Chapter 10
Thermal Face Recognition in Unconstrained Environments Using Histograms of LBP Features

Javier Ruiz-del-Solar, Rodrigo Verschae, Gabriel Hermosilla and Mauricio Correa

Abstract Several studies have shown that the use of thermal images can solve limitations of visible spectrum based face recognition methods operating in unconstrained environments. The recognition of faces in the thermal domain can be tackled using the histograms of Local Binary Pattern (LBP) features method. The aim of this work is to analyze the advantages and limitations of this method by means of a comparative study against other methods. The analyzed methods were selected by considering their performance in former comparative studies, in addition to being real-time—10 fps or more—to require just one image per person, and to being fully online (no requirements of offline enrollment). Thus, in the analysis the following local-matching based methods are considered: Gabor Jet Descriptors (GJD), Weber Linear Discriminant (WLD) and Local Binary Pattern (LBP). The methods are compared using the UCHThermalFace database. The use of this database allows evaluating the methods in real-world conditions that include natural variations in illumination, indoor/outdoor setup, facial expression, pose, accessories, occlusions, and background. In addition, the fusion of some variants of the methods was evaluated. The main conclusions of the comparative study are: (i) All analyzed methods

J. Ruiz-del-Solar (✉)
Department of Electrical Engineering, Universidad de Chile, Av. Tupper 2007, 837-0451 Santiago, Chile
e-mail: jruizd@ing.uchile.cl

J. Ruiz-del-Solar · R. Verschae · M. Correa
Advanced Mining Technology Center, Universidad de Chile, Av. Tupper 2007, 837-0451 Santiago, Chile

R. Verschae
e-mail: rodrigo@verschae.org

M. Correa
e-mail: macorrea@ing.uchile.cl

G. Hermosilla
Escuela de Ingeniería Eléctrica, Pontificia Universidad Católica de Valparaíso, Valparaiso, Chile
e-mail: gabriel.hermosilla@ucv.cl

S. Brahnam et al. (eds.), *Local Binary Patterns: New Variants and Applications*, Studies in Computational Intelligence 506, DOI: 10.1007/978-3-642-39289-4_10,

perform very well under the conditions in which they were evaluated, except for the case of GJD that has low performance in outdoor setups; (ii) the best tradeoff between high recognition rate and fast processing speed is obtained by LBP-based methods; and (iii) fusing some methods or their variants improve the results up to 5 %.

1 Introduction

The recognition of human faces in unconstrained environments has attracted increasing interest in the research community in recent years. Several studies have shown that the use of thermal images can solve limitations of visible-spectrum based face recognition, such as invariance to variations in illumination and robustness to variations in pose [33, 34], which are two of the major factors affecting the performance of face recognition systems in unconstrained environments [32]. This is thanks to the physical properties of thermal technology (long-wave infrared spectrum, 8–12 μm), and the anatomic characteristics of the human body:

- thermal sensors collect the energy emitted by a body instead of the reflected light, and the emissivity of human skin is between 8 and 12 μm,
- thermal sensors are invariant to changes in illumination; they can even work in complete darkness, and
- the anatomic and vascular information that can be extracted from thermal images is unique to each individual [12].

In addition, in recent years, the price of thermal cameras has decreased significantly, and their technology has improved, obtaining better resolution and quality, and the fixed pattern noise that was produced by old thermal cameras has been eliminated using non-uniformity correction techniques (NUC) [27, 28]. Thus, the interest in the use of thermal technology in face recognition applications has increased in recent years. Nevertheless, thermal face images still have undesirable variations due to (i) changes in ambient temperature, (ii) modifications of the metabolic processes of the subjects, (iii) camera susceptibility on extrinsic factors such as wind, and (iv) variable sensor response overtime when the camera is working for long periods of times [9, 12, 35].

In this general context, the aim of this article is to carry out a comparative study of thermal face-recognition methods in unconstrained environments. The results of this comparative study are intended to be a guide for developers of face recognition systems. This study concentrates on methods that fulfill the following requirements: (i) *Full online operation*: No offline enrollment stages. All processes must run online. The system has to be able to build the face database incrementally from scratch; (ii) *Real-time operation*: The recognition process should be fast enough to allow real-time interaction in case of HRI (Human-Robot Interaction) applications, and to search large databases in a reasonable time (a few milliseconds depending on the application and the size of the database; we think at least 10fps is a good minimum

requirement); (iii) *Single image per person problem*: One thermal face image of an individual should be enough for his/her later identification. Databases containing just one face image per person should be considered. The main reasons for this are savings in storage and computational costs, and the impossibility of obtaining more than one face image from a given individual in certain situations; and (iv) *Unconstrained environments*: No restrictions on environmental conditions such as illumination, indoor/outdoor setup, facial expression, scale, pose, resolution, accessories, occlusions, and background are imposed.

Thus, in this study three local-matching methods are selected by considering their fulfillment of the previously mentioned requirements, and their good performance in former comparative studies of face recognition methods [29–32, 37]. Two local-matching methods, namely, histograms of LBP (Local Binary Pattern) features [3] and Gabor Jet Descriptors with Borda count classifiers [29] are selected based on their performance in the studies reported in [29, 32]. The third local-matching method, histograms of WLD (Weber Linear Descriptor) features, which was recently proposed in [11], has shown very good performance in face detection applications, and is used here in face recognition.

The comparative study is carried out using the UCHThermalFace database.[1] This database was specially designed to study the problem of unconstrained face recognition in the thermal domain. The database incorporates thermal images acquired in indoor and outdoor setups, with *natural* variations in illumination, facial expression, pose, accessories, occlusions, and background. This database will be made public for future comparative studies, which is also a contribution of this paper.

This comparative study intends to be a complement to the recently published comparative study on visible-spectrum face recognition methods in unconstrained environments [32].

The paper is structured as follows: Related works are outlined in Sect. 2. The methods under analysis are described in Sect. 3. In Sect. 4 the comparative analysis of these methods in the UCHThermalFace database, which includes the analysis of their combined use, are presented. Finally, in Sect. 5 results are discussed, and conclusions are given.

2 Related Work

Several comparative studies of thermal face recognition approaches have been developed in recent years [33, 34, 36]. Most of the developed approaches make use of appearance-based methods, such as PCA (Principal Component Analysis), LDA (Linear Discriminant Analysis), and ICA (Independent Component Analysis), which project face images into a subspace where the recognition is carried out. These methods achieve a ∼95 % recognition rate in experiments that do not consider real-world

[1] The UCHThermalFace database is available for download at http://vision.die.uchile.cl/dbThermal/.

conditions (unconstrained environmental conditions), as in [33, 34, 36], or when using the Equinox thermal face database [15]. The Equinox database is *de facto* standard database in thermal face recognition. It consists of indoor images of 91 individuals, captured with 3 different expressions and 3 different illumination conditions.

