N. Al-Hassan, S. A. Jassim and H. Sellahewa, "Construction of dictionaries to reconstruct high-resolution images for face recognition," presented at the 2013 International Conference on Biometrics, ICB 2013, Madrid; Spain, 2013.

Construction of Dictionaries to Reconstruct High-Resolution Images for Face Recognition

Nadia AL-Hassan   Sabah A. Jassim   Harin Sellahewa

Department of Applied Computing, University of Buckingham

Buckingham, MK18 1EG, UK

{Nadia.al-hassan; sabah.jassim; harin.sellahewa}@buckingham.ac.uk

Abstract

This paper presents an investigation into the construction of over-complete dictionaries to use in reconstructing a super resolution image from a single input low-resolution image for face recognition at a distance. The ultimate aim is to exploit the recently developed Compressive Sensing (CS) theory to develop scalable face recognition schemes that do not require training. Here we shall demonstrate that dictionaries that satisfy the Restricted Isometry Property (RIP) used for CS can achieve face recognition accuracy levels as good as those achieved by dictionaries that are learned from face image databases using elaborate procedures.

1. Introduction

Developing face recognition schemes that perform well in uncontrolled conditions, such as in surveillance situations and at a distance, is a tough challenge due to many factors including difficulties in determining a model for image degradation that encompasses various conditions. The performance of face recognition schemes whose feature extraction is based on using training sets of images is strongly dependent on how representative of the real world is the training set.

Face resolution enhancement is deemed to be necessary for accurate face recognition in many surveillance scenarios, where there is always a large distance between the camera and the objects (people) of interest or where a cheaper, lower resolution camera is used. Super-resolution (SR) is an inverse problem used as a pre-processing technique for face recognition, which deals with the recovery (reconstruction) of a high-resolution image from a given low-resolution image by a procedure that involves up-sampling, using criteria for estimating the missing details, followed by procedures to filter out distortions such as noise and blurring effects. Mathematically, the SR problem can be modeled as a solution of the matrix equation:

$$y = \text{HBx} + \text{n} \qquad (1)$$

where y is the known low-resolution (LR) image, H and B are the decimation and blurring matrices respectively; x is the unknown high-resolution (HR) image and n is additive noise. SR image reconstruction is generally a severely ill-posed problem because of the insufficient number of LR images, ill-conditioned registration, unknown blurring operator and the solution from the reconstruction constraint is not unique. Therefore, various regularization techniques are proposed to stabilize the inversion of this ill-posed problem [1], [2], [3]. However, the performance of these algorithms degrades rapidly when the desired magnification factor is large or the number of the input images is small.

The emerging paradigm of compressive sensing (CS) has been recently investigated as a method for improving signal/image resolution [4], [5], [6], [7]. Compressive sensing is a novel paradigm of signal sampling that exploits the empirical observation that many types of signals and images can be quite accurately approximated by sparse expansion in terms of suitable bases by a relatively small number of non-zero coefficients. The basic premise of CS is that sparse signals can be recovered from far fewer samples than those required by the classical Shannon-Nyquist Theorem, which specifies that, to avoid losing information when capturing/reconstructing a signal, one must sample at least two times faster than the signal bandwidth.

The basic principle of sparse coding (sparse representation) assumes that a natural signal can be compactly expressed and represented efficiently as a linear combination of atom (vector) signals from a pre-specified over-complete dictionary, where most of the linear coefficients are zeros. A dictionary [8] is a generalization of vector space basis represented as an over-complete matrix $m \times n$, where $m \ll n$ whose columns form a pool of bases whereby any vector in R^n can have multiple representations in terms of different bases in R^m.

The sparse coding problem to a column signal $x \in R^m$ can be described as:

$$\alpha_0 = \min_{x} \|\alpha\|_1 \; s.t. \|x - D\alpha\|_2 \qquad (1)$$

where α_0 is the sparse presentation of x and $D = \{d_1, d_2, \ldots \ldots d_n\} \in R^{m \times n}$ is an over-complete dictionary.

In the past few years, many people have been using CS on single image SR problems to reconstruct a high-resolution image. In [5], [6], [7], Yang, et al. proposed a new method to reconstruct a super resolved image based on sparse representations from an appropriately chosen pair of over-complete dictionaries whose columns are closely related to patches of high and low resolution images. In particular, the high and low dictionaries, D_H and D_L, are created by simply randomly sampling raw patches from the HR and the corresponding LR training images from a database of similar statistical nature to the input image, where the LR images are generated from the HR training images database by blurring and down-sampling. A sparse representation α_0 of the input LR image patch y is computed using the low-dictionary D_L built from LR images, and the HR image patch x is estimated using the given sparse representation with respect to the high-dictionary D_H. Back project iterative method is used to remove possible artifacts from the sparse representation stage and to eliminate the reconstruction errors in the estimated HR image (see also [9], [10]).

