Visual Information Measurement with Quality Assessment*

Jinjian Wu, Guangming Shi, Man Zhang, and Guanmi Chen
School of Electronic Engineering, Xidian University, Xi'an, Shannxi, China

***Abstract*—The quantity of visual information measurement is significant for many perception-oriented signal processing system. The classical Shannon theory, which based on the probability of signal, is useful to measure the quantity of channel information. However, it fails to accurately measure the quantity of visual information of a given image. Image quality refers to the subjective perception on the visual information that an image carried. An image with high quality carries more information than that of a low quality image. Thus, the quality of an image can effectively represent its quantity of visual information. In this paper, we propose a novel visual information measurement, and verify it with quality assessment. Firstly, a dictionary is learned from natural images, in which the content change of each atom is calculated to present its quantity of visual information. Then, a testing image is represented by the dictionary, and the sparse coefficients for each local block in the image are acquired. Finally, according to the sparse coefficients, the quantity of visual information is measured. The information measurement result is verified with the subjective quality score. Experimental results on a large amount of images demonstrate the accuracy of the proposed method for quantity of visual information measurement.**

***Index Terms*—Visual Information Measurement, Quality Assessment, Sparse Representation, Dictionary**

I. INTRODUCTION

Visual information measurement is meaningful in may perception-oriented visual signal processing system [1], [2], such as visual saliency estimation, image segmentation, image compression, image reconstruction, and so on. However, limited by the understanding and knowledge on the visual perception of the human brain, how to accurately measure visual information (which is consistent with subjective perception) is still an open problem.

In the past half century, the Shannon information theory was proposed and became the standard for information measurement [3]. Generally speaking, if one event occurs often, the event provides less information; on the contrary, a rare event provides much more information. According to the probability distribution of events, the Shannon entropy is proposed to measure the quantity of information. The Shannon entropy is successful in channel coding. However, Shannon entropy only considers the probability of the events, which mainly represents the probability distribution and cannot effectively represent the visual information of the events themselves. To this end, the weighted entropy is proposed to improve Shannon entropy [4]. The weighted entropy takes both the probabilities and weights (importances) of events into account for quantity of information measurement. However, it is difficult to automatically weight the visual importances of events for information measurement.

Image quality refers to the comprehensive perception on the visual content of a given image [1]. Generally speaking, distortion will degrade the visual content of an image, namely, distortion will decrease the quantity of visual information of an image. Thus, for a given scene, a high quality means it provides much information; and a low quality means its visual content is degraded and it provides less information. According to the analysis above, we try to measure the visual information via quality assessment in this paper.

In the past decade, a large amount of image quality assessment (IQA) algorithms have been proposed. These existing IQA algorithms are usually classified into three types according to the availability of the reference data: Full-Reference (FR) [5], [6], which need the whole reference image; Reduced-Reference (RR) [7], [8], [9], for which only a limited amount of reference data is required; and No-Reference (NR) [10], [11], [12], which measures the quality with only the distorted image. Since the FR and RR IQA algorithms need information from the reference, and there are always no reference data available for visual information measurement, we focus on the NR IQA algorithm for visual information measurement. Moreover, the existing NR IQA algorithms can only qualitatively measure whether the visual information is sufficient for scene perception. Therefore, a novel visual information measurement is required, which can quantitatively measure the visual information.

In this paper, a sparse representation based visual pattern index is proposed to measure visual information, and is further verified with visual quality. Though different images are composed with colorful and varied visual contents, they share the same basic atoms (i.e., basic visual patterns) [13]. Therefore, we firstly extract blocks from natural images for basic atoms clustering, and an overcomplete dictionary composed with these basic atoms is learned. Then, for a given image, each block is sparse represented with the overcomplete dictionary, and the visual information of each block is computed according to the sparse coefficients. Finally, by pooling the visual information of all blocks, the quantity of visual information of the image is acquired. In order to demonstrate the effectiveness of the proposed method, we verify the measurement results with the subjective quality assessment result (i.e., the Mean

*This work was supported by the National Natural Science Foundation of China (No. 61401325), the Ph.D. Programs Foundation of Ministry of Education of China (20130203130001), and the Young Talent fund of University Association for Science and Technology in Shaanxi (No. 20150110).

