www.ietdl.org

Published in IET Image Processing
Received on 10th August 2011
Revised on 3rd October 2012
Accepted on 26th October 2012
doi: 10.1049/iet-ipr.2012.0308

ISSN 1751-9659

Optimised image retargeting using aesthetic-based cropping and scaling

Yun Liang[1,2], Zhuo Su[2,3,4], Chuntao Wang[1], Dong Wang[1,2], Xiaonan Luo[2,3,4]
[1]School of Information, South China Agricultural University, Guangzhou 510642, People's Republic of China
[2]National Engineering Research Center of Digital Life, Guangzhou 510006, People's Republic of China
[3]State-Province Joint Laboratory of Digital Home Interactive Application, Guangzhou 510006, People's Republic of China
[4]School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510006, People's Republic of China
E-mail: sdliangyun@163.com

Abstract: Image retargeting is a critical technique in displaying images on devices with different resolutions. This study presents a new image retargeting algorithm based on aesthetic-based cropping and scaling. A composite measurement is first constructed under the guidelines of composition aesthetics in photographing. An aesthetic-based cropping is proposed to yield an optimal candidate retargeted image with maximum aesthetic value computed via a constructed composite measurement. The optimal candidate is uniformly scaled to obtain the retargeted image of target size. Some subjective and objective assessments demonstrate that the proposed scheme significantly improves the aesthetics of retargeted images while preserving the important objects. It also achieves better performance in terms of aesthetics than a number of conventional image retargeting approaches.

1 Introduction

Image retargeting is a technique resizing an original image to an aesthetic version of target size. It is mandatory for displaying images on small devices with different resolutions, and for building content-aware media manipulation tools. It aims at preserving the important objects and enhancing the visual aesthetics of an image while changing its size.

A simple way for image retargeting is to scale down or crop the original image. However, the uniformly scaling method usually squeezes the important objects in image retargeting, and introduces distortions on important objects. The cropping method generally loses information out of the cropping windows. To remedy the weakness of simple approaches, some intelligent cropping methods [1–4] have been proposed, which improve the retargeting performance by preserving the important contents. Among them, the saliency-based cropping (SBC) [2] is a popular one. However, it still loses the important information from the clipping window, for example, the situation in Fig. 1*b* where parts of the cartoons are not contained in the retargeted image.

The image retargeting methods can be roughly classified into three categories, that is, the discrete, continuous and multi-operators image retargeting methods. The discrete approaches [5–9] shrink the image to the target size by removing the pixels or patches with least importance, such as the improved seam carving (SC) [6] and the shift map (SM) [8]. Although they lead to excellent results for images with blur backgrounds, they either introduce visual artifacts (see the body of the right cartoon in Fig. 1*c*) or lose small objects (see Fig. 1*d*). The continuous schemes [10–15] take image retargeting as the process of mapping an image from the original size to the target one, for example, the scale-and-stretch (SNS) [10] is a representative for this category. These methods usually define an optimal function to homogeneously scale down the important areas to preserve the important contents while diffusing the distortions in deformation into the unimportant areas. They, however, cannot highlight the important objects sufficiently and cannot preserve the shape of important objects even at the cost of extra deformation distortions (see Fig. 1*e*). The third category [16, 17] utilises multi-operators to resize an image, such as the multi-operator (MULTIOP) [16]. That is, they combine the cropping, scaling and SC method together to generate retargeted images. However, it is difficult to decide when and how to use the operator in MULTIOP, namely, how to effectively integrate the corresponding operators to form MULTIOP. For example, the MULTIOP cannot preserve the aspect ratio of important objects for using SC operator, for example, the moon in Fig. 1*f* has become an ellipse rather than a circle as shown in the original image.

The aforementioned retargeting methods only focus on preserving the important objects and the image integrality without taking the aesthetics of retargeted images into account, and it would be produce an undesirable result because the aesthetics plays an important role in human vision. Recently, some aesthetic-based image retargeting methods [18, 19] are proposed. For example, the optimising photo composition (OPC) [18] introduced the composition aesthetics in photographing into image retargeting.

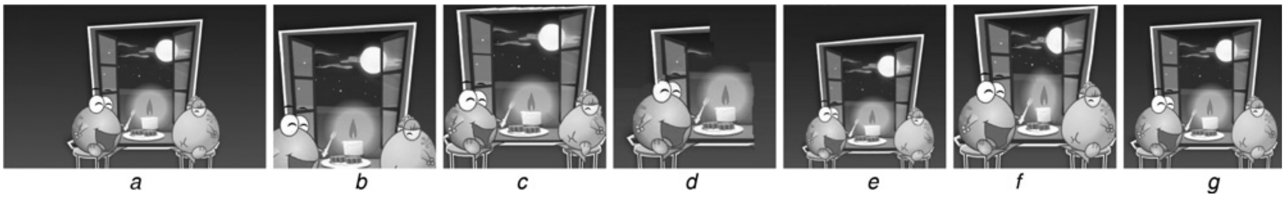


Fig. 1 *Comparisons of different retargeting approaches*

The retargeted images *b–g* are computed by methods of SBC, SC, SM, SNS, MULTIOP and Ours

Although it has significantly improved the visual effects of retargeted images, it does not take into account the rule of centre [20] and the prominent vertical and horizontal lines [21] and, thus, sometime they cannot bring preferable visual effect. This then motivates us to develop a new image retargeting algorithm that further incorporates the mentioned two rules.

In this paper, we present a new image retargeting algorithm using the aesthetic-based cropping and scaling, which can preserve the important objects while enhancing the aesthetics in image retargeting. The aesthetic-based cropping is defined by some composition aesthetics in photographing to measure the aesthetics of candidate retargeted result. In more detail, we first generate quite a lot candidate retargeted images by defining a series of clipping windows with target aspect ratio, then choose the optimal candidate with the maximum aesthetic value computed via the composite measurement, and finally uniformly scale the optimal candidate to the target size. Fig. 1*g* illustrates the retargeted image yielded via our scheme. Extensive assessments demonstrate that the proposed scheme significantly improves the aesthetics of retargeted images while preserving the important objects. The superiority over the conventional image retargeting schemes is also demonstrated in both the subjective and objective evaluations on Section 4.

