



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
SCHOOL OF ENGINEERING

**FACE RECOGNITION USING ADAPTIVE  
DICTIONARIES AND SPARSE  
FINGERPRINT CLASSIFICATION  
ALGORITHM**

**TOMÁS ANTONIO LARRAIN ARELLANO**

Thesis submitted to the Office of Research and Graduate Studies  
in partial fulfillment of the requirements for the degree of  
Master of Science in Engineering

Advisor:

DOMINGO MERY QUIROZ, PH.D.

Santiago de Chile, June 2015

© MMXV, TOMÁS ANTONIO LARRAIN ARELLANO



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
SCHOOL OF ENGINEERING

**FACE RECOGNITION USING ADAPTIVE  
DICTIONARIES AND SPARSE  
FINGERPRINT CLASSIFICATION  
ALGORITHM**

**TOMÁS ANTONIO LARRAIN ARELLANO**

Members of the Committee:

DOMINGO MERY QUIROZ, PH.D.

KARIM PICHARA BAKSAI, PH.D.

PATRICIO DE LA CUADRA BANDERAS, PH.D.

GONZALO ACUÑA LEIVA, PH.D.

CRISTIÁN TEJOS NUÑEZ, PH.D.

Thesis submitted to the Office of Research and Graduate Studies  
in partial fulfillment of the requirements for the degree of  
Master of Science in Engineering

Santiago de Chile, June 2015

© MMXV, TOMÁS ANTONIO LARRAIN ARELLANO

*Para el Tata Humberto y la Keka*

## **ACKNOWLEDGEMENTS**

This work was supported in part by Fondecyt grant 1130934 from CONICYT-Chile and in part by Seed Grant Program of The College of Engineering at the Pontificia Universidad Catolica de Chile and the College of Engineering at the University of Notre Dame.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS . . . . .	iv
LIST OF FIGURES . . . . .	vi
LIST OF TABLES . . . . .	vii
ABSTRACT . . . . .	viii
RESUMEN . . . . .	ix
1. INTRODUCTION . . . . .	1
2. PROPOSED METHOD AND TESTING METHODOLOGY . . . . .	4
2.1. Training . . . . .	5
2.2. Testing . . . . .	7
2.2.1. Adaptive Dictionary Selection . . . . .	7
2.2.2. Fingerprint . . . . .	9
2.2.3. Classification . . . . .	11
2.3. Testing Methodology . . . . .	12
3. EXPERIMENTAL RESULTS . . . . .	14
3.1. Experiments under different lighting conditions . . . . .	14
3.2. Experiments with subjects with different expressions . . . . .	15
3.3. Experiments with real occlusion and face expressions . . . . .	16
3.4. Implementation . . . . .	18
3.5. Parameters Sensitivity Analysis . . . . .	18
4. CONCLUSION . . . . .	21
References . . . . .	22

## LIST OF FIGURES

2.1 Overview of the proposed method. . . . .	4
2.2 Grid Example . . . . .	5
2.3 Dictionaries Example . . . . .	8
2.4 Fingerprints for different subjects . . . . .	9
2.5 Examples of $\mathbf{f}_t$ vectors . . . . .	12
2.6 Examples of the databases . . . . .	13
3.1 Example of the Tan-Triggs algorithm . . . . .	18
3.2 Sensitivity graphs . . . . .	19

## LIST OF TABLES

3.1 Comparison on Yale . . . . .	15
3.2 Comparison on MPIE . . . . .	15
3.3 Comparison on ORL . . . . .	16
3.4 Comparison on FWM . . . . .	17
3.5 Comparison on AR . . . . .	17
3.6 Comparison on AR <sub>x</sub> . . . . .	18

## ABSTRACT

Unconstrained face recognition is still an open problem, as state-of-the-art algorithms have not yet reached high recognition performance in such environments. This paper addresses this problem by proposing a new approach called Sparse Fingerprint Classification Algorithm (SFCA). In the training phase, for each enrolled subject, a grid of patches is extracted from each subject's face images in order to construct representative dictionaries. In the testing phase, a grid is extracted from the query image and every patch is transformed into a binary sparse representation using the dictionary, creating a fingerprint of the face. The binary coefficients vote for their corresponding classes and the maximum-vote class decides the identity of the query image. Experiments were carried out on five widely-used face databases. The results demonstrate that SFCA is able to deal with a larger degree of variability in ambient lighting, pose, expression, occlusion, face size and distance from the camera than other current state of the art algorithms.

**Keywords:** face recognition, fingerprint, sparse representation.

## RESUMEN

El reconocimiento facial en ambientes no controlados sigue siendo un problema abierto, dado que los algoritmos del estado del arte no han alcanzado un alto porcentaje de aciertos en estas situaciones. Este documento ataca este problema proponiendo un nuevo método llamado *Sparse Fingerprint Classification Algorithm* (SFCA). En la fase de entrenamiento, para cada sujeto enrolado, se extrae una grilla de parches de cada imagen de su cara para construir diccionarios representativos. En la fase de prueba, una grilla es extraída de la imagen en cuestión y cada parche es convertido en una representación rala (*sparse*) binaria usando los diccionarios, con esto se crea un *fingerprint* de la cara. Los coeficientes binarios votan por su clase correspondiente y la clase con mayor cantidad de votos decide la identidad de la imagen en cuestión. Se realizaron experimentos a lo largo de cinco bases de datos conocidas; los resultados muestran que SFCA es capaz de lidiar con variaciones de luz ambiente, posición de la cara, expresión, oclusión, tamaño de la cara y distancia a la cámara de mejor manera que otros algoritmos vigentes del estado del arte.

**Palabras Claves:** reconocimiento facial, fingerprint, representaciones sparse.

## 1. INTRODUCTION

Face recognition has been a very active area of research in computer vision, making many important contributions since the 1990s. In recent years the emphasis of face recognition research has shifted to dealing with unconstrained conditions, including variability in ambient lighting, pose, expression, face size, occlusion Wei et al. (2014) and distance from the camera Phillips et al. (2011). In the last few years, many approaches have been proposed to deal with the aforementioned problems (see for example Taigman et al. (2014)).

Algorithms based on Sparse Representation Classification (SRC) have been widely explored recently Wright et al. (2009). In the sparse representation approach, a dictionary is built from the gallery images, and matching is done by reconstructing the query image using a sparse linear combination of the dictionary. The identity of the query image is assigned to the class with the minimal reconstruction error. Many variations of this approach were recently proposed. In Wagner et al. (2012), registration and illumination are simultaneously considered in the sparse representation. In Deng et al. (2012), an intra-class variant dictionary is constructed to represent the possible variation between gallery and query images. In J. Wang et al. (2014b), sparsity and correlation are jointly considered. In Jia et al. (2012) and Wei et al. (2012), structured sparsity is proposed for dealing with occlusion and illumination. In Deng et al. (2013), the dictionary is assembled by the class centroids and sample-to-centroid differences. In J. Chen & Yi (2014), SRC is extended by incorporating the low-rank structure of data representation. In Jiang et al. (2013), a discriminative dictionary is learned using label information. In Ptucha & Savakis (2013), a linear extension of graph embedding is used to optimize the learning of the dictionary. In Qiu et al. (2014), a discriminative and generative dictionary is learned based on the principle of information maximization. In Shi et al. (2014), a sparse discriminative analysis is proposed using the  $\ell_{2,1}$ -norm. In Xu et al. (2011a), a sparse representation in two phases is proposed. In Y. Chen et al. (2010), sparse representations of patches distributed in a grid manner are used. In Mery & Bowyer (2014), the construction of the dictionary is with

patches that are randomly located on the face image. These variations improve recognition performance significantly as they are able to model various corruptions in face images, such as misalignment and occlusion.

Other approaches with comparable performance are based on the similarity between features extracted from regions of the gallery images and from the query image Tan et al. (2009). Recently, one novel approach proposed a new representation of the face image that is a sequence of forehead, eyes, nose, mouth and chin in a natural order Wei et al. (2013).

In a related field, ‘audio fingerprints’ are now widely used to represent audio signals for matching. Different methods are used to extract an audio fingerprint, such as wavelet transform Kamaladas & Dialin (2013); Baluja & Covell (2007), the Fourier Transform Ouali et al. (2014) or entropy based Ibarrola & Chavez (2006). These algorithms are very robust in terms of ambient noise and volume. Fingerprinting is a way to create a database with reduced information about the signal, but preserving distinctive elements, so that it is easier to search over the database to find the closest match. Commercial uses of the fingerprinting approach have been developed by companies like Shazam for its mobile application A. Wang et al. (2003), Microsoft to detect duplicates on audio sets Burges et al. (2005b), and other companies to monitor audio in a radio broadcast Allamanche et al. (2001); Camarena-Ibarrola et al. (2009). It is known that fingerprinting in audio is an effective method of recognizing songs. Since face images can be interpreted as signals, we demonstrate in our work that the fingerprinting concept can also be used in face recognition.

