Multimodal Face Recognition using Hybrid Correlation Filters

Anamika Dubey, Abhishek Sharma
Electrical Engineering Department, Indian Institute of Technology Roorkee, India
{ana.iitr, abhisharayiya}@gmail.com

***Abstract*-This article introduces novel hybrid correlation filter for multimodal face recognition in IR and visible spectrum. In the proposed approach, wavelet based image fusion is used to fuse the face images of a person in visible and IR spectrum. Obtained fused images are then used to synthesize HMACE (Hybrid Minimum Average Correlation Energy) and HMACH (Hybrid Maximum Average Correlation Height) filters. These hybrid correlation filters contained the information captured in both the spectrums so recognition of IR as well as CCD image is possible from the same filter. A direct consequence of this approach is the reduction in storage space by 50%. Thorough experimentation is carried out in which face images of different subjects under different pose, illumination and expression were tested using the proposed filters and the results are encouraging.**

Keywords- Thermal, visible, face recognition, wavelet fusion, correlation filter

1. Introduction

Human face recognition has a myriad of applications ranging from security systems, defense applications, intelligent machines to smart human computer interaction and still new applications are coming into picture everyday. But face recognition is a challenging task owing to three main reasons- Illumination variation, Pose variation and Expression variation [1]. It has been over four decades but still the complete solution to these problems is yet to be found. Face recognition is basically a problem related to pattern recognition which can be broadly classified under two categories – Feature-Based recognition and Correlation Pattern Recognition (CPR). There are several feature based methods for face recognition [1]. In the field of CPR some of the most popular correlation filters are MACE [2], MACH [3] and their variants [4 5]. These filters are robust to noise and variations such as - Illumination and view point. These filters have been used with great success in war field scenario for battle tank, truck, missile etc. detection, identification and tracking. Due to their robustness and accuracy these filters can be very effective for face recognition applications. A comparative analysis of feature based methods and CPR demonstrated that CPR is superior than feature based method for face recognition [6]. A comparison for face recognition in IR domain has also shown that CPR is better than feature based method [7]. But unfortunately, very less amount of work has been done in field of CPR application for face recognition. This article is a novel attempt to apply CPR for bimodal face recognition (IR and Visible).

IR images of a person are illumination invariant and they can also be captured in total darkness. This property of IR images has attracted many researchers to work on IR images for illumination invariant face recognition. The images captured by a CCD and IR camera are in different spectrum so the information content of both the images is different and complimentary. Naturally, fusion of these complimentary information contents will result in better performance. Wavelet based fusion is one of the most popular methods for image fusion because of its localization and multi-resolution property [8 9].

This article introduces novel hybrid correlation filters for face recognition in IR and visible spectrum. In the proposed approach, wavelet based image fusion is used to fuse the face images of a person in visible and IR spectrum. Obtained fused images are then used to synthesize MACE and MACH filters. These hybrid correlation filters contained the information captured in both the spectrums so recognition of IR as well as CCD image is possible from the same filter. A direct consequence of this approach is the reduction in storage space by 50%. Thorough experimentation is carried out in which face images of a person under different pose, illumination and expression were tested using the proposed filters and the results are encouraging.

The article is organized in different sections. Section 2 gives a brief description of MACE and MACH filters. Section 3 discusses the proposed methodology and steps involved in a detailed manner. Section 4 carries the details of experimentation. Conclusion and future work are described in Section 5 and 6 respectively.

2. Background theory

The earliest correlation filters were simple matched filters [5]. These matched filters have a profound disadvantage of being extremely noise sensitive and they cannot provide very sharp peaks. Moreover, as many number of matched filters are required as the number of training images are used. These limitations demanded for more noise tolerant and sharp peak yielding filters. MACE and MACH are two such filters which are most frequently used CPR filters owing to their excellent noise tolerance and very high discriminating power.

2.1 Minimum Average Correlation Energy (MACE) filters

MACE filters are obtained by forcing the average cross-correlation plane energy to minimum for the training images with hard constraints at the origin of the plane to yield specific value. It implies that resulting filter gives cross-correlation plane which resembles delta function. However, in practice we do not get the exact delta function but resulting peaks are very sharp and provide a good measure for discrimination between authentic and impostor images. The MACE filter $\boldsymbol{h}$ in frequency domain is -

$$\boldsymbol{h} = \boldsymbol{D}^{-1}\boldsymbol{X}(\boldsymbol{X}^{+}\boldsymbol{D}^{-1}\boldsymbol{X})^{-1}\boldsymbol{u} \qquad (1)$$

Assume we have N training image each having P number of pixel. Then $\boldsymbol{X}$ is P×N matrix of the Fourier transform of the training images arranged in a lexicographic manner. $\boldsymbol{D}$ is P×P diagonal matrix of the average power spectrum of training images. And $\boldsymbol{u}$ is the pre-specified value at the origin of the correlation plane (generally it is taken to be 1 for true class images and 0 for false class images).

