Multi-Sensory Face Biometric Fusion (for Personal Identification)

Paper #: 20

Abstract

The objective of this work is to recognize faces using sets of images in visual and thermal spectra. This is challenging because the former is greatly affected by illumination changes, while the latter frequently contains occlusions due to eye-wear and is inherently less discriminative. Our method is based on a fusion of the two modalities. Specifically: we examine (i) the effects of preprocessing of data in each domain, (ii) the fusion of holistic and local facial appearance, and (iii) propose an algorithm for combining the similarity scores in visual and thermal spectra in the presence of prescription glasses and significant pose variations, using a small number of training images (5-7). Our system achieved a high correct identification rate of 97% on a freely available test set of 29 individuals and extreme illumination changes.

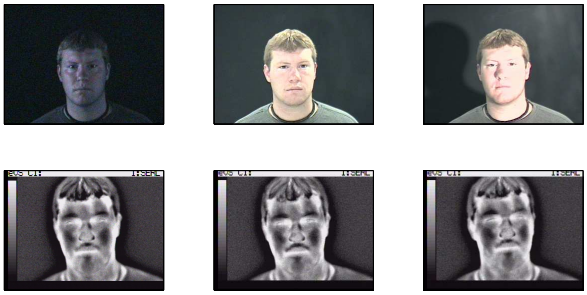


Figure 1: ***Invariance:*** *Illumination changes have a dramatic effects on images acquired in the visible light spectrum (top row). In contrast, thermal imagery (bottom row) shows remarkable invariance.*

1 Introduction

Variations in head pose and illumination are the most challenging aspects of face recognition. In practice, the effects of changing pose are usually less problematic and can oftentimes be overcome by acquiring data over a time period e.g. by tracking a face. Consequently, image sequence or image set matching has recently gained a lot of attention in the literature [2] [11] [29] and is the paradigm adopted in this paper as well. In contrast, illumination is much more difficult to deal with: the illumination setup is in most cases not possible to control and its physics difficult to accurately model.

Thermal spectrum imagery is useful in this regard as it is virtually insensitive to illumination changes, see Fig. 1. On the other hand, it lacks much of the individual, discriminating facial detail contained in visual images. In this sense, the two modalities can be seen as complementing each other. The key idea behind the system presented in this paper is that robustness to extreme illumination changes can be achieved by *fusing* the two. This paradigm will further prove useful when which we consider the difficulty of recognition in the presence of prescription glasses.

1.1 Mono-sensor based techniques

Optical sensors. Among the most sensors used in face biometric systems is the optical imager. This is driven by its availability and low-cost. An optical imager captures the light reflectance of the face surface in the visible spectrum. The visible spectrum provides features that depend only on surface reflectance. Thus, it is obvious that the face appearance changes according to the ambient light. In order to overcome the lighting, pose and facial expression changes, a flurry of face recognition algorithms, from the two well-known broad categories, appearance-based and feature-based methods, has been proposed [25]. Appearance-based methods find the global properties of the face pattern and recognize the face as a whole. In contrast, feature-based methods [24] [21] [12] explore the statistical and geometrical properties of facial features like eyes and mouth. The face recognition performance depends on the accuracy of facial feature detection. Moreover, local and global lighting changes cause existing face recognition techniques for the visible imagery to perform poorly.

Infrared sensors. Recent studies have proved that face recognition in the thermal spectrum offers a few distinct advantages over the visible spectrum, including invariance to ambient illumination changes [38] [32] [26] [31]. This is due to the fact that a thermal infrared sensor measures the heat energy radiation emitted by the face rather than the light reflectance. A thermal sensor generates imaging features that uncover thermal characteristics of the face pattern. Indeed, thermal face recognition algorithms attempt to take advantage of such anatomical information of the hu-

man face as unique signatures.

Appearance-based face recognition algorithms applied to thermal IR imaging consistently performed better than when applied to visible imagery, under various lighting conditions and facial expressions [22] [30] [32] [28]. Further performance improvements were achieved using decision-based fusion [32]. In contrast to other techniques, Srivastana *et al.* [33], performed face recognition in the space of Bessel function parameters. First, they decompose each infrared face image using Gabor filters. Then, they represent the face by a few parameters by modeling the marginal density of the Gabor filter coefficients using Bessel functions. This approach has been improved by Buddharaju *et al.* [8]. Recently, Friedrich *et al.* [17] shown that IR-based recognition is less sensitive to changes in 3D head pose and facial expression.

