Reduced-Reference Image Quality Assessment Based on Edge Preservation

Maria G. Martini[1], Barbara Villarini[2,⋆], and Federico Fiorucci[2]

[1] WMN Research Group, Kingston University London
Penrhyn Road, KT12EE, London, UK
[2] DIEI, University of Perugia
via Duranti, Perugia, Italy

Abstract. Assessing the subjective quality of processed images through an objective quality metric is a key issue in multimedia processing and transmission. In some scenarios, it is also important to evaluate the quality of the received images with minimal reference to the transmitted ones. For instance, for closed-loop optimisation of image and video transmission, the quality measure can be evaluated at the receiver and provided as feedback information to the system controller. The original images - prior to compression and transmission - are not usually available at the receiver side, and it is important to rely at the receiver side on an objective quality metric that does not need reference or needs minimal reference to the original images.

The observation that the human eye is very sensitive to edge and contour information of an image underpins the proposal of our reduced reference (RR) quality metric, which compares edge information between the distorted and the original image.

Results highlight that the metric correlates well with subjective observations, also in comparison with commonly used full-reference metrics and with a state-of-the-art reduced reference metric.

Keywords: Edge detection, image quality assessment, reduced-reference, quality index, mean opinion score (MOS), SSIM.

1 Introduction

The quality of images can be assessed based on a variety of methodologies. Subjective tests, such as the mean opinion score (MOS) [1], are difficult to reproduce. In addition, they are expensive, time consuming and typically cannot be implemented algorithmically. On the other side, pure objective measurements such as mean square error (MSE) and peak signal-to-noise ratio (PSNR) do not adequately represent the subjective quality [2].

For both performance assessment and on-the-fly system adaptation, it is crucial in image/video compression and transmission to develop an objective quality

⋆ B. Villarini was with Kingston University London, UK.

L. Atzori, J. Delgado, and D. Giusto (Eds.): MOBIMEDIA 2011, LNICST 79, pp. 31–45, 2012.

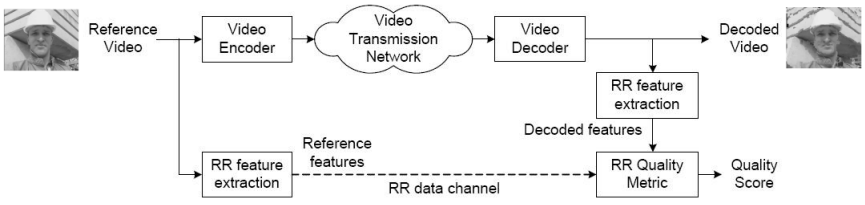

Fig. 1. Reduced reference scheme

assessment metric which accurately represents the subjective quality of compressed and corrupted images. Objective methods based on subjective measurements are based either on a perceptual model of the Human Visual System [3], or on a combination of relevant parameters tuned with subjective tests [4] [5].

It is also important to evaluate the quality of the received images with minimal reference to the transmitted ones [6]. For closed loop optimisation of image/video transmission, the image quality measure can be provided as feedback information to a system controller [7]. The original images - prior to compression and transmission - are not usually available at the receiver side, hence an objective quality metric that does not need reference or needs minimal reference to the original images is required.

The human eye is very sensitive to edge and contour information of an image, *i.e.*, the edge and contour information give a good indication of the structure of an image and it is critical for a human to capture the scene [8]. Some works in the literature proposed full-reference (FR) image quality metrics considering edge structure information. For instance in [9] the structural information error between the reference and the distorted image is computed based on the statistics of the spatial position error of the local modulus maxima in the wavelet domain.

We propose here a reduced reference quality metric which compares the edge information between the distorted image and the original one. We consider the Sobel operator [10] for edge detection, due to its simplicity and efficiency. Further details on this choice are reported in the following Section.

2 Edge Detection

There are many methods to perform edge detection. The most used among these can be grouped into two categories: gradient and Laplacian. The gradient method detects the edges by finding the maximum and minimum in the first derivative of the image. The Sobel method is an example of these. A pixel location is declared an edge location if the value of the gradient exceeds a threshold. Edges will have higher pixel intensity values than those surrounding it. Once a threshold is set, the gradient value can be compared to the threshold value and an edge is detected when the threshold is exceeded. When the first derivative is at a maximum, the

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

Fig. 2. Sobel masks

second derivative is zero. As a result, an alternative to finding the location of an edge is to locate the zeros in the second derivative. This method is known as the Laplacian, since it makes use of the Laplacian operator.

