

Objective Quality Prediction of Image Retargeting Algorithms

Yun Liang, Yong-Jin Liu, and Diego Gutierrez

Abstract—Quality assessment of image retargeting results is useful when comparing different methods. However, performing the necessary user studies is a long, cumbersome process. In this paper, we propose a simple yet efficient *objective* quality assessment method based on five key factors: i) preservation of salient regions; ii) analysis of the influence of artifacts; iii) preservation of the global structure of the image; iv) compliance with well-established aesthetics rules; and v) preservation of symmetry. Experiments on the RetargetMe benchmark, as well as a comprehensive additional user study, demonstrate that our proposed objective quality assessment method outperforms other existing metrics, while correlating better with human judgements. This makes our metric a good predictor of subjective preference.

Index Terms—Image retargeting, quality assessment, similarity and aesthetic measure, symmetry

1 INTRODUCTION

IMAGE retargeting, which adjusts an image into arbitrary sizes such that the image can be displayed on screens of different sizes, has received much attention in recent years [1], [2]. Many retargeting methods have been proposed, although a single method that works well on any image still does not exist. Instead, different images favor different retargeting algorithms (Fig.1), which makes it difficult to predict a priori which method will work best on which image. An efficient objective quality assessment for image retargeting would thus be useful to select the best result from a pool of retargeted results given a single input image, without the need to rely on costly user studies. Furthermore, this could also assist in developing new image retargeting strategies.

Existing image retargeting quality assessment methods can be coarsely classified into subjective [3], [4] and objective methods [5], [6], [7]. Subjective methods are usually cumbersome and time-consuming, since they require repeated votes by many participants over relatively large combinations of results. Therefore, similar to existing image quality metrics in other domains (see [8] for a recent review), objective methods are a desired tool for assessing retargeted results fast and automatically. However, as summarized in the next section, existing state-of-the-art objective methods are limited in their underlying image analysis. In this paper, we propose a novel objective metric



Fig. 1: Different input images favor different retargeting methods. By subjective evaluation [4], nonhomogeneous warping (WARP [22]) produces better results than shift-maps (SM [25]) in (a), while SM performs better than WARP in (b). For the first image, SM outputs an unrealistic compression of the back of the car, whereas for the second one it preserves the heart shape better than WARP.

for quality assessment of retargeted images based on five key critical factors that define image quality for a retargeted result, selected by carefully analyzing existing retargeting methods and their outcomes. These factors are the following (see Fig. 2):

- **Preservation of salient regions:** These salient regions dictate viewing patterns when looking at an image. Alterations to such patterns in the retargeted image should be minimized.
- **Influence of introduced artifacts:** Newly introduced artifacts are one of the most damaging factors when judging the quality of retargeted results.
- **Preservation of the global structure:** Changes in the global structure of the image usually yield

Y. Liang is with the College of Mathematics and Informatics, South China Agricultural University, Guangzhou, China (E-mail: sdliangyun@163.com).

Y.-J. Liu is with the TNLIS, Department of Computer Science and Technology, Tsinghua University, Beijing, China. Corresponding author (E-mail:liuyongjin@tsinghua.edu.cn).

D. Gutierrez is with the Graphics and Imaging Lab, Universidad de Zaragoza, Spain (E-mail:diegog@unizar.es).

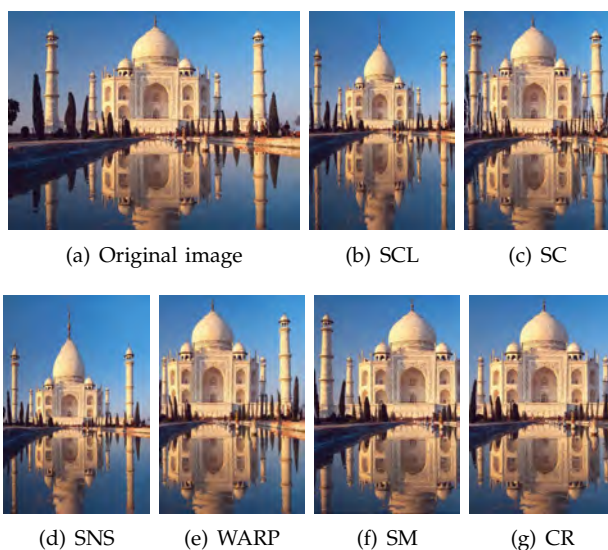


Fig. 2: Original image (a), plus the results of six popular retargeting methods (b through g): Linear scaling (SCL), seam-carving (SC [21]), scale-and-stretch (SNS [23]), non homogeneous warping (WARP [22]), shift-maps (SM [25]) and cropping (CR). For this particular image, it can be seen how some results do not preserve all salient regions (WARP, SM, CR); introduce artifacts (SCL, SNS); do not preserve the global structure (SC, WARP, SM); alter the original composition (SCL, WARP, SM, CR); or break the original symmetry (WARP, SM).

incorrect inter-relationships between objects, thus altering semantics.

- **Aesthetics:** Well-known rules for image composition [9], [11] might be changed in the retargeted results.
- **Preservation of symmetry:** Broken symmetric features are quickly detected as undesired artifacts by human observers.

The first three factors define the *similarity* component of our metric, or how well the original content is preserved in the result. In previous work [3], similarity is compared by analyzing the shape of saliency maps resulting from eye-tracking experiments; our objective metric removes the need for these experiments, which require specialized equipment, providing a definition of similarity that closely addresses the three main goals of image retargeting algorithms: preserving content, limiting artifacts and preserving structure [4]. Note that since global structure information may not contain any clear salient region (see Fig. 3), both factors need to be computed separately. Last, we rely on the field of computational aesthetics to define well-known compositional rules that can be objectively measured [12], [13], [9], and go beyond a simple preservation of symmetry. All these factors are explained in detail in Section 3.

In our quality assessment metric, we first measure the preservation of saliency by computing the changes of the salient areas and the color variations; a weighted bidirectional similarity metric is used

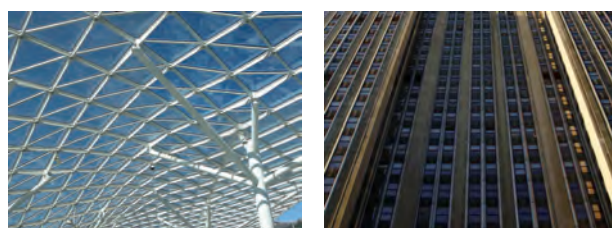


Fig. 3: Two examples of images without a clear salient region, but with well-defined structure.

to measure the influence of artifacts, while global structure preservation is estimated by means of a combination of two existing image similarity metrics (SSIM [14] and HDR-VDP2 [15]). Secondly, an aesthetic metric is defined by measuring the changes in image composition using several well-established rules of aesthetics. Last, a symmetry measure is proposed to evaluate the how well symmetric structures are preserved. The final quality assessment is the result of an optimally weighted linear combination of the partial similarity, aesthetic and symmetry metrics. Extensive experimental results on the RetargetMe benchmark [4], as well as a comprehensive user study, demonstrate that our objective metric¹ outperforms other objective quality assessment methods [6], [16], [17], [18], [19], [20], showing a much higher agreement with human preference.

2 RELATED WORK

Many different image retargeting algorithms have been published over the past few years, such as seam carving [21], nonhomogenous warping [22], scale-and-stretch [23], patchwise scaling [24], multi-operators [17], optimized resizing [26], content-aware resizing [27], [28] or symmetry summarization [29]. Most of them share the common strategy of defining some sort of “energy” map over the 2D manifold of the image, which encodes what areas of the images need to be preserved during retargeting. This energy can be defined in terms of image gradients, salient areas, aesthetic considerations or a combination of several of them. We refer the reader to some recent courses [1], [2], [30] and surveys [31] for a complete overview of existing algorithms, and focus here on quality metrics.

Several quality assessment techniques have been proposed to compare image retargeting methods. **Subjective methods** are based on analyzing the preference of participants in carefully designed user studies. A publicly available benchmark called RetargetMe, composed of a large number of original and retargeted images computed by several state-of-art methods, was presented in [4], along with the first in-depth study of the subjective preference of a large

1. Our code will be available at <http://cg.cs.tsinghua.edu.cn/people/~Yongjin/yongjin.htm>.

number of participants. Later, eye tracking was used to compare image retargeting methods through the examination of gaze fixations and viewing patterns by Castillo et al. [3]. The working hypothesis is that changing patterns and fixation points reflect unwanted changes in the salient regions of the original and the retargeted image, which should be preserved during retargeting. These subjective methods work well and offer valuable insights, but they are costly to carry out.