1.2 Multi-sensor based techniques

As the surface of the face and its temperature have nothing in common, one would state that the extracted cues from both sensors are not redundant and yet complementary. Several attempts have been made in face recognition based on the fusion of different types of data from multiple sensors. Face recognition algorithms based on the fusion of visible and thermal IR images demonstrated higher performance than individual image types [7] [27] [10] [18]. Biometric systems that integrate face and speech signals [4], the face and fingerprint information [20], and the face and the ear images [9] improved the accuracy in personal identification.

Recently, Heo *et al.* [19] proposed two types of visible and thermal fusion technique, the first fuses low-level data while the second fuses matching outputs. Data fusion was implemented by applying pixel-based weighted averaging of co-registered visual and thermal images. Decision fusion was implemented by combining the matching scores of individual recognition modules. To deal with occlusions caused by eyeglasses in thermal imagery, they used a simple ellipse fitting technique to detect the circle-like eyeglass regions in the IR image and replaced them with an average eye template. Using a commercial face recognition system, FaceIt, they demonstrated improvements in recognition accuracy.

2 Method Details

In the sections that follow we explain our system in detail, the main components of which are conceptually depicted in Fig. 2.

2.1 Matching image sets

In this paper we deal with face recognition from *sets* of images, both in the visual and thermal spectrum. We will show how to achieve illumination invariance using a combination of simple data preprocessing (§2.2), local features (§2.3) and modality fusion (see §2.4). Hence, the requirements for the basic set-matching are that of (i) some pose generalization and (ii) robustness to noise. We compare two image sets by modelling the variations within a set using a linear subspace and comparing two subspaces by finding the most similar modes of variation within them.

The modelling step is a simple application of Principal Component Analysis (PCA) without mean subtraction. In other words, given a data matrix $\mathbf{d}$ (each column representing a rasterized image), the subspace is spanned by the eigenvectors of the matrix $\mathbf{C} = \mathbf{d}\mathbf{d}^T$ corresponding to the largest eigenvalues; we used 5D subspaces.

The similarity of two subspaces U_1 and U_2 is quantified by the cosine of the smallest angle between two vectors confined to them:

$$\rho = \cos\theta = \max_{\mathbf{u}\in U_1} \max_{\mathbf{v}\in U_2} \mathbf{u}^T\mathbf{v}. \quad (1)$$

The quantity ρ is also known as the first canonical correlation. It is this implicit "search" over entire subspaces that achieves linear pose interpolation and extrapolation, by finding the most similar appearances described by the two sets. The robustness of canonical correlations to noise is well detailed in [6].

Further appeal of comparing two subspaces in this manner is contained in its computational efficiency. If $\mathbf{B}_1$ and $\mathbf{B}_2$ are the corresponding orthonormal basis matrices, the computation of ρ can be rapidly performed by finding the largest singular value of the 5×5 matrix $\mathbf{B}_1^T\mathbf{B}_2$ [6].

2.2 Data preprocessing & feature extraction

The first stage of our system involves coarse normalization of pose and brightness. We register all faces, both in the visual and thermal domain, to have the salient facial features aligned. Specifically, we align the eyes and the mouth due to the ease of detection of these features (e.g. see [3] [5] [13] [16] and [35]). The 3 point correspondences, between the detected and the canonical features' locations, uniquely define an affine transformation which is applied to the original image. Faces are then cropped to 80×80 pixels, as shown in Fig. 3.

Coarse brightness normalization is performed by band-pass filtering the images. The aim is to reduce the amount of high-frequency noise as well as extrinsic appearance variations confined to a low-frequency band containing little discriminating information. Most obviously, in visual imagery, the latter are caused by illumination changes [1].

