ISSN 2347-4289

Fusion Of Multiexposure And Multifocus Images To Maximize Image Information

Rashmi S, Deepu R

Dept. of computer science & Engineering. Maharaja institute of technology, Mysore, India
Dept. of computer science & Engineering. Maharaja institute of technology, Mysore, India

Abstract: The objective of image fusion is to combine relevant information from multiple images into single image. The proposed technique fuses the multi-exposure images into a high quality image, avoiding the intermediate step of converting it into high dynamic range image. The acquisition pipeline is simplified by skipping the physically based high dynamic range assembly. The advantage includes avoiding the camera response curve calibration and increased computational efficiency. The technique blends multiple exposures using simple quality measures like saturation and contrast. The blending is done in multi resolution fashion to account for better results. Extending this to multi-focus image stack, the goal is to select only those pixels which are in good focus and thus are sharp. An efficient and low complex approach for fusion of multi focus images in the DCT domain based on variance is proposed. A non-reference objective image fusion metrics based on mutual information is used to for evaluation of fusion performance.

Keywords*:* HDR; multi-exposure; multi-focus;

I. Introduction

Image processing techniques primarily focus upon enhancing the quality of an image or a set of images and to derive the maximum information from them. Image Fusion is such a technique of producing a superior quality image from a set of available images. It is the process of combining relevant information from two or more images into a single image wherein the resulting image will be more informative and complete than any of the input images. A lot of research is being done in this field encompassing areas of Computer Vision, Automatic object detection, Image processing, parallel and distributed processing, Robotics and remote sensing. The ratio of highest to the lowest luminance is defined as dynamic range. A very wide range of luminance is often seen in real world scenes, sometime goes beyond 10 orders of magnitude. The conventional digital capture and display devices, which have limited dynamic range of only 2 orders of magnitude, face a problem in reproducing these scenes. To address this problem, the most common solution is to take a sequence of images of the same scene which are low dynamic range in nature but captures all the radiance information under different exposure intervals. There exist two solutions to overcome this problem. The first solution involves estimating the camera response function to recover the true radiance of the original image from the image stack and the radiance map needs to tone mapped to be displayed on LDR reproduction media. This method is computationally expensive and time consuming, even though it provides the satisfying results. To overcome this, a new method called exposure fusion has been proposed. Exposure fusion avoids the intermediate step of creating the radiance map instead produces HDR like images at a lower computational cost. The same technique can be extended to produce all in focus image for multi focus images. Exposure fusion simplifies the acquisition pipeline. Since the proposed technique is not physically based, there is no need to worry about the calibration of camera response curve and exposure time of each photograph need not be monitored. Another method for fusing the multi focus images in DCT domain, which is less complex and energy efficient is proposed. This technique is ideal for use in visual sensor networks which reduces the amount of data in the network transmission and at the same time produces a new image suitable for both visual perception and further processing. Discrete cosine transform (DCT) is an important transform in image processing. The lower frequency region contains large DCT coefficients and hence provides great energy preserving property.

II. Related work

A. Multi exposure image fusion

A high dynamic range image from a set of low dynamic range images which are captured with normal camera is assembled to high dynamic range imaging. In order to linearize the intensities, the camera specific response curve should be recovered. The input sequences are used to compute the calibration step and their exposure settings. The HDR images cannot be directly displayed due to limited dynamic range of most display devices.To fit the dynamic range of display devices, tone mapping is used to compress the dynamic range. With different advantages and disadvantages, many different tone mapping operators have been suggested. A spatially uniform remapping of intensity is applied in global operators to compress the dynamic range. Speed is the main advantage, but it fails in reproducing a pleasing image. A spatially varying remapping, i.e., for different regions in the image, the mapping changes is applied in local tone mapping operators. Pleasing images are produced in this technique but sometimes the results looks unnatural. A wide variety of techniques have been employed by operators to compress the dynamic range: from compression in the gradient domain, to bilateral filtering, in which image is decomposed into edge-aware low and high frequency components. Burt et al. [1] have already proposed to use image fusion in similar fashion. However, this method is more flexible by using simple image measures, such as contrast and saturation. Goshtasby [2] proposed a method that blends

ISSN 2347-4289

the multiple exposures, its drawback lies in dealing with object boundaries. Grundland et al. [3] uses pyramid decomposition to cross dissolve between two images using different quality measures.

