MEASURING BIOMETRIC SAMPLE QUALITY IN TERMS OF BIOMETRIC INFORMATION

Richard Youmaran and Andy Adler

University of Ottawa,
School of Information Technology and Engineering,
Ontario, Canada

ABSTRACT

This paper develops a new approach to understand and measure variations in biometric sample quality. We begin with the intuition that degradations to a biometric sample will reduce the amount of identifiable information available. In order to measure the amount of identifiable information, we define biometric information as the decrease in uncertainty about the identity of a person due to a set of biometric measurements. We then show that the biometric information for a person may be calculated by the relative entropy $D(p\|q)$ between the population feature distribution q and the person's feature distribution p. The biometric information for a system is the mean $D(p\|q)$ for all persons in the population. In order to practically measure $D(p\|q)$ with limited data samples, we introduce an algorithm which regularizes a Gaussian model of the feature covariances. An example of this method is shown for PCA, Fisher linear discriminant (FLD) and ICA based face recognition, with biometric information calculated to be 45.0 bits (PCA), 37.0 bits (FLD), 39.0 bits (ICA) and 55.6 bits (fusion of PCA and FLD features). Based on this definition of biometric information, we simulate degradations of biometric images and calculate the resulting decrease in biometric information. Results show a quasi-linear decrease for small levels of blur with an asymptotic behavior at larger blur.

1. INTRODUCTION

Biometric sample quality is a measure of the usefulness of a biometric image [9]. One recent development is the significant level of interest in standards for measurement of biometric quality. For example, ISO has recently established a biometric sample quality draft standard [9]. According to [9], biometric sample quality may be considered from the point of view of character (inherent features), fidelity (accuracy of features), or utility (predicted biometrics performance). A general consensus has developed that the most important measure of a quality metric is its utility – images evaluated as higher quality must be those that result in better identification of individuals, as measured by an increased separation of genuine and impostor match score distributions. The nature of biometric sample fidelity has seen little investigation, although for specific biometric modalities, algorithms to measure biometric quality have been proposed. For example, the NFIQ algorithm [12] is a widely used measure for fingerprint image quality.

One current difficulty is that there is no consensus as to what a measure of biometric sample fidelity should give. In this paper, we propose a new approach to measure this quantity, based on an information theoretic framework. We begin with the intuitive observation that a high quality biometric image is more useful to identify the individual than a low quality image. This suggests that the quantity of identifiable information decreases with a reduction in quality. Given a way to measure the decrease in information caused by a given image degradation, one can measure the associated decrease in biometric information.

Measuring biometric information content is related to many issues in biometric technology. For example, one of the most common biometric questions is that of uniqueness, eg. to what extent are fingerprints unique? From the point of view of identifiability, one may be interested in how much identifiable information is available from a given technology, such as video surveillance. In the context of biometric fusion [11] one would like to be able to quantify the biometric information in each system individually, and the potential gain from fusing the systems. Additionally, such a measure is relevant to biometric cryptosystems and privacy measures. Several authors have presented approaches relevant to this question ([1, 6, 11, 14]). In this paper we elaborate an approach to address this question based on definitions from information theory [2]. We define the term "biometric information" as follows:

biometric information (BI): the decrease in uncertainty about the identity of a person due to a set of biometric measurements.

In order to interpret this definition, we refer to two instants: 1) before a biometric measurement, t_0, at which time we only know a person p is part of a population q, which may be the whole planet; and 2) after receiving a set of measurements, t_1, we have more information and less uncertainty about the person's identity.

Based on these measures, we then define the information loss due to a degradation in image quality, as the relative change

in BI. The degradation process is modeled by H, which maps the original high quality images F to G. For the case with no degradation, we measure the image from a person p_F as part of a population q_F, while in the presence of degradation H, we obtain a person's image p_G as part of population q_G.
This paper then develops a mathematical framework to measure biometric information for a given system and set of biometric features. In practice, there are limited numbers of samples of each person, which makes our measure ill-conditioned. In order to address this issue, we develop a stable algorithm based on a distribution modeling and regularization. We then use this algorithm to analyze the biometric information content of three different face recognition algorithms, and to measure biometric quality loss due to a degradation.