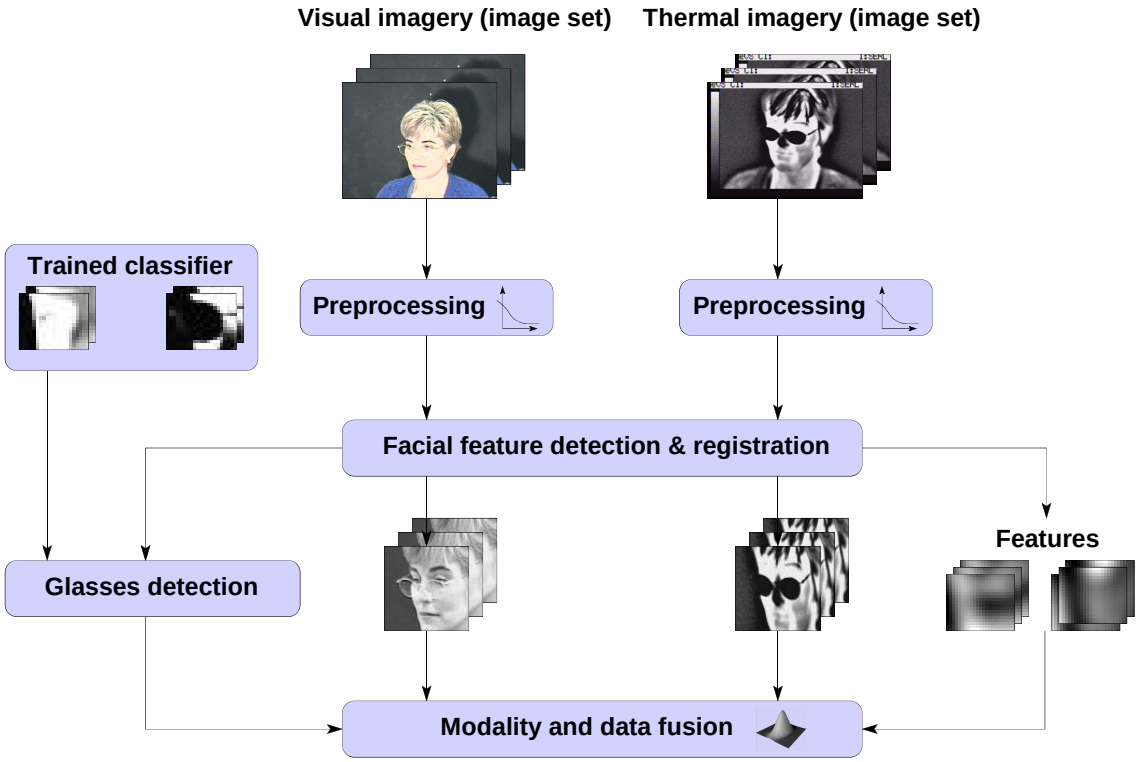


Figure 2: ***System overview:*** *Our system consists of three main modules performing (i) data preprocessing and registration, (ii) glasses detection and (iii) fusion of holistic and local face representations using visual and thermal modalities.*

Figure 3: ***Registration:*** *Shown is the original image in the visual spectrum with detected facial features marked by yellow circles (left), the result of affine warping the image to the canonical frame (centre) and the final registered and cropped facial image.*

We consider the following type of a band-pass filter:

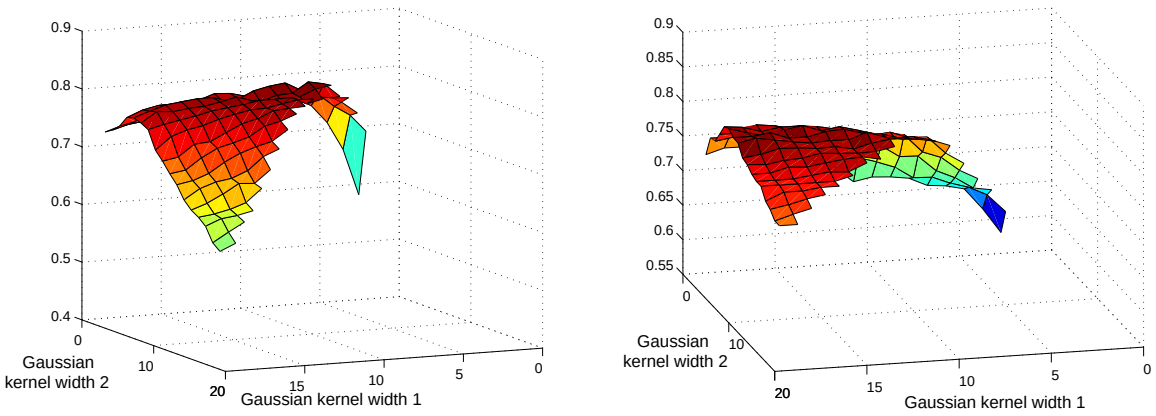
$$\mathbf{I}_F = \mathbf{I} * \mathbf{G}_{\sigma=W_1} - \mathbf{I} * \mathbf{G}_{\sigma=W_2}, \quad (2)$$

which has two parameters - the widths W_1 and W_2 of isotropic Gaussian kernels. These are estimated from a small training corpus of individuals in different illuminations. Fig. 4 shows the recognition rate across the corpus as the values of the two parameters are varied. The optimal values were found to be 2.3 and 6.2 for visual data; the optimal filter for thermal data was found to be a *low-pass* filter with $W_2 = 2.8$ (i.e. W_1 was found to be very large). Examples are shown in Fig. 5. It is important to note from Fig. 4 that the recognition rate varied smoothly with changes in kernel widths, showing that the method is not very sensitive to their exact values, which is suggestive of good generalization to unseen data.

