ASPECT RATIO SIMILARITY (ARS) FOR IMAGE RETARGETING QUALITY ASSESSMENT

Yabin Zhang[‡]*, Weisi Lin*[‡]*, Xinfeng Zhang*[‡] *Yuming Fang*[⋆] *Leida Li*[†]

[‡] School of Computer Engineering, Nanyang Technological University, Singapore
[⋆]School of Information Technology, Jiangxi University of Finance and Economics, China
[†]School of Information and Electrical Engineering, China University of Mining and Technology, China

ABSTRACT

During the past few years, there have been various kinds of content-aware image retargeting methods proposed for image resizing. However, the lack of effective objective retargeting quality metric limits the further development of image retargeting. Different from the traditional image quality assessment, the quality degradation of the retargeted images is mainly caused by the geometric changes due to retargeting. In this paper, we propose a practical approach to reveal the geometric changes during image retargeting, and design an Aspect Ratio Similarity (ARS) metric to predict the visual quality of the retargeted image. The experimental results on the widely used dataset show that the proposed metric outperforms the state of the arts.

***Index Terms*—** Image retargeting quality assessment, aspect ratio similarity

1. INTRODUCTION

Image retargeting studies can be categorized in two types: discrete and continuous approaches [1], based on whether they view the image as discrete pixels or the continuous signal. The typical discrete approaches include the manual Cropping (CR), Seam-Carving (SC) [2], Shift-Map(SM) [3] and Multi-Operator (MO) [4], while uniform Scaling in one dimension (SCL), non-homogeneous Warping (WARP) [5], Streaming Video (SV) [6] and Scale-and-Stretch (SNS) [7] are representative continuous methods. Both kinds of content-aware approaches remove pixels or warp the image to the targeted size according to the visual importance of image contents. The aim of the content-aware image retargeting is to preserve visually important contents and structures (i.e. to reduce information loss), and at the same time to limit the visual distortions in the retargeted images [1, 8, 9]. CR and SCL are two basic methods based on geometric constraints without considerations of image content. In retargeted images by CR, there is only information loss occurring, while in images produced by SCL, visual distortions due to squeezing or stretching degrade the image quality. Since the size reduction during retargeting is inevitable, most content-aware methods try to remove or shrink the less important contents, and thus achieve better overall performances by balancing information loss and visual distortions.

In traditional Image Quality Assessment (IQA) [10, 11, 12], the image grids are assumed to be intact and well-aligned, and the quality degradations are mostly intensity-related, while in Image Retargeting Quality Assessment (IRQA), intensity values are normally persevered in high quality but the image grids are reformulated by different retargeting methods. To measure the quality of the retargeted image, Rubinstein *et al.* [9] conducted a comprehensive study of different retargeting methods. In [13], Liu *et al.* explored the global structures and local correspondence to measure the visual quality. Fang *et al.* [14] exploited a Structural Similarity (SSIM) map to measure the quality of the preserved structural information in retargeted images. In [8], Hsu *et al.* obtained the dense correspondence between original and retargeted images for quality evaluation of retargeted images based on the variation of the flow field. The major drawback is that the removed or squeezed image contents in original image are allowed to be matched with pixels in the retargeted image, which is not necessary and also inhibit the estimation of real geometric change during image retargeting.

In this paper, the visual quality of the retargeted image is obtained based on the estimation and evaluation of the geometric change during image retargeting. We treat the retargeting process as the forward resampling from the original image and the proposed approach can effectively reveal the geometric change during the process of retargeting the original image to the retargeted one. Unlike the well-aligned images in IQA, the quality evaluation is based on the geometric relationship for IRQA. The geometric change is the major reason for quality degradation of the retargeted image. We utilize the Aspect Ratio (AR) change in the local region as an effective feature to measure the visual quality degradation. We design an Aspect Ratio Similarity metric (ARS) to measure the amount of geometric-related quality degradations with the assistance of visual importance map for perceptual quality of the retargeted image.

To summarize, our major contributions are a general geometric change estimation approach to reveal the geometric change during image retargeting and an effective ARS metric to predict the visual quality of the retargeted image.

The remainder of the paper is organized as follows. In section 2, we present the overview of the proposed image retargeting quality assessment metric. Section 3 introduces the geometric change estimation in details and Section 4 presents the ARS metric. Experimental results are reported in Section 5 and Section 6 concludes the paper.

