Incorporating Angular Ratio Images into Two-Frame Stereo Algorithms

Pablo Arturo Martínez González and Mario Castelán

Robótica y Manufactura Avanzada, CINVESTAV – Unidad Saltillo
Carretera Saltillo-Monterrey Km. 13.5, C.P. 25900, Ramos Arizpe, Coah.
Mexico

{pablo.martinez, mario.castelan}@cinvestav.edu.mx

Abstract. Light Transport Constancy (LTC) asserts that the reflectance ratio obtained from two different illumination variations remains constant for any given view of the observed scene. LTC was proposed in [21] 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 can be applied in order to obtain a disparity map. Unfortunately, a ratio image may be subject to saturated pixel values, and this may diminish the quality of disparity maps. To solve this problem, as a first contribution in this work, we propose a post-processing operation based on slope angles related to the ratio values. Experiments show that new angular ratio images are more robust and deliver improved disparity maps. A second contribution of this paper consists in performing an experimental evaluation of angular ratio images under the standard test bed for two-view stereo algorithms, i.e., using different aggregation and optimization approaches. The results of our research are consistent with previously reported conclusions for two-view stereo surveys. It means that LTC may benefit from a vast variety of existent methods to solve the two-view stereo problem.

Keywords. Light Transport Constancy, two-frame stereo, ratio images.

Incorporación de las imágenes de relación angular en algoritmos de estéreo binocular

Resumen. La Constancia de Transportación de la Luz (LTC) establece que la relación de reflectancia obtenida de dos diferentes variaciones en iluminación permanece constante para cualquier vista dada de la escena observada. En [21] LTC fue propuesta como una restricción de rango para resolver el problema de la correspondencia en estéreo de múltiples vistas. En estéreo binocular, el escenario más simple para LTC requiere solamente dos variaciones en iluminación y una sola fuente de luz. Bajo este escenario, la restricción de rango puede ser formulada a través de las imágenes de relación y los algoritmos estéreo estándar son aplicados con el objeto de obtener un mapa de disparidad. Desafortunadamente, una imagen de relación puede ser sujeta a valores de pixeles saturados, los cuales pueden disminuir la calidad de los mapas de disparidad. Para superar este problema, como una primera contribución en este artículo presentamos una operación de post-procesado basada en los ángulos de pendiente relacionados a los valores de relación. Los experimentos muestran que las nuevas imágenes de relación son más robustas y ofrecen mejores mapas de disparidad. Como una segunda contribución, realizamos evaluación experimental de las imágenes de relación angular bajo una cama de pruebas estándar para algoritmos de estéreo binocular, i.e., usando diferentes enfoques de agregación y optimización. Los resultados de esta investigación son consistentes con conclusiones previamente reportadas en estudios sobre estéreo. Esto significa que LTC puede beneficiarse de una vasta variedad de métodos existentes para el problema de estéreo binocular.

Palabras clave. Constancia de Transportación de la Luz, estéreo binocular, imágenes de relación.

1 Introduction

Acquiring a three-dimensional surface of objects is an important problem in computer vision because a 3D surface simplifies modeling of the object's appearance. 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 input images prone to errors. With respect to this, lighting manipulation represents a way to pose constraints on image-based shape recovery techniques. For instance, the light intensity can be regulated in order to obtain a 3D shape. This is the core idea of Light Fall-off Stereo (LFS) [12], where a number of images are gathered from a stationary camera as the illumination source moves away from the scene. Based on the inverse square law for light intensity, ratio images are directly related to the scene depth from the perspective of the light source. Controlling the geometric position of the light source is another option to attack the problem. The Photometric Stereo Method (PSM) [6, 22] is a classical technique in this respect. Here, a single camera captures images while the light source moves around the object in a fixed pose.

When more than one camera is required, binocular stereo (two-frame stereo) is obtained using the image-based 3D shape recovery method with the simplest setting. Here, only two cameras are needed to capture a still scene, 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, calculation of correspondences becomes a difficult task. In binocular stereo, the manipulation of lighting has also been proposed. For example, structured light patterns may be projected over the surface of an object [17]. 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 Helmholtz stereopsis. This method allows matching arbitrary Bidirectional Reflectance Distribution Functions (BRDF) and uses reflectance function reciprocity as an invariant [13, 24]. By collocating point light sources with each camera, it is possible to record reciprocal pairs using two different lighting conditions. Due to the reciprocity, the reflected light to the cameras will be equal. This method, however, requires the light sources to be collocated with respect to the optical center of each camera.

Other approaches combine photometric and geometric cues. For instance, in multiple-view photometric stereo, a number of images of an object are obtained from multiple viewpoints under varying lighting directions. Here, the silhouette of the object is used to recover camera motion. The correspondence problem, however, is not solved by means of illumination variations [10, 18].

Recently, Light Transport Constancy (LTC) [21] 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 (changes in light intensity). LTC asserts 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 can 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 disparity maps.

On the other hand, unlike classical grayscale (or color) image pairs which usually assume brightness/color constancy, ratio images rely on LTC and have demonstrated to provide improved disparity maps. For these reasons, Wang *et al*. pointed out a potential use of ratio images in future as the most feasible and robust way to deal with the two-frame stereo problem. Using ratio images for such a task may borrow ideas from the extensive literature related to two-frame stereo algorithms [4, 8, 11, 16]. For example, from the taxonomy of Scharstein and Szelisky [16], different matching costs, aggregation support, and optimization approaches can be applied to make calculation of dense disparity maps from ratio images more robust.

1.1 Aim and Contribution

The aim of this paper is to provide a new insight into the use and performance of the LTC constraint in binocular stereo approaches. The topic has been practically unexplored due to the original multi-view formulation of the LTC constraint, which differs from the binocular approach in the methodologies used to address the correspondence problem. With respect to our objective, the contribution of this paper is twofold. First, we introduce a post-processing operation based on the slope angles related to ratio values, which we call the *angular ratio image*. This

Table 1. Notation used in this paper

Expression	Meaning
x_i	The i_{th} pixel in an image
C_j	The j_{th} camera
$I_{C_j}(x_i)$	Reflected intensity (grayscale value) in the direction of C_j from the pixel at x_i
$L(x_i)$	Incident light source intensity at pixel position x_i
$R(x_i, L, C_j)$	Reflectance function (BRDF) at pixel position x_i
$k(x_i)$	Variation factor between two different illumination intensities at pixel position x_i
I_{CV}	The matrix formed by the grayscale values registered by the j_{th} camera under the k_{th} lighting variation and for the pixel position x_i
$I_{C_j V_k}$	The grayscale value observed by the j_{th} camera under the k_{th} lighting variation and for the pixel position x_i
$R(v_1, v_2, x_i) = k_i$	Ratio image expressed as a function of parameters v_1 and v_2 (two different light source intensities) and the pixel position x_i
$k'_1 = |\tan^{-1}(k_1)|$	Angular variation factor between the two lighting variations registered by the camera C_1
$k'_2 = |\tan^{-1}(k_2)|$	Angular variation factor between the two lighting variations registered by the camera C_2
$R'(v_1, v_2, x_i) = k'_i$	Angular ratio image as a function

operation attempts to overcome the unavoidable problem of either highly saturated or too dark ratio values in traditional ratio images. Experiments show that new angular ratio images are more robust and deliver improved disparity maps in comparison with their traditional counterparts.