Opinion Score, MOS). By removing the energy (content change) of an image from its quantity of visual information, its qualitative information content (i.e., image quality) is calculated and verified with MOS value. Experimental results on three large IQA databases demonstrate the effectiveness and accuracy of the proposed measurement.

The result of this paper is organized as follows. The novel visual information measurement is introduced in Section II. In Section III, the effectiveness of the proposed method is verified on a large amount of images. Finally, conclusion is drawn in Section IV.

II. Quality Assessment and Visual Information Measurement

It is well known that images are composed by some basic visual patterns (basic atoms), and the primary visual cortex is highly adapted to extract these patterns for processing [14]. With these basic visual patterns, a dictionary which is effective to represent visual content can be created. But how to find out these basic visual patterns is still an open problem. In this work, we extract blocks from natural images for clustering, and the clustering centers are considered as the basic visual patterns of images. For a large set of patches $\{\mathcal{P}_i\}_{i=1}^{N}$, their clustering centers can be calculated as,

$$\mathcal{C} = \arg\min_{\mathcal{C},x} ||\mathcal{P} - \mathcal{C}x||_2^2 + \lambda||x||_1, \;\; s.t. \; ||\mathcal{C}_j||_2^2 \leq 1 \tag{1}$$

Where $\mathcal{C}$ is an overcomplete dictionary, which is composed by the clustering centers $\mathcal{C}_j|j = 1, 2 \cdots M$; x is the sparse coefficient vector; $||\cdot||_2^2$ is the l_2 norm; λ is a fixed parameter which aims to balance the sparsity of the solution and the fidelity; and $||\cdot||_1$ is the l_1 norm.

With (1), the dictionary $\mathcal{C}$ can be learned with an iterative update procedure, for which $\mathcal{C}$ and x are computed as follows,

1) The initial $\mathcal{C}$ is defined as a Gaussian random matrix, where each column is normalized as a unit vector.
2) With the current $\mathcal{C}$, updating x according to the following formulation,

$$x = \arg\min_{x} ||\mathcal{P} - \mathcal{C}x||_2^2 + \lambda||x||_1 \tag{2}$$

3) With the updated x, updating $\mathcal{C}$ according to the following formulation,

$$\mathcal{C} = \arg\min_{\mathcal{C}} ||\mathcal{P} - \mathcal{C}x||_2^2 \;\; s.t. \; ||\mathcal{C}_j||_2^2 \leq 1 \tag{3}$$

4) Return to step 2 until convergency is achieved.

In order to create the dictionary, 9 nature images from the LIVE database are chosen for basic visual pattern clustering, as shown in Fig 1. Image patches with size of 8×8 are extracted from these images. With the clustering procedure introduced above, an overcomplete dictionary which composed by basic visual patterns is acquired, as shown in Fig. 2. As can be seen, these basic visual patterns are quite similar to the local receptive fields in the primary visual cortex [14], which represent the high-level features of images.

Fig. 1: Nature images for basic visual pattern clustering.

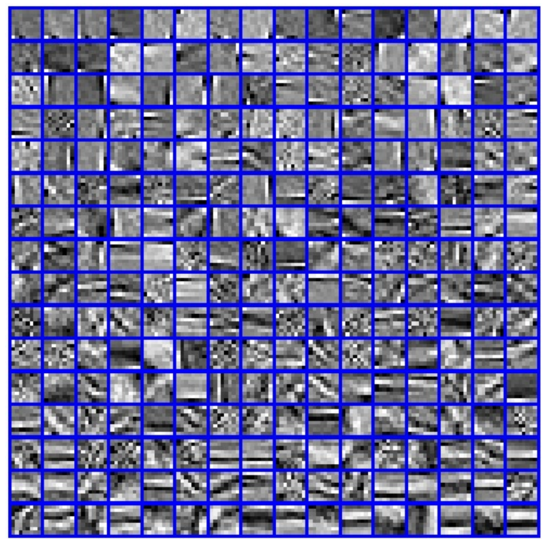


Fig. 2: Overcomplete dictionary learned from nature images.

