Real Time Stereo Matching for Radiometric Changes

Merlin George
Department of Computer Science and Engineering,
Sree Chitra Thirunal College of Engineering,
Trivandrum, Kerala, India

Rejimol Robinson R. R
Department of Computer Science and Engineering,
Sree Chitra Thirunal College of Engineering,
Trivandrum, Kerala, India

ABSTRACT

Most of the existing stereo matching algorithms assume that the corresponding pixels have the same intensity (color) in both images. But in real world situations, image color values are often affected by various radiometric factors such as illumination direction, illuminant color, and imaging device changes. Hence, we cannot fully depend on the raw color recorded by a camera. So, the assumption of color consistency does not hold well for stereo images in real scenes. Thus, the performance of most conventional stereo matching algorithms would be severely degraded under radiometric variations. The main focus of this work is on illumination invariant stereo matching by generating illumination invariant images from stereo image data using a non-iterative normalisation in log RGB space. The actual stereo matching is done using the similarity measure, Normalized Cross-Correlation (NCC) which is the standard statistical method for determining similarity which itself is invariant to linear brightness and contrast variations. In this work we propose a novel method for error analysis by dividing disparity into uniform and discontinuity regions. The proposed algorithm was evaluated using standard datasets and the results are comparable to state-of-art techniques in the literature.

General Terms

Stereo Matching, Computer Vision, Image Processing

Keywords

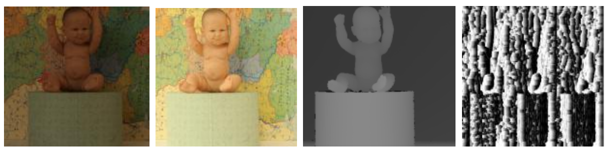
Disparity Map, Stereo Correspondence, Invariant Images, Radiometric Variations, Illumination Invariant, Color Normalisation.

1. INTRODUCTION

The main objective of stereo matching is to obtain 3D information by finding the correct correspondence between images captured from different point of views or at different times. However, finding the accurate correspondence is not an easy task since there exists a number of difficulties, such as occluded regions, texture less regions, and object boundaries. This issue has been an important area of research in the past several decades, and considerable progress has been made with respect to the problem surrounding stereo matching algorithms. Efforts toward this end have resulted in numerous stereo algorithms that perform relatively well for the images in the Middlebury database [1], [2]

These algorithms are based on a common assumption that corresponding pixels have similar color values, an assumption we refer to as color consistency. However, it should be noted that these methods do not hold good for stereo images which do not have similar corresponding color values. Nonetheless, a few studies have been performed in order to solve this problem.

In a real scenario, various factors prevent two corresponding pixels from having the same color value. One major factor is a radiometric change, which includes lighting geometry, illuminant color, and camera device changes between stereo images. Different color values are obtained when the same scene is viewed under a different lighting geometry; the reason for this is that the intensity at each point is determined by the angle between the direction of the incident light and the direction of the surface normal in a Lambertian model. After fixation of the lighting geometry, the object when viewed under different illuminant colors also produces different colors because there is a change in the spectral distribution of the reflected light from the object. Furthermore, color changes can also be induced by using a camera device or setting changes such as exposure variations because there are options to vary the total amount of photon that is incident to the camera. Common situations such as these can be a reason for practical problems in stereo images such as satellite images. Having said that, one should not completely trust the raw color recorded by a camera for purposes of matching, and it remains that the color consistency assumption is no longer valid for stereo images in real scenes. Under radiometric variations such as these, the performance of most stereo matching algorithms can be severely degraded. To prove this point with an example, the case shown in Fig. 1.1 in which a conventional method such as the Sum of Absolute Difference (SAD) method fails to produce correct depth map.



a) b) c) d)

Fig 1: The output of conventional SAD for illumination varying stereo images. (a) and (b) are the left and right Baby1 image with varying illumination.(c) is the ground truth (d) is the result using SAD method for images (a) and (b)

For the past few years, a large number of researches have been undertaken by Finlayson et al, to study methods for creating invariant images from a single digital image. An invariant image is one that is independent of illumination conditions. In this paper, we have implemented the method adopted by Finlayson et al. [3] for generating invariant images. This work finds an invariant image motivated by the assumptions of Lambertian surface and the imaging device is linear with respect to light intensity. But, the method still works well when these assumptions do not hold. And, the proposed method initially finds invariant images of both left and right images using non-iterative comprehensive normalisation in log space and the actual stereo matching using Normalized Cross correlation is done to these invariant color images which are approximately invariant to intensity and color of scene illumination.

