

Face Recognition in Uncontrolled Condition: Can Compressive Sensing and Super-Resolution Meet the Challenge?

Sabah A Jassim* and Nadia AL-Hassan

Department of Applied Computing, University of Buckingham, Buckingham, UK

Abstract

This paper is concerned with face recognition under uncontrolled conditions, e.g. at a distance in surveillance scenarios, and post-rioting forensic, whereby captured face images are severely degraded/blurred and of low resolution. This is a tough challenge due to many factors including difficulties in determining a model for image degradation that encompasses a range of realistic capturing conditions. We present the results of our investigations into recently developed Compressive Sensing (CS) theory to develop scalable face recognition schemes using a variety of over-complete dictionaries that construct super-resolved face images from any input low-resolution degraded face image. We shall demonstrate that deterministic as well as non-deterministic dictionaries that do not involve the use of face image information but satisfy some form of the Restricted Isometry Property (RIP) used for CS can achieve face recognition accuracy levels as good as, if not better than, those achieved by dictionaries, proposed in the literature, that are learnt from face image databases using elaborate procedures. We shall elaborate on how this approach helps in crime fighting and terrorism.

Keywords: Compressive sensing; Super-resolution; Face recognition; Forensic; RIP dictionaries

Introduction

Face recognition in uncontrolled conditions arise primarily in fighting crime through surveillance using CCTV cameras. In contrast to recognition in control conditions, which has witnessed significant improvement over the last 3 decades, very little progress has been made to attain acceptable recognition rates due to the degraded nature of the captured images. CCTV cameras are at a distant from the imaged scenes and thus capture small low resolution, blurred and low-quality face images. Image degradation results from a variety of recording conditions: subject on the move, unstable sensors, out of focus optical system, or abnormal weather and atmospheric conditions such as thermal waves. Image resolution enhancement is deemed necessary for face recognition in these cases.

In biometric systems, normally there is a set of feature vectors (one or more for each of the enrolled subjects) called the Gallery representing the digital templates that are obtained during the enrolment stage. For recognition under controlled conditions, these templates are extracted from good quality images of the highest possible resolution, and when a claimant presents him/her-self a reasonably good quality image is input to the same feature extraction procedure is applied and the output will be matched against all the gallery templates and identification is determined as that of the nearest neighbour. The situation is fundamentally different in the uncontrolled scenarios. Recognising faces when matching low-resolution (LR) degraded small images against a gallery of high-resolution good size face images need to incorporate some preprocessing resolution enhancing procedures. Various super-resolution methods have been developed with the aim of reconstructing a higher resolution version of the LR image. Hennings-Yeomans et al. [1] proposed to perform super-resolution and recognition simultaneously. The performance of this method depends on the training dataset of images. He and Zhang [2] developed an SR scheme that constructs a high-resolution face image, from a sequence of low-resolution images, to be processed by Gabor feature based recognition.

In recent years advances in compressive sensing (CS) theory

and sparse representation have been exploited to develop image/signal processing and analysis tools to be used in pattern recognition including face recognition from LR face images. In particular, the development of efficient l_1 -minimization procedures to find sparse solutions of certain under determined linear systems has led to the emergence of new SR schemes for the recovery of high quality super-resolved images from low resolution degraded images [3-7]. Yang et al. [5] proposed a method to reconstruct super-resolved image from a single low-resolution image using a pair of overcomplete dictionaries D_H and D_L whose columns are constructed, through a learning process, from a number of randomly selected patches of high and low resolution training datasets of face images. This pair of image-trained dictionaries will be referred to, thereafter, as the LD system. The main objective of this paper is to question the need for image-trained dictionaries for SR tasks, and in particular for face recognition in uncontrolled conditions. We shall describe a simple method to implicitly construct CS compliant dictionaries without using images and demonstrate that such non-adaptive dictionaries perform as well as the LD dictionary, if not better. We shall also demonstrate that this is not only true for our implicitly constructed dictionary, but rather for a number of different random dictionaries, and show that there are no visible differences in the quality of the super-resolved images obtained from all the investigated dictionaries. For completion, we also present the performance of a non-CS based iterative SR method, and of matching in low-resolution.

The rest of the paper is organized as follows. Sections 2 and 3 provide a brief review of Super resolution and Compressive Sensing respectively. In section 4, we shall discuss a recently designed CS

*Corresponding author: Sabah A Jassim, Department of Applied Computing, University of Buckingham, Buckingham, UK, E-mail: sabah.jassim@buckingham.ac.uk

Received July 29, 2013; Accepted August 29, 2013; Published September 04, 2013

Citation: Jassim SA, AL-Hassan N (2013) Face Recognition in Uncontrolled Condition: Can Compressive Sensing and Super-Resolution Meet the Challenge? J Biomet Biostat 4: 171. doi:10.4172/2155-6180.1000171

Copyright: © 2013 Jassim SA, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

approach to image RS using different types of dictionaries, and discuss the properties of these dictionaries that are relevant to the recovery of a sparse signal from a down-sampled degraded version of images. In section 5, we shall conduct experiments to compare the performance of a known face recognition scheme when applied to super-resolved images using the different types of dictionaries as well as to the original LR images. In the conclusion, section 6, we shall briefly describe the contribution of the paper and also highlight benefits of using certain types of implicitly constructed CS dictionaries conclusions.