2. METHODS

In this section we develop a approach to measure the effect of an image degradation model on biometric image quality. First, we develop an algorithm to calculate biometric information based on a set of features, using the relative entropy measure [5]. We explain our method in the following steps: A) measure requirements, B) relative entropy of biometric features, C) Gaussian models for biometric features and relative entropy calculations, D) regularization methods for degenerate features, E) regularization methods for insufficient data, and F) information loss due to degradation.

2.1. Measure requirements

In order to elaborate the requirements that a good measure of biometric information measure must have, we consider system that measures height and weight. These values differ within the global population, but also vary for a given individual, both due to variations in the features themselves and to measurement inaccuracies. We now wish to consider the properties a measure of biometric information should have:

1. If an intra-person distribution p is exactly equal to the inter-person q distribution, then there is no information to distinguish a person, and biometric information is zero.

2. As the feature measurement becomes more accurate (less variability), then it is easier to distinguish someone in the population and the biometric information increases.

3. If a person has unusual feature values (i.e. far from the population mean), they become more distinguishable, and their biometric information will be larger.

4. The biometric information of uncorrelated features should be the sum of the biometric information of each individual feature.

5. Features that are unrelated to identity should not increase biometric information. For example, if a biometric system accurately measured the direction a person was facing, information on identity would be unchanged.

6. Correlated features such as height and weight are less informative. In an extreme example consider the height in inches and in cm. Clearly, these two features are no more informative than a single value.

Based on this definition, the most appropriate information theoretic measure for the biometric information is the relative entropy ($D(p\|q)$) [5] between the intra- ($q(\mathbf{f})$) and inter-person ($p(\mathbf{f})$) biometric feature distributions. $D(p\|q)$, or the Kullback-Leibler distance, is defined to be the "extra bits" of information needed to represent $p(\mathbf{f})$ with respect to $q(\mathbf{f})$. $D(p\|q)$ is defined to be

$$D(p\|q) = \int_{\mathbf{f}} p(\mathbf{f}) log_2 \frac{p(\mathbf{f})}{q(\mathbf{f})} d\mathbf{f} \qquad (1)$$

where the integral is over all N_F feature dimensions, $\mathbf{f}$. $p(\mathbf{f})$ is the probability mass function or distribution of features of an individual and $q(\mathbf{f})$ is the overall population distribution. A comment on notation: we use p to refer to both an individual, and the distribution of the person's features, while q represents the population and the distribution of its features.

This measure can be motivated as follows: the relative entropy, $D(p\|q)$, is the extra information required to describe a distribution $p(\mathbf{f})$ based on an assumed distribution $q(\mathbf{f})$ [5]. $D(p\|q)$ differs from the entropy, $H(p)$, which is the information required, on average, to describe features $\mathbf{f}$ distributed as $p(\mathbf{f})$. H is not in itself an appropriate measure for biometric information, since it does not account the extent to which each feature can identify a person p in a population q. An example of a feature unrelated to identity is the direction a person is facing. Measuring this quantity will increase H of a feature set, but not increase its ability to identify a person. The measure $D(p\|q)$ corresponds to the requirements: given a knowledge of the population feature distribution q, the information in a biometric feature set allows us to describe a particular person p.

2.2. Distribution modeling

In a generic biometric system, an image $\mathbf{F}$ is acquired, from which N_f features are measured, to create a biometric feature vector $\mathbf{f}$ for each person. For person p, we have N_p samples, while we have N_q samples for the population. For convenience of notation, we sort p's measurements to be the first grouping of the population. Defining $\mathbf{f}$ as an instance of random variable F, we calculate the population feature mean

$$\mu_q = \underset{q}{E}[F] \qquad (2)$$

where the feature mean of person p, μ_p, is defined analogously, replacing q by p. The population feature covariance is

$$\mathbf{\Sigma}_q = \underset{q}{E}\left[(F-\mu_q)^t(F-\mu_q)\right]. \quad (3)$$

The individual's feature covariance, $\mathbf{\Sigma}_p$, is again defined analogously. Features are calculated from a set of N_q images using three different feature extraction methods: Principle Component Analysis (PCA, also referred to as Eigenface features) [7] [13], Fisher linear discriminant (FLD) [10] and Independent Component Analysis (ICA) [8]. μ_p and μ_q are $N_f \times 1$ vectors of the population and individual mean distributions, while $\mathbf{\Sigma}_p$ and $\mathbf{\Sigma}_q$ are $N_f \times N_f$ matrices of the individual and population covariance matrices.