2. INVARIANT IMAGE FORMATION

2.1 Color Image formation Model

An image is made up of pixels, and each pixel reflects the light that has been received at a sensor in a digital camera or other optical sensor. This light that hits the sensor is made up of two main components, the illumination component that reflects the colour and intensity of the illuminant, and the reflectance component which captures the surface reflectance properties, namely the color of the object reflecting the light. These two components define the properties of the light that hits the sensor. The sensor then converts the light into a digital signal, adding in a certain amount of noise. The following equation describes an image taken by a linear imaging device [5]:

$$h_k^x = \int_{\omega} E(\lambda) \quad S^x(\lambda) Q_k(\lambda) d\lambda, \tag{1}$$

where h_k^x represents the kth sensor (color channel) response at a point x in the scene and λ is the wavelength. $E(\lambda)$ represents the spectral power distribution of the incident illuminant, $S^x(\lambda)$ represents the surface reflectance at a point x in the scene, and $Q_k(\lambda)$ stands for the spectral response of the kth sensor.In order to simplify this equation and remove the integral, we can assume that the camera sensor $Q_k(\lambda)$ behaves similar to a Direc delta function $Q_k(\lambda) = q_k \, \delta(\lambda - \lambda_k)$, where q_k represents the sensor strength $q_k = Q_k(\lambda_k)$.Now, (1) simply becomes

$$h_k^x = \mathrm{E}(\lambda_k) S^x(\lambda_k) \tag{2}$$

During the image acquisition process, the device responds in a linear fashion. However, for the compression of the dynamic range, there is a nonlinear transformation of the image data before the storage process. This process is called gamma correction [6] and it results in the raising of the value of each RGB response to a power function of an exponent γ value depending on the camera. Taking all these factors into consideration, the color image formation model at a pixel can be represented as follows[3]:

$$\begin{pmatrix} R_i \\ G_i \\ B_i \end{pmatrix} \rightarrow \begin{pmatrix} [a\rho_i R_i]^\gamma \\ [b\rho_i G_i]^\gamma \\ [c\rho_i B_i]^\gamma \end{pmatrix} \tag{3}$$

where each pixel has its own individual brightness factor ρ_i which depends on the angle between the direction of the light and direction of the surface normal at that point. Changing illumination color while fixing the lighting geometry would result in changes in the responses of three color channels by the global scale factor a, b and c, respectively.

2.2 Color Image Normalisation

In order to eliminate the effect of lighting geometry that depends only on direction of surface normal and the direction of light in the Lambertian model, chromaticity normalisation is commonly employed [3].

$$r_i = \frac{\rho_i R_i}{\rho_i R_i + \rho_i G_i + \rho_i B_i},$$

$$g_i = \frac{\rho_i G_i}{\rho_i R_i + \rho_i G_i + \rho_i B_i}, \tag{4}$$

$$b_i = \frac{\rho_i B_i}{\rho_i R_i + \rho_i G_i + \rho_i B_i},$$

(Clearly the ρ_i term cancels). We will denote the chromaticity normalisation carried out on the image *I* as *C(I).*

Let $\mu(R)$ denote the mean red pixel value for an image.

Assuming *N* pixels in an image:

$$\mu(R) = \sum_{i=1}^{N} R_i \tag{5}$$

Under a change in illuminant color, the mean becomes:

$$\mu(R)' = \frac{\sum_{i=1}^{N} aR_i}{N} = a\mu(R) \tag{6}$$

That is, the mean changes by the same scale factor *a*. Thus to cancel the effect of light colour on RGBs, we can apply the following transformation:

$$R_i' = \frac{aR_i}{a\mu(R)}, \; G_i' = \frac{aG_i}{b\mu(G)}, \; B_i' = \frac{aB_i}{c\mu(B)} \tag{7}$$

Equation (7) is commonly referred to as grey-world normalisation and we denote this normalisation by a function *G( )*such that the image *I* post greyworld normalisation is denoted *G(I).*

Neither (4) nor (7) by itself suffices to remove both lighting geometry and lighting colour change. Finlayson et al. [7] defined a third normalisation which they called Comprehensive normalisation, which can remove both dependencies.

