

RESEARCH

Open Access



# GMVP: gradient magnitude and variance pooling-based image quality assessment in sensor networks

Zhong Zhang and Shuang Liu\*

## Abstract

In this paper, we focus on image quality assessment (IQA) in sensor networks and propose a novel method named gradient magnitude and variance pooling (GMVP). The proposed GMVP follows a two-step framework. In this first step, we utilize gradient magnitude to compute the local quality, which is efficient and responsive to degeneration when the images are transmitted by sensor networks. In the second step, we propose a weighted pooling operation, i.e., variance pooling, which explicitly considers the importance of different local regions. The variance pooling operation assigns different weights to local quality map according to the variance of local regions. The proposed GMVP is verified on two challenging IQA databases (CSIQ and TID 2008 databases), and the results demonstrate that the proposed GMVP achieves better results than the state-of-the-art methods in sensor networks.

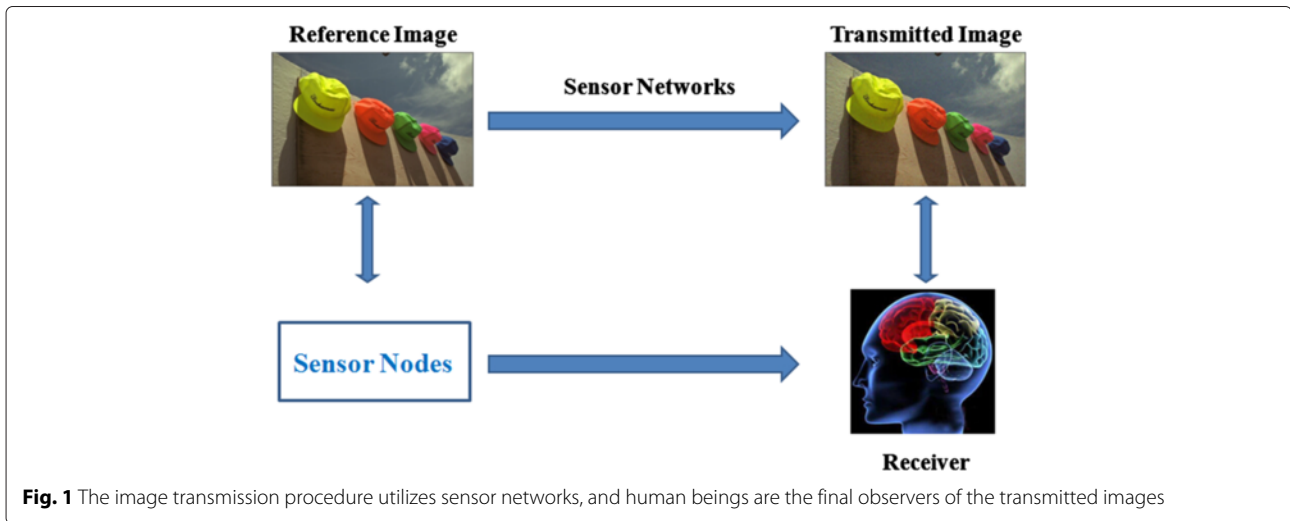
**Keywords:** Sensor networks, Image quality assessment, Variance pooling, Gradient magnitude

## 1 Introduction

With the rapid development of wireless communications and electronics, sensor networks have received much attention in research fields [1, 2]. The wireless sensor network (WSN) consists of a variety of sensors, such as video cameras, microphones, infrared badges and RFID tags, which drives the applications of WSN in the fields of surveillance systems, guiding systems, biological detection, habitat, agriculture, and health monitoring. There are a mount of images transmitted in sensor networks. Thus, finding ways to test the performance of sensor networks about the transmitted image quality has provoked great interests in research fields. In this paper, we focus on image quality assessment (IQA) for testing sensor network. Human beings are the final observers of the transmitted images, and therefore, they are entitled to evaluate the image quality as shown in Fig. 1. Hence, the target of IQA is to develop automatic methods that can predict image quality consistently with human subjective evaluation.

