Computer Vision and Image Understanding xxx (xxxx) xxx–xxx

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
1077-3142/

(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.

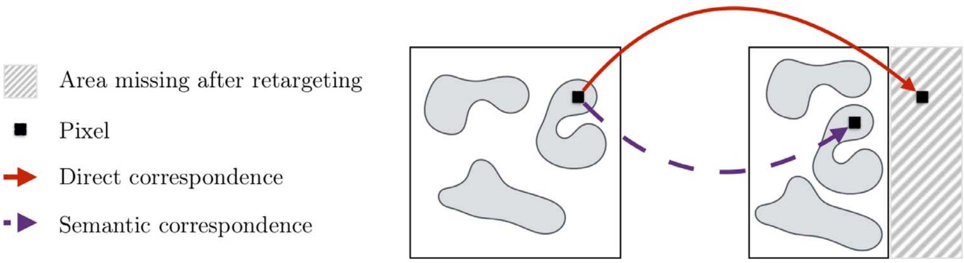

Fig. 2. Content matching in image retargeting. As in retargeting results the spatial location of pixels is not always preserved, a semantic correspondence must be employed so that we can compare the image contents.

(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