# Improve the spoofing resistance of multimodal verification with representation-based measures

Zengxi Huang[1], Zhen-Hua Feng[2], Josef Kittler[2], and Yiguang Liu[3, *]

[1] School of Computer and Software Engineering, Xihua University, China
[2] Centre for Vision, Speech and Signal Processing, University of Surrey, UK
[3] College of Computer Science, Sichuan University, China
luomu117@hotmail.com, lygpapers@aliyun.com

**Abstract.** Recently, the security of multimodal verification has become a growing concern since many fusion systems have been known to be easily deceived by partial spoof attacks, i.e. only a subset of modalities is spoofed. In this paper, we verify such a vulnerability and propose to use two representation-based metrics to close this gap. Firstly, we use the collaborative representation fidelity with non-target subjects to measure the affinity of a query sample to the claimed client. We further consider sparse coding as a competing comparison among the client and the non-target subjects, and hence explore two sparsity-based measures for recognition. Last, we select the representation-based measure, and assemble its score and the affinity score of each modality to train a support vector machine classifier. Our experimental results on a chimeric multimodal database with face and ear traits demonstrate that in both regular verification and partial spoof attacks, the proposed method significantly outperforms the well-known fusion methods with conventional measure.

**Keywords:** Multimodal Verification, Spoof Attacks, Representation-based Measure, Support Vector Machine.

## 1 Introduction

A generic biometric system has eight vulnerable points that can be exploited by an intruder to gain unauthorized access [1]. Among them, spoof attacks usually present a counterfeited biometric sample (e.g., a gummy fingerprint, a face image/video/mask) to a system sensor, which do not require knowledge about the system’s operational mechanism and internal parameters. Spoof attacks are also known as non-zero effort attacks, presentation attacks, and direct attacks. The concept of non-zero effort attacks is relative to zero effort attempts, where an imposter doesn’t fabricate the biometric trait of any specific client and merely presents his/her own biometric trait to the system. In the literature, an imposter is generally regarded as an intruder who performs zero effort attempts. In this paper, for clarity and terminological consistence, a legitimate claim, zero effort attempt, and non-zero effort attack are termed as genuine, imposter and spoof, respectively, together with their associated executor/sample/score.

Multimodal systems have been considered intrinsically more secure than unimodal systems based on the intuition that an intruder would have to spoof all the biometric traits to successfully impersonate the targeted client [2]. Such a belief has long been established disregarding the possibility that an intruder is falsely accepted by spoofing only a subset of the biometric traits. The vulnerability of multimodal systems to partial spoof attacks has been shown in the worst-case scenario, where the intruder is assumed to be able to replicate a subset of the biometric traits of a genuine client exactly. Under this assumption, Rodrigues [3] showed experimental results on chimeric multimodal databases with face and fingerprint that multimodal systems can be deceived easily by spoofing only a subset of the modalities, if the fusion rule is not designed with any anti-spoofing measure. Wild et al. [4] showed the sensitivity of multimodal systems to partial spoof attacks with real fake biometric databases.

Some efforts to enhance the security of multimodal systems against partial spoof attacks have already been reported. Rodrigues et al. [5] proposed a modification of the classic likelihood ratio (LLR) method that considers the possibility of spoof attacks and the degree of security to individual trait when modelling score distributions. However, these prior probabilities are application dependent and may not be time invariant, hence are quite difficult to quantify. Rodrigues et al. [3] also proposed the idea of using quality measures to protect against spoof attacks. Intuitively, a fake biometric sample is likely to be of inferior quality. However, biometric quality assessment is still an open issue to most biometrics. Besides, fake biometric sample is not necessarily to be inferior with the emerging image/video synthesis, 3D printing, and materials.

Liveness detection is another kind of approach used to improve the spoofing resistance for a given system. Marasco et al. [6] proposed a multimodal system that incorporates a liveness detection algorithm to reject spoofed samples. If a spoof attempt is indicated, the related modality matching score is ignored. Wild et al. [4] combined the recognition score and liveness measure at score level with a 1-median filtering scheme for enhanced tolerance to spoof attacks. Nevertheless, neither one of hardware-based and software-based liveness detection systems have shown acceptable performance and cost against spoof attacks. Physiological and behavioral characteristics are also employed to enhance multimodal verification security in [7].

