Pixel Based Fusion Methods for Concealed Weapon Detection

Dr. K. Indira¹, Manjunath C Lakkanavar², Reshma Verma³, Nagabhushan M⁴
¹Professor, MSRIT, ²³Asst. Professor, MSRIT, ⁴Associate Professor, MSRIT

ABSTRACT
Concealed Weapon Detection (CWD) is the detection of weapons underneath a person’s clothing which is an important obstacle for the security of general public as well as safety of public assets like airports and buildings. Concealed weapons such as handbags, knives and explosives are detected using manual screening procedures. It is desirable to detect the concealed weapons from a far off distance at airports and other secured places. A number of sensors with different phenomenology have been developed to observe objects underneath’s persons clothing. As no single technology provide improved performance in CWD applications, different image fusion schemes based on pixel level is proposed. Image obtained from visual camera does not reveal any information hidden under persons clothing whereas MWM image obtained from MWM (Millimeter Wave Imaging) sensor reveals clothing penetration underneath persons cloth but cannot identify the person. In this paper fusion of MWM image with visible image based on pixels is proposed. Experimental results reveal that fused image can identify the person with concealed weapons. Performance metrics such as standard deviation, entropy and cross entropy is calculated and from simulation results it is observed that PCA based fusion method is similar to DWT based fusion scheme.

Index Terms: MWM imaging, DWT, Principal Component Analysis, Standard Deviation, Entropy

I. INTRODUCTION
A weapon is any object that can do harm to another individual or group of individuals. This definition not only includes objects typically thought of as weapons, such as knives and firearms, but also explosives, chemicals, etc. so this harmful things need to be detected for the safety of general public as well as security of public assets like airports and buildings etc. The manual screening procedure which is being adopted sometimes gives wrong alarm indication, and fails when the object is not in the range of security personnel as well as when it is impossible to manage the flow of people through a controlled procedure. It also fails when identification of a person is essential, who is the victim of an accident in future. In recent times there are series of bomb blasts in Mumbai, Delhi, and Guhahati and bombs went off in buses and underground stations which killed many and left many injured and left the world in shell shock and the Indians in terror. This situation is not limited to India but it can happen or already happened anywhere and anytime in the world. People think bomb blasts can’t be predicted before handled. In all of these cases detecting concealed weapons is especially difficult when one wants to monitor an area where portable systems are not practicable. As no single sensor technology can provide acceptable performance, there is the requirement of multisensor fusion. Multisensor fusion is a technology to combine information from multiple sensors and sources to achieve improved accuracies and more specific inferences. As many sensors produce images, the term image fusion is used which generates a single image and contains accurate description than individual source images. The sensors used for image fusion must be accurately coaligned so that their images will be in spatial registration. A simple fusion technique is to take the average of source images pixel by pixel. In this paper fusion of visual and MWM images are considered. Visual image provides the outline and appearance of the people while MWM imaging technique detects concealed objects such as plastics, metal weapons, explosives, drugs etc. To identify the person with concealed weapons, fusion of Visual and MWM images is necessary. The main goal of this paper is to study different fusion techniques and evaluate the performance of concealed weapon detection system. The organization of the paper is as follows:

Literature survey is discussed in Section II. Section III deals with different fusion techniques. Qualitative metrics for evaluating the performance of concealed weapon detection system is discussed in section IV. Section V deals with simulation results and finally conclusion and future work is presented in Section VI.