It is defined as:

1. $I_0 = I$ Initialization
2. $I_{i+1} = G(C(I_i))$ Iteration step
3. $I_{i+1} = I_i$ Termination condition (8)

That is, chromaticity normalisation and greyworld normalisation are applied successively and repeatedly to an image until the resulting image converges to a fixed point.

The problem with (8) (unlike (4) or (7)) is that it is iterative. In order to avoid this, Finlayson et al. [4] defined a comprehensive normalisation without iteration. From the implementation point of view, it would be preferable to perform a non-iterative procedure than an iterative one. So this work depends on the non-iterative comprehensive normalisation procedure performed in log RGB space [3] for generating invariant images which are independent of illumination conditions.

2.3 Non-iterative Comprehensive normalisation in log-space

We begin with the simple observation that the lighting geometry and lighting colour processes which are multiplicative in RGB space become additive in log RGB space.

To simplify matters we incorporate γ into scalars: $a' \to a^{\gamma}$, $b' \to b^{\gamma}$, $c' \to c^{\gamma}$ and $\rho_i' \to \rho_i^{\gamma}$ so that (3) becomes :

$$\begin{pmatrix} R_i \\ G_i \\ B_i \end{pmatrix} \to \begin{pmatrix} a'\rho_i' R_i^{\gamma} \\ b'\rho_i' G_i^{\gamma} \\ c'\rho_i' B_i^{\gamma} \end{pmatrix} \tag{9}$$

Taking logarithms turns multiplications into additions so that our model of image formation becomes:

$$\begin{aligned} \log(R_i) &\to a'' + \rho_i'' + \gamma \log(R_i) \\ \log(G_i) &\to b'' + \rho_i'' + \gamma \log(G_i) \\ \log(B_i) &\to c'' + \rho_i'' + \gamma \log(B_i) \end{aligned} \tag{10}$$

where $a'' = \log a'$, $b'' = \log b'$, $c'' = \log c'$, and $\rho_i'' = \log \rho_i'$.

Representing log *RGB*s by R', G', and B' we have:

$$\begin{pmatrix} R_i' \\ G_i' \\ B_i' \end{pmatrix} \to \begin{pmatrix} a'' \\ b'' \\ c'' \end{pmatrix} + \rho_i'' \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \gamma R_i' \\ \gamma G_i' \\ \gamma B_i' \end{pmatrix} \tag{11}$$

Equation (11) tells us that in log RGB space, lighting geometry change only affects the length of the log *RGB* vector in the direction of *U* = (1,1,1).That is, the directions orthogonal to (1,1,1) are unaffected by brightness change. By applying some simple results from linear algebra, we can normalise a log RGB to remove brightness by projecting it onto the two-dimensional space which is orthogonal to the line that is spanned by *U*. We define a 3 x 3 projection matrix *Pr* for the space spanned by *U* and a complementary projection matrix $Pr^{\perp}$ = [*I-Pr*] for the space which is orthogonal to the space spanned by *U*:

$$\mathrm{Pr} = U^t (UU^t)^{-1} U = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \tag{12}$$

and so

$$Pr^{\perp} = I\text{-}Pr = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \tag{13}$$

where *t* denotes the matrix transpose operator and I denotes the 3 x 3 identity matrix. By definition these matrices have the property that $\mathrm{Pr} * (1,1,1)^t = (1,1,1)^t$ and $[1 - \mathrm{Pr}] * (1,1,1) = (0,0,0)^t$: To project a log RGB onto the space orthogonal to *U*, we simply multiply by the projection matrix $Pr^{\perp}$:

$$\begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix} \begin{pmatrix} R_i' \\ G_i' \\ B_i' \end{pmatrix} = \begin{pmatrix} \frac{2R_i'}{3} - \frac{G_i'}{3} - \frac{B_i'}{3} \\ \frac{2G_i'}{3} - \frac{R_i'}{3} - \frac{B_i'}{3} \\ \frac{2B_i'}{3} - \frac{R_i'}{3} - \frac{G_i'}{3} \end{pmatrix} =$$

$$\begin{pmatrix} R_i{}' - \frac{R_i' + G_i' + B_i'}{3} \\ G_i{}' - \frac{R_i' + G_i' + B_i'}{3} \\ B_i{}' - \frac{R_i' + G_i' + B_i'}{3} \end{pmatrix} \tag{14}$$

That is, we can remove dependency on lighting geometry by subtracting the mean log response at a pixel from each pixel.

