Computer Vision and Image Understanding 168 (2018) 172–181

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

A bi-directional evaluation-based approach for image retargeting quality assessment

Saulo A.F. Oliveira[*,a], Shara S.A. Alves[b], João P.P. Gomes[a], Ajalmar R. Rocha Neto[b]

[a] Federal University of Ceará, Department of Computer Science, Fortaleza, Ceará, Brazil
[b] Federal Institute of Ceará, Department of Teleinformatics, Fortaleza, Ceará, Brazil

ARTICLE INFO

Keywords:
Image retargeting
Feature extraction
Keypoint detection
Quality assessment

ABSTRACT

Image retargeting is a technique that adjusts input images into arbitrary dimensions (rows and columns) and simultaneously preserves regions of interest. Assess the image quality under varying aspect ratio is significantly more challenging since it requires content matching in addition to semantic content analysis. In this work, we propose an objective quality assessment algorithm for image retargeting, called bi-directional importance map similarity (BIMS). The key step in our approach is to assess quality in image retargeting through some features in a bi-directional way, all in a feature fusion framework. The motivation behind employing bi-directional features is because the nature of them is useful to estimate pertinent locations where we can analyze whenever relevant content is missing or any visual distortion arises. Our proposal was assessed on a well-known state-of-the-art dataset in which human viewers provided their personal opinions on the perceptual quality. Due to the experimental results obtained, we consider the BIMS is a good choice for quality assessment of retargeted images.

1. Introduction

In image processing, retargeting is a method which aims at adjusting input images into arbitrary dimensions and also preserving their regions of interest (ROIs). In other words, the idea is to resize an image while taking its content into consideration to preserve important regions and minimize distortions. The retargeting problem can be stated as follows. Let $\boldsymbol{I}$ be an input image of size $m \times n$, where m is the number of rows and n is the number of columns. Similarly, let $\boldsymbol{J}$ be an output image of size $m' \times n'$, where $m' < m$ and $n' < n$ for reduction. The objective is then to produce a new image $\boldsymbol{J}$ which will be a good representative of the original image $\boldsymbol{I}$. There is no clear definition or measure as to the quality of $\boldsymbol{J}$ being a good representative of $\boldsymbol{I}$. Roughly speaking, retargeting is in general applied to (*i*) preserve the input image content, (*ii*) preserve the input image structure, as well as to (*iii*) achieve a resulting artifact-free image (Shamir and Sorkine, 2009). Furthermore, the problem of image retargeting can be easily extended to videos as well (Guttmann et al., 2011). An example of a retargeting in which the original image had its width decreased by different retargeting algorithms is depicted in Fig. 1.

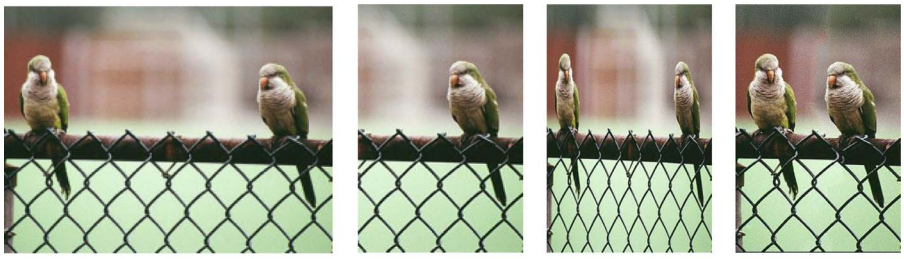
Popular image quality algorithms (IQA), such as the Peak Signal-to-Noise-Ratio (PSNR), the Structural Similarity Index (SSIM) (Wang et al., 2004), the Visual Information Fidelity Index (VIF) (Sheikh and Bovik, 2006), or even the Mean Absolute Error (MAE) can not be applied directly in retargeting applications because they require the sizes of the input (reference) and output (retargeted) images to be the same. As highlighted in Rubinstein et al. (2010), designing a quality metric for retargeting that compares image content under varying aspect ratio is significantly more challenging since the problem also demands semantic image analysis and content matching (see Fig. 2).

Image retargeting quality algorithms (IRQA) usually rely on creating a pixel correspondence mapping that indicates at each spatial location in the reference image how the content is preserved in the retargeted one. The image quality is then computed by applying some similarity criterion or distance measure (Liu et al., 2015) with respect to the content matching and maybe relevance. In this context, the usage of local descriptors to build this content matching have been successfully employed for this task because it becomes unnecessary steps of pre or post-processing, such as to adopt complex data structures, or even to solve global optimization problems.

Most IRQA that employed saliency models into the quality assessment have opted for the salient object detection task instead of the fixation prediction ones. In literature, a saliency model can be used for two different tasks: salient object detection and fixation prediction. The salient object detection task is considered a foreground-background segmentation problem while the fixation prediction task results is a

* Corresponding author.
E-mail addresses: sauloafoliveira@lia.ufc.br (S.A.F. Oliveira), shara.alves@ppget.ifce.edu.br (S.S.A. Alves), joaopaulo@lia.ufc.br (J.P.P. Gomes), ajalmar@ifce.edu.br (A.R. Rocha Neto).

https://doi.org/10.1016/j.cviu.2017.11.011
Received 31 January 2017; Received in revised form 4 November 2017; Accepted 25 November 2017
Available online 05 December 2017



(a) Input image. (b) Cropping. (c) Scaling. (d) Shift-map.

Fig. 1. Image retargeting example. The input image (a) and retargeted results from manual Cropping (b), Scaling (c) and Shift-map Pritch et al. (2009). As one can see, in (b) we have content discarding via cropping, while in (c) the scaling applied a uniformly distortion in each pixel. Finally, in (d) the Shift-map operator achieved the best result.

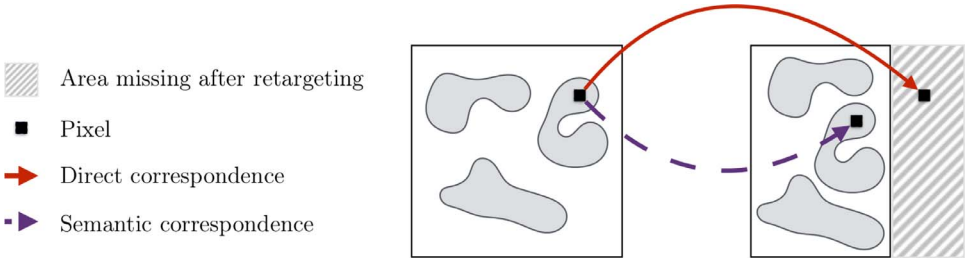


Fig. 2. Content matching in image retargeting. As in retargeting results the spatial location of pixels is not always preserved, a semantic correspondence must be employed so that we can compare the image contents.

(a) Input image. (b) Object detection. (c) Fixation prediction.

Fig. 3. An example of salient object detection (b) and fixation prediction (c) applications. The clearer (yellower) the region, the more relevant it is.

sparse blob-like salient regions map, see the clearest (yellowish) areas in Fig. 3. Although both types of saliency models are expected to be applicable interchangeably, their generated saliency maps actually demonstrate remarkably different characteristics due to their distinct purposes in saliency detection (Borji et al., 2015).