Objective methods have also been proposed, both in the general context of image manipulation and for image retargeting in particular. The edge histogram (EH) [18] and the color layout (CL) [19] define image distances to assess the content similarity in two images. The earth-mover's distance algorithm (EMD) [32] and SIFT-flow (SFlow) [20] were used in [4] to assess image retargeting methods, since they are robust to capturing the structural properties inherent in images. They performed well when measuring the preservation of salient regions but tend to underestimate the influence of artifacts. Ma et al. [7] further verified the efficiency of the above methods by fusing EH, CL and SIFT-flow, concluding that the combination of these methods performs better than any single one of them in isolation, for image retargeting quality assessment. Simakov et al. [16] proposed a bidirectional similarity to describe the coherence and completeness between input and output images. This method is efficient in measuring the influence of artifacts, but not in the preservation of salient regions. Rubinstein et al. [17] proposed a bi-directional warping distance based on the non-symmetric variant of dynamic time warping to measure the similarity between each row/column or patch of an image and its retargeted image. This method efficiently controls artifacts by optimizing the alignment of an original image with its retargeted image, but it does not take into account changes in the composition of the layout from an aesthetics perspective. Liu et al. [6] used a top-down simplified model of the human vision system to define a saliency-based image similarity metric in the CIE Lab color space. Fang et al. [33] applied a spatial pooling method to provide dense pixel correspondences in the reference and retargeted images, and applied the SSIM measure to each pair of corresponding pixels to define an overall image retargeting SSIM index. Hsu et al. [34] used a local variance of SIFT flow vector field to measure the geometric distortion of a retargeted image. These methods [6], [33], [34] can identify loss or distortions of salient regions in the original image, but again they do not take aesthetics into account. Last, used in the context of deblurring algorithms, Masia et al. [35] propose an image quality measure which combines the L_2 norm, the SSIM index and the HDR-VDP-2 metric. The SSIM index performs well in measuring the structure similarity between two images, while the HDR-VDP-

2 produces a good estimation on the overall quality of an image; since both of these aspects are important in image retargeting, we employ these two metrics as well when measuring structure preservation.

3 OBJECTIVE ASSESSMENT METHOD

We assess the quality of a retargeted image T with its original image I as reference. Our quality assessment consists of three parts: a similarity measure $Q_{sim}(I, T)$, an aesthetics measure $Q_{aes}(I, T)$, and a symmetry measure $Q_{sym}(I, T)$. $Q_{sim}(I, T)$ further consists of three subparts, dealing with the preservation of salient regions $Q_{sal}(I, T)$, the influence of artifacts $Q_{art}(I, T)$, and the preservation of the global structure $Q_{str}(I, T)$. Given that Q_{sym} only works on input images with symmetric features, we first define a general quality metric $Q(Q_{sal}, Q_{art}, Q_{str}, Q_{aes})$ (Eq.1), and then extend it to $Q'(Q, Q_{sym})$ to include symmetry (Eq.2). Both Q and Q' can be interpreted as distance functions between T and I , defined in the $[0, 1]$ domain, where 0 means $T = I$ and smaller value means better retargeted result.

To represent the function Q , we make the basic assumption that Q_{sal} , Q_{art} , Q_{str} and Q_{aes} show mutual preferential independence. In a deterministic preference structure, three attributes X_1, X_2, X_3 are preferentially independent of a fourth attribute X_4 if the preference between outcomes $\langle x_1, x_2, x_3, x_4 \rangle$ and $\langle x'_1, x'_2, x'_3, x_4 \rangle$ does not depend on the particular value x_4 for attribute X_4 [38]. Then, by the Debreu's theorem [39], the preference behavior can be described by minimizing the additive value function Q :

$$Q = \omega_{sal}Q_{sal} + \omega_{art}Q_{art} + \omega_{str}Q_{str} + \omega_{aes}Q_{aes} \quad (1)$$

For the images with symmetry features, a supplementary metric Q_{sym} is used:

$$Q' = \omega_Q Q + \omega_{sym}Q_{sym} \quad (2)$$

The optimal values of weights ω_{sal} , ω_{art} , ω_{str} , ω_{aes} , ω_{sym} and ω_Q will be specified in Section 4.1. In the following sections, we will introduce our three main components of the metric, namely similarity, aesthetics and symmetry.

3.1 Similarity Component

The similarity component evaluates to what extent the important contents and the structures are preserved, and to what extent some new artifacts are introduced. For this, we define three quality terms: Q_{sal} for salient region preservation, Q_{art} for the influence of artifacts, and Q_{str} for global structure preservation.

3.1.1 Q_{sal} for salient region preservation

We assume that preserving the salient regions of an original image is a desired feature of any image retargeting algorithm [3]. In the field of image retargeting, Castillo et al. [3] compute salient regions by

analyzing eye-tracking data. This method is obviously accurate, but time-consuming. Many methods have been proposed to *automatically* predict such salient regions (e.g [6], [23], [40]). In this paper, we follow the approach of Cheng et al. [41], who used global contrast differences and spatial coherence information to efficiently separate salient large-scale objects from their low-saliency surroundings. This method has been evaluated on the largest publicly available data set, consistently producing good results.

In principle, we could measure the preservation of salient regions by taking into account the change of the salient areas between the original image I and its retargeted result T . We measure such change as the normalized difference $Q_{area} = |S_I - S_T| / \max(S_I, S_T)$, where S_I and S_T represent the areas of the original and retargeted salient regions, respectively. However, this naive approach would not always produce good results, since sometimes the areas of salient regions may happen to be close, but the *content* of such regions might have changed during retargeting. We thus rely on common image understanding strategies, and also take into account variations in content as changes in the color histogram (in HSV space) of the region [42].

For an efficient representation, we quantize the HSV space and map it into a one-dimensional, 256-bin color histogram C as [43]. In their work, a quantizer Q_c^M (where $M = N_h \times N_s \times N_v$) is proposed to describe any HSV color by N_h hues, N_s saturations and N_v values (we use 16, 4 and 4 respectively). This way, each color in HSV space is assigned to a unique index in M dimensions. Given the quantized values of pixel p as $Q_c^M(h)$, $Q_c^M(s)$ and $Q_c^M(v)$, we can compute its color by $C(p) = Q_c^M(h) \times N_s \times N_v + Q_c^M(s) \times N_v + Q_c^M(v)$.

Let h_I and h_T be the color histograms in C space describing the salient regions of I and T . We measure their color difference as $Q_{color} = \frac{1}{2} \sqrt{\sum_{i=1}^{255} (h'_I - h'_T)^2}$, where h'_I and h'_T represent normalized histograms. Q_{sal} is finally defined as:

$$Q_{sal} = \omega_{area} Q_{area} + \omega_{color} Q_{color} \quad (3)$$

where $\omega_{area} + \omega_{color} = 1$ and $\omega_{area}, \omega_{color} > 0$. The optimal values of ω_{area} and ω_{color} are discussed in Section 4.1. Fig. 4 shows some examples of Q_{area} and Q_{color} for some retargeted results.

3.1.2 Q_{art} for artifacts influence

Visual artifacts in the form of distortions or broken structures may appear in the resulting retargeted image. The recently proposed bidirectional similarity metric [16] captures a quantitative measure of the potential visual artifacts introduced during retargeting, by comparing small patches from the original and retargeted images. Two images are considered to be visually similar if as many small patches as possible from one are included in the other, and vice versa.



Fig. 4: Example values of Q_{area} and Q_{color} for retargeted results. The left image of each sub-image is the original image while the right is its salient region by Cheng's method [41]. Multi-operator (MOP): $Q_{area} = 0.46$, $Q_{color} = 0.11$. Scale-and-stretch (SNS): 0.66 and 0.22. Shift-maps (SM): 0.49 and 0.73. Seam-carving (SC): 0.46 and 0.15.