Second, we perform an experimental evaluation of angular ratio images under the standard test bed for two-view stereo algorithms, i.e., under different aggregation and optimization approaches. To the best of our knowledge, this is the first work to report a detailed experimentation related to LTC under the well-known two-view stereo test bed. The results of this research are consistent with previously reported conclusions in stereo surveys, and this fact suggests that angular ratio images conserve some properties of intensity images and therefore are eligible to be put into any binocular stereo frameworks. In other words, LTC, in the form of pairs of angular ratio images, may benefit from a vast variety of existent methods for the two-view stereo problem.

The paper is organized as follows. In Section 2, LTC and use of angular measures for improving ratio images are explained; Section 3 presents experimental evaluation of angular ratio images based on the standard stereo taxonomy; finally, Section 4 gives conclusions and outlines further research directions.

2 Light Transport Constancy and Angular Ratio Images

For an easier comprehension of notation used in this article, Table 1 presents expressions referring to cameras, pixels, light source intensity variations, and ratio images. 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 [21], let us denote a particular point in the scene as x_i. This point will reflect light to 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 vectors
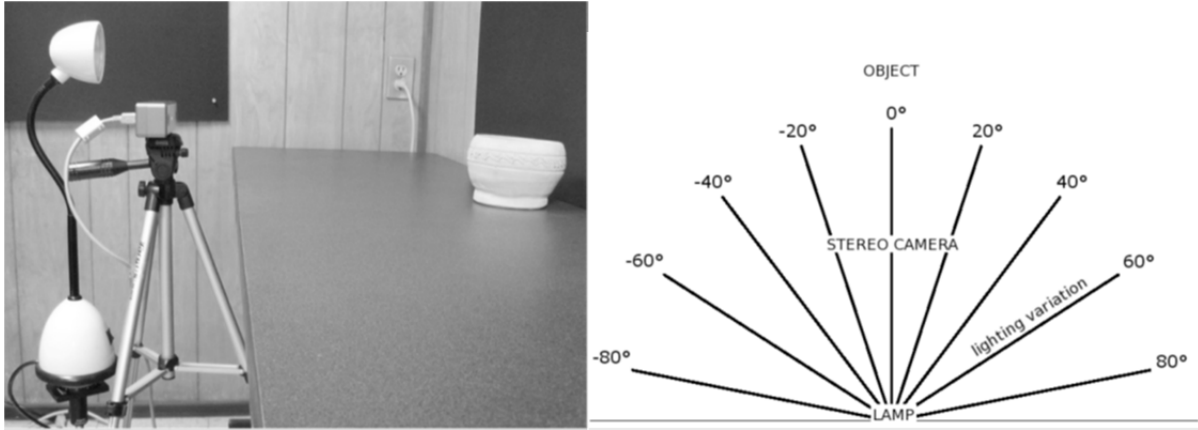
Fig. 1. Acquisition setting and illumination variation. The left image shows the acquisition setting: a stereo camera, a desk lamp, and an object. The $20W$ lamp was mounted on a rotating ruler on a tripod approximately $50cm$ away from the closest object to the camera, so that rotations could be measured. The right image shows a top-down sketch of the setting, where the degree of separation between each camera shot can be seen

in the direction of L and C_j. The Lambertian assumption states 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 hold in general 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 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 $k(x_i)$ because

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

Wang et al. [21] have shown how LTC can be used in multiple-view stereo to impose a rank constraint on the matrix

$$I_{CV} = \begin{pmatrix} I_{C_1V_1} & I_{C_2V_1} & \cdots & I_{C_mV_1} \\ I_{C_1V_2} & I_{C_2V_2} & \cdots & I_{C_mV_2} \\ \vdots & \vdots & \ddots & \vdots \\ I_{C_1V_m} & I_{C_2V_n} & \cdots & I_{C_mV_n} \end{pmatrix}, \qquad (2)$$

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 omit the

expression (x_i). However, each of the remaining equations in the paper is related to a single pixel at the position x_i. The matrix with the minimum rank is therefore sought, i.e., if LTC is observed through 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 holds only 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 equal at most to the number of light sources. Since I_{CV} will be corrupted with noise, it is impossible to calculate the 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 the minimum rank, as

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

where σ_i are singular values of I_{CV}. For multiple-view and multiple-lighting stereo, the minimum score is used as the matching cost.
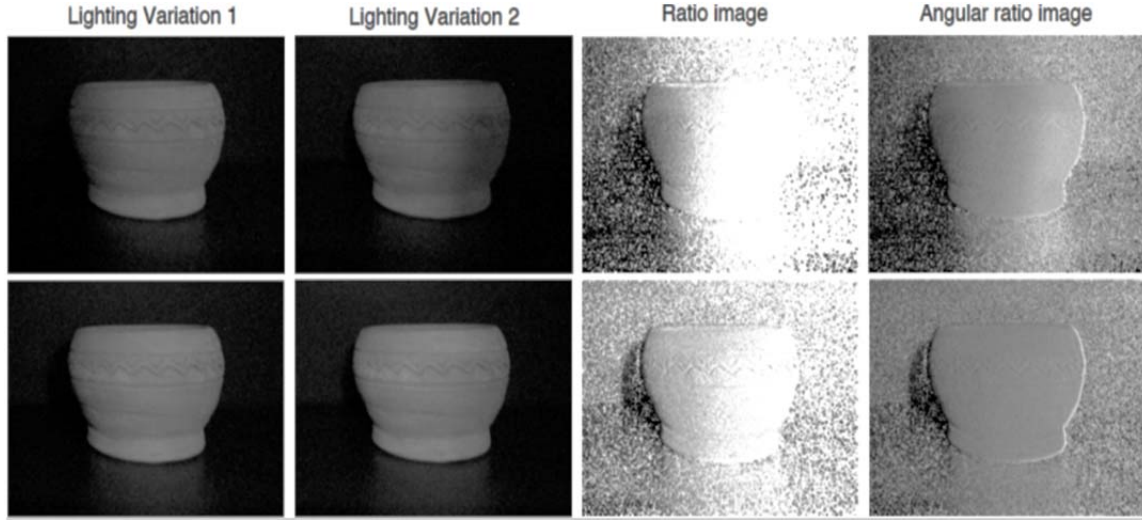
Fig. 2. Ratio images and angular ratio images. From top to bottom, the rows of the figure include lighting variation pairs (-60º, 60º) and (-40º, 40º). The left-camera images for lighting variations 1 and 2, the ratio images and the angular ratio images are placed column-wise
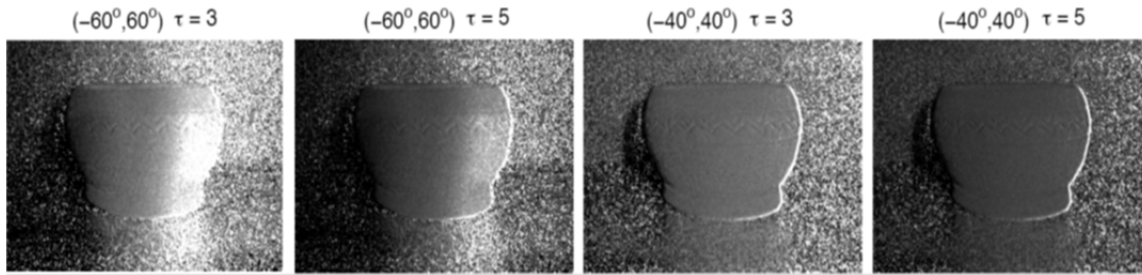