The aforementioned methods can be extended to the two-dimensions case. The Sobel operator performs a 2-D spatial gradient measurement on an image. The Sobel edge detector uses a pair of 3×3 convolution masks, one estimating the gradient in the x-direction (columns) and the other estimating the gradient in the y-direction (rows). The mask is then slid over the image, manipulating a square block of pixels at a time. The Sobel operator can hence detect edges by calculating the partial derivatives in 3×3 neighborhood. The main reason for using the Sobel operator is that it is relatively insensitive to noise and it has relatively smaller masks with respect to other operators such as the Roberts operator and the two-order Laplacian operator.

The partial derivatives in x and y directions are given as:

$$S_x = f(x+1, y-1) + 2f(x+1, y) + f(x+1, y+1) + \\ -[f(x-1, y-1) + 2f(x-1, y) + f(x-1, y+1)] \tag{1}$$

and

$$S_y = f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1) + \\ -[f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1)] \tag{2}$$

The gradient of each pixel is calculated according to $g(x, y) = \sqrt{S_x^2 + S_y^2}$ and a threshold value t is selected. If $g(x, y) > t$, the point is regarded as an edge point.

The Sobel operator can also be expressed in the form of two masks as shown in Figure 2: the two masks are used to calculate S_y and S_x, respectively.

3 Proposed Metric

Since structural distortion is tightly linked with edge degradation, our reduced reference (RR) quality metric compares edge information between the distorted image and the original one. We propose to apply Sobel filtering locally, only for some blocks of the entire image, after subsampling the images. The different steps for the calculation of the metric are reported in detail below, together with the relevant motivation.

Subdivision of the Image in Sub-windows. The first step consists of the subdivision of images in sub-windows. For instance, if images have size 512 x 768 we could subsample of a factor of 2 and consider 16×16 macroblocks of size 16×24 each, or we can subsample of a factor 1.5 and consider 18×16 macroblocks with size 19×32 each, as in the example reported in Figure 3.

Selection of the Visual Attention Area. In order to reduce the overhead associated with the transmission of side information, only 12 blocks are selected to represent the different areas of the images. The block pattern utilized for our tests is chosen after several investigations based on visual attention (VA). Various experiments have been proposed in the literature for VA modeling and salient region identification, aiming at the detection of salient regions in an image. Models on visual attention are often developed and validated by visual fixation patterns through eye tracking experiments [11] [12]. In [13] a framework is proposed in order to extend existing image quality metrics with a simple VA model. In this work a subjective Region Of Interest (ROI) experiment was performed, with 7 images, in which the viewers' task was to select within each image the region mostly attracting their attention. For simplicity, in this experiment, only rectangular-shaped ROIs were allowed. Considering the obtained ROI as a random value, it is possible to calculate the mean value and the standard deviation. It was observed that the ROI's center coordinates are around the image center for most of the images, and the mean values of the ROI dimensions are very similar in both x and y directions. This confirms that the salient region, which include the most important informative content of the image, is often placed in the center of the picture.

Following these observations we selected the block pattern as a subset of the ROI with a central symmetry, by minimizing the number of blocks in order to reduce the overhead associated to the transmission of side information. Figure 3 shows the block pattern considered below.

Edge Comparison. We propose to compare the edge structure of the blocks of the corrupted image to the structure of the corresponding blocks in the original image. For this purpose, we apply Sobel filtering locally in the selected blocks.

For each pixel in each block we obtain a bit value, where one represents an edge and zero means that there are no edges. If m and n are the block dimensions, we denote the corresponding blocks l in the original and in the possibly corrupted image as the $m \times n$ matrices O_l and C_l respectively, and the Sobel-filtered version of blocks l as the $m \times n$ binary matrices $SO_l = \mathcal{S}(O_l)$, with elements $so_{i,j}$, with $i = 1, ..., m, j = 1, ..., n$, and $SC_l = \mathcal{S}(C_l)$, with elements $sc_{i,j}$, with $i = 1, ..., m, j = 1, ..., n$. We denoted above with $\mathcal{S}()$ the Sobel operator. The similarity of two images can be assessed based on the similarity of the edge structures, *i.e.*, by comparing the matrices SO_l, associated to the filtered version of the block in the original image, and SC_l, associated to the filtered version of the block in the possibly corrupted image.