Since human beings are the ultimate consumers of retargeted images and therefore the image quality judges, it is necessary that the IRQA are related to the subjective evaluation criteria. A study conducted by Rubinstein et al. (2010) found that humans generally agree with each other on the quality of retargeted images, and some retargeting algorithms are consistently more favorable than others. Moreover, it was found that some IRQA are useful in assessing the visual quality of retargeted images. However, their correlations with subjective evaluations are not always consistent (Rubinstein et al., 2010). The Earth Mover's Distance (EMD) (Pele and Werman, 2009) and the Scale Invariant Feature Transform Flow (SIFT-Flow) (Liu et al., 2011a) generally agreed better with users' preferences under the study evaluation criteria. This was noticed through the stronger correlation with the subjective results. Furthermore, another very interesting study with respect to evaluating subjectively the quality of retargeted images is presented in Ma et al. (2012). In such study, it was observed that the human subjects are very sensitive to the distortion of the faces as well as the geometric structures, while they can tolerate more distortions on the natural scenery, especially on the texture regions. Although the present study has provided an insight on how to design an effective objective quality metric for evaluating retargeted images, the performances of the assessed metrics were not good enough. This was confirmed by analyzing some statistical correlation between the subjective scores and the algorithm outputs which were quite low. Recently, a study conducted in Ma et al. (2015) discussed how to design an effective IRQA considering the reference image content, retargeting scale, the shape distortion and content information loss, the HVS properties, among other descriptors. The authors of such study also highlighted that the assessed IRQA performances are still unsatisfactory. The statistical correlations between the subjective values and the IRQA outputs were not close, indicating that there is still room for improvements. Thus, an objective IRQA that outputs an approximation of the subjective evaluation is highly desirable.

This work tackles the problem of image retargeting quality assessment by proposing an IRQA based on a bi-directional approach in a fusion framework. The key step in our proposal is to extract features from the retargeting context and combine them through a fusion strategy so that we can predict the quality in the sense of the users' perceptions. The bi-directional approach is the manner we found to take into account the loss of relevant content, as well as, the introduction of visual artifacts in retargeting results. For that, we propose a set of four features, namely, two similarity scores with respect to importance maps (in a bi-directional way), the retargeting ratio, and the content matching information. It is noticeable that there are some previous works that have some of these features (e.g., feature extraction and metric fusion), however, employing them in a bi-directional assessment into the quality one, to the best of our knowledge, was never addressed before. Thus, the main novelty of this work is the proposal of a competitive IRQA which takes into account the loss of relevant content, as well as, the introduction of visual artifacts through a bi-directional image retargeting quality prediction paradigm. The remainder of this paper is organized as follows. In Section 2, we present our proposal in details. After that, we describe the carried out experiments in Section 3. In Section 4, we present some related work. Finally, in Section 5 we present some concluding remarks.

2. Proposal: The bi-directional importance map similarity

As stated before, our proposal, the bi-directional importance map

similarity (BIMS) relies on extracted features from the retargeting process alongside a bi-directional quality evaluation in a fusion framework. For sake of simplicity, consider the following description and notation to describe BIMS:

$$\mathrm{BIMS}(\boldsymbol{I}, \boldsymbol{J}) = \mathscr{F}(\alpha_{\boldsymbol{I}}, \alpha_{\boldsymbol{J}}, \rho, \beta), \tag{1}$$

where $\alpha_{\boldsymbol{I}}$ and $\alpha_{\boldsymbol{J}}$ are the bi-directional quality scores, ρ is the retargeting ration score, β is the keypoint matching score, and $\mathscr{F}(.,.,.,.)$ is the fusion function. The first $\alpha_{\boldsymbol{I}}$ and $\alpha_{\boldsymbol{J}}$ are related to the measurements of the loss of relevant content and the introduction of visual artifacts, respectively. The ρ feature is a simple measure which translates how much of the content was retargeted and the β feature is for measuring the keypoint matching between such images alongside the cardinality of keypoint sets. To better distinguish and appreciate our proposal, we present how each feature is extracted with its contribution to the final quality score and the fusion framework in details. We present them from this point on.

2.1. The bi-directional quality features: $\alpha_{\boldsymbol{I}}$ and $\alpha_{\boldsymbol{J}}$

The bi-directional quality evaluation works by assessing importance maps from each image, the original and the retargeted one, in terms of saliency map similarity. The bi-directional aspect of our proposal is that both image settings are analyzed to obtain quality scores, i.e., we use saliency similarity metrics to measure the loss of relevant content (from the original to the retargeted image) and the introduction of visual artifacts (from the retargeted to the original image).

For that purpose, we define four importance maps based on the original $\boldsymbol{I}$ and the retargeted image $\boldsymbol{I}$, namely:

- the original image importance map of $\boldsymbol{I}$, denoted by $\boldsymbol{S_I}$;
- the retargeted image importance map of $\boldsymbol{J}$, denoted by $\boldsymbol{S_J}$;
- the lossy importance map of $\boldsymbol{I}$, denoted by $\boldsymbol{S_I^M}$;
- the artifact introduction importance map of $\boldsymbol{J}$, denoted by $\boldsymbol{S_J^M}$.

All importance maps are obtained through the Salient Point Parzen Map (SPPM, Oliveira et al. (2016)) algorithm using the Speeded Up Robust Features (SURF, Bay et al. (2008)) as the keypoint detector and feature extractor, but with SPPM using distinct settings, see Table 1. The motivation behind adopting SPPM is because it is a single-parameter model for saliency map generation and despite its simplicity, achieved competitive results when compared to other existing methods (Oliveira et al., 2016). Furthermore, as we later compute a feature based on keypoint matching, β, we take advantage of such already computed keypoints from this previous step for the later feature extraction.

Let $\boldsymbol{P_I}$ and $\boldsymbol{P_J}$ be the set of detected keypoints by SURF from $\boldsymbol{I}$ and $\boldsymbol{J}$, respectively. Also, let $\boldsymbol{M}$ be the matching set containing the keypoints whose features have been matched between $\boldsymbol{P_I}$ and $\boldsymbol{P_J}$, i.e., $\boldsymbol{M} = \boldsymbol{P_I} \cap \boldsymbol{P_J}$ (in the feature space). In possession of such sets, we then derive four importance maps from SPPM (see Fig. 4), namely, $\boldsymbol{S_I}$, $\boldsymbol{S_J}$, $\boldsymbol{S_I^M}$, and $\boldsymbol{S_J^M}$, for latter similarity assessment through saliency quality metrics. Each map is a combination of keypoint origin space and the feature space, see Table 1.

Table 1
Importance map derivation through SPPM and its settings regarding the keypoint point origin and feature space.

Map	Size	Keypoint origin	Feature space	Keypoint set
$\boldsymbol{S_I}$	$m \times n$	Detected keypoints in $\boldsymbol{I}$.	$\boldsymbol{I}$	$\boldsymbol{P_I}$
$\boldsymbol{S_J}$	$m' \times n'$	Detected keypoints in $\boldsymbol{J}$.	$\boldsymbol{J}$	$\boldsymbol{P_J}$
$\boldsymbol{S_I^M}$	$m \times n$	Matched keypoints between $\boldsymbol{I}$ and $\boldsymbol{J}$.	$\boldsymbol{I}$	$\boldsymbol{M}$
$\boldsymbol{S_I^M}$	$m' \times n'$	Matched keypoints between $\boldsymbol{I}$ and $\boldsymbol{J}$.	$\boldsymbol{J}$	$\boldsymbol{M}$

The first map, $\boldsymbol{S_I}$ is the one that captures the original image state in a importance map. The second map, $\boldsymbol{S_J}$ is the one that captures the current (retargeted) importance state. The third and fourth maps, namely, $\boldsymbol{S_I^M}$ and $\boldsymbol{S_J^M}$ are special maps in which the feature space is related to $\boldsymbol{I}$ and $\boldsymbol{J}$, respectively. By computing the similarity score $\alpha_{\boldsymbol{I}}$ between $\boldsymbol{S_I}$ and $\boldsymbol{S_I^M}$ we are able to capture the content preservation in terms of keypoint matching. The greater the correspondence, the greater the similarity since this will indicate the level of keypoint preservation. The second similarity score $\alpha_{\boldsymbol{J}}$ is obtained between $\boldsymbol{S_J}$ and $\boldsymbol{S_J^M}$. From this point of view, the level of keypoint correspondence will indicate how new regions with strong features have arisen after the retargeting.