II. LITERATURE SURVEY AND REVIEW
A numbers of technologies are being developed for Concealed Weapon Detection (CWD). Tuzhi Xu and Q M Jonathan Wu [1] have developed image Fusion algorithm using double density dual tree complex wavelet transform for concealed weapon detection. They have proved that
their fusion result is better than other fusion algorithms. Pooja Pratihar and Arun Kumar Yadav [2] have discussed detection of concealed weapon using fusion methods for IR and visual images. J. Yang and R S Blum[3] has presented fusion of visual and non-visual images using Expectation – Maximum (EM) algorithm for CWD applications. They have proved that EM algorithm is better that pixel averaging, selection of maximum pixel and Laplacian Pyramid fusion algorithms. Zhiyun Xue and Rick. S.Blum [4] developed color image fusion algorithm to fuse color visual image and the corresponding IR image for CWD applications. The algorithm maintained high resolution of the visual image, incorporate any concealed weapon in the IR image and keep the natural color of the visual image. Karol etat[5] have presented a miniature prototype device for CWD using IR and visual cameras. Such miniature device could be mounted under the ceiling. Thomas Metziller , e tal [6] have demonstrated fusion of visual and Infrared sensory image using wavelet based fusion and Fuzzy Logic Approach fusion. Samir K Bandopadhyya e tal [7] have developed new algorithm using color visual image and a corresponding IR image for CWD application by the help of fusion technology. Hua Mei Chen e tal [8] have briefly reviewed sensor technologies for CWD applications. Among the various methods, passive Millimeter Wave Imaging(MWM) sensors offer the best near time potential for providing a non invasive method of observing metallic and plastic objects concealed underneath persons clothing. MWM cameras alone provide useful information about the detail and location of the individual being monitored. To enhance the practical values of passive MWM sensors, sensor fusion approaches using MWM and IR are investigated. D. Dinkar Rao Dongre [9] have proposed DWT based algorithm using visual image and MWM image. Zhong Zhang and Blum[10,11] have demonstrated fusion of visual and 94 GHz millimeter wave images using multiscale transform and region based fusion scheme for CWD applications. David Sheen e tal[12,13] have developed a novel cylindrical millimeter wave imaging technique and wide bandwidth three dimensional holographic microwave imaging technique for detection of metallic and non-metallic concealed weapons. Adel Slamanl[14] e tal have discussed image enhancement for reliable detection and identification of weapons concealed under varying layers of clothing using Automated Statistical Characterization and Partitioning of Environments (A’SCAPE). From literature review, it is clear that a single source image either visual or IR image when presented to a human operator, is difficult to recognize a weapon concealed underneath a persons clothing. Hence fusion is an important tool to detect and recognize the person with a weapon underneath his clothing. This paper discusses different pixel level image fusion technique.

III. PIXEL LEVEL IMAGE FUSION METHODS

In pixel level fusion process, a composite image is formed from several input images based on their respective pixels(picture elements). The data fusion process should carry all the useful and relevant information from input images to the composite image, should not introduce any additional inconsistencies and should be shift and rotational invariant. The fused image should have improved contrast and it should be easy for the user to detect, recognize and identify the targets. The simplest pixel level image fusion method is to take the average of the gray level images pixel by pixel, but this may produce undesired effects and reduced feature contrast. Different Pixel level image fusion methods are:

A. Principle Component Analysis Method

PCA[15] is a numerical procedure that transforms a number of correlated variables into a number of uncorrelated variables called principal components(PCs). The first PC accounts for much of the variance in the data and each succeeding component accounts for much of the remaining variance; the first PC is along the direction with maximum variance, the second component lie in the subspace perpendicular to the first component and this component points to the direction of maximum variance.

Steps for calculating Principle component coefficients:

i) The source images to be fused are arranged in two column vectors which yields a matrix Z of n x 2 dimensions.

ii) The empirical mean for each column is calculated and the mean vector has dimensions 2 x 1.

iii) The mean vector from each column of the data matrix Z is subtracted which yields a matrix X of dimension n x 2.

iv) The covariance C of the matrix Z is computed C==X'X.

v) The eigen vectors V of C and the corresponding eigen values is computed.

vi) Eigen vector V is sorted in decreasing order. The diagonal matrix D of eigen values is of dimension 2 x 2.

vii) The weights NPC$_1$ and NPC$_2$ such that (NPC$_1$ + NPC$_2$=1) in Equation (1) is computed using first column of V which corresponds to the largest eigen value.
\[ \text{NPC}_1 = \frac{V(1)}{\sum \text{V}} \quad \text{and} \quad \text{NPC}_2 = \frac{V(2)}{\sum \text{V}} \quad (1) \]

viii) Finally the fused Image is obtained defined in Equation (2)
\[ I_f = \text{NPC}_1 I_1 + \text{NPC}_2 I_2 \quad (2) \]
where \( I_1 \) is a visible image and \( I_2 \) is the MWM image.

B. Block Based image fusion using PCA

In block based image fusion, the input images are decomposed into blocks(I_{1k} and I_{2k}) of size \( m \times n \) where \( I_{1k} \) and \( I_{2k} \) are the \( k \)th block of \( I_1 \) and \( I_2 \). NPC_{1k} and NPC_{2k} are the principal components corresponding to the \( k \)th blocks.

The fusion of \( k \)th block of the image is defined in Equation (3)
\[ I_{fk} = \begin{cases} I_{1k} \quad \text{NPC}_{1k} > \text{NPC}_{2k} + \text{th} \\ I_{2k} \quad \text{NPC}_{1k} < \text{NPC}_{2k} - \text{th} \\ \frac{I_{1k} + I_{2k}}{2} \quad \text{others} \end{cases} \quad (3) \]
where \( \text{th} \) is the user defined threshold and \( (I_{1k}, I_{2k})/2 \) is the gray level averaging of corresponding pixels.