2 Aesthetic measurement based on photographic composition guidelines

Retargeting an image is to generate an aesthetic version with target size from the original image. Its key point is to exploit a suitable measurement to well reflect the aesthetic characteristics. In this section, we develop an aesthetic measurement based on a series of computable rules corresponding to the photographic composition guidelines. In Section 2.1, we first introduce the composition guidelines in photographing while utilising them in Section 2.2 to quantitatively define the aesthetic measurement.

2.1 Photographic composition guidelines

In photography, composition guidelines are generally exploited by photographers to yield some aesthetic photos. Recently, some works [22–26] demonstrate these composition guidelines, for example, rule of third to enhance the aesthetic quality of consumer photos. Although many composition guidelines have been introduced to shoot well-composed photos [24], only a limited set of such guidelines leads to prominent results in various images. Thus, we only deploy these limited set of guidelines to define the aesthetic measurement, which are classified into three rules, that is, the points rule, lines rule and area rule.

2.1.1 Points rule: The points rule denotes the rule following two well-known photographic composition guidelines, that is, the rule of thirds (see Figs. 2*a* and *b*) [18, 22–26] and the rule of centre (see Fig. 2*c*) [23]. By the rule of thirds, an image is divided into nine equal parts by two equally spaced horizontal lines and two equally spaced vertical lines (see the four white lines in Fig. 2*a* and *b*), and the four intersections formed by these four lines are called power points. By the rule of centre, the centre of an image is defined as the power point. Under the points rule, important objects are generally placed around the mentioned power points to form an aesthetic image. For example, the sun and boat in Fig. 2*a* are located in the neighbourhood of the power points, namely the line intersections. Similarly, the spider in Fig. 2*c* is also located near another power point, namely the centre of the image. It is clearly that placing important objects around the power points can greatly evoke aesthetic feelings and attract users' attention.

2.1.2 Lines rule: There may exists long edges representing the boundaries or the trend of important objects in an image,

Fig. 2 *Illustration of several composition rules*

Important objects, for example, the sun and the boat, the water lily, and the spider in *a–c*, respectively, are located around the 'power points' to form high-quality images while those in the high-quality images *d–f* obey the lines rule

a Rule of thirds
b Rule of thirds
c Rule of centre
d Dominant diagonal lines
e Horizontal lines
f Vertical lines

for example, the branch in Fig. 2*d* and the electrical lines in Fig. 2*e*, which are generally termed as prominent lines. According to composition guidelines in photography, the layout of prominent lines influences the dynamic effect and balance. To compute the layout of prominent lines, we denote some lines as feature lines, including the diagonal and back-diagonal lines (see the white lines in Fig. 2*d*), the two equally spaced horizontal lines and the central horizontal line (refer to the white line Fig. 2*e*), the two equally spaced vertical lines (see the white lines in Fig. 2*f*) and the central vertical line (see the grey line in Fig. 2*f*). In photographing, people are encouraged to locate prominent lines near the feature lines to produce aesthetic effects. Following this, we define the lines rule as the rule satisfying the aesthetic layouts of the mentioned eight feature lines to measure the aesthetics introduced by the mentioned eight feature lines. Under the lines rule, the prominent lines in an image should lie around the feature lines to enhance the aesthetics [24].

2.1.3 Area rule: In some situations, image retargeting yields a number of aesthetic versions from the original image, each of which obeys the points and lines rules but contains a different percentage of important objects. Nevertheless, to avoid losing important objects the larger the percentage of important objects included in the target image, the more favourable the target image is. Therefore it makes sense to define the area rule to describe the percentage of important objects contained in the target image. Under this rule, the target image with the largest percentage of important objects is selected as the desirable result.

2.2 Aesthetic measurement computation

For the points, lines and area rules mentioned above, we define an aesthetic measurement for each of them to compute the aesthetic value induced by them. By these measurements, we propose a composite aesthetic measurement that will be utilised in the aesthetic-based cropping.

2.2.1 Aesthetic measurement from points rule: As aforementioned, the points rule takes into account five power points, namely, the points P0, P1, P2, P3, P4. Since the points rule requires placing important objects around the power points, the Euclidean distances from the centres of important objects to the power points can be exploited to define the aesthetic measurement based on the points rule.

To obtain the aesthetic measurement of the points rule, we first compute its important objects. Here, we define important object as a visual conspicuous, continuous and homogeneous image component that attracts human attention. We use a saliency map to identify important object. After computing the saliency map of the original image such as Fig. 4*b*, we normalise it to get an importance map by a threshold such as Fig. 4*c*. Then, the n largest four-connected regions [27] of the importance map are taken as important objects. To control the number of important objects, the threshold is specified to the average of salient value of the whole saliency map, and n to be no bigger than eight.

For an important object S_i, we compute its centroid as the centre S_{ic}, such as S_{12} illustrated in Fig. 3*a*. We then calculate its total saliency value $F(S_i)$. Many approaches have been proposed to compute the saliency value, for example, the works [27, 28]. In [27], it performs well in

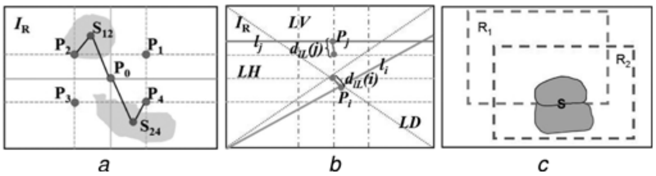


Fig 3 *Three rules of the aesthetic measurement computation*
a Points rule
b Lines rule
c Area rule