Other reported thermal face recognition approaches are based on the use of local-matching: Local Binary Pattern (LBP) [25] and Gabor Jet Descriptors (GJD) [2, 18]. In the Equinox thermal database, a recognition rate of ∼97 % for the LBP approach [25] and ∼80 % for the GJD approach [2, 18] has been reported. Methodologies based on global matching, such as Scale Invariant Features Transform (SIFT) [24], have also been used for thermal face recognition [17, 18]. These approaches are based on the use of local feature descriptors that are invariant to rotation, translation and scale changes. These local descriptors are used to match pairs of images by considering geometrical and probabilistic restrictions. In [17] the SIFT methodology is used to obtain the descriptors directly in the thermal face images, while in [18] they are computed in the vascular images generated by processing the thermal images. These approaches obtained a recognition rate that depended strongly on the database used; ∼80 % when using Equinox [18] and ∼95 % when using a non-public database [17]. There are some recently proposed methods, such as Local Phase Quantization (LPQ) [23, 26], which have given good results using visible images but have not been yet tested in thermal images.

Recent work uses vascular information of the face in order to develop thermal face recognition systems. This is accomplished by detecting thermal minutia points, and then matching them using a similar approach to the one used for fingerprint identity verification [7–10]. This kind of methodology achieves a ∼80 % recognition rate in a non-public database. In [4], an efficient approach for the extraction of physiological features from thermal face images is presented. The features represent the network of blood vessels under the skin of the face. This network is unique to each individual, and can be used to develop thermal face recognition systems. In [13] a similar approach based on thermal *faceprints* is presented. This approach uses new feature sets to represent the thermal face: the bifurcation points of the thermal pattern and the center of gravity of the thermal face region.

In addition, in [5, 14, 29, 38] methodologies based on the fusion of visible and thermal spectrum images are proposed. In [5, 14] standard appearance-based methods are used together with genetic algorithms for the analysis and fusion of visible and thermal data. The method achieved a recognition rate of ∼96 % in the Equinox database. In [1] two schemes of fusion, data fusion and decision fusion, are applied. The algorithm is designed to detect and replace eyeglasses with an eye template in the case of thermal images. Commercial face recognition software, *FaceIt*, is used in the evaluation of the fusion algorithm. In [20], the fusion of visual and thermal images using the Discrete Wavelet Transform (DWT) domain is described. The results of the experiments demonstrate that the fusion method is effective in terms of visual quality compared to conventional fusion approaches. In [2], Gabor filters are used to extract facial features in the thermal and visual domains. In [6] different levels of fusion between visual and thermal data are analyzed. In [29] the advantages of

combining thermal and visible face recognition are analyzed, and the recognition is achieved using a k-nearest neighbor classifier. The current study focuses on pure thermal-based methods. However, its results could be used in order to select the best thermal method to be used with methods that use visible images.

In [18] the authors presented a preliminary comparative study of thermal face recognition methods that did not consider real-world conditions (it did not use the UCHThermalFace database; it only used the Equinox database), nor the use of the recently proposed WLD descriptors. The current work is partly a subset of [19], but in the current work we only consider the fastest methods, thus SIFT (Scale-Invariant Feature Transform) nor the SURF (Speeded Up Robust Features) were considered because of the time processing constrains that these methods cannot fulfill, and we consider only UCHThermalFace which is a more difficult database. In addition to the results presented in [19], in the present work results obtained by combining the rankings obtained by LBP, WLD and their variants are presented.

3 Methods Under Comparison

As mentioned above, the methods under comparison were selected considering their fulfillment of the defined requirements (real-time—10 fps or more—fully online, just one image per person), and their performance in former comparative studies of face-recognition methods [31, 32, 37, 39] and in face detection applications [11].

3.1 LBP Histograms

Face recognition using histograms of LBP (Local Binary Pattern) features was originally proposed in [3], and has been used by many groups since then. In the original approach, three different levels of locality are defined: pixel level, regional level, and holistic level. The first two levels of locality are achieved by dividing the face image into small regions from which LBP features are extracted and histograms are used for efficient texture information representation. The holistic level of locality, i.e. the global description of the face, is obtained by concatenating the regional LBP extracted features. The recognition is performed using a nearest neighbor classifier in the computed feature space, using one of the three following similarity measures: histogram intersection, log-likelihood statistic, or Chi square. We implemented this recognition system, without considering preprocessing (cropping using an elliptical mask and histogram equalization are used in [3]), and by choosing the following parameters: (i) images divided in 10 (2×5), 40 (4×10) or 80 (4×20) regions, instead of using the original divisions which range from 16 (4×4) to 256 (16×16), and (ii) using the mean square error as the similarity measure, instead of the log-likelihood statistic, in addition to histogram intersection and Chi square.

3.2 Gabor Jet Descriptors

Local-matching approaches for face recognition in the visible spectrum are compared in [39]. The study analyzes several local feature representations, classification methods, and combinations of classifier alternatives. Taking into account the results of their study, the authors implemented a system that integrates the best possible choice at each step. That system uses Gabor Jet Descriptors as local features, which are uniformly distributed over the images, one wave-length apart. In each grid position of the test and gallery images, and at each scale (multiscale analysis), the Gabor jets are compared using normalized inner products, and these results are combined using the Borda-Count method. In the Gabor feature representation, only Gabor magnitudes are used, and 5 scales and 8 orientations of the Gabor filters are adopted. We implemented this system using all parameters described in [39] (i.e. filter frequencies and orientations, grid positions, face image size).

3.3 WLD Histograms

The WLD (Weber Linear Descriptor) descriptor [11] is inspired by Weber's Law, and computes a two dimensional histogram of local intensity (differential excitations) and orientations. First, for a given pixel, the differential excitation component is computed as the ratio between the relative intensity differences of a current pixel against its neighbors, and the intensity of the current pixel descriptor. The orientation component is the gradient orientation of the current pixel. Afterwards, the 2D histogram of differential excitations and orientations is built. Thus, like LBP, WLD is a dense descriptor computed for every pixel. In addition, WLD has much smaller pixel-granularity than SIFT, given that it is computed in a smaller region. In this work, the recognition is performed using a nearest neighbor classifier in the computed feature space using one of the following similarity measures between histograms: histogram intersection, Euclidean distance, or Chi square. As in the case of LBP, images are divided into a variable number of regions (10, 40 or 80), and the obtained histograms for each region are concatenated to obtain the descriptor. The 2D histograms are quantized to 8 orientations and 64 differential excitation values.

3.4 Notation: Methods and Variants

We use the following notation to refer to the methods and their variations: A–B–C. (i) A describes the name of the face-recognition algorithm: LBP—Histogram of LBP features, WLD—Histogram of WLD features, GJD—Gabor Jet Descriptors; (ii) B denotes the similarity measure or classification approach: HI—Histogram Intersection, XS—Chi square, EU—Euclidian Distance, BC—Borda Count; and (iii) C describes additional parameters: number of divisions in the case of the LBP-based and WLD-based methods.

4 Comparative Study Using the UCHThermalFace Database

The methods under study are analyzed considering real-world conditions that include indoor/outdoor setups and natural variations on facial expression, pose, accessories, occlusions, and background.

4.1 Database Description

The UCHThermalFace database (available for download at http://vision.die.uchile.cl/dbThermal/) is divided into three sets, *Rotation, Speech* and *Expressions*. The Rotation and Speech sets consist of indoor and outdoor thermal face images of 53 subjects obtained under different yaw and pitch angles, as well as a set of images captured while the subjects were speaking. The Expressions set consists of thermal images of 102 subjects captured in an indoor setup. The thermal images were acquired using a FLIR 320 TAU Thermal Camera,[2] with sensitivity in the range 7.5–13.5 μm, and a resolution of 324 × 256 pixels.

