Incorporating Angular Ratio Images into Two-Frame Stereo Algorithms

Pablo-Arturo Martínez-González and Mario Castelán*

Centro de Investigación y de Estudios Avanzados del I.P.N.
Robotics and Advanced Manufacturing Group,
Ramos Arizpe, Coahuila, 25900, México

Abstract. The Light Transport Constancy (LTC) establishes that the reflectance ratio obtained from two different illumination variations remains constant for any given view of the observed scene. In [1], LTC was proposed as a rank constraint for solving the correspondence problem in multiple view stereo. In two-frame stereo, the simplest setting for LTC requires only two illumination variations and a single light source. Under this scenario, the rank constraint can be formulated through ratio images and standard stereo algorithms be applied in order to obtain a disparity map. Unfortunately, a ratio image may be subject to saturated pixel values, which may diminish the quality of the disparity maps. To overcome this problem, we introduce a post-processing operation based on the slope angles related to the ratio values. Experiments show that the new angular ratio images are more robust and deliver improved disparity maps.

Key words: Light Transport Constancy, two-frame stereo, ratio images.

1 Introduction

Acquiring the three dimensional surface of objects is an important problem in Computer Vision, since a 3D surface simplifies the modeling of the appearance of the object. A 3D surface can be obtained using contact devices such as laser scanners. Other possibilities imply information provided by one or more cameras. This methodology is known as image-based 3D shape recovery. Although the image-based approach is appealing, the nature of the image acquisition process makes the input images prone to errors. In this sense, lighting manipulation represents a way to posing constraints on image-based shape recovery techniques. For instance, the intensity of the light can be regulated in order to obtain 3D shape. This is the core idea of Light Fall-off Stereo (LFS)[7], where a number of images is gathered from a stationary camera as the illumination source moves away from the scene. Based on the inverse square law for light intensity, the ratio images are directly related to scene depth from the perspective of the light source. Controlling the geometric position of the light source represents a

* This work has been supported by Project Conacyt Ciencia Bsica 61593.

Special Issue in Advances in Pattern Recognition
Research in Computing Science 44, 2009, pp. 21-28

different alternative to face the problem. The Photometric Stereo Method (PSM) [6] is the classical technique dealing with this consideration. Here, a single camera captures images while the direction of a light source moves around a fixed pose object.

When more than one camera is required, binocular stereo (two-frame stereo) is the image-based 3D shape recovery method with the simplest setting. Here, only two cameras capturing a still scene are needed and the correspondence problem is solved between the two views in order to obtain depth information. Unfortunately, when Lambertian reflectance and color/brightness constancy are not observed, the calculation of correspondences becomes a difficult task. In binocular stereo, the manipulation of lighting has also been proposed. For example, projective geometry laws may be applied if structured light patterns are projected over the surface of the object [5]. This is normally done using a projector, but colored laser rays can also be projected if more accurate results are needed. Another approach based on lighting variations is the Helmholtz stereopsis. This method allows matching of arbitrary Bidirectional Reflectance Distribution Functions (BRDF) and uses the reflectance function reciprocity as an invariant [8, 9] . By collocating point light sources with each camera it is possible to record reciprocal pairs using two different lighting conditions. Due to reciprocity the reflected light to the cameras will be equal. This method, however, requires the light sources to be collocated with the optical center of each camera.

Recently, the Light Transport Constancy (LTC) [1] has been proposed as a correspondence clue in multiple-view stereo. LTC is used to formulate a rank constraint matching cost when the scene is observed in several lighting variations (i.e., changes in light intensity). LTC establishes that the reflectance ratio obtained from two different illumination variations remains constant for any given view of the observed scene. LTC does not require the position of light sources to be precisely calibrated or even known. In two-frame stereo, the simplest setting requires only two illumination variations. Under this scenario, the rank constraint can be formulated through ratio images, and standard stereo algorithms be applied in order to obtain a disparity map. Unfortunately, a ratio image may be subject to saturated pixel values, noise and occlusions, which may diminish the quality of the disparity maps. To overcome this problem, in this paper, we introduce a post-processing operation based on the slope angles related to the ratio values.

The paper is organized as follows: in Section 2, the LTC is introduced; in Section 3, the use of angular measures for improving ratio images is explained; experiments are presented in Section 4 and finally conclusions are described in Section 5.

