Transformation-Aware Similarity Measurement for Image Retargeting Quality Assessment via Bidirectional Rewarping

Feng Shao, *Member, IEEE*, Zhenqi Fu, Qiuping Jiang, *Student Member, IEEE*, Gangyi Jiang, *Member, IEEE*, and Yo-Sung Ho, *Fellow, IEEE*

***Abstract*—Image retargeting is an effective way to adapt images for target displays with different aspect ratios and sizes. Meanwhile, effective image retargeting quality assessment (IRQA) is important for optimizing the image retargeting operations. In this paper, we propose a transform-aware similarity (TRASIM) measurement metric for IRQA, including bidirectional geometric distortion measurement, bidirectional information loss measurement, and global salient structure distortion measurement. The main innovation of the TRASIM is to build a universal framework to establish the similarity transformation via bidirectional rewarping to simulate different types of retargeting operators. Based on the similarity transformation, geometric distortion and content loss are measured to determine the retargeting quality. Experimental results on two widely used databases (CUHK and RetargetMe) indicate that the proposed TRASIM has higher consistency with subjective ranks, compared with the state-of-the-art IRQA metrics.**

***Index Terms*—Bidirectional rewarping, content loss, geometric distortion, image retargeting quality assessment (IRQA).**

I. Introduction

With the rapid popularity of various display devices, requirements for displaying image/video on different terminal devices with different resolutions are much urgent [1]–[3]. However, traditional cropping (CR) and scaling (SCL) operations cannot obtain satisfactory experience due to ignoring image content semantics. Content-aware image retargeting is an effective way to adapt display devices with different aspect ratios and sizes, while preserving the important image content with less distortion.

The issue of image/video retargeting has received much attention in recent years, and many retargeting operators have been proposed. The existing image retargeting operators can be roughly divided into two categories. The first category is discrete method based on directly removing or inserting pixels. CR, seam carving (SC) [4], and shift maps (SMs) [5] are well-known discrete approaches. The second category is continuous method that resize an image via continuous deformation operation. SCL, nonhomogeneous warping (WARP) [6], streaming video (SV) [7], and scale-and-stretch (SNS) [8] are representative continuous approaches. Unfortunately, there is no single method that can work well in preserving image content and structure, preventing artifacts for every image. Thus, it is meaningful to design a metric that can objectively evaluate the quality of retargeted images, so that the influence of different retargeting operators can be quantified [9].

Traditional image quality assessment (IQA) approaches, e.g., feature similarity index (FSIM) [10], structural similarity (SSIM) [11], and gradient magnitude similarity deviation (GMSD) [12], focus on measuring pixel-to-pixel similarity to capture image distortion. However, these methods cannot be directly employed to image retargeting quality assessment (IRQA) because the retargeted image has different resolutions with the original image. In fact, evaluating the image quality under different aspect ratios is extremely challenging because it requires to assess the content loss and geometric distortion after resolution adjustment. The early IRQA metrics, such as bidirectional similarity (BDS) [13], bidirectional warping (BDW) [14], SIFT-flow [15], and earth-Mover's distance (EMD) [16], measure the quality of retargeted images by establishing the correspondence between the original and retargeted images. Recent IRQA methods aim to solve the challenges from two pathways: 1) establishing the attribute correspondence and calculating the distance for measurement and 2) extracting features from the retargeted images and using these features to train a network for quality prediction.

Inspired by the work in [17] that registers the original image with the Markov random field (MRF) in the retargeted image, and also inspired by the work in [18] that establishes block warping between the retargeted image and the scaled original image, the approach in this paper to measure retargeting distortion lies in simulating different retargeting operators to regenerate the retargeted/original image from the original/retargeted image via bidirectional rewarping.

Although methods [17] and [18] have achieved good results, they still have the following limitations: 1) method [17] relies

Manuscript received March 5, 2019; accepted May 10, 2019. This work was supported in part by the Natural Science Foundation of China under Grant 61622109, in part by the Zhejiang Natural Science Foundation of China under Grant R18F010008, in part by the Natural Science Foundation of Ningbo under Grant 2017A610112, and in part by the K. C. Wong Magna Fund in Ningbo University. This paper was recommended by Associate Editor G. Nicosia. *(Corresponding author: Feng Shao.)*

F. Shao, Z. Fu, Q. Jiang, and G. Jiang are with the School of Information Science and Engineering, Ningbo University, Ningbo 315211, China (e-mail: shaofeng@nbu.edu.cn).

Y.-S. Ho is with the School of Information and Communications, Gwangju Institute of Science and Technology, Gwangju 500-712, South Korea (e-mail: hoyo@gist.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMC.2019.2917496

on correspondence relationship to measure aspect ratio and absolute size changes, but these features are not sufficient to characterize retargeting semantics and 2) method [18] first scales the original image to the same size with the retargeted image, and then analyze the distribution of deformities and losses, but scaling itself will introduce such information losses and distortions. In this paper, we propose a transform-aware similarity (TRASIM) measurement for IRQA by estimating similarity transformation between the original and retargeted images via a bidirectional way. The major innovation of our TRASIM model is to build a universal framework to establish the transformation relationship between the original and retargeted images to simulate different types of retargeting operators and derives the retargeting semantics from the similarity transformation matrices to evaluate the retargeting distortions.

Summary of the contributions of this paper are as follows.

1) We simulate different types of retargeting operators using a common rewarping operation and estimate the similarity transformation to build the relationship between the original and retargeted images via bidirectional rewarping.
2) To capture the geometric change, we propose a bidirectional geometric distortion (BDGD) metric based on the bias of similarity transformation matrices in a bidirectional way.
3) We propose a bidirectional information loss (BDIL) measurement based on the displacement of similarity transformation to quantify how much salient information is preserved in the retargeted image, and inversely quantify how much salient content can be recovered from the retargeted image.
4) A global salient structure distortion (GSSD) metric is proposed to measure the global structure consistency between the original and retargeted images based on similarity transformation matrices.

The remainder of this paper is organized as follows. In Section II, we introduce some related works and the motivation of rewarping for IRQA. We present the details of our method in Section III, and finally present the experimental results in Section IV and conclusion in Section V.

II. Background

In this section, we first make a brief overview of some related works on IRQA. Then, we discuss the motivation of using rewarping for IRQA in this paper.

A. Related Works

Strictly speaking, IRQA is an inverse process to interpret the retargeting modification imposed on the original images. Since the goal of image retargeting is to preserve important contents as much as possible while having less distortion, IRQA is to measure the content loss and structure degradation in retargeted images. However, since the retargeting operator by which the retargeted image is generated is usually unknown, directly measuring geometric distortion and content loss of the retargeted images are challenging.

In recent years, great promotions have been made for IRQA. As compared and discussed in [9], edge histogram (EH) [19] and color layout (CL) [20] use global image distance to estimate the dissimilarity between two images based on edge or color distribution. BDS [13] calculates a bidirectional mapping between the patches in the source and retargeted images by seeking the minimum sum of squared distances. BDW [14] is similar with BDS but takes asymmetric dynamic time warping distance as a measure. However, both BDW and BDS have poor correlation with the subjective rank scores due to ignoring the important contents. SIFT-flow [15] and EMD [16] use more accurate matching metrics which could well capture the image structural properties. Among these metrics, SIFT-flow and EMD demonstrate high consistency with subjective ranks.

Besides the above models, many special IRQA metrics have been proposed based on objective measurements on geometric distortion and content loss. Ma *et al.* [21] investigated various structural descriptors for evaluating perceptual quality of retargeted image. Liu *et al.* [22] extracted global geometric structures of two images and built the local pixel-to-pixel correspondence for assessment. Fang *et al.* [23] proposed a metric (IR-SSIM) by generating an SSIM map to obtain the structural information in retargeted images. Hsu *et al.* [24] measured local variance in SIFT-flow fields to depict geometric distortion and measured saliency loss as information loss. Zhang *et al.* [17] measured aspect ratio change of local blocks to interpret the geometric change. Karimi *et al.* [18] first established pixel-to-pixel correspondence between the scaled and retargeted images and defined shape feature, area feature, and aspect ratio feature to measure the deformities and losses. Liang *et al.* [25] combined elements from salient content, artifact, global structure, aesthetic, and symmetry to measure the retargeted image quality. Oliveira *et al.* [26] evaluated the quality of retargeted images by a bidirectional fusion framework. In our previous work [27], we used sparse representation framework to investigate how much structure and saliency features are preserved or changed between the original and retargeted images as a measure. Other relevant works that map features (or relative scores) into rank scores can be found in [28]–[34].

B. Motivation of Rewarping for IRQA

As a common operator in image retargeting, the essence of warping is to obtain an optimal transformation based on limited control points. The physically correct way to warp a scene with m control points is to follow the distortion energy [35]:

$$\varepsilon\left(P', P\right) = \min_{\rho \in P} \sum_{i=1}^{m} \left|\rho(\mathbf{p}_i) - \mathbf{p}_i'\right|^2 \tag{1}$$

where $\mathbf{p}_i'$ denotes the deformed position of $\mathbf{p}_i$, P is a set of similarity transformations ρ having the general form

$$\rho(\mathbf{p}) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \mathbf{p} = \begin{bmatrix} x \\ y \end{bmatrix} \tag{2}$$

where a, b, c, d, t_x, and t_y are the parameters to determine a unique similarity transformation.

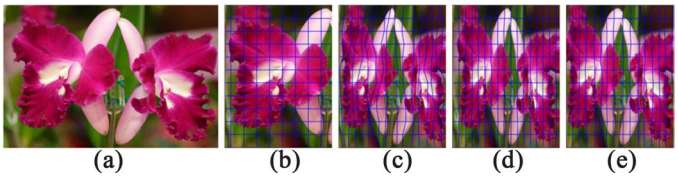


Fig. 1. Example of similarity transformations for typical retargeting operators. (a) Original. (b) CR. (c) SC. (d) SCL. (e) WARP.

