# A low cost incremental biometric fusion strategy for a handheld device

Lorène Allano, Sonia Garcia-Salicetti, Bernadette Dorizzi

# A Low Cost Incremental Biometric Fusion Strategy for a Handheld Device

Lorene Allano, Sonia Garcia-Salicetti and Bernadette Dorizzi

Telecom & Management SudParis – Institut Telecom
9 rue Charles Fourier, 91011 Evry Cedex, France
{Lorene.Allano, Sonia.Salicetti, Bernadette.Dorizzi }@it-sudparis.eu

**Abstract.** In this paper, we present some results on multimodality implementation resulting from the VINSI ("Vérification d'Identité Numérique Sécurisée Itinérante" for Secured Mobile Digital Identity Verification) French project. The VINSI handheld terminal allows identity verification in mobile conditions (airport gates) using two biometrics usable in future biometric passports (fingerprint and face). We propose an incremental fusion strategy aiming at improving the global performance of the combined system over each individual recognizer while optimizing the cost resulting from the fusion. Indeed, in this kind of application, time and complexity optimization is essential. To this aim, we split the fingerprint scores' range into different interest zones, on which we do not apply the same strategy depending on the relative quality of the modalities at hand. Results on a virtual database corresponding to VINSI applicative conditions (Combination of BIOMET fingerprints and FRGCv2 faces) show that this incremental fusion strategy allows the same improvement in performance as global fusion methods while significantly reducing the cost.

**Keywords:** Biometrics, Fusion, Incremental Strategy, Face, Fingerprint, Handheld Device, Cost Optimization, Multiple Classifiers System.

## 1 Introduction

The VINSI handheld terminal is a mobile secured terminal allowing the control of identity in mobile conditions. Two biometric modalities are implemented, namely face and fingerprint, on this mobile device which contains a highly secured part, named "security box", allowing the terminal itself to be safe and secured from the point of view of storage and processing of biometric data. This is in accordance with the specification of the new biometric passport which will contain, stored on a contactless chip, the fingerprint and face references of the owner of the passport. This way, this terminal could be used to verify a person's identity in airport gates at the

exit of planes or even more generally outdoors, in the framework of mobile customs for instance. Note that the terminal embeds not only the softwares but also the sensors (camera, fingerprint sensor) necessary for the verification stage.

In this article we want to explore the interest of using multimodality in this context. Indeed, multimodality is often presented as a way to increase the performance of the system, to provide an alternative when one modality alone cannot be used and to enhance robustness to forgeries.

However, this VINSI application deserves several remarks. In this mobility context where a lot of disturbance on images will come from illumination variations (indoor versus outdoor acquisitions), fingerprints will give much better results than face. Processing time and usage friendliness are also important issues. Indeed, the processing capabilities of such a mobile terminal are limited compared to those of a fixed PC and in verification applications using the terminal, acquisition of both face and fingerprints can appear fastidious both to the user and to the operator.

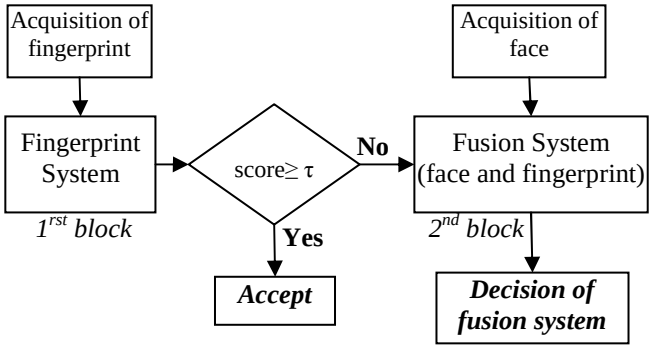


**Fig 1:** The proposed incremental biometric fusion strategy composed of two blocks: the fingerprint system and the fusion system

For this reason, we propose a novel incremental fusion strategy, based on the multistage approach proposed in [1], that we exploit for fusion of biometric modalities. In [1] Pudil et al. proposed a multi-experts system based on a serial architecture with a reject option. At each stage, the reject option consists in an impossible decision of the expert due to a low confidence in its decision. This multistage approach with reject option was also used in [2, 3] applied to handwritten digit and word recognition. In biometrics, Cordella et al. proposed in [4] a cascaded multi expert system for signature verification based on the same approach. Our idea is to used a multi-stage approach applied to multimodal biometrics, that is the fusion of two different biometrics. In most of multimodal systems, experts are fused in a parallel way that we call “global” fusion methods, that is all the available scores are systematically used. The architecture that we propose is a sequential fusion strategy composed of two blocks (see Fig 1). The first block corresponds to the fingerprint system alone, and the second, eventually used, is a fusion system of the two biometric scores (face and fingerprint). Therefore, the architecture has increasing complexity, since using the second block involves the acquisition of a second modality, face.

