Neuro-Fuzzy Based Enhanced Colour Video Fusion

S.Batmavady1, Arun2, and Sharanjith3

Pondicherry Engineering College, Department of ECE, Puducherry, India

Email: 1 s_batmavady@yahoo.co.in, 2 arun91@pec.edu, 3sharan_07@pec.edu

Abstract - Infrared sensors and visible light sensors are complimentary on reflecting human and surrounding information. In order to improve the performance of surveillance systems, the infrared and visible light videos can be fused. In the applications like surveillance, the fusion algorithm should be on a less complex platform and should take lesser time for processing. Colour representation always provides more visual information. In this paper, colour video fusion based on neuro-fuzzy logic is presented. Visible images/videos taken under poor lighting conditions need to be enhanced before fusing. Hence contrast limited adaptive histogram equalization (CLAHE) based enhancement of visible light images is discussed. The neuro-fuzzy logic approach is implemented for colour image and video fusion and is compared with other image fusion algorithms such as DWT and spatial averaging through some subjective and objective image fusion performance measures. Experimental results demonstrate that the neuro-fuzzy logic approach is effective in real-time colour image/video fusion applications.

Index Terms - contrast limited adaptive histogram equalization; entropy; image fusion; neuro-fuzzy logic; video fusion.

I. INTRODUCTION

Image fusion deals with integrating data obtained from different sources of information for intelligent systems. It provides output as a single image from a set of input images. The purpose is to extract and synthesize information from multiple images in order to produce a more accurate, complete and reliable composite image of the same scene so that it is more suitable for human or machine interpretation. It plays a vital role in the multi-sensor environments like surveillance systems where infrared (IR) and visible light sensors are utilized. To meet real time surveillance - demands, video representation is necessary which gives a complete spatial and temporal visual information as compared to only spatial visual information presented by still images. The video fusion is the much required work in the field of surveillance.

Infrared cameras sense radiation emitted by an object in the infrared spectrum which is not available in a visual image, but lose information, such as texture, color and geometric, which is available in normal visual camera. Infrared image generally allows a better contrast to be obtained between a person and environment. Hence fusion of IR and visual images can enhance features in both kinds of images. Application of image fusion in concealed weapon detection (CWD) is gaining significance in the general area of law enforcement. In visible image, the facial information of the person and surrounding are visible, while the hidden weapon is clearly visible in the IR image, as there is a temperature difference between weapon and the body. Thus the intelligent fusion of the information provided by both visible and infrared will increase the performance of surveillance system and concealed weapon detection.

Because of the importance of image fusion techniques, many image fusion algorithms have been proposed. Traditional image fusion methods contain the intensity-hue-saturation (IHS) transform principal component analysis (PCA), high pass filtering (HPF) [1], contrast pyramid based image fusion [2] etc. However, these methods often lose spectral information or high frequency information in the process and lead to significant color distortion in the fused image. The wavelet Transform (WT), due to its superior qualities in both spatial and spectral domains, has become the preferable algorithm for image fusion [3]. Wavelet transform has the advantage that it could easily be realized in hardware. Since the orthogonal filters of wavelet transform do not possess linear phase, biorthogonal wavelets were used for image fusion application in [14]. But real-time applications like surveillance requires lesser image data processing and faster computing which requires fast computing platform as well as less complex algorithms. Hence DWT is not the wise choice. Fuzzy logic approach (FLA) is a simple and yet efficient image fusion scheme, which is based on simple rules, which are easy to apply and takes lesser processing time. This approach forms an alternative to a large number of conventional approaches, which are based on a host of empirical relations. FLA has been effectively implemented for multi-sensor gray scale image fusion applications [4]. A strong point of FLA is that it permits the encoding of expert knowledge directly and a weak point is that it is difficult to design and tune the membership function that quantitatively defines the image sets. It has been found that artificial neural network learning techniques can automate this process and substantially reduce development time while improving performances. Hence, the combination of neural network and fuzzy logic, the neuro-fuzzy approach can be used for image fusion. Image fusion applications are highlighted in [5-7].