Fig. 3. Varying τ for ratio image generation. Normalization results for $\tau = 3$ and $\tau = 5$ are shown with lighting variations as indicated by the labels above the images

Let us now consider the simplest setting for LTC-based stereo, which is the case of interest in this paper: a single light source and two cameras. For each pixel in the left and right images, the intensities observed with the first lighting variation can be explained in terms of the intensities observed with 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:

$$\frac{I_{C_1V_2}}{I_{C_1V_1}} = \frac{I_{C_1V_1}k_1}{I_{C_1V_1}} = k_1, \qquad (4)$$

$$\frac{I_{C_2V_2}}{I_{C_2V_1}} = \frac{I_{C_2V_1}k_2}{I_{C_2V_1}} = k_2. \qquad (5)$$

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

$$I_{CV} = \begin{pmatrix} I_{C_1V_1} & I_{C_2V_1} \\ I_{C_1V_2} & I_{C_1V_2} \end{pmatrix}. \qquad (6)$$

Note that LTC holds only 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 $k_1 = k_2$. Therefore, minimizing the second singular value is equivalent to minimizing
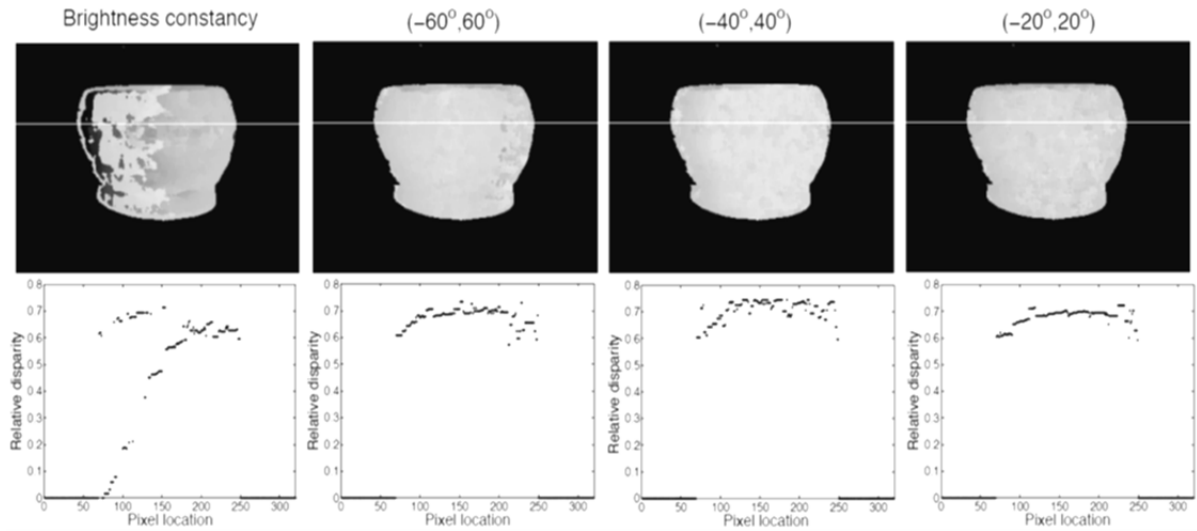
Fig. 4. Disparity results on brightness constancy and angular ratio images. The first row presents disparity maps, the second row shows the lines of interest with constant $y - axis$ along the disparity map

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 defined only 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 provide extensive ways to calculate dense disparity maps through ratio images.

In practice, unfortunately, the intensities of 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 most common. Moreover, in some regions of an 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, that is, much greater than 100%, and as a consequence, correspondence cannot be solved. The unwanted effect of these pixel values can be reduced if the ratios are redefined as:

$$k'_1 = |tan^{-1}(k_1)|, \qquad k'_2 = |tan^{-1}(k_2)|, \qquad (7)$$

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 a new ratio $R'(v_1, v_2, x_i) = k'_i$ which will be referred to as the *angular ratio image*.

2.1 Experiments on Angular Ratio Images

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

Before 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 hold for its corresponding right pair), i.e., using
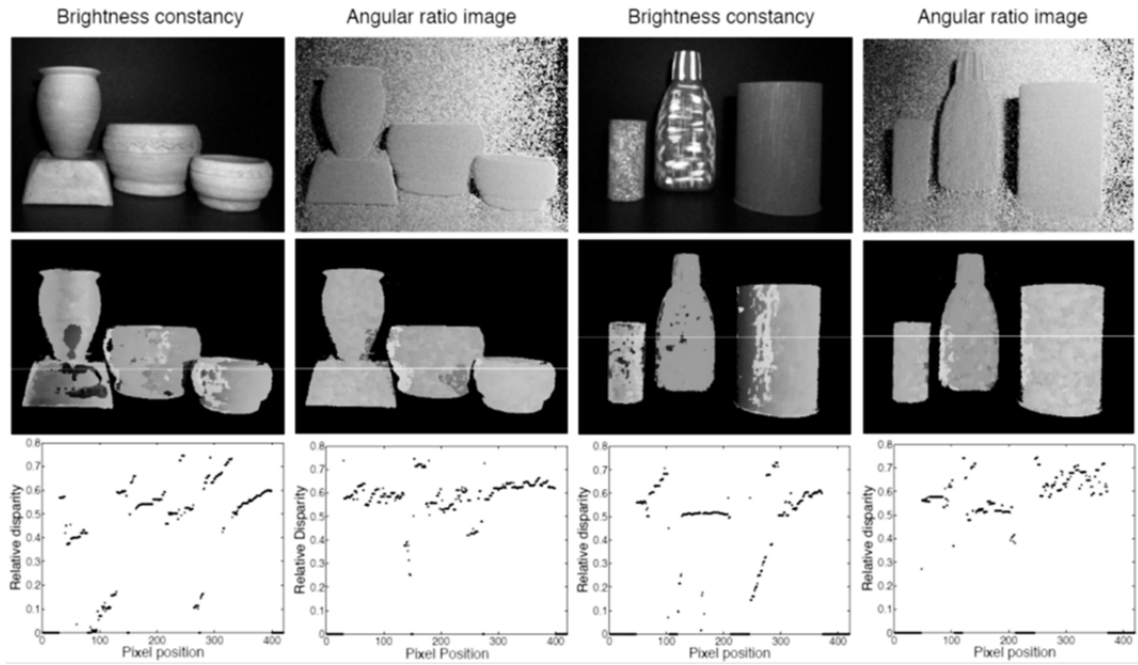
Fig. 5. Additional scenes. The first row presents input images; the second row, the obtained disparity maps; the third row, the lines of interest

Eqs. 4 and 5. Ratio images are stored after the following filter has been applied:

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

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

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

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 [21]. 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.

