


A Hybrid Quality Metric for Non-Integer Image Interpolation

Jinling Chen*, Yiwen Xu*, Kede Ma[†], Huiwen Huang*, and Tiesong Zhao* 

*College of Physics and Information Engineering, Fuzhou University, E-mail: t.zhao@fzu.edu.cn

[†]Center for Neural Science and Howard Hughes Medical Institute, New York University

Abstract—A great need of High-Resolution (HR) images has boosted the development of interpolation techniques. However, it is still a challenging task to objectively evaluate the perceptual quality of interpolated images, especially when the interpolation factor is a non-integer. To address this issue, we propose a hybrid quality metric for non-integer image interpolation that combines both reduced-reference and no-reference philosophies. To validate the proposed metric, we construct a non-integer interpolated image database and conduct a subjective user study to collect subjective opinions for each image. Experiments on the new database show that the proposed metric outperforms previous methods by a large margin.

Index Terms—Image quality assessment, image interpolation, high-resolution images, perceptual image processing.

I. INTRODUCTION

With the rapid development of High-Resolution (HR) displays, there has been a great desire for image interpolation techniques that generate HR images from Low-Resolution (LR) ones. The past decade has witnessed a growing number of image interpolation algorithms [1]–[7], but how to objectively compare their interpolation performance is not well resolved. The state-of-the-art objective Image Quality Assessment (IQA) metrics can be broadly classified into three categories: Full-Reference (FR), Reduced-Reference (RR) and No-Reference (NR) methods. Although FR-IQA metrics seem to be natural choices for image interpolation, Yang *et al.* [8] demonstrated that existing FR metrics fail to predict the perceptual quality of interpolated images. In real-world scenarios where the ground-truth HR images do not exist, RR- and NR-IQA metrics are more practical yet more challenging to design. Yeganeh *et al.* [10] extracted three sets of Natural Scene Statistics (NSS) for interpolated images, but the interpolation factor is strictly limited to the power of two. Fang *et al.* [11] evaluated the energy and texture similarity between LR and HR images. Recently, Ma *et al.* [12] presented an NR quality metric by supervised learning on a large set of interpolated images without knowledge of the corresponding ground-truth HR images.

All the above-mentioned quality metrics are constrained to integer interpolation factors and their extensions to account for non-integer factors are highly non-trivial. In this paper,

This research is supported by the National Natural Science Foundation of China (Grant 61671152).

we develop a hybrid quality metric for non-integer image interpolation, which combines the philosophies of RR and NR methods. Specifically, we first establish a patch correspondence between LR and HR images, and evaluate the energy and frequency similarity of the HR image using the corresponding LR image as reduced reference. Moreover, we measure the sharpness of the HR image in a no-reference manner. The overall hybrid quality metric is constructed by a weighted sum of the above three measures. To demonstrate the superiority of the proposed method, we construct one of the first non-integer interpolated image databases and source the ground-truth human opinions for each image. Experimental results on the new database show that our metric works well in visual quality prediction of both integer and non-integer interpolated images, and significantly outperforms state-of-the-art methods.

II. PROPOSED METHOD

The flowchart of the proposed method is summarized in Fig. 1, which includes a combination of three quality indexes. To employ these indexes, image matching shall be performed in the first step. However, it is impossible to establish a pixel-to-pixel correspondence between LR and HR images in the case of non-integer interpolation. Instead, we introduce a patch correspondence that approximately matches the patches between LR and HR images.

A. Patch Correspondence

Since the shapes of image contents are usually preserved in all directions during the interpolation process, the locations of image patches can be linearly matched. Specifically, for an $M_H \times N_H$ HR image patch and its corresponding $M_L \times N_L$ LR image patch, we have

$$M_L = \left\lfloor \frac{M_H}{f} \right\rfloor \quad \text{and} \quad N_L = \left\lfloor \frac{N_H}{f} \right\rfloor, \quad (1)$$

where f represents the interpolation factor. Due to the non-integer nature of f , the two patches can only be roughly matched. In this work, we set $M_H = N_H = 256$. If the size of the HR image is less than 256×256 , we set the whole image as the HR patch.

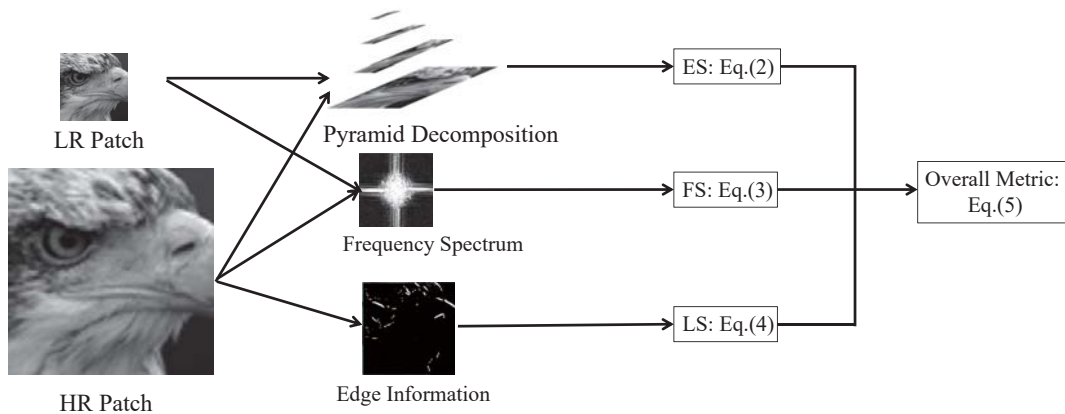


Fig. 1. The general flowchart of our approach.

