

FNTF:First No-reference Then Full-reference Image Quality Assessment Using Dark Channel

Xiaoxin Lv, Min Qin, Xiaohui Chen, Xiaowei Qin

Abstract—It is an indispensable step to faithfully evaluate and control the perceptual quality of digital visual applications such as image compression, image restoration and multimedia streaming. In this paper, we propose an efficient and effective two-step framework named First No-reference Then Full-reference (FNTF), to evaluate different kinds of noise, and distinguish the quality of distorted image, using features Natural Scene Statistics of Dark Channel(NSSDC), Average Distance of Matched Keypoints(ADMK) and Dark Channel Similarity Deviation(DCSD) we proposed. Dark Channel is a kind of natural statistics based on the key observation - most local patches in images contain some pixels whose intensity are close to zero in at least one color channel. Features extracted from Dark Channel can greatly represent the pollution level of the image and the kind of noise it suffered from. The average distance and dark channel similarity are sensitive to image distortions, while different local structures in a distorted image suffer different distance and degrees of similarity, respectively. This motivated us to explore global variation based local quality for overall image quality prediction. We find that our two-step framework can predict accurately perceptual image quality.

Index Terms—image quality; objective quality; visual perception; dark channel

I. INTRODUCTION

With the widespread usage of communication applications and great development of intelligent mobile devices, a large number of digital images are being captured and shared every day. However, many of them suffered from distortions because of limited access devices, storage media, and transmission equipment. Distortions make it difficult for human to extract and understand the information in images. Therefore, it is highly desired to develop useful approaches that can predict image quality consistently with human subjective evaluation and it is important and necessary to develop practical objective image quality assessment(IQA) models. Depending on whether original reference images are available, IQA models can be classified into three categories: full-reference(FR), where the reference image is available, reduced-reference(RR), where partial information of the reference image is available, and no-reference(NR), where the reference image is not available.

As for FR-IQA models, they have a common two-step stage. First, compute a local quality map (LQM), comparing the distorted image with the reference image via some similarity local function. Then, using some pooling strategies, an overall quality score is computed from the LQM. The most widely

used pooling strategy is average pooling, i.e., taking the average of local quality values as the overall prediction score. In most cases, an original reference image does not exist, making FR-IQA and RR-IQA an impossible method to be practically embedded into such application systems. So only NR-IQA can be used in this situation. At present, proposed NR methods can be divided into two categories: (1) Distortion-specific quality assessment methods assess a specific distortion regardless of other factors and score a distorted image accordingly. (2) This general-purpose quality assessment methods based on training-learning perform much better than other approaches, which this paper use..

The rest of the paper is organized as follows. In Section 2, we introduced the two-stage framework FNTF for image quality assessment. In section 3, we introduce Dark Channel briefly. In Section 4, we describe the features extracted from a distorted image. In Section 5, we evaluate the performance of the proposed method and conclude the paper in Section 6.

II. OVERVIEW OF THE METHOD

How to evaluate an image? Given an image which is heavily polluted by salt and pepper noise (or Gaussian noise), we don't need to compare it with the original one, or even to examine the details, we are able to judge that the picture is unqualified. It means that when judging a picture, human first of all assess the picture from its overall information and make direct decision. But it's difficult to distinguish and assess the picture which is slightly distorted by noise. Now we have to compare it with the original picture. Meanwhile, we find that the statistical feature of a picture, i.e., NSSDC feature can greatly reflect the overall information, and the average distance of the matched keypoints and dark channel similarity derivation can reflect the details. Based on the above description, we proposed a two-step framework FNTF to evaluate a picture. First we evaluate the picture from its overall outlook to get noise category introduced into an image, then compare it with the original one, thus we can get more reference from the details to make the final grading. Following are the major steps of our method:

The first step(framework step one) of the IQA is to process incoming distorted images using NSSDC processing, enabling a statistical analysis in the whole image to judge the category of noise introduced into the image, and we find that NSSDC features have a very good performance. The meaning of this step is just like we human evaluate a distorted picture without original reference one, i.e. this is a NR-IQA process.

The second step of our method(framework step two) is to compute the ADMK and DCSD features, which can better

Xiaoxin Lv, Min Qin, Xiaohui Chen, Xiaowei Qin are with Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230027, China (email: xin2011@mail.ustc.edu.cn(X.Lv), qinminss@mail.ustc.edu.cn, cxh@ustc.edu.cn(X.Chen), qinxw@ustc.edu.cn(X.Qin)).