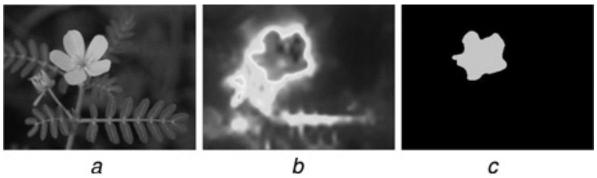


Fig. 4 *Generating important objects from saliency map*
a Original image
b Saliency map
c Important object

predicating human fixations on natural images, it is utilised in our scheme to calculate $F(S_i)$, which is expressed as

$$F(S_i) = \sum H(j), \quad j \in S_i$$

where $H(j)$ is the saliency value of pixel j computed by the method in [27]. Let $d_{\rm sp}(i)$ denote the minimum distance between an important object S_i and all the five power points $P_j(j = 0, \ldots, 4)$, that is

$$d_{\rm sp}(i) = \min\{\text{Euclid}(S_{ic}, P_j)\}, \quad j = 0, \ldots, 4$$

where Euclid(S_{ic}, P_j) is the Euclidean distance between S_{ic} and P_j. Then under the points rule, we construct the aesthetic measurement as

$$V_{\rm p} = \frac{\sum_{i=1}^{n} \left(1/\left(1 + d_{\rm sp}(i)\right)\right) F(S_i)}{\sum_{i=1}^{n} F(S_i)}$$

The larger $V_{\rm p}$ implies the better visual effect or the more favourable aesthetics.

Fig. 5 illustrates the computation of saliency value and the aesthetic measurement $V_{\rm p}$. Two candidate retargeted images are generated from the original image, as indicated by the white rectangles in Fig. 5*a*, respectively. Each candidate is then taken to obtain its saliency values and the corresponding $V_{\rm p}$. The results show that the candidate within Fig. 5*b* yields $V_{\rm p} = 0.0135$, whereas that within Fig. 5*d* has $V_{\rm p} = 0.2100$. This implies that the former is more preferable than the latter, which is clearly consistent with the shown visual effect and, thus, demonstrates the feasibility of (3).

2.2.2 Aesthetic measurement from lines rule: Recall that the lines rule uses eight feature lines, $\text{lf}_j(j = 1, \ldots, 8)$, as shown in Fig. 3*b*. They are divided into three categories, that is, the horizontal lines named LH, the vertical lines LV and the diagonal lines LD. Since the lines rule prefers to place the prominent lines of an image around any feature line, it makes sense to utilise the distance between the prominent

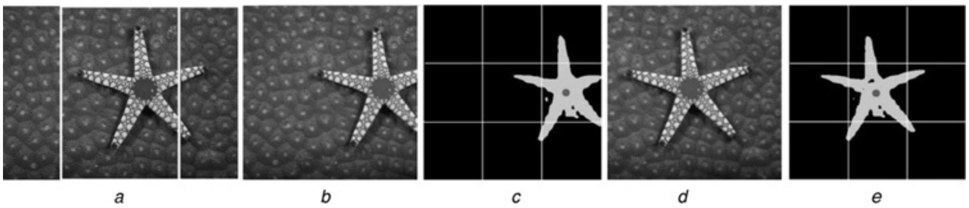


Fig. 5 *Illustration of aesthetic measurement from the points rule*

White rectangles in *a* yield the candidate retargeted images *b* and *d*, respectively. Their corresponding saliency values are shown in *b* and *e*, while the corresponding values of V_p are 0.0135 and 0.2100, respectively
a Original image
b Output of left rectangle
c Saliency of *b*
d Output of right rectangle
e Saliency of *d*

lines and the feature lines to define the aesthetic measurement from the lines rule.

Assume that n prominent lines, denoted as $l_i(i = 1, \ldots, n)$, are generated by the Hough transformation. We denote the slop of l_i is slp_i. Then $\text{slp}_i = 0$, $\text{slp}_i \to \infty$ and other values of slp_i imply that l_i belongs to LH, LV and LD, respectively. Via such classification, the distance between l_i and all feature lines is defined as the minimal one between l_i and the feature lines belonging to the same category as l_i, that is

$$\begin{aligned} d_1(i) &= \min\{d_1(l_i, \text{lf}_j)\} \\ &= \min\{\text{Euclid}(c_i, c_j)\}, \quad \begin{cases} \text{lf}_j \in \text{LH}, & \text{if } \text{slp}_i = 0 \\ \text{lf}_j \in \text{LV}, & \text{if } \text{slp}_i = \infty \\ \text{lf}_j \in \text{LD}, & \text{otherwise} \end{cases} \end{aligned}$$

where $c_i(i = 1, \ldots, n)$ and $c_j(j = 1, \ldots, 8)$ stand for the centre points of l_i and lf_j, respectively, and $\text{Euclid}(c_i, c_j)$ is the Euclidean distance between c_i and c_j.

Let $G_1(i)$ be the total salient value of l_i, then $G_1(i)$ is calculated as

$$G_1(i) = \sum H(j), \; j \in l_i$$

By taking the $d_1(i)$ and $G_1(i)$ into account, the aesthetic measurement from the lines rule can be defined as

$$V_1 = \frac{\sum_{i=1}^{n} \left(1/\left(1 + d_1(i)\right)\right) G(l_i)}{\sum_{i=1}^{n} G(l_i)}$$

Similar to V_p, the larger V_1 implies better aesthetics.

Fig. 6 illustrates the effects of V_1. The blue and red rectangles in Fig. 6*a*, generate two candidate retargeted images, namely the Figs. 6*b* and *d*, respectively. Their prominent lines are plotted in Figs 6*c* and *e* by the red lines. The white lines in Figs. 6*c* and *e* are the closest feature lines to these prominent lines. The aesthetic values of V_1 corresponding to Figs. 6*b* and *d* are 0.033 and 0.1500, respectively, which accords with the subjective assessment.

2.2.3 Aesthetic measurement from area rule: Let Si($i = 1, \ldots, n$) be the important object extracted from the candidate retargeted image (see Section 2.2.1 for the extraction approach). Denote R1 and R2 two candidate retargeted images containing parts of S_i, as shown in Fig. 3*c*. It is found that R_1 only contains part of S_i, whereas R_2 contains the whole S_i. Thus according to the area rule, the candidate represented with R_1 has better aesthetics than that represented with R_2. Although this example considers only one important object, it is straightforward to extended to the situations containing multiple important objects, that is, the area containing more important objects is more preferable in the sense of aesthetics. This observation motivates the aesthetic measurement from the area rule given below.

