Selecting Suitable Image Retargeting Methods with Multi-instance Multi-label Learning

Muyang Song[1], Tongwei Ren[1,2], Yan Liu[3], Jia Bei[1,2], and Zhihong Zhao[1,2]

[1] Software Institute, Nanjing University, China
[2] State Key Laboratory for Novel Software Technology, Nanjing University, China
[3] Department of Computing, The Hong Kong Polytechnic University, China
rentw@nju.edu.cn

Abstract. Althogh the diversity of mobile devices brings in image retargeting technique to effectively display images on various screens, no existing image retargeting method can handle all images well. In this paper, we propose a novel approach to select suitable image retargeting methods solely based on original image characteristic, which can obtain acceptable selection accuracy with low computation cost. First, the original image is manually annotated with several simple features. Then, suitable methods are automatically selected from candidate image retargeting methods using multi-instance multi-label learning. Finally, target images are generated by the selected methods. Experiments demonstrate the effectiveness of the proposed approach.

Keywords: Image retargeting, method selection, image characteristic analysis, multi-instance multi-label learning.

1 Introduction

With the popularization of multimedia applications on mobile devices, the requirements of displaying image on small screens with various aspect ratios increase significantly. Images captured by camera usually have much higher resolutions and fixed aspect ratios. They should be adapted to match the respective resolutions of screens of mobile devices. This problem is usually called image retargeting [7].

Much work has been devoted to image retargeting in the past few years. Seam carving calculates the energy of each pixel and iteratively removes the seams with the least energy [8]. Non-homogenous warping formulates image retargeting as a pixel relocation problem, and relocates all the pixels by solving sparse linear system [9]. Scale-and-stretch method represents the original image with square meshes, and adjusts the mesh vertices by quadratic programming [10]. Multi-operator method combines multiple operators, including cropping, scaling and seam carving, and finds the best operator sequence by maximizing the similarity between the original image and the target image [11]. Shift map method utilizes graph labeling to realize the pixel relocation in retargeting, and optimizes the relocation results by graph cut [12]. Streaming video method

K. Imamura et al. (Eds.): BHI 2013, LNAI 8211, pp. 418–426, 2013.

considers object position and edge preservation, and warps image content under the constraints [13].

These methods present prominent effectiveness in content-aware image resizing. Nevertheless, a comparative evaluation of current image retargeting methods found that no method can handle all images [14]. Each image retargeting method succeeds on some images but fails on others. Even multi-operator method, which attempts to combine multiple operators together by optimization, still fails on many cases. Therefore, to obtain high quality target images, selecting suitable retargeting methods for each image is important.

To solve this problem, an institutive strategy is generating target images with different image retargeting methods, and selecting the good results from them. However, it requires production of target images by all candidate image retargeting methods for each original image, thus incurring huge computing cost. Moreover, it is difficult to select high quality results, even if all target images have been generated. Manual selection is labor intensive and time consuming, and current automatic assessment methods are still far from human perception [14]. Hence, a better strategy is selecting the suitable methods from all candidate methods, and generating the target images by the selected methods.

In this paper, we propose a novel approach to select suitable image retargeting methods with multi-instance multi-label learning [15]. In our approach, the selection of image retargeting methods is directly based on the analysis of the original images characteristic, and no target image is required to generate before retargeting method selection. To the best of our knowledge, it is the first work about selecting suitable image retargeting method according to the characteristic of original image. Fig. 1 shows an overview of the proposed approach. First, several features are manually annotated to each original image to represent its characteristic. Then, suitable methods are automatically selected from the candidate image retargeting methods with multi-instance multi-label learning. Finally, the high quality target images are generated by the selected methods and provided to the user.

2 Image Retargeting Methods Selection

2.1 Image Characteristic Analysis

To select suitable image retargeting methods, we should first analyze the characteristic of original image. There are several ways to represent image characteristic, such as extracting visual features from image content, generating tags from the text co-occurring with image, and annotate the image either manually or automatically.

However, current research has only found several high-level features related to retargeting performance. For example, if the original image contains obvious geometric structures, retargeting methods that warp the content of it may cause artifacts. In addition, if the original image contains multiple foreground objects, retargeting methods that simply crop the image may lead to content loss.