Reflecting on the problems confronting unconstrained face recognition, and on the solutions proposed in recent years, we believe that there are some key ideas that should be present in new proposed solutions. First, if the face image is somehow occluded, it is clear that the occluded parts are not providing any information of the subject identity. For this reason, such parts should be automatically detected and should not be considered by the recognition algorithm. Second, in recognizing any face, there are parts of the face that are more relevant than other parts (for example birthmarks, moles or large eyebrows, to name but a few). For this reason, relevant parts should be subject-dependent, and could be found

using unsupervised learning. Third, the expression that is present in a query face image can be subdivided into sub-expressions, for different parts of the face (*e.g.*, eyebrows, nose, mouth). For this reason, when searching for similar gallery subjects it would be helpful to search for image parts in all images of the gallery instead of similar gallery images.

Inspired by these key ideas, this paper proposes a new method for face recognition that is able to deal with less constrained conditions. Two main contributions of our approach are:

- (i) A new representation for the gallery face images of a subject: this is based on representative dictionaries learned for each subject of the gallery, which correspond to a rich collection of representations of selected relevant parts that are particular to the subject's face.
- (ii) A new representation for the query face image: this is based on i) a discriminative criterion that selects the best test patches extracted from a grid of the query image and ii) a sparse fingerprint made with a binary sparse representation of the best patches.

Using these new representations, the proposed method (SFCA) can achieve high recognition performance under many conditions, as shown in our extensive experiments.

The method proposed in this article is based on Mery & Bowyer (2014) but with two important differences: i) the extraction of the patches is not random but using a square grid, and ii) the classification is a novel approach based on sparse fingerprint representations. These two differences are important and result in performance improvement on several of the tests presented later in this paper.

The rest of the thesis is organized as follows: in Section 2, the proposed method is explained in further detail. In Section 3, the experiments and results are presented. Finally, in Section 4, concluding remarks are given.

## 2. PROPOSED METHOD AND TESTING METHODOLOGY

Following a sparse representation methodology, in a learning stage, a grid of patches can be extracted from each training image, and a dictionary can be built for each class by concatenating its patches (stacking in columns). In the testing stage, several patches can be extracted and each of them can be classified using its sparse representation. The final decision is taken by using our proposed method. This baseline approach, however, shows three important disadvantages: i) The location information of the patch is not considered, *i.e.*, a patch of one part of the face could be erroneously represented by a patch of a different part of the face. This first problem can be solved by considering the  $(x, y)$  location of the patch in its description. ii) The method requires a huge dictionary for reliable performance, *i.e.*, each sparse representation process would be very time consuming. This second problem can be remedied by using only a part of the dictionary adapted to each patch. Thus, the whole dictionary of a class can be subdivided into sub-dictionaries, and only the ‘best’

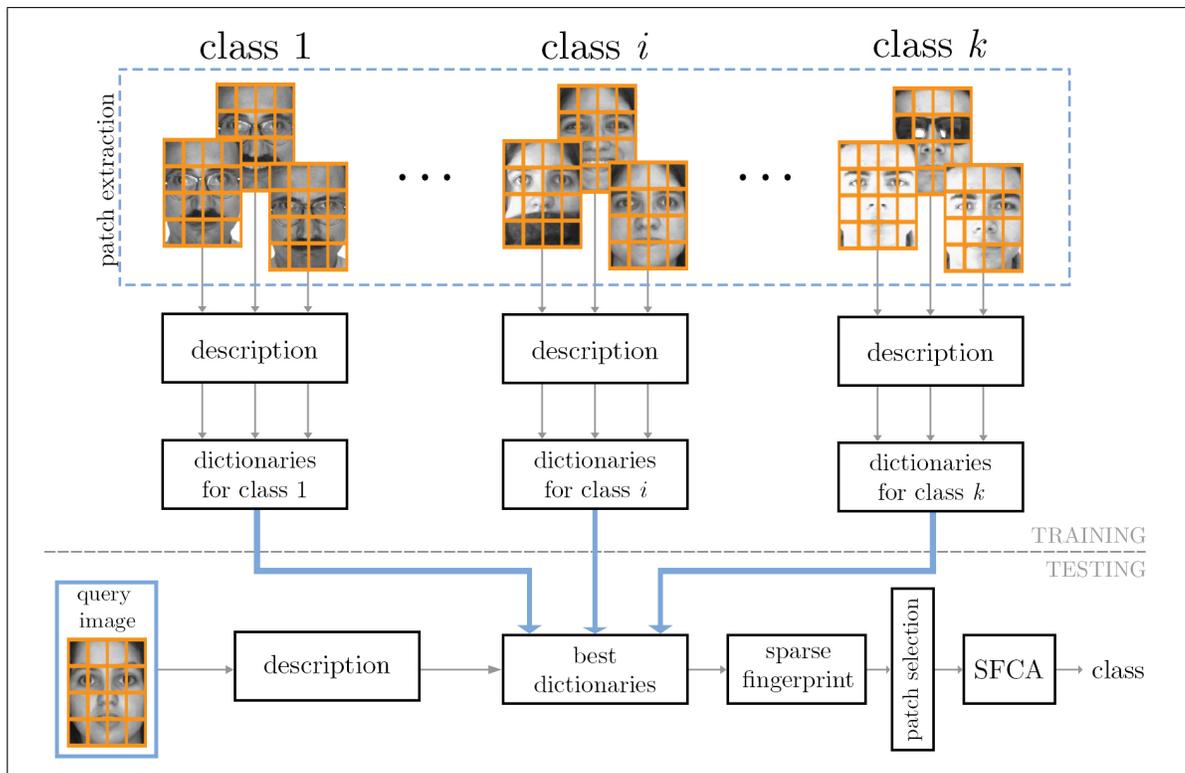


FIGURE 2.1. Overview of the proposed method.

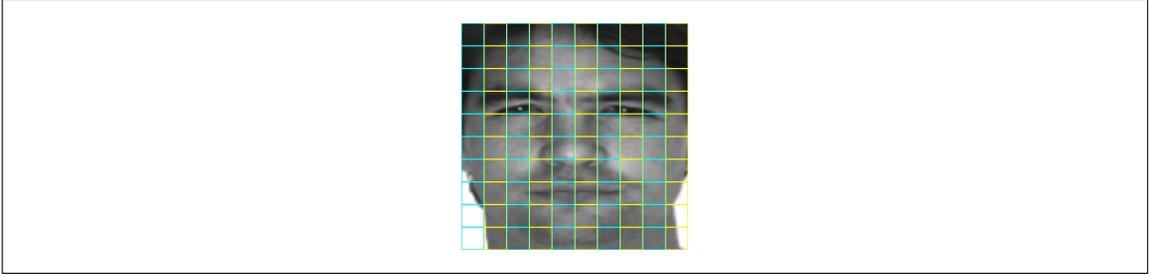


FIGURE 2.2. Example of a grid using  $m = 100$  patches (10 rows and 10 columns)

ones used to compute the sparse representation of a patch. iii) Not all query patches are relevant, *i.e.*, some patches of the face do not provide any discriminative information of the class (*e.g.*, patches over sunglasses or other kind of occlusion). This third problem can be addressed by selecting the query patches according to a score value. In this section we describe our approach taking into account the three mentioned improvements.

As illustrated in Figure 2.1, in the learning stage, for each class of the gallery, a grid of patches is extracted and described from their images (using both intensity and location features) to build representative dictionaries. In the testing stage, a square grid of test patches is extracted from the query image and described. For each test patch a dictionary is built concatenating the ‘best’ representative dictionary of each class. Using this adapted dictionary, each test patch is classified using the method proposed in this paper. Afterwards, the patches are selected according to a discriminative criterion. Finally, the query image is classified by applying SFCA for the selected patches. The training and the testing stages are explained in detail later in this section.