However, the original formulation of the metric does not take into account the influence of the salient areas in the images. Castillo et al. [3] analyzed eye tracking data from many viewers and found that relatively large artifacts outside the main salient areas tend to go unnoticed much longer than artifacts in the main salient regions. We thus modify the original bidirectional similarity metric to take into account the influence of saliency, defining our measure of artifacts influence as:

$$Q_{art} = 0.5 \frac{\frac{1}{N_I} \sum_{R \subset I} S_R \min_{Q \subset T} D(R, Q)}{\max_{R \subset I} (S_R \min_{Q \subset T} D(R, Q))} + 0.5 \frac{\frac{1}{N_T} \sum_{Q \subset T} S_Q \min_{R \subset I} D(Q, R)}{\max_{Q \subset T} (S_Q \min_{R \subset I} D(Q, R))} \quad (4)$$

where R and Q are 3×3 patches from the original and retargeted images respectively, N_I and N_T are the number of patches in such original and retargeted images, and D is the distance measure between two patches as defined in [16]. The saliency weights S_R and S_Q are given by the average of the saliency values of all pixels contained in patches R and Q . These saliency values are computed by the method of Cheng [41].

3.1.3 Q_{str} for structure preservation

A good retargeting method should also preserve the global structure of the original image as much as possible. However, measuring the preservation of the global structures between two images (before and after resizing) is challenging. On the one hand, it is well known that objective metrics working at pixel level, such as the L_2 norm, do not perform well when measuring higher level attributes like structure. On the other hand, more sophisticated metrics such as the Structural Similarity Index Measure (SSIM [14]) can only be used to measure quality degeneration after registration between the original and the modified image, usually without changes in content.

Since the original image and its retargeted result have different sizes, we first establish structure-aware dense pixel correspondences between the original and retargeted images. Instead of directly searching the whole retargeted image for each pixel of the original image, we use a structure-aware pixel mapping scheme relying on *scale-space theory* [44], [6]. Briefly (please refer to the original publications for further details), a convolution of a Gaussian function $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$ with image I is applied, and a difference-of-Gaussian image D is obtained as $D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$. The Gaussian image $G(x, y, \sigma) * I(x, y)$ is down-sampled by a factor of 2 and the process is repeated to produce an image scale space $(I, D^1, D^2, \dots, D^n)$. We follow [6] to set the parameters $k = \sqrt{2}$ and $n = \lfloor m \rfloor$ where m is the minimum number of (a, b, c, d) , while $(a \times b)$ is the original image size and $(c \times d)$ is the retargeted size.

Applying this to both our original image I and its retargeted image T , we obtain the image scale spaces $(I, D^1_{ori}, D^2_{ori}, \dots, D^n_{ori})$ and $(T, D^1_{ret}, D^2_{ret}, \dots, D^n_{ret})$, from which a hierarchical pixel match is performed. First, the pixels in D^n_{ori} and D^n_{ret} are matched and then propagated to level $(n - 1)$ as an inter-scale constraint. These pixel matches are further fine tuned in a 5×5 local window as an intra-scale constraint to give accurate pixel matches at level $(n - 1)$. This process is repeated until the I and T levels are reached. Note that the inter-scale constraint offers a consistent image structure correspondence in a top-down manner and the intra-scale constraint offers structure-aware accurate pixel matches.

After this process, the global structure similarity is evaluated by a weighted summation of local similarity windows from every pair of pixel correspondences. Based on the mapping relationship of pixels between an original image and a retargeted result, we define Q_{str} based on the SSIM [14], and the more recent version of the Visual Differences Predictor (VDP2 [15]). SSIM compares the structural similarity between two images by analyzing the degradation of structural information between corresponding win-

dows in images, while HDR-VDP-2 (from here on, VDP2) includes a model of human perception to predict the overall quality of an image, compared to a given reference. For both, larger values mean better results. To match the measures of Q_{sal} and Q_{art} , where smaller values mean better results, we compute $(1 - SSIM(p, p'))$ and $(1 - VDP2(p, p')/100)$ to evaluate the structure preservation. Let pixel p_i be the i^{th} pixel of the original image I and pixel p'_i its corresponding pixel in the retargeted image T , we then define Q_{str} as:

$$Q_{str}(I, T) = \omega_{ssim} \sum_{i=1}^{Nt} (1 - SSIM(p_i, p'_i)) + \omega_{vdp} \sum_{i=1}^{Nt} (1 - VDP2(p_i, p'_i)/100) \quad (5)$$

where $\omega_{ssim} + \omega_{vdp} = 1$, $\omega_{ssim}, \omega_{vdp} > 0$ (computed in section 4.1) and Nt is the numbers of pixels of the retargeted image T . Smaller Q_{str} means better retargeted result.

3.2 Aesthetics component

Inspired by the field of computational aesthetics [12], [13] and its applications in image composition [9], [11], we incorporate a computational measure of image aesthetics into our quality assessment metric. Specifically, we focus on two rules that have been empirically shown to increase image aesthetics in most cases [9]:

- Rule of thirds. This rule divides an image into nine parts of equal size by equally spacing two horizontal and two vertical lines. These lines define four intersection points as shown in the red points in Fig. 5 (left). When the salient regions lie near to them, the perceived aesthetics of an image generally improved (see the sun and boat in Fig. 6 (a)).
- Visual balance. The image is regarded to be visually balanced if the center of mass of all salient regions is close to the image center (see the spider in Fig. 6 (b)).

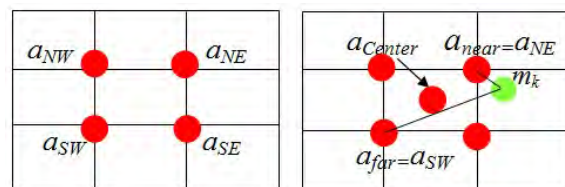
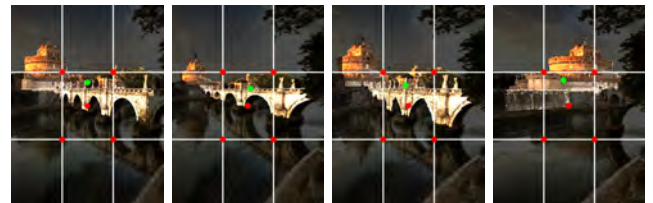


Fig. 5: Illustration of the Q_{aes} computation. Left: the rule of thirds and intersection points. Right: The power points are defined as the union of intersection points and the image center (a_{Center}). Supposed m_k is the center of a salient region, its nearest power point a_{near} is the a_{NE} while its farthest power point a_{far} is the a_{SW} .



(a) Rule of thirds (b) Rule of center

Fig. 6: Representative images for each aesthetic rule.



(a) MOP (b) SNS (c) SC (d) SM

Fig. 7: Example values of Q_{aes} for some retargeted results. The red points are the power points, while the green points are the centers of salient regions. By Q_{aes} , SNS ($Q_{aes}=0.30$) performs better than MOP ($Q_{aes}=0.38$), SC ($Q_{aes}=0.35$) and SM ($Q_{aes}=0.32$).

Note that there are many other (sometimes competing) rules devised to compute the aesthetic measure of an image, such as diagonal dominance, color-related rules, or rules defining different salient-region sizes. However, our goal is to obtain an aesthetics component that is as simple as possible and serves our goal in the context of image retargeting. Since we also need to integrate its value coherently with Q_{sal} and Q_{art} , we propose the following simple variant of the aesthetic score functions defined in [9].

Our aesthetic metric Q_{aes} consists of two parts, Q_{third} for rule of thirds and Q_{bal} for visual balance. Let the set $A = (a_{NE}, a_{NW}, a_{SE}, a_{SW})$ denote the power points in the rule of thirds, and a_{Center} be the image center (Fig. 5, right). The Q_{third} is computed as follows: In the retargeted image, we first compute the center position m_k of each of k salient regions such as the green point in Fig. 5. We then compute the positions of the power points in A , and find the nearest point a_{near} to m_k and the farthest point a_{far} to m_k as shown in Fig. 5, right. Then

$$Q_{third} = \frac{\|m_k - p(a_{near})\|_2}{\|m_k - p(a_{far})\|_2} \quad (6)$$

where $p(a_i)$ denote the position of a_i and $\|\cdot\|_2$ denotes the L_2 norm. Similarly, we define Q_{bal} by

$$Q_{bal} = \frac{1}{HL} \|m - p(a_{Center})\|_2 \quad (7)$$

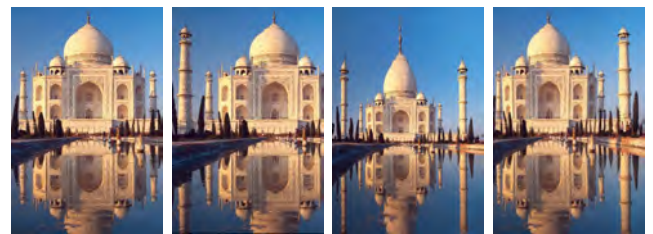
where m is the center of mass of all salient regions, HL means the half length of diagonal line of a retargeted image and is used to do normalization.

