

SALIENCY & STRUCTURE PRESERVING MULTI-OPERATOR IMAGE RETARGETING

Lingling Zhu, Zhibo Chen, Xiaoming Chen, Ning Liao

University of Science and Technology of China
CAS Key Laboratory of Technology in Geo-spatial Information Processing and Application System
Hefei 230027, China

ABSTRACT

Content-aware image retargeting has attracted substantial research interests in the related research community. However, so far there is still no method can preserve important image contents and structure well without introducing deformation. To address this problem, we propose a Saliency & Structure Preserving Multi-operator (SSPM) method. SSPM classifies images into three categories utilizing SIFT density to improve performance of saliency preservation, helping to mitigate negative influence from center-bias property of most existing saliency detection models. SSPM also employs different principles to improve structure preservation performance, including Earth Mover's Distance (EMD) and Gray-Level Co-occurrence Matrix (GLCM) to get optimal operator sequences for smart content-aware image retargeting. SSPM method not only can well preserve salient contents and structure, but also can greatly improve deformation resilience. Experimental results demonstrated that our method outperforms state-of-art image retargeting methods.

Index Terms— retargeting, saliency preservation, structure preservation, multi-operator

1. INTRODUCTION

Image retargeting is a technique that automatically adjusts input images into desired screen sizes while preserving important visual contents of the original image. Due to the diversity and versatility of display devices available today, image retargeting has drawn increasing research attention of the related research community. A well designed image retargeting algorithm would relief people from manually changing the aspect ratios of images in order to fit into different displays.

Although researchers have proposed quite some methods to automatically retarget images, it remains a challenging task. Usually single retargeting operators like scale [1], crop [2] and seam carving [3, 4] can't adapt to diverse image content types even by introducing saliency object information into the operators. For cropping, Cavalcanti added some image features in cropping algorithm, considering face detection and saliency to preserve contents that people may pay

attention to [2]. Cropping doesn't distort image and change image structure, but it would lead to poor results when salient objects are separated at the boundary of the image. For scaling, Zhu employed saliency to preserve important objects and scale much of the background. However, when it comes to complicated background, it will introduce disproportion of objects [1]. For seam carving, it performed well on smooth backgrounds [3], but it will introduce significant distortion when dealing with trees or backgrounds with complicated textures. Achanta introduced saliency in seam carving which can preserve important objects well [4]. However, when saliency detects unimportant areas, the method shows very poor performance. Wang presented mesh-based scale-and-stretch which employed gradient and saliency as its importance map [5]. This algorithm aims at reducing distortion at important areas, but it may result in distortion in other areas and may change the ratios of objects. We observed that retargeting methods using multiple operators [6, 7] such as scaling, cropping, seam carving, are a feasible approach to adapt to diverse image content. Rubinstein proposed multi-operator image retargeting method [6], which used Bi-Directional Warping (BDW) with a dynamic programming algorithm to find an optimal operator sequence in the resizing space formed by cropping, scaling and seam carving. However, BDW only considers the similarity of images but not human visual perception. Consequently, it may result in severe visual discomfort, such as disproportion and distortion of objects. Furthermore, methods [7] employed gradient and saliency as importance maps, but they may introduce deformation and destroy the structure of the image when the image background is complicated.

Thus, the key problem in multi-operator retargeting methods is to define the "importance map", which corresponds to the areas that draw human attention in images [8] and form the operator sequences as described in Section 2.1 according to image content property.

