PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
SCHOOL OF ENGINEERING

BLUR ADAPTIVE SPARSE REPRESENTATION OF RANDOM PATCHES FOR FACE RECOGNITION ON BLURRED IMAGES

DANIEL IGNACIO HEINSOHN MANETTI

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, January 2016

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
SCHOOL OF ENGINEERING

BLUR ADAPTIVE SPARSE REPRESENTATION OF RANDOM PATCHES FOR FACE RECOGNITION ON BLURRED IMAGES

DANIEL IGNACIO HEINSOHN MANETTI

Members of the Committee:
DOMINGO MERY QUIROZ, PH.D.
CRISTIÁN TEJOS NUÑEZ, PH.D.
GABRIEL HERMOSILLA VIGNEAU, PH.D.
LUIS CIFUENTES LIRA, 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, January 2016

Quiero dar las gracias a todos los que me acompañaron en los momentos difíciles y también en los momentos alegres. Agradecer a todos aquellos que tuvieron y/o desarrollaron la paciencia hacia mi ser y mis seres queridos. Para mi madre y padre, toda mi familia, mis amigos, pero en especial para mi Nonna Nena

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 Católica de Chile and the College of Engineering at the University of Notre Dame.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Although unconstrained face recognition has been widely studied over the last 25 years, state-of-the-art algorithms continue to have low performance on blurred images. In this paper, we address the problem of face recognition in blurred query images by using an algorithm that is a modification of the Adaptive Sparse Representation of Random Patches (ASR+) algorithm. This algorithm extracts random patches from the images of previously enrolled subjects to create representative dictionaries. ASR+ then builds an adaptive dictionary concatenating the 'best' dictionary of each subject. Finally, using this adaptive dictionary, each test patch is classified following the Sparse Representation Classification (SRC) methodology. In the testing stage, the query image is classified through patch voting. In this article, we propose bASR+ (blur ASR+) to create additional (synthetic) dictionaries with different levels of blurriness computed by using Gaussian blur kernels. Each blurred set of synthetic images will be described by a global 'sharpness metric'. When the query image arrives, the sharpness metric s of the query image is computed and compared to the sharpness metrics of all dictionaries. We then select the dictionary that has the most similar metric value to s. Thus, an adaptive dictionary is constructed from the selected blurred dataset, and ASR+ methodology is used for the recognition. We designed an experiment based on blurred face images that were created from out-of-focus pictures taken from face images of AR database displayed on a monitor. In our experiments, bASR+ outperforms state-of-the-art algorithms and commercial software, obtaining 85.6% of recognition in overly blurred face images.

Keywords: Face recognition, blurred face images, sparse representations, patches.

RESUMEN

A pesar de que el reconocimiento de caras, en ambientes no controlados, ha sido ampliamente estudiado en los últimos 25 años, el estado del arte en estos algoritmos continúan con bajos rendimientos en imágenes borrosas. En este documento se aborda el reconocimiento de rostros en imágenes que se encuentran borrosas, al ejecutar un algoritmo que es una modificación del algoritmo *Adaptive Sparse Representation of Random Patches* (ASR+). Este algoritmo extrae, al azar, parches de cada una de las imágenes de sujetos previamente enlistados. Luego, ASR+ construye un diccionario adaptivo el cual concatena los 'mejores' diccionarios de cada sujeto. Finalmente, usando este diccionario, cada parche de la imagen de prueba es clasificado acorde a la metodología de *Sparse Representation Classification* (SRC). En la fase de prueba, la imagen en cuestión es clasificada por mayoría de votos de los parches previamente clasificados. En este artículo proponemos bASR+, o *blur* ASR+, que crea diccionarios con distintos niveles de desenfoque sintético a partir de distintos niveles de desenfoque gaussiano. Cada set de imágenes borrosas será descrito por una métrica de nitidez. Cuando se inicia el reconocimiento de la imagen de prueba, la métrica de nitidez s de la imagen es calculada y comparada a la métrica de nitidez de cada uno de los set de imágenes. Luego, se selecciona el diccionario que más se parece a la métrica s. Así, un diccionario adaptivo es contruído a partir del set de imágenes borrosas y la metodología ASR+ continúa para realizar el reconocimiento. Se diseñaron experimentos basados en imágenes borrosas, que fueron creadas tomando fotografías des-enfocadas, a las imágenes de cara de sujetos de la base de datos AR, desplegadas en un monitor. En nuestros experimentos, bASR+ sobrepasa a los algoritmos del estado del arte y *software* comerciales obteniendo reconocimientos de 85.6% en imágenes sumamente borrosas.

Palabras Claves: reconocimiento facial, imágenes de caras borrosas, representaciones sparse, parches.

1. INTRODUCTION

Face recognition is still an active area of research in computer vision, making important contributions since the 1990's. Nowadays, face recognition algorithms are dealing with unconstrained conditions which simulate real world environments. In unconstrained images, there are variations on ambient lightning, pose, expression, face size, and distance from the camera ,Phillips et al. (2011). Other challenging characteristics present in these unconstrained conditions are occlusion and/or blur. In this research, we address one of the biggest problems: face recognition in blurred images.

There are many different types of blurriness in real world environments ($e.g.$, relative motion, atmospheric turbulence, and out-of-focus lens). When it comes to recognition, many algorithms have been proposed. On the one hand, there are inverse methods based on deblurring, in which image restoration is performed. Among these algorithms, we can find blind deconvolution, Levin et al. (2011), non-blind deconvolution, Yuan et al. (2008), regularization methods on total variation, Rudin et al. (1992), and Tikhonov regularization, Tikhonov & Arsenin (1977). On the other hand, there are more direct methods based on features of the image that are invariant to blurriness, such as processing images in spatial and frequency domains ($e.g.$, Flusser & Suk (1998), Gopalan et al. (2012)).