This paper is enlightened by the fact that in a partial spoof attack, the recognition scores achieved from non-spoofed modalities are generally near the imposter score distribution center, given that they are also zero effort attempts from a unimodal viewpoint. Unlike the quality- and/or liveness-based methods that focus on the spoofed modalities, we propose to take advantage of non-spoofed modalities. To this end, we put forward a representation-based measure to gauge the affinity of a query sample to a claimed client. This is based on the assumption that a biometric sample would result in inferior sparse representation fidelity if it doesn’t lie in any subspace spanned by the samples from the same subject [8-10]. Note that, it is unlikely to exhaustively collect the representative samples per subject to construct a class specific overcomplete dictionary. We propose to build the dictionary together with samples from non-target subjects to collaboratively represent a query sample.

This affinity score could be an additional measure to a traditional verification method. However, we further consider sparse coding as a one-to-many comparison among the claimed client and non-target subjects, and hence explore other sparsity-based metrics for verification. We evaluate two measures, namely, sparse coding error (SCE) and sparse contribution rate (SCR), on a multimodal database with face and ear. Encouraging performance of SCE-based and SCR-based sum fusion methods evidently supports the usage of sparsity-based one-to-many comparisons in multimodal verification. However, SCR shows much more inferior performance in spoof attacks. Last, we assemble the proposed affinity score and SCE score of each modality as an input vector to train a support vector machine (SVM) classifier.

To validate the effectiveness of the proposed method, we construct a chimeric multimodal database with face and ear traits. The proposed method is compared with the well-known multimodal methods like LLR, SVM, and Sum fusion that are based on cosine similarity. The experimental results validate that in both no spoof and partial spoof cases, the proposed method significantly outperforms its competitors. For example, the traditional methods get the best equal error rates (EER) of 8.32% and 11.89% in no spoof and spoof cases, while our method achieves 0.27% and 2.12%. Apparently, the proposed method helps to increase the spoofing resistance of multimodal systems.

The remainder of the paper is structured as follows. We discuss the approaches to verification based on one-to-many match, and we review the existing methods using sparse coding in Section 2. In Section 3, we present the sparsity-based affinity and recognition measures, together with the proposed multimodal verification system. In Section 4, we describe our chimeric multimodal database and report the corresponding experimental results. The conclusion is drawn in Section 5.

## 2 Related Work

In a biometric verification system, an individual who desires to be recognized claims an identity and presents biometric samples. Then the system conducts a comparison to determine whether the claim is licit or not. Verification is used for positive recognition, where the aim is to prevent multiple people from using the same identity.

Typically, biometric verification systems conduct a one-to-one match that compares a query image against the gallery template(s), whose identity is being claimed. The comparison produces a similarity score. The system accepts the claim if the score is higher than an operating threshold, otherwise rejects it. The operating threshold is determined in the training phase based on the genuine and imposter score distributions. However, it is unlikely to collect all the representative samples of a client that cover all possible variations, for example, expression, pose, illumination, aging, and occlusion in face. Under such circumstances, it cannot be guaranteed that no imposter score is higher than the predefined operating threshold. The system is at a risk of being cracked by intruders. Therefore, the one-to-one match solely based on a predetermined operating threshold is problematic.

Two decades ago, Verlinde et al. [11] proposed a one-to-many match biometric verification method using a k-NN classifier. To the best of our knowledge, this is one of the first attempts to consider non-target subjects for verification in the test phase. Nevertheless, the inferior comparison algorithm like k-NN could probably account for the rare use of one-to-many match in verification. Cohort-based score normalization also takes advantage of non-target subjects but serves the traditional one-to-one match verification [12]. In recent years, we have witnessed the great success of sparse coding techniques in biometric recognition [13-15]. The sparse representation-based classification (SRC) conducts one-to-many comparisons in a sparse coding procedure and is naturally applicable to biometric identification. Note that, along with the initial research of SRC-based face identification in [13], a metric called sparse concentration index (SCI) was applied to reject outliers, i.e. the subjects who do not appear in dictionary.

Inspired by the success of SRC identification and sparsity-based outlier verification, SRC-based comparison has been introduced in speaker verification. In [16], GMM mean supervector is used as feature of an utterance. The $L_1$-norm value of the representation coefficients associated with the claimed identity is used as genuine score, while the $L_1$-norm of the coefficients of each other non-target subject are imposter scores. Based on a similar idea, Li et al. [17] created the dictionary using the total variability i-vectors and evaluated three sparsity-based measures for speaker verification, which achieved better results than a SVM baseline.

