Fusion of Infrared and Visual Images Using Bacterial Foraging Strategy

RUTUPARNA PANDA AND MANOJ KUMAR NAIK

Department of Electronics and Telecommunication Engg.
VSS University of Technology, Burla-768018, INDIA.
Phone: 91-663-2431857; Fax: 91-663-2430204
Email: r_ppanda@yahoo.co.in

Abstract – This paper presents new methods for fusion of the visual and thermal images for pattern recognition. Researchers have suggested different fusion schemes to find out pattern vectors for object detection and recognition. The different fusion schemes are — data fusion, decision fusion etc. These schemes have been proposed in different way to improve the performance. Hence, here we propose three new methods for fusing the visual and infrared (IR) images. The proposed new methods are – fusion using information content from Gray Level Co-occurrence Matrix (GLCM), fusion using wavelet energy signature and fusion by maximizing wavelet energy signature using *E. coli* bacteria foraging strategy (EBFS). In the third method, we consider information fusion as an optimization problem and then solve it using EBFS as a search algorithm. Finally, we compare the results using the contrast signature from GLCM and observed that the later scheme using EBFS shows better results than other two methods.

Keywords – Pattern recognition, Evolutionary computation, Bacteria foraging, Wavelet theory.

1 Introduction

Despite of significant research in the field of object recognition, there is a practical challenge due to different lightening conditions. In the case of poor lightening condition, many algorithms fail to detect the object correctly. Let us take an example of face recognition, where different face recognition algorithms show better recognition rate under good lightening conditions. But, when the image poses bad lightening condition, the algorithm fails [1]. This led to the development of thermal infrared image based recognition schemes [2]. Using thermal infrared (IR) images for pattern recognition, one can improve the recognition accuracy [3]. This, in turn, warrants us to develop new efficient techniques for information fusion. The fused images, thus, formed can be used by standard face recognition algorithms for better recognition.

Thermal IR spectrum comprising of mid-wave IR ($3-5\,\mu m$) and long-wave IR ($8-12\,\mu m$) bands have been suggested as an alternative source of information for detection and recognition of faces. Generally, the Thermal IR sensors measure heat energy emitted (not reflected) from the objects. So, this property of the Thermal IR sensor can be used to take image at low illumination conditions or even in the total darkness, where visual recognition techniques fail. As in the case of face recognition, the thermal IR captures the heat generated by the blood vessels in face. It may be noted that every human being has different signature. So this can be used as a feature for classification of the face image.

Recently, Passino [4] has reported a new distributed optimization technique known as *E. coli* bacteria foraging strategy (EBFS) for solving non-gradient optimization problems. In his paper [4], the author has explained the biology and physics underlying the foraging behavior of *E. coli* bacteria that are present in human intestines. In addition to foraging behavior, these bacteria also exhibit other behaviors like aerotaxis, thermotaxis and phototaxis. It is interesting to note here that Chemotaxis is a foraging behavior of these bacteria, which can be used to solve non-gradient optimization problems. This kind of foraging behavior of bacteria can be easily

simulated by a digital computer. Note that a chemotactic step may be a tumble followed by a run or else a tumble followed by another tumble. This enables a cell to move in a right direction of increasingly nutrient gradient given a patch area of nutrients. Swarming is also a bacterial foraging strategy where cell released attractants is used for signaling other cells so that they swarm together. Other important steps involved in bacterial foraging strategy are – reproduction, elimination and dispersal. In the reproduction step, the least healthy bacteria die because they could not get much nutrient during their lifetime of foraging. But, the healthiest bacteria each split into two bacteria and are stored in the same location. The elimination and dispersal step should immediately follow a reproduction step. In an elimination-dispersal step, any bacteria can be eliminated from the population by dispersing it to a random location. Usually the frequency of chemotactic steps is greater than the frequency of reproduction steps.

This EBFS is a kind of evolutionary computation (EC) method and can be used to solve different non-gradient optimization problems. In this paper, we have been motivated to use the EBFS for finding the optimal wavelet coefficients for maximizing the energy of the fused images.

In this paper, we propose three fusion schemes for fusion of visual and thermal images using the spatial information possess by the original image to generate the fused image. First, we propose a method based on the information content using the co-occurrence matrix signature. Second, we present a method using the energy signature obtained from 2-D wavelet transformed images. The motivation is due to the given fact that the energy content of the wavelet coefficients gives better means of texture classification. Third, we propose a method using bacteria foraging strategy algorithm to maximize the energy of fused approximation coefficients derived from the wavelet transformed images. Finally, we compare the results using a gold standard, i.e. contrast and inverse different moment (IDM) of co-occurrence signature of the fused image. Third method being an optimization scheme gives better fused image compared to the other two given methods. All three proposed methods require some information from GLCM or wavelet decomposition, so the name suggested as Information Fusion.

The organization of the paper is as follows : Section 2 is the preliminary section, where GLCM, Wavelet theory and EBFS are introduced. The EBFS algorithm has also been presented in Section 2. Section 3 deals with three new proposed schemes for information fusion. Experimental results are produced in Section 4. Concluding remarks are given in Section 5.

