

Redaktør: Patrick Bours, Norwegian Information Security Laboratory (NISLab),  
Gjøvik University College

Norwegian Information Security Conference  
Norsk Informasjonssikkerhetskonferanse

# NISK 2010

Gjøvik University College, Gjøvik  
23.–24. november 2010

## Program Chair

Patrick Bours HiG

## Program Committee

|                     |            |                       |                    |
|---------------------|------------|-----------------------|--------------------|
| Kristian Gjøsteen   | NTNU       | Vladimir Oleshchuk    | UiA                |
| Tor Hellesest       | UiB        | Anders Paulshus       | Conax              |
| Erik Hjelmås        | HiG        | Ragnar Soleng         | UiT                |
| Audun Jøsang        | UNIK       | Nils Kalstad Svendsen | HiG                |
| Martin Gilje Jaatun | SINTEF IKT | Svein Willassen       | Svein Willassen AS |
| Stig F. Mjølnes     | NTNU       | Eli Winjum            | FFI                |
| Leif Nilsen         | UNIK       | Andre Årnes           | HiG                |

**Norwegian Information Security Conference**  
**Norsk Informasjonssikkerhetskonferanse**

**NISK 2010**

Gjøvik University College, Gjøvik  
23-24 November 2010

**Program Chair**

Patrick Bours                      HiG

**Program Committee**

|                       |                    |
|-----------------------|--------------------|
| Kristian Gjøsteen     | NTNU               |
| Tor Helleseeth        | UiB                |
| Erik Hjelmås          | HiG                |
| Audun Jøsang          | UNIK               |
| Martin Gilje Jaatun   | SINTEF IKT         |
| Stig F. Mjølnes       | NTNU               |
| Leif Nilsen           | UNIK               |
| Vladimir Oleshchuk    | UiA                |
| Anders Paulshus       | Conax              |
| Ragnar Soleng         | UiT                |
| Nils Kalstad Svendsen | HiG                |
| Svein Willassen       | Svein Willassen AS |
| Eli Winjum            | FFI                |
| Andre Årnes           | HiG                |

© NISK-stiftelsen og Tapir Akademisk Forlag, 2010

ISBN 978-82-519-2705-5

Det må ikke kopieres fra denne boka ut over det som er tillatt etter bestemmelser i «Lov om opphavsrett til åndsverk», og avtaler om kopiering inngått med Kopinor.

*Redaktør: Patrick Bours, Norwegian Information Security Laboratory (NISlab), Gjøvik University College*

*Tapir Akademisk Forlag har som målsetting å bidra til å utvikle gode læremidler og alle typer faglitteratur. Vi representerer et bredt fagspekter, og vi gir ut ca. 100 nye titler i året. Vi samarbeider med forfattere og fagmiljøer i hele landet, og våre viktigste produktområder er:*

*Læremidler for høyere utdanning  
Fagbøker for profesjonsmarkedet  
Vitenskapelig publisering*

Forlagsredaktør for denne utgivelsen:  
Lasse.Postmyr@tapirforlag.no

Tapir Akademisk Forlag  
7005 TRONDHEIM  
Tlf.: 73 59 32 10  
Faks: 73 59 32 04  
E-post: post@tapirforlag.no  
www.tapirforlag.no

# Preface

Welcome to NISK 2010, the third edition of the Norwegian Information Security Conference. After the initial NISK conference in Agder and its follow up in Trondheim, it will now take place in Gjøvik on the 23<sup>rd</sup> and 24<sup>th</sup> of November. As before the conference will take place in combination with NIK and NOKOBIT. NISK2010 is sponsored by NISnet, the resource network of Norwegian Information Security researchers funded by the Norwegian Research Council.

This year we had 27 high quality submissions from 8 different institutes. Of those one was withdrawn and one came in too late. The remaining 25 were reviewed by 2 members of the Program Committee each and from their feedback 14 papers were selected for presentation. This means that the acceptance rate of 56% is very close to the 58% from last year. All 14 papers will get a 30 minutes timeslot for presenting the ideas. Out of the 14 papers, 8 are authored or co-authored by PhD students and 1 is co-authored by master students.

We are glad to announce that Dr. Mike Bond from the Computer Laboratory at the University of Cambridge accepted the invitation as a keynote speaker. The title of his presentation is *Chip and Empiricism: Breaking EMV, with proof*. In May 2010 Mike Bond presented the controversial paper *Chip and PIN is broken*, which he co-authored with Steven J. Murdoch, Saar Drimer, and Ross Anderson, at USENIX Security. The paper described how an EMV card can be used to make purchases at Point-of-Sale without knowing the correct PIN. During the subsequent publicity, demonstrations of the technique deployed against the live banking system aired on various European television channels.

I would like to thank all the members of the Program Committee for their valuable input in the reviewing process. Furthermore I would like to thank the organizers of NIK, Erik Hjelmås and of NOKOBIT, Tom Røise for the pleasant cooperation and last but certainly not least I would like to thank Kari Lauritzen for all the help with the practical organization of the three conferences.

