# Evaluation of Keypoint Descriptors for Gender Recognition

Florencia Soledad Iglesias, María Elena Buemi,
Daniel Acevedo, and Julio Jacobo-Berlles

Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales,
Departamento de Computación, Argentina
iglesiasflorencia@gmail.com, {mebuemi,dacevedo,jacobo}@dc.uba.ar

**Abstract.** Gender recognition is a relevant problem due to the number and importance of its possible application areas. The challenge is to achieve high recognition rates in the shortest possible time. Most studies are based on Local Binary Patterns (LBP) and its variants to estimate gender. In this paper, we propose the use of Binary Robust Independent Elementary Features (BRIEF), Oriented FAST and Rotated BRIEF (ORB) and Binary Robust Invariant Scalable Keypoints (BRISK) in gender recognition due to their good performance and speed. The aim is to show that ORB and BRISK are faster than LBP but allow to achieve similar recognition rates, which makes them suitable for real-time systems. For the best of our knowledge, it has not been studied in literature.

**Keywords:** Gender recognition, LBP, Keypoint Descriptors.

## 1 Introduction

Image processing and computer vision play an important role in the study of human biometric identification attributes. Gender, age and ethnicity are aspects that identify individuals and allow the improvement of various applications. Human-Computer Interaction systems, surveillance, content-based indexing and searching, biometric systems, demographic research, market research and construction of targeted advertisements are some areas where this attributes would be useful because they require reliable user information to provide service correctly. The gender recognition problem has been studied from different viewpoints. Most studies are based on the analysis of faces because these are one of the less invasive biometric characteristics; recently, head-shoulders combination was investigated showing good accuracy in gender estimation [8]. In several studies it was observed that intentionally adding distorted images to the training data allows classifiers to be more robust [3]. The present challenge focuses on getting elevated rates of gender recognition with the fastest response times.

Fig. 1 shows the structure of a typical gender recognizer. It can be divided into three consecutive steps which are *Face Detection*, *Feature Extraction* and *Gender Classification*.

E. Bayro-Corrochano and E. Hancock (Eds.): CIARP 2014, LNCS 8827, pp. 564–571, 2014.

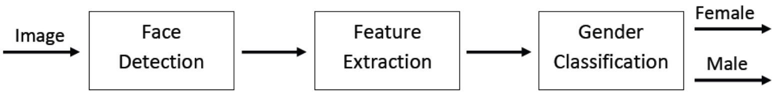


**Fig. 1.** A Tipical Gender Recognition System Structure

A method used to locate human faces in an image (in *Face Detection* stage), is the Viola-Jones algorithm [17] that provides high success rates and a low computational cost, making it feasible for use in real time. Various techniques have been analyzed for *Feature Extraction*, from those based on dimensionality reduction (LDA, ICA, PCA) [5,10] to the most recent ones that have attempted to improve robustness to rotations, brightness variation and noise [1,3]. Among the most widely used techniques, are Local Binary Pattern (LBP) and its variants LBP Histogram Fourier (LBP-HF) and Rotation Invariant Uniform LBP ($LBP^{uri}$) [1,14,19,20]. Gender recognition is a binary classification problem; the subject can be classified as either male or female. The most widely used classifiers for solving the *Gender Classification* stage, are Support Vector Machine(SVM) and Adaboost. SIFT [9] and SURF [4] are two keypoint detectors and descriptor algorithms that are commonly used in computer vision to extract image descriptors because they provide good performance at a low computational cost. They inspired the creation of a group of methods that generate binary feature vectors; BRIEF [6], ORB [15] and BRISK [7] are three of them. Several interesting investigations that test and analize these methods have been published since 2011 [11]. The studies have shown that these techniques are faster than SIFT and SURF, achieving high success rates in recognition problems. The aim of this paper is to evaluate the use of BRIEF, ORB and BRISK in face gender recognition and show that ORB and BRISK are more suitable to use in real-time gender recognizers than LBP, which for the best of our knowledge, has not been studied in the literature. The recognizers are generated combining this methods with SVM and Adaboost. In our experiments we analyze and compare their accuracy and response time to observe which has the best ratio between recognition and time performance.