2 Preliminaries

2.1. Gray Level Co-occurrence matrices

The Co-occurrence matrix introduced by Haralick et al. [5], originally called gray-tone spatial dependency matrices, define textural properties of images. The co-occurrence matrix (also known as Gray Level Co-occurrence Matrix (GLCM)) is a directional histogram constructed by counting the occurrence of pairs of pixels separated by some vector displacements.

Let I be an image whose pixel grey levels are in the range 0, 1, . . ., L-1. Let us take an integer valued displacement vector $\bar{d} = (p,q)$, which specifies the relative position of the pixels at coordinates (x,y) and $(x+p, y+q)$. A GLCM is a $L \times L$ matrix whose (i,j) element is the number of pairs of pixels of I in relative position $\bar{d}$ such that the first pixel has gray level i and the second pixel has gray level j. So the GLCM matrix M involves counts of pairs of neighboring pixels. Then M is formed for each of four quantized directions 0, 45, 90, and 135 degrees. So GLCM matrix can be represented as $\mathrm{M}(p,q)$ or $\mathrm{M}(\bar{d},\theta)$, where $\bar{d}$ refers to displacement distance and θ refers to a particular angle. There are simple relationships existing among certain pairs of the estimated GLCM $\mathrm{M}(\bar{d},\theta)$. Let $\mathrm{M}^T(\bar{d},\theta)$ denote the transpose of matrix $\mathrm{M}(\bar{d},\theta)$. Then $\mathrm{M}(\bar{d},0^0) = \mathrm{M}^T(\bar{d},180^0)$; $\mathrm{M}(\bar{d},45^0) = \mathrm{M}^T(\bar{d},225^0)$; $\mathrm{M}(\bar{d},90^0) = \mathrm{M}^T(\bar{d},270^0)$; and $\mathrm{M}(\bar{d},135^0) = \mathrm{M}^T(\bar{d},315^0)$. Thus, the knowledge of $\mathrm{M}(\bar{d},180^0)$, $\mathrm{M}(\bar{d},225^0)$, $\mathrm{M}(\bar{d},270^0)$, and $\mathrm{M}(\bar{d},315^0)$ add nothing to the specification of the texture.

A number of texture features may be extracted from the GLCM [5,6]. Some of the important texture features computed from the GLCM are:

- Angular Second Moment (ASM):

$$\text{ASM} = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\{M(i,j)\}^2 \quad (1)$$

ASM is a measure of homogeneity of an image. For an image, constant gray levels means higher ASM.

- Contrast:

$$\text{Contrast} = \sum_{n=0}^{L-1} n^2 \left\{ \sum_{i=1}^{L}\sum_{j=1}^{L} M(i,j) \right\}, \qquad |i-j| = n \quad (2)$$

This measure of local intensity variation will favor contributions from $M(i,j)$ away from the diagonal, i.e. $i \neq j$.

- Correlation:

$$\text{Correlation} = \frac{\left\{ \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} (i \times j) \times M(i,j) - (\mu_x \times \mu_y) \right\}}{\sigma_x \times \sigma_y} \quad (3)$$

This is a measure of gray level linear dependence between the pixels at the specified positions relative to each other.

- Variance:

$$\text{Veriance} = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1} (i-\mu)^2 M(i,j) \quad (4)$$

Variance puts relatively high weights on the elements that differ from the average value of $M(i,j)$, and treated as a measure of heterogeneity.

- Inverse Different Moment (IDM):

$$\text{IDM} = \sum_{i-0}^{L-1}\sum_{j=0}^{L-1} \frac{1}{1+(i-j)^2} M(i,j) \quad (5)$$

IDM is inversely related to contrast, and also known as Local Homogeneity.

- Entropy:

$$\text{Entropy} = -\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} M(i,j) \log M(i,j) \equiv H_{xy} \quad (6)$$

This is a measure of the randomness of the intensity distribution.

- Information Measure (IM):

$$\text{IM} = \frac{H_{xy} - H_{xy}^1}{\max\left[H_x, H_y\right]} \quad (7)$$

This is the measure of information content; depend on the upper value of the gray scale value. Where

- $H_{xy}^1 = -\sum_{i=0}^{L-1}\sum_{j=0}^{L-1} M(i,j) \log\left(M_x(i) M_y(j)\right)$
- $M_x(i) = \sum_{j=0}^{L-1} M(i,j)$
- $M_y(j) = \sum_{i=0}^{L-1} M(i,j)$
- means: μ, μ_x, μ_y
- standard deviations: σ_x, σ_y
- entropies: H_x, H_y

2.2. Wavelet Theory and Signature

Usually, the 2-D discrete wavelet transformation is computed by applying different filter bank to the image [7]:

$$L_n(b_i,b_j) = \left\lfloor H_x \otimes \left\lfloor H_y \otimes L_{n-1} \right\rfloor_{\downarrow 2,1} \right\rfloor_{\downarrow 1,2} (b_i,b_j) \quad (8)$$

$$D_{n1}(b_i,b_j) = \left\lfloor H_x \otimes \left\lfloor G_y \otimes L_{n-1} \right\rfloor_{\downarrow 2,1} \right\rfloor_{\downarrow 1,2} (b_i,b_j) \quad (9)$$

$$D_{n2}(b_i,b_j) = \left\lfloor G_x \otimes \left\lfloor H_y \otimes L_{n-1} \right\rfloor_{\downarrow 2,1} \right\rfloor_{\downarrow 1,2} (b_i,b_j) \quad (10)$$

$$D_{n3}(b_i,b_j) = \left\lfloor G_x \otimes \left\lfloor G_y \otimes L_{n-1} \right\rfloor_{\downarrow 2,1} \right\rfloor_{\downarrow 1,2} (b_i,b_j) \quad (11)$$

