A Sparse Representation Approach to Face Matching across Plastic Surgery

Gaurav Aggarwal, Soma Biswas, Patrick J. Flynn and Kevin W. Bowyer
Department of Computer Science and Engineering
University of Notre Dame, Notre Dame
{gaggarwa, sbiswas, flynn, kwb}@nd.edu

Abstract

Plastic surgery procedures can significantly alter facial appearance, thereby posing a serious challenge even to the state-of-the-art face matching algorithms. In this paper, we propose a novel approach to address the challenges involved in automatic matching of faces across plastic surgery variations. In the proposed formulation, partwise facial characterization is combined with the recently popular sparse representation approach to address these challenges. The sparse representation approach requires several images per subject in the gallery to function effectively which is often not available in several use-cases, as in the problem we address in this work. The proposed formulation utilizes images from sequestered non-gallery subjects with similar local facial characteristics to fulfill this requirement. Extensive experiments conducted on a recently introduced plastic surgery database [17] consisting of 900 *subjects highlight the effectiveness of the proposed approach.*

1. Introduction

Traditionally, research in the area of face recognition has been concentrated on matching faces across changes in illumination and pose [21]. The tremendous improvement in capability of automatic algorithms and commercial systems to address these challenges [12][15] has led to interest in other exciting and challenging avenues like matching faces in low resolution [5], face matching in newborns [4], etc. Face matching across plastic surgery variations is another such challenging avenue that is starting to interest researchers [17]. Technological advancements have led to availability of speedy and affordable plastic surgery procedures. Social pressure and decreased costs are driving more and more people to undergo such procedures to correct feature defects and remove birth marks [1]. As a result, recognizing individuals across plastic surgery procedures is becoming a pertinent problem for law-enforcement agencies [17].

Figure 1. A few sample images from the plastic surgery database [17] to illustrate the difficulty of the task. Top row: Pre-surgery images; Bottom row: Corresponding post-surgery images. In addition to changes due to plastic surgeries, the images differ in terms of imaging conditions and imaging sensor. These problems have challenged automatic face recognition for decades.

Recently, Singh et al. [17] introduced plastic surgery as a new dimension to face recognition. They report a significant degradation ($25 - 30\%$ in rank-1 accuracy) in the matching performance of various existing face matching algorithms when confronted with pre- and post-plastic surgery face images. In this paper, we build on their efforts and introduce a novel approach customized to deal with the challenges of matching faces across variations caused by plastic surgeries. The proposed approach bridges the performance gap to a significant degree on the same database of 900 individuals. Example pre-surgery and post-surgery images of a few subjects from the database are shown in Figure 1.

1.1. Motivation

Facial plastic surgeries are typically performed either

- locally to correct defects, anomalies or to improve general skin texture, e.g., to correct congenital defects such as cleft lip and palate, to improve nose structure, chin, etc., or

- globally to reconstruct the complete facial structure. For example, for patients with severe burns.

Though facial plastic surgeries can be mis-used by criminals to avoid law-enforcement, typically the goal of these surgeries in not to create a new identity. In such cases, both local and global surgeries may result in varying amount of change in relative positioning of facial features and texture. Though the overall face appearance changes, the resulting face typically resembles the original face in a part-wise manner. Unfortunately, these appearance variations are enough to cause most face matching approaches to show significant degradation in performance [17]. Based on these observations, we propose to use a part-wise approach to deal with the challenges posed by these subtle variations in facial appearance. The part-wise framework is combined with the sparse representation approach [20] to improve face matching performance across plastic surgery variations.

It is known that facial features like nose, eyes, lips, etc. play an important role in human and automatic face matching. Holistic approaches [21] to face matching characterize entire face as one entity. Due to this, even changes to parts of the face may lead to very different overall facial characterization, making such approaches non-ideal for the challenges we address here. Therefore, we propose a part-wise approach that is based on the intuition that appearance of one or more facial features may not change much across plastic surgery procedures. In such a part-wise framework, the proposed approach exploits recent successes of sparse representations for face matching [20]. Most sparse representation based face matching approaches require several images of each subject in the gallery. More often than not, one does not have multiple images per person in the gallery as is the case with the plastic surgery database we use in this work. With just one image per subject in the gallery, we overcome this challenge by using sequestered training face images (not pre- and post-plastic surgery images). For each facial part of each gallery subject, we identify most similar facial parts from the training images and use them in the absence of multiple images per subject in the gallery. Even though entire faces from training data do not resemble those in the gallery at all, faces typically show higher part-wise resemblance. Conceptually, this approach shares similarities with *simile classifiers* [10] recently proposed for the task of face verification. The simile classifiers are binary classifiers trained to recognize the similarity of faces, or regions of faces to a few reference people. For example, an unseen face might be described as having a mouth that looks like Barack Obama's and a nose that looks like Owen Wilson's [10].