The effect of illuminant colour can be removed in a similar way. However, rather than dealing with log RGB vectors we must operate on the vector of all log red (or green, or blue) responses. Considering all red responses, we can write:

$$\begin{pmatrix} a'' + \gamma R_1' \\ a'' + \gamma R_2' \\ : \\ a'' + \gamma R_N' \end{pmatrix} \to \begin{pmatrix} \gamma R_1' \\ \gamma R_2' \\ : \\ \gamma R_N' \end{pmatrix} + a'' \begin{pmatrix} 1 \\ 1 \\ : \\ 1 \end{pmatrix} \tag{15}$$

It follows that to remove dependence due to illumination colour we need to project responses onto the space orthogonal to the vector $(1,1,\ldots,1)^t$. This can be achieved similarly to the intensity normalisation, by defining a projection matrix *Pc* which projects onto the space spanned by $(1, 1,\ldots, 1)^t$ and its complement $Pc^{\perp}$ which projects onto the orthogonal space:

$$\mathrm{Pc} = \begin{pmatrix} \frac{1}{N} & \frac{1}{N} & \cdots & \frac{1}{N} \\ \frac{1}{\mathrm{N}} & \frac{1}{N} & \cdots & \frac{1}{N} \\ : & : & & : \\ \frac{1}{\mathrm{N}} & \frac{1}{N} & \cdots & \frac{1}{N} \end{pmatrix} \tag{16}$$

$$Pc^{\perp} = I\text{-}Pc = \begin{pmatrix} \frac{\mathrm{N-1}}{N} & -\frac{1}{N} & \cdots & -\frac{1}{N} \\ -\frac{1}{\mathrm{N}} & \frac{\mathrm{N-1}}{N} & \cdots & -\frac{1}{N} \\ : & : & & : \\ -\frac{1}{\mathrm{N}} & -\frac{1}{N} & \cdots & \frac{\mathrm{N-1}}{N} \end{pmatrix} \tag{17}$$

From this projection matrix, we can see that, to implement this normalisation we only need to subtract the mean log red value from all log red pixel values and subtract the mean log green and log blue values from the log green and log blue pixel values.

To remove lighting geometry and illuminant colour both at the same time we just apply the projectors (13) and (17) consecutively to the log image. This operation is easy to write if we think of an N pixel image as an N x 3 matrix of log RGBs. Let us denote this image as $Y\gamma$ .So for the application of the lighting geometry and light colour normalisations we can write:

$$\gamma Y = (I - Pc) Y\gamma (I - \mathrm{Pr}) \tag{18}$$

where Y represents the normalised image. To implement (18),

we simply subtract row means from rows and then column means from columns.

It is understood from projection theory that matrix $[I-\mathrm{Pr}]$ and $[I-Pc]$ are both idempotent. So (18) becomes

$$\gamma Y=(I-Pc)(I-Pc)Y\gamma(I-\mathrm{Pr})(I-\mathrm{Pr}) \tag{19}$$

This removes shading or light colour completely.

Till now we have not attempted to remove the dependence of γ. Also we have not yet considered its effect on our approach to remove lighting geometry and colour. Since γ term is a multiplicative factor in log space, it is easy to see that it does not affect the operations of the projectors involved in the procedure. Therefore the dependencies due to lighting geometry and illuminant colour are removed regardless of γ.

But our normalised image still depends on γ, and so, to make the procedure fully comprehensive we should remove the effect of gamma. By definition, the mean of all elements in Y must be zero (if the mean of the rows and columns are individually 0 then the overall mean must also be 0).But γ cannot be removed by dividing by the mean. Rather we must use a second order statistic. For that we can use variance. The variance of all the elements in Y can be calculated as:

$$\sigma^2(\gamma Y)=\frac{trace((\gamma Y)^t(\gamma Y))}{3N} \tag{20}$$

where Y has N rows and 3 columns (there are N pixels in the image) and trace( ) is the sum of the diagonal elements of a matrix. It follows that gamma can be removed by dividing by the standard deviation.

$$\frac{\gamma Y}{\sigma(\gamma Y)}=\frac{\gamma Y}{\gamma\sigma(Y)}=\frac{Y}{\sigma(Y)} \tag{21}$$

In summary, for non-iterative comprehensive normalisation, we took the log of the RGB image. Then at each pixel we subtracted the pixel mean (of the log R, G, and B responses).We then subtracted the mean of all the resulting red responses from each pixel and the mean green and blue channel responses for the green and blue colour channels. The result is an image independent of the light colour and lighting geometry. Further dividing by the standard deviation of the resultant image renders the representation independent of gamma.