There are three kinds of IQA models in terms of the availability of a reference image: full reference (FR) models where the pristine reference image is available, reduced reference (RR) models where only a small fraction of reference information is available, and no reference (NR) models where the reference image is unavailable. This paper only discusses FR-IQA models which can be widely used to evaluate the performance of image transmission system, e.g., sensor networks, by measuring the quality of their output images. Generally speaking, FR-IQA models can be classified into two types. The first one is built under a bottom-up framework [3–5] which simulates the various processing stages in the visual pathway of human visual system (HVS), including just noticeable differences [6], visual masking effect [7], etc. Nevertheless, HVS is too intricate to construct an accurate bottom-up FR-IQA framework. The second one constructs a top-down framework [8–11] which designs to model the overall function of HVS according to some global assumption. Recent studies [8, 9] have demonstrated the effectiveness of these kinds of methods, and thus, many approaches follow the top-down framework. The structural similarity (SSIM) [12], as a representative approach of top-down model, is based on the assumption that HVS is highly adapted to extract the structural information from the visual scene,

\*Correspondence: shuangliu.tjnu@gamil.com  
College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China



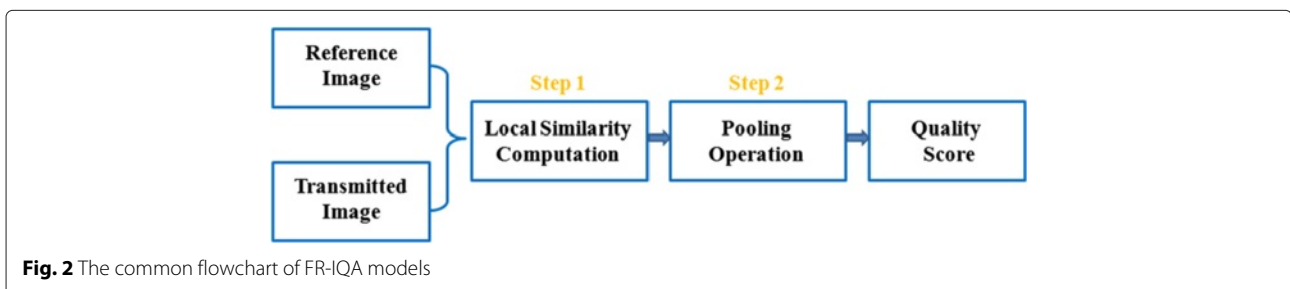
**Fig. 1** The image transmission procedure utilizes sensor networks, and human beings are the final observers of the transmitted images

and thus, a measurement of SSIM should provide a good approximation of image quality. The improvements of SSIM, for example, multi-scale structural similarity (MS-SSIM) [13], three-component weighted SSIM (3-SSIM) [14], and information-weighted SSIM [15] also employ the same assumption and achieve better results than original SSIM. Moreover, information fidelity criteria (IFC) [16] and visual information fidelity (VIF) [17] regard HVS as a communication channel. The subjective image quality is predicted by computing how much the information in the reference image is preserved in the transmitted one.

From another point of view, many FR-IQA models consist of two stages [15, 18, 19] as shown in Fig. 2. The first step is local similarity computation which is calculated by locally comparing the transmitted image with the reference image according to some similarity function. Considering the computational complexity, many approaches adopt image gradient as a measurement feature [20–22] due to effectively capturing image local structure which is incentive to HSV. Most gradient-based FR-IQA models [8, 9] are inspired by SSIM [12]. They first compute the similarity between the gradients of the reference image and transmitted image and then compute some additional information, such as the difference of gradient orientation and luminance similarity, to combine

with the gradient similarity. The second stage is the pooling operation which obtains a single overall quality score calculated from local similarity computation. The pooling operations, which aggregate similarity map or vectors into a single score or one vector, are widely used in many fields, such as image quality assessment [23], image classification [24, 25], human action recognition [26, 27]. The common used pooling operation is the average pooling, i.e., calculating the average of all local quality values as the final quality score. However, the average pooling treats each local region in an image equally, which neglects the local contrast information of the reference image. As a result, some weighted pooling operations, including visual attention [28], assumed visual fixation [29], and distortion-based weighting [30], have been proposed and achieve better performance than the average pooling.

In this paper, we propose a novel FR-IQA model named gradient magnitude and variance pooling (GMVP) for testing sensor networks. First, we utilize gradient magnitude, i.e., Sobel filter, to compute the local quality, which is responsive to artifacts introduced by compression, blur or additive noise, etc. In addition, natural images usually have diverse local structures which reflects the degree of importance of different local regions. Based on the consideration, we propose a novel pooling operation, i.e.,



**Fig. 2** The common flowchart of FR-IQA models

variance pooling, which assigns different weights according to the variance of local regions. Our method is verified on two challenging IQA databases, and the experimental results demonstrate that the proposed GMVP achieves higher prediction accuracy than that of previous methods on image quality assessment.