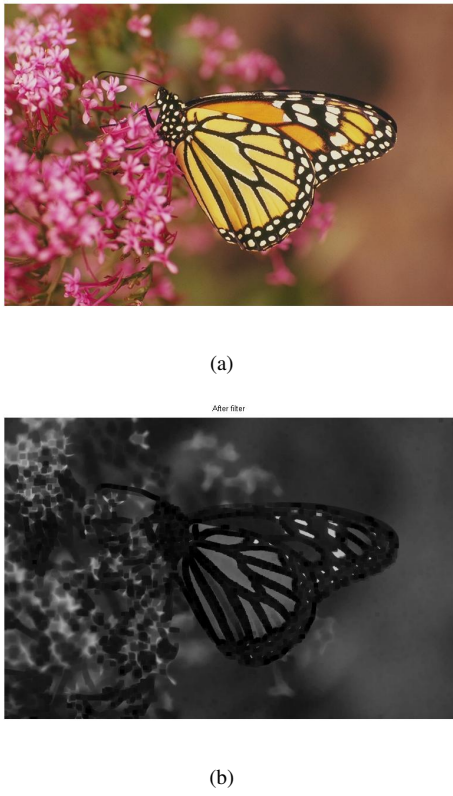


Fig. 1. (a): reference image. (b): Dark Channel of (a).

reflect the local information of the picture, the same as we compare the reference and distorted images and score the latter. It's obvious that this part is a FR-IQA process. The features talked above is covered in more detail later.

The final stage is prediction of the image quality score. we extract 4 features from a distorted image. So the feature vector of an image has 4 dimensions. Given the 4-dimensional feature vector, we can produce the final image quality score.

III. WHAT IS DARK CHANNEL

The dark channel prior [1] is based on the following important observation on almost every image: In most of the nonsky patches of images, there are some pixels whose intensity are very low and close to zero in at least one color channel. Equivalently, the minimum intensity of dark channel in such a patch is close to zero. To formally describe this observation, for an image J , its dark channel J^{dark} is given by

$$J^{dark}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in (r,g,b)} J^c(y) \right) \quad (1)$$

where J^c is a color channel of J and $\Omega(x)$ is a local patch centered at x . A dark channel is the result of two minimum operators: $\min_{c \in (r,g,b)}$ is conducted on each pixel, and $\min_{y \in \Omega(x)}$ is a minimum filter. The minimum operators are commutative, Fig. 1. shows the result image.

In our experiment, we find that not only the NSS features extracted from Dark Channel are very useful in IQA, but also it can reflect edge information of an image, this motivates us to utilize Dark Channel to improve our FNTF-IQA model.

IV. FEATURE EXTRACT

A. NSSDC

The first step of FNTF model is founded on perceptually relevant NSS features extracted from the Dark Channel of a given image(perhaps distorted image) that effectively capture the essential low order statistics of natural images. The NSSDC model that we use begins by preprocessing the Dark Channel by processes of local mean removal and divisive normalization:

$$J^{\hat{dark}}(i, j) = \frac{J^{dark}(i, j) - \mu(i, j)}{\sigma(i, j) + 1} \quad (2)$$

where $i \in 1, 2 \dots M$, $j \in 1, 2 \dots N$, M and N are the image dimensions, and

$$\mu(i, j) = \sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} J^{dark}(i+k, j+l) \quad (3)$$

$$\sigma(i, j) = \sqrt{\sum_{k=-K}^K \sum_{l=-L}^L w_{k,l} [J^{dark}(i+k, j+l) - \mu(i, j)]^2} \quad (4)$$

estimate the local mean and contrast, respectively, where

$\omega = v_{k,l} | k = -K, \dots, K, l = -L, \dots, L$ is a 2D circularly symmetric Gaussian weighting function sampled out to 3 standard deviations ($K = L = 3$) and rescaled to unit volume.