The Rotation set contains 22 images per subject, each one corresponding to a different rotation angle acquired in indoor and outdoor sessions (see experimental setup in Fig. 1). In both cases, indoor and outdoor, the distance from the subject to the thermal camera was fixed at 120 cm, and the thermal camera was situated at position P6 (see Fig. 1). The face images were acquired while subjects were observing positions 1–11 (see Fig. 1), which correspond to the following rotation angles: R1: (yaw = −15°, pitch = 15°), R2: (yaw = 0°, pitch = 15°), R3: (yaw = 15°, pitch = 15°), R4: (yaw = −30°, pitch = 0°), R5: (yaw = −15°, pitch = 0°), R6: (yaw = 0°, pitch = 0°), R7: (yaw = 15°, pitch = 0°), R8: (yaw = 30°, pitch = 0°), R9: (yaw = −15°, pitch = −15°), R10: (yaw = 0°, pitch = −15°), R11: (yaw = 15°, pitch = −15°). Figure 2 shows an example of 11 thermal face images corresponding to one individual of the database, acquired under different yaw and pitch angles in an indoor session. In addition to the rotation set, a video sequence was captured while each subject was observing point P6 and speaking the word, "Pa-ra-le-le-pi-pe-do", in the indoor and in the outdoor session. Later on, three frames were randomly selected from the video sequence of each individual in each session (indoor and outdoor). These images form the Speech set, which essentially contains images with different facial expressions.

The Expression set is captured in a different setup. In this setup subjects observe frontally the camera at a fixed distance of 150 cm. First, images were acquired while subjects were expressing three different expressions "Happy", "Sad", and "Angry". In addition a video sequence was captured while each subject was speaking different vowels. Later on, three frames were randomly selected from the video sequence of each individual.

[2] http://www.flir.com/cvs/cores/uncooled/products/tau/

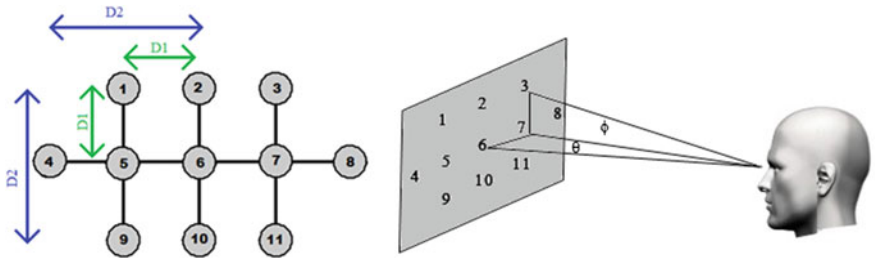

Fig. 1 Image acquisition setup at different yaw (θ) and pitch (ϕ) angles. The distance between the individual and the observed point P6 is 120 cm. D1 is 32.15 cm and D2 is 69.28 cm. See text for details

Fig. 2 Example of the 11 thermal face images of the rotation set of one individual of the UCHThermalFace database, captured at the indoor session

In summary, for the Rotation and Speech sets, 14 indoor and 14 outdoor subsets are defined in order to carry out face recognition experiments. For the indoor session and the outdoor session, 11 subsets correspond to the different yaw-pitch combinations of the Rotation set (subsets R1 to R11), and 3 to the different images captured in the Speech set (subsets S1 to S3). In the Expressions sets, 3 expressions ("Happy", "Sad", and "Angry"; E1 to E3) and 3 vowels (V1 to V3) subsets are defined. The experiments reported in the next section make use of the 34 defined subsets. In each experiment a given subset is used as a test set, and a second one as a gallery set.

All images for a given subject of the indoor subset were captured during the same session. The same applies for the outdoor session. The indoor subset and the outdoor subset were captured in different days. It is important to note that this does not make the problem less interesting because of the following reasons: (1) in some applications the subjects may need to be recognized during the same session (e.g. in the case of service robots), (2) in the case of thermal images, the appearance of the subject does not change much over time—further analysis is part of future work, (3) in the case of thermal images, the clothes do not bring much information

(in the results shown below, in most cases best results are obtained when clothing information is not available), and (4) for other existing databases, such as the Equinox database, the images were also captured during a single session. The main features affecting the recognition are features such as moustaches, expressions, eyeglasses and haircuts; this can be partly simulated by the occlusion and expression results presented below. Also note that even if the images where captured during the same session, the recognition rates are nevertheless low, in particular for high rotation angles.

4.2 Description of Experiments

In order to evaluate the face recognition methods under analysis, six kinds of experiments were carried out: (i) variable window size, (ii) partial face occlusions, (iii) indoor versus outdoor galleries, (iv) facial expressions, and (v) variable distance. In all experiments face images are aligned using the annotated eye position; faces are aligned by centering the eyes in the same relative positions, at a fixed distance of 42 pixels between the eyes, except in the case of the variable distance experiments where the distanced between the eyes is decreased accordingly with the reduction in the image resolution. The experiments are:

Variable Window Size. The effect of using different window sizes in the performance of the methods is analyzed. Increasing the size of the windows corresponds to adding or removing different amounts of background to the region being analyzed, given that we are not decreasing the scale of the faces (there is no change in image resolution). Thus, experiments were performed including window sizes of 81×150 pixels, 100×185 pixels, and 125×225 pixels (see Fig. 3 for examples).

Partial Face Occlusions. In order to analyze the behavior of the different methods in response to partial occlusions of the face area, images were divided into 10 different regions (2 columns and 5 rows), and one of the regions was randomly selected and its pixels set to 0 (see Fig. 4 for some examples). This experiment analyzes the behavior of the different methods in response to these partial occlusions.

Indoor versus Outdoor Galleries. The performance of face recognition methods depends largely on environmental conditions, particularly under the indoor or outdoor conditions. In these experiments the test and gallery images correspond to images taken in an indoor session or in an outdoor session. When the test images are indoor images, then the gallery images are outdoor images, and vice versa. The outdoor images were captured in summer (with high temperatures up to 30 °C), and at times the faces, as well as the camera, were receiving direct sunlight.

Facial Expressions. The effect of having different facial expressions in the subjects is analyzed. The experiments consider subjects with different facial expressions, as well as subjects speaking different vowels.

