Seeing People in the Dark: Face Recognition in Infrared Images

Gil Friedrich and Yehezkel Yeshurun

School of Computer Science , Tel-Aviv University, Israel

Abstract. An IR image of the human face presents its unique heat-signature and can be used for recognition. The characteristics of IR images maintain advantages over visible light images, and can be used to improve algorithms of human face recognition in several aspects. IR images are obviously invariant under extreme lighting conditions (including complete darkness). The main findings of this research are that IR face images are less effected by changes of pose or facial expression and enable a simple method for detection of facial features. In this paper we explore several aspects of face recognition in IR images. First, we compare the effect of varying environment conditions over IR and visible light images through a case study. Finally, we propose a method for automatic face recognition in IR images, through which we use a preprocessing algorithm for detecting facial elements, and show the applicability of commonly used face recognition methods in the visible light domain.

1. Introduction

State-of-the-art algorithms for face recognition in the visible light domain achieve a remarkable high level of recognition under controlled environmental conditions. However, these algorithms perform rather poorly under variable illumination, position and facial expression. On the other hand, all known computer algorithms developed for face recognition seem to significantly deteriorate in their performance even when minor alterations to basic conditions are imposed (Ref. [4]), in sharp contrast to the phenomenal ability of the human brain to correctly identify human faces (Ref. [3]) under many variations and even after long periods of time or when the face is partially disguised.

While most existing Computer Vision algorithms are naturally based on visible light, heat signature is vastly used in the animal kingdom for various tasks. We have thus set out to explore the potential benefits of using Infra Red (IR) data for object detection and recognition, and specifically for face detection and recognition. Our IR camera was sensitive to 5-12 nm, as in general body heat temperature is around 9.2 nm.

The main finding of this report is that IR based face recognition is more invariant than CCD based one under various conditions, specifically varying head 3D orientation and facial expressions.

The authors would like to thank the AMN fund, Center of excellence for applied Geometry and the Ministry of Science for partial funding of this research

Modifying both facial expressions and head orientation cause direct 3D structural changes, as well as changes of shadow contours in CCD images, which deteriorate the accuracy of any classification method. In an IR image this effect is greatly reduced.. In the following we demonstrate this finding through a case study of two faces, and then proceed to show the benefits for a full fledged face recognition algorithm for a set of 40 persons.