The coefficient $J^{\hat{dark}}$ has been observed to reliably tend towards a unit normal Gaussian characteristic when computed from dark channel of natural images that have suffered little or no apparent distortion. This model, however, is violated when the images do not derive from a natural scene(e.g. computer graphics) or when natural images are subjected to unnatural distortions. From our experiment it turns out that the generalized Gaussian distribution effectively captures the behavior of $J^{\hat{dark}}$ computed from natural and distorted versions of them. The generalized Gaussian distribution (GGD) with zero mean is given by:

$$f(x; \alpha, \sigma^2) = \frac{\alpha}{2\beta\Gamma(\frac{1}{\alpha})} \exp\left(-\left(\frac{|x|}{\beta}\right)^\alpha\right) \quad (5)$$

where $\beta = \sigma \sqrt{\frac{\Gamma(\frac{1}{\alpha})}{\Gamma(\frac{3}{\alpha})}}$, and $\Gamma(\bullet)$ is the gamma function:

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt \quad a > 0 \quad (6)$$

The parameters of the $GGD(\alpha, \sigma)$, can be reliably estimated using the moment-matching based approach proposed in [4]. The signs of the transformed $J^{\hat{dark}}$ have been observed to follow a fairly regular rules. However, this correlation structure will be destroyed by distortions. We compute NSSDC features on LIVE database. Fig. 2. show the distribution of $(\alpha, \log \sigma^2)$.

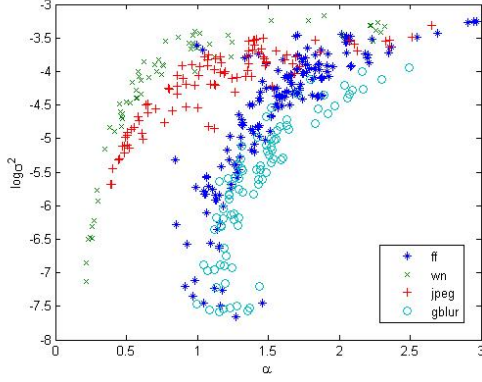


Fig. 2. the distribution of NSSDC feature tested on LIVE database.

B. ADMK

To better assess image quality compared with its reference image, we propose another robust feature named Average Distance of Matched Keypoints (ADMK), following are abstract of the major stages of computation used to generate ADMK feature¹:

1. *Fine keypoints*: The first stage of computation finds all keypoints of images.

2. *Compute SIFT descriptor*: After finding keypoints of an image, we can get SIFT descriptor of each keypoint, which is sensitive to local distortion.

3. *Average distance*: we first define the average distance of the matched keypoints (ADMK) as follow:

$$ADMK = \frac{1}{N} \sum_{i \in \text{matched keypoints}} dis(i) \quad (7)$$

where dis is the Euclidean distance of keypoints matched in two images.

From methods proposed before, it is easy to find that the features extracted from the overall image can not evaluate images very well, the reason is that details of an images should be consider much more than overall features in IQA model, the simplest detail features of an image are points and edges obviously, so this section and next section we introduce the point feature and edge feature. Based on our experiment, we find that Both of them are robust feature which can be used to evaluate the quality of image.

C. DCSD

The boundaries of the objects in images contains various information, which not only represent the outline of the object but are suitable for analysis in machine. Therefore boundary extraction becomes the main step in many applications, such as object separation, location of plate number, object detection and so on. Changes of the objects boundaries can well reflect the quality changes of an image, meanwhile boundary is also an important benchmark for human to judge an image. In this paper, we applied dark channel to represent the boundaries of

objects, the reasons are: (1) Normal operator can only detect the edge in one or several certain directions, such as Prewitt and Sobel operators, they are not suitable for processing the textures of an image, thus having less universality. (2) On one hand the dark channel extract the boundaries, on the other hand enhance the effect of boundaries, which is important for setting a better grading function and which can not be fulfilled by other operators. And follows are the process of feature extracting from Dark Channel.

Given reference and distorted images, the Dark Channel of ref and dis are computed as follows

$$dark_{ref}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in (r,g,b)} ref^c(y) \right) \quad (8)$$

$$dark_{dis}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in (r,g,b)} dis^c(y) \right) \quad (9)$$

With the Dark Channel images $dark_{ref}$ and $dark_{dis}$ in hand, the Dark Channel Similarity (DCS) map is computed as follows:

$$DCS(x) = \frac{2dark_{ref}(x)dark_{dis}(x) + c}{dark_{ref}^2(x) + dark_{dis}^2(x) + c} \quad (10)$$

Dark Channel Similarity Mean (DCSM):

$$DCSM = \frac{1}{N} \sum_{x=1}^N DCS(x) \quad (11)$$

and Dark Channel Similarity Deviation (DCSD):

$$DCSD = \sqrt{\frac{1}{N} \sum_{x=1}^N (DCS(x) - DCSM)^2} \quad (12)$$

where N is the total number of pixels in the image. The reason of using this feature is that we find that the pixel-wise Dark Channel Similarity (DCS) between the reference and distorted images combined with a novel pooling strategy the standard deviation of the DCS map can predict perceptual image quality accurately. The resulting DCSD algorithm is much faster than many other IQA methods, and performs highly competitive prediction accuracy.