The rest of this paper is organized as follows. We present the proposed GMVP in Section 2, including Sobel similarity and variance pooling. Section 3 shows the experimental results which outperform the state-of-the-art methods on the two publicly IQA databases. Finally, in Section 4, we conclude this paper.

## 2 Gradient magnitude and variance pooling

### 2.1 Sobel similarity

Many gradient-based FR-IQA approaches utilize a similarity function to calculate gradient similarity [8, 10, 20]. In addition to gradient magnitude, these approaches also adopt other similarity features, for example, luminance similarity and structural similarity. Zhang et al. [8] combined phase congruency, which is a dimensionless measure of the significance of a local structure, with gradient magnitude. However, the computation of phase congruency is time consuming.

The proposed GMVP only utilizes gradient magnitude as the similarity feature to increase computational efficiency. The gradient magnitude is defined as the root mean square of image directional gradients along two orthogonal directions. We usually calculate the gradient by convolving an image with a filter, for instance Sobel, Prewitt, Roberts filters, or others [31, 32]. In GMVP, we adopt the Sobel filter, which is a simple and classic  $3 \times 3$  filter, to detect the gradient. The gradient along horizontal ( $x$ ) and vertical ( $y$ ) directions using Sobel filters is calculated by

$$f_x = \frac{1}{4} \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \quad (1)$$

$$f_y = \frac{1}{4} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \quad (2)$$

The reference image ( $ri$ ) and transmitted image ( $ti$ ) are filtered by Sobel operator, and then the gradient magnitudes of  $ri$  and  $ti$  at location  $(m, n)$  are calculated by

$$g_{ri}(m, n) = \sqrt{(ri \otimes f_x)^2(m, n) + (ri \otimes f_y)^2(m, n)} \quad (3)$$

$$g_{ti}(m, n) = \sqrt{(ti \otimes f_x)^2(m, n) + (ti \otimes f_y)^2(m, n)} \quad (4)$$

where  $g_{ri}(m, n)$  and  $g_{ti}(m, n)$  are the gradient magnitudes of  $ri$  and  $ti$  at location  $(m, n)$ , respectively.  $\otimes$  denotes the convolution operation. With the gradient magnitude

images  $g_{ri}$  and  $g_{ti}$ , the Sobel similarity map (SSM) is computed by

$$SSM(m, n) = \frac{2g_{ri}(m, n)g_{ti}(m, n) + T}{g_{ri}^2(m, n) + g_{ti}^2(m, n) + T} \quad (5)$$

where  $T$  is a positive constant. The SSM reflects the similarity between the reference image and transmitted image. Specifically, when  $g_{ri}(m, n)$  and  $g_{ti}(m, n)$  are the same, the  $SSM(m, n)$  is the maximal value 1.

Some example images about gradient magnitude and SSM are shown in Fig. 3. The first and second columns denote reference and transmitted images, respectively. The third and fourth columns are the gradient magnitudes of reference and transmitted images, respectively. The last column indicates the Sobel similarity map which is the input of pooling stage.