In this paper, pixel level fusion of enhanced colour visible light images and IR images and the extension of the algorithm for video fusion are discussed. The paper is organized as follows. In Section II, the enhancing of visible images taken under poor lighting conditions using Contrast Limited adaptive Histogram Equalization (CLAHE) algorithm is illustrated. In Section III, the Neuro-Fuzzy approach for color image fusion and video fusion is discussed. Experimental results are presented in Section IV and Section V the performance of this algorithm is compared with some traditional algorithms by certain objective performance.
III. Neuro-Fuzzy Based Color Image And Video Fusion

The main drawback of fuzzy logic approach in image fusion is the difficulty in design of fuzzy inference system (FIS). The determination of percentage of overlapping among fuzzy sets, appropriate membership function tuning and if-then rule definition are to be carried out by trial and error method only and needs a lot of time in design stage. A neuro-fuzzy network is a fuzzy inference system (FIS) in the body of an artificial neural network. In neuro-fuzzy logic neural networks are used for determination of all these fuzzy logic parameters and can be trained by the input from the sensors to tune the membership function parameters. In this paper, Adaptive Neuro-Fuzzy Inference System (ANFIS) available in the Matlab fuzzy logic toolbox is used to carry out neuro-fuzzy based fusion.

A. ANFIS structure

Adaptive Neuro Fuzzy Inference System (ANFIS) is a kind of simple feed forward neural network that is based on Tagaki-Sugeno fuzzy inference system. It applies a combination of the least-squares method and the backpropagation gradient descent method for training FIS membership function parameters to emulate a given training data set [10]. The subtractive clustering method is used to partition the data into groups called clusters, and generates a FIS with the minimum number rules required to distinguish the fuzzy qualities associated with each of the clusters.

The typical structure of ANFIS (two inputs, one output and two membership function) can be divided into three layers, as shown in Figure 2. $x(k)$ and $y(k)$ are the input pixel gray level value of IR and visible light images. The first layer will check the degree of membership of $x(k)$ and $y(k)$ and maps it to fuzzy sets $\Omega_{11}$ or $\Omega_{12}$ and $\Omega_{21}$ or $\Omega_{22}$ respectively. Second layer defines fuzzy if then rules and determines its relevance grade of each rule $w_i$ (four rules $i=1, 2, ..., 4$) which modified after training and learning [7]. The learning of fuzzy inference system is the adjustment about the antecedent parameters and consequent parameters. In this paper, use hybrid algorithm to estimate the antecedent parameters, and the adjustments of consequent parameters utilize the linear Least Mean Square algorithm.

II. Enhancing Of Visible Light Images Using CLAHE

In the case of night time surveillance application the visible images taken are under poor lighting conditions and hence need to be enhanced. The enhancement of these images are done by contrast limited adaptive histogram equalization (CLAHE) algorithm [8]. Adaptive histogram equalization (AHE) is an image enhancement technique used to improve contrast in images. It differs from ordinary histogram equalization in the respect that the adaptive method computes several histograms, each corresponding to a distinct section of the image i.e., it operates on small data regions rather than the entire image and uses them to improve the total contrast of the image. It is therefore suitable for improving the local contrast of an image and bringing out more detail. However, AHE has a tendency to over amplify noise in relatively homogeneous regions of an image. Contrast limited adaptive histogram equalization (CLAHE) prevents this by limiting the amplification. The contrast, especially in homogeneous areas, can be limited in order to avoid amplifying the noise which might be present in the image[9].

A. CLAHE algorithm [8]

1) The original image should be divided into contextual regions(tiles) which are continuous and non-overlapping. The size of each contextual region is $M \times N$.
2) The histograms of the contextual regions are calculated.
3) The histograms are clipped depending on the ‘Clip limit’ value.

‘Clip Limit’ is a contrast factor that prevents over-saturation of the image specifically in homogeneous areas and part of the histogram that exceeds the clip limit is redistributed equally among all gray levels.

4) Each tile’s contrast is enhanced by equalizing the clipped histogram of the tile.

Similar to uniform distribution the ‘Exponential distribution’ is used as the basis for creating the contrast transform function.

5) The neighbouring tiles are then combined using bilinear interpolation in order to eliminate artificially induced boundaries.

6) The algorithm is applied to individual color components (RGB) and the resultant enhanced components are combined to get enhanced color image.