The result of filtering visual data is further scaled by a smooth version of the original image:

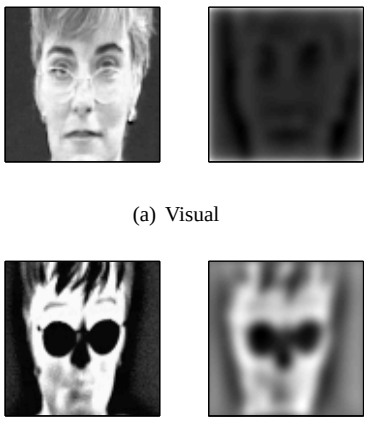
$$\hat{\mathbf{I}}_F(x,y) = \mathbf{I}_F(x,y)./(\mathbf{I} * \mathbf{G}_{\sigma=W_2}), \quad (3)$$

where $./$ represents element-wise division. The purpose of local scaling is to equalize edge strengths in dark (weak edges) and bright (strong edges) regions of the face; this is similar to the Self Quotient Image of Wang *et al.* [37]. This step further improves the robustness of the representation to illumination changes, see §3.



(a) Visual (b) Thermal

Figure 4: ***Band-pass filter:*** *The optimal combination of the lower and upper band-pass filter thresholds is estimated from a small training corpus. The plots show the recognition rate using a single modality, (a) visual and (b) thermal, as a function of the widths W_1 and W_2 of the two Gaussian kernels in* (2)*. It is interesting to note that the optimal band-pass filter for the visual spectrum passes a rather narrow, mid-frequency band, whereas the optimal filter for the thermal spectrum is in fact a* low-pass *filter.*



(a) Visual

(b) Thermal

Figure 5: ***Preprocessing:*** *The effects of the optimal band-pass filters on registered and cropped faces in (a) visual and (b) thermal spectra.*

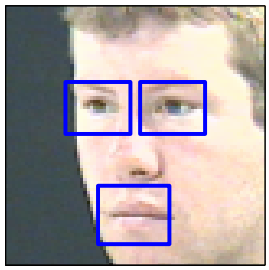


Figure 6: ***Features:*** *In both the visual and the thermal spectrum our algorithm combines the similarities obtained by matching the holistic face appearance and the appearance of three salient local features - the eyes and the mouth.*

2.3 Single modality-based recognition

We compute the similarity of two individuals using only a single modality (visual or thermal) by combining the holistic face representation described in §2.2 and a representation based on local image patches. These have been shown to benefit recognition in the presence of large pose changes [29].

As before, we use the eyes and the mouth as the most discriminative regions, by extracting rectangular patches centred at the detections, see Fig. 6. The overall similarity score is obtained by weighted summation:

$$\rho_{v/t} = \omega_h \cdot \rho_h + \omega_m \cdot \rho_m + (1 - \omega_h - \omega_m) \cdot \rho_e, \quad (4)$$

where ρ_m, ρ_e and ρ_h are the scores of separately matching, respectively, the mouth, the eyes and the entire face regions, and ω_h and ω_m the weighting constants.

The optimal values of the weights were estimated from the offline training corpus. For the visual spectrum we obtained $\omega_e = 0.3$, while the mouth region was found not to improve recognition (i.e. $\omega_m = 0.0$). The relative magnitudes of the weights were found to be different in the ther-

mal spectrum, both the eye and the mouth region contributing equally to the overall score: $\omega_m = 0.1,\ \omega_h = 0.8$.

2.4 Fusing modalities

Until now we have focused on deriving a similarity score between two individuals given sets of images in either thermal or visual spectrum. A combination of holistic and local features was employed in the computation of both. However, the greatest power of our system comes from the fusion of the two modalities.

Given ρ_v and ρ_t, the similarity scores corresponding to visual and thermal data, we compute the joint similarity as:

$$\rho_f = \omega_v(\rho_v) \cdot \rho_v + (1 - \omega_v(\rho_v)) \cdot \rho_t. \quad (5)$$

Notice that the weighting factors are no longer constants, but *functions*. The key idea is that if the visual spectrum match is very good (i.e. ρ_v is close to 1.0), we can be confident that illumination difference between the two images sets compared is mild and well compensated for by the visual spectrum preprocessing of §2.2. In this case, visual spectrum should be given relatively more weight than when the match is bad and the illumination change is likely more drastic.

The function $\omega_v \equiv \omega_v(\rho_v)$ is estimated in three stages: first (i) we estimate $\hat{p}(\omega_v, \rho_v)$, the probability that ω_v is the optimal weighting given the estimated similarity ρ_v, then (ii) compute $\omega^(\rho_v)$ in the maximum a posteriori sense and finally (iii) make an analytic fit to the obtained marginal distribution. Step (i) is challenging and we describe it next.

