Applying Fusion in Thermal Face Recognition

Jan Váňa*, Štěpán Mráček*, Ahmad Poursaberi+, Svetlana Yanushkevich+, Martin Drahanský*

* Faculty of Information Technology
Brno University of Technology
Brno, Czech Republic
{ivanajan,imracek,drahan}@fit.vutbr.cz
+Department of Electrical and Computer Engineering
University of Calgary
Calgary, AB, Canada
{apoursab,syanshk}@ucalgary.ca

Abstract: Face recognition based on thermal images has minor importance in comparison to visible light spectrum recognition. Nevertheless, in applications such as livelyness detection or fever scan, thermal face recognition is used as a stand-alone module, or as part of a multi-modal biometric system. This paper investigates combinations of many methods, used for thermal face recognition, and introduces some new and modified algorithms, which have not been used in the area as of yet. Moreover, we show that the best method is always limited to a certain database (input data). In order to address this problem, the multi-algorithmic biometric fusion, based on the logistic regression, is deployed.

1 Introduction

Face recognition is widely used in access control and surveillance biometric systems. Standard visible spectrum cameras are used as he sensors in these applications. Recently, there has been an increased interest in face recognition in the thermal infra-red spectrum, because it solves some problems of the visible spectrum recognition. Thermal images are remarkably invariant to light conditions, and, most importantly, provide higher level of liveness detection. On the other hand, intra-class variability is higher due to variations in environment and face temperature, emotions and health state [ABB06, Che03].

According to recent studies, there are several algorithmic approaches used for thermal face recognition [ABB06, HRdSVC12]. The appearance based methods [SS06, FY02, Che03] deal with suche aligned images as a matrix of numbers. They are computationally efficient, but have some problems with head pose variance. Another approach is to extract features using Gabor filter bank [BPK04]. In [BPTB07], authors present thermal face recognition by vascular network extraction.

Some papers use a fusion of visible and infrared face images at the decision or feature level [AH06, FY02, SVN08]. Contrary to other papers, which describe multi-algorithmic fusion [KS08], we evaluate our proposed algorithm using images from several databases. The

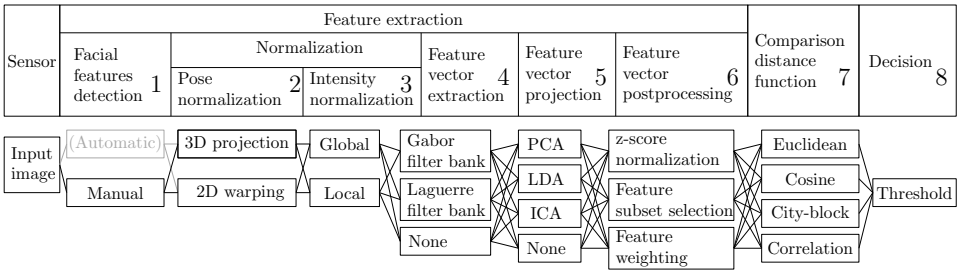


Figure 1: The overview of a generalized recognition pipeline. Any path from the left to the right yields a possible recognition algorithm.

theoretical background of a general multi-biometric fusion, especially the link between the correlation and variance of both the impostor and genuine distribution between the employed recognition methods, is described in [PB05].

In contrast, we perform the fusion by combining both appearance approaches and feature extraction approaches for infrared images. In addition, we present a new approach to geometrically normalize images before performing the feature extraction. This normalization uses a 3D model back-projection in order to deal with different head poses.

2 Thermal face recognition

Almost every biometric system consists of four major parts – sensor, feature extraction, comparison, and the final decision. In order to address specific parts of the process, we use a slightly more complex pipeline structure (see Figure 1). The numbers in the following list refer directly to the numbers in Figure 1.