## 2.1. Training

In this stage, we use a set of  $N$  face images of each of the  $K$  subjects, where  $\mathbf{I}_j^i$  denotes image  $j$  of subject  $i$  (for  $i = 1 \dots K$  and  $j = 1 \dots N$ ). In each image  $\mathbf{I}_j^i$ ,  $m$  patches of size  $a \times a$  pixels are extracted using a grid  $\mathbf{G}_m$ . This grid has equal number of rows and columns, like the one illustrated in Figure 2.2. The patches are denoted as  $\mathcal{P}_{hw}^{ij}$  (for  $h, w = 1 \dots \sqrt{m}$ ) and distributed according to:

$$\mathbf{G}_m = \begin{array}{|c|c|c|} \hline \mathcal{P}_{11} & \cdots & \mathcal{P}_{1\sqrt{m}} \\ \hline \vdots & \ddots & \vdots \\ \hline \mathcal{P}_{\sqrt{m}1} & \cdots & \mathcal{P}_{\sqrt{m}\sqrt{m}} \\ \hline \end{array} \quad (2.1)$$

The center of the patch is also a relevant variable and will be denoted as  $(x_{hw}^{ij}, y_{hw}^{ij})$ . In this work, the description of a patch  $\mathcal{P}$  is defined as a vector:

$$\mathbf{y} = f(\mathcal{P}) = [\mathbf{z}; \alpha x; \alpha y] \in \mathbb{R}^{d+2} \quad (2.2)$$

where  $d$  is the number of pixels of the patch and  $\mathbf{z} \in \mathbb{R}^d$  is a descriptor of patch  $\mathcal{P}$  made by stacking vertically the columns of  $\mathcal{P}$ ,  $(x, y)$  are the image coordinates of the center of patch  $\mathcal{P}$ , and  $\alpha$  is a weighting factor between description (given by  $\mathbf{z}$ ) and location (given by  $(x, y)$ ). Using (2.2) all  $m$  extracted patches of image  $j$  of subject  $i$  are described as  $\mathbf{y}_{hw}^{ij} = f(\mathcal{P}_{hw}^{ij})$  ( $h$  and  $w$  denotes the position of the patch in the grid  $\mathbf{G}_m$ ). Thus, for subject  $i$  an array with the description of all patches is defined as  $\mathbf{Y}^i = \{\mathbf{y}_{hw}^{ij}\} \in \mathbb{R}^{(d+2) \times nm}$ . The description  $\mathbf{Y}^i$  of subject  $i$  is clustered using a k-means algorithm in  $Q$  clusters that will be referred to as *parent* clusters:

$$\mathbf{c}_q^i = \text{kmeans}(\mathbf{Y}^i, Q) \quad (2.3)$$

for  $q = 1 \dots Q$ , where  $\mathbf{c}_q^i \in \mathbb{R}^{(d+2)}$  is the centroid of parent cluster  $q$  of subject  $i$ . We defined  $\mathbf{Y}_q^i$  as the array with all samples  $\mathbf{y}_{hw}^{ij}$  that belong to the parent cluster with centroid  $\mathbf{c}_q^i$ .

In order to select a reduced number of samples, each parent cluster is clustered again in  $R$  child clusters:

$$\mathbf{c}_{qr}^i = \text{kmeans}(\mathbf{Y}_q^i, R) \quad (2.4)$$

for  $r = 1 \dots R$ , where  $\mathbf{c}_{qr}^i \in \mathbb{R}^{(d+2)}$  is the centroid of child cluster  $r$  of parent cluster  $q$  of subject  $i$ . All centroids of child clusters of subject  $i$  are arranged in an array  $\mathbf{D}^i$  (orange rectangle in Figure 2.3), and specifically for parent cluster  $q$  are arranged in a matrix:

$$\mathbf{A}_q^i = [\mathbf{c}_{q1}^i \dots \mathbf{c}_{qr}^i \dots \mathbf{c}_{qR}^i] \in \mathbb{R}^{(d+2) \times R} \quad (2.5)$$

Thus, this arrangement contains  $R$  representative samples of parent cluster  $q$  of subject  $i$  as illustrated in Figure 2.3. The set of all centroids of child clusters of subject  $i$  ( $\mathbf{D}^i$ ), represents  $Q$  representative dictionaries with  $R$  descriptions  $\{\mathbf{c}_{qr}^i\}$  for  $q = 1 \dots Q$ ,  $r = 1 \dots R$ .

## 2.2. Testing

In the testing stage, the task is to determine the identity of the query image  $\mathbf{I}^t$  given the model learned in the previous section. This stage consists of the following three steps.

### 2.2.1. Adaptive Dictionary Selection

A grid of patches is extracted from the query image, and described using (2.2), in the same way as for a training image. A subset of the patches is then selected according to a criterion explained later in this section. For each selected query-image patch, the nearest parent-cluster  $\hat{q}^i$  is found for each subject  $i$  of the gallery by computing the minimum distance to the corresponding child-cluster centroids of each one (*i.e.* the distance to each  $\mathbf{c}_{qr}^i$ ). Using (2.6) the nearest parent-cluster are selected:

$$\hat{q}^i = \underset{q}{\operatorname{argmin}} \|\mathbf{c}_{qr}^i - \mathbf{y}\|_2 \quad (2.6)$$

Finally, the adaptive dictionary for each patch is constructed by the concatenation of the parent clusters that contains the nearest child cluster centroid of each subject.

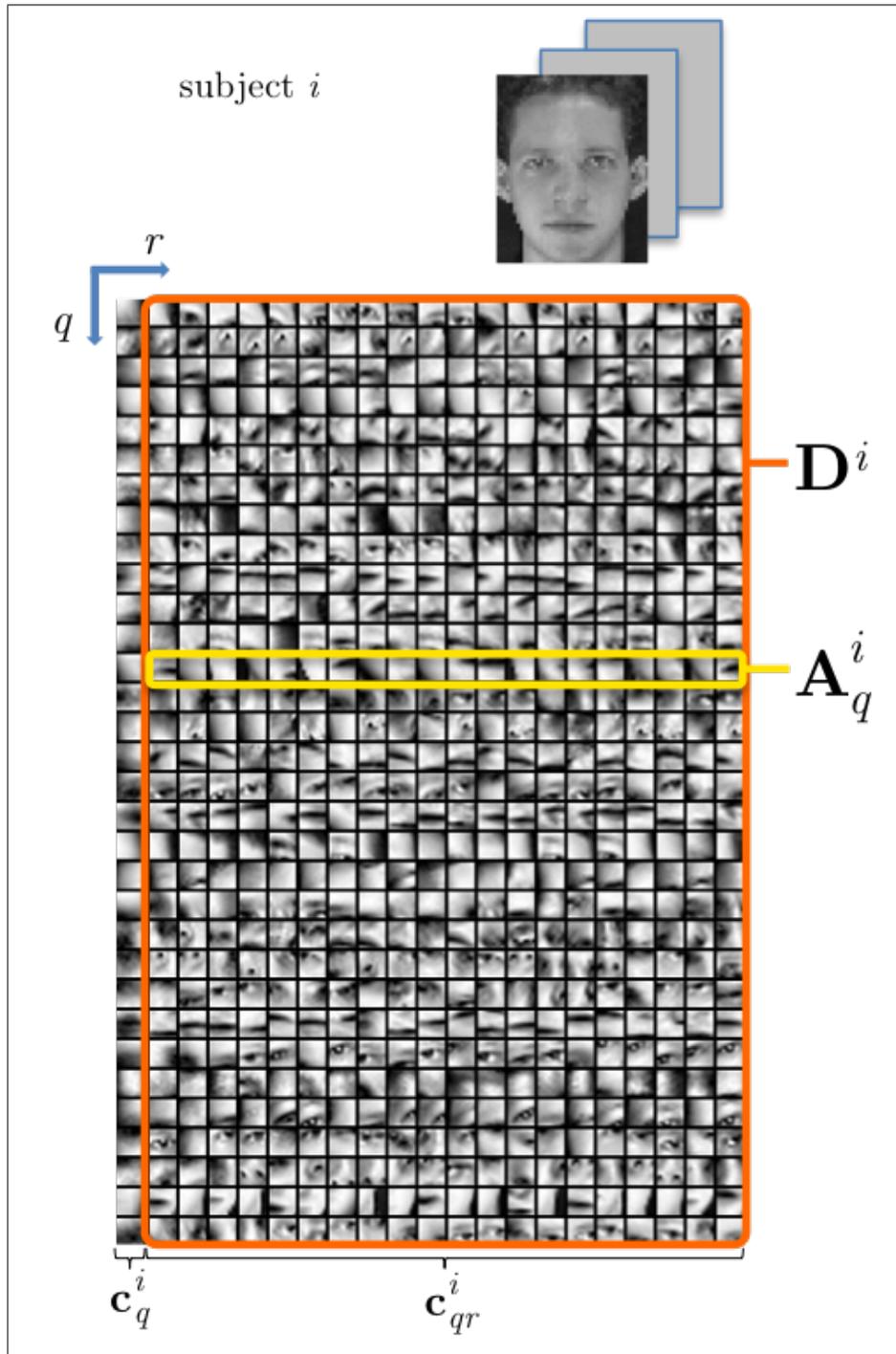


FIGURE 2.3. Dictionaries of subject  $i$  for  $Q = 32$  parent clusters and  $R = 20$  child clusters. Left column shows the centroids  $\mathbf{c}_q^i$  of parent clusters. Right columns (orange rectangle called  $\mathbf{D}^i$ ) shows the centroids  $\mathbf{c}_{qr}^i$  of child clusters.  $\mathbf{A}_q^i$  is row  $q$  of  $\mathbf{D}^i$ , *i.e.*, the centroids of child clusters of parent cluster  $q$ .