V. EXPERIMENTS AND RESULTS

We tested our two-step IQA on the famous LIVE IQA database[5], which consists of 29 reference images and 779 distorted images that span various distortions, including JPEG and JPEG2000 compression, white noise(WN) and Gaussian blur(Blur). Each distorted images has an associated human difference mean opinion scores(DMOS), which represents the subjective quality of the image.

FNTF is a training based model, so the overall database needs to be partitioned into two sets: training set and testing set. Training set helps us train the classification and prediction models, and the testing set is used to evaluate the performance of our FNTF methods. In order to get the valid evaluation results, the process of database partition was performed with several rules. (1) the training and testing sets must be separated by content. (2) the training set consists of 80% of the reference images and corresponding distorted images, and the 20% reference image left and corresponding distorted

¹The first 2 steps are to compute matched keypoints, you can learn about SIFT to better understand this part.



(a)



(b)

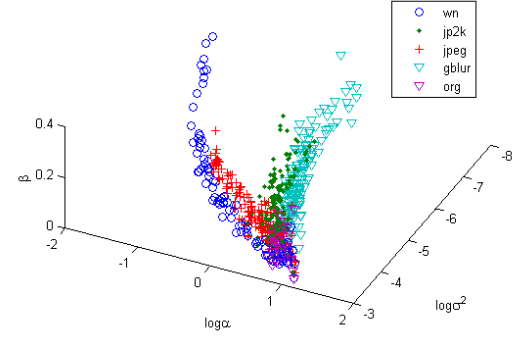
Fig. 3. (a): reference image.(b): dark channel image, we can see that dark channel emphasizes the boundary of object in an image.

images constituted the test set.(3) repeating 80% training-20% testing randomly 1000 times on the LIVE IQA database and assessed the performance each time under rule 1 and 2. The final algorithm performance was assessed by the median performance evaluation indices across the 1000 iterations.

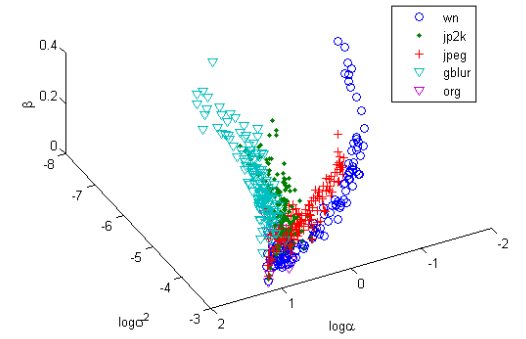
The performance evaluation indices consist of the Spearman's Rank Ordered Correlation Coefficient(SROCC), the Linear Correlation Coefficient(LCC) and the Root Mean Squared Error(RMSE) of the objective quality scores(our scores) and the subjective quality scores(DMOS). A value close to 0 for RMSE and a value close to 1 for SROCC and LCC indicate better correlation with human perception.

A. Correlation of each feature vector with human perception

We build the testing algorithm and conduct the performance evaluation correspondingly on each distortion category, as well as all distortion categories using each feature vector. The median experimental results on the FNTF model over 1000 iterations of random 80-20% train-tests using one feature is shown in tables 2-4, and table 1 shows the feature vector used in this part.



(a)



(b)

Fig. 4. (a): 3-D scatter plot between shape, scale parameters obtained from both NSSDC and DCSD. We conduct the classification on LIVE IQA database - JPEG 2000, JPEG, White Noise, and Gaussian Blur were introduced in each image at different degrees of severity (b): 3-D scatter plot between shape, scale parameters obtained from both NSSDC and DCSD from another visual angle.

TABLE I
FEATURES USED FOR FNTF IQA

Feature vector	Feature description
$f_{NSSDC}(f_1-f_2)$	coefficients of Natural Scene Statistics of Dark Channel(overall feature)
$f_{ADOMK}(f_3-f_5)$	Dark Channel Similarity Deviation,Average Distance Of Matched Keypoints and Keypoints ratio(local feature)

TABLE V
MEDIAN LCC ACROSS 1000 TRAIN TEST TRIALS ON THE LIVE-IQA DATABASE.