2 Light transport constancy

The Light Transport Constancy states that the percentage of light reflected by a surface patch (the BRDF) remains constant for any given viewing direction of

a static scene. Following the explanation in [1], let us call a particular point in the scene x_i. This point will reflect light to the cameras C_1 and C_2 according to $I_{C_j}(x_i) = L(x_i)R(x_i, L, C_j)$, where $I_{C_j}(x_i)$ is the reflected intensity in the direction of C_j from the point x_i, $L(x_i)$ is the incident light intensity at point x_i, and $R(x_i, L, C_j)$ is the reflectance function or BRDF at point x_i, indexed by the vectors in the direction of L and C_j. The Lambertian assumption asserts that the reflected light is equal in the directions of C_1 and C_2, i.e., the BRDF is shared and $R(x_i, L, C_1) = R(x_i, L, C_2)$. Thus we have $I_{C_1}(x_i) = I_{C_2}(x_i)$. However this relation will not in general hold for arbitrary (non-Lambertian) BRDFs. Light transport constancy assumes that the surface reflectance function, $R(x_i, L, C_j)$, remains constant under variable illumination. If we vary the lighting conditions, so that the incident illumination varies by a factor of $k(x_i)$, then the observed reflected light, $I_{C_j}(x_i)$, will also vary by a factor of $k(x_i)$, as

$$I_{C_j}(x_i) = k(x_i)L(x_i)R(x_i, L, C_j) \tag{1}$$

Wang et al [1] have shown how LTC can be used in multiple-view stereo to impose a rank constraint on the matrix

$$I_{CV} = \begin{pmatrix} I_{C_1V_1} & I_{C_2V_1} & \cdots & I_{C_mV_1} \\ I_{C_1V_2} & I_{C_2V_2} & \cdots & I_{C_mV_2} \\ \vdots & \vdots & \ddots & \vdots \\ I_{C_1V_n} & I_{C_2V_n} & \cdots & I_{C_mV_n} \end{pmatrix}, \tag{2}$$

where $I_{C_jV_k}$ is the observed grayscale value by the j_{th} camera under the k_{th} lighting variation. Note that, for the sake of simplicity, we have omitted the notation (x_i). However, each of the remaining equations in the paper is related to a single pixel at position x_i. The matrix with minimum rank is therefore sought, i.e., if LTC is observed through the different camera viewing positions and lighting variations, then the dimension of the column space of the matrix I_{CV} should be minimal. The rank constraint only holds when the number of light sources is less than both the number of lighting variations and the number of cameras. Then the rank of I_{CV} is at most the number of light sources. Since I_{CV} will be corrupted with noise, it is impossible to calculate rank exactly. The Singular Value Decomposition of I_{CV} may be used for rank approximation. A matrix with most of their energy in the first few principal components is preferred and moments can be used to approximate the notion of minimum rank, as

$$\mathcal{M} = \sum_i i \cdot \sigma_i^2 \Big/ \sum_i \sigma_i^2, \tag{3}$$

where σ_i are the singular values of I_{CV}. For multiple-view and multiple-lighting stereo, the minimum score is used as matching cost.

3 Improved ratio Images and the two-frame case

Let us now consider the simplest setting for LTC-based stereo: a single light source and two cameras. For each pixel in the left and right images, the intensities

observed by the first lighting variation can be explained in terms of the intensities observed by the second lighting variation, as $I_{C_1V_2} = I_{C_1V_1}k_1$ and $I_{C_2V_2} = I_{C_2V_1}k_2$. Therefore the relation between lighting variations is given by the ratio:

$$\frac{I_{C_1V_1}}{I_{C_1V_2}} = \frac{I_{C_1V_1}}{I_{C_1V_1}k_1} = k_1, \tag{4}$$

$$\frac{I_{C_2V_1}}{I_{C_2V_2}} = \frac{I_{C_2V_1}}{I_{C_2V_1}k_2} = k_2. \tag{5}$$

The matrix of intensities, I_{CV} can be now defined as

$$I_{CV} = \begin{pmatrix} I_{C_1V_1} , I_{C_2V_1} \\ I_{C_1V_2} , I_{C_2V_2} \end{pmatrix} \tag{6}$$

Note that LTC only holds if the second singular value of I_{CV} is zero. This means that the minimum rank of I_{CV} is one (the number of light sources) if and only if $k_1 = k_2$. Minimizing the second singular value is therefore equivalent to minimizing Eq. 3. The so-called *ratio image* is a function $R(v_1, v_2, x_i) = k_i$, where v_1 and v_2 are two different lighting variations. The ratio image is only defined for the two-frame/two lighting variations scenario and at most two ratio images can be recorded for a given stereo image pair, as in Eq. 4 and Eq. 5. Note how the minimization of Eq. 3 can also be carried out using a simple absolute difference matching cost over the ratio image pair. In this sense, a wide variety of two-frame stereo algorithms [2] provides many ways to calculate dense disparity maps through ratio images.

