Thermal Face Recognition over Time using Sparse Representation Approach

José Luis Verdugo*, Gabriel Hermosilla*, Gonzalo Farías* and Sebastián Fingerhuth*
jose.verdugo.a@mail.ucv.cl, gabriel.hermosilla.ucv.cl, gonzalo.farias@ucv.cl, sebastian.fingerhuth@ucv.cl
*Escuela de Ingeniería Eléctrica, Pontificia Universidad Católica de Valparaíso
Avenida Brasil 2147, Valparaíso, Chile

***Abstract*—Thermal Face Recognition over time is a difficult challenge due to faces varies with different factor such as metabolism or ambient conditions. Thus, the aim of this work is to improve recognition rates of thermal faces acquired in time lapse mode, since the results available in other articles are not entirely satisfactory in this modality, which is mainly due to the large variation in the thermal characteristics of the faces in time lapses. To improve the recognition rates the approach called "Sparse Representation" was chosen. This method represents an input image as a linear combination of a dictionary composed of images of different subjects and a vector of sparse coefficients. The results are obtained using the two sets of UCHThermalFace database. The method shows high performance in the time lapse for thermal images.**

Index Terms*—Face Recognition, Sparse Representation, Thermal, Time Lapse.

I. Introduction

Nowadays, face recognition is an increasingly important task, mainly in the area of security. Currently, there is an increasing number of devices that perform face recognition, for example some airports have implemented recognition systems to check passengers or institutions that have laboratories that use face recognition to restrict access to certain areas. The face recognition systems have a great importance; therefore the recognition must be robust against different factors that may affect the recognition rates, and thus, for example, offer more security in access systems. Usually face recognition is performed by visible images where the major factor affecting the performance of recognition systems are the illumination conditions and variations in pose. The natural solution to improve the recognition rate in the visible spectrum is the use of thermal images of faces for recognition due the thermal spectrum in not affected by illumination. Some contributions in this regard are described in [1][2][3]. However, the drawback with the use of thermal faces for recognition is that the thermal characteristics of the face vary considerably with time: acquisition time of the gallery image and the test images [4]. These variations are due to different factors, such as human physiology (inhale, exhale, body temperature regulation), diseases causing anomalous changes in body temperature (fever), environment, because the thermal characteristics are different between a sunny and a rainy day. All these factors are non-controllable, and must be analyzed when the face recognition in the thermal domain is performed [5]. In addition, the thermal camera has a susceptibility on extrinsic factors such as wind, and a time-variable sensor response when the camera is working over long periods of time [2][4][6]. In order to solve this high variation in the thermal characteristics of the faces over time the methodology called Sparse Representation was selected. This methodology was chosen because of the good recognition rates of this method on occluded visible imagery, as can be seen in [7], where recognition rates bordering the 90.3% for occlusion blocks of 40% of the total surface are shown. This methodology also get satisfactory results when it is applied on thermal imagery on modality "same session" [8]. After studying the existing results, the choice of the method is based on considering the face regions with greater variability in the thermal properties as occluded regions, which means we consider the variations produced by the time lapse problem as an occlusion in the thermal face.

II. The Sparse Representation method

The Sparse Representation method [7] is a relatively new approach to pattern recognition, which is based on the scientific concept where certain visual perceptions are the result of a sparse representation of visual patterns using highly redundant visual neurons [9]. In addition, the method has been study in other areas such as human perception and biology. In our case the Sparse Representation of a signal input (image) is generated as a linear combination of elements and a sparse vector, where most of the coefficients are zero, and the non-zero coefficients correspond to images in the dictionary which belong to the same subject represented by the input image, i.e. the problem of face recognition is now transformed into a mathematical optimization problem. Figure 1 shows the schema of Sparse Representation. As proposed in [7], the face recognition algorithm using Sparse Representation allows obtaining the sparse coefficient vector by minimizing the L1 norm of the following equation:

$$x_1 = \arg\min{}_x ||y - Ax||_1 \tag{1}$$

Where y represents the testing image, A is the training images, x is the vector of coefficients and x_1 is the sparsest solution desired. The L1 norm minimization of equation (1) is performed by an optimized algorithm called: Primal Augmented Lagrangian Method (PALM), published in [10].