Studer and Christen, in [11], presented the K-SVD method for generating the learning dictionary to up sample the LR image. It is an iterative algorithm that alternates between a sparse coding of the training samples for a current dictionary and dictionary update step, such that it better fits the data. The K-SVD algorithm performs a singular value decomposition (SVD) for each of the K different sub-matrices, and hence the name K-SVD (K is used as the number of columns in the dictionary; and SVD will be defined in section 2).

In [12], Biswas et al. proposed an approach for matching non-frontal LR image with HR frontal gallery images by using Multidimensional Scaling (MDS) technique to transform the feature of LR and HR images such that the distance between them approximate the distances had the probe image been taken in the same conditions as the gallery. They used SIFT descriptors at fiducial locations as the features for performing recognition. Suitable scale changes in SIFT descriptor computation are made to make the comparison across resolution feasible.

In recent work [13], we considered the problem of matching LR probe face images with HR gallery images. We proposed to reconstruct the feature vector of the LR probe image by using dictionaries constructed from the high frequency training face images and compared the method with state-of-the art adaptive learning dictionary based SR method [6]. Where learning dictionary based on images database help wavelet-based features of super-resolved LR face images from SD cameras to achieve similar levels of accuracies with those of costly HR images captured from HD cameras.

In this paper, we investigate CS based sparse representation of a sample image in order to reconstruct a HR image for face recognition by taking different approaches to construct dictionaries as an alternative to the learning dictionary (LD) approach that is based on images (see section 3.3). In particular, our focus is on constructing dictionaries without relying on training images. We shall evaluate the performance of different dictionaries in terms of the quality of their super-resolved images, face recognition accuracy, as well as their CS properties.

The rest of the paper is organized as follows. Mathematical Properties of CS is presented in section 2. Restricted Isometry Property of dictionaries is reviewed in section 3. Experiments and results are discussed in section 4, and finally, section 5 is devoted to conclusions.

2. Properties of CS Dictionaries

The main concern of CS is the recovery of a sparse signal from a relatively small number of measurements of the signal. A vector x is k-sparse if at most k coefficients x_i are non-zero.

Two closely related properties of underdetermined matrices have been identified in the CS literature in relation to recovery of sparse signals: the Null Space Property (NSP) and the Restricted Isometry Property (RIP).

The Null Space Property (NSP):

An $m \times n$ matrix (i.e. dictionary) $A \in R^{m \times n}, m \ll n,$ is said to be an NSP of order k if every 2k columns of A are linearly independent. This definition shows that NSP matrices can be formed as a concatenation of different bases as long as certain conditions are satisfied. The following well-known theorem is straight forward, [14], but for clarity, we repeat its proof below.

Theorem 1. If A is NSP of order k matrix then $Ax = b$ has a unique k-sparse solution.

Assume that there are two k-sparse signals x_1 and $x_0 \in R^n$ with $A x_1 = A x_0$, which implies $A (x_1 - x_0) = 0$. However, $(x_1 - x_0)$ is 2k-sparse, so there is a linear dependency between 2k columns of A, contradiction due to A is NSP.

This theorem shows how to construct the unique solution to the L_0 − minimisation problem. Unfortunately, this is a non-convex optimization problem, which in contrast to the L_2 − minimisation problem (least-squares) is computationally intractable. Fortunately, the corresponding L_1 − minimization problem can be solved efficiently and under certain slightly stronger condition than the NSP, it has a unique solution that is also the solution of the corresponding L_0 − minimization problem. The required sufficient condition that the dictionary must satisfy, in this case, is the Restricted Isometry Property first introduced by Candes and Tao [14]. RIP is defined as follows:

The Restricted Isometry Property (RIP):

A matrix A of size $m \times n$ is said to have the RIP of order k if there exists a Restricted Isometry Constant (RIC) $0 < \delta_k < 1$ such that:

$$(1 - \delta_k)\|x\|_2 \leq \|Ax\|_2 \leq (1 + \delta_k)\|x\|_2 \qquad (3)$$

for all k-sparse vectors x.

RIP requires that all column sub-matrices of A of k columns must be well conditioned. This is a very restrictive condition and quite difficult to check. Calculating the condition number of the matrix A is a simpler way of ensuring recovery of sparse vectors. The condition number of A can be calculated using the Singular Value Decomposition (SVD) of A.

Singular Value Decomposition (SVD):

The SVD of any matrix $R^{m \times n}$ $where$ $m < n$, is of the form $A = USV^T$ where U and V are unitary-matrix matrices of size $m \times m$ and $n \times n$, respectively, and S is a $m \times n$ rectangular matrix with non-negative real numbers on the diagonal and zero elsewhere. The columns of each of U and V form orthonormal bases.

