International Journal of Computational Intelligence and Informatics, Vol. 3: No. 1, April - June 2013

Multi-camera Image Quality Measure in Video Images using Sub-pixel Allocation Algorithm

A Niranjil Kumar
Department of Electronic Communication Engineering,
P.S.R.Rengasamy college of Engineering for women,
Sivakasi, Tamil Nadu, India.

C Sureshkumar
Department of Computer Science Engineering,
J.K.K.Nattraja College of Engineering and Technology,
Komarapalayam, Tamil Nadu, India

***Abstract*- Multi-camera applications are numerous and each application has its specific means of acquisition representation and display. The quality of the perceived multi view video image is dependent on the means of presentation. The most of the fundamental problem in MIQM (Multi-camera Image Quality Measure) is finding the image quality by reducing the distortions. Distortions in multi-camera system can be classified into geometric and photometric distortions. Geometric distortion in multi-camera system is structural disparity such as discontinuity and misalignment in the observed image due to geometric error. Geometric error can occur during mapping which may include rotation and translation. Photometric distortion in single camera is degradation in perceptual feature that are known to attract visual attention such as noise blur and blocking artifacts. MIQM is comprised of three index measures; luminance and contrast index, spatial motion index and edge-based structural index. We propose multi-camera image quality measure is a combination of these three index measure that captures the impact of distortions on multi view perception. The measure was designed to capture the visual effects of artifacts introduced at the acquisition and pre compositing process to predict the composed image quality. By reducing the distortions the quality of video image is improved by sub-pixel allocation algorithm.**

Keywords-MIQM, Image Quality, Sub Pixel Allocation Algorithm, Geometric Distortions, Photometric Distortions, Video Processing.

I. INTRODUCTION

The multiple-camera setup, multiple-camera mode of production, or multi-camera is a method of filmmaking and video production. Several cameras either film or professional video cameras are employed on the set and simultaneously record a scene. While doing so video images are subject to a wide variety of distortions during acquisition process, compression, storage, transmission and reproduction. Any of which may result in a degradation of visual quality. For application in which images are ultimately to be viewed by human beings, the only "correct" method of quantifying visual image quality is through subjective evaluation. In practice however, subjective evaluation is usually too inconvenient, time-consuming and expensive.

Synthetic visual artifacts in images share similar acquisition apparatus and pre-compositing processing block. So the single quality measure that would capture the perceived quality of all multi-camera applications is impossible [1]. Image distortion types in multi camera systems are investigated where distortion is classified as either geometric or photometric. In photometric distortion the intensity values varies without changing location of the original edge at that same time geometric distortion the intensity varies with changing the location of the edges. A new algorithm that characterizes the type of distortion in a given image captured by a multi camera system is proposed and evaluated. It is based on the Edge Intensity Summation (EIS). It is only used to detect the type of distortion but not used to reduce the distortion [2]. The use of structural similarity as an alternative motivating principle for the design of image quality measure so, they used SSIM (Structural Similarity Index Measure). One problem with this method is that the resulting SSIM index map often exhibits undesirable "blocking" artifacts [3].

Quality assessment of multimedia content is achievable either through subjective tests or through objective metrics. Subjective tests are used to correct and reliable quality. Objective metrics predicting is the perceived quality of image. A quality metric for the assessment of stereo pairs is using the fusion of 2D quality metrics and the depth information. The naturalness, viewing experience is studies related to 3D displays. So the "viewing experience", "naturalness", "presence" of the image is difficult to obtain [4]. Two different metrics are used in quality assessment in free view point. They are full reference metric and no reference metric. Full reference metric is used to measure the fidelity of the image. No reference metric is used to measure the error in the image. The framework is applied in both the metrics. In full reference comparison an explicit ground truth reference is required but it is not needed in no reference metrics. Here Root Mean Square Error (RMSE) occurs across the entire image. Any errors at distinct image features cannot measure [5].

