Multi-operator Image Retargeting based on Automatic Quality Assessment

Zhongyan Qiu¹,³, Tongwei Ren¹,²,∗, Yan Liu⁴, Jia Bei¹,², Yang Yang²
¹State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China
²Software Institute, Nanjing University, Nanjing, China
³School of Software, Nanchang University, Nanchang, China
⁴Department of Computing, The Hong Kong Polytechnic University, Hong Kong, China
rentw@nju.edu.cn

Abstract—Image retargeting aims to avoid visual distortion while retaining important image content in resizing. However, no single image retargeting method can handle all cases. In this paper, we propose a novel multi-operator image retargeting approach, which utilizes an efficient and human perception based automatic quality assessment in operator selection. First, we calculate the importance map and distortion map for quality assessment. Then, we construct the resizing space and assess the performance of each operator in iterative width and/or height reduction. Finally, we select the optimal operator sequence by dynamic programming and generate the target image. Experiments demonstrate the effectiveness of the proposed approach.

Keywords—image retargeting; multi-operator retargeting; automatic quality assessment

I. INTRODUCTION

The prevalence of mobile devices requires image display on small screens with various aspect ratios. To provide good view experience, content-aware image resizing technique is highly demanded to retain important image content while reducing visual distortion. Such image resizing technique is usually referred to as image retargeting [1].

In the past decade, numerous of image retargeting methods have been proposed [2], [3], [4]. They usually calculate the energy map(s) to represent the attribute(s) of original images, and expand and/or shrink the low energy content to fit the required sizes and aspect ratios. These methods obtain good performance in generating target images from different original images. However, a comparative evaluation of current image retargeting methods shows that no single method can handle all cases, even they try to generate the best target images with experiential strategy or optimization [5]. Each single image retargeting operator may succeed on some original images but fail on others. To achieve better results, Rubinstein et al. combined multiple operators, including cropping, scaling and seam carving, and optimized the operator sequence by maximizing the similarity between original image and target image [6]. Multi-operator method effectively improves retargeting performance and generates good target images in most cases [5], but it has some drawbacks. First, to optimize operator sequence, multi-operator method uses bi-directional warping based similarity measurement between the original image and the intermediate result in each operator selection, which requires high time cost. For example, if horizontally resizing an image in size of \( w \times h \) with \( n \) operators, the time complexity of multi-operator method is \( O(w^3hn) \). Furthermore, bi-directional warping only considers the similarity of image representation but not real user experience in image viewing. It leads to low agreement with human perception and failure in some cases [5].

To address this problem, we propose a novel multi-operator image retargeting approach using human-like quality assessment. In operator selection, we provide an automatic quality assessment method instead of bi-directional warping, which requires low time cost and refers human perception based assessment criteria. Figure 1 shows an overview of our approach. First, we derive two assessment criteria from real user requirement and detect importance map and distortion map for quality assessment. Then, we construct the resizing space by attempting all the operators in each operation and automatically assess their performance. With the iterative image width and/or height reduction, we select the optimal operation sequence by dynamic programming and finally generate the target image.

II. RELATED WORK

A. Image Retargeting

To satisfy the requirement of image display on mobile devices, amounts of image retargeting methods are proposed in the past few year. Chen et al. proposed an automatic cropping method to cover most important objects [7]. Setlur et al. proposed a recomposition based method, which decomposes the original image into objects and background, scales them independently and recombines objects and background to generate the retargeted image [8]. Liu et al. proposed a fisheye-view warping method, which warps the original image with the pre-defined non-linear functions [9]. Avidan et al. proposed seam carving method, which iteratively removes the 8-connected pixels with the lowest energy [2]. Wolf et al. proposed non-homogenous warping method, which relocates image pixels by solving sparse linear system [10]. Wang et al. proposed scale-and-stretch method, which represents the original image with square meshes and adjusts the mesh vertices by quadratic programming [3]. Shi et al. proposed
a grid warping method, which represent the original image with rectangle grids and optimizes the warping of grids in the same row or column to accelerate resizing [11]. Pritch et al. proposed shift map method, which relocate image pixels by graph labeling and optimizes the relocation results by graph cut [4]. Huang et al. proposed a real time image retarget method, which parallelizes seam carving and implements it in real time [12]. Krähenbühl et al. proposed streaming video method, which warps image content under the constraints of object position and edge preservation [13]. Ren et al. proposed a constrained region warping method, which calculates multiple energy maps and constrain region warping with these maps [14].