1. The detection of facial features involves the localization of important facial landmarks. These landmarks are used in subsequent steps. In this paper, we are using manually annotated data, because precise detection of facial features is still a challenge. Moreover, we are focusing on algorithm performance, rather than detection accuracy.
2. The face can be normalized into some predefined position by affine transformation, based on the detected points (*2D-warping*). The other solution (*3D-projection*) involves a 3D model that is adapted to the input image.
3. Some of our testing databases contain the IR images, captured in dynamic range mode. Due to this fact, intensity normalization is needed.
4. The feature vector extraction is a simple vectorization of the normalized image. The other possibility is the application of a filter bank. Either the Gabor filter [Lee96] or the Laguerre-Gaussian filter banks [JN95] may be used.
5. In order to reduce space dimensionality, as well as to increase inter-class variability and/or reduce redundancy, the feature vector may optionally be passed onto some statistical projection technique. The PCA, LDA, and ICA [TP91, BHK97, Hyv99]

were used for this purpose.

6. Feature vector post-processing involves additional processing of the feature vector. For example, a selection of the best components (in terms of recognition performance) may be used. Individual components may obtain weights, or the components may be normalized.
7. A comparison of two processed faces (feature vectors) is accomplished by calculating the distance between them. Any distance-metric function may accomplish this task. In our paper, we are using Euclidean, cosine, city-block (Manhattan, sum of absolute differences), and correlation metric.
8. The final decision is simply a thresholding of the achieved comparison score.

2.1 Pose normalization

The pose normalization is a function that transforms points $\mathbf{p_R}$ of the raw input image, $\mathbf{R}$, to its new positions, $\mathbf{p_G}$, of the normalized image, $\mathbf{G}$. This transformation tries to eliminate intra-class variance, caused by a different head pose. Geometrical transformation usually needs an additional information, such as a minimal face bounding box, or the locations of some important facial landmarks.

This paper presents two methods for geometrical normalization: 2D-warping and 3D-projection (see Figure 2(a)). Both use a set of important facial landmarks, $\mathbf{F} = (f_1, f_2, ...f_n)$. Let $\mathbf{p}_{\mathbf{I},f}$ be the coordinates of a landmark f in an image $\mathbf{I}$. The position of $\mathbf{p}_{\mathbf{R},f}$ is obtained during facial feature detection, while the $\mathbf{p}_{\mathbf{G},f}$ is the same for all input images.

The 2D-warping using affine transformation is a well-known normalization approach. The three important facial points (typically eyes and mouth or nose) are selected as landmarks, $\mathbf{F}$. The affine transformation matrix, $\mathbf{AM}$, can be computed by solving a system of three equations (see [Rat95]). This matrix is used to map an input image to the normalized template as follows:

$$\mathbf{G}[\mathbf{AM} \cdot \mathbf{p_R}] = \mathbf{R}[\mathbf{p_R}], \tag{1}$$

where $\mathbf{R}[\mathbf{p_R}]$ is the pixel intensity at position $\mathbf{p_R}$ of image $\mathbf{R}$.

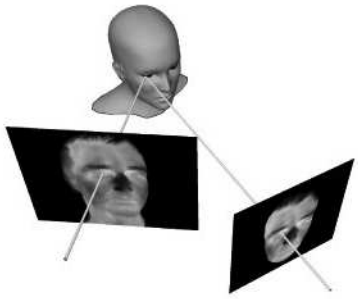
Human heads have an irregular ellipsoid-like 3D shape, therefore, the 2D-warping method works well, when the head is scaled or rotated in the image plane. In case of any other transformation, the normalized face is deformed. Possible solution is to normalize the head, using 3D head model [Bla06]. This is a complex approach, and needs some manual assistance.

The proposed 3D-projection method works with the average 3D model of a human head, M, consisting of vertices $\mathbf{v}_M$. A geometrical transformation, $\tau(\mathbf{v}_M)$, can be applied to each vertex. $\tau(\mathbf{v}_M)$ usually consists of translation, rotation and scaling. A transformed model point $\tau(\mathbf{v}_M)$ can be perspectively projected by $\vartheta(\tau(\mathbf{v}_M))$ to the 2D plane. The coordinates $\mathbf{v}_{M,f}$ of each landmark f have to be known within the model M.

