

Subtraction-Based Forward Obstacle Detection Using Illumination Insensitive Feature for Driving-Support

Haruya Kyutoku¹, Daisuke Deguchi², Tomokazu Takahashi³,
Yoshito Mekada⁴, Ichiro Ide¹, and Hiroshi Murase¹

¹ Graduate School of Information Science, Nagoya University,
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan

² Information and Communications Headquarters, Nagoya University,
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601, Japan

³ Faculty of Economics and Information, Gifu Shotoku Gakuen University,
1-38, Nakauzura, Gifu, Gifu, 500-8288, Japan

⁴ School of Information Science and Technology, Chukyo University,
101, Tokodachi, Kaizu-cho, Toyota, Aichi, 470-0393, Japan
kyutoku@murase.m.is.nagoya-u.ac.jp

Abstract. This paper proposes a method for detecting general obstacles on a road by subtracting present and past in-vehicle camera images. The image-subtraction-based object detection approach can be applied to detect any kind of obstacles although the existing learning-based methods detect only specific obstacles. To detect general obstacles, the proposed method first computes a frame-by-frame correspondence between the present and the past in-vehicle camera image sequences, and then registers road surfaces between the frames. Finally, obstacles are detected by applying image subtraction to the registered road surface regions with an illumination insensitive feature for robust detection. Experiments were conducted by using several image sequences captured by an actual in-vehicle camera to confirm the effectiveness of the proposed method. The experimental results shows that the proposed method can detect general obstacles accurately at a distance enough to avoid them safely even in situations with different illuminations.

1 Introduction

Nowadays, development of driving support systems has become an active research topic. Forward obstacle detection from in-vehicle camera images is one of the most important techniques in ITS. This technique is essential to realize a collision-warning system that can reduce the number of traffic accidents significantly. Accordingly, a number of techniques to detect forward obstacles have been developed. Still, most of them detect only specific obstacles such as pedestrians or vehicles. These approaches can improve the detection accuracy by learning image features of the obstacles. However, they cannot detect obstacles that are not considered in the learning phase. Thus, detection of general obstacles is a challenging task because various obstacles exist on a road.

Recently, in-vehicle camera image databases such as Google Street View¹ have become a familiar service. Moreover, the development of higher speed wireless telecommunications and rapid growth of data storage capacities are remarkable. This allows us to collect and store large amount of in-vehicle camera images for driving support purposes.

We proposed a method for detecting forward general obstacles, based on image subtraction between an in-vehicle camera image and its past image captured at the same location [1]. However, this method did not consider the difference of illumination conditions between the present and the past images. This would degrade the detection accuracy in practical situations because the illumination conditions are not the same between the different days/times. Therefore, in this paper, we propose a subtraction-based forward obstacle detection method using an illumination insensitive local feature, and try to improve the accuracy of obstacle detection even when the illumination conditions have changed.

This paper is organized as follows: Section 2 describes related work. Section 3 explains details of the proposed method. Experimental results and discussions are described in Section 4. Finally, Section 5 summarizes the paper.

2 Related Work

Various techniques have been proposed to detect forward obstacles, using different kinds of sensors including millimeterwave radar, stereo, monocular, or infrared camera [1][2][3][4]. However, millimeterwave radars cannot be used for detecting small obstacles due to their low spatial resolutions. On the other hand, stereo cameras still have hurdles with their calibration [5]. It may be difficult to obtain accurate point correspondences between the images. Meanwhile, the techniques using infrared cameras are specialized for pedestrian detection to assist a driver's visibility at nighttime.

Obstacle detection methods with a monocular camera has been actively researched. However, most of these methods are learning-based methods, so they detect only specific obstacles such as pedestrians or vehicles [3][4]. Here, one of the most effective approaches to detect general obstacles is the subtraction-based approach [1]. This method detects obstacles by comparing present and past in-vehicle camera images. Assuming that the past sequence does not include any obstacle, the detection can be done by a background subtraction strategy. The advantage of this method over the learning-based methods is that it does not require to learn image features of the obstacles beforehand, so it can handle arbitrary objects. Again, since this method uses simple local features in the subtraction operation, such as brightness value of each pixel, the detection accuracy is strongly affected by the difference of illumination conditions between the present and the past sequences. Unlike the background subtraction using a fixed camera [6], the illumination conditions between the input and the background images are quite different in the case of subtraction between the present

¹ Google Street View, <http://maps.google.com/>

and the past in-vehicle camera images because they are usually taken on different days/times. Therefore, the proposed method introduces an illumination insensitive local feature for the subtraction. As another approach, the optical-flow-based approach can also detect general obstacles, but it has a difficulty when detecting obstacles that are not moving.