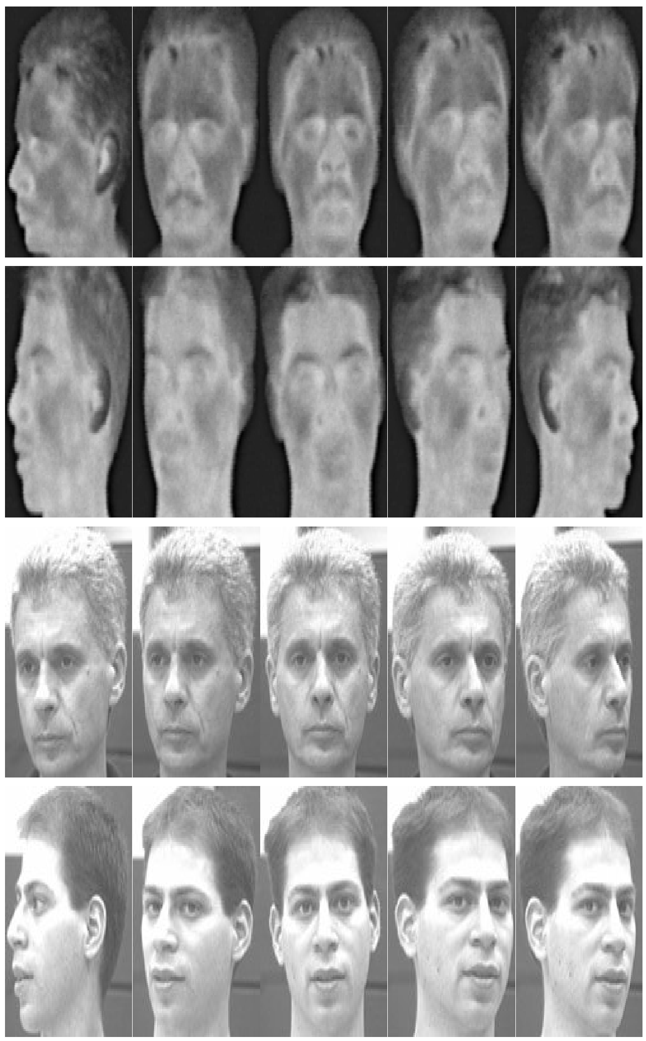

Figure 1 Examples of IR (top) and CCD (bottom)images with head orientation variations

2 IR vs. CCD Image Invariance

2.1 Head Orientation

3D position variations naturally give rise to 2D image variations. The hypothesis we were making, based on the type of observation depicted in Fig. 1, is that the effect of 3D head rotation in IR images might be less prominent.
3D position variations naturally give rise to 2D image variations. The hypothesis we were making, based on the type of observation depicted in Fig. 1, is that the effect of 3D head rotation in IR images might be less prominent.
This section tests this hypothesis quantitatively over the presented test case. The comparison is done by comparing the ratio of Euclidean distances between images of the same class and the distances of images from the two different classes. First, we align all images using two manually selected interest points (later we present an automatic method to accomplish this task in IR images). Then the average face is found and subtracted from each image. Finally, the matrix of euclidean distances between every pair is calculated. In order to rule out the possibility that the mere difference between the IR and CCD images is the basic resolution, we have repeated the test with CCD images blurred to attain the same resolution of the IR images.
Let $\bar{d}_{In}$ and $\bar{d}_{Ext}$ be the averages of Euclidean distances within and between the classes respectively. Table 1 summarizes the ratio $\frac{\bar{d}_{In}}{\bar{d}_{Ext}}$.

	Face 1	Face 2
CCD	0.522	0.751
Blurred CCD	0.521	0.738
IR	0.384	0.532

Table1. Ratio of Euclidian distance within and between classes (Figure1 images)

This result clearly indicates that IR face images are less effected by face 3D orientation in comparison to CCD images. Blurring the CCD images indeed made a small difference, but clearly did not bring the CCD corresponding within/between ratio to what is achieved for IR images.

2.2 Facial Expression

As was the case for head orientation, based on observation similar to Fig.2, we have set out to examine the relative effect of varying facial expressions on their corresponding IR images.

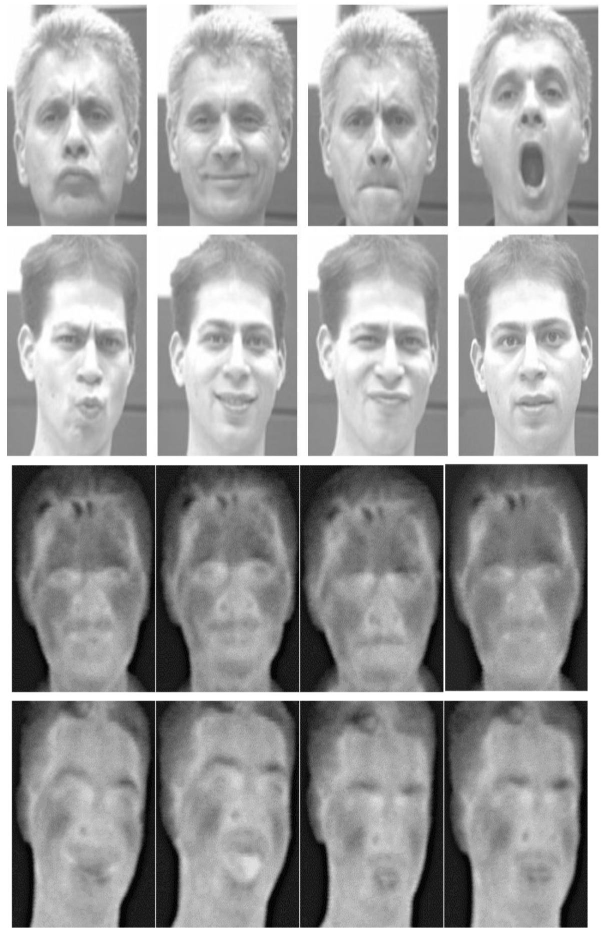

Figure 2 Examples of CCD (top) images and IR (bottom) with facial expressions variations