One important general difficulty with direct information theoretic measures is that of data availability. Distributions are difficult to estimate accurately, especially at the tails; and yet $log_2\left(p(\mathbf{f})/q(\mathbf{f})\right)$ will give large absolute values for small $p(\mathbf{f})$ or $q(\mathbf{f})$. Instead, it is typical to fit data to a model with a small number of parameters. The Gaussian distribution is the most common model; it is often a good reflection of the real world distributions, and is analytically solvable in entropy integrals. Another important property of the Gaussian is that it gives the maximum entropy for a given standard deviation, allowing such models to be used to give an upper bound to entropy values. Based on a Gaussian model for p and q, $D(p\|q)$can be written as:

$$D(p\|q) \quad = \quad k\left(\alpha + trace\left((\mathbf{\Sigma}_p+\mathbf{T})\mathbf{\Sigma}_q{}^{-1}-\mathbf{I}\right)\right) \quad (4)$$

where $\alpha = ln\frac{|\mathbf{\Sigma}_q|}{|\mathbf{\Sigma}_p|}$, $\mathbf{T} = (\mu_p-\mu_q)^t(\mu_p-\mu_q)$ and $k = log_2\sqrt{e}$.

This expression calculates the relative entropy in bits for Gaussian distributions $p(\mathbf{f})$ and $q(\mathbf{f})$. This expression corresponds to most of the desired requirements for a biometric information measure introduced in the previous section:

1. If person's feature distribution matches the population, $p = q$; this yields $D(p\|q) = 0$, as required.
2. As feature measurements improve, the covariance values, $\mathbf{\Sigma}_p$, will decrease, resulting in a reduction in $|\mathbf{\Sigma}_p|$, and an increase in $D(p\|q)$.
3. If a person has feature values far from the population mean, $\mathbf{T}$ will be larger, resulting in a larger value of $D(p\|q)$.
4. Combinations of uncorrelated feature vectors yield the sum of the individual $D(p\|q)$ measures. Thus, for uncorrelated features f_1 and f_2, where $\{f_1, f_2\}$ represents concatenation of the feature vectors, $D(p(f_1)\|q(f_1))+D(p(f_2)\|q(f_2)) = D(p(\{f_1,f_2\})\|q(\{f_1,f_2\}))$
5. Addition of features uncorrelated to identity will not change $D(p\|q)$. Such a feature will have an identical distribution in p and q. If U is the set of such uncorrelated features, $[\mathbf{\Sigma}_p]_{ij} = [\mathbf{\Sigma}_q]_{ij} = 0$ for i or $j \in U$, and $i \neq j$, while $[\mathbf{\Sigma}_p]_{ii} = [\mathbf{\Sigma}_q]_{ii}$ and $[\mu_q]_i = [\mu_p]_i$. Under these conditions, $D(p\|q)$ will be identical to its value when excluding the features in U. One way to understand this criterion is that if the distributions for q and p differ for features in U, then those features can be used as a biometric to help identify a person.
6. Correlated features are less informative than uncorrelated ones. Such features will increase the condition number (and thus reduce the determinant) of both $\mathbf{\Sigma}_p$ and $\mathbf{\Sigma}_q$. This will decrease the accuracy of the measure $D(p\|q)$. In the extreme case of perfectly correlated features, $\mathbf{\Sigma}_p$ becomes singular with a zero determinant and $D(p\|q)$ is undefined. Thus, our measure is inadequate in this case. In the next section, we develop an algorithm to deal with this effect.