Algorithms based on Sparse Representation Classification (SRC) have been widely used, Wright & a.Y. Yang (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 that has the least reconstruction error. Several 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 Wang et al. (2014), sparsity and correlation are jointly considered. In Jia et al. (2012) and Wei et al. (2012), structured sparsity is proposed for dealing with the problem of occlusion and illumination. In Deng et al. (2013), the dictionary is assembled by the class centroids and sample-to-centroid difference. 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 by 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 $l_{1,2}$-norm. In Xu et al. (2011), 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), patches that do not give information ($e.g.$, occluded parts) are not considered in the recognition. Moreover, distinguished parts from images ($e.g.$, birthmarks, moles, eyebrows, etc.) that are subject-dependant are found with unsupervised learning. Also, featured extraction is performed randomly in order to correct misalignment. These variations improve recognition performance significantly as they are able to model various corruptions in face images, such as misalignment and occlusion.

As shown in previous works, Aggarwal et al. (2012), sharpness metric has been proven to be sensitive to blurriness. Inspired by these key ideas, this paper proposes a new method for face recognition in blurred images when dealing with unconstrained conditions. Our main contributions are: (i) implementation and use of additional dictionaries with different levels of blurriness, and (ii) description of these dictionaries by a global sharpness metric. Using these new upgrades based on blur Adaptive Sparse Representation of Random Patches (bASR+), the proposed method can achieve high recognition performance under certain complex conditions with a high level of blurriness. In addition, we developed a new experimental methodology based on blurred face images that were created from out-of-focus pictures taken from sharp face images displayed on a monitor.

The rest of the paper 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

We propose an algorithm based on ASR+, *i.e., Adaptive Sparse Representation of Random Patches*,Mery & Bowyer (2014). Original ASR+ consists of two stages: learning and testing. In the learning stage, for each subject of training dataset, several random patches are extracted from their face images in order to construct representative dictionaries. In the testing stage, random test patches of the query image are extracted. For each test patch, a dictionary is built concatenating the 'best' representative dictionary of each subject. Using this adaptive dictionary, each test patch is classified following the Sparse Representation Classification (SRC) methodology ,Wright & a.Y. Yang (2009). Finally, the query image is classified by patch voting. Our proposed approach, that we call 'blur ASR+' or bASR+, consists of adding new dictionaries containing patches of the faces with different levels of blurriness to the training stage, so when a query blurred image is tested, the algorithm should recognize the face by selecting the dictionary with the same level of blurriness. Thus, when the query image is tested, the algorithm should work at the same level of blurriness, as shown in Fig. 2.1.

The algorithm consists of two stages: learning and testing. They will be explained in further detail in the following sections.

2.0.1. Learning

In this first stage, also called training, a set of n face images of k subjects is available, where $\mathbf{I}_j^i$ denotes image j of subject i (for $i = 1 \ldots k$ and $j = 1 \ldots n$). The goal of this step is to characterize the subjects of the gallery by computing synthetic blurred images, a sharpness metric, and dictionaries.

2.0.1.1. Synthetic blurred images

We create B sets of images with different levels of blurriness b, for $b = 0 \ldots B$. Blurred training images are achieved by filtering the original sharp face image with a Gaussian mask of

$$\sigma_b = \frac{b}{4}, \qquad \text{for } b = 0 \ldots B. \tag{2.1}$$

for each blur level b. In which σ_b represents the standard deviation of the Gaussian function in pixels. In this approach, $b = 0$ means the original image. In our experiments, we used $B = 14$. Thus, fourteen new sets of synthetic images with different levels of blurriness are created.