Iterative density estimate. The principal difficulty of estimating $\hat{p}(\omega_v, \rho_v)$ is of practical nature: in order to obtain an accurate estimate (i.e. a well-sampled distribution), a prohibitively large training database is needed. Instead, we employ a heuristic alternative. Much like before, the estimation is performed using the offline training corpus.

Our algorithm is based on an iterative incremental update of the density, initialized as uniform over the domain $\omega, \rho \in [0, 1]$. We iteratively simulate matching of an unknown person against a set gallery individuals. In each iteration of the algorithm, these are randomly drawn from the offline training database. Since the ground truth identities of all persons in the offline database is known, for each $\omega = k\Delta\omega$ we can compute the separation i.e. the difference between the similarities of the test set and the set corresponding to it in identity, and that between the test set and the most similar set that does *not* correspond to it in identity. Density $\hat{p}(\omega, \rho)$ is then incremented at each $(k\Delta\omega, \rho^{p,p})$ proportionally to $\delta(k\Delta\omega)$ after being passed through the sigmoid function.

Input: visual data $d_v(person, illumination)$,
thermal data $d_t(person, illumination)$.
Output: density estimate $\hat{p}(\omega, \rho_v)$.

1: Init
$\hat{p}(\omega, \rho_v) = 0$,

2: Iteration
for all illuminations $i,\ j$ **and** persons p

3: Iteration
for all $k = 0, \ldots, 1/\Delta\omega,\ \omega = k\Delta\omega$

5: Separation given ω
$\delta(k\Delta\omega) = \min_{q \neq p}[\omega\rho_v^{p,p} + (1-\omega)\rho_t^{p,p} - \omega\rho_v^{p,q} + (1-\omega)\rho_t^{p,q}]$

6: Update density estimate
$\hat{p}(k\Delta\omega, \rho_v^{p,p}) = \hat{p}(k\Delta\omega, \rho_v^{p,p}) + \text{sig}(C \cdot \delta(k\Delta\omega))$

7: Smooth the output
$\hat{p}(\omega, \mu) = \hat{p}(\omega, \mu) * \mathbf{G}_{\sigma=0.05}$

8: Normalize to unit integral
$\hat{p}(\omega, \rho) = \hat{p}(\omega, \rho) / \int_\omega \int_\rho \hat{p}(\omega, \rho) d\rho d\omega$

Figure 7: ***Offline:*** *Optimal fusion training algorithm.*

Fig. 7 summarizes the proposed offline learning algorithm. An analytic fit to $\hat{p}(\omega_v)$ in the form $(1+\exp(a))/(1+\exp(a/\rho_v))$ is shown in Fig. 8.

2.5 Dealing with glasses

The appeal of using the thermal spectrum for face recognition stems mainly from its invariance to illumination changes, in sharp contrast to visual spectrum data. The exact opposite is true in the case of prescription glasses, which appear as dark patches in thermal imagery, see Fig. 5. The practical importance of this can be seen by noting that in the US in 2000 roughly 96 million people, or 34% of the total population, wore prescription glasses [36].

In our system, the otherwise undesired, gross appearance distortion that glasses cause in thermal imagery is used to help recognition by detecting their presence. If the subject is not wearing glasses, then both holistic and all local patches-based face representations can be be used in recognition; otherwise the eye regions in thermal images are ignored.

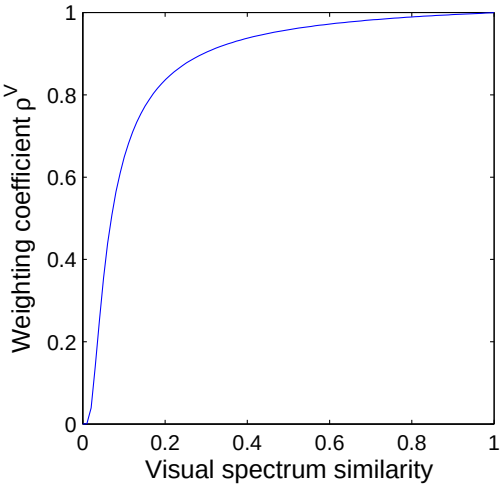


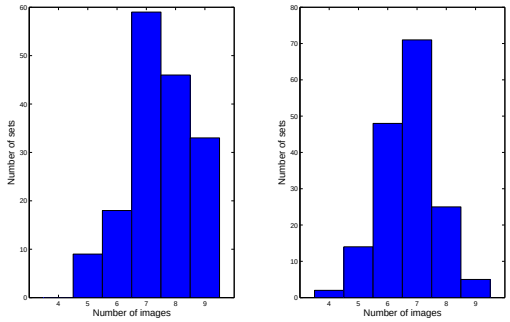
Figure 8: ***Modality fusion:*** *The contribution of visual matching, as a function of the similarity of visual imagery. A low similarity score between image sets in the visual domain is indicative of large illumination changes and consequently our algorithm leant that more weight should be placed on the illumination-invariant thermal spectrum.*