In this paper, we propose a novel importance map by combining three operators, including cropping, scaling and seam carving, to deal with varieties of images. There are two major contributions in this paper: First, due to the center-bias property of most existing out-standing saliency models, we introduce SIFT density features which help to classify

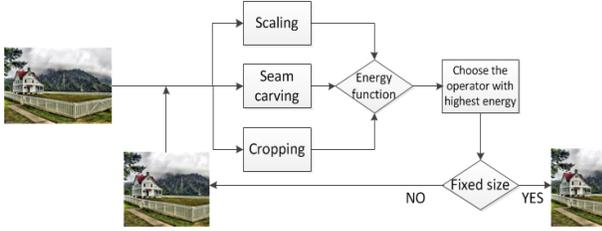


Fig. 1. Flow diagram of SSPM retargeting method.

images into three categories and improve the performance of saliency detection. Second, we use SIFT features to measure the preservation of image structure. For the matched SIFT features in the retargeted image and the original one, we use Earth Mover’s Distance (EMD) to measure the structural similarity. For unmatched SIFT features, we regard them as seriously destroyed structures, and use the ratio of them in our energy function to punish the operator bringing in such distortion. The proposed Saliency & Structure Preserving Multi-operator (SSPM) method not only can well preserve salient contents and structure, but also can significantly improve the deformation resilience.

In Section 2, we introduce SSPM retargeting method. We show our subjective experiment results in Section 3. The conclusion and future work are presented in Section 4.

2. SALIENCY & STRUCTURE PRESERVING MULTI-OPERATOR (SSPM) RETARGETING METHOD

We took advantages of three simple operators, including cropping (CR), scaling (SCL) and seam carving (SC), to resize images. The flow diagram of SSPM method is shown in Fig. 1. To get the optimal operator sequence, we define an effective energy function, as illustrated in Fig. 2, which aims to preserve as many salient objects as possible and the structure information. In particular, to improve the accuracy of saliency detection, we first employed SIFT features and classified images into three different categories: 1) saliency in both boundaries of the image; 2) saliency in one side of the image boundaries; and 3) saliency in the center of the image. Then we set different energy functions, such as saliency preservation and structure preservation, to get multi-operator sequences.

2.1. Retargeting

For a given image I of size (m, n) , retargeted image I' of size (m, n') . The multi-operator sequences are $O\{o_1, o_2, \dots, o_k\}$, $k = \frac{n-n'}{c}$. o_i is chosen from one of the three operators (scaling, cropping, seam carving), c represents the quantity each operator resizes the width of the image. Thus k represents resizing an image from (m, n)

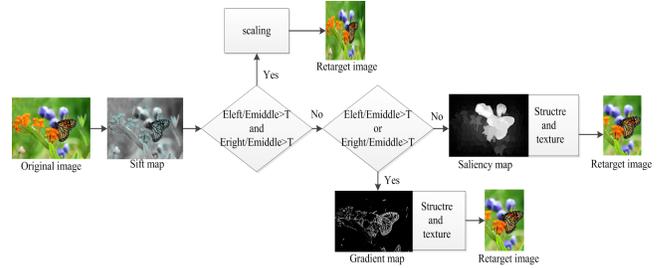


Fig. 2. The decision tree and definition of the energy function of SSPM retargeting method.

to (m, n') , it needs k operators. The ratios and orders of operators will definitely influence the quality of retargeted images.

In addition, we improved the operators of cropping and seam carving. For cropping, we choose the cropping window with the highest gradient value. For seam carving, Avidan [3] proposed to preserve image contents with high gradient values, which may lead to distortion in important contents. We enhanced this algorithm by introducing saliency map to avoid seams across salient objects. With these three operators, we can adapt to diverse image contents. However, how to define the energy function to get the optimal operator sequence remains unsolved. We introduce it in the next section.

2.2. Energy Function

2.2.1. Saliency Preservation

As mentioned earlier, how to measure the importance map is the key part in content-aware image retargeting. Traditional Saliency model [9] implies center-bias with the assumption that important objects are always in the center of images, which will cause problems in case of important objects distributed at the boundary of images.

To address this problem, we divide the image into three parts: 1) I_{left} represents 25% of the image on the left; 2) I_{middle} represents 50% of the image in the center; and 3) I_{right} represents 25% of the image on the right. Then we compute the density of SIFT features of the three parts in the image, since SIFT can detect key features in images, which are invariant to image translation, scaling, and rotation. We use E_{left} , E_{right} and E_{middle} to respectively represent the densities of SIFT features in different parts of the image. When E_{left} or $E_{right} / E_{middle} > T$ (T is the threshold of SIFT features), we assumed important objects are distributed in image boundary. In this paper, we experimentally set T to 1.5.

