ORIGINAL PAPER

No-reference image quality assessment of authentically distorted images with global and local statistics

Milosz Rajchel[1] · Mariusz Oszust[1]

Abstract

The development of digital image processing techniques requires reliable image quality assessment (IQA) methods. Since images acquired by a camera often contain various distortions and their non-distorted versions are not available, a no-reference IQA (NR-IQA) technique should be used. Many popular methods are developed to assess artificially distorted images, available in benchmark databases. In this paper, a new large benchmark database, containing naturally distorted images captured with a digital camera, is introduced along with a new NR-IQA metric. The method uses a wide spectrum of local and global image features and their statistics to address a diversity of distortions. Among 80 employed features, 56 are introduced to the IQA for the first time, while the remaining statistics are used to further improve the quality prediction performance of the method. The obtained perceptual feature vector is used to provide a quality model with support vector regression technique. The experimental comparison of the method with the state-of-the-art IQA measures on the database reveals its superiority in terms of correlation with human scores.

1 Introduction

The recent growth of digital image processing techniques results in the need for the development of automatic image quality assessment (IQA) methods, aiming to replace tests with human subjects [3]. Since image processing often alters the visual content, IQA methods often support parameter tuning and comparison of various techniques designed for image restoration, enhancement, acquisition, storage, or transmission. Depending on the availability of a distortion-free reference image, IQA metrics are divided into no-reference (NR), full reference (FR), and reduced reference (RR) techniques [3]. Full reference metrics are commonly used. However, they can be employed in cases in which a reference image is available, reducing their applicability. Reduced reference methods, in turn, use some properties of a reference image. In practice, the NR techniques are desired, despite their challenging development.

In order to predict the objective quality of an image, NR-IQA methods often mimic properties of the human visual system (HVS) or use features sensitive to image distortions. Here, to model natural scene statistics (NSS), different domains are taken into account, e.g., spatial [14], DCT [21], or wavelet [15]. Also, perceptual features are identified [9,16,27,31] or multiple cues are found [28,30]. Another direction is to find interesting parts of an image for a description [1,16] or to describe its all pixels [7,8]. Such approaches require bridging the image features with the perceptual quality. Hence, the support vector regression (SVR) [14–16,21,31] or neural networks [10] are often used for this purpose. A different approach can be found in techniques that are based on deep learning, in which feature extraction and learning steps are fused. Since they require large databases for training or suffer from an architecture devoted to image recognition tasks, in these methods, image patches [2,6], objective scores of FR-IQA methods [6,12], fine-tuning [29], or handcrafted features [12] are typically employed.

Most of the IQA measures are designed to evaluate artificially distorted images, present in popular benchmarks. Such a benchmark contains images distorted by several distortion types with different severity of degradation. Since an

✉ Mariusz Oszust
  marosz@kia.prz.edu.pl

[1] Department of Computer and Control Engineering, Rzeszow University of Technology, W. Pola 2, 35-959 Rzeszow, Poland

image in those databases is contaminated by only one distortion type [20,25], there are also databases with multiply contaminated images [23], but natural distortions are still seldom addressed [4]. Therefore, in this paper, a novel database with real distortions is introduced. Since the state-of-the-art IQA methods may exhibit inferior performance on images with multiple and authentic distortions, it is important to introduce new methods for the quality prediction of such images. Therefore, in this paper, along with the database, a new NR-IQA method is proposed. In the method, many novel quality-aware features that deliver a global and local description of an image are introduced. Furthermore, to improve the quality prediction, some of well established in IQA literature quality indicators and their modifications are employed. Local features detect and describe interesting regions of an image from the HVS point of view, while global features are sensitive to overall noise or image degradation. Finally, the SVR is used to map the obtained feature vector to subjective scores and provide a quality model used to obtain objective scores for assessed images. The contributions of this work can be summarized as follows:

1. A novel IQA database with authentically distorted images and subjective scores is introduced.
2. A set of novel quality-aware features is proposed.
3. Modifications of popular features to better address natural distortions are considered.
4. A novel NR-IQA method that incorporates a reasonably small number of features and provides superior image quality prediction performance in comparison with the state-of-the-art methods is developed and described.
5. Experiments regarding the comparative evaluation of the method with related techniques on the new database are conducted and reported.

The rest of this paper is arranged as follows: In Sect. 2, a novel IQA database is introduced and compared with widely recognized benchmarks. Then, in Sect. 3, the proposed NR technique is described. In Sect. 4, a comparative evaluation of the technique with the state-of-the-art NR-IQA measures is presented. Section 5 concludes the paper and indicates future directions.