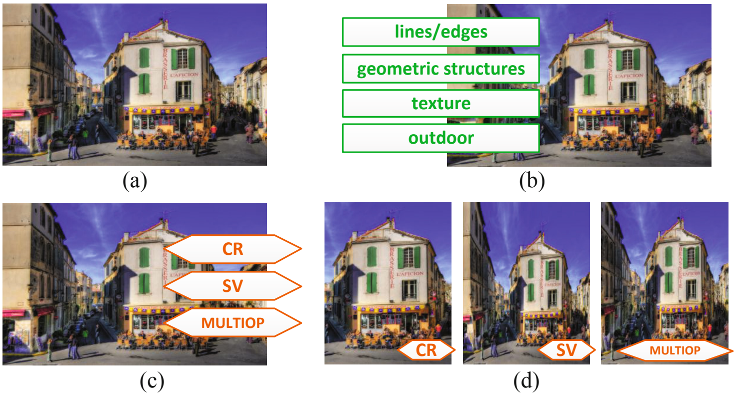

Fig. 1. An overview of the proposed approach. (a) Original image. (b) Manually annotated features to represent image characteristic. (c) Automatically selected image retargeting methods with multi-instance multi-label learning. (d) Target images generated by the selected methods.

These high-level features are hard to extract automatically. Hence, in the proposed approach, we designate some easy-to-find features, such as face and line, and ask the users to manually annotate these features to represent original image characteristic accurately.

2.2 Selection Using Multi-instance Multi-label Learning

According to the manually annotated features, we select suitable image retargeting methods for a given image. Considering each image may have multiple features and multiple suitable retargeting methods, we formulate the selection of suitable image retargeting methods as a multi-instance multi-label learning problem. Compared to traditional learning framework, such as multi-instance learning and multi-label learning, multi-instance multi-label learning provides more natural problem representation and leads to better performance [15].

Let $F_i = \{f_{i,1}, f_{i,2}, \ldots, f_{i,x_i}\}$ be the feature set of the original image i, and $M_i = \{m_{i,1}, m_{i,2}, \ldots, m_{i,y_i}\}$ be the suitable retargeting methods of image i, the selection of suitable image retargeting methods for image i can be considered as finding the relationship $\varphi_i : F_i \to M_i$.

We treat an image feature as an instance and a suitable retargeting method as a label. In this way, the selection of suitable image retargeting methods can be represented as a multi-instance multi-label problem. Let F denote the set of all image features and M the set of all the candidate image retargeting methods, the selection of suitable image retargeting methods can be formulated to learn a function $\varphi : 2^F \to 2^M$ from the given training data set $\{(F_1, M_1), (F_2, M_2), \ldots, (F_n, M_n)\}$, where F_i and M_i are the feature set and suitable retargeting method set of image i respectively.

We solve the problem with MIMLSVM algorithm [15]. We first collect all F_i from the training data and put them into a data set F_{train}. Then, we carry out k-medoids clustering on F_{train} using Hausdorff distance [16]. After the clustering process, the data set F_{train} is divided into k partitions. We calculate the Hausdorff distance between F_i and the medoid of each partition, and transform F_i to a k-dimensional vector δ_i, whose i-th component is the distance between F_i and the medoid of the i-th partition. We assume $\varphi(F_i) = \varphi^*(\delta_i)$, where $\varphi^* : \delta \rightarrow 2^M$ is a function learning from the data set $\{(\delta_1, M_1), (\delta_2, M_2), \ldots, (\delta_n, M_n)\}$. In this way, the selection of suitable image retargeting methods is transformed into a multi-label learning problem, and solved with the MLSVM algorithm [17]. The multi-label learning problem is further decomposed into multiple independent binary classification problems. In each problem, one label is processed with SVM.

With the partition medoids and function φ^*, suitable image retargeting methods can be automatically selected based on annotated features of a given image, and target images can be further generated with the selected retargeting methods.

3 Experiments

3.1 Dataset

We verify the proposed approach on RetargetMe dataset [18]. It contains 37 original images with manually annotated features, including lines/edges, faces/people, texture, foreground objects, geometric structures, symmetry, outdoors, and indoors. Each original image has eight corresponding target images generated by seam carving (SC) [8], non-homogeneous warping (WARP) [9], scale-and-stretch (SNS) [10], multi-operator (MULTIOP) [11], shift-maps (SM) [12], streaming video (SV) [13], uniform scaling (SCL) and manual cropping (CR), respectively. It also provides manual evaluation results of target image quality in two versions, reference version and no-reference version, with each version containing the number of votes each target image gained. In each version, 210 participants voted the better target image in paired comparisons, and each target image could obtain up to 63 votes [14]. The organization of this dataset is suitable for our experiments.

3.2 Experiment Results

In our experiments, for each original image, if the number of votes of a target image is not less than 80% of the highest vote of all the target images generated from it, we treat the corresponding image retargeting method as a suitable method for this original image.