Variable Distance. The sensitivity of the methods to the distance between the subject and the camera is analyzed. By decreasing the resolution of the thermal

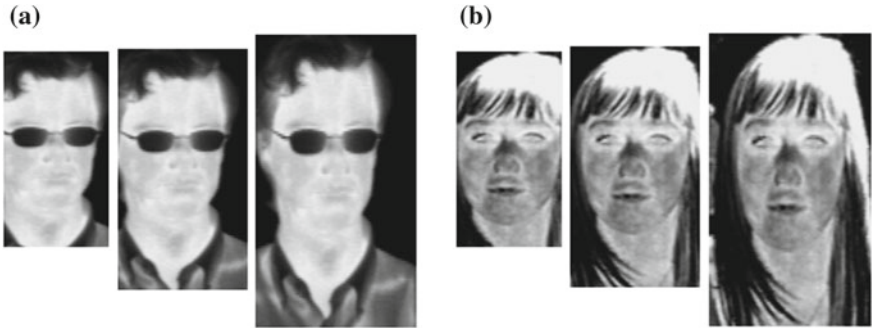

Fig. 3 Examples of faces with different cropping sizes (UCHThermalFace database). **a** Indoor session: window size (in pixels): 81 × 150, 100 × 185, 125 × 225. **b** Outdoor session: window size (in pixels): 81 × 150, 100 × 185, 125 × 225

Fig. 4 Example of Images with Partial Occlusion (UCHThermalFace database). *Left* Indoor session: window size 125 × 225. *Right* Outdoor session: window size 125 × 225

images, the effect of having variable distances between the subject and the camera is simulated.

The Expressions set (E1–E3, V1–V3) is used in the facial expressions experiments, while in all other experiments the Rotation and Speech datasets are used (R1–R11, S1–S3).

4.3 Recognition Results

The performance of the different methods is evaluated using the top-1 recognition rate. In all experiments the rotation subset R6, without any occlusion, is selected as a gallery set because it contains clean frontal views of the faces. Naturally, in the indoor experiments the indoor R6 subset is used, while in the outdoor experiments the outdoor R6 subset is employed.

In Tables 1, 2, 3 and 4, the top-1 recognition rate is given separately for the two different sets, Rotation and Speech. For each category a mean value is calculated: the mean recognition rate over all rotation subsets for the rotation set, and the mean recognition rate over all speech subsets. In addition, the average between these two results is given. Table 7 reports average results obtained using the Expressions set.

Variable Window Size. Table 2 shows the performance of the different methods when different window sizes are used in the test and gallery sets. In these experiments all gallery and test sets correspond to indoor images.

It can be concluded from the experiments that the best window size depends on the method. For GJD-BC the best size is 125×225, while for WLD-X-X and LBP-X-X the best size is 81×150. These results are consistent with the ones obtained in [32] for GJD-BC and LBP-X-X methods. As noted in that study, GJD-BC works better with windows that contain large portions of background. The reason is three-fold: (i) the Gabor-filters encode information about the contour of the face, (ii) large regions allow the use of large filters, which encode large-scale information, and (iii) the Borda Count classifier may reduce the effect of regions (or jets) that are not relevant for the recognition. On the contrary, for the LBP-X-X variants, adding some background does not help, and in most cases reduces the performance. We can observe that WLD variants have similar behavior compared to that of LBP variants in terms of optimal window size. This seems to be due to the similar kind of analysis of the face information that both methods carry out, which basically computes histograms of local features over non-overlapping regions of the face area. In the cases where the histograms include background information (as in larger croppings of the image), the discriminability of the methodology decreases.

In terms of the best overall performance, WLD-X-X methods variants are the ones with the highest top-1 recognition rates, followed GJD-BC and LBP-X-X obtain similar performance.

GJD-BC variants are also robust to facial expressions. One of the variants (GJD-BC, 100×185 pixels) obtains a 100 % top-1 recognition rate in the experiments reported in Table 1. This robustness is achieved thanks to the use of an ensemble of classifiers, which is implemented by the Borda Count voting of the Gabor-jets. WLD-X-X and LBP-X-X variants also show some invariance to facial expression variations. When a window size of 81×150 pixels is used, they achieve a top-1 recognition rate of 95.6 and 92.5 %, respectively.

For each method, the best results in top-1 recognition rate are obtained by the following variants (considering different parameters and window sizes):

- WLD-EU-80, 81×150 pixels: 91.8 %
- WLD-HI-80, 81×150 pixels: 91.5 %
- GJD-BC, 125×225 pixels: 91.4 %
- LBP-HI-80, 81×150 pixels: 88.5 %

In the case of the WLD two variants are selected, either because the performance of the two variants is very similar. For the next experiments, results for only these four variants are presented and analyzed.

Table 1 Experiment using different windows sizes

Methods	Rotation (%)												Speech	Average (%)
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	Mean (%)	Mean (%)	
LBP-EU-80 (81 × 150)	54.72	77.36	54.72	50.94	81.13	100	90.57	39.62	73.58	92.45	84.91	72.73	81.76	77.24
LBP-HI-80 (81 × 150)	**73.58**	**88.68**	**73.58**	**64.15**	**96.23**	**100**	**96.23**	**50.94**	**90.57**	**100**	**96.23**	**84.56**	**92.45**	**88.51**
LBP-XS-80 (81 × 150)	64.15	90.57	73.58	62.26	88.68	100	92.45	41.51	86.79	100	81.13	80.10	89.94	85.02
GJD-BC (81 × 150)	66.04	96.23	75.47	50.94	92.45	100	88.68	33.96	81.13	98.11	69.81	77.53	98.11	87.82
WLD-EU-80 (81 × 150)	**77.36**	**96.23**	**90.57**	**64.15**	**96.23**	**100**	**92.45**	**71.70**	**90.57**	**98.11**	**90.57**	**87.99**	**95.60**	**91.80**
WLD-HI-80 (81 × 150)	**73.58**	**96.23**	**88.68**	**64.15**	**94.34**	**100**	**96.23**	**66.04**	**90.57**	**100**	**92.45**	**87.48**	**95.60**	**91.54**
WLD-XS-80 (81 × 150)	73.58	94.34	90.57	64.15	88.68	100	92.45	62.26	92.45	96.23	84.91	85.42	94.97	90.20
LBP-EU-80 (100 × 185)	45.28	75.47	41.51	45.28	81.13	100	69.81	30.19	77.36	90.57	66.04	65.69	81.76	73.73
LBP-HI-80 (100 × 185)	67.92	86.79	67.92	54.72	98.11	100	92.45	39.62	94.34	100	84.91	80.62	93.08	86.85
LBP-XS-80 (100 × 185)	71.70	90.57	62.26	56.60	96.23	100	94.34	33.96	90.57	100	81.13	79.76	91.82	85.79
GJD-BC (100 × 185)	83.02	98.11	84.91	50.94	94.34	100	96.23	37.74	79.25	98.11	79.25	81.99	100	91.00
WLD-EU-80 (100 × 185)	79.25	90.57	79.25	54.72	88.68	100	94.34	47.17	86.79	100	84.91	82.33	91.82	87.08
WLD-HI-80 (100 × 185)	69.81	88.68	81.13	56.60	92.45	100	96.23	45.28	90.57	100	86.79	82.50	92.45	87.48
WLD-XS-80 (100 × 185)	81.13	92.45	84.91	60.38	96.23	100	92.45	49.06	90.57	98.11	94.34	85.42	92.45	88.94
LBP-EU-80 (125 × 225)	39.62	73.58	37.74	24.53	66.04	100	62.26	24.53	58.49	84.91	49.06	56.43	81.76	69.10
LBP-HI-80 (125 × 225)	58.49	84.91	58.49	37.74	86.79	100	83.02	39.62	79.25	96.23	75.47	72.73	86.16	79.44
LBP-XS-80 (125 × 225)	60.38	90.57	58.49	39.62	90.57	100	90.57	39.62	84.91	98.11	79.25	75.64	90.57	83.11
GJD-BC (125 × 225)	**84.91**	**98.11**	**84.91**	**49.06**	**96.23**	**100**	**92.45**	**41.51**	**86.79**	**100**	**83.02**	**83.36**	**99.37**	**91.37**
WLD-EU-80 (125 × 225)	62.26	92.45	69.81	49.06	88.68	100	84.91	33.96	79.25	96.23	81.13	76.16	89.94	83.05
WLD-HI-80 (125 × 225)	58.49	88.68	71.70	47.17	90.57	100	92.45	41.51	88.68	100	86.79	78.73	93.08	85.91
WLD-XS-80 (125 × 225)	67.92	94.34	73.58	50.94	92.45	100	94.34	47.17	88.68	98.11	86.79	81.30	95.60	88.45