In possession our four importance maps, $\boldsymbol{S_I}$, $\boldsymbol{S_J}$, $\boldsymbol{S_I^M}$, and $\boldsymbol{S_J^M}$, we now compute, in our context, the so-called bi-directional similarity through some saliency quality metrics. Let $\Psi(.,.)$ be a similarity score function for saliency maps, the similarity scores, $\alpha_{\boldsymbol{I}}$ and $\alpha_{\boldsymbol{I}}$, are given by:

$$\alpha_{\boldsymbol{I}} = \Psi(\boldsymbol{S_I}, \boldsymbol{S_I^M}). \tag{2}$$

$$\alpha_{\boldsymbol{J}} = \Psi(\boldsymbol{S_J}, \boldsymbol{S_J^M}). \tag{3}$$

In the following we present some similarity score functions for saliency maps $\Psi(.,.)$ that can be employed in our proposal.

2.2. Saliency map similarity

To compute the similarity between saliency maps, a simple technique such as the Mean Absolute Error (MAE) computed on the normalized by range importance maps could be employed as suggested by Borji et al. (2015). As in retargeting applications, unlike in other applications, the relative saliency values of different image regions are important (Bylinskii et al., 2016). Therefore, metrics such as the Pearson's Correlation Coefficient (CC) (Borji et al., 2013) and the Similarity (sometimes called histogram intersection) (SIM) (Li et al., 2015) are more suitable for this purpose. The CC metric is symmetric and penalizes false positives and negatives equally. High positive pixel-wise CC values occur where both maps have values of similar magnitudes at the same locations (Bylinskii et al., 2016). The CC metric is defined as follows

$$CC(\boldsymbol{I}, \boldsymbol{J}) = \frac{\sigma(\boldsymbol{I}, \boldsymbol{J})}{\sigma(\boldsymbol{I})\,\sigma(\boldsymbol{J})}, \tag{4}$$

where $\sigma(.,.)$ stands for the covariance of two images, and $\sigma(.)$ is the standard deviation of a given image (as a single row vector). Conversely, the SIM is a very sensitive metric to missing values, and penalizes maps that fail to account for all of the ground truth density (Bylinskii et al., 2016). The SIM metric is defined as

$$SIM(\boldsymbol{I}, \boldsymbol{J}) = \sum_{x=1}^{m}\sum_{y=1}^{n} \min(\overline{\boldsymbol{I}}(x,y), \overline{\boldsymbol{J}}(x,y)), \tag{5}$$

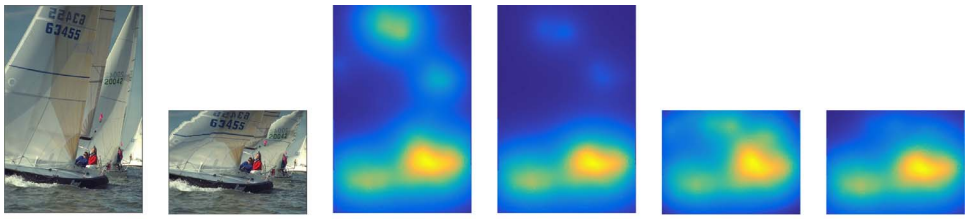
where the images $\overline{\boldsymbol{I}}$, $\overline{\boldsymbol{J}}$ are normalized versions of $\boldsymbol{I}$, and $\boldsymbol{J}$, respectively, in such a way that $\sum_{x=1}^{m}\sum_{y=1}^{n}\overline{\boldsymbol{I}}(x,y) = 1$ and $\sum_{x=1}^{m}\sum_{y=1}^{n}\overline{\boldsymbol{J}}(x,y) = 1$ (as probability distributions). Additionally, we also adopted the Complement of the Mean Absolute Error (CMAE) as similarity criterion

$$CMAE(\boldsymbol{I}, \boldsymbol{J}) = 1 - MAE(\boldsymbol{I}, \boldsymbol{J}) = 1 - \frac{1}{m \times n}\sum_{x=1}^{m}\sum_{y=1}^{n}\left|\widetilde{\boldsymbol{I}}(x,y) - \widetilde{\boldsymbol{J}}(x,y)\right|, \tag{6}$$

where $|.|$ stands for the absolute operator, and the images $\widetilde{\boldsymbol{I}}, \widetilde{\boldsymbol{J}}$ are normalized in range in such a way that $0 \le \widetilde{\boldsymbol{I}}(x,y), \widetilde{\boldsymbol{J}}(x,y) \le 1,\ \forall\ (x,y) \in \boldsymbol{I}, \boldsymbol{J}$.

2.3. The retargeting ratio feature: ρ

The retargeting ratio information is regarding to the retargeted



(a) $\boldsymbol{I}$. (b) $\boldsymbol{J}$. (c) $\boldsymbol{S_I}$. (d) $\boldsymbol{S_I^M}$. (e) $\boldsymbol{S_J}$. (f) $\boldsymbol{S_J^M}$.

Fig. 4. Input images (a)–(b) and their importance maps (c)–(f) from SPPM.

Table 2
Relationship between the keypoint detection and the matching for BIMS.

Relation	Case	Description
$\lvert\boldsymbol{P_I}\rvert = \lvert\boldsymbol{P_J}\rvert = \lvert\boldsymbol{M}\rvert$	Ideal.	All keypoints were matched and indicating that occurred a full match, i.e., the SPPM importance maps are similar.
$\lvert\boldsymbol{P_I}\rvert > \lvert\boldsymbol{P_J}\rvert \approx \lvert\boldsymbol{M}\rvert$	Tolerable.	Missing keypoints in $\boldsymbol{I}$ implies that some interest regions were distorted, but no new serious artifacts were introduced.
$\lvert\boldsymbol{P_J}\rvert > \lvert\boldsymbol{P_I}\rvert \approx \lvert\boldsymbol{M}\rvert$	Almost tolerable.	New unmatched keypoints in $\boldsymbol{J}$ represent possible new artifacts caused by serious distortions.
$\lvert\boldsymbol{P_I}\rvert, \lvert\boldsymbol{P_J}\rvert > \lvert\boldsymbol{M}\rvert$	Worst.	Combination from the two aforementioned acceptable cases, i.e., some regions were distorted and new artifacts were introduced in $\boldsymbol{J}$.

image size ($m' \times n'$) with respect to the original size ($m \times n$). It is a simple measure which translates how much of the content was retargeted and is given by

$$\rho = \frac{\min(m', m)\min(n', n)}{\max(m', m)\max(n', n)}. \tag{7}$$

The reasoning behind such feature is because is more likely to have content loss and emergence of visual artifacts when the retargeting process occurs exceedingly, either for reduction or enlargement. This feature can be interpreted as the probability of producing an sloppy result, which later will be weighted in the fusion framework.

