

# Face Recognition Using TOF, LBP and SVM in Thermal Infrared Images<sup>\*</sup>

Ramiro Donoso Floody<sup>1,2</sup>, César San Martín<sup>1,2</sup>, and Heydi Méndez-Vázquez<sup>3</sup>

<sup>1</sup> Center for Optics and Photonics, University of La Frontera, Chile

<sup>2</sup> Information Processing Laboratory, DIE, University of La Frontera, Chile

<sup>3</sup> Advanced Technologies Application Center, CENATAV, Cuba

**Abstract.** In this work, Binary Local Patterns (LBP), Support Vector Machine (SVM) and Trade-off (TOF) correlation filter are evaluated in face recognition tasks using thermal infrared imagery. The infrared technology has a particular kind of noise called non-uniformity and correspond to a fixed pattern noise superimposed at the input image, degrading the quality of the scene. Non-uniformity varies over time very slowly, and in many applications, depending of the technology used, can be assumed constant for at least several hours. Additionally, additive Gaussian noise (variable over time) is generated by the associated electronics. Both kind of noise affect the performance of classifiers in face recognition applications using infrared technology and must be considered. The comparison of performance of each method considering fixed and variable over time noise leads allow to conclude that SVM is more robust under both kind of noise.

**Keywords:** Face Recognition, Infrared Thermal Imaging, SVM, LBP and TOF.

## 1 Introduction

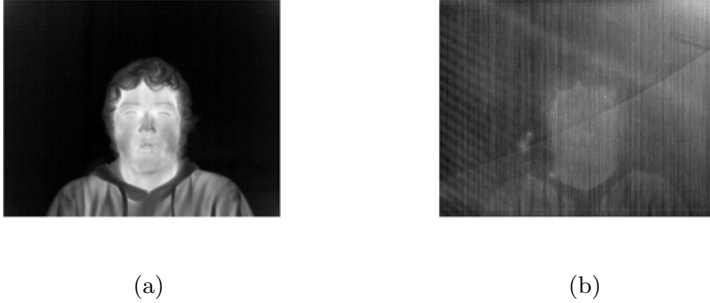
Actually, there are many works on face recognition [7], which mainly uses the visible spectral range, although there are some works like Ghiass et al. [3], that present a state of art in terms of facial recognition in the infrared spectral range. Additionally, the work of Kong et al [4] presents a comparison between the advantages and disadvantages of the techniques used in both spectral ranges. In face recognition problem, the efficiency of classification process in visible range depends principally of the face angle view, occluding objects, distance between face and camera, facial expressions, and mostly the light or existing lighting in the environment. In this work, an evaluation between TOF classification methods [9], LBP [6,1] and SVM [11] using infrared imagery is presented. The infrared thermal images correspond to the range  $8 - 14\mu\text{m}$ , i.e., infrared emission being independent of any light source. In particular, human skin has an emissivity close to 1 (see [8]), which represents a unique thermal signature for each subject.

Infrared technology present a non-uniformity in the output scene when a flat-image is captured. This effect is know as fixed pattern noise, i.e., remains

---

<sup>\*</sup> This work was partial supported by *Center for Optics and Photonics* FB0824/2008.

constant in time. For example, Figure 1a shows a capture without noise and in Figure 1b a capture with fixed pattern noise using CEDIP JADE UC camera is presented. In effect, the original pattern capture is represented in Figure 1b while the Figure 1a is obtained by correcting the non-uniformity at the scene using a black body at two different temperatures (two-point calibration). This correction requires stopping the capture process and the use of black bodies as a reference. The aim of this work is to recognize individuals using images as Figure 1b without the need to correct or eliminate the fixed pattern noise.



**Fig. 1.** Infrared images captured using the CEDIP JADE camera, a) free noise image and b) image with fixed pattern noise. The goal is to recognize the subject in an image as b).

This work is organized as follows. In section 2 the algorithms TOF and LBP are reviewed, and what is a support vector machine (SVM) is briefly explained. In section 3 the used comparison methods for the algorithms mentioned above is presented. In section 4 the experiment and results is presented. Finally, some discussions and conclusions are reported in section 5 and 6, respectively.

## 2 Methodology

In this section, three traditional methods of face recognition are presented: Trade-Off (TOF) correlation filter, Local Binary Patterns (LBP) based, and Support Vector Machine (SVM).