Assume that $\text{Area}(S_i)$ and $\text{Area}(S_i')$ are the area of the important object S_i and S_i' included in the original image and retargeted image, respectively. n denotes the number of important objects of the original image. Then the quantitative measurement of the area rule is constructed as follows

$$V_s = \frac{\sum_{i=1}^{n} \text{Area}(S_i')}{\sum_{i=1}^{n} \text{Area}(S_i)}, \quad i = 1, \ldots, n(n \geq 1)$$

The bigger the V_s, the better the aesthetics of the candidate retargeted image is.

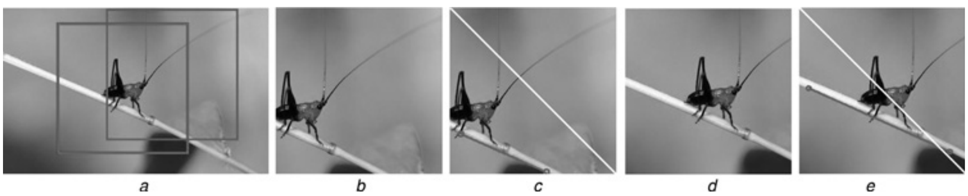


Fig. 6 *Illustration of aesthetic measurement from the lines rule*

We compute the candidate retargeted images *b* and *d* by the white rectangles in *a*. The grey lines in *c* and *e* describes their prominent lines, and the white lines are the closest feature lines to these prominent lines. The corresponding values of V_1 are 0.033 and 0.1500, respectively, for Figs. 5*b* and 15*d*
a Original image
b Output of left rectangle
c Lines of *b*
d Output of right rectangle
e Lines of *d*

doi: 10.1049/iet-ipr.2012.0308

Fig. 7 illustrates the effects of V_s. The left and right rectangles in Fig. 7*a*, generate two candidate retargeted images in Figs. 7*c* and *e*, respectively. Their saliency maps are plotted in Figs. 7*d* and *f*, and their corresponding values of V_s are 0.83 and 0.91, respectively. It is clear that Fig. 7*e* has the larger value of V_s and thus preferable aesthetics, which is consistent with the visual effects shown in Figs. 7*c* and *e*.

2.2.4 Total aesthetic measurement: Since the aesthetic measurements V_p, V_l and V_s reflect different characteristics of aesthetics, they can be further integrated to form one composite measurement. We assign different weights, say ω_p, ω_l and $\omega_s \in [0, 1]$, for the individual measurement V_p, V_l and V_s, respectively, and design the composite measurement as

$$V_{\text{comp}} = \frac{\omega_p V_p + \omega_l V_l + \omega_s V_s}{\omega_p + \omega_l + \omega_s}$$
$$\text{s.t.} \quad \sum \omega_p + \omega_l + \omega_s = 1$$

where 's.t.' stands for 'subject to', and the weights ω_p, ω_l and ω_s are constants in the range [0,1]. It is worth pointing out that $\omega_l = 0$ if no prominent lines exist in the candidate retargeted image.

3 Image retargeting algorithm using aesthetic-based cropping and scaling

Let I be the original image of size $h \times w$ and O the candidate retargeted image of size $h' \times w'$. Then image retargeting is formulated as follows

$$f : \max_{O} V_{\text{comp}}(O), \quad \text{s.t.} \quad I(h, w) \to O(h', w')$$

where $V_{\text{comp}}(O)$ is the composite measurement for the candidate retargeted image O. That is, the objective of image retargeting is to maximise $V_{\text{comp}}(O)$ while computing a version of I with the size of $h' \times w'$.

To achieve the optimisation we develop a new image retargeting algorithm exploiting the aesthetics-based cropping and scaling, which is shown in Fig. 8. The proposed scheme includes three stages, that is, defining the clipping windows, compute the optimal candidate retargeted image from a number of candidates, and uniformly scale the optimal candidate to the target size, respectively. The details are presented below.

3.1 Defining the clipping windows

In the interest of obtaining the retargeted image with best visual effect, we generate a number of candidate retargeted images $O_i (i = 1, 2, \ldots)$, that have different image sizes but the target aspect ratio as $r' = w'/h'$. To avoid the possibility of overflowing the original image and losing information of the original image, the width of each clipping window should neither too large nor too small. In our scheme, the smallest width of clipping windows, denoted as w'_S, is set to be the largest width of all important objects. The largest width of clipping windows, denoted as w'_B, is determined as $r'h$ for the case $r' < w/h$, otherwise w'_B is set to be w. Thus, the width of each clipping window is calculated as

$$w'_c = w'_S + k\Delta_{w'}, \quad k = 0, 1, \ldots, \lfloor (w'_B - w'_S)/\Delta_{w'} \rfloor$$

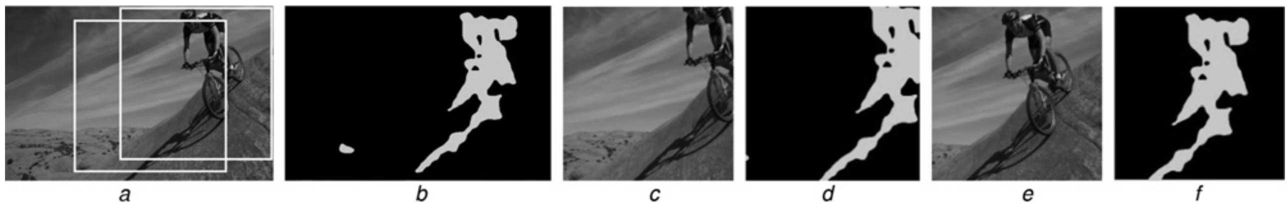


Fig. 7 *Illustration of the aesthetic measurement from the area rule*

Red and blue rectangles in *a* yield the candidate retargeted images *c* and *e*, respectively. Their saliency maps and given in *d* and *f*, respectively. The aesthetic meansurement V_s for *c* is 0.83 while that for *e* is 0.91

a Original image
b Saliency of *a*
c Output of left rectangle
d Saliency of *c*
e Output of right rectangle
f Saliency of *e*