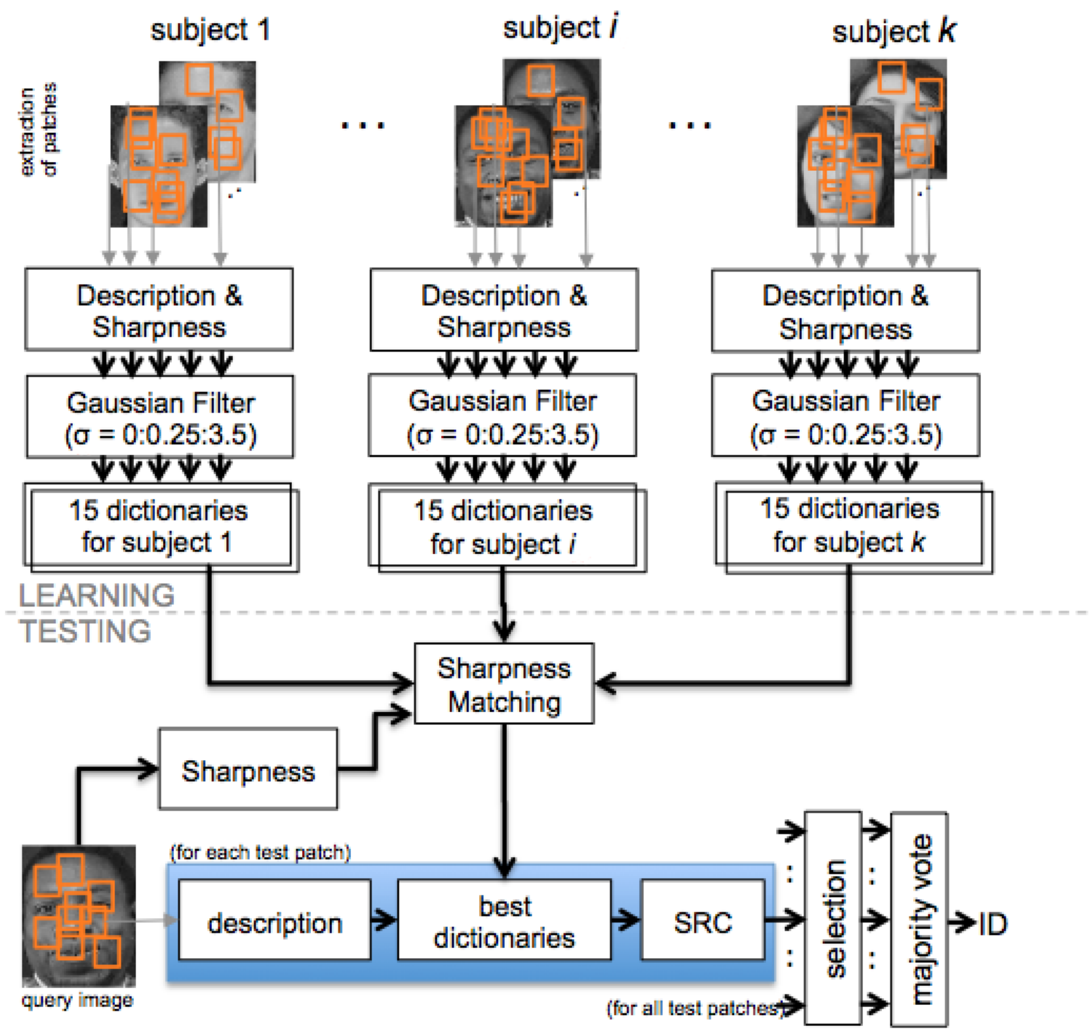

FIGURE 2.1. Overview of proposed method bASR+

2.0.1.2. Sharpness metric

A global sharpness metric for each image dataset is calculated. Sharpness metric is obtained following a Single Value Decomposition (SVD) approach, Zhu & Milanfar (2009). In this approach, the sharpness metric of an image $\mathbf{I}$ is computed in three steps:
1) gradient is computed from images using a 2D convolution:

$$\mathbf{G} = [\mathbf{G}_x\ \mathbf{G}_y] \quad \begin{aligned} \mathbf{G}_x &= \mathbf{I} * [-1\ 0\ 1] \\ \mathbf{G}_y &= \mathbf{I} * [-1\ 0\ 1]^{\mathsf{T}} \end{aligned} \tag{2.2}$$

2) SVD method is computed in order to get singular values s_1 and s_2 along with their respective directions v_1 and v_2:

$$\mathbf{G} = \mathbf{USV}^T = \mathbf{U} \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} [\ v_1 \quad v_2\ ]^T \tag{2.3}$$

3) The sharpness metric of image **I** is:

$$s = s_1(\mathbf{I}) \tag{2.4}$$

that has been proven to be useful as a sharpness metric, Aggarwal et al. (2012), Zhu & Milanfar (2009).

Then, $\mathbf{s}^b$ is a vector created out of every singular value $s_j^{ib} = s_1(\mathbf{I}_j^{ib})$, that is from image j of subject i at blurriness b.

$$\mathbf{s}^b = \begin{bmatrix} s_1^{1b} \\ \vdots \\ s_n^{1b} \\ s_1^{2b} \\ \vdots \\ s_n^{2b} \\ \vdots \\ s_1^{kb} \\ \vdots \\ s_n^{kb} \end{bmatrix} \quad \text{for } b = 0 \dots B \tag{2.5}$$

In order to characterize each set of blurriness we compute:

$$\hat{s}^b = \text{median}(\mathbf{s}^b) \tag{2.6}$$

where s^b is considered as a representative sharpness metric of this set of images. The values are stored in array $\hat{\mathrm{s}} = [\hat{s}^0 \dots \hat{s}^B]^\mathsf{T}$.

2.0.1.3. Dictionaries

Afterwards, the dictionaries of each subject of the gallery are defined. In each $\mathbf{I}_j^{ib}$, m patches $\mathcal{P}_{jp}^{ib}$ of size $w \times w$ pixels (for $p = 1 \dots m$) are randomly extracted. These patches are centred at $(x_{jp}^{ib}, y_{jp}^{ib})$. The description of a patch p is defined as a vector:

$$\mathbf{y} = f(\mathcal{P}) = [\ \mathbf{z}\ ;\ \alpha x\ ;\ \alpha y\ ] \in \mathbb{R}^{d+2} \tag{2.7}$$

where $z = g(\mathcal{P}) \in \mathbb{R}^d$ is a descriptor of patch $\mathcal{P}$, (x, y) are the image coordinates of the center of patch $\mathcal{P}$, and α is a weighting factor between description and location. Using (2.7), all extracted patches are described as $\mathbf{y}_{jp}^{ib} = f(\mathcal{P}_{jp}^{ib})$. Thus, for subject i an array with the description of all patches is defined as $\mathbf{Y}^{ib} = \{\mathbf{y}_{jp}^{ib}\} \in \mathbb{R}^{(d+2)\times nm}$ (for $j = 1 \ldots n$, $b = 0 \ldots B$ and $p = 1 \ldots m$). The description $\mathbf{Y}^{ib}$ of subject i is clustered using a k-means algorithm in Q clusters. These Q clusters will be referred as $parent$ clusters:

$$\mathbf{c}_q^{ib} = \text{kmeans}(\mathbf{Y}^{ib}, Q) \tag{2.8}$$

for $q = 1 \ldots Q$, where $\mathbf{c}_q^{ib} \in \mathbb{R}^{(d+2)}$ is the centroid of parent cluster q of subject i at blurriness level b. We define $\mathbf{Y}_q^{ib}$ as the array with all the samples $\mathbf{y}_{jp}^{ib}$ that belong to the parent cluster with centroid $\mathbf{c}_q^{ib}$.

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

$$\mathbf{c}_{qr}^{ib} = \text{kmeans}(\mathbf{Y}_q^{ib}, R) \tag{2.9}$$

for $r = 1 \ldots R$, where $\mathbf{c}_{qr}^{ib} \in \mathbb{R}^{(d+2)}$ is the centroid of child cluster r of parent cluster q of subject i at blurriness level b. All centroids of child clusters of subject i are arranged in an array $\mathbf{D}^{ib}$ (orange rectangle in figure 2.2). Parent cluster q is arranged in a matrix:

$$\bar{\mathbf{A}}_q^{ib} = [\mathbf{c}_{q1}^{ib} \ldots \mathbf{c}_{qr}^{ib} \ldots \mathbf{c}_{qR}^{ib}] \in \mathbb{R}^{(d+2)\times R} \tag{2.10}$$

Thus, this arrangement contains R representative samples of parent cluster q of subject i as illustrated in Figure 2.2. The set of all centroids of child clusters of subject i ($\mathbf{D}^{ib}$), represents Q representative dictionaries with R descriptions $\{\mathbf{c}_{qr}^{ib}\}$ for $q = 1...Q,\ r = 1...R$.