### 2.1 Correlation Filter

Correlation is a measure commonly used to characterize the similarities between a reference pattern and a test pattern. This concept is used frequently in recognition applications, presenting a greater importance degree on the use of cross-correlation for get the relative position of the object. The cross-correlation is given by:

$$c(\tau_x, \tau_y) = \int \int T(f_x, f_y) R^*(f_x, f_y) \exp^{j2\pi(f_x \tau_x + f_y \tau_y)} df_x df_y, \quad (1)$$

$$= IFT \{T(f_x, f_y) R^*(f_x, f_y)\}, \quad (2)$$

Where  $R(f_x, f_y)$  and  $T(f_x, f_y)$  are the 2D Fourier transforms of the reference pattern and the test pattern, respectively. In general, the use of a correlation filter for face recognition, as shown in Figure 2. Consists in to apply a filter (in particular, in this work we used the filter TOF 2.1) on subject image under study, and observe which is the highest amplitude (high peak (equation 1)) of the resulting image cross-correlation, and then calculate a similarity measure (PSR or PCE [5]) to classify a subject as valid or not.

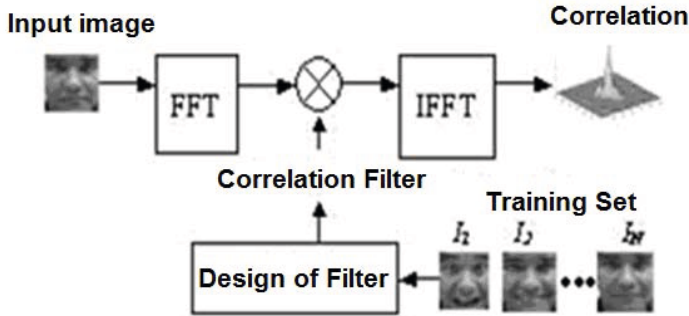


Fig. 2. Correlation filters in face classification

**Optimal Tradeoff Filter (TOF).** TOF algorithm (3) [9] provides a compromise between the MACE filter characteristics [2], which increases the peak of positive results, and the NTC filter, which aims to reduce the output noise variance. In this case it is necessary to assume the presence of white noise in order to approximate the matrix C to identity matrix, and the commitment between the two is defined by  $\alpha$  in the form:

$$\text{TOF filter: } h = T^{-1}X(X * T^{-1}X)^{-1}u, \tag{3}$$

$$\text{Trade-off: } T = \alpha MACE + (1 - \alpha)NTC. \tag{4}$$

### 2.2 Local Binary Patterns-(LBP)

The use of LBP in face recognition was introduced by Ahonen et al. in their work [1] and different variations have appeared after Marcel et al. (2007) [6]. The work of Socolinsky et al. [10] shows good results using methods based on appearance. By the above, we were decided to explore the LBP algorithm, which is a texture descriptor based on appearance, which is rather robust and less influenced by possible alignment problems of infrared images.

As shown in Figure 3, the original LBP, each neighborhood (of 3x3) is thresholding according to the central pixel ( $g_c$ ) thereof, and the result is considered as a binary number called LBP code. Then, the image is divided into rectangular regions, and for each region is calculated the histogram of the LBP code. Finally, the histogram of each region is concatenated into one that represents the image of the face. The Chi-square similarity measure [1] is used to compare the

histograms of two different images, and thereby discriminate whether a subject is considered valid or impostor. The method of nearest neighbor was used to make the classification (identification). The TEER (Threshold of Equal Error Rate) was found to verification problem, and the value (obtained by means of training set, that corresponding to the comparison with the measure Chi-square similarity) was used as a threshold in the test set.

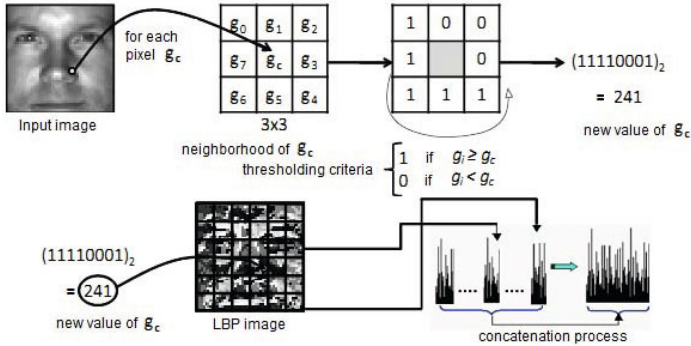


Fig. 3. Original Local-Binary Patterns [8]

### 2.3 Support Vector Machine (SVM)

Support Vector Machines [11] allow to solve classification problems. In its original form were made to solve classification problems with only two classes, although there are different methods for use in multiclass problems. In this work we chose the original version, due to it adapts to the verification problem, where the first class represents a subject, and the second class represents the other subjects (impostors). A support vector machine assumes that if we have a set of data not linearly separable of  $n$  dimension (which belong to two classes), then they are linearly separable in the  $(n + 1)$  dimension. The SVM can use different types of kernels, which greatly promotes the search for solutions to problems of classification. These kernels can be: linear, polynomial, RBF, sigmoid, etc. In particular, this paper each image was represented as a column vector where each element contains the intensity of the pixel, which can be classified as a method of appearance.