To further analyze if such transformation can quantify the geometric distortion and information loss induced by different retargeting operators, we simulate the similarity transformation for different retargeting operators. Four typical operators, CR, SC, SCL, and WARP are considered here.

1) For CR operator that selects important image content by a cropping window, if partial or whole content of a selected grid in the original image is discarded, the similarity transformation can be easily solved by mapping all control points to the same boundary [as shown in Fig. 1(b)].
2) For SC operator that removes pixels by seams, similar to the CR operator, the importance of the grid determines the number of seams in the grid, which can be easily solved by finding the matching control points between the two images [as shown in Fig. 1(c)].
3) For SCL or WARP operator that in fact interpolates the retargeted pixels in a homogeneous or inhomogeneous way, it can be approximately solved by a linear transformation based on the observation that the continuous approaches are similar with the discrete approaches in the subpixel level [as shown in Fig. 1(d) and (e)].

Therefore, as analyzed above, since almost all the retargeting operators can be accurately or approximately solved by the similarity transformation, it is motivated to further consider if we can establish a unified framework to clarify the transform relationship between the original and retargeted images. In this paper, we formulate IRQA as a rewarping issue, and derive the retargeting semantics directly from the transformation. The motivation using rewarping to simulate different retargeting operators mainly comes from three aspects.

1) Compared with CR, SC, SCL, SNS, MULTIOP, or other retargeting operators, warping operator will be comparatively straightforward that only needs a transformation matrix to describe the deformation of all pixels within a grid, while different constraints imposed on the grid can change the manifestation of the similarity transformation.
2) No matter whether discrete or continuous retargeting approaches, the essence of these approaches is to preserve important content of the original image. Thus, it can be simulated via rewarping that distributes more deformations on less important regions while less distortion on more important regions.
3) For geometric distortion, the bias of the similarity transformation can reflect the structure changes imposed on the grids. For information loss, large grids usually preserve more important content, while small grids remove more information. Our BDGD measurement computes the scale and rotation changes in similarity transformation, while BDIL metric calculates the important information loss during retargeting. Thus, by deriving the retargeting semantics from the transformation, the structure and completeness of important content are well considered in our bidirectional measurement.

III. TRASIM Metric

In this paper, we propose a TRASIM measurement metric for IRQA, as shown in Fig. 2. The primary motivation of TRASIM is to interpret different image retargeting operators by a unified framework. Different from the previous methods that only establish pseudo-mapping relationship between the original and retargeted images, we simulate the general warping operator (represented as similarity transformation) to produce different retargeted images via bidirectional rewarping. Our method can be divided into the following steps.

1) We apply SIFT-flow to build forward matching from original image to retargeted image, and backward matching from retargeted image to original image. With the matching rule, similarity transformation matrices can be estimated from the SIFT-flow fields for the grids on original image or retargeted image and are used to regenerate the retargeted and original images via bidirectional rewarping.
2) After obtaining the similarity transformation matrices, we calculate BDGD to represent the geometric distortion by the distance between the similarity transformation matrices and their benchmark, and calculate BDIL to represent the important content loss during retargeting.
3) In order to comprehensively evaluate retargeting distortions, the object-level measurement is crucial. Thus, we calculate GSSD to measure the global structure and information preservation via the consistent of similarity transformation matrices and the preservation of saliency object.
4) We consider to fuse BDGD, BDIL, and GSSD to obtain the final perceptual quality via a prediction function trained by support vector regression (SVR), with higher accuracy.

Next, steps of the proposed method are demonstrated in details. In order to make it easier for understanding, some important variables and notations are presented in Table I.

A. Problem Formulation

Different from image retargeting in which the retargeted image is the output, in this situation, both the source and retargeted images are known in advance but the retargeting rule about how to generate the retargeted image is unknown (e.g., warping, SC, CR, or SCL). To simulate the retargeting rule, we aim to solve an inverse warping issue on how to regenerate the original or retargeted image from its opposite. To be distinguished with the warping rule in image retargeting, we take inverse warping as a rewarping issue. Given the original and retargeted images $\mathbf{I} = \{I_{\text{org}}, I_{\text{ret}}\}$, a set of grids are selected from the original image, denoted as $\mathbf{V}_{\text{org}} = \{\mathbf{v}_{\text{org}}^k, k = 1, \ldots, m\}$, in which each grid is determined

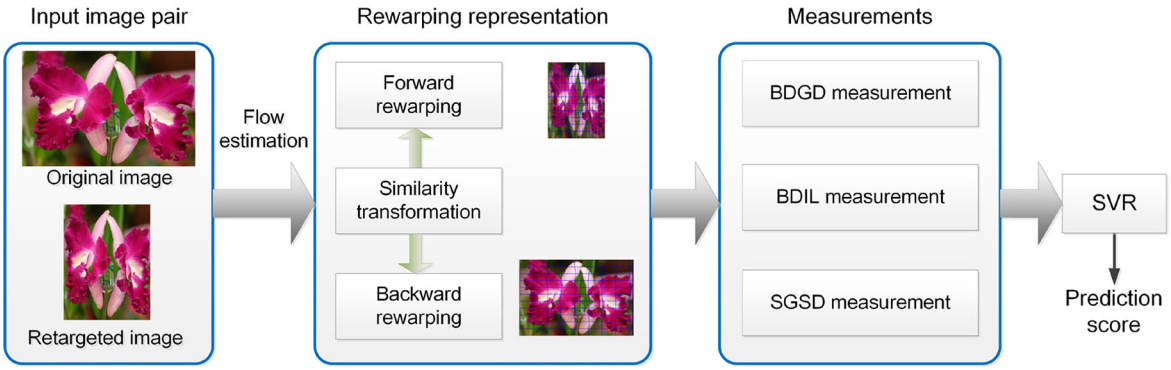


Fig. 2. Framework of the proposed TRASIM measurement method.

TABLE I
IMPORTANT VARIABLES AND NOTATIONS

Symbol	Definition
I_{org}	Original image
I_{ret}	Retargeted image
$\mathbf{v}_{org}^{k}$	Selected grid in the original image
$\vec{\mathbf{v}}_{org}^{k}$	The estimated grid of $\mathbf{v}_{org}^{k}$ in the original image
$\tilde{\mathbf{v}}_{org}^{k}$	The matched grid of $\mathbf{v}_{org}^{k}$ in the original image
$\vec{\mathbb{Z}}_k$	Forward rewarping rule
$\mathbf{P}_{org}^{k}$	Forward similarity transformation matrix
$\mathbf{v}_{ret}^{k}$	Selected grid in the retargeted image
$\vec{\mathbf{v}}_{ret}^{k}$	The estimated grid of $\mathbf{v}_{org}^{k}$ in the retargeted image
$\tilde{\mathbf{v}}_{ret}^{k}$	The matched grid of $\mathbf{v}_{org}^{k}$ in the retargeted image
$\overleftarrow{\mathbb{Z}}_k$	Backward rewarping rule
$\mathbf{P}_{ret}^{k}$	Backward similarity transformation matrix

with four vertices (m is the number of grids in the original image). Also, a set of grids can be selected from the retargeted image, denoted as $\mathbf{V}_{\text{ret}} = \{\mathbf{v}_{\text{ret}}^{k}, k = 1, \ldots, m'\}$($m'$ is the number of grids in the retargeted image). The purpose of the image warping is to find the warping rule $\vec{\mathbb{Z}}_k$ so that the retargeted grid $\vec{\mathbf{v}}_{\text{ret}}^{k}$ has the minimum shape distortion with the matched grid $\tilde{\mathbf{v}}_{\text{ret}}^{k}$, i.e.,

$$\vec{\mathbf{v}}_{\text{ret}}^{k} = \vec{\mathbb{Z}}_k\left(\mathbf{v}_{\text{org}}^{k}\right). \tag{3}$$

As an inverse issue, if the original grid $\mathbf{v}_{\text{org}}^{k}$ and the matched retargeted grid $\tilde{\mathbf{v}}_{\text{ret}}^{k}$ are known, the warping rule $\vec{\mathbb{Z}}_k$ can be estimated to interpret the warping process. In this paper, such process is defined as a forward rewarping to distinguish the warping operator in image retargeting. Since the forward rewarping is not reversible in transformation, by recovering the original image from the retargeted image, another backward rewarping is also performed to measure the degree of important content preservation, which is defined as

$$\overleftarrow{\mathbf{v}}_{\text{org}}^{k} = \overleftarrow{\mathbb{Z}}_k\left(\mathbf{v}_{\text{ret}}^{k}\right). \tag{4}$$

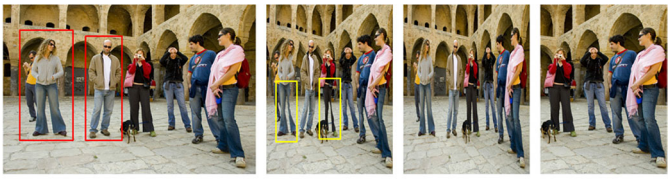


Fig. 3. Example of retargeted images with typical geometric distortion and content loss.

In the practical image warping, $\vec{\mathbb{Z}}_k$ and $\overleftarrow{\mathbb{Z}}_k$ are solved by similarity transformation. As will be demonstrated later, similarity transformation matrix is useful to capture the retargeting distortion which is significantly different from the pixel-wise similarity metrics.

