SPCA: A No-reference Image Quality Assessment based on the Statistic Property of the PCA on Nature Images

Yun Zhang, Chao Wang, Xuanqin Mou
Institute of Image Processing and Pattern Recognition, Xi'an Jiaotong University, China; xqmou@mail.xjtu.edu.cn

ABSTRACT

Despite the acceptable performance of current full-reference image quality assessment (IQA) algorithms, the need for a reference signal limits their application, and calls for reliable no-reference algorithms. Most no-reference IQA approaches are distortion specific, aiming to measure image blur, JPEG blocking or JPEG2000 ringing artifacts respectively. In this paper, we proposed a no-reference IQA algorithm based on the statistic property of principal component analysis on nature image, named SPCA, which does not assume any specific type of distortion of the image. The method gets statistics of discrete cosine transform coefficients from the distort image’s principal components. Those features are trained by $\nu$-support vector regression method and finally test on LIVE database. The experimental results show a high correlation with human perception of quality (average over 90% by scores of SROCC), which is fairly competitive with the existing no-reference IQA metrics.

Keywords: No-reference Image Quality Assessment, Natural Scene Statistics, Principal Component Analysis (PCA), Discrete Cosine Transformation, Support Vector Regression

1. INTRODUCTION

Measurement of image quality is crucial for many image-processing algorithms, such as acquisition, compression, restoration, enhancement, and reproduction. Traditional image quality assessment (IQA) is full-reference algorithms which interpret image quality as similarity with a “reference” image. The obvious limitation of such approach is that the reference image may not be available to the IQA algorithm. The field of no-reference, or blind IQA, has attracted more and more attentions. Considerable full-reference algorithms have already established a significant correlation with subjective evaluation of image quality. However, still few no-reference image quality assessment algorithms (NR IQA) can both ensure the high reliability and generic performance now. Currently, NR IQA metrics can be divided into two major categories: 1) metrics based on specific distortion [1-3], 2) metrics based on natural image statistics (NIS) [4-6]. But, practically speaking, it is far from impossible to comprehensively quantify each type of distortion on a real degraded image by the first type of metrics. By contrast, NSS based NR IQA try to utilize the explicit mathematical properties and corresponding intuitive biological properties of natural image as the features to distinguish it from the distort image.

Natural images are statistically redundant. During the long time of sensory neurons evolution, the light reflected off the natural image reaches the finite photoreceptors, called ganglion cells, on the retina and has been effectively transmitted by the optic nerves through lateral geniculate nucleus (LGN) to the visual cortex to be “seen”. The neurons in such the earliest stage of visual pathway are assumed to be adapted through developmental processes to the statistical properties of the natural image signals to which they are exposed. Attneave (1954) and Barlow (1961) proposed that information theory could efficiency. This assumption has been verified from different angles of view. For example, the classical receptive field can be modeled by the Laplacian of Gaussian (LOG) function to produce the basic primitive structures of images [7]. The responses of simple cells in primary visual cortex can be modeled by Gabor-like filters with the constraint of sparsity or independency to learn the oscillation feature of images [8, 9]. The works of M. Zhang et al. [10] and W.F. Xue et al. [11] have shown that the features of early vision are suited for the design of IQA metrics. In this paper, we focus on the principle components (PCs) of natural images. the PCs are the basic statistical description of the visual appearance of our surroundings. Neurobiology researchers found that the unsupervised single cell learning rules of Hebbian type mostly lead to an extraction of the PCs. So the mechanisms of the PCs extraction are supposed to be a fundamental principle for the formation of receptive fields in the mammalian early visual system [1]. Computed very fast, PCA are widely be used as the preprocessing tools of dimension reduction, whitening and anti-aliasing. Even though
PCA has been an active area of research, we are not aware of any previous work which seeks to discover the PCs’ characters of the distorted images. NIS are used for both full-reference IQA and reduced-reference IQA, however neither of them explicitly characterize distorted image statistics (DIS) and utilize them on IQA. Recently, Mookht et.al [10] try to use the DIS to identify the distortion types which is meaningful for the specific distortion based IQA. Saad et al [11], achieved a competitive NR IQA based on the distribution difference between natural image and distorted image on discrete cosine transform (DCT) domain.

