A Novel Approach for Increased Depth Perception in Stereo Images

Anju. J[1], Avani Nath. N. J[2]

[1]PG Scholar, Department of Electronics and Communication Engineering
TKM Institute of Technology, Kollam, Kerala, India

[2]Assistant Professor, Department of Electronics and Communication Engineering
TKM Institute of Technology, Kollam, Kerala, India

Abstract: ***Stereoscopy is a technique to create the illusion of depth from given two- dimensional images. Most stereoscopic methods present two offset images separately to the left and right eye of the viewer. These two dimensional images are then combined in the brain to give the perception of 3D depth. Binocular Just-Noticeable Difference (BJND) is a visibility threshold defined for the stereo image and it has been modeled using psychophysical experiments. It describes the sensitivity of the human visual system to luminance changes, in stereo images. In this algorithm, the BJND model is utilized for enhancing the sharpness of stereo images. An over enhancement problem in the sharpness enhancement of stereo image is also defined and an effective solution is found within an optimization framework. The algorithm can provide sharpness enhanced stereo images without excessive distortion.***

Keywords: Binocular Just- Noticeable Difference (BJND), Human Visual System (HVS), image enhancement, stereo image

1. Introduction

Three dimensional stereoscopic displays can provide additional information to the viewers, compared to the two-dimensional displays. Such displays are widely used for entertainment, medical imaging, etc. One well-documented problem in the 3D image perception is that the 3D shape and scene layout are often distorted. In some applications, such as cinema, the distortions are not necessarily a serious problem for the designer, but in other applications such as medical imaging, they can have consequences. So, quality enhancement for stereo image is very important for the 3D display applications.

Human depth perception can be increased by enhancing the sharpness of the stereo image. Therefore, among the different types of quality enhancement processes, sharpness enhancement is a crucial component in most of the 3D display applications. The conventional approaches for the sharpness enhancement of stereo images neglect inter differences of the luminance values between the left and right images. The inter difference between a stereo image pair may produce visual fatigue. So for developing a sharpness enhancement algorithm [1] for stereo images this inter difference problem should be considered.

The stereo data can be obtained in a number of ways. The most common approach is to use a combination of left and right viewpoint images. Quality of the stereo content can be improved by enhancing the left and right images independently. However, cross-view image processing [2] is clearly advantageous because humans do not perceive left and right images independently.

There have been very limited researches in the area of specialized sharpness enhancement techniques for the stereo images. Any image enhancement method [3] [4] [5] [6] [7] developed for the planar images can be applied separately to the left and right images so that the quality of the stereo image gets enhanced. To improve the depth perception of a planar image, the depth value can be unsharp masked as described in [8]. But these enhancement techniques do not consider the inter difference between the left and right images and so the enhanced stereo image may not be pleasing for the observer. The rest of this paper is organized as follows: In Section 2, the proposed algorithm is described in detail. Experimental results and discussions are presented in Section 3 and Section 4 concludes this paper.

2. Method

Based on the fact that the HVS is more sensitive to closer objects, the technique simply controls the strength of the sharpness increase according to estimated disparity (horizontal shift) values. In this algorithm, the left and right view images of the stereo image pair are enhanced using adaptive unsharp masking [9] which is advanced over the classic unsharp masking algorithm for sharpness enhancement. The disparity information is also extracted from the stereo image pair and it is utilized for the visibility threshold (BJND) estimation. The over enhanced pixel pairs are then extracted and the optimization algorithm called Levenberg- Marquardt algorithm [13] is applied to produce the stereoscopic content without excessive distortion. A generalized block diagram for the entire algorithm is shown in Figure1.