Indoor session. Top-1 recognition rate Rotation and Speech test sets. The best variant of each method under analysis is presented in bold type

Table 2 Best results windows size

Methods	Rotation (%)												Speech	Average (%)
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	Mean (%)	Mean (%)	
LBP-HI-80 (81 × 150)	86.79	86.79	83.02	60.38	98.11	100	96.23	73.58	88.68	98.11	90.57	87.48	90.57	89.03
GJD-BC (125 × 225)	62.26	98.11	71.70	39.62	88.68	100	94.34	35.85	60.38	96.23	67.92	74.10	86.79	80.45
WLD-EU-80 (81 × 150)	62.26	73.58	79.25	47.17	84.91	100	84.91	58.49	77.36	84.91	64.15	74.27	72.96	73.62
WLD-HI-80 (81 × 150)	69.81	86.79	92.45	60.38	96.23	100	94.34	66.04	84.91	96.23	79.25	84.22	96.23	90.23

Outdoor session. Top-1 recognition rate. Rotation and Speech test sets

Table 3 Partial occlusion; Indoor session

Methods	Rotation (%)												Speech	Average (%)
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	Mean (%)	Mean (%)	
LBP-HI-80 (81 × 150)	60.38	77.36	62.26	58.49	96.23	100	90.57	47.17	83.02	100	81.13	77.87	86.79	82.33
GJD-BC (125 × 225)	71.70	86.79	66.04	35.85	71.70	100	83.02	26.42	66.04	92.45	64.15	69.47	93.71	81.59
WLD-EU-80 (81 × 150)	75.47	96.23	86.79	62.26	96.23	100	92.45	71.70	84.91	98.11	90.57	86.79	93.08	89.94
WLD-HI-80 (81 × 150)	75.47	94.34	86.79	62.26	94.34	100	94.34	71.70	90.57	100	90.57	87.31	96.23	91.77

Top-1 recognition rate. Rotation and Speech test sets

Table 4 Partial occlusion; Outdoor Session

Methods	Rotation (%)												Speech	Average (%)
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	Mean (%)	Mean (%)	
LBP-HI-80 (81 × 150)	73.58	73.58	75.47	52.83	90.57	100	79.25	52.83	81.13	92.45	67.92	76.33	81.13	78.73
GJD-BC (125 × 225)	49.06	79.25	52.83	24.53	81.13	100	77.36	24.53	56.60	94.34	49.06	62.61	83.02	72.82
WLD-EU-80 (81 × 150)	52.83	71.70	71.70	39.62	81.13	100	81.13	47.17	64.15	81.13	50.94	67.41	71.70	69.56
WLD-HI-80 (81 × 150)	73.58	88.68	90.57	60.38	94.34	100	90.57	66.04	86.79	92.45	79.25	83.88	87.42	85.65

Top-1 recognition rate. Rotation and Speech test sets

Table 2 shows the performance of the four best variants of the methods for the defined Rotation and Speech sets in the outdoor case. Best results are obtained by WLD-HI-80. Second best performance is obtained by LBP-EU-80, and the third place is taken by GJD-BC. In the case of WLD-HI-80 and LBP-HI-80, the performance is also similar in the indoor and outdoor cases. Interestingly, in the case of WLD-EU-80, a variant that uses the Euclidian distance, the performance decreases greatly in the outdoor case. The same happened for GJD-BC, whose performance decreased about 11 percentage points in the outdoors setting. These results indicate that GJD-BC does not behave appropriately in outdoor conditions.

Partial Face Occlusions. Table 3 shows the methods' sensitivity to partial occlusions of the face area for the indoor case, while Table 4 shows the method's sensitivity for the outdoor case. In the indoor and outdoor case we observe that the best performance by WLD-X-80. In both cases, the top-1 recognition rate decreases by about 2–5 % with occlusions of 10 % of the face area, which is considered very good behavior. LBP-HI-80 and GJD-BC show lower performance, and they are more affected by occlusions (the top-1 recognition rate decreases by 6–11 %).

Indoor versus Outdoor Galleries. In Table 5 face recognition experiments that use an indoor gallery set together with outdoor test sets are reported. Conversely, in Table 6 experiments that use an outdoor gallery set together with indoor test sets are reported. It can be clearly observed that in these cases all methods under comparison decrease their performance dramatically compared with previous experiments. In all cases the top-1 recognition rate is very low. Interestingly, the best performance is achieved by GJD-BC.

Facial Expressions. In Table 7 face recognition experiments that use the Expressions sets are reported. The table shows the average top-1 recognition rate of six experiments. In each experiment one set is chosen as gallery and one as test. The sets are E1–E3 and V1–V3. It can be observed that best results are obtained by LBP-X-X, followed by GJD-BC and WLD-X-X.

Variable Distance. In Table 8 face recognition experiments that correspond to the following subject-camera distances 1.2, 1.69, 2.4, 3.39 and 4.8 m are shown. In these experiments the image resolution decreases with the distance in a factor of $\sqrt{2}$. For instance, in the case of GJD-BC, where the distance is 1.69 m, the corresponding resolution is 88x159 pixels. It can be observed that all methods are robust to variations in the subject-camera distance. The most robust method is and LBP-X-X, which is almost not affected by changes in resolution, in particular for higher distances (lower resolutions).

From Tables 1 and 2 it can be observed that for all methodologies the performance decreases as the yaw rotation increases. For low rotations ($\pm 15°$), in the indoor case (Fig. 2a) the performance of all methods is very similar, while in the outdoor case GJD-BC and WLD-EU-80 clearly show a lower performance than other methods, methods which all have a similar recognition rate. All other methods show worst performance for large rotations, with WLD-HI-80 working.