Next, the model alignment according to the image $\mathbf{I}$ is required. It is aimed at finding a

transformation, $\tau_{\mathbf{I}}(\mathbf{v}_M)$, of model points, $\mathbf{v}_M$, such that each $\vartheta(\tau_{\mathbf{I}}(\mathbf{v}_{M,f}))$ is equal to $\mathbf{p}_{\mathbf{I},f}$. The normalized image is computed by the assignment $\mathbf{G}[\vartheta(\tau_{\mathbf{G}}(\mathbf{v}_M)] = \mathbf{R}[\vartheta(\tau_{\mathbf{R}}(\mathbf{v}_M)]$ for each visible point $\mathbf{v}_M$. In other words, we project the texture of the image, R, to the aligned model, M, according to $\mathbf{R}$, and then align the model with $\mathbf{G}$. Finally, the texture from model to resulting image, $\mathbf{G}$, is projected (see Figure 2(b)).

(a) Geometric pose normalization methods: Original image (left), 2D-warp (middle), and 3D-projection (right).



(b) Visualization of the 3D-projection pose normalization method.

Figure 2: Pose normalization.

2.2 Intensity normalization

Each pixel of a thermal image represents a temperature in the corresponding area. Backward mapping from pixel intensity to temperature is possible only when calibration data is available. Since some parts of our databases were captured in the dynamic range mode, we could not reconstruct absolute temperatures, and only the relative temperature differences were retrieved. For that reason, the intensity normalization is needed in image processing. It is accomplished via histogram equalization of face regions of interest. e Global histogram equalization, (*Global*), and local histogram equalization, (*Local*), were performed (see Figure 3).

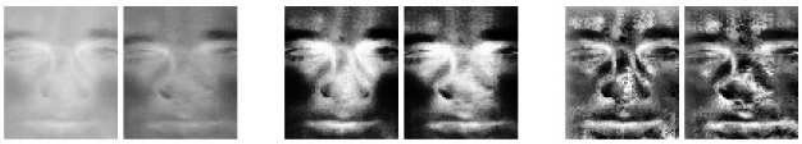


Figure 3: Comparison of e intensity normalization methods: original images are on the left; results of global histogram equalization are in the center, and results of local histogram equalization are on the right.

2.3 Feature extraction

Feature extraction involves the transformation of the input image data into some vector representation. The simplest approach is vectorization. The image is considered as a matrix $\mathbf{I}$ with r rows and c columns. The vectorization transforms $\mathbf{I}$ to the feature vector $\mathbf{fv} \in \mathbb{R}^{r \cdot c}$, by simple row/column concatenation.

In this paper, we apply filter banks to feature extraction. This approach has been previously described in [YJW07, SSC09]. The normalized image is convolved with a bank of 2D filters, which are generated, using some kernel function with different parameters. Each response of this convolution forms a final feature vector.

The Gabor filter bank is one of the most popular banks [Lee96]. We employed the Laguerre-Gaussian filter bank as well, due to its good performance in the face recognition area [JN95]. In case of the Gabor bank, we use 3 different scales and 4 different orientations. The Laguerre-Gaussian filter bank consists of 4 filters, generated by 2 kernel functions, using both real and imaginary parts.

2.4 Statistical projections

The statistical projection methods linearly transform the input feature vector from an m-dimensional space into an n-dimensional space, where $n < m$. They are usually integrated into the feature extraction part of the pipeline, but we have decided to split feature extraction and statistical projections. We have used the following methods:

- Principal component analysis (PCA, Eigenfaces)
- PCA followed by linear discriminant analysis (LDA of PCA, Fisherfaces)
- PCA followed by independent component analysis (ICA of PCA)

Every projection method has a common learning parameter, which defines how much variability of the input space is captured by the PCA. This parameter controls the dimensionality of the output projection space. Let the k eigenvalues, computed during the PCA calculation, be denoted $e_1, e_2, \ldots, e_k$, $(e_1 > e_2 > \cdots > e_k)$. These eigenvalues directly represent the variability in each output dimension. If we want to preserve only 98% of variability, then only the first l eigenvalues, and corresponding eigenvectors are selected, such that their sum forms only 98% of the $\sum_{j=1}^{k} e_j$.

There is an optional step to perform a per-feature z-score normalization after the projection, such that each vector $\mathbf{fv}$ is transformed into $\mathbf{fv}' = \frac{\mathbf{fv} - \overline{\mathbf{fv}}}{\boldsymbol{\sigma}}$, where $\overline{\mathbf{fv}}$ is the mean vector and $\boldsymbol{\sigma}$ is the vector of standard deviations.