The proposed method needs past in-vehicle camera images with no obstacles. There are some techniques [7] for constructing and updating an in-vehicle camera image database. The database contains images corresponding to each road on a street map. We assume that past images could be obtained from such databases. Furthermore, some methods have been developed to remove obstacles from an in-vehicle camera image database [8][9], which could be used to obtain obstacle-free image sequences.

3 Proposed Method

The cruising speed and trajectory of each vehicle may be different between the sequences. To detect obstacles based on subtraction between the present and the past in-vehicle camera image sequences captured along the same road, the proposed method performs the following two processes:

1. Frame-Level and Pixel-Level Registrations

The cruising speed and the trajectory differ between the sequences. Therefore, this process first finds a frame in the past sequence that was taken at the same location as the present frame, and then finds correspondences of pixels in road surface regions between the present and the past frames.

The proposed method performs this process in the same manner as the method introduced in [1].

2. Subtraction-Based Obstacle Detection Using an Illumination Insensitive Local Feature

This process detects obstacles based on differences between the local features extracted from road surface regions of the present and the past frames. Here, we propose an illumination insensitive local feature.

3.1 Frame-Level and Pixel-Level Registrations

Here, we briefly explain the registration process. Detailed information of the process is described in [1].

Finding Frame Correspondences. To obtain the frame correspondences, a Dynamic Time Warping (DTW) method is applied to the present and the past image sequences by introducing a penalty calculated from the relationship between the two cameras. This penalty is defined as

$$\frac{1}{|e_x| + \alpha}, \quad (1)$$

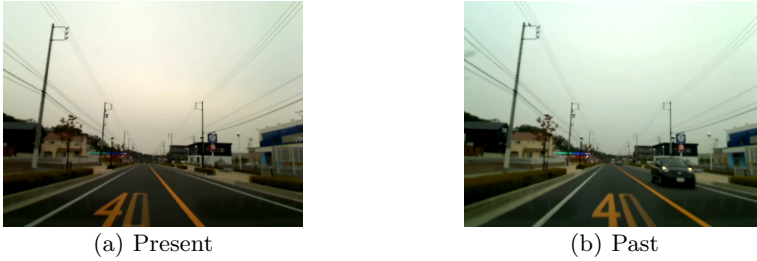


Fig. 1. Corresponding frames in the present and past image sequences

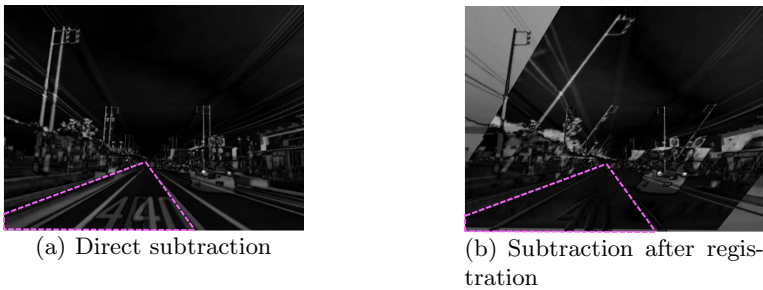


Fig. 2. The subtraction results of corresponding frames in Fig. 1

where e_x is the horizontal position of the epipole between the present and the past frame, and α is a positive constant. Based on the theory of the epipolar geometry analysis, the epipole is calculated from corresponding point-pairs between the frames. The penalty term is robust against the difference in running positions of a vehicle and the occlusions caused by obstacles.

Registration of Road Surfaces. This step performs registration of road surface regions between the corresponding frames. It is difficult to detect obstacles by directly applying an image subtraction operation between the corresponding frames because of the difference in the running positions between the frames. Figure 1 shows an example of corresponding frames obtained by the frame-level registration. The result of the image subtraction is shown in Fig. 2(a), where we can see a spatial gap. To compensate for the gap, this method performs the pixel-level registration of the road surface regions. The subtraction result after the registration is shown in Fig. 2(b). It is assumed that the road surface is flat, and the obstacle detection area is restricted to the road surface region. This step makes correspondences of pixels in the road surface regions between the present and the past images by a projective transformation. The projective transformation matrix is calculated from the epipolar lines and the road boundary lines.