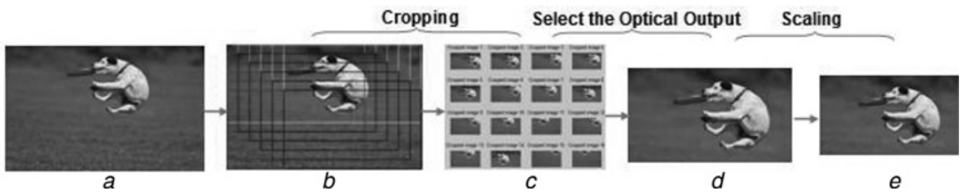


Fig. 8 *Overview of the proposed scheme*

a Original image
b Clipping windows
c Candidate retargeted images
d Optimal candidate
e Scaled retargeted image

where $\lfloor v \rfloor$ denotes the largest integer that is not larger than v and $\Delta_{w'}$ is equal to $(w'_B - w'_S)/N$, where N is an constant integer to control the step of width variation. In our experiments, we compute N as follows

$$N = \begin{cases} w'_B - w'_S, & \text{if } (w'_B - w'_S) \leq 20 \\ 20, & \text{if } (w'_B - w'_S) > 20 \end{cases}$$

3.2 Determining the optimal candidate retargeted image

For each clipping window of size $(r'w'_c) \times w'_c$, we align it at the top-left corner of the original image and then move it pixel by pixel along the horizontal, vertical and diagonal directions, respectively. In this process, each moved window yields a candidate retargeted image O_i. For the original image of size $h \times w$, $(h - h'_c + 1) \times (w - w'_c + 1)$ candidates will generate for the size $(r'w'_c) \times w'_c$. Thus, total $\lfloor (w'_B - w'_S)/\ \Delta_{w'} \rfloor + 1$window sizes results in $M = \left(\lfloor (w'_B - w'_S)/\Delta_{w'} \rfloor + 1 \right) \cdot \left((h - h'_c + 1) \times (w - w'_c + 1) \right)$ candidates, $O_i (i = 1, \ldots, M)$. We then calculate the aesthetic measurement $V_{\text{comp}}(O_i)$ for each candidate retargeted image, and select the one with the largest maximum measurement as the optimal candidate, namely O_{opt}.

3.3 Scale the optimal candidate to the target size

Owing to the intention to obtain the best aesthetics, as aforementioned, the size of the optimal candidate retargeted image O_{opt} is not necessarily the target size of $h' \times w'$. Thus, we need to uniformly scale the O_{opt} to the target size via the linear or bilinear interpolation method, which finally results in the desired retargeted image.

4 Experimental result and analysis

In this section, we evaluate the proposed aesthetic-based image retargeting algorithm. Diverse images are used as the original ones, for example, the images frequently used in other schemes and those from the RetargetMe benchmark [29]. We run the simulation on a PC with Intel Pentium(R) T2370 1.73 GHz CPU and 2 GB RAM. Under this setting, generating the retargeted image generally takes a few seconds, for example, obtaining the retargeted result of size 300×300 from the original image of size 400×300 needs about 10 s. We give some typical results in this section while summarising the others in the appendix.

To measure the quality of retargeted image, we employ both the subjective assessment based on image illustration and user study, and the objective evaluation using the approach of image retargeting assessment (OIRA) [30].

4.1 Subjective assessment on retargeted images

Fig. 9 illustrates the comparison between the proposed scheme and the simple scaling method (SCL). It is observed that the SCL distorted objects by squeezing them, and the proposed method has significantly better visual effect in preserving the shape of important objects than SCL. This is because our scheme obtains the final retargeted image by scaling the optimal candidate satisfying the target aspect ratio (see Section 3) while SCL uniformly resizes the original image to the target size.

The SBC [2] method performs better than some related cropping-based image retargeting algorithms. However, it may lose the important information, as shown in Figs. 10*b* and *e* where part of the boat and panda have been excluded, respectively, because of the fact that clipping windows are not large enough to cover all interested important objects. Compared to SBC, the proposed method can well contain the important objects (see Fig. 10*c*) and preserve the image balance (see Fig. 10*f*). This implies that the clipping window adopted in our method is much more reasonable than that of SBC.

In OPC [19], Liu *et al.* developed an aesthetic-based image retargeting algorithm by using some composite aesthetics. Although this scheme produces retargeted images with good visual effect, it does not take into account the vertical and horizontal prominent lines and the rule of centre and, thus, it may not result in good image balance (see Fig. 11*b*) or preserve the image integrality (see the hat in Fig. 11*e*). In contrast to OPC, our scheme exploits these mentioned lines rule and hence leads to better aesthetic effects, as shown in Figs. 11*c* and *f*.

The continuous image retargeting algorithms such as the streaming video (SV) method [14] and the SNS method

Figure 9 *Comparisons between the proposed method with SCL*

a and *d* are the original images, *b* and *e* are the retargeted results of the SCL method while *c* and *f* are computed by the proposed method

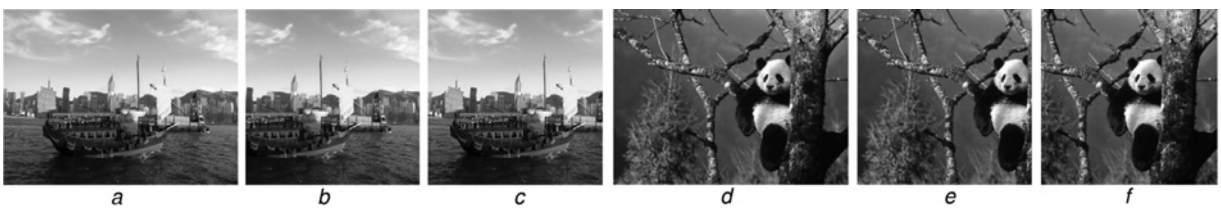


Fig. 10 *Comparisons between the proposed method and SBC [2]*

a and *d* are the original images, *b* and *e* are the retargeted results of the SBC method while *c* and *f* are calculated by the proposed method

