Comparison of Some Image Quality Approaches

Boris B. Parfenenkov, Maksim A. Panachev

Ural Federal University
{idlerboris,tiopox}@gmail.com

Abstract. This paper is devoted to image quality problem. We analyze advantages and disadvantages of existing methods. Classification of quality metrics into some groups has been done. Based on this classification, we formed proposition about prospects of using this methods in solving image quality problem.

Keywords: image processing, image fidelity, image quality, MSE, SSIM, VDP, anisotropic.

Introduction

The problem of quality assessment arises in many different subjects. From computer graphics, where rendering of complex scenes may had a lot of time. To bioinformatics and computer security, where quality and accuracy of images may safe human lives. Although, count of image editors, which must define image quality and may improve this, significantly increases.

Certainly, our eyes is good classifier of image fidelity, but there are a lot of different software systems, which must define quality of digital images. Consequently, using of human resource is not acceptable for this problem.

In this paper we are describe of existing methods of quality image assessment. Classification of quality metrics into some groups has been done. Although analyze of advantages and disadvantages of the most promising methods was performed. We tested these metrics on the collected set of images and selected metric, which we recommend to use for solving similar tasks.

Image quality assessment

Let there be two digital images: X – original, Y – test (distorted image – with possible defects). The challenge is to build algorithm, which have these two images and define quality assessment of test image.

Digital image may be represent with brightness matrix $I = (p_{i,j})_{H\times W}$, where $p_{i,j} \in [a,b] \cap \mathbb{Z}$, $H \; and \; W$ – height and width of image, respectively. Although, in some cases, we will consider image as one-dimensional signal $X = (x_{i\cdot W+j})_{H\cdot W}$.

Basic metrics

At the beginning, we consider classic metrics that came to computer vision from mathematical statistics. Mean squared error of images X and Y presented as $MSE(X,Y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$. Sometimes researchers considered root mean squared error $RMSE(X,Y) = \sqrt{MSE(X,Y)}$, which may be generalize on l_p: $d_p(X,Y) = (\sum_{i=1}^{N}|x_i - y_i|^p)^{\frac{1}{p}}$.

Peak signal-to-noise ratio is calculated based on MSE and often apply for measure of distortion when image was compressed. $PSNR = 10\log_{10}\frac{L^2}{MSE}$, where $L = (b - a) -$ is the dynamic range of allowable image pixel intensities (e.g., for image that have allocations of 8 b/pixel of gray-scale, $L = 255$).

However, these metrics are not best instruments for quality assessment of images [3], because they ignore features of human image perception.

Structural similarity metrics

In paper [7] was discussed reasons of creating metrics based on structural similarity. The main idea is that human able to extract some structure from image and perceive it, but not separately pixels. Therefore, metric, which can be measure amount and kind of structural information from image, can significantly increases image quality assessment.

The first result in this approach was metric SSIM (Structural SIMilarity), which computing as composition of: illumination ($l(X,Y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$), contrast ($c(X,Y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$) and structural comparison ($s(X,Y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$); where μ_x – expected value of brightness, σ_x – standard deviation, σ_{xy} – covariance of x and y, and C_1, C_2, C_3 – some constants, that obtained experimentally.

$$SSIM(X,Y) = l(X,Y) * c(X,Y) * s(X,Y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}.$$

For improve this metric we can compute weighted mean value of SSIM on local features (local feature is small part of image, which focuses people's attention; it is known, that using local features allow to discard noise and improve quality of metric [6]). Weight indicates the significance of this local feature.

MS-SSIM (Multi-scale SSIM) [8] allows to improve the image quality assessment. This metric used setp by step computation $c(X,Y)$ and $s(X,Y)$ for different resolutions. By using computation MS-SSIM for each local features we can get more impressive results [2].

SCSSIM [1] allows to evaluate the image quality by using correction for the structural features of original and test images.

Visible Differences Predictor

First this metric was developed by Scott Daly in his paper [10]. He analyzed, how to construct Human Visual System model for improving existing methods of image quality assessment.