Super Resolution

Recovering a high-resolution (HR) image from one or more low-resolution (LR) images is a challenging inverse problem that has been traditionally dealt with by iterative procedures of incremental enhancement, referred to as Super-resolution (SR). SR schemes, therefore, assume that an observed degraded LS small image y is the result of blurring and down-sampling applied to an ideal HR image x and corrupted by additive noise, i.e. x is a solution of a matrix equation:

$$y = SBx + \eta \tag{1}$$

where B is a point-spread function with a blurring effect, S is a down sampling function, and η is additive noise. The most common traditional non-CS based super-resolution techniques are variants of the Iterative Back Projection (IBP) SR scheme that can super-resolve a single or multiple input LR image(s). The standard single LR image IBP scheme works by first generating the initial HR image x_0 by decimating the pixels of y using Bi-cubic interpolation. For $n > 0$, calculate an error image x_e of the size of the $x_{(n-1)}$ image by the 3 steps procedure:

Convolute $x_{(n-1)}$ image with an appropriate degradation function,

Down sample the resulting image to obtain $y_{(n)}$, and

x_e is obtained from $(y - y_{(n)})$ by up-sampling.

The n th iteration output the n th version of the HR image simply by calculating $x_{(n)} = (x_{(n-1)} + x_e)$, representing the back projection of the difference $(y - y_{(n)})$ onto $x_{(n-1)}$. The SR scheme terminates either when the energy of the error term $(y - y_{(n)})$ is reduced below a certain threshold or the number of iterations reached a fixed maximum number [8]. Variants of the IBP expand each iteration by pack-projecting additional terms, representing high frequency information in x_0 e.g. using the Canny edge detection [9].

The main challenge in recovering x is the modelling of the unknown blurring function. Gaussian functions with different blurring effect have been considered as a suitable model for use in SR procedures, but they do not reflect severe degradation conditions seen in surveillance scenarios. A suitable model can be based on the use of atmospheric turbulence functions of different strengths (i.e., degradation functions that model environmental conditions caused by variation in temperature, wind speed and exposure time) which extends the effect of the Gaussian functions. In the frequency domain such functions are of the form:

$$H(u, v) = e^{-k(u^2 + v^2)^{5/6}} \tag{2}$$

Where k is a constant that reflects severity of blurring. We label degradation as severe if $k \in [0.045, 0.09]$; mild (similar to most Gaussian blurring functions) if $k \in [0.02, 0.04]$; and low if $k \in [0, 0.02]$. The small images in Figure 1 are the down-sampled degraded images after applying H for different values of k . Throughout the paper, we shall adopt this model of degradation for a number of k values in these ranges to test performance of face recognition from LR images.

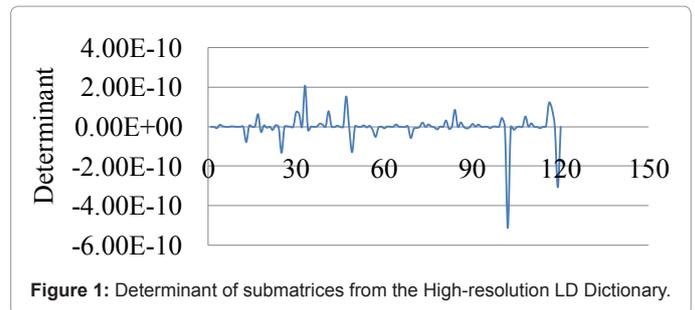


Figure 1: Determinant of submatrices from the High-resolution LD Dictionary.

Compressive Sensing

Images/videos and other media signals/objects have long been benefiting from frequency domain decompositions and dimension reduction methods that help express the original signals as superposition of certain bases functions. These methods yield sufficiently informative but sparse approximation of the original signals suitable for efficient implementation of a variety of processing and analyse tasks. Traditionally the input signals to these methods are uniformly sampled for the highest resolution afforded by the deployed sensors. It is natural to ask if this is necessary, when most data resulting from these methods are thrown away. Compressive sensing, also known as sparse recovery, is a novel paradigm of signal sampling that attempts to answer this question and greatly relaxes the stringent limitations of the conventional Shannon-Nyquist Sampling Theorem, for signals that can be approximated by a sparse expansion in terms of a suitable basis of waveforms. The underlying principle of CS is that the number of linear measurements needed to reconstruct a compressed signal should be proportional to the compressed size of the signal, not the uncompressed size. This suggests the use of generating sets of vectors belonging to different bases of functions. Bruckstein et al. [10], suggest that the concatenation of 2 bases one constructed from wavelet functions and the other from sinusoid functions would be of benefits for image processing/analysis tasks.

The central challenge for CS is the construction of non-adaptive relatively small number of linear measurements that can guarantee the recovery of a sparse or approximately sparse signal. Such a set of linear measurements are represented by rows of an over complete dictionary [11], i.e., an $m \times n$ matrix whose columns form a spanning set of m -dimensional vectors to be used to decompose the signal. Dictionaries generalize vector space basis, and are represented by overcomplete $m \times n$ matrices, ($m << n$), whose columns are expected to form a pool of \mathbb{R}^m bases. Consequently, vectors in \mathbb{R}^m can have multiple representations by the different bases each capturing different features, perhaps at different scales. A main premise of this work is that good CS dictionaries can be constructed implicitly from certain pools of bases by concatenation.

