Sensors & Transducers, Vol. 183, Issue 12, December 2014, pp. 240-244

Sensors & Transducers

Image Quality Assessment via Quality-aware Group Sparse Coding

[1] Minglei Tong, [2] Hong Han, [3] Shudong Chen
[1] School of Electronics and Information, Shanghai University of Electric Power, Shanghai, China
[2] Xi Dian University, Xi'an, 710126, Shaanxi, China
[3] Institute of Microelectronics of Chinese Academy of Sciences, Wuxi, China
E-mail: tongminglei@gmail.com, hanh@mail.xidian.edu.cn, chenshudong@ciotc.org

Received: 14 August 2014 /Accepted: 27 October 2014 /Published: 31 December 2014

Abstract: Image quality assessment has been attracting growing attention at an accelerated pace over the past decade, in the fields of image processing, vision and machine learning. In particular, general purpose blind image quality assessment is technically challenging and lots of state-of-the-art approaches have been developed to solve this problem, most under the supervised learning framework where the human scored samples are needed for training a regression model. In this paper, we propose an unsupervised learning approach that work without the human label. In the off-line stage, our method trains a dictionary covering different levels of image quality patch atoms across the training samples without knowing the human score, where each atom is associated with a quality score induced from the reference image; at the on-line stage, given each image patch, our method performs group sparse coding to encode the sample, such that the sample quality can be estimated from the few labeled atoms whose encoding coefficients are nonzero. Experimental results on the public dataset show the promising performance of our approach and future research direction is also discussed.

Keywords: Image quality assessment, Group sparse coding, Regression model, Supervised learning, Sparse dictionary.

1. Introduction

With the massive generating of the digital images from the ubiquitous use of digital imaging devices, various levels of quality images are becoming available. These images are either corrupted by imaging noise, motion blur, or the compression from storage and transmission, such as the JPEG, JPEG2000 etc. And it is exciting to observe the continuous technical progress [1, 2, 4-13] and the publicly released benchmark datasets [3, 6, 9] over the past decade in this area which motivates the work of this paper.

In the literature, a considerable number of IQA algorithms have been proposed, which exhibit substantial diversity in the methodologies and contexts being used. From the purpose point of view, the existing methods can be classified into two categories [12]: application-specific methods and general-purpose methods. The former one considers the characteristics of a known and fixed type of distortion e.g., JPEG compression, and achieves satisfactory results by quantifying the particular distortion. Saad et al [7, 8] first determine the distortion type of a test image, and then employ an associated distortion-specific quality metric to predict the quality of the given image. And the images features can be extracted which is fed into the regression model to estimate the subjective quality rate.

In this paper, we present a two-step method for general-purpose image quality assessment, with the full access to the reference image while without knowing the human score. The key novelty lies in the design of group-sparse coding mechanism to encode the image quality using a pre-trained dictionary covering different levels of distortions. More specifically, we trained a quality-aware dictionary using distorted image and the perfect reference one which accounts for different levels of distortion, and then weight differential mean assignment scores by the sparse coding coefficients to obtain the final visual quality values.

The proposed method is thoroughly validated on the CSIQ database [3]. The experimental results demonstrate that our method can produce promising prediction consistent with the subjective quality, and achieved competitive results with the state-of-the-art algorithm [12].

The rest of the paper is organized as follows. Related works are discussed in Section 1. Quality-aware dictionary training is introduced in Section 2. Section 3 details the proposed group-sparse coding methodology and quality prediction system. And experimental results are given in Section 4. Section 5 concludes this paper.

2. Quality-aware Dictionary Learning

2.1. Data Preparation

Our method starts with directly estimating the quality on image patches for blind image quality assessment. And the whole-image quality is generated by pooling the patch-wise qualities. To this end, a set of perfect reference images and the associated distorted images are required. In this paper, four types of distortions are included: Gaussian noise, Gaussian blur, JPEG compression and JPEG2000 compression. For each image, we generate its distorted versions of each type on 5 quality levels by controlling the noise standard deviation (for distortion of Gaussian noise), the support of blur kernel (for distortion of Gaussian blur), the resulted quality level (for distortion of JPEG compression) and the compression ratio (for distortion of JPEG2000 compression), respectively. Finally, we obtain a dataset of 120 distorted images and 10 reference images.