B. Multi focus image fusion

Image fusions which are performed on images in spatial domain have been focused by researchers till now. The multi decomposition algorithms are more popular. A multi scale transform is performed on each source image, and then combine all of these decomposition coefficients to produce a representation that is composite. A quantity called activity level measure is monitored to combine the source images. The quality of each source image is determined by activity level. Choosing the coefficients with larger activity level or a weighted average of the coefficients becomes the basis for integration. The inverse multi scale transform reconstructs the final fused image. The Laplacian, gradient, morphological pyramids, discrete wavelet transform and shift variant discrete wavelet transform are the examples of this approach. The drawback of these methods is time consuming, complex and is not suitable for real time applications. In visual sensor networks where the links between the nodes are wireless, the energy consumed for energy for communication is much larger than energy consumed for data processing. Therefore, the compression of data before transmission to other node becomes a major task in most sensor networks. In VSN, compression of images takes place at camera nodes and then sent to the fusion agent, from where the compressed fused images will be saved or transmitted to an upper node. Spatial domain fusion methods requires images to be initially decoded and transferred into spatial domain and after the fusion procedure is applied, the encoding of fused image is required. There is no need of encoding and decoding in DCT domain. The above mentioned issue of complexity reduction has been considered by Tang [4] and has proposed two image fusion techniques in DCT domain, namely DCT+Average and DCT+Contrast. Some undesirable side effects including blurring were seen due to simple method of averaging in DCT+Average method. In DCT+Contrast method, fusion criterion is based on contrast measure which needs to be calculated for every 63AC coefficients of blocks. Calculating the contrast measure for each coefficient makes the algorithm more complex and it also suffers from blocking artifacts.

III. Proposed methods

A. Multi exposure image fusion

Exposure fusion produces the desired image by retaining only the “best “parts in the multi-exposure image sequence. A set of quality measures guides this process, which is combined into a scalar-valued weight map. Input sequence is treated as a stack of images.by collapsing the stack using weighted blending produces the final image. Colorless regions due to under and overexposure could be seen in many images present in stack. Less weightage should be given to such regions, while areas with bright colors and details need to be preserved. Following measures are used to achieve this:

- Contrast: Laplacian filter is applied to the grayscale version of each image, and absolute value of the filter response is considered. Simple indicator C for contrast is produced. Important elements such as edges and texture are assigned high weight. The multi-focus fusion for extended depth-of-field follows similar measure.
- Saturation: The resulting colors become desaturated and eventually clipped as the photograph undergoes a longer exposure. Saturated colors make the image look vivid and are desirable. Taking standard deviation within the R, G and B channel, at each pixel gives the saturation measure S.
- Well-exposedness: The over exposed pixels are clamped to 1 and under exposed pixels are mapped to 0, according to camera response curve. The Guass curve weights each intensity g based on how close it is to 0.5.

$$W = e^{-\frac{(i-0.5)^2}{2\sigma^2}}, i \in [0,1]$$

Map combination The algorithm combines the information from different measures into scalar weighting map using multiplication for each pixel and power function controls the influence of each measure:

$$W_k = C_k(i,j)^{w_c} * S_k(i,j)^{w_s} * E_k(i,j)^{w_E}$$

Where W_k, S_k, E_k, C_k refer to the final weighting map, saturation map, well-exposedness map and contrast map respectively. P_S, P_E, P_C are the corresponding weighting parameters to control how much each measure contributes to the final map. The corresponding measure is 1 when any parameter is equal to 0 in the multiplication and thus will not make any difference. Naïve Fusion The next step is to fuse the multiple images after obtaining weighting maps. A weighted average across the N images at each pixel (i, j) is computed. The drawback of this approach is that transitions between the pixels are not smooth which makes the blurs the final result. The reason, the naïve averaging the image set could not provide seamless blending is due to varying weights. Pyramid Fusion To overcome the seam problem, a multi-resolution fusion technique is adopted .The image is transformed into pyramid representations and fusion is conducted on each level, and the final image is reconstructed from fused pyramid. The algorithm conducts a pyramid fusion to seamlessly blend multiple images. The weighting map for each image in the image stack is computed as described. The Laplacian pyramids for each input image and Gaussian pyramids for each weighting map are constructed. The Gaussian pyramid of weighting maps and Laplacian pyramid of input images on each level are multiplied and the results are summed together across the N images. The resultant is a new Laplacian pyramid. The fused Laplacian pyramid can used to reconstruct the final image. The pyramid fusion technique blends the image features instead of intensities, which solves the seam problem and produces the desired final image

B. Multi focus image fusion

When the limited depth of focus on a selected focus setting of a camera results in parts of an image being out of focus, multi focus image fusion is used. The solution is to fuse the

ISSN 2347-4289

number of multi focus images to create an all in focus image.DCT block analysis The given image is partitioned into nonoverlapping 8*8 blocks. The coding process begins with level shifting since 8 bit pixels of an image are in the range of 0:255.the input data is shifted to the range of 128:127 in order to reduce the amount of DC coefficients, so that it becomes distributed about zero. On each 8*8 block two dimensional DCT is performed. The step of quantization includes quantizing the large number of small coefficients to zeros making it a lossy compression. In order to have the high frequency coefficients with zero value to be grouped together at the end of rearranged array, quantized coefficients are rearranged in zigzag order. To encode the coefficients run-level procedure and Huffman encoders are used. For decoding, each block of data is decoded and then dequantized using quantized table. Using inverse DCT, the reconstructed coefficients are transformed back into image block. Variance as contrast measure in DCT domain

Two dimensional DCT transform of image g(x, y) is defined as:

$$d(i,j)=\frac{2\alpha(i)\alpha(j)}{N}\times\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}g(x,y)\times\cos\left[\frac{(2x+1)\pi i}{2N}\right]\times\cos\left[\frac{(2y+1)\pi i}{2N}\right]$$

Where i,j=0,1,2..N-1 and α(k)= 0.7071, if k=0 and 1otherwise.The inverse DCT can be calculated suing very similar formula. Mean value and variance of N*N block in the spatial domain is calculated as:

$$\mu=\frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}g(x,y)$$

$$\sigma^2=\frac{1}{N^2}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}g^2(x,y)-\mu^2$$

Using mathematical calculation, we can arrive at the variance of the block from DCT coefficients.

$$\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}g^2(x,y)=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}g(x,y)\ g(x,y)$$

$$=\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}g(x,y)\times\frac{2\alpha(i)\alpha(j)}{N}\times\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}d(i,j)\times\cos\left[\frac{(2x+1)\pi i}{2N}\right]\times\cos\left[\frac{(2y+1)\pi i}{2N}\right]$$

By interchanging the positions, we will have variance of an N*N block of pixels calculated from its DCT coefficients by sum of squared normalized AC coefficients of DCT block. In multi focus images, the more informative area is the focused area. In image processing applications, the variance value is assumed as contrast measure. Since the variance in the DCT domain can be easily calculated, this method best suffice the problem.

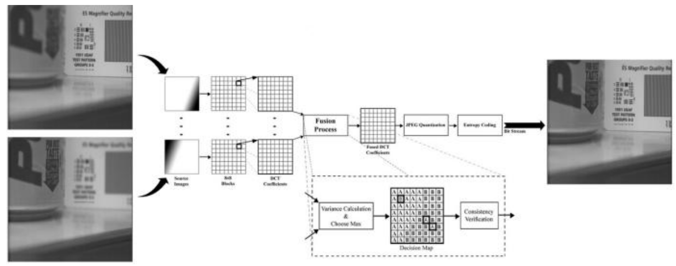

Fig.1 Proposed method

Consistency Verification (CV) Consistency verification that utilizes a majority filter is introduced by LI et al. The source image B contributes the center block and source image A contributes the majority of surrounding block, then the center sample is simply switched to the corresponding block in source image A. this method produces more efficient image both in quality and complexity.