If $D = \{d_1, d_2, \dots, d_n\} \in \mathbb{R}^{m \times n}$ is a suitable underdetermined dictionary then for any observed vector y , CS-based tools recover the sparsest solution of the equation: $y = Dx$, i.e. determine $\hat{x} \in \mathbb{R}^n$ such that:

$$\hat{x} = \min_y \|x\|_0 \text{ subject to } \|y - Dx\|_2 \tag{3}$$

Unfortunately, this l_0 -minimization problem, known as the (P_0) problem, is computationally NP-hard due to the need to exhaustively testing all m columns of D , indexed by a subset Ω of $\{1, 2, \dots, n\}$, such that $y = D \Omega z$ has a non-trivial solution $z \in \mathbb{R}^m$. However, if x is sparse and D satisfies certain properties, then a unique solution of the l_1 -minimisation (P_1) problem exists:

$$\hat{x} = \min_y \|x\|_1 \text{ subject to } \|y - Dx\|_2. \quad (4)$$

This is a convex optimisation problem which is amenable to linear programming. In fact, if x is small, then the Least Square (LS) method can be used to solve the corresponding the l_2 -minimisation (P2) problem:

$$\hat{x} = (\arg \min_{x \in D} \|x\|_2) = D^*(DD^*)^{-1}b.$$

In many applications such as when x is spiky, the LS solution is not suitable. However, the use of the l_1 -minimisation to recover the solution of (P₀) problem have been the subject of intense research. Bruckstein et al. [10], discuss two basic questions about (P₀): (1) Under what conditions, does it have a unique solution? and (2) Given a feasible solution, is there a simple test to verify that is a global minimizer? The uniqueness requirements are known to depend on certain parameters and properties of the matrix D .

The sparke of an $m \times n$ matrix D , denoted by $sp(D)$ is the minimum number of linearly dependent columns of D . It is clear that $sp(D) \leq m+1$. Equality occurs when D has a full row rank, and then D is said to be of full sparke.

Theorem 1 [11]

If every $(sp(D)-1)$ columns of D are linearly independent then every $(sp(D)/2)$ -sparse x can be recovered uniquely from Dx . //

Throughout we assume that these matrices are of full row ranks. Computing the sparke of a dictionary D may sound as difficult and NP-hard as solving the P_0 -problem, however, the absence of this property can be established statistically by testing a randomly selected large set of m -columns for independence. More importantly, this theorem provides an efficient strategy for the implicit construction of suitable CS-dictionaries that guarantee uniqueness of solutions. The main theme in this paper is to show that such dictionaries can be implicitly constructed by concatenating certain sets of \mathbb{R}^m bases.

The answer to the second question, above relates to the Null Space Property (NSP) An $m \times n$ dictionary D satisfies the NSP of order k if for each size k set $\Omega \subset \{1, \dots, n\}$ and nonzero vector $z \in \text{Ker}(D)$,

$$\|z_\Omega\|_1 \leq \|z_{\Omega^c}\|_1, \quad (5)$$

Where z_A is obtained from z by making 0 all coordinates not indexed by $A \subset \{1, \dots, n\}$.

Theorem 2 [12]

An $m \times n$ dictionary D satisfies NSP of order k if every k -sparse solution x can be recovered by l_1 -minimization. //

It is not difficult to show that if D satisfies NSP of order k then every k columns of D are linearly independent. Consequently, NSP of order $2k$ guarantee uniqueness by Theorem 1 while Theorem 2 provides a method for recovering the sparsest solution.

The Isometry Property (RIP) is a less stringent property than the NSP that was introduced by Candes and Tao [13], as sufficient for l_1 -recovery. An $m \times n$ dictionary D , $m \ll n$, is said to satisfy the RIP of order k if there is a constant $0 < \delta_k < 1$, such that for any k -sparse signal $x \in \mathbb{R}^n$:

$$(1 - \delta_k) \|x\|_2^2 \leq \|Dx\|_2^2 \leq (1 + \delta_k) \|x\|_2^2 \quad (6)$$

The smallest δ_k is called the restricted Isometry constant (RIC) of order k . If D satisfies RIP of order k , then any $2k$ -columns sub-matrix of D must be well-conditioned [13,14], (i.e. the ratio of its maximum to

its minimum singular values is small). Again, checking this property for all $2k$ -columns submatrices is computationally infeasible as it requires exhaustive check of all $\binom{n}{2k}$ submatrices. Again the non satisfaction of RIP, can be deduced by computing condition numbers of sufficiently large set of randomly selected $2k$ -submatrices.

Gan et al. [15], developed a STRIP performance bound in terms of the mutual coherence μ of the dictionary which is an indicator of the dependence between columns of the matrix. The coherence of a matrix provides information about the likelihood of guaranteed recovery of the sparse solution, and is defined as the largest absolute normalized inner product of distinct columns a_i and a_j of D i.e.,

$$\mu(D) = \max_{i < j \leq n} \frac{|\langle a_i, a_j \rangle|}{\|a_i\|_2 \|a_j\|_2} \quad (7)$$

It is not difficult to show that if D is a dictionary with unit column vectors and coherence μ then D satisfies RIP of order k with $\delta_k \leq (k - 1)\mu$, and if $n \gg m$ then $\frac{1}{\sqrt{m}} \leq \mu$.