2.2 Maximum Average Correlation Height (MACH)

MACH filters are obtained by minimizing the variance of the correlation plane between the training images while maximizing the average correlation output at the origin. These filters are found to be statistically optimally for at least Gaussian noise assumption. As a consequence, MACH filters are generally more distortion tolerant than MACE filters. The expression for MACH filter $\boldsymbol{h}$ in frequency domain is –

$$\boldsymbol{h} = \mathbf{S}^{-1}\mathbf{m} \qquad (2)$$

Where, $\mathbf{m}$ is a column vector containing Fourier coefficients of the mean of the training images and $\mathbf{S}$ is a P×P diagonal matrix which reflects the measure of the variation between the correlation planes given as –

$$\mathbf{S} = \frac{1}{\mathrm{N.P}}\sum_{i=1}^{N}(\mathbf{I}_i - \mathbf{M})(\mathbf{I}_i - \mathbf{M})^{*} \qquad (3)$$

I_i and M are P×P diagonal matrices which contain the Fourier coefficients of i^{th} training image and mean of the training images along the diagonal.

2.3 Wavelet Fusion

As discussed already, the objective of image fusion is to combine the complimentary information present in the IR and visible spectrum to produce a new image that contains the information content of both the spectrums Wavelet decomposition, i.e. Multi-resolution analysis is a popular method for image fusion because of these reasons - Firstly, it provides the opportunity to combine the information at different resolution levels. Secondly, there is a plethora of different wavelets having different properties for information extraction from the images. Thirdly, wavelet analysis provides a way to combine the information in the form of extracted coefficients at different resolutions which ensures better fusion result. A comprehensive tutorial for image fusion using wavelets is given in [8].

The most important step in image fusion based on wavelets is the process of coefficient combination; namely, the process of merger of coefficients in an appropriate way to obtain the desired qualities in the fused image. There are a lot of strategies for coefficient fusion, each having different properties [8]. In this work, simple averaging of the coefficients is done because it is desired to have the information content of both the spectrums in the fused image.

3. Proposed Methodology

As already mentioned our aim is to facilitate multimodal face recognition using wavelet based Multi-resolution image fusion approach. The proposed methodology can be divided primarily into three major steps (Fig. 1) -

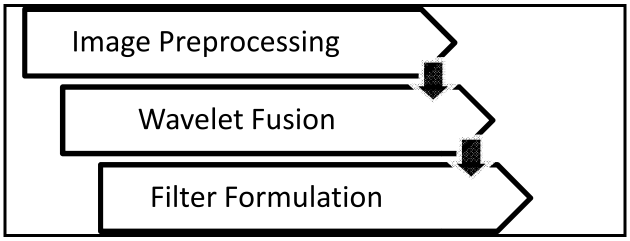

Fig 1 Proposed methodology

3.1 Image Preprocessing:

While carrying multi-resolution wavelet fusion it is required that both images must have their respective features aligned properly. However facial images obtained from different imaging modalities (like CCD and IR) usually differ drastically in terms of corresponding feature position. In order to eliminate these differences it is required to register respective images. 8 features points were selected in both IR and respective CCD image and '*affine*' transform was used to align corresponding facial features. Once the corresponding facial features are aligned, images are ready for wavelet fusion.

3.2 Wavelet Fusion:

This can be further subdivided into two steps -

a) **Wavelet Decomposition**

Multi-level 2D wavelet decomposition is carried out using a suitable wavelet basis function. For experimentation purpose we have used **db4** wavelet. One level wavelet decomposition, decomposes image into four frequency bands: low–low (LL), low–high (LH), high –low (HL) and high–high (HH). Thus, a DWT with N decomposition levels will have M = 3N + 1 such frequency bands. Fig. 2 shows 5 level decomposition of image using wavelet basis **db4**. Images were decomposed to 5 level and respective decomposition frequency bands were fused together. It was observed during experimentation that lower level decomposition (less than 5) introduces blocking effects and distortion around the edges while fusion. It has also been reported in [8] that higher level of decomposition leads to deterioration in fused image quality. Consequently, a 5 level decomposition is used in our experimentation for wavelet fusion.