The proposed approach is evaluated on the plastic surgery database introduced in [17]. Following the suggested evaluation protocol for the database, a significant improvement in rank-1 matching accuracy is observed. Effectiveness of the part-wise analysis without the use of sparse representation is also highlighted.

1.2. Organization of the paper

The rest of the paper is organized as follows. Section 2 briefly discusses a few related works from the literature. A brief discussion of sparse representation for face matching is provided in Section 3. Section 4 provides the details of the proposed part-wise sparse representation approach. The details of the conducted experiments along with results are provided in Section 5. The paper concludes with a brief summary and discussion.

2. Related work

Traditionally, face recognition research has focused primarily on developing novel characterizations and algorithms to deal with challenges posed by variations in acquisition conditions like illumination conditions and head pose with respect to the camera [21]. Tremendous success in dealing with these problems is probably one of the primary factors that has generated interest in new avenues in face matching that include matching faces across plastic surgery variations. Singh et al. [17] introduced matching across plastic surgery variations as a new dimension to face recognition discussing various ethical, social and engineering challenges. They observe that six existing appearance-, feature- and texture-based face matching algorithms show significant performance degradation on the plastic surgery database. The existing algorithms evaluated in [17] are: Principal Component Analysis (PCA) [19], Fisher Discriminant Analysis (FDA) [3], Local Feature Analysis (LFA) [13], Circular Local Binary Pattern (CLBP) [2], Speeded Up Robust Features (SURF) [9], and Neural Network Architecture-based 2-D Log Polar Gabor Transform (GNN) [18].

In this paper, we propose a part-wise sparse representation approach to address the challenge of plastic surgery variations. We provide pointers to a few significant works that have used sparse representation for recognition tasks. Wright et al. [20] show in their pioneering work that exploiting sparsity is critical for high-performance classification of high-dimensional data such as face images. They indicate that choice of feature become less critical than the number of features when the concept of sparsity is properly used. The principle of sparsity has since been used for several other biometrics and computer vision problems. Pillai et al. [16] use a similar approach to select and recognize individuals from iris images. One practical drawback of sparsity-based biometric recognition is the need for several images per subject in the gallery [20][16]. The plastic surgery database used for experimental evaluation used in our work consists of only one pre-surgery and one post-surgery image per subject making it difficult to directly apply sparsity framework to this task. To this end, we propose a part-wise sparse representation approach that chooses similar images from sequestered training data in a part-wise fashion to fulfill the requirement of multiple images per-subject in the gallery.

3. Sparse representation

The use of sparse representation for face recognition was introduced by Wright et al. [20]. Given a set of labeled training samples from k distinct classes, the task is to determine the class to which a new unseen test sample belongs. Let $A_i = [v_{i,1}, v_{i,2}, \ldots, v_{i,n_i}]$ be an $m \times n_i$ matrix of training images from the i-th class in which the n_i training samples are arranged as columns. Each column $v_{i,j}$ in matrix A_i can be the vectorized intensity image or some suitable characterization of the intensity image. One simple and effective approach to exploit the structure of the matrix A_i for face recognition is to model the samples from a single class as lying on a linear subspace. In other words, given enough samples from the i-th class , any new test sample y from the same class can be approximated to lie in the linear span of the columns of matrix A_i, i.e.,

$$y = \sum_{j=1}^{n_i} \alpha_{i,j} v_{i,j} \tag{1}$$

for some scalars $\alpha_{i,j} \in \mathbb{R}, j = 1, 2, \ldots, n_i$. Since the identity of the test sample is initially unknown, we define a new matrix A which is the concatenation of the training samples from all the classes

