On Multiview Analysis for Fingerprint Liveness Detection

Amirhosein Toosi, Sandro Cumani, and Andrea Bottino(✉)

Politecnico di Torino, Corso Duca Degli Abruzzi, 24, 10129 Turin, Italy
{amirhosein.toosi,sandro.cumani,andrea.bottino}@polito.it

Abstract. Fingerprint recognition systems, as any other biometric system, can be subject to attacks, which are usually carried out using artificial fingerprints. Several approaches to discriminate between live and fake fingerprint images have been presented to address this issue. These methods usually rely on the analysis of individual features extracted from the fingerprint images. Such features represent different and complementary views of the object in analysis, and their fusion is likely to improve the classification accuracy. However, very little work in this direction has been reported in the literature. In this work, we present the results of a preliminary investigation on multiview analysis for fingerprint liveness detection. Experimental results show the effectiveness of such approach, which improves previous results in the literature.

Keywords: Spoofing detection · Multiview approach · SVM · Multi task learning · Sparse reconstruction

1 Introduction

On September 2014, the new iPhone 6 was unveiled and released on sale. This device is equipped with a Touch ID fingerprint reader allowing users to unlock their device and to authenticate for on-line purchases. Two days after that launch, a group of German hackers showed how to bypass the Touch ID security system [1]. This is just one of the many possible examples of the vulnerability of fingerprint recognition systems, which is a severe issue due to the integration of such devices into a number of forensic, commercial and military applications [2]. The typical scenario depicts an adversary trying to gain unauthorized access by using the biometric traits of a person legitimately enrolled into the system. In the case of fingerprint recognition systems, these attacks are usually carried out using spoof artifacts, i.e. duplicated artificial fingerprints. Artificial fingerprints can be created filling a mold, obtained from a live or a latent fingerprint, with materials such as gelatine, silicone or Play-Doh [3]. It has been shown that the success rate of such spoof attacks can be up to 70% [4].

To address this problem, several methods have been developed to detect the *liveness* of a fingerprint image. Software-based approaches distinguish between live and fake fingerprint relying solely on the digital processing of images acquired

A. Pardo and J. Kittler (Eds.): CIARP 2015, LNCS 9423, pp. 143–150, 2015.
DOI: 10.1007/978-3-319-25751-8_18

from the device, and can be further divided into dynamic and static ones. Dynamic methods are based on the analysis of certain phenomena like skin deformation [5] and perspiration [6] on a temporal image sequence. However, these methods are not general, since their multi-temporal dimension makes them applicable in a minority of operative conditions. Static methods, on the contrary, focus their analysis on a single fingerprint image, which makes them more general and attractive. These methods can be, again, divided into two main categories. Holistic methods process the image as a whole to derive some discriminative global characteristics, such as the texture coarseness [7] or several first and second order statistics (like mean, energy, entropy, variance, skewness [7] or Gray-Level Co-Occurrence Matrices [8]). However, as shown in [9], their discriminative power is quite low, while better performance is given by local methods, which rely on mathematical descriptors summarizing texture features of small regions surrounding an image point. Global image descriptors can then be obtained by summing up the local descriptors into a histogram collected from the whole image or into multiple histograms obtained from image patches.

Several global image descriptors have been experimented in the context of spoof detection, such as basic [10] and multi-scale [11] Local Binary Pattern (LBP), Local Phase Quantization (LPQ) [12], Weber Local Descriptor (WLD) [13] and Binary Statistical Image Features (BSIF) [14]. Recently, Local Contrast Phase Descriptors (LCPD), a novel global descriptor specifically designed to deal with the characteristics of fingerprint images, has been proposed in [9]. Local Contrast Phase Descriptors (LCPD) is composed by a spatial-domain component, derived from WLD, and by a rotation invariant phase component, derived from LPQ.

All these descriptors provide complementary information or, equivalently, complementary views of the objects under analysis. Previous studies in the area of pattern recognition and machine learning have shown that the combination of features of different nature is usually a powerful method to improve the recognition accuracy of the final classifier. Despite that, such integration has not been fully analyzed yet in the context of fingerprint liveness detection. To the best of our knowledge, the only paper tackling this problem was [13], where the integration of WLD plus LPQ and LBP plus LPQ were analyzed. To this end, in this paper we present the preliminary results of an investigation aimed at detecting the fingerprint liveness by analyzing the integration, at feature level, of different attributes summarizing individual fingerprint images from different views.