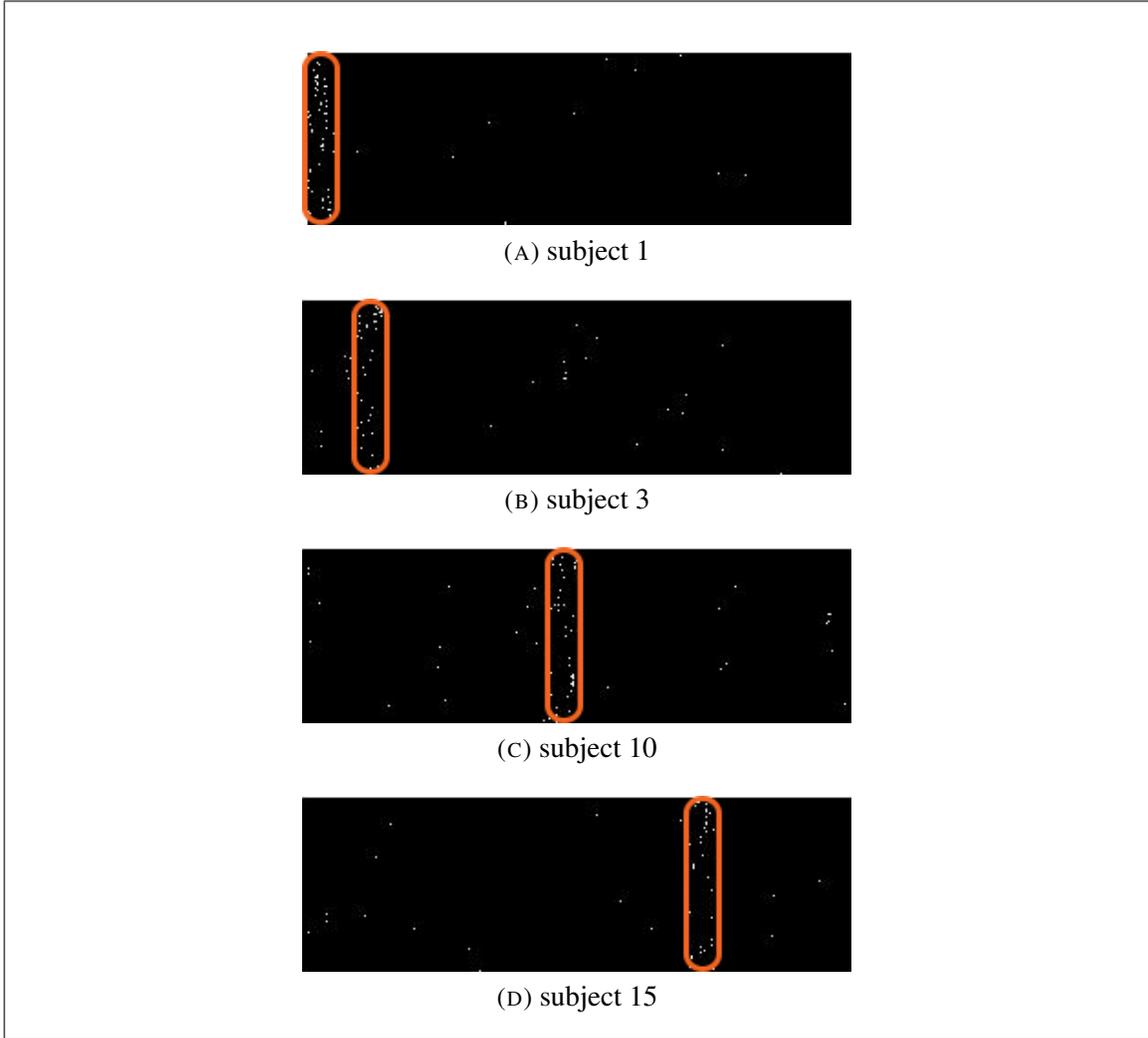


FIGURE 2.4. Fingerprints  $\mathbf{F}$  for different subjects. The Orange area shows that the biggest concentration of sparse coefficients are in the columns that correspond to the correct subject. Because of space considerations only 4 subjects are shown.

$$\mathbf{A}(\mathbf{y}) = [\mathbf{A}_{\hat{q}^1}^1 \dots \mathbf{A}_{\hat{q}^i}^i \dots \mathbf{A}_{\hat{q}^K}^K]^\top \in \mathbb{R}^{(d+2) \times KR} \quad (2.7)$$

### 2.2.2. Fingerprint

The main contributions of our work are in this step. Both the computation of the fingerprint as a method of face recognition, and the method used to classify face fingerprints, are novel contributions introduced in this paper.

The first step of computing the fingerprint of a patch  $\mathbf{y}$  is to look for a sparse representation of it. This is achieved by using the  $\ell_1$ -minimization approach, with the adaptive dictionary  $\mathbf{A}$  found for this patch using (2.7):

$$\begin{aligned} \hat{\mathbf{x}} &= \min_{\mathbf{x}} \|\mathbf{x}\|_1 \\ \text{s.t. } \mathbf{A}\mathbf{x} &= \mathbf{y} \\ \|\hat{\mathbf{x}}\|_0 &= L \end{aligned} \tag{2.8}$$

Note that the parameter  $L$  limits the number of sparse coefficients that appear on the sparse vector. Each sparse representation has exactly  $L$  atoms. In Figure 2.4, the fingerprint is computed using  $L = 1$  in a gallery of 20 subjects. Each patch is represented this way, transposed and then stacked vertically on a matrix called  $\mathbf{X}$  as shown on (2.9), where  $\mathbf{x}_{hw}$  is the sparse representation of the patch  $\mathbf{y}_{hw}^t$ .

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{11} \\ \mathbf{x}_{12} \\ \vdots \\ \mathbf{x}_{hw} \\ \vdots \\ \mathbf{x}_{\sqrt{m'}\sqrt{m'}} \end{bmatrix} \tag{2.9}$$

To simplify the notation, the rows of  $\mathbf{X}$  will be called  $\mathbf{x}_f$  ( $f = 1 \dots m'$ ). To every  $\mathbf{x}_f$  a previous filter is made using the sparsity concentration index (SCI). SCI of each patch is computed in order to evaluate how spread are its sparse coefficients. SCI is defined by:

$$S_f := \text{SCI}(\mathbf{x}_f) = \frac{k \max_i \|\delta_i(\mathbf{x}_f)\|_1 / \|\mathbf{x}_f\|_1 - 1}{k - 1} \tag{2.10}$$

where  $\delta_i(\mathbf{x}_f)$  is a vector of the same size as  $\mathbf{x}_f$  whose only nonzero entries are the entries in  $\mathbf{x}_f$  corresponding to subject  $i$ . The rows of  $\mathbf{X}$  that have a SCI higher than a threshold  $\theta$  form the selection matrix  $\mathbf{X}'$ . For each row of  $\mathbf{X}'$ , the highest sparse coefficient is set to

one and the other entries are set to zero. That way, the rows contains only one non-zero entry.

An refinement is made on  $\mathbf{X}'$  before it is turned into the final fingerprint  $\mathbf{F}$  of  $\mathbf{I}^t$ . A binarization of the vector is made as follows:

$$\mathbf{F}(x, y) = \begin{cases} 1 & \mathbf{X}'(x, y) \neq 0 \\ 0 & \mathbf{X}'(x, y) = 0 \end{cases} \quad (2.11)$$

In Figure 2.4 we can see fingerprints made with 20 subjects and a dictionary with  $Q = 10$  parent and  $R = 5$  child clusters. It is clear how the higher sparse coefficients are concentrated in the orange areas, that correspond to the identity of the image in question.

### 2.2.3. Classification

Once  $\mathbf{F}$  is computed we proceed to classify it. The first step is to vertically sum the columns of  $\mathbf{F}$  to obtain a one-dimensional vector with the accumulated sum of every sparse coefficient. This vector will be called  $\mathbf{f}_t$  (a graphic view of this is illustrated on Figure 2.5). The computation of this vector is done as follows:

$$\mathbf{f}_t(x) = \sum_{i=1}^{m'} \mathbf{F}(x, i) \quad (2.12)$$

Once  $\mathbf{f}_t$  is obtained, the classification is made according to:

$$\hat{i} = \underset{i}{\operatorname{argmax}} \|\delta_i(\mathbf{f}_t)\|_1 \quad (2.13)$$

It is worth mentioning that in (2.13),  $\|\cdot\|_1$  it is the same as  $\|\cdot\|_0$  since the vector  $\delta_i(\mathbf{f}_t)$  is binary. This means that the class that accumulates more sparse coefficients along the rows of  $\mathbf{F}$  will be chosen as the identity of the image  $\mathbf{I}^t$ . The vector  $\delta_i(\mathbf{f}_t)$  is the same used in (2.10).