- when $E_{left} / E_{middle} > T$ and $E_{right} / E_{middle} > T$

We assume important objects are distributed in both boundaries of the image and just use scaling (SCL) to

resize images.

- when $E_{left}/E_{middle} > T$ xor $E_{right}/E_{middle} > T$

$$E_{gradient} = \sum_{i=1}^n \sum_{j=1}^m g_{ij} \quad (1)$$

g_{ij} is the gradient value of pixel (i, j) in image with size (m, n) .

We assume important objects are distributed in one side of the image boundaries and employ gradient as one of our principles instead of saliency to select operators.

- when $E_{left}/E_{middle} \leq T$ and $E_{right}/E_{middle} \leq T$

$$E_{saliency} = \sum_{i=1}^n \sum_{j=1}^m s_{ij} \quad (2)$$

s_{ij} is the saliency value of pixel (i, j) in image with size (m, n) .

We assume important objects are distributed in the middle of the image and employ saliency as one of our principles to select operators.

In this way, we can reduce the effect of saliency center-bias problem, which has been validated by our experimental results presented in Section 3.

2.2.2. Structure Preservation

In this paper, we employed Earth Mover' Distance (EMD) algorithm to preserve the internal structure of image [10]. Although saliency can preserve important contents, it may lose too much background information and destroy the structure of original images. Thus, it is important to balance important contents and structure information.

We first detect SIFT features in both original image and retargeted images, then we match these features. However, due to the change of content in retargeted images, some SIFT features cannot find the corresponding one, we consider those features as information loss, which represent severe destruction of the structure in the original image. The information loss is measured as follows:

$$d_{loss} = \frac{a-b}{a} \quad (3)$$

d_{loss} measures the ratio of the unmatched features with the whole SIFT feature in the original image. And b is the number of matched SIFT features, a is the number of the whole SIFT features in the original image.

The process of measuring the distance of matched SIFT features between original image and retargeted image is briefed as follows [11]:

$$EMD(P, Q) = (\min_{\{f_{ij}\}} \sum_{i,j} f_{ij} d_{ij} / (\sum_{i,j} f_{ij}) \quad s.t \quad f_{ij} \geq 0$$

$$\sum_j f_{ij} \leq P_i \quad \sum_i f_{ij} \leq Q_j \quad \sum_{i,j} f_{ij} = \min(\sum_i P_i, \sum_j Q_j) \quad (4)$$

Here, Where EMD represents the distance between any two features P, Q in the source image and the retargeted image. We set S to determine whether any two features are similar, if $EMD < S$, then feature P and Q are matched. We call d_{ij} the ground distance between bin i and bin j in feature P and Q . And f_{ij} denotes the flow that represents the amount transported from the i th supply to the j th demand.

Then, the energy of the structure of the image is defined as:

$$E_{structure} = d_{loss} + EMD \quad (5)$$

Therefore, by choosing the retargeted image with smallest $E_{structure}$, the proposed SSPM method can well preserve structure of images and improve deformation resilience, confirmed by our experimental results later.

2.2.3. Texture Preservation

Gray-Level Co-occurrence Matrix (GLCM) can measure the complexity and distribution of textures. Energy of the texture is defined as the entropy of GLCM is:

$$E_{texture} = - \sum_{i=1}^k \sum_{j=1}^k G(i, j) \log G(i, j) \quad (6)$$

Where $G(i, j)$ represents GLCM. We chose the operator which got the smallest $E_{texture}$, which represents that the operator introduces least noise in textures. In this way, we can improve deformation resilience too.