# Table of Content

## NISK 2010

### Keynote

|   |   |
|---|---|
| Chip and Empiricism: Breaking EMV, with proof ..... | 1 |
| <i>Mike Bond</i>                                    |   |

### Session 1: Crypto

|  |    |
|--|----|
| Coercion-Resistant Receipts in Electronic Elections .....                                | 3  |
| <i>Håvard Raddum</i>   |    |
| Algebraic Attack on the Second class of Modified Alternating $\vec{k}$ -Generators ..... | 12 |
| <i>Mehdi M. Hassanzadeh, Tor Hellesteth</i>  |    |
| Formal Verification of Reductions in Cryptography .....                                  | 21 |
| <i>Kristian Gjøsteen, George Petrides, Asgeir Steine</i>                                 |    |

### Session 2: Biometrics

|  |    |
|--|----|
| Accelerometer-Based Gait Analysis, A survey .....                                | 33 |
| <i>Mohammad Omar Derawi</i>  |    |
| Sift Based Recognition of Finger Knuckle Print .....                             | 45 |
| <i>Baptiste Hemery, Romain Giot, Christophe Rosenberger</i>                      |    |
| Evaluation of Biometric Systems: An SVM-Based Quality Index .....                | 57 |
| <i>Mohamad El-Abed, Romain Giot, Christophe Charrier, Christophe Rosenberger</i> |    |

### Session 3: Hardware / Security

|  |     |
|--|-----|
| WinSCard Tools: a software for the development and security analysis of transactions with smartcards ..... | 69  |
| <i>Sylvain Vernois, Vincent Alimi</i>  |     |
| Robustness of TRNG against Attacks that Employ Superimposing Signal on FPGA Supply Voltage .....           | 81  |
| <i>Knut Wold, Slobodan Petrović</i>  |     |
| Security and trust for mobile phones based on virtualization .....   | 93  |
| <i>Chrystel Gaber, Jean-Claude Paillès</i>   |     |
| Non-Invasive Reverse Engineering of the Relative Position of Bus Wires .....                               | 104 |
| <i>Geir Olav Dyrkolbotn</i>  |     |

**Session 4: Information Security Management**

A Dynamic Approach to Security Management ..... 110  
*Jose J. Gonzalez, Finn Olav Sveen*  
Enhancing Credibility of a Dynamic Model Involving Hidden Behavior ..... 122  
*Jaziar Radianti, Jose J. Gonzalez*

**Session 5: Biometrics / Forensics**

Secure and Privacy Preserving Management of Biometric Templates ..... 134  
*Vincent Alimi, Rima Belguechi, Christophe Rosenberger*  
Storage and Exchange Formats for Digital Evidence ..... 146  
*Anders O. Flaglien, Aleksander Mallasvik, Magnus Mustorp, André Årnes*

**Author Index** ..... 159

# Evaluation of Biometric Systems: An SVM-Based Quality Index

Mohamad El-Abed, Romain Giot, Christophe Charrier and Christophe Rosenberger

GREYC laboratory

ENSICAEN – CNRS – University of Caen – Basse-Normandie

14000 Caen, France

{mohamad.elabed, romain.giot, christophe.charrier, christophe.rosenberger}@greyc.ensicaen.fr

## Abstract

Biometric quality assessment is an active research field from recent years. Quality information could be used during the enrolment step and it can also be incorporated in multimodal or soft biometrics approaches. We present in this paper an SVM-based method to compute the quality of a biometric sample using two information. The first one is based on the image quality and the second is pattern-based metrics using the SIFT keypoints extracted from the image. Experimental results on three large and significant face databases show the efficiency of the proposed quality method in predicting matching performance illustrated by the Equal Error Rate (EER).

## 1 Introduction

Biometric-based authentication systems constitute one of the most promising candidate for either replacing or enhancing traditional systems based on secret such passwords and/or token such ID cards. Many biometric modalities have been proposed in the last decade [1, 2] for authentication issues. Such systems do not provide a 100% reliable answer and it is quite impossible to obtain such a response. This uncertainty is due from the variations of human characteristics (such as occlusions for iris recognition systems [3]), environment factors (i.e., variation of acquisition conditions such as illuminations for face recognition systems [4]) and cross-device matching [5]. These kinds of variations may deeply affect biometric raw data quality. Low quality samples increase the enrolment failure rate, and decrease the system performance [4]. Therefore, quality assessment is considered as a crucial factor required in both the enrolment and verification phases. Using quality information, the bad quality samples can be removed during enrolment or rejected during verification. Such information could also be used in soft biometrics or multimodal approaches [6].

In this paper, we present an SVM-based method, to assess biometric raw data, which is based on the use of image quality and pattern-based metrics. The paper is organized as follows: section two describes previous works in the state of the art concerning biometric raw data quality. In section three, we detail the developed method. Section four illustrates the experimental results on three large face databases. We conclude and give some perspectives of this work in section five.