This strategy was also proposed by Marcialis et al. in [5] for a multimodal biometric system composed of fingerprint and face. The major difference between our

architecture and the one in [5] is the second stage in which the authors use only the second system while we applied fusion of the two modalities. This guarantees better results, as performance of fusion when performed in an optimal way are better than those of any of the two modalities alone [11].

Our incremental fusion strategy will allow increasing system performance to the same level as the one reached by global fusion methods (which always use the two modalities) while using only the additional modality (in this case face) in the situations when this can bring an improvement over using fingerprint alone.

We first describe briefly our experimental monomodal context (systems and database), then we present the global fusion schemes that are commonly used in the literature before introducing our new incremental fusion strategy. Comparative results are finally presented allowing putting forward the interest of this incremental approach.

## 2 Bimodal Database and Evaluation Protocol

Experiments have been realized on publicly available databases suitable in the context of this application because we did not have at disposal during the time of the project a database acquired on VINSI terminal.

In the VINSI application, reference images stored on the ID document are acquired with an optical sensor while the platform embedded sensor which will be used for test, is a capacitive sensor. The BIOMET database allows therefore evaluating the fingerprint system in conditions close to those of the application because it includes fingerprint images acquired with an optical sensor as well as images acquired with a capacitive sensor. BIOMET database contains for each person 6 optical fingerprint images and 6 capacitive fingerprint images. Thus, we can compute in the best case 36 matching scores per person for the fingerprint system. However, as a quality control is applied in order to select images with good enough quality to be used by the system, we selected 124 persons among all BIOMET users having enough genuine scores (30 scores in average).

FRGCv2 [7] database has been chosen for the face system evaluation. This database includes important lighting variations and blurred images as will be the case in the VINSI application. We chose to use the “Experiment 4” protocol associated to FRGCv2. It is a degraded conditions protocol with controlled images (indoor, controlled light, controlled background) as references and uncontrolled images (outdoor, uncontrolled light, blurred images …) for test. As for fingerprints, the face modality also has a quality control that selected images that have good enough quality to be used by the face system. We thus selected from FRGCv2, 124 persons with in average 30 genuine face matching scores per person, so that we have the same number of persons and the same number of matching scores per person for both biometric modalities.

In order to realize fusion experiments we considered a database of virtual persons obtained by combining a person of the BIOMET database with a person of the FRGCv2 database. This can be considered as valid as fingerprint and face are two uncorrelated modalities [8]. For each person, each of the 30 fingerprint matching

scores is coupled with a face matching score. This permits to have 124 virtual persons with in average 30 pairs of genuine scores (one fingerprint score associated to one face score).

The fingerprint system used in this study is NFIS2 [9] (NIST Fingerprint Image Software) developed by NIST (National Institute of Standards and Technology). The Face system has been developed by our team and is based on PCA (Principal component analysis) [10].

Score distributions provided by both systems and presented in Fig 2, allow noticing that performance of both systems are very different. Client and impostor scores distributions are well separated for the fingerprint system as shown in Fig 2 (left).Because of this sharp separation, the fingerprint system considered individually can take the correct decision in most of the cases. On the contrary, scores distributions of clients and impostors for the face system have an important overlap, as shown in Fig 2 (right). This will induce a lot of decision errors when considering the face system individually.

## 3 Global Fusion Methods

The aim of all fusion methods is to take a final decision, that is to separate the 2 classes (client and impostor) in the 2-dimensional space of scores (face and fingerprint scores).

In this section we will present global fusion methods, that is, methods that always fuse all the available scores. Most of these methods are described in [11].

