

A new fusion framework for multispectral face recognition in the texture space

by M. Akhloufi* and A. Bendada*

* *Computer Vision and Systems Laboratory, Laval University, Quebec City (Quebec) G1V 0A6, Canada, akhloufi@gel.ulaval.ca, bendada@gel.ulaval.ca*

Abstract

In this work, we introduce a new fusion framework for multispectral face recognition in the textures space. Active and passive infrared imaging modalities are used and comparison with visible face recognition is performed.

The proposed texture space is based on the use of LBP (Local Binary Pattern) and LTP (Local Ternary Pattern) techniques. Also, a new adaptive texture descriptor is presented. This last descriptor called LATP (Local Adaptive Ternary Pattern) is less sensitive to noise in near uniform regions and permit to overcome some of the limitations of LBP and LTP descriptors. The obtained multispectral data are fused using a multi-scale fusion scheme in order to get interesting face characteristics. The obtained results are promising and show an increase in recognition performance when texture channels are fused in a multi-scale fusion scheme.

1. Introduction

Face recognition is an area of computer vision that has attracted a lot of interest from the research community. A growing demand for robust face recognition software in security applications has driven the development of interesting approaches in this field.

A large quantity of research in face recognition deals with visible face images [1,2]. In the visible spectrum the illumination changes represent a significant challenge for the recognition system. Illumination change introduces a lot of errors during the recognition phase. Another challenge for face recognition in the visible spectrum involves the changes in facial expressions. Facial expression can lead to a poor performance of the face recognition system in visible images. To avoid these problems, researchers propose the use of 3D face recognition [3] and infrared face recognition [4,5]. Infrared face recognition is a growing area of research. Many of the techniques used in infrared face recognition are inspired from their visible counterparts. Known techniques used in visible face image recognition are also used with infrared images, like Eigenfaces or Fisherfaces [4,5]. More recently in [6,7,8] physiological information extracted from high temperature regions in thermal face images were used in infrared face recognition.

The vast majority of infrared face recognition techniques are based on linear approaches used in visible face recognition. Recent work has been conducted using non linear dimensionality reduction [9] and Bayesian techniques for infrared face recognition with promising results [10].

Texture analysis have been widely studied in the literature. Many methods have been proposed in order to handle machine vision problems where texture features serve as a cue for classification, segmentation and recognition. Various statistical descriptors have been proposed for the measure of image textures [11,12]. Since its introduction, Local Binary Pattern (LBP) texture descriptor has been successfully used in texture classification [13]. Many work was done using LBP for visible spectrum face recognition [14-18].

In this work, we introduce a new framework for multispectral face recognition in the textures space. Active and passive infrared imaging modalities are used and comparison with visible face recognition is performed. Two multispectral face recognition databases were used in our experiments: Equinox Database (Visible, SWIR, MWIR, LWIR) and Laval University Multispectral Database (Visible, NIR, MWIR, LWIR).

The proposed texture space is based on the use of LBP (Local Binary Pattern) and LTP (Local Ternary Pattern) [18] techniques. Also, a new adaptive texture descriptor is presented. This descriptor uses statistical data extracted from nearest neighbours in order to define the texture spectral kernel. This descriptor called LATP (Local Adaptive Ternary Pattern) [25] is less sensitive to noise in near uniform regions and permit to overcome some of the limitations of LBP and LTP descriptors.

The obtained multispectral data are then fused using a multi-scale fusion scheme in order to get interesting face characteristics. The following multi-scale fusion techniques were used: Laplacian pyramid, FSD pyramid, Ratio pyramid, Contrast pyramid, Gradient pyramid, Morphological pyramid, Discrete wavelets transform (DWT) and Haar wavelets (SIDWT). The multi-scale fusion is performed in different texture channels and between different modes.

Tests were conducted using two multispectral infrared face databases: Equinox multimodal face database [19] and a new infrared multispectral face database we have developed recently in order to evaluate infrared face recognition techniques in a close to real world situations [9].