---

*This paper was presented at the NISK-2010 conference.*

## 2 State of the art

Quality assessment is receiving more and more attentions in biometrics. Many quality algorithms have been developed mainly for the fingerprint modality [7, 8], face [4, 9], iris [3], voice [10] and signature signals [11]. These works have demonstrated that the performance of biometric systems is heavily affected by the quality of the acquired biometric data. Tabassi *et al.* present in [7] a method based on the measurement of the matching scores to assess fingerprint quality. The proposed method uses a black box composed of two modules, *feature extraction* and *neural network*, which associate image quality into five classes according to the quartiles of genuine matching score distribution. He *et al.* [12] present a hierarchical model to compute the biometric sample quality at three levels: database, class and image quality levels. The method is based on the quantiles of genuine and impostor matching score distributions. However, their model could not be used directly on a single capture (i.e., requires a pre-acquired database). Zhang *et al.* present in [4] an asymmetry-based quality assessment method of face images. The method uses SIFT descriptor for quality assessment. The presented method has shown its robustness against illumination and pose variations. Another asymmetry-based method is presented in [13]. However, asymmetry-based methods suppose the asymmetry hypothesis hence, could not be used for the others types of modalities. Other efforts [5, 14, 15] have also been focused on the incorporation of biometric quality information to multimodal fusion approaches. Fierrez-Aguilar *et al.* [14] used an adaptive score fusion scheme based on automatic quality estimation in the spatial frequency domain. Poh *et al.* present in [15] an overview of methods for quality-based multimodal biometric fusion using Bayesian networks and propose a novel fusion classifier incorporating both the quality and device information simultaneously.

The works done in quality assessment are very few in comparison to performance ones. To our knowledge, most of the existing quality algorithms are modality and matcher dependent. The others, based on the genuine and impostor matching score distributions, could not be used directly on a single capture (i.e., they require a large number of captures for the same person in order to constitute its genuine score distribution). Therefore, the main contribution of this paper is the definition of a method which can be considered as independent from the used matching system. It detects in a reasonable accuracy three types of real alterations that may deeply affect the global performance of the most widely used matching systems. The presented method is not based on asymmetry hypothesis. Thus, it may be used for several types of modalities and can be used directly on a single capture after training the model.

## 3 Developed Method

The developed method is designed to assess biometric raw data. The method principle is illustrated in Fig.1: it is based on the use of image quality and pattern-based metrics. The method constitutes a vector of both information and uses the Support Vector Machine (SVM) [16] classifier to assign a class for an image.

### No-reference image quality

The development of general-purpose no-reference approaches to image quality assessment (NR-IQA) still lags recent advances in full-reference methods. Additionally, most no-reference or blind approaches are distortion-specific, meaning they assess only a specific type of distortion assumed present in the test image (such as blockiness, blur, or



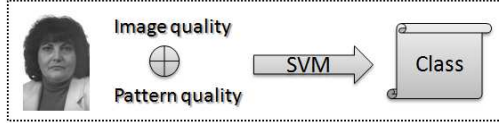


Figure 1: General scheme of the proposed method

ringing). This limits their application domain. Other approaches rely on training a machine learning algorithm. These methods however, are only as effective as the features used to train their learning machines. The used NR-IQA method in this paper is the BLIINDS index introduced by Saad *et al.* [17]. This index is based on a DCT framework entirely. This makes it computationally convenient, uses a commonly used transform, and allows a coherent framework. The BLIINDS index is defined from four features that are then pooled together: 1) a contrast feature  $v_1$ , 2) a structure distortion based feature  $v_2$  and 3)4) two anisotropic based measures  $v_3, v_4$ . Contrast is a basic perceptual attribute of an image. One may distinguish between global contrast measures and ones that are computed locally (and possibly pooled into one measure post local extraction). In resort to computing contrast based on the  $k^{th}$  local DCT patch as follows:

$$c^k(x) = \frac{1}{N} \sum_{i=1}^N \frac{x_{AC}^i}{x_{DC}} \quad (1)$$

where  $N$  is the patch size,  $x_{DC}$  represents the DC coefficient and the set  $\{x_{AC}^i | i = 1 : N\}$  represents the AC coefficients. Then, the local contrast scores from all patches of the image are then pooled together by averaging the computed values to obtain a global image contrast value  $v_1$ :

$$v_1 = \frac{1}{M} \sum_{i=1}^M c^i(x) \quad (2)$$

where  $M$  is the number of local patches. Structure features are derived locally from the local DCT frequency coefficients computed on a patch  $k$ . They are based on statistical traits of the DCT histogram for which the DC coefficient is ignored. To measure these statistical traits of the DCT histograms of the patch  $k$ , its kurtosis is computed to quantify the degree of its peakedness and tail weight:

$$\kappa^k(x_{AC}) = \frac{E(x_{AC} - \mu^4)}{\sigma^4} \quad (3)$$

where  $\mu$  is the mean of  $x_{AC}$ , and  $\sigma$  is its standard deviation. Then the resulting values for all patches are pooled together by averaging the lowest tenth percentile of the obtained values to compute the global image kurtosis value  $v_2$ . As degradation processes damage a scenes directional information, anisotropy measure, which is a directionally dependent quality of images, is computed using the Renyi Entropy on DCT image patches along four different orientations  $\theta = 0, 45, 90, 135$  in degrees. Each patch consists of the DCT coefficients of oriented pixel intensities. We discard the DC coefficient, since the focus is on directional information. Let the DCT coefficients of  $k^{th}$  patch of orientation  $\theta$  be denoted by  $P_\theta[k, j]$ , where  $j$  is the frequency index of the DCT coefficient. Each DCT patch is then subjected to a normalization of the form :