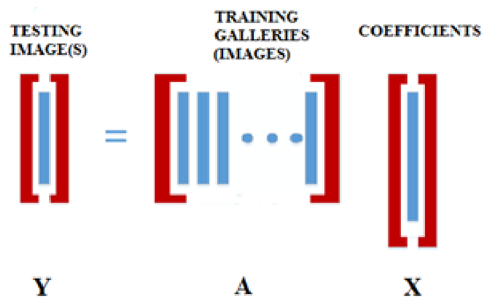

Fig. 1. Schematic of the Sparse Representations applied to images

TABLE I
SPARSE REPRESENTATION ALGORITHM [7]

1	**Input:** n training images, partitioned into k classes, $A_1, \ldots, A_k$, and a test sample $\mathbf{y}$.
2	Normalize the training samples to have unit $l2$ norm and set: $A = [A_1, \ldots, A_k]$
3	Solve the $l1$ minimization problem by linear programming. $x_1 = \arg\min_x \|\|y - Ax\|\|_1$
4	Compute the residuals $r_j(y) = \|\|y - A\delta_j(x_1)\|\|_2$, for $j = 1, \ldots, k$
5	**Output:** $\mathrm{id}(y) = \arg\min_{j=1,\ldots,k} r_j(y)$

The face recognition is performed after obtain the sparse coefficients. To perform the recognition, we define k functions $\delta j, k$ represents the number of subjects, and δ_j is a vector who preserves the coefficients for the $j-$th subject and the remainder becomes zero, i.e.:

$$\delta_j(x) = [0 \ldots 0\, x_j^T\, 0 \ldots 0]^T, \text{ for } j = 1, \ldots, k \tag{2}$$

Once that δ_j are obtained for each subject, it is performed the multiplication between each δ_j and the corresponding training class to obtain the new representation of the test image $\hat{y}_j$ equation 3.

$$\hat{y}_j = A\delta_j(x_1) \tag{3}$$

Finally, we use the L2 norm with the test image and the estimated image to compute the minimum of residuals. The argument of residuals minimum is indicative of the identity returned by the system (see equation 3). The pseudo code is shown in Table 1.

$$\mathrm{id}(y) = \arg\min r_j(y)\,, \text{ for } j = 1, \ldots, k \tag{4}$$

where

$$r_j(y) = ||y - \hat{y}_j||_2\,, \text{ for } j = 1, \ldots, k \tag{5}$$

Fig. 2. Images extracted form the set time-lapse, corresponding to the subject one. Form left to right the acquisition sessions are: 1, 12, 27 and 31.

III. EXPERIMENTS

To evaluate the effectiveness of the Sparse Representation method on thermal imagery, two types of experiments were performed. In the first of these experiments, we analyzed the behavior of the methods with thermal images varying over time. We tried to find a relationship between the number of training images per subject and the recognition rate obtained, to obtain the best tradeoff between the optimal images number required to obtain high recognition rates. The second experiment consists of analyzing the robustness of the method against images with rotation angles of the face.

A. Database description

To perform the experiments, we use two different datasets: UCHThermalTemporalFace and UCHThermalFace.

UCHThermalTemporalFace: the database consists of 350 thermal images of frontal faces cropped and aligned to 150x81 pixels. The images correspond to 7 different subjects taken in different days, obtaining 50 images per subject. Each of the images of a particular subject, was captured on different acquisition session during a time range of 69 days. For the acquisition the camera FLIR TAU 320 was used, with sensitivity in the range of 7.5-13.5μm and a resolution of 324x256 pixels [12], located approximately 1.1m from the test subjects as indicated in [11]. Figure 2 shows an example of a subject in different days of capture.

The use of this particular data set is very significant, since this set is directly related to the problem studied in this research: the use of thermal face recognition with images acquired in time lapse. Thus, the results of tests performed on this set, allow having a preliminary idea of how the algorithm behaves to changes in the thermal characteristics of the faces that occur between each acquisition session.

UCHThermalFace: the database consists of 1166 images of 150x81 pixels resolution. These images correspond to 53 different subjects and 22 images per subject. These 22 images are divided into two groups, where 11 images are acquired in a controlled environment (Indoors) and 11 in an uncontrolled environment (Outdoors). In each of these 11 images (Indoor and Outdoor), the face has a different orientation angle (yaw) and tilt (pitch). Figure 3 shows images with some of the angles of rotation used in this set. For the acquisition the same camera previously mentioned, the FLIR TAU 320, was used. These images were acquired in the same session. We use only the rotation subset to perform the experiments.

Fig. 3. These images corresponding to a subject of the roation set. Images have inclination angle 0°, ans -30°,-15°, 0°, 15°and 30° for yaw angle.