$$\begin{aligned} A_{m\times n} &= [A_1, A_2, \ldots, A_k] \\ &= [v_{1,1}, \ldots, v_{1,n_1} | \ldots | v_{k,1}, \ldots, v_{k,n_k}] \end{aligned} \tag{2}$$

where $n = \sum_{i=1}^{k} n_i$.With this definition of A, y in (1) can be written as

$$y = Ax \tag{3}$$

where $x = [0, \ldots, 0, \ldots, \alpha_{i,1}, \ldots, \alpha_{i,n_i}, \ldots, 0, \ldots, 0]^T$ is a coefficient vector whose entries are zero except those corresponding to the i-th class. Given a new test sample y and matrix A, the objective is to obtain x that is informative about the identity of the test sample to aid in recognition task. There exist several decomposition methods for this. Sparse representation approaches find the sparsest solution to the linear systems of equations $y = Ax$. This follows from the intuition that test sample y can be sufficiently represented using only the samples from its true class, which naturally leads to sparse x if the total number of classes in A is large. The more sparse the recovered x is, the better it indicates the identity of the unlabeled test sample, which leads to solving the following optimization problem

$$\hat{x} = arg \min_{x \in \mathbb{R}^n} ||x||_1 \quad \text{subject to} \quad Ax = y \tag{4}$$

Actually, l^0-minimization instead of l^1-minimization provides the sparsest solution but it leads to a NP-hard problem even for approximation. Recent developments in compressed sensing [8] show that if the solution sought is sparse enough, the solution to the l^0-minimization problem is equal to that of the l^1-minimization problem. Readers are encouraged to refer to [20] for details. The optimization problem in (4) is known as Basis Pursuit (BP) and can be solved in polynomial time by standard linear programming methods [6].

3.1. Applicability to plastic surgery data

One of the main requirements of the sparse representation based approaches is the availability of multiple samples for each class in matrix A. The plastic surgery dataset [17] consists of only one pre-surgery and one post-surgery image for each subject in the dataset. To this end, we propose to use sequestered training data with no identity overlap to fulfill the requirement. Holistically, the faces in plastic surgery images do not look similar to the training face images in any way since those identities are not present in the training data, but the images show much higher similarity when considered at a more local level. For example, eyes of a person may resemble those of some other person. This intuition forms the basis of the proposed training approach. For each gallery identity, we populate matrix A with the images with most similar facial characteristics from the sequestered training data. This is done in a part-wise manner. The following section describes the details of the proposed framework.

4. Proposed framework

The proposed approach consists of the following steps

1. Localization of face and primary facial features.
2. Generation of training Matrix A (for each facial part).
3. Sparse Recognition (for each facial part).

We now provide details of each of these steps.

4.1. Localization of face and primary facial features

Like any other modern face matching algorithm, given a face image, we first need to locate the face region and various primary facial features like eyes, eyebrows, nose, mouth, etc. This is done automatically using publicly

Figure 2. A few example images illustrating automatic facial feature localization obtained using an ASM-based library known as STASM [11].

available Active Shape Model (ASM) library known as STASM [11].

Here, we provide brief description of STASM for the sake of completeness. Traditional ASM [7] involves an iterative process to locate facial landmarks. The search starts from the mean shape aligned to the face in the test image determined by a global face detector. STASM improves traditional ASM by incorporating a few simple but effective extensions that include (a) fitting more landmarks than are actually needed, (b) selectively using two-dimensional templates in ASM model instead of one-dimensional templates, and (c) relaxing the shape model where advantageous. Figure 2 shows a few example images from the plastic surgery dataset with automatically localized landmarks.

4.2. Generation of training Matrix A

Using the landmark locations automatically determined by STASM, we crop several facial regions that include the two eyes, nose, mouth and eyebrows. Figure 3 shows the cropped facial regions for two subjects obtained from pre-surgery and post-surgery face images. As can be seen, the appearance of the two subjects changes significantly from a holistic point of view. On the other hand, appearance appears to be better preserved for individual face regions (especially eyes and eye-brows in this example) across the plastic surgery procedures. Although we use the mentioned six regions in this investigation, other regions can easily be added in the framework. The rest of the processing including generation of the training matrix and recognition using the principle of sparsity, is performed individually for each facial region.

