

# A New Similarity Measure with Deformation Detection of Visual Salient Regions for Image Retargeting

Canlin Li, Fubao Zhu and Rijian Su

*School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450000, China*  
*lcl\_zju@aliyun.com*

## **Abstract**

*Measurement of image similarity is a fundamental issue in both image processing and computer vision and is very important for image retargeting applications, which have drawn much attention in recent years. In this paper, we propose a new similarity measure with deformation detection of visual salient regions for image retargeting. According to the fact that the human visual system is sensitive to content changes as well as visual effect variations, the proposed approach combines content similarity and visual effect similarity. Firstly, a content correspondence between original and retargeted images is established, which aims to assess their content similarity. Secondly, deformation detection of visual salient regions between two images is conducted to measure their visual effect similarity. Real images have been used to test the proposed method and very good results have been achieved, validating it.*

**Keywords:** *Similarity Measure; Image Retargeting; Content Similarity; Visual Effect Similarity; Deformation Detection*

## **1. Introduction**

Image retargeting is a technique that adjusts input images into arbitrary sizes and simultaneously preserves the salient regions of the input images. Image retargeting has drawn much attention in image and vision research in recent years [1]. Due to the increase in the variety of commonly used display devices, and the prevalent use of mobile devices as available means for image intake, image needs to be adapted to different resolutions and aspect ratios. This problem further increases with the explosion of image and video content on the web.

Measurement of image distance or similarity is a fundamental issue in both image processing and computer vision [2]. An obvious and accurate way is the subjective measurement based on the human perception. A widely used subjective measure computes mean opinion scores (MOSs) from the human ratings [3, 4]. But this method is time-consuming and not suitable for practical use. Objective measures by computer programs whose evaluations are in close agreement with human judgment have been extensively studied in the past. Early work about objective measures characterized the similarity of two images of same size using peak signal to noise ratio (PSNR) and mean squared errors (MSE) [5]. A structural similarity framework [6], called SSIM, was proposed based on the assumption that human vision system (HVS) is highly adapted for extracting statistic structural information.

Measurement of image similarity is very important for image retargeting applications. Firstly, it can be used to dynamically monitor and adjust image quality. Secondly, it can be used to optimize algorithms and parameter settings of image retargeting systems. Thirdly, it can be used to benchmark image retargeting systems and algorithms. However, many of the

well-established image similarity metrics such as PSNR and structural similarity cannot meet the needs of image retargeting application because they work on images of the same size or aspect ratio only. The similarity measure such as SIFT flow [7] and EMD [8] and intensity-related distance metrics such as BDS [9] and BDW [10] have been applied to image retargeting, but they achieved unsatisfying performance [11] because these similarity metrics mainly focus on local features and show smaller overall correspondence with the users.

According to the postulate [12] in cognitive science that the human visual system is sensitive to content changes and visual effect variations, the similarity measure for image retargeting should take content similarity as well as visual effect similarity into account. Content similarity for image retargeting mainly involves the number and layout of the salient contents in the image, and visual effect similarity mainly focuses on shape preservation of visual salient regions.

In this paper, a new objective similarity measure method for image retargeting is proposed. The method is based on a combination of content similarity and visual effect similarity. At an abstract level, the proposed method can be outlined in two steps. First, a content correspondence between two images is established, which aims to assess the content similarity in original and retargeted images. Secondly, deformation detection of visual salient regions or contents between two images is conducted to measure the visual effect similarity of them. Experimental results show that the objective quality values closely match the subjective scores evaluated by users, indicating that our proposed objective metrics are congruent with human perception mechanism.

The remainder of this paper is organized as follows. Section 2 illustrates the framework of the proposed similarity measure for image retargeting. Section 3 explores measuring content similarity of original and retargeted images, and Section 4 elaborates visual effect similarity. Section 5 focuses on the proposed similarity measure with deformation detection of visual salient regions for image retargeting. Section 6 provides the experimental results. Finally, the paper is concluded in Section 7.

## **2. Framework of the Proposed Similarity Measure**

We illustrate the framework of similarity measure for image retargeting proposed by us in Figure 1.

As we can see from Figure 1, the proposed similarity measure for image retargeting mainly includes two parts, where one is content similarity which consists of the sub-steps of finding seams, dividing image into regions and matching regions, and the other is visual effect similarity which is characterized by deformation detection of visual salient regions between original and retargeted images. Here deformation detection mainly focuses on over-squeezing, over-stretching, distortion of proportion and distortion of geometric structure. The detailed procedure of the approach is as follows.

