

Efficiently Capturing Object Contours for Non-Photorealistic Rendering

Jiyoung Park¹ and Juneho Yi²

¹ Computer Graphics Research Team, Digital Content Research Division,
Electronics and Telecommunications Research Institute, Daejeon 305-700, Korea
jiyp@etri.re.kr

² School of Information and Communication Engineering, Sungkyunkwan University,
Suwon 446-740, Korea
jhui@ece.skku.ac.kr

Abstract. Non-photorealistic rendering (NPR) techniques aim to outline the shape of objects and reduce visual clutter such as shadows and inner texture edges. As the first phase result of our entire research, this work is concerned with a structured light based approach that efficiently detects depth edges in real world scenes. Depth edges directly represent object contours. We exploit distortion of the light pattern in the structured light image along depth discontinuities to reliably detect depth edges. However, in reality, distortion along depth discontinuities may not occur or be large enough to detect depending on the distance from the camera or projector. For practical application of the proposed approach, we have presented a novel method that guarantees the occurrence of the distortion along depth discontinuities for a continuous range of object location. Experimental results show a great promise that the technique can successfully provide object contours to be used for non-photorealistic rendering.

Keywords: depth edges, structured light, non-photorealistic rendering.

1 Introduction

Depth edges directly outline shape contours of objects [2-5]. Unfortunately, there have been reported few research results that only provide depth discontinuities without computing 3D information at every pixel in the input image of a scene. On the other hand, most effort has been devoted to stereo vision problems in order to obtain depth information. In fact, stereo methods for 3D reconstruction would fail in textureless regions and along occluding edges with low intensity variation [6, 7]. Recently, the use of structured light was reported to compute 3D coordinates at every pixel in the input image [8, 9]. However, the fact that this approach needs a number of structured light images makes it hard to be applicable in realtime.

One notable technique was reported recently for non-photorealistic rendering [10]. They capture a sequence of images in which different light sources illuminate the scene from various positions. Then they use shadows in each image to assemble a depth edge map. This technique was applied to stylized rendering highlighting boundaries between

geometric shapes. Although very attractive, it only works where shadows can be reliably created. In contrast, our method is shadow free. In addition, by a slight modification of the imaging system so that it can capture white and structured images at the same time, it can be easily applied to dynamic scenes where the camera moves.

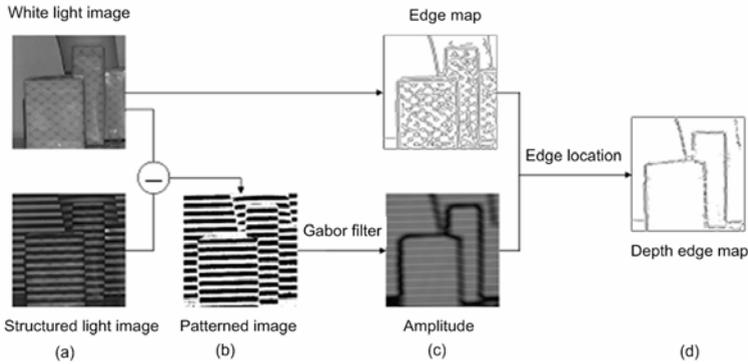


Fig. 1. Illustration of the basic idea to compute a depth edge map: (a) capture of a white light image and structured light image, (b) patterned image, (c) detection of depth edges by applying a Gabor filter to the patterned image with edge information from the white light image, (d) final depth edge map

The eventual goal of this research is to produce a depth edge map that conveys object contours to be used for non-photorealistic rendering. This work describes the first stage research result that is concerned with reliably capturing depth edges. We present a structured light based framework for reliably capturing depth edges in real world scenes without dense 3D reconstruction.

We have illustrated in Fig. 1 the basic idea that detects depth edges. First, as can be seen in Fig. 1 (a), we project a white light and structured light in a row onto a scene where depth edges are to be detected. The structured light contains a special light pattern. In this work, we have used simple black and white horizontal stripes with the same width. Vertical stripes can be used with the same analysis applied to horizontal stripes. We capture the white light image and structured light image. Second, we extract horizontal patterns simply by differencing the white light and structured light images. We call this difference image ‘patterned image’. Refer to Fig. 1 (b). Third, we identify depth edges in the patterned image guided by edge information from the white light image. We exploit distortion of light pattern in the structured light image along depth edges. Since the horizontal pattern can be considered a periodic signal with specific frequency, we can easily detect candidate locations for depth edges by applying a Gabor filter to the patterned image [11]. The amplitude response of Gabor filter is very low where distortion of light pattern occurs. Fig. 1 (c) illustrates this process. Last, we accurately locate depth edges using edge information from the white light image, yielding a final depth edge map as in Fig. 1 (d).