With the overcomplete dictionary $\mathcal{C}$, each patch (p) from a test image can be sparse represented,

$$p = \mathcal{C}x = \sum_{i=1}^{M} x_i\mathcal{C}_i, \;\; s.t. \; ||p - \mathcal{C}x||_2^2 \leq \epsilon \tag{4}$$

where x is the sparse coefficients vector, and ϵ is a small constant.

Generally speaking, an image patch with no or limited distortion corresponds to one basic visual pattern in the dictionary $\mathcal{C}$ and has little correlation with other basic visual patterns. From the perspective of sparse coefficients, the sparse coefficient vector from (4) for an image patch with no or limited distortion has only one large value, while other coefficients approach to 0. Therefore, for an image patch with no or limited distortion, its energy is concentrated when projecting on to the dictionary $\mathcal{C}$. On the contrary, if an image patch is with serious distortion, it correspond to several basic visual patterns rather than only one. As a result, its energy is dispersed. According to the analysis above, we can conclude that the quality of each patch is highly correlated to the degree of energy distribution.

The energy ($\mathcal{E}$) of the sparse coefficients of an image patch are calculated as,

$$\mathcal{E} = \sum_{i=1}^{M} ||x_i||_2^2. \tag{5}$$

In this paper, we suggest the energy of each patch as its quantity of visual information. By pooling the energies of all patches, the quantity of visual information $\mathcal{V}_I$ of an image is acquired,

$$\mathcal{V}_I = \frac{1}{N}\sum_{j=1}^{N} \mathcal{E}_j, \tag{6}$$

where N is the patch number of the test image.

In order to verify the accuracy of the proposed visual information metric, the quality assessment is employed. Different image patches possess different structural information, which present different content change values. Thus, we firstly normalize the energy of each patch with its content change value,

$$q = \frac{\mathcal{E}}{\sigma}, \tag{7}$$

where σ is the content change value (i.e., variance) of patch p,

$$\sigma = \frac{1}{P}\sum_{i=1}^{P} (\mathcal{I}_i - \bar{\mathcal{I}})^2, \tag{8}$$

where P is the pixel number in a patch, $\mathcal{I}_i$ is the pixel value, and $\bar{\mathcal{I}}$ is the mean pixel value of an 8×8 image patch.

Then, the quality of the whole image ($\mathcal{Q}$) is calculated as the sum of all patches,

$$\mathcal{Q} = \frac{1}{N}\sum_{j=1}^{N} q_j. \tag{9}$$

III. Experimental Result

In this section, the performance of the proposed visual information measurement is firstly illustrated. Then, the proposed visual information measurement is further verified with subjective quality assessment scores.

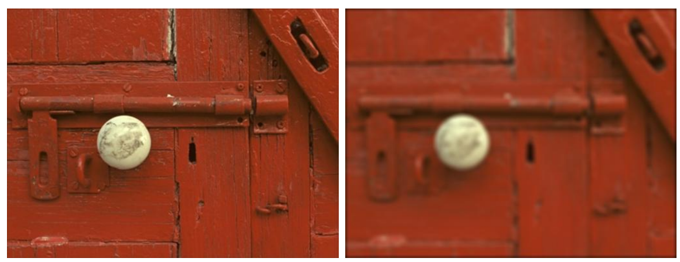
As shown in Fig. 3, the *Hats* image is distorted by Gaussian Blur noise. With the increase of noise level from (a) to (d), much more detailed visual information is erased. In other words, the quantities of visual information for the four images are decreased. The computed quantities of visual information $\mathcal{V}_I$ for the four distorted images are 2.43, 1.80, 1.00, and 0.39, respectively (which are also gradually decreased). Thus, the proposed information measurement can accurately measure the quantities of information for these images.