As discussed in previous works, geometric distortion and information loss are two primary retargeting distortions. Fig. 3 gives an example of different retargeted images generated by SC, SCL, and CR operators. As we can observe, the retargeted image obtained by cropping suffers from foreground objects removal (as shown in the red rectangle), while seam carving will introduce serious discontinuities and content loss (as shown in the yellow rectangle). The retargeted image obtained by scaling will squeeze the objects uniformly. To capture such visual quality changes, the goal of IRQA motivated in this paper is to simulate and measure the degradation process of different retargeting operators based on the following considerations: 1) similarity transformation is a straightforward way to capture the local geometric distortion of a grid; 2) retargeting should preserve the important content as much as possible; and 3) global structure distortion in the salient regions should be minimized. Therefore, we define the overall quality of retargeted images by combining the above discussed factors

$$\begin{aligned} Q = & \underbrace{M_1\left\{\vec{\mathbb{Z}}_k, \overleftarrow{\mathbb{Z}}_k\right\}}_{\text{BDGD}} + \underbrace{M_2\left\{\vec{\mathbf{v}}_{\text{org}}^{k}, \overleftarrow{\mathbf{v}}_{\text{ret}}^{k}\right\}}_{\text{BDIL}} \\ & + \underbrace{M_3\left\{\vec{\mathbb{Z}}_k, \overleftarrow{\mathbb{Z}}_k, \vec{\mathbf{v}}_{\text{org}}^{k}, \overleftarrow{\mathbf{v}}_{\text{ret}}^{k}\right\}}_{\text{GSSD}} \end{aligned} \tag{5}$$

where $M_1\{\cdot\}$, $M_2\{\cdot\}$, and $M_3\{\cdot\}$ denote different quality pooling operators for BDGD, BDIL, and GSSD measurements. In the

next, we will demonstrate how to derive the BDGD, BDIL, and GSSD measurements from the similarity transformation.

B. SIFT-Flow Estimation

First of all, to build the matching relationship between the source image (grids) and the retargeted image (grids), the pixel-to-pixel correspondence should be first established, and then is converted to a series of transformations. To obtain such correspondence, we extract dense SIFT descriptor for each pixel in the original and the retargeted image by SIFT-flow algorithm [15], which is calculated by minimizing the following energy function:

$$\begin{aligned} E(\mathbf{x}) = &\sum_{\mathbf{p}} \min\big(\|s_1(\mathbf{p}) - s_2(\mathbf{p}+\mathbf{x}(\mathbf{p}))\|_1, t\big) \\ &+ \sum_{\mathbf{p}} \eta(|u(\mathbf{p})| + |v(\mathbf{p})|) \\ &+ \sum_{(\mathbf{p},\mathbf{q})\in\varepsilon} \min(\alpha|u(\mathbf{p}) - u(\mathbf{q})|, d) \\ &+ \min(\alpha|v(\mathbf{p}) - v(\mathbf{q})|, d) \end{aligned} \tag{6}$$

where $\mathbf{x}(\mathbf{p})$ denotes the SIFT-flow vector of a pixel $\mathbf{p}$, parameters t and d denote the thresholds to limit the amount of maximum error, η and α are the weights, and $u(\mathbf{p})$ and $v(\mathbf{p})$ are the horizontal and vertical components of the flow vector $\mathbf{x}(\mathbf{p})$. If the resolution of the retargeted image is smaller than that of the original image, it is a possible condition that multiple pixels in the original image are mapped into a same location in the retargeted image, or no pixel is mapped into a location in the retargeted image. To minimize such effects as much as possible, we add two topological assumptions in the SIFT-flow estimation.

Ordering Assumption: If pixel $\mathbf{A}$ is located in the left of pixel $\mathbf{B}$, its matched pixel $\mathbf{A}'$ will not be located in the right of the matched pixel $\mathbf{B}'$ in another image.

Uniqueness Assumption: Each pixel in the retargeted image will has its unique correspondence point in the original image.

As a result, the algorithm outputs an SIFT-flow vector $\mathbf{x}(\mathbf{p}) = [u'(\mathbf{p}), v'(\mathbf{p})]^T$ for each pixel $\mathbf{p}$, where $u'(\mathbf{p})$ and $v'(\mathbf{p})$ are the horizontal and vertical displacement vectors. The properties of SIFT-flow can be found in [15] and [17]. It should be noted that our similarity transformation depends on accurate dense correspondence to construct the forward and backward rewarping relationship. In this paper, we do not focus on designing the dense correspondence method, but aim to provide a universal similarity transformation framework for IRQA. Therefore, if we can design the dense correspondence method more accurately, the accuracy of similarity transformation will be further improved. This topic will be studied in our future work.

C. Similarity Transformation

As the purpose of rewarping, considered in this paper, is to regenerate the retargeted/original image from the original/retargeted image, we convert the issue into a field of similarity transformation, as done in warping-based retargeting operations [35]. The goal of forward rewarping is to warp a regular grid in the original image into its corresponding grid of the retargeted image, while SIFT-flow vector maps a single pixel positions into other position. In our procession, transformations are represented via a set of 2×3 affine matrices, which can express the change of scaling, rotation, and translation, defined as

$$\rho_k(\mathbf{p}) = \begin{bmatrix} a & b & t_x \\ c & d & t_y \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{P}_k \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \tag{7}$$

Similar with the warping rule in image retargeting methods [35], [36], four vertices are used to locate a grid. For a grid $\mathbf{v}^k_{\text{org}}$ in the original image and the matched grid $\tilde{\mathbf{v}}^k_{\text{org}}$ in the retargeted image, an optimal similarity transformation equals to find the minimum dissimilarity distance for four vertices in a grid between the estimated and matched grids

$$\min_{\rho\in P} \sum_{i=1}^{4} \left\| \rho_k\big(\mathbf{p}^i_k\big) - \tilde{\mathbf{p}}^i_k \right\|^2. \tag{8}$$

In the methods [35] and [36], the energy function is solved by least-squares linear regression as the target grid is unknown. Since the source and target grids are known in our method, to eliminate the influence of inaccurate matching, we adopt M-estimator [37] to estimate the transformation. Particularly, to improve the robustness of grid transformation, instead of using four vertices to estimate the transformation, all pixels in the matched grids are used to give a robust solution. For backward rewarping, the solution of similarity transformation is the same with forward rewarping.

With the estimated similarity transformations, we can efficiently rewarp the source grid to its target location and approximatively recover the source image from the retargeted image. Fig. 4 shows the forward rewarped and backward rewarped images for different retargeting operators. For CR operator without geometric distortion, the estimated forward and backward rewarped images are almost the same with the reference ones. For SCL operator, the estimated forward grids are relatively regular and the estimated backward grids can cover the original image. For SC and WARP operators, that introduce serious geometric distortion and content loss, there are obvious holes and grids deformation in rewarped images. In general, the difference between the forward rewarped and source retargeted images is very small, and content loss can be well revealed in the backward rewarped images, which indicates that rewarping is a feasible way to reveal the intrinsic relationship between the source and retargeted images.

D. Bidirectional Geometric Distortion

As known, the purpose of warping is to assign more distortions on less important regions and less distortions on more important regions. The estimated similarity transformation matrices can reflect the degree of distortion imposed on the grids. Considering six parameters in a similarity transformation matrix, un-uniform scale parameters [a and d in (7)] will have a negative influence on geometric distortion and content loss, and the rotation parameters [b and c in (7)] will

Fig. 4. Forward rewarped and backward rewarped images for different retargeting operators. (a) Original. (b) CR. (c) SCL. (d) SC. (e) WARP. (f) Forward rewarped image and the error map (b). (g) Backward rewarped image and the error map (b). (h) Forward rewarped image and the error map (c). (i) Backward rewarped image and the error map (c). (j) Forward rewarped image and the error map (d). (k) Backward rewarped image and the error map (d). (l) Forward rewarped image and the error map (e). (m) Backward rewarped image and the error map (e).

change the geometric structure, while translation parameters [t_x and t_y in (7)] itself do not produce any geometric distortion and content loss. Therefore, the benchmark transformation matrix defined in this paper is

$$\mathbf{P}_B = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \end{bmatrix}. \tag{9}$$

Since scale and rotation parameters can reflect the geometric distortion, we compute the distance between a similarity transformation matrix and its benchmark to capture the degree of geometric distortion imposed on a grid. Instead of computing the absolute difference between two matrices, the distance calculated from the similarity transformation is composed of two components: 1) scale and rotation change and 2) uniformity

$$\zeta(\mathbf{P}_k, \mathbf{P}_B) = \underbrace{(a-1)^2 + (d-1)^2 + b^2 + c^2}_{\zeta_{\text{SR}}:\text{Scale and Rotation}} + \underbrace{\left\{(a-d)^2 + (|b| + |c|)^2\right\}}_{\zeta_{\text{UN}}:\text{Uniformity}}. \tag{10}$$

Here, ζ_{SR} represents the scale and rotation change while ζ_{UN} represents the uniformity. We further illustrate the role of two terms in measuring the distance by four representative operations, as shown in Fig. 5. Testing results are shown in Table II (due to limited space of the table, we only show the scale and rotation parameters). We can make the following observations.

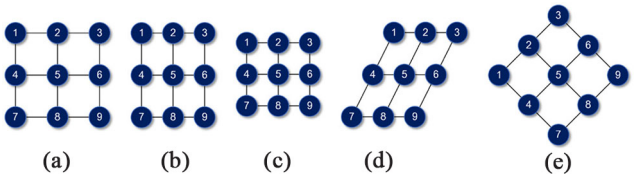


Fig. 5. (a) Original grid. (b) Horizontal scaling. (c) Horizontal and vertical scaling. (d) Horizontal scaling and rotation. (e) Anti-clockwise rotation.

TABLE II
GRID MODELS FOR ζ_{SR} AND ζ_{UN} MEASUREMENTS

image	(a)	(b)	(c)	(d)	(e)
P	$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$	$\begin{bmatrix}0.5 & 0\\ 0 & 1\end{bmatrix}$	$\begin{bmatrix}0.5 & 0\\ 0 & 0.5\end{bmatrix}$	$\begin{bmatrix}0.5 & 0.5\\ 0 & 1\end{bmatrix}$	$\begin{bmatrix}\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{bmatrix}$
ζ_{SR}	0	0.25	0.50	0.50	1.17
ζ_{UN}	0	0.25	0	0.50	2.00

1) For operation in (c), the value of ζ_{UN} is 0 which indicates the transformation is uniform.
2) The ζ_{SR} and ζ_{UN} in operation (d) is larger than (b). It is reasonable because the deformation in operation (d) is larger than (b).
3) For operation in (e), both the ζ_{SR} and ζ_{UN} are significantly larger than the others, indicating serious deformation.