We will present two types of global fusion methods:

- Score combination after a normalization step. Combination could be a simple average of the scores or a weighted average. Normalization of scores as well as of weighted coefficients if needed, requires a learning phase for the normalization parameters and the weights. Combination methods normalize each type of score separately before combining them.
- Classification with a learning phase. In this paper, we will present only two methods that model the bi-dimensional distribution of the scores provided by the fingerprint and face systems. These methods model the two types of scores together by estimating their joint distribution as already presented in [12].

We will now detail two score normalization schemes for combination methods:

- MinMax normalization method.
- Normalization method using posterior probabilities.

The MinMax score normalization method is very simple. The aim is to rescale each system's scores in the same interval. Indeed, as we can see in Fig. 2 (left and right), fingerprint scores vary from 0 to 1000 while face scores vary from -1 to 1. MinMax normalization permits to rescale each system's scores in the [0,1] interval. Normalization parameters are estimated on a dedicated training set.

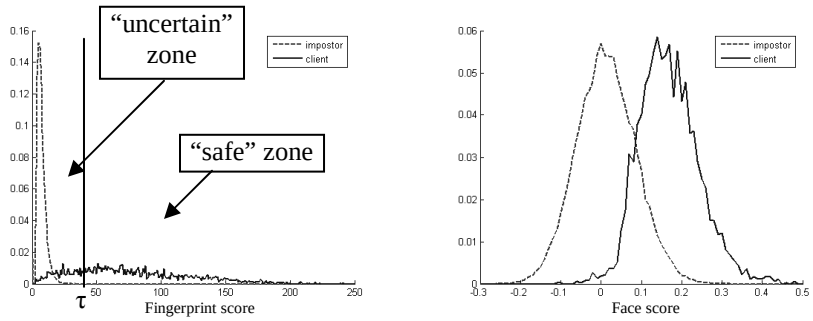


**Fig 2 :** Scores distributions of the fingerprint system (left) and face system (right). The "safe" and "uncertain" zones for the fingerprint system (first block of the architecture) are reported in Fig. 2 (left).

Normalization using posterior probabilities consists in transforming each system's scores in posterior probabilities to be a client. To this aim, an assumption of normal distribution for the two classes (client and impostors) is made. This assumption permits to estimate the class conditional densities of each system by using a Gaussian component whose parameters are learnt on a dedicated training set. Bayes rule is then use to compute the posterior probabilities.

This normalization also allows rescaling each system's scores in the [0,1] interval. It generally turns out to be more precise than MinMax normalization presented previously because it relies on a model of each class and not only on a simple rescaling of scores. However, the model requires a training set representative of the application population and large enough to learn the Gaussian parameters accurately.

We will now present two fusion methods based on a bi-dimensional scores distribution model:

- Joint class conditional densities estimation with assumption of joint normal distribution.
- Estimation of joint class conditional densities based on Gaussian Mixture Model (GMM) [13].

These two bi-dimensional fusion methods consider face and fingerprint scores jointly and not separately as done by combination methods. Their objective is to estimate the joint class conditional densities in order to estimate the probability to be a client knowing the scores pair (face and fingerprint). This probability is the posterior probability given by the Bayes rule.

The difference between the two bi-dimensional fusion methods is the way the two class conditional densities are estimated. The first method is based on a normal assumption of the two-dimensional class conditional densities. The second method is less restrictive on the shape of the joint density of scores, since it is based on Gaussian Mixture Model (GMM) [13].

## 4 Performance of Global Fusion Methods

We run fusion experiments using the global fusion methods presented above. We also tested another combination method using a weighted sum. For this method, scores are normalized using MinMax method before being weighted. Weights are learnt in order to minimize the Half Total Error Rate (HTER) on the training set, by an exhaustive search of weights such as their sum equals 1.

In Table 1, the Equal Error Rate (EER) of each fusion method is reported as well as their relative improvement compared to face and fingerprint systems. The 90% confidence intervals for EER values are also indicated in square brackets. Confidence Intervals are estimated using a parametric method as described in [14]. Performance is estimated using a leave-one-out cross-validation protocol [15] on the persons. In practice, for each person, we learn the fusion parameters on the 123 remaining ones in order to produce the result associated to this person. All the global fusion methods improve the performance of the fingerprint system considered individually. Fusion by MinMax normalization and average of scores which is the simplest method does not improve significantly performance of the fingerprint system alone. All other schemes bring a big improvement on the performance compared to the best unimodal systems (fingerprint) by 42% to 45%. We notice that when scores are weighted before combining them (after normalization), performance is in the same range as those obtained with more complex methods involving density estimation (GMM, normalization using posterior probabilities, joint density model). It could be explained by the fact that the size of the training set is too small to correctly estimate densities for models like GMM.