2.2. Patch Quality Generation

Let r_i denote the patch of the reference image, and d_i the corresponding distorted version. We are motivated to calculate the similarity between r_i and d_i such that the perceptual quality can be estimated. We chose the FSIM metric [14]:

$$q_i = \frac{2PC(r_i)PC(d_i)+t_1}{PC(r_i)^2+PC(d_i)^2+t_1} \times \frac{2G(r_i)G(d_i)+t_1}{G(r_i)^2+G(d_i)^2+t_1}, \quad (1)$$

where PC is the phase congruency [12] and G is the gradient magnitude at the center of r_i, respectively, and t_1 and t_2 are the positive constants for numerical stability.

We also perform the normalization step as the post-processing on q_i as in [12] based on the observation that the mean values of the lowest 10 % predicted quality scores shows much better linearity to the human subjective scores. In particular, we divide s_i by a constant C such that the average quality of all patches in an image will equal to the percentile pooling result:

$$C_i = \frac{q_i}{C},\ C = \frac{\sum_{i\in\omega} q_i}{10\sum_{i\in\omega_p} q_i}, \quad (2)$$

where ω denotes the set of patch indices of an image, and ω_p the set of indices of the 10 % lowest quality patches. C serves as a factor of the normalization to reduce the bias. By doing so, we obtain the quality score at the patch level C_i between [0,1], which is associated with each distorted image patch d_i.

2.3. Quality-aware Dictionary Learning

After obtaining the patch quality measurement C_i from each image, we uniformly quantize the patch into L levels. And then clustering 1 is performed to each quantized group whose quality is set from 0 to 1. It is called such procedure as quality aware clustering in [12].

To guide the clustering procedure, we use a high pass filter to extract the structural feature of d_i:

$$h_\sigma(s) = 1_{s=0} - \frac{1}{\sqrt{2\pi\sigma}}\exp(-\frac{s^2}{2\sigma^2}), \quad (3)$$

where s is the radius of the filtering window. By convoluting h over the images, the detailed structures will be enhanced. In particular, three scales (σ = 0.5, 2.0, 4.0) are chosen to be concatenated to form the feature instance f_i for clustering and Euclidean distance is used to measure the distance between two feature instances. In this paper, we set quality group level up to L = 10 and cluster centroid number as K = 30 in each quality group. Fig. 1 illustrates three clusters of patches on the worst quality level and the best quality level respectively. The summary of the offline clustering process is described in Algorithm 1.

Algorithm 1. Quality-aware dictionary training using un-scored training samples.

Input: A set of reference images and the distorted versions at different levels.
Output: A quality-aware dictionary.
Step 1: Calculate the similarity between the distorted image patch and the perfect version, which is regarded as the quality approximation.
Step 2: Divide the distorted samples into multiple levels and perform clustering within each quality level to construct a quality-aware dictionary.

3. Quality-aware Group-sparse Coding

3.1. Sparse Coding

Given the trained quality-aware dictionary consisting of various quality levels atoms, we are motivated to perform sparse coding to discover the most relevant atoms with the test image patch. Specifically, we transform the image patch into feature space as a linear combination of sparse coding basis functions i.e. the dictionary atoms. Sparse modelling is one of the most successful recent signal processing paradigms. The basic idea of sparse coding is simple: a vectorized image patch x can be represented in terms of a linear superposition of basis c_i:

$$x = \sum_{i=1}^{n} r_i c_i = C * r \,, \tag{4}$$

where C denotes the quality aware basis dictionary. In [5], this model is introduced as a possible explanation of the emergence of orientation selective cells in the primary visual cortexV1; the matrix representing C corresponds to neural connections.