This paper is organized as follows: in Section 2, we give an overview of the methods used to generate the gender recognizers. Section 3 shows the experimental results and comparisons between the recognizers studied. Finally, in Section 4 we expose our conclusions and discuss the future works.

## 2 Feature Extraction and Gender Classification

In this paper we focus only on the *Feature Extraction* and *Gender Classification* stages to generate the gender recognizers (see Fig. 1). We categorize the feature extraction methods used into two groups: those based on LBP variations and those that generate descriptors from a set of points of interest such as BRIEF, ORB and BRISK. The keypoint detection algorithms suggested for these techniques [6,7,15] detect only a few points per face image and the descriptors generated using these points have little descriptive power. Wang et al. [18] studied gender recognition using SIFT descriptors and found the same problem; they implemented a "dense" version of SIFT as a solution. The same idea is used in this paper: points are extracted from the face images using a regular grid, creating a "dense" version of BRIEF, ORB and BRISK. Faces are previously reshaped to 62 x 62 pixels; the same number of keypoints are extracted for each one. The classification algorithms used to estimate the gender of faces are SVM with a linear kernel and Adaboost. The input of each classifier is the set of descriptors extracted from the images.

### 2.1 LBP

The original LBP operator [13] labels the pixels of an image with a binary number defined as

$$LBP(x,y) = \sum_{n=0}^{7} S(I_n - I(x,y))2^n \tag{1}$$

where $I_n$ with $n = 0, 1, \ldots, 7$ are the neighbors of $I(x,y)$ and the thresholding function $S(z)$ is 1 if $z \geq 0$ and 0 otherwise. A local binary pattern is called uniform if it contains at most two bitwise transitions from 0 to 1 (or viceversa) when the binary string is considered circular.

$LBP^{uri}$ [13] is built rotating circularly each LBP binary code into its minimum value

$$LBP^{uri} = \min_i ROR\,(LBP, i) \tag{2}$$

where $ROR(x,i)$ denotes the circular bitwise right rotation of bit sequence $x$ by $i$ steps. LBP-HF is another rotation invariant image descriptor based on uniform LBP [2,13]. The feature vector consist of three LBP histogram values (all zeros, all ones, non-uniform) and the Fourier magnitude spectrum is defined as

$$|H(n,u)| = \sqrt{H(n,u)\overline{H(n,u)}} \tag{3}$$

where $H(n,\cdot)$ is the DFT of *n-th* row of the histogram $h_I(U_P(n,r))$:

$$H(n,u) = \sum_{r=0}^{P-1} h_I(U_p(n,r))e^{-i2\pi ur/P} \tag{4}$$

and $U_P(n,r)$ denotes a specific uniform LBP pattern.

In order to characterize a face more efficiently it is also necessary to store spatial information with any variants of LBP. To do this, a face image is divided into small regions from which LBP descriptors are extracted and concatenated into a single feature vector [1].

### 2.2 BRIEF

The BRIEF descriptor [6] is a description of an image generated from a set of points of interest. For each keypoint $k$, a patch $P_k$ of size $S \times S$ around $k$ is considered. The feature vector of $P_k$ is constructed from a set of pairwise intensity comparisons. To do that, a test $\tau_k$ on $P_k$ is defined as

$$\tau_k\left(P_k, x, y\right) = \begin{cases} 1 & \text{if } p_k\left(x\right) < p_k\left(y\right) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $p_k(x)$ and $p_k(y)$ are the intensity in a smoothed version of $P_k$ at points $x$ and $y$, respectively. The descriptor of $P_k$ is defined as a vector of $n$ binary tests:

$$f_n\left(P_k\right) = \sum_{i=1}^{n} 2^{i-1} \tau_k(P_k, x_i, y_i) \tag{6}$$

### 2.3 ORB

The ORB descriptor [15] is based on BRIEF. The idea is to steer the BRIEF descriptor according to the orientation of keypoints. For each keypoint $k$, a set of $n$ binary tests at location $(x_i; y_i)$ define a $2 \times n$ matrix:

$$S_k = \begin{pmatrix} x_1 \cdots x_n \\ y_1 \cdots y_n \end{pmatrix} \tag{7}$$

Using the patch orientation $\theta$, calculated from the intensity centroid of the patch and the corresponding rotation matrix $R_\theta$, a "steered" version $S_\theta$ of $S_k$ is built: $S_\theta = R_\theta S_k$. The steered BRIEF descriptor is calculated as $f_n\left(P_k\right)$ using only the points in $S_\theta$:

$$g_n\left(P_k, \theta\right) = f_n(P_k) \qquad \text{for } (x_i, y_i) \in S_\theta. \tag{8}$$

where $P_k$ is the patch around $k$ and $f_n$ is the BRIEF descriptor, defined in equation (6).