The experimental analysis commences with Figure 2, where the difference between ratio and angular ratio images can be appreciated. From top to bottom, the rows of Figure 2 show the lighting variation pairs $(-60°, 60°)$ and $(-40°, 40°)$. The left-camera image for lighting variations 1 and 2, the ratio image $(\tau = 1)$, and the angular ratio image are shown row-wise. Recall that ratio and angular ratio images are calculated for a single view (i.e., left or right camera images) and two variations. In Figure 2, 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 the main problem in ratio image generation, i.e., choosing an optimum cut
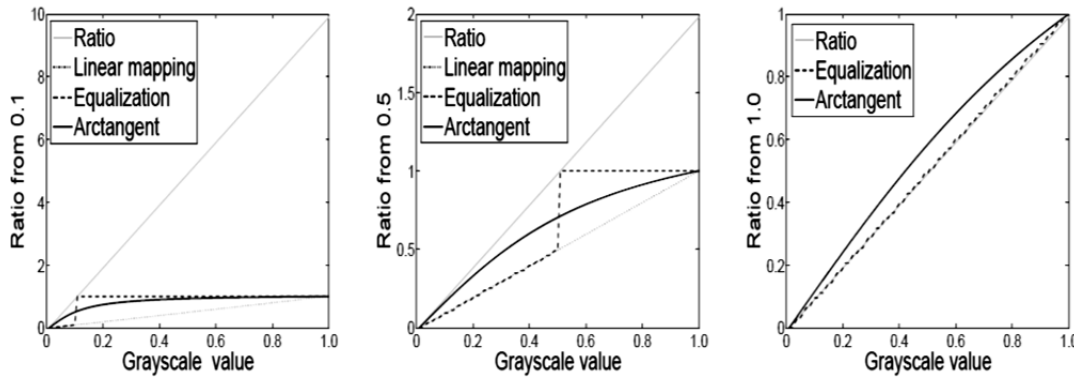
Fig. 6. Different fitting functions. The behavior of different adjustment methods is presented as a function of grayscale values, i.e., grayscale values ranging from 0 to 1 ($x - axis$) are divided by a fixed grayscale value in order to obtain a ratio ($y - axis$). From left to right, the results for the fixed grayscale values of 0.1, 0.5, and 1, respectively, are given

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, since the figure shows that $\tau = 5$ favors the ratio obtained from the pair $(-60°, 60°)$, but over-darkens the ratio obtained from the pair $(-40°, 40°)$.

As far as the disparity results are concerned, these are demonstrated in Figure 4. Here, the first row includes disparity maps, while the second row shows the lines of interest with constant $y - axis$ along the disparity map. For visualization purposes, a mask has been applied for isolating the objects of interest from the background. The first column demonstrates the results for brightness constancy, i.e., neither ratio images nor angular ratio images are used here. Instead, a usual grayscale image left/right image pair is used as stereo input. The rest of the columns present results from the angular ratio images with lighting variation pairs of $(-60°, 60°)$, $(-40°, 40°)$, and $(-20°, 20°)$. As expected, there is an improvement in disparity calculation for the angular ratio images over the brightness constancy, that is, the disparities are located over more continuous regions. Interestingly, there is little difference between the results related to angular ratio images, which suggests that the angles may represent a robust way to obtain similar disparity maps through different lighting

variations. Additional scenarios are presented in Figure 5, where again, the disparity results for angular ratio images are better than for those relying on the brightness constancy assumption.

We explore the use of two alternative adjustment functions for ratio values, namely, the standard histogram equalization for ratio images and a linear mapping which fits the elements of vector $\mathbb{x}$ into the upper and lower boundaries b_u and b_l, respectively. This linear mapping is given by the formula:

$$
\begin{aligned}
&f(\mathbb{x}, u_b, l_b) \\
&= \left(\frac{b_l - b_u}{\min(\mathbb{x}) - \max(\mathbb{x})} \right) x_i \\
&+ \left(b_u - \left(\frac{b_l - b_u}{\min(\mathbb{x}) - \max(\mathbb{x})} \right) \max(\mathbb{x}) \right),
\end{aligned}
\tag{10}
$$

where x_i is the i_{th} element of $\mathbb{x}$. Figure 6 shows the behavior of different adjustment methods as a function of grayscale values, i.e., grayscale values ranging from 0 to 1 ($x - axis$) are divided by a fixed grayscale value in order to obtain a ratio ($y - axis$). From left to right, Figure 6 presents the results for the fixed grayscale values of 0.1, 0.5, and 1, respectively. The purpose of this figure is to provide a perspective of possible ratios which can be obtained from a fixed value. The linear mapping, equalization, arctangent and
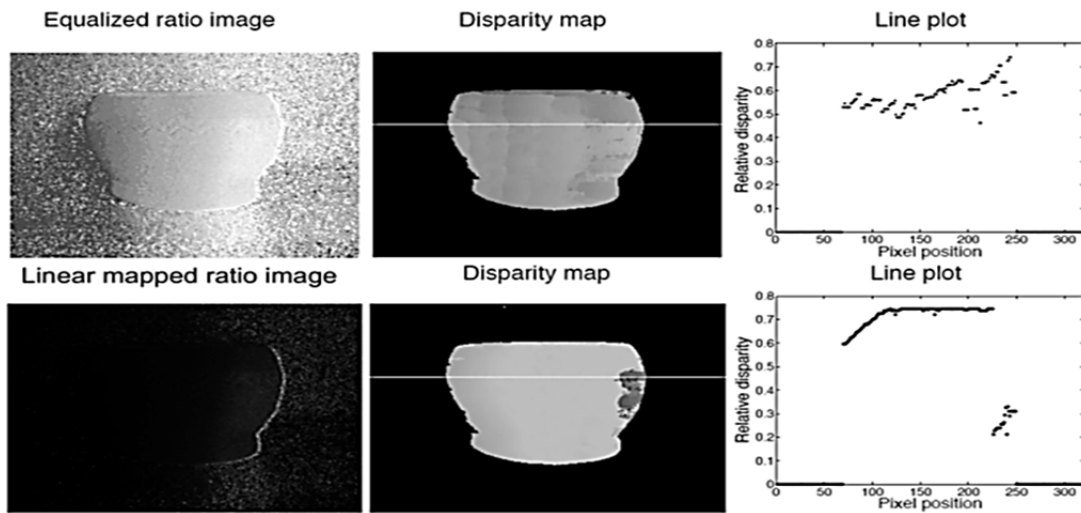
Fig. 7. Results on a histogram-equalized ratio image and a linear mapped ratio image. First row, from left to right: an equalized ratio image, a disparity map, a line of interest along the disparity map. Second row, from left to right: a linear mapped ratio image, a disparity map, a line of interest along the disparity map

ratio functions are shown as dark dotted, dark dashed, dark solid and gray solid plots, respectively. The arctangent function was bounded from $(0, \pi)$ to $(0,1)$ for visualization purposes.

