Light Source Intensity Adjustment for Enhanced Feature Extraction

Francisco J. Castro-Martínez*, Mario Castelán, and Ismael López-Juárez

Centro de Investigación y de Estudios Avanzados del I.P.N.
Robotics and Advanced Manufacturing Group,
Ramos Arizpe, Coahuila, 25900, México
{francisco.castro,mario.castelan,ismael.lopez}@cinvestav.edu.mx

Abstract. We explore the automatic adjustment of an artificial light source intensity for the purposes of image-based feature extraction and recognition. Two histogram-based criteria are proposed to achieve this adjustment: a two-class separation measure for 2D features and a Gaussian distribution measure for 2.5D features. To this end, the light source intensity is varied within a fixed interval as a camera captures one image for each intensity variation. The image that best satisfies the criteria for feature extraction is tested on a neural-network based recognition system. The network considers information related to both 2D (contour) and 2.5D shape (local surface curvature) of different objects. Experimental tests performed during different times of the day confirm that the proposed adjustment delivers improved feature extraction, extending the recognition capabilities of the system and adding robustness against changes in ambient light.

Keywords: Object recognition, neural networks, feature extraction.

1 Introduction

It is well known that illumination is a highly important factor in computer vision tasks. Basically, if knowledge is to be obtained from visual data, little can be inferred from a scene whose objects are illuminated with a too intense or too low light source. The impact of illumination in computer vision has been explored for many purposes. The Photometric Stereo Method (PSM) [12] may be the best (and probably oldest) example to get some benefit from changes in light source direction in order to obtain 3D information of the surface observed by the camera. More recently, stereopsis has also exploited the idea of illumination variations in order to redefine the correspondence problem under the Light Transport Constancy (LTC) constraint [11], with encouraging results. Variations in intensity of the light source have also proved to be useful in obtaining the 3D

* The authors wish to thank CONACyT for the financial support given to the project through grant No. 61373 and also for the scholarship granted to Francisco J. Castro-Martínez during his Master studies.

J.A. Carrasco-Ochoa et al. (Eds.): MCPR 2010, LNCS 6256, pp. 50–60, 2010.

surface of objects, as shown in [10], where Light Fall-off Stereo (LFS) is employed to estimate 3D surface as a light source increasingly departs from the illuminated object. Researchers on the field of robotics have also shown interest in the automatic positioning of light sources so as to maximize luminance and contrast in the image for the purposes of tracking [8].

Studies have demonstrated that the appearance of convex surfaces such as human faces is more dependable on changes in direction and intensity of the light source than on changes in facial pose. [9]. In general, many computer vision problems make assumptions on (generally Lambertian) reflectance properties of objects as well as good illumination conditions, i.e., a negligible effect of ambient light and inter-reflections. In reality, even when the reflectance of a material can be approximated with a Lambertian reflectance model, the intensity of the light source plays a key role in revealing appropriate levels of reflectance for each pixel, therefore facilitating the task of feature extraction from pixel values.

In this paper, we do not address the problem of 3D shape recovery nor automatic light source positioning, rather, we focus on the adjustment of the light source intensity for the task of 2D and 2.5D feature extraction. To this end, two main criteria based on gray level histogram are used. In a first criterion, we draw on the ideas of Otsu's method [13] in order to decide whether a light source intensity is suitable for background segmentation. The second criterion searches for the best Gaussian distribution of a gray level histogram and is related to the construction of shape-index histograms for local surface orientation estimation [5]. This improved-from-illumination feature extraction benefits recognition rates in a neural network system.

The organization of the paper is as follows: in Section 2, the object recognition system is roughly described. Section 3 provides an explanation of the two main criteria used for feature extraction and light source intensity regulation. Experiments on the behavior of these criteria as well as recognition performance are depicted in Section 4. Finally, conclusions are given in Section 5.

2 The Object Recognition System

The object recognition system consists of a neural network which has been trained with feature vectors reflecting 2D (contour) and 2.5D (local surface curvature) shape of the objects. Roughly, the main concepts related to this recognition system are explained below.