However, distortion along depth discontinuities may not occur or be sufficient to detect depending the distance from the camera or projector. For practical application of the proposed approach, it is essential to have a solution that guarantees the occurrence

of the distortion along depth discontinuities irrespective of object location. Fig. 2 shows an example situation. Along the depth edges between objects A and B, C and D, the distortion of pattern almost disappears. This makes it not feasible to detect these depth edges using a Gabor filter.

We propose a method based on a single projector-camera system that guarantees the occurrence of the distortion for a continuous range of object location. Based on a modeled imaging geometry of camera, projector, object, and its mathematical analysis, we first compute the exact ranges of object location where detection of distortion is not feasible. We simply use several structural light images with different width of horizontal stripes. We have used a general purpose LCD projector, however, an infrared projector can be employed with the same analysis in order to apply the method to humans. Experimental results have confirmed that the proposed method works very well for shapes of human hand and body as well as general objects.

The remaining of this paper is organized as follows. In section 2, we describe the application of Gabor filter to detect depth edges in a patterned image. Section 3 present our method that guarantees the occurrence of the distortion along depth discontinuities for a continuous range of object location. We report our experimental results in section 4. Finally, conclusions and future work are discussed in section 5.

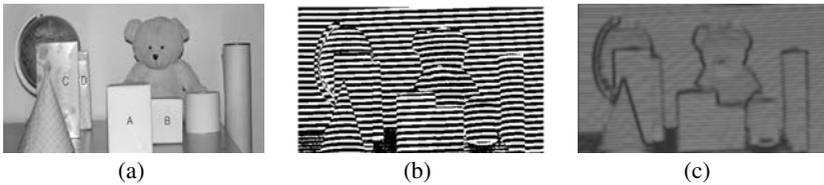


Fig. 2. Problem of disappearance of distortion along depth edges depending on the distance of an object from the camera and projector: (a) white light image (b) patterned image (c) amplitude response of Gabor filter. Along the depth edges between objects A and B, C and D, in the patterned image (b), the distortion of pattern almost disappears. This makes it not feasible to detect these depth edges using a Gabor filter.

2 Detecting Depth Edges

We detect depth edges by projecting structured light onto a scene and exploit distortion of light pattern in the structured light image along depth discontinuities. In order to exploit distortion of light pattern, we use 2D Gabor filtering that is known to be useful in segregating textural regions [1, 12]. We first find candidate depth edges by applying a Gabor filter to the patterned image. We then accurately locate depth edges using edge information from white light image.

2.1 The Use of Gabor Filter

Since a horizontal pattern can be considered a spatially periodic signal with specific frequency, we can easily detect candidate locations for depth edges by applying a

Gabor filter to the patterned image. A 2-D Gabor filter is an oriented complex sinusoidal grating modulated by a 2-D Gaussian function, which is given by

$$G_{\sigma,\phi,\theta}(x,y) = g_{\sigma}(x,y) \cdot \exp[2\pi j\phi(x\cos\theta + y\sin\theta)] \quad (1)$$

$$\text{where } g_{\sigma} = \frac{1}{2\pi\sigma^2} \exp[-(x^2 + y^2)/2\sigma^2]$$

The frequency of the span-limited sinusoidal grating is given by ϕ and its orientation is specified as θ . $g_{\sigma}(x,y)$ is the Gaussian function with scale parameter σ . Decomposing $G_{\sigma,\phi,\theta}(x,y)$ into real and imaginary parts gives

$$G_{\sigma,\phi,\theta}(x,y) = R_{\sigma,\phi,\theta}(x,y) + jI_{\sigma,\phi,\theta}(x,y) \quad (2)$$

$$\text{where } R_{\sigma,\phi,\theta}(x,y) = g_{\sigma}(x,y) \cdot \cos[2\pi\phi(x\cos\theta + y\sin\theta)]$$

$$I_{\sigma,\phi,\theta}(x,y) = g_{\sigma}(x,y) \cdot \sin[2\pi\phi(x\cos\theta + y\sin\theta)]$$