Where $\otimes$ denotes the convolution operator, $\downarrow 2,1 (\downarrow 1,2)$ represent subsampling along the rows (columns), and $L_0 = I(\vec{x})$ is the original image. H is the lowpass filter and G is the bandpass filter. L_n is obtained by lowpass filtering and known as low resolution image at scale n. The *detail images* D_{ni} are obtained by bandpass filtering in a specific direction and contain the directional information at scale n. The original image I is represented by a set of subimages at several scales: $\{L_d, D_{ni}\}_{i=1,2,3,\,n=1....d}$ which is a multiscale representation of depth d of the image I.

For the texture classification in Wavelet domain we need energy signature as a prime component [8,9] and it is given as:

- For a subimage L_n, containing *N* coefficients is defined as

$$E_n = \frac{1}{N}\sum_{j,k}\left(L_n\left(b_j, b_k\right)\right)^2,\ n = 1,.....d \quad ,(12)$$

- And for a subimage D_{ni} containing *N* coefficients is defined as

$$E_{ni} = \frac{1}{N}\sum_{j,k}\left(D_{ni}\left(b_j, b_k\right)\right)^2,\ \ n = 1,....d, i = 1,2,3 \quad (13)$$

The wavelet energy signature reflects the distribution of energy in frequency axis over scale and along different orientations. This is a measure of dispersion of the wavelet coefficients.

2.3. E. *coli* Bacteria Foraging Strategy (EBFS)

E. *coli* Bacteria foraging strategy [4] is an optimization process, inspired from biological behavior of bacteria. Basic idea of foraging reveals the fact that animals take actions to maximize the energy obtained per unit time spent for foraging. The foraging theory is based on search of nutrients in a way that maximizes their energy intake *E* per unit time *T* spent for foraging, and tries to maximize a function like $\frac{E}{T}$. Foraging involves finding such patches, deciding whether to enter a patch and search for food, and whether to continue searching for food in the current patch or to go and find another patch that hopefully has a higher quality and quantity of nutrients than the current patch. The bacteria can move in two different ways; it can run (swim for a period of time) or it can tumble, and it alternates between these two modes in its entire lifetime. After a tumble, the cells are generally pointed in random direction, but slightly bias towards the previous traveling zone. A bacteria comes under three different stages in its life time – chemotaxis ; reproduction ; and elimination & dispersal event.

Chemotaxes: The motion patterns that the bacteria will generate in the presence of chemical attractants and repellents are called chemotaxes. This helps the other bacterium to follow the root. Next, suppose that the bacterium happens to encounter a nutrient gradient. The change in the concentration of the nutrient triggers a reaction such that the bacterium will spend more time swimming and less time tumbling. As long as it travels on a positive concentration gradient, it will tend to lengthen the time for swimming, up to a certain point. The swimming or tumbling is done by the decision-making mechanisms. Here it performs a type of sampling, and it remembers the concentration a moment ago, compares it with a current one, and makes decisions based on the difference.

To represent a tumble, a unit length random direction, say $\phi(j)$, is generated; this will be used to define the direction of movement after a tumble. In particular, we let

$$\theta^i(j+1,k,l) = \theta^i(j,k,l) + C(i)\phi(j) \quad (14)$$

Where $\theta^i(j,k,l)$ represent location of the *i*th bacterium and $C(i)$ denote a basic chemotactic step size. When there is a cell-to-cell signaling via an attractant, bacteria swarm together. And it can be methodically treated as combined cell-to-cell attraction and repelling effects. That is

$$\begin{aligned} J_{cc}(\theta, P(j,k,l)) &= \sum_{i=1}^{S} J_{cc}^i\left(\theta, \theta^i(j,k,l)\right) \\ &= \sum_{i=1}^{S}\left[-d_{attract}\exp\left(-w_{attract}\sum_{m=1}^{p}\left(\theta_m - \theta_m^i\right)^2\right)\right] \\ &\quad + \sum_{i=1}^{S}\left[h_{repellent}\exp\left(-w_{repellent}\sum_{m=1}^{p}\left(\theta_m - \theta_m^i\right)^2\right)\right] \end{aligned} \quad (15)$$

where *S* is the total number of bacterium, p is the number of parameters to be optimized, *J* is the cost function and $d_{attract}$, $w_{attract}$, $h_{repellent}$, $w_{repellent}$ are different coefficients that are properly chosen.

Reproduction: After some chemotaxis steps it compares all the nutrient concentration where bacteria are present. Where ever it finds the higher nutrient concentration, at that place each bacterium reproduces an exact copy of its own. With low nutrient concentration, the bacterium should die.