The remainder of the paper is organized as follows. Section 2 outlines our approach and Section 3 presents and discusses the experimental results. Finally, conclusions are drawn in Section 4.

2 Multiview Approaches to Fingerprint Liveness Detection

When tackling this work, our main research questions were the following. Which are the global descriptors most suited to tell live from fake fingerprints? Which

are the best combinations of different descriptors and how can they be effectively combined to improve the classification accuracies? As we already stated in the introduction, in this paper we will provide some preliminary answers to these questions. These answers are supported by the results of our experiments, which show the effectiveness of multiview approaches in developing anti-spoofing software systems.

In the following subsections we will first discuss two possible methods to (i) integrate, at feature level, different attributes (i.e. feature types) extracted from fingerprint images, and (ii) to classify them into live and fake fingerprints. Then, we will describe the set of individual attributes we found most suitable for the problem in analysis.

2.1 Support Vector Machines (SVM) Based Classification

A simple but effective way of combining multiple representations of the same sample is to concatenate the characteristic vector of each representation. Denote $y = [y^1, \ldots, y^K]$ a test sample described under K *tasks*, where each task represents a different view of the sample (i.e. for images, tasks can be colour histograms, edges, local descriptors and so on). Each task $y^k \in \mathbb{R}^{m_k}$, and each sample $y \in \mathbb{R}^m$, where $m = \sum_{k=1}^{K} m_k$.

The samples are then fed to a linear SVM for classification. The choice of linear SVMs was mainly motivated by the properties of the datasets used in our experiments and by the good accuracy the linear kernel achieves. Indeed, linear SVMs tend to be less prone to overfitting, due to the lower complexity of the separation surface. The dimensionality of the input space is sufficiently high to ensure that the linear classifier is able to properly separate the classes (as we will show in the results section). Furthermore, linear SVMs provide huge benefits in terms of time and memory requirements, since the separation hyperplane can be computed offline and scoring reduces to a simple dot-product in feature space. Finally, SMVs provide a good alternative to feature selection, on the condition that the regularization coefficient is properly chosen. The main motivations behind feature selection are the removal of nuisance dimensions and the reduction of overfitting issues. The presence of the regularization term in the SVM objective function tends to favour simpler separation surfaces, thus mitigating the problems of overfitting, especially in presence of large dimensional vectors, thus improving the generalization capabilities of the model [15].

2.2 Multi-Task Joint Sparse Reconstruction Classification (MTJSRC)

Multi-Task Joint Sparse Reconstruction Classification (MTJSRC), introduced in [16], combines multi-task learning and classification based on sparse representation. In brief, sparse coding aims at representing a signal as a linear combination of a set of reference samples enforcing sparsity in the coefficient set. Multi-task or multi-view learning aims at jointly estimating models from multiple representations of the same data.

Suppose we have a training set $X^k = [X_1^k, \ldots, X_J^k]$ for task k, where J is the number of classes and $X_j^k \in \mathbb{R}^{m_k \times n_j}$, with n_j the number of training samples for class j. Given a test sample y, we can reconstruct each of its representation modalities y^k from the corresponding training set X^k as:

$$y^k = \sum_{j=1}^{J} X_j^k w_j^k + \epsilon^k \qquad k = 1, \ldots, K$$

where w_j^k are the reconstruction coefficients associated to class j and task k and ϵ_k is the residual for the kth modality. Defining $w_j = [w_j^1, \ldots, w_j^K]$ as the representation coefficients for the jth class across the different tasks, the multi-task joint sparse representation can be obtained from the solution of the following least square regression problem:

$$\min_W \frac{1}{2} \sum_{k=1}^{K} \left\| y^k - \sum_{j=1}^{J} X_j^k w_j^k \right\|_2^2 + \lambda \sum_{j=1}^{J} \|w_j\|_2 \tag{1}$$

where $W = [w_j^k]_{j,k}$ and $\lambda \sum_j \|w_j\|_2$ is a regularization term.