VDP receives 2 input images and generates output differences map (each pixel has value, which describes how different the pixels of corresponding input images). Schema of work VDP presents on Fig. 1. One of the advantages of this model is possibility to get prediction of local differences between images (on the pixel level), while methods described previously provided a single value for the entire image. Although, the basis of this metric are components that are already recognized in the computer vision: CSF [11], Cortex transform [4, 11], psychometric Weibull function [9]. One of the disadvantages of VDP is non-use information about color, and work only with brightness.

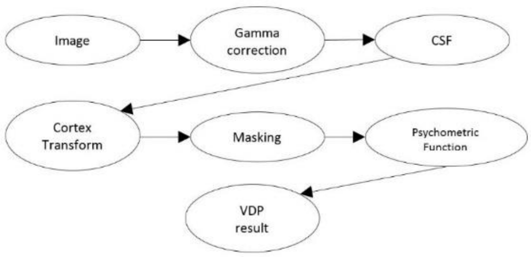

Fig. 1. Schema of Visual Differences Predictor

Anisotropy

Metrics, which described above, based on original image, and quality of test image is defined in relation to it. However, there are a variety of tasks, in which we don't have original image. Such tasks are appear in different research areas related to image quality assessment in real time. For example quality assessment of rendering complex scene in which we don't have template of image, or quality assessment of photograph that made by medical device for analyze reliability of the data, and etc.

Quality and entropy are related subjects, however, noise and information of image cannot be separated from each other (noise also has some information). For example, human with good eyesight can easily distinguish a clear object even when the image is noisy. However, analytically, entropy increases with sharpness but, in general, there is not a fair correlation when images are noisy. Hence entropy by itself is not a good indicator of image quality. And in paper [5] metric based on anisotropy was proposed, which can be represent by following:

$$LMQ(X) = B * log_{10}A(X),$$

where $A(X)$ – anisotropy of image and B is a constant that must be determined to fix the range of operative values (experimentally good results were obtained for $B = 20$).

Analyze image fidelity metrics

For analyze image fidelity metrics we collect set of images, which has been obtained by various deformations of original image. Then these images were sorted by decrease quality in terms of human perception (this sort was done using quality image assessment by three independent people). Result set of images in the sort order present on Fig. 2 (presented reduced copies, real images have a size 256x256 pixels). Names of the images correspond to deformation types: 1) shift($\boldsymbol{n}$) – shift brightness of all pixels by $\boldsymbol{n}$ (i.e. image becomes lighter); 2) noise($\boldsymbol{n}$) – add Gaussian noise (the more n – the more noise); 3) jpeg(n) – image after compression by JPEG for different sizes of block (the less n – the more defects); 4) blur(n) – Gaussian blur (the more n, the more blur).

Fig. 2. Images for testing image fidelity metrics

Then testing of 4 metrics (PSNR, LMQ, variation of SSIM, VDP) was conducted. Each metric returns evaluation image deformation – i.e. result that inverse to quality and takes values from 0 (quality image) to 1 (deformed image). The resulting graph of the image deformation (vertical axis) of the image (horizontal axis) presented on Fig. 3. Highline that main criteria for correctness of metric is not absolute deformation value but result graph has been directed upwards (i.e. for any pair of images, deformation value of left image must be greater than deformation value of right image).

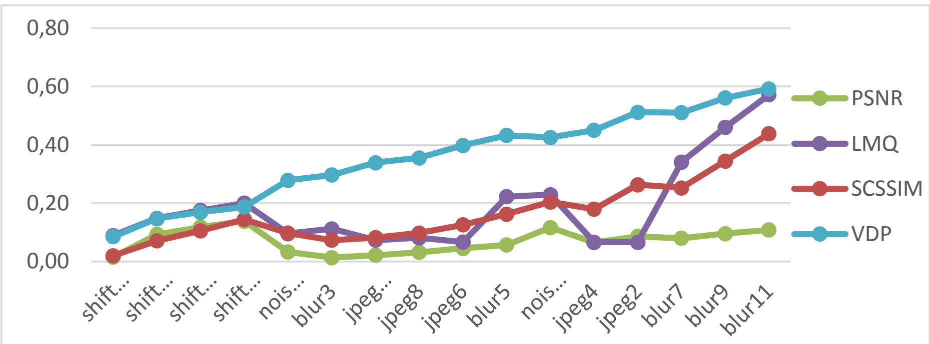

Fig. 3. Graph for image quality assessment by metrics: PSNR, LMQ, SCSSIM, VDP.