2 New database with authentically distorted images

In practice, most digital images captured by a mobile camera contain multiple distortions of various types. To reliably assess them, IQA measures should be developed taking into account their characteristics or be trained on images captured in similar conditions. The scarcity of such databases in the literature resulted in only one measure devoted to

such images, published together with a benchmark [4]. It is worth noticing that the most databases in this field cover many distortion types or their mixture obtained by artificial contamination of reference images [19,23]. For pictures captured by a camera, reference images are not available, hence the inability of using best performing FR-IQA techniques. In this paper, a new relatively large authentically distorted database of images assessed by human subjects is introduced. The database is created using images available in the Beautiful Rzeszow (BR) dataset, which contains 3000 images of 50 tourist attractions in Rzeszow, Poland [18]. Since the BR dataset is designed for image recognition problems, each attraction is associated with one good quality image serving as its base representation and 60 photographs captured at a different time of a day (day and night) and seasons (Spring, Autumn, and Winter). For this work, 1500 images were selected and assessed by human subjects using a paired comparison (PC) or pair-wise sorting methodology [19]. In the subjective test according to the methodology, three images are displayed, and then, an observer selects a better image between two distorted ones, considering the third (base) image. Consequently, the observers were not assigning scores, but a better, worse, and equal image in the pair obtained 1, 0, and 0.5 points, respectively. Finally, the points for images were averaged, considering the number of comparisons in which an image took part, and used to obtain the MOS. As reported by authors of TID2008 and TID2013 [19], the used methodology is easier to be applied in tests and more convenient for observers. In tests, each image was evaluated 2–13 times by 22 participants, while a participant assessed 100 randomly selected pairs of images. One full test lasted less than 15 min, which gives around 9 seconds per image pair. All tests were in line with the VQEG recommendations [24].

The comparison of the introduced database with other benchmarks is shown in Table 1, while best- and low-quality images are shown in Fig. 1. Among compared databases, CID2013 and LIVE In the Wild [4] databases contain authentic distortions. However, the human scores in the second dataset come from crowdsourcing, making it less reliable due to an uncontrolled manner of data collection. The main differences between the existing databases and the proposed BR database are: (i) large number of authentically distorted images, (ii) subjective scores obtained in a laboratory setting, and (iii) availability of images which were used as supporting information on observed objects for human observers. The base images could be helpful for the development of future FR-IQA measures which, similarly to human observers, would not compare pixels between images to indicate the objective quality but take into account a relationship between image content (i.e., an object and background) and its quality. Nowadays, there exist full reference methods which assess

Table 1 Comparison of IQA benchmark datasets

DB	Year	Type	No. of ref. im.	No. of test im.	Method.	No. of subj.	Dist. levels	No. of dist. per im.	Dist. types	Format	No. of ratings	Res.
A57	2007	Graysc.	3	54	NA	7	NA	NA	6	BMP	NA	512×512
BID	2011	Color	NA	585	SS	180	NA	1–5	5	JPG	11	NA
CID2013	2013	Color	NA	480	ACR-DR	188	NA	NA	12–14	JPG	31	1600×1200
CSIQ	2010	Color	30	866	NA	35	4–5	1	6	PNG	5–7	512×512
DRIQ	2012	Color	26	78	NA	9	NA	NA	3	PNG	NA	512×512
HTI	2011	Color	12	60	NA	18	NA	NA	1	NA	NA	512×768
IBBI	2011	Color	12	60	NA	18	NA	NA	1	NA	NA	321×481
IRCCyN/DIBR	2011	Color	3	96	NA	43	NA	NA	3	NA	NA	1024×768
IRCCyN/MICT	2008	Color	14	168	NA	27	NA	NA	2	BMP	NA	768×512
IVC(i)	2005	Color	10	185	DSIS	15	5	1	4	BMP	15	512×512
LIVE	2006	Color	29	779	SS	29	5	1	5	BMP	23	768×512
LIVE In the Wild	2015	Color	NA	1162	SS	>8100	NA	NA	NA	JPEG, BMP	175	500×500
MDID	2016	Color	20	1600	PCR	192	4	1–4	5	BMP	33–35	512×384
MDIQ	2012	Color	15	405	SS	37	4	2	3	BMP	18–19	1280×720
MICT	2008	Color	14	168	SS	16	6	1	2	BMP	16	768×512
TID2008	2008	Color	25	1700	PC	838	4	1	17	BMP	33	384×512
TID2013	2013	Color	25	3000	PC	971	5	1–2	24	BMP	~30	512×384
TUD1	2010	Color	8	16	NA	12	NA	NA	1	BMP	NA	NA
TUD2	2010	Color	11	44	NA	20	NA	NA	1	BMP	NA	NA
VCL@FER	2011	Color	23	522	NA	118	NA	NA	4	NA	NA	NA
WID/BA	2009	Graysc.	10	120	NA	17	NA	NA	2	PGM	NA	512×512
WID/Enrico	2007	Graysc.	5	100	NA	16	NA	NA	10	BMP	NA	512×512
WID/FSB	2009	Graysc.	5	210	NA	7	NA	NA	6	BMP	NA	512×512
WID/MW	2009	Graysc.	12	120	NA	14	NA	NA	2	BMP	NA	512×512
WIQ	2009	Graysc.	7	80	NA	30	NA	NA	1	BMP	NA	512×512
BR	**2019**	**Color**	**50**	**1500**	**PC**	**22**	**NA**	**NA**	**NA**	**JPEG**	**2–13**	**up to 1024×768**