The diagonal entries $S_{i,i}$, $denoted$ by σ_i, are the non-zero singular values of A arranged in decreasing order. In fact, σ_i are the square roots of the non-zero eigenvalues λ_i of AA^T, for $1 \leq i \leq r$.

The condition number of the matrix A, defined as:

$$Condition(A) = \frac{\sigma_{max}}{\sigma_{min}}$$

where σ_{max} and σ_{min} represent the maximum and minimum singular values of A. A high condition number points to an ill-conditioned matrix, whereas a low condition number points to a well-conditioned matrix [2], [15]. This will be a measure of the strength of the RIP.

3. RIP Dictionaries – Examples

In this section, we describe examples of different, random as well as deterministic, dictionaries and discuss their RIP strength. We shall investigate the use of these dictionaries, as alternatives to the existing methods of using training images to create dictionaries, for CS based super-resolution.

3.1. Gaussian Random Matrix (GRM)

A widely used RIP dictionary is the Gaussian Random Matrix. For Gaussian Random Matrix, the entries $x_{i,j}$ of the matrix A of size $m \times n$ are independently sampled from a normal distribution $x_{i,j} \sim N(0, 1/m)$ with expectation of $\mu = 0$ and variance $\sigma^2 = 1/m$, where the normal distribution function is as follows:-

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad (4)$$

In order to reconstruct a HR image from a single LR image for face recognition via sparse representation, two over-complete Gaussian random matrices D_H and D_L of size 25×512 and 100×512 respectively have been created from a zero mean Gaussian distribution with variance $1/25$, and $L_2 - norm$ used to normalize each column. Therefore, the over-complete dictionary D_H satisfies the sufficient property for CS and its condition number (i.e. ratio between largest and minimum singular value) is known to be of high RIP strength. In fact, the Gaussian dictionary, generated in our experiments, has a reasonably low condition number $\frac{5.3466}{3.6789}$ =1.453.

3.2. Toeplitz-Circular Random Matrix (T-CRM)

The Gaussian matrix provides an optimal condition for the minimal number of required samples for sparse recovery. However, some CS applications often do not allow the use of "completely" random matrices, but put certain physical constraints on the measurement process and limit the amount of randomness that can be used. This leads then to a structured random measurement matrix. Bajwa et al. [16] have shown that Toeplitz-structured matrices are sufficient to recover undersampled sparse signals. These matrices are important in applications such as wireless communications and radar [17], [18]. Toeplitz and Circular matrices of the size $k \times n$ are of the form, respectively:

$$T = \begin{bmatrix} t_n & t_{n-1} & \cdots & t_1 \\ t_{n+1} & t_n & \cdots & t_2 \\ & & \vdots & \\ t_{n+k-1} & t_{n+k-2} & \cdots & t_k \end{bmatrix},$$

and

$$C = \begin{bmatrix} t_n & t_{n-1} & \cdots & t_1 \\ t_1 & t_n & \cdots & t_2 \\ & & \vdots & \\ t_{n-1} & t_{n-2} & \cdots & t_k \end{bmatrix}$$

where every left-to-right descending diagonal is constant. These matrices satisfy the RIP. The high-dictionary D_H of size 25×512 and the low-dictionary D_L of size 100×512 are generated as Toeplitz-Circular Random Matrices where the first row consist of standard Gaussian random variables with zero mean and standard deviation equal one, and the rest of the rows are permuted versions of the first row as indicated in the above format. The high-dictionary satisfies the compressive-sensing RIP property with a similar strength to that of Gaussian dictionaries. In fact, the condition number of the T-CRM high dictionary generated for our experiments is equal to $\frac{5.25290}{3.6146} = 1.453$.

3.3. Image Learning Dictionary (LD)

Learning dictionary is based on the work of Yang et al. [5], [6] referred to in the introduction.

The LD approaches for SR uses two appropriately chosen over-complete dictionaries, D_H and D_L, that are created by simply randomly sampling raw patches from HR and the corresponding LR training images from a database of similar statistical nature to the input image. The LR training images are generated by blurring and down-sampling HR training images.

For example, Yang et al. prepared dictionaries using flower images as training data to generate high-quality images from LR images with simple texture. In another experiment, they created dictionaries relying on animal images to reconstruct LR input images with complex textures.

We note that exhaustively testing Yang's LD dictionary D_H, for NSP property is not realistic since there are

$$\binom{512}{25} = \frac{512\,!}{25!\ (512-25)!}$$

such 25×25 submatrices of the dictionary. As an indicator, however, we opted to evaluate the determinants of more than a hundred randomly selected 25×25 submatrices. The results are displayed in Figure 2, below, which illustrates that although in theory, the dictionary may satisfy NSP but practically and computationally this is not the case because the determinant of most of these submatrices are very close to zero. Moreover, the condition number of the learning dictionary is very large, nearly 1.00E+15, which makes this dictionary slightly ill-conditioned with extremely low RIP strength.

