Distributed Computing and Artificial Intelligence: 9th International Conference. Advances in Intelligent and Soft Computing, Volumen 151. Springer, 2012. 129-139

DOI: http://dx.doi.org/10.1007/978-3-642-28765-7_17

Variability Compensation using NAP for Unconstrained Face Recognition

Pedro Tome, Ruben Vera-Rodriguez, Julian Fierrez and Javier Ortega-García

Abstract The variability presented in unconstrained environments represents one of the open challenges in automated face recognition systems. Several techniques have been proposed in the literature to cope with this problem, most of them tailored to compensate one specific source of variability, e.g., illumination or pose. In this paper we present a general variability compensation scheme based on the Nuisance Attribute Projection (NAP) that can be applied to compensate for any kind of variability factors that affects the face recognition performance. Our technique reduces the intra-class variability by finding a low dimensional variability subspace. This approach is assessed on a database from the NIST still face recognition challenge "The Good, the Bad, and the Ugly" (GBU). The results achieved using our implementation of a state-of-the-art system based on sparse representation are improved significantly by incorporating our variability compensation technique. These results are also compared to the GBU challenge results, highlighting the benefits of adequate variability compensation schemes in these kind of uncontrolled environments.

1 Introduction

Most biometric technologies are able to provide satisfactory matching performance in controlled situations where the user is cooperative and data acquisition conditions and environment can be controlled. However, in many applications, biometric data is acquired in less than ideal conditions, such as uncontrolled and unconstrained face recognition scenarios [5]. The low performance of biometrics technologies in

Pedro Tome · Ruben Vera-Rodriguez · Julian Fierrez · Javier Ortega-Garcia
Biometric Recognition Group - ATVS, Escuela Politecnica Superior
Universidad Autonoma de Madrid
Avda. Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain
e-mail: {pedro.tome, ruben.vera, julian.fierrez, javier.ortega}@uam.es

these relatively uncontrolled situations has limited their deployment, therefore, a significant improvement in recognition performance in less controlled situations is one of the main challenge facing biometric technologies.

In the particular case of face recognition in uncontrolled scenarios there are numerous sources of variation, which can be known or unknown, affecting the performance. Hence, there is a need for developing methods capable of identifying and compensating/removing these variability sources in order to guarantee the robustness of the system in unconstrained and uncontrolled real environments.

In the present paper, a variability compensation approach based on Nuisance Attribute Projection (NAP) is presented for face recognition. In this field, to our knowledge only V. Štruc *et al.* in [13] have analysed such a normalization technique for illumination invariant face recognition based on NAP, which removes the illumination induced artifacts in two controlled scenarios. In our case, the proposed NAP compensation approach is used not only to compensate illumination variations, but also other variability factors. In particular, we study the uncontrolled scenario provided by the NIST - GBU still face recognition challenge, which consists of three partitions called the Good, the Bad, and the Ugly [6].

The performance of the proposed variability compensation scheme is evaluated on a state-of-the-art system based on sparse representation [14]. Results achieved show that variability compensation using NAP in combination with this system is a very interesting approach in uncontrolled face recognition environments.

The paper is structured as follows. Sect. 2 briefly describes the variability factors found in the GBU challenge. Sect. 3 describes the variability compensation approach using NAP. Sect. 3 describes the recognition system based on sparse representation. Sect. 5 presents the experimental protocol followed and the performance evaluation. Sect. 6 presents the experimental results obtained, and finally conclusions are drawn in Sect. 7.

2 Variability in Unconstrained Environments

Face recognition in unconstrained environments is a very challenging problem which has attracted increasing attention from the research community.

Some the recent studies in this field are the Multiple Biometric Grand Challenge (MBGC 2009) [7] and the Face Recognition Vendor Test (FRVT 2006) [8], whose focus of research is shifting to recognizing faces taken under less constrained conditions. As a result of the evolution of this NIST challenges a new competition called GBU has been defined, which consists of three partitions called the Good, the Bad, and the Ugly. The Good partition consists of pairs of face images of the same person that are easy to match (based on FRVT 2006 top performers); the Bad partition contains pairs of face images of a person that have average matching difficulty; and the Ugly partition concentrates on difficult to match face pairs. Fig. 1 shows an example of these three partitions and their respective histograms of match and non-match scores.