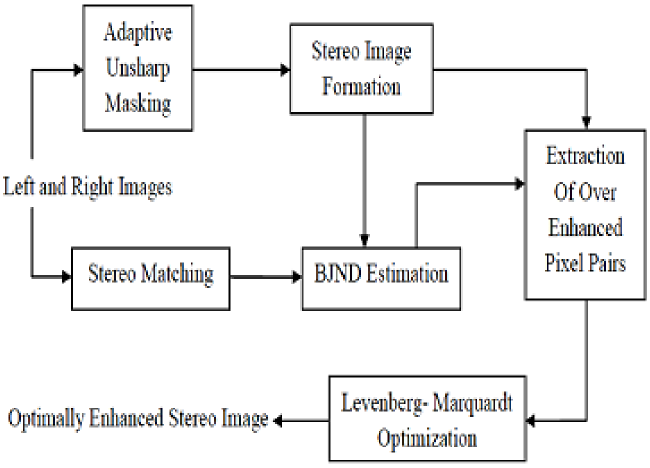

Figure 1: Block diagram for the proposed system

2.1 Adaptive Unsharp Masking

The objective of the adaptive filter is to emphasize the medium-contrast details in the input image more than large-contrast details, so as to avoid overshoot effects in the output image. The adaptive filter does not perform a sharpening operation in smooth areas, and therefore the overall system is more robust to the presence of noise in the input. The adaptive unsharp masking technique [9] accomplishes the dual objectives of avoiding noise amplification as well as excessive overshoot in the detail areas. So it is more appropriate for the stereo image enhancement.

The same algorithm [9] is applied to both the left and right images of a stereo pair and thus the sharpness enhanced stereo image is obtained. But, unfortunately there occurs some over enhancement due to the inter difference problem between the left and right images. The over enhancement in the sharpness enhancement of stereo images needs to be controlled by some post processing based on the Human Visual System (HVS) characteristics.

2.2. Disparity Estimation

Disparity estimation is a correspondence problem. The correspondence between the two images can be determined by either matching features or by operating on or matching of all small patches of gray values. Feature matching can be used as a preprocessing step, for the extraction of appropriate features from the images such as corners and edges. After obtaining the features, the correspondence problem is first solved for the spatial locations at which the features occur, from which next the full disparity field can be deduced. Typically, a matching cost is computed at each pixel [10] for all disparities under consideration.

Local matching requires defining a matching score. Here, a matching score that is a sum of absolute intensity differences (SAD) [11]and a gradient based measure (GRAD) are used [1]. The absolute difference measure assumes brightness constancy for corresponding pixels. It can be defined as,

$$C_{SAD}(i,j,d) = \sum |I_l(x,y) - I_r(x+d,y)| \quad (1)$$

Where 'i' and 'j' represent the pixel locations, 'd' the disparity level, 'I_l' and 'I_r' represent the corresponding intensity levels of the left and right images. Similarly, the gradient based matching score can be defined as,

$$C_{GRAD}(i,j,d) = \sum |\nabla_x I_l(x,y) - \nabla_x I_r(x+d,y)| + \sum |\nabla_x I_l(x,y) - \nabla_y I_r(x+d,y)| \quad (2)$$

Where ∇_x and ∇_y are the horizontal and vertical gradients, respectively. An optimal weighting 'w' between C_{SAD} and C_{GRAD} is determined by maximizing the number of reliable correspondences with a winner-take-all optimization (choosing the disparity with the lowest matching cost). Then, the resulting dissimilarity measure is given by,

$$C(i,j,d) = (1-w)C_{SAD}(i,j,d) + wC_{GRAD}(i,j,d) \quad (3)$$

The efficiency of the stereo matching is evaluated by comparing the obtained disparity map with the ground truth disparity map. The time taken for the matching process is also a major factor in considering the efficiency of stereo matching.

2.3. Overview of BJND

The visibility threshold, Binocular Just Noticeable Difference, for stereo image is related to the inter difference between the left and right views [12] that a human can recognize. A human cannot realize a distortion when viewing the stereo image if the distortion in one viewpoint image is less than the BJND. The BJND describes the minimum distortion in one view that evokes perceptual differences in the stereo image. Many psychophysical experiments, has showed that the BJND is dependent on the well-known HVS characteristics of luminance adaptation and contrast masking.