3.2. Group Sparse Coding

Since the given image patch is assumed at a certain quality level and shall be mainly correlated with a few clusters aggregating at a certain quality range, rather than evenly (and sparsely) correlated with all quality levels. This observation motivates us to enforce the encoding coefficients have a grouping of its components, and the components within a group are likely to be either all zeros or all nonzeros. A favorable approach in the literature is to use the mixed $l_{2,1}$-regularization. In general the un-weighted unconstrained form is:

$$\|r\|_{2,1} + \frac{1}{2\mu}\|C * R - x\| \,, \tag{5}$$

Not the $l_{2,1}$ Norm $\|r\|_{2,1}$ is defined as:

$$\|r\|_{2,1} = \sum_{i-1}^{s} \|r_{g_i}\|_2 \,, \tag{6}$$

where g_i is the index set corresponding to the i-th group for a subset of the coefficients of r. In our case, the subset is non-overlapped in the sense that quality level has no overlap regarding clusters. Given such formulation, we perform the new alternating direction method (ADM) algorithm to obtain the group-sparsity coefficient vector r. Compared with the simple nearest neighboring method, and sparse coding mechanism, the group-sparsity model is believed to capture the quality pattern more robustly and compactly. The algorithm of our approach is detailed in Algorithm 2.

Algorithm 2. Quality-aware group sparse coding for test image quality assessment.
Input: the test image and the learned quality-aware dictionary.
Output: The quality assessment rate with the test image.
Step 1: Perform group sparse coding to obtain the nonzero entries associated with the dictionary atoms at different quality levels.
Step 2: Obtain the patch quality score by weighting over the nonzero entries and the atoms.
Step 3: Perform weighted pooling to obtain the final whole-image quality rate from patch-wise quality predictions.

3.3. From Patch Quality to Image Quality by Weighted Pooling

There exist many pooling strategies such as max pooling and pooling of percentile, which is also widely applied in other image research topic such as deep convolution networks. We adopt a very simple weighting strategy to obtain the final image-wise quality prediction from the patch-wise quality predictions:

$$q^{patch} = \frac{\sum_{l=1}^{10} 0.1 * l \sum_{i=1}^{30} |r_{l_i}|}{\sum_{k=1}^{300} |r_k|} \,, \tag{7}$$

$$q^{image} = \frac{\sum_{k=1}^{m} q_k^{patch}}{m} \,, \tag{8}$$

where l denotes the quality level from 1 to 10, and g_i denotes the within quality group of the size 30. The first equation calculate the patch-wise quality as weighed by the encoding coefficients associated with the L = 10 × K = 30 atoms of the given dictionary. And the second equation calculates the image-wise quality by averaging all patches of the image.

4 Experimental Results

4.1 Experimental Protocol

The performance of the proposed approach is evaluated regarding its prediction capability for subjective ratings of image quality. A subjective quality/distortion score, i.e., the mean opinion score (MOS) or difference mean opinion score (DMOS) [10], is assigned to validate the algorithms. The database used in our experiment is from CSIQ [3]. The CSIQ database is composed of 30 original images and their distorted counterparts by using six types of distortions on five different distortion levels. To evaluate the performance, two correlation coefficients between the prediction results and the subjective scores are adopted: the Spearman rank order correlation coefficient (SROCC), which is related to the prediction monotonic, and the Pearson correlation coefficient (PCC), which is related to the prediction linearity. A good BIQA method will demonstrate a big (close to 1) correlation coefficient with the subjective score MOS or a small (close to -1) correlation coefficient with DMOS. Specifically, in this paper, we adopt the DMOS score as provided by the used database.

For illustration, Fig. 1 plots the scatter distribution for the prediction quality score and the DMOS values on CSIQ database.

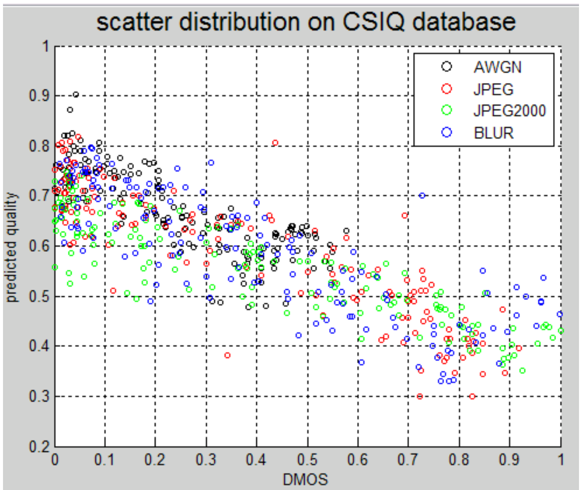


Fig. 2. DMOS vs. Predicted quality scatter plot for JPEG, JPEG2000, AWGN and Gaussian Blur. A linear correlation is presented.