2.3. Regularization for degenerate features

In order to guard against numerical instability in our measures, we wish to extract a mutually independent set of N_w "important" features ($N_w \leq N_f$). To do this, we use the principal component analysis (PCA) [7] to generate a mapping ($\mathbf{U}^t : F \rightarrow W$), from the original biometric features F ($N_f \times 1$) to a new feature space W of size $N_w \times 1$. The PCA is calculated from a Singular Value Decomposition (SVD) of the feature covariance matrix, such that

$$\mathbf{U}\mathbf{S}_q\mathbf{U}^t = svd(cov(F)) = svd(\mathbf{\Sigma}_q) \quad (5)$$

Since $\mathbf{\Sigma}_q$ is positive definite, $\mathbf{U}$ is orthonormal and $\mathbf{S}_q$ is diagonal. We choose to perform the PCA on the population distribution q, rather than p, since q is based on far more data, and is therefore likely to be a more reliable estimate. The values of $\mathbf{S}_q$ indicate the significance of each feature in PCA space. A feature j, with small $[\mathbf{S}_q]_{j,j}$ will have very little effect on the overall biometric information. We use this analysis in order to regularize $\mathbf{\Sigma}_q$, and to reject degenerate features by truncating the SVD. We select a truncation threshold of j where $[\mathbf{S}_q]_{j,j} < 10^{-10}[\mathbf{S}_q]_{1,1}$. Based on this threshold, $\mathbf{S}_q$ is truncated to be $N_w \times N_w$, and $\mathbf{U}$ is truncated to $N_f \times N_w$. Using the basis $\mathbf{U}$ calculated from the population, we decompose the individual's covariance into feature space Y:

$$\mathbf{S}_p = \mathbf{U}^t\mathbf{\Sigma}_p\mathbf{U} \quad (6)$$

where $\mathbf{S}_p$ is not necessarily a diagonal matrix. However, since p and q describe somewhat similar data, we expect $\mathbf{S}_p$ to have a strong diagonal component. Note that the PCA analysis used here is not the same as that for eigenface features [13].

Based on this regularization scheme, (4) may be rewritten in the PCA space as:

$$D(p\|q) = k\left(\beta + trace\ \mathbf{U}\left((\mathbf{S}_p+\mathbf{S}_t)\mathbf{S}_q{}^{-1}-\mathbf{I}\right)\mathbf{U}^t\right) \quad (7)$$

where $\beta = ln\frac{|\mathbf{S}_q|}{|\mathbf{S}_p|}$ and $\mathbf{S}_t = \mathbf{U}^t\mathrm{T}\mathbf{U}$

2.4. Regularization for insufficient data

The expression developed in the previous section solves the problem of ill-poseness of $\mathbf{\Sigma}_q$. However, $\mathbf{\Sigma}_p$ may still be singular in the common circumstance in which only a small number of samples of each individual are available. Given N_p images of an individual from which G features are calculated, $\mathbf{\Sigma}_p$ will be singular if $G \geq N_p$, which will result in $D(p\|q)$ diverging to ∞. In practice, this is a common occurrence, since most biometric systems calculate many hundreds of features, and there are only rarely more then ten of samples of each person. In order to address this issue, we develop an estimate which may act as a lower bound using the following assumptions:

1. Estimates of feature variances are valid $[\mathbf{S}_p]_{i,i}$ for all i.
2. Estimates of feature covariances $[\mathbf{S}_p]_{i,j}$ for $i \neq j$ are only valid for the most important L features, where $L < N_p$.

Features which are not considered valid based on these assumptions, are set to zero by multiplying $\mathbf{S}_q$ by a mask $\mathbf{M}$, where

$$\mathbf{M}_{i,j} = \begin{cases} 1, & \text{if } i = j \text{ or } (i < L \text{ and } j < L); \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

Using (8), $[\mathbf{S}_p]_{i,j} = (\mathbf{M}_{i,j})[\mathbf{U}^t \mathbf{\Sigma}_p \mathbf{U}]_{i,j}$. This expression regularizes the intra-person covariance, $\mathbf{\Sigma}_p$, and assures that $D(p\|q)$ does not diverge. To clarify the effect of this regularization on $D(p\|q)$, we note that intra-feature covariances will decrease $|\mathbf{\Sigma}_p|$ toward zero, leading a differential entropy estimate diverging to ∞. We thus consider this regularization strategy to generate a lower bound on the biometric information. The selection of L is a compromise between using all available measurements (by using large L) and avoiding numerical instability when $\mathbf{S}_p$ is close to singular (by using small L).

2.5. Average information of a system

This section has developed a measure of biometric information content of a biometric feature representation of a single individual with respect to the feature distribution of the population. The biometric information will vary between people; those with feature values further from the mean have larger biometric information. Using this approach, the biometric information content of a biometric system is calculated as the average information across all people in the system at a specific L.