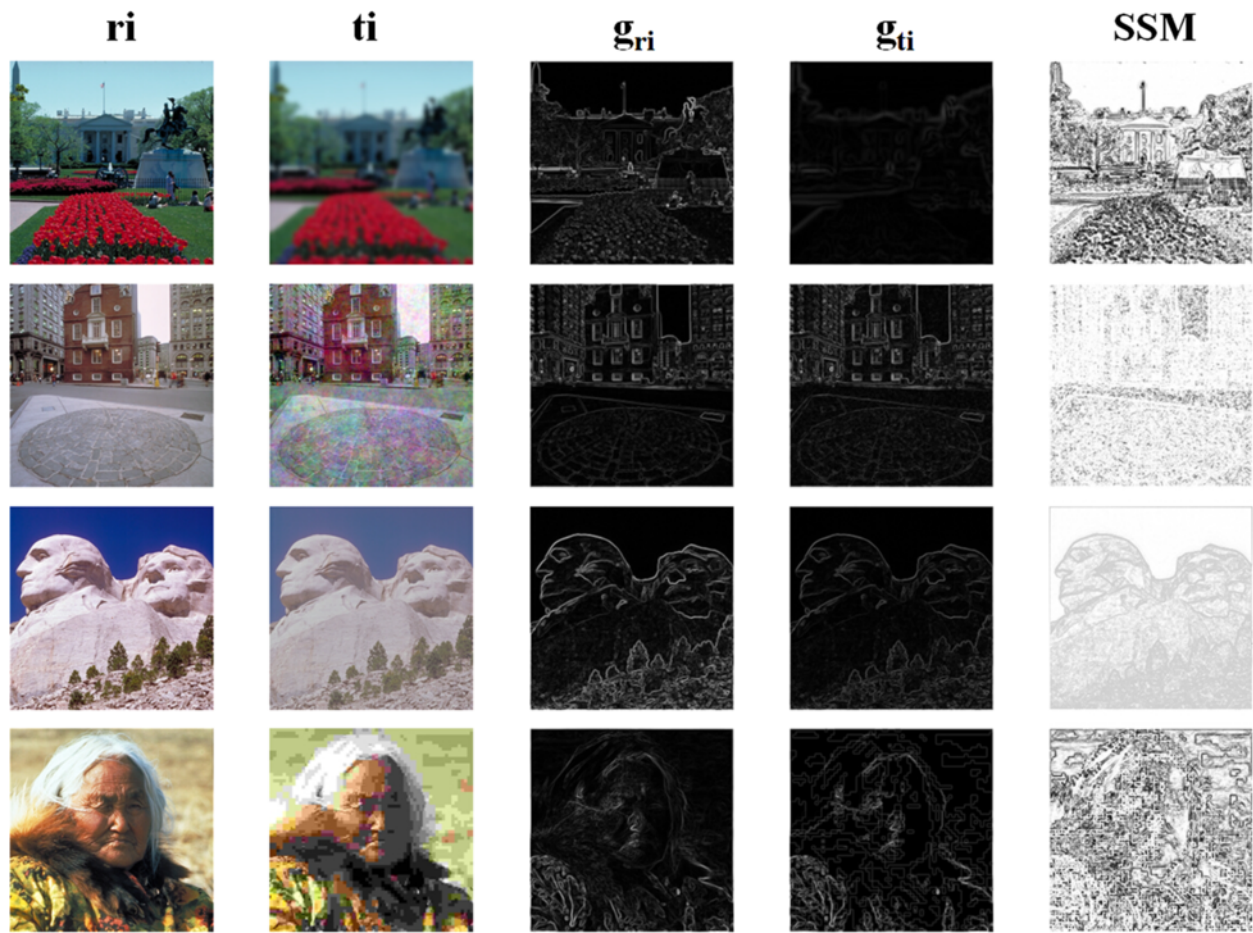
### 2.2 Variance pooling

The final quality score can be obtained from SSM via pooling operation. The most commonly used pooling operation is average pooling, i.e., averaging all the SSM values as the final quality score. However, this pooling operation equally treats each SSM value, i.e., each SSM value is assigned the same weight 1 regardless of what the local structure is. It fails to capture the local contrast information in the reference image. Figure 4b shows the variance map of the  $3 \times 3$  region of Fig. 4a. From this figure, we can see that the local regions with high contrast contain much structure information, and therefore, they should contribute more to evaluate the image quality. In order to exploit the local contrast information of the reference image, we define the variance of a local region as a weight in the pooling stage

$$w(m, n) = \frac{1}{R+1} \sum_{p=m-R}^{m+R} \sum_{q=n-R}^{n+R} (ri(p, q) - \theta(m, n))^2 \quad (6)$$

where  $\theta(m, n) = \frac{1}{R+1} \sum_{p=m-R}^{m+R} \sum_{q=n-R}^{n+R} ri(p, q)$  is the mean value of reference image  $ri$  at location  $(m, n)$ , and  $R$  is the radius of the local region. Here, we set  $R$  to 1, i.e., we calculate the weight in a  $3 \times 3$  local region. Furthermore, the pooling weight  $w$  should be normalized using the maximum value of  $w$ .