Glasses detection We detect the presence of glasses by building representations for the left eye region (due to the symmetry of faces, a detector for only one side is needed) with and without glasses, in the thermal spectrum. The foundations of our classifier are laid in §2.1. Appearance variations of the eye region with out without glasses are represented by two 6D linear subspaces, see Fig. 9 for example training data. Patches extracted from a set of thermal imagery of a novel person is then compared with each subspace. The presence of glasses is deduced when the corresponding subspace results in a higher similarity score. We obtain close to flawless performance on our data set (also see §3 for description), as shown in Fig. 10.

The presence of glasses severely limits what can be achieved with thermal imagery, the occlusion heavily affecting both the holistic face appearance as well as that of the eye regions. This is the point at which our method heavily relies on decision fusion with visual data, limiting the contribution of the thermal spectrum to matching using mouth appearance only i.e. setting $\omega_h = \omega_e = 0.0$ in (4).

3 Empirical Evaluation

We evaluated the described system on the *"Dataset 02: IRIS Thermal/Visible Face Database"* subset of the *Object Tracking and Classification Beyond the Visible Spectrum (OTCBVS)* database[1], freely available for download at `http://www.cse.ohio-state.edu/OTCBVS-BENCH/`. Briefly, this database contains 29 individuals, 11 roughly matching poses in visual and thermal spectra and large illumination variations (some of these are exemplified in Fig. 1).



(a) Visual (b) Thermal

Figure 11: ***Training sets:*** *Shown are histograms of the number of images per person used to train our algorithm. Depending on the exact head poses assumed by the user we typically obtained 7-8 visual spectrum images and typically a slightly lower number for the thermal spectrum.*

Our algorithm was trained using all images in a single illumination in which all 3 salient facial features could be detected. This typically resulted in 7-8 images in the visual and 6-7 in the thermal spectrum, see Fig. 11.

3.1 Results

A summary of the recognition results in shown in Tab. 1. Firstly, note the poor performance achieved using both raw visual as well as raw thermal data. The former is suggestive of challenging illumination changes present in our data set. This is further confirmed by significant improvements gained with both band-pass filtering and the Self-Quotient Image (of respectively 35% and 47%). On the other hand, the reason for low recognition rate of raw thermal imagery is twofold: we have previously argued that the two main limitations of this modality are the inherently low discriminative power and occlusions caused by prescription glasses. The addition of the glasses detection module of §2.5 is of little help at this point - some benefit is gained by steering away from misleadingly good matches between any two people wearing glasses, but it is limited in extent as a very

[1]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.

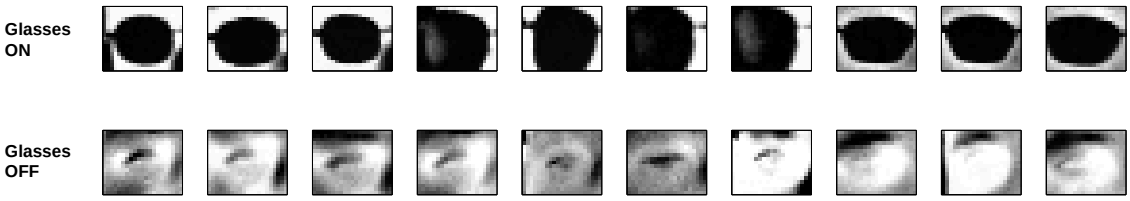
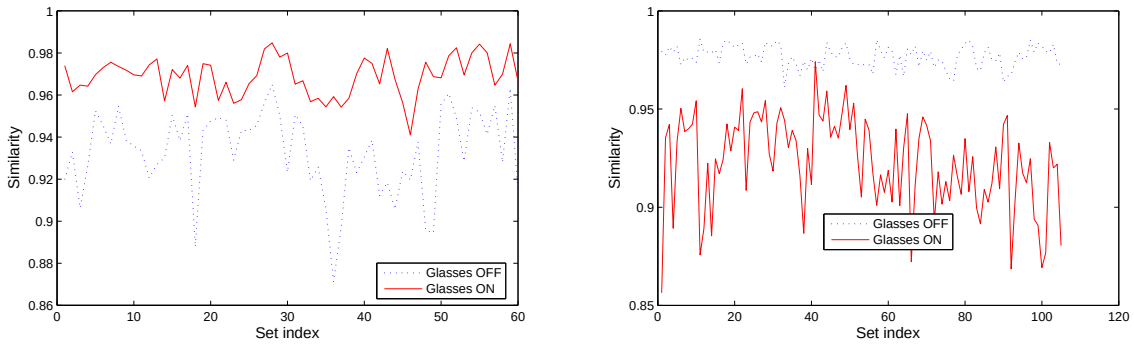