2.5 Feature selection and weighting

An optional processing, after the application of the statistical projection methods, is the feature weighting. We can consider the decorrelated feature vectors, since the PCA, ICA, or LDA was applied. Suppose that we have a set $\mathbf{FV}$ of all pairs of feature vectors, $\mathbf{fv}_j$, and their corresponding class (subject) labels, id_j: $\mathbf{FV} = \{(id_1, \mathbf{fv}_1), (id_2, \mathbf{fv}_2), \ldots, (id_n, \mathbf{fv}_n)\}$.

The individual feature vector components, $fv_{j_1}, fv_{j_2}, \ldots, fv_{j_m}$, of the vector $\mathbf{fv}_j$ do not have the same discriminative ability. While some component may have positive contribution to the overall recognition performance, the other component may not. The optimization criterion in feature weighting is achieved EER with specific feature weights $\mathbf{w} = w_1, \ldots, w_m$.

We have implemented and evaluated two possible feature-weighting techniques. The first option is to consider individual feature vector components as a standalone one-dimensional feature vectors, and then to evaluate the equal error rate EER_i for each specific component i, separately. The resulting discriminative ability, that we call *EER potential* of the component i, is:

$$\text{EER potential}_i = 1 - EER_i \tag{2}$$

The individual EER potentials are linearly transformed, such that the smallest value becomes 0 and the greatest is transformed to 1. These new values EER'_i directly forms a weight vector $\mathbf{w} = EER'_1, \ldots, EER'_m$.

The second option is to make an assumption that the appropriate feature vector component has stable values across the different scans of the same subject, and, on the other hand, the mean value of a specific component across different subject differs as much as possible. Let the intra-class variability of feature component i be denoted as $intra_i$, and expresses the mean of standard deviations of all measured values, for the same subjects. The inter-class variability of component i is denoted as $inter_i$, and expresses the standard deviation of means of the measured values for the same subject. The resulting *discriminative potential* of component i, therefore, can be expressed as follows:

$$\text{Discriminative potential}_i = inter_i - intra_i \tag{3}$$

Further, in our tests, we have used only the discriminative potential, because it yields almost the same result as an EER potential, but due to its definition, it is much faster to compute.

While, EER calculations for each feature vector component have to be performed for the EER potential evaluation, the discriminative potential can be computed only within one iteration over a training set. For example, our implementation of a discriminative potential in C++ takes about 2 seconds to compute discriminative potential of 1000 feature vectors (images) of length 4080 in contrast to 20 seconds that are needed for an EER potential calculation (Measured on an Intel Core i3 CPU, 2GB of RAM).

3 Evaluation

We use four databases to evaluate the performance of the described methods. First, two databases were captured at Brno University of Technology (BUT database I and BUT database II). They are not public, and together contain more than 600 images of 60 persons. The BUT database II contains high intra-class variability, caused by various head poses and facial expressions. The other two databases are Equinox [equ] and the Notre Dame University database [CFB03, FBP03]. Equinox contains 243 scans of 74 subjects, while the Notre Dame database consists of 2292 scans of 82 subjects.

The evaluation scenario was as follows. We divided each database into three equal parts. Different data subjects were present in each part. The first portion of the data was used for training of the projection methods. The second portion was intended for an optional calculation of the z-score normalization parameters, feature weighting, and training of fusion classifiers. The last part was used for evaluation.

To ensure that particular results of the employed methods are stable and reflect real performance, the following cross-validation process was selected. The database was divided into three parts randomly, where all three parts had equal number of subjects. This random division and subsequent evaluation was processed n times, where n depends on the size of the database. The smaller databases (BUT I, BUT II, and Equinox) were cross-validated 10 times, while the Notre-dame database was cross-validated 3 times. The performance of a particular method was reported as the mean value of the achieved equal error rates (EERs).

3.1 Comparison of standalone methods

The evaluation procedure involves comparing two different normalization approaches: 2D-warping and 3D model projection. The achieved results are given in Table 1. Based on these baseline tests, several conclusions are deduced. In most tests, the best combination of the statistical projection method and distance comparison function is the independent component analysis (ICA) and cosine metric. Thus, in further tests, we focus mainly on this combination.