The *FuzzyARTMAP neural network* is based on the Adaptive Resonance Theory (ART) which was developed by Stephen Grossberg and Gail Carpenter at Boston University. In Fuzzy ARTMAP there are two modules ART_a and ART_b and an inter-ART module "map field" that controls the learning of an associative map from ART_a recognition categories to ART_b recognition categories [1]. The map field module also controls the match tracking of ART_a vigilance parameter. A mismatch between Map field and ART_a category activated by input **a** and ART_b category activated by input **b** increases ART_a vigilance by the minimum amount needed for the system to search for, and if necessary, learn a new ART_a

category whose prediction matches the ART_b category. The search initiated by the inter-ART reset can shift attention to a novel cluster of features that can be incorporated through learning into a new ART_a recognition category, which can then be linked to a new ART prediction via associative learning at the Map field. The algorithm uses a preprocessing step, called complement coding which is designed to avoid category proliferation. Similar to ART-1, a vigilance parameter measures the difference allowed between the input data and the stored pattern. Therefore this parameter is determinant to affect the selectivity or granularity of the network prediction. For learning, the FuzzyARTMAP has 4 important factors: Vigilance in the input module (ρ_a), vigilance in the output module (ρ_b), vigilance in the Map field (ρ_{ab}) and learning rate (β). These were the considered factors in this research with values of $\rho_a = 0.93$, $\rho_b = 1$, $\rho_{ab} = 0.95$ and $\beta = 1$.

The method obtains the object's contour using metric properties such as the perimeter, area and centroid information in order to form the so-called *Boundary Object Function (BOF)*[2]. The BOF is a 2D descriptor vector that contains the euclidean distance between the object's contour and its centroid. The vector is formed by 180 elements obtained from the measurement of the distance between a contour point and the centroid every two degrees. The BOF and its starting point is easily determined for geometrical figures such as circles, but in complex shapes the procedure is more involved. The starting point is important since this is also a reference for the Neural Network Pattern Recognition system.

The Shape-From-Shading method (SFS) [3] consists primarily of obtaining the orientation of the surface due to local variations in brightness that is reflected by the object, in other words, the intensities of the greyscale image is taken as a topographic surface. The surface normal representation has also been called the 2.5D sketch by Marr [6]. SFS is known as an ill-posed problem, causing ambiguity between what has a concave and convex surface, which is due, among other things, to poorly illuminated surface patches [4]. In this paper, we make use of the particular SFS method proposed in [5], the surface normal is rotated in accordance with a local surface curvature measure known as the *Shape Index* (SI) [7]. This measure provides an idea of how patches over the surface correspond to degrees of concavity and convexity. The SFS method in [5] has proved to be useful for object recognition when using SI histograms. In this work, SI-histograms are also used to build feature vectors encoding 2.5D information for the neural network.

3 Improved Image Feature Extraction

In this section we describe the methods used to extract relevant information for 2D and 2.5D shape classification. Both methods are based on intensity histogram analysis. The distribution of the intensities in the images is then helpful to determine the goodness of each image for the purposes of 2D and 2.5D feature extraction. We commence with the description of the *two-class separation criterion* for BOF. The idea here is to provide the BOF with a feature extraction method able to deliver correct information, i.e., a good segmentation between

object and background that allows the classifier to sharply determine the different 2D shapes (contour) of the objects. To this end, we borrow ideas from Otsu's method [13] to automatically perform histogram shape-based image thresholding, reducing a graylevel image to a binary image. The algorithm assumes that the image to be thresholded contains two classes of pixels (e.g. foreground and background) then calculates the optimum threshold separating those two classes so that their combined intra-class variance is minimal. Let us define the global and intra-class variances as

$$\sigma_G^2 = \sum_{i=0}^{L-1}(i - m_G)^2 P_i \quad \text{and} \quad \sigma_B^2 = P_1(m_1 - m_G)^2 + P_2(m_2 - m_G)^2, \tag{1}$$