	Face 1	**Face 2**
CCD	0.817	0.518
Blurred CCD	0.763	0.511
IR	0.350	0.408

Table 2 - Ratio of Euclidian distance within and between classes (Figure2 images)

Table 2 summarizes the comparison results, which show that indeed facial expression variations are less prominent in IR images.
Several explanations to these results can be found by analyzing the type of information revealed in IR images with comparison to CCD images.

First, changing face parameters modifies the skin surface, thus creating additional image contours. By observing the CCD images, one can see that several parts of the human's face change their shadow pattern, as the expression is altered. Also important is the fact that in the IR image several elements on the face's skin, like pigmentation, are hidden, since they are exactly of the same temperature as the rest of the skin. As a result the skin in the IR image creates a more unified pattern. When the face's expression changes the rather unified pattern in the IR image creates only a slight change between two images, whereas the rich pattern in the CCD image increases the distinction between the two images, reducing the possibility of correct identification. The trivial explanation that IR based images have a better within/between ratio is just an artifact that reflects the fact that IR images “blur” images by definition (due to a more uniform heat distribution than visibly patterns distribution) is ruled out by using blurred CCD images.
In conclusion, face images taken in the IR range inherently contain characteristics that improve correct identification under varying conditions.

3 DETECTION AND RECOGNITION ALGORITHMS

Having obtained the basic finding contributing to facial image invariance, we have implemented an algorithm for human face recognition of infrared images.
Forty face images were used in this study. Each face was captured in 10 different images. All faces were frontal, slightly varying in angle, face size (distance from camera's lens) and expression. The images of each face were divided into a training sample set (2-3 images per face) and the testing sample set. All the images were preprocessed to obtain automatic segmentation, alignment and normalization. Finally, the processed faces are used for an Eigenfaces (Ref. [1]) based classification. We emphasize that the main motivation behind the implementation is the relative lack of data regarding the characteristics of IR face images and the validation of our IR specific segmentation and detection approach, and thus we have chosen to use a common recognition methodology.

3.1 Segmenting Background Temperature

A crucial initial stage in any face processing system involves segmentation of target (faces) and background. Looking for IR specific characteristics, we have found out that the typical distribution of gray (heat) level associated with faces, is markedly different from the typical distribution of the background heat levels. This might be related to some of he human body heat distribution, as opposed to the relatively narrow range of temperature found at “room temperature”, and was found to be quite consistent under various environmental circumstances.
This observation gives rise to the following background removal algorithm:

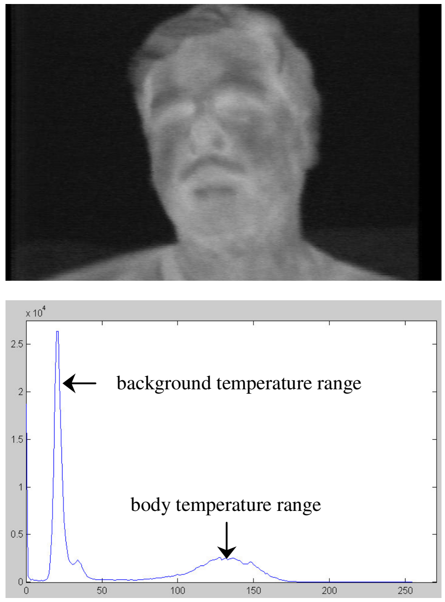

Figure 3. Gray-level Histogram of IR face image

Let Nbody, Nback be the number of pixels and Maxbody, Maxback be the respective maximal value in the histogram, then,

$$\frac{N_{body}}{N_{back}} \approx \frac{1}{2} \quad \text{but} \quad \frac{Max_{body}}{Max_{back}} \approx \frac{1}{10}$$