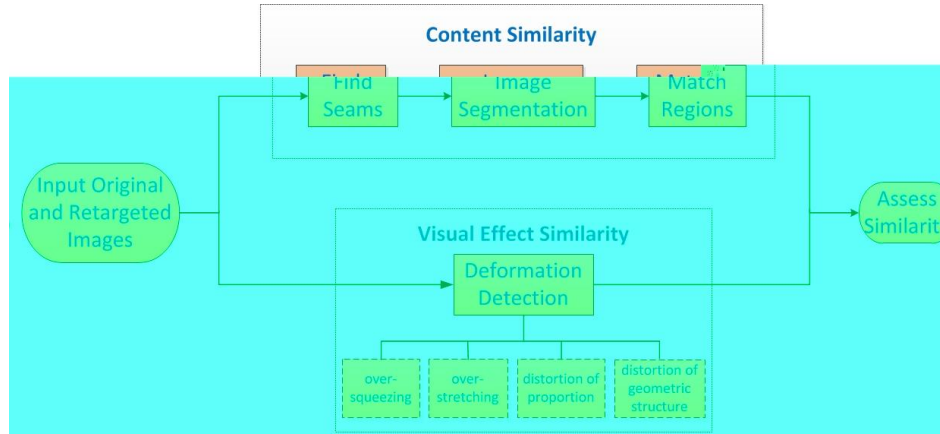


Figure 1. Similarity Measure for Image Retargeting

### 3. Measure of Content Similarity

#### 3.1. Find Seams

The definition of seams in the image is the same as in [13]. That is, a vertical seam is an 8-connected path of pixels in the image from top to bottom, containing one, and only one, pixel in each row of the image. Formally, let  $I$  be an  $n \times m$  image and define a vertical seam to be:

$$s^x = \{s_i^x\}_{i=1}^n = \{(x(i), i)\}_{i=1}^n, \text{ s.t. } \forall i, \forall j, |x(i) - x(i-1)| \leq 1 \quad (1)$$

where  $x$  is a mapping  $x: [1, \dots, n] \rightarrow [1, \dots, m]$ . Similarly, a horizontal seam can be defined. The pixels of the path of seam  $s$  (e.g. vertical seam  $\{s_i\}$ ) will therefore be  $I_s = \{I(s_i)\}_{i=1}^n = \{I(x(i), i)\}_{i=1}^n$ .

For assessing the retargeting image similarity, the images are divided into several regions by virtue of the seams in our approach. Therefore, the question is how to choose the seams in the image? We use the following simple energy function as in [13].

$$e_1(I) = \left| \frac{\partial}{\partial x} I \right| + \left| \frac{\partial}{\partial y} I \right| \quad (2)$$

Given the given energy function, we can define the energy of a seam as  $E(s) = E(I, s) = \sum_{i=1}^n e_1(I(s_i))$ .

Intuitively, our goal is to find noticeable pixels or seams that divide image into regions. That is to say, we should look for the optimal seam  $s^*$  that maximizes this seam energy:

$$s^* = \max_s E(s) = \max_s \sum_{i=1}^n e_1(I(s_i)) \quad (3)$$

The optimal seam can be found using dynamic programming. The first step is to traverse the image from the second row to the last row and compute the cumulative maximum energy  $M$  for all possible connected seams for each entry  $(i, j)$ :

$$M(i, j) = e_1(I(s_i)) + \max(M(i-1, j-1), M(i-1, j), M(i-1, j+1)) \quad (4)$$

At the end of this process, the maximum value of the last row in  $M$  will indicate the end of the maximal connected vertical seam. Hence, in the second step we backtrack from this maximum entry on  $M$  to find the path of the optimal seam. The definition of  $M$  for horizontal seams is similar.

### 3.2. Divide Image into Regions

For every image of the similarity measure, the optimal vertical seam and the optimal horizontal seam can be determined using the above method. Then the image can be segmented into four regions with the optimal vertical seam and the optimal horizontal seam. Moreover, for every region, the optimal vertical seam and the optimal horizontal seam can also be determined using the above method. Every region can further be segmented into four sub-regions with the optimal vertical seam and the optimal horizontal seam in the region as illustrated in Figure 2 and Figure 3, which are for the original image  $I_{or}$  and the retargeted image  $I_{re}$  respectively. The image segmentation is recursive and from coarse to fine and can be carried out until the sub-region has been relatively small. Generally, the image is divided into 16 sub-regions, up to 64 sub-regions.