## 3 Evaluation Methods

To measure the fixed and variable noise tolerance, the following evaluation methods are used (see [7]):

**False Accepted Rate (FAR).** The false accepted rate is calculated by the relation between the number of accepted impostor subjects and the total number of impostor subjects, i.e.:

$$FAR = \left( \frac{\text{impostors accepted as valid subjects}}{\text{total number of impostor subjects}} \right) \quad (5)$$

**False Rejected Rate (FRR).** The false rejected rate is calculated by the relation between the number of rejected valid subjects and the total number of valid subjects, i.e.:

$$FRR = \left( \frac{\text{valid subjects rejected}}{\text{total number of valid subjects}} \right) \quad (6)$$

**Equal Error Rate (EER).** As shown in Figure 4, the *EER* value is obtained when  $FRR = FAR$ . Note that while more smaller is this value, better is the system to classify.

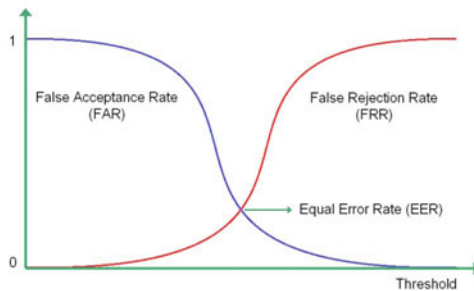


Fig. 4. EER as the intersection of FAR and FRR curves [7]

Note that these three indicators (*EER*, *FAR* and *FRR*) are proportional inversely to the performance of facial recognition methods, i.e., while more closer to zero are these indicators, the methods obtain a better classification.

## 4 Experimentation and Results

### 4.1 Experiment Description

To evaluate the noise tolerance of the three methods mentioned in section 2, we performed the following experiment. A training set and test set from a database was created, which contains two separate sets of faces, the expression set E (joy, anger, surprise, etc.) and vocalization set V (vowels or words). The database has 102 subjects, 6 images per subject (subdivided into two sets: 3 images for the set E and 3 for the set V). The training set was of 34 subjects, i.e., 68 subjects will be impostors for the system. For each subject belonging to training set, a training subset was defined, which consists of 3 images of a set (E or V) as hits and 1 image (the same set) of each one of the other subjects in the training set

(as rejections). As test subset to this subject was considered the 3 images in the group (E or V) that weren't considered in the training set (as hits) and all images of impostor subjects (as rejections).

To train the system, the complete and original images were used, i.e., free noise and the original size (320x240px). And then, case-by-case the noise tolerance (fixed and variable pattern) was evaluated, adding to the original image a Gaussian noise simulated by a normal distribution with mean 0 and variance  $t\%$  of  $2^8$ , where  $t = \{10, 20, 30, 40, 50, 80\}$  for fixed noise, and  $t = \{1, 5, 10\}$  for temporal noise. Recall that in each case were used two training subsets (one E and one with V) and two test subsets for each subject belonging to the training set. Moreover, 5 exercises for each combination to obtain better statistics were performed.

### 4.2 Results

Below, the Table 1 is presented, which contains the performance of the methods with images in their original condition, i.e., free noise. Following, the tables show the results of the experiment described above. Remember that the images are classified into two sets (E and V), so the results are given separately for each one. In Tables 2, 3 and 4 are presented the results of classification with fixed pattern noise. Moreover, in Tables 5, 6 and 7 is shown the results considering the variable white noise.