© 2012 ACEEE
DOI: 01.IJR TET.7.1.27
The third layer calculates the total output of ANFIS based on fuzzy if-then rules.

\[ z_i = p_i x(k) + q_i y(k) + r_i, \quad i = 1, \ldots, 4 \]  

where \( p_i, q_i, r_i \) are the consequent parameters. The output fused pixel value corresponding to the input pixel values \( x(k) \) and \( y(k) \) is calculated as the weighted average of each \( z_i \).

\[ z(k) = \sum_{i=1}^{4} w_i z_i / \sum_{i=1}^{4} w_i \]  

**B. Image fusion algorithm**

Adaptive Neuro-Fuzzy Inference System (ANFIS) in the Matlab fuzzy logic toolbox is used to carry out the following neuro-fuzzy based fusion algorithm.

1) The first image (visible light image) is read in variable \( I_1 \) and its size \((r_1 * c_1 * 3)\) is found out.
2) The second image (IR image) is read in variable \( I_2 \) with same size of \( I_1 \), as both are registered pairs.
3) Variables \( I_1 \) is colour images which is in three dimensions and hence the Red-Blue-Green (RGB) color components of size \((r_1 * c_1)\) are separated.
4) Individual color component of visible light image and grayscale converted IR image are in a 2D matrix form where each pixel value is in the range from 0-255 using gray colormap.
5) The images are converted in column form which has \( C = r_i * c_1 \) entries. This is done to easily pass each corresponding pixel value of \( I_1 \) and \( I_2 \) to fuzzy inference.
6) A training data is formed, which is a matrix with three columns and entries in each column are from 0 to 255 in steps of 1.
7) A check data, which is a matrix of pixels of two input images in column format is formed.
8) The number and type of membership function for both input images are decided. Here three membership functions are used as a compromise between quality and processing time. If the number of membership functions is increased, it may provide more quality fused output but time required for fusion is higher.

All the type of membership functions like ‘Gaussian’, ‘Trapezoidal’, ‘Triangular’ etc. provide quite similar results but as the computation complexity for triangular membership function[7] is less, the fusing time is also less. Hence the membership function selected is triangular.

1) For training, FIS structure is generated by \texttt{genfis1} command with training data, number and type of membership functions as input.
2) To start training, \texttt{anfis} command is used which inputs generated FIS structure and training data and returns trained data.
3) For num = 1 to \( C \) in steps of 1, fuzzification is applied using generated FIS structure with check data and trained data as input which returns output image in column format.
4) The column form is converted to matrix form and the fusion for other components is also repeated.
5) The fused outputs of all components are combined to produce the fused color output.

**C. Enhanced color video fusion algorithm**

The video fusion is just an extension of previous image fusion algorithm. The frames are extracted from sensor video inputs and are fused. But fusing of all corresponding IR and visible image frames needs more time for processing. Hence only alternative frames are fused and remaining frames are dropped. The fused frames are replicated to achieve the same frame rate and viewer cannot find much difference as it is beyond their visual perception.

![Figure 3 video fusion block diagram](image)

The algorithm is as follows:

1) Visible and IR videos are taken as inputs.
2) Video frames are extracted from each video.
3) The visible light image is enhanced, if required.
4) IR and visible frames are then fused alternatively.
5) Each fused frame output is replicated.
6) The fused frames are combined to obtain fused video output.

**IV. EXPERIMENTAL RESULTS**

Two sets of images (a) taken under poor lighting conditions are used as shown in Figure 4. First image has a dark background and second has a dark foreground. These images are enhanced using CLAHE algorithm and shown in (b). The CLAHE algorithm is found to be effective for both the cases.
B. Neuro-fuzzy based colour fusion

In order to illustrate the efficiency of the proposed image fusion method for CWD application, a set of images are used in the experimental tests and the corresponding fusion results are shown in Figures 5. These test images are selected to test the fusion algorithm under various different conditions including different positions for the person, different clothing, different positions of the weapon, and different shapes of the weapon. In all these cases, the visual and IR images considered are registered images. The following images for CWD application are taken from [11].

EXAMPLE 1:

EXAMPLE 2:

EXAMPLE 3:

EXAMPLE 4:

C. Video fusion

The IR and visible light videos are downloaded from [12]. The fusions of four uniformly spaced frames are shown in Figure (7):
The fused output can be qualitatively analyzed by measuring the attributes like entropy and processing time. Entropy of an image is a measure of the average information content [7]. More the entropy of an image, more is the information content included in the image. It is defined as

\[ EN = -\sum_{i=1}^{m} p_i \log p_i \]  

(3)

where \( p_i \) is the probability of gray level \( i \) and the range of \( i \) is \([0, m]\).