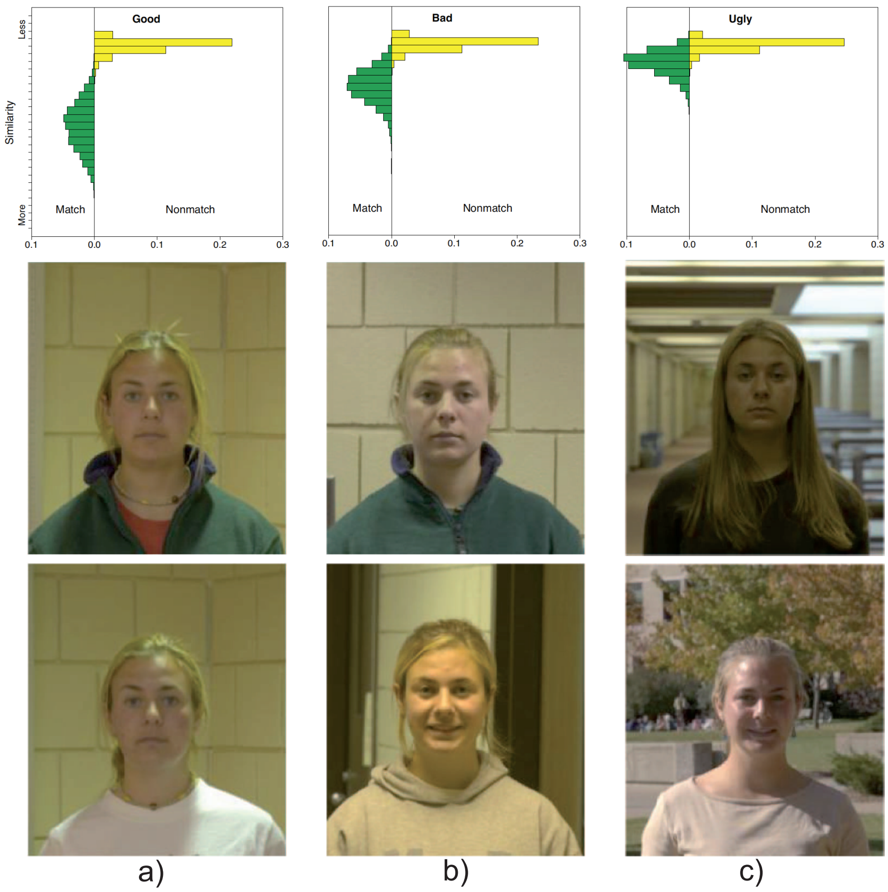

Fig. 1 GBU image samples and histograms of match and non-match distributions for the *a)* Good, *b)* the Bad, and *c)* the Ugly partitions with the relative frequency of similarity scores in horizontal axes. Extracted from [6].

Various techniques have been presented in the literature to compensate the variability present in these kind of scenarios [11, 12, 4, 15]. However, most of these techniques are focused on an isolated variability source, e.g., illumination, pose compensation, etc.

In the present paper, a variability compensation approach is presented, using the Nuisance Attribute Projection (NAP) to remove the variability induced in uncontrolled face recognition systems.

3 NAP for Variability Compensation

3.1 Nuisance Attribute Projection (NAP)

Nuisance Attribute Projection (NAP) is a powerful technique traditionally used in the field of speaker recognition for compensation of channel effects regardless of its source [10, 9], which are assumed to lie in a low dimensional variability subspace. In others fields like biometrics at a distance and unconstrained environments, the variability sources are mostly unknown and mixed, hence, we seek to understand to what extent variability compensation techniques as NAP are useful.

Consider a dataset X of n image vectors of size N pixels, where $X \in \Re^{n \times N}$. The NAP technique tries to remove any unwanted distortion in the images as follows:

$$X' = P(X - M), \tag{1}$$

where X' denotes the new data whose component in the variability subspace is removed, M denotes a matrix containing in each of its columns the global mean of the images in X and P stands for the $n \times n$ projection matrix:

$$P = I - VV^T = I - \sum_{i=1}^{d} v_i v_i^T \,. \tag{2}$$

Here, I denotes the $n \times n$ identity matrix, v_i represents the i^{th} direction of the variability subspace base V of size d defined by NAP.

