A No-Reference Image Sharpness Metric Based on Large-Scale Structure

Bingzhi Pei\textsuperscript{1}, Xiaolin Liu\textsuperscript{1,}\textsuperscript{*} and Zicheng Feng\textsuperscript{1}

\textsuperscript{1}College of Mechatronic Engineering and Automation, National University of Defense Technology, 109Deya Road, Changsha, China

*xlchangsha@sina.com

Abstract. Most of no-reference image sharpness metrics suffer interference of tiny textures and the contrast around edges. Based on large-scale structure, this paper proposes a novel metric to avoid this problem. Firstly prominent edges are extracted by a weighted least-squares filter. Then, we calculate the statistics proportion of edges with different width and the mean value of the maximum gradients from the 51th to 250th as the features of the image. Finally, the support vector regression is used to get the relationship model between the features and the subjective assessment result. The metric can reduce the influence of textures, contrast and the content of images effectively, accordingly enhancing the assessment accuracy.

1. Introduction

In contrast to machines, assessing the sharpness of an image is easy for humans. However, subjective image quality assessment is easily interfered by evaluators and observation conditions. Besides, it is inefficient and can not be embedded in digital image processing systems. All of these shortcomings narrow their applicability. To tackle this challenge, full-reference (FR) quality assessment is considered. In practice, however, FR algorithms are only applicable when both sharp image and blurred image can be captured available at the same physical location and imaging conditions. Thus, no-reference objective image quality assessment is widely used.

In spatial domain, no-reference image sharpness metrics mainly focus on pixel intensity and edge width. Image sharpness can be estimated through the ratio of the grey difference between two edge ends and the edge width, as in [1]. But because the method is dependent on image contrast, it can not accurately evaluate the sharpness of different images. In [2], the notion of just noticeable blur (JNB) based on the characteristics of human visual system is introduced. On this basis, the authors in [3] propose the method of cumulative probability of blur detection (CPBD). The probability of detecting blur is calculated through the edge width and the JNB width, and the cumulative probability histogram is used to evaluate the sharpness of the image.

In frequency domain, metrics mainly associate sharpness with transform coefficients. In [4], the concept of local phase coherence (LPC) is proposed. For a sharp image, the phases of steerable pyramid coefficients exhibit a consistent relationship across scales, and the phases of the finest scale coefficients can be well predicted from the coarser scale coefficients. Therefore, image sharpness can be evaluated by the reduction of LPC. In [5], a generalized Gaussian distribution (GGD) is used to fit the distribution of the normalized steerable pyramid coefficients, and the model parameters can describe the image quality. The distribution of DCT coefficients can also be fitted by GGD as in [6]. In [7], the distribution of Fourier coefficients is similar to a negative exponential distribution.
the exponential parameter and the total variation of the block represents image sharpness. However, due to the methods in [6] and [7] using overlapped blocks, they are time-consuming.

In most cases, edge width can represent image sharpness. The wider the edges, the more blurred the image. However, sometimes there is no obvious increase in edge width after an image is blurred, while gradients of the edge decrease. Besides, it’s considered that textures and image content can interfere with evaluation accuracy.

To avoid these problems, a new metric based on large-scale structure is proposed. After extracting prominent edges by a weighted least-squares filter (WLS), we combine the statistics proportion of edges with different width and the mean value of the maximum gradients from the 51th to 250th to evaluate image sharpness. The flow chart of the proposed method is shown as figure 1. Denote the proportion of edges by \( p_{prop} \), and the mean value by \( GRAD_{in} \).

![Flow chart](image_url)

**Figure 1.** The flow chart of the proposed metric.

Extracting the large-scale structure reduces the interference of textures; calculating \( p_{prop} \) reduces the interference of content; and adding \( GRAD_{in} \) avoids the problem that sometimes there is no obvious increase in edge width after an image is blurred.

This paper is organized as follows. Section 2 mainly introduces the details of the proposed method. Performance results are presented in Section 3. A conclusion is given in Section 4.

2. No-reference sharpness metric

2.1. Edge width

In this paper, an edge point is defined as the pixel with the largest gradient value along the normal direction of the edge. On both sides of the edge point, when the pixel’s gradient value is no longer decreasing or is the last one that is bigger than 10% of the edge point’s gradient, the pixel is the end of the edge. Edge width is considered to be the distance between two ends. Figure 2 illustrates the width measurement by showing the variance of gradient values along the normal direction of two consecutive edges. The black diamonds represent the edge points, the red circles represent the edge ends, and the distances between the adjacent red circles represent the edge width. Sobel operator is chosen to calculate gradients.

When calculating the edge width, in order to avoid the interpolation operation in a certain direction, the normal direction is divided into 0° direction, 45° direction, 90° direction and 135° direction according to the phase of the gradient at the edge point.

In most cases, edge width can represent the sharpness of an image. The more blurred the image, the wider the edges, as shown in figure 3. In order to reduce the effect of the number of edges, this paper selects the statistics proportion of edges with different width as the sharpness feature of an image.
Figure 2. An illustration of the width measurement.

Figure 3. Edges become wider after the image is blurred.