3. PROPOSED WORK

This section provides an overview of the algorithm used in the work. Fig. 2 depicts the overview. The idea behind the process of an illumination invariant stereo matching is to post process input image data by forming a logarithm of a set of chromaticity coordinates, and then project the resulting 2-dimensional data in a direction orthogonal to a special direction, that best describes the effect of lighting change. We perform a non-iterative log normalisation of the image data for finding that special invariant direction. It will give an invariant image that is independent of lighting without any need for a calibration step or special knowledge about an image. The last and final step is stereo matching using Normalized cross correlation which itself is invariant to linear brightness and contrast variations.

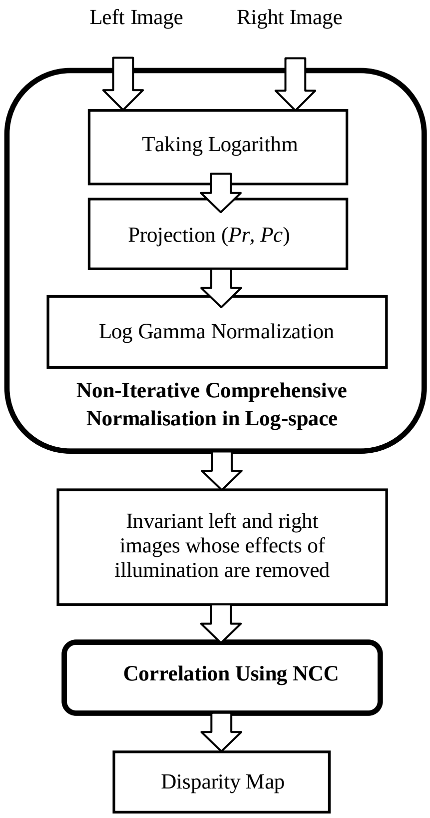


Fig 2: Overview of the proposed Algorithm

4. STEREO MATCHING USING NCC

Disparity estimation algorithms can be categorized into local and global approaches [2]. Local approach determines the disparity of a pixel based on the support window similarity. The local approach usually has low-computation complexity and storage requirement, and has been frequently adopted by real-time implementations. Global approach determines the disparity of all the pixels in an image as a whole by optimizing a global energy function. However, the optimization is usually complex and extremely computation intensive. But they usually achieve much better performance than simple algorithms. However, simple algorithms are very much faster than complex algorithms. As a result, most real-time applications have adopted simple algorithms to trade the performance for speed. Hence, we have focused on a local method, Normalized Cross Correlation (NCC) for stereo correspondence since its normalisation, both in the mean and the variance makes it relatively insensitive to radiometric gain and bias.

Let I_L and I_R are corresponding pixel values in the left and right invariant images respectively. NCC [10] is a well-known similarity measure between two pixels with neighbours that is defined by the equation

$$NCC(u,v) = \frac{\left(\sum_{u,v}\left[I_L(u,v) - \bar{I}_L\right] \times \left[I_R(u+d,v) - \bar{I}_R\right]\right)}{\left(\sqrt{\sum_{u,v}\left|I_L(u,v) - \bar{I}_L\right|^2} \times \sqrt{\sum_{u,v}\left|I_R(u+d,v) - \bar{I}_R\right|^2}\right)} \quad (22)$$

where $\bar{I}_L$ and $\bar{I}_R$ are the mean intensity values of pixels and d is the disparity. The summation is over a window which is centered on the pixel to match. Equation (22) assumes that the matches are made along a scan line. Matching a pixel from left to right image requires the computation of the matching cost with the disparity d varying from its minimum value to its maximum value. The lowest cost is taken as the best match.

Simply applying this NCC to raw stereo images in a naive fashion does not work well because the various radiometric changes caused by ρ, a, b, c, and γ are not taken into consideration [6]. The proposed method finds remedy to this problem by transforming the nonlinear relationship that exists between corresponding pixel color values into a linear one by employing log-chromaticity color space.