The second conclusion concerns the comparison of the 3D-projection pose normalization and simple 2D-warping. Except the Notre-dame database, the recognition results, expressed in EER, are always better, when 3D-projection normalization was employed. The decrease of recognition performance for 3D-projection in Notre-dame is caused by the fact that 3D projection performs better on images with higher head pose variance, while it has slightly worse performance on frontal images, because of an imprecise alignment.

After the evaluation of the standalone methods in their basic form, we have selected the best candidates for further improvement: PCA and ICA, using the cosine distance function. For both PCA and ICA, we have evaluated the recognition performance, without any additional feature vector processing. Moreover, the z-score normalization of feature vector components and weighting, based on discriminative potential, have been added.

Table 1: The EER (in %) for various single methods applied on the Equinox and Notre-dame databases, and selected results from both BUT databases, given various pose normalizations, statistical projections, and distance functions.

Projection	Warp2D				Projection3D			
	Euclidean	city-block	cosine	correlation	Euclidean	city-block	cosine	correlation
Database: Equinox, PCA selection threshold: 0.98								
PCA	7.27	6.40	5.90	5.69	5.60	5.17	5.00	5.12
LDA	7.77	6.76	6.31	6.42	5.93	5.69	5.51	5.60
ICA	7.46	6.93	**4.21**	4.44	6.66	6.33	**4.06**	4.15
Database: Equinox, PCA selection threshold: 0.9999								
PCA	6.69	6.15	5.26	5.17	6.24	5.67	5.51	5.55
LDA	7.76	7.18	6.84	7.05	6.62	6.68	6.50	6.61
ICA	7.77	7.21	**3.95**	4.13	7.63	7.36	**4.22**	4.31
Database: Notre-dame, PCA selection threshold: 0.98								
PCA	24.11	19.45	23.02	23.03	25.66	19.83	24.46	24.47
LDA	26.84	26.11	26.50	26.54	28.5	27.90	28.01	28.10
ICA	21.98	21.67	**8.22**	8.33	22.34	22.06	**10.29**	10.32
Database: Notre-dame, PCA selection threshold: 0.9999								
PCA	24.24	20.62	22.85	22.86	25.28	21.22	23.83	23.80
LDA	27.27	26.54	26.73	26.81	28.47	27.94	27.61	27.69
ICA	21.80	21.47	**7.64**	7.69	23.46	23.14	**10.24**	**10.23**
Database BUT I, PCA selection threshold: 0.9999								
ICA	9.86	9.25	8.06	8.07	6.89	6.32	5.66	5.83
None	6.11	**5.40**	6.11	6.22	3.87	**3.28**	3.95	4.12
Database BUT II, PCA selection threshold: 0.98								
ICA	20.16	20.01	**16.73**	16.82	17.68	17.42	**13.42**	13.55

The results are presented in Table 2. This table shows that the feature vector component normalization, and subsequent weighting have the most impact on the PCA. On the other hand, there is no significant improvement for the ICA, when the z-score normalization or weighting is present. This is due to the fact that the used FastICA algorithm requires a preliminary whitening of training data, so that their correlation matrix equals unity [Hyv99]. This means that in the case of the ICA, the data is already normalized.

Table 2: The EER (in %) for PCA and ICA after z-score normalization and feature weighting. PCA selection threshold is 0.98.

Database	Projection	Warp2D			Projection3D		
		Plain	Z-score	Weighting	Plain	Z-score	Weighting
BUT I	PCA	8.82	6.83	5.41	6.59	4.61	**3.88**
BUT I	ICA	7.29	7.07	5.70	4.00	5.00	3.89
BUT II	PCA	19.82	17.09	16.05	17.40	12.86	13.18
BUT II	ICA	16.58	17.03	16.23	**12.53**	12.81	13.08
Equinox	PCA	3.87	2.37	**2.08**	6.35	3.49	4.03
Equinox	ICA	3.65	2.51	2.40	3.69	3.25	3.66
Notre-dame	PCA	24.36	**7.42**	9.11	22.23	8.47	8.83
Notre-dame	ICA	7.73	7.46	9.15	9.10	8.09	10.31

Other evaluated properties of the recognition pipeline are the intensity normalization and feature extraction, using filter banks. We have evaluated all possible combinations of filter banks and pose normalization techniques, using PCA and ICA on all four databases. These results are given in Table 3. Neither the Gabor nor the Laguerre filter bank, followed by the weighted PCA or ICA, outperformed plain feature extraction. However, the filter banks are used in an additional fusion approach (see section 3.2).