respectively, where L is the number of gray levels (tipically 256), m_G is the average gray level in the image and P_i is the probability of the i_{th} gray level in the image. The probabilities P_1 and P_2 of the two potential classes C_1 and C_2 are defined respectively as $P_1(k) = \sum_{i=0}^{k} P_i$ and $P_2(k) = 1 - P_1(k)$ with $0 < k < L - 1$. The average probabilities m_1 and m_2 of C_1 and C_2 are defined respectively as $m_1(k) = \frac{1}{P_1(k)} \sum_{i=0}^{k} iP_i$ and $m_2(k) = \frac{1}{P_2(k)} \sum_{i=k+1}^{L-1} iP_i$.

Note that the optimal classes C_1 and C_2 are separated by the k_{th} gray level. Therefore, the optimal threshold in Otsu's method is the gray level value that makes σ_B^2 maximal. The relationship between global and intra-class variances is given by $\eta(k) = \sigma_B^2(k)/\sigma_G^2$. This means that the intra-class variability must be evaluated for every gray level as $\sigma_B^2(k) = (m_G P_1(k) - m(k))^2 / P_1(k)(1 - P_1(k))$, with $m(k)$ being the cumulative average probability up to the k_{th} gray level. The optimal threshold is then defined as k^* satisfying

$$\sigma_B^2(k^*) = max(\sigma_B^2(k)). \tag{2}$$

For the total number of changes in light source intensity, the criterion

$$\eta_b(n) = \frac{\sigma_B^2(k^*, n)}{\sigma_G^2(n)}, 1 < n < N, \tag{3}$$

is recorded for every n_{th} intensity variation, where N is the number of variations. The n_{th}^* image maximizing the criterion is then selected using the equation $\eta_b(n^*) = max(\eta_b(n))$. Finally, the optimal threshold k^* is used to segment the selected image. Once the contour of the object is at hand, the BOF neural network feature vector can be built.

In the following, the *Gaussian distribution criterion* for SFS is described. As far as 2.5D shape is concerned, shape index histograms are constructed for each image in order to represent local surface shape for the different objects. The aim of this representation is to allow the recognition system to determine whether the surface of the observed object corresponds to a pyramidal or curved one (see Fig. 1 (left)). As this data is calculated from surface normal estimations obtained through a geometrical SFS method [5], it is important to guarantee that the light source illuminating the objects minimizes unwanted artifacts in the image. Typically, these artifacts appear due to specular reflections and as

a consequence of overly illuminated patches on the surface of the object. Other undesired effects may be caused by the opposite situation, i.e., by poorly illuminating surface patches. In any case, a light source intensity too high or too low leads to pixel values located near the limits of the gray level range. As a result, surface orientation may be wrongly estimated due to the ambiguity of such gray level values. An appropriate distribution of gray levels among the objects in the observed image is therefore sought in order to provide a solution to reducing the number of ambiguously valued pixels. To overcome this problem, we propose selecting the image whose gray level histogram best fits a Gaussian distribution with mean μ and standard deviation σ. The criterion can be defined as

$$\eta_s = \sum_{x=1}^{L} \left|\left| P(x) - \left(\frac{1}{\sigma\sqrt{2\pi^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \right) \right|\right|^2, \tag{4}$$

where $P(x)$ is the probability of the gray level x to occur in the image. The criterion $\eta_s(n), 1 < n < N$, where N is the total number of changes in light source intensity is recorded for each intensity variation. Finally, the n^*_{th} image best satisfying the criterion is selected as the optimal image for building the SFS neural network feature vector, as shown by the equation $\eta_s(n^*) = min(\eta_s(n))$.