Suppose a data matrix X has n_{C_j} sample images from the j^{th} class, whose labels of the classes are $C_1, C_2, ..., C_r$, then, for each of these images we can write:

$$x_{C_{j,k}} = x'_{C_{j,k}} + e_k, \tag{3}$$

where C_j represents the class label of the image, k denotes the index of the image in the j^{th} class, $x'_{C_{j,k}}$ stands the variability-free part of $x_{C_{j,k}}$, and e_k represents the vector encoding the variability effects for the k^{th} image of the j^{th} class. Fig. 2*a)*, describes graphically the procedure.

Assuming the unwanted variability effects inside each class coincide and they can be modelled by a Gaussian distribution, then, the base of the variability subspace defined by the matrix V, can be estimated from the first d eigenvectors (NAP directions $v_i (i = 1, 2, ..., d)$) of the matrix Σ_t:

$$\Sigma_t = \sum_{j=1}^{r} \sum_{k=1}^{n_{C_j}} (X_{C_{j,k}} - \mu_{C_j})(X_{C_{j,k}} - \mu_{C_j})^T, \tag{4}$$

where the mean value of each of the r classes μ_{C_j} $(j = 1, 2, ..., r)$ represents a variability-free estimate of an image from the j class. This is typically done by using Principal Component Analysis (PCA). Fig. 2*b)* shows the eigenvalues of the associated eigenvectors for NAP estimated variability subspace.

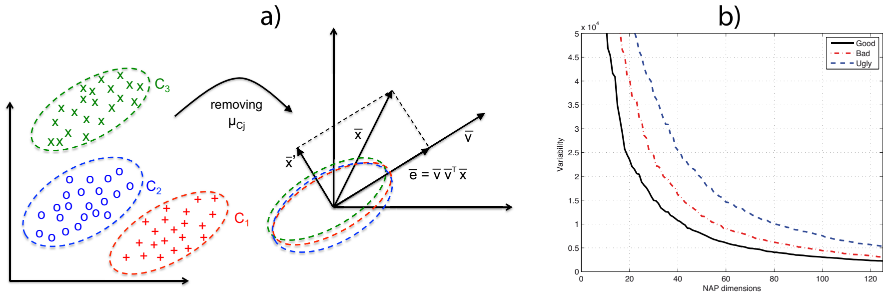

Fig. 2 *a)* Schematic illustration of the NAP technique in a 2-dimensional space. Firstly, every distribution is centred in the origin by removing the global mean of each class μ_{C_j}. The input data- and variability-dependent sample vector ($\bar{x}$) is pair up into two components ($\bar{x}'$) and ($\bar{e}$). ($\bar{e}$) stands for the component in the variability subspace, and ($\bar{x}'$) is the resulting compensated sample vector. The vector $\bar{v}$ represents the first eigenvector of the estimated variability subspace. *b) Eigenvalues of three target datasets: Good, Bad, Ugly.*

3.2 Removing Variability Effects

In the case considered in this paper, we have separated the channels YCbCr of the images and the NAP compensation scheme has been applied over the luminance (Y) component of images.

Consider an input data set in X from which we estimate the NAP directions corresponding to the unknown variability in the unconstrained scenarios. In the GBU database considered only four factors are controlled: subject aging, pose, change in camera, and variations among faces. Other factors, such as: illumination, indoor/outdoor, distance, ... are considered as unwanted variability factors. These factors do not always affect in the same level, making the problem of their compensation even more challenging. Any input image x is compensated with respect to the estimated variability effects by projecting away a number of directions in the NAP subspace. Fig. 2*a)* illustrates graphically the procedure. The compensation procedure is described by:

$$x' = P(x - \mu) = (I - VV^T)(x - \mu), \tag{5}$$

where μ represents the global mean of the images in X, I denotes the identity matrix and V stands for the NAP compensation matrix. To effectively remove the effects of annoying variability, the data matrix X must be constructed in such a way as to include the highest available number of images captured in different variability conditions.