B. Image Retargeting Assessment

With the increasing of image retargeting methods, quality assessment of image retargeting attracts much attention. Manual evaluation is widely used to obtain assessment results close to human perception. Rubinstein et al. comparatively studied of eight representative image retargeting methods [5] and provided a public dataset as a benchmark for further image retargeting assessment research [15]. Castillo et al. used eye-tracking data to evaluate different image retargeting methods [16].

Manual evaluation requires high labor cost and time cost, which is not suitable for large scale evaluation or rapid response applications. In recent years, some automatic quality assessment methods for image retargeting are proposed. Liu et al. used top-down manner to organize global and local features in assessment [17]. Ren et al. approximately computed the quality factors extracted from user perception [18] and further improved the criteria and assessment with salient region similarity measurement [19]. Hua et al. proposed a similarity measurement for image resizing, which uses SIFT feature in original image and target image matching [20]. Ma et al. built a large scale dataset and studied the existing subjective and objective quality assessment methods for image retargeting [21].

III. IMAGE RETARGETING USING QUALITY ASSESSMENT

A. Assessment Criteria and Energy Map

When view a target image, a user prefers to obtain the entire information for precise understanding and insensible artifacts for good experience. Hence, important content retainment (ICR) and visual artifact reduction (VAR) can reflect the real user requirement in retargeting and they are widely used as image retargeting assessment criteria [19].

To assess these two criteria, we detect the energy maps of the original image. For using multiple energy maps for different criteria can obtain better performance than a combined energy map [14], we utilize two energy maps, importance map $E_I$ and distortion map $E_D$, for important content retainment and visual artifact reduction assessment, respectively. In importance map detection, considering salient content usually attracts more attention of the viewers, we calculate the saliency of image content to represent its importance [22]. In distortion map detection, for higher gradient regions are more sensitive to distortion, we calculate the gradient of image content to represent its sensitivity of distortion.

B. Automatic Quality Assessment

With the criteria and energy map, we resize the original image by iteratively reducing its width and/or height. In the following, we only illustrate the procedure of reducing image width, and image height reduction can be handled similarly.

Assume the original image $I_O$ is in size of $w \times h$, and the target image $I_T$ is in size of $w' \times h$. In ICR assessment, we calculate the sum of importance energy of the retained pixels:

$$s_{ICR} = \sum_{p_{ij} \in I_O} R(i,j)E_I(i,j),$$

where $p_{ij}$ is the pixel in the coordinate of $(i,j)$; $R(i,j)$ denotes whether $p_{ij}$ is retained after resizing, which equals 1 if $p_{ij}$ is retained and 0 otherwise.

In VAR assessment, we utilize a method similar to [3]. We represent the original image with quad meshes and assess the distortion of image as the sum of the deformation assessment of each mesh. Considering only image width is reduced,
we simplify the deformation measurement of each mesh as follows:

$$\delta(k) = \sum_i \left( (x'_m - x'_n) - \eta(k)(x_m - x_n) \right)^2,$$

where \(x_m\) and \(x_n\) are the x coordinates of any two vertexes of the \(k\)th mesh in \(I_D\), \(x'_m\) and \(x'_n\) are the x coordinates of the corresponding vertexes in the resized image, and \(\eta(k)\) is calculated as follows:

$$\eta(k) = \frac{\sum_i (x'_m - x'_n)(x_m - x_n)}{\sum_i (x_m - x_n)^2}.$$

Hence, the score of VAR assessment is calculated as follows:

$$s^{VAR} = \sum_k E_D(k)\delta(k),$$

where \(E_D(k)\) is the mean of distortion energy of the pixels in the \(k\)th mesh.