4 Experiments

Eight different hand-made pieces were used for experimental tests. The different pieces attempt to emulate common objects in a manufacturing environment. The pieces can be observed in Figure 1 (left). As shown in the figure, the combination of features in the set consists of four 2D shapes (square, triangle, cross, star) and two 2.5D shapes (curved, pyramidal). The acquisition platform is shown in Figure 1 (right). The platform is conformed by a halogen lamp connected to a dimmer in order to regulate light source intensity. An extra halogen lamp was used for experiments considering a fixed light source intensity. A camera was synchronized to capture an image along each intensity variation. The gain of the camera was set to zero and the auto-adjustment feature was off in order to isolate the effect of illumination only to the influence of the lamps. The tests

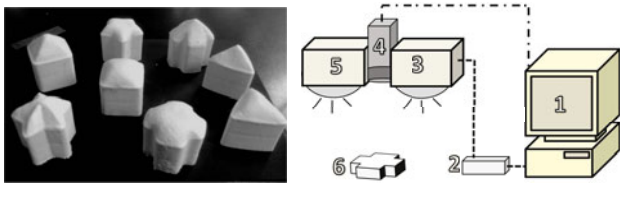

Fig. 1. The image acquisition platform. The eight different pieces used for the experiments (left). The acquisition platform (right) shows a diagram of the different devices (labeled as numbers) used for capturing images: one fixed(5), one variable(3) light source, one camera(4), a dimmer(2), a computer(1) and one piece(6).

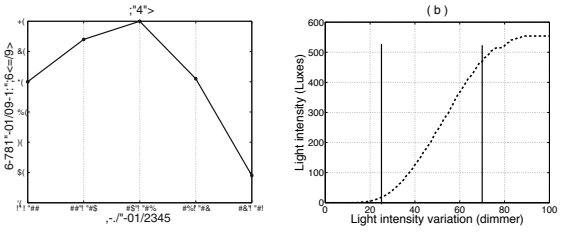

Fig. 2. Behavior of ambient light and artificial light source intensities. In (a), the intensity of ambient light is shown, in luxes, as a function of daytime interval. In (b), the intensity of the artificial light source, in luxes, is shown as a function of the resistance of the dimmer. The linear response of the light is bounded between the two vertical lines in the diagram.

were carried out during five intervals of time in one day. The duration of each interval was of two hours, starting at 9 and finishing at 19 hrs. A light meter was used to measure the luminance of both ambient light (i.e., the amount of light received around the scene area, coming from bulbs in the working area as well as from natural external light) and the artificial light source (the halogen lamps).

This information is shown in Figure 2. The left diagram of the figure depicts a plot of the ambient light intensity behavior during the five time intervals. Note that the peak intensity was reached at the sunniest time (13-15 hrs. interval), while a considerably lower intensity was reached near the evening (17-19 hrs. interval). The right diagram of the figure depicts a plot of the lamp intensity against resistance of the dimmer, for the 9-11 hrs. interval. In the plot, the ambient light is subtracted. Note that only the linear response, bounded between two vertical lines, was taken into account for the experiments. Therefore, only ten equally spaced dimmer values (from 25 to 70) were synchronized with the camera.

4.1 Optimality Analysis for the Criteria η_b and η_s

The image acquisition procedure involved in the experiments is described in this section. One single piece was placed on three main locations: the center, the top-left and bottom-right corners of the viewing plane of the camera, over a dark card. At the center, the piece was arbitrarily rotated on its own axis twice for a total number of four events. For each event, an image was taken 10 times from the synchronized-with-dimmer camera. From the ten image set, the corresponding criteria η_b and η_s were recorded. In the next figures, results on a total number of 1600 images = 8 pieces $\times$ 4 events $\times$ 10 variations in intensity light $\times$ 5 intervals of time are shown. We start the analysis in Figure 3, where the evolution of the 2D shape criterion η_b is shown as a function of light intensity variation. From left to right, results over the 2D shapes of square, triangle, cross and star are presented. For each shape, the average of all the corresponding light source intensity variations during the five intervals of time are depicted.