B. Energy Similarity

It has been discovered that energy information is capable of capturing global image degradations [10], [11]. In addition, the energy falloff trends for natural images are approximately straight lines in logarithmic scale. These motivate us to measure global visual information degradations through the Energy Similarity (ES) between LR and HR patches. Specifically, we decompose the image patches into steerable pyramids [9]. For each patch, the energy at the j th scale $E(j)$ is calculated by the sum of squared coefficients. After that, we compute the energy difference $\Delta E(j)$ between adjacent scales. The ES from the energy falloff curves between LR and HR patches can be computed by

$$ES = \|\Delta \mathbf{E}_L - \Delta \mathbf{E}_H\|_2, \quad (2)$$

where $\Delta \mathbf{E}_L = [\Delta E(1)_L, \dots, \Delta E(s-1)_L]^T$ is the energy difference vector of the LR patch and s is the number of scales. $\Delta \mathbf{E}_H$ is defined similarly. ES is normalized to $[0, 1]$ over an image set.

C. Frequency Similarity

We evaluate the Frequency Similarity (FS) between LR and HR patches in Fourier domain. Specifically, we first pad the LR patch with zeros to match the size of the HR patch and use Two-Dimensional Discrete Fourier Transform (2D-DFT) to obtain the frequency spectra, based on which the FS is computed by

$$FS = \frac{1}{2} \left(\frac{\mathbf{H}_L^T \mathbf{H}_H}{\|\mathbf{H}_L\|_2 \|\mathbf{H}_H\|_2} + \frac{\mathbf{V}_L^T \mathbf{V}_H}{\|\mathbf{V}_L\|_2 \|\mathbf{V}_H\|_2} \right), \quad (3)$$

where \mathbf{H}_L , \mathbf{V}_L , \mathbf{H}_H , and \mathbf{V}_H represent the horizontal and vertical spectrum components of LR and HR patches, respectively. FS lies in $[0, 1]$ with a larger value indicating better performance.

D. Local Sharpness

In addition to the above RR-based features, we also incorporate the idea of NR measurement because humans are able to judge the HR image quality without access to its corresponding

TABLE I
PERFORMANCE EVALUATION ON THE PROPOSED NON-INTEGERS DATABASE

Image	WIND			Proposed		
	PLCC	KRCC	SRCC	PLCC	KRCC	SRCC
Beauty	0.87	-0.18	-0.14	0.62	0.50	0.67
Bongo	0.82	-0.27	-0.26	0.95	0.85	0.94
Buildings	0.86	-0.08	-0.08	0.91	0.72	0.84
Carrot	0.57	-0.16	-0.13	0.97	0.61	0.74
Eagle	0.72	-0.27	-0.27	0.89	0.67	0.77
Flowers	0.87	-0.11	-0.22	0.89	0.51	0.62
Fruit	0.52	-0.03	-0.04	0.89	0.66	0.75
Horse	0.55	0.36	0.42	0.97	0.90	0.96
House	0.86	-0.16	0.20	0.87	0.42	0.58
Human	0.89	0.05	0.02	0.90	0.32	0.45
Leaves	-0.06	-0.18	-0.18	0.86	0.81	0.86
Monarch	0.71	0.04	0.00	0.83	0.59	0.71
Rhinoceros	0.91	-0.16	-0.22	0.96	0.51	0.64
Teacher	0.60	-0.20	-0.16	0.94	0.82	0.92
Wheel	0.83	-0.24	-0.28	0.95	0.87	0.94
Average	0.70	-0.10	-0.11	0.89	0.65	0.76

ground-truth reference. A major problem of HR images is that they are blurrier than typical natural images, whose degree of blurriness increases with the interpolation factor. Here we exploit this fact and compute the Local Sharpness (LS) of the HR patch. We adopt the Sobel operator using two 3×3 kernels to calculate the horizontal and vertical derivative approximations, denoted by \mathbf{G}_x and \mathbf{G}_y , respectively, based on which the gradient magnitude at location (i, j) can be computed by $\mathbf{G}(i, j) = \sqrt{\mathbf{G}_x(i, j)^2 + \mathbf{G}_y(i, j)^2}$. Finally, the LS of the HR patch is calculated by

$$LS = \frac{1}{M_H N_H} \sum_{i,j} \mathbf{G}(i, j). \quad (4)$$

LS lies in $[0, 1]$ with a larger value indicating sharper edge information.