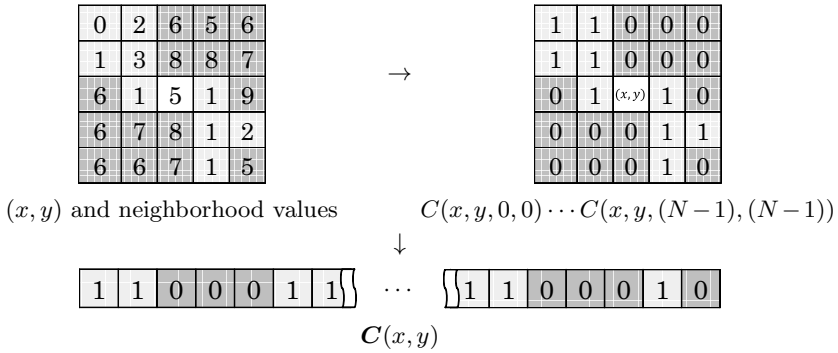


Fig. 3. Example of Census Transform ($N = 2$)

3.2 Subtraction-Based Obstacle Detection Using an Illumination Insensitive Feature

The proposed method detects obstacles by the subtraction of road surface regions after the registration process. In this paper, we define an “obstacle” as the image difference compared to a past road surface where obstacles do not exist. For each location in the road surface region, the method calculates the difference of local features between the present and the past images, and then significantly different regions are detected as obstacles. Here, the features for measuring the difference is very important because simple features, such as the brightness, are easily affected by the illumination condition changes. Hence, we propose an illumination insensitive local feature that is robust for different illumination conditions.

Illumination Insensitive Local Feature. Census Transform [10] is a method for extracting local features from an image, which encodes the relationship between a pixel and its neighbors. The method was first proposed to calculate the cost of stereo matching. Similar to Local-Binary-Patterns (LBP) [11], the local feature extracted by the Census Transform is robust against the illumination changes [12]. The Census Transform constructs a vector $C(x, y)$ for each pixel. The LBP uses the pixels only along a ring around the target pixel, although the Census Transform uses all pixels in a $(2N + 1) \times (2N + 1)$ region around the target pixel. Thus, the vector consists of $4N(N + 1)$ elements. Each element is calculated as

$$C(x, y, i, j) = \begin{cases} 1 & I(x, y) > I(x + i, y + j) \\ 0 & \text{otherwise} \end{cases}, \quad (2)$$

where $-N \leq i \leq N$ and $-N \leq j \leq N$ except $i = j = 0$. $I(x, y)$ represents the pixel value at position (x, y) . Figure 3 shows an example of the construction of a census transform vector. The difference between the two vectors is calculated based on Hamming distance. However, this representation may be unsuitable for

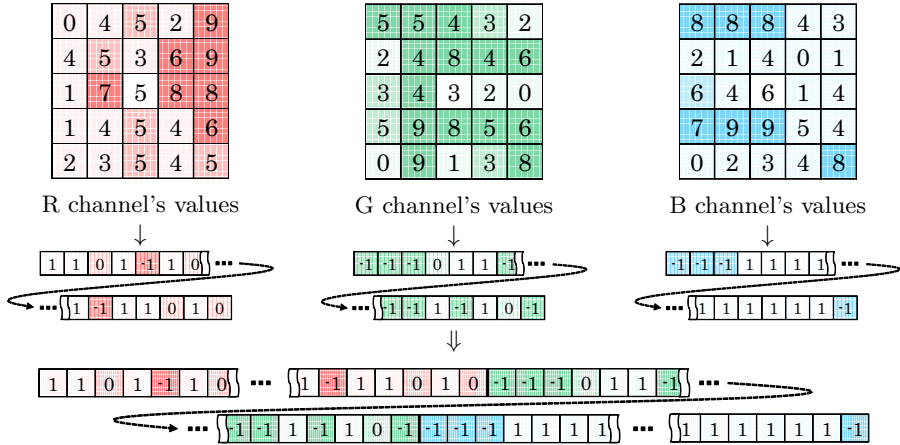


Fig. 4. Example of modified Census Transform ($N = 2$ and $\beta = 0$)

feature extraction from a road surface region, because pixel values are almost equal in most parts of the road surface. Therefore, this paper introduces trinary representation [13] that is used for LBP. Thus, instead of Eq. (2), we use the following representation.