C. Spatial Frequency

Spatial Frequency(SF)[15] measures the overall information present in an image. For an image \( I \) of dimension \( M \times N \), row frequency and column frequency are defined in Equation (4) and Equation(5).
\[ RF = \sqrt{\frac{1}{MN} \sum_{j=0}^{M-1} \sum_{i=0}^{N-1} [I(i, j) - I(i, j-1)]^2} \quad (4) \]
\[ CF = \sqrt{\frac{1}{MN} \sum_{j=0}^{M-1} \sum_{i=0}^{N-1} [I(i, j) - I(i-1, j)]^2} \quad (5) \]
Spatial frequency is defined by Equation(6)
\[ SF = \sqrt{RF^2 + CF^2} \quad (6) \]
The spatial frequency based weighted image fusion is defined in Equation (7)
\[ I_f = \text{NSF}_1 + I_1 + \text{NSF}_2 * I_2 \quad (7) \]
Where the weights NSF_1 and NSF_2 are defined in Equation(8)
\[ \text{NSF}_1 = \frac{SF_1}{SF_1 + SF_2} \quad \text{NSF}_2 = \frac{SF_2}{SF_1 + SF_2} \quad (8) \]

D. Block based image fusion using SF

The images are decomposed into blocks \( I_{1k} \) and \( I_{2k} \), the normalized SFs for each block are computed. NSF_{1k} and NSF_{2k} are the normalized SFs of \( I_{1k} \) and \( I_{2k} \), the \( k \)th block of the fused image[15] is defined in Equation(9)
\[ I_{fk} = \begin{cases} I_{1k} \quad \text{NSF}_{1k} > \text{NSF}_{2k} + \text{th} \\ I_{2k} \quad \text{NSF}_{1k} < \text{NSF}_{2k} + \text{th} \\ \frac{I_{1k} + I_{2k}}{2} \quad \text{otherwise} \end{cases} \quad (9) \]

E. Discrete Wavelet Transform based image fusion

Wavelet theory is an extension of Fourier Theory and is an alternative to the Short Time Fourier Transform(STFT). In Fourier analysis the signal is decomposed into sine waves of different frequencies whereas in wavelet analysis, the signal is decomposed into scaled(dilated or expanded) and shifted(translated) version of mother wavelet or function. A wavelet is a small wave and satisfies two basic properties:
(i) Time integral must be zero
\[ \int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (10) \]
(ii) Square of wavelet integrated over time is unity
\[ \int_{-\infty}^{\infty} \psi^2(t) dt = 1 \quad (11) \]
Wavelet transform for 1D signal \( f(x) \) onto a basis of wavelet functions is defined in Equation(12)
\[ W_{ab} (f(x)) = \int_{x=\infty}^{x=\infty} f(x) \psi_{a,b}(x) dx \quad (12) \]
Basis is obtained by translation and dilation of mother wavelet as defined in Equation(13)
\[ \psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (13) \]
The mother wavelet would localize in both spatial and frequency domain and has to satisfy zero mean constraint. In Discrete Wavelet Transform(DWT), the dilation factor \( a = 2^n \) and translation factor is \( b = n2^m \), where \( m \) and \( n \) are integers. The information flow for one level of 2-D image [16,17] decomposition is illustrated in Fig1. Wavelet, separately filters and down samples the 2-D (image) data in the vertical and horizontal directions(separable filter bank). The input (source) image is \( I(x,y) \) is filtered by Low Pass Filter L and High pass filter H in horizontal direction and then down sampled by a factor of 2 to create the coefficient matrices \( I_L(x,y) \) and \( I_H(x,y) \). The coefficient matrices \( I_L(x,y) \) and \( I_H(x,y) \) are both low pass and high pass filtered in vertical directions and down sampled by a factor of 2 to create subbands[16]: \( I_{L1}(x,y), I_{L2}(x,y), I_{H1}(x,y), \) and \( I_{H2}(x,y) \).
\( I_{L1}(x,y) \) contains the average image information corresponding to low frequency band of multiscale decomposition. It is a smoothed and
subsampled version of the source image \( I(x,y) \), \( I_{1lh}(x,y) \), \( I_{1hh}(x,y) \) and \( I_{10}(x,y) \) are detailed subimages containing horizontal, vertical and diagonal directions. Multiresolution is achieved by applying the same algorithm to the lowpass coefficients from previous decomposition.