IV. Results

Multi exposure fusion result

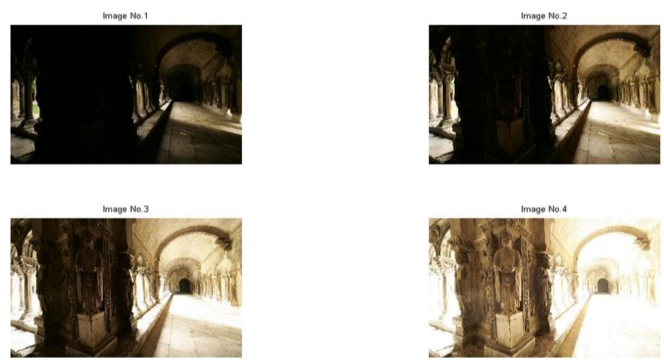

Fig.2 Input image Stack

Fig.3 Resultant image from Naive fusion.

ISSN 2347-4289

Fig.4 Resultant image from Pyramid fusion.

The proposed unoptimized software implementation performs the exposure fusion in seconds. The proposed methods nables user have more control over the fusion process, as weighting of quality measures can be adjusted by user.Multi focus fusion result Mutual information gives the amount of information contained in fused image. A new quality metrics is used which is non-reference method. The feature information F contained in input images A and B are individually calculated and combined to get feature mutual information [5].

$$I_{FA} = \sum_{f,a} P_{FA}\ (x,y,z,w) log_2 \frac{P_{FA}(x,y,z,w)}{P_F(x,y).P_A(z,w)}$$
$$I_{FB} = \sum_{f,b} P_{FB}\ (x,y,z,w) log_2 \frac{P_{FB}(x,y,z,w)}{P_F(x,y).P_A(z,w)}$$

$$FMI_F^{AB} = I_{FA} + I_{FB}$$

Table 1 FMI comparison of image set

Dataset	Pepsi	airplane	clock
DCT+variance	0.9144	0.9467	0.9142
DCT+variance+CV	0.9246	0.9494	0.9152

Conclusion The technique blends images in a multi-exposure sequence, guided by simple quality measures like saturation and contrast. This is done in a multi-resolution fashion to account for the brightness variation in the sequence. Quality is comparable to existing tone mapping operators. The approach is controlled by only a few intuitive parameters, which can be updated at near-interactive rates in the unoptimized implementation. A new DCT based fusion technique for multi-focus images is implemented. The method is based on the definition of variance in DCT domain. Simplicity of the proposed method makes it appropriate for real-time applications. Furthermore, utilization of variance in the proposed algorithm, as a proper contrast measure in multi-focus images, leads to better quality of the fused image. The proposed method out performs both in terms of quality and complexity reduction.

References

[1] P. J. Burt, K. Hanna, and R. J. Kolczynski. Enhanced image capture through fusion. In Proceedings of the Workshop on Augmented Visual Display Research, pages 207–224. NASA– Ames Research Center., Dec. 1993.

[2] A. Goshtasby. Fusion of multi-exposure images. Image and Vision Computing, 23:611–618, 2005.

[3] M. Grundland, R. Vohra, G. P. Williams, and N. A. Dodgson. Cross dissolve without cross fade: Preserving contrast, color and salience in image compositing. Computer Graph-ics Forum, 25(3):577–586, Sept. 2006.

[4] TANG, A contrast based image fusion technique in the DCT domain, Digital Signal Processing 14(3), 2004, pp. 218–226.R

[5] Mohammad Bagher Akbari Haghighat, Ali Aghagolzadeh, Hadi Seyedarabi, A non-reference mage fusion metric based on mutual information of image features, Computers and Electrical Engineering 37(5), 2011, pp. 744–756.Y.