Elimination & Dispersal: This is another important event, assist to chemotaxis step. It keeps track on the bacteria and see whether they are appropriately placed or not. If not then it places a bacterium in an arbitrary food space for new beginning of search. From a broad prospective, elimination and dispersal are parts of the population-level long-distance motile behavior.

2.4. EBFS Algorithm

1. Initialization
   - Chose S number of bacteria for the chemotaxis step as the number of population.
   - Then determine the number of parameters to be optimized p.
   - Then determine the number of chemotaxis steps N_c, number of reproduction steps, N_{re} and the number of elimination & dispersal steps as N_{ed}.
   - Then determine the maximum length of swimming of a bacterium while hill climbing as N_s.
   - Also determine the chemotactic step size for swimming $C(i)$ as $i = 1,2,....,S$.
   - Then choose the $d_{attract}$, $w_{attract}$, $h_{repellent}$, $w_{repellent}$ parameters that helps in swarming with appropriate value.
   - Initial value of $\theta^i, i = 1,2,....,S$ must be chosen, so that these are randomly distributed across the domain of the optimization problem.
   - Initially $j = k = l = 0$, where j,k,l parameter determine how many steps it already moves in chemotaxis, reproduction and elimination & dispersal event.
   - Define elimination and dispersal probability p_{ed}.

   For the given algorithm, note the fact that updates to the θ^i automatically results in updates to P, where P represent the position of each member in the population of the S bacteria at the respective step.
2. Iterative algorithm
   A. Elimination-dispersal loop: $l = l + 1$
   B. Reproduction loop: $k = k + 1$
   C. Chemotaxis loop: $j = j + 1$
      a. For $i = 1,2,...S$, take a chemotactic step for bacterium i as follows.
      b. Compute $J(i, j, k, l)$. Let $J(i, j, k, l) = J(i, j, k, l) + J_{cc}(\theta^i(j, k, l), P(j, k, l))$ (i.e., add on the cell-to-cell attractant effect to the nutrient concentration).
      c. Save the value $J(i, j, k, l)$ as J_{last} for the next step, to get a better cost via a run.
      d. Tumble: Generate a random vector $\Delta(i) \in \Re^p$ with each element $\Delta_m(i), m = 1,2,...p$, a random number on [-1, 1].
      e. Move: Let $\theta^i(j+1,k,l) = \theta^i(j,k,l) + C(i)\frac{\Delta(i)}{\sqrt{\Delta^T(i)\Delta(i)}}$. This results in a step of size $C(i)$ in the direction of the tumble for bacterium i.
      f. Compute $J(i, j+1, k, l) = J(i, j, k, l) + J_{cc}(\theta^i(j+1,k,l), P(j+1,k,l))$.
      g. Swim:
         i. Let $m = 0$, as a counter for swim length.
         ii. While $m < N_s$
         - Let $m = m + 1$
         - If $J(i, j+1.k.l) < J_{last}$ (if doing better), then $J_{last} = J(i, j+1, k, l)$ and let $\theta^i(j+1,k,l) = \theta^i(j+1,k,l) + C(i)\frac{\Delta(i)}{\sqrt{\Delta^T(i)\Delta(i)}}$ and use this $\theta^i(j+1,k,l)$ to compute the new $J(i, j+1, k, l)$ as f.
         - Else, let $m = N_s$ End of the while loop.
      h. Move to the next bacterium $(i+1)$ if $i \neq S$, to step b. till $i == S$.
   D. Verify the $j < N_c$, if yes then go to C.
   E. Reproduction:
      a. For the given k and l, and each $i = 1,2,...S$, let $J^i_{health} = \sum_{j=1}^{N_c+1} J(i, j, k, l)$ be the health of bacterium i. Sort bacteria and chemotactic

parameters $C(i)$ in ascending order of the cost J_{health} (higher cost means lower health).

b. The $S_r = \frac{S}{2}$ bacteria with higher cost will die and other S_r with best value split, that means the exact replica of the lower cost will be generated and placed in the same location as their parents.

F. If $k < N_{re}$, go to step B.

G. Elimination-dispersal: For $i = 1,2,...S$, with probability $p_{ed,}$ eliminate and disperse each bacterium by keeping the population constant. This can be achieved by randomly placing the bacterium in the search space.

H. If $l < N_{ed}$, then go to step A; otherwise end.

3 Information Fusion

In biometrics we have to find the particular pattern and recognize it against many more available patterns. Recently, researchers have shown more interest to develop some schemes for information fusion. Examples include – face and fingerprint [10, 11], fusion of face and hand geometry [10], fusion of face and speech data [12] etc. The combined use of visual and IR image data makes a visible means of improving performance in face recognition [2]. It is interesting to note here that the face recognition algorithm applied to fused images of visual and thermal images demonstrate better result than the visual face recognition or thermal face recognition alone [13]. In [14], it has been shown that the fused image improves the recognition performance. In [15], authors have demonstrated the data fusion and decision fusion for robust face recognition and shown that the recognition performance improves significantly. Huang and Jing [17] have discussed Multi-focus image fusion using pulse coded Neural Networks. Multiscale fusion algorithms using Pyramid, DWT and Iterative DWT are presented in [18]. This motivates us to make a fusion scheme, taking some textural information from both images and then fuse it to get a new image, which contains information from both images. For this reason, we describe information fusion as a method for fusing visual and thermal images and make the fused image more suitable for recognition.