2.4. The keypoint matching feature: β

In an ideal retargeting scenario, in which no distortion happened, each keypoint in $\boldsymbol{P_I}$ has a match in $\boldsymbol{P_J}$, i.e., $\boldsymbol{P_I} = \boldsymbol{P_J} = \boldsymbol{M}$. However, this important situation is not always achieved because new keypoints are generated (or lost) during the retargeting process. This condition must be taken into account and a penalization must be employed since not taking into account such unmatched points would be inappropriate to a suitable quality assessment. Thus, we introduce a way of penalizing this misbehavior by analyzing the keypoint matching alongside the cardinality of keypoint sets as

$$\beta = \frac{2\,\lvert\boldsymbol{M}\rvert}{\lvert\boldsymbol{P_I}\rvert + \lvert\boldsymbol{P_J}\rvert}. \tag{8}$$

It is presented in Table 2 the relationship between the detected keypoint and the level of correspondence.

2.5. Fusion strategy

The fusion strategy for IQA and IRQA is not new in the field. As more than one score is yield, it is necessary a successful strategy to combine these scores into a single one. There are several oft-used fusion models such as, simple averaging, product, linear addition, and non-linear addition for this task. The benefits of adopting machine learning in image quality evaluation was successfully demonstrated in some works, such as Liu et al. (2015), Ma et al. (2016) and Zhang and Kuo (2014). Since there are subjective scores from which we can learn from, employing fusion models based on machine learning strategies sounds plausible. We adopted a Constrained Linear Least Squares for the linear regression, while for the non-linear regression we adopted the Relevance Vector Machine (RVM) (Tipping, 2001).

2.6. BIMS algorithm

A diagrammatic representation for BIMS is shown in Fig. 5 and the pseudo-code for bi-directional importance map similarity is presented in Algorithm 1 in the Appendix A.

3. Experiments and discussion

We validate BIMS by exploring the level of agreement between the objective scores yield from BIMS and subjective scores from a reference dataset through correlations coefficients and errors. Furthermore, we demonstrate the advantage of combining some similarity criteria and fusion strategies through the same correlations coefficients and errors to obtain the best configuration of BIMS. We describe the dataset and the validation methodology in the following.

3.1. Dataset

In order to assess BIMS, we carried out some simulations by using images from the CHUK dataset provided in Ma et al. (2012). The CHUK dataset has 57 images (with 171 retargeted images) and the results for subjective tests in terms of Mean Opinion Scores (MOS), as well. In such dataset, the viewers provided their personal opinions on the perceptual quality, the level of shape distortion, and the level of content information loss.

3.2. Experiment setup

In our simulations, the h parameter for SPPM used to generate the importance maps is fixed for all experiments and set as $h = 9$. This value was obtained by previous tests by trial and error. We also highlight that the we carried out a slightly modification on the original SPPM algorithm so that it can perform faster by employing a vector quantization method through the watershed (the watershed ridge lines being the boundaries for each group) into the map generation. This modification is called FSPPM and it is described in details in Appendix B.

With respect to the saliency similarity metric and fusion strategies, we adopted the ones described in Subsection 2.2, namely, CC, CMAE, and SIM for similarity criteria, and the fusion strategies: averaging (AVG), multiplication (MULT), linear regression (LIN) and, non-linear regression (NLIN)[1].

[1] For non-linear regression, we adopted the Relevance Vector Machine (RVM) (Tipping, 2001) model with a Gaussian kernel and we also optimized the parameter σ. by carrying out a grid search with σ values ranging from $[2^2, \ldots, 2^{-7}]$

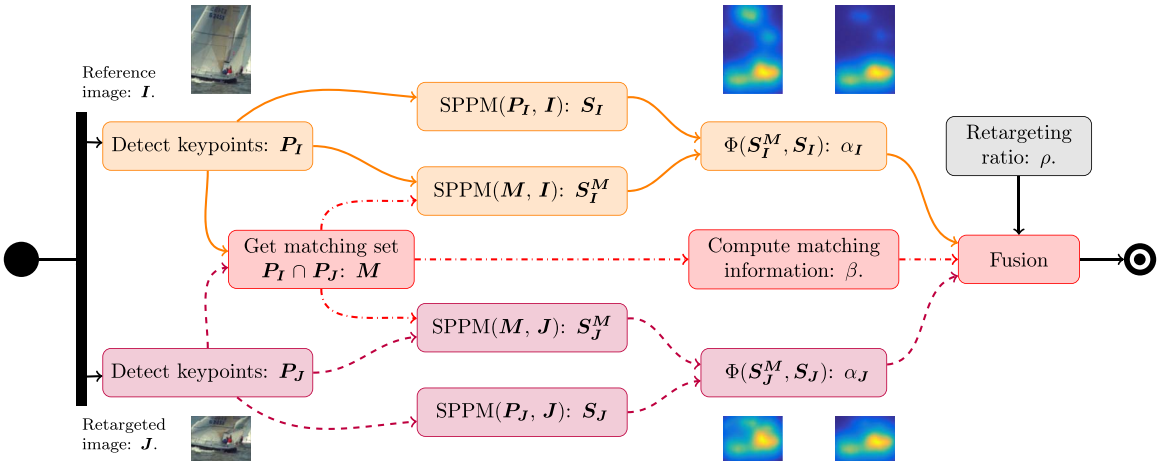


Fig. 5. The bi-directional importance map similarity Flowchart. The lines represent the component flows. The solid line represents the original image, the dotted line is for the matching points and matching information, while the dashed line represents the retargeted image.

As for the training dataset in linear regression and non-linear regression, we carried out 30 executions with an overlapping between training and test set. i.e., we trained with 50% of the MOS and tested on thew whole data. The IRQA that require learning from MOS, such as GLS and NR, vary a lot in terms of training/test methodology. They trained using 90% and 80% and tested in 10% and 20%, respectively. Thus, for comparison purposes, we also included the average of 30 executions of such non-overlapping methodology by using 80% for training and 20% for test.

3.3. Validation

Since there are important subjective components such as image distortion in the evaluation, a subjective test must be employed. This subjective test aims at giving a useful reference point on how an objective method agrees with a typical human subject. So, to compute the level of agreement among the scores, four statistical measurements were employed as suggested by the video quality experts group (VQEG) HDTV (VQEG, 2003). This strategy was accomplished in Ma et al. (2012), Zhang et al. (2016) and Liu et al. (2011b). The four statistical measurements are the linear correlation coefficient (LCC), the Spearman rank-order correlation coefficient (SRCC), the root mean square prediction error (RMSE) and the outlier ratio (OR). In all cases, higher correlation coefficients represent higher agreement, while lower RMSE and OR values indicate smaller errors between the two scores, therefore a better performance. For a proper metric assessment, each BIMS score was mapped according to the following five-parameter logistic function

$$f(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{\exp(\beta_2 (x - \beta_3))} \right) + \beta_4 x + \beta_5, \tag{9}$$

where $\{\beta_i\}_{i=1}^{5}$ are regression model parameters which are determined by minimizing the sum of squared differences between scores.