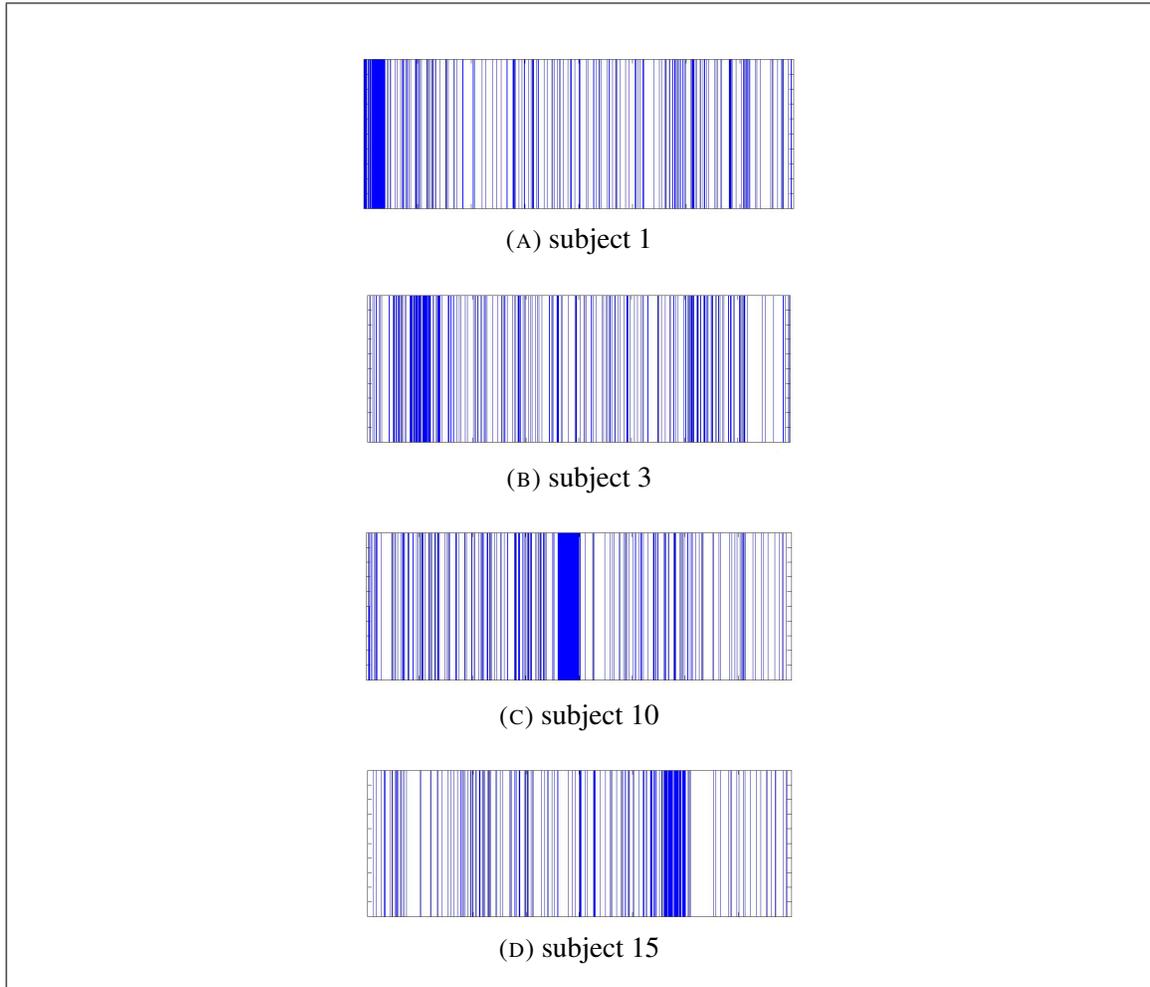


FIGURE 2.5. Different  $\mathbf{f}_t$  vector of query images that correspond to subjects 1, 3, 10 and 15 (of 20). Here we can see how the sparse coefficients are concentrated on the area that corresponds to the correct identity of  $\mathbf{I}^t$

### 2.3. Testing Methodology

We evaluate the performance of our SFCA approach by comparison with a number of recently published algorithms. We compare to each algorithm using the database and the experimental protocol (number of sample images for the learning) used in the paper about that algorithm.

In the databases, there were  $K'$  subjects and more than  $N$  images per subject. All images were resized to  $100 \times 100$  pixels and converted to a grayscale image if necessary.

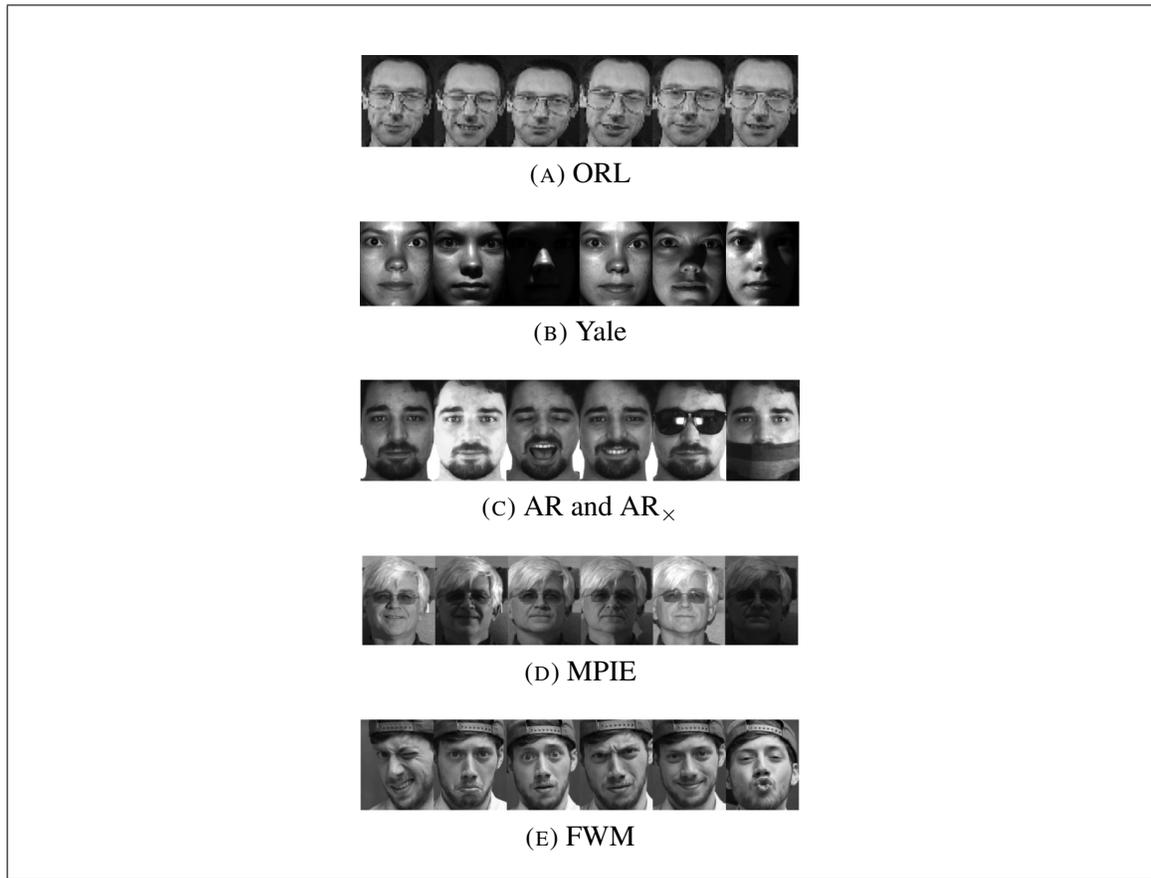


FIGURE 2.6. Examples of the databases used in our experiments

In each dataset, we collected all available images for each subject, *e.g.*, gallery images, different aging, illumination conditions, expressions, camera distances, etc.. We defined the following protocol: from these  $K'$  subjects, we randomly selected  $K \leq K'$  subjects. From each selected subject,  $N$  images were randomly chosen for training and one for testing. In order to obtain a better confidence level in the estimation of face recognition accuracy, the test was repeated 50 times by randomly selecting new  $K$  subjects and  $N$  images for training and one for testing each time. The performing metric  $\eta$  is the average of this 50 experiments.

The following presents the most important results, along with a comparison with well-known methods. The comparison with the ASR+ method are with rounded results, since the results in Mery & Bowyer (2014) are presented that way.

### 3. EXPERIMENTAL RESULTS AND IMPLEMENTATION

In our experiments we used 5 well-known databases. In Figure 2.6 there are 6 example faces of one subject of each database. The method was tested in three different conditions: lighting, expression and real occlusion. The results of these methods are extracted directly from the paper they were published, this explain the selection of the amount of training images used on each database. The testing method explained in the previous paragraph it was found to be a sufficiently robust and randomized way to measure a certain method, in order to compare it with any other form of performance measurement.

#### 3.1. Experiments under different lighting conditions

Two of the five databases used have varied lighting conditions. The first is the original and extended ‘Yale Database B’ (Lee et al. (2005)) (known as Yale). It consists of 38 subjects with 64 different images taken with many variations of lighting conditions. In this case, we use the Tan-Triggs illumination normalization (Tan & Triggs (2010)) that obtains better results than the raw images. (An example of what Tan-Triggs does can be seen in Figure 3.1). The other database is the ‘Multi-PIE’ database (Gross et al. (2010)) (will be called MPIE from now on). It contains more than 750,000 images taken from 337 subjects in four different sessions showing different expressions under 15 viewpoints and 19 illumination conditions. In our experiments, we used the frontal viewpoint only with all illuminations, expressions and sessions. All face images were cropped using the same fixed coordinates, thus the horizontal and vertical alignment of the faces varies between images. The results of this experiments can be seen in Tables 3.1 and 3.2. For Yale, our algorithm outperforms every method but ASR+, that only wins in two of the six experiments, and equals in the other five. In the case of MPIE, SFCA outperforms or equals all the other methods in the table. With  $N = 20$  and  $N = 30$  training images the results are 100%, with no misclassified images in any of the iterations of the experiment.