There are several features to note in Figure 6. First, the ratio function assigns values between 0 and 1 only if the fixed value is 1. This means that many pixels are prone to become saturated or too dark. Second, the remaining functions are able to bind the ratio values between 0 and 1. Third, the arctangent function appears as an interpolated plot between both equalization and linear mapping, allowing a more continuous mapping between different ratio values, thus avoiding the probabilities of assigning the extreme ratio values. Only when the fixed value tends to one (right-most diagram), the ratio function (which is identical to linear mapping) and the equalization function seem suitable for value assignment.

The last step of the analysis in this section is presented in Figure 7. Here, the top and bottom rows display the results using histogram equalization and linear mapping, respectively. In this figure, a standard histogram equalization operation was performed on a raw ratio image with the variation$(-60°, 60°)$.

Despite the equalized ratio image looks similar to its angular ratio image counterpart (top right corner of Figure 2), the disparity results are rather different, favoring the use of angular ratio images again. Similarly, the results of linear mapping present a deep discontinuity at the end of the depth line plot. Note that the linear mapped image is much darker than both the histogram equalized image and the angular ratio image.

3 Angular Ratio Images and the Standard Stereo Taxonomy

In [16], Scharstein and Szeliski proposed taxonomy developed to study and classify a wide variety of two-view stereo methods. Their work has become a benchmark for testing and reviewing novel stereo approaches. Recently, some papers also used this taxonomy to evaluate the robustness of stereo algorithms on changes in illumination and color [9, 11]. Roughly, the taxonomy establishes the principle that any stereo algorithm can be divided into the following main steps:

- **Matching cost.** A cost of correspondence is calculated in order to determine pixel

Fig. 8. Four different scenarios used in the experimental section. Left images of stereo pairs (light source direction at 0º) are given. From left to right: Fps, Bot, Gift, and Edy
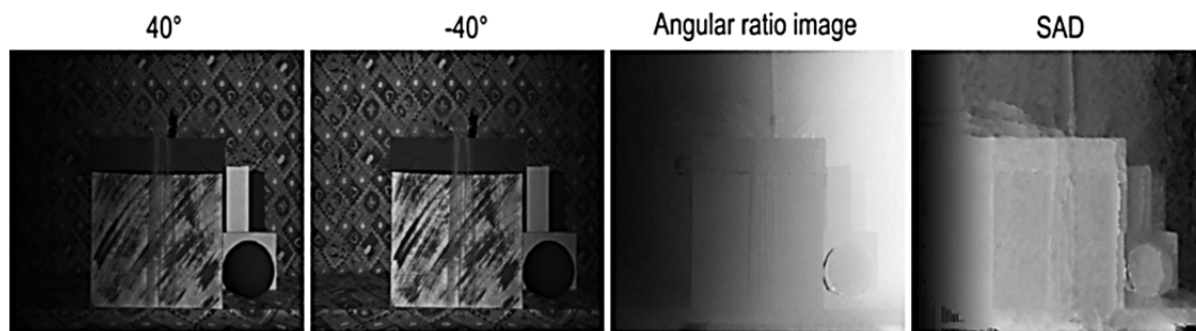

Fig. 9. The standard aggregation approach in scenario Gift. From left to right, the first two pictures are left images of the stereo pair Gift with illumination variations of $40°$ and $−40°$, respectively. The third and fourth images correspond to the angular ratio image and the recovered disparity map using the standard approach: 9×9 square window aggregation

disparity. Sum of Absolute Differences (SAD) and Sum of Square Differences (SSD) are typical examples of matching cost functions.

− **Aggregation**. The initial costs of correspondence are spatially aggregated over support regions. Square windows of fixed and varying shape and size are typical examples used as aggregating regions.

− **Optimization.** A disparity for each pixel is chosen as the result of minimizing a local or global objective function. Graph cut and dynamic programming are amongst the most popular optimization approaches.

− **Refinement.** The generated disparity maps are post-processed in order to remove errors, i.e., filling regions where disparity could not be determined.

The above steps can be combined into a specific sequence. Different stereo algorithms consist of different sequences of steps. For example, in local algorithms (window-based), the calculation of disparity at a given point depends only on the intensity values within a finite window (i.e., the aggregation step). These algorithms usually make implicit the softness assumptions due to aggregation. Typically, only matching cost with aggregation is used in these approaches [2, 23]. Global algorithms, on the other hand, make explicit the softness assumptions and solve the optimization problem. These algorithms do not usually include the aggregation step. Instead, they assign the disparity that minimizes the objective function. In some cases, this idea is realized with the help of a function that combines data of the first step with regularization terms [3, 14, 15].

The purpose of this section is to explore the influence of different aggregation and optimization approaches on disparity maps delivered from pairs of angular ratio images. The outcome of this experiment is relevant, since it may show if angular ratio images can be incorporated into standard two-view stereo frameworks. Following the experimental protocol in [16], we compare the results of applying the following aggregation methods to ratio images: shiftable window [1, 20], iterated binomial [5], regular diffusion [15], and membrane diffusion [15]. Likewise, we studied the performance of the following optimization methods: dynamic programming [1], scanline optimization [16], graph cut [3], simulated annealing [7].

Figure 8 illustrates different scenarios used in the experiments. Stereo image pairs were gathered under the same acquisition setting as explained in Section 2.1. From left to right, the scenarios are named Fps, Bot, Gift, and Edy. Fps and Bot were taken on a plain black background.

A $20W$ halogen desk lamp was used as a light source.

The lamp was located proximately $50cm$ away from the closest object to the camera.

These scenarios were used in Section 2.1. For scenarios Gift and Edy, a textured background was used. A $150W$ halogen reflector functioned as a light source. The reflector was located approximately $80cm$ away from the closest object to the camera. In Figure 8, only the left image with central illumination configuration (i.e., $0°$ according to the diagram in Figure 1 from each stereo pair is presented.