4 Face Verification System - SRC

A system based on recent works in sparse representation for classification purposes (SRC) [2, 14] has been adopted as face verification core.

Essentially, this kind of systems span a face subspace using all known training face images, and for an unknown face image they try to reconstruct the image sparsely.

The motivation of this model is that given sufficient training samples of each person, any new test sample for this same person will approximately lie in the linear span of the training samples associated with the person.

Once a new test image y is acquired, it can be represented using samples from the database by the linear equation $y = Ax_0$, where matrix A defines our training data and x_0 represents the sparse solution.

According to the assumption that images from a given subject are sufficient to represent themselves, the solution x_0 in the linear equation $y = Ax_0$ should be very sparse. This can be approximately recovered by solving the following noise-aware l^1-minimization problem:

$$\widehat{x}_1 = \operatorname{argmin}_x \|x\|_1 \ \ subject\ to \ \ \|Ax - y\|_2 \leq \varepsilon. \tag{6}$$

To recognize a probe test image, the SRC algorithm identifies the class by computing the minimum among the residuals reconstructed per class. The robust performance of the SRC algorithm has been proved experimentally on face datasets with noises and occlusions.

The solution of equation (6), was approximated, in an efficient way, via basis pursuit using linear programming by considering L1-norm instead of L0-norm. To this end, the available package provided in [1] was used.

5 Experimental Protocol

The experiments are carried out on the The Good, the Bad, and the Ugly (GBU) database [6] included in the last still face recognition challenge from NIST - National Institute of Standard and Technology. The GBU challenge problem consists of three partitions with are called the Good (face pairs easy to match), the Bad (face pairs with average matching difficult), and the Ugly (face pairs difficult to match). Each partition consists of two sets of images, a target set and a query set, each of which contains 1,085 images from 437 distinct subjects. The distribution of image counts per person in the target and query sets are 117 subjects with 1 image; 122 subjects with 2 images; 68 subjects with 3 images; and 130 subjects with 4 images.

For the experiments in this paper, we use the segmented datasets provided by MBGC - Multiple Biometric Grand Challenge [7] compressed to 20KB with 120 pixels between the centers of the eyes. The faces were normalized following the ISO norm described in [3], from a size of 408×528 to size 168×192 pixels.

The baseline system consists on the application of the SRC algorithm with a single preprocessing stage to normalize the face illumination by histogram equalization (HQ) over the band of luminance (Y) from YCbCr color space. In the experiments described here, we have used as features the downsampled images, whose good performance combined with SRC is demonstrated in [14]. In our case the downsampling ratio is $1/8$ obtaining feature vectors of 504 dimensions.

The performance of the evaluated system is computed not using the same experimental protocol described by the GBU challenge. We also use a one-to-one matching, but using prior information of the target sets in order to compensate the variability.

In the experimental protocol we consider two experiments. In the first experiment, a NAP compensation matrix is generated for each partition of the database using only the target images. In the second experiment, a global NAP compensation matrix NAPgbu is generated for the three partitions together using all the target images. In both cases we evaluate the performance of our recognition system using two different NAP dimensions of variability (d), low compensation $d = 5$ and high compensation, $d = 125$.

6 Results

The experiments have two different goals, namely: *i*) study the benefits of variability compensation schemes in uncontrolled environments, and *ii*) show the efficiency of NAP-based variability compensation when considering multiple uncontrolled sources of variability.

Our baseline algorithm based on sparse representation achieved better results than those obtained from the LRPCA-face GBU baseline algorithm [6] at a false accept rate (FAR = 0.001). On the Good partition, the base verification rate (VR) is 0.88, for the Bad partition, the VR in 0.29, and the VR in Ugly is 0.06. Table 1 shows the comparative results.

6.1 Experiment 1: NAP over each partition

The performance of the NAP compensation scheme is first analysed scenario by scenario. Results achieved for the *Good* partition are represented in Fig. 3*a*. In this case the compensation of few dimensions ($d = 5$) is much better than using many dimensions ($d = 125$). This is due to the fact that data are more or less clear of unwanted variability so the compensation of many dimensions leads to a discriminative information loss. As can be seen in Fig. 2*b)*, the eigenvalues of the *Good* partition decrease faster than for the other partitions, meaning that the variability is concentrated in the first dimensions.