2. Texture descriptors

The texture descriptors used in this work are LBP-Like textures. The following sections give more details on how to compute these descriptors.

2.1. Local Binary Pattern (LBP)

LBP was introduced by Ojala *et al.* [13]. It is a gray-scale invariant texture measure computed from the analysis of a 3x3 local neighborhood over a central pixel. The LBP is based on a binary code describing the local texture pattern. This code is built by thresholding a local neighborhood by the gray value of its center.

The eight neighbors are labeled using a binary code {0, 1} obtained by comparing their values to the central pixel value. If the tested gray value is below the gray value of the central pixel, then it is labeled 0, otherwise it is assigned the value 1:

$$P_i' = \begin{cases} 0 & \text{if } P_i < P_0 \\ 1 & \text{otherwise} \end{cases} \quad (1)$$

P_i' is the obtained binary code, P_i is the original pixel value at position i and P_0 is the central pixel value.

With this technique there is 256 (2^8) possible patterns (or texture units). The obtained value is then multiplied by weights given to the corresponding pixels. The weight is given by the value 2^{i-1} . Summing the obtained values gives the measure of the LBP (Fig. 1):

$$I_{LBP} = \sum_{i=1}^8 P_i' 2^{i-1} \quad (2)$$

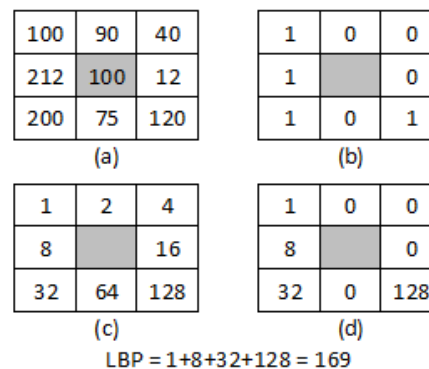


Fig. 1. The LBP texture descriptor computation steps

The central pixel is replaced by the obtained LBP value (169 in the example of figure 1). A new LBP image is constructed by processing each pixel and its 3x3 neighbors in the original image. LBP can be easily extended to include a larger neighborhood area, leading to a different result.

2.2. Local Ternary Pattern (LTP)

LBP has been successfully used by the research community in texture classification. However LBP is sensitive to noise in near uniform regions. In order to solve this problem a new texture descriptor was introduced recently in [18]. This Descriptor is called local ternary pattern (LTP). Like the LBP, it gives an invariant texture measure computed from the analysis of a local neighborhood over a central pixel.

LTP extends LBP to 3-valued codes, in which local pixels having their gray levels within an interval between a user defined threshold $-t$ and $+t$ when compared with the central pixel are assigned a value 0. The pixels above $+t$ when compared to the central pixel are assigned a value 1 and the ones below $-t$ when compared to the central pixel are assigned a value -1. Equation (3) shows how to compute the LTP:

$$P'_i = \begin{cases} 1 & \text{if } P_i \geq P_0 + t \\ 0 & \text{if } |P_i - P_0| < t \\ -1 & \text{if } P_i \leq P_0 - t \end{cases} \quad (3)$$

t is a user defined threshold. This makes the resulting descriptor less sensitive to noise but no longer strictly invariant to gray level transformations.

Figure 2 shows the result of applying LTP to an image using a threshold $t = 5$. In order to get rid of the negative part of the LTP, the LTP above is split in two LBP channels. The upper pattern (LTPU) is obtained by replacing the negative values with 0. The lower pattern (LTPL) is obtained by first replacing the 1 with 0 and then changing the negative values to 1. Figure 2 shows the resulting upper and lower patterns (LTPU and LTPL).