Inverse 2-D wavelet transform is used to reconstruct the image \( I(x,y) \) from subimages \( I_{lh}(x,y) \), \( I_{hh}(x,y) \) and \( I_{0}(x,y) \) using column upsampling and filtering using Lowpass and Highpass filters at each subimage and Row sampling and filtering using Lowpass and Highpass filters of the resulting image and summation of all matrices to yield \( I(x,y) \)

\[
I(x,y) = \text{IDWT} \{ \phi \{ \text{DWT} \{ I(x,y), \text{DWT} \{ I_1(x,y) \} \} \} \}
\] (14)

The flow diagram for wavelet based image fusion is shown in Fig.2. In wavelet based fusion scheme, the source images \( I_1(x,y) \) corresponding to visual image and \( I_2(x,y) \) corresponding to MWM image are decomposed into approximation and detailed coefficients at required level using DWT. The approximation and detailed coefficients of both images are combined using fusion rule [18]. The fused image is obtained using inverse discrete wavelet transform (IDWT) as defined in Equation(14)

\[
I_f(x,y) = \text{IDWT} \{ \phi \{ \text{DWT} \{ I_1(x,y), \text{DWT} \{ I_2(x,y) \} \} \} \}
\] (14)

**Fig 1: One Level of 2D Image Decomposition**

The fusion rule used in this paper are i) simple averages of approximation coefficient and picking the detailed coefficient in each subband with largest magnitude ii)Max of the approximation coefficients of visual image and retaining the detailed coefficients of visual image iii) Replacing the approximation coefficient of visual image with the approximation coefficient of MWM image

**IV. PERFORMANCE EVALUATION**

The different performance metric evaluated when reference image is not available are:

i) Standard Deviation (STD)[19]: measures the contrast in the fused image. An image with high contrast would have a high standard deviation.

\[
SD = \sqrt{\sum_{i=0}^{L} (i - \bar{i})^2 h_f(i)}
\] (15)

Where \( h_f(i) \) is the normalized histogram of the fused image \( I(x,y) \) and \( L \) is the number of frequency bins in the histogram.

ii) Entropy[19]: measures the information content of a fused image and also measures the sensitivity to noise and other unwanted rapid fluctuations

\[
He = -\sum_{i=0}^{L} h_f(i) \log h_f(i)
\] (16)

iii) Cross Entropy[20]: Evaluates similarities in information content between input images and fused images. Overall cross entropy of the source images \( I_1 \) and \( I_2 \) and the fused image \( I_f \) is defined in Equation(17) and (18).

\[
CE(I_1, I_f; I_f) = \frac{CE(I_1 : I_f) + CE(I_2 : I_f)}{2}
\] (17)

Where \( CE(I_1 : I_f) = \sum_{i=0}^{L} h_{1f}(i) \log \left( \frac{h_{1f}(i)}{h_{1f}(i)} \right) \) and

\[
CE(I_2 : I_f) = \sum_{i=0}^{L} h_{2f}(i) \log \left( \frac{h_{2f}(i)}{h_{2f}(i)} \right)
\] (18)

iv) Fusion MI[21] measures the degree of dependence of the two images. A larger measure signifies a better quality. If the joint histogram between \( I_1(x,y) \) and \( I_f(x,y) \) is defined as \( h_{1f}(i,j) \) and \( I_2(x,y) \) as \( h_{2f}(i,j) \) then MI between the source and fused images is defined in Equation (19) and

\[
FMI = MI_{I_1 : I_f} + MI_{I_2 : I_f}
\] (20)

\[
MI_{I_1 : I_f} = \sum_{i=0}^{M} \sum_{j=0}^{N} h_{1f}(i,j) \log \left( \frac{h_{1f}(i,j)}{h_{1f}(i,j) h_{1f}(i,j)} \right)
\] (19)

\[
MI_{I_2 : I_f} = \sum_{i=0}^{M} \sum_{j=0}^{N} h_{2f}(i,j) \log \left( \frac{h_{2f}(i,j)}{h_{2f}(i,j) h_{2f}(i,j)} \right)
\] (20)

v) Fusion Quality Index(FQI)[21] indicates that the fused image contains all the information from the source and fusion image
source images and is in the range of 0 to 1 and is defined as in Equation (21).
\[
FQA = \sum_{w} c(w)Q(I_w, I_f, |w|) + (1 - c(w))Q(I_w, I_f, |w|) \quad (21)
\]
where \( \lambda(w) = \frac{\sigma_{I_1}^2}{\sigma_{I_1}^2 + \sigma_{I_2}^2} \) and \( C(w) = \max(\sigma_{I_1}^2, \sigma_{I_2}^2) \)

are computed over window, \( c(\omega) \) is a normalized version \( C(\omega) \) and \( Q(I_1, I_f, |\omega|) \) is the quality index over a window for a given source image and fused image.