### 2.4 BRISK

The idea is similar to BRIEF. In BRISK [7], the characteristic direction of each keypoint is identified to allow for orientation-normalized descriptors. The BRISK descriptor uses a pattern for sampling the neighborhood of a keypoint $k$. Considering the set $A_k$ of all sampling point pairs centered at $k$, two subsets are defined, one of short-distance pairings $S_k = \{(p_i, p_j) \in A_k : \|p_j - p_i\| < \delta_{max}\}$ and another one of $\ell$ long-distance pairings $L_k = \{(p_i, p_j) \in A_k : \|p_j - p_i\| > \delta_{min}\}$, where $\delta_{max}$ and $\delta_{min}$ are distance thresholds. For each pair $(p_i, p_j) \in A_k$, the local gradient $g\left(p_i, p_j\right)$ is estimated by

$$g\left(p_i, p_j\right) = \left(p_j - p_i\right) \frac{\widetilde{I}\left(p_j\right) - \widetilde{I}\left(p_i\right)}{\left\|p_j - p_i\right\|^2} \tag{9}$$

where $\widetilde{I}(p_i)$ and $\widetilde{I}(p_j)$ are the smoothed intensity values at points $p_i$ and $p_j$, respectively. Iterating through the point pairs in $L_k$, the overall characteristic pattern direction of the keypoint $k$ is estimated as

$$g = (g_x, g_y)^T = \frac{1}{\ell} \sum_{(p_i,p_j)\in L_k} g(p_i, p_j) \tag{10}$$

For the formation of the descriptor, BRISK applies the sampling pattern rotated by $\alpha = arctan2(g_x, g_y)$ around the keypoint $k$. The bit-vector descriptor $d_k$ is generated by performing all the short distance intensity comparisons of point pairs $(p_i^\alpha, p_j^\alpha) \in S_k$ (*i.e.* in the rotated pattern), such that each bit $b$ corresponds to:

$$b = \begin{cases} 1 & \text{if } \widetilde{I}(p_j^\alpha) > \widetilde{I}(p_i^\alpha) \\ 0 & \text{otherwise} \end{cases} \quad \forall (p_i^\alpha, p_j^\alpha) \in S_k \tag{11}$$

## 3 Experimental Results

Combining the feature extraction methods with the classifiers, we built the gender recognizers. The structure *feature extractor + classifier* is used in this paper to differentiate the recognizers; for example *ORB + SVM* indicates that SVM and ORB are the methods used; when using LBP, the size of the blocks taken is given in brackets. To evaluate their capability and performance, two databases were considered: FERET [12] and FEI [16]. All images correspond to front faces of different people whose ages are between 18 and 40. In the experiments, we used a subset of 300 faces of FERET and a subset of 200 faces of FEI. 50 faces of FEI were extracted for testing the recognizers; the remaining images were used for training. Similarly, 200 faces of FERET were used for training and 100 for testing. The number of men and women in all cases is the same. Some examples from the FERET database are shown in Fig. 2.

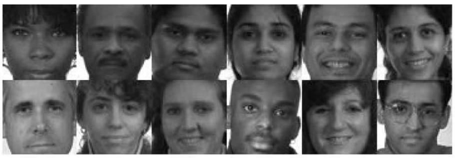

**Fig. 2.** Examples of frontal faces from the FERET database

We analyze the accuracy of each gender recognizer training and testing them with each database. Fig. 3 shows the recognition results achieved using BRIEF, BRISK and ORB as well as to extend the results obtained with the recognizers based on LBP and LBP-HF that present the highest rates; above each bar is the percentage of recognition. The recognition rates obtained are more elevated with FEI than with FERET, however this database has more variety in the physical features of the faces. The best performance is achieved with the