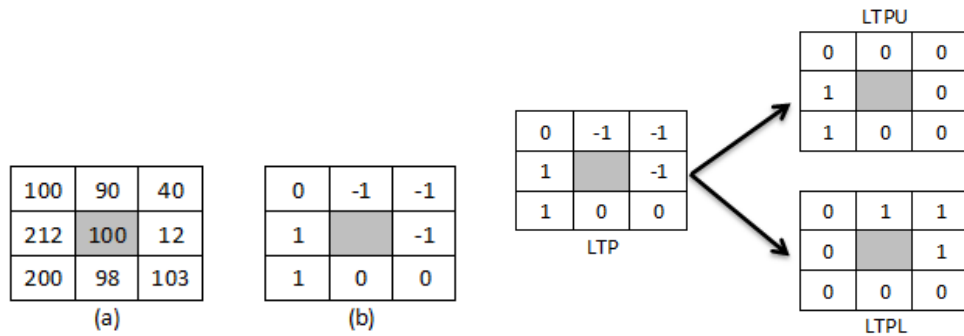


Fig. 2. The LTP texture descriptor

The resulting pixel is computed as for LBP by multiplying the resulting kernels with the local weights. Figure 3 shows how to compute the LTPU and LTPL values. This processing gives two texture images enhancing different characteristics of the image texture.

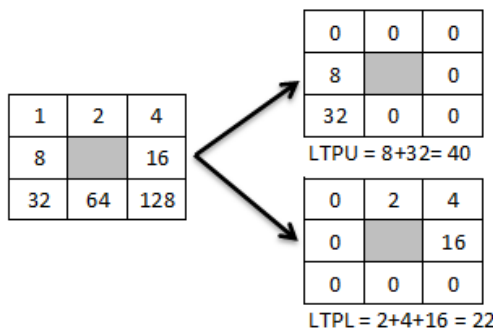


Fig. 3. The LTP Upper and Lower image computation

2.3. Local Adaptive Ternary Pattern (LATP)

LTP is an interesting descriptor since it permits to overcome some of the limitations of the LBP. However, in order to obtain good results the choice of the threshold is very important. In face recognition application, the best performing threshold for LTP can vary depending on the face database and the tests scenarios. The ideal solution is to find ways to compute this threshold automatically from the available image data. Previous work in the area of image segmentation [20], [21] used local statistics in order to compute an adaptive threshold. Local adaptive thresholding provides more robustness to changes in illumination. This later characteristic is very useful in face recognition. In this work, we introduce an adaptation of similar approaches in order to compute the threshold necessary for LTP processing. The proposed approach called Local Adaptive Ternary Pattern (LATP) is given by equation (4):

$$P_i = \begin{cases} 1 & \text{if } P_i \geq (\mu + k\sigma) \\ 0 & \text{if } |P_i| < (\mu + k\sigma) \\ -1 & \text{if } P_i \leq (\mu - k\sigma) \end{cases} \quad (4)$$

Where μ is the mean of the local region, σ is the standard deviation and k is a constant.

The proposed approach is adaptive and use simple statistics to compute the threshold. The resulting image enhances the most important face features and reduce the effect of illumination changes, thus making this technique suitable for face recognition. Figure 4 below show the result of applying the proposed technique in a local image area. In order to get rid of the negative part of the LATP, we use a similar approach to LTP (described above). The LATP is split in two LBP channels. The upper pattern (LATPU) is obtained by replacing the negative values with 0. The lower pattern (LATPL) is obtained by first replacing the 1 with 0 and then changing the negative values to 1. Figure 4 shows the resulting upper and lower patterns (LATPU and LATPL).