4.2. Performance Evaluation

In the implementation, we partition the 130 training images into overlapped patches of size 8×8. In total, 161,181 patches are extracted for training. In feature extraction, we set the three scales of high pass filters as σ = 0.5; 2.0; 4.0. For clustering, we quantize the quality into L = 10 levels; that is, q_l is from 0.1 to 1 with step length 0.1. On each quality level, K = 30 clusters are clustered by using the clustering algorithm. Examples of the atoms consisting of the quality-aware dictionary are illustrated in Fig. 2.

We evaluate our method with the state-of-the-art [12] whose parameters are set strictly according to the original paper. The experimental results in terms of SROCC are reported in Table 1. Note that both two evaluated methods do not need the human scored images for learning. As observed from the table, our method performs competitively with the state-of-the-art [12] on the CSIQ dataset.

Table 1. The SROCC comparison between the proposed method and the QAC method [12] learning from human scored images on CSIQ database[3]. AWGN denotes additive white Gaussian noise, GB denotes Gaussian blur, and JP2K denotes JPEG2000 compression.

Distortion	QAC	Proposed method
JP2K	0.8645	0.8681
JPEG	0.8774	0.8703
AWGN	0.8232	0.8198
GB	0.8187	0.8283

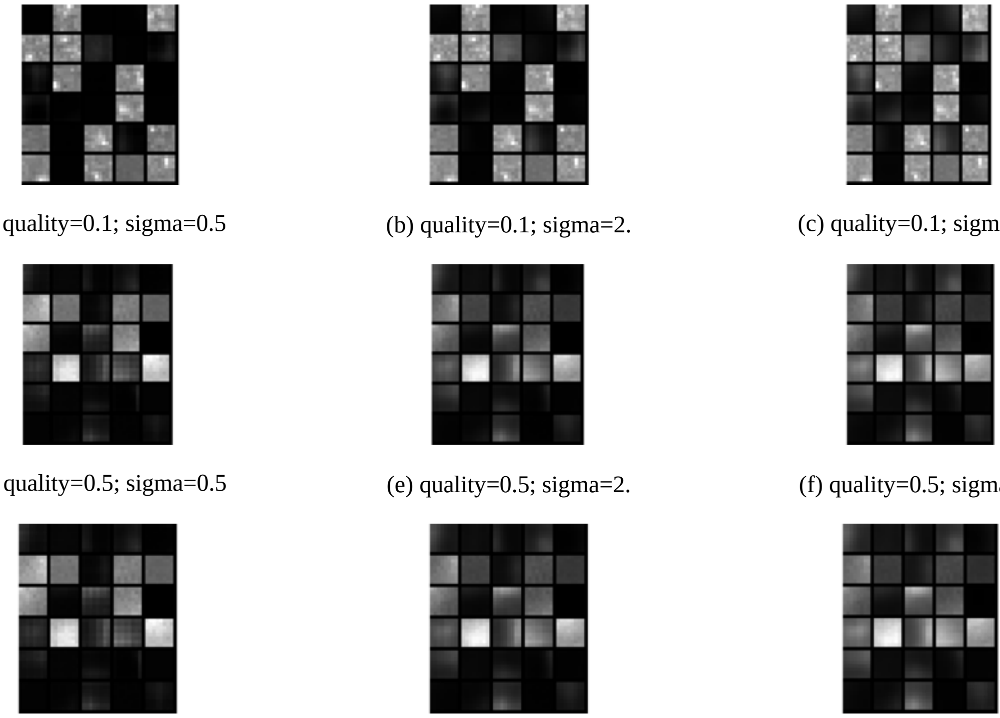
5. Conclusion and Future Work

We propose a new method by combining the layered dictionary with the sparse coding mechanism to predict the subjective image quality metrics. The key novelty lies in the design of sparse coding mechanism to encode the image quality using a pre-trained dictionary covering different levels of distortions. Overall, the proposed approach provides a fast and robust performance for image quality assessment. One interesting future research point is how to generate the final quality measurement of the test image given the estimated quality of each image patch.