TABLE 3.1. Comparison on Yale ( $K = 38$ ).

	Method ( $X$ )	$\eta_X$ [%]	$\eta_{SFCA}$ [%]
$N = 4$	ASRC (J. Wang et al. (2014a))	77	<b>86</b>
$N = 5$	ASRC	77	<b>91</b>
$N = 6$	ASRC	83	<b>92</b>
$N = 7$	ASRC	83	<b>95</b>
$N = 10$	L21FLDA (Shi et al. (2014))	84	<b>99</b>
	ASR+ (Mery & Bowyer (2014))	<b>99</b>	
$N = 15$	LC-KSVD (Jiang et al. (2013))	95	99
	ASR+	<b>100</b>	
$N = 16$	DLRR (J. Chen & Yi (2014))	96	99
	ASR+	<b>100</b>	
$N = 20$	L21FLDA	94	<b>100</b>
	ASR+	<b>100</b>	
$N = 30$	L21FLDA	97	<b>100</b>
	ASR+	<b>100</b>	
$N = 32$	LGE-KSVD (Ptucha & Savakis (2013))	96	<b>100</b>
	DLRR	99	
	ASR+	<b>100</b>	
$N = 33$	InfoMax (Qiu et al. (2014))	96	<b>100</b>
	LC-KSVD	97	
	ASR+	<b>100</b>	

TABLE 3.2. Comparison on MPIE ( $K = 68$ ).

	Method ( $X$ )	$\eta_X$ [%]	$\eta_{SFCA}$ [%]
$N = 10$	L21FLDA (Shi et al. (2014))	86	<b>100</b>
	ASR+ (Mery & Bowyer (2014))	98	
$N = 12$	DLRR (J. Chen & Yi (2014))	94	<b>100</b>
	ASR+	96	
$N = 20$	L21FLDA	92	<b>100</b>
	ASR+	<b>100</b>	
$N = 30$	L21FLDA	95	<b>100</b>
	ASR+	<b>100</b>	

### 3.2. Experiments with subjects with different expressions

The databases with different facial expressions are the ORL database (Samaria & Harter (1994)) and the ‘Face We Make’ (Miranda (2011)) (also called FWM) database.

TABLE 3.3. Comparison on ORL ( $K = 40$ ).

	Method ( $X$ )	$\eta_X$ [%]	$\eta_{SFCA}$ [%]
$N = 1$	RNS (Borgi et al. (2014))	88	<b>89</b>
$N = 2$	ASRC (J. Wang et al. (2014a))	82	<b>96</b>
$N = 3$	Bayes (Ouarda et al. (2013))	79	<b>98</b>
	L21FLDA (Shi et al. (2014))	82	
	ASRC	89	
	ASR+ (Mery & Bowyer (2014))	94	
$N = 4$	ASRC	93	<b>99</b>
$N = 5$	ASRC	96	<b>100</b>
	L21FLDA	93	
	LRC (Naseem et al. (2010))	94	
	ASR+	99	
$N = 7$	L21FLDA	97	<b>100</b>
	ASR+	99	
$N = 8$	PCA-LDA (Verma & Sahu (2013))	96	<b>100</b>
$N = 9$	LRC	99	<b>100</b>

ORL consists of 40 subjects with 10 different images taken with very small variation of lighting, face expressions and face details (glasses / no glasses). FWM contains images from 224 subjects (140 women and 84 men) with 10 different expressions that convey feelings related to common emoticons, *e.g.*, :) smile, :-O surprised, :( sad, etc. In both databases it is shown in Tables 3.3 and 3.4 that our method outperforms every other method in the comparison. It is worth saying that in ORL, the result with only 2 training images is more than 14% better than ASRC and with  $N = 3$  training images outperforms ASR+ by more than 4%. With 9 training images the average results is no less than 100%, not a single image was misclassified. In FWM we can appreciate that SFCA outperforms by more than 28% to DICW with  $N = 1$  and the difference is larger than 17% in every experiment with this algorithm.

### 3.3. Experiments with real occlusion and face expressions

The databases that are used to test this condition are AR and AR<sub>×</sub>. The images of AR MB98 (1998) were taken from 100 subjects (50 women and 50 men) with different facial expressions, illumination conditions, and occlusions with sunglasses and scarf (cropped

TABLE 3.4. Comparison on FWM ( $K = 55$ ).

	Method ( $X$ )	$\eta_X$ [%]	$\eta_{SFCA}$ [%]
$N = 1$	DICW (Wei et al. (2013))	62	<b>91</b>
$N = 3$	DICW	76	<b>98</b>
$N = 5$	DICW	77	<b>99</b>
$N = 8$	DICW	82	<b>100</b>
	ASR+ (Mery & Bowyer (2014))	97	

TABLE 3.5. Comparison on AR ( $K = 100$ ).

	Method ( $X$ )	$\eta_X$ [%]	$\eta_{SFCA}$ [%]
$N = 5$	LC-KSVD (Jiang et al. (2013))	94	<b>97</b>
	ASR+ (Mery & Bowyer (2014))	95	
$N = 7$	DLRR (J. Chen & Yi (2014))	94	<b>99</b>
	ASRC (J. Wang et al. (2014a))	95	
	ASR+	98	
$N = 9$	DLRR	90	<b>99</b>
	ASR+	97	
$N = 13$	SSRC (Deng et al. (2013))	99	<b>100</b>
	ASR+	<b>100</b>	
$N = 20$	DKSVD (Zhang & Li (2010))	95	<b>100</b>
	LC-KSVD	98	
	ASR+	<b>100</b>	

version is used). The number of images per subject is 26. We distinguish between AR and  $AR_{\times}$ : In AR, training and testing images are selected randomly from the 26 available images; whereas in  $AR_{\times}$ , training images are selected randomly from the images with no disguise, and testing from the images with disguise. Figure 2.6c shows an example of this images.

The results of Table 3.5 show that SFCA works well with real occlusions (sunglasses and scarves). The results with 5, 7 and 9 training images show improvements over other published methods. In the experiments of  $AR_{\times}$  (Table 3.6), SFCA shows improvements over many methods, and ASR+ has the best results.

TABLE 3.6. Comparison on  $AR_{\times}$  ( $K = 100$ ).

	Method ( $X$ )	$\eta_X$ [%]	$\eta_{SFCA}$ [%]
$N = 8$	LRC Naseem et al. (2010)	61	97
	$\ell_{struct}$ Jia et al. (2012)	94	
	SEC-MRF Zhou et al. (2009)	97	
	MLERPM Weng et al. (2013)	98	
	DICW Wei et al. (2013)	99	
	ASR+ Mery & Bowyer (2014)	<b>100</b>	

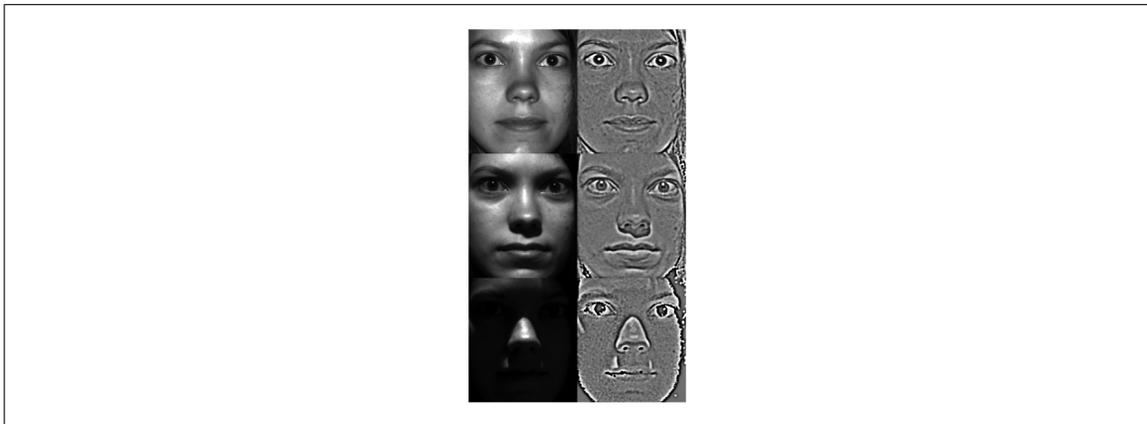


FIGURE 3.1. Example of how Tan-Triggs normalization works in different lighting conditions on the Yale database

### 3.4. Implementation

The machine used to perform the experiments was a MacBook Pro OS X 10.9.4 processor 2.5 GHz Intel Core i5 with 4 cores and memory of 4 GB RAM 1600 MHz DDR3. The algorithm is implemented in Python programming language. NumPy Dubois et al. (1996), SciPy Jones et al. (2001–), scikit-learn Pedregosa et al. (2011), OpenCV Bradski (2000) and SPAMS Mairal et al. (2010) libraries are used.