Based on Equation (1) and (4), we calculate the overall score of image quality by integrating the assessment results in ICR and VAR:

$$s = \omega_{ICR} \cdot s_{ICR} + \omega_{VAR} \cdot s_{VAR},$$

where \(s\) is the total assessment result of image quality; \(\omega_{ICR}\) and \(\omega_{VAR}\) are the nonnegative weights of ICR and VAR respectively, and \(\omega_{ICR} + \omega_{VAR} = 1\).

C. Multi-operator Retargeting

In image retargeting, we combine multiple operators to reduce image width. Assume \(N\) kinds of operators are available, represented as \(\Phi = \{\phi_1, \phi_2, \ldots, \phi_N\}\), and each operator reduces one pixel of image width per time (in our experiments, we reduce five pixels per time to improve efficiency). Image retargeting can be formulated as a problem to find an operation sequence \(\Phi = \{\phi_1, \phi_2, \ldots, \phi_{\Delta w}\}\), here \(\Delta w = w - w'\). All the possible operation sequences construct an n-dimension resizing space, and each monotonic path in resizing space can generate a unique target image. To obtain good performance, we attempt all the operators in each operation, automatically assess the quality of generated intermediate results, and find the optimal operation sequence with dynamic programming.

We treat the original image as the initial result, and apply all the operator on it to reduce one pixel of its width. To the result generated by each operator, we assess its quality using Equation (5). Then, the width of each intermediate result is further reduced by all the operators, and the quality of generated results are assessed. Similar to the primary multi-operator method, we consider the ratio of operators are more important than their order in sequence. To the intermediate results generated by different operation sequence with same operator ratio, we only retain the result with highest assessment score. It simplifies the problem complexity from \(O(n^{\Delta w})\) to dynamic programing solvable.

In this way, we iteratively reduce image width and generate the target image in the required size.

Though image quality assessment is required for each intermediate result, it can be iteratively updated instead of completely recalculated. To the \(i\)th operation, ICR assessment result is only influenced by the removed pixels. So the score of ICR assessment can be calculated as follows:

$$s^i_{ICR} = s^{i-1}_{ICR} - \sum_{P_i} R(i, j)E_I(i, j),$$

where \(P_i\) is the set of the pixels removed in the \(i\)th operation.

In VAR assessment, the result is also only influenced with the meshes whose pixels are removed. To accelerate assessment, we store the VAR assessment result of each mesh, and update the VAR score as follows:

$$s^i_{VAR} = s^{i-1}_{VAR} + \sum_{M_i} E_D(k)(\delta(k)^i - \delta(k)^{i-1}),$$

where \(m_k\) is the \(k\)th mesh, and \(M_i\) is the set of the meshes whose pixels are removed in the \(i\)th operation.

Based on Equation (6) and (7), image quality assessment score is updated in each operation:

$$s^i = s^i_{ICR} + s^i_{VAR},$$

where \(s^i\) is the quality assessment score of the intermediate result after \(i\)th operation. The time complexity of quality assessment in our approach is only \(O(whn)\), which is much lower than bi-directional warping used in primary multi-operator retargeting.

With dynamic programming, we select operation sequence to generate the result with highest quality assessment score:

$$\Phi^* = \arg\max_{\Phi} \delta^w(\Phi),$$

where \(\delta^w(\Phi)\) is the quality assessment score of the generated result after \(\Delta w\) operations with the operation sequence \(\Phi\). Using the selected optimal operation sequence, we can generate the final target image.

IV. EXPERIMENTS

A. Dataset

We verify the proposed approach on RetargetMe dataset [15]. RetargetMe dataset contains 80 original images with various attributes. Each original image has all or partial of eleven target images generated by nonhomogeneous warping (WARP), seam carving (SC), scale-and-stretch (SNS), multi-operator (MULTIOP), shift-maps (SM), streaming video (SV), uniform scaling (SCL), manual cropping (CR), energy-based deformation (LG), quadratic programming (QP) and object size adjusted (OSA), respectively [15].