There are a number of efficient sparse recovery algorithms that have been developed including the Homotopy method (LARS) and the Iteratively Reweighted Least Square method (IRLA) [12].

Super-Resolution by Compressive Sensing

CS-based image SR schemes exploit the fact that image signals, including degraded images, can be well-approximated by a sparse expansion in terms of suitable bases constructed from waveforms such as sinusoidal curves, wavelets, and chirplets. In this section we first briefly describe CS-based approach to super-resolve low resolution degraded images using various underdetermined dictionaries that are assumed to satisfy RIP. We list a number of dictionary construction including the LD pair of images-trained dictionaries adopted by Wang et al. [4], as well as random generated pairs of dictionaries, and a new construction strategy that is independent of training images but designed to implicitly be of full sparke. We shall test the strength of RIP, using the statistical tests described above.

CS-based image SR schemes

A CS-based SR scheme for images require 2, preferably RIP, dictionaries: A Low resolution matrix D_L of size 100×512 and a High resolution matrix D_H of size 25×512 . The input to this scheme is a degraded low resolution small image L_r , and the output is super-resolved to double the size image that is meant to be of "high quality". The L_r image is first resized by decimating its pixels and Bi-cubic interpolation to obtain double the size image LR which is also degraded in the same way as L_r is. Three spatial filters, designed to highlight edges in different directions, are applied to the LR image to obtain 3 edge-highlighted versions. The four images are then subdivided into blocks of size 5 and each block is transformed into column vectors of size 25. For each block location, concatenate the corresponding column vectors in the 4 versions to create a column vector of $100 = 4 \times 25$ by concatenation. In order to avoid the appearance of blocking artefacts, the image will be subdivided into overlapping blocks.

Initialise a HR image of the same size of the LR image for the super-resolved image. The 5×5 blocks are then processed iteratively as follows:

1. Let y be corresponding 100-dimensional vector.
2. Find the sparse solution z of the underdetermined equation $y = D_L z$
3. Calculate the 25-dimentional HR patch x using the matrix

multiplication $x=D_H z$,

4. Back-projection the 2-dimensional 5x5 version of x onto the existing HR.

5. In the rest of this section we describe the pairs of Dictionaries D_L and D_H for the various dictionary construction strategies adopted in this paper.

The image-based learnt dictionary

This is the LD pair of image-learnt dictionaries proposed by Yang et al. [5,6] who also used for super-resolution based face recognition. The columns of these dictionaries are dependent on 5x5 image patches randomly selected from a large training set of good quality high resolution images that exhibit similar statistical characteristics of the images to be tested for matching. The D_H and D_L dictionaries are created as follows:

A sufficiently large number of high resolution (HR) images (here Face images) are selected and each divided into patches of 5x5 pixels. Patches overlap.

Randomly sampled raw patches from the training HR images are transformed into normalised vectors to be added as columns of the D_H dictionary. The number of columns is set to 512 representing the size of the SR image.

Generate a set LR of blurred versions of the HR images, and create 3 other filtered versions, and the columns of D_L are constructed in similar way as in step 2, but by concatenating the patches from the LR images and their 3 filtered version. Again the columns are to be normalised.

Random dictionaries

Randomly constructed matrices that satisfy the Restricted Isometry Condition include Gaussian, Toeplitz and Circular random Matrices. For Gaussian Random Matrix (GRM), the entries x_{ij} of the CS matrix of size $m \times n$ are independently sampled from a normal distribution $x_{ij} \sim N(0,1/m)$, the l_2 -norm was used to normalize each columns in the dictionary. In order to recover super resolved image from a single LR image for face recognition via sparse representation, two overcomplete dictionaries D_H, D_L of size 25×512 and 100×512 respectively have been generated from a zero mean Gaussian distribution with variance $1/25$.

Toeplitz-Circular Random measurement matrix (TCRM) are another class of RIP dictionaries that have been widely used. Bajwa et al. [16], have shown that Toeplitz-structured matrices are sufficient to recover undersampled sparse signals. Toeplitz and Circular matrices of the size $k \times n$ are respectively of the form:

$$T = \begin{bmatrix} t_n & t_{n-1} & \dots & t_1 \\ t_{n+1} & t_n & \dots & t_2 \\ \vdots & \vdots & \ddots & \vdots \\ t_{n+k-1} & t_{n+k-2} & \dots & t_k \end{bmatrix}, \text{ and } C = \begin{bmatrix} t_n & t_{n-1} & \dots & t_1 \\ t_1 & t_n & \dots & t_2 \\ \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & \dots & t_k \end{bmatrix}$$

For image reconstruction, the D_H and D_L dictionaries are generated as TCRM matrices, by selecting the first row using the standard Gaussian distribution and the rest of the rows are permuted versions of it as required above.