2. OVERVIEW

Given the original and retargeted images in Fig. 1(a) and (c), we first find out the geometric change during image retargeting. We treat the image retargeting as a forward resampling process from the original image and the geometric change is revealed by the reverse estimation of the resampling locations for all the pixels from the retargeted image. The geometric change is visualized by distributing the pixels from the retargeted image to the original image as shown in Fig. 1(b). Based on the revealed geometric change, we can establish the local geometric change from the original image to the retargeted image, as the block pairs shown in Fig. 1(d). The Aspect Ratio Similarity (ARS) of the block pair indicates its information loss and visual distortion during the image retargeting process, where the block with high ARS is preserved in high quality as the lime route and the block with low ARS suffers serious information loss or visual distortion as the the red route. Most image retargeting methods try to hold the same aspect ratio of visual important regions and transfer the undesired and inevitable degradations to the visual unimportant regions. Therefore, with a visual importance pooling strategy, we can effectively predict the overall visual quality for the retargeted image.

3. GEOMETRIC CHANGE ESTIMATION

The major barrier that prevents us from assessing the quality degradations of retargeted images is the complicated geometric relationship between the original and retargeted images. We observed that almost all the quality degradation in the retargeted image is related to the geometric change. Here we treat the image retargeting process with the forward resampling and estimate the geometric change during image retargeting. Based on what we estimate, the visual quality of the retargeted image can be effectively evaluated.

3.1. Problem formulation

To obtain the substantial change during the image retargeting, it is necessary to have in-depth understanding of different retargeting methods. The discrete methods like SC generalize CR by choosing pixels for removal judiciously, while the continuous methods like SNS can be regarded as extensions of SCL. The continuous methods usually first project the retargeted image grid backward onto original image grid and then generate the retargeted image by resampling the original image via subpixel-level interpolations. The SCL is the simplest one where each resampling point location goes through the same scaling transformation along either the x or y axes, while other sophisticated non-homogeneous methods preserve visually important content through dense sampling and shrink or even remove other the visual unimportant content through sparse or zero sampling.

In the perspective of retargeting by sampling, the continuous methods are actually the generalizations of discrete methods by extending the pixel-level resampling to the subpixel-level resampling, and thus in a reverse way it is feasible to use the discrete methods to approximate the continuous methods by replacing their interpolation choices with nearest neighbour interpolations, which only introduce negligible intensity-related artifacts. It is difficult to establish the relationship with subpixel-level matching for continuous retargeting methods, since the computation increases explosively and we need to switch modes for different types. Therefore, we choose to depict both discrete and continuous methods as the resampling of the original image at pixel-level, so it is possible to reversely estimate the resampling locations in the original image for all the pixels from the retargeted image, thus reveal the undergone geometric change.

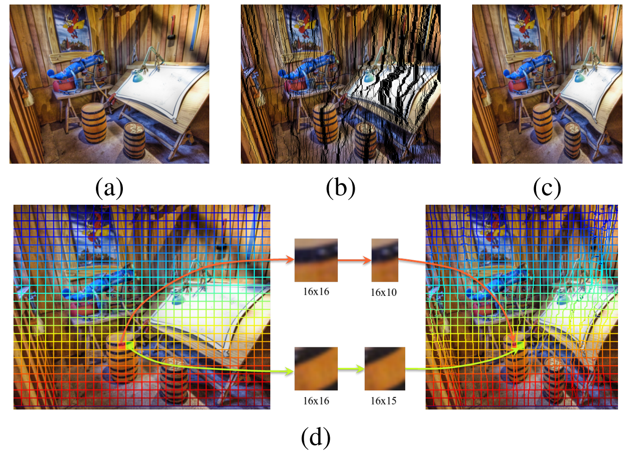

Fig. 1. (a) the original image; (b) visualized geometric change estimation results, where each pixel from the retargeted image is distributed onto its resampling location in the original image; (c) the retargeted image; (d) the geometric change from the regular blocks in the original image (left) to the retargeted blocks in the retargeted image (right), the block along the lime route is preserved in high quality while the block along the red route is squeezed seriously.