Figure 9: ***Appearance models:*** *Shown are examples of glasses-on (top) and glasses-off (bottom) thermal data used to construct the corresponding appearance models for our glasses detector.*



(a) Glasses ON

(b) Glasses OFF

Figure 10: ***Glasses detection results:***.

discriminative region of the face is lost. Furthermore, the performance improvement by optimal band-pass filtering in thermal imagery is much more modest than with visual data (35% vs. 8%). Finally, fusion of holistic and local appearance offered a small, yet statistically significant improvement. The real power of our method becomes apparent when the two modalities are fused. In this case the role of the glasses detection module is much more prominent, drastically decreasing the average error rate (from 10% to 3%).

4 Conclusion

In this paper we described a system for personal identification based on a face biometric that uses cues from visual and thermal imagery. The two modalities are shown to complement each other, their fusion providing good illumination invariance and discriminative power between individuals. Prescription glasses, a major difficulty in the thermal spectrum, are reliably detected by our method, restricting the matching to non-affected face regions. Finally, we examined how different preprocessing methods affect recognition in the two spectra, as well as holistic and local

Representation		**Rec.**
Visual	Holistic raw data	0.58
	Holistic, band-pass	0.78
	Holistic, SQI	0.85
	Mouth+eyes+holistic data fusion, SQI	0.87
Thermal	Holistic raw data	0.74
	Holistic raw w/ glasses detection	0.77
	Holistic, low-pass	0.80
	Mouth+eyes+holistic data fusion, low-pass	0.82
Proposed thermal + visual fusion	w/o glasses detection	0.90
	w/ glasses detection	*0.97*

Table 1: ***Recognition results:*** *Shown is the average rank-1 recognition rate using different face representations across all combinations of illuminations.*

feature-based face representations. The proposed method was shown to achieve a high recognition rate (97%) using only a small number of training images (5-7) in the presence of large illumination changes.

Our results suggest several possible avenues for improvement. We intend to make further use of the thermal spectrum, by not only detecting the glasses, but also by segmenting them out. This is challenging across large pose variations, such as those contained in our test set. Another research direction we would like to pursue is that of synthetically enriching the training corpus to achieve increased robustness to pose differences between image sets (c.f. [23] [34]).

References

[1] Y Adini, Y. Moses, and S. Ullman. Face recognition: The problem of compensating for changes in illumination direction. *PAMI*, 19(7):721–732, 1997.

[2] O. Arandjelović and R. Cipolla. Face recognition from video using the global shape-illumination manifold. *ECCV*, 2006. (to appear).

[3] O. Arandjelović and A. Zisserman. Automatic face recognition for film character retrieval in feature-length films. *CVPR*, 1:860–867, 2005.

[4] S. Ben-Yacoub, Y. Abdeljaoued, and E. Mayoraz. Fusion of face and speech data for person identity verification. *IEEE Trans. on Neural Networks*, 10(5):1065–1074, 1999.

[5] T. L. Berg, A. C. Berg, J. Edwards, M. Maire, R. White, and David A. Forsyth Yee Whye Teh, Erik Learned-Miller. Names and faces in the news. *CVPR*, 2004.

[6] Å. Björck and G. H. Golub. Numerical methods for computing angles between linear subspaces. *Mathematics of Computation*, 27(123):579–594, 1973.

[7] R. Brunelli and D. Falavigna. Personal identification using multiple cues. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(10):955–966, 1995.

[8] P. Buddharaju, I. Pavlidis, and I. Kakadiaris. Face recognition in the thermal infrared spectrum. In *IEEE International Workshop on Object Tracking and Classification Beyond the Visible Spectrum*, 2004.

[9] K. Chang, K. W. Bowyer, S. Sarka, and B. Victor. Comparison and combination of ear and face image in appearance-based biometrics. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(9):1160–1165, 2003.

[10] X. Chen, P. Flynn, and K. Bowyer. Visible-light and infrared face recognition. In *Workshop on Multimodal User Authentication*, pages 48–55, 2003.

