Extraction and fusion of spectral parameters for face recognition

Bertrand Boisier, Bastien Billiot, Zeguidi Abdessalem, Pierre Gouton, Jon Yngve Hardeberg

Extraction and fusion of spectral parameters for face recognition

B. Boisier *, B. Billiot *, Z. Abdessalem *, P. Gouton * , J. Y. Hardeberg **

* Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

** Gjøvik University College, P.O Box 191, N-2802 Gjøvik, Norway

ABSTRACT

Many methods have been developed in image processing for face recognition, especially in recent years with the increase of biometric technologies. However, most of these techniques are used on grayscale images acquired in the visible range of the electromagnetic spectrum.

The aims of our study are to improve existing tools and to develop new methods for face recognition. The techniques used take advantage of the different spectral ranges, the visible, optical infrared and thermal infrared, by either combining them or analyzing them separately in order to extract the most appropriate information for face recognition.

We also verify the consistency of several keypoints extraction techniques in the Near Infrared (NIR) and in the Visible Spectrum.

Keywords: Face recognition, Visible, Near Infrared, Short Wave Infrared, Feature extraction, feature descriptor

I. INTRODUCTION

Face recognition have been, for the last thirty year, a center of interest due to the fact that there much need in security applications, more precisely, in surveillance and identity control, which needs to automatically identify people in a simple and unobtrusive way beyond the limits of biometric control devices such as fingerprints or iris recognition.

Face recognition rely on several techniques which have grown in maturity these last years [1] . A current challenge in face recognition is to improve existing algorithms to have better results and extract more reliable information about face expression and different people.

This study was made to check if other features from a large range of spectral wavelengths can be found and can provide any improvements concerning face recognition.

Recent studies have shown advantages and disadvantages of using the visible and infrared spectrum for face recognition, visible range pictures are more consistent in detected features whereas infrared, mainly thermal infrared, can make an improvement particularly in harsh environments. [2-4]

Face recognition reduce accuracy when acquiring data in an uncontrolled environment so it is important to have a sufficient amount of characteristics to recognize a face.

Our study uses three acquisitions devices: a digital SLR camera, a near-infrared camera (800-2600 nm) and a thermal camera (12000 nm). We want here to extract better information by merging spectral bands and by checking if some complementary features in each band can be found.

Several face recognition techniques are based on feature extraction algorithms, in order to characterize them by describing area or Points Of Interests (POIs). We have implemented several methods to compare the efficiency and adaptability of such algorithms in different wavelength area. These tests allow finding the more accurate method to use within features detection algorithm and to optimize the multiple information we can extract from several wavelength ranges. For this study, we make a database of images of 15 people, with 4 different stands with one reference image, in the visible, Near-Infrared (NIR)-Short Wave Infrared (SWIR), and Long Wave Infrared (LWIR) ranges.

Figure 1. Example of acquired picture database for one person

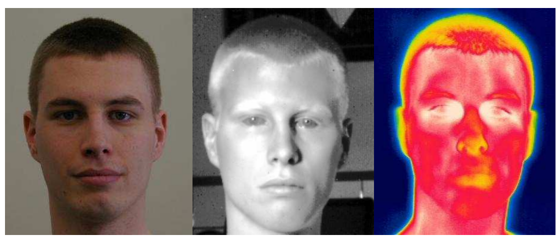

Figure 2. From left to right : visible acquired face, NIR-SWIR acquired face, LWIR acquired image (thermal infrared)

II. KEYPOINT EXTRACTION USING FEATURE DETECTORS

Subsequently, algorithms are being compared within the same picture in SWIR and in visible. In order to extract features from these two pictures acquired with approximately the same range of information in it. Spatial resolution has been conserved with each sensor resolution. The consistency and amount of features extracted under varying conditions of acquisition have been compared to check the robustness of facial features extraction.

With concordance with a paper from Mikolajczyk et al. [5] which evaluates the performances of several features detectors, we implemented a Hessian-Affine and a Harris-Affine detector to evaluate features in every spectral band. Acquired pictures admit a small variation of posing angle and face expression variations so performances of such a feature extraction kernel is sufficient to extract parameters.
The Hessian-Affine and the Harris-Affine detectors are feature detectors known as affine-invariant detectors; it allows detecting identifiable characteristics points and regions, which remain invariant to affine transformation which is also invariant to affine illumination changes [6]. These affine detectors rely on the combination of corner point detection, multi-scale analysis through Gaussian scale-space and affine normalization using an iterative affine shape adaptation algorithm.