Unlike the classical grayscale (or color) image pair which usually assumes brightness/color constancy, the ratio images rely on the light transport constancy and have proved to provide improved disparity maps [1]. For these reasons, Wang et al. have pointed out the potential use of ratio images in the future as the most feasible and robust way to approach the two-frame stereo problem. Using ratio images for such a task may borrow ideas from the vast literature related to two-frame stereo algorithms [2, 4, 3]. For example, from the the taxonomy of Scharstein and Szelisky [2], different matching costs, aggregation support and optimization approaches can be applied to add robustness in the calculation of dense disparity maps from ratio images.

In practice, unfortunately, the intensities of the pixels do not necessarily observe $k_1 = k_2$. This is due to several reasons, among which insufficient lighting and a poor camera response are the most common. Moreover, in some regions of the image, i.e., where specularities and edges occur, the ratio is likely to be either a value close to zero or an overly saturated value, i.e., much greater than 100%. The unwanted effect of these pixel values can be reduced if the ratios are redefined as:

$$k_1' = |\tan^{-1}(k_1)|, \qquad k_2' = |\tan^{-1}(k_2).|, \tag{7}$$

where $|\cdot|$ is the absolute value. The angular data constrain the ratio values from the interval $[0, \infty)$ to the interval $[0°, 90°]$. We can now define the new ratio as $R'(v_1, v_2, x_i) = k_i'$ which will be referred to as the *angular ratio image*.

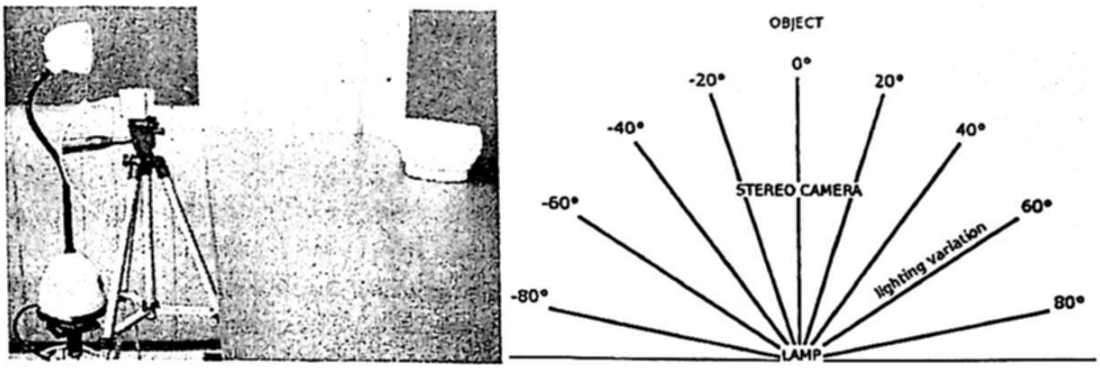

Fig. 1. The general setting and the illumination variation. The left image shows the general acquisition setting: the stereo camera, the desk lamp and the object. The lamp was mounted on a rotating ruler on a tripod, so that rotations could be measured. The right image shows a top-down sketch of the setting, where the degree of separation between each camera shot can be seen.

4 Experiments

For image acquisition, a Bumblebee stereo camera, $9cm$ baseline was used. The size of the grayscale images was 640×480 pixels. A halogen-bulb desk lamp was mounted on a rotating ruler in order to capture illumination variations around a range of 180^o, with 20^o increment as shown in Figure 1 (right). A picture of the general setting is shown in Figure 1 (left).