$$\tilde{P}_\theta[k, j] = \frac{P_\theta[k, j]^2}{\sum_{j=1}^N P_\theta[k, j]^2} \quad (4)$$

where  $N$  is the size of the oriented  $k^{th}$  patch. Finally the associated Renyi entropy  $R_{\theta}^k$  is computed as

$$R_{\theta}^k = \frac{1}{1-\beta} \log_2 \left( \tilde{P}_{\theta}[k, j]^{\beta} \right) \quad (5)$$

where  $\beta > 1$ . Finally the two measures of anisotropy  $v_3$  and  $v_4$  are defined as

$$v_3 = \sigma(E(R_{\theta}^k)) \text{ and } v_4 = \max(E(R_{\theta}^k)), \forall k, \forall \theta \quad (6)$$

Due to the fact that the perception of image details depends on the image resolution, the distance from the image plane to the observer, and the acuity of the observers visual system, a multiscale approach is applied to compute the final global score as:

$$\text{BLIINDS} = \prod_{i=1}^M v_1^{\alpha_1^i} v_2^{\alpha_2^i} v_3^{\alpha_3^i} v_4^{\alpha_4^i} \quad (7)$$

constraints by  $\sum_{j=1}^4 \sum_{i=1}^M \alpha_j^i = 1$  and where  $M$  represents the number of decomposition level used.

### Pattern-based quality

The pattern-based metrics used to contribute to quality assessment use the Scales Invariant Feature Transform (SIFT) descriptors described in [18]. SIFT algorithm consists of four major stages: 1) scale-space extrema detection, 2) keypoint localization, 3) orientation assignment and 4) keypoint descriptor. In the first stage, potential interest points are identified, using a difference-of-Gaussian function, that are invariant to scale and orientation. In the second stage, candidate keypoints are localized to sub-pixel accuracy and eliminated if found to be unstable. The third stage identifies the dominant orientations for each keypoint based on its local image patch. The keypoint descriptor in the final stage is created by sampling the magnitudes and orientations of the image gradients in a neighborhood of each key-point and building smoothed orientation histograms that contain the important aspect of the neighborhood. Each local descriptor is composed on a 4x4 array (histogram). To each coordinate of this array, an 8 orientation vector is associated. A 128-elements vector is then built for each keypoint. In other words, each image  $im$  is described by a set of invariant features  $X(im) = \{k_i = (s_i, sc_i, x_i, y_i) \mid i = 1 : N(im)\}$  where  $s_i$  is the 128-elements SIFT invariant descriptor computed near keypoints  $k_i$ ,  $(x_i, y_i)$  its position in the original image  $im$ ,  $sc_i$  its scale and  $N(im)$  the number of detected keypoints for image  $im$ . The features extracted are invariant to image scaling and rotation, and partially invariant to change in illumination and 3D camera viewpoint. From these features, we use the following four metrics that we considered as potentially interesting (see section 4) to contribute to quality assessment:

- keypoints: the number of keypoints detected from image  $im$ ;
- DC coefficient: DC coefficient of the matrix  $M_s$ , with  $N(im)$  rows and 128 columns, related to SIFT invariant descriptor for  $s_i$ ,  $i = 1 : N(im)$  where  $N(im)$  is the number of detected keypoints for image  $im$ ;
- scales: mean and standard deviation of scales related to the keypoints detected from image  $im$ .

## SVM classification

In order to assign a class to an image from both image quality and pattern-based information, we use the Support Vector Machine (SVM). From all existing classification schemes, a Support Vector Machine (SVM)-based technique has been selected due to high classification rates obtained in previous works [19] and to their high generalization abilities. The SVMs were developed by Vapnik et al. [16] and are based on the structural risk minimization principle from statistical learning theory. SVMs express predictions in terms of a linear combination of kernel functions centered on a subset of the training data, known as support vectors (SV). In our study, the input patterns has five dimensions (BLIINDS, number of keypoints detected, DC coefficient, mean and standard deviation of scales) and can belong to ten different classes defined as follows (see table 1):

- class 1 illustrates a reference image;
- classes 2 to 10 illustrate 3 types of alterations and 3 levels for each type (see section experimental results for details about the introduced alterations).