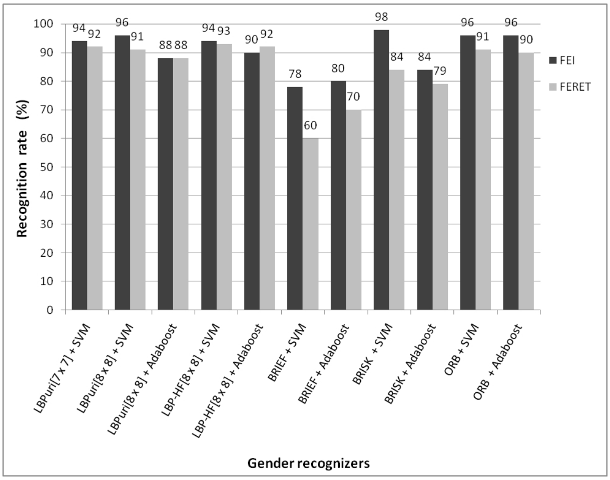


**Fig. 3.** Recognition rates achieved by different gender recognizers

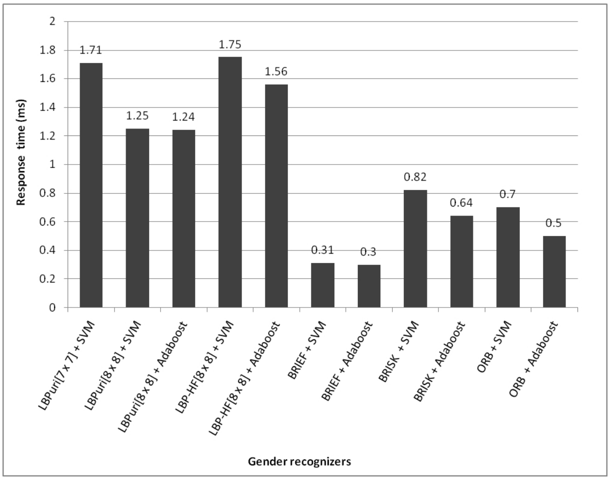


**Fig. 4.** Response time of different gender recognizers

recognizer *ORB+Adaboost* and with the ones that use the SVM classifier, except when using the feature extractor BRIEF; they show between 96% and 94% accuracy with FEI and between 91% and 93% with FERET. *BRIEF+SVM* and *BRIEF+Adaboost* provide rates considerably lower than those presented by the other methods. This experiment allows us to observe that *ORB+SVM* and *ORB+Adaboost* achieved similar rates to those obtained using LBP variants.

The feature extraction is a time-consuming task so it is important to investigate which are the most computationally efficient methods. To do this, we study the response time of each recognizer, calculating the average time per face. The computer used in our experiments has an Intel Core i5-2310 processor running at

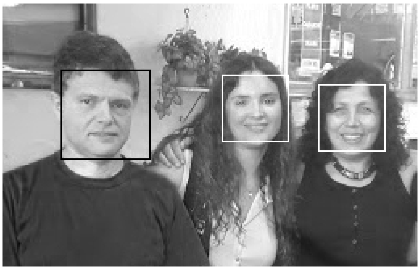

**Fig. 5.** Experimental results of gender recognition using ORB and Adaboost. Black and white rectangles indicate the subjects are recognized as one male and two females, respectively.

2.90 GHZ and 4 GB of RAM memory. Fig. 4 shows that the fastest methodologies are the ones using BRIEF, ORB and BRISK, being those based on LBP the slowest ones. It can also be seen that Adaboost is a faster classifier than SVM.

Analyzing the results of both experiments, we can see that although BRIEF is the fastest method, the recognition rate achieved by the recognizer is poor compared with the other tested methods. LBP variants and ORB reach high and similar levels of recognition performance, but differ significantly in time performance; ORB proves to be the fastest among these techniques. *ORB + Adaboost* is the "best" gender recognizer studied because it has the best ratio between recognition performance and time response. Fig. 5 shows and output visualization of gender estimation using this recognizer.