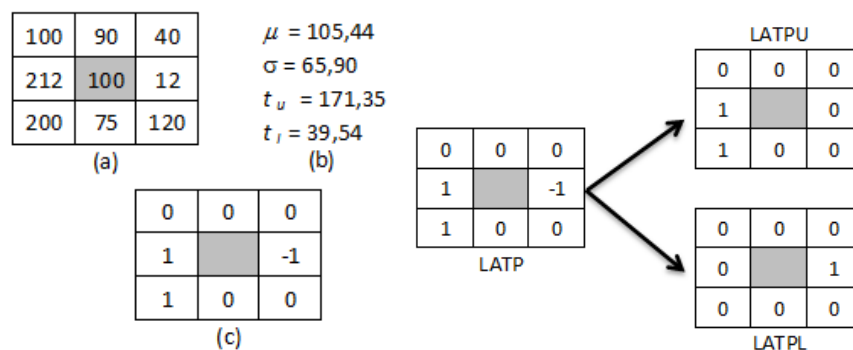


Fig. 4. The LATP texture descriptor

The resulting pixel is computed by multiplying the resulting kernels with the local weights. Figure 5 show how to compute the LATPU and LATPL values. This processing gives two texture images with enhanced face characteristics suitable for face recognition..

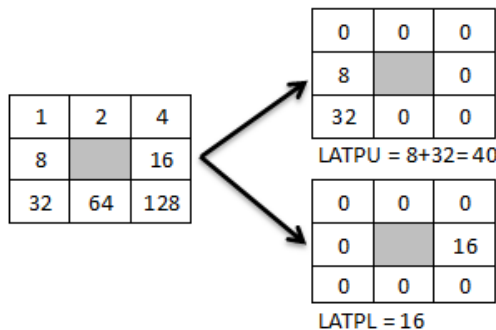


Fig. 5. The LATP Upper and Lower image computation

3. Face recognition

Face learning and recognition techniques based on linear and non linear dimensionality reduction were implemented. Dimensionality reduction techniques are a set of mathematical techniques used for representing available data in a low dimensional space. The obtained representation is a compressed version of the original data with the most important characteristics preserved for further processing. Dimensionality reduction seek to represent a set of data as a p -dimensional manifold embedded in an m -dimensional space (with $p < m$) [13]. Non linear dimensionality reduction for subspace learning was implemented in order to obtain a face space that best detects the essential manifold structure for face recognition. In the

infrared spectrum, variations can occur between face images of the same individual due to pose, metabolic, time changes, etc. In the visible and active infrared spectrums variations can occur due to facial expressions, pose, etc. By using non linear techniques, unwanted variations resulting from these changes may be eliminated or reduced. The following approaches are available [9-10, 22-26]: PCA, KPCA, LGE, TLGE, KGE, LDA, KDA, LPP, TLPP, KLPP, LSDA.

4. Face extraction and alignment

For face learning and recognition, the face region was extracted and normalized. The normalization is based on the extracted eyes and mouth positions. The normalized face images can be used for face recognition and for multispectral image fusion. Two different approaches were developed. The first approach operate in visible and active infrared spectrums (NIR and SWIR) and the second approach operates in the thermal infrared spectrums [26, 27]. Figures 6 and 7 give the algorithms developed in order to extract, align and normalize the face area of interest in different spectrums.

```

1. Edge extraction of the original image (Canny edge
detection);
2. Morphological filtering of the edge image;
3. Projection profiles (horizontal and vertical) for face
region extraction;
4. Eyes extraction:
  a. Extract ROI of the upper part of the face;
  b. Left eye extraction:
    i. Analyze left area (Left ROI);
    ii. Local Otsu thresholding;
    iii. Connected components analysis;
    iv. Circular Hough Transform;
    v. Extract left eye candidate position.
  c. Right eye extraction:
    i. Analyze right area (Left ROI);
    ii. Do the same steps as for the left eye;
    iii. Extract right eye candidate position.
  d. Analyze the left and right eyes positions:
    i. Distance analysis of eyes candidate positions.
5. Mouth extraction:
  a. Extract ROI of the lower part of the face;
  b. Local Canny edge extraction;
  c. Horizontal projection profile for mouth area
localization;
  d. Extract center of the area as the center position of
the mouth.