For the purpose of salient information preservation, the destructions of geometric changes are much more significant in the salient areas than in the un-salient areas. Then, the forward geometric distortion (FGD) and backward geometric distortion (BGD) are computed with the significance values as weights

$$f_{\mathrm{FGD}} = \sum_{\mathbf{v}^k_{\mathrm{org}}} S_{\mathbf{v}^k_{\mathrm{org}}} \cdot e^{-\zeta\left(\mathbf{P}^k_{\mathrm{org}}, \mathbf{P}_B\right)} \Big/ \sum_{\mathbf{v}^k_{\mathrm{org}}} S_{\mathbf{v}^k_{\mathrm{org}}} \tag{11}$$

$$f_{\mathrm{BGD}} = \sum_{\mathbf{v}^k_{\mathrm{ret}}} S_{\mathbf{v}^k_{\mathrm{ret}}} \cdot e^{-\zeta\left(\mathbf{P}^k_{\mathrm{ret}}, \mathbf{P}_B\right)} \Big/ \sum_{\mathbf{v}^k_{\mathrm{ret}}} S_{\mathbf{v}^k_{\mathrm{ret}}} \tag{12}$$

where $S_{\mathbf{v}^k_{\mathrm{org}}}$ denotes the average saliency score of grid $\mathbf{v}^k_{\mathrm{org}}$ in the original image, and $S_{\mathbf{v}^k_{\mathrm{ret}}}$ denotes the average saliency score of grid $\mathbf{v}^k_{\mathrm{ret}}$ in the retargeted image. Larger f_{FGD} and f_{BGD} values mean better results. In this paper, we use hierarchical saliency (HS) method [38] to estimate the saliency values because it demonstrates a remarkable consistency with the subjective observation (the influence of different saliency models will be analyzed in Section IV-F). Here, we measure BDGD with two reasons: 1) if a salient grid in the original image is not salient in the retargeted image, it will have worse BGD result opposite to better FGD result, reflecting the incompleteness of salient content preserved in the retargeted image and 2) for some types of retargeting operators, e.g., SC and CR, information loss due to directly discarding unimportant contents may lead to incomplete forward warping, which is different from the information loss caused by other scaling and warping operators.

TABLE III
DISTANCE COMBINATION ANALYSIS ON CUHK DATABASE

distance combination		PLCC	SRCC	RMSE	OR
ζ_{SR}	ζ_{UN}				
√		0.6307	0.5798	10.8186	0.0293
	√	0.6587	0.6343	10.1585	0.0117
√	√	0.6948	0.6514	9.7102	0.0117

Based on the above analysis, we further test the performance of FGD measurements (f_{FGD}) using different distance combination on the CUHK dataset. As illustrated in Table III, the distances are both capable to promote the overall performance.

E. Bidirectional Information Loss

As discussed, geometric change can lead to information loss if partial information is preserved or discarded. A large distance between the estimated and benchmark transformation matrices usually means large information loss for the grid. However, the geometric distortion measurements are not always useful in evaluating the information loss. For example, if a salient region is wholly discarded, the BGD measurement cannot correctly reflect such type of information loss. Therefore, in order to characterize different roles of information loss and geometric distortion in measuring retargeted image quality, and accurately evaluate how much salient information is preserved or discarded in the retargeted image, we calculate the change of salient areas in both original and retargeted images via forward and backward rewarping.

Let $\vec{\mathbf{v}}^k_{\mathrm{ret}}$ be the rewarped grid in the retargeted image from the original grid $\mathbf{v}^k_{\mathrm{org}}$, the forward information loss (FIL) is measured by the ratio between the areas of the rewarped grids and the original grids

$$f_{\mathrm{FIL}} = \sum_{\mathbf{v}^k_{\mathrm{org}}} S_{\mathbf{v}^k_{\mathrm{org}}} \cdot \frac{M_{\mathrm{area}}\left(\vec{\mathbf{v}}^k_{\mathrm{ret}}\right)}{M_{\mathrm{area}}\left(\mathbf{v}^k_{\mathrm{org}}\right)} \Big/ \sum_{\mathbf{v}^k_{\mathrm{org}}} S_{\mathbf{v}^k_{\mathrm{org}}} \tag{13}$$

where $M_{\mathrm{area}}(\cdot)$ denotes the area of a grid (computed by the number of pixels in the grid). Obviously, larger value of f_{FIL} indicates more salient content is retained in the retargeted image. Particularly, $M_{\mathrm{area}}(\vec{\mathbf{v}}^k_{\mathrm{ret}}) = 0$ reflects the original grid; $\mathbf{v}^k_{\mathrm{org}}$ is directly discarded in the retargeted image.

From another aspect, the goal of backward rewarping is to measure how much image content can be correctly recovered from the retargeted image. Toward this end, we first regenerate the original image from the retargeted image based on the previous estimated similarity transformation matrices $\{\overleftarrow{\mathbb{Z}}_k\}$. Then, pixel-wise similarity measurement between the original and the recovered image is calculated by

$$\tau(\mathbf{p}_i) = \frac{2\tilde{I}_{\mathrm{org}}(\mathbf{p}_i) \cdot I_{\mathrm{org}}(\mathbf{p}_i) + R}{\tilde{I}_{\mathrm{org}}(\mathbf{p}_i)^2 + I_{\mathrm{org}}(\mathbf{p}_i)^2 + R} \tag{14}$$

where R is a small constant to avoid the denominator being zero. The similarity score falls in the range of [0, 1], where 0 indicates no similarity between two pixels and 1 denotes perfect similarity between two pixels.

Similarly, the backward information loss (BIL) is measured by using the saliency values as modulation for quality pooling

$$f_{\mathrm{BIL}} = \sum_{\mathbf{p}_i} S_{\mathbf{p}_i} \cdot \tau(\mathbf{p}_i) / \sum_{\mathbf{p}_i} S_{\mathbf{p}_i}. \tag{15}$$

To match f_{FGD} and f_{BGD} defined in (11) and (12) in which larger values mean better results, the high values of $M_{\mathrm{area}}(\vec{\mathbf{v}}^k_{\mathrm{ret}})/M_{\mathrm{area}}(\mathbf{v}^k_{\mathrm{org}})$ and $\tau(\mathbf{p}_i)$ denote more important image content is preserved in retargeting process.

It should be noted that design of our BDIL is inspired by the bidirectional salient information loss measurement (BDSIL) in [30], but has different properties in the following aspects.

1) For FIL, BDSIL estimates the saliency map of retargeted image from the original image according to SIFT-flow. The final FIL is computed as the ratio between the sums of saliency values of the retargeted and original images, while the FIL in BDIL is measured by the ratio between the areas of the rewarped and original grids.
2) For BIL, BDSIL generates a recovered image which has the same resolution with the original image from the retargeted image by SIFT-flow, while in BDIL, we regenerate the original image from the retargeted image based on the estimated similarity transformation matrices.

F. Global Salient Structure Distortion

Besides the above local measurements, we also propose a GSSD to measure the global structural information preservation. Adding GSSD measurement is mainly based on the following considerations: 1) if the similarity transformation matrices among adjacent grids in an object are quite inconsistent, the grids may suffer from inconsistent deformation, while the above BDGD measurement cannot capture such structure deformation and 2) the retargeted image should preserve saliency object information of the original image as much as possible.

To extract the salient object, we first employ the HS algorithm [38] to estimate the saliency map of original image. Then, the top 80% of total saliency values are considered as important objects in the image to be preserved. Thus, to detect structure consistency within an object, we define GSSD as

$$f_{\mathrm{SSD}} = \frac{1}{2}\left(\frac{\sum_{\mathbf{v}^k_{\mathrm{org}} \in O_{\mathrm{org}}} S_{\mathbf{v}^k_{\mathrm{org}}} \cdot \left(\sum_{k' \in \mathbb{N}_k} \sigma\left(\mathbf{v}^k_{\mathrm{org}}, \mathbf{v}^{k'}_{\mathrm{org}}\right) \cdot e^{-\|\mathbf{P}_k - \mathbf{P}_{k'}\|_2^2}\right)}{\sum_{\mathbf{v}^k_{\mathrm{org}} \in O_{\mathrm{org}}} S_{\mathbf{v}^k_{\mathrm{org}}} \cdot \left(\sum_{k' \in \mathbb{N}_k} \sigma\left(\mathbf{v}^k_{\mathrm{org}}, \mathbf{v}^{k'}_{\mathrm{org}}\right)\right)} + \frac{M_{\mathrm{area}}(O_{\mathrm{ret}})}{M_{\mathrm{area}}(O_{\mathrm{org}})}\right) \tag{16}$$

$$\sigma\left(\mathbf{v}^k_{\mathrm{org}}, \mathbf{v}^{k'}_{\mathrm{org}}\right) = \begin{cases} 1, & \text{if } \mathrm{mask}_{\mathbf{v}^k_{\mathrm{org}}} = \mathrm{mask}_{\mathbf{v}^{k'}_{\mathrm{org}}} \\ 0, & \text{otherwise} \end{cases} \tag{17}$$

where $\mathbb{N}_k$ is a set of neighboring meshes to $\mathbf{v}^k_{\mathrm{org}}$; $\mathbf{P}_k$ and $\mathbf{P}_{k'}$ are the similarity transformation matrices for grids $\mathbf{v}^k_{\mathrm{org}}$ and $\mathbf{v}^{k'}_{\mathrm{org}}$, respectively; $\mathrm{mask}_{\mathbf{v}^k_{\mathrm{org}}}$ and $\mathrm{mask}_{\mathbf{v}^{k'}_{\mathrm{org}}}$ represent the object indexes for grids $\mathbf{v}^k_{\mathrm{org}}$ and $\mathbf{v}^{k'}_{\mathrm{org}}$, respectively; and O_{org} and O_{ret} are the salient objects in the original image and retargeted image, respectively.