# 3 The Proposed Method

## 3.1 Affinity metric

In this section, we present a representation-based measure to gauge the affinity of a query sample to a claimed client, based on the assumption that a biometric sample would result in inferior sparse representation fidelity if it doesn't lie in the subspace spanned by the samples from the same subject [8, 9]. Note that, it is unlikely to exhaustively collect the representative samples per subject to construct a class specific overcomplete dictionary. A feasible way is to use non-target subjects to collaboratively represent the query samples [18].

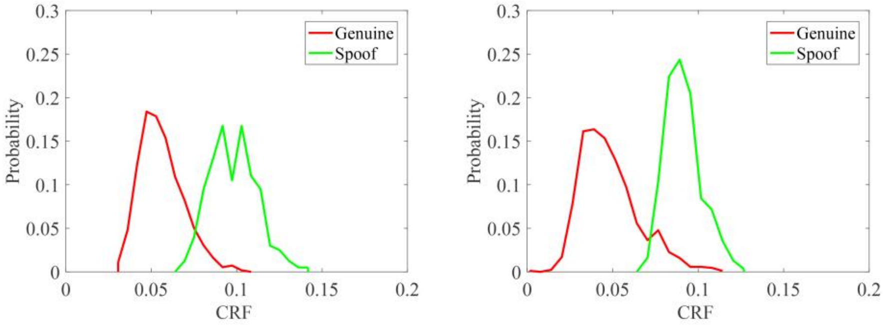
Therefore, we select a number of non-target subjects together with the claimed client. Their gallery samples/features are used to construct an over-complete dictionary $\boldsymbol{A} = [\boldsymbol{A}_c, \boldsymbol{A}_b] \in R^{M\times N}$ $(M << N)$. The first sub-dictionary $\boldsymbol{A}_c = [\boldsymbol{a}_{c,1}, \boldsymbol{a}_{c,2}, \cdots, \boldsymbol{a}_{c,n}] \in R^{M\times n}$ is composed of the gallery samples of the claimed client, which is a dynamic part of the dictionary. The other sub-dictionary $\boldsymbol{A}_b = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \cdots, \boldsymbol{a}_{(N-n)}] \in R^{M\times(N-n)}$ consists of the samples of non-target subjects. Without any specific instructions, $\boldsymbol{A}_b$ is fixed for all identity verification processes. Given a query sample $\boldsymbol{y}$, if it is from a genuine client and isn't of inferior quality, $\boldsymbol{y}$ should lie in a subspace spanned by $\boldsymbol{A}_c$. In this context, $\boldsymbol{y}$ can be sparsely represented by $y = \boldsymbol{A\alpha}$ with high fidelity (see the genuine distribution in Fig. 1), where $\boldsymbol{\alpha} \in R^N$ is the coefficient vector. A sparse solution of $\boldsymbol{\alpha}$ can be obtained by the following optimization problem [13]:

$$\hat{\boldsymbol{\alpha}} = \arg\min \|\boldsymbol{\alpha}\|_1 \text{ s.t. } \|\boldsymbol{y} - \boldsymbol{A\alpha}\|_2 < \varepsilon , \tag{1}$$

where $\|\cdot\|_1$ denotes the $L_1$-norm, and $\varepsilon > 0$ is a positive constant.

In a partial spoof attack, a query sample of non-spoofed modalities is unlikely to lie in any subspace spanned by the dictionary samples given that the non-target subjects are confidential. In this context, only a solution with inferior collaborative representation fidelity (CRF), described in Eq. (2), can be found by optimizing Eq. (1).

$$F\left(\boldsymbol{y}\right)=\left\|\boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\alpha}}\right\|_2 . \tag{2}$$



(a) Face distribution in ear spoof case

(b) Ear distribution in face spoof case

**Fig. 1.** CRF distributions in partial spoof attacks.

Fig.1 shows the CRF distributions on a chimeric multimodal database using face and ear, detailed in Section 4. When the ear of a client is spoofed, the intruder needs to show his/her face or an arbitrary face to complete the biometric data enrollment. Such arbitrary face is unlikely to be from the non-target subjects since the combination of the overcomplete dictionary is confidential. In this context, the non-spoofed face is an outlier that does not lie in the subspace spanned by $\boldsymbol{A}$ and hence leads to an inferior CRF score, see in Fig. 1(a). When the face is spoofed, we see similar CRF distribution of the non-spoofed ears in Fig. 1(b). From the perspective of the client, CRF score can be used to represent the affinity of the query sample to it.