The CID2013, LIVE Challenge, and the introduced BR dataset contain images with authentic distortions
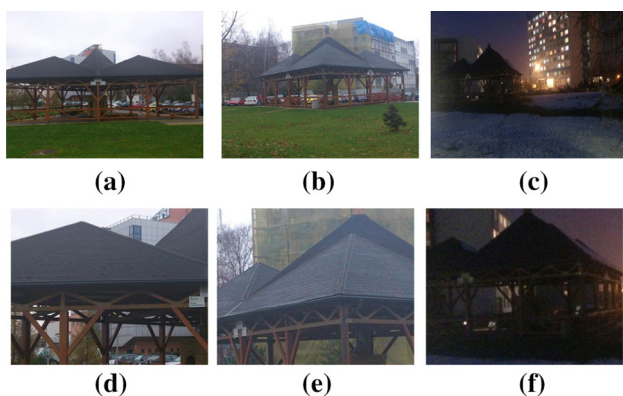The characteristics of BR dataset are written in bold

Fig. 1 Exemplary images of the same object (**a**, **b**, **c**) and their magnifications to highlight differences in quality (**d**, **e**, **f**)

images, disregarding rotation or scale differences between a reference and its distorted equivalent [11].

3 Proposed NR approach

In the literature, for the IQA of artificially distorted images, a variety of approaches have been used [3]. However, as reported by Ghadiyaram and Bovik [4], to reliably assess images captured in real conditions, a set of diverse features is required. Hence, in the measure devoted to the evaluation of such images, FRIQUEE, a bag of various features in several color spaces are employed [4]. In FRIQUEE, after transformation of an image into RGB, LMS, and CIE LAB color spaces, a set of feature maps is obtained by applying diverse operations, among which steerable pyramid decomposition (applied from legacy NR-IQA method C-DIIVINE), the difference of Gaussians, or the Laplacian decomposition is used [4]. Then, statistics or perceptual models are created using specified color channels or filtered images.

In the method introduced in this paper, a set of novel features that describe an image taking into local and global characteristics is proposed. It is assumed that to address various natural distortions diverse quality-aware features would be more suitable than homogeneous quality indicators employed by methods designed for the IQA of images with controlled distortion levels. The new method, namely *QUality Evaluator of Authentically Distorted Images* (QUEADI), uses novel 56 features, including sharpness, brightness, image invariants and moments, or statistics of FREAK descriptors. Also, the feature vector of QUEADI contains statistics of SURF features [16], a small set of well-established quality indicators that are often used in the literature, and statistics of proposed modifications of SURF features. The usage of 24 popular features or statistics is justified by the need for the development of the best possible image quality model, additionally improving the

IQA performance of QUEADI. It is worth noticing that the method without these features outperforms the state-of-the-art NR techniques (Sect. 4). As local features calculated for descriptors of interesting, from the HVS point of view, image regions, as well as some new perceptual statistics can successfully model distortions, they should be considered. Also, their usage can be motivated by the local presence of distortions in real images, contrary to artificially distorted images in which distortions are often uniformly distributed. The QUEADI calculates the following statistics for an image, filtered image, or (SURF/FREAK) descriptors: skewness, kurtosis, entropy, histogram variance, sample mean, and standard deviation. The skewness $s(v)$ is calculated as $s(v) = \overline{(v - \bar{v})^3}/sd(v)^3$, where $\bar{v}$ denotes the sample mean $(m(v) = \bar{v})$ and sd the standard deviation. Here, $sd(v) = \sqrt{\overline{((v - \bar{v})^2)}}$. The v denotes processed image or descriptors. The kurtosis, in turn, is obtained as $k(v) = \overline{(v - \bar{v})^4}/sd(v)^4 - 3$. The entropy $e(v) = -\sum_i p_i(v) \log_2 p_i(v)$, where p_i is the histogram counts for v. The histogram variance is defined as $hvar(v) = \sum_v ((h(v) - \bar{h})^2$, where $h(v)$ denotes the histogram of v normalized to unit sum [10]. The following moments for an image are considered: 0th moment and eight central moments $((0, 2), (0, 3), (1, 1), (2, 1), (1, 2), (2, 0), (2, 1), (3, 0))$, while ten moment invariants are determined using relationships between individual moments. Furthermore, the ratio between bright and dark pixels, sharpness, or mean and median of kurtosis are used. In QUEADI, also popular perceptual features are incorporated to provide a performance gain of the obtained quality model. Therefore, features such as a variance of Asymmetric Generalized Gaussian Distribution (AGGD) [14], the histogram variance of gradient magnitude (GM) and relative gradient magnitude (RM) [10], independency distribution of LOG conditioned on image gradient [27], statistics of an image or an image filtered with Prewitt operators, and their 64-dimensional SURF descriptors obtained for local features detected using the determinant of the Hessian [16]. These features are extended by using other color spaces than proposed by their authors or introducing completely new features. For example, QUEADI employs the 128-dimensional version of the SURF or statistics of the binary descriptor (FREAK) on keypoints detected by the FAST technique [17]. The calculation of statistics of feature vectors obtained with the second descriptor together with other introduced features, such as moments and invariants of an image, can be considered as novel contributions of this work. The features used by QUEADI to produce the quality model are summarized in Table 2.