2.3. SSPM Retargeting Method

The above discussion formed our energy function: salient object detection $E_{saliency}$, structure preservation $E_{structure}$ and texture preservation $E_{texture}$. The energy function E_{SSPM} is formed as follows: When images satisfy the condition stated in Section 2.2.1, we divide images into three categories and use three different energy functions to choose operators to form the optimal operator sequence:

- Saliency distributed in both boundaries of the image , we just scale image .

- Saliency distributed in one of the image boundaries,

$$E_{SSPM} = \alpha E_{gradient} - \beta E_{structure} - \gamma E_{texture} \quad (7)$$

- Saliency distributed in the middle of the image,

$$E_{SSPM} = \alpha E_{saliency} - \beta E_{structure} - \gamma E_{texture} \quad (8)$$

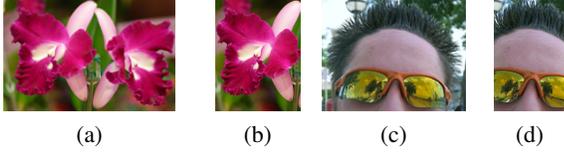


Fig. 3. Images retargeted by cropping. (a), (c) are original images, (b), (d) are images retargeted by cropping.

As illustrated in Fig. 1, each time we use CR, SCL and SC to resize image I by c columns respectively, then compare E_{SSPM} of the retargeted image I'_{CR} , I'_{SCL} and I'_{SC} , choose the operator with highest E_{SSPM} . And repeat the process until the fixed size, record the operator sequences.

3. EXPERIMENTS AND RESULTS

3.1. Experiment Setup

In order to validate our SSPM method from observers' perspective, we organized subjective quality assessment to build the retargeting image dataset. We conducted a no-reference survey as participants' preferences are independent of whether or not they are aware of the original image. The whole dataset with 80 images in [12] were used (3 images are removed as no result from method MULTIOP). And we also chose the paired comparisons technique that the participants are shown two retargeted images at one time, then they were asked to pick the one they prefer.

Rubinstein in [12] showed that cropping (CR), stream-video (SV) and multi-operator (MULTIOP) usually outperformed the rest of other methods such as warping, seam carving and so on. However, among them, CR may cause severe perceptual quality degradation in real application by cropping salient objects, as shown in Fig. 3. Therefore, cropping is excluded from our survey process. We compared SSPM method with SV and MULTIOP methods.

Given out a set of 77 images (each has three comparison methods tested), the total number of possible paired comparisons is: C_3^2 per image \times 77 images = 231 comparisons, for each participant in our survey. And a total of 25 participants took part in the test and the final votes of each image are: C_3^2 per image \times 25 participants = 75 votes. Then we count the votes of MULTIOP, SV and SSPM respectively. The result of subjective assessment will be illustrated in next section.

3.2. Subjective Results

We count the votes of every image on the three method and rank the three comparison methods. The method with the most votes is ranked as 1, with least votes is ranked as 3. Then we count the rank of 77 images, detailed information is shown in Fig. 4. As we can see, SSPM get the most rank 1 and the least rank 3, which significantly outperforms other methods



Fig. 4. The rank of 77 images obtained by MULTIOP, SV and SSPM.

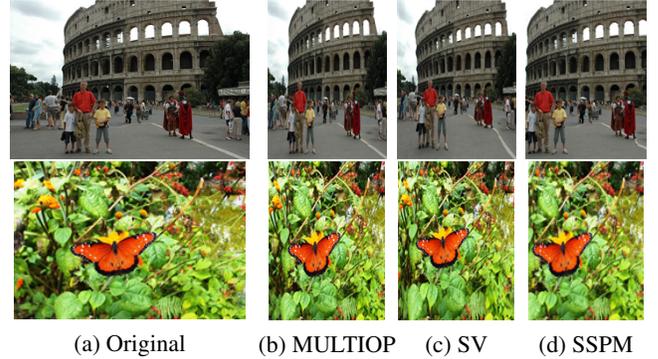


Fig. 5. Comparison of three methods.

as it clearly more coincides with human visual perception. Some results are shown in Fig. 5.