Iteratively constructed full spark dictionaries

Full-sparked dictionaries is a class of full row rank overcomplete $m \times n$ dictionaries, where $m \ll n$, so that each m -columns sub-matrix is a basis of \mathbb{R}^m . Here we describe an example on how to construct

such matrices by starting with an invertible $m \times m$ matrix and iteratively appending a set of image independent linearly independent m -column vectors in \mathbb{R}^m while maintaining the full spark property after every addition. One way to maintain the full spark is to insist that every new column can only be generated by the full columns of the previous inserted submatrices.

Our generic full spark dictionaries, referred to as LID, is of the form:

$$D = [A_{p_1}, A_{p_2}, \dots, A_{p_k}, C(A_{p_{k+1}})]$$

where for $i=1, \dots, k+1$, the p_i 's are distinct real numbers >1 , and

$$A_{p_i} = \begin{pmatrix} 1 & \frac{1}{p_i} & \frac{1}{p_i^2} & \dots & \frac{1}{p_i^{m-1}} \\ \frac{1}{p_i} & 1 & \frac{1}{p_i} & \dots & \frac{1}{p_i^{m-2}} \\ & & \vdots & & \\ \frac{1}{p_i^{m-1}} & \frac{1}{p_i^{m-2}} & \frac{1}{p_i^{m-3}} & & 1 \end{pmatrix}$$

Note that $k = \lfloor n/m \rfloor$ and the last sub-matrix of D is simply the first $(n - km)$ columns. Then, the $m \times n$ LID dictionary is obtained from the resulting matrix after normalising its columns using the l_2 -norm.

For our experimental purposes the LID1 high-dictionary D_H is generated from using integers $p_i > 1$. For simplicity, the low-dictionary D_L was created from a Standard Gaussian Random Matrix (GRM).

Measuring RIP strength of D_H dictionaries

In this section, we present the result of statistical test of the "strength" of the RIP for the LD and LID1 dictionaries. The various statistical tests on full spark property of a dictionary or the condition numbers of m -submatrices is conducted on a randomly selected sample of 100 submatrices. For the *full spark* property, we evaluated the determinants, as indicator of linear independence, for more than a hundred randomly selected sample of 25×25 submatrices of the corresponding D_H dictionaries. Although in theory, the LD dictionary may statistically satisfy NSP of order 12, Figure 2 shows that the determinant of most 25×25 submatrices is so small (almost zero) and hence the full spark property is not satisfied. In contrast, Figure 3 confirms that the LID1 is indeed fully spark.

The next experiment to calculate another RIP indicator, namely the condition number of 25×25 submatrices of 4 of the DH dictionaries for LD, LID1, GRM and the TCRM. These condition numbers are expected

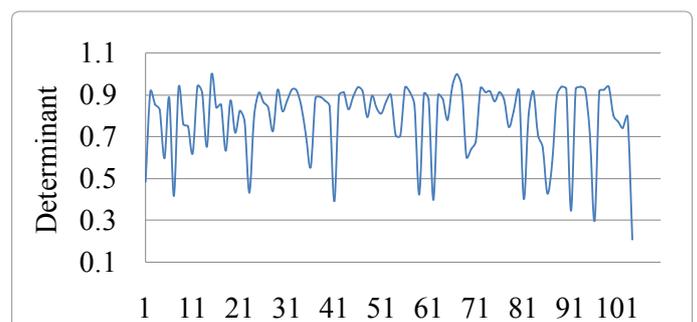


Figure 2: Determinant of submatrices from the High-resolution LID1 dictionary.