Fig. 11 *Comparisons of the proposed method with the OPC method [19]*

a and *d* are the original images, *b* and *e* are the retargeted results of the OPC method while *c* and *f* are calculated by the proposed method

Fig. 12 *Comparisons of the proposed method with SV [14] and SNS [10]*

a and *d* are the original images, *b* and *e* are the retargeted results of the SNS method while *c* and *f* are computed by the proposed method

[10], are quite effective to avoid losing important contents of the original image, whereas they are deficient in preserving the aspect ratio of important objects. For example, the aspect ratio of the important object – the girl – in Fig. 12*b* generated via SV has been changed because of the fact that SV fixes a global scaling factor for the entire image. Also, the heart in Fig. 12*e* obtained via SNS has become unsymmetrical since SNS allows the scaling factor to vary within a predefined range. Compared to SV and SNS, our method can well preserve the target aspect ratio of retargeted images and, thus, lead to preferable aesthetic effects (see Figs. 12*c* and *f*).

Moreover, we compare our method with some other image retargeting methods such as SC [6], MULTIOP [16], non-homogeneous warping (WARP) [12] and SM [8], which are included in the RetargetMe benchmark [29]. The SC, a discrete image retargeting algorithm, would introduce the discontinuity of important objects, for example, the girls in Fig. 13*b*. The MULTIOP, which works by incorporating the scaling and cropping in SC, cannot completely eliminate the discontinuity, such as the white lines illustrated in Fig. 13*c*. The WARP can preserve image integrality, while it generally cannot emphasise the salient areas, for example, the girls in Fig. 13*d*. The SM may lose important objects for excluding important pixels out of image boundary, as shown in Fig. 13*e*. In contrast to these methods, the proposed method can well preserve important objects and significantly enhance the composition aesthetics, as demonstrated in Fig. 13*f*.

To further evaluate the effectiveness of our method, we utilise user study to compare the proposed scheme method with those such as SNS, SC, SCL, SM, SV, WARP and MULTIOP. These compared schemes have been contained in the RetargetMe benchmark which collects many original images and their corresponding retargeted images generated via the mentioned seven schemes, respectively. In the user study, we take 37 images in the RetargetMe benchmark as the original images and invite 120 users aged from 18 to 45 to do pair-wise comparisons. In more detail, we first employ the proposed method to generate 37 retargeted images while taking the corresponding retarget images yielded by the seven compared method for assessment. We then require five users to examine each pair of the retargeted images $(O_{pi}, O_{ci})(i = 1, \ldots, 37)$, where O_{pi} is a retargeted image from our scheme and O_{ci} is one of the compared methods. In every examination, each of the five users would determine which retargeted image of the pair (O_{pi}, O_{ci}) is better and the dominant determination is taken as the final evaluation.

Table 1 summarises the result of user study, where 'better' (see the bold values), 'similar' and 'worse' mean that the retargeted image from the proposed method is better than, similar to, and worse than the one from any compared method, respectively. It is observed that our scheme is better in most cases than other compared schemes. Also, it is found that our scheme is similar to the MULTIOP which is better results than the other six compared schemes.

4.2 Objective metric to assess retargeted images

As different users may have quite different evaluation, the subjective assessment may lead to unfair results, especially in the case of insufficient users. In this subsection, we use an objective metric, namely, the OIRA [30], to compare our method with other available conventional and state-of-art methods. In OIRA, the objective evaluation score is

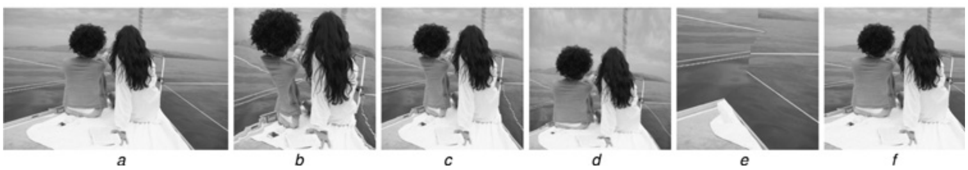


Fig. 13 *Comparisons of the proposed method with some image retargeting methods presented in RetargetMe benchmark*

a is the original image, *b*–*f* are the retargeted results computed by the methods of SC, MULTIOP, WARP, SM and Ours

Table 1 Result of user study

Algorithm	Better	Similar	Worse
SNS	**22**	11	4
SC	**23**	8	5
SCL	**25**	7	5
SM	**21**	10	6
SV	**18**	10	9
WARP	**22**	6	9
MULTIOP	**17**	12	8

calculated as follows

$$\mathrm{Sim}(I^0_{\mathrm{ori}}, I^0_{\mathrm{ret}}) = \frac{\#_{\mathrm{ver}}}{pn(I^0_{\mathrm{ori}}) + pn(I^0_{\mathrm{ret}})} \cdot \frac{1}{\#_{\mathrm{edge}}} \cdot \times \sum_{i=1}^{\#_{\mathrm{edge}}} \mathrm{SSIM}(v_0(e_i), v_1(e_i))$$

where I^0_{ori} and I^0_{ret} are the original and retargeted images, respectively, $pn(I)$ is the number of pixels in image I, e_i is

Table 2 Objective metric to compare our method with others

Image	Ours	MULTIOP	SC	SCL	SM	SNS	SV	WARP
Car1	0.918	0.827	0.837	0.820	0.797	0.752	0.789	0.803
DNKYgirl	0.799	0.707	0.749	0.645	0.746	0.700	0.691	0.753
Getty	0.899	0.718	0.788	0.721	0.736	0.694	0.706	0.728
Jon	0.785	0.738	0.711	0.670	0.715	0.701	0.710	0.698
lotus	0.785	0.718	0.739	0.681	0.706	0.698	0.699	0.698
San	0.905	0.807	0.771	0.789	0.833	0.709	0.738	0.770
set	0.743	0.689	0.683	0.657	0.598	0.605	0.619	0.662
surfer	0.917	0.851	0.888	0.819	0.857	0.706	0.762	0.784
tower	0.897	0.749	0.757	0.711	0.738	0.662	0.696	0.710
woman	0.909	0.758	0.777	0.734	0.714	0.758	0.724	0.767
Johanneskirche	0.918	0.816	0.825	0.789	0.817	0.713	0.766	0.795
brick	0.754	0.787	0.801	0.707	0.738	0.701	0.640	0.759