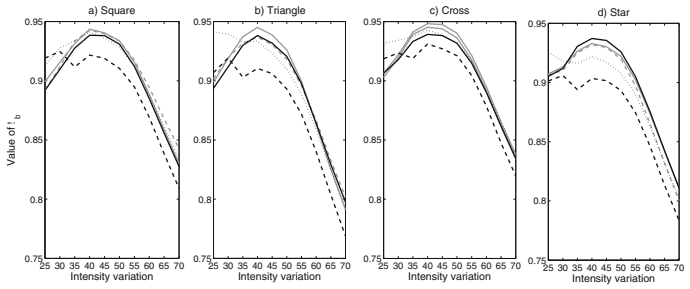

Fig. 3. Evolution of η_b as a function of dimmer values and for the different time intervals. Results on the 2D shapes are labeled accordingly on each diagram. The different intervals of time are represented using the following lines: 9 - 11 hrs., solid gray; 11 - 13 hrs., dashed gray; 13 - 15 hrs., solid black; 15 - 17 hrs., dotted black and 17 - 19 hrs., dashed black.

The different intervals are represented using the following lines: 9 - 11 hrs., solid gray; 11 - 13 hrs., dashed gray; 13 - 15 hrs., solid black; 15 - 17 hrs., dotted black and 17 - 19 hrs., dashed black. From the figure, it is noticeable that the highest value of η_b, and therefore the optimal criterion for binary segmentation happens near the dimmer resistance area of 40 - 50, with a relatively steady behavior for the first three intervals of time. A less predictable behavior is nonetheless shown during the last two intervals of time (dashed and dotted black lines), which may be explained by the rapid descent of ambient light during those hours of the day. Note how, for all the cases, the value of η_b tends to descend once a maximum value has been reached.

The evolution of the 2.5D shape criterion η_s is shown in Figure 4. Results during different intervals of time are presented in a similar way as in Figure 3. Here, the diagrams are organized in accordance with the curved and pyramidal surface features of the pieces. Although the overall behavior appears to increase once a minimum has been reached, this minimum seems to require a bigger amount of light intensity for the last two intervals of time. Such effect may be again explained as a consequence of the lack of ambient light towards the evening. Note how, in the figure, the different shapes of the pieces seem to be responsible of most of the variability in both the minimum value of η_s and the required light source intensity. This fact suggests that the adjustment of the lightning requirements of the scene depends on the particular 2D and 2.5D features of the piece rather than using a fixed illumination for every object and for all times of the day.

4.2 Recognition

In order to show the importance of a dynamic selection of the criteria η_b and η_s, we compare the performance of a fixed light source intensity with an empirical fixed threshold for segmentation against a variable light source intensity with optimal citeria η_b and η_s as well as optimal threshold k^* (see Eq. 2). The fixed light

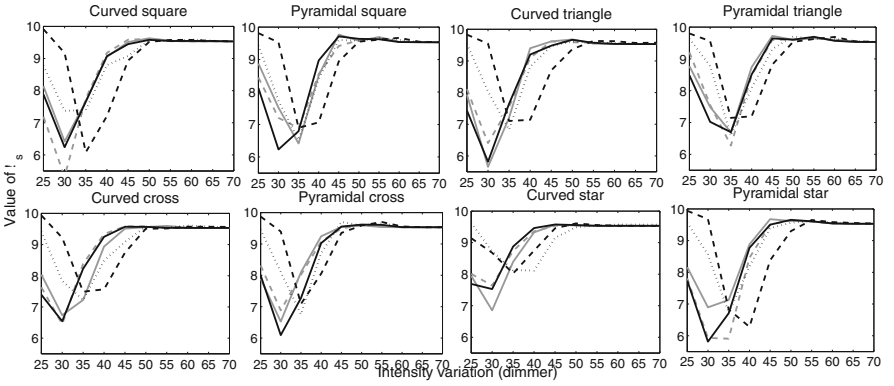

Fig. 4. Evolution of η_s as a function of dimmer values and for the different time intervals. Results on the 2.5D shapes are labeled accordingly on each diagram. The different intervals of time are represented as in Figure 3.