Finally, we obtain Q_{aes} as $Q_{aes} = 0.5Q_{third} + 0.5Q_{bal}$. Note that the measure Q_{aes} is normalized in the range $[0, 1]$, and smaller values mean better compliance with aesthetics rules, and better retargeted result (such as the SNS result in Fig. 7).

3.3 Symmetry component

Symmetry is one of the most important structural features in images. In fact, broken symmetries are usually one of the easiest-to-spot artifacts in retargeted images [4] (see Fig. 8). Recently, some retargeting methods have been specially designed to preserve symmetric features [29].

To take symmetry into account, we first detect the symmetry regions. According to [29], the symmetric



(a) CR (b) SM (c) SNS (d) WARP

Fig. 8: Example values of Q_{sym} for image Taj Mahal. The CR ($Q_{sym} = 0.37$) and the SNS ($Q_{sym} = 0.55$) methods preserve image symmetry better than WARP ($Q_{sym} = 0.59$) and SM ($Q_{sym} = 0.57$).

regions in an image usually satisfy the following features: first, they possess similar distinguishing, invariant and stable properties; second, their intensities are very similar or satisfy some monotonic transformation. Therefore, the areas covering "windows" in Fig.9 (denoted by red ellipses in Fig.9 (c)) are symmetric regions for possessing similar properties under some affine transformation. Many methods can be used to identify symmetric regions such as Harris corners [46], or SIFT points [47]. We use maximally stable extremal regions (MSER) [48], which have been recently shown to be effective at detecting symmetric regions for retargeted images [29]. The MSER method can efficiently identify regions with similar content even after affine transformations which often occur in image retargeting. However, the regions detected by the MSER are some ellipses with overlapped or trivial regions such as Fig.9 (b). Therefore, we use adaptive mean-shift clustering [49] to group detected regions, and select as symmetric regions the clusters with the largest covering areas (Fig.9 (c)).

Suppose the final symmetric regions of a retargeted result T are given by $R = \{r_i(c_i, u_i, v_i), i = 1, 2, \dots, N_s\}$, where r_i describes an ellipse and (c_i, u_i, v_i) are its center, major axis vector and minor axis vector, respectively. For example, the red ellipse in the yellow rectangle in Fig.9 (c) is a symmetric region, with the black point as its center, the blue line as its major axis, and the green line as its minor axis. We then evaluate the image quality introduced

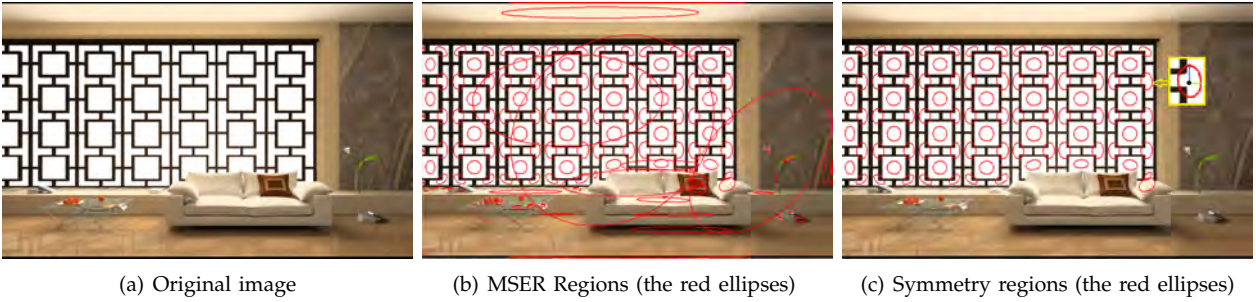


Fig. 9: Symmetric regions detection. After obtaining MSER regions such as (b), we cluster them by the adaptive mean-shift clustering to select the major cluster as symmetry regions such as (c) while discarding the trivial and overlapped regions.

by symmetry preservation of T by accumulating the minimum symmetry distances of all its symmetric regions. First, we define the symmetry distance D_{sym} from its symmetric region r_m to r_n by:

$$D_{sym}(r_m, r_n) = \frac{1}{N_m} \sqrt{\sum_{i=1}^{N_m} (P_i(r_m) - P_i(F(r_n)))^2} + \sqrt{\left(\frac{|u_m - u_n|}{u_m + u_n} - \frac{|v_m - v_n|}{v_m + v_n} \right)^2} \quad (8)$$

where N_m is the number of pixels in r_m , $P_i(x)$ is the i^{th} pixel in region x , u_m and v_m are the axis vectors of r_m , u_n and v_n are the axis vectors of r_n , and $F(r_n)$ is an affine transformation to make r_n the same size as r_m .

If two regions satisfy the symmetry relationship, they should have similar shape and content. Therefore, our symmetry distance is evaluated by two similarity measures, namely the intensity similarity (the first term in Eq.8) and the shape similarity (the second term in Eq.8). The intensity similarity, computed by the intensity difference, describes the content consistency between them. For the regions satisfying the symmetry relationship usually formed by affine transformation, we first transform r_n by F to the size of r_m when compute the symmetry distance from r_m to r_n . Similarly, the r_m will be transformed to the size of r_n when computing the symmetry distance $D_{sym}(r_n, r_m)$ from r_n to r_m . The shape of a symmetric region is determined by its major and minor axes; we measure the shape similarity by the length differences between the axis vectors as shown in Eq.8.

For each symmetry region of a retargeted image, we compute the symmetry distances to all other symmetry regions, and select its minimum symmetry distance. This minimum symmetry distance is related to the most similar region that forms a symmetry relationship with it. Then, we define the Q_{sym} , which describes the symmetry preservation of the whole retargeted image by accumulating all the minimum

symmetry distances of its symmetric regions, as:

$$Q_{sym} = \frac{1}{N_s} \sum_{r_m \in R} \min_{r_n \in R} D_{sym}(r_m, r_n) \quad (9)$$

where R is the set of symmetry regions of a retargeted image T , and N_s is the number of symmetric regions in R . Smaller Q_{sym} values mean fewer differences between symmetric regions, which leads to better symmetry preservation (see Fig.8).

3.4 Final measure

To use our objective metric, the user simply indicates whether symmetry must be taken into account, given a pair of original and retargeted images. This can also have been previously tagged: For example, the RetargetMe dataset [4] already contains information about symmetry. Then we define

$$(Q, Q') = \begin{cases} Q' & \text{if the image has symmetric features} \\ Q & \text{otherwise} \end{cases} \quad (10)$$

where Q is defined in Eq.(1) and Q' is defined in Eq.(2).

4 RESULTS AND EVALUATION

To evaluate our results, we use the 37 images in the RetargetMe database [4] in our experiment. These images are classified into six image types, including lines/edges (25 images), faces/people (15 images), foreground objects (18 images), geometric structures (16 images), symmetry (6 images) and texture (6 image). Note that one image may belong to several types. For each image, there are eight retargeting results produced by state-of-the-art methods, including: simple cropping (CR), multi-operator (MOP) [17], streaming video (SV) [50], shift map (SM) [25], non-homogenous warping (WARP) [22], seam carving (SC) [21], simple scaling (SCL) and scale-and-stretch (SNS) [23].

As observed in [4], the quality of retargeted results cannot be represented in a linear scale and the paired comparisons technique was suggested to replace the traditional ranking methods for quality assessment. We follow the method proposed in [4] to use the

Kendall correlation coefficient [51] to measure the degree of agreement between an objective assessment and the subjective assessment, which is computed as follows: First, for an original image I , we build a subjective score vector $s = (s_1, s_2, \dots, s_8)$ for the eight retargeting methods based on the subjective scores in RetargetMe, i.e., s_i is the number of times the retargeted result T_i (computed by the i^{th} retargeting method) is selected as the best by users. Second, we build an objective score vector $o = (o_1, o_2, \dots, o_8)$ for I using our objective assessment method where $o_i = Q(I, T_i)$ represents an objective quality measure. Third, we rank s and o after sorting them. s is sorted in descending order since the higher s_i the better quality T_i , while o is sorted in ascending order since a smaller o_i means better quality. Finally, we compute the Kendall correlation coefficient τ between the two ranked vectors to obtain the rate of agreement between the objective and the subjective assessments:

$$\tau = \frac{1}{m} \left(\sum_{i=1}^m N_i^+ - \sum_{i=1}^m N_i^- \right) \quad (11)$$

where $m = c_n^2$ is the number of compared pairs given the ($n = 8$) retargeting methods. For a given pair i of entries in the ranking, if it is a concordant pair in s and o , we set $N_i^+ = 1$ and $N_i^- = 0$, otherwise $N_i^+ = 0$ and $N_i^- = 1$. By accumulating N_i^+ (N_i^-) we obtain the number of concordant (discordant) pairs over all the pairs of entries in the ranking. Note that $-1 \leq \tau \leq 1$. A higher τ means a better agreement between the objective and the subjective assessments.