### 3.5. Parameters Sensitivity Analysis

To further analyze our method, sensitivity analyses were made over four of the more important parameters. This was made in order to tune the parameters that have more impact in the performance of the algorithm.

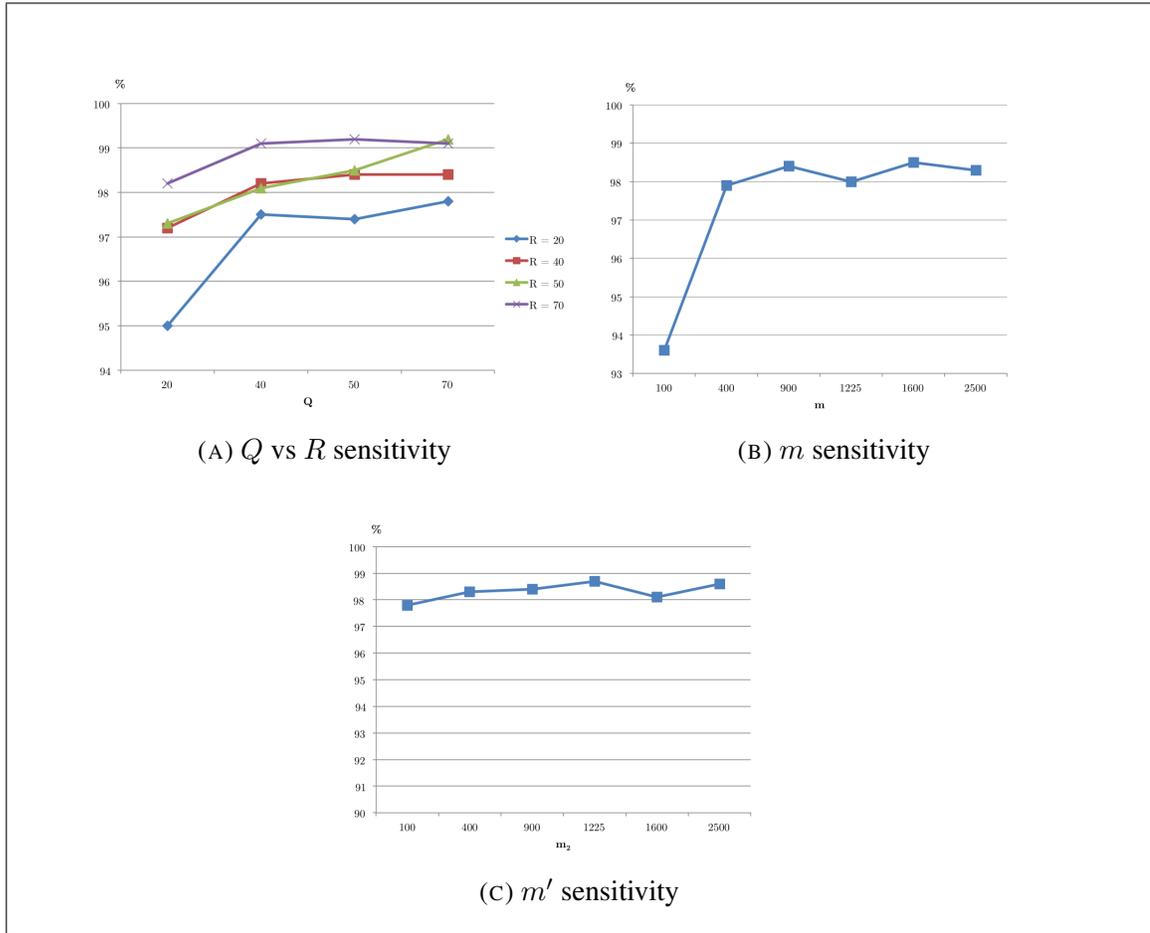


FIGURE 3.2. Sensitivity analyses for the most important parameters of the model.  $Q$ ,  $R$  and  $m$  have more influence in the final result than the number of patches in the grid,  $m'$ .

To perform this analysis, a random test was made over the AR database with  $K = 20$  subjects and  $N = 4$  images to compute the dictionary. The same set of subjects and pictures was used in every experiment, to more directly reflect the change due only to a parameter.

The values used for the sensitivity study were  $m = 1225$  patches for the training grid,  $m' = 900$  for the testing grid, both with patches of  $20 \times 20$  pixels. The weighting coefficient  $\alpha$  for the center of the patch was 0.5,  $Q = 50$  parent clusters,  $R = 40$  child clusters,  $L = 4$  atoms for the  $\ell_1$ -minimization constrain and a thershold of  $\theta = 0.1$  for the SCI selection.

The parameters analyzed for sensitivity were the following:

(i) Analysis of  $Q$  vs  $R$ : These two parameters are the number of parent clusters ( $Q$ ) and child clusters ( $R$ ). Since both parameters are closely tied in with the definition of the dictionary, we perform tests varying both to evaluate the behavior of the method. Figure 3.2a gives the results of this experiment. We can appreciate that if both values are low, the performance of the method is poor, but performance increases considerably as either parameter is increased.

(ii) Analysis of  $m$ :

The parameter  $m$  defines the number of patches over the training grid to compute the dictionaries. Figure 3.2b shows the importance of extracting a large number of patches.  $m = 100$  shows a poor performance in comparison with the values over 400.

(iii) Analysis of  $m'$ : After evaluating the behaviour of this parameter the conclusion is that it is much less important than the others. From  $m' = 100$  to  $m' = 2500$  the performance of the algorithm only varies less than 1%, and every time is considerably high.

## 4. CONCLUSION

We introduced a new approach to face recognition, the Sparse Fingerprint Classification Algorithm. SFCA has demonstrated high accuracy under a large number of different conditions, such as variations in ambient light, pose, occlusion, size of the face and distance from the camera. SFCA's simplicity and effectiveness are due to it working with a binary sparse matrix. Advantages over previous methods are that SFCA doesn't require sparse reconstruction and is based only on the sparse coefficient vector.

We have extensively evaluated SFCA and compared it with other state-of-art methods. The approach to the evaluation experiments with SFCA, using the same datasets as used in evaluating other state-of-the-art methods, is meant to ensure its robustness and shows that SFCA achieves improved accuracy in face recognition under variations in ambient lighting, pose, expression, face size, occlusion and distance from the camera. From a total of 33 different experiments, SFCA outperforms or equals the methods in comparison in 30, and being outperformed only in 3.

Analysing the results of the algorithm, the strengths of it are manily two: it works as an all-around method, with good performance in many different situations such as the one tested and it does not need many training images to obtain good results. The experiments that shows a weak point of the method were those when all training images have occlusion (*i.e.*  $AR_{\times}$ ), where other three methods works better. Working in a way to eliminate the information of the occluded patches (assuming these are the ones that produce the errors) from the training phase effectively, could help to overcome these situations.

The novel approach of the fingerprints used here differs from similar concepts used in audio processing because the fingerprint itself carries information about the subject that it belongs to. In this way there is no need to have a query database and make searches to identify the class of the fingerprint. Using only sparse binary matrices, a subject face image can be classified correctly with high accuracy.

## References

- Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12), 2037–2041.
- Allamanche, E., Herre, J., Hellmuth, O., Fröba, B., Kastner, T., & Cremer, M. (2001). Content-based identification of audio material using mpeg-7 low level description. In *Ismir*.
- Baluja, S., & Covell, M. (2007, April). Audio fingerprinting: Combining computer vision data stream processing. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing* (Vol. 2, p. II-213-II-216).
- Boiman, O., Shechtman, E., & Irani, M. (2008). In defense of nearest-neighbor based image classification. In *2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1–8).
- Borgi, M., Labate, D., El'arbi, M., & Ben Amar, C. (2014, May). Regularized Shearlet Network for face recognition using single sample per person. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (p. 514-518). doi: 10.1109/ICASSP.2014.6853649
- Bradski, G. (2000). *Dr. Dobb's Journal of Software Tools*.
- Burges, C. J., Plastina, D., Platt, J. C., Renshaw, E., & Malvar, H. (2005b, March). Using audio fingerprinting for duplicate detection and thumbnail generation. In *2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (Vol. 3, p. iii/9-iii12 Vol. 3). doi: 10.1109/ICASSP.2005.1415633
- Burges, C. J., Plastina, D., Platt, J. C., Renshaw, E., & Malvar, H. S. (2005a). Using audio fingerprinting for duplicate detection and thumbnail generation. In *2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (Vol. 3, pp. iii–9).
- Camarena-Ibarrola, A., Chávez, E., & Tellez, E. S. (2009). Robust radio broadcast monitoring using a multi-band spectral entropy signature. In *Progress in pattern recognition*,

*image analysis, computer vision, and applications* (pp. 587–594). Springer.