The 3D video is represented in the format of monoscopic color video augmented by per-pixel depth map and the encoded with H.264 encoder. To optimize the encoding performance they test different bit budges for the color video and the depth and measure the quality by virtual view quality metrics. This virtual view quality metrics are divided into subjective metric and objective quality metric. The depth information is to be less important than color distortion in subjective metric. But objective quality metric gives equal grades. The quality estimation method is not perfect and better metrics are needed [6]. For combining two or more images into a large image mosaic a multi resolution spline technique is used. The algorithm like pyramid and fast algorithm are included. It is used to decrease the image sample intensity with iteration so that the bandwidth is also reduced. The problem in applications of photo mosaics is joining two images so that the edge between them is not visible. Due to the difference in camera positions or in image processing, the gray level differences are frequently unavoidable [7].

Taking these problems into considerations, we propose a new algorithm called sub-pixel allocation algorithm in luminance and contrast index measure in order to find the distortion in the video image. By using this algorithm we found that as we got better result when compare with previous related work.

II. PROPOSED SYSTEM

In the proposed system, the distortions present in an image are reduced by using sub-pixel allocation algorithm. Here we used three index measures. The first one is luminance and contrast index, spatial motion index and edge based structural index. Distortions in multi-camera system can be classified into geometric and photometric distortions.

Photometric distortions occur due to blurring, blocking artifacts, misalignment and over lapping in an image. Photometric distortion can be intrinsic due to the acquisition device or extrinsic due to applications, such as loss compression, transmission over error prone channels, or image enhancements. Quantifying the impact of these distortion types on perceptual quality is essential to the improvements or developments of new video or image applications, and hence has motivated the development of contemporary image and video quality metrics. In multi-camera systems, photometric distortions are the visible variations in brightness levels and color gamut across the entire displayed image. The source of this variation can be the non-uniformity between individual camera properties or the post processing applications, such as compression. Human perception is sensitive to abrupt local changes in images. Photometric distortions are especially obvious around overlapping and content rich areas of the captured images. They are reduced by using luminance and contrast index and edge based structural index.

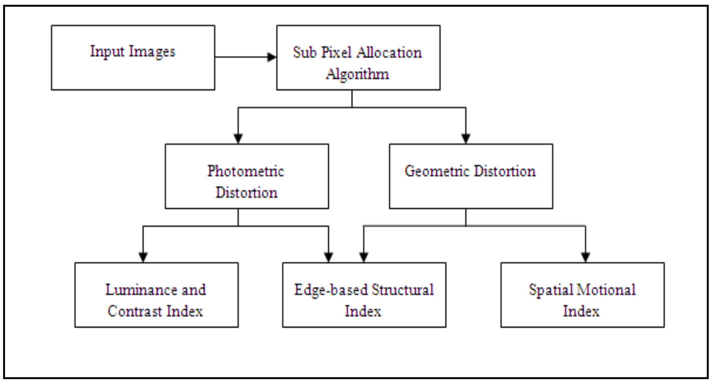

Figure 1. Flow chart of proposed method

Geometric distortions occur during mapping which may include rotation and translation. In multi-camera systems, a scene captured by N cameras can vary with each individual camera’s position and orientation. Geometric distortions are the visible misalignments, discontinuities, and blur in the processed image. These distortions could result from noticeable calibration errors between adjacent cameras, affine/linear corrections, and error in scene geometry estimations. In manually built multi-camera arrays, these errors could also result from the mismatch in the vertical and horizontal directions among images and irregular camera rotations. There are two types of geometric distortions planar and perspective distortions. Planar distortions can occur during the mapping, which may include rotation and translation. Perspective distortions can occur in the mapping from the 3-D world to the 2-D plane of the image. In multi-camera systems such errors can also occur when mapping a certain camera plan to another reference camera plane in the systems.

The proposed model assumes that single view image perceptual distortions caused by spatial scaling, rotation and translation are significant. This assumption is true for multi view images, where discontinuities, misalignments, blur and double images can results in catastrophic distortions. Geometric distortions are reduced by using spatial motion index and edge based structural index Figure 1. The enhancement process is done by using scaling, translation and rotation effects in edge based structural index to obtain a clear result. So the quality of the multi-camera system is increased.