3.4. Discussion

As for the BIMS validation, we summarize the combinations of similarity criteria and fusion strategies with respect to the provided MOS in terms of agreement in Table 3. In this case, we explore the similarity criteria between maps using CC, CMAE and SIM and the four fusion strategies described in Section 2.4. Furthermore, it is depicted in Fig. 6 maps from SPPM for the Volleyball image and their respective BFMS scores.

Table 3
Combinations of feature origin and fusion strategies for BIMS. AVG stands for averaging, MULT stands for multiplication, LIN is linear regression, and NLIN is for non-linear regression using RVM.

Similarity score	Fusion	LLC	SRCC	RMSE	OR
CMAE	**AVG**	**0.5510**	**0.5426**	**11.6163**	**0.0170**
	MULT	0.5244	0.5033	11.5654	0.0409
	LIN$^{50/100}$	**0.5895**	**0.5632**	**10.9105**	**0.0101**
	NLIN$^{50/100}$	**0.6376**	**0.6213**	**10.3978**	**0.0229**
	LIN$^{80/20}$	0.5413	0.5131	10.9489	0.0203
	NLIN$^{80/20}$	**0.6229**	**0.6167**	**10.1169**	**0.0230**
CC	AVG	0.5278	0.5449	11.6556	0.0351
	MULT	0.5232	0.5110	11.5055	0.0462
	LIN$^{50/100}$	**0.5825**	**0.5751**	**10.9755**	**0.0106**
	NLIN$^{50/100}$	**0.5884**	**0.5851**	**10.9138**	**0.0283**
	LIN$^{80/20}$	**0.6094**	**0.5745**	**10.3726**	**0.0185**
	NLIN$^{80/20}$	**0.5838**	**0.6064**	**10.7355**	**0.0269**
SIM	AVG	0.5419	0.5347	11.5606	0.0175
	MULT	0.5169	0.4985	11.6061	0.0368
	LIN$^{50/100}$	**0.5772**	**0.5476**	**11.0278**	**0.0108**
	NLIN$^{50/100}$	**0.6234**	**0.6060**	**10.5502**	**0.0123**
	LIN$^{80/20}$	0.5491	0.5066	10.9460	0.0183
	NLIN$^{80/20}$	**0.6503**	**0.6283**	**10.0484**	**0.0137**

In Table 4, we summarize the agreement between the values for BIMS scores and provided MOS. We highlight that BIMS achieved competitive results in terms of subjective evaluation by being placed at second place. Although, we have not been able to achieve the best ranking in all criteria, there are some details that must be taken into account in this evaluation. As LLC is used to measure how the our scores are close to the MOS, while the SRCC is employed to indicate the correlation between the score rankings, two conclusions can be drawn by analyzing the Table 4: (1) most images were predicted correctly, and (2) the ranking with respect to the image quality scores in BIMS was preserved more in contrast to the other metrics.

Concerning the LLC, even achieving the third place in BIMS best configuration, we noticed that combinations in Table 3 reinforce that the extracted features are not as representative as they could be. Since the fitting through the supervised fusion strategies can not predict all that correctly, this behavior suggests that the linear relationship between the explanatory variables (α_I, α_J, β, and ρ) and the dependent (MOS) is not a linear function. Thus, as in BIMS, we extract our explanatory variables from SPPM, we also support that further improvement in SPPM keypoint detector may, in fact, improve the performance in BIMS.

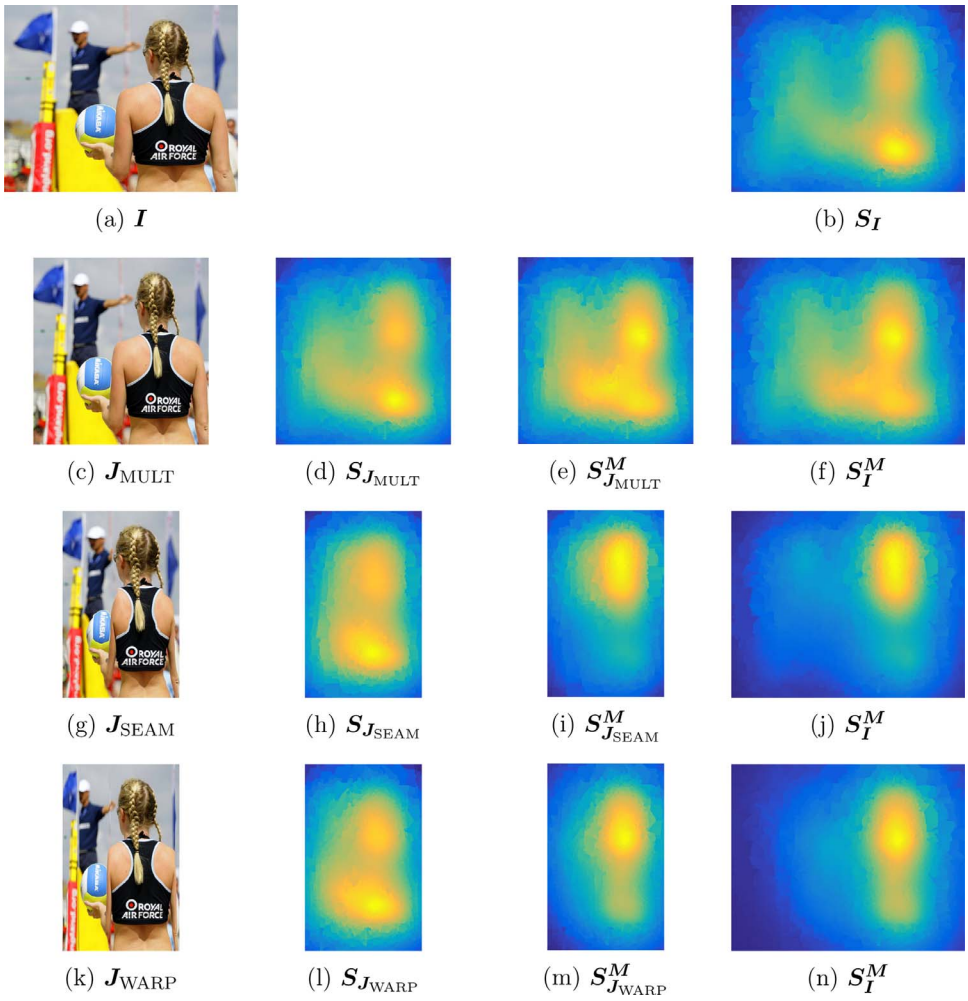


Fig. 6. Maps from FSPPM for Volleyball in BIMS. The input image (a) and retargeted results from Seam Carving (c), Multi-operator (g), and Warping (k), respectively. The BIMS score using SIM and Non-linear and the MOS for (c)-(g)-(k) are, rounded, 69/74, 30/29, and 41/38, respectively.

Table 4
Performance comparison on CUHK retargeting database. The scores were taken directly from Ma et al. (2015), except for PGD, NR, ARS, and FMID which were taken directly from their publications.