```

Fig. 6. Pseudo-code for active IR / Visible face extraction algorithm.

```

1. Histogram stretching for image intensity normalization of
the original image;
2. SVM classifier for face extraction;
3. High temperature areas are extracted (these areas contain
internal eyes regions);
  a. If in an ROI around estimated eyes positions we have 2
hot areas:
    i. An ROI is selected close to the detected areas;
    ii. Canny edge detector is applied;
    iii. Circular Hough Transform is performed in order
to extract the candidate eyes positions.
  b. If no hot areas are found, we expect the presence of
eyeglasses:
    i. Canny edge detector is applied in a larger area
around estimated eyes region;
    ii. Circular Hough Transform is performed in order
to extract the eyeglasses;
    iii. The central position of the eyeglasses is
selected as candidate eyes positions.
  c. Analyze the left and right eyes positions:
    i. Distance analysis of eyes candidate positions.
4. Mouth extraction (the same procedure as for active IR is
applied):
  a. Extract ROI of the lower part of the face;
  b. Local Canny edge extraction;
  c. Horizontal projection profile for mouth area
localization;
  d. Extract center of the area as the center position of
the mouth.

```

Fig. 7. Pseudo-code for thermal IR face extraction algorithm.

The obtained positions of eyes and mouth give three positions that are used to extract the parameters of an affine transformation matrix to a normalized triangle position in a 128x128 image. A region of interest is then extracted relative to the normalized triangle. This region keeps the most interesting face areas for further processing. The extracted face images are used for face learning, face recognition and multispectral data fusion (after this step all the obtained images are co-aligned).

5. Multi-scale fusion

The proposed framework implements a multi-scale fusion scheme (Fig. 8) in order to evaluate the performance of multispectral face fusion in face recognition. The following multi-scale fusion techniques are used [28, 29, 30]:

- Laplacian pyramid;
- FSD pyramid;
- Ratio pyramid;
- Contrast pyramid;
- Gradient pyramid;
- Morphological pyramid;
- Discrete wavelets transform (DWT);
- Haar wavelets (SIDWT);
- Averaging.

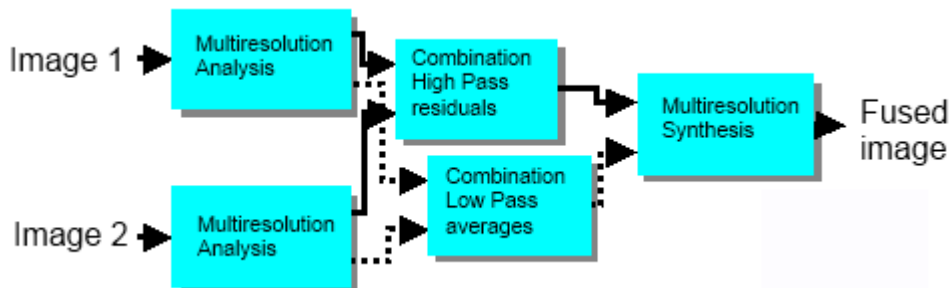


Fig. 8. Multi-scale fusion scheme.

The multi-scale fusion can be performed in different texture channels (Fig. 9) and between different texture images of different spectrums (Fig. 10).



Fig. 9. Example of LATP Lower and Upper image fusion for SWIR using DWT, MWIR using SIDWT and NIR using FSD.



Fig. 10. Example of inter-spectrums fusion: SWIR-LWIR using Gradient Pyramid, Visible-SWIR using PCA and Visible-MWIR using Laplacian Pyramid.

6. Experimental results

Multiple tests have been conducted using multi-scale fusion between Lower and Upper spectral texture images obtained using LTP and LTP and also between different spectrums. Different face recognition algorithms were used. 1000 tests for each scenario were conducted and average results analyzed.