b) **Merging DWT coefficients**

After decomposition of both images into DWT coefficients it is required to merge the coefficients in such a manner that information of both imaging modalities can be extracted. Fusion of coefficients is allowed only at same resolution level. As already stated, it is required to have the information present in both the spectrums. Hence, simple DWT coefficient averaging scheme is used to fuse the DWT coefficients so DWT coefficients of fused image are average of the DWT coefficients of IR and CCD face images. Then inverse Wavelet transform is applied to get fused image. Last two rows i.e. 4-6 in Fig. 4 show some of the fused images of a subject.

3.3 Hybrid Correlation Filter Formulation

Hybrid Correlation filters termed as HMACE and HMACH were obtained using the fused images as discussed in section 2. Here, "H" stands for "Hybrid" which indicates that the hybrid filters contain the information present in both the spectrums. Formulated hybrid filters are then tested for real and imposter images. Fig. 3 demonstrates the overall process. The classification is done on the basis of PSR value [5] of the peak in output correlation plane.

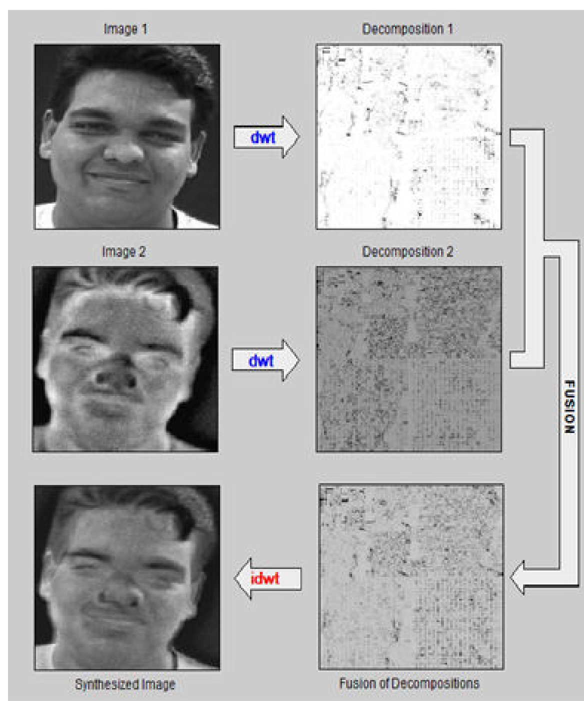

Fig 2 Five level wavelet fusion using MATLAB.

4. Experimentation

To access the validity and efficiency of proposed algorithm, we have extensively tested the proposed algorithm on IRIS Thermal/Visible face dataset [10]. HMACE and HMACH filters corresponding to 10 different subjects (refer Fig 5) were created. For each subject, 18 IR and 18 CCD images (including pose and expression) were used to create 18 fused images. These fused images were zero padded to 512 × 512 dimensions to ensure proper correlation. Finally using these padded images, correlation filters were formulated. Fig 4 shows fused images generated from CCD and IR images for one subject from the figure it is clear that the visual information from both the spectrum is present in fused images.

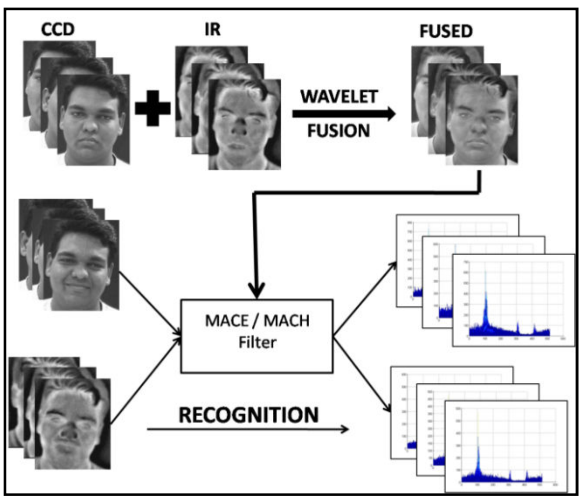

Fig 3 Flow diagram of proposed methodology.

Fig 4 Rows 1-2 are CCD images, 2-4 are IR images and 4-6 are fused images.

