depth" in Figure 5. Contrarily, in Figure 6(g) the computer-generated strokes follow the orientation of the disparity layers and the geometric structure of the desk is clearly discernible.

The benefit of the image enhancement step described in section 2.2.1 can be recognized by comparing the result in Figure 7(g) with the corresponding original and contrast-enhanced images in Figure 3. According to the disparity layers in Figure 7(e), most parts of the person's body (except head, neck and hands) are represented by a single layer and thus results in a uniform stroke orientation. However, a sense of depth is conveyed by variations in the stroke density that follow the brightness pattern of the enhanced image in Figure 3(b).

4 Summary

We have proposed a method to generate stylized images derived from real scenes. Our approach employs the output of a stereo matching algorithm and utilizes tone stylization and object outlining techniques to synthesize images with a drawn-like appearance. The layered output of the stereo matcher is used to orient the strokes and in this way preserve the perspective impression of the image. In addition, we generate outlines that further enhance the image outlook and help to better apprehend the discrete scene structure.

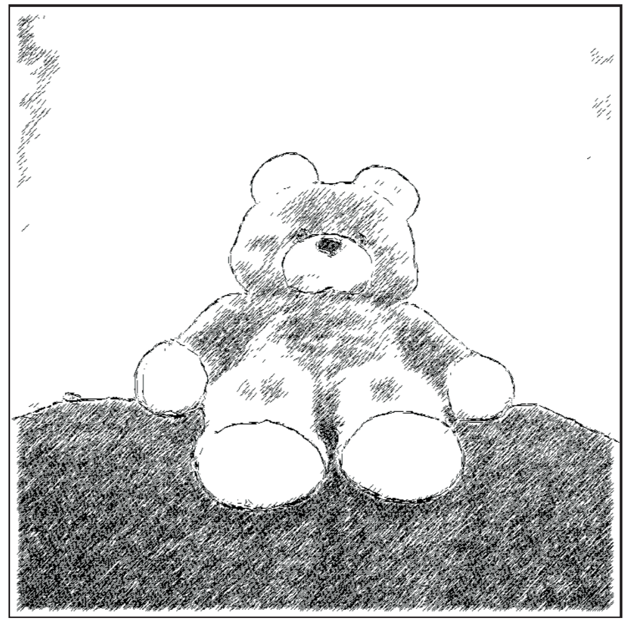


Figure 5: Example of the drawn-like result without disparity information.

Acknowledgments

We wish to thank Michel Bleyer for providing the disparity maps. This research has been funded by the Austrian Science Fund (FWF) (project no. P15663) and by the Austrian Federal Ministry for Education, Science, and Culture, as well as the European Social Fund (ESF) under grant 31.963/46-VII/9/2002.

References

BLEYER, M., AND GELAUTZ, M. 2005. Graph-based surface reconstruction from stereo pairs using image segmentation. In *Proceedings of SPIE*, vol. 5665, 288–299.

COMANICIU, D., AND MEER, P. 2002. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*, 5, 603–619.

HERTZMANN, A., AND ZORIN, D. 2000. Illustrating smooth surfaces. In *Proceedings of SIGGRAPH '00*, 517–526.

HERTZMANN, A. 1998. Painterly rendering with curved brush strokes of multiple sizes. In *Proceedings of SIGGRAPH '98*, 453–460.

LI, N., AND HUANG, Z. 2003. A feature-based pencil drawing method. In *Proceedings of GRAPHITE '03*, 135–140.

MAO, X., NAGASAKA, Y., AND IMAMIYA, A. 2001. Automatic generation of pencil drawing from 2d images using line integral convolution. In *Proceedings of CAD/Graphics 2001*, 240–248.

MARKOVIC, D., STAVRAKIS, E., AND GELAUTZ, M. 2005. Parameterized sketches from stereo images. In *Proceedings of SPIE*, 783–792.

PRAUN, E., HOPPE, H., WEBB, M., AND FINKELSTEIN, A. 2001. Real-time hatching. In *Proceedings of SIGGRAPH '01*, 581–586.