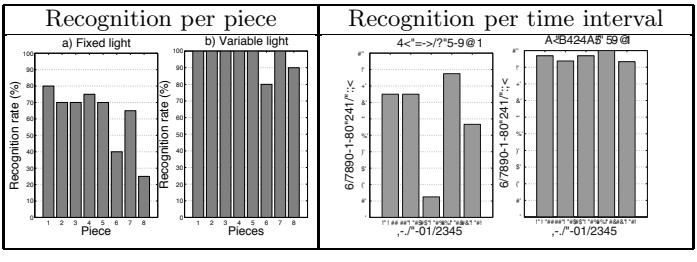

Fig. 5. Recognition rates under fixed and variable light source intensities. The two panels contain bar diagrams with results concerning fixed and variable light source intensity. For the fixed light, a fixed threshold for BOF segmentation was also used. Recognition rates per piece (left) and per time interval (right) are shown.

was registered to emit around 100 luxes (without ambient light) at the start of the experiments, which can be comparable to using the dimmer value of 40 on the variable light source. The recognition experiment is roughly explained as follows. For each event, ten intensity variations are synchronized with a camera shot. The image with optimal criterion η_b and threshold k^* is selected, its contour extracted, and its BOF feature vector built. Once the optimal segmentation is at hand, the background is removed from all of the ten images and the optimal criterion η_s sought in order to select the best image for building the SFS feature vector. Finally, a single image is captured using the fixed light source and a fixed threshold. The image is segmented and its BOF feature vector is generated. Background removal is then performed and the SFS feature vector built. A piece is successfully recognized when both 2D and 2.5D features are correctly determined. For example, the system should be able to determine whether the piece is a curved square or a pyramidal cross, among other possibilities in the set.

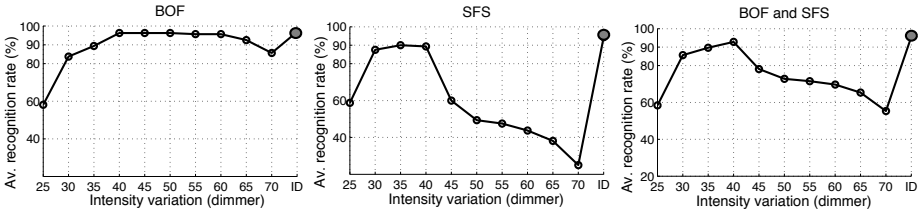

Fig. 6. Average recognition rates for all time intervals per light source intensity variation. From left to right, results concerning BOF, SFS and BOF+SFS are shown as a function of light source intensity variation, respectively. The recognition rate from optimal values η_b and η_s is shown as a gray circle at the end of the line plot.

Recognition rates per piece and per time interval are shown in Figure 5 (left) and Figure 5 (right), respectively. In the figure, the different pieces are labeled as follows: 1-2, square; 3-4, triangle; 5-6, cross; and 7-8, star. Even and odd numbers correspond to curved and pyramidal shapes, respectively. The figure reveals that both fixed and variable approaches struggle to recognize more complex shapes such as crosses and stars, with a particular emphasis in pyramidal shapes. Nonetheless, the advantage of varying light source intensity over a fixed light is clearly demonstrated. As far as the time interval analysis is concerned, it is evident that the fixed light intensity approach encounters more difficulty during the 13 - 15 hrs. period, which may be caused by the peak in ambient light during that interval of time. On the contrary, recognition rates for the variable light intensity approach remain relatively steady along the five intervals of time. This fact justifies the need of an automatic adjustment of both light source intensity and optimal threshold to the particular features of the observed piece and the particular illumination conditions of the different hours of the day.