The obtained 80-dimensional feature vector is further reduced to the length of 13 using the principal component analysis technique (PCA) since some information can be

Table 2 Summary of features used in the proposed NR-IQA measure

Feature number	Input	Color space	Feature	Number of features
f1	GM map [10]	Weighted sum (YCbCr)	$hvar$	1
f2	RM map [10]	Weighted sum (YCbCr)	$hvar$	1
f3	RM map [10]	Weighted sum (YCbCr)	$hvar$	1
f4	AGGD [14]	Grayscale	Variance	1
f5	GM map [10]	Grayscale	$hvar$	1
f6	RM map [10]	Grayscale	$hvar$	1
f7	RM map [10]	Grayscale	$hvar$	1
f8–f12	LOG conditioned on Grad [27]	Grayscale	Independency distribution [27]	5
f13–f18	SURF (64), filtered image [16]	Grayscale	$s, m, k, hvar, sd, e$	6
f19–f24	SURF (128), filtered image	Grayscale	$s, m, k, hvar, sd, e$	6
f25–f30	SURF (64), image [16]	Grayscale	$s, m, k, hvar, sd, e$	6
f31–f36	SURF (128), image	Grayscale	$s, m, k, hvar, sd, e$	6
f37–f41	FREAK (512), filtered image	Grayscale	$s, m, k, hvar, sd, e$	5
f42–46	FREAK (512), image	Grayscale	$s, m, k, hvar, sd, e$	5
f47	Image	Grayscale	m	1
f48	Image	Grayscale	sd	1
f49	Image	Grayscale	Image sharpness	1
f50	Image	RGB	Image brightness	1
f51–f52	Image	YCbCr	Dark/bright pixels	2
f53–f71	Image	Binary image	Invariants and moments	19
f72–f73	Image	YCbCr	$m, median(k)$	2
f74–f79	Image	HSV, LAB	$m, median(S, V, L)$	6
f80	Image	YCbCr	Image sharpness (Y)	1

The numbers of 56 quality-aware features introduced in this paper are written in bold

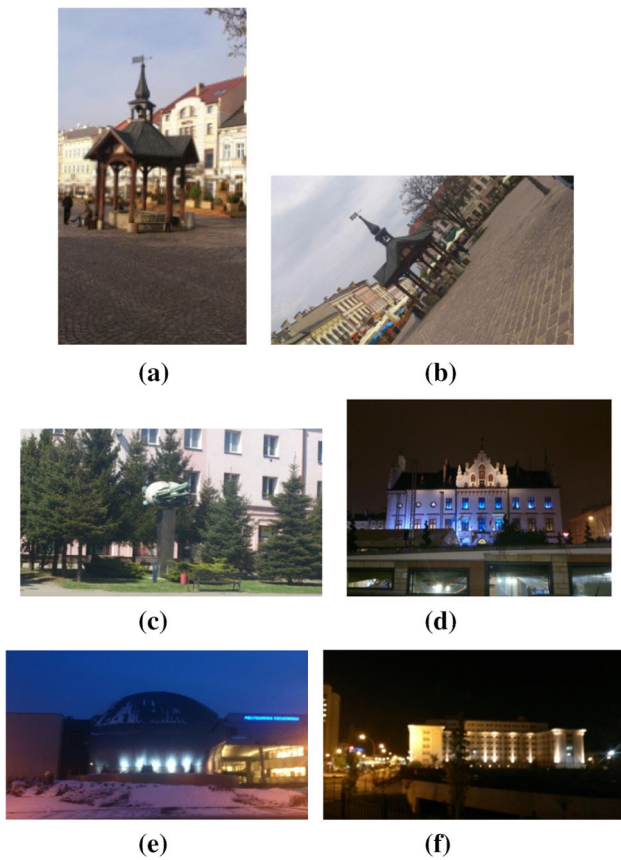**Fig. 2** Exemplary images and their corresponding MOS: **a** 1.0, **b** 0.7500, **c** 0.6667, **d** 0.4167, **e** 0.1667, **f** 0.0

Fig. 3 Values of exemplary features for images of a different distortion severity shown in Fig. 2; (Fig. 2a is denoted as '1,' Fig. 2b as '2,' ..., Fig. 2f as '6'

Fig. 4 Block diagram of the introduced method

redundant. Interestingly, it is significantly shorter than the vector extracted by FRIQUEE (828 values).