Before the ratio values are converted into image values (grayscale) a normalization operation has to be performed. Let $R(v_1, v_2, x_i)$ be the ratio values obtained from a pair of left images with lighting variations (v_1, v_2) (the same observations holds for its corresponding right pair), i.e., using Eqs. 4 and 5. A ratio image is stored after the following filter is applied:

$$R(v_1, v_2, x_i) = \begin{cases} \tau & \text{if } k_i \geq \tau \\ k_i & \text{otherwise} \end{cases} \tag{8}$$

where τ is a cut ratio value. Once the high values have been filtered, the final image is generated from the normalized values:

$$R(v_1, v_2, x_i) = \frac{R(v_1, v_2, x_i)}{\tau}. \tag{9}$$

Note that, once Eq. 8 is applied, τ becomes the maximum value of $R(v_1, v_2, x_i)$. As far as angular ratio images are concerned, i.e. $R'(v_1, v_2, x_i)$, Eq. 8 is not required, since Eq. 9 can be directly applied with $\tau = \pi/2$ (radians). Once a ratio (or angular ratio) image is generated, dense disparity maps are calculated using the standard platform developed in [1]. The *sum of absolute differences* (SAD) and the mean filter 9×9 window were respectively used as matching cost and aggregation support parameters. The disparity maps were finally calculated under the *winner-takes-all* (WTA) criteria.

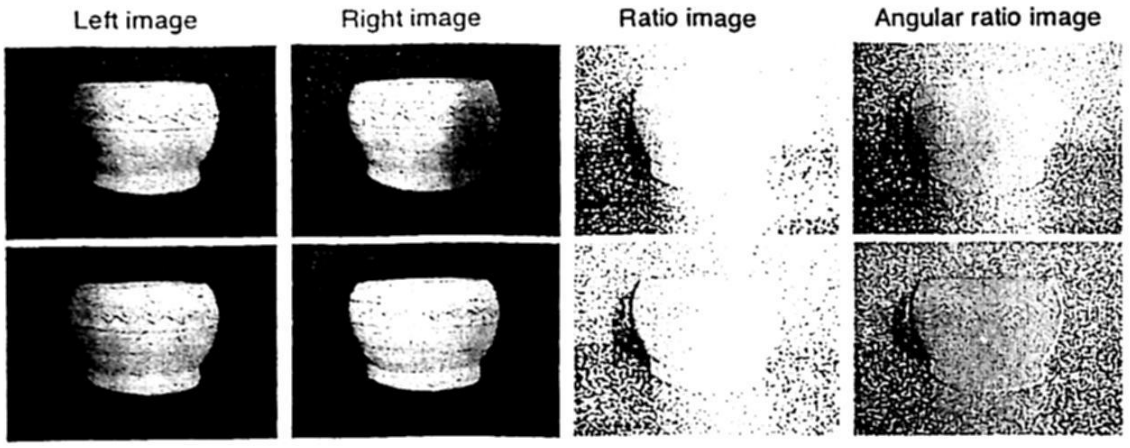

Fig. 2. Ratio images and angular ratio images. From top to bottom, the different rows of the figure show the lighting variation pairs $(-60°, 60°)$ and $(-40°, 40°)$. The left-camera image, right-camera image, ratio image and angular ratio image are shown row-wise.

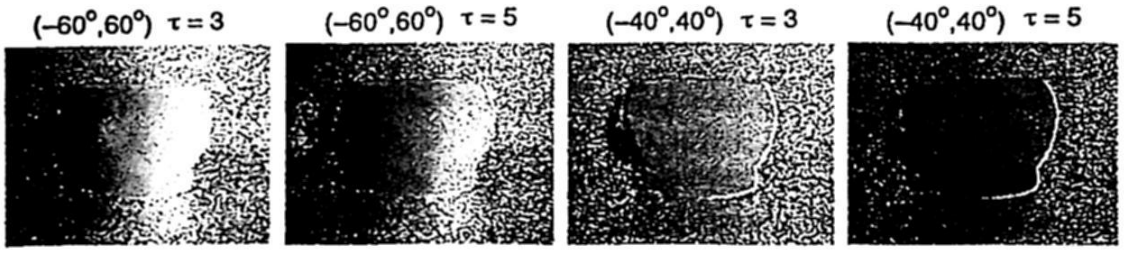

Fig. 3. Varying τ for ratio image generation. Normalization results for $\tau = 3$ and $\tau = 5$ are shown with lighting variations as depicted in the labels of the figure.