Table 3: The EER (in %) for various applications of filter banks and intensity normalization approaches. The PCA selection threshold was set to 0.98. The results are reported for weighted PCA and ICA.

Database	Normalization	Projection	Global normalization			Local normalization		
			No bank	Gabor	Laguerre	No bank	Gabor	Laguerre
BUT I	Warp2D	wPCA	5.06	7.63	7.41	8.60	11.69	10.33
		wICA	5.03	7.76	7.48	9.06	12.14	10.90
	Projection3D	wPCA	**3.06**	4.62	4.15	5.47	7.19	7.92
		wICA	3.22	4.92	4.25	5.62	7.76	8.28
BUT II	Warp2D	wPCA	16.79	19.26	18.56	20.50	20.98	21.97
		wICA	16.58	19.47	19.09	20.88	21.44	22.39
	Projection3D	wPCA	**12.04**	17.81	12.78	16.13	20.47	17.46
		wICA	12.32	18.01	13.08	16.70	20.47	18.11
Equinox	Warp2D	wPCA	2.77	6.30	6.45	5.35	8.95	6.01
		wICA	**2.71**	6.27	6.31	5.49	8.60	6.27
	Projection3D	wPCA	3.76	7.76	6.42	4.97	9.24	5.99
		wICA	4.05	8.08	6.22	5.49	9.17	6.53
Notre-Dame	Warp2D	wPCA	8.30	14.69	14.87	7.23	13.30	13.24
		wICA	**8.19**	14.63	15.14	21.94	13.33	13.40
	Projection3D	wPCA	9.22	16.23	15.87	9.69	15.47	14.54
		wICA	13.88	16.23	16.38	22.20	15.41	14.57

3.2 Biometric fusion

It follows from previous experiments that no single combination of the feature extraction method, and no distance function exist that outperform all other combinations on every database. The best solution is always limited to a certain data. As a solution to the posed problem, we propose an approach, based on biometric fusion.

In this paper, we are using score-level fusion. Each employed method (feature extraction and distance function) provides the resulting comparison score. The task of the fusion component is to combine these input scores into a final one, that is finally thresholded.

In order to compare and fuse the scores that come from different methods, the normalization, to a certain range, has to be performed. We use the following score normalization: the score values are linearly transformed, such that the genuine mean (the score, obtained from comparing the same subjects) is 0, and the impostor mean (score, obtained from comparing different subjects) is 1. Note that at this point, individual scores may have negative values. This does not matter in the context of the score-level fusion, since these values represent the positions within the classification space, rather than the distances between two scans.

Score fusion is, in fact, a binary classification problem. Based on an input score vector, $\mathbf{s} = (s_1, s_2, \ldots, s_n)$, the fusion classifier has to decide, whether the score vector belongs to a genuine user or impostor. We have implemented the following fusion methods: Support Vector Machine (SVM) with linear and sigmoid kernel, Linear Discriminant Analysis (LDA), and Logistic Regression. Initially, all of these methods have been applied to the Notre-dame database, but no significant differences were noticed. Thus, in further tests, we concentrate on the logistic regression only. The overview of the recognition pipeline that involves logistic regression fusion of 10 individual feature extraction and comparison methods, is shown in Figure 4.