**Table 1.** Performance, EER, FAR y FRR for original images

Method	Accuracy	Set E			Set V			
		EER	FAR	FRR	Accuracy	EER	FAR	FRR
<b>TOF</b>	94.77	5.88	6.39	8.33	98.69	3.92	3.70	10.78
<b>LBP</b>	97.71	1.96	2.15	3.27	99.35	1.96	2.15	3.27
<b>SVM</b>	97.69	2.28	2.29	5.88	97.19	2.70	2.75	9.80

**Table 2.** Performance of TOF method with fixed pattern noise

exercise	var	Set E			Set V		
		10%	20%	30%	var 10%	var 20%	var 30%
<b>1</b>		74.5098	64.7059	60.4575	85.9477	82.0261	78.1046
<b>2</b>		65.0327	56.8627	52.2876	84.9673	81.6993	76.7974
<b>3</b>		59.8039	49.3464	49.0196	83.3333	82.0261	76.7974
<b>4</b>		56.8627	47.3856	47.7124	82.0261	80.3922	76.1438
<b>5</b>		52.2876	45.4248	45.7516	84.3137	79.7386	74.5098
<b>Mean</b>		61.6993	52.7451	51.0457	84.1176	81.1765	76.4706

**Table 3.** Performance of LBP method with fixed pattern noise

exercise	Set E			Set V		
	var 10%	var 20%	var 30%	var 10%	var 20%	var 30%
1	96.7320	96.0784	96.0784	97.0588	97.0588	98.0392
2	97.0588	96.7320	96.0784	97.7124	97.0588	97.3856
3	96.7320	96.7320	96.0784	97.7124	97.3856	97.0588
4	96.7320	96.0784	96.0784	97.3856	97.3856	97.0588
5	97.3856	96.4052	95.7516	96.7320	97.0588	98.0392
Mean	96.9281	96.4052	96.0130	97.3202	97.1895	97.5163

**Table 4.** Performance of SVM method with fixed pattern noise

exercise	Set E			Set V		
	var 10%	var 20%	var 30%	var 10%	var 20%	var 30%
1	98.1466	98.0034	97.3952	98.1466	98.3970	98.3469
2	98.1895	97.9748	97.7601	98.2038	98.4901	98.4901
3	98.1394	97.9963	97.8174	98.1823	98.5258	98.5616
4	98.1752	98.0034	97.9176	98.1609	98.4471	98.4686
5	98.2038	98.0893	97.7315	98.1537	98.4686	98.5187
Mean	98.1709	98.0135	97.7243	98.1695	98.4657	98.4772

**Table 5.** Performance of TOF method with variable white noise

exercise	Set E			Set V		
	var 1%	var 5%	var 10%	var 1%	var 5%	var 10%
1	94.7712	93.4641	88.2353	97.3856	94.4444	91.5033
2	94.4444	91.1765	86.9281	97.3856	95.4248	89.2157
3	94.1176	92.1569	88.2353	97.0588	96.0784	88.8889
4	94.7712	93.4641	88.2353	97.0588	95.7516	88.5621
5	94.1176	93.7908	89.2157	97.3856	94.7712	89.8693
Mean	94.4444	92.8105	88.1699	97.2549	95.2941	89.6079

**Table 6.** Performance of LBP method with variable white noise

exercise	Set E			Set V		
	var 1%	var 5%	var 10%	var 1%	var 5%	var 10%
1	96.7320	96.7320	95.0980	98.0392	97.7124	97.0588
2	97.0588	96.4052	95.4248	98.6928	97.3856	96.7320
3	96.7320	97.0588	95.0980	97.7124	97.0588	97.0588
4	96.7320	96.7320	95.4248	98.3660	96.4052	96.7320
5	97.0588	97.0588	95.0980	99.0196	96.7320	97.0588
Mean	96.8627	96.7974	95.2287	98.3660	97.0588	96.9281

**Table 7.** Performance of SVM method with variable white noise

exercise	Set E			Set V		
	var 1%	var 5%	var 10%	var 1%	var 5%	var 10%
<b>1</b>	97.8317	98.0607	98.1466	97.5097	97.9891	98.2038
<b>2</b>	97.8102	98.0750	98.1967	97.4524	98.0106	98.1967
<b>3</b>	97.8460	98.0321	98.1895	97.4882	97.9963	98.2181
<b>4</b>	97.8818	98.0249	98.2110	97.5025	98.0106	98.1609
<b>5</b>	97.8603	97.9963	98.1466	97.4667	97.9677	98.1251
<b>Mean</b>	97.8460	98.0378	98.1780	97.4839	97.9948	98.1809

## 5 Discussion

As shown in Table 1 the LBP method succeeds in obtaining the best performance for face recognition with free noise images in E set, obtaining a performance of 97.71%, closely followed by SVM with 97.69%. As is expected, the LBP method also gets the lowest values for the indexes EER (1.96%), FAR (2.15%) and FRR (3.27%). When is considered the V set of images, again the Table 1 shows that the LBP method gets the best performance with 99.35% followed by TOF correlation filter with 98.69%. Note that despite the TOF filter gets a better performance than SVM, the SVM gets the lowest value to EER, FAR and FRR.