RetargetMe dataset also provides two versions of manual evaluation results, reference version and no-reference version, for the target images generated with eight methods of 37 original images. The difference of two manual evaluation...
versions is whether the original image was shown as the reference in evaluation. In each version, 210 participants were invited to vote the better target image in paired comparisons, and each target image could obtain up to 63 votes at most [15].

B. User Study

To illustrate the performance of our approach, we compare our approach with the existing image retargeting methods in three experiments. In the experiments, we invite 15 volunteers in age of 20 to 38 to manually evaluate the generated target images. These volunteers include eight males and seven females, with occupation of undergraduate, graduate, worker and official. During the evaluation, each target image generated by our approach is required to be pairwisely compared with all other compared target images by three volunteers, and the dominant judgment is treated as the final evaluation. Considering the users usually view the target images without the reference of original images in real applications, we only carry out no-reference evaluation in our experiments. To facilitate comparison, we retarget each original image to the same size of the target images provided in RetargetMe dataset.

In the first experiment, we compare the proposed approach with three operators combined in our approach. Similar to the primary multi-operator retargeting method, we combine scaling, cropping and seam carving in our approach. For 77 original images in RetargetMe dataset have the corresponding target images of these three methods, we use these 77 original images in the first experiment. Figure 2 illustrates some comparison examples, and Table I shows the result of user study, in which “Better”, “Similar” and “Worse” denote the number of original images whose results generated by our approach are better, similar and worse than/to the corresponding methods, respectively. It is found that our approach is outperformance than the existing methods. It shows that both the human-like quality assessment and multiple-operator combination play the effective roles in our approach to improve the retargeting performance.

We also compare the efficiency of our approach and the primary multi-operator method. We implement the primary multi-operator and our approach with Matlab, and retarget the 71 original image. It takes 14.6 minutes to retarget an image with the primary multi-operator method in average and only 57.3 seconds with our approach. The efficiency of our approach can be further improved by implementing our approach with C++ or other more efficient language.

To further demonstrate the performance of our approach, we compare our approach with other methods in the third experiments. To reduce manual evaluation labor, we use the 37 original images with manual evaluation results in RetargetMe dataset, and compare our approach with the top 3 target images with the highest votes for each original image. Figure 4 illustrates some comparison examples, and Table III shows the result of user study. It is found that our approach obtains very similar performance to the top 1 result, and slightly better than the top 2 result and the top 3 result. Considering the top 3 results are generated by different methods on different original images, it shows that our approach is outperformance than the existing methods.

C. Discussion

In experiment, we also find some limitations of our approach. For example, our approach is influenced by the accuracy of saliency detection. When the saliency detection is inaccurate, it will lead to poor performance. Figure 5(a) and (b) show an example of low quality target image generated with inaccurate importance map.

In addition, our approach depends on the performance of the combined operators. If all the combined operators cannot
handle the image well, our approach will also generate poor result. Figure 5(c) and (d) show a failure example of our approach when scaling, cropping and seam carving cannot handle the original image.

In the future, our work will focus on searching better operator combinations to handle different cases. We will also extend our approach and apply it in video retargeting.

V. CONCLUSION

This paper presents a novel multi-operator image retargeting method using automatic quality assessment. Compared to primary multi-operator method, the proposed approach utilizes an efficient and human perception based automatic quality assessment method, which obviously improves retargeting efficiency and effectiveness.

ACKNOWLEDGMENT

The authors want to thank the anonymous reviews for helpful suggestion. This paper is supported by Natural Science Foundation of China (61202320), Natural Science Foundation of Jiangsu Province (BK2012304), and National Undergraduate Innovation Program (201210284031).
Figure 4. Examples of comparison to top 3 results. (a) Original images. (b) The top 1 results. (c) The top 2 results. (d) The top 3 results. (e) Our results.

Figure 5. Examples of failure. (a) and (c) Original images. (b) and (d) Our results.

REFERENCES