4. CONCLUSIONS AND FUTURE WORK

In this paper, we present Saliency & Structure Preserving Multi-operator (SSPM) retargeting method, which well accords with human perception. Considering the center-bias problem of saliency, we first compared the density of SIFT features in different parts of the images, then employed different principles to get optimal operator sequences. Our principles contain saliency, structure and texture preservation to improve deformation resilience and preserve important objects and structure. Our approach was tested on 77 images and 25 participants. Objective assessment confirmed that our SSPM method outperforms other methods.

So far, there is no saliency algorithm can accurately detect the important objects that attract human attention. It's necessary to explore more accurate saliency models applicable for retargeting scenarios. Besides, larger datasets would be more desirable to validate efficiency of proposed retargeting algorithms.

5. ACKNOWLEDGMENT

This work was supported by the NSFC (Natural Science Foundation of China) under Grant 61571413, and National 973 Program of China under Grant 2015CB351803, and in part by NSFC Grant 61390514 and Intel ICRI MNC.

6. REFERENCES

- [1] Tanchao Zhu, Weiming Wang, Ping Liu, and Yongming Xie, “Saliency-based adaptive scaling for image retargeting,” in *Computational Intelligence and Security (CIS), 2011 Seventh International Conference on*. IEEE, 2011, pp. 1201–1205.
- [2] Claudio SVC Cavalcanti, Herman Martins Gomes, and José Eustáquio Rangel de Queiroz, “Combining multiple image features to guide automatic portrait cropping for rendering different aspect ratios,” in *Signal-Image Technology and Internet-Based Systems (SITIS), 2010 Sixth International Conference on*. IEEE, 2010, pp. 66–73.
- [3] Shai Avidan and Ariel Shamir, “Seam carving for content-aware image resizing,” in *ACM Transactions on Graphics (TOG)*. ACM, 2007, vol. 26, p. 10.
- [4] Radhakrishna Achanta and Sabine Susstrunk, “Saliency detection for content-aware image resizing,” in *Image Processing (ICIP), 2009 16th IEEE International Conference on*. IEEE, 2009, pp. 1005–1008.
- [5] Yu-Shuen Wang, Chiew-Lan Tai, Olga Sorkine, and Tong-Yee Lee, “Optimized scale-and-stretch for image resizing,” in *ACM Transactions on Graphics (TOG)*. ACM, 2008, vol. 27, p. 118.
- [6] Michael Rubinstein, Ariel Shamir, and Shai Avidan, “Multi-operator media retargeting,” in *ACM Transactions on Graphics (TOG)*. ACM, 2009, vol. 28, p. 23.
- [7] Zhongyan Qiu, Tongwei Ren, Yan Liu, Jia Bei, and Yang Yang, “Multi-operator image retargeting based on automatic quality assessment,” in *Image and Graphics (ICIG), 2013 Seventh International Conference on*. IEEE, 2013, pp. 428–433.
- [8] Daniel Vaquero, Matthew Turk, Kari Pulli, Marius Tico, and Natasha Gelfand, “A survey of image retargeting techniques,” in *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 2010, pp. 779814–779814.
- [9] Bowen Jiang, Lihe Zhang, Huchuan Lu, Chuan Yang, and Ming-Hsuan Yang, “Saliency detection via absorbing markov chain,” in *Computer Vision (ICCV), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1665–1672.
- [10] Yossi Rubner, Carlo Tomasi, and Leonidas J Guibas, “The earth mover’s distance as a metric for image retrieval,” *International Journal of Computer Vision*, vol. 40, no. 2, pp. 99–121, 2000.
- [11] Michael Pele, Ofir Werman, “Fast and robust earth mover’s distances,” in *Computer vision, 2009 IEEE 12th international conference on*, 2009.
- [12] Michael Rubinstein, Diego Gutierrez, Olga Sorkine, and Ariel Shamir, “A comparative study of image retargeting,” in *ACM transactions on graphics (TOG)*. ACM, 2010, vol. 29, p. 160.