## 4 Conclusions

In this paper many combinations of *feature extractor + classifier* for gender recognition have been compared from the point of view of recognition accuracy and response time. The feature extractors considered here were: Uniform Rotation Invariant LBP, LBP-HF, BRIEF, BRISK ORB, and the classifiers were Adaboost and SVM. Recognition performance was poor for BRIEF, specially in the case of the FERET database. It was better for BRISK and even better for ORB, as well as for the LBP methods. For these latter ones, time performance is poor compared with BRIEF, BRISK and ORB. For classification and time performance, the combination *ORB+AdaBoost* is the best. Future work aims at the implementation of a real-time system using the recognizer *ORB+Adaboost* and its application to mobile devices, as well as to extend the studies to consider age recognition problem.

## References

1. Ahonen, T., Hadid, A., Piettikäinen, M.: Face description with local binary patterns: Application to face recognition. IEEE Transactions on pattern Analysis and Machine Intelligence 28(12) (2006)
2. Ahonen, T., Matas, J., Piettikäinen, M., He, C.: Rotation invariant image description with local binary pattern histogram fourier features. Scandinavian Conference on Image Analysis, SCIA (2009)

3. Andreu, Y., García-Sevilla, P., Mollineda, R.A.: Face gender classification: A statistical study when neutral and distorted faces are combined for training and testing purposes. Image and Vision Computing 32(1), 27 (2014)
4. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part I. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
5. BekiosCalfa, J., Buenaposada, J.M., Baumela, L.: Revisiting linear discriminant techniques in gender recognition. IEEE Transactions on pattern Analysis and Machine Intelligence 33(4), 858–864 (2011)
6. Calonder, M., Lepetit, V., Strecha, C., Fua, P.: BRIEF: Binary robust independent elementary features. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part IV. LNCS, vol. 6314, pp. 778–792. Springer, Heidelberg (2010)
7. Leutenegger, S., Chli, M., Siegwart, R.Y.: Brisk: Binary robust invariant scalable keypoints. In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 2548–2555. IEEE (2011)
8. Li, M., Bao, S., Dong, W., Wang, Y., Su, Z.: Head-shoulder based gender recognition. In: IEEE International Conference on Image Processing (ICIP), pp. 2753–2756 (2013)
9. Lowe, D.G.: Object recognition from local scale-invariant features. In: The Proceedings of the Seventh IEEE International Conference on Computer Vision 1999, vol. 2, pp. 1150–1157 (1999)
10. Lu, L., Shi, P.: A novel fusion-based method for expression-invariant gender classification. In: IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2009, pp. 1065–1068 (2009)
11. Muja, M., Lowe, D.: Fast matching of binary features. In: 2012 Ninth Conference on Computer and Robot Vision (CRV), pp. 404–410 (2012)
12. Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.J.: The feret database and evaluation procedure for face recognition algorithms. Image and Vision Computing 16(5), 295–306 (1998)
13. Pietikäinen, M., Hadid, A., Zhao, G., Ahonen, T.: Computer Vision Using Local Binary Patterns. In: Computational Imaging and Vision, vol. 40, Springer (2011)
14. Ramón-Balmaseda, E., Lorenzo-Navarro, J., Castrillón-Santana, M.: Gender classification in large databases. In: Alvarez, L., Mejail, M., Gomez, L., Jacobo, J. (eds.) CIARP 2012. LNCS, vol. 7441, pp. 74–81. Springer, Heidelberg (2012)
15. Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: Orb: an eficient alternative to sift or surf. In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 2564–2571. IEEE (2011)
16. Thomaz, C.E., Giraldi, G.A.: A new ranking method for principal components analysis and its application to face image analysis. Image and Vision Computing 28(6), 902–913 (2010)
17. Viola, P., Jones, M.J.: Robust real-time face detection. International Journal of Computer Vision 57(2), 137–154 (2004)
18. Wang, J.G., Li, J., Yau, W.Y., Sung, E.: Boosting dense sift descriptors and shape contexts of face images for gender recognition. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 96–102 (2010)
19. Wang, J.G., Wang, H.L., Ye, M., Yau, W.Y.: Real-time gender recognition with unaligned face images. In: 2010 The 5th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 376–380 (June 2010)
20. Ylioinas, J., Hadid, A., Pietikäinen, M.: Combining contrast information and local binary patterns for gender classification. In: Image Analysis, pp. 676–686 (2011)