2.0.2. Testing

The goal of this stage is to determine the identity of the subject in the query image $\mathbf{I}^t$. As this query image could be degraded ($i.e.$, out of focus blur, movement blur, down sampled), we need to first match the query image sharpness with the sharpness $\mathbf{s}$ of the different levels of blurriness of dictionaries $\mathbf{D}^{ib}$.

2.0.2.1. Matching

In this part of the training, we start by extracting the sharpness metric from query image $\mathbf{I}^t$ as $s^t = s_1(\mathbf{I}^t)$ in the same way we did in the previous stage with equation (2.4). For each level of blurriness in $\hat{\mathbf{s}}$, we look for the one which minimizes the distance between sharpness values:

$$b^t = \underset{b}{\text{argmin}} ||s^t - \hat{s}^b|| \tag{2.11}$$

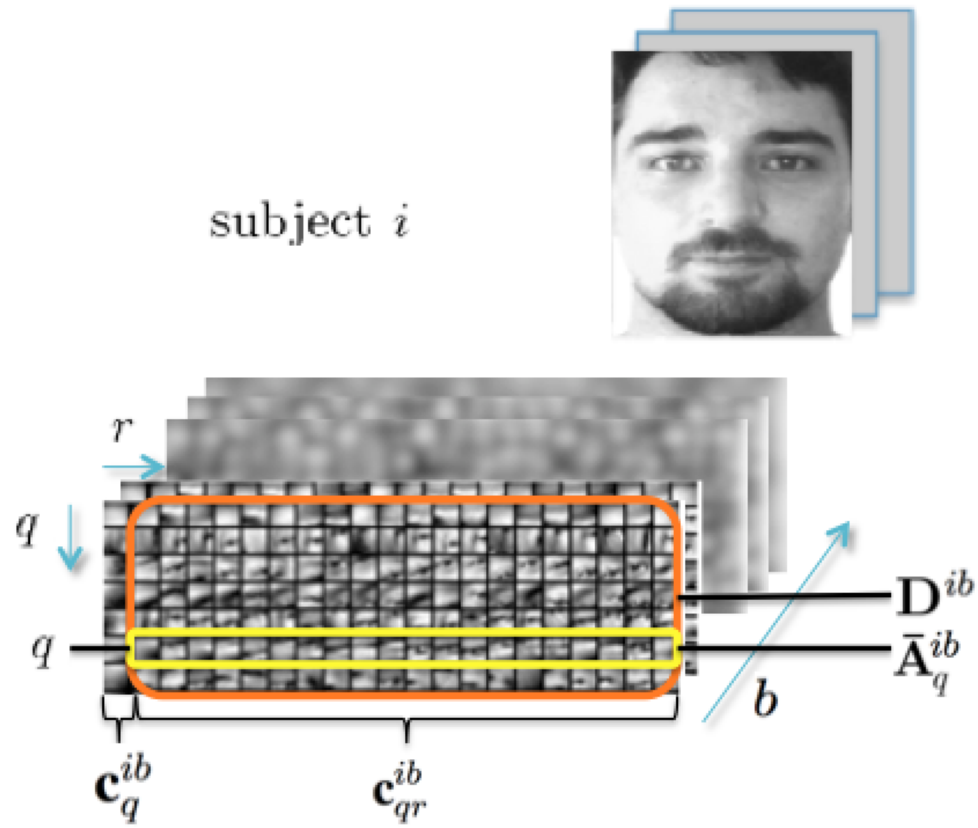

FIGURE 2.2. Dictionaries with 15 different levels of blurriness b of subject i for $Q = 32$ (only for $q = 1 \ldots 7$ is shown) and $R = 20$. Left column shows the centroid $\mathbf{c}_q^{ib}$ of parent clusters. Right columns (orange rectangle called $\mathbf{D}^{ib}$) show the centroids $\mathbf{c}_{qr}^{ib}$ of child clusters. $\bar{\mathbf{A}}_q^{ib}$ is row q of $\mathbf{D}^{ib}$, $i.e.$, the centroids of child clusters of parent cluster q.

where index b^t represents the level of blurriness in which the algorithm should work. Henceforth, for simplicity of further explanation, we will assume that the arrangement $\bar{\mathbf{A}}_q^{ib}$ will continue as shown below:

$$\bar{\mathbf{A}}_q^{i} = \mathbf{D}_q^{ib} \text{ for } b = b^t. \tag{2.12}$$

2.0.2.2. Testing Methodology

From the test image, S test patches $\mathcal{P}_p^t$ of size $w \times w$ pixels are selected, extracted, and described using (2.7) as $\mathbf{y}_p^t = f(\mathcal{P}_p^t)$ (for $p = 1 \ldots S$). The selection criterion of a test patch will be explained later in this section. For each selected test patch with description

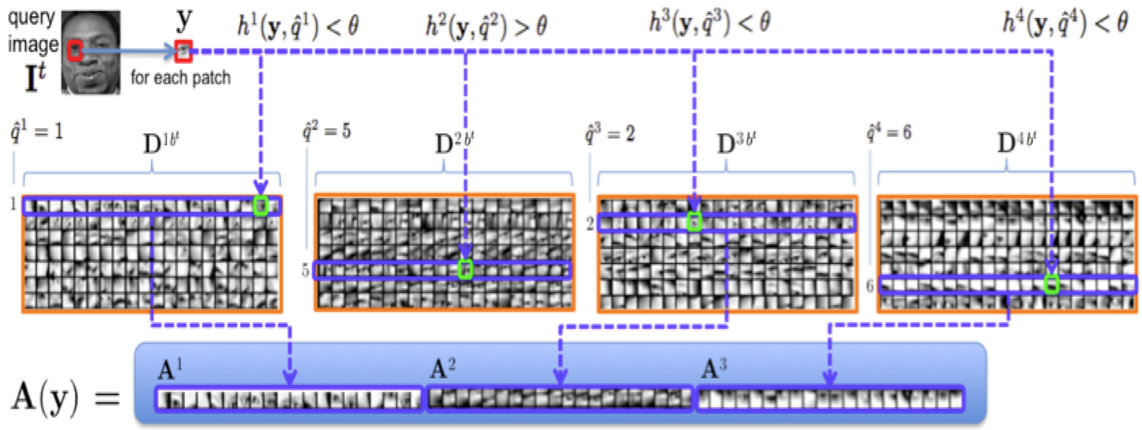

FIGURE 2.3. Adaptive dictionary $\mathbf{A}$ of patch $\mathbf{y}$ working at blurriness level b^t. In this example there are $k = 4$ subjects in the gallery. For this patch only $k' = 3$ subjects are selected. Dictionary $\mathbf{A}$ is built from those subjects by selecting all child clusters of a parent cluster (see blue rectangles) which have a child with the smallest distance to the patch (see green squares). In this example, subject 2 does not have a child cluster that is similar enough to patch $\mathbf{y}$, *i.e.*, $h^2(\mathbf{y}, \hat{q}^2) > \theta$.