The introduced NR-IQA technique is based on a set of perceptual features whose usability for the assessment of authentically distorted images may require investigation. Therefore, several features and their values for images that were differently assessed by human subjects are shown in Figs. 2 and 3. Here, the severity of distortions, expressed by the values of subjective scores, is reflected by the visible difference between values of the presented features (Fig. 3). As shown in Fig. 4, SVR maps the feature vector into mean subjective scores and produces a quality model used for the prediction. Here, the SVR technique with the radial basis function (RBF) is employed using the popular LIBSVM library.

To show that all used features are important, a result of an experiment in which a given feature is removed from the feature vector is also presented. In the experiment, a typical protocol for the evaluation of NR-IQA measures is employed in which the database is randomly divided 100 times into the learning and training subsets (split 80/20%) and the median of the Spearman Rank Correlation Coefficient (SRCC) is reported [16,26,27,30]. As presented in Fig. 5, all features are
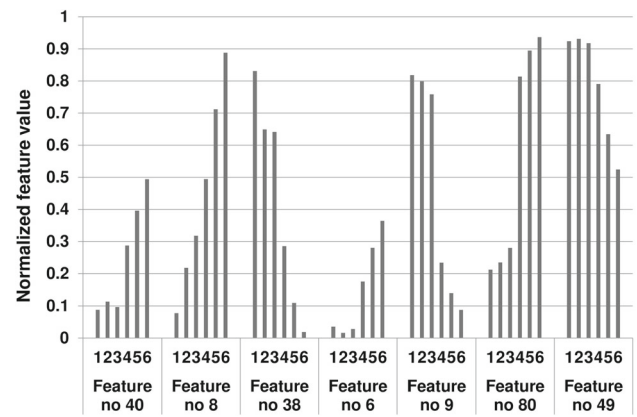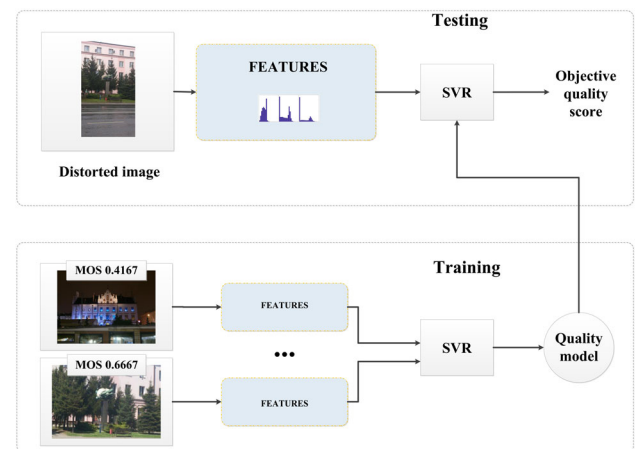
important. However, the features 53 and 54 (the introduced group of moments and invariants) are the most contributing to the performance of QUEADI, followed by the 2nd and 74th feature (the histogram variance of the relative gradient magnitude and the mean saturation). The 14 out of 20 most influential features are introduced in this paper. Their separate IQA performance is shown in Fig. 6, in which features are sorted from the most to the least influential, taking into account the results shown in Fig. 5. Interestingly, features that have the largest impact on the QUEADI do not offer outstanding separate performance. This indicates that many of them are complimentary and provide quality-aware information that is effectively used by the quality model. Therefore, such weaker features should not be avoided as they contribute to the overall quality prediction.

Since the method uses two modifications of features obtained from SURF descriptors, their separate IQA performance is reported in Fig. 7. It can be seen that the description of more detailed image regions in filtered images
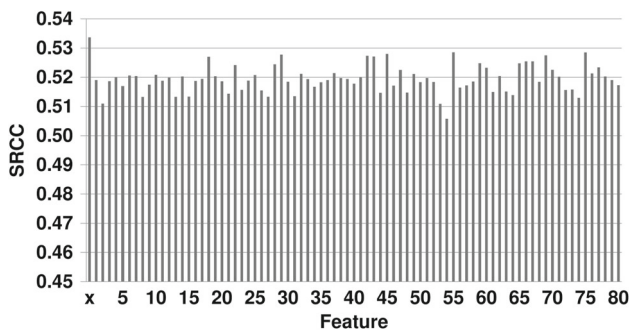
Fig. 5 Performance of the method in a case in which a given feature is removed from the feature vector. The result for the entire feature vector is denoted by '*x*'
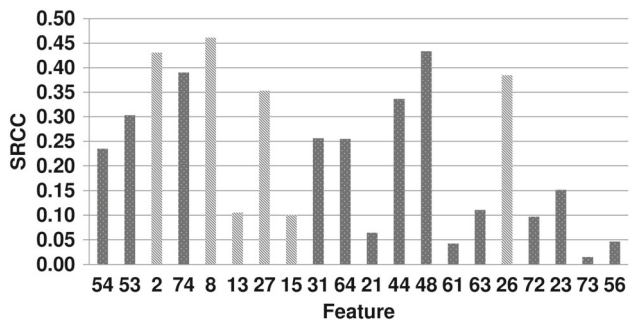