	JP2K	JPEG	NOISE	BLUR	FF	ALL
SSEQ	0.9350	0.9602	0.9572	0.9612	0.9240	0.9375
BRISQUE	0.9560	0.9451	0.9689	0.9579	0.9463	0.9484
SSIM	0.9714	0.9741	0.9524	0.9362	0.9566	0.9487
CURELET-IQA	0.9432	0.9305	0.9581	0.9539	0.9211	0.9328
FNTF	0.9736	0.9635	0.9775	0.9760	0.9713	0.9650

TABLE II
MEDIAN SROCC, LCC AND RMSE ACROSS 1000 TRAIN-TEST TRIALS
USING ONLY OVERALL FEATURES.

	SROCC	PCC	RMSE
JP2K	0.5844	0.6179	18.9818
JPEG	0.7515	0.7687	15.4724
NOISE	0.8746	0.8762	10.6117
BLUR	0.8367	0.8328	12.0778
FF	0.5550	0.6503	16.7781
ALL	0.6970	0.7026	16.4484

TABLE III
MEDIAN SROCC, LCC AND RMSE ACROSS 1000 TRAIN-TEST TRIALS
USING ONLY LOCAL FEATURES.

	SROCC	PCC	RMSE
JP2K	0.9512	0.9661	6.2837
JPEG	0.9176	0.9574	6.9519
NOISE	0.9299	0.9561	6.4379
BLUR	0.9439	0.9714	5.1837
FF	0.9527	0.9679	5.5162
ALL	0.9345	0.9539	6.9211

TABLE VI
MEDIAN RMSE ACROSS 1000 TRAIN TEST TRIALS ON THE LIVE-IQA
DATABASE.

	JP2K	JPEG	NOISE	BLUR	FF	ALL
SSEQ	8.5591	6.7403	6.3682	6.0324	8.4145	8.0236
BRISQUE	7.0791	7.8991	5.4452	6.2748	7.0753	7.3125
SSIM	5.7841	5.4649	6.6902	7.6692	6.4411	7.3102
CURELET-IQA	8.0375	8.9256	6.3067	6.5120	8.5857	8.3912
FNTF	5.5470	6.4646	4.6376	4.7657	5.2156	6.0507

TABLE VII
STANDARD DEVIATION OF SROCC, LCC AND RMSE ACROSS 1000
TRAIN-TEST TRIALS ON THE LIVE-IQA DATABASE.

	SROCC STD	PCC STD	RMSE STD
SSEQ	0.0185	0.0211	1.2315
BRISQUE	0.0083	0.0107	0.7879
SSIM	0.0074	0.0064	0.4739
CURELET-IQA	0.0312	0.0346	1.7099
FNTF	0.0070	0.0054	0.4976

From tables 1-3, we can conclude that the local features performed better than the overall features. However, the overall features are collectively a valuable and result-enhancing to local features, especially in distortion categories classification. Further, we also find that our method performs better for all kinds of noise considered, which is a common case of many other methods, such as SSEQ performs very poorly with FF, BRISQUE in JPEG, and curelet-IQA behaves badly with both JPEG and FF.

B. Comparison with other IQA approaches

We compared FNTF methods with other FR approaches(SSIM[6]) and NR-IQA methods(SSEQ[3], BRISQUE[4] and CURELET-IQA[2]), The results of this part are shown in Table 4-6. In order to make a valid comparison, we did a same random 80-20% train test for 1000 times to compute the median performance indices. The standard deviation(STD) of the indices were calculated to show the

TABLE IV
MEDIAN SROCC ACROSS 1000 TRAIN TEST TRIALS ON THE LIVE-IQA
DATABASE.

	JP2K	JPEG	NOISE	BLUR	FF	ALL
SSEQ	0.9256	0.9392	0.9607	0.9513	0.9109	0.9343
BRISQUE	0.9482	0.9307	0.9808	0.9615	0.9306	0.9517
SSIM	0.9700	0.9610	0.9217	0.9426	0.9392	0.9355
CURELET-IQA	0.9285	0.9087	0.9717	0.9465	0.9045	0.9293
FNTF	0.9562	0.9341	0.9647	0.9543	0.9567	0.9500

algorithm stability in table 7, and a higher LCC and SROCC with a lower STD mean excellent performance. Also, a lower RMSE indicates better performance.

C. Noise Classification

We also test our features on LIVE database in order to analysis noise classification performance. Noise Classification is very useful in the field of IQA, knowing the species of noise distorted a image, difficulty of quality assessment will be lowed by greatly. Table 8 shows the result of classification.

TABLE VIII
MEAN CONFUSION MATRIX FOR DISTORTION CLASSIFIER ACROSS 1000
TRAIN-TEST TRIALS.