We check these two features extractors with visible, and NIR-SWIR pictures acquired with approximately the same spatial resolution for the two sensors in order to keep the same ratio and pixel amount and information for features detection. Thermal pictures are not used with these algorithms; they give us a few keypoints that are not significantly large.

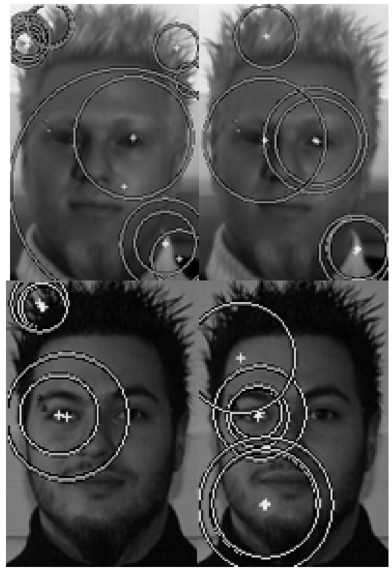

Figure 3. Display of several features extracted with Harris-Affine (first column), Hessian-Affine (second column)

A threshold of sensitivity is applied in order to display the most robust features extracted from a Hessian-Affine and a Harris-Affine detectors. The amount of features depends on this threshold and on the variability and robustness of picture variations.

	Hessian-A	Harris-A
IR	80	30
VIS	96	22

Figure 4. Number of keypoints acquired with Hessian-Affine and Harris-Affine for the same picture in both NIR and Visible

Owing to these two detectors are almost identical we describe the Hessian-Affine methods here because we have more features extracted. This detector chooses interest points based on the Hessian matrix at that point [6] :

$$H(x) = \begin{bmatrix} L_{xx}(x) & L_{xy}(x) \\ L_{xy}(x) & L_{yy}(x) \end{bmatrix}$$

$L_{xx}(x)$ is the second partial derivative in the "x" direction and $L_{xy}(x)$ is the mixed partial second derivative in the "x" and "y" directions. The derivatives are computed in the current iteration scale and thus are derivatives of an image smoothed by a Gaussian kernel:

$$L(x) = g(\sigma_I) \otimes I(x)$$

The derivatives must be scaled appropriately by a factor related to the Gaussian kernel σ_I^2.At each scale, POIs are simultaneously local extrema of both the determinant and trace of the Hessian matrix.

$$DET = \sigma_I^2 \left(L_{xx}L_{xy}(x) - L_{xy}^2(x)\right)$$
$$TR = \sigma_I(L_{xx} + L_{yy})$$

By this way, the Hessian-Affine detector responds well to textured scenes, more points are detected in the visible range with this accurate interest point detector.
Tests have been extended to some acquired pictures with simulated decreased quality of acquisition environment.

	100	80	60	40
IR	121	78	43	9
Vis	136	97	42	13

Figure 5. Acquired keypoints with a decreasing luminosity in percent

The Hessian-Affine algorithm is not efficient enough between affine illuminant changes. However in a weak variability and mastered environment the use of Hessian-Affine detectors allow to extract sufficient points.
This methods offers good results in characterizing keypoints in a little processing time to be affine invariant, it is really sensitive to case with huge texture and noise.

We want to check with several features descriptors such as SIFT in order to check if we have a larger number of features extracted and well-describe to be robust against affine transformation and projection.

III. KEYPOINTS DESCRIPTORS

We have to use a technique allowing the creation of a vector which is characteristic and determining vector which keypoints remains invariant in order to determine the most robust algorithm for face recognition in different spectral wavelength.

SIFT (Scale Invariant Feature Transform) [7] , which mix a keypoint dectector and a keypoint descriptor bring the best results. SIFT is mostly used in computer vision and in computer imaging so as to detect and describe points of interest of a bidimentional picture.

The SIFT algorithm is made in 4 steps. First, a detection of extrema in a scale space is made with the use of Difference of Gaussian (DoG) method. Only the strongest points detected by the DoG, are kept by eliminating features too close from the edge of the image and from other ones. Extrema which have a weak contrast (to avoid a noise sensibility) have been excluded too.
An orientation vector is described to allow a rotation invariance of keypoints. The gradient histogram is computed, in order to detect the primary direction of a keypoints.