Model Optimization. In [16], the authors proposed Accelerated Proximal Gradient (APG) for model optimization. A drawback of APG is that, to ensure convergence of the objective function in reasonable time, it requires proper selection of the gradient step at each iteration. The main issue in the optimization of (1) is the presence of a non–differentiable regularization term. Several approaches could be modified to handle this regularizer [17]. In practice, we observed that, for the task at hand, convergence can be easily achieved through the use of L–BFGS algorithm [18], provided that the regularizer is replaced by an ϵ–smoothed term (for small values of ϵ): $\sum_j \|w_j\|_2 \approx \sum_j \sqrt{\|w_j\|_2^2 + \epsilon}$

Classification. Once the optimal reconstruction coefficients for a test sample y have been computed, for each task k and each class j it is possible to compute the reconstruction error $\|y^k - X_j^k w_j^k\|_2$. A straightforward way to assign the sample label is then to pick the class minimizing the sum of the reconstruction errors over all the sample modalities:

$$label = \arg\min_j \sum_{k=1}^{K} \theta^k \|y^k - X_j^k w_j^k\|_2 \tag{2}$$

where $\Theta = \{\theta^k\}$ are values weighting the relative relevance of the different modalities in the final classification choice. Since our classification problem is binary, it is easy to verify that the label assignment in (2) corresponds to:

$$label = \begin{cases} 1 & \text{if } \sum_k \theta^k \left(\left\| y^k - X_1^k \right\|_2 - \left\| y^k - X_2^k \right\|_2 \right) < 0 \\ 2 & \text{if } \sum_k \theta^k \left(\left\| y^k - X_1^k \right\|_2 - \left\| y^k - X_2^k \right\|_2 \right) > 0 \end{cases} \tag{3}$$

Equation (3) can be interpreted as the fusion of K different systems with associated score function:

$$s_k(y^k) = -\theta^k \left(\left\| y^k - X_1^k \right\|_2 - \left\| y^k - X_2^k \right\|_2 \right) \ . \tag{4}$$

where higher (resp., lower) values of (4) tend to favour the hypothesis of sample belonging to class one (resp., two). We define the scoring function for the *fused* systems as:

$$s(\Theta, y) = \theta^0 + \sum_k \theta^k s_k(y^k)$$

where an additional term θ^0 is added to act as bias. Given a validation set, the weights are then estimated by training a Logistic–Regression (LR) classifier. The advantage of LR–based fusion with respect to the approach in [16] is that it both allows to improve the discriminative ability of the system, and, instead of simply providing class membership, it produces outputs which can be interpreted as log–likelihood ratios between class hypotheses [19]. Moreover, the LR objective function is convex, and can be easily trained using standard solvers as L–BFGS.

2.3 Feature Extraction

The individual fingerprint images have been characterized with the following attributes: Histogram of Oriented Gradients (HOG) [20], BSIF, LPQ, WLD and several variants of the LBP, such as patch-based or rotation invariant LBP [21].

The results of initial experiments, which we do not report for the sake of brevity, led us to exclude from the list of candidate attributes both HOG and the various LBP formulations since they were consistently providing lower accuracies than other candidates. As for the other attributes, we provide in the following a brief description and pieces of information on their computation.

BSIF are histograms of binary codes computed for each pixel. The pixel code is obtained by projecting local image patches onto a subspace learnt from natural images. From the results in [14], it can be deduced that variations of the local window size actually capture different characteristics of live and fake fingerprint images. Thus, we experimented different window sizes (from 3x3 to 17x17) as complementary attributes, each of which has dimension 4.096. LPQ codes are obtained first computing local phase information on a window surrounding each pixel, by means of different possible filters, and then extracting the quantized phase of selected frequency components. Histograms of LPQ codes in image patches are then computed and concatenated. Each LPQ attribute has size 256. WLD compute for each pixel the differential excitation (the ratio between the sum of neighboring pixel intensity and the intensity of the pixel itself) and the orientation of the pixel gradient. WLD features of the image can be computed at different image scales and then encoded into a histogram that contains, for each scale, 960 elements.

3 Results and Discussion

The accuracies of our liveness detection approach were assessed on LivDet 2011 dataset [3]. This dataset is one of the most used in the literature and, thus, allows for a comparison with a large number of methods. LivDet 2011 consists of four datasets of images acquired from different devices (*Biometrika*, *Digital Persona*, *Italdata* and *Sagem*). For each device, 2000 live images of different subjects and 2000 fake images obtained with different materials (such as gelatine, latex, PlayDoh, silicone and wood glue), were collected. Images were divided into a training and a test set, each containing an equal number of live and fake images, with fake images equally distributed among different materials. LivDet 2011 datasets were acquired using a consensual method [3], where the subject actively cooperated to create a mold of his/her finger, thus obtaining surrogates of better quality, i.e. more difficult to detect, than those created from latent fingerprints.