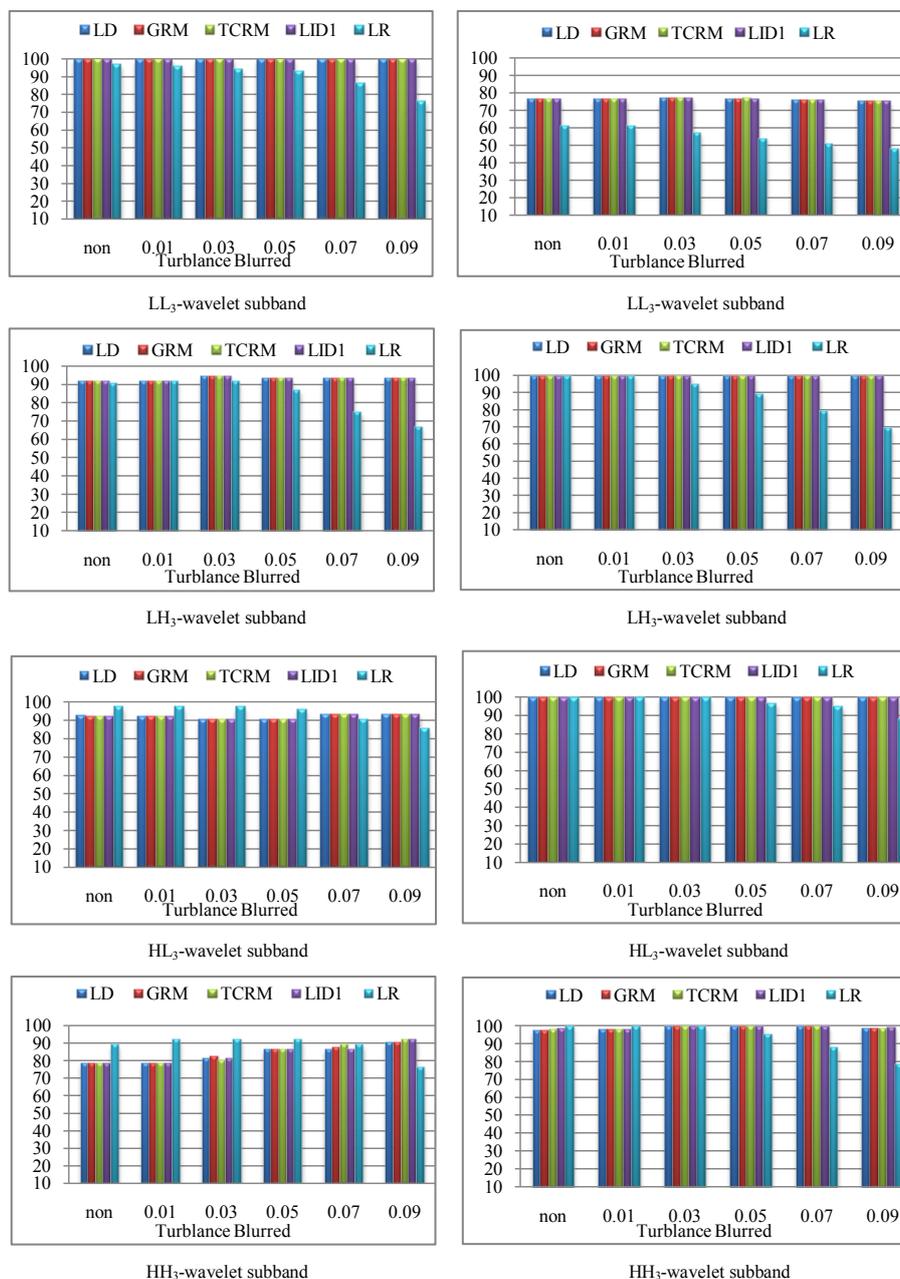


Figure 3: Recognition accuracy rates using different dictionaries and in comparisons with matching in low-resolution.

submatrices	Dictionaries							
	LID1		LD		GRM		TCRM	
	mean	std	mean	std	mean	std	mean	std
25×25	3.08	3.14	3.34E+16	1.79E+17	279.36	597.79	85.19	155.68
Full matrix	1.977		1.00E+15		1.43		1.453	

Table 1: Mean and Standard deviation for CN for a hundred random sub-matrices of different sizes.

to be bounded by RIC of order 2k, with k=12. Table 1, below, displays the mean and standard deviation of the condition numbers for 100 randomly selected submatrices and the condition number of the full size 25×512 matrix.

These results again demonstrate that the overcomplete LID1

Dictionary	LID	LD	GRM	TCRM
Row Rank	25	24	25	25
Coherence	0.9958	<0.2	0.7438	0.7318

Table 2: Row-rank and Coherence.

dictionary is well-conditioned in comparison to all others for the various submatrices but for the full matrix GRM and TCRM have similar condition numbers that are better than the LID1. Moreover, the condition number of the LD is extremely large for all cases, which make these dictionaries very ill conditioned. ,

Another test relates to calculating the row-rank and coherence values for the various dictionaries. It is well known that the highest

sparsity recovered signal for any dictionary $\mu = (1 + \text{row rank})/2$, and coherence μ must satisfy $0.2 = 1/\sqrt{m} \leq \mu \leq 1$. Again, results in Table 2 highlight the superiority of the LID1 dictionary.

In order to test the level of success of the CS-based SR schemes, in comparison to non-CS based SR such as the IISR and the usual Bi-cubic interpolation schemes, we display in Figure 4 an example of an original HR image its degraded and down-sampled versions, and the super-resolved images output from the various schemes. The degraded small images were obtained from the HR image by applying the degradation function in equation (6) for different k values ranging from mild to severe and then down sampled. In terms of image quality, the difference between the recovered HR image using the various dictionaries is not discernable by the human eye, but a noticeable improvement can be seen when the SR methods were used, including IISR, over the low-resolution images and the Bi-cubic interpolation method for every degradation value of k . PSNR values for the entire dataset of face images calculated between the output SR image and the original images confirm the same pattern, but we omit these results. As can be expected, regardless of the SR method used, the quality of super-resolved images decreases as the level of blurring increases. With increased level of blurring there

is no difference in image quality from different dictionaries. But the dictionary methods produced slight improvement on the IISR method, and superiority over the interpolation method at every level of blurring.

Face Recognition-Experimental Results

In this section, we test the accuracy of face recognition when different SR dictionary methods as well as the IISR method as well as the Bi-cubic interpolation method are used to reconstruct super resolved face image from a single LR image with different magnification blur. We use a simple but efficient wavelet-based face recognition scheme, whereby the training as well the matching image are wavelet decomposed to level 3 and each of the each of the subbands at level 3 (i.e. LL3, HL3, LH3 and HH3) is used as a face feature vector. The well lit sets of face images from the Extended Yale B database will form the bases of the experiments.

Experiments and results

The Extended Yale B database consists of 2,414 frontal-face images of 38 individuals. The cropped and normalized 192x168 face images were captured under various laboratory-controlled lighting conditions.