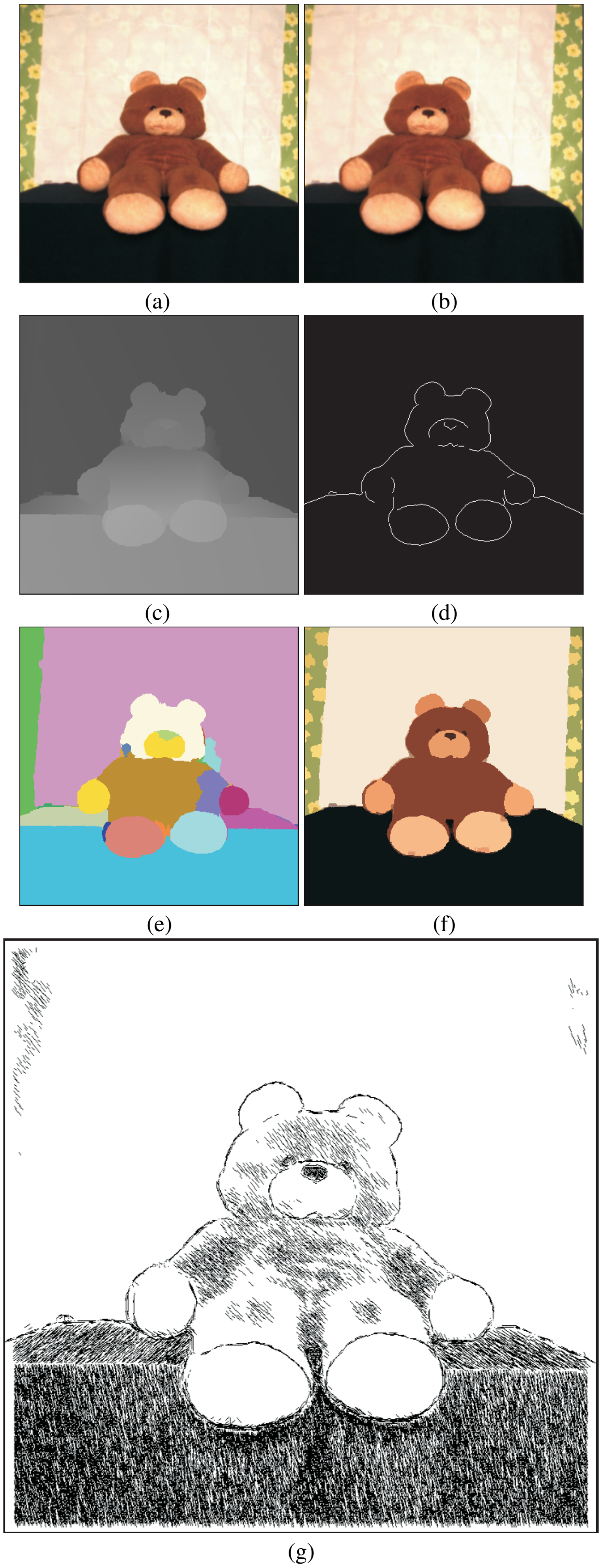


Figure 6: (a) Left camera image, (b) right camera image, (c) disparity image, (d) Edge Combination image, (e) disparity layers, (f) color segmented image, (g) drawn image.