Since the size of RetargetMe dataset is small, we randomly divide the dataset into ten groups according to the number of original images, seven groups contain the data of four original images and three groups contain the data of three original images. For each run, we use nine groups as training data and the other

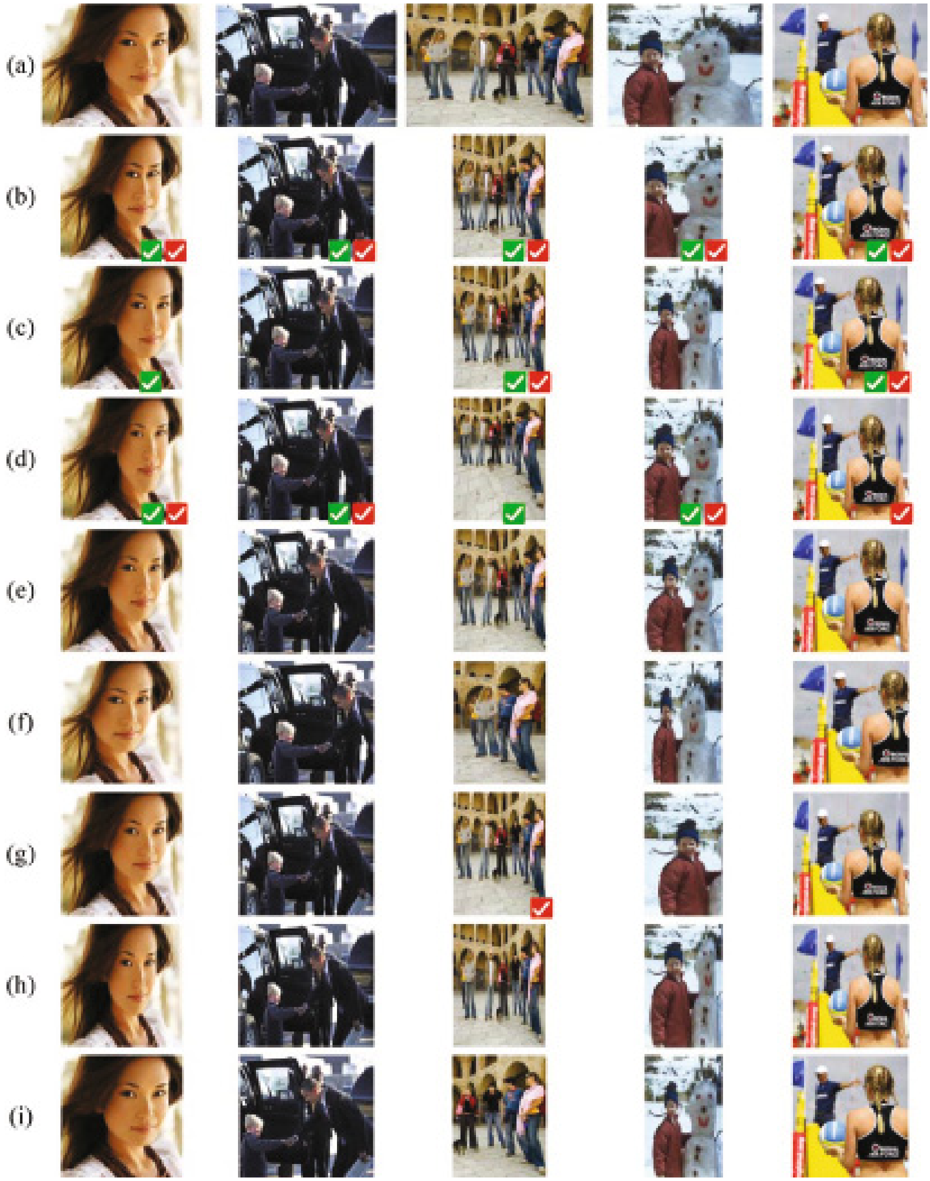

Fig. 2. Examples of our results. (a) Original images named Obama, Umdan, Jon and volleyball in RetargetMe dataset. (b)-(i) Target images generated by SC, WARP, SNS, MULTIOP, SM, SV, SCL and CR, respectively. The target images with green marks are ground truth, and the target images with red marks are selected by our approach.

one group as test data. Fig. 2 illustrates examples of the selection results with our approach. It shows that our approach can obtain the consistent selection results with manual evaluation results.