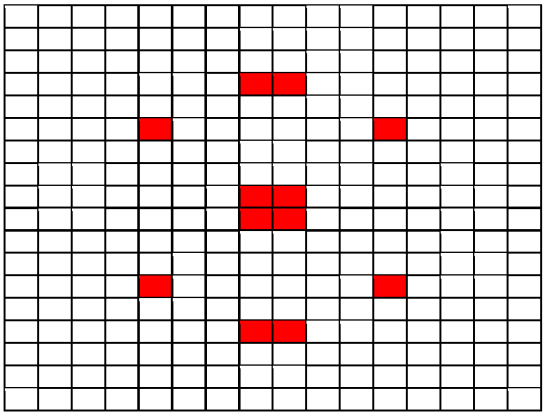

Fig. 3. Example of block pattern selected based on visual attention models

The following step is thus the quantification of the edge similarity between the reference and the processed images, done by summing the zeros and ones which are unchanged after compression or lossy transmission of the image.

Hence, for each block the similarity index can be computed as

$$I_{s,l} = n_l / p_l \tag{3}$$

where

$$n_l = p_l - \sum_{i=1}^{m} \sum_{j=1}^{n} |sc_{l,ij} - so_{l,ij}| \tag{4}$$

is the number of zeros and ones unchanged in the $l-th$ block and $p_l = m \times n$ is the total number of pixels in the $l-th$ block.

If N_b is the number of blocks in the selected block pattern, the similarity index I_s for an image is defined here as

$$I_s = \frac{1}{N_b} \sum_{l=1}^{N_b} I_{s,l} \tag{5}$$

For images decomposed in blocks of equal size, as considered here, the proposed quality index is thus:

$$I_s = \frac{1}{N_b} \sum_{l=1}^{N_b} \left(1 - \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} |sc_{l,ij} - so_{l,ij}|}{mn} \right) \tag{6}$$

3.1 Threshold Selection

The threshold value is an important parameter that depends on a number of factors, such as image brightness, contrast, level of noise, and even edge direction. The selection of the threshold in Sobel filtering is associated to the sensitivity of the filter to edges. In particular, the lower the value of the threshold, the higher the sensitivity to edges. Too high values of the threshold do not detect edges which are important for quality assessment. On the other side, if the value of the threshold is too small, large parts of the image are considered as edges, whereas these are irrelevant for quality assessment. The threshold can be selected following an analysis of the gradient image histogram. Based on this consideration and on the analysis of the performance of Sobel filtering for the images of the considered databases, the selected threshold value is $t = 0.001$.

3.2 Complexity

The selection of Sobel filtering results in a low complexity metric. The Sobel algorithm is characterized, in fact, by a low computational complexity and consequently high calculation speed. In [14] some edge detection techniques are compared for an application which uses a DSP implementation: the Sobel filter exhibits the best performance in terms of edge detection time in comparison with the other wavelet-based edge detectors. Sobel filtering has been implemented in hardware and used in different areas, often when real-time performance is required, such as for real-time volume rendering systems, and video assisted transportation systems [15] [16]. This makes the proposed metric suitable for real-time implementation, an important aspect when an image/video metric is used for the purpose of "on the fly" system adaptation as in the scenario considered here.

3.3 Overhead

In order to perform the proposed edge comparison, we should transmit the matrices composed of one's and zeros's in the reference blocks. By considering the pattern in Figure 3, this would result in the transmission of $19 \times 32 \times 12 = 7.29$ kbits per image. Note that the size of the original image (not compressed) is $3 \times 512 \times 768 \times 8 = 9.4$ Mbits.

Since side information is in our case composed of a large number of zeros appearing in long runs, it is possible to reduce the overhead by compressing the relevant data, *e.g.*, through run-length encoding, or to transmit only the positions of ones in the matrix.

In order to reduce both overhead and complexity, we propose to calculate the metric on the luminance frame only.

4 Simulation Set-Up and Results

In order to test the performance of our quality assessment algorithm, we considered publicly available databases.