G. Quality Fusion

With the estimated scores f_{FGD}, f_{BGD}, f_{FIL}, f_{BIL}, and f_{SSD}, we map the five-dimensional quality vectors to the associated quality score via a pretrained prediction function. In our method, we adopt the well-known SVR algorithm to train the model. In our experiment, we apply polynomial kernel to fuse the individual quality scores (different quality pooling schemes will be analyzed in the next section).

IV. Experimental Results and Analyses

A. Databases

For the experiment, we test our TRASIM method on two widely used IRQA datasets: 1) RetargetMe [9] and 2) CUHK [39]. The RetargetMe dataset [9] consists of 37 original images. For each original image, eight different retargeting operators are employed to generate retargeted images, including SCL, CR, SC [4], SM [5], WARP [6], SNS [8], SV [7], and multioperator (MULTI) [14]. There are 296 retargeted images in the RetargetMe database. The subjective rank score of each retargeted image is recorded as the number of times that it is preferred over others. Refer to [17], Kendall rank correlation coefficient (KRCC) is applied to estimate the model's performance, which is defined as

$$\mathrm{KRCC} = \frac{n_c - n_d}{0.5n(n-1)} \tag{18}$$

where n is the ranking length, and n equals 8 in RetargetMe dataset, n_c and n_d are the number of concordant and discordant pairs, respectively.

The CUHK dataset [39] includes 57 source images and 171 retargeted images generated from three retargeting operators. For each source image, the employed retargeting operators may not be the same because these operators are randomly selected from ten typical retargeting methods, including optimized seam carving and scale (SCSC) [40], energy-based deformation (ENER) [41], and other eight operators used in the RetargetMe database. The subjective study in CUHK is quite different from RetargetMe. Each retargeted image is provided an associated mean opinion score (MOS) value corresponding to five category quality scales. To evaluate model's performance, Pearson linear correlation coefficient (PLCC), Spearman rank order correlation coefficient (SRCC), root mean square error (RMSE), and outlier ratio (OR) between objective and subjective values are calculated. PLCC, RMSE, and OR measurements are obtained after nonlinear regression, which is defined as

$$f(x) = \beta_1 \cdot \left(\frac{1}{2} - \frac{1}{1 + \exp(\beta_2 \cdot (x - \beta_3))}\right) + \beta_4 \cdot x + \beta_5 \tag{19}$$

where β_1, β_2, β_3, β_4, and β_5 are parameters determined by minimizing the sum of squared differences between subjective and objective scores.

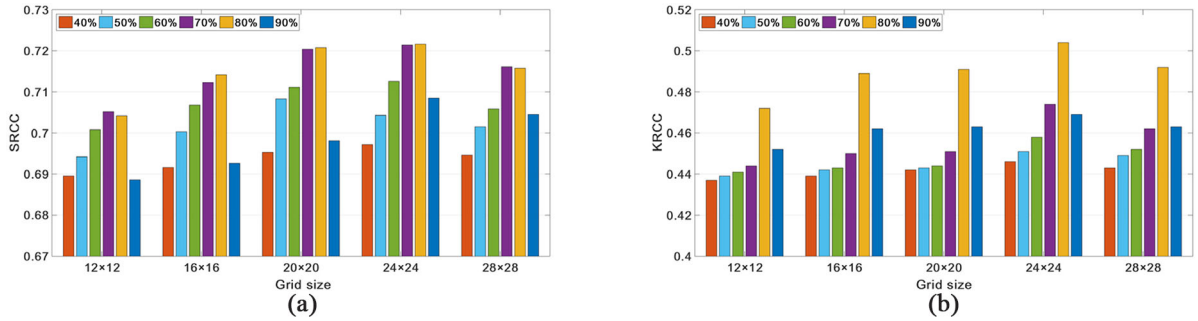


Fig. 6. Performance results based on different parameter combinations on (a) CUHK and (b) RetargetMe.

Since we use fivefold SVR to fuse the predicted quality scores, it is necessary to construct a training set to train the model. In the CUHK database, the dataset is randomly divided into five groups. For each group, it is adopted for testing and the remaining four group is selected for training. Such train-test procedure is repeated five rounds to obtain the mean prediction as the final performance measure. Since the rank score for each image in the RetargetMe dataset only denotes the relative quality against other seven retargeted images generated from a same source image, it cannot reflect the actual perceptual quality against all other images in the database. Therefore, we train an SVR model via using all images in the CUHK database, then the trained model is employed to predict the quality of retargeted images in the RetargetMe database. Note that if other rank learning models (e.g., learning to rank [42], rank SVM [43]) can be used, the generalization ability of our model can be further examined via cross-database evaluation.

B. Influence of Parameter Setting

In our method, we need to set the size of the similarity transformation grid and the percentage of total saliency values for important object extraction. In our method, the grid size will have certain impact on the similarity transformation, and hence influence the BDGD and BDIL measurements, while the percentage threshold will affect the number of important objects and hence influence the result of GSSD measurement. As shown in Fig. 6, we test different combinations of grid sizes and saliency percentages. The test results show that the choice of grid size and saliency percentage has certain impact on the quality prediction, and the combination of $\{24 \times 24, 80\%\}$ has a relatively high performance on both CUHK and RetargetMe databases.

C. Performance Comparisons With Other Methods

We compare our approach with 11 state-of-the-art IRQA methods, including BDS [13], EH [19], SIFT-flow [15], EMD [16], CSim [22], IR-SSIM [23], GLS [44], PGDIL [24], ARS [17], Karimi's method [18], two recent training-based methods (Jiang's method [27] and Chen's method [30]), and two recent ranking-based methods (Ma's method [31] and Chen's method [32]). Table IV gives comparisons of mean KRCC values on each subset and all the images of RetargetMe dataset, as well as standard deviations of KRCC values and p-value on all images. Table V gives comparisons of PLCC, SRCC, RMSE, and OR values on the CUHK dataset. We can make the following observations.

1) In the RetargetMe database, the results of our TRASIM model are better than other IRQA metrics on most image sets, especially in image sets of line edge, faces people, and geometric structure. The reason may be that the proposed transformation-driven approach is more effective to measure structure and shape distortion by comprehensively taking geometric distortion, information loss, and saliency distortion into account. The overall performance of our proposed TRASIM model is better than all comparison methods.
2) In the CUHK database, the proposed TRASIM outperforms most metrics and has almost identical performance with Chen's method [30].
3) Compared with two training-based methods [27], [30] that are similar with our method in quality fusion, our method still has the advantage in predicting the perceptual quality, i.e., higher assessment accuracy compared with Jiang's method [27] and lower feature dimension compared with Chen's method [30].
4) Compared with two ranking-based methods [31], [32] that rely on the learned ranking models, our method is better than Chen's method [32] on all images and most individual image sets except for foreground objects and symmetry. Besides, the evaluation results of our approach are better than Ma's method [31] since Ma's method works in a no-reference manner. Overall, combination of BDGD, BDIL, and GSSD measurements achieves great performance promotion for predicting the retargeted image quality.

D. Performance of Single Quality Component

In this section, the validity of five proposed quality components in the transformation-driven model are analyzed: the BDGD measurement (f_{FGD} and f_{BGD}), the BDIL measurement (f_{FIL} and f_{BIL}), and the GSSD measurement (f_{SSD}). Table VI shows the comparison results of these quality components. As shown in the table, the following observations can be obtained.

1) Each independent measurement is not always superior to other methods (e.g., ARS) in Tables IV and V on two databases because each measurement is complementary in capturing the structural distortion or content

TABLE IV
PERFORMANCE OF DIFFERENT METRICS ON RETARGETME DATABASE. THE TOP TWO OF EACH TYPE ARE IN BOLD AND THE BEST RESULTS ARE IN BLUE

Metric	Mean KRCC on each subset						Total		
	Line Edge	Faces People	Foreground Objects	Texture	Geometric Structure	Symmetry	Mean KRCC	Std KRCC	p-val
BDS [13]	0.040	0.190	0.67	0.060	-0.004	-0.012	0.083	0.268	0.107
EH [19]	0.043	-0.076	-0.079	-0.060	0.103	0.298	0.004	0.334	0.641
SIFT-flow [15]	0.097	0.252	0.218	0.161	0.085	0.071	0.145	0.262	0.031
EMD [16]	0.220	0.262	0.226	0.107	0.237	0.500	0.251	0.272	1e-5
CSim [22]	0.097	0.290	0.293	0.161	0.053	0.150	0.164	0.263	0.028
IR-SSIM [23]	0.309	0.452	0.377	0.321	0.313	0.333	0.363	0.271	1e-3
PGDIL [24]	0.431	0.390	0.389	0.286	0.438	0.523	0.415	0.296	6e-10
ARS [17]	**0.463**	0.519	0.444	0.330	**0.505**	0.464	0.452	0.283	1e-11
Karimi [18]	0.453	**0.589**	**0.564**	**0.494**	0.431	0.380	**0.494**	0.261	1e-10
Chen [30]	0.448	0.552	0.494	0.423	0.497	0.471	0.473	0.257	—
Ma [31]	0.229	0.273	0.182	0.218	0.252	0.484	0.477	—	—
Chen [32]	0.437	0.505	**0.536**	0.429	0.438	**0.536**	0.473	—	—
TRASIM	**0.465**	**0.600**	0.532	**0.476**	**0.522**	**0.524**	**0.504**	0.237	4e-15

TABLE V
PERFORMANCE OF DIFFERENT METRICS ON CUHK DATABASE. THE TOP TWO OF EACH TYPE ARE IN BOLD AND THE BEST RESULTS ARE IN BLUE

Metric	PLCC	SRCC	RMSE	OR
BDS [13]	0.2896	0.2887	12.922	0.2164
EH [19]	0.3422	0.3288	12.686	0.2047
SIFT-flow [15]	0.3141	0.2899	12.817	0.1462
EMD [16]	0.2760	0.2904	12.977	0.1696
CSim [22]	0.4374	0.4662	12.141	0.1520
GLS [44]	0.4622	0.4760	10.932	0.1345
PGDIL [24]	0.5403	0.5409	11.361	0.1520
ARS [17]	0.6835	0.6693	9.855	0.0702
Jiang [27]	0.644	0.616	10.763	—
Chen [30]	**0.7528**	**0.7514**	—	—
Ma [31]	0.5371	0.4926	—	0.1928
TRASIM	**0.7586**	**0.7216**	8.609	0.0103

loss for different retargeted images, while other methods measure geometric distortion and information loss, but combination of BDGD, BDIL, and GSSD in our model can achieve the best performance.