There are two advantages about the variance pooling. First, different local regions show different variance which considers microscopic structures of reference image. Specifically, when a local region is flat, its variance is low, while the variance is high at the boundaries of different regions. Second, the pooling weight  $w$  and SSM are complementary. Their joint distribution can better



**Fig. 3** Some example images about gradient magnitude and SSM from CSIQ database.  $ri$  reference image,  $ti$  transmitted image,  $g_{ri}$  gradient magnitude of  $ri$ ,  $g_{ti}$  gradient magnitude of  $ti$ ,  $SSM$  Sobel similarity map

characterize the difference between reference image and transmitted image. The final quality score is computed by

$$score = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N w(m,n)SSM(m,n) \quad (7)$$

Note that the higher quality score, the higher image transmission quality.

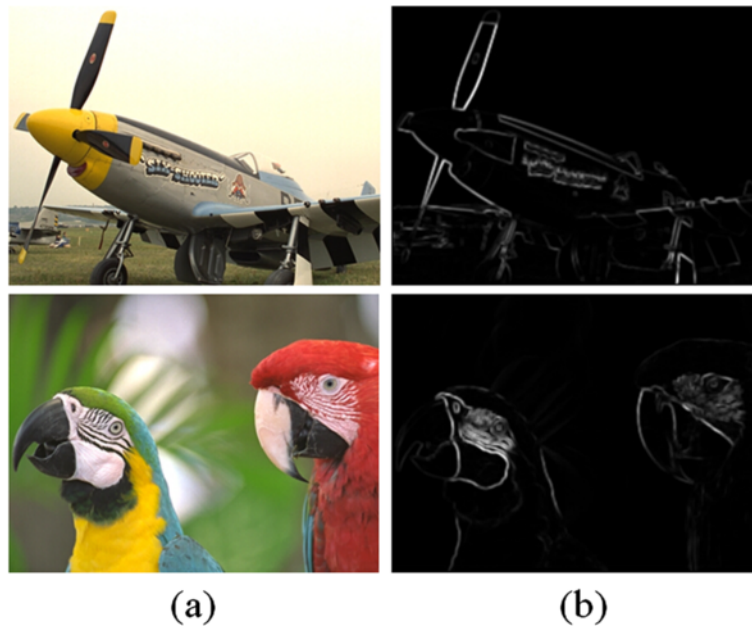
### 3 Experimental results

#### 3.1 Experimental setup

We verify the proposed GMVP on two publicly available databases: CSIQ database [33] and TID2008 database [34]. It should be noted that we consider the distorted images as transmitted images because the images will degenerate when they are transmitted by sensor networks. The CSIQ database consists of 886 transmitted images and 30 reference images. The transmitted images contains six types of distortions at five different distortion levels. Concretely, the six types of distortions involve

JPEG compression, JPEG 2000 compression, Gaussian blur (GB), additive white noise (AWN), additive pink Gaussian noise (APGN), and global contrast decrements (GCD). The TID 2008 database has 1700 transmitted images and 25 reference images with 17 kinds of distortions at 4 levels. Note that each image in the IQA databases has been assessed by human beings under controlled conditions and then assigned a quantitative quality score: mean opinion score (MOS) or difference MOS (DMOS).

For fair comparison, we employ three commonly used criteria to evaluate the proposed GMVP. The first criterion is the Pearson linear correlation coefficient (PLCC) between MOS and the objective scores after nonlinear regression. The second criterion is the Spearman rank-order correlation coefficient (SROCC) which measures the prediction monotonicity of an IQA approach. The last one is the root-mean-square error (RMSE) between MOS and the objective scores. We adopt the nonlinear regression proposed in [35].



**Fig. 4 a** Reference image and **b** its variance map of the  $3 \times 3$  region

### 3.2 Performance comparison

The choice of local region size has an impact on the GMVP performance. They are parameterized by a discrete set  $\{2, 3, 4, 5\}$ . We evaluate different values for RMSE on the CSIQ database. The value of RMSE is  $\{0.082, 0.070, 0.078, 0.094\}$ , and therefore, we choose 3 as the local region size. With the optimal parameter, Table 1 shows the comparative results on the CSIQ database according to PLCC, SROCC, and RMSE. We show the top three methods in boldface for each evaluation criterion. Note that higher PLCC and SROCC values, or a lower RMSE value,

indicate better performance. From this table, we can see that the proposed GMVP achieves the best results based on all three criteria. The proposed GMVP obtains better performance than the other gradient-based approaches, i.e., FSIM, G-SSIM, GSD, GS, and GMSD, because the proposed GMVP adopts weighted pooling operation which explicitly considers the local structure contribution.