4.1 Optimal parameter setting

As the parameters in Eq.2 depends on Eq.1, we first compute the optimal parameters of Eq.1, namely $(\omega_{sal}, \omega_{art}, \omega_{str}, \omega_{aes})$ and $(\omega_{area}, \omega_{color}, \omega_{ssim}, \omega_{hdp})$. To find an optimal set, we define a functional $\tilde{\tau} = (\omega_{sal}, \omega_{art}, \omega_{str}, \omega_{aes}, \omega_{area}, \omega_{color}, \omega_{ssim}, \omega_{hdp})$ in an 8-dimensional space X as follows. The constraints $\omega_{sal} + \omega_{art} + \omega_{str} + \omega_{aes} = 1$, $\omega_{area} + \omega_{color} = 1$, $\omega_{ssim} + \omega_{hdp} = 1$, and the values of these parameters that should be controlled in the range $[0, 1]$ form a hypercube H in X . Each point $p \in H$ defines a deterministic measure $Q(p)$ in Eq.1. Using $Q(p)$, we compute the Kendall correlation coefficient τ of the images in a training set. The value of the functional $\tilde{\tau}(p)$ is defined to be the average of all Kendall correlation coefficients for all the original images in a training set. Then the optimal parameters correspond to the position $p' \in H$ where $\tilde{\tau}(p')$ reaches the maximum.

To find the optimal point p' , we estimate the functional τ using the RBF interpolations. We sample H using the interval 0.05 in each dimension. For each sample point $s_i \in H$, we compute the functional value $\tilde{\tau}(s_i)$. The RBF interpolating function is

$$\tilde{\tau}(x) = \sum_{i=1}^n \omega_i \Phi(x - s_i) \quad (12)$$

where ω_i is the weight for each sample s_i , $x \in H$ and n is the number of sample points in H . We use the Gaussian radial basis function $\Phi(r) = e^{-r^2}$ due to its positive definite property. The weights ω_i are solved by the linear system from the interpolating constraints

$$\tilde{\tau}(s_j) = \sum_{i=1}^n \omega_i \Phi(s_j - s_i), \forall s_j \in H$$

Given the analytical form $\tilde{\tau}(x)$, we find its maximum value in H using the BFGS (Broyden-Fletcher-Goldfarb-Shanno) algorithm in multiple dimensions.

4.2 Comparison of objective metrics

We compare the proposed metric with six other objective methods, namely BDS [16], BDW [17], EH [18], CL [19], SFlow [20] and CSim [6]. BDS and DBW work similarly as the factor Q_{art} defined in Eq.4. EH and CL use signatures of fixed lengths regardless of image size to estimate the image quality. In the previous evaluation by Rubinstein and colleagues [4], SFlow and EMD [32] were found to yield a similar performance, since both use a dense SIFT descriptor. In this study, we choose SFlow as a representative for comparison. CSim is specially designed for assessing image retargeting methods, which simulates the human vision system in a top-down manner; i.e., in the scale space of images, the coarse level is used to evaluate the global structure correspondence and the fine level is used to evaluate the pixel similarity with the constraints of structure correspondence. In our proposed measure, we use the five key factors to assess the quality of a retargeted image: Q_{sal} and Q_{art} , Q_{str} evaluate the preservation of image content and structure, the factor Q_{aes} evaluates aesthetics, and Q_{sym} evaluates the preservation of symmetry feature.

To verify the performance of our metric and compare it to other objective methods, we apply leave-one-out cross validation (LOOCV) in the RetargetMe database. In each fold of our LOOCV, one original image and its eight retargeted results are used as the test set, with the remaining images as the training set. After 37 folds (the number of original images in RetargetMe), each image has thus been used as a test set once. The parameters of our metric are determined as described in Section 4.1.

To estimate how well the objective metrics agree with the participants' subjective preferences, we compute the correlation between rankings produced by each objective metric, and the subjective results from the RetargetMe paper [4]. We use the Kendall correlation coefficient [51] for all seven metrics using the test set. The results are summarized in Table 1, classified according to the image types defined in the RetargetMe database. We also compute a mean Kendall correlation coefficient using all the images in the test set (last column). It can be seen how our metric consistently produces the best results, being a good predictor of subjective users' preferences.

TABLE 1: The mean Kendall correlation coefficients of seven objective metrics, organized by the image types defined in the RetargetMe database. Our metric consistently yields the best results (highlighted in blue).

	Lines/edges	Faces/people	Texture	Foreground objects	Geometric structure	Symmetry	All
BDS	0.040	0.190	0.089	0.167	-0.004	-0.012	0.083
BDW	0.031	0.048	-0.009	0.060	0.004	0.119	0.046
EH	0.043	-0.076	-0.063	-0.079	0.103	0.298	0.004
CL	-0.023	-0.181	-0.089	-0.183	-0.009	0.214	-0.068
SFlow	0.097	0.252	0.161	0.218	0.085	0.071	0.145
CSim	0.091	0.271	0.188	0.258	0.063	-0.024	0.151
(Q, Q')	0.351	0.271	0.304	0.381	0.415	0.548	0.399

We further compare our metric against two other recent ones [33], [34]. Using again all 37 images in RetargetMe, the mean Kendall correlation coefficient of the IR-SSIM metric in [33] is 0.363, smaller than the 0.399 correlation achieved by our metric. Different from Table 1 in which eight retargeted results (CR, MOP, SV, SM, WARP, SC, SCL and SNS) are used to compute the Kendall correlation coefficients, only five retargeted results (MOP, SM, WARP, SC and SCL) are used in [34]. By using the same five retargeted results in RetargetMe, the mean values of the Kendall correlation coefficients of our metric and [34] (*ours*, [34]) in each image type are: lines/edges (0.552, 0.431), faces/people (0.533, 0.390), texture (0.500, 0.286), foreground objects (0.544, 0.389), geometric structure (0.600, 0.438), symmetry (0.567, 0.523) and all (0.567, 0.523). Our metric consistently yields better predictions.

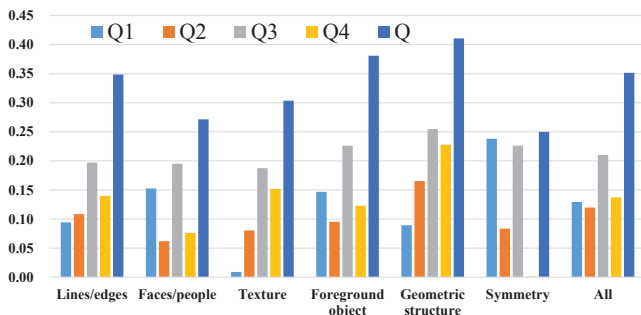


Fig. 10: The mean Kendall correlation coefficients of Q_1, Q_2, Q_3, Q_4 and Q for each image type and all the 37 images.

4.3 Validity of the components in Q and Q'

4.3.1 Components in Q

Four factors $Q_{sal}, Q_{art}, Q_{str}, Q_{aes}$ are considered in Q , as defined in Eqs.(3-6). To analyze the validity each factor, we define the following four alternative measures and compare them with Q :

$$\begin{cases} Q_1 = \omega_{art}Q_{art} + \omega_{str}Q_{str} + \omega_{aes}Q_{aes} \\ Q_2 = \omega_{sal}Q_{sal} + \omega_{str}Q_{str} + \omega_{aes}Q_{aes} \\ Q_3 = \omega_{sal}Q_{sal} + \omega_{art}Q_{art} + \omega_{aes}Q_{aes} \\ Q_4 = \omega_{sal}Q_{sal} + \omega_{art}Q_{art} + \omega_{str}Q_{str} \end{cases}$$

We repeat the LOOCV experiment as presented in Section 4.2, this time replacing the measure Q' by $Q_1,$

TABLE 2: Kendall coefficients of all images tagged as symmetric in the RetargetMe database.