The Gabor filtered output of an image $f(x,y)$ is obtained by the convolution of the image with the Gabor function $G_{\sigma,\phi,\theta}(x,y)$. Thus, its amplitude response can be computed as follows:

$$E_{\sigma,\phi,\theta}(x,y) = \sqrt{[R_{\sigma,\phi,\theta}(x,y) * f(x,y)]^2 + [I_{\sigma,\phi,\theta}(x,y) * f(x,y)]^2} \quad (3)$$

2.2 Referring to Edges in White Light Image

It is possible to accurately locate depth edges by combining the Gabor filter output and edge information from the white light image. In this work, we use a gradient based technique to detect edges in the white light image, although other methods could also be applied. Fig. 3 (b) represents the gradient magnitude of the intensity along the line in the white light image in Fig. 3(a). We only take the gradient magnitude where Gabor amplitude has low values. Fig. 3 (c) illustrates the regions of low Gabor amplitude. The accurate locations of depth edges are obtained by finding peak points of the gradient magnitude in these regions (see Fig. 3 (d)).

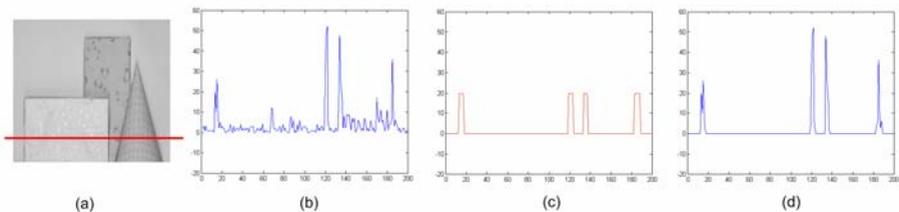


Fig. 3. Location of depth edges using edge information from the white light image: (a) white light image, (b) gradient magnitude along the line in (a), (c) regions of low Gabor amplitude (d) location of depth edges

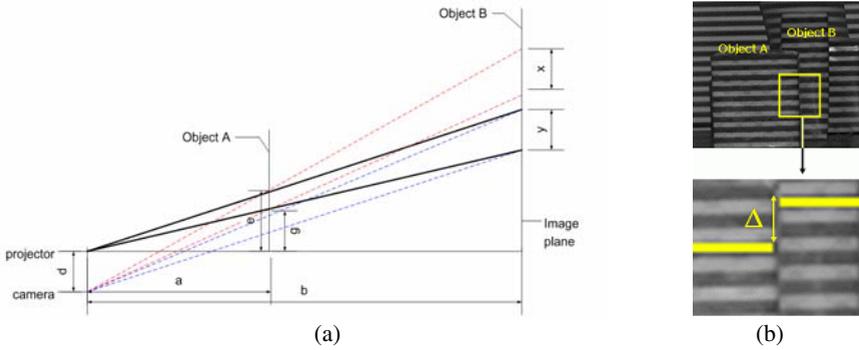


Fig. 4. Imaging geometry and the amount of distortion: (a) spatial relation of camera, projector and two object points viewed from the side, r : disparity in the image plane of the same horizontal stripe projected onto different object points, (b) the magnitude of pattern distortion, Δ , in a real image

3 Detectable Range of Depth Edges

We have described that we can easily detect depth edges by exploiting the distortion along depth discontinuities in the patterned image. However, as previously mentioned, the distortion may not occur or be sufficient to detect depending the distance of depth edges from the camera or projector. In this section, we present a method to guarantee the occurrence of the distortion for a continuous range of object location.

3.1 Reliably Detectable Distortion

In order to compute the exact range where depth edges can be detectable, we have modeled imaging geometry of camera, projector and object as illustrated in Fig. 4. The solid line represents a light ray from the projector. When structured light is projected onto object points A and B, they are imaged at different locations in the image plane due to different depth values. That is, distortion of horizontal pattern occurs along the depth discontinuity. The amount of distortion is denoted by Δ . From this model, we can derive the following equation using similar triangles:

$$\Delta = fd \left(\frac{1}{a} - \frac{1}{b} \right) = \left(\frac{fdr}{a(a+r)} \right) \tag{4}$$

However, the exact amount of Δ may not be measurable because we use simple black and white stripes with equal width and the amount of offset is periodic as it gets large. In order for a depth edge to be detectable by applying a Gabor filter, the measurable Δ , in the image plane should be above a certain amount. We have confirmed through experiments that 2/3 of the width of a horizontal stripe used is necessary for reliable detection of the distortion. Thus, the range of Δ for reliable detection of pattern distortion can be written as in equation (5).

$$2wk + \frac{2w}{3} \leq \Delta \leq 2wk + \frac{4w}{3}, \quad k = 0, 1, \dots \quad (5)$$

where w denotes the width of horizontal stripes. From equation (5), given the distance, r , between two object points and the separation, d , between the camera and the projector, we can compute the exact range of the foreground object point, A , from the camera where reliable detection of distortion is guaranteed. For practical application of the proposed approach, we need to guarantee the occurrence of distortion for a continuous range. Note that the width of horizontal stripes projected onto object locations A and B are the same in the image plane although they have different depth values. This is because the perspective effect of the camera and projector are canceled each other out.