5. ERROR ANALYSIS

The performance of a stereo matching algorithm is usually evaluated by the error rate of a disparity map when compared to a ground truth disparity map. Lower error rate implies higher performance. To analyse the error, a novel method is proposed here by dividing the disparity image into two regions: Uniform and Discontinuity. The uniform region needs a large window support for the effective estimation of disparity while discontinuity region requires the window support to be small to effectively represent the depth discontinuities. In the error analysis we divide the ground truth and estimated disparity into uniform and discontinuity regions using an edge mask and a uniform mask. The error is estimated as given below:

$$Error = \frac{1}{m*n}\left(\frac{|E(i,j) - G(i,j)|}{\max(G(i,j))}\right) \quad (23)$$

where $m*n$ is the total number of pixels, $E(i,j)$ is the estimated disparity and $G(i,j)$ is the ground truth disparity.

6. EXPERIMENTAL RESULTS

In this section, we present some experimental results on Baby1 stereo pairs with ground truth from the Middlebury Stereo Vision page [1] .The other parameters are window size and maximum disparity. In our experiments, a fixed window of size 7x7 is selected for the study. The maximum disparity value is the maximum value of the pixel in the ground truth map, divided by the scale factor. The estimated disparity map was compared with the ground truth and the error was calculated as described in the previous section.

There are three different exposures (indexed as 0, 1, 2) and three different light sources (indexed as 1, 2, 3) in each data set in [1], resulting in total of nine different images. However, we avoided the case of the same illumination (or exposure) combination for left and right images. For instance, between the 1/3 (left/right illumination) and the 3/1 (left/right illumination) cases, we experimented only on the 1/3 case because the properties of 1/3 and 3/1 cases are similar.

6.1 Light Source Changes

For testing the effects of changes in light source (illumination), we set index of exposure to 1 for all images and varied only the index of illumination from 1 to 3. Figs. 3(a) and 3(b) depict the Baby1 stereo images taken under extremely different illumination condition (the left and the right images have been taken at an index of illumination of 1 and 3, respectively). Fig 3(c) is the ground truth disparity map. Fig 3(d) is the disparity map of proposed algorithm for input stereo image pair in Figs. 3(a) and 3(b).

a) b) c) d)

Fig 3: Result of proposed algorithm for window size 7x7 on Baby1 image pair with varying illumination. (a) The left image with illumination (1)-exposure (1). (b) The right image with illumination (3)-exposure (1). (c) The ground truth disparity map. (d) The disparity map of proposed algorithm for input stereo image pair in (a) and (b)

Table1: Error values of the disparity map produced for Baby1 image pair by the proposed algorithm according to the different left/right image illumination combinations for exposure 1.

Error Analysis	**(Left Illumination/Right Illumination)**			
	(1/1)	**(1/2)**	**(1/3)**	**(2/3)**
Uniform region error	0.107542	0.104739	0.109191	0.094362
Discontinuity region error	0.0203056	0.0206676	0.0210402	0.0188883

6.2 Camera Exposure Changes

In order to test the effects of changes in camera exposure, we fixed the index of illumination to 1, and changed only the index of exposure from 0 to 2. Figs. 4(a) and 4(b) show the Baby1 stereo images which have undergone extremely different exposure conditions (the left and the right images have been captured with an index of exposure of 0 and 2, respectively). Fig 4(c) is the ground truth disparity map. Fig 4(d) is the disparity map of proposed algorithm for input stereo image pair in Figs. 4(a) and 4(b)

a) b) c) d)

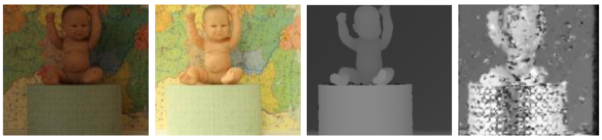
Fig 4: Results of test stereo algorithms for window size 7x7 on Baby1 image pair with varying exposure. (a) The left image with illumination (1)-exposure (0). (b) The right image with illumination (1)-exposure (2). (c) The ground truth disparity map. (d) The disparity map of proposed algorithm for input stereo image pair in (a) and (b)

Table 2: Error values of the disparity map produced for Baby1 image pair by the proposed algorithm according to the different left/right image exposure combinations for illumination 1.