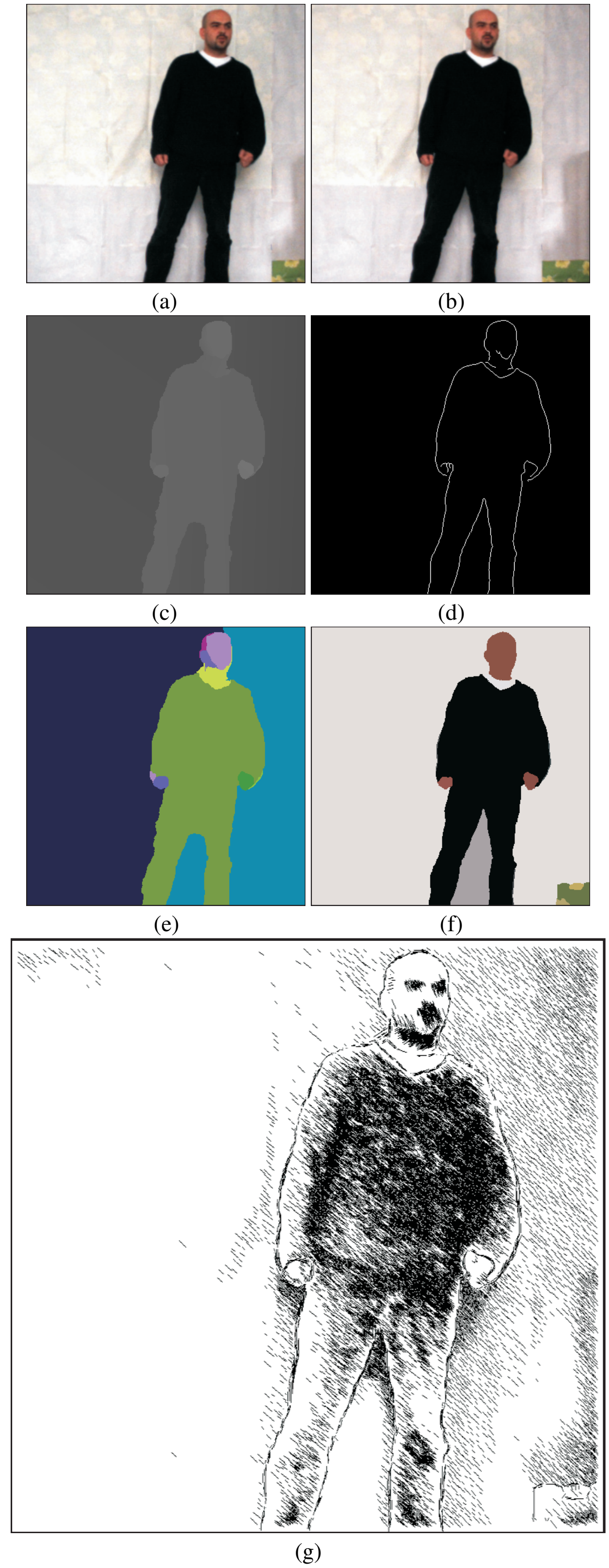


Figure 7: (a) Left camera image, (b) right camera image, (c) disparity image, (d) Edge Combination image, (e) disparity layers, (f) color segmented image, (g) drawn image.

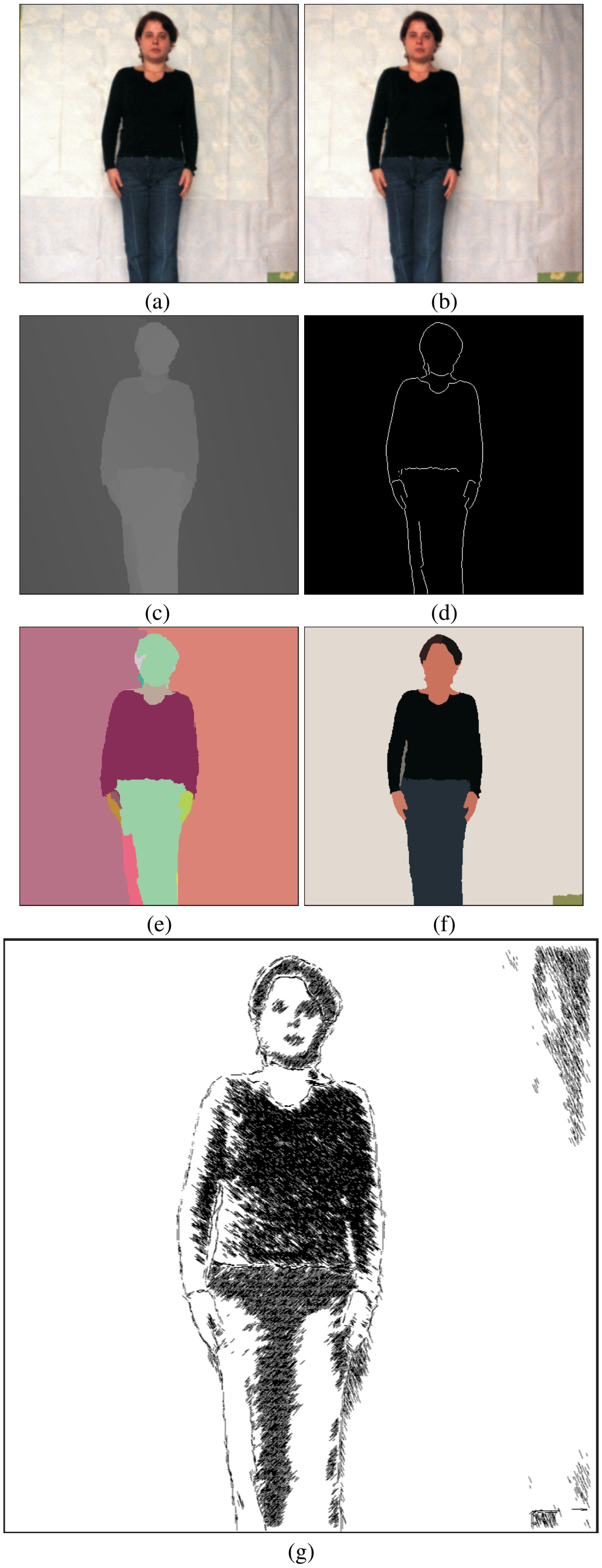


Figure 8: (a) Left camera image, (b) right camera image, (c) disparity image, (d) Edge Combination image, (e) disparity layers, (f) color segmented image, (g) drawn image.