III. SUB PIXEL ALLOCATION ALGORITHM

Video images are discredited into pixels. Each pixel corresponds to an integer-valued location. Integer-valued locations are not accurate enough for many applications, such as tracking, camera calibration, image registration and mosaic king, 3D reconstruction. To achieve better accuracy we need floating-point-valued locations i.e., sub-pixel localization.

A. General Idea

A model of the feature is developed to be localized. Conventional algorithm is applied on input image to detect feature up to pixel accuracy. Then iteratively model is matched with input image to localize detected feature with sub pixel accuracy. Most sub pixel algorithms require a good estimate of the location of the feature. Otherwise, the algorithms may be attracted to the noise instead of desired features**.**

B. Point Localization

A point usually occupies more than one pixel. A point does not have sharp edges. The edges are smooth or blurred. An appropriate model of point is 2D Gaussian (un normalized).

$$g(x,y) = exp(-\frac{x2^2+y2^2}{2\sigma^2}) \quad (1)$$

So, a point can be modelled by the 2D function M as follows.

$$M(x,y;A,B,\sigma,u,v) = A + Bexp(\frac{(x-u)^2-(y-v)^2}{2\sigma^2}) \quad (2)$$

M: intensity, (x, y): any location in the image. A: intensity of background (dark region). B: peak intensity of point (brightest region). (u, v): peak location, i.e., center of point σ: amount of spread of the Gaussian. In short-hand notation: M(x, μ)

x = (x, y)T : variable image location (3)

θ = (A,B, ¾, u, v)T : parameters of the point model. If the model M matches a point in image I perfectly,

then $M(X,\theta) = I(X)$ (4)

For all locations x within the model M. (u, v) gives the location of the point. Since u, v can be taken on floating-point values; they indicate a sub pixel location. To obtain a good match, compute error of match

$$E(\theta) = \sum_{X\in W}[M(X,\theta) - I(X)]^2 \quad (5)$$

where W is the extent of M (like a small window or template). Next, apply appropriate algorithm to find the μ that minimizes the error E(μ). The sub pixel location is the (u, v) of the optimal μ.

Method 1: Direct Solution- Do the usual thing: ${\partial E}/{\partial \theta}$ = 0. Then, rearrange the terms to try to obtain a set of equations that can be solved.

Method 2: Apply Optimization Algorithm- Some possible algorithms: Gradient descent. Compute ${\partial E}/{\partial \theta}$ .Then, change μ iteratively until it converges:

$$\theta(t+1) = \theta(t) - \eta\frac{\partial E}{\partial \theta} \quad (6)$$

C. Edge Localization

Edge localization can be performed in a similar manner. An edge is defined by a change of intensity: Derivation of Edge Model

(O, x, y) is the global coordinate system of the image. (O′, x′, y′) is the local coordinate system in which the edge i. A unit step edge is

$$U(x) = \begin{cases} 1\, if\ x > 0 \\ 0\, if\ otherwise \end{cases} \quad (7)$$

An ideal 2-D step edge S located at O′ in the coordinate system (O′, x′, y′) along the y′-axis is given by

$$S(x', y') = U(x') \quad (8)$$

A 2-D blurred edge F can be modeled by convolving the 2-D step edge S with a 1-D Gaussian G across the edge:

$$F(x', y', \sigma) = \int G(w;\sigma)S(x'-w, y')dw \quad (9)$$

Where, $G(w,\sigma) = \exp(\frac{w^2}{2\sigma^2})$ (10)

O′ is located at (u, v) of the global coordinate system. So, transform edge from local system to global system:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} (x-u)cos\Theta + (y-v)sin\Theta \\ -(x-u)sin\Theta + (y-v)cos\Theta \end{bmatrix} \quad (11)$$

Let the gray level on the darker side be A and the gray level on the brighter side be B. Then, the final edge model M is:

$$M(x,y,\theta) = A + BF(x',y',\sigma) \quad (12)$$

Where, $\theta = (u,v,\theta,\sigma,A,B)^T$ is the parameter vector. Now, we can compute the error of match E(μ) as

$$E(\theta) = \sum_{X\in W}[M(X,\theta) - I(X)]^2 \quad (13)$$

Where W is the extent of M. Next, apply appropriate algorithm to find the μ that minimizes the error E(μ). The sub pixel location is the (u, v) of the optimalμ. The orientation of the edge is perpendicular to μ of the optimal θ.