### 3.2 Sparsity-based recognition scores

We consider sparse coding as a competing comparison among the client and non-target subjects, and hence explore other two sparsity-based measures, namely, sparse coding error (SCE) and sparse contribution rate (SCR), for multimodal verification.

Since $\hat{\boldsymbol{\alpha}}$ is achieved in Eq. (1), the SCE value is calculated by

$$E\left(\boldsymbol{y}\right)=\left\|\boldsymbol{y}-\boldsymbol{A}_c\delta_c\left(\hat{\boldsymbol{\alpha}}\right)\right\|_2 , \tag{3}$$

where $\delta_c$ : $R^N \rightarrow R^N$ is the characteristic function that selects the coefficients associated with the claimed client.

The well-known SRC and most of its extensions identify a query sample based on comparing the SCEs of all classes in dictionary. Their superior classification performance validates that SCE is a good candidate to measure the correlation between a query sample and a specific class, as a distance score. Thus, it is reasonable to use SCE for verification.

Wright et al. [13] presented a metric called sparse concentration index (SCI) to reject outliers in face identification. Essentially, the SCI value depends on the class who contributes the most in sparse coding. Given a query sample that isn't an outlier, it generally belongs to the class with the maximal sparse contribution rate (SCR), as defined in Eq. (4). A large value of SCR obtained by a class indicates a greater possibility of the query sample belonging to this class. Therefore, SCR could possibly be used as a similarity score for verification.

$$R(\hat{\boldsymbol{\alpha}}) = \left\|\delta_c(\hat{\boldsymbol{\alpha}})\right\|_1 \Big/ \left\|\hat{\boldsymbol{\alpha}}\right\|_1 . \tag{4}$$

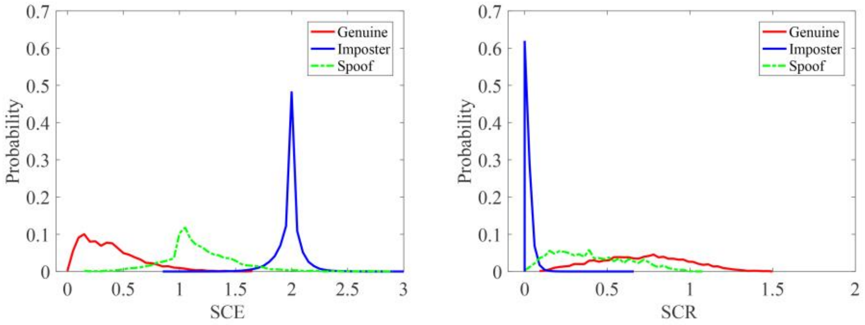


**Fig. 2.** The distributions of SCE and SCR with Sum fusion on our multimodal dataset.

Fig. 2 plots the distributions of SCE and SCR scores obtained on the proposed chimeric multimodal database of face and ear. For convenience to illustrate the effectiveness of SCE and SCR in multimodal verification, we use the Sum rule to fuse face and ear scores. As for SCE, the distribution centers of the genuine and imposter scores are far away from each other with little overlap. Although there is no a clear distribution center peak of the genuine SCR, the overlap is not evident as well. More experimental evidence supporting SCE and SCR is shown in Section 4. In addition, Fig. 2 also demonstrates that most spoof scores are located between the distribution centers of genuine and imposter scores. This implies that the multimodal fusion methods based on SCE or SCR are vulnerable to spoof attacks.

Some variants of SCE and SCR have been used in speaker verification and shown to achieve comparable performance with the traditional one-to-one verification. However, in our face and ear unimodal experiments, a genuine client might lose his/her chance to obtain an eligible SCE or SCR score in the competing comparison, owing to the variations in query samples. If it happens, the genuine score will be extremely low. It means that many licitly claimed clients could not pass the verification system by tuning a client specific operating threshold. Instead, more user cooperation will be necessary, which would degrade the user experience. Therefore, for high accuracy and user convenience of identity verification, sparsity-based one-to-many comparisons would be rather preferable in multimodal scenarios rather than in unimodal applications.