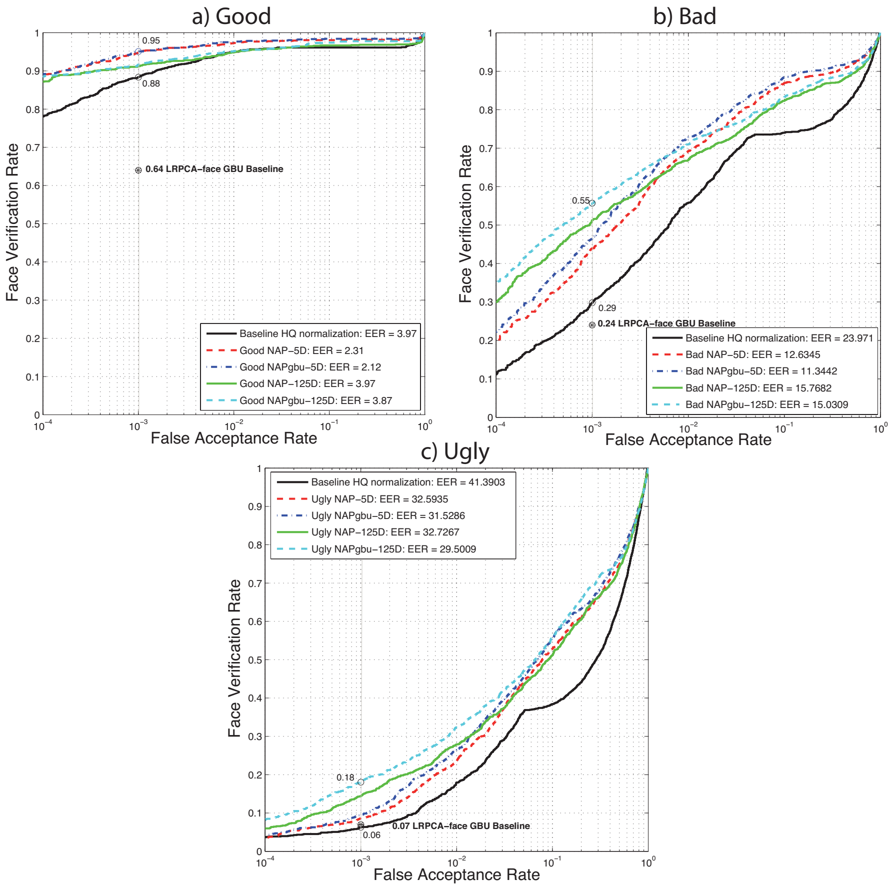

Fig. 3 ROC curves obtained for the three partitions: *a) Good* partition, *b) Bad* partition and *c) Ugly* partition. The verification rate for the LRPCA-face GBU baseline [6], our baseline system and the best NAP solution are highlighted at a FAR = 0.001.

Results achieved for the *Bad* partition are shown in Fig. 3*b*. In this case the variability increases, implying that there are more corrupted dimensions with variability. Therefore, in this case better results are obtained for the case of compensating more dimensions ($d = 125$) with NAP working at a FAR = 0.001. On the other hand, the EER of the system is better for the case of compensating less dimensions ($d = 5$), but in this case working at a much more permissive application FAR.

Results achieved for the *Ugly* scenario are shown in Fig. 3*c*. In this case, as in the *Bad* one, better results are obtained when ($d = 125$) dimensions are compensated. Here, the verification rate (VR) of the baseline system versus the compensated system with $d = 125$ improves from 0.06 to 0.14 at a FAR = 0.001. Table 1 summarises all the results achieved for this experiment.

Partition	LRPCA face [6]	Baseline	Best NAP Comp.	Relative Improvement (%)
Good	0.64	0.88	0.94	6.8
Bad	0.24	0.29	0.51	91.6
Ugly	0.07	0.06	0.14	133.3

Table 1 Results achieved in Experiment 1. Performance of the LRPCA-face [6] baseline system versus our baseline and best NAP compensation results, being respectively $d = 5, 125, 125$ for Good, Bad and Ugly partitions. Also, relative improvement in the verification rate reached by NAP compensation are highlighted at a FAR = 0.001.