Table 1: SVM classes definition

| Class       | Description                 | Alteration levels        |
|-------------|-----------------------------|--------------------------|
| 1           | reference or original image |                          |
| 2, 3 and 4  | blurring alteration         | 1, 2 and 3, respectively |
| 5, 6 and 7  | gaussian noise alteration   | 1, 2 and 3, respectively |
| 8, 9 and 10 | resize alteration           | 1, 2 and 3, respectively |

Suppose we have a training set  $\{\mathbf{x}_i, \mathbf{y}_i\}$  where  $\mathbf{x}_i$  is the training pattern and  $\mathbf{y}_i$  the label. For problems with two classes, with the classes  $\mathbf{y}_i \in \{-1, 1\}$ , a support vector machine [16, 20] implements the following algorithm. First, the training points  $\{\mathbf{x}_i\}$  are projected into a space  $\mathcal{H}$  (of possibly infinite dimension) by means of a function  $\Phi(\cdot)$ . The second step is to find an optimal decision hyperplane in this space. The criterion for optimality will be defined shortly. Note that for the same training set, different transformations  $\Phi(\cdot)$  may lead to different decision functions. A transformation is achieved in an implicit manner using a kernel  $K(\cdot, \cdot)$  and consequently the decision function can be defined as :

$$f(\mathbf{x}) = \langle w, \Phi(\mathbf{x}) \rangle + b = \sum_{i=1}^{\ell} \alpha_i^* \mathbf{y}_i K(\mathbf{x}_i, \mathbf{x}) + b \quad (8)$$

with  $\alpha_i^* \in \mathbb{R}$ . The values  $w$  and  $b$  are the parameters defining the linear decision hyperplane. We use in the proposed system a linear function as kernel function. In SVMs, the optimality criterion to maximize is the margin, that is to say, the distance between the hyperplane and the nearest point  $\Phi(\mathbf{x}_i)$  of the training set. The  $\alpha_i^*$  which optimize this criterion are obtained by solving the following problem :

$$\begin{cases} \max_{\alpha_i} \sum_{i=1}^{\ell} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{\ell} \alpha_i \alpha_j \mathbf{y}_i \mathbf{y}_j K(\mathbf{x}_i, \mathbf{x}_j) \\ \text{with constraints,} \\ 0 \leq \alpha_i \leq C, \\ \sum_{i=1}^{\ell} \alpha_i \mathbf{y}_i = 0. \end{cases} \quad (9)$$

where  $C$  is a penalization coefficient for data points located in or beyond the margin and provides a compromise between their numbers and the width of the margin (for this study  $C = 1$ ). In this paper, we use the RBF kernel with  $\gamma = 1/5$ :

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (10)$$

Originally, SVMs have essentially been developed for the two classes problems. However, several approaches can be used for extending SVMs to multiclass problems. The method we use in this communication, is called *one against one*. Instead of learning  $N$  decision functions, each class is discriminated here from another one. Thus,  $\frac{N(N-1)}{2}$  decision functions are learned and each of them makes a vote for the affectation of a new point  $\mathbf{x}$ . The class of this point  $\mathbf{x}$  becomes then the majority class after the voting.

## Discussion

The goal of the proposed method is to detect, in a reasonable accuracy, some real alterations which may deeply affect the most widely used matching systems. The presented method may be considered as independent from the used matching system. An example of its practical use is illustrated in figure 2. The method predicts the quality of the input image. Then, depending from the robustness of the used matching system against the predicted alteration, the matching system qualifies the image (good, medium or bad quality).



Figure 2: An example of use of the method

## 4 Experimental results

In this section, we detail the experimental protocol and the results we obtained. We present first the protocol used for the evaluation process. Then, we illustrate the effectiveness of the extracted metrics in detecting the introduced alterations. In order to validate the presented quality method, we use the SIFT matching algorithm as a use case. Finally, we show the efficiency of the proposed quality method in predicting the performance of SIFT matching system using the Equal Error Rate (EER) metric. This error rate corresponds to the point at which the FAR and FRR cross (compromise between FAR and FRR). It is widely used to evaluate the performance of biometric systems [21].

### Protocol

In this study, we use three benchmark databases. For each database, we made three types of alterations (blurring, gaussian noise and resize alterations) and three levels for each type. The introduced alterations are commonly realistic during the acquisition of biometric data which may deeply affect the global performance of biometric systems. Finally, we have 30 databases: 3 reference databases and 27 altered databases (i.e., 9 for each reference database):

### Reference databases

- FACES94 Database [22]: This database is composed of 152 individuals and 20 samples per individual. These images have been captured in regulated illumination and the variation of expression is moderated.
- ENSIB Database [23]: It is composed of 100 individuals and 40 samples per individual. Each sample corresponds to one pose from the left one to the right.
- FERET Database [24, 25]: It is composed of 725 individuals with from 5 to 91 samples per individual (the average value is 11). Each sample corresponds to a pose angle, illumination and expression.



Samples from  
FACES94



Samples from  
ENSIB



Samples from  
FERET

### Altered databases

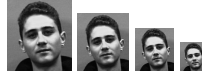
We generated 27 databases from the three reference databases: FACES94, ENSIB and FERET databases. For each database, we use three types of alterations (blurring, gaussian noise and resize) and three levels for each type. Fig. 3 shows these alterations on a sample from FACES94 database.



(a) blurring alteration



(b) gaussian noise alteration



(c) resize alteration

Figure 3: Alterations for a reference image from FACES94. From left to right, reference image then alteration level 1, 2 and 3.