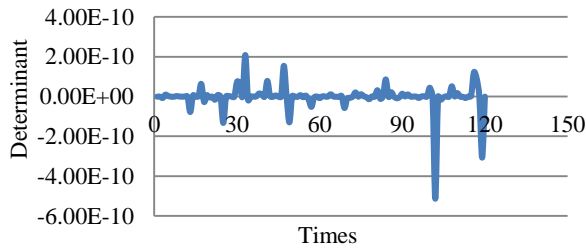These observations have motivated our search for other dictionaries with reasonable RIP strength.

Figure 2. Determinant of submatrices from the High-Learning Dictionary.

3.4. Constructed Deterministic Dictionaries

Our approach to constructing deterministic dictionaries is motivated first by the desire to build NSP dictionaries by concatenated bases and second by attempt to reduce the amount of randomness in the dictionary construction. We construct an over-complete dictionary of size $m \times n$ where $m \ll n$ by concatenating a carefully selected sets of linearly independent m-column vectors that are not only distinct but any pair of which satisfies the NSP property.

We shall refer to the proposed class of dictionaries as Linearly Independent Dictionary (LID). The constructed dictionary is a collection of bases, starting with an identity matrix and gradually adding $m \times m$ matrices with linear independent vectors of R^m and each non-zeros coefficients vector from the matrix is a linear combination of the other set of vectors. The proof that such dictionaries could have the NSP property is beyond the scope of this paper. We shall now describe such a construction. Moreover, discuss their properties.

Let,
$$\mathrm{Id} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \vdots & \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

For real numbers $p_i > 1$ $such$ $that$ $p_i \neq p_{i+1}$;

$$A_{p_i} = \begin{pmatrix} 1 & \frac{1}{p_i} & \frac{1}{p_i{}^2} & \dots & \frac{1}{p_i{}^{24}} \\ \frac{1}{p_i} & 1 & \frac{1}{p_i} & \dots & \frac{1}{p_i{}^{23}} \\ & & \vdots & & \\ \frac{1}{p_i{}^{24}} & \frac{1}{p_i{}^{23}} & \frac{1}{p_i{}^{22}} & \dots & 1 \end{pmatrix}, \text{ where } i = 1,2,..,k$$

and

$$C = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{12} \\ \frac{1}{2} & 1 & \frac{1}{2} & \dots & \frac{1}{11} \\ & & \vdots & & \\ \frac{1}{12} & \frac{1}{11} & \frac{1}{10} & \dots & 1 \\ & & \vdots & & \\ \frac{1}{m} & \frac{1}{m-1} & \frac{1}{m-2} & \dots & \frac{1}{m-11} \end{pmatrix}$$

Then, the $m \times n$ dictionary is as the following form:-

$$D = \begin{bmatrix} Id, A_{p_1}, A_{p_2}, \dots, A_{p_k}, C \end{bmatrix} \tag{5}$$

First we generated the high-dictionary D_H of size 25×512 using the integers numbers $2 \leq p_i \leq 20$, where $i = 1,2,\dots,19$.

To test the strength of evidence that such a dictionary satisfies the NSP property, we display in Figure 3 the determinants of more than a hundred 25×25 sub-matrixes. In comparison to the LD dictionary, all the determinant values of the example LID (i.e. LID₁) are comfortably away from zero. Moreover, the over-complete dictionary is well conditioned as its condition number is small and equal to $\frac{6.79695}{3.4363} = 1.977 < 2.5$.

On the other hand, we constructed another such an over-complete dictionary (i.e. LID₂) with real numbers $1 <$

$p_i < 2$. This dictionary seems to have a weak RIP property where the condition number is equal $\frac{13.82}{1.19} = 11.55$. This indicates that the RIP strength is p_i dependent.
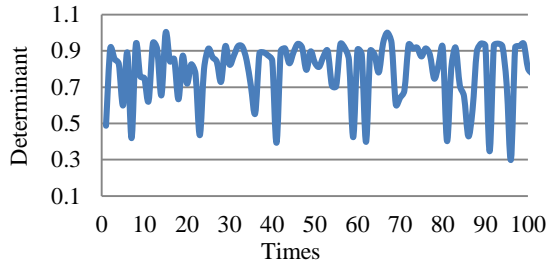

Figure 3. Determinant of submatrices from the Constructed dictionary with $2 \leq p_i \leq 20$

For simplicity, the low-dictionary D_L of size 100×512 in all the experiments is created from a Standard Gaussian Random Matrix (SGRM) with zero mean and variance of one and uses L_2 -norm to normalize each column in the dictionaries.

4. Experimental Work

A number of experiments were conducted on two publically available face databases to evaluate the suitability of the proposed dictionaries to super-resolve face images and for face recognition in particular. A brief description of the two databases and experiment protocols are given below before discussing the results in section 4.2.