(Although the background covers an area only 2 times larger than the rest of the image, its heat is presented in the histogram by a maximum 10 times larger then the maximum created by face heat). The heuristics works over the image histogram in the following way. Let

$$\{h_i\}_{i=0}^{255}$$

be the bin values in the gray-level histogram (256 possible values). Then,

$$Max, j_{\max} = Max\{h_i\}_{i=1}^{255} \tag{1}$$

and,

$$Med, j_{med} = Median(\{h_i > 0\}) \tag{2}$$

are the value and index of maximal and median respectively (for the median value we omit zero bins). Then we seek the first bin after jmax that has value smaller then median value,

$$k = \min_j \{ j > j_{\max}, h_j < Med \} \tag{3}$$

Now, bins {*0...k-1*} represent the background temperature gray-level values and are set to 0, thus completing the stage of background segmentation and removal. Notice that it is possible that objects other than faces will be left as well after the background removal, due to their heat distribution. This, however, is being taken care of in a subsequent stage.

3.2 Clustering

Background removal leaves areas possessing certain intensity distribution levels intact. These areas consist of faces and potentially other objects. Locating and clustering the image pixel corresponding to faces is the next stage of the algorithm. This goal, which is far from trivial with regular CCD images (Ref. [15], [18]), could be much simpler for IR images. We carry this out by using a narrowly tuned heat filtering (Ref. [17]), followed by removal of pixels with only a single 8-connected neighbor. Since sometimes facial elements are colder then the lower limit of the body temperature (brows, nose etc.), an additional flood-fill operation is performed. The algorithm takes a binary image *I* and starts 4-connected flooding from frame boundary pixels. At the end, flooded pixels J include all non-holes background pixels, therefore the new image Inew=(~J | I) is the original image with all the holes filled.

Figure 4 demonstrate the process.

A= original image

$$B_{ij} = \begin{cases} 1 & h_{low} < A_{ij} < h_{high} \\ 0 & elsewhere \end{cases}$$

J= FloodFill(B)
$C_{ij}=(\sim J_{ij} \mid B_{ij})$

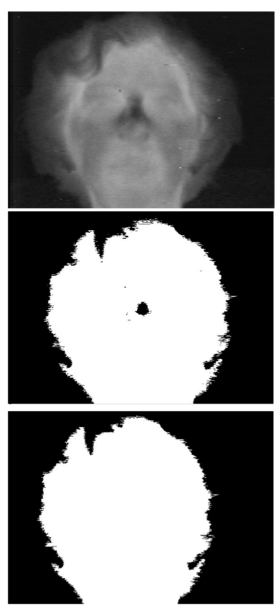

Figure 4. Flood-filling holes in a cluster

We finalize the clustering stage by filtering out clusters that are not elliptical, using the following simple metric. Let ,

$$e = (A, B, C, D, E) \tag{1}$$

be the ellipse coefficients and

$$r = (X^2, Y^2, XY, X, Y) \tag{2}$$

is calculated from the set of edge-point coordinates $\{X\},\{Y\}$. Ellipse parameters e should satisfy:

$$\left\| e \cdot r' - \vec{1} \right\| \longrightarrow Min \tag{3}$$

It follows that

$$e = (\vec{1} \cdot r) \cdot (r' \cdot r)^{-1} \tag{4}$$

From which we derive the radii $\{r_x, r_y\}$ and rotation angle θ. On our sample data, we have found the following thresholds (5) to be most useful:

$$-30^\circ \le \theta \le 30^\circ \quad 1 \le \frac{r_y}{r_x} \le 5 \tag{5}$$

In summation, this stage serves for two purposes. First to horizontally align the image, which is a significant for next section. Secondly, it segments the head image from the full body image, by heat filtering (which removes anything that is not skin surface including clothing) and then by elliptic fitting that removes parts of neck and shoulders).