As can be seen from the graph, the most correct results were obtained by metrics VDP and SCSSIM. Metric LMQ also gave a good result, but poorly handled with images, which are compressed using JPEG.

Conclusions

In this paper we considered principal image fidelity metrics, from basic that came to computer vision from mathematical statistics, to hybrid models that use modern knowledge of computer vision and information quality assessment. Described metrics split into the following groups: statistics (MSE, $\boldsymbol{l_p}$ norm, PSNR); structural similarity (SSIM, MSSIM, MS-SSIM, IW-SSIM, SCSSIM); human perception (VDP); anisotropy (LMQ). Advantages and disadvantages of these methods were considered.

In the result of this work we determine that the presence of the original image should be use VDP metric. If we don't have original image for quality assessment, then should be use LMQ, which is very costly from a computational point of view.

References

1. K. Gu, G. Zhai, X. Yang, and W. Zhang "An improved fullreference image quality metric based on structure compensation", APSIPA ASC, pp. 1 - 6, 2012
2. Z. Wang and Qiang Li "Information content weighting for perceptual image quality assessment", IEEE Transaction on Image Processing, vol. 20, no. 5, pp. 1185-1198, 2011
3. Z. Wang and A. C. Bovik "Mean squared error: Love it or leave it?-A new look at signal fidelity measures," IEEE Signal Processing Mag., vol. 26, no. 1, pp. 98-117, 2009
4. A. S. Lukin. "Improved Visible Differences Predictor Using a Complex Cortex Transform", 19-th International Conference on Computer Graphics GraphiCon'2009, pp. 145–150, 2009
5. Gabarda, Salvador, and Gabriel Cristóbal. "Image quality assessment through a logarithmic anisotropic measure", Photonics Europe. International Society for Optics and Photonics, pp. 70000J-70000J, 2008
6. Datta, Ritendra, Jia Li, and James Z. Wang. "Content-based image retrieval: approaches and trends of the new age", Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval, pp. 253-262, 2005.
7. Wang, Zhou, et al. "Image quality assessment: from error visibility to structural similarity", Image Processing, IEEE Transactions on 13.4, pp. 600-612, 2004
8. Wang, Zhou, Eero P. Simoncelli, and Alan C. Bovik. "Multiscale structural similarity for image quality assessment", Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on. Vol. 2, pp. 1398-1402, 2003
9. Wichmann, Felix A., and N. Jeremy Hill. "The psychometric function: I. Fitting, sampling, and goodness of fit", Perception & psychophysics 63.8, pp. 1293-1313, 2001
10. Daly, Scott J. "Visible differences predictor: an algorithm for the assessment of image fidelity", SPIE/IS&T Symposium on Electronic Imaging: Science and Technology. International Society for Optics and Photonics, pp. 2-15, 1993
11. Mannos, James, and David J. Sakrison. "The effects of a visual fidelity criterion of the encoding of images", Information Theory, IEEE Transactions on 20.4, pp. 525-535, 1974

Сравнение методов оценки качества изображений

Борис В. Парфененков, Максим А. Паначев

Уральский Федеральный Университет
{idlerboris,tiopox}@gmail.com

Аннотация. Рассматривается задача оценки качества цифрового изображения. Проанализированы преимущества и недостатки существующих методов. Проведена классификация метрик качества в некоторые группы. На основе анализа и тестирования методов на выбранных изображениях, сформировано утверждение о перспективности использования выбранных методов в решении поставленной задачи.

Ключевые слова. анализ изображений, метрики качества изображений, image fidelity, MSE, SSIM, VDP, анизотропия.