A descriptor of POIs is created by calculating the gradient magnitude and orientation for each pixel of the image in a neighborhood around each point of interest. The amplitude is weighted by a Gaussian centered on the point of interest to promote gradients near the center of the neighborhood. The amplitudes of the neighboring pixels are then combined in the histograms of orientation. The length of each arrow corresponds to the sum of the amplitudes of the gradients near the corresponding direction.
Every keypoints is determined by a 4x4 matrix containing the orientation histograms which are composed of 8 orientations. 128 elements describe each interest point. Information is normalized in order to become illuminant invariant.

A SIFT algorithm have been implemented to extract and descript features of many faces to compare the number of keypoints between each spectral range.

	Person N° 1	Person N° 2	Person N° 3	Person N° 4
Visible	174	105	164	352
IR	222	231	183	415

Figure 6. Number of SIFT features extracted from four different faces.

SIFT isn't completely invariant to affine transformation so that J-M. Morel and G. Yu develop a new method based on SIFT which upgrade the performance against affine transformation. This method is called Affine Scale Invariant Feature transform (ASIFT). [8]

	Person N° 1	Person N° 2	Person N° 3	Person N° 4
Visible	87	59	63	140
IR	151	151	122	244

Figure 7. Number of ASIFT features extracted from four different faces.

IV. MATCHING RESULTS

A matching of keypoints extracted from a face is implemented, to determine how many points can be matched with another to check the relevance and consistency of described features. LWIR pictures provide less POIs than visible and NIR-SWIR ones, however it is a reliable acquisition system which is invariant to illuminant change. ASIFT described points are shown here.

ASIFT	Face 1	Face 2	Matches
IR	643	634	122
VIS	489	432	94
THERM	250	254	54

Figure 8. Matching of a peer of face (reference face with no particular expression between face with closed eyes) and numbers of keypoints extracted.

The number of characteristic points in the infrared range is more important, an average of 242 points per person in the database have been extracted with ASIFT, whereas this number is half reduced in the visible spectrum, with an average of 151 points per person.

Features described by the ASIFT descriptor allow us to perform a better matching in the IR spectral range, a larger amount of invariant keypoints are described in the NIR-SWIR pictures than in the visible range.

V. CONCLUSION AND DISCUSSIONS

We aim this study to check for complementary information and efficiency of each features descriptors and keypoints detectors in order to mix information or to use it separately.
We admit that we can increase the robustness of face recognition using the near optical IR. It appears that well-descript features which doesn't stack within textures and noise variations can be exploited in a better way. Moreover with IR acquired pictures, we ensure a higher rate of recognition due to the fact that we get a higher number of features with some algorithms less sensible to textures and variations, so we operate a more efficient matching. Combining the features acquired in large and suitable spectral bandwidth can provide an enhancement to face recognition. Future work includes the selection and optimization of this appropriate information and the comparative of methods concurring with these ones. The uses of others descriptors and detectors that can provide greater results will be used for future tests in combination with all of the spectral data acquired.

REFERENCES:

1. W. Zhao and R. Chellappa. Image based face recognition, issues and methods. In B. Javidi,editor, Image Recognition and Classification, pages 375C402. Mercel Dekker, 2002.

2. F.J. Prokoski, History, current status, and future of infrared identification. Proceedings of the IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications,
Hilton Head, 2000.

3. Joseph Wilder, P. Jonathon Phillips, Cunhong Jiang, Stephen Wiener, Comparison of visible and infra-red imagery for face recognition. Proceedings of the 2nd International Conference on
Automatic Face and Gesture Recognition, Killington, VT, 1996, pp. 182–187.

4.D. Socolinsky, A. Selinger, J. Neuheisel, Face recognition with visible and thermal infrared imagery, Comput. Vis. Image Und. 91 (2003) 72–114.

5. K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir and L. Van Gool, A comparison of affine region detectors. In IJCV 65(1/2):43-72, 2005

6. K. Mikolajczyk, C.Schmid, An affine invariant interest point detector, Proceedings of the 7th European Conference on Computer Vision, Copenhagen, Denmark, vol. I, pp. 128–142.

7. D. G. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 2004.

8. J-M. Morel and G. Yu, ASIFT: A new framework for fully affine invariant image comparison, SIAM Journal on Imaging Science, 2009.