The experimental analysis commences with Figure 2, where the difference between ratio and angular ratio images can be appreciated. From top to bottom, the different rows of the figure show the lighting variation pairs $(-60^o, 60^o)$ and $(-40^o, 40^o)$. The left-camera image, right-camera image, ratio image ($\tau = 1$) and angular ratio image are shown row-wise. Recall that the ratio and angular ratio images are calculated for a single view (i.e., left or right camera images) and two variations. From the figure, it is noticeable that the angular ratio images reveal a more robust adjustment of values than the ratio images, where the cut value has set the ratios to saturated values. This is a main problem in ratio image generation, i.e., choosing an optimum cut value. This effect can be visualized in Figure 3, where different values of τ are applied to different lighting variations. Again, a generalized optimum value of τ is not clear, as the figure shows that $\tau = 5$ favors the ratio obtained from the pair $(-60^o, 60^o)$, but over-darkens the ratio obtained from the pair $(-40^o, 40^o)$.

As far as disparity results are concerned, these are depicted in Figure 4. Here, the first row presents disparity maps while the second row shows a line of interest with constant $y - axis$ along the disparity map. For visualization purposes, a

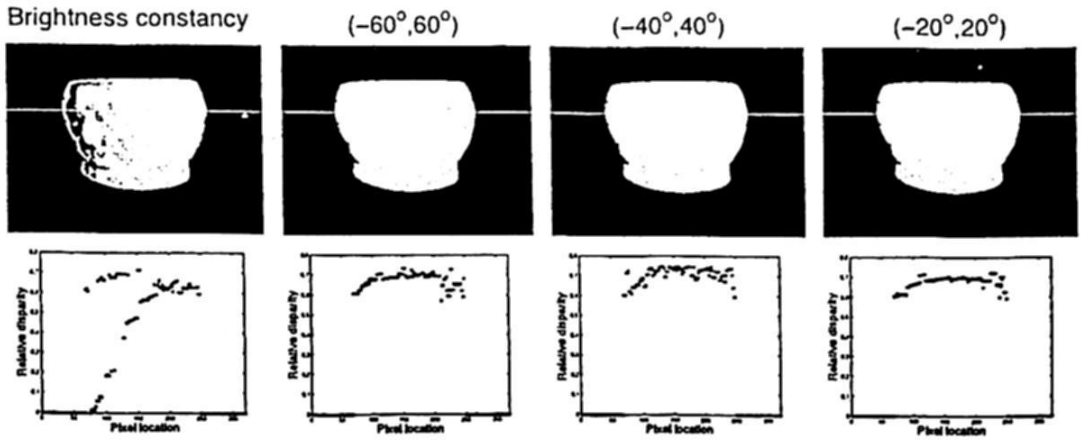

Fig. 4. Disparity results on brightness constancy and angular ratio images. The first row shows disparity maps, the second row shows a line of interest with constant $y - axis$ along the disparity map.

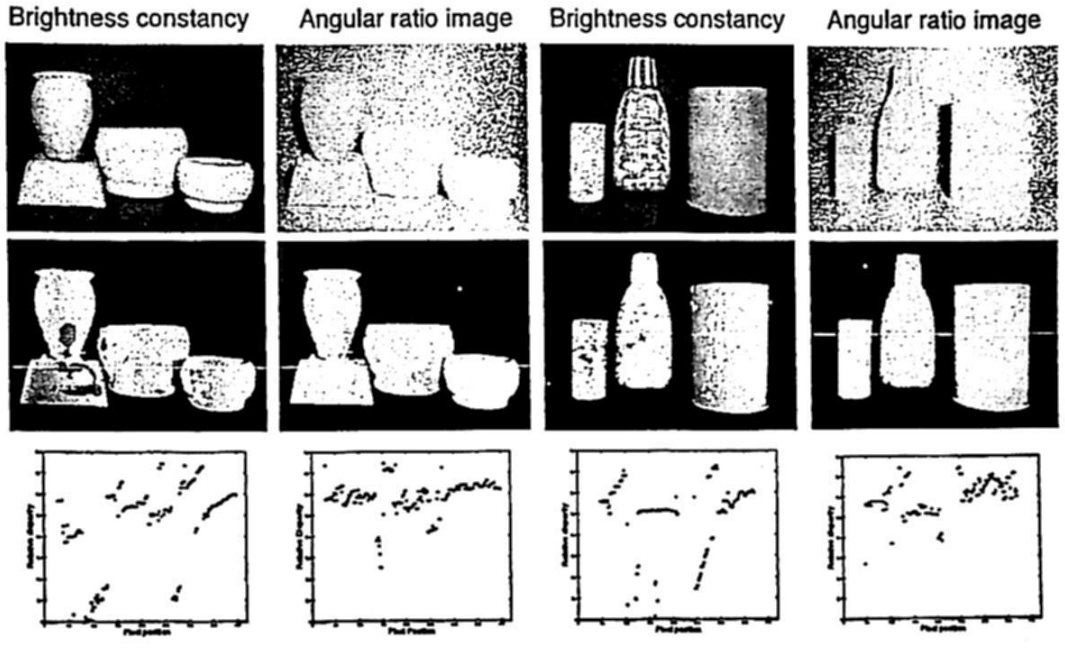

Fig. 5. Additional scenes. The first row presents the kind of input images used, the second row presents the obtained disparity maps. The third row presents a line of interest.