4.1. Experiment Protocol and Datasets

The UBHSD video database [19] contains 160 videos of 20 subjects recorded in two sessions. It includes videos captured in indoor and outdoor locations; two video recordings – one high-definition (HD) and one standard-definition (SD) – of a subject were captured at each location. Twelve frames are selected to capture a subject at four distance ranges. The database consists of blurred face images, faces with eyes closed and slightly varying poses. Each subject has 96 face images of size 128×128 pixels capturing a total of 1920 images. The gallery set G consists of 60 images (3 images per subject) in Session 1 and the probe set consists of 480 images (24 images per subject) from all four ranges, in both indoor and outdoor videos in Session 2.

The second database is the widely used Extended Yale B [20], [21] which consists of images of 38 subjects, each having 64 images size of each image is 192×168 pixels captured under different illumination conditions. The total number of images in the database is 2414. The images in the Extended Yale B are divided into five illumination subsets according to the direction of light source of the camera. We selected the P00A+000E+00 image of each subject for the

gallery set and the remaining images, which we blurred and down sampled by half, make up the probe set (i.e. LR images).

4.1. Results - Accuracy

Dictionary based SR is used on four different combinations of gallery/probe image resolutions to super-resolve LR images of size 64×64 – the LR images were obtained by down sampling the original images – to their original size of 128×128 pixels. Z-score normalized LH subband coefficients of Haar wavelet transform at decomposition level three is used as a feature vector for recognition [22]. Nearest-neighbor classification with City Block distance is used to classify probe images. In all the experiments, the learning dictionary (LD) is constructed using the training set, which contains three HR images of each subject from subset 1 of Extended Yale B database unattached from gallery/probe images.

The experimental results for UBHSD database in Table 1 show that, the SR by GRM, T-CRM and SR by learning dictionary produced marginally better accuracy rates in some cases and less by a similar amount in some other cases. Moreover, there is no significant difference in identification accuracy rates whether one use T-CRM, GRM, LD or LID_1 and LID_2 for dictionaries.

Dictionary methods on the Extended Yale B database were also tested. We observed similar results for the experiments on this database, but could not include details due to page limitations. Therefore, table 2 shows only the results on the well-lit face images in set1 and set2 and demonstrates there is no difference in identification accuracy rates between dictionary methods when the LH_3 subband features are used. In general, the experimental results in the two databases demonstrate that there is no significant difference in the identification accuracy rates between these methods of building the Dictionary.

G set	P set	SR	Gallery Image Range			
			R_1	R_2	R_3	R_4
SD$_{128}$	SD$_{64}$	LD	76.25	68.33	71.25	69.79
		GRM	76.04	68.33	71.45	69.79
		T-CRM	76.04	68.54	71.45	69.79
		LID$_1$	76.04	68.12	71.25	69.58
		LID$_2$	76.04	68.12	71.25	69.58
		BC	76.66	68.75	71.66	69.79
SD$_{64}$	HD$_{128}$	LD	75.62	71.04	71.25	67.91
		GRM	75.41	71.04	71.25	68.54
		T-CRM	75.41	71.04	71.45	68.54
		LID$_1$	75.41	71.04	71.25	68.33
		LID$_2$	75.41	71.04	71.04	68.33
		BC	75.62	71.25	71.25	68.75
HD$_{128}$	SD$_{64}$	LD	68.95	65.62	72.08	74.58
		GRM	68.75	65.41	72.08	75
		T-CRM	68.75	65.41	71.87	74.58
		LID$_1$	68.54	65.62	71.87	74.58

LID$_2$	68.54	65.41	71.87	74.58
BC	68.75	66.66	72.50	75

	LD	76.25	68.54	71.25	69.58
	GRM	76.04	68.12	71.66	69.79
SD$_{64}$ SD$_{64}$	T-CRM	76.04	68.12	71.66	70.00
	LID$_1$	76.04	68.12	71.66	69.16
	LID$_2$	76.25	67.91	71.66	69.79
	BC	76.04	68.95	71.45	69.79

Table 1. Recognition accuracy rates (%) for the UBHSD database using different dictionaries in the super-resolution process.

S	SB	LD	GRM	TCRM	LID$_1$	LID$_2$	BC
	LL$_3$	98.68	98.68	98.68	98.68	98.68	98.68
S$_1$	LH$_3$	92.10	92.10	92.10	92.10	92.10	92.10
	HL$_3$	92.10	90.78	92.10	92.10	90.78	92.10
	HH$_3$	78.94	78.94	78.94	78.94	78.94	78.94
	LL$_3$	78.72	78.72	78.72	78.72	78.72	77.63
S$_2$	LH$_3$	100	100	100	100	100	100
	HL$_3$	95.61	96.27	95.61	95.83	96.05	95.83
	HH$_3$	98.02	98.24	98.24	98.24	98.24	99.56

Table 2. Recognition accuracy rates (%) for the Extended Yale B database using different dictionaries in the super-resolution process.