IRQA	LLC	SRCC	RMSE	OR
FMID (2017) (Zhang et al., 2017b)	0.7974	0.7984	8.3780	0.0643
ARS (2016) (Zhang et al., 2016)	0.6835	0.6693	9.8550	0.0702
BIMS (SIM - $\mathrm{NLIN}^{80/20}$)	**0.6503**	**0.6283**	**10.0484**	**0.0135**
BIMS (CMAE - $\mathrm{NLIN}^{50/100}$)	**0.6376**	**0.6213**	**10.3978**	**0.0229**
BIMS (CC - $\mathrm{LIN}^{50/100}$)	**0.5884**	**0.5851**	**10.9138**	**0.0283**
BIMS (CC - $\mathrm{LIN}^{80/20}$)	**0.6094**	**0.5745**	**10.3726**	**0.0185**
BIMS (CMAE - AVG)	**0.5510**	**0.5426**	**11.6163**	**0.0170**
PGD (2014) (Hsu et al., 2014)	0.5403	0.5409	11.3610	0.1520
GIST (2015) (Ma et al., 2015; Oliva and Torralba, 2001)	0.5443	0.5114	11.3260	0.1579
NR (2016) (Ma et al., 2016)	0.5371	0.4926	-	0.1928
GLS (2014) (Zhang and Kuo, 2014)	0.4622	0.4760	10.9320	0.1345
CSim (2011) (Liu et al., 2011b)	0.4374	0.4662	12.1410	0.1520
EH (2001) (Messing et al., 2001)	0.3422	0.3288	12.6860	0.2047
EMD (2009) (Pele and Werman, 2009)	0.2760	0.2904	12.9770	0.1696
SIFT-Flow (2011) (Liu et al., 2011a)	0.3141	0.2899	12.8170	0.1462
BDS (2008) (Simakov et al., 2008)	0.2896	0.2887	12.9220	0.2164
PHOW (2007) (Bosch et al., 2007)	0.3706	0.2308	12.5400	0.2222

With respect the SRCC, the combinations between SIM with Averaging and SIM with Non-linear resulted in the best configurations of BIMS. Another worth mentioning behavior is the fact that BFMS, in most cases, have SRCC values higher than LLC values. This behavior shows that the ranking correlation remaining higher, even the prediction values no being as close as the MOS. By analyzing this important behavior in Table 4, we noticed that most metrics, in fact, have focused on guessing the actual MOS value more than capture the ranking essence in terms of agreement[2]. Additionally, it is also worth mentioning that the OR, which measures quantitatively the atypical predictions, were the lowest (20% of the OR value achieved by the second best method) among the State-of-the-art IRQA.

4. Related work

Several image distance-based IRQA have been proposed so far (Kasutani and Yamada, 2001; Liu et al., 2011a; Messing et al., 2001; Pele and Werman, 2009; Rubinstein et al., 2009; Simakov et al., 2008). Computational image distance metrics can predict human retargeting perception and have been widely used as image retargeting quality assessment (Rubinstein et al., 2010). In such algorithms, a cost of transforming the reference image into the retargeting one is represented as a distance (dissimilarity measurement). This cost is most related to the existing visual distortion (or lost content) in the retargeted image. Lower this cost, higher is the similarity between the images. The Edge Histogram (EH) (Messing et al., 2001) and Color Layout

[2] The ranking essence, in this context, means that for a given set of images, the metric actually assigns higher score values for images with higher perceptual quality, while lower values are assigned to the ones with lower perceptual quality.

(CL) (Kasutani and Yamada, 2001) are two low-level metrics in MPEG-7 standard. The EH metric represents the image in terms of edge distribution in the spatial domain and employs the L_1-norm distance to measure the feature distance between two images. The CL metric consists of computing the spatial distribution of colors in images and it measures the quality through a feature distance between two images. The EMD is a dissimilarity metric between two distributions defined as the minimal cost to transform one distribution into the other, while SIFT-Flow (Liu et al., 2011a) is an across-scene image alignment algorithm that employs SIFT (Lowe, 2004) for feature matching and assesses quality through a cost function that analyzes the pixel displacement.

In addition to the distance-based IRQA, another group of algorithms based on user perception have been developed over the past years (Fang et al., 2014; Hsu et al., 2014; Zhang et al., 2016). The foundation of such algorithms relies on assessing similarity based on relevant content preservation alongside a correspondence mapping between pixels. The content preservation can be measured by many factors, such as saliency loss, geometric distortion, structural information, among others. For that, a saliency model or image quality descriptors usually takes place in the process to assess the quality of retargeted images. The motivation behind employing saliency detection models into the IRQA framework is the fact that when observers look at the retargeted results, they will focus on the salient regions while ignoring the non-salient ones, thus, incorporating such saliency information sounds effective for this task. On the other hand, some quality descriptors are able to build an image representation in which information with respect to details, structure, geometry, among others, are highlighted so that the analysis between the ones extracted from the reference image and the one's from the retargeted one can also be employed to assess the quality.

The IR-SSIM (Fang et al., 2014) is a similarity metric based on SSIM and SIFT-Flow that measures the quality of the preserved structural information. In such IRQA, the authors exploited both bottom-up and top-down image saliency models to generate a global saliency map and by weighting a similarity map generated from SSIM and SIFT-Flow with this global saliency map, the image quality is then assessed. Similarly, the metric PGD (Hsu et al., 2014) also employed SIFT-Flow to create a dense correspondence between the reference and the retargeted images, however, the quality score is obtained by analyzing the perceptual geometric distortion and information loss after the retargeting. The perceptual geometric distortion is measured by the local variance from the SIFT-Flow field between the images, while the information loss is measured by evaluating the relationship of warping the saliency map generated from the reference image to the retargeted size. Recently, the Aspect Ratio Similarity (ARS) (Zhang et al., 2016) was proposed by formulating the geometric change estimation as a backward registration problem with Markov random field. Under the guidance of the geometric change, the authors evaluated the quality of retargeted images by the local block changes alongside a visual importance pooling strategy. Also, the NR (Ma et al., 2016), a no-reference (evaluation using only the retargeted image) IRQA was proposed. In NR, each image is represented as a feature vector embodying image characteristics which are employed to discriminate the perceptual quality of the retargeting in a rank learning way.

More recently, special attention has been given to the dense correspondence between the reference and retargeted images. As highlighted in Zhang et al. (2017a), all of the previous works treated the image and the flow from the dense correspondence field separately and, thus, not taken into consideration intrinsic geometric aspects from the images under analysis. For instance, the metric LGI proposed in Zhang et al. (2017a) explores a novel registration method for dense correspondence based on SURF alongside three scores: a local and a global similarity ones, and an information loss one. On a similar fashion, the metric proposed by Zhang et al. (2017b) denoted here by (FMID) arises from the incorporation of both fidelity measures and inconsistency detection into the quality assessment. In such work, the authors report more reliable scores by investigating the fidelity and detecting the inconsistency on three different levels, namely, the region-level, the patch-level, and the pixel-level. The gain with respect to quality in such investigation is mainly referred to the exploration of the correspondence information between reference and retargeted images, in which anomalies concerning shape distortion, foreground cropping, and discontinuity are assessed.

5. Conclusion

In this work, we proposed a novel image retargeting quality algorithm called bi-directional importance map similarity (BIMS). Our proposal is mainly based on computing similarity over importance maps in a bi-directional way, combined with some content correspondence information in a fusion strategy fashion. Since the saliency maps describe where one is looking in images, employing this behavior is useful to estimate the locations we can analyze how the information is preserved after the retargeting. In order to penalize whenever keypoints are missed and/or new ones arisen, we adopted a bi-directional approach to capture this behavior. We explored different combinations of computing similarity over saliency maps and fusion strategies to find a feasible combination in which our proposal can take advantage of it.