### Quality metrics behavior with alterations

In this section, we show the robustness of the used metrics in detecting alterations presented in the previous section. To do so, we use the Pearson's correlation coefficient between two variables defined in Eq. 11. It is defined as the covariance of the two variables divided by the product of their standard deviation.

$$Pearson(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} \quad (11)$$

In order to compute the correlation of the used metrics with the three types of alterations, we define for each type of alteration and for each metric  $p$  the variables as follows:

- $X_p = \{X_{pk} | k = 1 : 4\}$  where  $X_{p1}$  is the set of values of metric  $p$  for the reference databases images,  $(X_{p2}, X_{p3}, X_{p4})$  are the sets of values of metric  $p$  for the altered databases level 1, 2 and 3, respectively;
- Alteration levels are represented by the variable  $Y$  (1: for the reference databases, 2, 3 and 4: for the altered databases level 1, 2 and 3). More precisely,  $Y = \{y_k | y_k = 1 \text{ for } k = 1 : N, y_k = 2 \text{ for } k = N + 1 : 2N, y_k = 3 \text{ for } k = 2N + 1 : 3N \text{ and } y_k = 4 \text{ for } k = 3N + 1 : 4N\}$  where  $N$  is the size of the 3 reference databases.

Table 2 shows that our four pattern-based metrics (keypoints, DC coefficient, mean and standard deviation of scales) are pertinent in detecting the three types of alterations: blurring, gaussian noise and resize alterations. The image quality metric BLIINDS has shown to be efficient (with a correlation coefficient more than 0.8) in detecting blurring and gaussian noise alterations. For the resize alteration, BLIINDS has not shown to be efficient which is not a surprising result since resize alteration do not affect image quality.

Table 2: Pearson correlation coefficients between the proposed metrics and the three alterations

| Metrics               | $\rho_{blurring}(X, Y)$ | $\rho_{gaussian\ noise}(X, Y)$ | $\rho_{resize}(X, Y)$ |
|-----------------------|-------------------------|--------------------------------|-----------------------|
| <i>keypoints</i>      | 0.617                   | 0.446                          | -0.581                |
| <i>DC coefficient</i> | -0.651                  | 0.634                          | -0.588                |
| <i>mean scale</i>     | 0.821                   | -0.543                         | -0.423                |
| <i>std scale</i>      | 0.399                   | -0.342                         | -0.543                |
| <i>BLIINDS</i>        | 0.802                   | -0.821                         | -0.17                 |

## Biometric matching algorithm

The matching algorithm used in this paper is SIFT presented in [18]. The matching similarity principle used is described in our previous work in [26]. Each image  $im$  is described by a set of invariant features  $X(im) = \{k_i = (s_i, x_i, y_i) | i = 1 : N(im)\}$  where  $s_i$  is the 128-elements SIFT invariant descriptor computed near keypoints  $k_i$ ,  $(x_i, y_i)$  its position in the original image  $im$  and  $N(im)$  the number of detected keypoints for image  $im$ .

The verification between two images  $im_1$  and  $im_2$  corresponds to compute a similarity between two sets of features  $X(im_1)$  and  $X(im_2)$ . We thus use the following matching method which is a modified version of a decision criterion first proposed by Lowe [18]. Given two keypoints  $x \in X(im_1)$  and  $y \in X(im_2)$ , we say that  $x$  is associated to  $y$  iff:

$$d(x, y) = \min_{\{z \in X(im_2)\}} d(x, z) \text{ and } d(x, y) \leq C d(x, y') \quad (12)$$

where  $C$  is an arbitrary threshold,  $d(\cdot, \cdot)$  denotes the Euclidean distance between the SIFT descriptors and  $y'$  denotes any point of  $X(im_2)$  whose distance to  $x$  is minimal but greater than  $d(x, y)$ :

$$d(x, y') = \min_{\{z \in X(im_2), d(x, z) > d(x, y)\}} d(x, z) \quad (13)$$

In other words,  $x$  is associated to  $y$  if  $y$  is the closest point from  $x$  in  $X(im_2)$  according to the Euclidean distance between SIFT descriptors and if the second smallest value of this distance  $d(x, y')$  is significantly greater than  $d(x, y)$ . The significance of the necessary gap between  $d(x, y)$  and  $d(x, y')$  is encoded by the constant  $C$ . Then, we consider that keypoint  $x$  is matched to  $y$  iff  $x$  is associated to  $y$  and  $y$  is associated to  $x$ . Figure 4 illustrates an example of matching results resulting from an impostor and a genuine comparison. The number of associations is used here as a similarity measure.

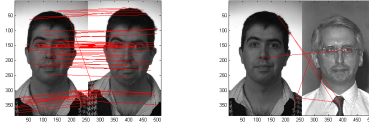


Figure 4: Example of matching results resulting from a genuine (on the left) and an impostor comparisons (on the right)

### Quality sets definition

The proposed SVM quality method predicts a class for an image. In order to quantify the efficiency of this method, we need first to define the quality sets for the used matching system. Depending from the used matching system, some alterations may not have an impact on it and others may deeply affect its performance. Therefore, we have tested the robustness of SIFT matching against the introduced alterations. We use the first image for the enrolment and the others for the test. We illustrate in table 3 the impact of adding altered images to the reference database images. We can clearly see that all the introduced alterations have shown an impact on system performance. Therefore, we have defined for this study the four quality sets as illustrated in table 4. Certainly, we may have choose another definition of sets that will more penalize the resize alteration (i.e., by putting label 9 in quality set IV) since it has the most impact on SIFT matching.