2) Among the five measurements, f_{FGD} is an essential component to measure geometric distortion whose performance is better than other four measurements, but f_{SSD} is also critical to evaluate geometric distortion and information loss via GSSD. To further demonstrate the advantage of GSSD measurement, we analyze the performance of TRASIM with GSSD or without GSSD in Table VII. It is observed that GSSD measurement can effectively improve the final performance. Compared with the methods mainly for measuring structure distortion [23], [24], [30], f_{SSD} still has comparative performance on two databases.
3) The f_{BIL} has the lower values for all retargeted images compared with f_{FIL} because the BIL has poor similarity for the hole areas after rewarping. Overall, the forward measurements (e.g., f_{FGD} and f_{FIL}) and the GSSD measurement (e.g., f_{SSD}) are relatively stable for most retargeting operators, while the backward measurements (e.g., f_{BGD} and f_{BIL}) have favorable effects for some specific retargeting operators.

Further, we provide an example to show the influence of each component in predicting the quality values of different retargeted images. Table VIII reports the quality values of the retargeted images in Fig. 7. Larger quality values mean better results. From the table, we find that each metric has its advantages and disadvantages in reflecting the influence of different retargeting operators. For the cropped image in (b), f_{BGD} is close to 1 because each grid in the retargeted image can find an almost same one in the original image (as clearly demonstrated in the backward rewarped image in Fig. 4), but such measurement in fact cannot reflect content loss. On the other hand, f_{BGD} has a lower value for the retargeted image generated by SCL because the image is seriously deformed after SCL.

E. Influence of Different Quality Pooling Methods

Since the final evaluation performance is highly dependent on the used quality pooling method, to further investigate the influence of different quality pooling methods on final quality prediction, refer to [26] and [27], we use seven linear or nonlinear quality pooling methods to fuse five individual quality components, including linear average, direct multiplication linear regression, logistic regression, SVR with linear kernel (linear-SVR), SVR with RBF kernel (RBF-SVR), and SVR with polynomial kernel (poly-SVR). The train-test procedure is the same with the above experiment. Test results are shown in Table IX, we can observe that the nonlinear models always perform better than the linear ones. For the first two average and multiply combination schemes without needing training process, the worst results are obtained among all

TABLE VI
PERFORMANCE OF SINGLE QUALITY COMPONENT ON CUHK AND RETARGETME DATABASES

Dataset	Criteria	f_{FGD}	f_{BGD}	f_{FIL}	f_{BIL}	f_{SSD}	TRASIM
CUHK	PLCC	0.6948	0.4899	0.5235	0.6311	0.5513	**0.7586**
	SRCC	0.6514	0.4787	0.3136	0.6303	0.5471	**0.7216**
	RMSE	9.7102	11.7740	11.5031	10.6316	11.2667	**8.6090**
	OR	0.0117	0.0175	0.0117	0.0294	0.0294	**0.0103**
RetargetMe	KRCC	0.436	0.344	0.141	0.284	0.371	**0.504**

TABLE VII
PERFORMANCE COMPARISONS OF THE GSSD MEASUREMENT

Dataset	Criteria	IR-SSIM [23]	PGDIL [24]	Chen [30]	f_{SSD}	W/O GSSD	W/ GSSD
CUHK	PLCC	—	0.5403	0.7528	0.5513	0.7348	0.7586
	SRCC	—	0.5409	0.7514	0.5471	0.7042	0.7216
	RMSE	—	11.361	—	11.2667	8.9132	8.6090
RetargetMe	KRCC	0.363	0.415	0.473	0.371	0.484	0.504

TABLE VIII
PREDICTED QUALITY SCORES OF DIFFERENT METRICS FOR RETARGETED IMAGES IN FIG. 7

Image	Metric	(b)	(c)	(d)	(e)
Umdan	f_{FGD}	0.6775 (1)	0.5948 (3)	0.5879 (4)	0.6124 (2)
	f_{BGD}	0.9911 (1)	0.4652 (2)	0.2110 (4)	0.3704 (3)
	f_{FIL}	0.6624 (1)	0.6189 (2)	0.5240 (4)	0.5568 (3)
	f_{BIL}	0.6368 (1)	0.5056 (2)	0.4999 (4)	0.5049 (3)
	f_{SSD}	0.8422 (1)	0.6013 (4)	0.7239 (2)	0.6395 (3)
	final	59.386 (1)	39.933 (4)	43.528 (2)	42.037 (3)
	rank	48 (1)	4 (4)	34 (2)	27 (3)
Butterfly	f_{FGD}	0.8240 (1)	0.7831 (2)	0.5970 (4)	0.7336 (3)
	f_{BGD}	0.9944 (1)	0.7659 (3)	0.1918 (4)	0.8892 (2)
	f_{FIL}	0.8294 (1)	0.7467 (2)	0.5182 (4)	0.6652 (3)
	f_{BIL}	0.7992 (1)	0.7152 (2)	0.4977 (4)	0.6375 (3)
	f_{SSD}	0.8905 (1)	0.7684 (2)	0.7182 (3)	0.6914 (4)
	final	65.372 (1)	54.243 (2)	44.024 (4)	53.003 (3)
	rank	52 (1)	44 (2)	6 (4)	10 (3)

the compared schemes. As expected, the best two evaluation results are obtained by applying SVR with RBF and polynomial kernels, and we utilize the widely used polynomial kernel in the experiment.

We also apply leave-one-out cross-validation with Poly-SVR on CUHK database for quality pooling. We use 170 images for training and test the rest one image. Thus, via leave-one-out cross-validation, each image is tested only once. As shown in Table X, compared with the adopted fivefold cross-validation strategy also with Poly-SVR, leave-one-out strategy perform better in predicting the quality, indicating the effectiveness of the derived f_{FGD}, f_{BGD}, f_{FIL}, f_{BIL}, and f_{SSD}.

F. Impact of Different Saliency Models

Acquiring saliency map is important for image retargeting to preserve visually important regions and is also very essential for IRQA to measure salient information. To further analyze the performance of different saliency models on the proposed method, we adopt six widely used saliency detection approaches to generate saliency map in the BDGD, BDIL, and GSSD measurements, including spectral residual (SR) [45] (#1), Itti's model [46] (#2), maximum symmetric surround (MSS) [47] (#3), graph-based visual saliency (GBVS) [48] (#4), minimum barrier (MB) [49] (#5), and HS [38] (#6). The comparison results are shown in Table XI. As can be observed, the saliency model will highly affect the quality prediction. Among these saliency models, HS (#6) produces the best results. The reason may be that HS (#6) can detect more complete and accurate salient objects compared with other algorithms (clearly shown in Fig. 7). It is assumed the saliency detection algorithm used in generated a retargeted image is unknown, the current TRASIM metric is not the optimal in selecting the saliency model, and further exploration on the saliency model can further promote our measurement.

G. Application to Multioperator Image Retargeting

Besides evaluating the quality of a retargeted image using our metric, another important application of IRQA is to guide the optimization of multioperator retargeting [50]. As discussed, if we hope to resize an image with a width of m into the desired one with a width of m' using a set of k operators ($m > m'$), there are $k^{(m-m')}$ possible combinations in the multioperator retargeting. Therefore, it requires an objective IRQA with high accuracy in the optimization function to find the optimal operator at each resizing iteration

$$\hat{\Gamma} = \arg\max_{\Gamma} \text{TRASIM}(\Gamma(\mathbf{x})) \tag{20}$$

where $\mathbf{x}$ is an input image signal, and $\Gamma(\cdot)$ indicates a particular image retargeting operator. On this basis, it is natural to embed our TRASIM metric into the multioperator framework.

In Fig. 8, we develop a multioperator method that employs an iterative algorithm to determine the sequence of retargeting operators by TRASIM. Three operators including simple CR, SCL, and SC are independently used and the best retargeting

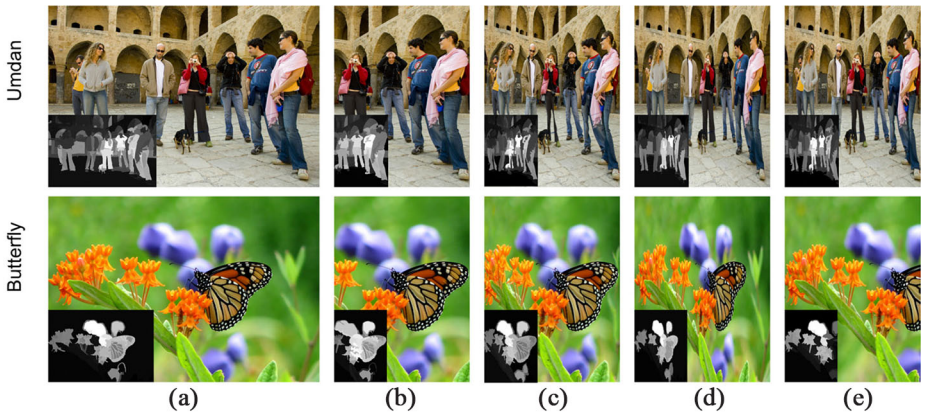


Fig. 7. Two original images and the corresponding retargeted images obtained by four typical retargeting operators. (a) Original. (b) CR. (c) SC. (d) SCL. (e) WARP.