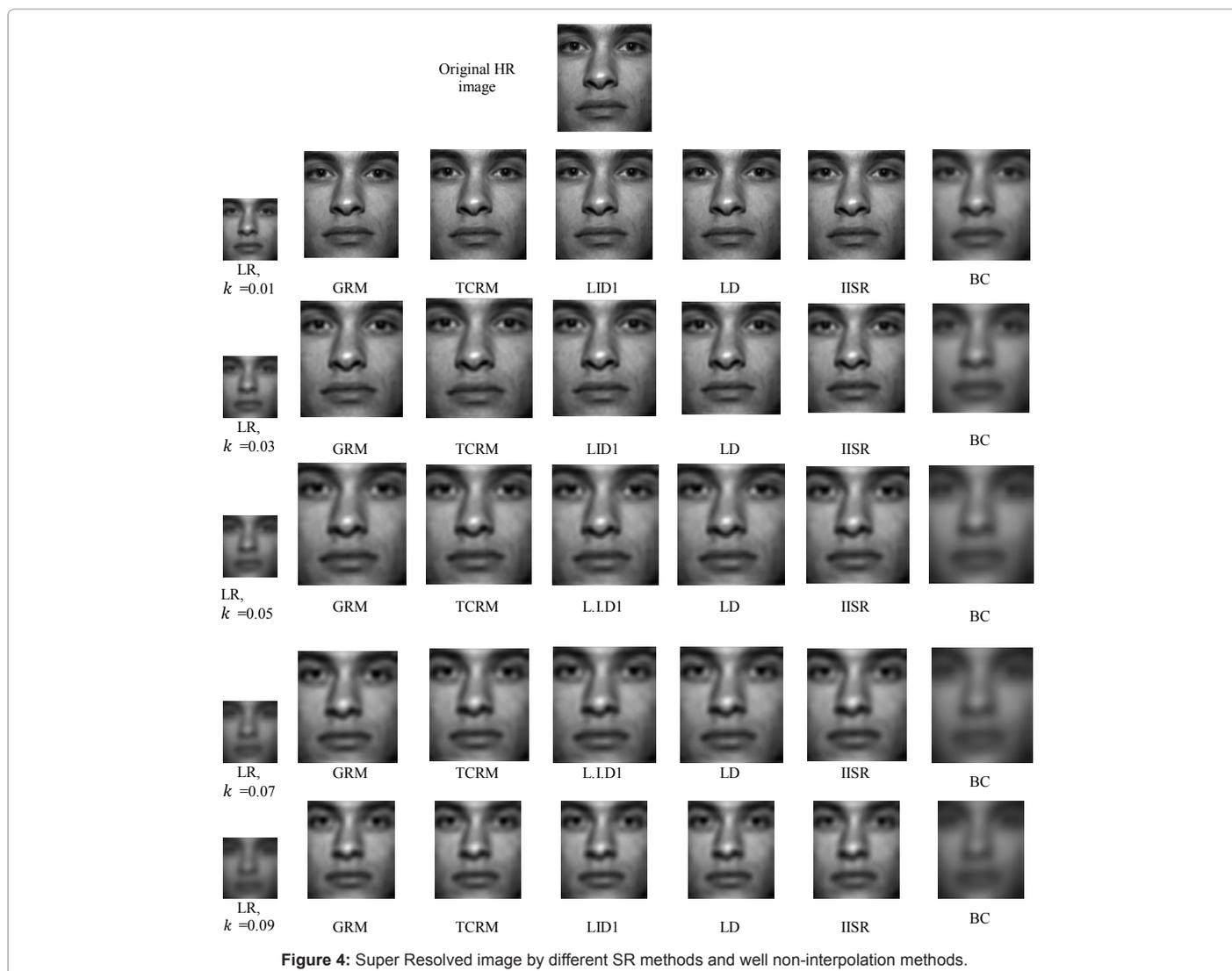


Figure 4: Super Resolved image by different SR methods and well non-interpolation methods.

For each subject, we selected the P00A+00E+00 image for the gallery set and the other images for testing. To construct LD that depends on images, three images for each subject were selected from the well-lit face images in set 1 that were not included in the gallery/test images. To simulate the intended uncontrolled scenarios, we first apply the degradation function defined by equation 2 for different values of k on the high resolution images in the reasonably lit sets (1 and 2) of the database. The low resolution degraded Lr images are finally obtained by down-sampling the degraded images by a factor of 2.

For feature extraction, each test Lr face image is first super-resolved using each of the above described dictionaries, the IISR scheme or simply by up-sampling and bi-cubic interpolation. On the resulting HR templates as well as the SR-resolved test images, we use the Z-score normalized coefficients of the subbands of the Haar wavelet decomposed face images at level 3. Matching is based on the City Block distance function.