Here we describe three methods of information fusion. The first method use information content of both images computed from the GLCM matrix. The second method is wavelet based, where the energy parameter is taken into consideration for fusing two images. Authors in [18] have claimed that iterative DWT gives better results than pyramidal and DWT algorithms. They have used iterative DWT algorithm for optimization. Multi-focus image fusion using PCNN has been reported in [19]. Image fusion using lifting wavelet transform with human visual features has been discussed in [20]. This has motivated us to develop an evolutionary computational algorithm for optimization of energy content. In this context, we propose a third method which is based on EBFS. It seems to be a challenging, yet, interesting problem. In fact, we do not find any precise mathematical model to obtain optimal wavelet coefficients to maximize the energy content of the fused image. Hence, we need a suitable evolutionary computation (EC) scheme to solve this problem. In this paper, we use bacteria foraging strategy for maximizing energy content of the fused image, considering the optimal approximation coefficients of one level of the wavelet transform.

3.1 Fusion using information content from GLCM matrix

This scheme can be expressed as a weighted sum of pixel intensity from both the images:

$$F(x,y) = a \times V(x,y) + b \times I(x,y) \qquad (16)$$

where $F(x,y)$ is the fused image formed from visual image $V(x,y)$ and infra red image $I(x,y)$. The parameters '*a*' and '*b*' are chosen such that the sum is equal to one. The algorithm for determining values for '*a*' and '*b*' is stated below:

Check that the dimension and intensity level of both visual and thermal images must be same.

Calculate the GLCM matrix of both the images.

Then using the corresponding GLCM matrix calculate the Information Measure of both the images using Eq.(7) and name it as *IM*1 (visual) and *IM*2 (thermal).

Then,

if (*IM*1>*IM*2)

$$a = \frac{IM1}{IM1 + IM2} \text{ and } b = \frac{IM2}{IM1 + IM2}$$

else

$$a = \frac{IM2}{IM1 + IM2} \text{ and } b = \frac{IM1}{IM1 + IM2}$$

From above discussions, it is seen that always we give more importance to the visual image for this scheme. In other words, the proposed fusion scheme is more biased towards the visual side irrespective of the fact that it also contains thermal data. Then, we generate the fusion image using the parameters '*a*' and '*b*' together with Eq.(16).

3.2 Fusion using Wavelet energy signature

As we know from [8], one of the most important texture feature using wavelet transformation is the energy signature given in Eqs.(12,13). Here we employ first level decomposition of the image. For our application, the equations (8-11) can be rewritten as:

$$L_1(b_i, b_j) = \lfloor H_x \otimes \lfloor H_y \otimes L_0 \rfloor_{\downarrow 2,1} \rfloor_{\downarrow 1,2} (b_i, b_j) \quad (17)$$

$$D_{11}(b_i, b_j) = \lfloor H_x \otimes \lfloor G_y \otimes L_0 \rfloor_{\downarrow 2,1} \rfloor_{\downarrow 1,2} (b_i, b_j) \quad (18)$$

$$D_{12}(b_i, b_j) = \lfloor G_x \otimes \lfloor H_y \otimes L_0 \rfloor_{\downarrow 2,1} \rfloor_{\downarrow 1,2} (b_i, b_j) \quad (19)$$

$$D_{13}(b_i, b_j) = \lfloor G_x \otimes \lfloor G_y \otimes L_0 \rfloor_{\downarrow 2,1} \rfloor_{\downarrow 1,2} (b_i, b_j) \quad (20)$$

Here we use the wavelet energy signature given in Eq.(12) as the required factor for fusing wavelet coefficients obtained from both visual and infrared images and are given as:

$$L_1^F = a \times L_1^V + b \times L_1^I \quad (21)$$

$$D_{11}^F = a \times D_{11}^V + b \times D_{11}^I \quad (22)$$

$$D_{12}^F = a \times D_{12}^V + b \times D_{12}^I \quad (23)$$

$$D_{13}^F = a \times D_{13}^V + b \times D_{13}^I \quad (24)$$

Where the superscripts *F, V, I* represent the fusion, visual and thermal images, respectively. And *a*, *b* are the parameters determined by using energy signature given as below:

i. Decompose the visual and thermal images with 2-D two-scale wavelet transform into 4 sub-images (one approximation image and other three detail images).

ii. Calculate the energy signature using Eq.(12) of all the 4 sub-images of both the visual and thermal images as $\xi_{L_1}^y, \xi_{D_{11}}^y, \xi_{D_{12}}^y, \xi_{D_{13}}^y$. Where *y* can be *either V or I*.

iii. Then calculate *a* and *b* (note that $a + b = 1$):