Table 5 Different gallery sets: indoor gallery set, outdoor test sets

Methods	Rotation (%)												Speech	Average (%)
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	Mean (%)	Mean (%)	
LBP-HI-80 (81 × 150)	24.53	30.19	16.98	20.75	32.08	32.08	16.98	9.43	22.64	20.75	11.32	21.61	22.01	21.81
GJD-BC (125 × 225)	30.19	39.62	24.53	18.87	35.85	54.72	30.19	7.55	28.30	33.96	16.98	29.16	42.77	35.97
WLD-EU-80 (81 × 150)	22.64	28.30	16.98	9.43	15.09	20.75	11.32	13.21	16.98	16.98	11.32	16.64	22.01	19.33
WLD-HI-80 (81 × 150)	28.30	35.85	28.30	13.21	33.96	33.96	18.87	11.32	26.42	35.85	18.87	25.9	32.08	28.99

Top-1 recognition rate. Rotation and Speech test sets

Table 6 Different gallery sets: outdoor gallery set, indoor test sets

Methods	Rotation (%)												Speech	Average (%)
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	Mean (%)	Mean (%)	
LBP-HI-80 (81 × 150)	13.21	15.09	3.77	11.32	24.53	24.53	13.21	7.55	24.53	24.53	15.09	16.12	20.13	18.13
GJD-BC (125 × 225)	22.64	28.30	13.21	5.66	22.64	35.85	24.53	5.66	24.53	33.96	22.64	21.78	35.85	28.82
WLD-EU-80 (81 × 150)	15.09	20.75	13.21	15.09	30.19	24.53	20.75	11.32	22.64	22.64	15.09	19.21	25.79	22.50
WLD-HI-80 (81 × 150)	15.09	16.98	11.32	9.43	18.87	18.87	20.75	15.09	16.98	22.64	18.87	16.81	20.13	18.47

Top-1 recognition rate. Rotation and Speech test sets

Table 7 Gallery/test Expressions sets: E1–E3, V1–V3

Methods	Average (%)
LBP-HI-80	94.9
GJD-BC	94.4
WLD-EU-80	93.9
WLD-HI-80	94.1

Average top-1 recognition rate of six experiments, in which 1 set is chosen as gallery and 1 as test

Table 8 Gallery set: R6; Test set: S1

Methods	Mean (%)				
	1.2 m	1.69 m	2.4 m	3.39 m	4.8 m
LBP-HI-80. Initial resolution (81 × 150)	92.5	94.3	94.3	92.5	92.5
GJD-BC. Initial resolution (125 × 225)	100	100	100	94.3	88.7
WLD-EU-80. Initial resolution (81 × 150)	96.2	98.1	92.5	86.8	66.0
WLD-HI-80. Initial resolution (81 × 150)	96.2	96.2	94.3	94.3	67.9

Top-1 recognition rate. The image resolution decreases with the distance in a factor of $\sqrt{2}$

Tables 9, 10 and 11[3] present results on the fusion of some of the presented methods and variants in the UCHThermalFace in the Indoor subset. The fusion was done at the ranking level using the following procedure: first the ranking of each of the considered methods was calculated and afterwards these rankings were combined. The fusion was done in three different ways as reported in the tables: using the average of the rankings, the minimum of the rankings, and the maximum of the rankings. After the rakings are fused, the top-1 recognition rates are obtained. From the tables the following can be observed:

- Using the average as fusion method gives best results in most cases, followed by fusing the methods using them maximum. Worst results are obtained using the minimum.
- In all cases using the minimum gives better results than using the worst performing single method (i.e. without using fusion).
- In the cases that the WLD descriptor is used together with the XS distance (WLD-XS-80 81 × 150), the fusion using the maximum and the fusion using the average give better results than in the cases when using any of variants alone. The observed improvement ranges from 1 to 5 % when compared to the best of the methods being combined, and from 5 to 9 % when compared to the worst of the methods being combined.
- Best results are obtained when combining WLD-XS-80 81 × 150 with WLD-HI-80 81 × 150. This is an interesting result because this could de implemented efficiently as the feature evaluation must be performed only once.

[3] In Tables 9, 10 and 11, unlike in previous tables, the average top-1 recognition results do consider using the set R6 as a test set.

Table 9 Method fusion: WLD and LBP

Input methods		Fusion		
WLD-X-80 (81 × 150)	LBP-X-80 (81 × 150)	Avg	Min	Max
EU: 86.79	EU: 70.00	77.17	79.06	76.98
EU: 86.79	HI: 83.02	86.42	85.09	85.85
EU: 86.79	XS: 78.11	83.58	82.45	83.40
HI: 86.23	EU: 70.00	77.74	79.25	77.17
HI: 86.23	HI: 83.02	85.28	83.96	84.91
HI: 86.23	XS: 78.11	82.08	82.08	82.08
XS: 83.96	EU: 70.00	81.32	77.74	80.38
XS: 83.96	HI: 83.02	**88.68**	83.77	**88.30**
XS: 83.96	XS: 78.11	**87.74**	80.19	**86.60**

Gallery set: R6, Test Set: Rotation sets. Average top-1 recognition rates over all subsets. In bold are shown the cases where the fusion gives better results than both input methods
Avg fusion by averaging the ranking of the two input methods. *Min* fusion by taking the minimum of the ranking of the two input methods. *Max* fusion by taking the maximum of the ranking of the two input methods

Table 10 Method fusion: WLD variants

Input methods		Fusion		
WLD-X-80 (81 × 150)	WLD-X-80 (81 × 150)	Avg	Min	Max
EU: 86.79	HI: 86.23	**87.36**	86.42	**87.74**
EU: 86.79	XS: 83.96	86.04	85.09	86.23
HI: 86.23	XS: 83.96	**89.43**	84.91	**89.43**

Gallery set: R6, Test Set: Rotation sets. Average top-1 recognition rates over all subsets. In bold are shown the cases where the fusion gives better results than both input methods
Avg fusion by averaging the ranking of the two input methods. *Min* fusion by taking the minimum of the ranking of the two input methods. *Max* fusion by taking the maximum of the ranking of the two input methods

Table 11 Method Fusion: LBP variants

Input methods		Fusion		
LBP-X-80 (81 × 150)	LBP-X-80 (81 × 150)	Avg	Min	Max
EU: 70.00	HI: 83.02	74.53	77.55	73.96
EU: 70.00	XS: 78.11	72.83	74.72	72.26
HI: 83.02	XS: 78.11	79.62	80.38	79.62

Gallery set: R6, Test Set: Rotation sets. Average top-1 recognition rates over all subsets
Avg fusion by averaging the ranking of the two input methods. *Min* fusion by taking the minimum of the ranking of the two input methods. *Max* fusion by taking the maximum of the ranking of the two methods

4.4 Computational Performance

The speed of the recognition process is an important constrain in many face recognition applications (e.g. Human-Robot-Interaction or identity identification using large

face databases). For this reason, in this section we present a comparative analysis of the selected methods in terms of processing time. In order to achieve this, we have evaluated the time required for feature extraction (FET: Feature Extraction Time), the time required for matching two feature vectors (MT: Matching Time), and the total processing time (PT: Processing time) required to recognize a face depending on the size of the database. Note that in the case of GJD, the total processing time is not linear but $n \log n$ on the size of the database. This is because of the way in which the Borda Count classifier works (After our experience the $\log n$ factor is relevant for large gallery databases, for example more than 1,000 images). All other methods are linear on the size of the gallery. Note also that in this analysis we are only considering the time required during operation, and not the time required to create the database.

The experiments were carried out on a computer running Windows 7 Ultimate (64-bits) with an Intel Core 2 duo CPU T5870 @2.00GHz GHz (4GB RAM) processor. For all methods, we used our own C/C++ implementations compiled as 32-bit applications.