$\mathbf{y} = \mathbf{y}_p^t$, a distance to each parent cluster q of each subject i of the gallery is measured:

$$h^i(\mathbf{y}, q) = \text{distance}(\mathbf{y}, \bar{\mathbf{A}}_q^i) \tag{2.13}$$

Different distance metrics were tested. The best performance, however, was obtained by $h^i(\mathbf{y}, q) = \min_r ||\mathbf{y} - \mathbf{c}_{qr}^i||$, which is the smallest distance to the centroids of child clusters r of parent cluster q as illustrated in Fig. 2.3. Normalizing $\mathbf{y}$ and $\mathbf{c}_{qr}^i$ to have unit l_2 norm, (2.13) can be rewritten as:

$$h^i(\mathbf{y}, q) = 1 - \max < \mathbf{y}, \mathbf{c}_{qr}^i > \text{for } r = 1 \dots R \tag{2.14}$$

where the term $< \bullet >$ corresponds to the scalar product that provides similarity (cosine of angle) between vectors $\mathbf{y}$ and $\mathbf{c}_{qr}^i$. The parent cluster that has the most minimal distance is searched:

$$\hat{q}^i = \underset{q}{\text{argmin}}\ h^i(\mathbf{y}, q) \tag{2.15}$$

where the minimum distance is $h^i(\mathbf{y}, \hat{q}^i)$. For patch $\mathbf{y}$, we select those gallery subjects that have a minimal distance less than a threshold θ in order to ensure similarity between the test patch and the representative patches of the subject. If k' subjects fullfill the condition $h^i(\mathbf{y}, \hat{q}^i) < \theta$ for $i = 1 \dots k$, with $k' \leq k$, we can build a new index $v_{i'}$ that indicates the index of the i'-th selected subject for $i' = 1 \dots k'$. For instance in a gallery with $k = 4$ subjects, if $k' = 3$ subjects are selected (*e.g.*, subjects 1, 3 and 4), then the indexes are

$v_1 = 1$, $v_2 = 3$, and $v_3 = 4$ as illustrated on Fig. 2.3. The selected subject i' for patch **y** has its dictionary $\mathbf{D}^{v_{i'}b^t}$, and the corresponding parent cluster is $u_{i'} = \hat{q}^{v_{i'}}$, in which child clusters are stored in row $u_{i'}$ of $\mathbf{D}^{v_{i'}b^t}$, *i.e.*, in $\mathbf{A}^{i'} := \bar{\mathbf{A}}^{v_{i'}}_{u_{i'}}$. Therefore, a dictionary for patch **y** is built using the best representative patches as follows (see Fig. 2.3):

$$\mathbf{A}(\mathbf{y}) = [\mathbf{A}^1 \dots \mathbf{A}^{i'} \dots \mathbf{A}^{k'}] \in \mathbb{R}^{(d+2)\times Rk'} \tag{2.16}$$

With this adaptive dictionary **A**, built for patch **y**, we can use SRC methodology, Wright & a.Y. Yang (2009). That is, we look for a sparse representation of patch **y** using the l_1-minimization approach:

$$\hat{x} = \text{argmin}||x||_1 \text{ subject to } \mathbf{Ax} = \mathbf{y} \tag{2.17}$$

The residuals are calculated for the reconstruction of the selected subjects $i' = 1 \dots k'$:

$$r_{i'}(\mathbf{y}) = ||\mathbf{y} - \mathbf{A}\delta_{i'}(\hat{x})|| \tag{2.18}$$

where $\delta_{i'}(\hat{x})$ is a vector of the same size as $\hat{x}$, in which only nonzero entries are the entries in $\hat{x}$ corresponding to class $v(i') = v_{i'}$. Thus, the class of selected test patch **y** will be the class that has the minimal residual, which would be:

$$\hat{i}(\mathbf{y}) = v(\hat{i}') \tag{2.19}$$

where $\hat{i}' = argmin_{i'} r_{i'}(\mathbf{y})$. Finally, the identity of the query subject will be chosen by the majority vote of the classes assigned to the S selected test patches $\mathbf{y}^t_p$, for $p = 1 \dots S$:

$$\text{identity}(\mathbf{I}^t) = \text{mode}(\hat{i}(\mathbf{y}^t_1), \dots \hat{i}(\mathbf{y}^t_p), \dots \hat{i}(\mathbf{y}^t_s)) \tag{2.20}$$

The selection of S patches of query image is as follows:
i) From query image $\mathbf{I}^t$, m patches are randomly extracted and described using (2.7): $\mathbf{y}^t_j$, for $j = 1 \dots m$, where $m \geq S$.
ii) Each patch $\mathbf{y}^t_j$ is represented by $\hat{x}^t_j$ using (2.17).
iii) The $Sparsity\ Concentration\ Index$ (SCI) of each patch is computed in order to evaluate how spread each of its sparse coefficients are, Wright & a.Y. Yang (2009). SCI is defined by

$$S_j := \text{SCI}(\mathbf{y}^t_j) = \frac{k \max\limits_i(\|\delta_{i'}(\hat{x}^t_j)\|_1)/\|\hat{x}^t_j\|_1 - 1}{k-1} \tag{2.21}$$

If a patch is discriminative enough, it is expected that its SCI is large. Note that we use k instead of k' because the concentration of the coefficients related to k classes must be measured.

iv) Array $\{S\}_{j=1}^{m}$ is sorted into a descending order of SCI values. The first s patches in this sorted list, in which SCI values are greater than a τ threshold, are then selected. If only S' patches are selected, where $S' < S$, then the majority vote decision in (2.20) will be taken with the first S' patches.