Meanwhile, the subjective visual quality (represented by the MOS value) is adopted to further demonstrate the effectiveness of the proposed information measurement. As shown in Fig. 3 (a)-(d), the MOS values for the four distorted images are 5.79, 5.24, 4.00, and 3.07, respectively; and the normalized

(a) M=5.79, $\mathcal{V}_I$=2.43, $\mathcal{Q}$=15.93 (b) M=5.24, $\mathcal{V}_I$=1.80, $\mathcal{Q}$=11.91

(c) M=4.00, $\mathcal{V}_I$=1.00, $\mathcal{Q}$=6.72 (d) M=3.07, $\mathcal{V}_I$=0.39, $\mathcal{Q}$=2.70

Fig. 3: The *Hats* image is distorted by Gaussian Blur noise, and the noise levels from (a)-(d) are gradually increased.



(a) M=5.69, $\mathcal{V}_I$=1.37, $\mathcal{Q}$=14.23 (b) M=3.13, $\mathcal{V}_I$=0.23, $\mathcal{Q}$=2.63

Fig. 4: The *Door Lock* image distorted by Gaussian Blur noise.

quantities of information (i.e., the computed quality score $\mathcal{Q}$) for them are 15.93, 11.91, 6.72, and 2.70. respectively. Both MOS and $\mathcal{Q}$ values are gradually decreased, which means $\mathcal{Q}$ performs consistently with MOS on this image set.

Moreover, another set of images are chosen to illustrate the performance of the proposed information measurement on different images. Fig. 4 shown a *Door Lock* image distorted by Gaussian Blur noise. Under the same noise level, Fig. 3 (a) and Fig. 4 (a) have similar visual quality (with MOS 5.79 VS. 5.69). However, the visual contents of the two images are quite different and their quantities of visual information are also different. For example, Fig. 3 (a) possesses much more visual contents than Fig. 4 (a), and the computed $\mathcal{V}_I$ can accurately represent this (with $\mathcal{V}_I$ 2.43 VS. 1.37 for the two images). With the increasing of noise level, the quantities of visual information for the two images (Fig. 3 (d) and Fig. 4 (b)) are decreased, and Fig. 3 (d) ($\mathcal{V}_I$=0.39) still has much more information than that of Fig. 4 (b) ($\mathcal{V}_I$=0.23). Therefore, the proposed method can effectively measure the quantities of

TABLE I: PERFORMANCE COMPARISON

DB	Cri. \ Alg.	JNB [15]	LPC [16]	MLV [17]	BRISQUE [18]	NIQE [19]	Proposed
TID	PLCC	0.693	0.857	0.858	0.804	0.830	0.880
	SRCC	0.667	0.856	0.855	0.799	0.815	0.875
	RMSE	0.846	0.604	0.602	0.697	0.654	0.558
CSIQ	PLCC	0.806	0.916	0.949	0.927	0.926	0.936
	SRCC	0.762	0.907	0.925	0.903	0.894	0.916
	RMSE	0.170	0.115	0.091	0.107	0.108	0.101
LIVE	PLCC	0.816	0.918	0.9429	–	0.943	0.960
	SRCC	0.787	0.939	0.9312	–	0.933	0.958
	RMSE	10.68	7.322	6.1514	–	6.131	5.192

visual information on different images

In order to further illustrate the effectiveness of the proposed information measurement, the normalized measurement is verified with the subjective visual quality on Gaussian Blur distorted images from 3 big databases, namely, TID [20], CSIQ [21], and LIVE [22]. Five classic and latest IQA methods, JNB [15], LPC [16], MLV [17], BRISQUE [18], and NIQA [19], are adopted for comparison. And their performances on these databases are verified by three often used criteria, namely, PLCC, SRCC, and RMSE [6].