2.6. Information loss due to degradation

In this section, we explore the effect of image degradation and the resulting decrease in biometric quality on the relative entropy measure. Intuitively, it is expected that image degradation changes the intra and inter person distribution of the face features resulting in a loss of biometric information. In general, image degradation is a non-linear process; however, in this paper we use a linear degradation model to explore its effect. Equation(9) represents the blur degradation model used to generate degraded features where h is a space invariant Gaussian operator of size $n \times n$ and $\sigma = 3$, $\mathbf{F}$ is the original image and $\mathbf{G}$ is the resulting degraded image.

$$G(x,y) = \sum_{\alpha} \sum_{\beta} F(\alpha,\beta) h(x-\alpha, y-\beta) \quad (9)$$

Features, $\mathbf{g}$, are then extracted from the degraded images $\mathbf{G}$ using three feature extraction methods given. We then compute the biometric information for the non-degraded distributions $(D(p(f)||q(f)))$ and for the degraded distributions $(D(p(g)||q(g)))$ using equation (7). Here $D(p(f)||q(f))$ represents the relative entropy between the individual and population distribution prior to degradation while $D(p(g)||q(g))$ is the relative entropy measure between the degraded individual and population distributions, respectively. From this, we calculate the normalized mean square distance characterizing the loss of information caused by the degradation model on the underlying features as:

$$\Delta BI = \frac{1}{N_{\mathbf{f}}} \sum_{i=1}^{N_{\mathbf{f}}} \frac{|D(p(\mathbf{f}_i)||q(\mathbf{f}_i)) - D(p(\mathbf{g}_i)||q(\mathbf{g}_i))|^2}{\sigma^2_{D_{\mathbf{f}}}} \quad (10)$$

where $\sigma^2_{D_f}$ is the variance of $D(p(\mathbf{f}_i)||q(\mathbf{f}_i)$. ΔBI measures the relative distance offset between the original and degraded distributions. ΔBI is a unitless measure, and may be interpreted as the fractional loss in BI due to a given image degradation.

In order to motivate this calculation, we initially considered calculating $D(p(\mathbf{g})\|q(\mathbf{f}))$ as a function of degradation. Surprisingly, this measure increases with decreasing quality. The reason is that a single person p is considered to have degraded images in a population q of high quality images. The algorithm seems to be saying: "Aha! I can recognize p. He always has a blurry face!". Therefore, it is necessary to compare a degraded person's image to the degraded population $D(p(\mathbf{g})\|q(\mathbf{g})$ in order to compensate for this effect.

3. FACE RECOGNITION

Information in a feature representation of faces is calculated using our described method for different individuals. Using the Aberdeen face database [4], we chose 18 frontal images of 16 persons, from which we calculate the PCA (eigenface) features using the algorithm of [7], the FLD (fisherfaces) using the method described in [3] and the ICA face features components using the FastIca algorithm [8]. For PCA, FLD and ICA feature decompositions, 288 independent vectors were computed, and the most significant 100 features used for subsequent analysis.

3.1. Biometric information calculations

After fitting the distributions of $p(\mathbf{f})$ and $q(\mathbf{f})$ to a Gaussian model, we initially analyze the biometric information in each PCA, FLD and ICA feature separately. The PCA and FLD relative entropy measures as function of a feature number are shown in Fig. 1, and show a gradual decrease from an initial peak at feature 2. The form of the curve can be understood from the nature of the PCA decomposition, which tends to place higher frequency details in higher number features. Since FLD features are calculated using PCA, these tend to contain similar amount of information. Also, since noise tends to increase with frequency, the biometric information in these higher numbered PCA features will be less.