Error Analysis	(Left Exposure/Right Exposure)			
	(0/0)	**(0/1)**	**(0/2)**	**(1/2)**
Uniform region error	0.167377	0.161286	0.154631	0.0988581
Discontinuity region error	0.0275638	0.0278884	0.0269872	0.0203633

6.3 Both illumination and Camera Exposure Changes

For testing the effects of changes in illumination and in camera exposure, we set index of exposure to 0 and index of illumination to 1 for left images and varied the index of illumination and exposure for right images. Figs. 5(a) and 5(b) depict the Baby1 stereo images taken under extremely different illumination and exposure condition (the left image was taken at an index of illumination and exposure of 1 and the right image with illumination index 3 and exposure index 2, respectively). Fig 5(c) is the ground truth disparity map. Fig 5(d) is the disparity map of proposed algorithm for input stereo image pair in Figs. 5(a) and 5(b).



a) b) c) d)

Fig. 5: Results of test stereo algorithms for window size 7x7 on Baby1 image pair with varying illumination and exposure. (a) The left image with illumination (1)-exposure (1). (b) The right image with illumination (3)-exposure (2). (c) The ground truth disparity map. (d) The disparity map of proposed algorithm for input stereo image pair in (a) and (b)

Table 3: Error values of the disparity map produced for Baby1 image pair by the proposed algorithm according to the different left/right image illumination and exposure combinations.

Error Analysis	(Left illumination/Right Illumination)			
(Left Exposure/ Right Exposure)	**(1/2)**		**(1/3)**	
	Uniform region error	Discontinuity Region Error	Uniform region error	Discontinuity Region error
(0/1)	0.158135	0.027148	0.160088	0.0275397
(0/2)	0.150862	0.0264031	0.153133	0.0267344

7. CONCLUSIONS AND FUTURE WORK

We have implemented a real time stereo matching technique for radiometric changes by finding left and right invariant images that is free of radiometric effects without any need for a calibration step or special knowledge about an image. This method produces promising results that are quite robust to various kinds of radiometric changes. The disparity map of the proposed method was evaluated using standard datasets and the results are comparable to state-of-art techniques in the literature. In future, the disparity map can be improved using some adaptive window methods whose complexity is low.

8. REFERENCES

[1] Middlebury Stereo Vision Page: http://vision.middlebury.edu/stereo

[2] D. Scharstein and R. Szeliski, “A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms,” Int’l J.Computer Vision, vol. 47, no. 1, pp. 7-42, 2002.

[3] Finlayson, Graham D. and Xu, Ruixia (2003) “Illuminant and Gamma comprehensive normalisation in log RGB space,” Pattern Recognition Letters, 24 (11). pp. 1679-1690. ISSN 0167-8655

[4] Finlayson, G.D., Xu, R., “Non-iterative comprehensive normalisation,” First European Conference on Color in Graphics, Image and Vision (CGIV), 2002

[5] Y.S. Heo, K.M. Lee, and S.U. Lee, “Illumination and Camera Invariant Stereo Matching,” Proc. IEEE Conf. Computer Vision and Pattern recognition, 2008.

[6] Yong Seok Heo, Kyoung Mu Lee, Sang Uk Lee, “Robust stereo matching using adaptive normalized cross correlation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 33, No. 4, April 2011.

[7] G. D.Finlayson, Bernt Schiele, and James L.Crowley, “Comprehensive Colour Image Normalization,” 5th European Conference on Computer Vision, page 475-490, 1998.

[8] Heiko Hirschmüller, Daniel Scharstein, “Evaluation of Stereo Matching Costs on Images with Radiometric Differences,” IEEE Trans. Pattern Anal. Mach. Intell. 31(9): 1582-1599 (2009)

[9] H. Hirschmuller and D. Scharstein, “Evaluation of Cost Functions for Stereo Matching,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2007.

[10] O. Faugeras, B. Hotz, H. Mathieu, T. Vie´ville, Z. Zhang, P. Fua, E.The´ron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proy, “Real Time Correlation-Based Stereo: Algorithm, Implementations and Applications,” Technical Report RR-2013, INRIA, 1993.

[11] A.S. Ogale and Y. Aloimonos, “Robust Contrast Invariant Stereo Correspondence,” Proc. IEEE Int’l Conf. Robotics and Automation, 2004.

[12] J. Zhang, L. McMillan, and J. Yu, “Robust Tracking and Stereo Matching under Variable Illumination,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2006.

[13] G.D. Finlayson, S.D. Hordley, and M.S. Drew, “Removing Shadows from Images,” Proc. European Conf. Computer Vision, 2002.