### 3.3 Multimodal Verification

The CRF score that measures the affinity of a query sample to its claimed client can be utilized to enhance the system's resistance to partial spoof attacks in a serial or parallel fusion mode. In a serial fusion mode, multimodal systems firstly examine the CRF scores of each modality to determine whether they are spoofed or not, and then conduct multimodal verification.

However, as shown in Fig. 1, the overlap of the genuine and the spoof CRF score distribution is still rather obvious. A hard CRF threshold would lead to high false acceptance rate (FAR), while a loose one may compromise crack the multimodal system. Furthermore, there is a high possibility that the non-spoofed modalities get inferior recognition scores along with inferior CRF scores from the same sparse coding. The CRF score and sparsity-based recognition score are complementary. Hence, it is worthwhile to combine them in a parallel way to achieve better performance.

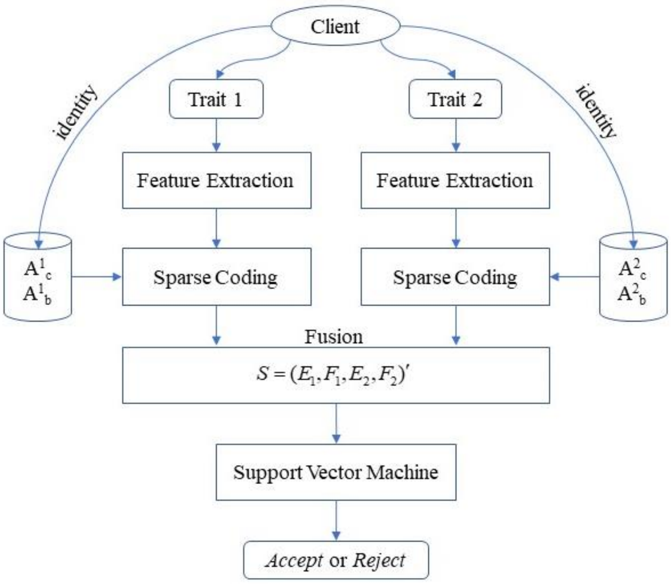


**Fig. 3.** An overview of the multimodal system architecture.

Two sparsity-based recognition scores, i.e., SCE and SCR, are introduced in Section 3.2. Both the Sum fusion methods based on them get promising verification performance in zero effort attempts, as shown by the distributions in Fig. 2. These results support the use of the sparsity-based one-to-many comparison in multimodal systems. On the other hand, SCR is much more inferior to SCE in spoof attacks. The detailed experimental results will be given in Section 4.

Last, we select the SCE and CRF scores of each modality to form a score vector for a verification claim. Suppose there are K modalities, $e_k$ and $f_k$ are the SCE and CRF scores of the k^th^ modality. The final score vector can be denoted by

$S = (e_1, f_1, e_2, f_2, \cdots, e_K, f_K)'$. In the training phase, we use genuine, imposter, and spoof score samples to train a SVM classifier with RBF kernel. For simplicity but without the loss of generality, an overview of system architecture with two modalities (K=2) is shown in Fig. 3 to illustrate the proposed method.

The chimeric multimodal database introduced in Section 4 contains 79 subjects with 7 gallery samples each. All these samples are used to form an overcomplete dictionary with 553 atoms. We don't have abundance data to discuss how to optimally select the non-target subjects in this paper. Note that, we ignore the issue of score normalization, given that the scores of face and ear are compatible in our experiments.

# 4 Experiments and Discussion

## 4.1 Databases

The proposed method is general for verification using multiple biometric traits. In this paper, we construct a chimeric multimodal database with publicly available face and ear databases. All the 79 subjects in USTB III ear database [19] are randomly paired with the first 79 subjects of AR face database [20]. For each subject, the 7 face images without occlusion of Session 1 are used as gallery samples, while the same type of 7 images of Session 2 are used as probe samples. The USTB III is a multi-view ear database with 20 images per subject. We use the same gallery and probe partition rule in [8, 9], where 7 ear gallery images and 13 ear probe images are selected for each subject. In our experiments, the 2 probe images per subject with extreme pose variation are discarded. For each subject on the multimodal database, in the gallery set, 7 face images are uniquely paired with the 7 ear images to form 79×7=553 multimodal samples. In the probe set, each face image is paired with all the ear images to form 79×7×11=6083 multimodal samples.