In order to calculate $D(p\|q)$ for all features, we are limited by the available information. Since $N_p = 18$ images are used to calculate the covariances, attempts to calculate $D(p\|q)$ for more than 17 features will fail, because $\mathbf{\Sigma}_p$ is singular. This effect is seen in the condition number (ratio of the largest to the smallest singular value) which was 4.82×10^3 for $\mathbf{S}_q$ and 1.32×10^{20} for $\mathbf{S}_p$. The relatively small condition number of $\mathbf{S}_q$ indicates that no features are degenerate for PCA, FLD and ICA face recognition features. However, $\mathbf{S}_p$ is severely ill-conditioned. To overcome this ill-conditioning, we introduced a regularization scheme based on a mask (equation 8) with a cut-off point L. This scheme is motivated by the diagonal structure of $\mathbf{S}_p$. To ensure convergence, the mask size L is set to a value smaller than N_p. Results for $D(p\|q)$ for PCA features for each person as a function of L are shown in Fig. 2 for N_p = 8, 12 and 18. In these curves, we observe a "hockey stick" shape. The relative entropy measure remains stable when $L < N_p$, but if $L \geq N_p$, we observe a dramatic increase in $D(p\|q)$ as the algorithm approaches a singularity of $\mathbf{\Sigma}_p$ and the ill-conditioning of $\mathbf{\Sigma}_q$. In order to produce an unique and stable estimate for $D(p\|q)$, it is necessary to choose a compromise between having an under-estimated ($L \ll N_p$) or an over-estimated ($L \geq N_p$) solution. We therefore recommend choosing $L = \frac{3}{4}N_p$, since a larger value of L puts the estimate in an unstable region of Fig. 2. Using this algorithm and value of L, we calculate the overall biometric information for different face recognition algorithms. For PCA features, the average $D(p\|q)$ is 45.0 bits and 37.0 bits for FLD features. If PCA and FLD features are combined (making 200 features in all), average $D(p\|q)$ is 55.6 bits (Fig. 3).

3.2. Degraded features

Using the degradation model described by equation (9), two different sets of images (S_1 and S_2) are generated. Each set of images is composed 16 people with 18 images per individual for a total of 288. S_1 is obtained by degrading half of each individual's face using different Gaussian operators of size $n \times n$ while S_2 is a set of images obtained as a result of blurring the entire face region. An example of images in S_1 and S_2 are seen in Fig. 4.

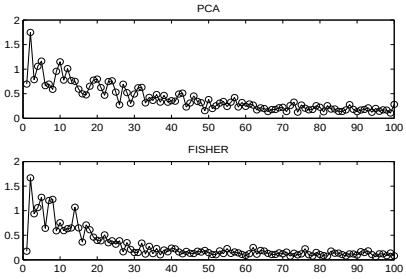

Fig. 1. Biometric information as a function of feature number (circles) for PCA (top) and FLD (bottom) feature decomposition.

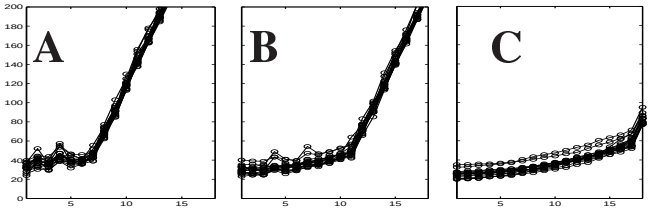

Fig. 2. $D(p\|q)$ (y-axis) vs L (x-axis) for each person. Each subfigure represents a different value of N_p: (A) 8, (B) 12 and (C) 18. The curves show that $D(p\|q)$ diverges as Σ_p becomes singular ($L \geq N_p$).

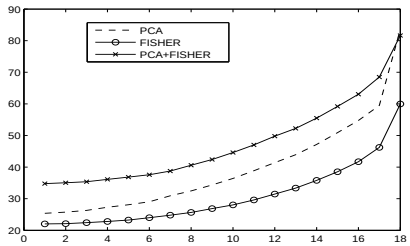

Fig. 3. Average $D(p\|q)$ (y-axis) vs L (x-axis) for $N_p = 18$. Each line represents the average of information calculated for a population of 16 individuals with 18 images each using PCA (middle), FLD (bottom) and a fusion of PCA and FLD features (top).

Using S_1 and S_2, new PCA, FLD and ICA features ($\mathbf{g}$) are extracted using the original (non-degraded) principal component vectors. From the degraded features, ΔBI is computed for the degraded individual and population distributions using equation(10). This measure represents the amount of information lost as a function of the degradation level. Fig.(5) shows ΔBI computed as function of the blur level for different images taken from S_1 and S_2. The x-axis represents 9 different levels (in increasing order) of Gaussian blur. As seen in Fig.(5), the relative information loss in an image increases with the amount of system degradation. Interestingly, ΔBI tends to reach a steady state after some level of degradation.