Acknowledgements

This research is supported by the supported by the National Nature Science Foundation of China (No. 61105016). The Innovation Program of Shanghai Municipal Education Commission (Grant No. 12YZ139).

.



(a) quality=0.1; sigma=0.5 (b) quality=0.1; sigma=2. (c) quality=0.1; sigma=4.

(d) quality=0.5; sigma=0.5 (e) quality=0.5; sigma=2. (f) quality=0.5; sigma=4.

(g) quality=1; sigma=0.5 (h) quality=1; sigma=2. (i) quality=1; sigma=4.

Fig. 2. Trained dictionary across 10 quality levels: level 0.1, 0.5 and 1.0 are illustrated at three filtering scales: 0.5, 2, 4 respectively.

References

[1]. J. Caviedes and S. Gurbuz, No-reference sharpness metric based on local edge kurtosis, in *Proceedings of the International Conference on Image Processing (ICIP'02)*, 2002, pp. 53-56.

[2]. P. Kovesi, Image features from phase congruency, *Journal of Computer Vision Research*, Vol. 1, Issue 3, 1999, pp. 1-26.

[3]. E. Larson and D. Chandler, Most apparent distortion: full reference image quality assessment and the role of strategy, *Journal of Electronic Imaging*, Vol. 19, Issue 1, 2010, Article ID 011006.

[4]. A. Moorthy and A. Bovik, Blind image quality assessment: From natural scene statistics to perceptual quality, *IEEE Transaction on Image Processing*, Vol. 20, Issue 12, 2011, pp. 3350-3364.

[5]. B. Olshausen and D. Field, Emergence of simple-cell receptive field properties by learning as parse code for natural images, *Nature*, Vol. 381, No. 6583, 1996, pp. 607-609.

[6]. N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, Tid2008-a database for evaluation of full-reference visual quality assessment metrics, *Advances of Modern Radioelectronics*, Vol. 10, Issue 10, 2009, pp. 30-45.

[7]. M. Saad, A. Bovik, and C. Charrier, A dct statistics-based blind image quality index, *IEEE Signal Processing Letters*, Vol. 17, Issue 6, 2010, pp. 583-586.

[8]. M. Saad, A. Bovik, and C. Charrier, Model-based blind image quality assessment using natural dct statistics, *IEEE Transaction on Image Processing*, Vol. 21, Issue 8, 2012, pp. 3339-3352.

[9]. H. Sheikh, Z. Wang, L. Cormack, and A. Bovik, Live image quality assessment database release 2, 2005, (http://vision.okstate.edu/?loc=csiq).

[10]. E. A. Silva, K. Panetta, and S. S. Agaian, Quantifying image similarity using measure of enhancement by entropy, in *Proceedings of the International Conference on Image Processing (ICIP'07)*, 2007, pp. 169-172.

[11]. H. Tang, N. Joshi, and A. Kapoor, Learning a blind measure of perceptual image quality, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR'11)*, 2011, pp. 305-312.

[12]. W. Xue, L. Zhang, and X. Mou, Learning without human scores for blind image quality assessment, *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'13)*, 2013, pp. 995-1002.

[13]. P. Ye, J. Kumar, L. Kang, and D. Doermann, Unsupervised feature learning framework for no-reference image quality assessment, Learning without human scores for blind image quality assessment, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR'12)*, 2012, pp. 1098-1105.

[14]. L. Zhang, L. Zhang, X. Mou, and D. Zhang, Fsim: a feature similarity index for image quality assessment, *IEEE Transactions on Image Processing*, Vol. 20, Issue 8, 2011, pp. 2378-2386.