There are so many tiny edges in texture area that will interfere with evaluation accuracy. As shown in figure 4(a) and figure 4(b), ripples on the surface of the lake bring about lots of edges, whose width are small and won’t increase obviously after the image is blurred. And these edges get close to each other. In addition, humans are more concerned with the change of prominent edges while ignoring tiny textures. Therefore, WLS is used to extract the large-scale structure of an image and filter textures, as equation (1):

$$I_{big_{-}stru} = WLS * I_{in}$$  \hspace{1cm} (1)

$I_{in}$ represents the input image, and $I_{big_{-}stru}$ represents the large-scale structure. The result is shown as figure 4(c). Canny operator is chosen to extract edge points of $I_{in}$ and $I_{big_{-}stru}$. It can be found that edges of ripples are erased after WLS is used, as shown in figure 4(b) and figure 4(d).

Figure 4. Large-scale structure and edge maps. (a) Blur image; (b) Edge map of the blur image; (c) Large-scale structure of the blur image; (d) Edge map of the large-scale structure.

For different images, in order to ensure that the extracted edge points are enough and their gradient values are relatively large, this paper chooses different Canny thresholds adaptively. $GRAD_{in}/3$ and $GRAD_{beg_{-}edge}/3$ are selected as threshold. For large blurriness images whose $GRAD_{in}$ is less than 100, textures have been obscured. Thus edge width can be calculated directly in $I_{in}$ without extracting the large-scale structure. And $3*GRAD_{in}/4$ is selected as threshold in order to prevent false edges around the true edges. The threshold is computed as equation (2):

$$T_{in} = \begin{cases} \frac{GRAD_{in}}{3} & GRAD_{in} > 100 \\ \frac{3*GRAD_{in}}{4} & GRAD_{in} \leq 100 \end{cases}$$  \hspace{1cm} (2)

$$T_{beg_{-}stru} = \frac{GRAD_{beg_{-}stru}}{3}$$

$$GRAD_{beg_{-}stru} = \frac{\sum_{i=1}^{N}(grad_{beg_{-}stru})}{200}$$

Since edge width of two images are different, we still calculate it in $I_{in}$, and only the common edge points are taken into consideration. Thus, the interference of textures can be reduced. It’s found that the WLS filter makes the positions of edge points drift. Therefore, when we extract the common edge points, for an edge point A with the position $(x_0, y_0)$ in $I_{beg_{-}stru}$, if there is an edge point B at the same...
position in $I_0$, B is considered as the common edge point. If B doesn’t exist, one edge point in $I_0$ is searched within the eight neighborhoods of $(x_0, y_0)$ as the common edge point. Otherwise, A is discarded.

Based on width, the obtained edges are divided into nine parts: less than or equal to 3, equal to 4 to 9, and equal to or greater than 11, and the statistics proportion of each part is calculated as sharpness features of the image, as equation (3):

$$prop_i = \frac{\sum_{count_i}}{\sum_{count_i}} \quad i = 3, 4, 5\ldots 11$$

2.2. Mean value of Maximum Gradients

It is difficult to evaluate image sharpness only via edge width, because in some cases, the edge width does not change after the image is blurred. However the gradients at edge points become smaller, as shown in figure 5. Thus, gradients can also describe image sharpness.

![Figure 5. Edge width remain unchanged but gradients become smaller after the image is blurred](image)

![Figure 6. The relationship between $GRAD_{in}$ and image scores.](image)

The relationship between $GRAD_{in}$ and image scores are shown as figure 6. So $GRAD_{in}$ is dealt as another sharpness feature of the image, as equation (4).

$$cont = 20\ln(GRAD_{in}) - 56.5$$

2.3. Support vector regression

The relationship model style between image features $prop$, $cont$ and subjective assessment scores is unknown. Linear regression model and RBF-SVR model are chosen to fit the relationship severally.

SVR maps the input samples to a high-dimensional space through an inner product kernel function, and then build a linear model in the high-dimensional space to estimate the regression model. The computation complexity is independent of the dimension of the samples, which avoids the dimension disaster. SVR can avoid over-fitting caused by too many adjustable parameters, so that the evaluation model has better adaptability.

RBF-SVR is proved to have better results via experiments.

3. Performance result

3.1. Experiment setup

This paper uses the LIVE database, the IVC database and the CSIQ database to verify the validity of the proposed method. There are 29 sets of blurred images in the LIVE database, 4 sets in the IVC database, and 30 sets in the CSIQ database. Each set contains 5 different sharpness images with the same content. Images are compressed after blurred in CSIQ database.

Because the LIVE database contains a large number of images which are uncompressed, it is used as the training set and IVC and CSIQ are used as the test set. For the LIVE database itself, 28 sets of images are used for training, and the remained set is used for testing. The process is carried out circularly, in order to avoid that different sharpness images with the same content appear in the training set and the test set at the same time, which leads to a non-objective evaluation result.
3.2. Results and analysis

The proposed method is compared with three classical methods: S3, CPBD and LPC-SI. Scatter plots and four criteria between the objective scores and the subjective scores are employed to evaluate the metric effectiveness. These criteria are SRCC (Spearman rank-order correlation coefficient), KRCC (Kendall's rank-order correlation coefficient), PLCC (Pearson linear correlation coefficient) and RMSE (root mean squared prediction error), as in [4]. SRCC and KRCC are used to evaluate prediction monotonicity, the greater the value, the better the monotonicity. PLCC and RMSE are used to evaluate prediction accuracy, the larger the PLCC and the smaller the RMSE, the higher the accuracy.