In this paper, we have introduced the various multi-camera applications and the different type of distortions affecting each one of them. Then, we studied two particular types of distortions that are unique to multi-camera images. We provided examples on how each can influence multi-camera image perceived quality. Then, we introduced a MIQM as a combination of three index measures.

We presented the derivation and reasoning for each index measure. We used sub pixel allocation algorithm to find the distortion in the image instead of bit allocation algorithm. And also we tuned our pixel parameter value to increase image quality and also we developed MIQM for scaling, translation and rotation images with this work. Finally, we compared MIQM against a database of multi-camera images. We ran a set of subjective tests to evaluate the quality of the images in the database, and the MOS (Mean Opinion Score) score was calculated for each image. The results and examples showed that MIQM outperforms SSIM and PSNR (Peak Signal to Noise Ratio) for multi-camera images quality assessment.

IV. Experimental Results

In our paper we have used video image from PETS data set. In a video sequence each frame is trained in order to find out the occurrence of error in an input image. The errors such as Gaussian, average, motion, and logarithm may occur in video. The input image with error is given in the Figure 2. Mapping of each input images into matrix form and set co-ordinates values in the (x, y) directions is done. Then all the images are separated into a set of macro blocks. After that the sub pixel algorithm is applied to find which type of distortion occurred in that pixel.

Figure 2. Input image with error

Figure 3. Enhanced image

Figure 4. Quality improved image

If the distortion is geometric then spatial motion index or edge based structural index is used or if the distortion is photometric then luminance and contrast index or edge based structural index is used to reduce the distortion. By using these three methods translation and rotation are done perfectly. Thus the quality of the image is improved by using sub pixel allocation algorithm is given in the Figure 4. To analyse the performance of proposed method SSIM, MSE (Mean Squared Error) and PSNR are calculated on various windows of an image frame in video. To obtain the relationship between SSIM and PSNR, the relationship between the SSIM and the MSE is derived and then we use that relationship to link SSIM to PSNR. The measure between two windows *x* and *y* of common size M×*N is.*

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (14)$$

Where μ_x the average of x;μ_y the average of y;σ_x^2 the variance of x; σ_y^2 the variance of y; σ_{xy} the covariance of x and y ; $c_1 = (k_1L)^2$, $c_2 = (k_2L)^2$ two variables to stabilize the division with weak denominator; L is the dynamic range of the pixel-values ($2^{\#bits\ per\ pixel} - 1$); $k_1 = 0.01$ and $k_2 = 0.03$ by default.

$$MSE(x,y) = \sigma_x^2 + \sigma_y^2 - 2\sigma_{xy} + (\mu_x - \mu_y)^2 \tag{15}$$

$$PSNR(x,y) = 10\log_{10}\left[\frac{255^2}{MSE(x,y)}\right] \tag{16}$$

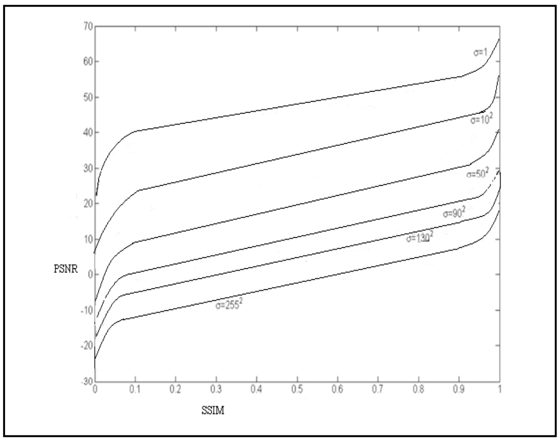

Figure 5. Performance Analysis of Proposed Method

The performance of proposed method is analysed in Figure 5. The graph is drawn by taking the variation of the PSNR as function of the SSIM for different values of σ_{xy} (covariance between x and y). When SSIM varies in the interval (0.1, 0.9), the curves are straight lines. It shows that the proposed work result is more efficient in order to improve the quality of video images.