To conclude the recognition analysis, Figure 6 presents a panorama of the average recognition rate as a function of light source intensity variation (value of dimmer). The figure is divided into three plots, where recognition on BOF (left), SFS (center) and BOF + SFS (right) are shown separately. For this figure, results using the fixed light source and fixed threshold are not included. Instead, we focused on showing the recognition rate per each light source intensity variation and using an optimal threshold k^* for segmentation, i.e., as if none of the criteria η_b and η_s had been calculated for image selection. However, at the end of each line plot, results on recognition with optimal criteria η_b and η_s are depicted with a gray circle. Note how for the BOF-based recognition, a stable high recognition rate is achieved from 40 to 60 dimmer values, which suggests that a good separation of object from background may be achieved from images obtained within a range of light source intensities. However, for results related to SFS-based recognition, a smaller region (from 30 to 40 dimmer values) shows a lesser recognition rate, which suggests that the extraction of accurate local surface curvature from images may be too sensitive to light source intensity fluctuations. Predictably, for the BOF + SFS case, only the dimmer value of 40 provides a recognition rate slightly higher than 90%, reducing the trustable

region to a single dimmer value. This observation suggests that fixing the light source intensity at a dimmer value of 40 may be sufficient for obtaining good recognition results. Nonetheless, although this fixed value of 40 shows a relatively steady recognition rate, ambient light is a factor which cannot be controlled. In other words, what may be useful in a sunny day, may not be so on a darker day. Interestingly, the advantage of using the optimal criteria η_b and η_s is revealed as the highest and most stable of all recognition rates. This justifies again the necessity of an automatic adjustment of the light source intensity if robustness against uncontrolled changes in ambient light is to be provided in recognition systems based on image feature extraction.

5 Conclusions

We have addressed the problem of compensating illumination requirements for the purposes of 2D and 2.5D shape feature extraction. To this end, we have proposed to adapt two histogram-based criteria to the particular needs of the observed scenario during different times of the day. The proposed criteria have shown to improve stability and robustness to ambient light changes. Particularly, this enhanced feature extraction from the automatic adjustment of light source intensity has extended the capabilities of a neural network based recognition system. As future work, we plan to investigate the inclusion of additional light sources as well as the adaptation of other criteria for different feature extraction tasks, i.e., the optimal illumination conditions for stereo matching and other 2D feature extraction such as corners and contours for a more complex class of scenes.

References

1. Carpenter, G.A., Grossberg, S., Markuzon, N., Reyno1ds, J.H., Rosen, D.B.: FuzzyARTMAP: A neural network architecture for incremental learning of analog multidimensional maps. IEEE Trans. on Neural Networks 3(5), 698–713 (1992)
2. Peña-Cabrera, M., Lopez-Juarez, I., Rios-Cabrera, R., Corona-Castuera, J.: Machine Vision Approach for Robotic Assembly. Assembly Automation 25(3), 204–216 (2005)
3. Horn, B.K.P.: Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View. PhD thesis, MIT (1970)
4. Brooks, M.: Two results concerning ambiguity in shape from shading. In: AAAI-83, pp. 36–39 (1983)
5. Worthington, P.L., Hancock, E.R.: Object Recognition Using Shape-from-Shading. IEEE Trans. on Pattern Analysis and Machine Intelligence 23(5), 535–542 (2001)
6. Marr, D., Nishihara, H.K.: Representation and Recognition of the Spatial Organization of Three Dimensional Shapes. Proc. Royal Society of London, B. 200, 269–294 (1978)
7. Koenderink, J., Van Doorn, A.: Surface shape and curvature scale. Image and Vision Computing 10, 557–565 (1992)
8. Collewet, C.: Modeling complex luminance variations for target tracking. In: Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, pp. 1–7 (2008)

9. Moses, Y., Adini, Y., Ullman, S.: Face Recognition: the Problem of Compensating for Changes in Illumination Direction. In: Proc. European Conference on Computer Vision, pp. 286–296 (1994)
10. Liao, M., Wang, L., Yang, R.: Gong. M, Light Fall-off Stereo. In: IEEE Conference on Computer Vision and Pattern Recognition (2007)
11. Wang, L., Yang, R., Davis, J.E.: BRDF Invariant Stereo Using Light Transport Con- stancy. IEEE Trans. on Pattern Analysis and Machine Intelligence 29(9), 1616–1626 (2007)
12. Woodham, R.J.: Photometric method for determining surface orientation from multiple images. Optical Enginnering 19(1), 139–144 (1980)
13. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Sys., Man., Cyber. 9, 62–66 (1979)