We also test the performance of the proposed GMVP on the TID2008 database. The results are shown in Table 2. The proposed GMVP achieves better results in all criteria. It is because the proposed GMVP not only utilizes the

**Table 1** Performance of the proposed GMVP and other methods on the CSIQ database

	PLCC	SROCC	RMSE
IFC [16]	0.837	0.767	0.144
GSD [11]	0.854	0.854	0.137
G-SSIM [9]	0.874	0.872	0.127
SSIM [12]	0.861	0.876	0.133
VIF [17]	0.928	0.919	0.098
GS [20]	0.896	0.911	0.116
MS-SSIM [13]	0.899	0.913	0.115
MAD [33]	0.950	0.947	0.082
IW-SSIM [15]	0.914	0.921	0.106
FSIM [8]	0.912	0.924	0.108
GMSD [22]	0.954	0.957	0.079
GMVP	0.962	0.958	0.070

**Table 2** Performance of the proposed GMVP and other methods on the TID2008 database

	PLCC	SROCC	RMSE
IFC [16]	0.203	0.568	1.314
GSD [11]	0.707	0.657	0.949
G-SSIM [9]	0.760	0.731	0.873
SSIM [12]	0.773	0.775	0.851
VIF [17]	0.808	0.749	0.790
GS [20]	0.842	0.850	0.723
MS-SSIM [13]	0.845	0.854	0.717
MAD [33]	0.829	0.834	0.751
IW-SSIM [15]	0.858	0.856	0.689
FSIM [8]	0.874	0.880	0.653
GMSD [22]	0.879	0.891	0.640
GMVP	0.882	0.897	0.628



gradient information, but also assigns different weights in the pooling stage. Once again, we prove the effectiveness of our algorithm on this database.

#### 4 Conclusions

This paper proposes a novel FR-IQA approach named GMVP to overcome the limitation of traditional average pooling operation. For computational efficiency, we adopt the Sobel filter to compute the local quality map. Then, we explicitly consider the local contrast information of reference images in the pooling stage. To this end, the variance of a local region is defined as a weight which is utilized to reflect the importance of local regions. The experimental results on CSIQ and TID 2008 databases show that the proposed GMVP achieves better results than previous approaches in testing sensor networks.

#### Competing interests

The authors declare that they have no competing interests.

#### Acknowledgements

This work is supported by the National Natural Science Foundation of China under Grant No. 61401309, No. 61501327, and No. 61401310, Natural Science Foundation of Tianjin under Grant No. 15JQJJC01700, and Doctoral Fund of Tianjin Normal University under Grant No. 5RL134 and No. 52XB1405.