**Table 1.** Performance of global fusion methods.

| | Average after MinMax normalization | Average of posterior probabilities | Client posterior probability after joint normal density estimation | Client posterior probability after joint density estimation by GMM | Weighted sum after MinMax normalisation (Fingerprint weight= 0.8 ; Face weight = 0.2) |
|---|---|---|---|---|---|
| EER (%) | 4.33 [3.97;4.68] | 2.65 [2.37;2.93] | 2.57 [2.29;2.84] | 2.57 [2.29 ;2.84] | 2.49 [2.22 ;2.76] |
| Relative improvement of EER compared to fingerprint system (EER=4.59% [4.23;4.95]) | 6% | 42% | 44% | 44% | 45.7% |
| Relative improvement of EER compared to face system (EER=13.87% [13.27;14.47]) | 69% | 81% | 81.5% | 81.5% | 82% |

## 4 Incremental Fusion Strategy

In Fig 2 (left) that represents the client and impostor distributions of fingerprint scores on the test set, we can distinguish two zones where the fingerprint system has different behaviors and performance: a "safe" zone for high fingerprint scores, that corresponds only to client scores, and an "uncertain" zone for low fingerprint scores, where client and impostor scores overlap.

The implementation of our incremental fusion strategy is therefore dependent on a threshold $\tau$ on the fingerprint score (see Fig 1). If the fingerprint score is higher than $\tau$ (belongs to the "safe" zone), the person is accepted, otherwise (the score is in the "uncertain" zone), we exploit the second block of our incremental fusion architecture, which is the fusion system combining both biometrics (see Fig 1).

We define the cost of the fusion scheme as the average time (in seconds) for acquisition and processing of the biometric data of one user. The longer is the identity verification process, the higher is the cost. The total time is composed of 5 steps: fingerprint acquisition, fingerprint processing, face acquisition, face processing and fusion of fingerprint and face. In Table 2, we give approximations of the time needed for each of these 5 steps in the framework of the VINSI application on the handheld terminal.

**Table 2.** Time estimations of each step of the verification process on the VINSI terminal

| Fingerprint acquisition | Fingerprint processing | Face acquisition | Face processing | Processing of scores fusion | Average time for one modality (fingerprint) | Average time for two modalities |
|---|---|---|---|---|---|---|
| 10 s | 1 s | 10 s | 4 s | 0.001 s | 11 s | 25 s |

Acquisition times are estimated empirically while processing times correspond to experiments on the mobile terminal. All the values given in Table 2 are average estimations and can vary from one user to the other and from one system to the other. Time for score fusion is not significant compared to all the other times. These values are only estimations allowing illustrating the cost gain using the incremental architecture.

The cost for global fusion methods is 25s. For the incremental fusion, for example if 30% of the users required two modalities, the cost will be 15,2s, that is a decrease of 40% compared to global fusion methods. Indeed, for these clients, the operator needs to acquire the face and process the fusion of both biometrics, that represents extra time cost mainly because of face acquisition, quality control and preprocessing. This extra time cost may be critical in certain applications where human flow is extremely important. The cost (c) is linearly related to the proportion (p) of clients requiring two modalities by the equation: $c=11*(1-p)+25*p$ .

The objective of our strategy is to minimize the cost while still keeping a good performance (at least as good as the global fusion one). For this purpose, we propose to determine the acceptance threshold ($\tau$) for fingerprint scores by the following formula:

$$\tau^* = \arg\min_{\tau} \left( FAR(\tau) \middle| FAR(\tau) > 0 \right) \quad (1)$$

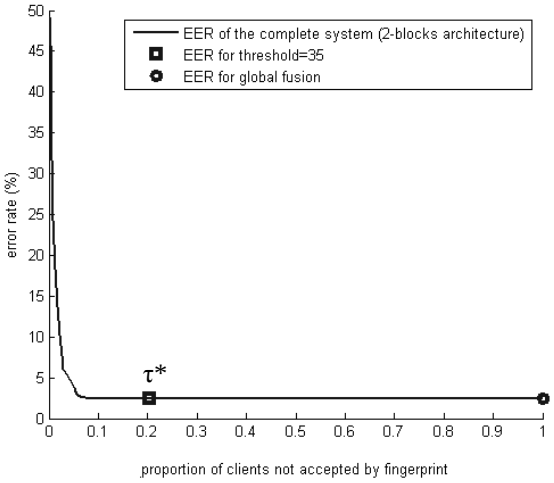


**Fig 3:** Error rate of the complete fusion system (2-blocks architecture) vs. the proportion of clients not accepted by fingerprint only at the first stage, on the test set.