As shown in Tables 2, 3 and 4, the performance of classifiers in general is better for set V than for set E, with differences exceeding 20% in accuracy comparing to the same noise (in the case of TOFF correlation filter) and minor variations in accuracy 0.0014% in case of SVM. Therefore, and according to the averages shown in Table 4, the SVM is more stable when classifying both sets of images, i.e. the accuracy obtained for the set E is very similar to that obtained for the set V.

As shown in Table 2, the TOF correlation filter obtains the worst results for the two sets (E and V) obtaining a minimum accuracy of 51.05% when considering a 30% of fixed noise. Moreover, the best accuracy to the same noise is achieved by the SVM (97.72%).

As shown in Tables 5, 6 and 7 the TOF correlation filter gets the worst performance, and the SVM gets the best performance again. Note that, the results obtained by SVM are usually slightly higher by about 1% to those obtained by LBP, except in the training set V, with variable noise of 1%, where LBP is higher than SVM at 0.8%.

For all the above and Table 1, we recommend the use of LBP method in faces recognition with free or very little noise; on the other hand we recommended the use of SVM for noisy images, due to SVM is robust to fixed and variable noise. Also, SVM gets low values for the indexes EER, FAR and FRR.

Finally, we note that in some cases to increase the percentage of noise, is produced a slight increase in accuracy, which could be considered as a possible inconsistency in the results, it's logical to think that while more noise the



analyzed image has, worse will be the performance of classification. But this behavior can be explained if is considered that with a slight noise added randomly to the image, will be incorporated a differentiation element between them, which is the responsible for the slight increase in accuracy obtained.

## 6 Conclusions

Due to the main factor of face recognition problem is the lighting, in this work was used infrared thermal imaging to make a comparison between the TOF method, LBP and SVM applied to this problem. In the previous section recommends the use of one method over another in certain cases. They also concluded that SVM is usually the most robust against noise from fixed and variable pattern over time.

Further work is needed to incorporate other methods in the study of face recognition in infrared images, it is also possible to extend the experiment and increasing the percentage of noise to see that finally the algorithms performance drops when this increase, and thus give more consistency to the data presented.

## References

1. Ahonen, T., Hadid, A., Pietikäinen, M.: Face Recognition With Local Binary Patterns. In: Pajdla, T., Matas, J(G.) (eds.) ECCV 2004. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004)
2. Casasent, D., Ravichandran, G.: Advanced distortion-invariant minimum average correlation energy (MACE) filters. *Applied Optics* 31, 1109–1116 (1992)
3. Ghiass, R.S., Bendada, A., Maldague, X.: Infrared Face Recognition: A Review of the State of the Art. In: 10th International Conference on Quantitative InfraRed Thermography (July 2010)
4. Kong, S.G., Heo, J., Abidi, B.R., Paik, J., Abidi, M.A.: Recent advances in visual and infrared face recognition - a review. *Computer Vision and Image Understanding* 97, 103–135 (2005)
5. Vijaya Kumar, B.V.K., Hassebrook, L.: Performance measures for correlation filters. *Applied Optics* 29(20), 2997–3006 (1990)
6. Marcel, S., Rodriguez, Y., Heusch, G.: On the Recent Use of Local Binary Patterns for Face Authentication. *International Journal on Image and Video Processing Special Issue on Facial Image Processing* (2007)
7. San Martin, C., Carrillo, R., Meza, P., Mendez, H., Plasencia, Y., García-Reyes, E., Hermosilla, G.: Recent Advances on Face Recognition using Thermal Infrared Images. InTech (2011)
8. Méndez, H., Martín, C.S., Kittler, J., Plasencia, Y., García-Reyes, E.: Face Recognition with LWIR Imagery Using Local Binary Patterns. In: Tistarelli, M., Nixon, M.S. (eds.) ICB 2009. LNCS, vol. 5558, pp. 327–336. Springer, Heidelberg (2009)
9. Refregier, R.: Optimal trade-off filter for noise robustness, sharpness of the correlation peaks, and horner efficiency. *Optics Letters* 32, 1933–1935 (1993)
10. Socolinsky, D., Selinger, A., Neuheisel, J.: Face recognition with visible and thermal infrared imagery. *Comput. Vis. Image Und.* 91, 72–114 (2003)
11. Theodoridis, S., Koutroumbas, K.: *Pattern Recognition* (2006)