Since the BJND is dependent on the distortion between the left and right images, the disparity information between two viewpoint images is utilized [12]. So, the correspondence matching between two views, i.e, stereo matching has to be performed to obtain the disparity information. Consider the left and right images with the disparity image corresponding to the left image, the BJND at the left view, is defined as follows

$$BJND_l(i,j,d) = BJND_l(bg_r(i+d,j), eh_r(i+d,j), n_r(i+d,j)) \quad (4)$$

bg_r – background luminance of right image
eh_r - the edge height of the right image
n_r - noise present in the right image
If there is no noise in the right view, BJND is equivalent to,

$$BJND_L(bg,eh) = T_{limit}(bg) + K(bg).eh \quad (5)$$

The background luminance 'bg' is obtained by averaging the luminance values in the 5 x 5 region at the corresponding pixel position. Edge Height,

$$eh = \sqrt{(E_h{}^2 + E_v{}^2)} \quad (6)$$

where E_k is defined as,

$$E_k(i,j) = (1/24)\sum_{h=1}^{5}\sum_{v=1}^{5} I(i-3+h, j-3+v).G_k(h,v) \quad (7)$$

where, G_k is the sobel edge[12] operator , k=h or v, I(i,j) denotes the luminance values at location (i,j) and 'T_{limit}' is the elevated threshold due to the contrast masking effect, and 'K(bg)' is the fitting function for the elevating factor T_{limit}(bg) and K(bg) are defined as,

$$T_{limit} = \begin{cases} 0.0027.(bg^2 - 96.bg) + 8, if\ 0 \le bg \le 48 \\ 0.0001.(bg^2 - 32.bg) + 1.7, if\ 48 \le bg \le 255 \end{cases} \quad (8)$$

and $K(bg)=-10^{-6}.(0.7.bg^2-32.bg)+0.07$ (9)

The above two functions in (8) and (9) are determined by using the psychophysical experimental data [12]. The BJND at the right view, i.e., $BJND_r$, can be similarly obtained. The over enhanced pixel pairs can be extracted from the sharpness enhanced stereo using the BJND values and the over enhancement is suppressed by using a Levenberg Marquardt algorithm.

2.4 Sharpness Enhancement Algorithm

Instead of developing a sharpness enhancement algorithm, the algorithm discussed here, post processes the enhanced image of each view in a way that the inter difference should not exceed the BJND. Let I_l and I_r denote the luminance components of the original left and right images, and $I_{l'}$ and $I_{r'}$ denote the corresponding luminance components of the sharpness-enhanced images, respectively.

Similar to the overshoot artifact problem in the sharpness enhancement of the 2-D image, there occurs an unnecessary increase in inter difference values, and can be considered as an overshoot artifact in the stereo image domain. This suggests that the sharpness enhancement needs to be controlled according to both the BJND and the stereo matching accuracy.

Let a=I_l(i,j) and b= I_r(i, j) be the original left and right image, and a'=I_l(i,j) and b'= $I_{r'}$(i, j) be the enhanced left and right image. The resultant luminance values are determined by solving the following equation

$$(a^*,b^*)_{i,j,d} = \textbf{argmin}\{(a-a')^2+(b-b')^2+\alpha(abs(a'-b')-BJND_{l'}(i,j,d))^2+\alpha(abs(a'-b')-BJND_{r'}(i,j,d))^2\} \quad (10)$$

Where 'α' is the weight for the constraint terms related to BJND. The constraint terms take effect only if the inter difference between two luminance values is larger than the BJND value. So, the luminance values are maintained if the inter difference is already lower than or equal to the BJND value. The aforementioned equation is solved using the Levenberg – Marquardt Algorithm [13] using p' and q' as the initial estimates.

A compensated right-view image can be reconstructed by using the left view image and the disparity image. All the techniques of the BJND estimation, the sharpness enhancement, and the proposed post processing are based on modification of the luminance components. Thus, the color space of the input images is converted to the hue–saturation–luminance space, (RGB to HSL) and the techniques are applied only to the L component. The resultant stereo images are finally obtained by converting the color space back to the RGB space.

3. Experimental Results

Performance of the algorithm is evaluated using the stereo image pairs available in [14]. The test stereo is shown in figure 2.

Figure 2:.Left and Right images of stereo pair

From the left and Right images,, the intensity components are extracted and adaptive unsharp masking is applied so that sharpness enhanced stereo pair can be produced. The original and the enhanced intensity components are shown in Figure 3.

Figure 3: Original and enhanced intensity components

The original and the enhanced stereo image are shown in Figure 4. Here an anaglyph type of stereo image is formed and the 3D depth can be viewed through an anaglyph glass.