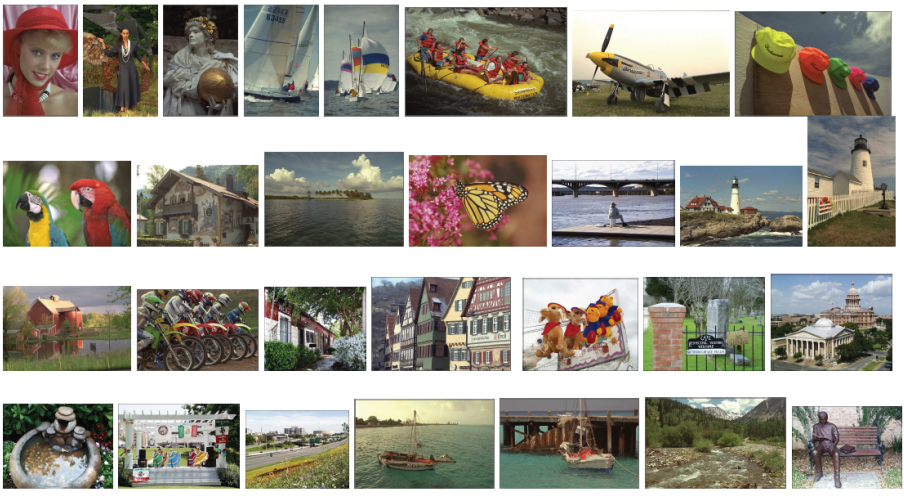

Fig. 4. Images in the LIVE [18] database

The first one [17] is provided by the Laboratory for Image & Video Engineering (LIVE) of the University of Texas Austin (in collaboration with the Department of Psychology at the same University). An extensive experiment was conducted to obtain scores from human subjects for a number of images distorted with different distortion types. The database contains 29 high-resolution (typically 768×512) original images (see Figure 4), altered with five types of distortions at different distortion levels: besides the original images, images corrupted with JPEG2000 and JPEG compression, white-noise, Gaussian blur and JPEG2000 compression and subsequent transmission over Rayleigh fading are considered. Our interest is in particular on the data representing transmission errors in the JPEG2000 bit stream using a fast-fading (FF) Rayleigh channel model, since our goal is to assess the quality of images impaired by both compression and transmission errors. Our quality metric is tested versus the subjective quality values provided in the database. Subjective results reported in the database were obtained with observers providing their quality score on a continuous linear scale that was divided into five equal regions marked with adjectives Bad, Poor, Fair, Good and Excellent. Two test sessions, with about half of the images in each session, were performed. Each image was rated by 20-25 subjects. No viewing distance restrictions were imposed, and normal indoor illumination conditions were provided. The observers received a short training before the session. The raw scores were converted into difference scores (between the test and the reference) and then converted to Z-scores [20], scaled back to 1 - 100 range, and finally a difference mean opinion score (DMOS) for each distorted image was obtained.

The second database, IRCCyN/IVC [19], was developed by the *Institut de Recherche en Communications et Cyberntique de Nantes*. It is a 512×512

Fig. 5. Images in the IRCCyN/IVC [19] database

pixels color images database. This database is composed by 10 original images and 235 distorted images generated by 4 different processing methods / impairments: JPEG, JPEG2000, LAR (Locally Adaptive Resolution) coding and blurring. Subjective evaluations were made at a viewing distance of 6 times the screen height, by using a Double Stimulus Impairment Scale (DSIS) method with 5 categories and 15 observers. The images in the database are reported in Figure 5.

With the aid of the databases above, we compare the performance versus subjective tests for the following metrics:

- Structure SIMilarity Index (SSIM) [4] (full reference);
- PSNR (full reference);
- Reduced-reference metric in [6];
- Proposed Sobel-based metric (reduced reference).

To apply the SSIM metric, the images have been modified according to [21].

We report our results in terms of scatter plots, where each symbol in the plot refers to a different image: Figures 6, 7, 8, and 9 report scatter plots for the metrics above in the case of compression according to the JPEG2000 standard and subsequent transmission over a fast fading channel.

The figures report, besides scatter plots, the linear approximation best fitting the data using the least-squares method, the residuals and the norm of residuals L for the linear model, i.e., $L = \sqrt{\sum_{i=1}^{N}(d_i)^2}$, where the residual d_i is the difference between the predicted quality value and the experimental subjective quality value for image i, and N is the number of the considered images. The values of the norms of residuals enable a simple numerical comparison among the different metrics. Note that in the case of the SSIM metric we have provided a non-linear approximation, better fitting the data.

A summary of the results for the LIVE database [18] in terms of norms of residuals is reported in Table 1.