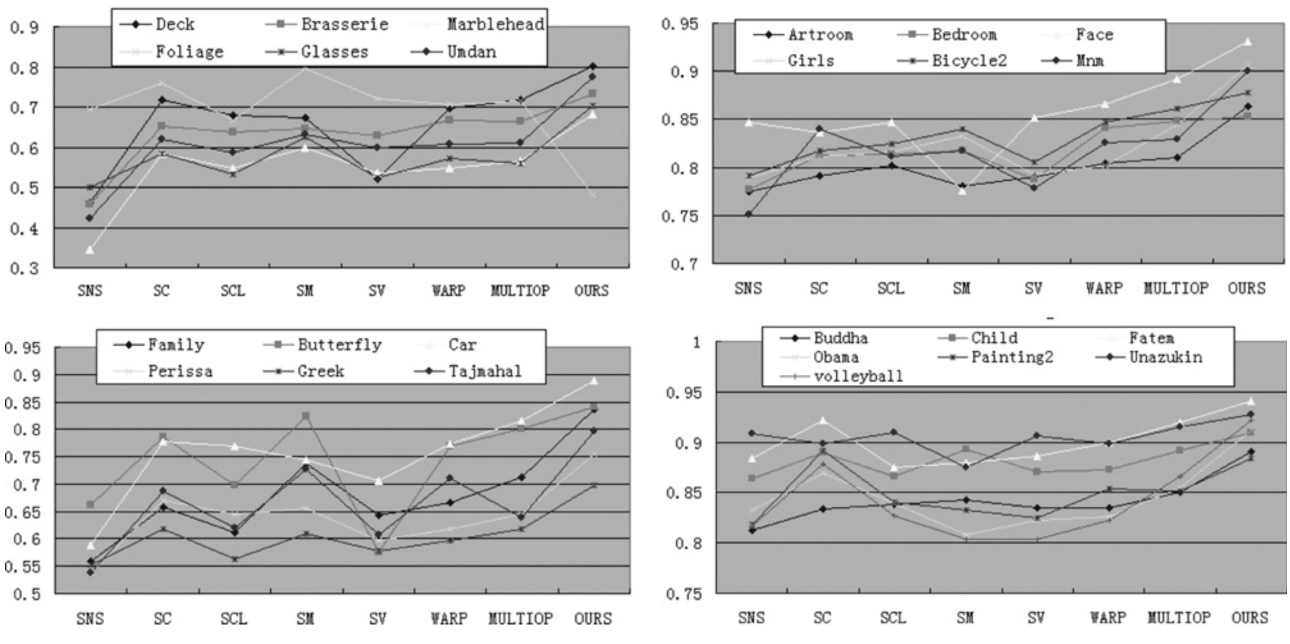
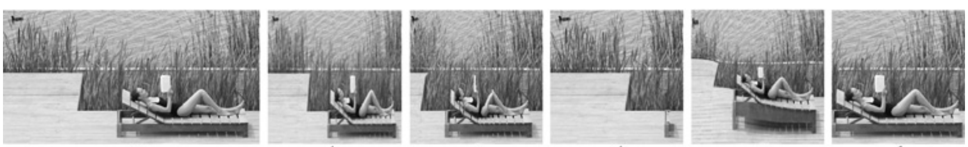


Fig. 14 *Comparisons of the proposed method with the seven image retargeting algorithms included in the RetargetMe benchmark, which is done under OIRA [30]*

y-coordinate presents the OIRA score



a b c d e f

Fig. 15 *Some retargeted results by different retargeting schemes*

a is the original image, *b–f* are the retargeted results computed by the methods of SLC, SC, SM, SNS and Ours

doi: 10.1049/iet-ipr.2012.0308

an edge belonging to a bipartite graph which serves as the correspondence of two geometric structures in I^0_{ori} and I^0_{ret}, $\#_{\mathrm{ver}}$ and $\#_{\mathrm{edge}}$ are the number of vertices and edges, and SSIM(.) denotes the SSIM metric in [31].

We still use the retargeted images in RetargetMe to do comparison. These retargeted images are computed by seven methods including SNS, SC, SCL, SM, SV, WARP and MULTIOP. Table 2 describes the metrics of ORIA for twelve images in RetargetMe. The metrics for the other 25 images in RetargetMe are shown on Fig. 14. It is observed that most of our retargeted results own bigger metric than the retargeted results of the seven methods in RetargetMe. According to the ORIA method, bigger metric means better result. So, we can conclude that our method is better than the seven retargeting methods in Retargete.

The proposed scheme performs worse for the images 'Brickhouse' and 'Foliage', which is because of that these images have no obvious salient areas.

5 Conclusions

By incorporating the developed composite measurement, we propose a new aesthetic-based image retargeting algorithm, which employs the aesthetic-based cropping to obtain the optimal candidate retargeted image with maximum composite measurement and uses the homogenous scaling to generate the retargeted image of target size. Both the subjective and the objective assessments demonstrate that the proposed scheme results in preferable aesthetics of retargeted images and obtain significant improvement in visual effect than SCL, SBC, OPC, SV, SNS, SC, MULTITOP, WARP and SM.

Although the proposed scheme achieves desirable aesthetics, it has slight weakness in losing small objects. As shown in Fig. 15*f*, the duck in the original image is not contained in the retargeted image. To preserve all objects including the important and trivial ones, however, it would degrade the preservation of important objects and even introduce distortions, as shown in Figs. 15*b*-–15*e*. In other words, it is more preferable to well preserve the important objects while losing the trivial ones.

6 Acknowledgments

This work was supported by National Science Fund of China (no. 61202293), NSFC-Guangdong Joint Fund (U0935004, U1135003), the National Key Technology R&D Program (2011BAH27B01), National Science Fund of China (no. 61103162) and the Fundamental Research Funds for the Central Universities of China (no. 1109021170001137105).