Figure 4: Original and the enhanced anaglyph stereo

From the sharpness enhanced stereo image, it is clear that there occurs some sort of over enhancement and it need to be post processed to obtain a stereo image without excessive distortion. The post processing described in this algorithm considers the visibility threshold Binocular Just Noticeable Difference (BJND). The estimation of the BJND depends on the stereo matching technique. So, the disparity estimation has to be performed prior to the BJND estimation. The disparity estimation gives information about the relative distances between the objects in the scene, and also the closeness of the different points to the eyes. Figure 5 shows the ground truth disparity map and the obtained disparity map.

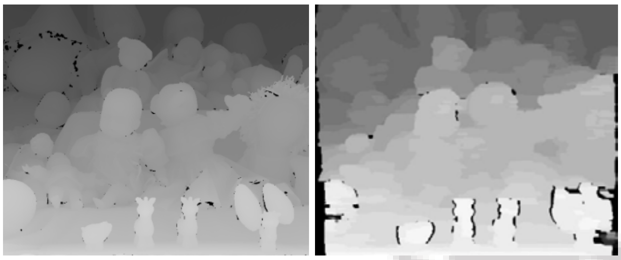

Figure 5: Ground truth disparity and obtained disparity map

The estimation of the visibility threshold BJND is an important step as it decides the amount of over enhancement. Normally the BJND values [12] are less than 15for both the left and right views.

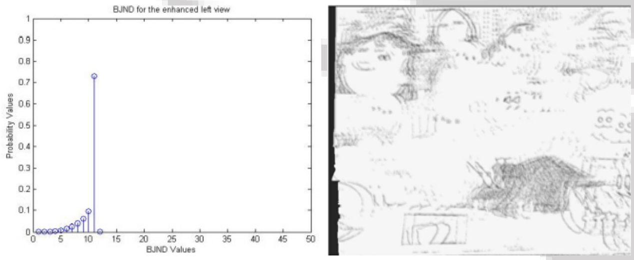

Figure 6: BJND Histogram and BJND image for stereo pair

Then the inter difference between the left and right views for the original stereo pair and the enhanced stereo pair are calculated and shown in Figure 7. It shows that the inter difference has increased in the enhanced stereo pair.

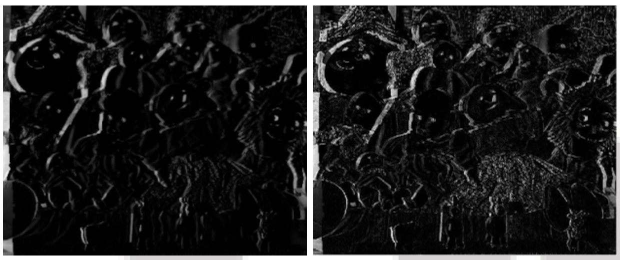

Figure 7: Interdifference for original and enhanced stereo

The pixel pairs with increased inter difference above the BJND are extracted and an optimization algorithm called Levenberg- Marquardt algorithm [13] is applied on the over enhanced pixel pairs. Thus the minimum distortions in one view that can be admitted for a stereo image pair are obtained, and the inter difference for the pair of stereo image are brought equal to or below BJND. Figure 8 shows the resultant stereo image.

Figure 8:.The original and resultant stereo pair

It is clear that resultant stereo image eliminates the over enhancement problem based on BJND and thus the technique provides subjectively pleasant stereo images.

4. Conclusion

Quality enhancement for stereo content is one of the most important techniques for 3D display applications. In 3D stereo setup two separate views are presented to each eye. A combination of left and right viewpoint images is widely used. The brain fuses these two images and uses the disparity between the two views as the binocular cue to perceive the depth. In most of the cases the 2D processing steps are applied directly to the 3D stereo content. But, inter difference between the left and right views is an additional attribute in stereo set-up which needs to be considered when mitigating the artifacts and enhancing the 3D stereo content. This sharpness enhancement technique considers an over enhancement problem, based on the visibility threshold, BJND. The sharpness of stereo image is increased using adaptive unsharp masking. Stereo matching is performed to extract the disparity information of the stereo content. The estimated BJND model decides the over enhancement and Levenberg -Marquardt algorithm has been implemented to reduce the over enhancement. Thus, sharpness-enhanced stereo images without excessive distortion over the BJND can be produced. The reliability of the algorithm depends on the accuracy of the stereo matching. So a better stereo matching technique [15] that can provide more accurate BJND values will be considered for future work.