For L_1^F, $a = \dfrac{\xi_{L_1}^V}{\xi_{L_1}^V + \xi_{L_1}^I}$ and

$$b = \frac{\xi_{L_1}^I}{\xi_{L_1}^V + \xi_{L_1}^I} \quad (25)$$

For D_{11}^F, $a = \dfrac{\xi_{D_{11}}^V}{\xi_{D_{11}}^V + \xi_{D_{11}}^I}$ and

$$b = \frac{\xi_{D_{11}}^I}{\xi_{D_{11}}^V + \xi_{D_{11}}^I} \quad (26)$$

For D_{12}^F, $a = \dfrac{\xi_{D_{12}}^V}{\xi_{D_{12}}^V + \xi_{D_{12}}^I}$ and

$$b = \frac{\xi_{D_{12}}^I}{\xi_{D_{12}}^V + \xi_{D_{12}}^I} \quad (27)$$

For D_{13}^F, $a = \dfrac{\xi_{D_{13}}^V}{\xi_{D_{13}}^V + \xi_{D_{13}}^I}$ and

$$b = \frac{\xi_{D_{13}}^I}{\xi_{D_{13}}^V + \xi_{D_{13}}^I} \quad (28)$$

Then using the parameters '*a*' and '*b*' and Eqs.(21-24), we get wavelet coefficients of the fused image.

iv. Then using the inverse wavelet transform we get the fused image.

3.3 Fusion by maximizing Wavelet energy signature using EBFS

From the above discussion, we know that the energy is a prime factor for texture classification in the wavelet domain. Hence, here we again consider the energy as our cost function in E. *coli* Bacteria Foraging Strategy

(EBFS). This procedure of fusion deviate much from the earlier methods. For the fusion purpose, we have to concentrate on the approximation coefficients (Eq.(17)) only, not the detail coefficients (Eqs.(18-20)). For this reason we apply the EBFS optimization technique on the approximation coefficients only for a better result. In the above procedure (second method described in section 3.2) '*a*' and '*b*' are computed from Eq.(25) and ,thus, are fixed for all the approximation coefficients of wavelet transformed images. On the other hand, here we propose an optimization approach where parameters '*a*' and '*b*' are independently decided for the approximation coefficients. The approximation coefficients of the fused image can mathematically be defined as :

$$L_1^F(x,y) = a(x,y)\times L_1^V + b(x,y)\times L_1^I,$$

$$\text{where } a(x,y)+b(x,y)=1 \tag{29}$$

But, for obtaining the other detail coefficients, parameters '*a*' and '*b*' are decided as per the previous method (second method given in section 3.2) by using Eqs.(26-28) and the detail coefficients for fused image can be computed by using Eq.(22-24).

The algorithm for computing the optimal parameters '*a*' and '*b*' for obtaining approximation coefficients for the fused image from the wavelet coefficients (approximation coefficients) of visual and thermal images is given below:

Algorithm

A. Decompose the given visual and thermal images into 2-D wavelet transform images to get four sub-images of same size (one approximation image and rest three detail images).
B. First take the approximation coefficients of the visual and thermal images and then do the following steps:
   a) Calculate the dimension of the approximation coefficients as $p = m \times n$, where m is the number of rows and n is the number of columns of the approximation coefficients.
   b) Bacteria representation: For EBFS optimization process, we have to represent the bacterial search dimension. Here the search dimension is equal to p. This means each coefficient represents itself as a search parameter.
   c) Cost function: For this problem, the cost function is the maximization of energy content of the fused approximation coefficients. This implies that after searching the visual and thermal approximation space we get the set of $a(x,y)$ and $b(x,y)$ values, where $x = 1,2,...,n$ and $y = 1,2,....,m$. The Cost Function can be written as:

$$J = energy_{\max}\left(L_1^F\right) = energy_{\max}\left(a(x,y)\times L_1^V + b(x,y)\times L_1^I\right), \tag{30}$$

   where the $a(x,y)$ and $b(x,y)$ parameters are obtained from the EBFS algorithm.
   d) Then taking bacterial cost function representation, use the Bacteria Foraging Strategy Algorithm to get optimal set of values $a_{opt}(x,y)$ and $b_{opt}(x,y)$.
   e) Then compute the approximation coefficients for the fused image by using the Eq.(29).
C. Then consider the detail coefficients of the thermal and visual images:
   a) Determine the *a* and *b* parameters using Eqs.(26-28).
   b) Then compute the detail coefficients for the fused image using Eqs.(22-24).
D. Then get the fused image out of the given visual and thermal images by evaluating the inverse Wavelet transform.

4 Experiments and Results

In this section, we discuss the images considered for the experiments, parameters for EBFS optimization technique and different results obtained from proposed fusion schemes. Finally, we compare the results using a gold standard.

Images for experiments

We consider some face images from the Equinox Database [16], which contain the visual and thermal images of a person with many modalities. But, for our experiment we only consider the LWIR (long-wave IR) images.

Parameters for EBFS Algorithm

As (from previous sections) we know that the search space $p = m \times n$ is too big, we use parameters such that it avoids complexity of algorithm to some extent. For this reason, we choose S=8, $N_c = 8$, $N_{re} = 5$, $N_{ed} = 3$, $N_s = 4$, $C(i) = 0.067$, $d_{attract} = 0.01$, $w_{attract} = 0.2$, $h_{repellent} = 0.01$, and $w_{repellent} = 10$. However, the above parameters can be changed to get better optimization. For example, if we increase the N_c, we get better optimization. Since the search space is too large, we minimize the parameters. If we increase the number of bacteria and other steps involved, the computational complexity increases. Here one can not vary the parameter $C(i)$, which is experimentally determined.

