Information content of biometric features

Andy Adler, Richard Youmaran, Sergey Loyka
University of Ottawa, Ontario, Canada
{adler,youmaran,lokya}@site.uOttawa.ca

1. Introduction

How much information is there in a face, or a fingerprint? In this paper we elaborate an approach to address this question based on information theory reasoning. In order to motivate our approach, we initially consider the properties that such a measure should have. Consider a soft biometric system which measures height and weight; furthermore, assume all humans are uniformly and independently distributed in height between 100–200 cm and weight between 100–200 lb. If a person's features were completely stable and could be measured with infinite accuracy, people could be uniquely identified from these measurements, and the biometric features could be considered to yield infinite information. However, in reality, repeated biometric measurements give different results due to measurement inaccuracies, and to short- and long-term changes in the biometric features themselves. If this variability results in an uncertainty of $\pm$ 5 cm and ± 5 lb, one simple model would be to round each measure measure to $105, 115, ..., 195$. In this case, there are 10×10 equiprobable outcomes, and an information content of $log_2(100) = 6.6$ bits.

Such an analysis is intrinsically tied to a choice of biometric features. Thus, it does not appear possible to answer "how much information is in a fingerprint?", but only "how much information is in the position and angle data of fingerprint minutiae?". Furthermore, for many biometrics, it is not clear what the underlying features are. Face images, for example, can be described by image basis features or landmark based features. To overcome this, we may choose to calculate the information in all possible features. In the example, we may provide height in inches as well as cm; however, in this case, a good measure of information must not increase with such redundant data. Additionally, the following issues associated with biometric features must be considered:

- Feature values are correlated. In the example given, taller people tend to be heavier.
- 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.
- Raw sample images need to be processed by alignment and scaling before features can be measured.
- Feature dimensionality may not be constant. For example, the number of available minutiae points varies.
- Feature space may not be bounded, linear or metric.

2. Relative Entropy of Biometric Measures

Given these considerations, it appears that the most appropriate information theoretic measure for the information content of a biometric feature is the relative entropy ($D(p\|q)$) [3] between the intra- ($q(\mathbf{x})$) and inter-person ($p(\mathbf{x})$) biometric feature distributions, defined as $D(p\|q) = \int p(\mathbf{x}) log_2 \left(p(\mathbf{x})/q(\mathbf{x})\right) d\mathbf{x}$. This measure can be motivated as follows: the entropy, $H(p)$, is the amount of information required on average to describe features $\mathbf{x}$ distributed as $p(\mathbf{x})$. However, H is not in itself an appropriate measure, since it includes information to describe a feature measurement more accurately than the intra-person distribution. On the other hand, the relative entropy, D, is the extra information required to describe a distribution p based on an assumed distribution q [3]. This corresponds nicely to the requirements: given a knowledge of the general, inter-person distribution, the information in a biometric feature set allows us to describe a particular individual.

If p and q are modeled as Gaussian with means μ_p, μ_q and covariances $\mathbf{\Sigma}_p$, $\mathbf{\Sigma}_q$, then the relative entropy in bits is

$$D(p||q) = \frac{1}{2} log_2 \left(\frac{|\mathbf{\Sigma}_q|}{|\mathbf{\Sigma}_p|} e^{trace\left(\left(\mathbf{\Sigma}_p + (\mu_p - \mu_q)^t (\mu_p - \mu_q) \right) \mathbf{\Sigma}_q^{-1} - \mathbf{I} \right)} \right) \qquad (1)$$

Using this measure we can calculate the entropy of a feature representation of face images.

3. Information in an eigenface representation of faces

In order to illustrate this approach, we calculate the information in a feature representation of faces. This result can then be compared to those that from our previous analysis of biometric encryption into face images [1], in which we found a limit to encryption of about 20 bits of key into face images. Using the Aberdeen face database [2] we chose 18 frontal images of 16 persons, from which to calculate the eigenface distribution of features. The distributions of p and q are fitted to a Gaussian model, and the information in each eigenface feature is calculated by averaging D for each intra-person distribution p. Results are shown in Fig. 1; there is a maximum information at eigenface 2 and a subsequent decrease. Summing over the first 100 eigenfaces, there are 41.7 bits, although inter-feature correlations will reduce this value. In order to calculate D for all features, we are limited by the available information. Since 18 images are used to calculate the covariances, attempts to calculate D for more than 17 features will fail because $\mathbf{\Sigma}_p$ is singular. We are thus currently unable to estimate the overall entropy for eigenface features.

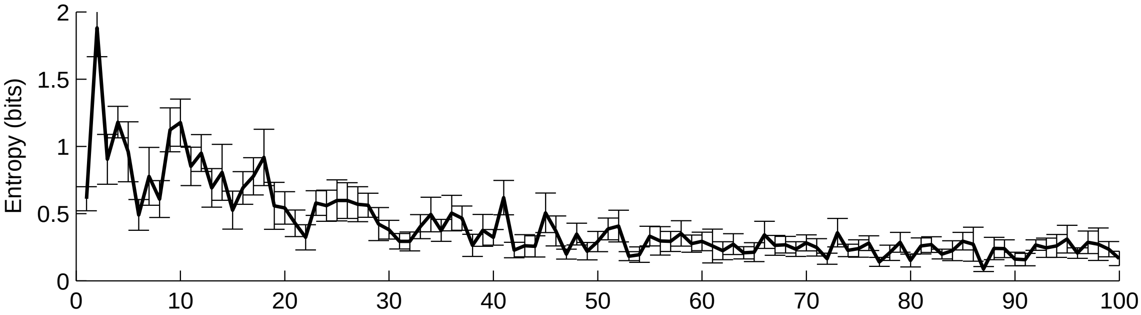

Figure 1. Eigenface feature information (bits) ($\pm SE$) as a function of eigenface number.

4. Discussion

This paper considers the measurement of information in a biometric feature representation, based on the *relative entropy* measure from information theory [3]. We anticipate that such a measure may help address many questions, such as the following:

- Uniqueness of biometric features: A common question is ”are fingerprints really unique?”. While Pankanti *et al.* [5] have recently provided a sophisticated analysis of this problem based on biometric feature distributions directly, a general approach based on information content would help address this question for other biometric modalities.
- Inherent limits to biometric template size requirements
- Feasibility of biometric encryption: Proposed biometric encryption systems use biometric data to generate keys [6], and thus the availability of biometric information limits the security of cryptographic key generation.
- Performance limits of biometric matchers: While some algorithms outperform others, it clear that there are ultimate limits to error rates, based on the information available in the biometric features.
- Privacy protection: It would be useful to quantify the threat to privacy posed by the release of biometric information, and also to be able to quantify the value of technologies to preserve privacy, such as algorithms to de-identify face images [4].

References

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

[2] 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

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

[4] Newton, E.M., Sweeney, L., Malin, B., “Preserving Privacy by De-Identifying Face Images”, *IEEE Trans. Knowledge Data Eng.* **17** 232–243, 2005

[5] Pankanti, S., Prabhakar, S., Jain, A.K., “On the Individuality of Fingerprints”, *IEEE Trans. Pat. Anal. Mach Intel.*, **24**:1010–1025, 2002

[6] Uludag, U., Pankanti, S., Prabhakar, S., Jain, A.K.: “Biometric Cryptosystems: Issues and Challenges”, *Proc. IEEE* **92**:948–960, 2004

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