Fig. 6 Separate SRCC performance of 20 most influential features ranked in order of importance (from the left to the right). The bars with the dark pattern denote features introduced in this paper

(features f19–f24) better reflects the quality of images than its original 64-dimensional version (features f13–f18). However, the lack of filtering considered in features f25–f36 is more important than it can be seen for images with artificial distortions [16]. Therefore, the 128-dimensional SURF versions (f31–f36), proposed in this paper, obtained worse results than shorter descriptors (f25–f30), but they are still better than 64-dimensional descriptors on filtered images. The figure also contains the SRCC performance of QUEADI with features introduced in this paper (f19–f24, f31–f80) to show that the quality model with only introduced features offers a promising performance in comparison with related approaches reported in Sect. 4 (Table 3).

The SRCC performance of the method in relation to the number of the PCA components is reported in Fig. 8. Here, the 13 components seem to be a good choice (SRCC equal to 0.5337); however, in many observed cases, the application of the PCA for dimensionality reduction seems reasonable. Without the PCA, the SRCC of the method is 0.5205.

4 Comparative evaluation

The performance of QUEADI and other NR-IQA methods is evaluated using a typical methodology [16,26,27,30]
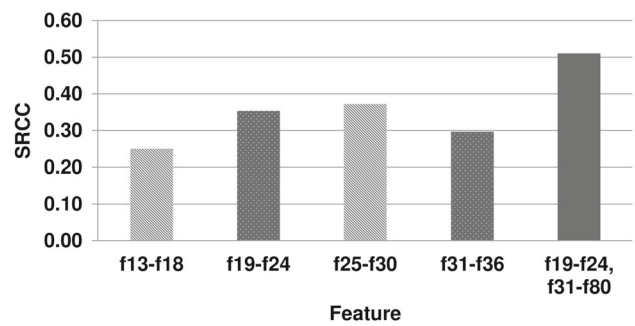

Fig. 7 SRCC performance of features based on SURF descriptors of different dimensionality obtained for original and filtered images. The result for all novel features is shown in the last bar. The bars with the dark pattern denote features introduced in this paper

Table 3 Performance evaluation on the BR image dataset

| NR measure	SRCC	KRCC	PCC	RMSE	Runtime (s)
HOSA	0.4860	0.3516	0.5017	0.2939	0.27
BRISQUE	0.4946	0.3595	0.5201	0.2916	0.10
FRIQUEE	0.4899	0.3558	0.5034	0.2930	44.9
GM-LOG	0.4974	0.3598	0.5139	0.2917	0.09
IL-NIQE	0.3869	0.2779	0.3991	0.3126	4.39
NIQMC	0.3000	0.2161	0.2955	0.3254	1.72
NOREQI	0.4988	0.3611	0.5169	0.2932	0.88
OG-IQA	0.4922	0.3586	0.4957	0.2952	0.07
dipIQ	0.0286	0.0210	0.1960	0.3344	1.77
MEON	0.2312	0.1664	0.2644	0.3274	0.11
QUEADI	**0.5337**	**0.3898**	**0.5460**	**0.2863**	**1.90**

The results for QUEADI are written in bold
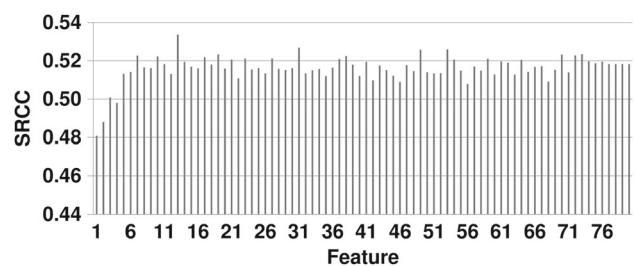
Fig. 8 Performance of the method with a different number of PCA components (from 1 to 80)

described in Sect. 3. However, apart from the SRCC, other criteria are also reported [22], such as Kendall Rank order Correlation Coefficient (KRCC), Pearson linear correlation Coefficient (PCC), and Root Mean Square Error (RMSE). They evaluate the prediction accuracy, monotonicity, and consistency of NR methods. The PCC and RMSE are calculated after a nonlinear mapping between the objective and subjective scores (Q_p and S): $Q_p = \beta_1 (0.5 - 1/(1 + \exp(\beta_2(Q - \beta_3)))) + \beta_4 Q + \beta_5$, where

$[\beta_1, \beta_2, \ldots, \beta_5]$ are parameters of the regression and Q and Q_p are the objective and obtained scores, respectively.