3.2 Extending the Detectable Range of Depth Edges

We use several structured light images with different width of horizontal stripes to extend the range where detection of distortion can be guaranteed. Fig. 5(a) depicts the relationship between Δ and a for any k and $k+1$. The marked regions in the horizontal axis, a , represent the ranges of the foreground object point, A , from the camera that correspond to reliably detectable distortion Δ in the vertical axis. We can see that there exist ranges where we cannot detect depth edges due to the lack of distortion depending the distance of a depth edge from the camera or projector. Therefore, to extend the range of detectable distortion, Δ , we use additional structured light whose spatial period is halved such as $w_2=2w_1$, $w_3=2w_2$, $w_4=2w_3$, \dots , as shown in Fig. 5(b). When n such structured light images are used, the range of detectable distortion, Δ , is expressed as follows.

$$\frac{2}{3}w_1 < \Delta < \frac{(n^2 + 3n)}{3}w_1 \quad (6)$$

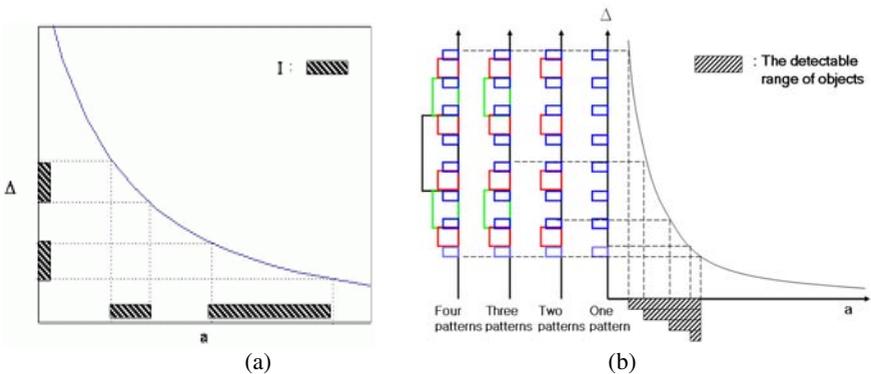


Fig. 5. The detectable range of depth edges (a) can be increased like (b) by projecting additional structured light with different width of stripes

3.3 Computation of the Detectable Range of Depth Edges

As shown in Fig. 6 (a), the detectable range of depth edges $[a_{\min}, a_{\max}]$ is computed in the following two steps: (1) Setting the maximum distance of the detectable range, a_{\max} ,

and the minimum distance between object points, r_{\min} , determines the width of stripes, w , in structured light image. r_{\min} can be set to different values depending on applications. (2) This w gives the minimum distance of the detectable range, a_{\min} , resulting in the detectable range of depth edges, $[a_{\min}, a_{\max}]$.

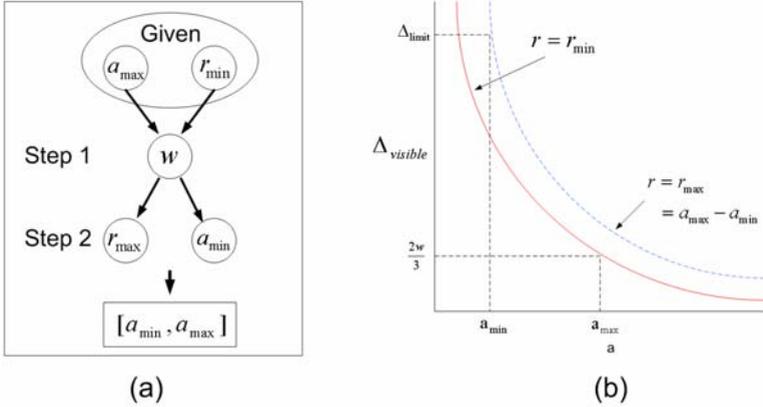


Fig. 6. Computation of the detectable range of depth edges: (a) computation process of $[a_{\min}, a_{\max}]$, (b) computation of a_{\min}

Step 1: Determination of the width of a stripe, w , in the structured light

First, we set a_{\max} to the distance from the camera to the farthest background. Given r_{\min} , w can be computed by equation (7) which is derived from equation (4).

$$w = \frac{3fd_1 r_{\min}}{2a_{\max}(a_{\max} + r_{\min})} \tag{7}$$

Thus, given a_{\max} and r_{\min} , we can compute the optimal width of stripes of the structured light.