The experimental results indicate that BIMS is promising for quality assessment. This is confirmed by the good correlation achieved in terms of subjective scores when compared with the other state-of-the-art algorithms. BIMS had a similar performance in three of the four subjective scores when compared to best metrics reported so far, namely, ARS and FMID, and significantly outperformed them in the OR score. Furthermore, all results suggest that the relationship between the extracted features by our approach and the mean opinion scores are not linear. Such behavior was confirmed by the non-linear configurations achieving better positions.

The simplicity of our proposal is intended to be embedded itself into retargeting algorithms as an attempt to obtain even more satisfactory retargeting results. Nowadays, we are exploring other methods to compute the importance maps, other descriptors to perform the matching between such keypoints, and an automatic way to choose suitable values for h as a mean to improve the current approach.

Acknowledgment

The authors acknowledge the support of CNPq (Grant 456837/2014-0), CAPES and Federal University of Ceará.

Appendix A. The Algorithm for bi-directional importance map similarity

Data: $\boldsymbol{I}$ (reference image with size $m \times n$),
$\boldsymbol{J}$ (retargeted image with size $m' \times n'$),
h (window size/kernel parameter for FSPPM),
Ψ (feature extraction function for FSPPM),
Φ (similarity score function), and
$\mathcal{F}$ (fusion strategy).
Result: quality value.

$\boldsymbol{P_I} \leftarrow \text{DetectKeypoints}(\boldsymbol{I})$.
$\boldsymbol{P_J} \leftarrow \text{DetectKeypoints}(\boldsymbol{J})$.
$\boldsymbol{M} \leftarrow \text{MatchKeyoints}(\boldsymbol{P_I}, \boldsymbol{P_J})$.
/*Importance map generation.*/
$\boldsymbol{S_I} \leftarrow \text{FSPPM}(\boldsymbol{P_I}, \boldsymbol{I}, h, \Psi)$.
$\boldsymbol{S_I^M} \leftarrow \text{FSPPM}(\boldsymbol{M}, \boldsymbol{I}, h, \Psi)$.
$\boldsymbol{S_J} \leftarrow \text{FSPPM}(\boldsymbol{P_J}, \boldsymbol{J}, h, \Psi)$.
$\boldsymbol{S_J^M} \leftarrow \text{FSPPM}(\boldsymbol{M}, \boldsymbol{J}, h, \Psi)$.
/*Similarity between map pairs.*/
$\alpha_{\boldsymbol{I}} \leftarrow \Phi(\boldsymbol{S_I}, \boldsymbol{S_I^M})$.
$\alpha_{\boldsymbol{J}} \leftarrow \Phi(\boldsymbol{S_J}, \boldsymbol{S_J^M})$.
/*Content matching information.*/
$\beta \leftarrow \dfrac{2|\boldsymbol{M}|}{|\boldsymbol{P_I}| + |\boldsymbol{P_J}|}$.
/*Retargeting ratio information.*/
$\rho \leftarrow \dfrac{\min(m, m')\ \min(n, n')}{\max(m, m')\ \max(n, n')}$;
/*Feature fusion.*/
return $\mathcal{F}(\alpha_{\boldsymbol{I}}, \alpha_{\boldsymbol{J}}, \beta, \rho)$.

Algorithm 1. The Algorithm for bi-directional importance map similarity.

Appendix B. The Algorithms for SPPM and FSPPM

We highlight we did not adopted any inner strategy for SPPM.

Data: $\boldsymbol{P}$ (the keypoint set),
$\boldsymbol{I}$ (the input image of size $m \times n$),
h (window size/kernel parameter), and
Ψ (feature extraction function).
Result: the importance map $\boldsymbol{S}$.

/*Initialize the importance map.*/
$\boldsymbol{S} \leftarrow \text{Zeros}(m, n)$.
/* Compute the descriptors for all keypoints $\in \boldsymbol{P}$*/
$\boldsymbol{Z} \leftarrow \Psi(\boldsymbol{P})$.
/*Compute importance map values.*/
for ***all*** *pixel* $(x, y) \in \boldsymbol{I}$ **do**
  $z \leftarrow \Psi(\boldsymbol{I}(x, y))$.
  $\boldsymbol{S}(x, y) \leftarrow \text{DensityEstimation}(z, \boldsymbol{Z}, h)$.
end
/*Normalize the result instead of using probability values.*/
$\boldsymbol{S} \leftarrow \dfrac{\boldsymbol{S} - \min(\boldsymbol{S})}{\max(\boldsymbol{S}) - \min(\boldsymbol{S})}$.
return $\boldsymbol{S}$.

Algorithm 2. Salient Point Parzen Map (SPPM).

Algorithm 3. The Algorithm for Fast Salient Point Parzen Map (FSPPM).

Data: $\boldsymbol{P}$ (the keypoint set),
$\boldsymbol{I}$ (the input image of size $m \times n$),
h (window size/kernel parameter), and
Ψ (feature extraction function).
Result: the importance map $\boldsymbol{S}$.

/*Quantization step.*/
$\boldsymbol{Q} \leftarrow$ DetectPrototypes($\boldsymbol{I}$).
/*Initialize he saliency map.*/
$\boldsymbol{S} \leftarrow$ Zeros(m, n).
for ***each*** *coordinate* (x_i, y_i) *of each prototype* $\boldsymbol{q_i} \in \boldsymbol{Q}$ **do**
  $z \leftarrow \Psi(\boldsymbol{I}(x_i, y_i))$.
  $p \leftarrow$ DensityEstimation($z, \boldsymbol{Z}, h$).
  /*Assigns the probability value of a prototype to all associated pixels.*/
  for ***each*** *pixel* (x_j, y_j) *belonging to the cluster associated to* $\boldsymbol{q_i}$ **do**
    $\boldsymbol{S}(x_j, y_j) \leftarrow p$.
  end
end
/*Normalize the result instead of using probability values.*/
$\boldsymbol{S} \leftarrow \dfrac{\boldsymbol{S} - \min(\boldsymbol{S})}{\max(\boldsymbol{S}) - \min(\boldsymbol{S})}$.
return $\boldsymbol{S}$.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.cviu.2017.11.011

References

Bay, H., Ess, A., Tuytelaars, T., Gool, L.V., 2008. Speeded-up robust features (SURF). Comput. Vision Image Understanding 110 (3), 346–359. http://dx.doi.org/10.1016/j.cviu.2007.09.014.

Borji, A., Cheng, M.-M., Jiang, H., Li, J., 2015. Salient object detection: a benchmark. IEEE Trans. Image Process. 24 (12), 5706–5722. http://dx.doi.org/10.1109/TIP.2015.2487833.

Borji, A., Sihite, D.N., Itti, L., 2013. Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study. IEEE Trans. Image Process. 22 (1), 55–69. http://dx.doi.org/10.1109/TIP.2012.2210727.

Bosch, A., Zisserman, A., Munoz, X., 2007. Image classification using random forests and ferns. 2007 IEEE 11th International Conference on Computer Vision, VOLS 1–6, IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, October 14–21, 2007. IEEE, 11th IEEE International Conference on Computer Vision, pp. 1863–1870.

Bylinskii, Z., Judd, T., Oliva, A., Torralba, A., Durand, F., 2016. What do different evaluation metrics tell us about saliency models? arXiv:1604.03605.

Fang, Y., Zeng, K., Wang, Z., Lin, W., Fang, Z., Lin, C.-W., 2014. Objective quality assessment for image retargeting based on structural similarity. IEEE J. Emerg. Sel. Top. Circuits Syst. 4 (1), 95–105. http://dx.doi.org/10.1109/JETCAS.2014.2298919.