3.3 Finding points of reference

There are quite a number of approaches in the face recognition literature for detection reference points in images (e.g eyes). We were looking for specific IR face image features that are characteristic and could be used in general. From comparing a vast database of IR images it was possible to conclude the following:

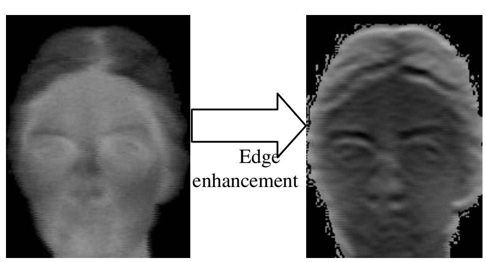

Figure 5. Edge enhancement used for emphasizing the forehead-brows-eyes pattern

Since the eyes are always hotter then their surroundings (even when the eyes are closed!) and the brows were significantly colder then their surroundings, it turned out the most prominent IR related feature that is robust over all the sample data we have used is the edge created by the difference between the eyes and the brows temperature. In order to utilize this property, we have applied edge enhancement (Fig 5), followed by thresholding (lower 1%) (Fig 6).

Figure 6. Lowest 1% values leaves mainly the brows within the image

And using an horizontal histogram to detect the brows (Figure 7).

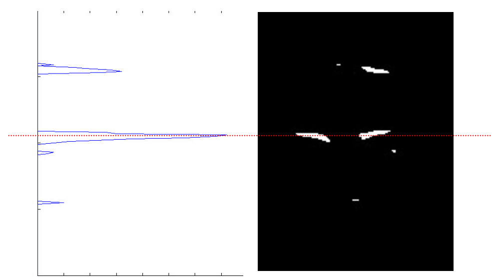

Figure 7. Horizontal histogram on low-values image used to localize the brows

Finally, by finding the average point of each of the two clusters, we get the two center points of the image. This is illustrated over the original face in Figure 8.

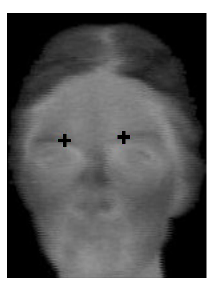

Figure 8. Points of interest

This method had proved to work properly for all our sample data. We finally align the face image using bilinear interpolation.

3.4 Face Recognition

The previous sections had shown that IR face images are less effected by specific factors such as face orientation and expression, and suggested methods for face detection and normalization. Still, the ability to correctly classify IR face images over a large number of faces needs to be proved. The main point that should be looked for, is whether the invariance gained using IR images and any other potential IR specific features, like face-specific heat spatial distribution compensate for the loss of specific visual information, like visible edges, skin patterns and texture and so on.

We have addressed this issue by implementing an Eigenfaces based classification algorithm (See [1], [2], [4] and [9], [10]). By comparing its performance over IR images with typical CCD based algorithms, we expected to quantify the relative amount of face-specific information within IR images.

The database we have used consisted of 40 different faces, under various head positions and facial expressions (see Fig 1 and 2 for typical examples). For each face 2-4 images were included in the training sample set (total of 96) images, and 5-10 images were tested for recognition quality (total of 250 images). In order to select the most significant Eigenfaces, a threshold of 99% was selected. This threshold yielded a set of 78 Eigenfaces. Some of the Eigenfaces used are shown in Fig. 9.