6.2 *Experiment 2: NAP over the whole partitions*

As mentioned before, in this experiment we generate a global NAP compensation matrix (NAPgbu) combining the three target datasets (Good, Bad, and Ugly) in order to demonstrate the potential of the proposed NAP approach on unconstrained environments. The main results are summarized in Table 2.

The difference between low and high compensation ($d = 5$ and $d = 125$) respectively, have the same behaviour than experiment 1 over the three datasets, as we can see in Fig. 3.

NAP compensation removes the intra-class variability by projecting away multiple dimensions of a low variability subspace. For this reason a subspace calculated using all possible target data is likely to improve the effect of the variability compensation. This is proved by observing who the NAPgbu compensation scheme achieves the best results over all scenarios (see Figs. 3 and 4). Note that in the *Good* partition both proposed schemes produce the same results due to the low influence of the variability source in this case.

As can be seen in Fig. 4 the *Bad* partition achieves the highest absolute improvement of VR going from 0.29 to 0.55 (at FAR = 0.001). This is possibly due to the fact that images contain a large amount of variability but still is possible to achieve reasonably good results with compensation.

Partition	LRPCA face [6]	Baseline	Best NAPgbu Comp.	Relative Improvement (%)
Good	0.64	0.88	0.95	7.9
Bad	0.24	0.29	0.55	89.6
Ugly	0.07	0.06	0.18	200

Table 2 Results achieved in Experiment 2. Performance of the LRPCA-face [6] baseline system versus our baseline and best NAPgbu compensation results, being respectively $d = 5, 125, 125$ for Good, Bad and Ugly partitions. Also, relative improvement in the verification rate reached by NAPgbu compensation are highlighted at a FAR = 0.001.

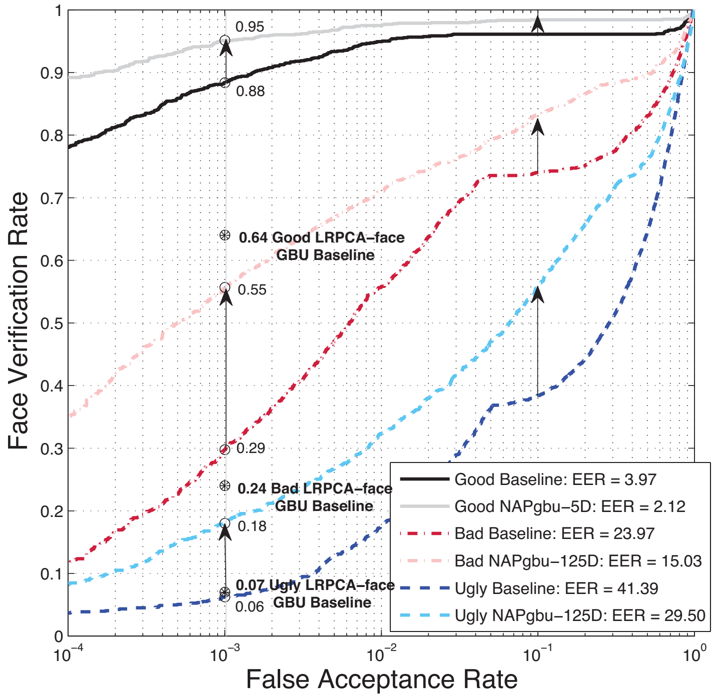

Fig. 4 ROC for the best NAP solutions vs. baseline systems on the three GBU partitions: Good, Bad and ugly. LRPCA-face GBU baseline [6] is also highlighted at a FAR = 0.001.

Finally, in the *Ugly* partition, the balance of discriminative information against noise is very low. Fig. 4 shows how the VR improves from 0.06 to 0.18 at FAR = 0.001, reaching better results than those presented in [6]. As can be seen in Table 2, the relative improvement of the verification rate in the *Ugly* partition in this experiment is higher (200%) than in others partitions.