$$C(x, y, i, j) = \begin{cases} 1 & I(x, y) > I(x + i, y + j) + \beta \\ -1 & I(x, y) < I(x + i, y + j) - \beta \\ 0 & \text{otherwise} \end{cases}, \quad (3)$$

where β is a small positive value that absorbs small changes of pixel values in a road surface region. We also consider that color information would work effectively for obstacle detection. Therefore, the proposed method extracts a vector from each color channel, and concatenates the vectors as shown in Fig. 4. We define the difference between two vectors as the sum of absolute differences of each element pair.

Integration of Results from Multiple-Frames To improve the reliability of the detection results, the proposed method integrates results obtained from successive frames. This reduces false detections caused by noises and registration errors. Here, we employ a simple strategy to integrate multiple-frame information as shown in Fig. 5.

First, the proposed method computes a spatio-temporal subtraction image by using the past successive T subtraction images (Fig. 5(a)), and then detects highly different pixels as obstacle candidates. Next, each candidate is tracked in the temporal direction, and tracking trajectories are obtained (Fig. 5(b)). Assuming that the vehicle and the obstacle move linearly for a very short time in the spatio-temporal subtraction image domain, candidates having straight-line trajectories are extracted as the final results (Fig. 5(d)).

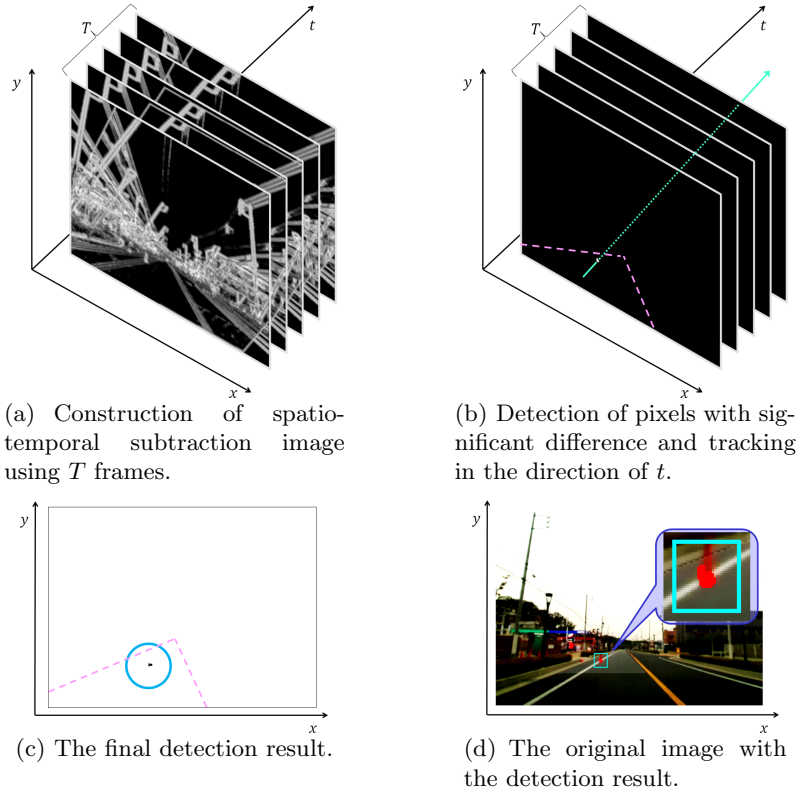


Fig. 5. Overview of integration of multiple-frame information

4 Experiments

To confirm the effectiveness of the proposed method, obstacle detection experiments were conducted by using in-vehicle camera images.