By using two different light source intensities, we aim to provide comparisons among different illumination scenarios. It is important to mention that this evaluation does not include a quantitative analysis, since no range scanner or structured-light equipment was available for the experiments.
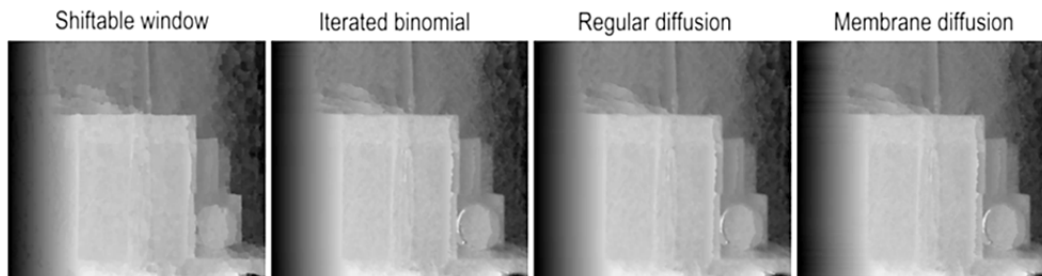

Fig. 10. Four different aggregation approaches in scenario Gift. The figure presents disparity maps calculated under angular ratio images using different aggregation methods: square shiftable window, iterated binomial, regular diffusion, and membrane diffusion
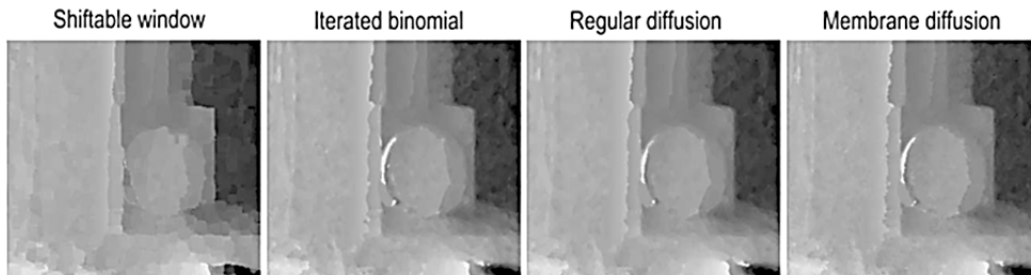

Fig. 11. Close-up images for regions of high discontinuity. From left to right, close-up images of a specific high discontinuity region of the disparity maps from Figure 10 are shown
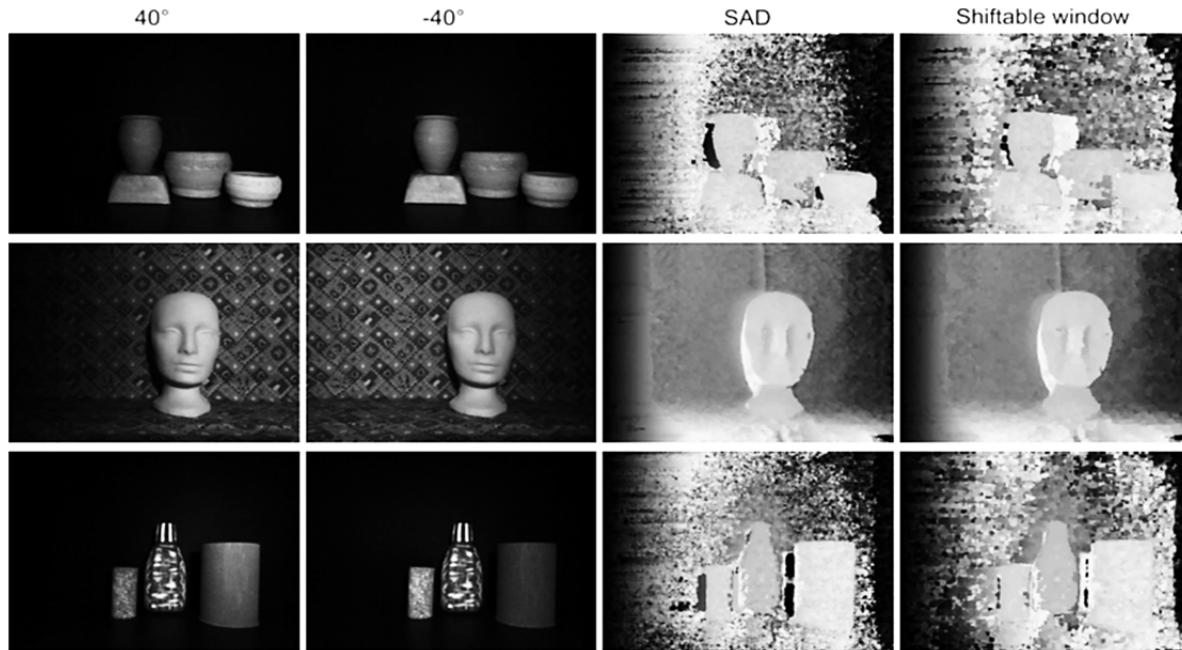
Fig. 12. The standard aggregation approach and the square shiftable window in scenarios Fps, Edy, and Bot**.** Scenarios Fps, Edy, Bot, are shown row-wise. From left to right, the first two columns show the left image from the stereo pair with $40°$ and $-40°$ light source direction. The last two columns show disparity maps calculated using 9×9 square aggregation window (standard approach) and square shiftable window over the angular ratio image pair, respectively

Nonetheless, we believe that qualitative results provide a fair idea of the performance of different methods. Also, no information about a public database suitable for LTC tests was available.

3.1 Evaluating Aggregation Methods on Angular Ratio Images

In this section, we present experiments aimed to compare the performance of different aggregation methods applied to angular ratio images. For all experiments, a matching cost of SAD and a WTA optimization were used. Following the suggested experiments in [16], we selected the following aggregation parameters: square shiftable window size 9×9, iterated binomial with 6 iterations, regular diffusion with 30 iterations, membrane diffusion with 150 iterations, and $\beta = 0.2$. In order to show a classical aggregation example, i.e., 9×9 square window aggregation, we start with Figure 9. It presents, column-wise, a left image with a $40°$ illumination variation, a left image with a $-40°$ illumination variation, an angular ratio