The results for the following NR methods with available source code are reported: FRIQUEE [4], NOREQI [16], OG-IQA [10], BRISQUE [14], GM-LOG [27], NIQMC [5], HOSA [26], IL-NIQE [30], dipIQ [12], and MEON [13]. The IL-NIQE and NIQMC do not require training images, and deep learning approaches (i.e., dipIQ and MEON) are already trained, and their source codes do not offer an opportunity to train them. Consequently, they are evaluated on the testing images, as the remaining methods. For a fair comparison, the SVR parameters of the learning-based methods are set to optimize their SRCC performance and all techniques assessed the same 100 subsets of testing images. The experiments are run on a machine with Intel Core i7-4790k 4.00 GHz, 16 GB RAM, Microsoft Windows 7 64bit, and MATLAB R2019a. The results are presented in Table 3. The table also contains the runtime comparison, reporting the average time of computing the feature vector by a given metric.

As reported, the proposed technique, QUEADI, clearly outperforms the compared methods on the new authentically distorted image database. The measure is better than the second-best technique (NOREQI) by a large margin. Other learning-based techniques perform closely to NOREQI, except for the deep learning methods. These methods were pretrained by their authors, and they cannot reliably predict the quality of images with authentic distortions. The proposed measure has an average time complexity; it is much faster than FRIQUEE and IL-NIQE and slower than worse performing OG-IQA, GM-LOG, or BRISQUE.

5 Conclusions

In this work, a novel IQA benchmark database with authentically distorted images and subjective scores is introduced along with a new NR-IQA measure (QUEADI). The QUEADI uses introduced diverse perceptual features and effectively maps them to subjective scores by the SVR to obtain a quality model. In the approach, global and local image features and their statistics are considered as well as image moments and invariants, or popular quality indicators with their modifications. The features are analyzed taking into account their impact on quality performance. The experimental comparison of the method with the state-of-the-art NR techniques on the introduced dataset reveals its superior performance, in terms of typically used evaluation criteria.

In the future, other features or statistics that can successfully describe the quality of authentically distorted images will be investigated.

The MATLAB code of the proposed QUEADI and the new IQA dataset is available at http://marosz.kia.prz.edu.pl/BR-QUEADI.html.

References

1. Alaei, A., Raveaux, R., Conte, D.: Image quality assessment based on regions of interest. Signal Image Video Process. **11**(4), 673–680 (2017). https://doi.org/10.1007/s11760-016-1009-z
2. Bosse, S., Maniry, D., Wiegand, T., Samek, W.: A deep neural network for image quality assessment. In: IEEE International Conference on Image Processing (ICIP), pp. 3773–3777 (2016). https://doi.org/10.1109/ICIP.2016.7533065
3. Chandler, D.M.: Seven challenges in image quality assessment: past, present, and future research. ISRN Signal Process. **2013**, 53 (2013). https://doi.org/10.1155/2013/905685
4. Ghadiyaram, D., Bovik, A.C.: Massive online crowdsourced study of subjective and objective picture quality. IEEE Trans. Image Process. **25**(1), 372–387 (2016). https://doi.org/10.1109/TIP.2015.2500021
5. Gu, K., Lin, W., Zhai, G., Yang, X., Zhang, W., Chen, C.W.: No-reference quality metric of contrast-distorted images based on information maximization. IEEE Trans. Cybern. **47**(12), 4559–4565 (2017). https://doi.org/10.1109/TCYB.2016.2575544
6. Kim, J., Lee, S.: Fully deep blind image quality predictor. IEEE J. Sel. Top. Signal **11**(1), 206–220 (2017). https://doi.org/10.1109/JSTSP.2016.2639328
7. Li, Q., Lin, W., Fang, Y.: No-reference image quality assessment based on high order derivatives. In: 2016 IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6 (2016). https://doi.org/10.1109/ICME.2016.7552997
8. Li, Q., Lin, W., Fang, Y.: No-reference quality assessment for multiply-distorted images in gradient domain. IEEE Signal Process. Lett. **23**(4), 541–545 (2016). https://doi.org/10.1109/LSP.2016.2537321
9. Li, Q., Lin, W., Xu, J., Fang, Y.: Blind image quality assessment using statistical structural and luminance features. IEEE Trans. Multimed. **18**(12), 2457–2469 (2016). https://doi.org/10.1109/TMM.2016.2601028
10. Liu, L., Hua, Y., Zhao, Q., Huang, H., Bovik, A.C.: Blind image quality assessment by relative gradient statistics and adaboosting neural network. Signal Process. Image **40**, 1–15 (2016). https://doi.org/10.1016/j.image.2015.10.005

11. Ma, K., Duanmu, Z., Wang, Z.: Geometric transformation invariant image quality assessment using convolutional neural networks. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 6732–6736. IEEE (2018)