Additionally, another method was proposed for face recognition "bASR+ (PSF)". Instead of using the sharpness metric, we intended to use a Point Spread Function (PSF) information to describe the global blurriness of dictionaries and query images. In this case, blind deconvolution was applied to each image for each of the fifteen generated datasets. As blur intensified, PSF matrix grew larger for better results. This gave us five different PSF matrix sizes. When query image was tested, five blind deconvolutions were applied and matching was performed to search for the minimum error between PSF of the same size. Afterwards, an adaptive dictionary was built out of the selected dataset. Lastly, the program proceeded as ASR+ have proceeded.

3. EXPERIMENTAL RESULTS AND IMPLEMENTATION

The algorithm was tested using a database that challenged face recognition algorithms based on changes on illumination, occlusion, and facial expression. In addition, and for sake of this investigation, we added images with different levels of blurriness to the database. Specifications and implementation on the database will be described in the following subsections. The performance of bASR+ algorithm was compared to the performance of different algorithms in two types of experiments. In the first experiment, we compared our approach with state-of-the-art algorithms. The second experiment compared bASR+ algorithm with commercial software that include face recognition algorithms, such as: Picasa ©, iPhoto © for OSx 10.9.5 (Mavericks) and, Photo © for OSx 10.10.5 (Yosemite), since these software should work on uncontrolled conditions.

3.0.3. Database

We tested bASR+ with the 'AR' database, Martinez & Benavente (June 1998), where images were presented with changing illumination conditions, facial expression ($e.j.$, yawn, smile, among others), and occlusion (such as scarf and sunglasses) of the subjects. This database included 26 images for each of the 100 subjects: 50 women and 50 men.

For this investigation, occluded images were not considered. Also, the first image of each subject was banned from the training phase, leaving a total of 13 images per subject. In addition, we created a new set of photos from the first image of each subject. These images were blurred manually with a Canon 60D at a 200mm focal length with five different levels of blurriness. Images were displayed on a Macbook Pro's 15-inch retina display, while the camera shot a single photo for each subject at a certain level of blurriness. This procedure was repeated four times, manually adding more blur to the lens, as shown in Fig. 3.1.

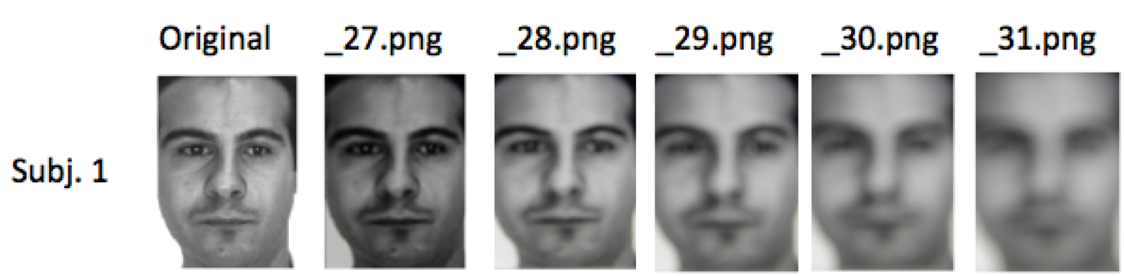

FIGURE 3.1. The image at the left is the first image of subject 1. From left to right are the ascending levels of blurriness of the original photo. The top of each photo represents the index of the photo in the new dataset.

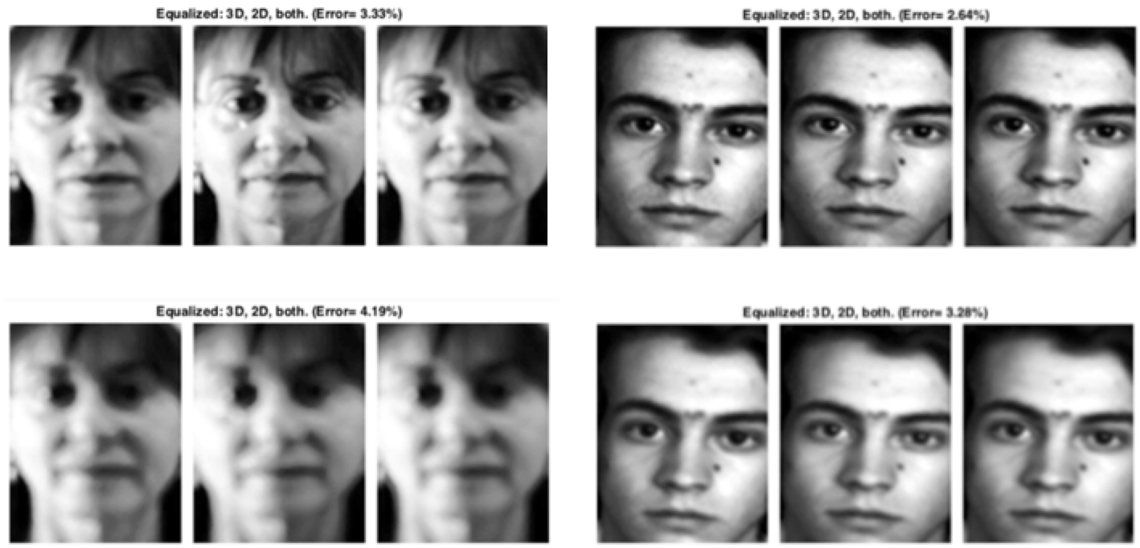

FIGURE 3.2. In each set of three images, the 3D picture is on the left, the 2D picture is in the center, and the the sum of both pictures divided by two is on the right. The percentage of error of this technique is a 3.33% for the top-left set, 4.19% for the bottom-left set, 2.64% for the top-right set, and 3.28% for the bottom-right set.

Hence, the new set contained five more photos per subject, which were intended to be used only in the testing phase.