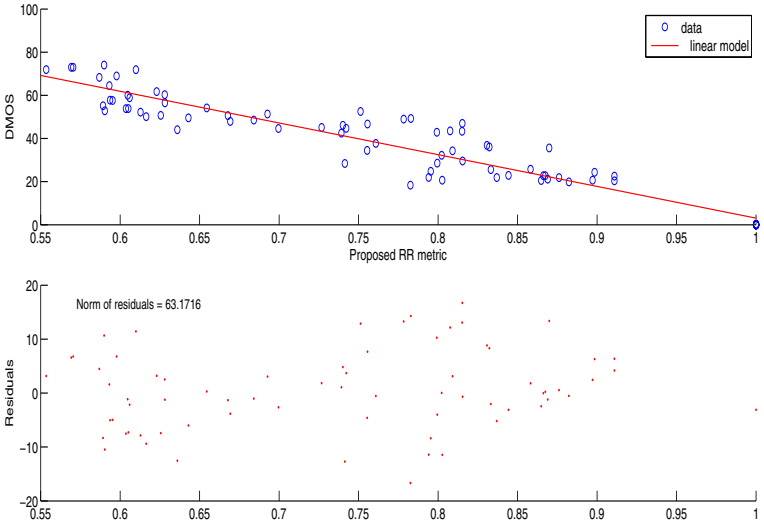

Fig. 6. Fast fading, LIVE image database [18] - Proposed metric. Above: scatter plot between difference mean opinion score and proposed metric. Below: residuals for the linear approximation and norm of residuals.

We can observe that our metric well correlates with subjective tests, with results comparable to those achieved by full reference metrics. Our metric outperforms the considered state-of-the-art reduced reference metric in all the considered scenarios, except JPEG2000 compression for images in the LIVE database. In the latter case, the benchmark reduced reference metric, based on the wavelet transform, provides a better performance.

However, for the same type of impairment (JPEG2000 compression) our metric performs slightly better than the benchmark one when the images in the IRCCyN/IVC database [19] are considered. The relevant results are reported in Table 2, where results for JPEG compression are also reported. Figures 10, 11, 12, and 13 present in detail the relevant results for the case of JPEG compression.

Table 1. Norm of residuals versus DMOS, LIVE image database [18]

	PSNR	RR [6]	Proposed RR	SSIM
JPEG 2000 + Fast fading	69.80	85.03	63.17	53.64
White noise	25.25	63.68	60.82	31.44
Gaussian blur	90.08	70.63	55.96	34.06
JPEG compression	82.83	115.06	83.04	65.58
JPEG2000 compression	70.23	63.94	84.33	74.06

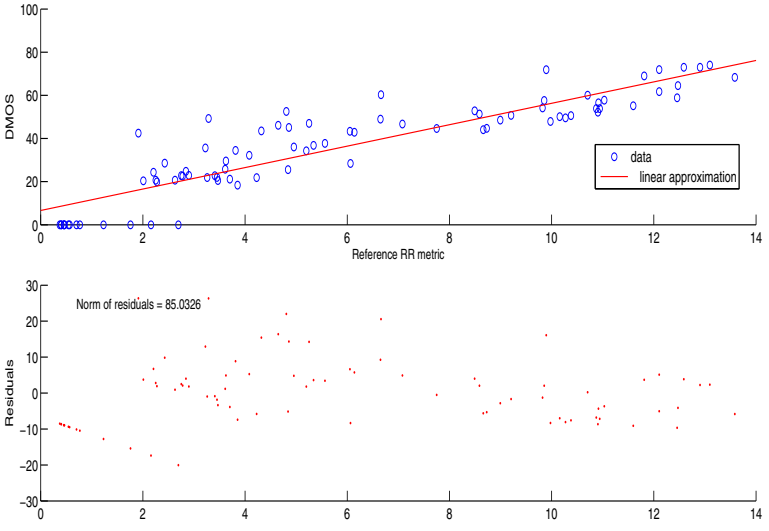

Fig. 7. Fast fading, LIVE image database [18] - Reduced reference metric in [6]. Above: scatter plot between difference mean opinion score and metric in [6]. Below: residuals for the linear approximation and norm of residuals.

We can observe that, in this case, our metric performs better not only with respect to the benchmark reduced-reference metric, but also with respect to PSNR, regardless of the need in the latter for full reference information.