Experiments were organized as follows. First, we optimized the parameters of each method, i.e. the parameter C of the linear SVM and the task weights for MTJSRC, with a 5-fold cross validation procedure on the training set. Then, we computed the classification capabilities of both individual and grouped attributes. A preliminary result, not detailed for the sake of brevity, is that the individual attributes perform consistently worse than their combination with other attributes, demonstrating the strength of multiview approaches. As for grouped attributes, we tested different combinations of the candidate attributes described in Section 2.3 and, for each candidate attribute, we tested different parameter settings. We found that the best results were obtained for WLD using three different image scales (referred in the results as W3), for BSIF using different windows size, 5x5 (B5), 15x15 (B15) and 17x17 (B17), while for LPQ we obtained similar results computing phase information with either Short Term Fourier Transform (LS) or Gaussian derivative quadrature filter pairs (LG).

Results are summarized in Table 1 where we report error rates for each method and dataset and for the attribute groups that obtained the best results; average error rates over all the datasets are reported as well. The baseline for benchmarking our results was the method [9], which combines feature selection, linear SVM and LCPD outperforming previous results in the literature (see [9]). Bold values in Table 1 are those improving or matching the baseline.

Based on the results, the following remarks can be drawn. We found several attribute groups improving the baseline, and an optimal average error of 5.2% was obtained on SVM with the combination is B5+B17+W3+LG, reducing of 8.8% the error rates of [9]. In general, SVM performs better than MTJSRC. However, this result deserves a closer look to the data, as Italdata accuracies stand out for being definitely higher than the average. In particular, MTJSRC appears to be severely penalized by the performances on this dataset. Indeed, the accuracies of the two methods on the other datasets are definitely comparable (their average difference being 0.4% in favour of SVM). Specifically, if we consider the last three groups in Table 1. i.e. those combining multiple BSIF features, their average accuracies over Biometrika, Digital Persona and Sagem dataset improve the corresponding accuracies of the baseline of 36.2% and 28.6% for,

Table 1. Performance comparision on LivDet 2011.

	Feat. chaining + linear SVM					MTJSRC				
Feature set	Biom	DigP	IData	Sag	Avg	Biom	DigP	IData	Sag	Avg
B5 + W3	5.7	**3.5**	**9.3**	3.7	**5.6**	8.7	4.5	18.6	3.9	8.9
B5 + W3 + LG	5.6	**3.4**	**9.0**	3.7	**5.4**	6.2	**4.3**	18.8	4.1	8.3
B5 + B15 + W3 + LF	**3.7**	**1.8**	13.0	2.8	**5.3**	**4.5**	**2.0**	18.2	**2.4**	6.8
B5 + B15 + W3 + LG	**3.8**	**1.7**	13.3	2.8	**5.4**	**4.6**	**2.2**	17.9	**2.5**	6.8
B5 + B17 + W3 + LG	**4.7**	**2.0**	11.4	**2.7**	**5.2**	**4.3**	**2.5**	19.1	**2.6**	7.1
Baseline(LCPD) [9]	4.9	4.2	11.0	2.7	5.7	4.9	4.2	11.0	2.7	5.7

respectively, SVM and MTJSRC. These last partial result seems also to suggest that performance is mainly related to the design of the feature set rather than to the choice of the classifier itself.

Hence, how can this behaviour on Italadata be interpreted? Actually, similar problems were reported in the literature ([14], [9]), and a possible explanation is that Italdata images seem to be more clear and less natural than those obtained with other sensors. The somewhat contradictory behaviour of this dataset in our experiments supports this conjecture. For SVM, groups scoring well on Italdata performed badly on other datasets, and the other way around. As for MTJSRC, we found that optimal λ values of the regularizer in (1) for Italdata penalized the other dataset accuracies and, again, the opposite.

Concluding, we think that, overall, our preliminary results highlights the benefits of tackling the fingerprint liveness detection problem with multiview approaches.

4 Conclusion and Future Work

We presented the initial results of our investigation on the application of a multiview approach to the problem of fingerprint liveness detection. This approach combines in various ways different and complementary representation modalities of the samples under analysis. Experimental results show the strength of such approach, which is capable of improving, on the same data, previous results in the literature. These preliminary outcomes are promising but, at the same time, highlight the fact that further studies are sorely needed to fully understand the different factors involved in the problem, which will be the objective of our future work.