image, and the obtained disparity map from the angular ratio image in scenario Gift. The results of applying the aggregation steps mentioned previously are shown as disparity maps in Figure 10. Here, square shiftable window, iterated binomial, regular diffusion and membrane diffusion are presented from left to right. At a first glance, Figures 9 and 10 reveal the fact that there is no significant difference between aggregation methods and the classical one. However, it can be noticed that the most successful aggregation method over regions of high discontinuity (located along the boundaries of the different boxes and the ball) is the square shiftable window. This observation is clearly consistent with the conclusions reported in [16], where it was shown that the best approach to deal with discontinuity errors was the square shiftable window. In order to emphasize the former observation, a close-up of the bottom-right corner area of different disparity maps of Figure 10 is shown in Figure 11. Here, it is noticeable that the shiftable window aggregation outperforms the rest of the
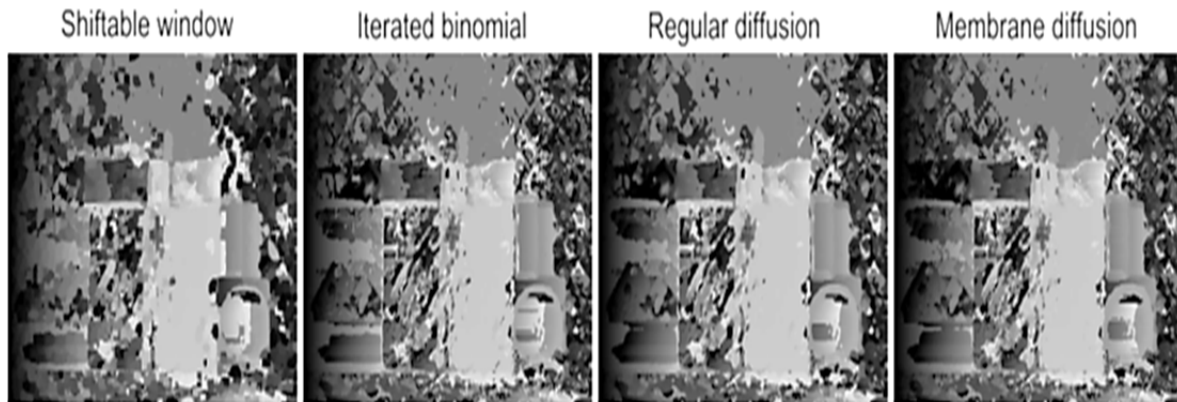
Fig. 13. Four different aggregation methods in scenario Gift using brightness constancy. From left to right, disparity maps recovered using square shiftable window, iterated binomial, regular diffusion, and membrane diffusion are shown. Note that no angular ratio images were used in this experiment

aggregation methods, i.e., the error surrounding the contour of the ball does not affect the disparities obtained from the square shiftable window approach. Besides, object-to-background transition is sharper for the square shiftable window result, which is more noticeable around the boundaries of the boxes, that is, the shiftable window presents a less diffuse transition around such areas. Note that the rest of the aggregation approaches present a rather softened object-to-background transition.

As far as the remaining scenarios are concerned, only comparisons between the standard 9×9 square window and the square shiftable window are presented in Figure 12, since no major changes were observed using iterated binomial, regular diffusion, and membrane diffusion. The results in scenarios Fps, Edy, and Bot are shown row-wise. Note that in scenario Fps, the square shiftable window helps to improve the disparity map by diminishing the error around the shaded area of the vases. A similar phenomenon can be noticed in Edy and Bot scenarios in the shaded area of the mannequin's head and the pots, respectively.

Finally, in order to demonstrate that LTC provides a substantial help to solve the correspondence problem in stereoscopy, we included Figure 13. The figure presents the results of different aggregation methods using only the brightness constancy constraint, i.e., no angular ratio images were used as inputs for these experiments. Instead, intensity images taken at $0°$ light source direction were used. From left to right, Figure 13 presents disparity maps recovered using square shiftable window, iterated binomial, regular diffusion and membrane diffusion. For all aggregation cases, only recovered disparity maps of less quality than those obtained with angular ratio images can be observed. Basically, this is due to the presence of irregularities in the recovered disparity maps. Such irregularities are not noticeable when angular ratio images are used (Figure 10). Recall that the purpose of Figure 13 is not to compare different aggregation methods, but to compare the results obtained from using LTC (in the form of angular ratio images) against a usage of brightness constancy (in the form of traditional grayscale images).

3.2 Evaluating Optimization Methods on Angular Ratio Images

In this section, we present experiments aimed to compare the performance of different optimization methods when applied to angular ratio images. Following the suggested experiments in [16], we used the following optimization parameters: dynamic programming with a softness weight

Fig. 14. Results using different optimization methods. From top to bottom, disparity maps obtained from angular ratio images in scenarios Edy, Fps, Gift, and Bot are shown. From left to right, the following optimization approaches are presented: dynamic programming, scan line optimization, graph cut, and simulated annealing

$\lambda = 20$ and occlusion cost 20 , scanline optimization with $\lambda = 20$, graph cut with $\lambda = 20$, simulated annealing with $\lambda = 20$ and 500 iterations.

Figure 14 shows four optimization approaches in different scenarios. From top to bottom, the figure presents disparity maps obtained from angular ratio images in scenarios Edy, Fps, Gift, and Bot. From left to right, the following optimization approaches are shown: dynamic programming, scan line optimization, graph cut, and simulated annealing. Observing the figure, interesting facts can be noticed. First, graph cut seems to group different regions of the disparity maps into clusters of disparity. These results resemble a clustered version of the standard approach (Figure 12); among the four methods, graph cut delivers the sharpest object segmentation. Unfortunately, this method is prone to errors due to self-shadowing in objects, as it can be seen mainly in scenarios Edy and Fps.

Dynamic programming, on the other hand, seems to cope well with self-shadowing, but the boundaries between different objects and the background are not defined as well as in the graph cut case. Scan-line optimization seems to be an intermediate option, between dynamic programming and graph cut.

Finally, the worst performance is delivered by simulated annealing, especially regarding the smoothness of the recovered depth maps. These results are consistent with the observations reported in [16], [3], and [19], where similar conclusions were derived for brightness constancy in grayscale images. Optimization methods were also applied to intensity images, i.e., to test the performance of brightness constancy against LTC and angular ratio images. The results of different optimization methods are presented in Figure 15; in these experiments, only the brightness constancy constraint was applied, i.e., no angular ratio images were used as inputs
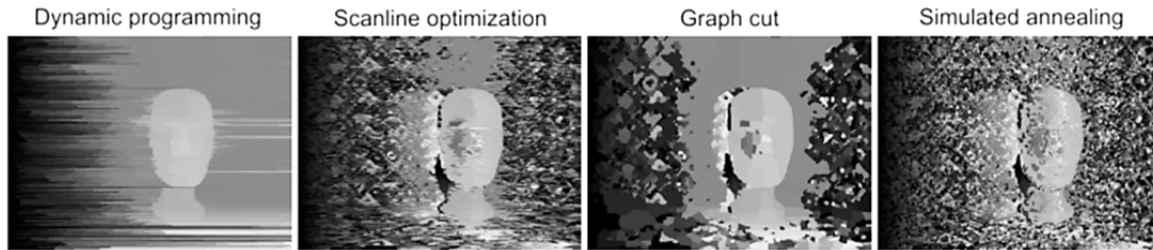
Fig. 15: Four different optimization methods in scenario Edy using brightness constancy. From left to right, disparity maps recovered by dynamic programming, scan line optimization, graph cut, and simulated annealing are shown. Note that no angular ratio images were used in this experiment

for these experiments. Instead, intensity images taken at 0° light source direction were used. From left to right, Figure 15 shows the disparity maps recovered using dynamic programming, scan line optimization, graph cut, and simulated annealing. For all aggregation cases, only poorly recovered disparity maps compared to those obtained with angular ratio images can be observed.