We calculate precision, recall and F1 measure of the results, where precision is the percent of the correctly selected methods in all selected methods, recall is the percent of the correctly selected methods in all suitable methods, and F1 measure is the harmonic mean of precision and recall. We also use hit-rate to denote the performance in real application of providing the target images generated by all selected methods to the user, where hit-rate is the percent of the images with at least one correctly selected method. We treat each group of data as test data in sequence, and calculate the mean values of precision, recall, F1 measure and hit-rate. The bottom row of Table 1 and 2 shows the performance of our approach for reference evaluation and no-reference evaluation, respectively.

We compare the proposed approach with automatic quality assessment based selection strategy. We choose representative automatic quality assessment methods for image retargeting, including bidirectional similarity (BDS), bidirectional warping (BDW), edge histogram (EH), color layout (CL), SIFT-flow (SIFTflow) and earth-mover's distance (EMD). RetargetMe dataset provides all assessment results of the above methods. We treat the reciprocal of distance between original image and target image as the assessment score, and select suitable image retargeting methods in a similar way to ground truth. For each original image, if the assessment score of a target image is not less than 80% of the highest score of all the target images, we treat the corresponding image retargeting method as a suitable method for this case. The comparison results show that our approach can obtain higher precision, F1 measure and hit-rate than the automatic quality assessment based selection approaches. In comparison of recall, the selection approach using EMD obtains higher recall than our approach. Under the analysis of its selected methods, we find it provides nearly all the candidate retargeting methods to user because the assessment scores of different target images are very close in many cases, leading to low precision and negatively influences its user experience in real applications. The influence of selection strategy also occurs on CL based approach for its assessment scores between SCL-generated target image and others differ greatly causing the low precision and recall of CL based approach. To avoid the bias, we apply a new selection criterion on CL and EMD based approaches, selecting the top 3 methods with the highest assessment scores. The seventh and eighth rows in Table 1 and 2 shows the performance of CL and EMD based approaches with the new selection strategy. It shows that our approach still outperforms them.

3.3 Discussion

In the experiments, we find some relationship between image features and corresponding suitable image retargeting methods. For example, if the original image contains only lines and geometric structures, CR and SM are very likely to be suitable methods for retargeting; however, if additional features are added, SV becomes a likely suitable while SM is no longer a one.

Table 1. Comparison with automatic quality assessment approaches on reference evaluation

	Precision	Recall	F1	Hit-rate
BDS	40.4%	43.8%	42.0%	89.2%
BDW	45.6%	29.5%	35.8%	64.9%
EH	34.7%	62.9%	44.7%	89.2%
CL	10.8%	3.81%	5.6%	10.8%
SIFTflow	56.0%	26.7%	36.2%	70.3%
EMD	37.4%	**87.6%**	52.4%	89.2%
CL*	35.1%	27.6%	30.9%	62.2%
EMD*	49.5%	52.4%	50.9%	89.2%
Our	**64.7%**	60.6%	**62.6%**	**94.6%**

Table 2. Comparison with automatic quality assessment approaches on no-reference evaluation

	Precision	Recall	F1	Hit-rate
BDS	46.5%	51.5%	48.9%	78.4%
BDW	48.5%	30.1%	37.1%	70.3%
EH	35.2%	65.0%	45.7%	81.1%
CL	13.5%	4.9%	7.2%	13.5%
SIFTflow	52.0%	25.2%	33.9%	62.2%
EMD	31.6%	**81.6%**	45.6%	89.2%
CL*	22.5%	24.3%	23.4%	56.8%
EMD*	45.9%	49.5%	47.6%	86.5%
Our	**58.5%**	53.4%	**55.8%**	**94.6%**

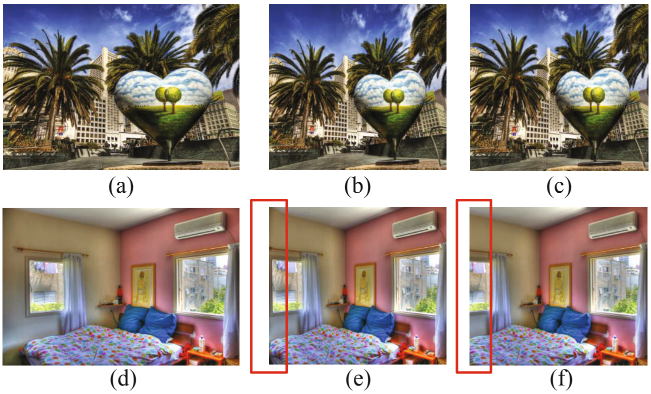

Fig. 3. Examples of inaccuracy of ground truth. (a) and (d) are the original images named Sanfrancisco and Bedroom. (b) and (c) are the target images selected by our approach by not in ground truth. (e) and (f) are the target images with obvious problems but in ground truth.