Table 12 shows the computed processing times of all methods under comparison in terms of feature extraction, and matching and processing times. In terms of feature extraction, LBP-X-80 is the fastest method, followed closely by WLD-X-80. The third fastest method, GJD-BC, has a feature extraction one order of magnitude slower than LBP-HI-80 and WLD-X-80.

In terms of Matching Time MT (time for pairs of images), the fastest methods are WLD-X-80, GJD-BC, and LBP-X-80, all of them with MT lower than 1 ms. When we consider the total processing time of the methods (PT), the method with the shortest processing time is LBP-X-80, independently of the size of the database. The second fastest methods are the WLD-X-80 variants. For large databases of large sizes (1,000 images in gallery), WLD-EU-80 and WLD-HI-80 are almost two times slower than LBP-X-80. GJP-BC is 4 to 20 times slower than LBP-X-80. In summary, there is a clear distinction on the speed of the methods, with LBP and WLD being the fastest methods, while GJD-BC standing in last position.

Table 12 Processing time

Method	FET (ms)	MT (ms)	PT (FET+MT) (ms)			
			1	10	100	1000
LBP-X-80 (81 × 150)	2.6	<1	3.6	3.6	12.9	122.3
WLD-EU-80 (81 × 150)	3.9	<1	4.9	5.9	23.9	202.9
WLD-HI-80 (81 × 150)	3.9	<1	4.9	6.9	26.9	229.9
WLD-XS-80 (81 × 150)	3.9	<1	4.9	4.9	16.9	130.9
GJD-BC (125 × 225)	95.94	<1	96.9	96.9	127.9	532.9

Time measures are in milliseconds. The experiments are those we carried out on a computer running Windows 7 Ultimate with an Intel Core 2 duo CPU T5870 @2.00GHz GHz (4GB RAM) processor *FET/MT*: feature extraction/matching time (ms). *PT*: processing time (ms). DB sizes (gallery) of 1, 10, 100 and 1,000 faces are included

5 Discussion and Conclusions

In this article, a comparative study of thermal-based face-recognition methods in unconstrained environments was presented. The analyzed methods were selected by considering their suitability for the defined requirements—real-time operation, just one image per person, fully online (no training), and robust behavior in unconstrained environments—and their performance in former studies. The comparative study was carried out using UCHThermalFace database. This database includes aspects such as yaw and pitch rotations, environment condition variations (indoor/outdoor) and facial expressions. In addition, occlusions were simulated. The methods under comparison are LBP histograms, Gabor Jet descriptors, and WLD histograms. Comments on the main results of this study, and some conclusions drawn from this work, follow.

Comments on the Size of the Face Region. Unlike the results of [32] for visible images, here the dependence of the methods on the size of the images is not large, and for the best working methods, the effect is rather low. This seems to be due to the mostly uniform background observed for thermal images, which in addition allows the methods to use more information about the contour of the face.

Comments on Alignment, Occlusions and Expressions. From our experiments we conclude that to a large degree only some of the analyzed methods are robust to inaccurate alignment, face occlusions, and variations in expressions. Accepting that these factors affect the face recognition process, their influence in the methods' performance is much lower than outdoor conditions or pose variations.

Comments on the Indoor/Outdoor Conditions. Most of the methods behave very well in natural, indoor conditions, as well as in outdoor conditions, with the one exception of GJD-BC, whose performance decreases considerably under outdoor conditions. This aspect should be further analyzed with additional experiments. In experiments where the test images are acquired in an outdoor setup and the gallery images are acquired in an indoor setup, or vice versa, the performance of all methods is very low. The reason seems to be the very different range of pixel values of thermal images acquired in indoor setups compared to images acquired in outdoor settings. The saturation observed in images acquired under outdoor conditions, because of the heat of the environment, may be the main reason for this. This eventually could be improved by a better calibration of the camera, or by using normalization algorithms, as in the case of images of the visible spectrum [31]. This aspect needs to be further analyzed.

Conclusions about the Fusion of the Methods. Fusion the methods give good results in all cases. It is interesting to note that in all cases fusing the methods gives better results than the worst performing of each the combined method, which indicate that the method can complement. We also think that it is rather interesting that even fusing different variants of WLD (variants that use the same features but different distance measures), the results also improve. This show that the distance measures being used could be improved, and that combining two or more distance measures could give even better results; this is left as future work.

Conclusions about the Performance of Methods. The question of which method is the best is a very difficult one. However, we could say that WLD-based methods are a good choice if real-time operation is needed as well as good recognition rates. LBP is fastest method, and it has good and stable performance which makes it a good choice is time constrains are present or if the size of the gallery is large. Also LBP is the most robust method to resolution changes, with the performance remaining almost the same for all tested resolutions—unlike the other tested methods.

Future Work. We believe that there is still room for improvement in many aspects of recognition of faces in unconstrained environments. The main open questions in the case of thermal images are: (i) how to achieve invariance to environment conditions such as temperature and saturation of the images because of the heating of the camera produced by long periods of operation or environmental conditions (e.g. outdoor operation under direct sun exposure), (ii) how to combine the use of different methods in order to achieve, at the same time, high recognition rates and processing speed, (iii) what is the influence of face resolution in the recognition process, (iv) what can be learned through a deeper analysis of the facial expression effect in the recognition of faces, (v) how to develop fast methods that are robust to rotations, and (vi) how to combine the output (ranking and features) of different methods. In addition, we will extend our comparative study by incorporating methods that use vein information in the recognition process (e.g. [17]), apply new texture classification methods (e.g. [21, 22]) in face recognition using visible and thermal images, as well as explore the use of information fusion approaches in multimodal recognition (e.g. [16]).

Acknowledgments This research was partially funded by the FONDECYT-Chile grant 1090250, by the FONDECYT-Chile grant 3120218, and by the Advanced Mining Technology Center.

References

1. Abidi, B., Huq, S., Abidi, M.: Fusion of visual, thermal, and range as a solution to illumination and pose restrictions in face recognition. In: Proceedings of IEEE Carnahan Conference on Security Technology, pp. 325–330, Albuquerque, NM (2004)
2. Ahmad, J., Ali, U., Qureshi, R.J.: Fusion of thermal and visual images for efficient face recognition using gabor filter. In: The 4th ACS/IEEE International Conference on Computer Systems and Applications, pp. 135–139, March 8–11. Dubai/Sharjah, UAE (2006)
3. Ahonen, T., Hadid, A., Pietikainen, M.: Face Description with local binary patterns: application to face recognition. IEEE Trans. Pattern Anal. Mach. Intell. **28**(12), 2037–2041 (2006)
4. Akhloufi, M., Bendada, A.: Thermal faceprint: a new thermal face signature extraction for infrared face recognition, In: CRV 2008: Fifth Canadian Conference on Computer and Robot Vision (Windsor, Ontario), pp. 269–272, 28–30 May 2008
5. Bebis, G., Gyaourova, A., Singh, S., Pavlidis, I.: Face recognition by fusing thermal infrared and visible imagery. Image Vis. Comput. **24**(7), 727–742 (2006)
6. Bhowmik, M.K., Bhattacharjee, D., Nasipuri, M., Basu, D.K., Kundu, M.: Optimum fusion of visual and thermal face images for recognition. In: Sixth International Conference on Information Assurance and Security (IAS 2010), pp. 311–316. IEEE Intelligent Transportation Systems Society, Atlanta, USA, 23–25 August 2010