This suggests that some features are unaffected by the degradation process and represent a lower bound of information measure of an individual distribution. PCA features extracted using the dominant eigenvalues of the system tend to be robust against blur since they preserve valuable information at a large degradation level.

Fig. 4. Degraded image obtained by applying a Gaussian blur to (b) a section of the original image (S_1) and to (c) the entire image (S_2).

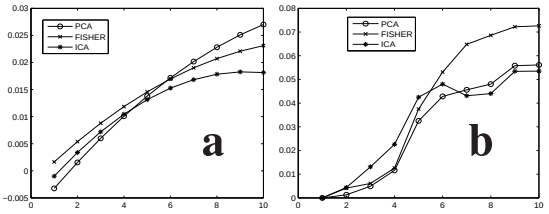

Fig. 5. Normalized mean square distance(y-axis) as a function of an increasing blur level (x-axis)for images taken from (a) S_1 and (b) S_2

4. DISCUSSION

This paper proposes a new approach to measure the changes in biometric sample quality resulting from image degradations. A definition of biometric information is introduced and an algorithm to measure it proposed, based on a set of population and individual biometric features, as measured by a biometric algorithm under test. Examples of its application were shown for different face recognition algorithms based on PCA (Eigenface), FLD (Fisherfaces) and ICA feature decompositions. Subsequently, we introduced a measure of information loss as a function of image degradation. It is shown that the normalized mean square distance measure (ΔBI), based on the relative entropy, increases with the blur level but reaches a steady state after some amount of degradation which suggests that some features are unaffected by this degradation process. In a general biometric system, the following issues associated with biometric features must be considered: 1) Feature distributions vary. Features, such as minutiae ridge angles may be uniformly distributed over 0–2π, while other features may be better modeled as Gaussian, 2) Raw sample images need to be processed by alignment and scaling before features can be measured, 3) Feature dimensionality may not be constant. While we have introduced a measure in the context of face recognition, we anticipate that such a measure may help address many questions in biometrics technology, such as: uniqueness of biometric features, inherent limits to biometric template size requirements, feasibility of biometric encryption, performance limits of biometric matchers, biometric fusion and privacy protection.

5. REFERENCES

[1] Adler, A., "Vulnerabilities in biometric encryption systems" *Audio- and Video-based Biometric Person Auth.* Tarrytown, NY, USA, Jul. 20–22, 2005

[2] Adler, A., Youmaran, R., Loyka, S., "Information content of biometric features" *Biometrics Consortium Conference* Washington, DC, USA, Sep. 19-21, 2005.

[3] Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J., "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection", *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* **19**:711 - 720, 1997.

[4] Craw, I., Costen, N.P., Kato, T., Akamatsu, S., "How should we represent faces for automatic recognition?", *IEEE Trans. Pat. Anal. Mach. Intel.* **21**725–736, 1999

[5] Cover, T.M., Thomas, J.A., *Elements of Information Theory* New York: Wiley, 1991

[6] Daugman, J., "The importance of being random: Statistical principles of iris recognition." *Pattern Recognition*, **36**:279–291, 2003.

[7] Grother,P.,"Software Tools for an Eigenface Implementation", National Institute of Standards and Technology, (2000) http://www.nist.gov/humanid/feret/+

[8] Hyvärinen, A., "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis". *IEEE Trans. Neural Networks* **10**:626–634, 1999.

[9] InterNational Committee for Information Technology Standards, Biometric Sample Quality Standard Draft (Revision 4), document number M1/06-0003, February 7, 2005.

[10] Li, S., Jain A. (Eds.). Handbook of Face Recognition. Springer, 2005.

[11] Ross, A., Jain, A., "Information Fusion in Biometrics", *Pattern Recognition Letters*, **24**:2115-2125, 2003

[12] Tabassi, E., Wilson C.R., Watson, C.I., "Fingerprint Image Quality", NISTIR 7151, August 2004

[13] Turk, M., Pentland, A., "Eigenfaces for recognition", *J. Cognitive Neuroscience*, **3**:71-86, 1991.

[14] Wayman, J.S., "The cotton ball problem", *Biometrics Conference*, Washington DC, USA, Sep. 20-22, 2004.