References

[1] Seung - Won Jung, Jae - Yun Jeong ,and Sung – Jea Ko, "Sharpness enhancement of stereo images using Binocular Just -Noticeable Difference, "*IEEE Trans. Image Process.*, vol. 21, no. 3,pp.1191-1199, March 2012

[2] Varuna De Silva, Anil Fernando, Stewart Worrall, Hemantha Kodikara Arachchi, and and Ahmet Kondoz, " Sensitivity Analysis of the Human Visual System for Depth Cues in Stereoscopic 3-D Displays", *IEEE Transactions On Multimedia*, vol. 13, no. 3,pp.498- 506, June 2011

[3] Joung-Youn Kim, Lee-Sup Kim, and Seung-Ho Hwang,"An Advanced Contrast Enhancement Using Partially Overlapped Sub-Block Histogram Equalization", *IEEE Transactions On Circuits And Systems For Video Technology*, vol. 11, no. 4,pp.475-483, April 2001

[4] S. Srinivasan and N. Balram, "Adaptive Contrast Enhancement Using Local Region Stretching",in *Proc.of ASID*, pp.152-155, New Delhi, October 2006
[5] Hasanul Kabir,Abdullah Al-adud, and Oksam Chae, " Brightness Preserving Image Contrast Enhancement Using Weighted Mixture of Global and Local Transformation Functions",*The International Arab Journal of Information Technology*, vol. 7, no. 4, pp.403-409,October 2010
[6] Tang, Eli Peli and Scott Acton," Image Enhancement Using a Contrast Measure in the Compressed Domain", *IEEE Signal Processing Letters*, vol. 10, No. 10,pp.289-292,October 2003
[7] Haidi Ibrahim,"Histogram Equalization with Range Offset for Brightness Preserved Image Enhancement", *International Journal of Image Processing* , vol.5,pp.599-609, 2011
[8] Thomas Luft, Carsten Colditz, and Oliver Deussen," Image Enhancement by Unsharp Masking the Depth Buffer",*SIGGRAPH* 2006
[9] Polesel, G. Ramponi, and V. J. Mathews, "Image enhancement via adaptive unsharp masking, " *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 505–510, March 2000.
[10] T. Kanade and M. Okutomi, "A stereo matching algorithm with an adaptive window: Theory and experiment,"in *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, Sacramento,CA, USA, Apr. 1991.
[11] Madain Perez, Francois Cabestaing, Olivier Colot and Pierre onnet , "A Similarity based Adaptive Neighborhood Method For Correlation - based Stereo Matching", *IEEE International Conference on Image processing* (ICIP) Singapore, 2004
[12] Y. Zhao, Z. Chen, C. Zhu, Y. P.Tan, and L.Yu, " Binocular Just–Noticeable Difference model for stereoscopic images", *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 19 – 22, January 2011.
[13] J. Nocedal, S. J. Wright, *Numerical Optimization*, 2nd ed. New York: Springer – Verlag, 2006.
[14] *Middlebury Stereo Vision*, [Online]. Available: http://vision.middlebury.edu/stereo
[15] Qingqing Yang, Dongxiao Li, Lianghao Wang and Ming Zhang, " Full – Image Guided Filtering for Fast Stereo Matching ", *IEEE Signal Processing Letters*, vol. 20 20, no. 3, pp.237- 241,March 2013

Author Profile

Anju. J received B. Tech. Degree in Electronics and Communication Engineering from Cochin University of Science and Technology (CUSAT) in the year 2011. She is now pursuing M. Tech. degree (CUSAT) from TKM Institute of Technology, Kerala.

Avani Nath. N. J received B. Tech. and M. Tech Degree in Electronics and Communication Engineering from Kerala University in the year 2009 and 2012 respectively. She is now working as an Assistant Professor in TKM Institute of Technology, Kerala.