Similarly, image fusion was carried out for all of the 10 subjects to formulate 10 different HMACE and HMACH filter. Each filter was tested against 40 true class images and 40 false class images. Corresponding PSR value plots were drawn and accordingly suitable thresholds were selected. Correct classification corresponds to the case when the PSR value of the test image exceeds the pre-calculated threshold. Percentage accuracy was calculated as the ratio of correct classification to the total number of test images for a particular subject.

One of the most important and useful characteristic of correlation filter is that they give acceptable recognition performance even with occluded images. In order to verify this property we tested the proposed hybrid filters against 25% occluded true class images (Fig 6). Accuracy corresponding to this experimentation has also been reported in the table as well as the PSR plot. Fig 7 and Fig 8 show the PSR value plot of 40 true class and 40 impostor images with and without occlusion when tested against the HMACE and HMACH filter for subject 1.

Table I and Table II report the percentage accuracy acquired by 10 different HMACE and HMACH filters corresponding to 10 different subjects tested against 40 true class images and 40 false class images with and without occlusion. The results demonstrate the excellent performance of proposed HMACE and HMACH filters. It should be noted that performance of proposed filters degraded slightly with occlusion showing its superiority over feature based methods in case of occlusion. Moreover, use of single filter for CCD and IR image has also resulted in reduction of requirement memory space by 50%. This reduction in storage space is especially important when a field deployable unit is to be constructed in which limited memory space is available. The reported accuracy clearly indicates that image fusion has resulted in efficient hybrid correlation filter formulation for simultaneous IR and CCD image recognition using a single filter.

Fig 5 Ten different Test subjects

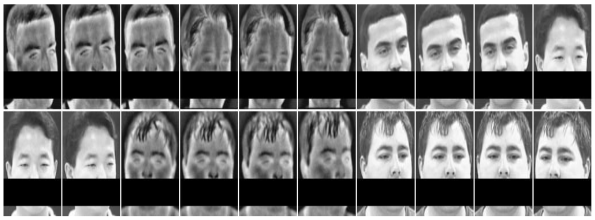

Fig 6 Occluded images of different subjects.

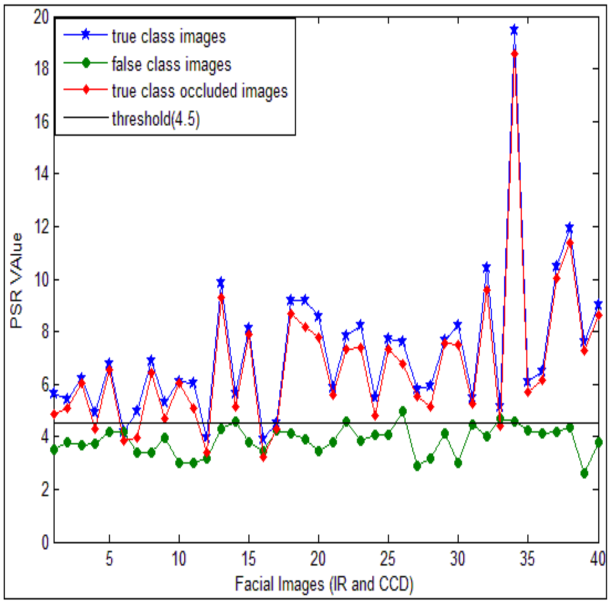

Fig 7 PSR Value plot for HMACE filter (Subject 1)

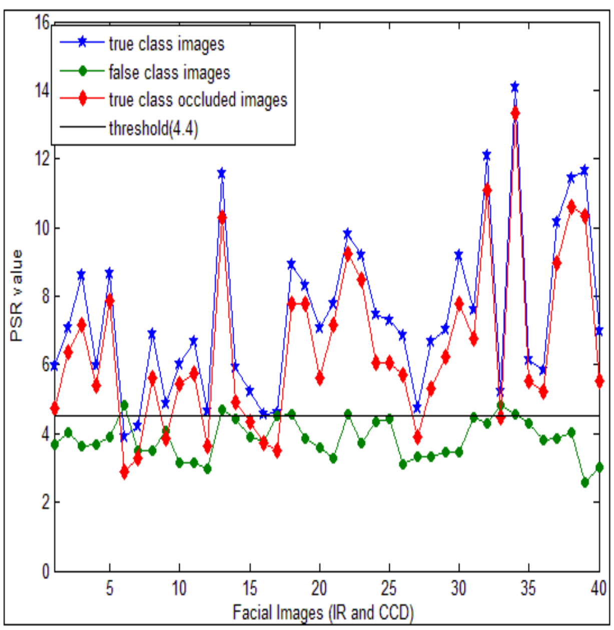

Fig 8 PSR Value plot for HMACH filter (Subject 1)