mask has been applied for isolating the objects of interest from the background. The first row presents results for brightness constancy, i.e., neither ratio nor angular ratio images are used here. Instead, the usual grayscale image left/right image pair is used as stereo input. The rest of the columns present results from angular ratio images with lighting variation pairs of $(-60^o, 60^o)$, $(-40^o, 40^o)$ and $(-20^o, 20^o)$. As expected, there is an improvement in disparity calculation for the angular ratio images over the brightness constancy, i.e., the disparities are

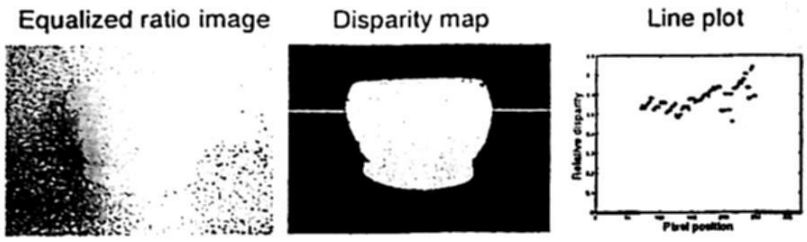

Fig. 6. Results on a histogram-equalized ratio image. From left to right: equalized ratio image, disparity map, line of interest along the disparity map.

located over more continuos regions. Interestingly, there is little difference between the results related to angular ratio images, which suggests the angles may represent a robust way to obtain similar disparity maps through the different lighting variations. Additional scenarios are presented in Figure 5, again, the disparity results for the angular ratio images outperform those relying on the brightness constancy assumption.Finally, we explore using histogram equalization as an alternative adjustment function for ratio images and the outcome of the experiment is shown in Figure 6. Here, a standard histogram equalization operation was performed on the raw ratio image with variation $(-60^o, 60^o)$. Despite the equalized ratio image look similar to its angular ratio image counterpart (top right corner of Figure 2), the disparity results are rather different, favoring again the use of angular ratio images.

5 Conclusions

A new way to define ratio images for LTC two-frame stereo has been presented. The new angular ratio image does not require cut ratio values and suggests to be robust to illumination variations. This facilitates experimenting with known two-frame stereo algorithms under the LTC stereo setting. As future work, we plan to study different aggregation and optimization methods on these images.

References

1. L. Wang, R. Yang, and J. E. Davis, BRDF Invariant Stereo Using Light Transport Constancy. *IEEE Trans. Pattern Anal. and Mach. Intell.*, Vol. 29, No. 9, pp. 1616–1626, (2007).
2. D. Scharstein and R. Szelisky. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *Int. Journal of Comp. Vis.*, 47(1):7–42, 2002.
3. H. Hirschmüller and D. Scharstein. Evaluation of Stereo Matching Costs on Images with Radiometric Differences. *IEEE Trans. Pattern Anal. and Mach. Intel.*, 31(9):582–1599, 2009.
4. M. Gong, R. Yang, L. Wang and M. Gong. A Performance Study on Different Cost Aggregation Approaches Used in Real-Time Stereo Matching. *Int. Journal of Comp. Vis.*, 75(2):283–296, 2007.
5. D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. *Proc. IEEE ICCV*, 2003.
6. R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Enginnering* 19(1):139–144, 1980.
7. M. Liao, L. Wang, R. Yang and M. Gong. Light Fall-off Stereo. *Proc. IEEE CVPR* 1–8, 2007.
8. S. Magda, D. J. Kriegman, T. Zickler, and P. N. Belhumeur. Beyond Lambert: Reconstructing surfaces with arbitrary BRDFs. *Proc. IEEE ICCV*, 2, pp. 391, 2001.
9. T. E. Zickler, P. N. Belhumeur, and D. J. Kriegman. Helmholtz stereopsis: exploiting reciprocity for surface reconstruction. *Int. Journal of Comp. Vis.*, 49:215, 2002.