References

1. International Business Time: iphone 6 touch id fingerprint scanner hacked days after launch (2015). http://tinyurl.com/px33mre. (Accessed June 01, 2015)
2. Marasco, E., Ross, A.: A survey on antispoofing schemes for fingerprint recognition systems. ACM Comput. Surv. **47**(2), 28:1–28:36 (2014)

3. Yambay, D., Ghiani, L., Denti, P., Marcialis, G., Roli, F., Schuckers, S.: Livdet 2011 - fingerprint liveness detection competition 2011. In: 2012 5th IAPR International Conference on Biometrics (ICB), pp. 208–215, March 2012
4. Matsumoto, T., Matsumoto, H., Yamada, K., Hoshino, S.: Impact of artificial "gummy" fingers on fingerprint systems. In: Proceedings of SPIE **4677**, January 2002
5. Antonelli, A., Cappelli, R., Maio, D., Maltoni, D.: Fake finger detection by skin distortion analysis. IEEE Transactions on Information Forensics and Security **1**(3), 360–373 (2006)
6. Schuckers, S.A.C., Parthasaradhi, S.T.V., Derakshani, R., Hornak, L.A.: Comparison of classification methods for time-series detection of perspiration as a liveness test in fingerprint devices. In: Zhang, D., Jain, A.K. (eds.) ICBA 2004. LNCS, vol. 3072, pp. 256–263. Springer, Heidelberg (2004)
7. Abhyankar, A., Schuckers, S.: Fingerprint liveness detection using local ridge frequencies and multiresolution texture analysis techniques. In: 2006 IEEE International Conference on Image Processing, pp. 321–324, October 2006
8. Nikam, S., Agarwal, S.: Co-occurrence probabilities and wavelet-based spoof fingerprint detection. Int. Journal of Image and Graphics **09**(02), 171–199 (2009)
9. Gragnaniello, D., Poggi, G., Sansone, C., Verdoliva, L.: Local contrast phase descriptor for fingerprint liveness detection. Pattern Recognition **48**(4), 1050–1058 (2015)
10. Nikam, S., Agarwal, S.: Texture and wavelet-based spoof fingerprint detection for fingerprint biometric systems. In: First International Conference on Emerging Trends in Engineering and Technology, ICETET 2008, pp. 675–680, July 2008
11. Jia, X., Yang, X., Cao, K., Zang, Y., Zhang, N., Dai, R., Zhu, X., Tian, J.: Multi-scale local binary pattern with filters for spoof fingerprint detection. Information Sciences **268**, 91–102 (2014)
12. Ghiani, L., Marcialis, G., Roli, F.: Fingerprint liveness detection by local phase quantization. In: ICPR 2012, pp. 537–540, November 2012
13. Gragnaniello, D., Poggi, G., Sansone, C., Verdoliva, L.: Fingerprint liveness detection based on weber local image descriptor. In: IEEE BIOMS 2013, pp. 46–50, September 2013
14. Ghiani, L., Hadid, A., Marcialis, G., Roli, F.: Fingerprint liveness detection using binarized statistical image features. In: 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1–6, September 2013
15. Burges, C.J.C.: A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery **2**, 121–167 (1998)
16. Yuan, X.T., Yan, S.: Visual classification with multi-task joint sparse representation. CVPR 2010, 3493–3500, June 2010
17. Schmidt, M., Fung, G., Rosales, R.: Fast optimization methods for l1 regularization: a comparative study and two new approaches. In: Kok, J.N., Koronacki, J., Lopez de Mantaras, R., Matwin, S., Mladenič, D., Skowron, A. (eds.) ECML 2007. LNCS (LNAI), vol. 4701, pp. 286–297. Springer, Heidelberg (2007)
18. Liu, D.C., Nocedal, J.: On the limited memory BFGS method for large scale optimization. Math. Program. **45**(3), 503–528 (1989)
19. Brummer, N.: Fusion of heterogeneous speaker recognition systems in the STBU submission for the NIST speaker recognition evaluation 2006. IEEE Transactions on Audio, Speech, and Language Processing **15**(7), 2072–2084 (2006)
20. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: CVPR 2005, vol. 1, pp. 886–893, June 2005
21. Nanni, L., Lumini, A., Brahnam, S.: Survey on LBP based texture descriptors for image classification. Expert Systems with Applications **39**(3), 3634–3641 (2012)