Table 3: Effect of alterations on each database: values of EERs (%).

|                           | FACES94 | ENSIB | FERET |
|---------------------------|---------|-------|-------|
| original db               | 0.29    | 10.41 | 26    |
| blurring alteration       | 1.65    | 15.54 | 29.61 |
| gaussian noise alteration | 0.76    | 13.3  | 27.81 |
| resize alteration         | 9.74    | 18.67 | 34.17 |

Table 4: Category of quality

| Quality set | Predicted quality class by SVM (refer to table 1 for class definition) | Description |
|-------------|--|-------------|
| I           | I  | good        |
| II          | 2, 5 and 8   | fair        |
| III         | 3, 6 and 9   | poor        |
| IV          | 4, 7 and 10  | very poor   |

### Assessing the SVM-Based quality method

According to Gother *et al.* [27], biometric quality measurement algorithms should predict the matching performance. That is, a quality measurement algorithm takes a biometric

raw data, and produces a class or a scalar related to error rates associated to that sample. Therefore, we illustrate the efficiency of the proposed SVM-based method using the quantification of the global performance of biometric systems illustrated by the Equal Error Rate (EER). EER is defined as the rate when both False Acceptance Rate (FAR) and False Reject Rate (FRR) are equal: The lower EER, the more accurate the system is considered to be [28]. In order to validate the proposed SVM-based quality method, we have proceed, for each benchmark database, as follows:

- Learning the SVM model: we learn the SVM model using the images from the 10 databases (i.e., the reference database and its 9 altered ones). In order to avoid the skew due to the fact of using the same data for learning and validation, we have operated a 10 folds cross-validation, 10 times with shuffling the data each time in order to have different partitions and averaging the results. The mean accuracy (standard deviation) for the 10 produced SVM models, for each database, are 83.13% (0.0075), 91.6% (0.0049) and 80.45% (0.0043) for FACES94, ENSIB and FERET databases, respectively;
- Quality sets definition: we have considered four quality sets as defined in table 4;
- EER value for each quality set: in order to quantify the effectiveness of our quality method in predicting system performance, we have put each image to a quality set, according to table 4, using its predicted label by our method. Then, we have calculated the EER value for each quality set. The effectiveness of the method is quantified by how well our quality method could predict system performance among the defined quality sets.

Table 5 illustrates the EER values for each quality set among the three used benchmark databases (FACES94, ENSIB and FERET). The proposed SVM-based quality method has shown to be efficient in predicting SIFT matching performance. The increase of EER values as quality degrades indicate vulnerability of low quality face images. From table 5 we can also conclude:

- for FACES94 and ENSIB databases, there were no significant differences between the EERs of the reference database (i.e., which is considered containing images of good quality) and the set predicted as good quality by our method.
- for FERET database, there was a difference of 5.88%. This variation was due to the difficulties of FERET database which contains altered images by resolution. Despite this, our method have shown to be efficient in predicting system performance.

Table 5: EER values for the reference databases and the mean of EER (standard deviation) values for each quality set among the three databases

| Database | EER values (%)     |                  |                  |                  |                    |
|----------|--------------------|------------------|------------------|------------------|--------------------|
|          | reference database | set I (good)     | set II (fair)    | set III (poor)   | set IV (very poor) |
| FACES94  | 0.29               | 0.4744 (0.0423)  | 0.6843 (0.0147)  | 1.8078 (0.0176)  | 5.7983 (0.0149)    |
| ENSIB    | 10.41              | 10.6397 (0.0343) | 13.2912 (0.0144) | 16.5495 (0.0387) | 17.787 (0.0015)    |
| FERET    | 26                 | 31.88 ( 0.0265)  | 32.23 (0.0745)   | 32.52 (0.012)    | 34.37 (0.0347)     |



## 5 Conclusion and perspectives

Quality assessment is becoming an important factor to take into account when developing and evaluating biometric systems. In this paper, we have presented an SVM-based quality method based on the use of image quality and pattern-based metrics. The developed method has shown its efficiency in predicting system's performance illustrated by the Equal Error Rate (EER). The method can be considered as independent from the used matching system. For example, if the used matching system tolerate resize alteration, we could modify the quality sets definition by putting their labels (i.e., labels 8, 9 and 10) in quality set I (i.e., we consider that the altered image, by resize alteration, may be considered as of good quality for this type of matching system). On the other hand, the method is not based on asymmetry hypothesis. Thus, it may be used for several types of modalities (such as hand veins, iris, fingerprint ...). The method also could be applied directly on a single capture after training the model.

For the perspectives, we would like to quantify the efficiency of the method on other types of modalities (such as fingerprint) to test if the method can be considered as modality-independent. We also intend to compare it with others quality algorithms existing in the state of the art for iris and fingerprint modalities.