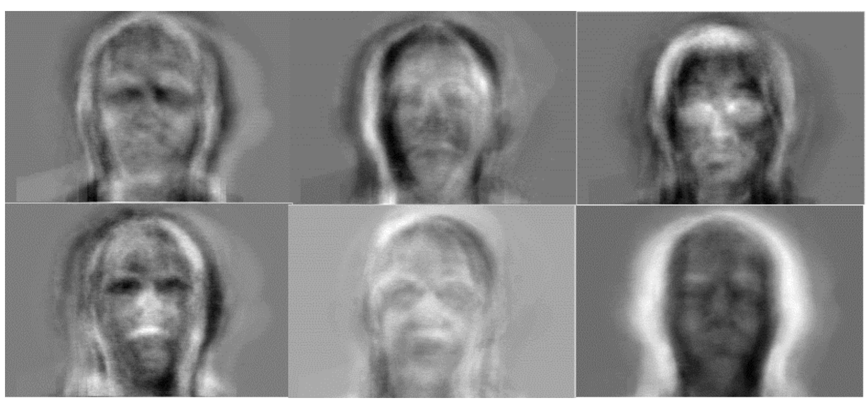

Figure 9. Examples of IR eigenfaces

Failing to recognize an image consisted of one of the following:

1. **Non-Face.** (Threshold $t_1=1.0$). [1]
2. **Reject.** (Threshold $t_2=0.10$). [2]
3. **Misclassify.**

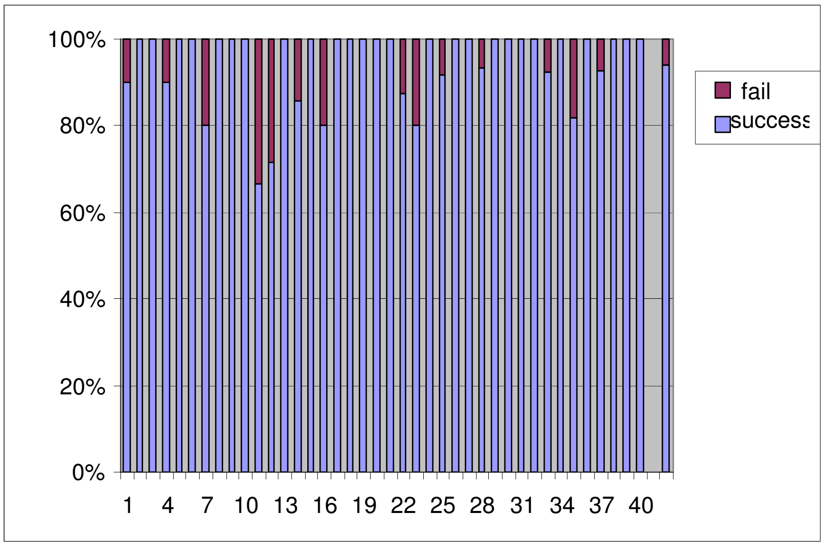

Figure 10. Recognition results by face indexes

Figure 10 plots success rates for each of the 40 faces used, and the total success rate, which is 94%, in the last column. This rate is comparable to the typical recognition rate reported for CCD based systems (94%-98%), and is clearly superior to the typical recognition rate of CCD based systems where the sample, like in our case, consists of highly varying pose and facial expressions (around 80%). Out of the 40 faces tested, In 26 all face instances were correctly identified, in 10 faces one image classification failed, and in 4 faces two images classification failed. In all misclassification errors, the 2nd and 3rd candidates were the right ones, meaning that a more sophisticated classification algorithm (which is not the main goal of this paper) might improve the recognition results even further.

---

1 Let d_0 be the maximal distance of a training sample set image to the average image, and d be the distance of the tested image, then if $d>(1+t_1)d_0$ then the image is a non-face.

2 Let ε be the minimal distance of the tested image with any image from the training sample set. Then if $\varepsilon>t_2$ it is categorized as an unfamiliar face.

4. CONCLUSIONS

IR based face processing methods have recently gained more attention ([19], [20]) and tested with standard face recognition algorithms ([9], [10]). This field of research still calls for further analysis in order to make optimal use of IR specific features. Some of the advantages of using IR images for face recognition (like invariance under illuminating conditions) are self evident, and some others (automatic detection of faces based on heat signature) are less self evident but natural.
However, IR specific research will undoubtedly yield new features and methods that could immensely facilitate the task, comparing to CCD based methods. In this regard, the main finding of this research is that IR specific feature are significant in eliminating some pose and facial expression variance, thus increasing significantly the typical performance of face recognition algorithms over highly variable pose and expression face images. Future research along the same lines might involve similar analysis for time varying and aging effects, where IR images might prove useful.