To verify the blurring procedure described earlier, a blurred image was obtained in three steps: *i)* the first original (sharp) face image of an AR subject was displayed on a computer monitor, *ii)* a digital single-lens reflex camera (SLR) was set out-of-focus, *iii)* a picture of the monitor was captured by the camera. Thus, we have real out-of-focus face images of all subjects of AR database. Nevertheless, it is an out-of-focus image of a 2D object (a blurred picture of a face image displayed on a monitor) instead of a 3D object (a blurred picture of a real face). We call these two kinds of images I_{2D} and I_{3D}, respectively. In real cases, all images are I_{3D}; however, in our experiments, the images were I_{2D}. In order to evaluate the similarity of these two kinds of images, we conducted the following experiment: from two different subjects (relatives of the authors), we took five I_{2D} and I_{3D} images (see Fig. 3.2) and we measured the residual RMS after registration, Szeliski (2006). The registered image computed from I_{2D} is called I^r_{2D}. In this experiment, the error between I_{3D} and I^r_{2D} was 2.64% $\sim$ 4.19% only.

TABLE 3.1. A) Comparison of our algorithm bASR+ with different face recognition algorithms ($k = 20, n = 4$).

Method	Original	_27.png	_28.png	_29.png	_30.png	_31.png
bASR+	100%	100%	99.8%	99.6%	98.5%	92.3%
bASR+(PSF)	98%	100%	100%	99.7%	98.6%	95.1%
ASR+	100%	100%	100%	98.3%	82.4%	42.4%
SRC	100%	100%	99.9%	98.5%	81.8%	45.3%
LBP	100%	100%	99.9%	98.8%	82.4%	42.3%

3.0.4. Protocol and Implementation

The protocol of bASR+ experiments consists on k subjects selected randomly from the database, for $k \leq K$. For each selected subject, n images were selected randomly and the new set of five images per subject were used as testing. In order to obtain a better confidence level in the face recognition accuracy, the test was repeated 100 times by randomly selecting new k subjects and n images each time. The accuracy reported in the first experiments is the average over the 100 tests.

We used open source libraries for the implementation of bASR+. In the case of k-means we used VLFeat, Vedaldi & Fulkerson (2010), whereas sparse representation was executed with SPAMS library[1]. When using 'AR' database, the algorithm worked well with parent clusters $Q =$ 80 and child clusters $R =$ 50. The numbers of patches used were $m =$ 800, and the location weighting factor used was $\alpha =$ 4. The size of patches $w =$ 16 pixels, and the threshold for minimum distance between test patch and child cluster is $\theta =$ 0.05. Threshold for **SCI** $\tau =$ 0.1 and number of selected patches $s =$ 300, as shown in Mery & Bowyer (2014).

3.0.5. Experiments

The first experiments consisted of comparing our approach to state-of-the-art algorithms such as ASR+, Mery & Bowyer (2014), SRC Wright & a.Y. Yang (2009), LBP Ahonen et al. (2006), and a variation of our approach using Point Spread Functions (PSF) characteristics, called bASR+ (PSF). These experiments were done with different numbers of subjects (k) and training images (n): A) $k = 20$, $n = 4$, B) $k = 40$, $n = 9$, and C) $k = 100$, $n = 13$.

Given the results shown in tables 3.1, 3.2 and 3.3, we observed a wide difference between our approach and state-of-the-art algorithms such as ASR+, SRC, and LBP in those images with high levels of blurriness ($e.g.$, images *_30.png and *_31.png). When

[1] SPArse Representation Software available on http://spams-devel.gforce.inria.fr

TABLE 3.2. B) Comparison of our algorithm bASR+ with different face recognition algorithms ($k = 40, n = 9$).

Method	Original	_27.png	_28.png	_29.png	_30.png	_31.png
bASR+	99.6%	100%	100%	100%	99.9%	93.2%
bASR+(PSF)	84.5%	100%	100%	99.2%	97.5%	88.4%
ASR+	100%	100%	100%	99.8%	86.4%	40.2%
SRC	100%	100%	100%	99.8%	86%	39.4%
LBP	100%	100%	100%	99.8%	85.4%	41.4%

TABLE 3.3. C) Comparison of our algorithm bASR+ with different face recognition algorithms ($k = 100, n = 13$).

Method	Original	_27.png	_28.png	_29.png	_30.png	_31.png
bASR+	99.8%	100%	100%	100%	99.3%	85.6%
bASR+(PSF)	85.1%	100%	100%	97.8%	90.2%	62.5%
ASR+	100%	100%	100%	99.2%	71.8%	26.3%
SRC	100%	100%	100%	99.2%	71.2%	26.3%
LBP	100%	100%	100%	99.1%	71.1%	26.6%

comparing our two methods, bASR+ and bASR+(PSF), the first worked best while working with more subjects in out-of-focus blurred images, while the later method, in theory, would work better having different types of blurs ($e.g.$, out-of-focus, motion, artificial, and zoom, among others) proper from uncontrolled photo conditions.

Our second experiment consisted of comparing our approach, bASR+, to commercial software that included a face recognition application. In this specific experiment, all of the subjects were considered with all their training images, manually classifying each subject with its class.

TABLE 3.4. Comparison of our algorithm bASR+ with different commercial software with face recognition algorithms ($k = 100, n = 13$).
* : In more than 50% of the images, no face was detected by the software.

Method	Original	_27.png	_28.png	_29.png	_30.png	_31.png
bASR+	99.8%	100%	100%	100%	99.3%	85.6%
Picasa ©	86%	71%	76%	5%*	0%*	0%*
iPhoto ©	58%	21%	3%*	0%*	0%*	0%*
Photo ©	94%	87%	34%	8%*	0%*	0%*

As shown in table 3.4, face recognition in tested commercial software leaves much to be desired. These software are commonly used to store uncontrolled photographs where photo quality is poor though they were not capable of matching our approach when a full face was given. Also, we can appreciate that there was an improvement on face recognition on Apple OS, where Photo © exceeds levels of recognition from its previous version.