	JP2K	JPEG	NOISE	BLUR	FF
JP2K	0.7002	0.0776	0.0749	0.0012	0.1461
JPEG	0.1516	0.7223	0.0688	0.0006	0.0567
NOISE	0.0366	0.0582	0.7144	0.0048	0.0098
BLUR	0	0	0.0189	0.9808	0.0003
FF	0.2661	0.746	0.1231	0.0128	0.5234

From the result, we find that blur is hard to be confused with other distortion categories while the FF is the most likely one to be confused, especially by JP2K, but it is easy to find FF is JP2K followed by wireless packet loss, and exhibited blur, blocking and ringing.

D. Image Quality Assessment

Finally, we produce a score function to evaluate image quality using machine learning method. Fig.5. are some examples which show that our method correspond to the virtual perception, compared with LIVE database DMOS.

VI. CONCLUSION

We proposed an efficient general-purpose image quality assessment (IQA) framework called First NR(No-Reference) Then FR(Full-Reference) (FNTF), which utilizes NSSD-C(Natural Scene Statistics of Dark Channel) to get overall information of tested image, then extracts Average distance of matched keypoints(ADMK) and Dark Channel Similarity Deviation (DCSD)features to respect the local detail information. From those features we can predict the image quality accurately. Now we are testing and proving that FNTF delivers quality prediction performance that is competitive with top-performing FR and NR IQA models. Recently we are working towards creating effective and speed-up NR image quality prediction models.

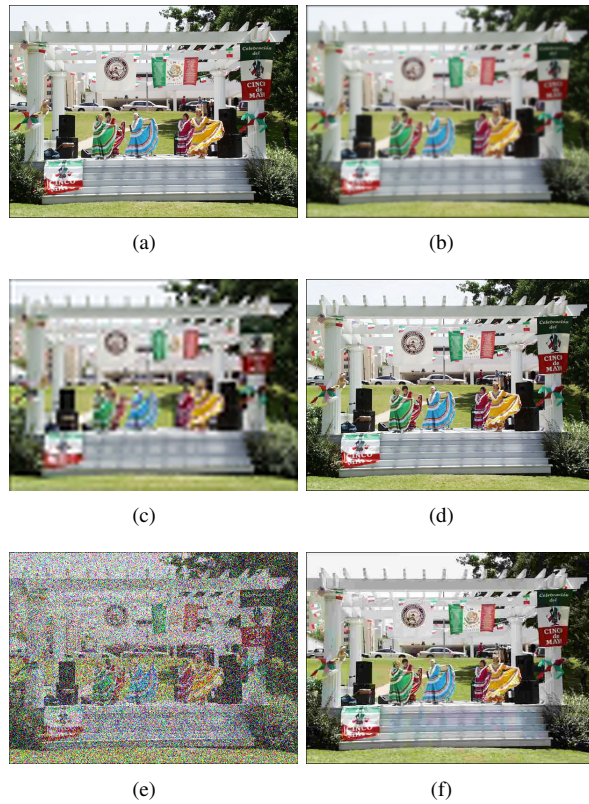


Fig. 5. result of FNTF.

- (a): reference image, DMOS=0, our score=9.8831.
 (b): gaussian blur image, DMOS=38.3902, our score=37.4533.
 (c): ff image, DMOS=63.2545, our score=67.2231.
 (d): jp2k image, DMOS=25.9001, our score=18.8920.
 (e): wn image, DMOS=65.4286, our score=57.4628.
 (f): jpeg image, DMOS=54.0497, our score=60.3491

VII. ACKNOWLEDGEMENTS

This work is supported by the National Basic Research Program of China (973 Program) granted No.2013CB329004.

REFERENCES

- [1] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(12):2341–2353, 2011.
- [2] Lixiong Liu, Hongping Dong, Hua Huang, and Alan C Bovik. No-reference image quality assessment in curvelet domain. *Signal Processing: Image Communication*, 29(4):494–505, 2014.
- [3] Lixiong Liu, Bao Liu, Hua Huang, and Alan Conrad Bovik. No-reference image quality assessment based on spatial and spectral entropies. *Signal Processing: Image Communication*, 29(8):856–863, 2014.
- [4] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. No-reference image quality assessment in the spatial domain. *Image Processing, IEEE Transactions on*, 21(12):4695–4708, 2012.
- [5] Hamid R Sheikh, Zhou Wang, Lawrence Cormack, and Alan C Bovik. Live image quality assessment database release 2, 2005.
- [6] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13(4):600–612, 2004.