The cross entropy reflects the pixel difference between two images and hence it can be used to evaluate the fused image. If \( p_i \) and \( q_i \) represent the probability of \( i^{th} \) gray level of images \( I_1 \) and \( I_2 \) respectively, then the cross entropy between images \( I_1 \) and \( I_2 \) namely \( CEN \) is defined as:

\[ CEN(I_1, I_2) = \sum_{i=1}^{m} p_i \log(p_i / q_i) \]  

(4)

If \( CEN(I_1, I_2) \) and \( CEN(I_1, I_1) \) denote the cross entropy between the original image and the fused image, the composite cross entropy can be calculated as:

\[ CEN(I, I_1, I_2) = \frac{CEN(I, I_1) + CEN(I, I_2)}{2} \]  

(5)

Generally, if the entropy is larger and the cross entropy is less, the fusion algorithm is better [7]. But for more accurate analysis the integrative entropy also used to evaluate performance. For determination of integrative entropy the absolute cross entropy should be calculated. The absolute cross entropy between image \( I_1 \) and \( I_2 \) is defined as:

\[ ACEN(I_1, I_2) = \sum_{i=1}^{m} |p_i \log(p_i / q_i)| \]  

(6)

The average cross entropy between three images \( (I, I_1, I_2) \) can be calculated as:

\[ ACEN(I, I_1, I_2) = (ACEN(I, I_1) + ACEN(I, I_2))/2 \]  

(7)

The integrative entropy is defined as:

\[ IEN(I, I_1, I_2) = EN(I) - ACEN(I, I_1, I_2) \]  

(8)
For the fused images, if the entropy is larger and the average cross entropy is less, the fusion algorithm is superior; otherwise, it is poor. Similarly, the larger the integrative entropy, the more effective is the fusion algorithm [13]. The performance of fuzzy based fusion algorithm is compared and analysed with other fusion algorithms like spatial averaging and Discrete Wavelet Transform (DWT) IR image pairs. The results are tabulated in Table I.

<table>
<thead>
<tr>
<th>fusion algorithm</th>
<th>fusion time (seconds)</th>
<th>entropy</th>
<th>composite cross entropy</th>
<th>integrative entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial averaging</td>
<td>0.080990</td>
<td>5.032</td>
<td>3.231827</td>
<td>1.6228</td>
</tr>
<tr>
<td>Fuzzy based</td>
<td>0.500362</td>
<td>6.605</td>
<td>1.194599</td>
<td>5.1142</td>
</tr>
<tr>
<td>DWT</td>
<td>7.217778</td>
<td>6.585</td>
<td>1.910851</td>
<td>4.2368</td>
</tr>
</tbody>
</table>

Integrative entropy is the measurement of effective information present in the fused image. Spatial averaging fusion method is the simplest technique and hence it has minimum integrative entropy and the least processing time. DWT based fusion has the maximum processing time which is more than the fuzzy based fusion. But in CWD and surveillance systems, processing time is an integral parameter and it should be minimal. It is inferred from the results that the fuzzy based fusion gives an appreciable performance without compromising the information content.

VI. Conclusion

In this paper, the Neuro-fuzzy logic is applied for image fusion in which the artificial neural network learning techniques are used for determination of fuzzy parameters. It aids in reducing the time taken for framing efficient rules while improving the performance. The Neuro-fuzzy logic approach is implemented for colour image fusion and colour video fusion. The Neuro-fuzzy logic approach is compared with other image fusion algorithms such as DWT and spatial averaging through some subjective and objective image fusion performance measures. Experimental results show that the Neuro-fuzzy logic approach is effective in real-time colour image/video fusion applications as it provides good quality fused image and the time required for fusion is also less. The DWT based image fusion requires higher fusion time. Image fusion based on special averaging technique is the simplest fusion algorithm and hence has the least fusion time but the resultant fused image is of poor quality. Hence Neuro-fuzzy logic approach is found to be optimal and effective for colour image/video fusion. It is hoped that this algorithm can be modified and implemented on a parallel computing platform so that the fusion time can be further reduced. The algorithms for object tracking can be combined with video fusion algorithm for better machine interpretation and to improve performance of surveillance system.

REFERENCES