Received: 24 August 2015 Accepted: 2 November 2015

Published online: 12 January 2016

#### References

- Q Liang, X Cheng, S Huang, D Chen, Opportunistic sensing in wireless sensor networks: theory and application. *IEEE Trans. Comput.* **63**(8), 2002–2010 (2014)
- L Zhao, Q Liang, Hop-distance estimation in wireless sensor networks with applications to resources allocation. *EURASIP J. Wirel. Commun. Netw.*, 1–8 (2007). (Article ID 084256)
- Z Wang, A Bovik, Modern image quality assessment. *Synth. Lect. Image, Video, and Multimed. Process.* **2**(1), 1–156 (2006)
- W Lin, C Kuo, Perceptual visual quality metrics: a survey. *J. Vis. Commun. Image Represent.* **22**(4), 297–312 (2011)
- L Zheng, S Wang, Q Tian, Coupled binary embedding for large-scale image retrieval. *IEEE Trans. Image Process.* **23**(8), 3368–3380 (2014)
- J Lubin, in *Proceedings of International Broadcasting Convention*. A human vision system model for objective picture quality measurements, (1997), pp. 498–503
- J Ross, H Speed, Contrast adaptation and contrast masking in human vision. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **246**(1315), 61–70 (1991)
- L Zhang, L Zhang, X Mou, D Zhang, Fsim: a feature similarity index for image quality assessment. *IEEE Trans. Image Process.* **20**(8), 2378–2386 (2011)
- G Chen, C Yang, S Xie, in *Proceedings of IEEE International Conference on Image Processing*. Gradient-based structural similarity for image quality assessment, (2006), pp. 2929–2932
- D Kim, H Han, R Park, Gradient information-based image quality metric. *IEEE Trans. Consum. Electron.* **56**(2), 930–936 (2010)
- G Cheng, J Huang, C Zhu, Z Liu, L Cheng, in *Proceedings of IEEE International Conference on Image Processing*. Perceptual image quality assessment using a geometric structural distortion model, (2010), pp. 325–328
- Z Wang, A Bovik, H Sheikh, E Simoncelli, Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
- Z Wang, E Simoncelli, A Bovik, in *Proceedings of IEEE Asilomar Conference on Signals, Systems and Computers*. Multiscale structural similarity for image quality assessment, (2003), pp. 1398–1402
- C Li, A Bovik, in *Proceedings of IS&T/SPIE Electronic Imaging*. Three-component weighted structural similarity index, (2009)
- Z Wang, Q Li, Information content weighting for perceptual image quality assessment. *IEEE Trans. Image Process.* **20**(5), 1185–1198 (2011)
- H Sheikh, A Bovik, GD Veciana, An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Trans. Image Process.* **14**(12), 2117–2128 (2005)
- H Sheikh, A Bovik, Image information and visual quality. *IEEE Trans. Image Process.* **15**(2), 430–444 (2006)
- L Zhang, L Zhang, X Mou, D Zhang, in *Proceedings of IEEE International Conference on Image Processing*. A comprehensive evaluation of full reference image quality assessment algorithms, (2012), pp. 1477–1480
- L Zheng, S Wang, Q Tian, Lp-norm IDF for scalable image retrieval. *IEEE Trans. Image Process.* **23**(8), 3604–3617 (2014)
- A Liu, W Lin, M Narwaria, Image quality assessment based on gradient similarity. *IEEE Trans. Image Process.* **21**(4), 1500–1512 (2012)
- C Li, A Bovik, Content-partitioned structural similarity index for image quality assessment. *Signal Process. Image Commun.* **25**(7), 517–526 (2010)
- W Xue, L Zhang, X Mou, A Bovik, Gradient magnitude similarity deviation: a highly efficient perceptual image quality index. *IEEE Trans. Image Process.* **23**(2), 684–695 (2014)
- Z Wang, X Shang, in *Proceedings of IEEE International Conference on Image Processing*. Spatial pooling strategies for perceptual image quality assessment, (2006), pp. 2945–2948
- N Murray, F Perronnin, in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Generalized max pooling, (2014), pp. 2473–2480
- L Zheng, S Wang, Z Liu, Q Tian, Fast image retrieval: query pruning and early termination. *IEEE Trans. Multimed.* **17**(5), 648–659 (2015)
- Z Zhang, C Wang, B Xiao, W Zhou, S Liu, Attribute regularization based human action recognition. *IEEE Trans. Inf. Forensics and Security.* **8**(10), 1600–1609 (2013)
- Z Zhang, C Wang, B Xiao, W Zhou, S Liu, Cross-view action recognition using contextual maximum margin clustering. *IEEE Trans. Circ. Syst. Video Technol.* **24**(10), 1663–1668 (2014)
- A Moorthy, A Bovik, Visual importance pooling for image quality assessment. *IEEE J Sel. Top. Signal Proc.* **3**(2), 193–201 (2009)
- Y Tong, H Konik, F Cheikh, A Tremeau, Full reference image quality assessment based on saliency map analysis. *J Imaging Sci. Tech.* **54**(3), 1–14 (2010)
- J Park, K Seshadrinathan, S Lee, A Bovik, Video quality pooling adaptive to perceptual distortion severity. *IEEE Trans. Image Process.* **22**(2), 610–620 (2013)
- S Coleman, B Scotney, S Suganthan, Multi-scale edge detection on range and intensity images. *Pattern Recog.* **44**(4), 821–838 (2011)
- E Nezhadarya, R Ward, in *Proceedings of IEEE International Conference on Image Processing*. An efficient method for robust gradient estimation of RGB color images, (2009), pp. 701–704
- E Larson, D Chandler, Most apparent distortion: full-reference image quality assessment and the role of strategy. *J Electron. Imaging.* **19**(1), 1–21 (2010)
- N Ponomarenko, V Lukin, A Zelensky, K Egiazarian, M Carli, F Battisti, Tid2008-a database for evaluation of full-reference visual quality assessment metrics. *Adv. Mod. Radioelectron.* **10**(4), 30–45 (2009)
- Video quality experts group and others, in *Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II (FR\_TV2)*, (2003)