E. The Hybrid Metric

To construct an overall quality metric, we employ the simple average pooling to obtain the image-level ES, FS, and LS from

TABLE II
PERFORMANCE EVALUATION ON THE INTEGER DATABASE [10]

factor	WIND			Proposed		
	PLCC	KRCC	SRCC	PLCC	KRCC	SRCC
2	0.79	0.49	0.61	0.54	0.32	0.42
4	0.79	0.55	0.69	0.72	0.44	0.55
8	0.66	0.38	0.50	0.53	0.22	0.29
Average	0.75	0.47	0.60	0.60	0.33	0.42

patch-level measures and combine them using a weighted sum

$$Q = \alpha \cdot \text{ES} + \beta \cdot \text{FS} + \gamma \cdot \text{LS}, \quad (5)$$

where $\{\alpha, \beta, \gamma\}$ are parameters to be fitted with $0 < \alpha, \beta, \gamma < 1$ and $\alpha + \beta + \gamma = 1$.

III. EXPERIMENTAL RESULTS

To validate the proposed hybrid quality metric, we build a database of non-integer interpolated images and conduct a subjective user study to obtain the human opinions of image quality. Our database consists of 25 original HR images. We obtain LR images by bilinearly subsampling the HR images with three non-integer interpolation factors. The interpolated HR images are generated using a number of state-of-the-art interpolation algorithms [1]–[7]. In summary, there are a total 480 images in the database. During subjective testing, HR images of the same content are randomly displayed on the same screen together with their corresponding original image. Participants are asked to give a quality score between 0 to 10 to each HR image using the original image as reference. The Mean Opinion Score (MOS) of each HR image is computed by the average of the subjective scores over 22 valid participants.

We divide the proposed database into training and test sets, which contain 210 and 270 images, respectively, with no content overlap. We fit the weights in Eq. (5) on the training set, resulting in $\alpha = 0.18$, $\beta = 0.79$, and $\gamma = 0.03$. Table I shows the Pearson Linear Correlation Coefficient (PLCC), Spearman Rank-order Correlation Coefficient (SRCC), and Kendall Rank-order Correlation Coefficient (KRCC) results compared with WIND [10], a conventional integer-based quality model. From the table, we see that the proposed method achieves significantly better performance than WIND. This may be because WIND depends on the exact pixel-to-pixel correspondence. However, when a non-integer factor is utilized, only a handful of pixels are directly copied from the LR image to its HR version. Therefore, the pixel-level correspondence cannot be established, leading to the failure of WIND. By contrast, our hybrid quality metric bypasses the pixel-level correspondence and works well for non-integer interpolation. It is worth noting that our metric still works for integer interpolation, as shown in Table II, whose performance is close to that of WIND.

IV. CONCLUSION

We propose a hybrid quality metric to evaluate the perceptual quality of non-integer interpolated images. To the best of

our knowledge, it is the first attempt toward quality assessment for non-integer image interpolation. By dropping the pixel-to-pixel correspondence and combining RR and NR features, the proposed method works well for both integer and non-integer interpolation.

REFERENCES

- [1] X. Li, and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521-1527, 2001.
- [2] B.-D. Choi, and H. Yoo, "Design of piecewise weighted linear interpolation based on even-odd decomposition and its application to image resizing," *IEEE Trans. Consum. Electron.*, vol. 55, no. 4, pp. 2280-2286, 2009.
- [3] W. Dong, L. Zhang, R. Lukac, and G. Shi, "Sparse representation based image interpolation with nonlocal autoregressive modeling," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1382-1394, 2013.
- [4] X. Liu, D. Zhao, R. Xiong, S. Ma, W. Gao, and H. Sun, "Image interpolation via regularized local linear regression," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3455-3469, 2011.
- [5] L. Zhang, and X. Wu, "An edge-guided image interpolation algorithm via directional filtering and data fusion," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2226-2238, 2006.
- [6] Z. Wei, and K.-K. Ma, "Contrast-guided image interpolation," *IEEE Trans. Image Process.*, vol. 22, no. 11, pp. 4271-4285, 2013.
- [7] J. Lu, M. Xia, W. Li, W. Guo, and K. Yang, "An interpolation algorithm using center coordinates of pixels for converting SDTV images to HDTV images," *Optik*, vol. 124, no. 21, pp. 5251-5253, 2013.
- [8] C.-Y. Yang, C. Ma, and M.-H. Yang, "Single-image super-resolution: A benchmark," *Europ. Conf. Comput. Vis. (ECCV)*, 2014, pp. 372-386.
- [9] Eero P. Simoncelli and William T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," *IEEE International Conference on Image Processing (ICIP)* 1995, pp. 444-447.
- [10] H. Yeganeh, M. Rostami, and Z. Wang, "Objective quality assessment of interpolated natural images," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4651-4663, 2015.
- [11] Y. Fang, J. Liu, Y. Zhang, W. Lin, and Z. Guo, "Quality assessment for image super-resolution based on energy change and texture variation," *IEEE Int. Conf. Image Process. (ICIP)*, 2016, pp. 2057-2061.
- [12] C. Ma, C.-Y. Yang, X. Yang, and M.-H. Yang, "Learning a no-reference quality metric for single-image super-resolution," *Comput. Vis. Image Underst.*, vol. 158, pp. 1-16, 2016.