In this paper, we focus on the existing differences of the statistical properties of the PCA between distorted image and the nature image, then a NR IQA metric SPCA is proposed. In the rest of the paper, we will begin with the study of statistical property of the PCA on distort images. After a trail experimental analysis, we figure out the general routines of proposed SPCA metrics. Then in section three, we explicitly discuss the experiment setup and the methodology of support vector regression to get the qualified SPCA score. In section 4, we present the competitive NR IQA performance of SPCA to SROCC, LCC and RSME. All of those results point out that the PCA is a valuable tool for extract distorted images statistical features with fast speed and high quality result.

2. STATISTICAL PROPERTY OF THE PCA ON DISTORT IMAGES

2.1 Prerequisite notions

The idea of PCA is to convey the most information about a set of data given a limited number of linear descriptors. The data table to be analyzed by PCA comprised \( I \) observations described by \( J \) variable and it is represented by the \( I \times J \) matrix \( X \). The matrix \( X \) has rank \( L \), where \( L \leq \min(I, J) \). The date table will be preprocessed before the analysis generally. The columns of \( X \) will be centered so that the mean of wach column is equal to 0 i.e., \( X^T1 = 0 \), where \( 0 \) is a \( J \) by 1 vector of zeros and \( 1 \) is an \( I \times 1 \) vector of ones.

To extract the most important information from the data table \( X \), PCA computes new variable called principal components which are obtained as linear combinations of the original variables. The first principal component is required to have the largest possible variance. The second component is computed under the constraint of being orthogonal to the first component and has the largest possible variance. The other components are computed like wise. Formally, if \( X \) is a rectangular matrix, its singular value decomposition (SVD) is as follows:

\[
X = P\Delta Q^T
\]
where \( \mathbf{P} \) is the \( I \times L \) matrix of left singular vectors, \( \mathbf{Q} \) is the \( J \times L \) matrix of right singular vectors, and \( \mathbf{\Delta} \) is the diagonal matrix of singular values, which is the diagonal matrix of the eigenvalues of \( \mathbf{X}^T \mathbf{X} \) and \( \mathbf{X} \mathbf{X}^T \). In PCA, the components are obtained from the SVD of data table \( \mathbf{X} \). The matrix \( \mathbf{Q} \) can be interpreted as a projection matrix because multiplying \( \mathbf{X} \) by \( \mathbf{Q} \) gives the values of projections of the observation on the principal components.

### 2.2 Property of the PCA on Distort Images

To study the statistic properties on distort images, we first pick up 12 natural images (including buildings, landscapes, people, plants and human face etc.) with six different general distortion types, say, additive white gaussian noise (awgn), gaussian blur (gb), high frequency noise (hfn), quantization noise (quan), JPEG (jpeg) and JPEG2000 (jp2k), as shown in Figure 1. To quantitatively evaluate the inherent regularities of their statistic properties, the distortion levels of each type correspond to same levels of subjective assessment scores. For example, four distortion level of awgn are chosen (0.002/0.005/0.02/0.07) accordant with subjective scores of “Bad”, “Poor”, “Fair” and “Good” (ITU-R five-point quality scale). The other distortions’ levels are chose likewise as described in Figure 1. To setup the data table \( \mathbf{X} \) of original and distorted images, \( m \times m \) image patches \( (x,y) \) are random sampled on each picture for certain amount of times. Then, we will get the \( \mathbf{\Delta} \) in equation (1) which is actually the square root of the eigenvalues of matrix \( \mathbf{X}^T \mathbf{X} \). At the same time, we get the principal component matrix \( \mathbf{Q} \) which is the corresponding eigenvector of \( \mathbf{X}^T \mathbf{X} \). Take one of the natural image and its distorted images group as an instance as illustrate in Figure 2.