Image name	Q	Q_{sym}	Q'
Johan	0.643	0.643	0.714
St. Angelo	0.643	0.143	0.714
Buddha	-0.071	0.214	0.357
Foliage	0.500	0.143	0.571
Glasses	-0.143	0.786	0.714
Taj Mahal	-0.071	0.143	0.214
Mean	0.250	0.345	0.548

Q_2, Q_3, Q_4 and Q respectively. Then, we compare the Kendall correlation coefficients obtained from Q_1, Q_2, Q_3, Q_4 and Q with the original images as reference images in the RetargetMe benchmark. The results are summarized in Fig. 10, showing that Q performs consistently and significantly better than all the other choices.

4.3.2 Components in Q'

The measure Q' consists of two factors, namely Q (analyzed above) and Q_{sym} , which is designed for images with symmetric features. As described in Section 3.3, our symmetry detection relies on the user's indication about whether or not symmetry is present in the image. We have also tried using suitable automatic state-of-the-art methods [52], [53]: By applying these methods to the data set under Symmetry Detection from the RealWorld Images Competition 2013 [54], the detection rates of [52] and [53] are 65.80% and 69.60% respectively. While reasonably good for automatic algorithms, these percentages are still too low for an applied method like ours.

To study how mistakenly choosing Q' (or Q) in images with clear symmetry (or lack of) will impact our quality assessments, we first repeat the LOOCV experiment presented in Section 4.2, this time replacing the measure Q' by Q and Q_{sym} . Then we compute the Kendall correlation coefficients for all the images in the RetargetMe dataset tagged as symmetric, using the three options: Q', Q and Q_{sym} . The results are summarized in Table 2, showing a much better performance of Q' for these symmetric subset, as expected.



Fig. 11: The new dataset of 17 images for our user study, to analyze the applicability of our metric to a different data set.

TABLE 3: Kendall coefficients of six images that are not tagged as symmetric in the RetargetMe database.

Image name	Q	Q_{sym}	Q'
DKNYgirl	0.643	0.429	0.500
Brick_house	0.571	-0.143	0.071
Butterfly	0.857	0.214	0.214
Car1	0.643	0.214	0.571
Painting2	0.714	0.357	0.286
Surfers	0.643	-0.214	-0.286
Mean	0.679	0.143	0.226

Similarly, Table 3 shows the Kendall correlation coefficients of six images not tagged as symmetric in the database, using Q and Q' . This time, again as expected, Q yields better results than wrongly using Q' .

5 APPLICATION TO NOVEL DATASETS

To demonstrate the applicability of our objective metric to *different* image datasets, we select 17 new images from the RetargetMe database *that lack subjective scores*². Figure 11 shows this new dataset. Each image has eight retargeted results by the same methods presented before: CR, MOP, SV, SM, WARP, SC, SCL and SNS.

Parameters. We compute objective scores using the proposed measures in Section 3, with parameters determined as specified in Section 4.1 on the training set of all 37 original images in RetargetMe. These parameters are:

$$\begin{cases} Q_1 = 0.66Q_{art} + 0.17Q_{str} + 0.17Q_{aes} \\ Q_2 = 0.45Q_{sal} + 0.38Q_{str} + 0.17Q_{aes} \\ Q_3 = 0.8Q_{sal} + 0.1Q_{art} + 0.1Q_{aes} \\ Q_4 = 0.31Q_{sal} + 0.03Q_{art} + 0.66Q_{str} \\ Q = 0.45Q_{sal} + 0.38Q_{art} + 0.1Q_{str} + 0.07Q_{aes} \end{cases} \quad (13)$$

2. There are 80 images in the RetargetMe database. Only 37 of those (the ones we used in Section 4) have subjective preference scores; the rest of the images were not used in the RetargetMe publication and thus do not have subjective scores.

and

$$Q' = 0.13Q + 0.87Q_{sym} \quad (14)$$

where the $(\omega_{area}, \omega_{color})$ for Q_2 , Q_3 , Q_4 and Q are $(0.5, 0.5)$, $(0.6, 0.4)$, $(0.3, 0.7)$, and $(0.6, 0.4)$, respectively. The $(\omega_{ssim}, \omega_{hdp})$ for Q_1 , Q_2 , Q_4 and Q are $(0.9, 0.1)$, $(0.9, 0.1)$, $(0.8, 0.2)$ and $(0.9, 0.1)$, respectively.

Participants and experiment description. Sixty-one university students with normal color vision (28 female and 33 male), aged 18 to 33, participated in this user study. The experiment was run on a desktop computer with a 23.6-inch monitor at a 1920×1080 resolution. Before taking the test, and similar to the procedure followed in the RetargetMe paper, all participants were instructed about what image retargeting is, and what their role in the test would be. For this, apart from oral instructions, they were shown an image (not included in the subsequent test) as well as a series of retargeted results. The participants only advanced to the real test once they successfully completed some easy examples first (not taken into account in the final analysis).

A pair of images were simultaneously displayed side-by-side on the screen, with a black background. The left image was always the original image (from the 17 used in the test), while the right image was one of its eight retargeted results. Therefore there are a total $17 \times 8 = 136$ image pairs, which were displayed in random order. We followed a variant of the ITU-R five-point quality scale [55] and, for each pair, asked participants to rate the retargeted result by choosing a score in five intervals: 1-5 (bad), 6-10 (poor), 11-15 (fair), 16-20 (good) and 21-25 (excellent). All participants completed the rating of the 136 pairs.

Results. The raw scores provided by participants were first normalized and converted into z-scores (a.k.a. standard scores or normal scores). Then all scores were re-scaled to fit in the range $[0, 100]$. After outlier removal by the interval method [55], mean opinion scores (MOSs) were computed for each retargeted image. The higher MOS indicates the better perceived quality. For each retargeted image, an objective score was also computed (Eqs. 13 and 14).

To measure the performance of our proposed metric, a nonlinear mapping between objective (X) and

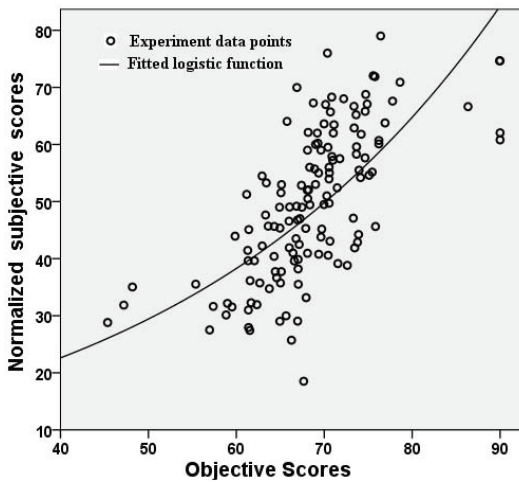


Fig. 12: The fitting curve using a logistic function in SPSS.

TABLE 4: Results using with four different metrics (A, B, C and D). Our proposed quality measure yields the best performance and is consistent with respect to subjective assessment. Please refer to the text for details.

Our measure	Correlation coefficients			(Outlier Metric D)
	Metric A	Metric B	Metric C	
Q_1	0.443	0.941	0.444	0.074
Q_2	0.577	0.950	0.586	0.066
Q_3	0.592	0.951	0.612	0.096
Q_4	0.424	0.942	0.406	0.096
(Q, Q')	0.656	0.963	0.721	0.051

subjective (Y) scores was applied with a logistic function:

$$Y = \frac{1}{a + b \times e^{-cX}}$$

where $a = 0.012$, $b = 1.196$ and $c = 0.073$ which were optimized in SPSS software. The resulting curve (shown in Figure 12) fits the data well ($R^2 = 0.656$).

Four metrics were further used for evaluation: Metric A is the correlation coefficient between objective/subjective scores after nonlinear regression analysis. Metric B is the correlation coefficient between objective/subjective scores after variance-weighted regression analysis. Metric C is the Spearman rank-order correlation coefficient between the objective/subjective scores. Metric D is the outlier ratio of the predictions after the nonlinear mapping. For metrics A, B and C, higher is better; for D, lower is better. The results are summarized in Table 4:

- Our metric (Q, Q') yields the best performance compared to the other measures (Eqs. 13 and 14).
- It also shows good consistency with respect subjective assessments.