Table 2. Norm of residuals versus MOS, IRCCyN/IVC image database [19]

	PSNR	RR [6]	Proposed RR	SSIM
JPEG compression	6.60	7.29	5.11	3.75
JPEG2000 compression	5.30	5.42	5.27	5.28

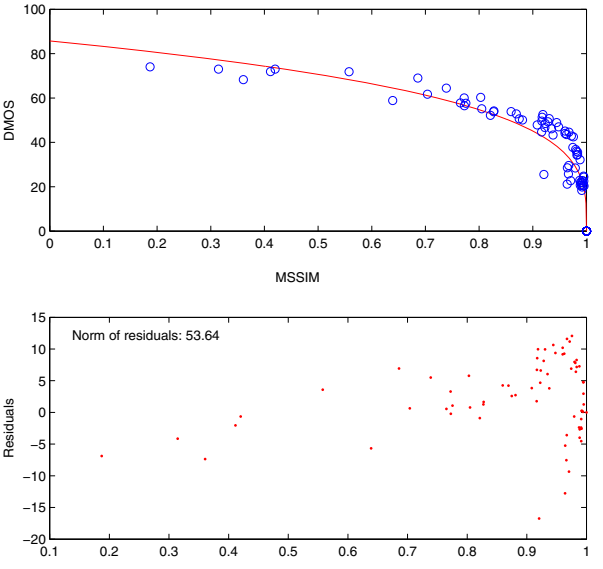

Fig. 8. Fast fading, LIVE image database [18] - MSSIM. Above: scatter plot between difference mean opinion score and MSSIM. Below: residuals for the considered approximation and norm of residuals.

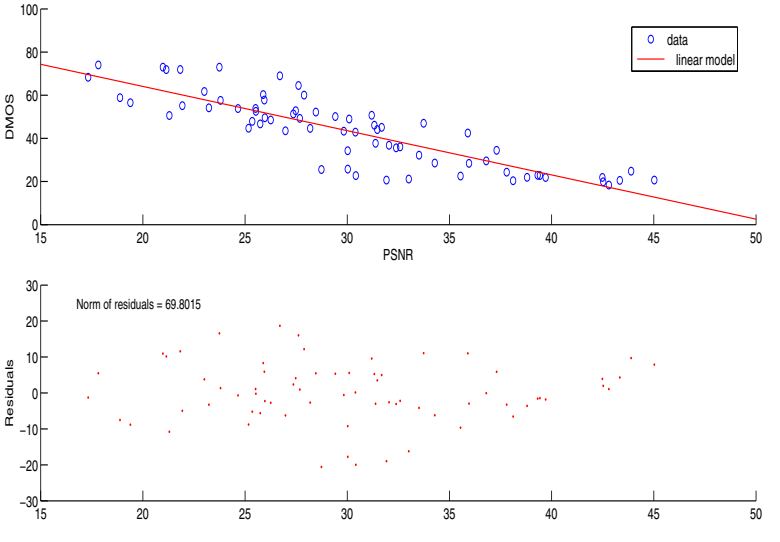

Fig. 9. Fast fading, LIVE image database [18] - PSNR. Above: scatter plot between difference mean opinion score and PSNR. Below: residuals for the linear approximation and norm of residuals.

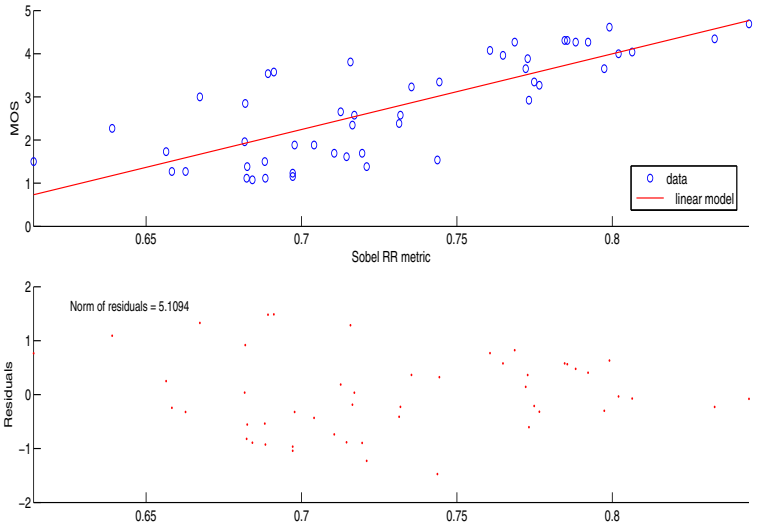

Fig. 10. JPEG compression, IRCCyN/IVC image database [19] - Proposed metric. Above: scatter plot between mean opinion score and proposed metric. Below: residuals for the linear approximation and norm of residuals.