4 Conclusions

We have focused on a potential issue in LTC-based two-frame stereo algorithms: generating ratio images suitable for standard stereo algorithms. To this end, we have shown that the arctangent function can transform the original ratio image values into a stabilized range of values, outperforming alternative adjustment functions such as histogram equalization or linear mapping. We have also fulfilled an experimental evaluation of performance of the standard benchmarking method in two-frame stereo algorithms over several angular ratio image scenarios. The main results of this evaluation confirm that aggregation and optimization techniques in two-view stereo algorithms applied to angular ration images present a behavior comparable to the behavior of aggregation and optimization approaches applied to grayscale images. The former fact suggests that the standard taxonomy for binocular stereo may be applied to angular ratio images, keeping specific properties of different aggregation and optimization approaches as when applied on grayscale imagery. As LTC is an emerging tool

whose potential use in the stereo vision area is mainly related to the two-frame case, this study represents the first attempt to validate the penetration of ratio images into the extensive two-frame stereo vision literature. Further promising directions of research include a rigorous study and, possibly, an appropriate automatization of light source intensity and position for two-frame stereo vision using LTC. This is motivated by the idea that an automatic control of the intensity of a light source based on information provided by cameras, i.e., the number of pixels out of normal ratio values, can deliver more appropriate disparity maps in accordance with the specific illuminations needs of the observed scene.

Acknowledgements

References

1. **Bobick, A.F. & Intille, S.S. (1999).** Large occlusion stereo. *International Journal of Computer Vision*, 33(3), 181–200.

2. **Bolles, R.C., Baker, H.H., & Hannah, M.J. (1993).** The JISCT stereo evaluation. *Image Understanding Workshop* (263–274). San Francisco, CA.: Morgan Kaufmann Publishers.

3. **Boykov, Y., Veksler, O., & Zabih, R. (2001).** Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11), 1222–1239.

4. **Brown, M.Z., Burschka, D., & Hager, G.D. (2003).** Advances in computational stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8), 993–1008.

5. **Burt, P.J. & Adelson, E.H. (1983).** The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4), 532–540.

6. **Georghiades, A.S., Belhumeur, P.N., & Kriegman, D.J. (2001).** From few to many: illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6), 643–660.

7. **Geman, S. & Geman, D. (1984).** Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6), 721–741.

8. **Gong, M., Yang, R., Wang, L., & Gong, M. (2007).** A Performance Study on Different Cost Aggregation Approaches Used in Real-Time Stereo Matching. *International Journal of Computer Vision*, 75(2), 283–296.

9. **Heo, Y.S., Lee, K.M., & Lee, S.U. (2008).** Illumination and Camera Invariant Stereo Matching. *IEEE Conference on Computer Vision and Pattern Recognition*. Anchorage, Alaska, 1–8.

10. **Hernandez, C., Vogiatzis, G., & Cipolla, R. (2008).** Multiview Photometric Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3), 548–554.

11. **Hirschmüller, H. & Scharstein, D. (2009).** Evaluation of Stereo Matching Costs on Images with Radiometric Differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(9), 1582–1599.

12. **Liao, M., Wang, L., Yang, R., & Gong, M. (2007).** Light Fall-off Stereo. *IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, USA, 1–8.

13. **Magda, S., Kriegman, D.J., Zickler, T., & Belhumeur, P.N. (2001).** Beyond Lambert: Reconstructing surfaces with arbitrary BRDFs. *Eighth IEEE International Conference on Computer Vision*, Vancouver, Canada, 2, 391–398.

14. **Marroquin, J., Mitter, S., & Poggio, T. (1987).** Probabilistic solution of ill-posed problems in computational vision. *Journal of the American Statistical Association*, 82(397), 76–89.

15. **Scharstein, D. & Szeliski, R. (1998).** Stereo matching with nonlinear diffusion. *International Journal of Computer Vision*, 28(2), 155–174.

16. **Scharstein, D. & Szelisky, R. (2002).** A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision*, 47(1-3), 7–42.

17. **Scharstein, D. & Szeliski, R. (2003).** High accuracy stereo depth maps using structured light. *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Wisconsin, USA, 1, 195–202.

18. **Seitz, S.M., Curless, B., Diebel, J., Scharstein, D., & Szeliski, R. (2006).** A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York, USA, 1, 519–528.

19. **Szeliski, R. & Zabih, R. (1999).** An experimental comparison of stereo algorithms. In Springer: Berlin, *International Workshop on Vision Algorithms,* Keyra, Greece, 1–19.

20. **Tao, H., Sawhney, H., & Kumar, R. (2001).** A global matching frame-work for stereo computation. *Eighth IEEE International Conference on Computer Vision*, Vancouver, Canada, 1, 532–539.

21. **Wang, L., Yang, R., & Davis, J.E. (2007).** BRDF Invariant Stereo Using Light Transport Constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9), 1616–1626.

22. **Woodham, R.J. (1980).** Photometric method for determining surface orientation from multiple images. *Optical Engineering,* 19(1), 139–144.

23. **Zabih, R. & Woodfill, J. (1994).** Non-parametric local transforms for computing visual correspondence. *Third European Conference on Computer Vision*, Stockholm, Sweden, 2, 151–158.

24. **Zickler, T.E., Belhumeur, P.N., & Kriegman, D.J. (2002).** Helmholtz stereopsis: exploiting reciprocity for surface reconstruction. *International Journal of Computer Vision*, 49(2-3), 215–227.

Pablo Arturo Martinez González obtained BSc from Technological Institute of Fresnillo Zacatecas (ITSF) in 2007 and MSc in Robotics and Advanced Manufacturing from the Research Center for Advanced Studies of National Polytechnic Institute (CINVESTAV-Saltillo) under the supervision of Dr. Castelán in 2010. Currently, he occupies a teaching position at Polytechnic University of San Luis Potosí (UPSLP). At present, his research interests are related to the area of computer vision for robotics applications.

Mario Castelán obtained a BSc from the University of Veracruz (UV) in 1999 and a MSc in Artificial Intelligence from the University of Veracruz and the National Laboratory of Advanced Informatics (LANIA) in 2002. He obtained a Ph. D. in Computer Science from the University of York, U.K., in 2006. Currently, he is a full-time researcher at the Robotics and Advanced Manufacturing Research Group of CINVESTAV-Saltillo. Dr. Castelán has published over thirty papers in refereed international journals and international conferences in the area of computer vision, 3D shape analysis, statistical learning, and vector space analysis for robotics applications.