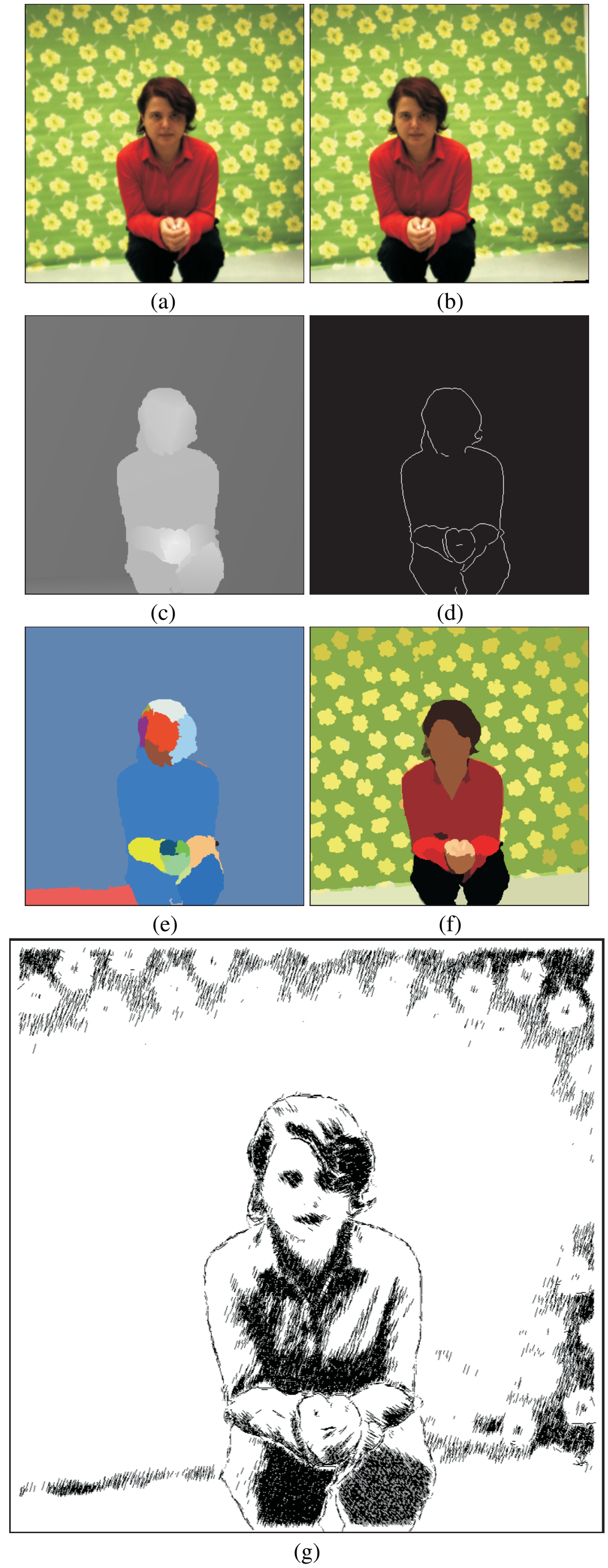


Figure 9: (a) Left camera image, (b) right camera image, (c) disparity image, (d) Edge Combination image, (e) disparity layers, (f) color segmented image, (g) drawn image.

SAITO, T., AND TAKAHASHI, T. 1990. Comprehensible rendering of 3-d shapes. *Computer Graphics 24*, 4, 197–206.

SALISBURY, M., WONG, M., HUGHES, J., AND SALESIN, D. 1997. Orientable textures for image-based pen-and-ink illustration. In *Proceedings of SIGGRAPH '97*, 401–406.

SATO, Y., FUJIMOTO, T., MURAOKA, K., AND CHIBA, N. 2004. Stroke-based suibokuga-like rendering for three-dimensional geometric models - ten and shun touches. *The Journal of the Society for Art and Science 3*, 4, 224–234.

SCHARSTEIN, D., AND SZELISKI, R. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision 47*, 1/2/3, 7–42. http://www.middlebury.edu/stereo/.

SECORD, A., HEIDRICH, W., AND STREIT, L. 2002. Fast primitive distribution for illustration. In *Proceedings of EGRW '02*, 215–226.

SHIRAISHI, M., AND YAMAGUCHI, Y. 2000. An algorithm for automatic painterly rendering based on local source image approximation. In *Proceedings of NPAR '00*, 53–58.

SOUSA, M. C., AND BUCHANAN, J. W. 1999. Computer-generated graphite pencil rendering of 3d polygonal models. In *Proceedings of Eurographics '99*, 195–207.

WINKENBACH, G., AND SALESIN, D. H. 1994. Computer-generated pen-and-ink illustration. In *Proceedings of SIGGRAPH '94*, 91–100.

XU, H., AND CHEN, B. 2004. Stylized rendering of 3d scanned real world environments. In *Proceedings of NPAR '04*, 25–34.