Guttmann, M., Wolf, L., Cohen-Or, D., 2011. Content aware video manipulation. Comput. Vision Image Understanding 115 (12), 1662–1678. http://dx.doi.org/10.1016/j.cviu.2011.05.010.

Hsu, C.-C., Lin, C.-W., Fang, Y., Lin, W., 2014. Objective quality assessment for image retargeting based on perceptual geometric distortion and information loss. IEEE J Sel Top Signal Process 8 (3), 377–389. http://dx.doi.org/10.1109/JSTSP.2014.2311884.

Kasutani, E., Yamada, A., 2001. The MPEG-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval. 2001 International Conference on Image Processing, Vol I, Thessaloniki, Greece, October 07–10, 2001. IEEE Signal Processing Soc; IEEE, IEEE International Conference on Image Processing (ICIP), pp. 674–677.

Li, J., Xia, C., Song, Y., Fang, S., Chen, X., 2015. A data-driven metric for comprehensive evaluation of saliency models. 2015 IEEE International Conference on Computer Vision (ICCV). IEEE International Conference on Computer Vision, pp. 190–198. http://dx.doi.org/10.1109/ICCV.2015.30.

Liu, A., Lin, W., Chen, H., Zhang, P., 2015. Image retargeting quality assessment based on support vector regression. Signal Process.-Image Commun. 39 (B, SI), 444–456. http://dx.doi.org/10.1016/j.image.2015.08.001.

Liu, C., Yuen, J., Torralba, A., 2011. Sift flow: dense correspondence across scenes and its applications, pattern analysis and machine intelligence. IEEE Trans. 33 (5), 978–994.

Liu, Y.-J., Luo, X., Xuan, Y.-M., Chen, W.-F., Fu, X.-L., 2011. Image retargeting quality assessment. Comput. Graphics Forum 30 (2), 583–592. http://dx.doi.org/10.1111/j.1467-8659.2011.01881.x.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60 (2), 91–110.

Ma, L., Deng, C., Lin, W., Ngan, K.N., Xu, L., 2015. Retargeted image quality assessment: Current progresses and future trends. Visual Signal Quality Assessment. Springer, pp. 213–242.

Ma, L., Lin, W., Deng, C., Ngan, K.N., 2012. Image retargeting quality assessment: a study of subjective scores and objective metrics. IEEE J Sel Top Signal Process 6 (6), 626–639. http://dx.doi.org/10.1109/JSTSP.2012.2211996.

Ma, L., Xu, L., Zhang, Y., Yan, Y., Ngan, K.N., 2016. No-reference retargeted image quality assessment based on pairwise rank learning. IEEE Trans. Multimed. 18 (11), 2228–2237. http://dx.doi.org/10.1109/TMM.2016.2614187.

Messing, D., van Beek, P., Errico, J., 2001. The MPEG-7 colour structure descriptor: Image description using colour and local spatial information. 2001 International Conference On Image Processing, Vol I, Proceedings, IEEE International Conference on Image Processing ICIP, Thessaloniki, Greece, October 07–10, 2001. IEEE Signal Processing Soc; IEEE, International Conference on Image Processing (ICIP 2001), pp. 670–673.

Oliva, A., Torralba, A., 2001. Modeling the shape of the scene: a holistic representation of the spatial envelope. Int. J. Comput. Vis. 42 (3), 145–175. http://dx.doi.org/10.1023/A:1011139631724.

Oliveira, S.A.F., Neto, A.R.R., Gomes, J.P.P., 2016. Towards fixation prediction: A non-parametric estimation-based approach through key-points. Proceedings of 2016 5th Brazilian Conference on Intelligent Systems (BRACIS 2016). pp. 391–396. http://dx.doi.org/10.1109/BRACIS.2016.75.

Pele, O., Werman, M., 2009. Fast and robust earth mover's distances. 2009 IEEE 12th International Conference on Computer Vision (ICCV), IEEE International Conference on Computer Vision, Kyoto, Japan, September 29-OCT 02, 2009. IEEE; IEEE Comp Soc, 12th IEEE International Conference on Computer Vision, pp. 460–467. http://dx.doi.org/10.1109/ICCV.2009.5459199.

Pritch, Y., Kav-Venaki, E., Peleg, S., 2009. Shift-map image editing. 2009 IEEE 12th International Conference On Computer Vision (ICCV), IEEE International Conference on Computer Vision, Kyoto, Japan, September 29–October 02, 2009. IEEE; IEEE Comp Soc, 12th IEEE International Conference on Computer Vision, pp. 151–158. http://dx.doi.org/10.1109/ICCV.2009.5459159.

Rubinstein, M., Gutierrez, D., Sorkine, O., Shamir, A., 2010. A comparative study of image retargeting. ACM Trans. Graph. (Proc. SIGGRAPH Asia) 29 (6), 160:1–160:10.

Rubinstein, M., Shamir, A., Avidan, S., 2009. Multi-operator media retargeting. ACM Transactions on Graphics (TOG). Vol. 28. ACM, pp. 23.

Shamir, A., Sorkine, O., 2009. Visual media retargeting. ACM SIGGRAPH ASIA 2009 Courses, SIGGRAPH ASIA '09. ACM, New York, NY, USA, pp. 11:1–11:13. http://doi.acm.org/10.1145/1665817.1665828.

Sheikh, H., Bovik, A., 2006. Image information and visual quality. IEEE Trans. Image Process. 15 (2), 430–444. doi:10.1109/TIP.2005.859378.

Simakov, D., Caspi, Y., Shechtman, E., Irani, M., 2008. Summarizing visual data using bidirectional similarity. Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, IEEE, pp. 1–8.

Tipping, M., 2001. Sparse bayesian learning and the relevance vector machine. J. Mach. Learn. Res 1 (3), 211–244. http://dx.doi.org/10.1162/15324430152748236.

VQEG, 2003. Final report from the video quality experts group on the validation of

objective models of video quality assessment. Phase II (FR_TV2). ftp://ftp.its.bldrdoc.gov/dist/ituvidq/Boulder_VQEG_jan_04/VQEG_PhaseII_FRTV_Final_Report_SG9060E.doc.

Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E., 2004. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. 13 (4), 600–612. doi:10.1109/TIP.2003.819861.

Zhang, B., Sander, P.V., Bermak, A., 2017. Registration based retargeted image quality assessment. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1258–1262. http://dx.doi.org/10.1109/ICASSP.2017.7952358.

Zhang, J., Kuo, C.-C.J., 2014. An objective quality of experience (qoe) assessment index for retargeted images. Proceedings of the 22Nd ACM International Conference on Multimedia, MM '14. ACM, New York, NY, USA, pp. 257–266. http://dx.doi.org/10.1145/2647868.2654922.

Zhang, Y., Fang, Y., Lin, W., Zhang, X., Li, L., 2016. Backward registration-based aspect ratio similarity for image retargeting quality assessment. IEEE Trans. Image Process. 25 (9), 4286–4297. http://dx.doi.org/10.1109/TIP.2016.2585884.

Zhang, Y., Ngan, K.N., Ma, L., Li, H., 2017. Objective quality assessment of image retargeting by incorporating fidelity measures and inconsistency detection. IEEE Trans. Image Process. 26 (12), 5980–5993. http://dx.doi.org/10.1109/TIP.2017.2746260.