Results

The results consist of fusion images, co-occurrence matrix images, and performance comparison using table format. We compare the quality of fused image so formed using the Co-occurrence signature, Contrast and Inverse Different Moment (IDM). The contrast and IDM are already defined in section 2.1 and are given as:

$$\text{Contrast} = \sum_{n=0}^{L-1} n^2 \left\{ \sum_{i=1}^{L} \sum_{j=1}^{L} M(i,j) \right\}, \qquad |i-j| = n$$

$$\text{IDM} = \sum_{i-0}^{L-1} \sum_{j=0}^{L-1} \frac{1}{1+(i-j)^2} M(i,j)$$

The contrast and IDM are inversely proportional to each other. When the contrast is more, the image has greater information content.

The original images considered without histogram equalization is shown in Figure 1. Fused images obtained using proposed techniques are displayed in Figure 2.

Figure 1: Original images.

Figure 2: Fused images (VI-Visual Image I, VII-Visual Image II, T-Thermal Image).

The corresponding co-occurrence matrix images of the fused images (shown in Figure 2) are displayed in Figure 3.

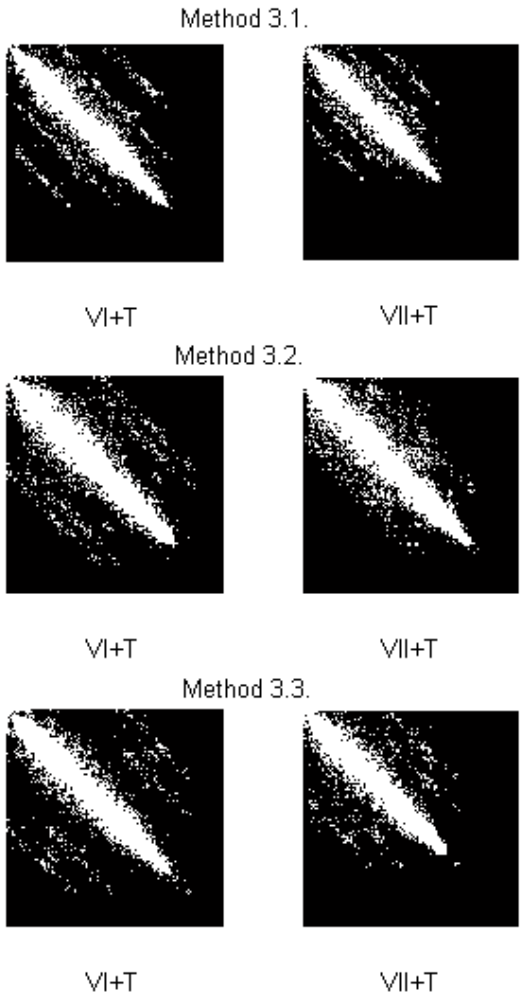

Figure 3 : The Co-occurrence Matrix Images.

Then a performance evaluation has been carried out using contrast and IDM as the gold standards to make it more convincing. The results are compared in Table 1.

Table 1: Performance Comparison of proposed fusion schemes.

	Fusion of Visual Image I and Thermal Image (Contrast, IDM)	Fusion of Visual Image II and Thermal Image (Contrast, IDM)
Method 3.1.	108.63, 0.46563	77.722, 0.48502
Method 3.2.	126.21, 0.4303	96.924, 0.43845
Method 3.3.	142.81, 0.41591	98.366, 0.43396

From Table 1, we can see that fusion based on optimal wavelet energy signature using EBFS gives better results. Further, the proposed method (Method 3.3) gives us good tracking record of energy parameter of fusion. This has been shown in Figure 4.

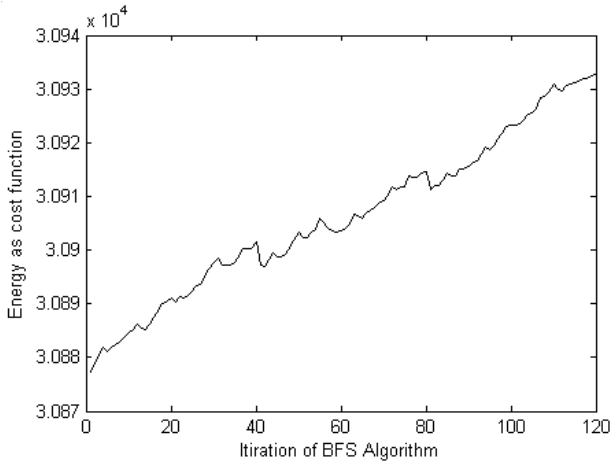

Figure 4: Cost Function of BFS Algorithm.

In the second experiment, we compare the results of proposed fusion schemes applied to the histogram equalize images. The images after histogram equalization are considered for this experiment and are shown in Figure 5.

Figure 5: (E-Equalize, V-Visual, T-Thermal)

Table 2: Comparison of proposed fusion schemes using histogram equalized images.