4.1 Experimental Condition

To acquire images used for the experiments, we mounted a web camera “Logi-cool Qcam® Pro 9000” on the windshield of a vehicle as shown in Fig. 6. The image size was 640×480 pixels, and the frame rate was 15 fps. We prepared a total of six sequences by driving along the same 150 meters road six times. Four of them including obstacles were used as the present sequences, and the others which did not include any obstacle were used as the past sequences. The total number of frames in the present sequences was 879, and 329 of these sequences included obstacles in the distances from 12 to 80 m. The obstacles were a pedestrian crossing the road, a street-parking vehicle, a forward running vehicle, a pylon, and a ball. On the other hand, the past sequences were captured under



(a) From outside (b) From inside

Fig. 6. Car-mounted in-vehicle camera



(a) Present (b) Past

Fig. 7. Example of large illumination change between present and past frames

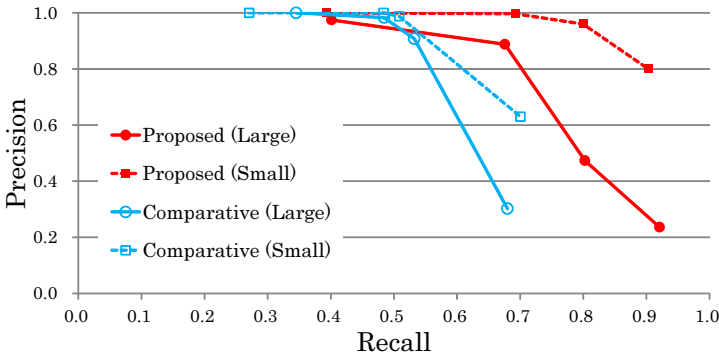


Fig. 8. The precision-recall curves of the obstacle detection

different illumination conditions that had small and large differences from the illumination condition of the present sequences as shown in Fig. 7.

We evaluated the detection accuracy by the following criteria:

$$\text{Precision rate} = \frac{\# \text{ of true-positives}}{\# \text{ of true-positives} + \# \text{ of false-positives}},$$

$$\text{Recall rate} = \frac{\# \text{ of true-positives}}{\# \text{ of obstacles}},$$

$$\text{F-value} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Here, 4-connected pixels of the detection result were merged to one region, and used for calculating the above criteria. The true-positives were counted for each obstacle in the image, even if multiple detected regions existed on an obstacle. On the other hand, the false-positives were counted for each detected region. We compared the detection accuracy between the proposed method and the method described in [1]. The comparative method detects obstacles based on simple subtraction of brightness values after the registration of road surface regions. We set the parameters used in the proposed method as $N = 3$ and $\beta = 5$. As for

Table 1. The accuracy comparison when maximum F-values were obtained

Illumination change	Method	F-value	Precision	Recall
Small	Proposed	0.87	0.96	0.80
	Comparative	0.67	0.99	0.51
Large	Proposed	0.77	0.89	0.68
	Comparative	0.67	0.91	0.53

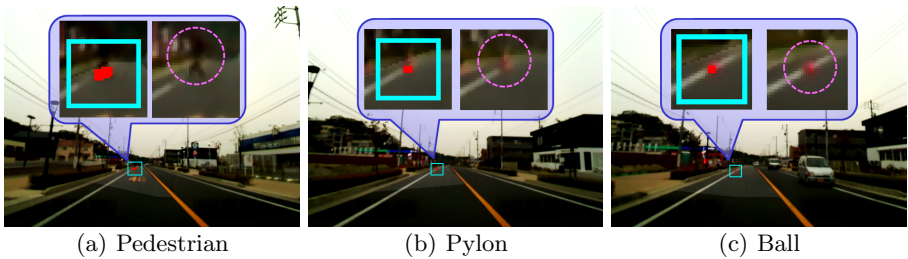


Fig. 9. Examples of detection results for obstacles approximately 50m ahead. The rectangles represent the detected regions. The pair of images in the balloon in each image is the close-ups of the detected region (left) and its original image (right).

the integration parameter T , we used the same value for the proposed method and the comparative one as $T = 5$. We determined these values based on the results of preliminary experiments.

4.2 Results

Figure 8 shows the precision-recall curves of obstacles closer than 80 m. The solid lines show the accuracy when large illumination change occurred, while the dashed lines show the accuracy in the case of small illumination change. Table 1 shows the accuracy comparison when the maximum F-values were obtained in Fig. 8. In addition, Fig. 9 shows the detection results of distant obstacles.

4.3 Discussion

From Table 1, we confirmed that the proposed method provided higher F-values than those of the comparative methods in both cases when small and large illumination changes exist between the image sequences. Moreover, as Fig. 8 shows, for example, in case when the recall rate is about 0.7, the proposed method keeps high precision rates over 0.9 even if the sequences had a large illumination difference, while the recall rates of the comparative methods dropped to less than 0.8. These results proved the effectiveness of the proposed method.