3.2. Proposed geometric change estimation approach

Given the original image I_{org} and the retargeted image I_{ret}, the geometric change estimation aims to find resampling location in I_{org} for each pixel at p_{ret} from I_{ret}. The matching energy function Eq. (1) contains *data* term and *smoothness* term, where the *data* term computes the likelihood that the pixel from I_{ret} is resampled at p_{org} in I_{org} and the *smoothness* term is used to constrain the flow field to be consistent. $f(p)$ denotes the feature for pixel at location p in I_{ret} or I_{org}. $p = (x, y)$ is the pixel coordinate in I_{ret} and $w(p) = (u(p), v(p))$ is the corresponding flow vector pointing to I_{org}, where $p_{org} = p + w(p)$ is the corresponding resampling location in I_{org}. ϵ denotes the four-connected neighbourhood of pixel p, and t and d are the thresholds for truncated L1 norms in the *data* term and *smoothness* term to tolerate the possible outliers and the discontinuities.

$$\begin{aligned} E(w) = \sum_{p} \min\left(\|f_{ret}(p) - f_{org}(p + w(p))\|_1, t\right) \\ \sum_{(p,q)\in\varepsilon} \min\left(\alpha\left|u(p) - u(q)\right|, d\right) \\ + \min\left(\alpha\left|v(p) - v(q)\right|, d\right) \end{aligned} \tag{1}$$

We adopt energy function minimization implementation in [15], which is a dual-layer loopy belief propagation based algorithm and utilize a coarse-to-fine scheme to speed up the optimization. The geometric change estimation results are shown in Fig. 2. The column (b) is the retargeted images and the column (c) is the visualized geometric change estimation results, where the geometric change is visualized by distributing the pixels from the retargeted image to the original image. It is obvious that we can estimate the cropping window ideally for CR and find out the uniform squeeze for SCL. We can effectively recover the removed seams for SC and estimate the seriously warped region for WARP. Generally, the proposed approach can effectively estimate the geometric change for images retargeted by different methods. The column (c) shows the retargeted blocks in the retargeted image, which are the important evidence for the measurement of information loss and visual distortion in the retargeted image.

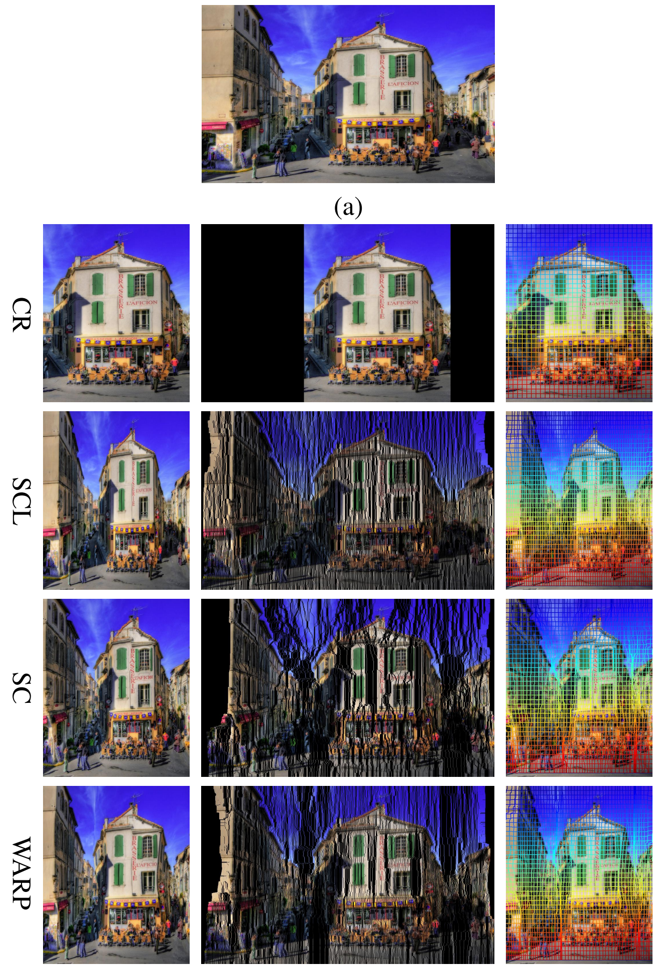

Fig. 2. (a) the original image (b) images retargeted by CR, SCL, SC and WARP; (c) the geometric change estimation results; (d) retargeted blocks in the retargeted image.

4. ASPECT RATIO SIMILARITY METRIC (ARS)

The aim and substantial change of image retargeting is the Aspect Ratio (AR) change for the whole image, and the global AR change is achieved by the local geometric change collectively across the image. The geometric change is the strong evidence about the information loss and visual distortion, where the resampling density is related to the information loss and the regularity of the forward resampling corresponds to the visual distortions. Therefore, the visual quality of the retargeted image can be assessed based on the measurement of the local geometric change. Here, we exploit the AR change of the local blocks as the feature to measure the local geometric change.