Inter-spectral fusion in the texture space gave very high performances when the SWIR spectrum was fused with the visible and LWIR spectrums. 99% success rate was obtained for SWIR-Visible and SWIR-LWIR for Equinox database using DWT, SIDWT. Contrast Pyramid gave the same performance for SWIR-LWIR fusion. These results were obtained using non linear subspace learning techniques : LSDA, LGE and LPP. For Laval university database, the best performance for inter-spectral fusion was obtained for the Visible-MWIR fusion using DWT and Contrast Pyramid with respectively 95% and 93% success rate. With SIDWT, Visible-NIR gave the best result with 95% success rate. These results were obtained using LDA face recognition technique.

Intra-spectral fusion was conducted by fusing Upper and Lower textures images obtained using LTP and LTP. In equinox database the best performance was obtained with SWIR using LTP. PCA face recognition technique was the best performing with Contrast Pyramid (97% success) and SIDWT (96% success). For Laval University, the best performance was obtained for visible spectrum using LTP and LDA face recognition. 92% success was obtained using Contrast Pyramid and SIDWT.

7. Conclusion

This work present a new fusion framework for multispectral face recognition in the textures space. Inter- and Intra-spectral multi-scale fusion is developed in texture space. Also, a new approach for face extraction and alignment in active and passive IR spectrums is presented. This approach permit the alignment of face images thus permitting inter-spectral fusion.

The obtained results show that the proposed framework permit the increase of face recognition performances. The highest performance is obtained for inter-spectral fusion particularly when SWIR is used. The proposed scheme for Intra-spectral fusion lead to a very interesting results. It show that high recognition rates can be achieved when only one spectrum is available for face recognition when using the proposed texture descriptors.

REFERENCES

- [1] A.A. Ross, K. Nandakumar, and A.K. Jain, Handbook of Multibiometrics, Springer, 2006.
- [2] A.K. Jain, P. Flynn, and A.A. Ross, Handbook of Biometrics, Springer, 2007.
- [3] A. Bronstein, M. Bronstein, and R. Kimmel., "Three-dimensional face recognition", International Journal of Computer Vision, vol. 64, pp. 5-30, 2005.
- [4] S. Kong, J. Heo., B. Abidi, J. Paik, and M. Abidi, "Recent advances in visual and infrared face recognition: a review", Computer Vision & Image Understanding, vol. 97, pp. 103-135, 2005.
- [5] M.A. Akhloufi, A. Bendada, and J.C. Batsale, "State of the art in Infrared face recognition", QIRT Journal, vol.5, n.1, pp.3-26, June 2008.
- [6] P. Buddharaju, I.T. Pavlidis and P. Tsiamyrtzis, "Pose-Invariant Physiological Face Recognition in the Thermal Infrared Spectrum", IEEE Conference on Computer Vision and Pattern Recognition Workshop, New York , pp. 53-53, 2006.
- [7] M.A. Akhloufi, A. Bendada, "Thermal Faceprint: A new thermal face signature extraction for infrared face recognition", Proceedings of the 5th Canadian Conference on Computer and Robot Vision (CRV 2008), Windsor, ON, Canada, May 2008.
- [8] M.A. Akhloufi, A. Bendada, "Infrared Face Recognition Using Distance Transforms", Proceedings of the 5th International Conference on Image and Vision Computing (ICIVC 2008), Vol.30, pp. 160-163, Paris, France, July 04-06, 2008.
- [9] M.A. Akhloufi, A. Bendada, and J.C. Batsale, "Multispectral face recognition using non linear dimensionality reduction", Proceedings of SPIE Visual Information Processing XVIII conference, Volume 7341, SPIE Defense, Security and Sensing, 7341-19, Orlando, Florida, USA, 2009.
- [10] M.A. Akhloufi and A. Bendada, "Probabilistic Bayesian framework for infrared face recognition", In proc. International Conference on Machine Vision, Image Processing, and Pattern Analysis 2009, pp 66-70, December 2009.
- [11] M.A. Akhloufi, X. Maldague, W. Ben Larbi, "A New Color-Texture Approach for Industrial Products Inspection," Journal of Multimedia (JMM), Vol. 3, Issue. 3, pp. 44-50, Academy Publisher, ISSN: 1796-2048, July 2008.
- [12] M.A. Akhloufi, W. Ben Larbi, X. Maldague, "Framework for color-texture classification in machine vision inspection of industrial products", Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, Montréal, Qc, Canada, October 2007, pp.1067-1071.