TABLE IX
PERFORMANCE RESULTS USING DIFFERENT QUALITY POOLING METHODS ON CUHK AND RETARGETME DATABASES

Dataset	Criteria	Average	Multiply	Linear regression	Logistic regression	Linear-SVR	RBF-SVR	Poly-SVR
CUHK	PLCC	0.5796	0.6150	0.6890	0.7235	0.7429	0.7542	**0.7586**
	SRCC	0.5695	0.5693	0.6548	0.6862	0.6858	0.7202	**0.7216**
	RMSE	11.0134	10.6462	9.4523	9.0274	8.7746	8.6179	**8.6090**
	OR	0.0234	0.0175	0.0196	0.0117	0.0117	0.0105	**0.0103**
RetargetMe	KRCC	0.380	0.376	0.458	0.472	0.473	0.494	**0.504**

TABLE X
COMPARISON OF DIFFERENT CROSS-VALIDATION STRATEGIES

Metric	PLCC	SRCC	RMSE	OR
Leave-one-out	0.7628	0.7260	8.7295	0.0058
5-fold	0.7586	0.7216	8.6090	0.0103

result is selected based on TRASIM metric at each iteration. The final retargeting results produced by our TRASIM-guided multioperator are shown in Fig. 9. Our multioperator achieves the best retargeting results compared with CR, SCL, SC, and MULTI [14]. Additionally, we also find that CR operator is selected more frequently during the multioperator retargeting process. The reason might be that CR can preserve image structural information better than SCL or SC, especially when the cropped region contains most of salient objects. Note that the results in Fig. 9 are not optimized in terms of individual operator selection, iteration times, and operator sequences. The solutions for these optimizations can be found in [14] and [50]. Therefore, our TRASIM metric can be well embedded into the existing multioperator framework for designing more powerful retargeting operators.

H. Computational Complexity Analysis

We further conduct a time complexity comparison for the proposed method and ARS metric [17] with the same computer environment (CPU: Intel Core i5-6500@3.2-GHz, memory: 8.00 GB, Win 10+ MATLAB R2014b). In the experiment, the original image size is 1024×813 and the

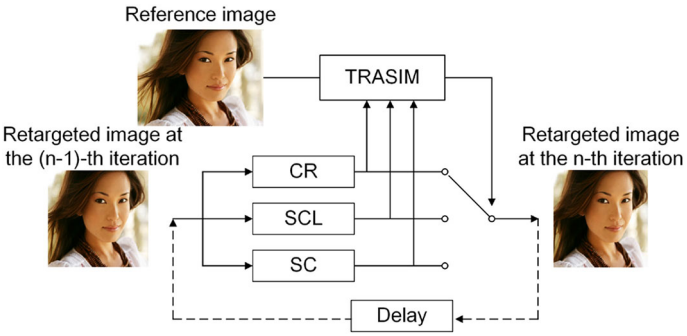


Fig. 8. Diagram of TRASIM-guided multioperator image retargeting algorithm.

retargeted image size is 768×813. Testing results are shown in Table XII (due to the limitation of space in the table, we use "DC" and "F-EF" to represent "dense correspondence" and "feature extraction and fusion," respectively). From the table, we can observe that the computational complexity of TRASIM is higher than ARS. The main reason is that TRASIM applies SIFT-flow algorithm for forward and backward matching between the retargeted and original images, which is time consuming. However, TRASIM can obtain better evaluation results than ARS, as shown in the previous experiments.

I. Limitations

In this paper, bidirectional rewarping for quality assessment of retargeted images are considered. Although the proposed

TABLE XI
PERFORMANCE OF TRASIM USING DIFFERENT SALIENCY MODELS ON CUHK AND RETARGETME DATABASES

Dataset	Criteria	#1	#2	#3	#4	#5	#6
CUHK	PLCC	0.6682	0.6745	0.7138	0.7322	0.7246	**0.7586**
	SRCC	0.6413	0.6616	0.6832	0.6951	0.6876	**0.7216**
	RMSE	9.8543	9.7748	9.1881	8.8640	8.9753	**8.6090**
	OR	0.0175	0.0151	0.0128	0.0106	0.0117	**0.0103**
RetargetMe	KRCC	0.3063	0.3147	0.3784	0.4492	0.4846	**0.5039**

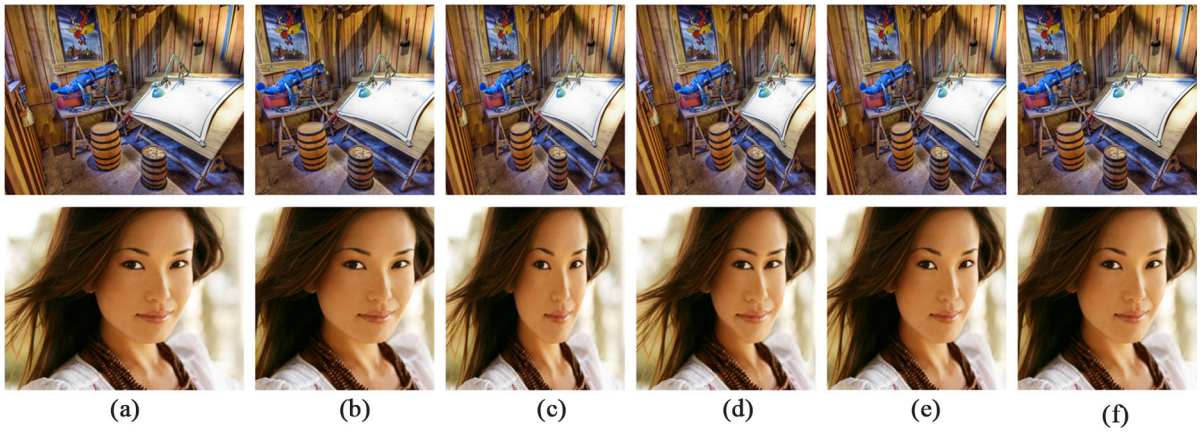


Fig. 9. Retargeting results. (a) Reference images. (b) Retargeted images by CR. (c) Retargeted images by SCL. (d) Retargeted images by SC. (e) Retargeted images by MULTI [14]. (f) Retargeted images by TRASIM-guided multioperator.

TABLE XII
COMPUTATIONAL COMPLEXITY OF DIFFERENT METRICS

Metric	Time Cost (Sec)		
	DC	F-EF	All
ARS	26.2933	0.2146	26.5079
TRASIM	49.1985	2.3767	51.5752

method demonstrates prominent performance in comparison with other IRQA approaches, it still has several limitations.

1) Our TRASIM model is highly dependent on the accuracy of SIFT-flow algorithm to establish the rewarping relationship. Nevertheless, SIFT-flow has its limitation in obtaining structural information, especially in smooth areas, leading to incorrect transformation for these areas.
2) Since we use rewarping rule to simulate the retargeting modification, it may encounter the common limitations of warping-based retargeting approaches that is it may fail to measure the shapes of important structures.
3) For the limitation of image retargeting databases, only key shape distortion and information loss are considered. In fact, each subset generated by different retargeting operators has its distinctive characteristic in presenting the retargeting semantics. Therefore, combination of universal and specific retargeting features may be a feasible solution.

V. CONCLUSION

In this paper, a TRASIM measurement model for IRQA is proposed. For this goal, we apply similarity transformation to establish and clarify the relationship between the original and retargeted images via bidirectional rewarping. To acquire the final quality measurement, we measure BDGD, BDIL, and GSSD to reflect geometric distortion and content loss. Compared with state-of-the-art IRQA metrics, the TRASIM achieved satisfactory evaluation results on both CUHK and RetargetMe datasets. In the future, we plan to solve the transformation with more accurate irregular grid partition, and concentrate on accurate detection of retargeting semantics.

REFERENCES

[1] F. Shao, W. Lin, W. Lin, Q. Jiang, and G. Jiang, "QoE-guided warping for stereoscopic image retargeting," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4790–4805, Oct. 2017.
[2] B. Li, L.-Y. Duan, C.-W. Lin, T. Huang, and W. Gao, "Depth-preserving warping for stereo image retargeting," *IEEE Trans. Image Process.*, vol. 24, no. 9, pp. 2811–2826, Sep. 2015.
[3] W. Lin and C.-C. J. Kuo, "Perceptual visual quality metrics: A survey," *J. Vis. Commun. Image Represent.*, vol. 22, no. 4, pp. 297–312, 2011.
[4] M. Rubinstein, A. Shamir, and S. Avidan, "Improved seam carving for video retargeting," *ACM Trans. Graph.*, vol. 27, no. 3, Nov. 2008, Art. no. 6.
[5] Y. Pritch, E. Kav-Venaki, and S. Peleg, "Shift-map image editing," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Kyoto, Japan, 2009, pp. 151–158.
[6] L. Wolf, M. Guttmann, and D. Cohen-Or, "Non-homogeneous content-driven video-retargeting," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2007, pp. 1–6.
[7] P. Krähenbühl, M. Lang, A. Hornung, and M. H. Gross, "A system for retargeting of streaming video," *ACM Trans. Graph.*, vol. 28, no. 5, 2009, Art. no. 126.
[8] Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee, "Optimized scale-and-stretch for image resizing," *ACM Trans. Graph.*, vol. 27, no. 5, 2008, Art. no. 118.
[9] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir, "A comparative study of image retargeting," *ACM Trans. Graph.*, vol. 29, no. 6, 2010, Art. no. 160.

[10] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.

[11] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[12] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb. 2014.

[13] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani, "Summarizing visual data using bidirectional similarity," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2008, pp. 1–8.