References

1. M. Turk and A. Pentland (1991). Eigenfaces for recognition, *J. Cog. Neuroscience*, vol. 3, no. 1, pp. 71-86.
2. M. Kirby and L. Sirovich (1990). Application of the karhunen-loeve procedure for the characterization of human faces, *IEEE Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103-108.
3. R. Chellappa, C. Wilson, and S. Sirohev (1995). *Human and machine recognition of faces: A survey*, in Proceedings of IEEE, May 1995, vol. 83, pp. 705-740.
4. P. Phillips, H. Wechsler, J. Huang, and P. Rauss (1998). *The FERET database and evaluation procedure for face recognition algorithms,* Image and Vision Computing, vol. 16, no. 5, pp. 295-306.
5. L. Wiskott, J-M. Fellous, N. Kruger, and C. von der Malsburg (1997). *Face recognition by elastic bunch graph matching*, IEEE Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 775-779.
6. K. Etemad and R. Chellappa (1997). *Discriminant analysis for recognition of human face images*, Journal of the Optical Society of America, vol. 14, pp. 1724-1733.
7. B. Moghaddam and A. Pentland (1997). *Probabalistic visual recognition for object recognition,* IEEE Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 696-710.
8. M. Weiser (1991). *The computer for the 21st century,* Scientific American, vol. 265, no. 3, pp. 66-76.
9. R. Cutler (1996). *Face recognition using infrared images and eigenfaces.* http://research.microsoft.com/~rcutler/face/face.htm.
10. J. Wilder, P.J. Phillips, C. Jiang, and S. Wiener. *Comparison of Visible and Infra-Red Imagery for Face Recognition,* Proceedings of the 2nd International Conference on Automatic Face and Gesture Recognition, Killington, VT, pp.182-187, October 1996
11. F. Prokoski. *History, Current Status, and Future of Infrared Identification*, IEEE Workshop on Computer Vision behind the Visible Spectrum: Methods and Applications (CVBVS 2000), pp 5-14.
12. D. Reisfeld and Y. Yeshurun. *Preprocessing of Face Images: Detection of Features and Pose Normalization*, Computer Vision and Image Understanding, Vol. 71 No. 3 pp 413-430, Sep 1998.

13. M. Irani and P. Anandan. *Robust Multi-Sensor Image Alignment*, International Conference on Computer Vision , Mumbai, January *1998.*
14. N. Intrator, D. Reisfeld, Y. Yeshurun**.** *Face Recognition using a Hybrid Supervised/Unsupervised Neural Network*, Pattern Recognition Letters 17:67-76, 1996.
15. Zhang, Y.J. *Evaluation and Comparison of Different Segmentation Algorithms*, PRL(18), No. 10, October 1997, pp. 963-974.
16. Zhou, Y.T, Venkateswar, V. and Chellappa R. *Edge Detection and Linear Feature Extraction Using a 2-D Random Field Model*, PAMI(11), No. 1, January 1989, pp. 84-95.
17. Harley, R.L., Wang, C.Y, Kitchen, L. and Rosenfeld, A. *Segmentation of FLIR Images: A Comparative Study*, SMC(12), No. 4, July/August 1982, pp. 553-566, or DARPA82 (323-341).
18. John F. Haddon and James. F. Boyce. *Image Segmentation by Unifying Region and Boundary Information*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(10), pp. 929-948, October, 1990.
19. J. Dowdall, I. Pavlidis, and G. Bebis. *A Face Detetcion Method Based on Multi-Band Feature Extraction in Near IR Spectrum*, IEEE Workshop on Computer Vision Beyond the Visible Spectrum. Kauai, December 2001.
20. Socolinsky, D., Wolff, L., Neuheisel, J., and Eveland, C. *Illumination Invariant Face Recognition Using Thermal Infrared Imagery*, Computer Vision and Pattern Recognition, Kauai, December 2001.