As mentioned in Section 3, one needs multiple images per subject to populate training matrix A for the sparse recognition process to function effectively. The plastic surgery database has just one pre-surgery and one post-surgery image per subject. Therefore, we make use of sequestered training data with no identity overlap to populate the training matrix. The basic intuition is to look for facial regions from the training face images that closely resemble those in pre-surgery gallery images. For each gallery image, the closely resembling facial regions from the training data are then used to populate training matrix A as if they belonged to the corresponding gallery identity. Following this protocol, class-wise sub-matrices A_i in the training matrix $A = [A_1, A_2, \ldots, A_k]$ from (3) consists of one pre-surgery gallery image and training images that closely resemble the gallery image as follows

$$A_i = [v_i, t_{i,1}, t_{i,2}, \ldots, t_{i,n_i-1}] \tag{5}$$

where $t_{i,1}, t_{i,2}, \ldots, t_{i,n_i-1}$ are from the training images such that

$$||t_{i,l} - v_i|| < ||t_{i,l} - v_j||, \tag{6}$$

$\forall l \in 1, \ldots, n_i - 1$ and $\forall j \in 1, \ldots, k, j \neq i$. Note that we omit the second subscript from the v's as there is just one sample per subject.

As mentioned earlier, each v_i (or $t_{i,l}$) can simply be the vectorized intensity image of the corresponding facial part or some suitable characterization of the intensity image. In all the experiments described in this paper, we use a standard PCA-based representation [19] to characterize each facial region. Sequestered non-plastic surgery training data is used to generate the PCA space.

4.3. Sparse recognition

The identity is determined independently for each facial part as follows. Given a test sample y and the training matrix A, the sparse coefficient vector $\hat{x}$ is obtained by solving (4). Final classification is performed by determining which class present in A best represents the test sample using the recovered $\hat{x}$. Representation error for the i-th class is computed by reconstructing test sample using the samples belonging to that class only as follows.

$$e_i(y) = ||y - A\hat{x}_i||, \tag{7}$$

where $\hat{x}_i = [0, 0 \ldots, 0, \hat{x}_{i,1}, \ldots, \hat{x}_{i,n_i}, 0, \ldots, 0]^T \in \mathbb{R}^n$. Since in our framework, sparse recognition is performed individually for each part, one needs to fuse evidence from all the six parts for overall recognition. For fusion, we use simple sum rule such that overall sparse representation error

Figure 3. Pre-surgery and post-surgery images for two subjects along with their corresponding cropped facial regions. Although the holistic appearance appears to change significantly due to plastic surgeries for these subjects, appearance seems to be better preserved for several of the cropped facial regions.

$e_i(y)$ for a test sample y for the i-th class is

$$e_i(y) = \sum_{p=1}^{6} e_i^p(y) \quad (8)$$

where $e_i^p(y)$ corresponds to the sparse representation error for the p-th part. The test sample is classified to the class that produces smallest representation error.

5. Experiments

We now describe the details of the experimental evaluations performed along with the results obtained to evaluate the efficacy of the proposed approach. The performance of the proposed approach is compared against all six algorithms evaluated in [17].

5.1. Datasets

Singh et al. [17] procured a dataset by downloading pre-surgery and post-surgery images from the web (www.locateadoc.com and www.surgery.com). A list of URLs to images of 900 subjects is publicly available that is used to download images to conduct experiments presented in this paper. There is one pre-surgery and one post-surgery image for each subject resulting in a total of 1800 images. A few images from the dataset are shown in Figure 1.

To facilitate applicability of the principle of sparsity to this task, the proposed framework requires a sequestered set of face images that have no identity overlap with the plastic surgery dataset. Here, we use a subset of still face images from the Multi-Biometric Grand Challenge (MBGC) dataset [14]. The images used are captured in reasonably good illumination with subjects in frontal pose (Figure 4). The imaging conditions and sensor quality in these images is quite different from the images in plastic surgery database. These images are also used for generation of PCA-space for characterization of facial parts.