	Fusion of Equalize Visual Image I and Equalize Thermal Image (Contrast, IDM)	Fusion of Equalize Visual Image II and Equalize Thermal Image (Contrast, IDM)	Fusion of Equalize Visual Image III and Equalize Thermal Image (Contrast, IDM)
Method 3.1.	412.86, 0.28774	287.97, 0.30159	186.3, 0.31952
Method 3.2.	748.52, 0.2264	479.63, 0.23495	293.54, 0.24186
Method 3.3.	762.14, 0.18659	498.77, 0.19939	299.79, 0.22311

The results are compared in Table 2. From Table 2, it is observed that significant improvement can be achieved by considering equalized images. So one can equalize given visual and IR images before fusion. From Table 2, we again see that the proposed EBFS method (section 3.3) gives better results than other two methods.

5 Conclusions

In this paper, three new methods have been proposed for fusion of the visual and IR images considering the advantages and disadvantages of visual and thermal images used for pattern recognition. Considering some images from the Equinox Database for experiments, results are compared. The contrast signature of GLCM has been considered given the fact that it gives us better means for performance evaluation. It has been shown that the fusion by maximizing wavelet energy signature using EBFS gives better results compared to the other two proposed methods, which is quite obvious. But, the EBFS method has some disadvantage while considering time as the prime factor. One has to choose EBFS parameters judiciously in order to reduce computational complexity. Finally, we conclude that the proposed EBFS method can be used for information fusion for different biometrics applications.

References:

[1] Y. Adini, Y. Moses, and S. Ullman, Face recognition: The problem of compensating for changes in illumination direction, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, 1997, pp.721-732.

[2] S.G. Kong, J. Heo, B.R. Abidi, J. Paik, and M.A. Abidi, Recent advances in visual and infrared face recognition – a review, Computer Vision and Image Understanding, Vol. 97, 2005, pp.103-135.

[3] F. Prokoski, History, current status, and future of infrared identification, Proc. IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications, 2000, pp.5-14.

[4] K.M. Passino, Biomimicry of Bacterial Foraging for Distributed Optimization and Control, IEEE Control Sys. Mag., 2002, pp.52-67.

[5] R.M. Haralik, K. Shanmugam, and I. Dinstein, Textural Features for Image Classification, IEEE Trans. on Systems, Man and Cybernetics, Vol. SMC-3, 1973, pp.610-621.

[6] R.W. Conners, M.M. Trivedi, and C.A. Harlow, Segmentation of a High-Resolution Urban Scene Using Texture Operators, Computer Vision, Graphics, and Image Processing, Vol. 25, 1984, pp.273-310.

[7] S. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, IEEE Trans. Pattern Anal. Machine Intell., Vol. 11, 1989, pp.674-693.

[8] G. Van de Wouwer, P. Scheunders, and D. Van Dyck, Statistical Texture Characterization from Discrete Wavelet Representations, IEEE Trans. on Image Processing, Vol. 8, No. 4, 1999, pp.592-598.

[9] T. Chang, and C.C. Jay Kuo, Texture Analysis and Classification with Tree-Structured Wavelet Transform, IEEE Trans. on Image Processing, Vol. 2, No. 4, 1993, pp.429-441.

[10]A. Ross, and A. Jain, Information fusion in biometrics, Pattern Recognition Letters, Vol. 24, 2003, pp.2115-2125.

[11]L. Hong, and A. Jain, Integrating faces and fingerprints for personal identification, IEEE Trans. Patt. Anal. Mach. Intell., Vol. 20, No. 12, 1998, pp. 1295-1307.

[12]S. Ben-Yacoub, Y. Abdeljaoued, and E. Mayoraz, Fusion of face and speech data for person identity verification, IEEE Trans. Neural Networks, Vol. 10, No. 5, 1999, pp. 1065-1074.

[13]X. Chen, P. Flynn, and K. Bowyer, Visible-light and infrared face recognition, Proc. of Workshop on Multimodal User Authentication, 2003, pp.48-55.

[14] J. Wilder, P.J. Phillips, C. Jiang, and S. Wiener, Comparison of visual and infrared imagery for face recognition, Proc. Int. Conf. Automatic Face and Gesture Recognition, 1996, pp.182-187.

[15] J. Heo, S. Kong, B. Abidi, and M. Abidi, Fusion of Visual and Thermal Signatures with Eyeglass Removal for Robust Face Recognition, IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum in conjunction with CVPR 2004, , Washington D.C., 2004, pp. 94-99.

[16] http://www.equinoxsensors.com/products/HID.html

[17] W. Huang,Z. Jing, Multi-focus image fusion using PCNN, Pattern Recognition Letters., vol. 28, (9), 2007, pp.493-500.

[18] Y. Zheng, Multiscale fusion algorithms: Pyramid, DWT and Iterative DWT, 12th International Conference on Information Fusion, 2009, pp. 1060-1065.

[19] Z. Wang, Y.Ma, J. ju, Multi-focus image fusion using PCNN, Pattern Recognition, vol.43,(6), 2010, pp.2003–2016.

[20] X, Gu, Z. Peng, Z. Chen, L. Yuan, B. Huang, Image fusion using lifting wavelet transform with human visual features, Proc SPIE, 2010, pp 1975-1979.