Table I: Recognition accuracy by HMACE filter

Filter	Threshold	Accuracy (Without Occlusion)	Accuracy (With Occlusion)
HMACE 1	4.5	91.25	83.75
HMACE 2	4.3	90.00	85.00
HMACE 3	4.45	92.50	82.50
HMACE 4	4.25	87.50	81.25
HMACE 5	4.6	92.50	83.75
HMACE 6	4.15	91.25	81.25
HMACE 7	4.7	93.75	85.00
HMACE 8	4.05	86.25	80.00
HMACE 9	4.35	87.50	82.50
HMACE 10	4.6	92.50	80.00
Average	-	**90.5**	**82.5**

Table II: Recognition accuracy by HMACH filter

Filter	Threshold	Accuracy (Without Occlusion)	Accuracy (With Occlusion)
HMACH 1	4.4	87.50	80.00
HMACH 2	4.5	88.75	82.50
HMACH 3	4.1	90.00	80.00
HMACH 4	4.50	86.25	77.50
HMACH 5	4.5	85.00	78.75
HMACH 6	4.0	87.50	81.25
HMACH 7	4.5	92.50	83.25
HMACH 8	4.2	86.25	80.00
HMACH 9	4.6	85.00	77.50
HMACH 10	4.4	91.25	76.25
Average	**-**	**88**	**79.7**

5. Conclusion

In this article a novel approach for multimodal facial recognition in IR and Visible light domain based on the formulation of hybrid correlation filter is proposed. The proposed methodology requires single a correlation filter for recognizing both IR as well as CCD images. Proposed HMACE and HMACH filters were extensively tested on both IR and CCD images under sufficient pose and expression variation. Test set contained both true class and false class image with and without occlusion. The robustness of filter is demonstrated by its satisfactory performance even in case of occluded images. The reported results indicate the usability of the proposed algorithm in field of multimodal face recognition. Moreover, the proposed methodology results in reduction of 50% memory space requirement since a single filter is now required for recognition of IR as well as CCD images.

6. Future Work

In the future work we are planning to use Machine learning tools to formulate a more suitable discriminating quantity (like PSR used in the article) which can enhance the performance of formulated hybrid filters.

7. References

[1] W. Zhao, R. Chellappa, A. Rosenfeld, "Face Recognition: A Literature Survey," *ACM Computing Surveys*, 35(4), 2003, 399–458.

[2] A. Mahalanobis, B. V. K. Vijaya Kumar, D. Casasent, ''Minimum average correlation energy filters,'' *Applied Optics, 26, 1987, 3633–3640*

[3] A.Mahalanobis, B.V.K. VijayaKumar, S. Song, S.R.F. Sims, J.F. Epperson, "Unconstrained correlation filters," *Applied Optics*," 33, 1994, 3751–3759

[4] B. V. K. Vijaya Kumar, ''Tutorial survey of composite filter designs for optical correlators,'' *Applied Optics, 31, 1992, 4773–4801*

[5] Abhijit Mahalanobis , B. V. K. Vijaya Kumar, "Correlation Pattern Recognition" *Cambridge University Press, February 2006*

[6] Martin David Levine, Yingfeng Yu, "Face recognition subject to variations in facial expression, illumination and pose using correlation filters," *Computer Vision and Image Understanding* 104, 2006, 1–15.

[7] J. Heo, M. Savvides, B.V.K. Vijayakumar, Performance evaluation of face recognition using visual and thermal imagery with advanced correlation filters," in proc. *IEEE CVPR,* 2005.

[8] G. Pajares, J. M. de la Cruz, **"**A wavelet-based image fusion tutorial**",** *Pattern Recognition Society, Elsevier B.V., 2004.*

[9] Heng Ma1, Chuanying Jia, and Shuang Liu, "Multisource Image Fusion Based on Wavelet Transform", *International Journal of Information Technology, Vol. 11, No. 7, 2005*

[10] IEEE OTCBVS WS Series Bench; DOE University Research Program in Robotics under grant DOE-DE-FG02-86NE37968; DOD/TACOM/NAC/ARC Program under grant R01-1344-18; FAA/NSSA grant R01-1344-48/49; Office of Naval Research under grant #N000143010022.