7 Conclusions

In the present work, a variability compensation approach based on Nuisance Attribute Projection has been presented and used to improve a state-of-the-art face recognition system based on sparse representation. The efficiency of this approach has been studied considering the three different challenge partitions designed by NIST for the still face recognition challenge "The Good, the Bad, and the Ugly" (GBU). In all cases, the baseline system performance is higher than the one achieved in the baseline algorithms from GBU challenge [6]. Furthermore, when the proposed compensation variability approach based on NAP is applied, the system performance improves significantly.

The application of NAP compensation using the whole partitions in a combined form is also analysed, highlighting the benefits of adequate variability compensation schemes in these kind of uncontrolled environments.

Acknowledgements Pedro Tome is supported by a FPU Fellowship from Univ. Autonoma de Madrid and Ruben Vera-Rodriguez is supported by a Juan de la Cierva Fellowship from Ministerio de Ciencia e Innovacion. This work has been partially supported by projects Bio-Challenge (TEC2009-11186), Contexts (S2009/TIC-1485), TeraSense (CSD2008-00068) and "Cátedra UAM-Telefónica".

References

1. E. Candes and J. Romberg. l1-magic: Recovery of sparse signals via convex programming. In *http://www.acm.caltech.edu/l1magic/*, 2005.
2. K. Huang and S. Aviyente. Sparse representation for signal classification. pages 609–616, 2006.
3. ISO/IEC JTC 1/SC 37 N 504. Biometric data interchange formats part 5: Face image. 2004.
4. S. Z. Li, R. Chu, S. Liao, and L. Zhang. Illumination invariant face recognition using near-infrared images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(4):627–639, 2007.
5. S. Z. Li, B. Schouten, and M. Tistarelli. *Handbook of Remote Biometrics for Surveillance and Security*, chapter Biometrics at a Distance: Issues, Challenges, and Prospects, pages 3–21. Springer, 2009.
6. P. J. Phillips, J. R. Beveridge, B. A. Draper, G. Givens, A. J. O'Toole, D. S. Bolme, J. Dunlop, Y. M. Lui, H. Sahibzada, and S. Weimer. An introduction to the good, the bad, amp; the ugly face recognition challenge problem. In *Int. Conf. on Automatic Face Gesture Recognition and Workshops (FG 2011)*, pages 346 –353, march 2011.
7. P. J. Phillips, P. J. Flynn, J. R. Beveridge, W. T. Scruggs, A. J. O'Toole, B. David, K. W. Bowyer, B. A. Draper, G. H. Givens, Y. M. Lui, H. Sahibzada, J. A. Scallan, Iii, and S. Weimer. Overview of the multiple biometrics grand challenge. In *Proceedings of the Third International Conference on Advances in Biometrics*, ICB '09, pages 705–714, Berlin, 2009.
8. P. J. Phillips, W. T. Scruggs, A. J. O'Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott, and M. Sharpe. FRVT 2006 and ICE 2006 large-scale experimental results. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99, 2009.
9. A. Solomonoff, W. M. Campbell, and C. Quillen. Nuisance attribute projection. *Speech Communication, Elsevier Science BV*, 2007. Amsterdam, The Netherlands.
10. A. Solomonoff, C. Quillen, and W. M. Campbell. Channel compensation for svm speaker recognition. In *In Proceedings on Odyssey: The Speaker and Language Recognition Workshop*, pages 41–44, Toledo, Spain, 2004.
11. P. Tome, J. Fierrez, F. Alonso-Fernandez, and J. Ortega-Garcia. Scenario-based score fusion for face recognition at a distance. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 67 – 73, June 2010.
12. P. Tome, J. Fierrez, M. C. Fairhurst, and J. Ortega-Garcia. Acquisition scenario analysis for face recognition at a distance. In *6th International Symposium on Visual Computing (ISVC)*, November-December 2010.
13. V. Štruc, B. Vesnicer, F. Mihelič, and N. Pavešić. Removing illumination artifacts from face images using the nuisance attribute projection. In *Proceedings of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'10)*, pages 846–849, Dallas, Texas, USA, March 2010.
14. J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(2):210–227, 2 2009.
15. X. Zhang and Y. Gao. Face recognition across pose: A review. *Pattern Recognition*, pages 2876–2896, 2009.