To simulate the worst-case partial spoof attacks, in a face spoof case, we replace the ear part of a multimodal sample with the image of USTB II ear database (77 subjects, 4 images per subject) [19], In an ear spoof case, we replace the face part with the image of Georgia Tech face database (GT, 50 subjects, 8 images per subject) [21]. Finally, we get 77 subjects, 28 face spoof multimodal samples per subject, and 50 subjects, 88 ear spoof multimodal samples per subject.

In the experiments, we use the features of gallery samples of all 79 subjects to construct the overcomplete dictionary. The SCE, SCR, and CRF scores are derived from the comparison between one-sample and one-set. The numbers of genuine, imposter and spoof score samples are 6083, 474474 (6083×78), and 6556, respectively. As for the competing methods using cosine similarity, we empirically select the best match score from each comparison, hence their score sample numbers are the same.

## 4.2 Settings

The 2D-DCT method is applied for feature extraction of face and ear images, since it is fast, general, and without specific training. The DCT coefficients are scanned in a zigzag manner starting from the top-left corner of the entire transformed image to form a feature vector with 200 dimensions.

The proposed multimodal method uses SVM with RBF kernel (sigma=0.25). It is compared with the Sum fusion methods of SCE and SCR, denoted by SUM(sce) and SUM(scr), respectively. The competing multimodal methods include the well-known LLR [22], SVM [23], and Sum fusion methods, which use cosine similarity and are respectively denoted by LLR(cos), SVM(cos), and SUM(cos). SVM(cos) also uses RBF kernel (sigma=1).

Without specific instructions, half of the genuine, imposter and spoof scores are randomly selected for training, and the remainder are for testing. To alleviate the imbalance of training samples, SVM-based classifiers use 1/10 imposters to train. The LLR(cos) uses half of all kinds of samples to fit Gaussian mixture models for score distribution estimation. We run all experiments 5 times, the results presented here are based on the average from these 5 runs.

### 4.3 Results

The metrics like false acceptance rate (FAR), false rejection rate (FRR), equal error rate (EER), and the receiver operating characteristic (ROC) curves are generally used to evaluate methods in regular verification. The spoof FAR (SFAR) is specifically used to note the FAR in spoof attacks.

In the first part of the experiments, we train all the learning-based classifiers without considering the spoof samples, namely no spoof training. Fig. 4 plots the ROC curves of all competing methods in regular verification. The methods with sparsity-based metrics are observed to be significantly better than the methods with traditional metric. Among the former methods, SUM(scr) is obviously inferior to SUM(sce) and the proposed method. The ROC curves and the EERs summarized in Table 1 do not show evident advantage of our method when compared with SUM(sce).

**Table 1.** Performance in terms of EER (%).

| Training | Testing | SUM(cos) | SVM(cos) | LLR(cos) | SUM(sce) | SUM(scr) | Ours |
|---|---|---|---|---|---|---|---|
| Genuine<br>Imposter | Regular | 11.83 | 6.632 | 6.85 | 0.20 | 0.39 | 0.18 |
| | Spoof attacks | 12.44 | 22.05 | 21.04 | 8.73 | 28.26 | 4.13 |
| Genuine<br>Imposter<br>Spoof | Regular | 11.83 | 8.79 | 8.32 | 0.20 | 0.39 | 0.27 |
| | Spoof attacks | 12.44 | 11.89 | 12 | 8.73 | 28.26 | 2.12 |

Fig.5(a) demonstrates that all these methods without spoof training are vulnerable to partial spoof attacks. Both the EERs of LLR(cos) and SVM(cos) increase by about 15%, and even that of SUM(scr) soars to 28.26%. On the other hand, our method achieves a 4.13% EER, which is less than half of the second best.

In the second part of the experiments, all the learning-based classifiers are trained with genuine, imposter and spoof samples, namely spoof training. We can see from Table 1 that, compared with the former experiments of spoof attacks, both LLR(cos) and SVM(cos) get about 10% improvements, while the EER of ours reduces by half, down to 2.12%. The overwhelming advantage of our method can be seen vividly with the ROC curves plotted in Fig. 5(b). It is quite promising provided that the experiments here are in the worst-case spoof conditions where the fake score distribution of the spoofed modalities is identical to that of genuine.