We also find some limitations of our approach, such as selection precision and recall are not very high. One possible reason of this limitation is the small size of RetargetMe dataset, which cannot provide enough training data for selection. Another reason is the ground truth of suitable retargeting methods is not accurate enough. To avoid bringing in subjective bias, we use the same selection strategy for each original image to determine the ground truth. However, it leads to inaccuracy in some situations. Fig. 3 illustrates the examples of inaccuracy of ground truth. The top row shows several target image selected by our approach but not in ground truth. However, we can find these target images all have high quality. In addition, the bottom row shows several target images in the ground truth but not selected by our approach. Nevertheless, we can find these target images have obvious problems.

4 Conclusion

In this paper, we propose an image retargeting method selection approach based on the characteristics of original image. The proposed approach formulates the selection of suitable image retargeting methods as a multi-instance multi-label learning problem, and automatically select the suitable image retargeting methods for a given image based on several simple features of the original image. Compared to target image selection with automatic quality assessment, the proposed approach requires less computing cost and obtains higher consistency with manual evaluation.

Our future work will focus on improving the proposed approach with enhanced dataset, e.g. enlarge the dataset and re-label the ground truth of suitable retargeting method manually. We will also consider the possibility to extend the approach to video retargeting method selection.

Acknowledgments. 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).

References

1. Smith, T.F., Waterman, M.S.: Identification of Common Molecular Subsequences. J. Mol. Biol. 147, 195–197 (1981)
2. May, P., Ehrlich, H.C., Steinke, T.: ZIB Structure Prediction Pipeline: Composing a Complex Biological Workflow through Web Services. In: Nagel, W.E., Walter, W.V., Lehner, W. (eds.) Euro-Par 2006. LNCS, vol. 4128, pp. 1148–1158. Springer, Heidelberg (2006)
3. Foster, I., Kesselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, San Francisco (1999)
4. Czajkowski, K., Fitzgerald, S., Foster, I., Kesselman, C.: Grid Information Services for Distributed Resource Sharing. In: 10th IEEE International Symposium on High Performance Distributed Computing, pp. 181–184. IEEE Press, New York (2001)

5. Foster, I., Kesselman, C., Nick, J., Tuecke, S.: The Physiology of the Grid: an Open Grid Services Architecture for Distributed Systems Integration. Technical report, Global Grid Forum (2002)
6. National Center for Biotechnology Information, `http://www.ncbi.nlm.nih.gov`
7. Cho, S., Choi, H., Matsushita, Y., Lee, S.: Image Retargeting Using Importance Diffusion. In: IEEE International Conference on Image Processing, Cairo (2009)
8. Rubinstein, M., Shamir, A., Avidan, S.: Improved Seam Carving for Video Retargeting. In: ACM International Conference on Computer Graphics and Interactive Techniques, New York (2008)
9. Wolf, L., Guttmann, M., Cohen-Or, D.: Non-homogeneous Content-driven video-retargeting. In: IEEE International Conference on Computer Vision, Rio de Janeiro (2007)
10. Wang, Y.S., Tai, C.L., Sorkine, O., Lee, T.Y.: Optimized Scale-and-stretch for Image Resizing. In: ACM International Conference on Computer Graphics and Interactive Techniques in Asia, Singapore (2008)
11. Rubinstein, M., Shamir, A., Avidan, S.: Multi-operator Media Retargeting. In: ACM International Conference on Computer Graphics and Interactive Techniques, New Orleans (2009)
12. Pritch, Y., Kav-Venaki, E., Peleg, S.: Shift-map Image Editing. In: IEEE International Conference on Computer Vision, Kyoto (2009)
13. Krähenbühl, P., Lang, M., Hornung, A., Gross, M.: A System for Retargeting of Streaming Video. In: ACM International Conference on Computer Graphics and Interactive Techniques in Asia, Yokohama (2009)
14. Rubinstein, M., Gutierrez, D., Sorkine, O., Shamir, A.: A Comparative Study of Image Retargeting. In: ACM International Conference on Computer Graphics and Interactive Techniques in Asia, Seoul (2010)
15. Zhou, Z.H., Zhang, M.L., Huang, S.J., Li, Y.F.: Multi-instance Multi-label Learning. Artificial Intelligence 176(1), 2291–2320 (2012)
16. Edgar, G.A.: Measure, Topology, and Fractal Geometry. Springer, Berlin (1990)
17. Boutell, M.R., Luo, J., Shen, X., Brown, C.M.: Learning Multi-label Scene Classification. Pattern Recognition 37(9), 1757–1771 (2004)
18. RetargetMe Dataset, `http://people.csail.mit.edu/mrub/retargetme/`