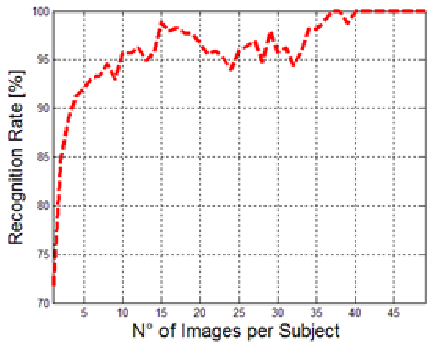

Fig. 4. Recognition rate versus number of training galleries.

B. Time-Lapse Experiment

The experiment consists in obtaining the best recognition rate for different sets over time. To achieve high performance we tried to find best tradeoff between performance and training images, varying the quantity of training sets and test sets until a high recognition rate is achieved. We used the set UCHThermalTemporalFace database previously described.

The evaluation methodology creates subsets of the database corresponding each one to the acquisition days, i.e., we created 50 subsets (S1 to S50, one per acquisition day). The experiment uses different sets to generate the training set (dictionary) for the Sparse Representation, for example for the first case we used the set S1 to training and the remaining sets to test, for the second case, we used S1 and S2 for training and the remaining for testing, and for example for the last case we used from set S1 to S49 to train and set S50 for testing. Thus, we vary the training set (dictionary) for the Sparse Representation algorithm obtaining different recognition rates for each training sets. The images are sequentially selected, in the order in which they were acquired. The results can be seen in the Figure 4.

Figure 4 shows a curve with an increasing rate until the use of 5 images per subject in training. From this point the curve tends to vary mainly in the range 94% to 98% of recognition rate. This is due to the effect of increasing the number of training images.

To correctly estimate the optimal quantity of images, we performed a new test. We randomly choose between 3 to 10 images per subject from the entire set to generate new training subsets. For each training images set, the experiment was repeated 100 times to compute the average of the recognition rates. Results can be seen in the Table 2. From Table 2, we choose the case with 5 training images because there is a good tradeoff between recognition rate and number of images, in addition, the performance is significantly lower when less than 5 images are used. After this amount, the recognition rate is almost constant. In real application, sometimes it is very difficult to obtain more than 5 images per subject.

TABLE II
RECOGNITION RATES USING RANDOM TRAINING SET OF UCHTHERMALTEMPORALFACE

	Number of images per subject							
	3	4	5	6	7	8	9	10
Average Recogition Rate	86,29	91,46	**94,06**	94,64	95,28	95,95	97,39	96,61

C. Rotation Experiment

In this experiment, we used the Rotation set of UCHThermalFace database [11] in indoor and outdoor conditions, previously described. The experiment analyzes the behavior of the algorithm on images having different rotation angles for unconstrained environment using different numbers of images to generate the training sets for the methodology.

The experiment takes the 11 subsets of rotations, from R1 to R11 (see details in [11]), where each rotation set represents a different pitch an yaw angle, for example the R6 set has a rotation of 0 (frontal image) and R1 has -15 for both pitch and yaw angles. We use different training sets to perform the face recognition, combining 1, 3 and 5 images per subject to generate the dictionary for the Sparse Representation method. Later, we computed the average recognition rate of each recognition rate obtained from the training sets and test sets. The results are shown in the Table 3 and Table 4, for indoor and outdoor case respectively.

The results obtained to apply the Sparse Representation method for indoor case is presented in Table 3. The result shows an improvement in the average recognition rate when the number of images for training increases, which means there is more information of the face rotation angles. However, when only one image per subject is used in this experiment, the Sparse Representation has very low result in comparison with the results obtained in other publications (see [11] for more details).

For the outdoor case, Table 4 shows a different result in comparison with the indoor case, with 3 images per subject the performance of the method is better. This is due to the saturated images caused by direct exposure to sun radiation or a bad calibration of the thermal camera in outdoor case. For this reason the methodology with 3 images create a better dictionary than with 5 images.

In addition, from Tables 3 and 4, we can see that using rotation angles between -15°, 0°, and 15° yaw angles (R5, R6, R7), the dictionary obtained allows to perform good face recognition rates for indoor and outdoor cases.

TABLE III
RECOGNITION RATES USING SET ROTATION-INDOOR SET OF UCHTHERMALFACE

Number of Images per subject	Training Galleries	Average Recogn. rate %
1	R6	57,65
2	R5-R6-R7	80,42
3	R4-R6-R8	79,72
4	R4-R5-R6-R7-R8	**87,74**

TABLE IV
RECOGNITION RATES USING SET ROTATION-OUTDOOR SET OF UCHTHERMALFACE

Number of Images per subject	Training Galleries	Average Recogn. rate %
1	R6	50,00
2	R5-R6-R7	**73,58**
3	R4-R6-R8	71,93
4	R4-R5-R6-R7-R8	70,75

IV. CONCLUSION AND FUTURE WORK

In this article, we studied the thermal face recognition over time using the sparse recognition method. The method was analyzed with the UCHThermalFace database in two modalities: variant in time and rotation subsets (same session).