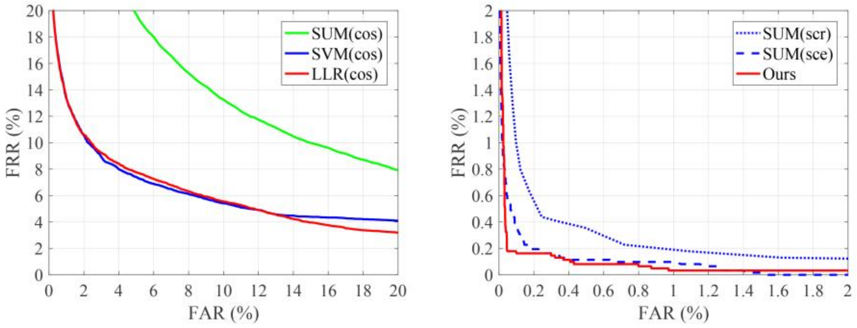
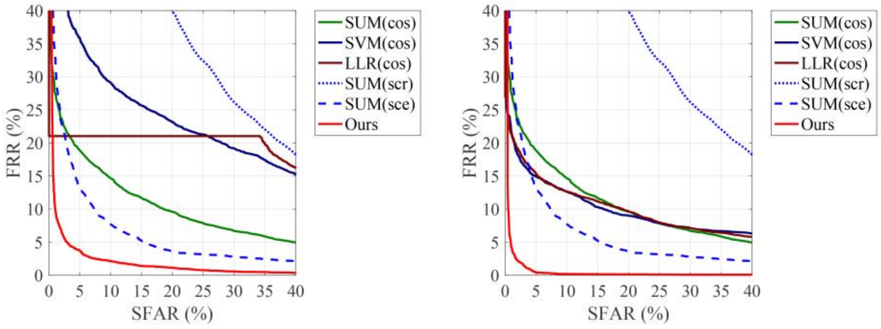


**Fig. 4.** Performance in regular verification.



(a) No spoof training.

(b) Spoof training.

**Fig. 5.** Performance in partial spoof attacks.

Although LLR(cos) and SVM(cos) also exhibit obvious improvements, they encounter obvious accuracy decline in regular verification, see Table 1. These results show again that the spoof training may bring about unacceptable performance degradation in regular identity verification [2]. As for the proposed method, the EER increases from 0.18% to 0.27%, which is still very low. Above all, the proposed method is able to achieve very low EER in both regular verification and partial spoof attacks.

## 5 Conclusion

In this paper, aiming to improve the multimodal system's resistance to partial spoof attacks, we proposed the use of collaborative representation fidelity with non-target subjects to measure the affinity of a query sample to a claimed client. We further considered sparse coding as a competing comparison among the claimed client and non-target subjects, and hence explored two sparsity-based measures associated with individual subjects for recognition. The encouraging performance evidently supports the use of sparsity-based one-to-many comparisons in multimodal systems. However, based on their performance in spoof attacks, only the representation-based one is selected as recognition score. Last, two types of representation-based scores for each modality are assembled to train a SVM classifier.

The proposed method was compared with well-known multimodal methods like LLR, SVM, and Sum fusion methods, using the cosine similarity measure, on a chimeric multimodal database of face and ear traits. The experimental results demonstrate that in both regular verification and partial spoof attacks, the proposed method overwhelmingly outperforms its competitors. The proposed method is a general model for combining multiple biometric traits. In the future work, we plan to evaluated with more biometric traits like palmprint, iris, and with real spoofed data. We believe the method can be further enhanced by using more robust feature extraction method like CNN-based, and advanced multimodal joint sparse coding techniques [24].

## Acknowledgements

This work was partly supported by the National Natural Science Foundation of China (61602390, 61860206007, 61532009, 61571313, 61605054), the EPSRC Programme Grant (FACER2VM) EP/N007743/1, EPSRC/dstl/MURI project EP/R018456/1, Chinese Ministry of Education (Z2015101), Department of Science and Technology of Sichuan Province (2017RZ0009, 2017FZ0029, and 18GJHZ0138), and by funding under 2016CDLZ-G02-SCU from Sichuan University and Lu-Zhou city.