5.1 Quality-Driven Image Retargeting

Our objective metric can be used to guide the design of new image retargeting methods. Previous works (e.g., [17], [26]) have shown that combining multiple operators often leads to better results than using a single operator. For instance, in [17], the bidirectional

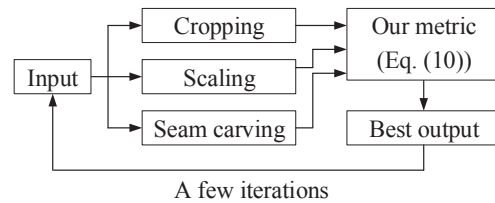


Fig. 13: A simple multi-operator scheme using our metric (Eq. (2)).

warping (BDW) measure is used to select an optimal sequence of several candidate operators. In Section 4.2 we have shown how our metric outperforms BDW, and thus can be used to design new multi-operator methods. Here, we present a simple proof-of-concept combination scheme that uses our metric.

As shown in Figure 13, we choose three simple retargeting operators: cropping, scaling and seam carving. Given an original image as input, the retargeted image is generated in similar fashion as previous multi-operator schemes: In each iteration, the three operators are applied independently and the best result evaluated by our metric is chosen as the input for next iteration. Since the metric outperforms previous ones, good results as usually achieved with very few iterations. Figure 14 shows an example using only three iterations and compared against eight other retargeting methods. Note that our simple scheme is not optimized in terms of choice of individual operators, number of iterations or error threshold, and determination of optimal sequences of operators. A full treatment of these optimizations and comparison with [17], [26] is beyond the scope of this paper, but our example serves as a proof of concept.

6 CONCLUSIONS AND FUTURE WORK

We have presented a simple yet effective *objective* quality metric for image retargeting. Five key factors are considered: preserving salient regions, reducing artifacts, controlling global structure, satisfying image aesthetics, and maintaining symmetry. Our experiments show that our method consistently and significantly outperforms other objective methods. We have also shown how it correlates better with users' subjective preferences by means of a leave-one-out cross validation test, and an additional user study. This indicates that our metric can indeed be used as a good predictor of subjective quality assessment, without the need to perform cumbersome user studies.

There are two limitations of our method, which we hope to address in future work. One is the speed of the evaluation; to evaluate a retargeting operation from 1024×754 to 768×754 takes around 50 seconds. While this is still orders of magnitude faster (and simpler) than running user studies, it rules out online applications. The second one has to do with the symmetry detection. Since state-of-the-art symmetry detection methods [52], [53] have reported success



Fig. 14: Example result of our multi-operator iterative scheme using our metric and three iterations. The numbers show the result of our metric. When compared to eight retargeting methods including multi-operator (MOP) [17], our new multi-operator scheme has the best result.

rates of less than 70%, we chose to rely on user's input to indicate whether an image has symmetric features. Nevertheless, if a breakthrough symmetry detection method is proposed in the future, we can easily incorporate it in our pipeline.

Although our metric offers excellent results, for some images there may exist some overlap between some components of our metric. Future work could analyze the possible influence of this overlap in the optimal parameter settings, further improving the metric. In addition, we would like to test our metric on *video* retargeting. This could be done on a per-frame basis by using our metric and adding a term for temporal consistency, although more sophisticated methods could be devised that leverage all the information at once instead. We believe this kind of automatic predictors of image quality will have an important rule in a near future, where retargeting operations along many dimensions (size, color, disparity...) will need to be applied to visual content, to adapt it to the characteristics and limitations of the many kinds of existing computational displays [37]. For instance, an additional disparity-preservation term could be devised to extend our method to evaluate stereo retargeting (e.g., [56], [36]). Given our modular approach, we hope that our proposed metric can be used as the starting point for these and possibly other cases, such as light field retargeting [10].

ACKNOWLEDGEMENTS

This work was partially supported by the National Science Fund of China (61202293, 61521002, 61322206, U1301253), the Science and Technology Planning Project of Guangdong Province (2015A020209124, 2014A050503057), the TNList Cross-discipline Foundation, and the Spanish Ministry of Science and Technology (project LIGHTSLICE). Diego Gutierrez was additionally funded by a Google Faculty Research Award.

REFERENCES

[1] A. Shamir and O. Sorkine, "Visual Media Retargeting," Course Notes on ACM SIGGRAPH ASIA 2009, Dec. 16-19, Japan, 2009.
 [2] F. Banterle, A. Artusi, E. Eisemann, D. Gutierrez, K. Myszkowski, T. Aydin, R. Mantiuk and P. Diddy, "Multidimensional Image Retargeting," Course Notes on SIGGRAPH Asia 2011, Dec.13-15, Hong Kong, 2011.

[3] S. Castillo, T. Judd and D. Gutierrez, "Using eye-tracking to assess different image retargeting methods," Proc. ACM SIGGRAPH Symp. Applied Perception in Graphics and Visualization (APGV 11), pp.7-14, 2011.
 [4] M. Rubinstein, D. Gutierrez, O. Sorkine and A. Shamir, "A comparative study of image retargeting," ACM Transactions on Graphics (ACM SIGGRAPH Asia 2010), Vol. 29, No. 6, Article No.160, 2010.
 [5] M. Decombas, F. Dufaux, E. Renan and B. Pesquet-Popescu, "A new object based quality metric based on SIFT and SSIM," IEEE Intl. Conf. Image Processing (ICIP), pp.1493-1496, 2012.
 [6] Y.J. Liu, X. Luo, Y.M. Xuan, W.F. Chen and X. Fu, "Image retargeting quality assessment," Computer Graphics Forum (Eurographics 2011), Vol.30, No.2, pp. 583-592, 2011.
 [7] L. Ma, W. Lin, C. Deng and K.N. Ngan, "Study of subjective and objective quality assessment of retargeted images," IEEE Intl. Symp. on Circuits and Systems (ISCAS 12), pp.2677-2680, 2012.
 [8] G. Lavoué and R. Mantiuk, "Quality assessment in computer graphics", In Visual Signal Quality Assessment, C. Deng, L. Ma, W. Lin and K.N. Ngan ed., pp. 243-286, 2015.
 [9] L. Liu, R. Chen, L. Wolf and D. Cohen-Or, "Optimizing photo composition," Computer Graphics Forum (Eurographics 2010), Vol.29, No.2, pp. 469-478, 2010.
 [10] C. Birkbauer and O. Bimber, "Light field retargeting", Computer Graphics Forum, Vol.31, pp. 295-303, 2012.
 [11] F.L. Zhang, M. Wang and S.M. Hu, "Aesthetic image enhancement by dependence aware object re-composition," IEEE Trans. on Multimedia, Vol. 15, No. 7, pp. 1480-1490, 2013.
 [12] R. Datta, D. Joshi, J. Li and J.Z. Wang, "Studying aesthetics in photographic images using a computational approach," Lecture Notes in Computer Science (ECCV 06), pp. 288-301,2006.
 [13] Y. Luo and X. Tang, "Photo and video quality evaluation: focusing on the subject," Lecture Notes in Computer Science (ECCV), pp. 386-399, 2008.
 [14] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Processing, Vol. 13, No. 4, pp.600-612, 2004.
 [15] R. Mantiuk, J. K. Kim, A. G. Rempel, W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," ACM Transactions on Graphics (ACM SIGGRAPH 2011), Vol. 30, No. 4, Article No. 40, 2011.
 [16] D. Simakov, Y. Caspi, E. Shechtman and M. Irani, "Summarizing visual data using bidirectional similarity," Computer Vision and Pattern Recognition (CVPR'08), pp.1-8, 2008.
 [17] M. Rubinstein, A. Shamir and S. Avidan, "Multi-operator media retargeting," ACM Transactions on Graphics (ACM SIGGRAPH 09), Vol. 28, No. 3, Article No. 23, 2009.
 [18] B.S. Manjunath, J.R. Ohm, V.V. Vasudevan and A. Yamada, "Color and texture descriptors," IEEE Trans. on Circuits and Systems for Video Technology, Vol.11, No.6, pp. 703-715, 2001.
 [19] E. Kasutani and A. Yamada, "The MPEG-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval," IEEE Intl. Conf. Image Processing (ICIP 01), pp.674-677, 2001.
 [20] C. Liu, J. Yuen, A. Torralba, J. Sivic and W.T. Freeman, "SIFT flow: dense correspondence across different scenes," European Conference on Computer Vision (ECCV 08), pp.28-42, 2008.
 [21] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing," ACM Transactions on Graphics (ACM SIGGRAPH 07), Vol. 26, No. 3, Article No. 10, 2007.