4.2. Results – Image quality

The second stage of the experiments is to compare the performance of SR by different dictionaries with well-known Bicubic-interpolation (BC) in terms of enhancing the resolution of the LR input image; the reconstruction fidelity was quantified using PSNR measure. The original images of the database are used as the ground truth to calculate the PSNR values. It can be seen from Figure 4 that, there is a small difference in quality of images between super resolutions by different dictionaries independent of the tested images. Moreover, SR by LD slightly outperforms the well-known interpolation method at each distance range and for each sessions and different conditions (indoor and outdoor).
Figure 5 illustrates PSNR results for the Extended Yale B database. The results show that, the BC method produced lower quality image than dictionary methods. In addition, LD based on images produce better quality images than SR methods by dictionary not depend on images. In general, our experiments show that the slightly better improvement in resolutions comes from the SR by dictionary based on images but this improvement is no significant. This leads us to the final stage of our work. A sever degradation function (atmospheric turbulence) has been used to obtain LR test image and the different dictionary methods are used to reconstruct its HR image. This model of degradation can even take into account environmental conditions that cause degradations. This model has a form:

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

Where a constant k depends on the nature of the turbulence, in this case severe turbulence when $k = 0.05 \ or \ k = 0.07$. In subjective tests, as can be seen in Figure 6, the reconstructed images by different dictionaries methods are quite similar to each other. This leads to conclude that a training images database is not necessary to create the dictionaries for image super resolution.
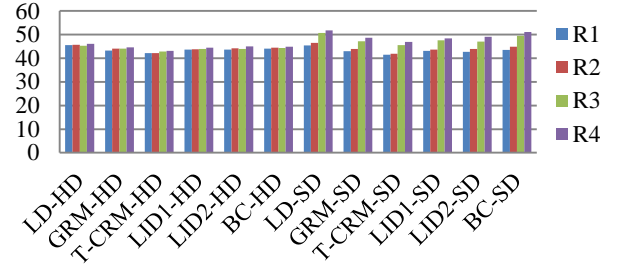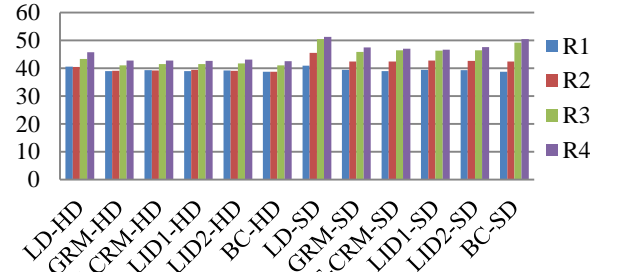

HD and SD video data from indoor, session 1
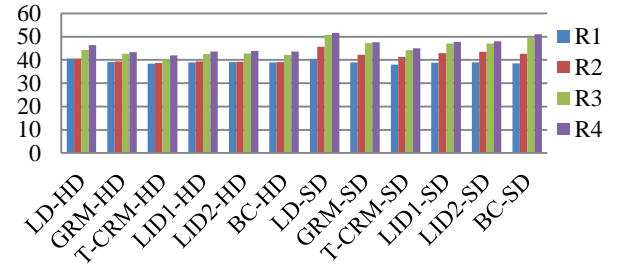
HD and SD video data from outdoor, session1
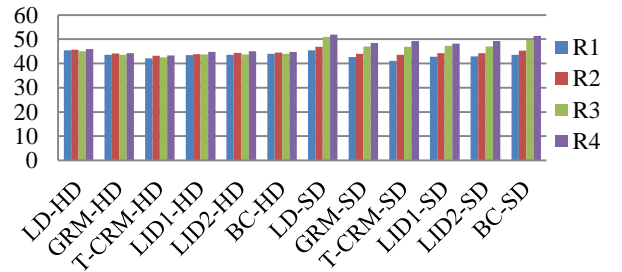
HD and SD video data from indoor, session 2