Additionally, the proposed method could detect small obstacles such as a ball or a pylon at a distance of more than 44 m, as seen in Fig. 9. It was confirmed that distant obstacles could be detected by the proposed method. Especially,

Fig. 9(c) demonstrates the remarkable ability of the proposed method. The proposed method successfully detected a ball with a diameter of only 20 cm at a distance of 46 m. This means that small obstacles could be detected far enough to avoid safely even when cruising at a speed of 60 km/h.

5 Summary

We proposed the method for detecting general obstacles on a road by subtracting the present and the past in-vehicle camera images with an illumination insensitive local feature. Compared to the existing learning-based approach that detects only specific obstacles, the proposed image-subtraction-based approach can detect any kind of obstacle. To achieve this, the proposed method first performed frame-level and pixel-level registrations of the present and the past image sequences, and then detected obstacles based on the subtraction of the road surface regions with an illumination insensitive local feature.

To demonstrate the effectiveness of the proposed method, obstacle detection experiments were carried out by applying the proposed method to several in-vehicle camera image sequences. The experimental results show that the proposed method can detect general obstacles accurately at a distance far enough to avoid them safely even under large difference of illumination conditions.

Acknowledgement. Parts of this research were supported by JST CREST and MEXT Grant-in-Aid for Scientific Research. This work was developed based on the MIST library (<http://mist.murase.m.is.nagoya-u.ac.jp/>).

References

1. Kyutoku, H., Deguchi, D., Takahashi, T., Mekada, Y., Ide, I., Murase, H.: On-road obstacle detection by comparing present and past in-vehicle camera images. In: Proc. 12th IAPR Conf. on Machine Vision Applications, pp. 357–360 (2011)
2. Liu, F., Sparbert, J., Stiller, C.: IMM-PDA vehicle tracking system using asynchronous sensor fusion of radar and vision. In: Proc. 2008 IEEE Intelligent Vehicles Symposium, pp. 168–173 (2008)
3. Nishida, K., Kurita, T.: Boosting with cross-validation based feature selection for pedestrian detection. In: Proc. 2008 IEEE World Congress on Computational Intelligence, pp. 1251–1257 (2008)
4. Mitsui, T., Fujiyoshi, H.: Object detection by joint features based on two-stage boosting. In: Proc. 9th IEEE Int. Workshop on Visual Surveillance 2009, pp. 1169–1176 (2009)
5. Collado, J.M., Hilario, C., de la Escalera, A., Armingol, J.M.: Self-calibration of an on-board stereo-vision system for driver assistance systems. In: Proc. 2006 IEEE Intelligent Vehicles Symposium, pp. 156–162 (2006)
6. Kawanishi, Y., Mitsugami, I., Mukunoki, M., Minoh, M.: Background image generation by preserving lighting condition of outdoor scenes. *Procedia—Social and Behavioral Sciences* 2, 129–136 (2010)

7. Sato, J., Takahashi, T., Ide, I., Murase, H.: Change detection in streetscapes from GPS coordinated omni-directional image sequences. In: Proc. 18th IAPR Int. Conf. on Pattern Recognition, vol. 4, pp. 935–938 (2006)
8. Boehm, J.: Multi-image fusion for occlusion-free facade texturing. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 35, 867–872 (2004)
9. Uchiyama, H., Deguchi, D., Takahashi, T., Ide, I., Murase, H.: Removal of moving objects from a street-view image by fusing multiple image sequences. In: Proc. 20th IAPR Int. Conf. on Pattern Recognition, pp. 3456–3459 (2010)
10. Zabih, R., Woodfill, J.: Non-parametric Local Transforms for Computing Visual Correspondence. In: Eklundh, J.-O. (ed.) *ECCV 1994, Part II. LNCS*, vol. 801, pp. 151–158. Springer, Heidelberg (1994)
11. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distributions. *Int. Journal on Computer Vision* 29, 51–59 (1996)
12. Hirschmuller, H., Scharstein, D.: Evaluation of stereo matching costs on images with radiometric differences. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 31, 1582–1599 (2009)
13. Tan, X., Triggs, B.: Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans. on Image Processing* 19, 1635–1650 (2010)