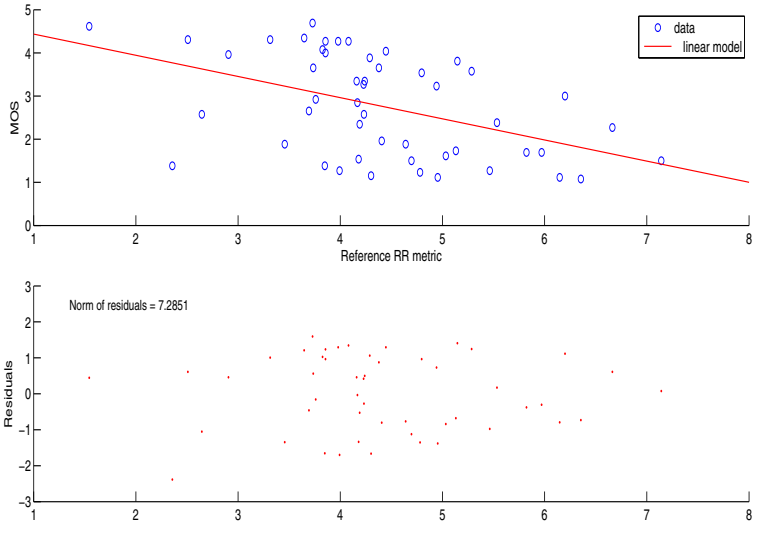

Fig. 11. JPEG compression, IRCCyN/IVC image database [19] - Reduced reference metric in [6]. Above: scatter plot between mean opinion score and metric in [6]. Below: residuals for the linear approximation and norm of residuals.

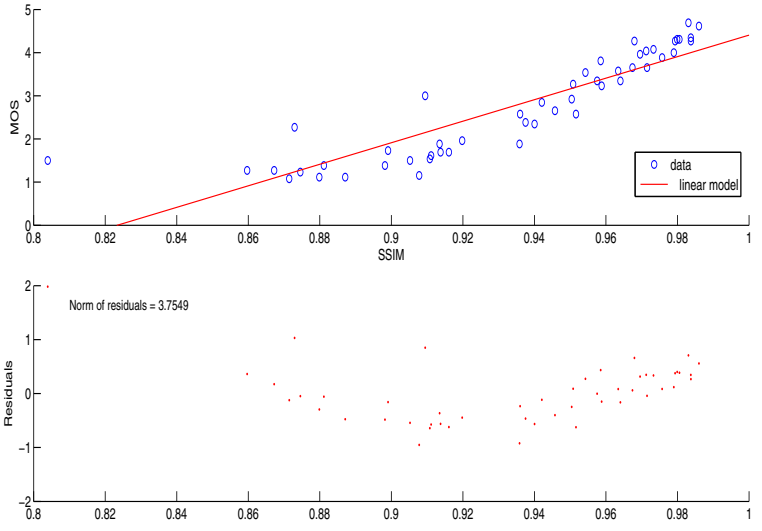

Fig. 12. JPEG compression - IRCCyN/IVC image database [19], MSSIM. Above: scatter plot between mean opinion score and MSSIM. Below: residuals for the linear approximation and norm of residuals.

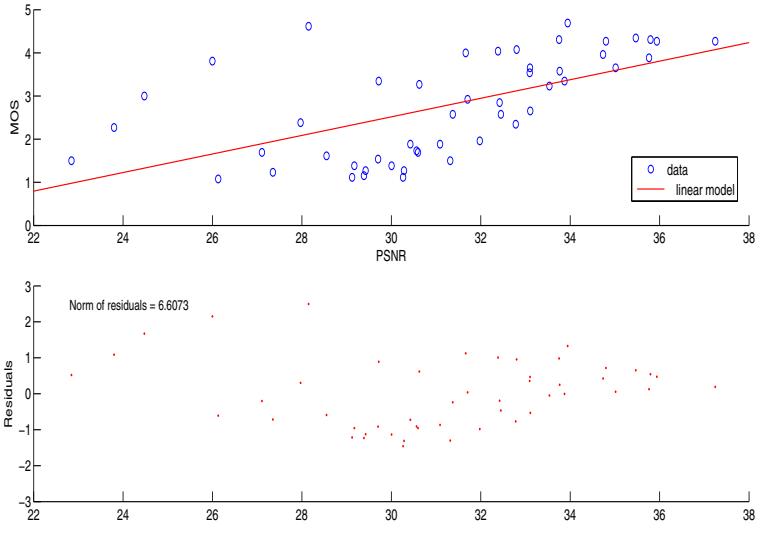

Fig. 13. JPEG compression - IRCCyN/IVC image database [19] - PSNR. Above: scatter plot between mean opinion score and PSNR. Below: residuals for the linear approximation and norm of residuals.