The block pair of the regular partitioned block in the original image and the corresponding block in retargeted image is utilized to calculate the local AR similarity scores. When the ARS score is close to 1, the block content in original image is generally kept in high quality in retargeted image, while when ARS score is close or equal to zero, it indicates that the retargeted block is suffering from serious information loss and distortion or even removed totally. As shown in Fig. 1(d), the original image is divided into 16×16 regular blocks and their corresponding retargeted blocks (more results are provided in Fig. 2(c)) are established in retargeted image based on the revealed geometric change, then we calculate the maximal width w_{ret} and maximal height h_{ret} of each retargeted block. The width and height ratio changes can be denoted as $R_w = w_{ret}/w_{ori}$ and $R_h = h_{ret}/h_{ori}$. The ARS score for each block can be formulated as follows:

$$S_{AR} = \frac{2 \cdot R_w \cdot R_h + C}{R_w^2 + R_h^2 + C} \cdot e^{-\alpha((R_w + R_h)/2 - 1)^2} \tag{2}$$

where C is a small positive constant to increase the stability, and α is the parameter of information loss penalty degree.

In the Eq. (2), the former term is the exact aspect ratio

Metric	mean KRCC for each Attribute						Total			
	Line Edges	*Faces People*	*Foreground Objects*	*Texture*	*Geometric Structure*	*Symmetry*	mean KRCC	std KRCC	LCC	*p*-val
SIFT flow	0.097	0.252	0.218	0.161	0.085	0.071	0.145	**0.262**	0.227	0.031
EMD	0.220	0.262	0.226	0.205	0.237	0.500	0.251	0.272	0.274	1e-5
IR-SSIM	0.309	0.452	0.377	0.321	0.313	0.333	0.363	0.271	0.439	1e-3
PGDIL	0.431	0.390	0.389	0.286	0.438	**0.523**	0.415	0.296	0.468	6e-10
Proposed ARS	**0.463**	**0.519**	**0.444**	**0.330**	**0.505**	0.464	**0.452**	0.283	**0.567**	**1e-11**

Table 1. Performance of different metrics on MIT RetargetMe dataset.

similarity between the original and retargeted blocks. Since the aspect ratio miss part of the absolute size changes, we also utilize the Gaussian function of the mean ratio to take account into the absolute block size change influence. To obtain the perceptual quality for the whole image, we utilize the saliency detection method [16] specifically designed for image retargeting as the visual importance map. The visual quality score for the retargeted image is defined as Eq. (3) by pooling the ARS of each block with the visual importance map.

$$ARS = \sum_m \sum_n S_{mn} \cdot V_{mn} \Bigg/ \sum_m \sum_n V_{mn} \tag{3}$$

where S_{mn} is S_{AR} for block (m, n), V_{mn} is the sum of the visual importance value for block (m, n), and (m, n) is the block coordinate in the original image.

5. EXPERIMENTAL RESULTS

To demonstrate the effectiveness of the proposed metric, we present rank correlation results of the proposed metric on the benchmark MIT RetargetMe dataset [17]. There are 37 images and their retargeted images are generated by eight typical methods including CR, SCL, SC [2], MO [4], SM [3], SNS [7], SV [6], and WARP [5] with either width or height dimension reduction by 25% (23P) or 50% (14P). There are six major image attributes provided for better insights: *Lines/Edges, Face/People, Foreground Objects, Texture, Geometric Structures* and *Symmetry*, and each image owns one or more attributes. The proposed metric ARS is compared with SIFT flow [15], EMD [18], IR-SSIM [14] and PGDIL [8]. We have used the same saliency detection method [16] as IR-SSIM and PGDIL. The correlations between the objective and subjective scores are measured by the Kendall Rank Correlation Coefficient (KRCC) [9]:

$$\text{KRCC} = \frac{n_c - n_d}{0.5n(n-1)} \tag{4}$$

where n is the length of the ranking (here $n = 8$), n_c is the number of concordant pairs and n_d is the number of discordant pairs from all the pairs.

We give the means and standard deviations of the rank correlations as well as the *p*-value and linear correlation-coefficient (LCC) in Table 1. The proposed ARS can obtain statistically better performance than the state-of-the-art methods. The major reason is that the revealed geometric change captures the major influence of image retargeting, so the evaluation based on the geometric change has obvious advantages over other dense correspondence based IRQA methods. The relative lower performance in *Symmetry* subset shows the limitation of the proposed metric in the measurement for the visual distortion of global structures.