To begin with, we put the eigenvalue matrix both of the original image, denoted as \( \mathbf{\Delta}_o \) hereinafter, and of distorted images, denoted as \( \mathbf{\Delta}_d \) together to show their differences. Because there are six distortion types and four levels for each, twenty four \( \mathbf{\Delta}_d \)s and one \( \mathbf{\Delta}_o \) are drawn in the same coordinates at the leftmost in Figure 2. The thickest solid black line stands for the distribution of \( \mathbf{\Delta}_o \). The other color each belongs to a specific type of distortion and each line style stands for the same level of subject evaluation score. Three type of \( \mathbf{\Delta}_d \)s (awgn, hfn and quan) are mostly above the \( \mathbf{\Delta}_o \), while other three (gb, jpeg and jp2k) are mostly beneath the \( \mathbf{\Delta}_o \). Such a phenomenon can be explained intuitively. For one thing, distortions like awgn, hfn and quan are whitening processes, which make the pixel values more and more uncorrelated with each other. For another thing, distortions like gb, jpeg and jp2k are to get rid of the “detailed” information so that most of the original images are projected to the subspace composed of the first \( n \) principal complements. Whatever the distortion types, the higher the distortion level is the farther the distance between the corresponding \( \mathbf{\Delta}_d \) line and \( \mathbf{\Delta}_o \) line.

Next, we will analysis the characters the eigenvector matrix \( \mathbf{Q} \). Actually \( \mathbf{Q} \) is a \( m^2 \)-dimension orthogonal square matrix. If we restore each vector of \( \mathbf{Q} \) to an \( m \times m \) image patch and rearrange them together, we get the right part of the result in Figure 2. In the middle is the rearranged eigenvectors “image patches”. From the left to right and top to bottom, each patch stands for the first principal component, the second and so on. The rightmost of figure 2 is the same style of image patches combination of the twenty four distorted images eigenvectors. One of the properties of the principle components
is sinusoids, i.e. PCA actually perform some kind of Fourier analysis. Take a look at the eigenvectors pattern of the original image in figure 2, which is very similar to a discrete Fourier transform (DCT) of the image patches. It just because of the random sampling of $i(x,y)$ and computation of the covariance matrix $X^TX$ that $Q$ is not completely exact sinusoids. In contrast, the eigenvectors patterns of the distorted images haven’t shown such properties. Aside from the first $n$ principal components, the other doesn’t exhibit the typical sinusoids pattern, especially those located in middle and rear parts.

Such a property can be further proven in figure 3. Here we just take two distorted images of the same type but different levels to make it more explicit. The leftmost is the enlarged eigenvectors pattern of the original image, then follows the low level and high level distorted images of $jp2k$. The orange boxes on three pattern images point out the parts of most different principle components. Assuming that each principle component patch of naturel image should be highly correlated with DCT, we take one patch from the same position in three orange boxes and do DCT on them respectively, and the 3D view of their DCT coefficients are demonstrated in the rightmost of figure 3. Apparently, the DCT coefficients patch of the original image is extremely similar to a Dirac function, while the other two are rather kinds of oscillations everywhere. This can also be concisely explained in the histograms of the three DCT coefficient patches at the bottom of figure 3. Evidently, the histogram of the original image is much more concentrated than those of the distorted image. Moreover, the higher the distortion level is, the more uniformly the histogram distributes.