That is, $\tau^*$ corresponds to the threshold giving the lower non zero FAR (False Acceptance Rate). It follows that $\tau^*$ also corresponds to the maximum value of the impostor scores.

In order to determine this value, we used an independent development set of fingerprint data extracted from the same BIOMET database. On this development set, the corresponding acceptance threshold estimated by Eq. (1) is 35. It roughly corresponds to 70% of client scores and 0% of impostor scores in the "safe" zone (that is, 70% of clients are only processed by the first block of the architecture (fingerprint)) and 100% of impostor scores and 30% of client scores in the "uncertain" zone. It corresponds to an average cost of 15,2s (0,7*11+0,3*25) on the development set, that is a decrease of 45% of the cost compared to global fusion.

This value of the threshold corresponds on the test set to 80% of client scores and 0% of impostor scores in the "safe" zone and to 20% of client scores and 100% of impostor scores in the "uncertain" zone. It gives an average cost of 13,8s (0,8*11+0,2*25) on the test set.

With this strategy, we do not insure the optimality of the choice of the threshold (as defined in Eq. (1)) on the test database but due to the fact that the impostor score distribution drops sharply in its upper tail, the corresponding values of cost and error rates on the test set will be very similar to those obtained on the development set. The important point is that the strong decision power of the fingerprint system in the "safe" zone (no impostors) is kept on the test database when using this strategy for choosing the threshold on the development set.

More generally, we can study the variation of the Error Rate in function of the cost when the threshold $\tau$ is varying. In Fig 3 the complete system (2-blocks architecture) error rate is presented depending on the proportion of clients needing to be processed by the second stage. Two special cases are considered, namely proportion=1, indicated in Fig 3 by the mark (o). In this case, the two scores are always fused through a global fusion strategy, corresponding to a cost of 25s. The second value

indicated in Fig 3 by the mark (□) correspond to the proportion=0.2 (τ*=35). We can observe that for all the values in the interval [0.2, 1], the error rate roughly remains constant. It is even true for lower values (until p=0.08). Another way to determine the threshold could be to find the lowest cost associated to the incremental strategy whose performance is still equivalent to that of the global strategy on the development dataset (threshold for p=0.08). But the main drawback of this threshold optimization is that it is not robust in generalization. As we can see in Fig 3, a small variation of threshold around 0.08 could imply a large degradation of performance. On the contrary, our choice of the threshold as defined in Eq. (1) is robust to small variations as explained above.

In the second block of our incremental architecture, the fusion method used to combine fingerprint and face can be of two sorts: it can be a global fusion method as described in Section 3; in this case, the parameters of the fusion scheme are learnt on the complete training dataset (all the available scores). But it could also be a local fusion method whose parameters are learnt only on the scores of the training set which belong to the "uncertain" zone. This direction will be studied in future work.

## 4 Conclusion

In this paper, we realized fusion experiments in the context of a mobile terminal aiming at performing identity verification using fingerprint and face. We implemented global fusion methods where fingerprint and face are combined for each test access. We noticed a significant improvement of performance due to fusion: 40% to 45% of relative improvement for global fusion methods compared to the best modality (fingerprint). We also implemented an incremental fusion strategy where the second biometric modality (face in this example) is used only when the fingerprint system alone is not sufficient to take a reliable decision. The aim of our incremental fusion strategy is to minimize the cost of using several modalities while keeping the same level of performance than global fusion methods. In the VINSI application, the cost is due to acquisition of biometric modalities and their processing. It is estimated in our strategy by the proportion of clients for which we need to acquire the second modality (face). With our incremental strategy we have shown that 80% of clients can be authenticated using fingerprint alone without the need of the face modality leading to a decrease of 55% of the cost compared to global fusion methods. This property is very important particularly in the context of a mobile identity verification application through a handheld device like the VINSI terminal, in which computational constraints are significant and real time response is mandatory. The only drawback of our incremental strategy is that if we do not use a second modality, we cannot guarantee that the fingerprint has not been forged, while it is possible at low effort as explained in [16]. Nevertheless, in this kind of application, a human operator is physically present and can therefore do a visual control thanks to the face image that is stored in the passport and can be visualized on the screen of the VINSI terminal. The big cost improvement that results from our incremental fusion strategy is due in particular to the fact that the client and impostor scores of the fingerprint system are very well separated and thus that fusion can be largely driven by the fingerprint