V. Conclusion

In the previous related works scaling, rotation, translation is not done perfectly but it is difficult to measure the quality of multi-camera. In this paper, we introduced sub-pixel allocation algorithm in luminance and contrast index in order to improve quality of the video images by reducing the distortion. In order to test the performance of MIQM, We conducted an extensive subjective quality assessment study. First, we produced a database of multi camera images generated using the techniques described earlier in this research, where various combinations of geometric and photometric distortions were applied. In this experiment, a total of 10 images, were evaluated, and the raw scores for each subject were processed to give mean opinion scores and a difference mean opinion score for each distorted image. The test images had varying types and levels of distortions. We calculated MSE, SSIM and PSNR. SSIM and PSNR are designed to capture the quality of single view images contradict with the actual perceived quality of multi-camera images. Thus, by using sub-pixel allocation algorithm the quality of image is improved by reducing the distortions.

References

[1] Mashhour Solh, Member, IEEE, and Ghassan AlRegib, Senior Member, IEEE, “Multicamera Image Quality Measures”, IEEE Transaction on Image Processing, vol. 21, No. 9, Sep. 2012.

[2] M. Solh and G. AlRegib, “Characterization of image distortions in multi-camera systems”, in Proc. 2nd Int. Conf. Immersive Telecommun., May 2009, pp. 1–6.

[3] S. Leorin, L. Lucchese, and R. Cutler, “Quality assessment of panorama video for videoconferencing applications”, in Proc. IEEE Workshop Multimedia Signal Process, Nov. 2005, pp. 1–4.

[4] R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, “High-speed digital-to-RF converter”, U.S. Patent 5668842, Sept. 16 1997.

[5] P. Campisi, A. Benoit, P. Callet, and R. Cousseau, “Quality assessment of stereoscopic images”, in Proc. EUSIPCO, Sep. 2007, pp. 2110–2114.

[6] J. Starck, J. Kilner, and A. Hilton, “Objective quality assessment in free-viewpoint video production”, in Proc. 3DTV, 2008, pp. 225–228.

[7] A. Tikanmaki, A. Gotchev, A. Smolic, and K. Miller, “Quality assessment of 3D video in rate allocation experiments”, in Proc. IEEE Symp.Consumer Electron., Apr. 2008, pp. 1–4.

[8] P. J. Burt and E. H. Adelson, “A multiresolution spline with application to image mosaics”, ACM Trans. Graph., vol. 2, no. 4, Oct. 1983, pp. 217–236.

[9] Baker, H., Bhatti, N., Tanguay, D., et al, "Understanding Performance in Coliseum, an Immersive Videoconferencing System", ACM Trans. on Multimedia Computing, Communications and Applications (TOMCCAP), 2005.

[10] Baker, P., A. Ogale, C. Fermüller, "The Argus Eye", IEEE Robotics and Automation Magazine, 11:4, 2004.

[11] Bax, M.R., "Real-time lens distortion correction:3D video graphics cards are good for more than games", Stanford Electrical Engineering and Computer Science Research Journal, 2004.

[12] Bolles, R., H.H. Baker, D.H. Marimont, "Epipolar- Plane Image Analysis: An approach to determining structure from motion", Intl. Jrnl. Computer Vision, 1987.

[13] H. Richard and A. Zisserman, "Multiple View Geometry in Computer Vision", Cambridge, U.K: Cambridge Univ. Press, 2000.

[14] C. Hewage, S. Worrall, S. Dogan, and A. Kondoz, "Prediction of stereoscopic video quality using objective quality models of 2-D video", Electron. Lett., vol. 44, no. 16, pp. 963–965, Jul. 2008.

[15] L. Cui and A. R. Allen, "An image quality metric based on corner,edge and symmetry maps", in Proc. British Mach. Vis. Conf., 2008, pp. 1–10.

[16] F. J. Canny, "A computational approach to edge detection", IEEE Trans .Pattern Anal. Mach. Intel., vol. 8, no. 6, Jun. 1986, pp. 679–698.

[17] Yang, J.C., M. Everett, C. Buehler, L. McMillan,"A real-time distributed light field camera", in Proc. 13th Eurographics Workshop on Rendering, Eurographics Association, 2002.