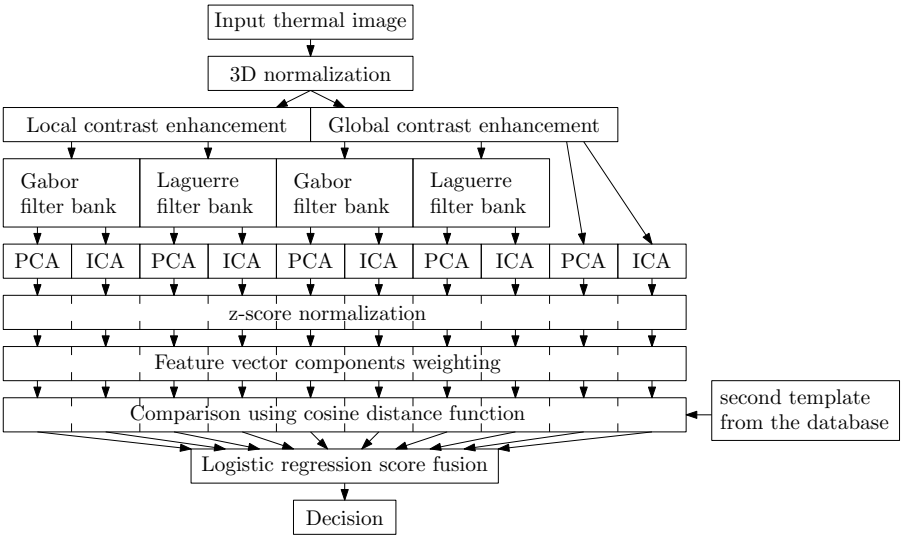


Figure 4: Fusion pipeline.

Fusion results from all databases are shown in Table 4. The fusion is always better than any of the employed methods alone. It is not clear, whether the best statistical projection is the weighted PCA or ICA. On the other hand, it is clear that fusion at the score-level outperforms both the PCA and ICA.

Contrary to previous evaluations shown in Tables 1, 2, and 3, we also added two important indicators of performance. These are False Non-Match Rates (FNMR) at specific False Match Rates (FMR). These measures show the probability of a genuine user being rejected, if we assume that 0.1% of impostors are wrongly accepted. The FNMRs were measured at FMRs = 0.1% and 1%.

Table 4: Evaluation of fusion based on logistic regression. For every fusion test, all individual components of the resulting fusion method were evaluated separately. The best component were compared with the overall fusion, and improvement was reported. The numbers represent the achieved EER in %. The FNMR at FMR = 0.1% and FMR = 1% are $FNMR_1$ and $FNMR_2$, correspondingly.

Database, normalization	Best single method name	Best single method EER	Fusion	improvement	$FNMR_1$	$FNMR_2$
BUT I, 2D	PCA, Laguerre filter, global norm.	4.31	4.09	5.10%	18.58	10.18
BUT I, 3D	ICA without any filter	2.70	2.21	18.15%	6.02	3.13
BUT II, 2D	PCA, Laguerre filter, global norm.	14.96	12.61	15.71%	54.47	37.44
BUT II, 3D	PCA, Laguerre filter, global norm.	9.70	8.69	10.41%	40.40	24.79
Equinox, 2D	ICA without any filter	2.28	1.06	53.51%	8.35	1.21
Equinox, 3D	ICA without any filter	2.68	1.90	29.10%	10.76	4.08
Notre-dame, 2D	PCA without any filter	6.70	5.99	10.60%	35.11	14.45
Notre-dame, 3D	PCA without any filter	8.22	7.61	7.42%	39.39	21.59

4 Conclusion

We have compared several existing methods of thermal face recognition. Additionally, we proposed normalization by 3D projection, feature extraction by the Laguerre-Gaussian filter bank, and feature vector component weighting by the discriminative potential.

3D normalization significantly improved performance within databases, where the head pose variation is high. For example, the experiment that used the BUT I database, fusion approach led to the EER decreasing from 4.09% to 2.21%. The discriminative potential is a good alternative to known weighting methods, in terms of computational complexity and performance improvement.

This paper confirms that not a single method performs equally well in all considered scenarios. Therefore, to maximize the performance of the thermal face recognition, a multi-algorithmic fusion must be applied.

5 Acknowledgment

This work was partially supported by the following grants: NATO CBP.EAP.CLG 984 "Intelligent assistance systems: multisensor processing and reliability analysis", ED1.1.00/02.0070 "The IT4Innovations Centre of Excellence", FIT-S-11-1 "Advanced secured, reliable and adaptive IT", and GD102/09/H083 "Information Technology in Biomedical Engineering".

The vast majority of this work emerged during the research visit of J. Váňa and Š. Mráček (Brno University of Technology) to the Biometric Technologies Laboratory, University of Calgary, supported by NATO Collaborative linkage grant CBP.EAP.CLG 984.