Chai, Z., Sun, Z., Mendez-Vazquez, H., He, R., & Tan, T. (2014, Jan). Gabor Ordinal Measures for Face Recognition. *IEEE Transactions on Information Forensics and Security*, 9(1), 14-26. doi: 10.1109/TIFS.2013.2290064

Chen, J., & Yi, Z. (2014). Sparse representation for face recognition by discriminative low-rank matrix recovery. *Journal of Visual Communication and Image Representation*, 25(5), 763–773.

Chen, Y., Do, T. T., & Tran, T. D. (2010). Robust face recognition using locally adaptive sparse representation. In *2010 IEEE International Conference on Image Processing (ICIP)* (pp. 1657–1660).

Deng, W., Hu, J., & Guo, J. (2012). Extended SRC: Undersampled face recognition via intraclass variant dictionary. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(9), 1864–1870.

Deng, W., Hu, J., & Guo, J. (2013). In defense of sparsity based face recognition. In *2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 399–406).

Dubois, P. F., Hinsen, K., & Hugunin, J. (1996, May/June). Numerical python. *Computers in Physics*, 10(3).

Gross, R., Matthews, I., Cohn, J., Kanade, T., & Baker, S. (2010). Multi-pie. *Image and Vision Computing*, 28(5), 807–813.

Ibarrola, A., & Chavez, E. (2006, July). A robust entropy-based audio-fingerprint. In *2006 IEEE International Conference on Multimedia and Expo* (p. 1729-1732). doi: 10.1109/ICME.2006.262884

Jia, K., Chan, T.-H., & Ma, Y. (2012). Robust and practical face recognition via structured sparsity. In *Computer vision—eccv 2012* (pp. 331–344). Springer.

Jiang, Z., Lin, Z., & Davis, L. S. (2013). Label consistent k-svd: learning a discriminative dictionary for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11), 2651–2664.

Jones, E., Oliphant, T., Peterson, P., et al. (2001–). *SciPy: Open source scientific tools for Python*. Retrieved from <http://www.scipy.org/> ([Online; accessed 2014-12-22])

- Kamaladas, M. D., & Dialin, M. M. (2013). Fingerprint extraction of audio signal using wavelet transform. In *Signal processing image processing & pattern recognition (icsipr), 2013 international conference on* (pp. 308–312).
- Lee, K.-C., Ho, J., & Kriegman, D. (2005). Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5), 684–698.
- Li, Z., Zhang, Q., Duan, X., & Zhao, F. (2014, April). Face recognition based on regression analysis using frequency features. In *2014 IEEE International Conference on Information Science and Technology (ICIST)* (p. 192-195). doi: 10.1109/ICIST.2014.6920363
- Lu, J., Tan, Y.-P., Wang, G., & Yang, G. (2013, June). Image-to-Set Face Recognition Using Locality Repulsion Projections and Sparse Reconstruction-Based Similarity Measure. *IEEE Transactions on Circuits and Systems for Video Technology*, 23(6), 1070-1080. doi: 10.1109/TCSVT.2013.2241353
- Mairal, J., Bach, F., Ponce, J., & Sapiro, G. (2010). Online learning for matrix factorization and sparse coding. *The Journal of Machine Learning Research*, 11, 19–60.
- Martínez, A., & Benavente, R. (1998, Jun). *The AR Face Database* (Tech. Rep. No. 24). Bellatera: Computer Vision Center. Retrieved from <http://www.cat.uab.cat/Public/Publications/1998/MaB1998>
- Mery, D., & Bowyer, K. (2014). Face Recognition via Adaptive Sparse Representations of Random Patches. In *2014 IEEE Workshop on Information Forensics and Security (WIFS)*.
- Miranda, D. (2011). *The Face We Make*. <http://thefacewemake.org>.
- Naseem, I., Togneri, R., & Bennamoun, M. (2010, Nov). Linear Regression for Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(11), 2106-2112. doi: 10.1109/TPAMI.2010.128
- Ouali, C., Dumouchel, P., & Gupta, V. (2014). A robust audio fingerprinting method for content-based copy detection. In *2014 IEEE International Workshop on Content-Based Multimedia Indexing (CBMI)* (pp. 1–6).
- Ouarda, W., Trichili, H., Alimi, A., & Solaiman, B. (2013, Dec). Combined local features selection for face recognition based on Naïve Bayesian classification. In

- 2013 international conference on hybrid intelligent systems (his) (p. 240-245). doi: 10.1109/HIS.2013.6920489
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... others (2011). Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research*, 12, 2825–2830.
- Phillips, P. J., Beveridge, J. R., Draper, B. A., Givens, G., O’Toole, A. J., Bolme, D. S., ... Weimer, S. (2011). An introduction to the good, the bad, & the ugly face recognition challenge problem. In *2011 IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG)* (pp. 346–353).
- Ptucha, R., & Savakis, A. (2013). LGE-KSVD: Flexible Dictionary Learning for Optimized Sparse Representation Classification. In *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 854–861).
- Qiu, Q., Patel, V., & Chellappa, R. (2014, Nov). Information-theoretic dictionary learning for image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(11), 2173-2184. doi: 10.1109/TPAMI.2014.2316824
- Samaria, F. S., & Harter, A. C. (1994). Parameterisation of a stochastic model for human face identification. In *1994 Proceedings of the IEEE Workshop on Applications of Computer Vision*, (pp. 138–142).
- Shi, X., Yang, Y., Guo, Z., & Lai, Z. (2014). Face recognition by sparse discriminant analysis via joint  $l_{2,1}$ -norm minimization. *Pattern Recognition*, 47(7), 2447–2453.
- Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1701–1708).
- Tan, X., Chen, S., Zhou, Z.-H., & Liu, J. (2009). Face recognition under occlusions and variant expressions with partial similarity. *IEEE Transactions on Information Forensics and Security*, 4(2), 217–230.
- Tan, X., & Triggs, B. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, 19(6), 1635–1650.

- Verma, T., & Sahu, R. (2013, March). PCA-LDA based face recognition system amp; results comparison by various classification techniques. In *2013 IEEE International Conference on Green High Performance Computing (ICGHPC)* (p. 1-7). doi: 10.1109/ICGHPC.2013.6533913
- Wagner, A., Wright, J., Ganesh, A., Zhou, Z., Mobahi, H., & Ma, Y. (2012). Toward a practical face recognition system: Robust alignment and illumination by sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(2), 372–386.
- Wang, A., et al. (2003). An Industrial Strength Audio Search Algorithm. In *Ismir* (pp. 7–13).
- Wang, J., Lu, C., Wang, M., Li, P., Yan, S., & Hu, X. (2014a). Robust Face Recognition via Adaptive Sparse Representation. *IEEE Transactions on Cybernetics*(99), 1.
- Wang, J., Lu, C., Wang, M., Li, P., Yan, S., & Hu, X. (2014b, Dec). Robust face recognition via adaptive sparse representation. *IEEE Transactions on Cybernetics*, *44*(12), 2368-2378. doi: 10.1109/TCYB.2014.2307067
- Wei, X., Li, C.-T., & Hu, Y. (2012). Robust face recognition under varying illumination and occlusion considering structured sparsity. In *2012 IEEE International Conference on Digital Image Computing Techniques and Applications (DICTA)* (pp. 1–7).
- Wei, X., Li, C.-T., & Hu, Y. (2013). Face recognition with occlusion using dynamic image-to-class warping DICW. In *2013 IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)* (pp. 1–6).
- Wei, X., Li, C.-T., Lei, Z., Yi, D., & Li, S. (2014, Dec). Dynamic Image-to-Class Warping for Occluded Face Recognition. *IEEE Transactions on Information Forensics and Security*, *9*(12), 2035-2050. doi: 10.1109/TIFS.2014.2359632
- Weng, R., Lu, J., Hu, J., Yang, G., & Tan, Y.-P. (2013, Dec). Robust Feature Set Matching for Partial Face Recognition. In *2013 IEEE International Conference on Computer Vision (ICCV)* (p. 601-608). doi: 10.1109/ICCV.2013.80
- Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., & Ma, Y. (2009). Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(2), 210–227.

Xu, Y., Zhang, D., Yang, J., & Yang, J.-Y. (2011a). A two-phase test sample sparse representation method for use with face recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 21(9), 1255–1262.

Xu, Y., Zhang, D., Yang, J., & Yang, J.-Y. (2011b). A Two-Phase Test Sample Sparse Representation Method for Use With Face Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 21(9), 1255–1262.

Zhang, Q., & Li, B. (2010). Discriminative K-SVD for dictionary learning in face recognition. In *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2691–2698). IEEE.

Zhou, Z., Wagner, A., Mobahi, H., Wright, J., & Ma, Y. (2009, Sept). Face recognition with contiguous occlusion using Markov random fields. In *2009 IEEE International Conference on Computer Vision* (p. 1050-1057). doi: 10.1109/ICCV.2009.5459383