Figure 4, below show results for the different subbands and each of set 1 and set 2 face images. The various charts display the accuracy rate at each level of degradation function. As can be seen, there is no significant difference in identification accuracy rates, between the different dictionaries methods. Moreover, the accuracy rates seem to be maintained at the same level for different degradation level. In comparisons to the method of matching the LR images with down-sampled gallery images, the performance of the dictionary based methods are far more superior and much more apparent as the image quality deteriorates from mild to severe degradation. However, the picture is surprisingly different for mild degradation (i.e. $k < 0.07$) when we use the HL3 or HH3 subband as feature vectors for set1 images. We attribute this to the effect of variation in direction of light source between the 2 sets on the significant HL3 and HH3 coefficients associated with high frequency features (e.g. vertical and diagonal edges, respectively). The light in set1 is centrally perpendicular to the face, but is slightly off the centre in set 2, and this result in the appearance of vertical and diagonal artefacts in set 2 images only, albeit very faint. Mild degradations in the LR images remove these artefacts in in all LR versions of set 2 images, but it also smoothes the horizontal facial features associated with significant coefficients in LH3 subband for all the LR images. This helps maintain the performance of the HL3 and HH3 subband schemes for mild blurring.

The differences in the performance of all the schemes on the two sets can be attributed to the fact that the LL3 subband being the approximation of the spatial domain image, it better approximates the spatial domain of the better quality images in set 1 than in set 2. But that would also means that the non-LL subbands (known as the detail subbands) of the images in set 1 retain less information than retained by the corresponding subbands in set 2. We argue that this is why all schemes perform better on set 2 than on set 1.

Finally, we note that the observed pattern of variation in the performance of face recognition, using different sub-bands is consistent with known results for wavelet-based face recognition without degradation [17].

Conclusion

We investigated the RIP property for random, and deterministic, constructions of CS overcomplete dictionaries as well as an existing learnt dictionary that trained on a set of high-resolution face images. These dictionaries were used to generate super resolved image with the aim of using for face recognition in uncontrolled conditions where the

input is degraded blurred LR image with a wide range of degradation. These results effectively support the use of SR based techniques that employ CS dictionaries for recovering super-resolved images that are suitable for face recognition. More importantly, that there is no need for using image sets for training dictionaries, because non-adaptive dictionaries perform equally well if not better in some cases. In order to find possible explanation, we conducted a number of tests of numerical matrix parameters relevant to the RIP condition. We note that the learning image-based dictionary is highly ill conditioned and far from satisfying the RIP related conditions. Perhaps the use of image patches with the same statistical parameters of general face image patches compensate for the lack of RIP properties.

Future work will focus on developing implicit construction of RIP dictionaries from a single basis with actions of certain finite groups on \mathbb{R}^m , and report on a new construction approach that result in RIP matrices implicitly satisfying known bounds on singular values. Such schemes are useful for revocable face biometrics. Moreover, we should also test our future developed schemes on a database of genuine CCTV face images.

References

1. Hennings-Yeomans PH, Baker S, Kumar BV (2008) Simultaneous super-resolution and feature extraction for recognition of low-resolution faces. In proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition.
2. He J, Zhang D (2010) Face super-resolution reconstruction and recognition from low-resolution image sequences. Computer Engineering and Technology (ICCET), 2nd International Conference on 2: 620-624.
3. Zeyde R, Protter M, Elad M (2010) On Single Image Scale-Up using Sparse-Representation. Springer, Curves and Surfaces 711-730.
4. Wang Y, Lucey S, Cohn JF (2008) Enforcing Convexity for Improved Alignment with Constrained Local Models. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit 2008: 1-8.
5. Yang J, Wright J, Huang TS, Yi Ma (2010) Image super-resolution via sparse representation. IEEE Transactions on Image Processing 11: 2861-2873.
6. Yang J, Tang H, Yi Ma, Huang T (2008) Face hallucination via sparse coding. Image Processing. 15th IEEE International Conference on 1264-1267.
7. Wang S, Zhang L, Liang Y, Pan Q (2012) Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis. Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on 2216-2223.
8. AL-Hassan N, Jassim SA, Sellaheewa H (2012) Enhancing face recognition at a distance using super resolution. ACM, MMSec 123-132.
9. Makwana RR, Mehta ND (2013) Single Image Super-Resolution VIA Iterative Back Projection Based Canny Edge Detection and a Gabor Filter Prior. International Journal of Soft Computing and Engineering (IJSCE) 3: 379-384.
10. Bruckstein AM, Donoho DL, Elad M (2009) From sparse solutions of systems of equations to sparse modeling of signals and images. SIAM Review 51: 34-81.
11. Rubinstein R, Bruckstein AM, Elad M (2010) Dictionaries for sparse representation modelling. Proceedings of the IEEE 6: 1045-1057.
12. Fornasier M, Rauhut H (2011) Compressive Sensing. In Scherzer O (ed). Handbook of Mathematical Methods in Imaging, pages, Springer.
13. Candes EJ, Tao T (2006) Near-Optimal Signal Recovery from Random Projections: Universal Encoding Strategies. IEEE Transaction on Information Theory 12: 5406-5425.
14. Baraniuk R, Davenport M, DeVore R, Wakin M (2008) A simple proof of the restricted isometry property for random matrices. Constr Approx 28: 253-263.
15. Gan L, Ling C, Do TT, Tran TD (2009) Analysis of the statistical restricted isometry property for deterministic sensing matrices using Stein's method. Citeseer.

16. Bajwa WU, Haupt JD, Raz GM (2007) Toeplitz-Structured Compressed Sensing Matrices. IEEE 14th Workshop on 294-298.
17. Al-Obaydy W, Sellahewa H (2011) On using high-definition body worn cameras for face recognition from a distance. Springer-Verlag 6583: 193-204.