In general terms there are some comments about the results of time lapse experiment. When we analyzed frontal images (unchanged at the angles of orientation and inclination), the Sparse Representation algorithm obtains high recognition rates. We also observed when increasing the number of training images recognition rates tend to increase. The best tradeoff between number of training images and recognition rate was obtained with 5 images per subject, achieving 94% of recognition rate in average.

For rotation experiment, it can be seen that there is a considerable decrease in recognition rates compared to the recognition rates obtained in time lapse. The rotation of the face affects the sparse methods because it is necessary to use a higher number of images to generate an accurate dictionary which represents the test image. One solution to this problem would be to train a sequence of images covering all possible rotations of the face. In addition, comparing indoor and outdoor results, a decrease can be observed in the performance due to the saturation in the thermal images. By comparing the obtained results with [11], the performance is significantly less for Sparse Representation method.

For future work, we will create a new public database of thermal images over time, because there are not databases available at the moment for the research community which allows study the time lapse problem. In addition, we want to study the current thermal methods for face recognition [13][14] that includes different variations in time of the thermal capture acquires in unconstrained environment (indoor-outdoor session) and analyze the internal configuration of the thermal cameras to maintain the thermal image without variations in time.

ACKNOWLEDGMENT

This research was funded by the FONDECYT-Chile Grant 11130466.

REFERENCES

[1] Socolinsky, D. A., Wolff, L. B., Neuheisel, J. D., & Eveland, C. K.(2001). *Illumination invariant face recognition using thermal infrared imagery.* InComputer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on (Vol. 1, pp. I-527). IEEE.

[2] Chen, X., Flynn, P. J., & Bowyer, K. W. (2003, October).*PCA-based face recognition in infrared imagery: Baseline and comparative studies.* In Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on (pp. 127-134). IEEE.

[3] Selinger, A., & Socolinsky, D. A. (2004, June). *Face recognition in the dark.* InComputer Vision and pattern Recognition Workshop, 2004. CVPRW'04. Conference on (pp. 129-129). IEEE.

[4] Socolinsky, D. A., & Selinger, A. (2004, August). *Thermal face recognition over time.* In Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on (Vol. 4, pp. 187-190). IEEE.

[5] Socolinsky, D. A., & Selinger, A. (2004, June). *Thermal face recognition in an operational scenario.* In Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on (Vol. 2, pp. II-1012). IEEE.

[6] P. Buddharaju, I. Pavlidis, I. Kakadiaris. *Physiology-Based Face Recognition.* In Proc. of the IEEE Conf. on Advanced Video and Signal Based Surveillance, Lake Como, Italy, 2005, pp. 354-359.

[7] Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., & Ma, Y. (2009). *Robust face recognition via Sparse Representation.* . Pattern Analysis and Machine Intelligence, IEEE Transactions on, 31(2), 210-227.

[8] Buyssens, P., & Revenu, M. (2010, September). *Ir and visible face identification via Sparse Representation.* In Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on (pp. 1-6). IEEE.

[9] Olshausen, B. A., & Field, D. J. (1997). *Sparse coding with an overcomplete basis set: A strategy employed by V1?,* Vision research, 37(23), 3311-3325.

[10] Yang, A. Y., Zhou, Z., Balasubramanian, A. G., Sastry, S. S., & Ma, Y. (2013). *Fast-Minimization Algorithms for Robust Face Recognition.* Image Processing, IEEE Transactions on, 22(8), 3234-3246.

[11] Hermosilla, G., Ruiz-del-Solar, J., Verschae, R., & Correa, M. (2012). *A comparative study of thermal face recognition methods in unconstrained environments.* Pattern Recognition, 45(7), 2445-2459. Available at: http://www.captura.uchile.cl/bitstream/handle/2250/17109/Hermosilla_Gabriel.pdf?sequence=1

[12] IR Camera FLIR TAU 320, [Online]. Available at: www.flir.com/cvs/cores/uncooled/products/tau

[13] B. Zhang, Y. Gao, S. Zhao and J. Liu. *Local Derivative Pattern Versus Local Binary Pattern: Face Recognition With High-Order Local Pattern Descriptor.* vol. 19, no. 2, pp. 533-544, 2010.

[14] N. Dalal and B. Triggs. *Histograms of Oriented Gradients for Human Detection.* 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1, pp. 886-893.