5.2. Evaluation

In [17], 360 subjects (40% of the dataset) are used for training while the remaining 540 subjects are used for testing. The non-overlapping train-test partitioning is repeated

Figure 4. A few example images from the MBGC dataset.

ten times. The accuracy is reported in terms of Cumulative Match Characteristic (CMC) curves and rank-1 accuracy. To facilitate easy comparison, we follow the same evaluation protocol. In our set up, we do not need images from the plastic surgery dataset for training. Therefore, we randomly sample 540 subjects several times from the entire dataset and report average performance accuracy. We start with showing performance of each of the six parts individually followed by overall performance obtained by their fusion.

Figure 5 shows the individual CMC curves corresponding to the six facial parts. Table 1 shows the corresponding accuracies in terms of rank-1 performance. Rank-1 recognition accuracy of the parts vary in $21.5\% - 40.9\%$ range with eyes, nose and mouth regions performing better than the eyebrows. Note that even with PCA-based characterization (used in conjunction with sparse recognition), four out of the six facial regions outperform the performance of PCA-based matching (performed on the entire face) as reported in [17].

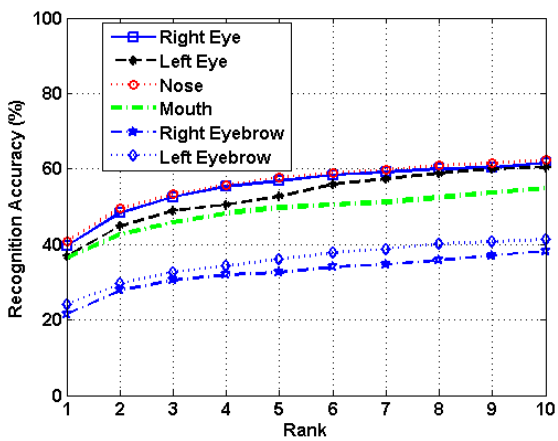

Figure 5. CMC plots demonstrating the recognition performance for each of the six facial regions used in the proposed framework.

Figure 6 shows the CMC obtained by performing simple sum-based fusion of evidence from the six facial parts. Fusion performance obtained without using sparse recognition for each part is also shown. CMCs of six algorithms evaluated in [17] are approximately reproduced from the publication. Corresponding rank-1 accuracies are shown in Table 2. Since the six facial parts contribute independent

Facial Part	Rank-1 Performance
Right Eye	39.5%
Left Eye	36.8%
Nose	40.9%
Mouth	36.4%
Right Eyebrow	21.5%
Left Eyebrow	24.2%

Table 1. Rank-1 recognition accuracy for each of the six facial regions used in the proposed framework.

information, the fused performance is much better than the one observed with any of the facial parts. The proposed part-wise sparse recognition-based approach significantly outperforms the rest by a significant margin. The good performance can be attributed to both the part-wise framework and the principle of sparsity. In fact, as shown in Figure 6 and Table 2, simple fusion of the six parts with PCA-based characterization (without sparse recognition) outperforms holistic PCA by over 15% in terms of rank-1 accuracy.

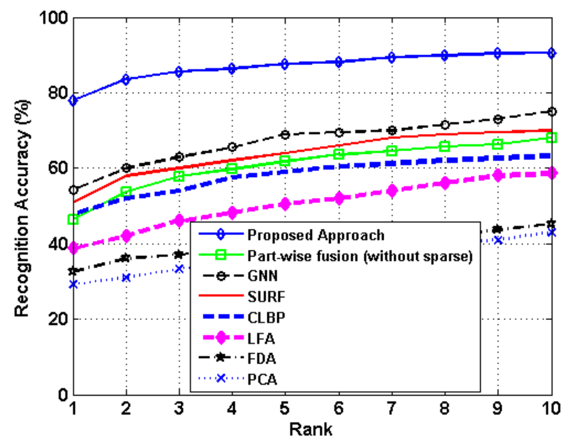

Figure 6. CMC plot demonstrating the recognition performance obtained by fusing evidence from the six facial regions used in the proposed framework. Fused performance of the six parts without using the sparse recognition framework is also shown. The two CMC curves are compared against the six algorithms evaluated in [17].