![Scatter plots for three databases](image)

**Figure 7.** The scatter plots between the objective scores and the subjective scores for the four metrics on three databases.

The scatter plots for the four metrics on three databases are presented in figure 7. The left of the subtitle is the metric, and the right represents the database. Each point represents a test image, whose ordinate represents the subjective score and abscissa represents the objective score. The closer the point to the diagonal, the better the consistency between the objective score and the subjective score. It can be found from figure 7 that the proposed metric has good monotonicity and accuracy. When the image quality is poor, CPBD and LPC-SI cannot distinguish the image quality, which is obvious in LIVE database and CSIQ database. This is because these two metrics only consider the edge width. However, the proposed metric overcomes this problem by extracting large-scale structure and using $GRAD_{\alpha}$. The scatter plots of S3, CPBD and LPC-SI are obtained from [4].

Table 1 summarizes the four criteria of the four metrics on three databases. Table 2 shows the runtime of the four metrics when processing a 1024*1024 image. The computer configuration is i7 CPU, 8 GB RAM, Windows7 64-bit and Matlab 7.10. It can be found from table 1 that the proposed metric has good monotonicity and accuracy on LIVE and CSIQ, and the criteria are better than others. For IVC, the metric performs averagely. The proposed method is short-time running and takes most of the time in extracting large-scale structure. S3 performs excellently, but it’s time-consuming because of too much fitting in overlapped blocks. The data of S3, CPBD and LPC-SI are obtained from [4].
Table 1. Four Criteria of the four metrics on three databases

| Database       | Measure  | SRCC      | KRCC      | PLCC      | RMSE      |
|----------------|----------|-----------|-----------|-----------|-----------|
| LIVE database  | S3       | 0.9517    | 0.8157    | 0.9494    | 4.9503    |
|                | CPBD     | 0.9271    | 0.7714    | 0.9024    | 6.7943    |
|                | LPC-SI   | 0.9501    | 0.7994    | 0.9219    | 6.1092    |
|                | Proposed | 0.9553    | 0.8219    | 0.9611    | 4.3606    |
| CSIQ database  | S3       | 0.9058    | 0.7290    | 0.9106    | 0.1184    |
|                | CPBD     | 0.8790    | 0.6905    | 0.8822    | 0.1349    |
|                | LPC-SI   | 0.8931    | 0.7022    | 0.9061    | 0.1212    |
|                | Proposed | 0.9212    | 0.7663    | 0.9508    | 0.0888    |
| IVC database   | S3       | 0.8691    | 0.7090    | 0.9274    | 0.4269    |
|                | CPBD     | 0.7744    | 0.6105    | 0.8012    | 0.6832    |
|                | LPC-SI   | 0.9202    | 0.7831    | 0.9574    | 0.3295    |
|                | Proposed | 0.8322    | 0.6773    | 0.9312    | 0.4165    |

Table 2. Runtime of the four metrics for images of 1024*1024 resolution

| Metric         | S3     | CPBD   | LPC-SI | Proposed Metric |
|----------------|--------|--------|--------|-----------------|
| Runtime (s)    | 142.5  | 11.3   | 4.37   | 4.35            |

4. Conclusion
This paper proposes a novel no-reference image sharpness metric based on large-scale structure, which can reduce the interference of textures, contrast and content of images. After prominent edges are extracted by a weighted least-squares filter, the statistics proportion of edges with different width and the mean value of the maximum gradients from the 51th to 250th are calculated and combined to construct the evaluation model using SVR. The metric has shown relatively good evaluating performance for most images. However, for images that are compressed after blurred, the metric is not so effective and needs for further improvements.

Reference
[1] Wang X, Baofeng T, Chao L and Dongcheng S 2008 Blind image quality assessment for measuring image blur CISP 08 1 467-470.
[2] Rony F and Liına J K 2009 A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB) IEEE Trans. Image Processing 18 717 – 728.
[3] Niranjana D N and Liına J K 2011 A no-reference image blur metric based on the cumulative probability of blur detection (CPBD) IEEE Trans. Image Processing 20 2678 – 83.
[4] Hassen R, Wang Z and Magdy S 2013 Image sharpness assessment based on local phase coherence IEEE Trans. Image Processing 22 2798-2810.
[5] Anush K M and Alan C B 2011 Blind image quality assessment: From natural scene statistics to perceptual quality IEEE Trans. Image Processing 20 3350-64.
[6] Michele A, Alan C B and Christophe C 2012 Blind image quality assessment: A natural scene statistics approach in the DCT domain IEEE Trans. Image Processing 21 3339-52.
[7] Cuong T V, Thien D P and Damon M C 2012 S3: A spectral and spatial measure of local perceived sharpness in natural images IEEE Trans. Image Processing 21 934-945.