## References

1. Ratha N. K., Connell J. H., Bolle R. M.: An analysis of minutiae matching strength. In: International Conference on Audio-and Video-Based Biometric Person Authentication, pp. 223-228. Springer, Berlin (2001)
2. Biggio, B., Fumera, G., Marcialis, G. L., & Roli, F.: Statistical meta-analysis of presentation attacks for secure multibiometric systems. IEEE transactions on pattern analysis and machine intelligence 39(3), 561-575 (2017)
3. Rodrigues R. N., Ling L. L., Govindaraju V.: Robustness of multimodal biometric fusion methods against spoof attacks. Journal of Visual Languages & Computing 20(3), 169-179 (2009)
4. Wild P., Radu P., Chen L., et al.: Robust multimodal face and fingerprint fusion in the presence of spoofing attacks. Pattern Recognition 50, 17-25 (2016)
5. Rodrigues R. N., Kamat N., Govindaraju V.: Evaluation of biometric spoofing in a multimodal system. In: IEEE International Conference on Biometrics: Theory Applications & Systems, pp. 1-5 (2010)

6. Marasco, E., Johnson, P., Sansone, C., & Schuckers, S.: Increase the security of multibiometric systems by incorporating a spoofing detection algorithm in the fusion mechanism. In: International Workshop on Multiple Classifier Systems, pp. 309-318. Springer (2011)
7. Bhardwaj I., Londhe N. D., Kopparapu S. K.: A spoof resistant multibiometric system based on the physiological and behavioral characteristics of fingerprint. Pattern Recognition 62, 214-224 (2017)
8. Huang Z., Liu Y., Li C., et al.: A robust face and ear based multimodal biometric system using sparse representation. Pattern Recognition 46(8), 2156-2168 (2013)
9. Huang Z., Liu Y., Li C., et al.: An adaptive bimodal recognition framework using sparse coding for face and ear. Pattern Recognition Letters 53, 69-76 (2015)
10. Song, X., Feng, Z.H., Hu, G., Kittler, J. and Wu, X.J.: Dictionary integration using 3D morphable face models for pose-invariant collaborative-representation-based classification. IEEE Transactions on Information Forensics and Security, 13(11), 2734-2745 (2018)
11. Verlinde P., Cholet G.: Comparing decision fusion paradigms using k-NN based classifiers, decision trees and logistic regression in a multi-modal identity verification application. In: AVBPA, 188-193 (1999)
12. Merati, A., Poh, N., & Kittler, J.: User-specific cohort selection and score normalization for biometric systems. IEEE Transactions on Information Forensics and Security 7(4), 1270-1277 (2012)
13. Wright J., Yang A.Y., Ganesh A. et al.: Robust face recognition via sparse representation. IEEE transactions on pattern analysis and machine intelligence 31(2), 210-227 (2009)
14. Cheng H., Liu Z., Yang L., Chen X.: Sparse representation and learning in visual recognition: Theory and applications. Signal Processing 93(6), 1408-1425 (2013)
15. Shao, C., Song, X., Feng, Z.H., Wu, X.J. and Zheng, Y.: Dynamic dictionary optimization for sparse-representation-based face classification using local difference images. Information Sciences, 393, 1-14 (2017)
16. Kua J., Ambikairajah E., Epps J., Togneri R.: Speaker verification using sparse representation classification. In: IEEE ICASSP, pp. 4548-4551. Prague, Czech Republic (2011)
17. Li M., Zhang X., Yan Y., Narayanan S.: Speaker verification using sparse representations on total variability i-vectors. In: 12th Annual Conference of the International Speech Communication Association, pp. 2729-2732. Florence, Italy (2011)
18. Zhang L, Yang M, Feng X.: Sparse representation or collaborative representation: Which helps face recognition? In: ICCV, Barcelona, Spain, pp. 471-478 (2011)
19. University of Science & Technology Beijing (USTB), http://www1.ustb.edu.cn/resb/, last, accessed Jan. 2016
20. Martinez A. M., Benavente R.: The AR Face Database. CVC Technical Report 24 (1998).
21. Georgia Tech Face Database, http://www.anefian.com/research/face_reco.htm, last, accessed June 2016
22. Figueiredo T., Jain A. K.: Unsupervised learning of finite mixture models. IEEE transactions on pattern analysis and machine intelligence 24(3), 381-396 (2002)
23. Liu, Y., You, Z., Cao, L.: A novel and quick SVM-based multi-class classifier. Pattern Recognition 39(11), 2258-2264 (2006)
24. Yuan X. T., Liu X., & Yan, S.: Visual classification with multitask joint sparse representation. IEEE Transactions on Image Processing, 21(10), 4349-4360 (2012).