References

[ABB06] M. Akhloufi, A. Bendada, and J. Batsale. State of the art in infrared face recognition. *Quantitative InfraRed Thermography Journal*, 5(1):3–26, 2006.

[AH06] O. Arandjelovic and R. Hammoud. Multi-Sensory Face Biometric Fusion (for Personal Identification). In *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, page 128. IEEE, June 2006.

[BHK97] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs.Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.

[Bla06] V. Blanz. Face Recognition based on a 3D Morphable Model. In *Proc. of the 7th Int. Conference of Automatic Face and Gesture Recognition*, pages 617–622, 2006.

[BPK04] P. Buddharaju, I. Pavlidis, and I. Kakadiaris. Face Recognition in the Thermal Infrared Spectrum. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04)*, pages 133–138, 2004.

[BPTB07] P. Buddharaju, I. Pavlidis, I. Tsiamyrtzis, and I. Bazakos. Physiology-based face recognition in the thermal infrared spectrum. *IEEE transactions on pattern analysis and machine intelligence*, 29(4):613–626, April 2007.

[CFB03] X. Chen, P. Flynn, and K. Bowyer. Visible-light and infrared face recognition. *ACM Workshop on Multimodal User Authentication*, pages 48–55, 2003.

[Che03] X. Chen. PCA-Based Face Recognition in Infrared Imagery: Baseline and Comparative Studies. In *Workshop on Analysis and Modelung of Faces and Gestures*, number April, pages 127–134. Citeseer, 2003.

[equ] Equinox, Multimodal face database: http://www.equinoxsensors.com/products/HID.html.

[FBP03] P. Flynn, K. Bowyer, and P. Phillips. Assessment of time dependency in face recognition: An initial study. *Lecture Notes in Computer Science*, 2688:44–51, 2003.

[FY02] G. Friedrich and Y. Yeshurun. Seeing People in the Dark : Face Recognition in Infrared Images. In *BMCV '02 Proceedings of the Second International Workshop on Biologically Motivated Computer Vision*, pages 348–359. Springer-Verlag London, UK, 2002.

[HRdSVC12] G. Hermosilla, J. Ruiz-del Solar, R. Verschae, and M. Correa. A comparative study of thermal face recognition methods in unconstrained environments. *Pattern Recognition*, 45(7):2445–2459, July 2012.

[Hyv99] A Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, 10(3):626–34, January 1999.

[JN95] G. Jacovitti and A. Neri. Multiscale image features analysis with circular harmonic wavelets. In *Proceedings of SPIE*, volume 2569, pages 363–374. SPIE, September 1995.

[KS08] A. Kumar and T. Srikanth. Online personal identification in night using multiple face representations. In *19th International Conference on Pattern Recognition, ICPR 2008*, pages 1–4, 2008.

[Lee96] T. Lee. Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):959–971, 1996.

[PB05] N. Poh and S. Bengio. How Do Correlation and Variance of Base-Experts Affect Fusion in Biometric Authentication Tasks? *IEEE Transacations on Signal Processing*, 53(11):4384–4396, 2005.

[Rat95] B. Raton. Standard Mathematical Tables and Formulae. In D. Zwillinger, editor, *Affine Transformations*, pages 265–266. 1995.

[SS06] A. Selinger and D. Socolinsky. Appearance-based facial recognition using visible and thermal imagery: a comparative study. Technical report, Equinox Corporation, 2006.

[SSC09] Y. Su, S. Shan, and X. Chen. Hierarchical ensemble of global and local classifiers for face recognition. *Image Processing, IEEE*, 18(8):1885–1896, 2009.

[SVN08] R. Singh, M. Vatsa, and A. Noore. Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition. *Pattern Recognition*, 41(3):880–893, March 2008.

[TP91] M. Turk and A. Pentland. Face recognition using eigenfaces. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 591, pages 586–591. IEEE Computer Society Press, 1991.

[YJW07] Y. Yao, X. Jing, and H. Wong. Face and palmprint feature level fusion for single sample biometrics recognition. *Neurocomputing*, 70(7-9):1582–1586, March 2007.