[11] T-J. Chin and D. Suter. A new distance criterion for face recognition using image sets. *ICCV*, pages 549–558, 2006.

[12] I. J. Cox, J. Ghosn, and P. N. Yianilos. Feature-based face recognition using mixture-distance. pages 209–216, 1996.

[13] D. Cristinacce, Cootes T. F., and I Scott. A multistage approach to facial feature detection. *BMVC*, 1:277–286, 2004.

[14] J. Dowdall, I. Pavlidis, and G. Bebis. A face detection method based on multi-band feature extraction in the near-ir spectrum. In *IEEE Workshop on Computer Vision Beyond the Visible Spectrum*, 2001.

[15] C. Eveland, D. Socolinsky, and L. Wolff. Tracking human faces in infrared video. *Image and Vision Computing*, 21:579–590, 2003.

[16] P. F. Felzenszwalb and D. Huttenlocher. Pictorial structures for object recognition. *IJCV*, 61(1):55–79, 2005.

[17] G. Friedrich and Y. Yeshurun. Seeing people in the dark: face recognition in infrared images. In *2nd BMCV*, 2003.

[18] J. Heo, B. Abidi, S. G. Kong, and M. Abidi. Performance comparison of visual and thermal signatures for face recognition. In *Biometric Consortium Conference*, Arlington, VA, September.

[19] J. Heo, S. Kong, B. Abidi, and M. Abidi. Fusion of visual and thermal signatures with eyeglass removal for robust face recognition. In *IEEE International Workshop on Object Tracking and Classification Beyond the Visible Spectrum*, 2004.

[20] L. Hong and A. Jain. Integrating faces and fingerprints for personal identification. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(12):1295–1307, 1998.

[21] T. Kanade. Picture processing by computer complex and recognition of human faces. Technical report, 1973.

[22] S. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi. Recent advances in visual and infrared face recognition - a review. *Computer Vision and Image Understanding*, 97:103–135, 2004.

[23] A. M. Martinez. Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class. *PAMI*, 24(6):748–763, 2002.

[24] P. S. Penev. Dimensionality reduction by sparsification in a local-features representation of human faces. Technical report, The Rockefeller University, 1999.

[25] P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone. Face recognition vendor test 2002. Technical report, National Institute of Standards and Technology, 2003. Evaluation Report, 1-56.

[26] F. Prokoski. History, current status, and future of infrared identification. In *IEEE Workshop on Computer Vision Beyond the Visible Spectrum*, Hilton Head.

[27] A. Ross and A. Jain. Information fusion in biometrics. *Pattern Recognition Letters*, 24(13):2115–2125, 2003.

[28] A. Selinger and D. Socolinsky. Appearance-based facial recognition using visible and thermal imagery: A comparative study. Technical Report 02-01, Equinox Corporation, 2002.

[29] J. Sivic, M. Everingham, and A. Zisserman. Person spotting: video shot retrieval for face sets. *CIVR*, 2005.

[30] D. Socolinsky and A. Selinger. Comparative study of face recognition performance with visible and thermal infrared imagery. In *International Conference on Pattern Recognition*, pages 217–222, 2002.

[31] D. Socolinsky and A. Selinger. Thermal face recognition in an operational scenario. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004.

[32] D. Socolinsky, A. Selinger, and J. Neuheisel. Face recognition with visible and thermal infrared imagery. *Computer Vision and Image Understanding*, pages 72–114, 2003.

[33] A. Srivastana and X. Liu. Statistical hypothesis pruning for recognizing faces from infrared images. *Image and Vision Computing*, pages 651–661, 2003.

[34] K. K. Sung and Tomaso Poggio. Example-based learning for view-based human face detection. *PAMI*, 20(1):39–51, 1998.

[35] L. Trujillo, G. Olague, R. Hammud, and B. Hernandez. Automatic feature localization in thermal images for facial expression recognition. *Joint IEEE International Workshop on Object Tracking and Classification Beyond the Visible Spectrum*, 3, 2005.

[36] T. C. Walker and R. K. Miller. *Health Care Business Market Research Handbook.* Norcross (GA): Richard K. Miller & Associates, Inc., fifth edition, 2001.

[37] H. Wang, S. Z. Li, and Y. Wang. Face recognition under varying lighting conditions using self quotient image. *AFG*, 2004.

[38] L. B. Wolff, D. A. Socolinsky, and C. K. Eveland. Quantitative measurement of illumination invariance for face recognition using thermal infrared imagery. In *IEEE Workshop on Computer Vision Beyond the Visible Spectrum*, 2001.