4. CONCLUSION

In this paper, we presented bASR+, an algorithm that is able to recognize faces automatically in less constrained conditions, including excessively blurred photos. Our proposal consisted of: $i)$ several dictionaries with different levels of blurriness obtained by digitally blurring the training images chosen from the gallery, and $ii)$ use of sharpness metric for matching blurriness between query image and dictionaries. These two main adjustments made the algorithm more robust to blurred conditions.

The goal of these experiments was to prove that our algorithm, bASR+, works in uncontrolled conditions, especially in images with high levels of out-of-focus blurriness. No blur robust state-of-the-art algorithm was found for comparing our method, so we used commercial software that had a face recognition algorithm application. These software had poor recognition rates when it came to blurred images, even though they are expected to perform in uncontrolled conditions. Another drawback was although a full face was shown, the software was not always capable of finding a face. On the other hand, when updating OS, Apple software, from iPhoto © to Photo ©, performance improved considerably, yet it was not even close to matching the performances of bASR+ and state-of-the-art algorithms.

In terms of future work, we would like to extend to robustness on other types of blurriness, such as ambient interference and motion.

References

Aggarwal, G., Biswas, S., Flynn, P. J., & Bowyer, K. W. (2012). Predicting good, bad and ugly match Pairs. *Proceedings of IEEE Workshop on Applications of Computer Vision*, 153–160. doi: 10.1109/WACV.2012.6163007

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. doi: 10.1109/TPAMI.2006.244

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 *Image processing (icip), 2010 17th ieee international conference on* (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. doi: 10.1109/TPAMI.2012.30

Deng, W., Hu, J., & Guo, J. (2013). In defense of sparsity based face recognition. In *Computer vision and pattern recognition (cvpr), 2013 ieee conference on* (pp. 399–406).

Ferzli, R., Karam, L. J., & Caviedes, J. (2005). A robust image sharpness metric based on kurtosis measurement of wavelet coefficients. *Proc. of Int. Workshop on Video Processing and Quality Metrics for Consumer Electronics*, *12*(3).

Flusser, J., & Suk, T. (1998). Degraded image analysis: an invariant approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(6), 590–603. doi: 10.1109/34.683773

Gopalan, R., Taheri, S., Turaga, P., & Chellappa, R. (2012). A blur-robust descriptor with applications to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(6), 1220–1226. doi: 10.1109/TPAMI.2012.15

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. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *35*(11), 2651–2664.

Levin, A., Weiss, Y., Durand, F., & Freeman, W. T. (2011). Computer Science and Artificial

Intelligence Laboratory Technical Report; Efficient Marginal Likelihood Optimization in Blind Deconvolution. *Optimization*.

Martinez, A., & Benavente, R. (June 1998). The AR face database. *CVC Technical Report 24*.

Mery, D., & Bowyer, K. (2014). Face Recognition via Adaptive Sparse Representations of Random Patches. *IEEE Workshop on Information Forensics and Security (WIFS 2014), Atlanta, Dec. 3-5.*.

Phillips, P. J., Beveridge, J. R., Draper, B. A., Givens, G., OŤoole, A. J., Bolme, D. S., ... Weimer, S. (2011). An introduction to the good, the bad, & the ugly face recognition challenge problem. *2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, FG 2011*, 346–353. doi: 10.1109/FG.2011.5771424

Ptucha, R., & Savakis, A. (2013). Lge-ksvd: Flexible dictionary learning for optimized sparse representation classification. In *Computer vision and pattern recognition workshops (cvprw), 2013 ieee conference on* (pp. 854–861).

Qiu, Q., Patel, V. M., & Chellappa, R. (2014). Information-theoretic dictionary learning for image classification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *36*(11), 2173–2184.

Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, *60*, 259–268. doi: 10.1016/0167-2789(92)90242-F

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.

Stainvas, I., & Intrator, N. (2000). Blurred face recognition via a hybrid network architecture. *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, *2*, 1–9. doi: 10.1109/ICPR.2000.906198

Szeliski, R. (2006). Image Alignment and Stitching: A Tutorial. *Foundations and Trends in Computer Graphics and Vision*, *2*, 1–104. doi: 10.1561/0600000009

Tikhonov, A., & Arsenin, V. (1977). *Solutions of ill-posed problems*. Vh Winston.

Vedaldi, A., & Fulkerson, B. (2010). VLFeat - An open and portable library of computer vision algorithms. *Design*, *3*, 1–4. Retrieved from `http://vision.ucla.edu/ brian/papers/vedaldi10vlfeat.pdf` doi: 10.1145/1873951.1874249

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*, 372–386.

doi: 10.1109/TPAMI.2011.112

Wang, J., Lu, C., Wang, M., Li, P., Yan, S., & Hu, X. (2014). Robust face recognition via adaptive sparse representation. *Cybernetics, IEEE Transactions on*, *44*(12), 2368–2378.

Wei, X., Li, C.-T., & Hu, Y. (2012). Robust face recognition under varying illumination and occlusion considering structured sparsity. In *Digital image computing techniques and applications (dicta), 2012 international conference on* (pp. 1–7).

Wright, J., & a.Y. Yang. (2009). Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(March), 210–227. Retrieved from `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4483511` doi: 10.1109/TPAMI.2008.79

Wright, J., Ma, Y., & Mairal, J. (2009). Sparse representation for computer vision and pattern recognition. *Proceedings of IEEE*(March), 1031–1044. Retrieved from `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5456194`

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

Yuan, L., Sun, J., Quan, L., & Shum, H.-Y. (2008). Progressive inter-scale and intra-scale non-blind image deconvolution. In *Acm transactions on graphics (tog)* (Vol. 27, p. 74).

Zhao, W., Chellappa, R., Phillips, P., & Rosenfeld, a. (2003). Face recognition: A literature survey. *Acm Computing Surveys*, *35*(4), 399–458. Retrieved from `http://dl.acm.org/citation.cfm?id=954342` doi: 10.1145/954339.954342

Zhu, X., & Milanfar, P. (2009). A no-reference sharpness metric sensitive to blur and noise. *2009 International Workshop on Quality of Multimedia Experience, QoMEx 2009*, *2*(3), 64–69. doi: 10.1109/QOMEX.2009.5246976