modality. Anyhow, this incremental fusion strategy can also be applied with profit to other application contexts reflecting other behaviors of the modalities at hand.

**Acknowledgments.** This work was funded by the French ministry of Industry.

## References

1. Pudil, P., Novovicoca, J., Blaha, S., Kittler, J.: Multistage Pattern Recognition with Reject Option. In: Proc. 11th IAPR Internat. Conf. on Pattern Recognition, The Hague. IEEE Computer Society Press, Los Alamos, pp. 92--95 (1992)
2. Alpaydin, E., Kaynak, C.: Cascading Classifiers. Kybernetika, vol. 34(4), 369--374 (1998)
3. Mandhavanath, S., Govindaraju, V.: Serial Classifier Combination for Handwritten Word Recognition. In: Proceedings of the Third international Conference on Document Analysis and Recognition (Volume 2) - Volume 2 (August 14 - 15, 1995). ICDAR. IEEE Computer Society, Washington, DC, 911 (1995)
4. Cordella L. P., Foggia P., Sansone C., Tortorella F., Vento M.: A Cascaded Multiple Expert System for Verification. Multiple Classifier Systems 2000, pp. 330-339 (2000)
5. Marcialis G. L., Roli F.: Serial Fusion of Fingerprint and Face Matchers. MCS 2007, pp. 151-160 (2007)
6. Garcia-Salicetti, S., Beumier, C., Chollet, G., Dorizzi, B., Leroux-Les Jardins, J., Lunter, J., Ni, Y., Petrovska-Delacretaz, D.: BIOMET: a Multimodal Person Authentication Database Including Face, Voice, Fingerprint, Hand and Signature Modalities". In: Audio and Video-Based Biometric Person Authentication, 4th International Conference (AVBPA'03), Lecture Notes in Computer Science (LNCS), Vol. 2688, pp. 845--853, Guildford, Angleterre (2003)
7. FRGCv2 database, http://www.frvt.org/FRGC/
8. Dorizzi, B., Garcia-Salicetti, S., Allano, L.: Multimodality in Biosecure : Evaluation on Real vs. Virtual Subjects. In: Proc. of ICASSP. 14-19 May 2006, Toulouse, France (2006)
9. NIST Fingerprint system NFIS2, http://www.itl.nist.gov/iad/894.03/fing/fing.html
10. Turk, M., Pentland, A.: Eigenfaces for Recognition. In: Journal of Cognitive Neuroscience. Vol 3, No. 1. 71—86 (1991)
11. Ross, A.A., Nandakumar, K., Jain, A.K.: Handbook of multibiometrics. Springer (2006)
12. Allano, L., Morris, A.C., Sellahewa, H., Garcia-Salicetti, S., Koreman, J., Jassim, S., Ly-Van, B., Wu, D., Dorizzi, B.: Non intrusive multi-biometrics on a mobile device: a comparison of fusion techniques. In :Proc of SPIE 2006, conference on Biometric Techniques for Human Identification III, Orlando, Florida, USA, 17-21 April, (2006)
13. Reynolds, D.A., Rose, R.C.: Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models. In: IEEE Transactions on Speech and Audio Processing, vol. 3, No. 1, Jan (1995)
14. Bolle, R.M., Ratha, N.K., Pankanti, S.: Error Analysis of pattern recognition systems – the subset bootstrap. In: Computer Vision and Image Understanding,, vol. 93 (2004)
15. Devijver, P.A., Kittler, J.: Pattern Recognition: A Statistical Approach. Prentice-Hall, London (1982)
16. Matsumoto, T., Matsumoto, H., Yamada, K., Hoshino, S.: Impact of Artificial Gummy Fingers on Fingerprint Systems. In: Proceedings of SPIE, vol. 4677, Optical Security and Counterfeit Deterrence Techniques IV (2002)