12. Ma, K., Liu, W., Liu, T., Wang, Z., Tao, D.: dipIQ: blind image quality assessment by learning-to-rank discriminable image pairs. IEEE Trans. Image Process. **26**(8), 3951–3964 (2017). https://doi.org/10.1109/TIP.2017.2708503

13. Ma, K., Liu, W., Zhang, K., Duanmu, Z., Wang, Z., Zuo, W.: End-to-end blind image quality assessment using deep neural networks. IEEE Trans. Image Process. **27**(3), 1202–1213 (2018). https://doi.org/10.1109/TIP.2017.2774045

14. Mittal, A., Moorthy, A.K., Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE Trans. Image Process. **21**(12), 4695–4708 (2012). https://doi.org/10.1109/TIP.2012.2214050

15. Moorthy, A.K., Bovik, A.C.: Blind image quality assessment: from natural scene statistics to perceptual quality. IEEE Trans. Image Process. **20**(12), 3350–3364 (2011). https://doi.org/10.1109/TIP.2011.2147325

16. Oszust, M.: No-reference image quality assessment using image statistics and robust feature descriptors. IEEE Signal Process. Lett. **24**(11), 1656–1660 (2017). https://doi.org/10.1109/LSP.2017.2754539

17. Oszust, M.: Optimized filtering with binary descriptor for blind image quality assessment. IEEE Access **6**, 42917–42929 (2018). https://doi.org/10.1109/ACCESS.2018.2860127

18. Oszust, M., Padjasek, J., Kasprzyk, P.: An approach to vision-based localisation with binary features for partially sighted people. Signal Image Video Process. **11**(7), 1261–1269 (2017). https://doi.org/10.1007/s11760-017-1083-x

19. Ponomarenko, N., Jin, L., Ieremeiev, O., Lukin, V., Egiazarian, K., Astola, J., Vozel, B., Chehdi, K., Carli, M., Battisti, F., Kuo, C.C.J.: Image database TID2013: peculiarities results and perspectives. Signal Process. Image **30**, 57–77 (2015). https://doi.org/10.1016/j.image.2014.10.009

20. Ponomarenko, N., Lukin, V., Zelensky, A., Egiazarian, K., Carli, M., Battisti, F.: TID2008—a database for evaluation of full-reference visual quality assessment metrics. Adv. Mod. Radioelectron. **10**, 30–45 (2009)

21. Saad, M.A., Bovik, A.C., Charrier, C.: Blind image quality assessment: a natural scene statistics approach in the DCT domain. IEEE Trans. Image Process. **21**(8), 3339–3352 (2012). https://doi.org/10.1109/TIP.2012.2191563

22. Sheikh, H.R., Sabir, M.F., Bovik, A.C.: A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans. Image Process. **15**(11), 3440–3451 (2006). https://doi.org/10.1109/tip.2006.881959

23. Sun, W., Zhou, F., Liao, Q.: MDID: a multiply distorted image database for image quality assessment. Pattern Recognit. **61**, 153–168 (2017). https://doi.org/10.1016/j.patcog.2016.07.033

24. Video Quality Experts Group: Final report from the video quality experts group on the validation of objective models of video quality assessment, phase II (fr_tv2). https://www.itu.int/ITU-T/studygroups/com09/docs/tutorial_opavc.pdf (2003). Accessed 20 July 2019

25. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004). https://doi.org/10.1109/tip.2003.819861

26. Xu, J., Ye, P., Li, Q., Du, H., Liu, Y., Doermann, D.: Blind image quality assessment based on high order statistics aggregation. IEEE Trans. Image Process. **25**(9), 4444–4457 (2016). https://doi.org/10.1109/TIP.2016.2585880

27. Xue, W., Mou, X., Zhang, L., Bovik, A.C., Feng, X.: Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features. IEEE Trans. Image Process. **23**(11), 4850–4862 (2014). https://doi.org/10.1109/TIP.2014.2355716

28. Xue, W., Zhang, L., Mou, X.: Learning without human scores for blind image quality assessment. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition, pp. 995–1002 (2013). https://doi.org/10.1109/CVPR.2013.133

29. Zeng, H., Zhang, L., Bovik, A.C.: A probabilistic quality representation approach to deep blind image quality prediction. arXiv:1708.08190

30. Zhang, L., Zhang, L., Bovik, A.C.: A feature-enriched completely blind image quality evaluator. IEEE Trans. Image Process. **24**(8), 2579–2591 (2015). https://doi.org/10.1109/TIP.2015.2426416

31. Zhou, W., Yu, L., Zhou, Y., Qiu, W., Xiang, J., Zhai, Z.: Blind screen content image quality measurement based on sparse feature learning. Signal Image Video Process. **13**(3), 525–530 (2019). https://doi.org/10.1007/s11760-018-1378-6