HD and SD video data from outdoor, session 2

Figure 4. A comparison of PSNR values for super resolved images in the different distance ranges are found.
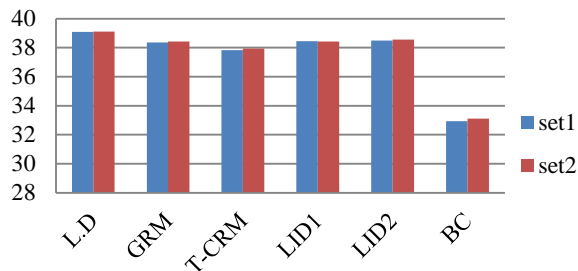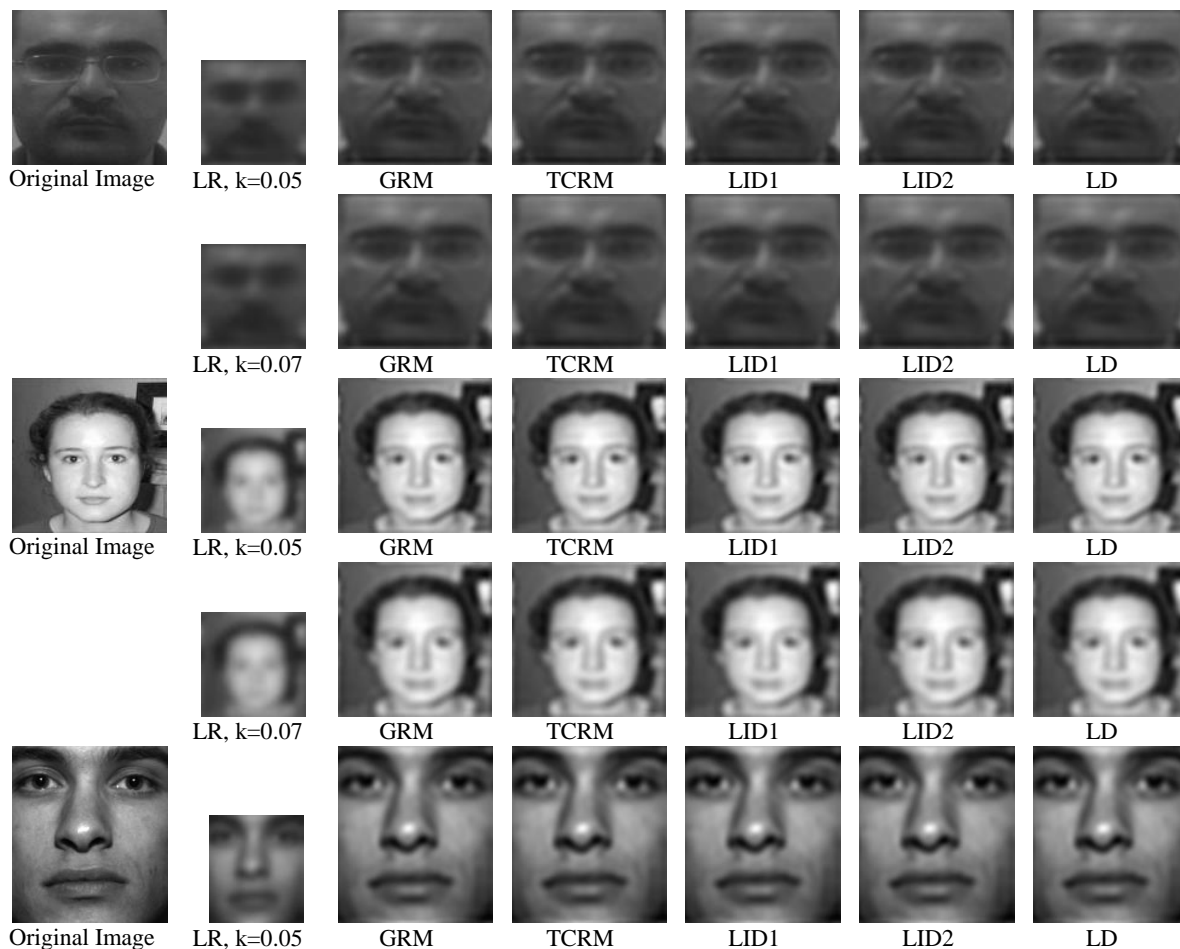

Figure 5. Comparison of PSNR values for super resolved images in Extended Yale B database.

5. Conclusion

We investigated the use of random and deterministic CS-based over-complete dictionaries for SR enhancement of low-resolution degraded images for face recognition in uncontrolled conditions and/or captured at a distance, and compared the results with that of using an existing learning dictionary that is trained on a set of high-resolution face images and their low-resolution versions obtained by blurring and down sampling. The experimental results demonstrated that the non-adaptive (i.e. independent of images) dictionaries perform as well as the LD dictionary in terms of recognition accuracy and quality of recovered images. The significance of this work, that will be further developed in the future, is that CS can be used to design scalable face recognition schemes in unsupervised manner. Considerations of the fundamental CS properties of NSP and RIP, which ensure k-sparse vectors, the experimental results raise some fundamental questions about the importance of these properties for pattern recognition and on the practical implications of non-compliance with theoretical CS hypothesis. These issues need to be further investigated in the future.
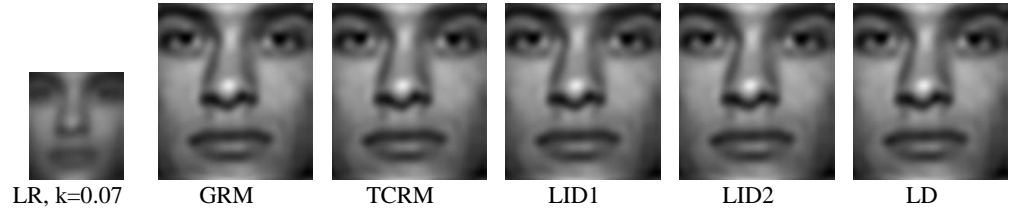