The performances are listed in Tab. I (No BRISQUE performances on LIVE, since it trains on this database). As can been seen, the proposed method performs the best on both TID and LIVE databases. On CSIQ database, the proposed method performs a slight worse than that of MLV, while better than other methods. Therefore, the proposed method performs highly consistently with subjective perception.

According to the analysis above, we can conclude that the proposed information measurement can accurately measure the quantities of visual information.

IV. Conclusion

In this paper, we have designed a visual information measurement. The basic visual patterns were firstly clustered from a large amount of natural image patches, and an overcomplete dictionary was learned. Then, each test patch was projected onto the dictionary, and the energy of each patch was calculated with the sparse coefficient. By pooling all patches from an image, its quantity of visual information was acquired. The proposed method was verify with subjective quality assessment on a large amount of images from three often used IQA databases. Experimental results demonstrate that the proposed measurement performs consistent with subjective perception.

References

[1] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *J. Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, May 2011.

[2] J. Wu, F. Qi, G. Shi, and Y. Lu, "Non-local spatial redundancy reduction for bottom-up saliency estimation," *Journal of Visual Communication and Image Representation*, vol. 23, no. 7, pp. 1158–1166, 2012.

[3] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical J.*, vol. 27, pp. 379–423, Jul. 1948.

[4] S. Guiau, "Weighted entropy," *Reports on Mathematical Physics*, vol. 2, no. 3, pp. 165–179, Sep. 1971.

[5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[6] J. Wu, W. Lin, G. Shi, and A. Liu, "Perceptual quality metric with internal generative mechanism," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 43 –54, Jan. 2013.

[7] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 41–52, Jan. 2012.

[8] R. Soundararajan and A. Bovik, "RRED indices: Reduced reference entropic differencing for image quality assessment," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 517 –526, Feb 2012.

[9] J. Wu, W. Lin, G. Shi, and A. Liu, "Reduced-reference image quality assessment with visual information fidelity," *IEEE Transactions on Multimedia*, vol. 15, no. 7, pp. 1700–1705, Nov. 2013.

[10] P. Ye and D. Doermann, "No-reference image quality assessment using visual codebooks," *IEEE Transactions on Image Processing*, vol. 21, no. 7, pp. 3129 –3138, Jul. 2012.

[11] L. Li, H. Zhu, Y. Gao, and J. Qian, "Referenceless measure of blocking artifacts by tchebichef kernel analysis," *Signal Processing Letters, IEEE*, vol. 21, no. 1, pp. 122–125, 2014.

[12] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, "No-reference image sharpness assessment in autoregressive parameter space," *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 3218 –3231, Oct. 2015.

[13] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1838 –1857, Jul. 2011.

[14] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 13, pp. 607 –609, Jun. 1996.

[15] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb)," *IEEE Transactions on Image Processing*, vol. 18, no. 4, pp. 717 –728, Jul. 2009.

[16] R. Hassen, Z. Wang, and M. Salama, "Image sharpness assessment based on local phase coherence," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2798 –2810, Jul. 2013.

[17] K. Bahrami and A. C. Kot, "A fast approach for no-reference image sharpness assessment based on maximum local variation," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 751 –755, Jun. 2014.

[18] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695 –4708, Dec. 2013.

[19] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209 –201, Mar. 2013.

[20] N. Ponomarenko and K. Egiazarian. Tampere image database 2008 tid2008. [Online]. Available: http://www.ponomarenko.info/tid2008.htm.

[21] E. C. Larson and D. M. Chandler. Categorical image quality (csiq) database. [Online]. Available: http://vision.okstate.edu/csiq.

[22] H. R. Sheikh, K. Seshadrinathan, A. K. Moorthy, Z. Wang, A. C. Bovik, and L. K. Cormack. Image and video quality assessment research at live. [Online]. Available: http://live.ece.utexas.edu/research/quality/.