Thus, it can be concluded that the principle components of natural image and distorted image are distributed very differently. Such kind of the distinction can be perfectly interpreted by means of discrete Fourier transformation. Therefore, we proposed a PCA based NR IQA metric to effectively utilize such statistic characteristics to evaluate the image distortion. A general routine of SPCA is:

1. Randomly sample the distorted image $I(x,y)$ to get the data table $X$.
2. Perform SVD of $X$ as in Eq. (1) and get $Q$.
3. Restore each column of $Q$ to a principal component patch $u(x,y)$ and have 2D-DCT on them get $d(x,y)$.
4. Extract a sequence of statistics from each $d(x,y)$ and use support vector regression (SVR) to get an objective image quality evaluation.
3. EXPERIMENT AND METHODOLOGY

3.1 Get PCs on Distort Images

To test the function of SPCA, LIVE IQA database was used which contains 29 reference images and five distortion types: 1) JPEG2000 (227 images), 2) JPEG (233 images), 3) White noise (174 images), 4) Gaussian blur (174 images), and 5) Fast-fading Rayleigh channel distortions (174 images). Firstly, the grayscale-distorted image \( I(x,y) \) is randomly sampled with patches \( i(x,y) \) of \( 17 \times 17 \) pixels for \( 5 \times 10^4 \) times to form a \( I \times J \) data table \( X \). Consequently, \( I = 17 \times 17 = 289 \), and \( J = 5 \times 10^4 \). Then perform SVD on \( X \) and get \( \Delta \) and \( Q \) as shown in figure 4. The eigenvalues are \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{289} > 0 \), and the unit eigenvectors are \( Q = [u_1, u_2, \ldots, u_{289}] \) , where \( u_i = (u_{1i}, u_{2i}, \ldots, u_{289i})^T \), \( i = 1, 2, \ldots, 289 \). Then, each \( u_i \) is resized to \( 17 \times 17 \) patch again and performed with 2D-DCT transformation, finally we get the distorted image’s eigenvectors DCT transformation matrix of \( D = [d_1, d_2, \ldots, d_i, \ldots, d_{289}] \).

3.2 Extract the static features

To adequately describe the statistic properties of the distorted image’s eigenvectors DCT transformation matrix \( D \), the features of distorted image \( I(x,y) \) can be extracted as follows:

\[
f_t = [\Delta, \text{var}(D), \text{ske}(D), \text{kur}(D)]
\]  

(2)

In equation (2), \( \Delta = (\lambda_1, \lambda_2, \ldots, \lambda_t) \), \( 1 < t < 289 \) is the square root of eigenvalues of \( X^T X \); \( \text{var}(D), \text{ske}(D) \) and \( \text{kur}(D) \) are the variance, skewness and kurtosis of the each vectors of \( D \), which are defined as:

\[
\text{var}(D) = [\text{var}(d_1), \text{var}(d_2), \ldots, \text{var}(d_i), \ldots, \text{var}(d_t)]
\]

\[
= [E(d_1 - E(d_1))^2, \ldots, E(d_i - E(d_i))^2, \ldots, E(d_t - E(d_t))^2]
\]  

(3)

Equation (3) explains the variance of \( D \) is composed of the 289 2D variance of \( d_i \), which is the DCT on corresponding eigenvector \( u_i \) as mentioned previously, where \( i = 1, 2, \ldots, t, 1 < t < 289 \). Actually \( \text{var}(D) \) is a \( t \)-element vector of the second moment DCT patch \( d_i \). Similarly, \( \text{ske}(D) \) and \( \text{kur}(D) \) are the skewness and kurtosis to DCT patch \( d_i \), which are defined as

\[
\text{ske}(D) = \left[ \frac{E(d_1 - E(d_1))^3}{\sqrt{\text{var}(d_1)}}, \ldots, \frac{E(d_i - E(d_i))^3}{\sqrt{\text{var}(d_i)}}, \ldots, \frac{E(d_{289} - E(d_{289}))^3}{\sqrt{\text{var}(d_{289})}} \right]
\]  