- [13] T. Ojala, M. Pietikainen., and D. Harwood, "A comparative study of texture measures with classification based on feature distributions", *Pattern Recognition* 29 (1), 1996, pp. 51-59.
- [14] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Recognition with Local Binary Patterns", In *proc. of Europ. Conf. on Comp. Vision (ECCV 2004)*, LNCS 3021, pp. 469–481, 2004.
- [15] G. Zhang, X. Huang, L. Stan, Y. Wang and X. Wu, "Boosting Local Binary Pattern (LBP)-based face recognition", In *Proc. Chinese conference on biometric recognition, Advances in biometric - Person authentication 2004*, vol. 3338, pp. 179-186, 2004.
- [16] H. Yang and Y. Wang, "A LBP-based Face Recognition Method with Hamming Distance Constraint", In *Proc. of Fourth International Conference on Image and Graphics 2007 (ICIG 2007)*, pp. 645-649, 2007.
- [17] T. Ahonen, A. Hadid and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 2037-2041, December, 2006.
- [18] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions", In *Proc. IEEE International Workshop on Analysis and Modeling of Faces and Gestures 2007 (AMFG 2007)*, 2007.
- [19] Equinox, "Multimodal face database", <http://www.equinoxsensors.com/products/HID.html>.
- [20] W. Niblack, *An Introduction to Digital Image Processing*, pp. 115-116, Englewood Cliffs, N.J., Prentice Hall, 1986
- [21] J. Sauvola, T. Seppänen, S. Haapakoski, and M. Pietikainen, "Adaptive Document Binarization", *4th International Conference on Document Analysis and Recognition*, Ulm, Germany, August, 1997.
- [22] L. Journaux, I. Foucherot, and P., Gouton, "Operational Comparison of Dimensionality Reduction Techniques applied on Multispectral Satellite Images," *Proc. 4th Int. Conf. Signal-Image Technology & Internet-Based Syst.*, 476-484, 2006.
- [23] X. He and P. Niyogi, "Locality Preserving Projections", *Proc. of Advances in Neural Information Processing Systems (NIPS) 16*, 2003.
- [24] D. Cai, X. He, Y. Hu, J. Han and T. Huang, "Learning a Spatially Smooth Subspace for Face Recognition", *Proc. IEEE Conf. Computer Vision and Pattern Recognition Machine Learning (CVPR'07)*, pp. 1-7, July 2007.
- [25] M.A. Akhloufi, and A. Bendada, " Locally adaptive texture features for multispectral face recognition", *Proc. of the 2010 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2010)*, October 2010.
- [26] M.A. Akhloufi, and A. Bendada, "A new framework for face recognition in and beyond the visible spectrum", *Proc. of the 2010 IEEE International Conference on Systems, Man, and Cybernetics (SMC 2010)*, October 2010.
- [27] M.A. Akhloufi, and A. Bendada, "Multispectral Infrared Face Recognition: a comparative study", *Proc. of the 10th Quantitative Infrared Thermography (QIRT) conference (QIRT 2010)*, July 2010.
- [28] O. Rockinger, and T. Fechner, "Pixel-level image fusion: The case of image sequences", In: *Proc. SPIE*, Vol. 3374, pp. 378-388, 1998.
- [29] F. Sadjadi, "Comparative Image Fusion Analysis", *Proc. of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, Vol. 03, p.8, 2005.
- [30] Y. Wang, and Z. Chen, " Performance analysis of hierarchical fusion approaches for CCD/IR images", *Proc. of the SPIE International symposium on instrumentation and control technology*, vol. 5253, pp. 996-1000, 2003.