7 References

1 Santella, M., DeCarlo, A.D., Salesin, D., Cohen, M.: 'Gaze-based interaction for semiautomatic photo cropping'. Proc. Conf. on Human Factors in Computing Systems, 22–27 April 2006
2 Suh, B., Ling, H., Bederson, B., Jacobs, D.: 'Automatic thumbnail cropping and its effectiveness'. Proc. 16th Annual ACM Symp. on User Interface Software and Technology, 2–5 November 2003
3 Chen, L.Q., Xie, X., Fan, X., Ma, W.Y., Zhang, H.J., Zhou, H.Q.: 'A visual attention model for adapting images on small displays', *ACM Multimedia Syst*., 2003, **9**, (4), pp. 353–364
4 Xiao, J., Zhang, X.M., Cheatle, P., Gao, Y.L., Atkins, C.B.: 'Mixed-initiative photo collage authoring'. Proc. Multimedia, 2008, pp. 509–518
5 Avidan, S., Shamir, A.: 'Seam carving for content-aware image resizing', *ACM Trans. Graph*., 2007, **26**, (3), pp. 1–10
6 Rubinstein, M., Shamir, A., Avidan, S.: 'Improved seam carving for video retargeting', *ACM Trans. Graph*., 2008, **27**, pp. 1–9
7 Achanta, R., Süsstrunk, S.: 'Saliency detection for content-aware image resizing'. Proc. 2009 IEEE Int. Conf. on Image Processing, Cairo, Egypt, 7–10 November 2009
8 Pritch, Y., Kav-Venaki, E., Peleg, S.: 'Shift-map image editing'. Proc. 12th IEEE Int. Conf. on Computer Vision, Kyoto, Japan, 27 September–2 October 2009
9 Cho, T.S., Butman, M., Avidan, S., Freeman, W.T.: 'The patch transform and its applications to image editing'. Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 24–26 June 2008
10 Wang, Y.S., Tai, C.L., Sorkine, O., Lee, T.-Y.: 'Optimized scale-and-stretch for image resizing', *ACM Trans. Graph*., 2008, **27**, (5), pp. 1–8
11 Guo, Y.W., Liu, F., Shi, J.: 'Image retargeting using mesh parametrization', *IEEE Trans. Multimedia*, 2009, **11**, pp. 856–867
12 Wolf, L., Guttmann, M., Cohen-Or, D.: 'Non-homogeneous content-driven video retargeting'. Proc. 11th IEEE Int. Conf. on Computer Vision, Rio de Janeiro, Brazil, 14–20 October 2007
13 Zhang, G.X., Cheng, M.M., Hu, S.M., Martinx, R.R.: 'A shape-preserving approach to image resizing'. Proc. 17th Pacific Conf. on Computer Graphics and Applications, Jeju, Korea, 7–9 October 2009
14 Krähenbü, H.L.P., Lang, M., Hornung, A., Gross, M.: 'A system for retargeting of streaming video', *ACM Trans. Graph*., 2009, **28**, (5), pp. 1–10
15 Karni, Z., Freedman, D., Gotsman, C.: 'Energy based image deformation', *Comput. Graph. Forum*, 2009, **28**, (5), pp. 1257–1268
16 Rubinstein, M., Shamir, A., Avidan, S.: 'Multi-operator media retargeting', *ACM Trans. Graph*., 2009, **28**, (3), pp. 1–11
17 Dong, W.M., Zhou, N., Paul, J.C., Zhang, X.P.: 'Optimized image resizing using seam carving and scaling', *ACM Trans. Graph*., 2009, **28**, (5), pp. 1–10
18 Liu, L.G., Chen, R.J., Wolf, L., Cohen-Or, D.: 'Optimizing photo composition', *Comput. Graph. Forum*, 2010, **29**, (2), pp. 469–478
19 Liu, L., Jin, Y., Wu, Q.: 'Realtime aesthetic image retargeting'. Proc. Eurographics Workshop on Computational Aesthetic in Graphics, Visualization, and Imaging, 2010, pp. 1–8
20 Grill, T., Scanlon, M.: 'jPhotographic composition' (Watson-Guptill, 1990)
21 Krages, B.: 'jPhotography: the art of composition' (Allworth Press, 2005)
22 Datta, R., Joshi, D., Li, J., Wang, J.Z.: 'Studying aesthetics in photographic images using a computational approach'. Proc. ECCV, 2006
23 Luo, Y., Tang, X.: 'Photo and video quality evaluation: focusing on the subject'. Proc. European Conf. Computer Vision (ECCV), 2008, pp. 386–399
24 Cheng, B., Ni, B., Yan, S., Tian, Q.: 'Learning to photograph' (ACM Multimedia, 2010), pp. 192–130
25 Li, C., Loui, A., Chen, T.: 'Towards aesthetics: a photo quality assessment and photo selection system'. Proc. ACM Multimedia, 2010
26 Bhattacharya, S., Sukthanker, R., Shah, M.: 'A framework for photo-quality assessment and enhancement based on visual aesthetics'. Proc. ACM Multimedia, 2010
27 Harel, J., Koch, C., Perona, P.: 'Graph-based visual saliency'. Proc. 20th Annual Conf. on Neural Information Processing Systems, Vancouver, Canada, 4–9 December 2006
28 Yan, J., Zhu, M., Liu, H.X., Liu, Y.C.: 'Visual saliency detection via sparsity pursuit', *IEEE Signal Process. Lett*., 2010, **17**, (8), pp. 739–742
29 Rubinstein, M., Gutierrez, D., Sorkine, O., Shamir, A.: 'A comparative study of image retargeting', *ACM Trans. Graph*., 2010, **20**, (6), pp. 1–10
30 Liu, Y., Luo, X., Xuan, Y.M., Chen, W.F., Fu, X.L.: 'Image retargeting quality assessment'. Computer Graphics Forum (Eurographics'11), 2011, vol. 30, no. 2, pp. 583–592
31 Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: 'Image quality assessment: from error visibility to structural similarity', *IEEE Trans. Image Process*., 2004, **13**, (4), pp. 600–612