| Original Image | LR, k=0.05 | GRM | TCRM | LID1 | LID2 | LD |

| LR, k=0.07 | GRM | TCRM | LID1 | LID2 | LD |

| Original Image | LR, k=0.05 | GRM | TCRM | LID1 | LID2 | LD |

| LR, k=0.07 | GRM | TCRM | LID1 | LID2 | LD |

| Original Image | LR, k=0.05 | GRM | TCRM | LID1 | LID2 | LD |

| LR, k=0.07 | GRM | TCRM | LID1 | LID2 | LD |

Figure 5. Comparison between SR approaches by different dictionaries.

Acknowledgments

References

[1] S. Farsiu, E. Michael, and M. Peyman. Fast and robust multiframe super resolution. IEEE Transactions on Image processing, 13(10):1327-1344, 2004.

[2] V. Bannore. Iterative-Interpolation Super-Resolution Image Reconstruction. Springer, 19-50, 2009.

[3] M.V. Zibetti, F.S. Bazan, and J. Mayer. Determining the regularization parameters for super-resolution problems. Signal Processing, 88(12):2890-2901, 2008.

[4] R. Zeyde, M. Protter, and M. Elad. On Single Image Scale-Up using Sparse-Representation. Springer, curves and surfaces, 711-730, 2010.

[5] J. Yang, and et al. Image super-resolution as sparse representation of raw image patches. In Proc. IEEE Computer Society Conference on computer Vision and Pattern Recognition, 1-8, 2008.

[6] J. Yang, and et al. Image super-resolution via sparse representation. IEEE Transactions on Image Processing, 11(19):2861-2873, 2010.

[7] J. Yang, and et al. Face hallucination via sparse coding. Image Processing, 15th IEEE international Conference on, 1264-1267, 2008.

[8] R. Rubinstein, A.M. Bruckstein, and M. Elad. Dictionaries for sparse representation modeling. Proceedings of the IEEE, 98(6):1045-1057, 2010.

[9] A. Ganesh, J. Wright, A.Y. Yang, and S.S. Sastry. Robust Face Recognition via Sparse Representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2(31): 210-227, 2009.

[10] A.Y. Yang, J. Wright, and Y. Ma. Feature selection in face recognition: A sparse representation perspective. IEEE Transactions Pattern Analysis and Machine Intelligence, 2007.

[11] C. Studer, and S. Christen. Dictionary Learning for Super-Resolution. Semester Project in Information Technology and Electrical Engineering. 2010.

[12] S.Biswas, G. Aggarwal & P.J.Flynn. Pose-robust recognition of low-resolution face images. *IEEE*, (pp. 601-608),2011.

[13] N. Al-Hassan, S.A. Jassim, and H. Sellahewa. Enhancing Face Recognition at a Distance using Super Resolution. ACM, MMSec, p. 123-132, 2012.

[14] E.J. Candes and T. Tao. Near-Optimal Signal Recovery from Random Projections: Universal Encoding Strategies?. IEEE Transaction on Information Theory. 52(12): 5406-5425, 2006.

[15] Z. Chen, and J.J. Dongarra. Condition numbers of Gaussian random matrices. SIAM Journal on Matrix Analysis and Applications. 27(3): 603-620, 2005.

[16] W.U. Bajwa, J.D. Haupt, and G.M. Raz. Toeplitz-Structured Compressed Sensing Matrices. IEEE 14th Workshop on SSP'07. 294-298, 2007.

[17] H. Rauhut. Compressive Sensing and Structured Random Matrices. Theoretical foundations and numerical methods for sparse recovery. 9: 1-92, 2010.

[18] Y. Lei, and P. Jean. Toeplitz-structured Chaotic Sensing Matrix for Compressive Sensing. 229-233, 2010.

[19] W. Al-Obaydy, and H. and Sellahewa. On using high-definition body worn cameras for face recognition from a distance. Springer-Verlag, 193-204, 2011

[20] K.C. Lee, J. Ho, and D.J. Kriegman. Acquiring Linear Subspaces for Face Recognition under Variable Lighting. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(5):684-698, 2005.

[21] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman. From few too many: Illumination cone models for face recognition under variable lighting and pose. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6):643-660, 2001.

[22] H. Sellahewa, S.A. and Jassim. Image-Quality-Based Adaptive Face Recognition. Instrumentation and Measurement, IEEE Transactions on, 59(4):805-813, 2010.