6. CONCLUSION

In this paper, we have proposed an ARS metric based on the estimation and evaluation of the geometric change during image retargeting. We treated the image retargeting as the forward resampling from the original image and proposed a practical approach to estimate the undergone geometric change. We observed that almost all the quality degradation in the retargeted image is related to the geometric change. Therefore,with the revealed geometric change, we designed an ARS metric to effectively predict the perceptual quality of the retargeted image by exploiting the local block-based aspect ratio similarity scores with a visual importance pooling strategy. Compared to other state-of-the-art methods in the widely used dataset, the ARS metric yields statistically better results in the prediction accuracy.

7. ACKNOWLEDGEMENT

This research was carried out at the Rapid-Rich Object Search (ROSE) Lab at Nanyang Technological University, Singapore. The ROSE Lab is supported by the National Research Foundation, Singapore, under its Interactive Digital Media (IDM) Strategic Research Programme.

8. REFERENCES

[1] Ariel Shamir and Olga Sorkine, "Visual media retargeting," in *International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH ASIA 2009, Yokohama, Japan, December 16-19, 2009, Courses Proceedings*, 2009.

[2] Michael Rubinstein, Ariel Shamir, and Shai Avidan, "Improved seam carving for video retargeting," *ACM Trans. Graph.*, vol. 27, no. 3, 2008.

[3] Yael Pritch, Eitam Kav-Venaki, and Shmuel Peleg, "Shift-map image editing," in *IEEE 12th International Conference on Computer Vision, ICCV 2009, Kyoto, Japan, September 27 - October 4, 2009*, 2009, pp. 151–158.

[4] Michael Rubinstein, Ariel Shamir, and Shai Avidan, "Multi-operator media retargeting," *ACM Trans. Graph.*, vol. 28, no. 3, 2009.

[5] Lior Wolf, Moshe Guttmann, and Daniel Cohen-Or, "Non-homogeneous content-driven video-retargeting," in *IEEE 11th International Conference on Computer Vision, ICCV 2007, Rio de Janeiro, Brazil, October 14-20, 2007*, 2007, pp. 1–6.

[6] Philipp Krähenbühl, Manuel Lang, Alexander Hornung, and Markus H. Gross, "A system for retargeting of streaming video," *ACM Trans. Graph.*, vol. 28, no. 5, 2009.

[7] Yu-Shuen Wang, Chiew-Lan Tai, Olga Sorkine, and Tong-Yee Lee, "Optimized scale-and-stretch for image resizing," *ACM Trans. Graph.*, vol. 27, no. 5, pp. 118, 2008.

[8] Chih-Chung Hsu, Chia-Wen Lin, Yuming Fang, and Weisi Lin, "Objective quality assessment for image retargeting based on perceptual geometric distortion and information loss," *J. Sel. Topics Signal Processing*, vol. 8, no. 3, pp. 377–389, 2014.

[9] Michael Rubinstein, Diego Gutierrez, Olga Sorkine, and Ariel Shamir, "A comparative study of image retargeting," *ACM Trans. Graph.*, vol. 29, no. 6, pp. 160, 2010.

[10] Weisi Lin and C.-C. Jay Kuo, "Perceptual visual quality metrics: A survey," *J. Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.

[11] Zhou Wang and Alan C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures," *Signal Processing Magazine, IEEE*, vol. 26, no. 1, pp. 98–117, Jan 2009.

[12] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[13] Yong-Jin Liu, Xi Luo, Yuming Xuan, Wenfeng Chen, and Xiaolan Fu, "Image retargeting quality assessment," *Comput. Graph. Forum*, vol. 30, no. 2, pp. 583–592, 2011.

[14] Yuming Fang, Kai Zeng, Zhou Wang, Weisi Lin, Zhijun Fang, and Chia-Wen Lin, "Objective quality assessment for image retargeting based on structural similarity," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 4, no. 1, pp. 95–105, 2014.

[15] Ce Liu, Jenny Yuen, and Antonio Torralba, "SIFT flow: Dense correspondence across scenes and its applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 978–994, 2011.

[16] Yuming Fang, Zhenzhong Chen, Weisi Lin, and Chia-Wen Lin, "Saliency detection in the compressed domain for adaptive image retargeting," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 3888–3901, 2012.

[17] "RetargetMe Benchmark," [Online].http://people.csail.mit.edu/mrub/retargetme.

[18] Ofir Pele and Michael Werman, "Fast and robust earth mover's distances," in *IEEE 12th International Conference on Computer Vision, ICCV 2009, Kyoto, Japan, September 27 - October 4, 2009*, 2009, pp. 460–467.