5 Conclusion

We proposed in this paper a perceptual reduced reference image quality metric which compares edge information between portions of the distorted image and of the original one by using Sobel filtering. The algorithm is simple and has a low computational complexity. Results highlight that the proposed metric well correlates with subjective observations, also in comparison with commonly used full-reference metrics and with state-of-the-art reduced-reference metrics. Hence, it appears suitable for the assessment of the quality experienced by the user without the need for full reference information.

References

1. Pinson, M.H., Wolf, S.: Comparing subjective video quality technologies. In: Proc. of SPIE Video Communication and Image Processing, Lugano, Switzerland (September 2003)
2. Eskicioglu, A.M., Fisher, P.S.: Image quality measures and their performance. IEEE Transactions on Comms. 43, 2959–2965 (1995)
3. Pinson, M.H., Wolf, S.: A new standardized method for objectively measuring video quality. IEEE Transactions on Broadcasting 50(3), 312–322 (2004)
4. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error measurement to structural similarity. IEEE Trans. Image Processing 13(4), 600–612 (2004)
5. Sheikh, H.R., Sabir, M., Bovik, A.C.: A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans. Image Processing 15(11), 3440–3451 (2006)
6. Wang, Z., Simoncelli, E.P.: Reduced-reference image quality assessment using a wavelet-domain natural image statistic model. In: Human Vision and Electronic Imaging, pp. 149–159 (March 2005)
7. Martini, M.G., Mazzotti, M., Lamy-Bergot, C., Huusko, J., Amon, P.: Content adaptive network aware joint optimization of wireless video transmission. IEEE Communications Magazine 45(1), 84–90 (2007)
8. Marr, D., Hildreth, E.: Theory of edge detection. Proceedings of the Royal Society of London. Series B (1980)
9. Zhang, M., Mou, X.: A psychovisual image quality metric based on multi-scale structure similarity. In: Proc. IEEE International Conference on Image Processing (ICIP), San Diego, CA, pp. 381–384 (October 2008)
10. Woods, J.: Multidimensional Signal, Image and Video Processing and Coding. Elsevier (2006)
11. Yarbus, A.L.: Eye Movements and Vision. Plenum Press, New York (1967)
12. Privitera, C.M., Stark, L.W.: Algorithms for defining visual regions-of-interest: comparison with eye fixations. IEEE Trans. Pattern Anal. Mach. Intell. 22(9), 970–982 (2000)
13. Engelke, U., Zepernick, H.: Framework for optimal region of interest-based quality assessment in wireless imaging. Journal of Electronic Imaging 19(1), 011 005–1 – 011 005–13 (2010)
14. Musoromy, Z., Bensaali, F., Ramalingam, S., Pissanidis, G.: Comparison of real-time DSP-based edge detection techniques for license plate detection. In: Sixth International Conference on Information Assurance and Security, Atlanta, GA, pp. 323–328 (August 2010)

15. Zhou, W., Xie, Z., Hua, C., Sun, C., Zhang, J.: Research on edge detection for image based on wavelet transform. In: Proceedings of the 2009 Second International Conference on Intelligent Computation Technology and Automation, Washington, DC, USA, pp. 686–689 (2009)
16. Kazakova, N., Margala, M., Durdle, N.G.: Sobel edge detection processor for a real-time volume rendering system. In: Proc. of the 2004 International Symposium on Circuits and Systems (ISCAS 2004), pp. 913–916 (May 2004)
17. Seshadrinathan, K., Soundararajan, R., Bovik, A.C., Cormack, L.K.: LIVE video quality assessment database (2010), http://live.ece.utexas.edu/research/quality/live_video.html
18. Sheikh, H.R., Wang, Z., Cormack, L., Bovik, A.C.: Live image quality assessment database (2008), http://live.ece.utexas.edu/research/quality
19. Callet, P.L., Autrusseau, F.: Subjective quality assessment IRCCyN/IVC database (2005), http://www.irccyn.ec-nantes.fr/ivcdb/
20. van Dijk, A.M., Martens, J.B., Watson, A.B.: Quality assessment of coded images using numerical category scaling. In: Proc. SPIE, vol. 2451, pp. 99–101 (1995)
21. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: The SSIM index for image quality assessment (2008), http://www.ece.uwaterloo.ca/~z70wang/research/ssim/#usage