Step 2: The minimum of the detectable range, a_{\min}

Given w from step 1, we can compute a_{\min} that corresponds to the upper limit of Δ , Δ_{limit} , as shown in Fig. 6 (b). After determining Δ_{limit} and r_{\max} , a_{\min} can be computed by equation (9). r_{\max} denotes the maximum distance between object points in the range $[a_{\min}, a_{\max}]$ that guarantees the occurrence of the distortion along depth discontinuities. Clearly, the distance between any two object points is bounded by $(a_{\max} - a_{\min})$, i. e., $r_{\max} = a_{\max} - a_{\min}$. Thus Δ_{limit} becomes:

$$\Delta_{\text{limit}} = \frac{fdr_{\max}}{a_{\min}(a_{\min} + r_{\max})} = \frac{fd(a_{\max} - a_{\min})}{a_{\max}a_{\min}} \tag{8}$$

On the other hand, $\Delta_{\text{limit}} = \frac{n^2 + 3n}{3} w_1$ from equation (6). Solving for a_{min} gives:

$$a_{\text{min}} = \frac{fd_1 a_{\text{max}}}{fd_1 + \frac{(n^2 + 3n)w_1}{3} a_{\text{max}}}. \quad (9)$$

This way, we can employ structured light of optimal spatial resolution so that we are guaranteed to detect depth edges of all object points located in the range $[a_{\text{min}}, a_{\text{max}}]$, and apart from each other no less than r_{min} and no more than r_{max} .

4 Experimental Results

For capturing structured light images, we have used a HP xb31 DLP projector and Cannon IXY 500 digital camera. Fig. 7 shows the result of depth edge detection using three structured light images with different width of horizontal stripes. Fig. 7 (a) and (b) display the front and side views of the scene, respectively. All the objects are located within the range of 2.4m ~ 3m from the camera. Setting $f=3\text{m}$, $d=0.173\text{m}$, $a_{\text{max}}=3\text{m}$ and $r_{\text{min}}=0.1\text{m}$, w_1 and a_{min} are determined as 0.0084m and 2.325m, respectively. That is, the detectable range of depth edges becomes $[2.325\text{m}, 3\text{m}]$ and the length of the range is 0.675m. Thus, the widths of stripes of the three structured light that guarantee the detection of depth edges in this range are w_1 , $2w_1$ and $4w_1$.