## References

- [1] Y. Chen and A. Jain, "Beyond minutiae: A fingerprint individuality model with pattern, ridge and pore features," in *ICB09*, 2009.
- [2] R. Giot, M. El-Abed, and C. Rosenberger, "Keystroke dynamics with low constraints svm based passphrase enrollment," in *IEEE Third International Conference on Biometrics : Theory, Applications and Systems (BTAS)*, 2009.
- [3] E. Krichen, S. Garcia Salicetti, and B. Dorizzi, "A new probabilistic iris quality measure for comprehensive noise detection," in *BTAS07*, 2007.
- [4] G. Zhang and Y. Wang, "Asymmetry-based quality assessment of face images," in *ISVC '09: Proceedings of the 5th International Symposium on Advances in Visual Computing*, 2009.
- [5] N. Poh, J. Kittler, and T. Bourlai, "Quality-based score normalization with device qualitative information for multimodal biometric fusion," *SMC-A*, 2010.
- [6] K. Kryszczuk, J. Richiardi, and A. Drygajlo, "Impact of combining quality measures on biometric sample matching," in *BTAS'09: Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems*, 2009.
- [7] E. Tabassi and C. Wilson, "A novel approach to fingerprint image quality," in *ICIP05*, 2005.
- [8] M. Vatsa, R. Singh, A. Noore, and S. Singh, "Quality induced fingerprint identification using extended feature set," in *BTAS'08: Proceedings of the 2nd IEEE international conference on Biometrics: Theory, applications and systems*, 2008.
- [9] J. Sang, Z. Lei, and S. Z. Li, "Face image quality evaluation for ISO/IEC standards 19794-5 and 29794-5," in *ICB '09: Proceedings of the Third International Conference on Advances in Biometrics*, 2009.
- [10] A. Harriero, D. Ramos, J. Gonzalez-Rodriguez, and J. Fierrez, "Analysis of the utility of classical and novel speech quality measures for speaker verification," in *ICB '09: Proceedings of the Third International Conference on Advances in Biometrics*, 2009.

- [11] S. Muller and O. Henniger, "Evaluating the biometric sample quality of handwritten signatures," in *ICB07*, 2007.
- [12] Q. He, Z. Sun, T. Tan, and Y. Zou, "A hierarchical model for the evaluation of biometric sample quality," in *ICPR08*, 2008.
- [13] X. Gao, S. Li, R. Liu, and P. Zhang, "Standardization of face image sample quality," in *ICB07*, 2007.
- [14] J. Fierrez Aguilar, Y. Chen, J. Ortega Garcia, and A. Jain, "Incorporating image quality in multi-algorithm fingerprint verification," in *ICB06*, 2006.
- [15] N. Poh, T. Bourlai, J. Kittler, L. Allano, F. Alonso-Fernandez, O. Ambekar, J. Baker, B. Dorizzi, O. Fatukasi, J. Fierrez, H. Ganster, J. Ortega-Garcia, D. Maurer, A. A. Salah, T. Scheidat, and C. Vielhauer, "Benchmarking quality-dependent and cost-sensitive score-level multimodal biometric fusion algorithms," *Trans. Info. For. Sec.*, 2009.
- [16] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 1995.
- [17] M. Saad, A. C. Bovik, and C. Charrier, "A DCT statistics-based blind image quality index," *IEEE Signal Processing Letters*, 2010.
- [18] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, 2004.
- [19] G. Lebrun, C. Charrier, O. Lezoray, C. Meurie, and H. Cardot, "Fast pixel classification by svm using vector quantization, tabu search and hybrid color space," in *the 11th International Conference on CAIP*, 2005.
- [20] B. Schölkopf and A. J. Smola, *Learning with kernels : support vector machines, regularization, optimization, and beyond*, 2002.
- [21] "Information technology biometric performance testing and reporting," ISO/IEC 19795-1, Tech. Rep., 2006.
- [22] U. of Essex, "Faces94 database, face recognition data," 1994. [Online]. Available: <http://cswww.essex.ac.uk/mv/allfaces/faces94.html>
- [23] B. Hemery, C. Rosenberger, and H. Laurent, "The ENSIB database : a benchmark for face recognition," in *International Symposium on Signal Processing and its Applications (ISSPA), special session "Performance Evaluation and Benchmarking of Image and Video Processing"*, 2007.
- [24] P. Phillips, H. Moon, S. Rizvi, and P. Rauss, "The FERET evaluation methodology for face recognition algorithms," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2000.
- [25] P. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The FERET database and evaluation procedure for face recognition algorithms," *Journal of Image and Vision Computing*, 1998.
- [26] P.-O. Ladoux, C. Rosenberger, and B. Dorizzi, "Hand vein verification system based on sift matching," in *the 3rd IAPR/IEEE International Conference on Biometrics (ICB)*, 2009.
- [27] P. Grother and E. Tabassi, "Performance of biometric quality measures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007.
- [28] A. J. Mansfield and J. Wayman, "Best practices in testing and reporting performance of biometric devices," BWG, Tech. Rep., 2002.