- [22] L. Wolf, M. Guttman and D. Cohen-Or, "Non-homogeneous content-driven video retargeting," Intl. Conf on Computer Vision (ICCV 07), pp.1-6, 2007.
- [23] Y.S. Wang, C.L. Tai, O. Sorkine and T.Y. Lee, "Optimized scale-and-stretch for image resizing," ACM Transactions on Graphics (ACM SIGGRAPH Asia 08), Vol 27, No. 5, Article No. 118, 2008.
- [24] Y. Liang, Y.J. Liu, X. Luo, L. Xie and X. Fu, "Optimal scaling factor assignment for patchwise image retargeting," IEEE Computer Graphics and Applications, 33(5), pp. 68-78, 2013.
- [25] Y. Pritch, E. Kav-Venaki and S. Peleg, "Shift-map image editing," Intl. Conf on Computer Vision (ICCV), pp. 151-158, 2009.
- [26] W. Dong, N. Zhou, J.C. Paul and X. Zhang, "Optimized image resizing using seam carving and scaling," ACM Transactions on Graphics (ACM SIGGRAPH Asia 09), Vol. 28, No. 5, Article No. 125, 2009.
- [27] C.H. Chang and Y.Y. Chuang, "A line-structure-preserving approach to image resizing," Computer Vision and Pattern Recognition (CVPR 12), pp.1075-1082, 2012.
- [28] S.S. Lin, I.C. Yeh, C.H. Lin and T.Y. Lee, "Patch-based Image Warping for Content-Aware Retargeting," IEEE Trans. Multimedia, Vol. 15, No. 2, pp.359-368, 2013.
- [29] H. Wu, Y.S. Wang, K.C. Feng, T.T. Wong, T.Y. Lee and P.A. Heng, "Resizing by symmetry-summarization," ACM Transactions on Graphics (ACM SIGGRAPH Asia 10), Vol. 29, No. 6, Article No. 159, 2010.
- [30] A. Shamir, A. Sorkine-Hornung and O.Sorkine-Hornung, "Modern Approaches to Media Retargeting," Course Notes on SIGGRAPH ASIA 2012, Nov.28-Dec.1, Singapore, 2012.
- [31] D. Vaquero, M. Turk, K. Pulli, M. Tico and N. Gelfand, "A survey of image retargeting techniques", In Proc. SPIE 7798, Applications of Digital Image Processing XXXIII, 2010.
- [32] O. Pele and M. Werman, "Fast and robust earth mover's distances," Intl. Conf on Computer Vision (ICCV 09), pp. 460-467, 2009.
- [33] Y. Fang, K. Zeng, Z. Wang, W. Lin, Z. Fang and C.W. Lin, "Objective Quality Assessment for Image Retargeting Based on Structural Similarity", IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Vol. 4, No.1, pp. 95-105, 2014.
- [34] C.C. Hsu, C.W. Lin, Y. Fang and W. Lin, "Objective Quality Assessment for Image Retargeting Based on Perceptual Geometric Distortion and Information Loss", IEEE Journal on Selected Topics in Signal Processing, Vol. 8, No.3, pp. 377-389, 2014.
- [35] B. Masia, L. Presa, A. Corrales and D. Gutierrez, "Perceptually-Optimized Coded Apertures for Defocus Deblurring", Computer Graphics Forum, Vol. 31, No.6, pp. 1867-1879, 2012.
- [36] B. Masia, G. Wetzstein, C. Aliaga, R. Raskar and D. Gutierrez, "Display Adaptive 3D Content Remapping", Computers & Graphics, Vol. 37, No.8, pp. 983-996, 2013.
- [37] B. Masia, G. Wetzstein, P. Didyk and D. Gutierrez, "A survey on computational displays: Pushing the boundaries of optics, computation, and perception", Computers & Graphics, Vol. 37, No.8, pp. 1012-1038, 2013.
- [38] S.J. Russel and P. Norvig. Artificial Intelligence: A Modern Approach. 2nd ed., Prentice Hall, 2003.
- [39] G. Debreu, "Topological methods in cardinal utility theory," In Mathematical Methods in the Social Sciences, pp.16-26, Standord University Press, 1960.
- [40] A. Borji and L. Itti, "State-of-the-art in visual attention modeling," IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol.35, No.1, pp. 185-207, 2013.
- [41] M. M. Cheng, G. X. Zhang, N. J. Mitra, X. L. Huang, S. M. Hu, "Global Contrast based Salient Region Detection," Computer Vision and Pattern Recognition (CVPR 11), 409-416, 2011.
- [42] C.L. Novak and S.A. Shafer, "Anatomy of a color histogram," Computer Vision and Pattern Recognition (CVPR 92), pp.599-605, 1992.
- [43] J.R. Smith and S.F. Chang, "VisualSEEK: a fully automated content-based image query system," ACM Multimedia, pp. 87-98, 1996.
- [44] T. Lindeberg, "Scale-space theory: a basic tool for analyzing structures at different scales," Journal of Applied Statistics, Vol.21, No.1-2, pp. 225-270, 1994.
- [45] P. Jonas. Photographic Composition Simplified. Amphoto Publishers, 1976.
- [46] C. Harris, M. Stephens. "A combined corner and edge detector", In proc. the 4th Alvey Vision Conference, pp. 147-151, 1998.
- [47] D.G. Lowe. "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision, 60(2), 91-110, 2004.
- [48] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally stable extremal regions", In British Machine Vision Conference (BMC 02), pp. 384-393, 2002.
- [49] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, pp. 603-619, 2002.
- [50] P. Krahenbuhl, M. Lang, A. Hornung and M. Gross, "A system for retargeting of streaming video," ACM Transactions on Graphics (ACM SIGGRAPH Asia 09), Vol. 28, No. 5, Article No.126, 2009.
- [51] M.G. Kendall, "A new measure of rank correlation," Biometrika, Vol.30, No.1-2, pp. 81-93, 1938.
- [52] M. Park, K. Brocklehurst, R.T. Collins and Y. Liu, "Deformed Lattice Detection in Real-World Images Using Mean-Shift Belief Propagation," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 31, No. 10, pp. 1804-1816, 2009.
- [53] S. Liu, T.T. Ng, K. Sunkavalli, M.N. Do, E. Shechtman and N. Carr, "PatchMatch-based Automatic Lattice Detection for Near-Regular Textures," Intl. Conf on Computer Vision (ICCV 15), 2015.
- [54] J. Liu, G. Slota, G. Zheng, Z. Wu, M. Park, S. Lee, I. Rauschert, and Y. Liu, "Symmetry detection from real world images competition 2013: Summary and results," IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW), pp.200-205, 2013.
- [55] ITU: ITU-R Recommendation BT.500-11. Methodology for the subjective assessment of the quality of television images. International Telecommunication Union: Geneva, 2002.
- [56] S.S. Lin, C.H. Lin, S.H. Chang, and T.Y. Lee, "Object-Coherence Warping for Stereoscopic Image Retargeting," IEEE Trans. on Circuits and Systems for Video Technology, Vol. 24, No. 5, pp. 759-768, 2014.



Yun Liang is an associate professor of College of Information in South China Agricultural University. She received the M.S. and Ph.D. degree in Information Science and Technology from Sun Yat-sen University, China, in 2005 and 2011 respectively. Her research interests include image processing, computer vision and machine learning.



Yong-Jin Liu received his Ph.D degree from the Hong Kong University of Science and Technology, Hong Kong, China, in 2004. He is now an Associate Professor with the TNList, Department of Computer Science and Technology, Tsinghua University, China. He is the corresponding author of this paper. His research interests include computational geometry, computer graphics and computer-aided design and pattern analysis.



Diego Gutierrez is a Professor at the Universidad de Zaragoza, in Spain, where he's the founder and director of the Graphics and Imaging Lab. His research focuses on computational imaging, computational light transport, and applied perception. He is currently an Editor-in-Chief of ACM Transactions on Applied Perception, and an Associate Editor of ACM Transactions on Graphics, Computers & Graphics, and Presence.