V. SIMULATION RESULTS

Visual and the corresponding MWM image for two image data set is considered for fusion for identifying the person with concealed weapon. Visual and MWM images for dataset 1 and dataset 2 are shown in Figs. 3 and 5. In both cases, the concealed weapon is clearer in MWM image, but cannot identify the person with concealed weapon. Hence fusion of visual and MWM image is considered for both the datasets.

![Fig 3: Dataset1: Original gray and MWM image](image)

![Fig 4: Fusion result for Dataset1: (a) PCA (b) Block PCA (c) Spatial Frequency](image)

![Fig 5: Fusion result using DWT (a) Average of approximate coefficients (b) Maximum of approximate coefficients (c) Retaining the approximate coefficients of MWM image](image)

![Fig 6: Dataset2: Original visual and MWM image](image)

![Fig 7: Fusion result for Dataset 2: (a) PCA (b) Block PCA (c) Spatial Frequency](image)

![Fig 8: Fusion result using DWT (a) Average of approximate Coefficients (b) Maximum of approximate coefficients (c) Retaining the approximate coefficients of MWM image](image)
shown in Figures 4, 5, 7 and 8 for both the datasets. For block based PCA, block sizes of 4x4, 8x8,16x16,32x32 is implemented and the fused image had no blocking artifacts for 32x32 and hence the performance metric is applied for this block size. From the simulation results, it is clear that all the pixel based fusion algorithms are able to detect the concealed weapon and also identify the person. The performance of image fusion algorithm is evaluated using the performance metric which are tabulated in Tables 1 and 2 for dataset 1 and data set 2.

![Table 1: Performance Metric for Dataset1](image1)

<table>
<thead>
<tr>
<th>Methods</th>
<th>Standard Deviation</th>
<th>Entropy</th>
<th>Cross Entropy</th>
<th>FM</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA</td>
<td>75.301</td>
<td>7.570</td>
<td>0</td>
<td>2.99</td>
</tr>
<tr>
<td>Block_PC_A(32x32)</td>
<td>65.496</td>
<td>7.539</td>
<td>0</td>
<td>2.82</td>
</tr>
<tr>
<td>Spatial Frequency</td>
<td>65.829</td>
<td>7.547</td>
<td>-3.4185</td>
<td>3.00</td>
</tr>
<tr>
<td>DWT(avg)</td>
<td>65.446</td>
<td>7.542</td>
<td>0.767</td>
<td>2.73</td>
</tr>
<tr>
<td>DWT(max)</td>
<td>88.144</td>
<td>7.556</td>
<td>-0.019</td>
<td>2.92</td>
</tr>
<tr>
<td>DWT(app)</td>
<td>59.984</td>
<td>6.748</td>
<td>1.262</td>
<td>2.54</td>
</tr>
</tbody>
</table>

![Table 2: Performance Metric for Dataset2](image2)

<table>
<thead>
<tr>
<th>Methods</th>
<th>Standard Deviation</th>
<th>Entropy</th>
<th>Cross Entropy</th>
<th>FM</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA</td>
<td>63.13</td>
<td>7.53</td>
<td>0</td>
<td>2.89</td>
</tr>
<tr>
<td>Block_PC(32x32)</td>
<td>62.04</td>
<td>7.47</td>
<td>0</td>
<td>2.73</td>
</tr>
<tr>
<td>Spatial Frequency</td>
<td>58.52</td>
<td>7.43</td>
<td>-5.47</td>
<td>3.12</td>
</tr>
<tr>
<td>DWT(avg)</td>
<td>59.76</td>
<td>7.49</td>
<td>-0.16</td>
<td>2.82</td>
</tr>
<tr>
<td>DWT(max)</td>
<td>70.63</td>
<td>7.40</td>
<td>-0.20</td>
<td>3.12</td>
</tr>
<tr>
<td>DWT(app)</td>
<td>60.22</td>
<td>7.40</td>
<td>-0.17</td>
<td>2.87</td>
</tr>
</tbody>
</table>

VI. CONCLUSION AND FUTURE WORK

Pixel level fusion algorithms using PCA, Block based PCA, SF and wavelet transform using different fusion rules is implemented for Dataset1 and Dataset2. Different performance metrics have been evaluated for the fusion algorithms. Image fusion using Wavelet with fusion rule as maximum of approximate coefficients shows better performance while in other metrics PCA shows better performance. The limitation of the paper is that it is very difficult to get visual image and MWM image simultaneously from two different sensors. In future, fusion algorithms can be applied for visual and IR images, but the drawback of IR imaging is low penetration.

REFERENCES