6. Summary and Discussion

In this paper, we built on the efforts of Singh et al. [17], who introduced face recognition across plastic surgeries as a new dimension to face recognition. The proposed framework exploits the advantages of part-wise analysis with the recently popular sparse recognition approach to deal with the challenges posed by plastic surgery variations. The proposed formulation relies on training images from sequestered non-gallery subjects to fulfill the multiple image

Algorithm	Rank-1 Performance
PCA [19]	29.1%
FDA [3]	32.5%
LFA [13]	38.6%
CLBP [2]	47.8%
SURF [9]	50.9%
GNN [18]	54.2%
Fusion of PCA-features (without use of sparsity)	**46.4%**
Proposed Approach	**77.9%**

Table 2. Rank-1 recognition accuracy obtained by fusing evidence from the six facial regions used in the proposed framework. Fused performance of the six parts without using the sparse recognition framework is also shown. Comparison against the six algorithms evaluated in [17] is also shown.

requirement of sparse recognition method. Closely resembling training images with facial parts most similar to each subject in the gallery are identified and used in the absence of multiple gallery images. We show that the proposed algorithm significantly bridges the performance gap earlier observed when matching faces across plastic surgeries, compared to matching normal face images. As supported by the experimental evaluations, the good performance of the proposed approach is attributed to both part-wise analysis and sparse recognition technique.

References

[1] American Society for Aesthetic Plastic Surgery: Statistics http://www.surgery.org/media/statistics, 2008. 1

[2] T. Ahonen, A. Hadid, and M. Pietikinen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, 2006. 2, 7

[3] P. N. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(7):711–720, 1997. 2, 7

[4] S. Bharadwaj, H. S. Bhatt, R. Singh, M. Vatsa, and S. K. Singh. Face recognition for newborns: A preliminary study. In *IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–6, 2010. 1

[5] S. Biswas, K. W. Bowyer, and P. J. Flynn. Multidimensional scaling for matching low-resolution facial images. In *IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–6, 2010. 1

[6] S. Chen, D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. *SIAM Review*, 43(1):129–159, 2001. 3

[7] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active Shape Models-their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995. 4

[8] D. Donoho. For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution. *Communications on Pure and Applied Math*, 59(6):797–829, 2006. 3

[9] P. Dreuw, P. Steingrube, H. Hanselmann, and H. Ney. Surfface: Face recognition under viewpoint consistency constraints. In *British Machine Vision Conference*, 2009. 2, 7

[10] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar. Attribute and Simile Classifiers for Face Verification. In *IEEE International Conference on Computer Vision (ICCV)*, Oct 2009. 2

[11] S. Milborrow and F. Nicolls. Locating facial features with an extended active shape model. In *Proceedings of European Conference on Computer Vision*, 2008. http://www.milbo.users.sonic.net/stasm. 4

[12] A. J. O'Toole, P. J. Phillips, F. Jiang, J. H. Ayyad, N. Penard, and H. Abdi. Face recognition algorithms surpass humans matching faces over changes in illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9):1642–1646, 2007. 1

[13] P. Penev and J. Atick. Local feature analysis: A general statistical theory for object representation. *Network: Computation Neural Systems*, 7:477–500, 1996. 2, 7

[14] P. J. Phillips, P. J. Flynn, J. R. Beveridge, W. T. Scruggs, A. J. O'Toole, D. S. Bolme, K. W. Bowyer, A. Draper, Bruce, G. H. Givens, Y. M. Lui, H. Sahibzada, J. A. Scallan, and S. Weimer. Overview of the multiple biometrics grand challenge. In *International Conference on Biometrics*, pages 705–714, 2009. 5

[15] 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*, 32(5):831–846, 2010. 1

[16] J. K. Pillai, V. M. Patel, and R. Chellappa. Sparsity inspired selection and recognition of iris images. In *Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems*, pages 184–189, 2009. 2, 3

[17] R. Singh, M. Vatsa, H. S. Bhatt, S. Bharadwaj, A. Noore, and S. S. Nooreyezdan. Plastic surgery: a new dimension to face recognition. *IEEE Transactions on Information Forensics and Security*, 5(3):441–448, 2010. 1, 2, 3, 5, 6, 7

[18] R. Singh, M. Vatsa, and A. Noore. Face recognition with disguise and single gallery images. *Image and Vision Computing*, 27(3):245–257, 2009. 2, 7

[19] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991. 2, 4, 7

[20] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210–227, 2009. 2, 3

[21] W.-Y. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399–458, 2003. 1, 2