7. Buddharaju, P., Pavlidis, I.: Multi-spectral face recognition—fusion of visual imagery with physiological information. In: Hammoud, R.I., Abidi, B.R., Abidi, M.A. (eds.) Face Biometrics for Personal Identification: Multi-Sensory Multi-Modal Systems, pp. 91–108. Springer, Berlin (2007)
8. Buddharaju, P., Pavlidis, I., Manohar, C.: Face Recognition Beyond the Visible Spectrum, Advances in Biometrics: Sensors, Algorithms and Systems, pp. 157–180. Springer, Berlin (2007)
9. Buddharaju, P., Pavlidis, I., Kakadiaris, I.: Physiology-based face recognition. In: Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, pp. 354–359. Lake Como, Italy (2005)
10. Buddharaju, P., Pavlidis, I., Kakadiaris, I.: Pose-invariant physiological face recognition in the thermal infrared spectrum. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 53–60. New York, USA (2006)
11. Chen, J., Shan, S., He, Ch., Zhao, G., Pietikäinen, M., Chen, ChX, Gao, W.: WLD: a robust local image descriptor. IEEE Trans. Pattern Anal. Mach. Intell. **32**(9), 1705–1720 (2010)
12. Chen, X., Flynn, P., Bowyer, K.W.: PCA-based face recognition in infrared imagery: baseline and comparative studies. In: IEEE International Workshop on Analysis and Modeling of Faces and Gestures, pp. 127–134. Nice, France (2003)
13. Cho, S.Y., Wang, L., Ong, W.L.: Thermal imprint feature analysis for face recognition. In: IEEE International Symposium on Industrial Electronics 2009, ISIE 2009, pp. 1875–1880 (2009)
14. Desa, S., Hati, S.: IR and visible face recognition using fusion of kernel based features. In: The 19th International Conference on Pattern Recognition (ICPR 2008), pp. 1–4. Tampa, Florida, USA, 8–11 December 2008
15. Equinox, 2011. Equinox Database. http://www.equinoxsensors.com/products/HID.html. Accessed July 2011
16. He, M., Horng, S.-J., Fan, P., Run, R.-S., Chen, R.-J., Lai, J.-L., Khan, M., Sentosa, K.: Performance evaluation of score level fusion in multimodal biometric systems. Pattern Recognit. **43**(5), 1789–1800 (2010)
17. Hermosilla, G.,Loncomilla, P., Ruiz-del-Solar, J.: Thermal Face Recognition using Local Interest Points and Descriptors for HRI Applications. Lecture Notes in Computer Science, vol. 6556 (RoboCup, Symposium 2010), pp. 25–35
18. Hermosilla, G., Ruiz-del-Solar, J., Verschae, R., Correa, M.: Face recognition using thermal infrared images for human-robot interaction applications: a comparative study. In: The 6th IEEE Latin American Robotics Symposium—LARS 2009, Valparaíso, Chile (CD Proceedings), 29–30 October
19. Hermosilla, G., Ruiz-del-Solar, J., Verschae, R., Correa, M.: A comparative study of thermal face recognition methods in unconstrained environments. Pattern Recognit. **45**(7), 2445–2459. ISSN 0031–3203 (July 2012). doi:10.1016/j.patcog.2012.01.001
20. Kwon, O.K., Kong, S.G.: Multiscale fusion of visual and thermal images for robust face recognition. In: Proceedings of IEEE International Conference on Computer Intelligence for Homeland Security and Personal Safety, vol. IV, pp. 112–116, FL, March 2005
21. Lan, W.: Face recognition system based on spatial constellation model and support vector machine. Master's Thesis, Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology (2010)
22. Lategahn, H., Gross, S., Stehle, T., Aach, T.: Texture classification by modeling joint distributions of local patterns with Gaussian mixtures. IEEE Trans. Image Processing **19**(6), 1548–1557 (2010)
23. Lei, Z., Li, S.Z.: Fast multi-scale local phase quantization histogram for face recognition. Pattern Recognit. Lett. **33**(13), 1761–1767. ISSN 0167–8655 (2012). doi:10.1016/j.patrec.2012.06.005
24. Lowe, D.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
25. Mendez, H., San Martín, C., Kittler, J., Plasencia, Y., García, E.: Face recognition with LWIR imagery using local binary patterns. LNCS, vol. 5558, pp. 327–336 (2009)

26. Nanni, L., Lumini, A., Brahnam, S.: High performance set of features for biometric data. Int. J. Autom. Ident. Technol **2**(1), 1–7 (2010)
27. Narendra, P.: Reference-free nonuniformity compensation for IR imaging arrays. Proc. SPIE **252**, 10–17 (1980)
28. Narendra, P., Foss, N.: Shutterless fixed pattern noise correction for infrared imaging arrays. Proc. SPIE **282**, 44–51 (1981)
29. Pop, F.M., Gordan, M., Florea, C., Vlaicu, A.: Fusion based approach for thermal and visible face recognition under pose and expressivity variation. In: Roedunet Int. Conf. (RoEduNet), pp. 61–66 (2010)
30. Ruiz-del-Solar, J., Navarrete, P.: Eigenspace-based face recognition: a comparative study of different approaches. IEEE Trans. Syst. Man Cybernet.Part C **35**(3), 315–325 (2005)
31. Ruiz-del-Solar, J., Quinteros, J.: Illumination compensation and normalization in eigenspace-based face recognition: a comparative study of different pre-processing approaches. Pattern Recognit. Lett. **29**(14), 1966–1979 (2008)
32. Ruiz-del-Solar, J., Verschae, R., Correa, M.: Recognition of faces in unconstrained environments: a comparative study. EURASIP Journal on Advances in Signal Processing, special issue, Recent Advances in Biometric Systems: A Signal Processing Perspective, vol. 2009, Article ID 184617, 19 pages (2009). doi:10.1155/2009/184617
33. Selinger, A., Socolinsky, D.: Appearance-based facial recognition using visible and thermal imagery: a comparative study. Tech. Rep., Equinox Corporation (2001)
34. Socolinsky, D., Selinger, A.: A comparative analysis of face recognition performance with visible and thermal infrared imagery. In: Proceedings of the International Conference on Pattern Recognition (ICPR), Quebec, Canada (2002)
35. Socolinsky, D., Selinger, A.: Thermal face recognition over time. In: Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), vol. 4, pp. 187–190 (2004)
36. Socolinsky, D., Wolff, L., Neuheisel, J., Eveland, C.: Illumination invariant face recognition using thermal infrared imagery. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2001)
37. Tan, X., Chen, S., Zhou, Z.H., Zhang, F.: Face recognition from a single image per person: a survey. Pattern Recognit. **39**, 1725–1745 (2006)
38. Yoshitomi, Y., Sung-Ill K., Kawano, T., Kilazoe, T.: Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face. In: Proceedings of 9th IEEE International Workshop on Robot and Human Interactive Communication, pp. 178–183 (2000)
39. Zou, J., Ji, Q., Nagy, G.: A comparative study of local matching approach for face recognition. IEEE Trans. Image Process. **16**(10), 2617–2628 (2007)