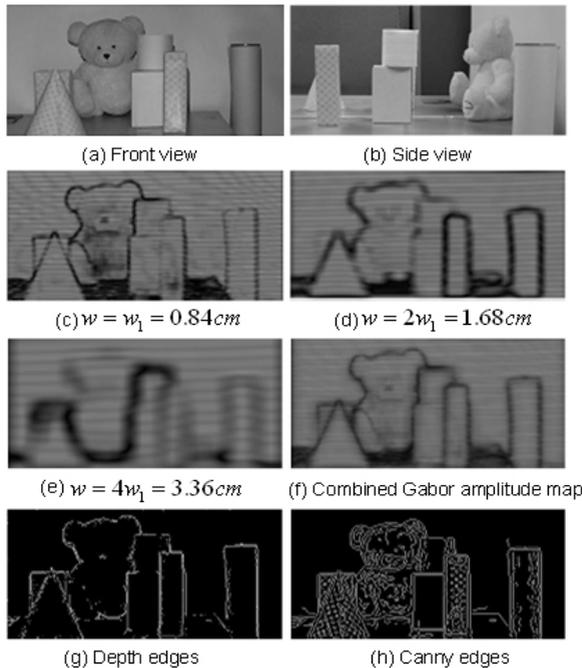


Fig. 7. Detecting depth edges using a single camera and projector

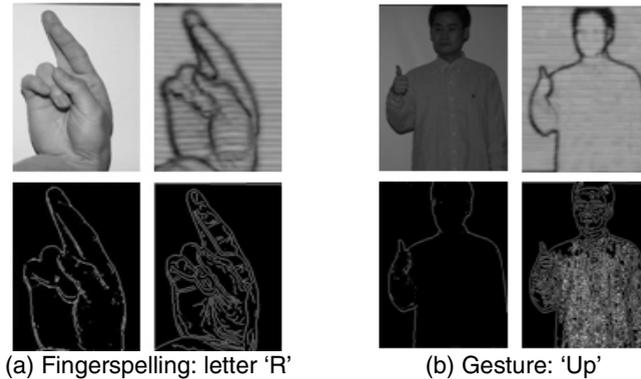


Fig. 8. (a) Detection of depth edges in the case of hand configurations (b) Detecting human body contours. Clockwise from top left: white light image, Gabor amplitude map, depth edges and canny edges.

Fig. 7 (c)–(e) show Gabor amplitude maps in the three cases. Each Gabor amplitude map shows that we cannot detect all the depth edges in the scene using a single structured light image. However, combining the results from the three cases, we can obtain the final Gabor amplitude map as in Fig. 7 (f) where distortion for detection is guaranteed to appear along depth discontinuities in the range of [2.325m, 3m]. Finally, we can get the depth edge map as in Fig. 7 (g). The result shows that this method is capable of detecting depth edges of all the objects located in the detectable range. We have also compared the result with the output of the traditional Canny edge detector (Fig. 7 (h)) where cluttered inner texture edges are also detected. The proposed method accurately detects depth edges by effectively eliminating inner texture edges of the objects.

Fig. 8 (a) shows detection of depth edges in the case of hand configuration. We have also applied our method to human body scenes. Fig. 8 (b) shows the result of detecting of human body contours. Our method effectively suppress inner texture details.

5 Conclusions

We have presented a structured light based approach for effectively suppressing inner texture details with only depth edges contained. We have strategically projected structured light and exploited distortion of light pattern in the structured light image along depth discontinuities. Through a modeled imaging geometry and mathematical analysis, we have also described a method that can guarantee the occurrence of the distortion along depth discontinuities for a continuous range of object location. The second phase of the research using the proposed method is currently under progress for non-photorealistic rendering.

Acknowledgement

This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Biometrics Engineering Research Center (BERC) at Yonsei University.

References

1. Bovik, A.C., Clark, M., Geisler, W.S.: Multichannel Texture Analysis Using Localized Spatial Filters. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 55-73 (1990)
2. Cass, T.A.: Robust Affine Structure Matching for 3D Object Recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 1264-1265 (1998)
3. Chen, Y., Medioni, G.: Object Modelling by Registration of Multiple Range Image. *Image and Vision Computing*, pp. 145-155 (1992)
4. Loncaric, S.: A survey of shape analysis techniques. *Pattern Recognition*, pp. 983-1001 (1998)
5. Weiss, I., Ray, M.: Model-based recognition of 3D object from single vision. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 116-128 (2001)
6. Frohlinghaus, T., Buhmann, J.M.: Regularizing phase-based stereo. In: *Proc. of 13th International Conference on Pattern Recognition*, pp. 451-455 (1996)
7. Hoff, W., Ahuja, N.: Surfaces from Stereo: Integrating Feature Matching, Disparity Estimation, and Contour Detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 11(2), 121-136 (1989)
8. Lee, S., Choi, J., Kim, D., Jung, B., Na, J., Kim, H.: An Active 3D Robot Camera for Home Environment. In: *Proc. of 4th IEEE Sensors Conference* (2004)
9. Scharstein, D., Szeliski, R.: High-Accuracy Stereo Depth Maps Using Structured Light. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1, 195-202 (2003)
10. Raskar, R., Tan, K.H., Feris, R., Yu, J., Turk, M.: Non-photorealistic Camera: Depth Edge Detection and Stylized Rendering Using Multi-Flash Imaging. In: *Proc. of ACM SIGGRAPH Conference*, Vol. 23, pp. 679-688 (2004)
11. Feris, R., Turk, M., Raskar, R., Tan, K., Ohashi, G.: Exploiting Depth Discontinuities for Vision-based Fingerspelling Recognition. *IEEE Workshop on Real-Time Vision for Human-Computer Interaction* (2004)
12. Ma, W., Manjunath, B.S.: EdgeFlow: a technique for boundary detection and image segmentation. *IEEE Trans. on Image Processing* 9, 1375-1388 (2000)