(4)

\[
\text{kur}(D) = \left[ \frac{E(d_1 - E(d_1))^4}{\sqrt{\text{var}(d_1)}}, \ldots, \frac{E(d_i - E(d_i))^4}{\sqrt{\text{var}(d_i)}}, \ldots, \frac{E(d_{289} - E(d_{289}))^4}{\sqrt{\text{var}(d_{289})}} \right]
\]  

(5)
3.3 Training SVR to get the objective image evaluation

Having the feature vector \( f_I \) of each distort image \( I(x,y) \), support vector regression (SVR) is utilized produce a quality index. The \( \nu \)-SVR is applied to perform the regression to learn the mapping from the feature space to subjective quality. The radial basis kernel function (RBF) \( k(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right), \gamma > 0 \) was chose to construct the SPCA. There are three parameters, \((\xi, \gamma, \nu)\), are to be determined, in which \( \xi \) stands for the cost of the slack parameter to control the outlier, \( \gamma \) is the variance of RBF to control the data mapping, and \( \nu \) is to control the number of support vectors. The grid-search optimization with 10-folder cross-validation is used to get the best \((\xi, \gamma, \nu)\) of \( \nu \)-SVR, which performed effectively and reliably.

4. RESULTS AND DISCUSSIONS

We randomly select 23 original images’ associated distorted versions from LIVE IQA database for training, and 6 (different) original images’ associated distorted versions for testing. Such a train-test combination is conducted 1000 times with random permutations of the LIVE database, in order to ensure that the algorithm is robust across all contents and distortion severities. Overall speaking, the median value of Spearman’s rank ordered correlation coefficient (SROCC) between the SPCA and the differential mean opinion score (DMOS) of the subjective quality was 89.7216 mean= 89.5311\%, std.dev=3.178\%). Table I shows the median value of SROCC over 1000 trials for SPCA. To show more clearly, the SROCC and LCC result are putting together according to the distortion types in the left of figure 5 and the right part is the perform between PSNR and SPCA on RMES. Obviously, our SPCA performed very high on LIVE database.

Generally speaking, the properties of the nature scene in our visual world greatly influence the design of human visual system (HVS). A major challenge for the early stages of the HVS is handling the huge data load it receives, while its information capacity is limited. Because such capacity highly depends on the characteristics of the input, it is important to note that natural images are not random, but show a large degree of structure. Such property has been featured on IQA algorithm by different models such as wavelet or DCT decomposition, Gabor-like filter, divisive normalization transformation (DNT). Our recent research proves that image primitive features produced in the HSV earliest processing stage can be simulated by the functions of retinal ganglion cells and lateral geniculate nucleus (LGN) cells. So we proposed a full reference IQA called Non-Shift Edge Ratio (NSER) to detect the zero-crossings by Laplacian of Gaussian (LOG) filter to reveal the primitive structure of image, which has achieved a remarkable performance. In this paper, we firstly shed light on the PCA on naturel and distorted image to figure out a NR-IQA. The classical approach has shown a disappointed result on PCA features which are no interesting as neurophysiological models. By contrast, our research in this paper are thinking about the PCs i.e. the eigenvectors’ Fourier analysis of sampling data matrix’s, which can be intuitively understood as to simulate the formation of receptive fields in the mammalian early visual system. Of
course, the primary research in the paper still leaves much room for improvement, such as the choosing of statics features from the DCT of distorted image's principal components and the optimization of the SVR training process. In fact, SPCA presented here is just a scratch on the surface of possible application about PCA statistical properties, and we strongly believe that this field is worth further extensively exploring.

5. CONCLUSION

In this paper, we proposed a no-reference IQA algorithm based on the statistic property of principal component analysis on nature image, which does not assume any specific type of distortion of the image. The experimental results show a high correlation with human perception of quality (averagely over 90% by scores of SROCC), which is fairly competitive with the existing no-reference IQA metrics.

6. ACKNOWLEDGEMENT

The work of this paper is partially supported by the National Natural Science Foundation of China (No. 90920003 and No. 60472004), which is gratefully acknowledged.

REFERENCES