[14] M. Rubinstein, A. Shamir, and S. Avidan, "Multi-operator media retargeting," *ACM Trans. Graph.*, vol. 28, no. 3, 2009, Art. no. 23.

[15] C. Liu, J. Yuen, and A. Torralba, "SIFT flow: Dense correspondence across scenes and its applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 978–994, May 2011.

[16] O. Pele and M. Werman, "Fast and robust earth mover's distances," in *Proc. IEEE Int. Conf. Comput. Vis. ICCV)*, 2009, pp. 460–467.

[17] Y. Zhang, Y. Fang, W. Lin, X. Zhang, and L. Li, "Backward registration-based aspect ratio similarity for image retargeting quality assessment," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4286–4297, Sep. 2016.

[18] M. Karimi, S. Samavi, N. Karimi, S. M. R. Soroushmehr, W. Lin, and K. Najarian, "Quality assessment of retargeted images by salient region deformity analysis," *J. Vis. Commun. Image Represent.*, vol. 43, pp. 108–118, Feb. 2017.

[19] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, pp. 703–715, Jun. 2001.

[20] E. Kasutani and A. Yamada, "The MPEG-7 color layout descriptor: A compact image feature description for high-speed image/video segment retrieval," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2001, pp. 674–677.

[21] L. Ma, L. Xu, H. Zeng, K. N. Ngan, and C. Deng, "How does the shape descriptor measure the perceptual quality of the retargeting image?" in *Proc. ICME Workshops*, 2014, pp. 1–6.

[22] Y.-J. Liu, X. Luo, Y. Xuan, W. Chen, and X.-L. Fu, "Image retargeting quality assessment," *Comput. Graph. Forum*, vol. 30, no. 2, pp. 583–592, 2011.

[23] 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 Trans. Emerg. Sel. Topics Circuits Syst.*, vol. 4, no. 1, pp. 95–105, Mar. 2014.

[24] 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 J. Sel. Topics Signal Process.*, vol. 8, no. 3, pp. 377–389, Jun. 2014.

[25] Y. Liang, Y.-J. Liu, and D. Gutierrez, "Objective quality prediction of image retargeting algorithms," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 2, pp. 1099–1110, Feb. 2017.

[26] S. A. F. Oliveira, S. S. A. Alves, J. P. P. Gomes, and A. R. R. Neto, "A bi-directional evaluation-based approach for image retargeting quality assessment," *Comput. Vis. Image Understand.*, vol. 168, pp. 172–181, Mar. 2018.

[27] Q. Jiang, F. Shao, W. Lin, and G. Jiang, "Learning sparse representation for objective image retargeting quality assessment," *IEEE Trans. Cybern.*, vol. 48, no. 4, pp. 1276–1289, Apr. 2018.

[28] B. Yan, B. Bare, K. Li, J. Li, and A. C. Bovik, "Learning quality assessment of retargeted images," *Signal Process. Image Commun.*, vol. 56, pp. 12–19, Aug. 2017.

[29] A. Liu, W. Lin, H. Chen, and P. Zhang, "Image retargeting quality assessment based on support vector regression," *Signal Process. Image Commun.*, vol. 39, pp. 444–456, Nov. 2015.

[30] Z. Chen, J. Lin, N. Liao, and C. W. Chen, "Full reference quality assessment for image retargeting based on natural scene statistics modeling and bi-directional saliency similarity," *IEEE Trans. Image Process.*, vol. 26, no. 11, pp. 5138–5148, Nov. 2017.

[31] L. Ma, L. Xu, Y. Zhang, Y. Yan, and K. N. Ngan, "No-reference retargeted image quality assessment based on pairwise rank learning," *IEEE Trans. Multimedia*, vol. 18, no. 11, pp. 2228–2237, Nov. 2016.

[32] Y. Chen, Y.-J. Liu, and Y.-K. Lai, "Learning to rank retargeted images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 4743–4751.

[33] Y. Zhang, W. Lin, Q. Li, W. Cheng, and X. Zhang, "Multiple-level feature-based measure for retargeted image quality," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 451–463, Jan. 2018.

[34] Y. Zhang, K. N. Ngan, L. Ma, and H. Li, "Objective quality assessment of image retargeting by incorporating fidelity measures and inconsistency detection," *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5980–5993, Dec. 2017.

[35] G.-X. Zhang, M.-M. Cheng, S.-M. Hu, and R. R. Martin, "A shape-preserving approach to image resizing," *Comput. Graph. Forum*, vol. 28, no. 7, pp. 1897–1906, 2009.

[36] Y. Guo, F. Liu, J. Shi, Z.-H. Zhou, and M. Gleicher, "Image retargeting using mesh parametrization," *IEEE Trans. Multimedia*, vol. 11, no. 5, pp. 856–867, Aug. 2009.

[37] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 138–156, 2000.

[38] Q. Yan, L. Xu, J. Shi, and J. Jia, "Hierarchical saliency detection," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Portland, OR, USA, 2013, pp. 1155–1162.

[39] L. Ma, W. Lin, C. Deng, and K. N. Ngan, "Image retargeting quality assessment: A study of subjective scores and objective metrics," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 6, pp. 626–639, Oct. 2012.

[40] W. Dong, N. Zhou, J.-C. Paul, and X. Zhang, "Optimized image resizing using seam carving and scaling," *ACM Trans. Graph.*, vol. 28, no. 5, 2009, Art. no. 125.

[41] Z. Karni, D. Freedman, and C. Gotsman, "Energy-based image deformation," *Comput. Graph. Forum*, vol. 28, no. 5, pp. 1257–1268, 2009.

[42] H. Li, "A short introduction to learning to rank," *IEICE Trans. Inf. Syst.*, vol. 94, no. 10, pp. 1854–1862, 2011.

[43] A. Elisseeff and J. Weston, "A kernel method for multi-labelled classification," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2002, pp. 681–687.

[44] J. Zhang and C.-C. J. Kuo, "An objective quality of experience (QoE) assessment index for retargeted images," in *Proc. ACM Int. Conf. Multimedia (MM)*, Nov. 2014, pp. 257–266.

[45] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2007, pp. 1–8.

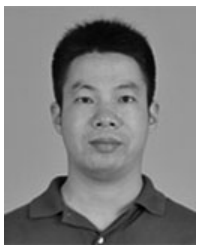
[46] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.

[47] R. Achanta and S. Süsstrunk, "Saliency detection using maximum symmetric surround," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2010, pp. 2653–2656.

[48] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2006, pp. 545–552.

[49] J. Zhang, S. Sclaroff, Z. Lin, X. Shen, B. Price, and R. Mech, "Minimum barrier salient object detection at 80 FPS," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2015, pp. 1404–1412.

[50] Y. Fang, Z. Fang, F. Yuan, Y. Yang, S. Yang, and N. N. Xiong, "Optimized multi-operator image retargeting based on perceptual similarity measure," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 11, pp. 2956–2966, Nov. 2017.



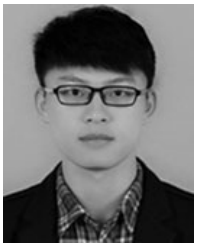
Feng Shao (M'16) received the B.S. and Ph.D. degrees in electronic science and technology from Zhejiang University, Hangzhou, China, in 2002 and 2007, respectively.

He is currently a Professor with the Faculty of Information Science and Engineering, Ningbo University, Ningbo, China. In 2012, he was a Visiting Fellow with the School of Computer Engineering, Nanyang Technological University, Singapore, for six months. He has published over 100 technical articles in refereed journals and proceedings in the areas of three-dimensional (3-D) video coding, 3-D quality assessment, and image perception.

Dr. Shao was a recipient of the "Excellent Young Scholar" Award by NSF of China in 2016.

Zhenqi Fu received the B.S. degree in electronic information engineering from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2016. He is currently pursuing the M.S. degree in signal and information processing with Ningbo University, Ningbo, China.

His current research interests include image/video processing and quality assessment.



Qiuping Jiang (S'17) received the Ph.D. degree in signal and information processing from Ningbo University, Ningbo, China, in 2018.

He is currently an Associate Professor with the School of Information Science and Engineering, Ningbo University. From 2017 to 2018, he was a visiting student with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His current research interests include image processing, visual perception modeling, and computer vision.

Dr. Jiang was a recipient of the JVCI 2017 Best Paper Award Honorable Mention as the first author. He is a Reviewer for several prestigious journals and conferences, such as the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON IMAGE PROCESSING, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE TRANSACTIONS ON MULTIMEDIA, IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS, ICME, and ICIP.

Gangyi Jiang (A'04–M'14) received the M.S. degree in electronic engineering from Hangzhou University, Hangzhou, China, in 1992 and the Ph.D. degree in electronic engineering from Ajou University, Suwon, South Korea, in 2000.

He is currently a Professor with the Faculty of Information Science and Engineering, Ningbo University, Ningbo, China. His current research interests include digital video compression and multiview video coding.

Yo-Sung Ho (SM'06–F'16) received the B.S. and M.S. degrees in electronic engineering from Seoul National University, Seoul, South Korea, in 1981 and 1983, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara, Santa Barbara, CA, USA, in 1990.

He joined the Electronics and Telecommunications Research Institute (ETRI), Daejon, South Korea, in 1983. From 1990 to 1993, he was with Philips Laboratories, Briarcliff Manor, NY, USA, where he was involved in development of the advanced digital high-definition television system. In 1993, he rejoined ETRI as a Technical Staff and was involved in development of the Korean DBS digital television and high-definition television systems. Since 1995, he has been with the Gwangju Institute of Science and Technology, Gwangju, South Korea, where he is currently a Professor of the Information and Communications Department. His current research interests include digital image and video coding, image analysis and image restoration, advanced video coding techniques, digital video and audio broadcasting, three-dimensional video processing, and content-based signal representation and processing.