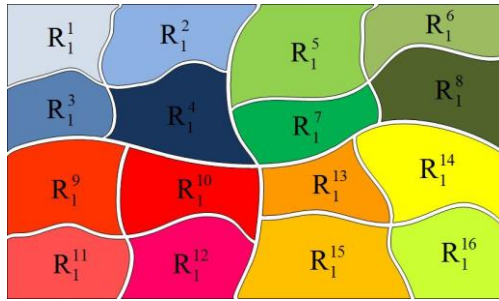


Figure 2. Divided the Original Image  $I_{or}$  into Sixteen Regions

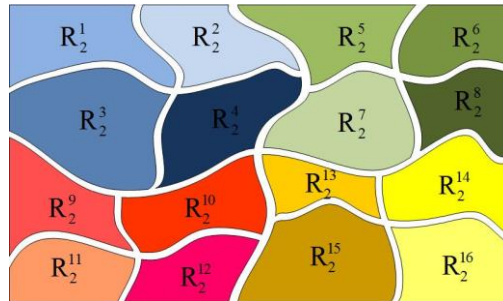


Figure 3. Divided the Retargeted Image  $I_{re}$  into Sixteen Regions

### 3.3. Match Regions

After getting the same amount of regions or sub-regions in the original image  $I_{or}$  and retargeted image  $I_{re}$  by virtue of image segmentation, the region matching is performed between these two images. We define the corresponding regions between these two images as: two regions with the same layout and position in respective image, such as  $R_1^{10}$  and  $R_2^{10}$ . For every region in the image  $I_{or}$ , we define its candidate matched region set  $cs1$  in the image  $I_{re}$  as the set including the corresponding region and its 8-connected neighbors, and its candidate matched region set  $cs2$  in the image  $I_{re}$  as the set including the corresponding region and its 24-connected neighbors. For example, for the region  $R_1^{10}$  in the image  $I_{or}$ , its candidate matched region set  $cs1$  is  $\{R_2^3, R_2^4, R_2^7, R_2^9, R_2^{10}, R_2^{11}, R_2^{12}, R_2^{13}, R_2^{15}\}$ . We adopt the following

region matching strategy. For every region in the image  $I_{or}$ , we look for the nearest neighbor in its candidate matched region set cs1 or cs2 using SIFT flow [7] method.

### 3.4. Calculate Content Similarity

During measuring content similarity of images, for every region  $R_1^k$  in the original image  $I_{or}$ , if its nearest neighbor in the retargeted image  $I_{re}$  is just right the corresponding region  $R_2^k$ , we assign a weight  $w_{cs0}$  to the matched region pair  $CCS(R_1^k, R_2^k)$ . Otherwise, if the nearest neighbor of  $R_1^k$  in the image  $I_{re}$  belongs to cs1- $\{R_2^u\}$ , for example, it is known as  $R_2^u$ , and we will assign a weight  $w_{cs1}$  to the matched region pair  $CCS(R_1^k, R_2^u)$ , and if the nearest neighbor of  $R_1^k$  in the image  $I_{re}$  belongs to cs2-cs1, for instance, known as  $R_2^v$ , we assign a weight  $w_{cs2}$  to the matched region pair  $CCS(R_1^k, R_2^v)$ . Here  $w_{cs2} < w_{cs1} < w_{cs0}$ . Generally, let  $w_{cs0} = 1$ ,  $w_{cs1} = 0.5$ ,  $w_{cs2} = 0.25$ . For measuring content similarity, we accumulate the weight of every matched region pair. Thus, based on the matched region pairs between  $I_{or}$  and  $I_{re}$ , we define content similarity measure for image retargeting as shown in formula (5).

$$CSM_{(I_{or}, I_{re})} = w_{cs0} \sum CCS(R_1^k, R_2^k) + w_{cs1} \sum_{R_2^u \in cs1 - \{R_2^k\}} CCS(R_1^k, R_2^u) + w_{cs2} \sum_{R_2^v \in cs2 - cs1} CCS(R_1^k, R_2^v) \quad (5)$$

We note that similarity measure  $CSM_{(I_{or}, I_{re})}$  for image retargeting places more importance on the exactly matched region pair where two regions have the same layout and position in respective image beyond other matched region pair.

## 4. Measure of Visual Effect Similarity

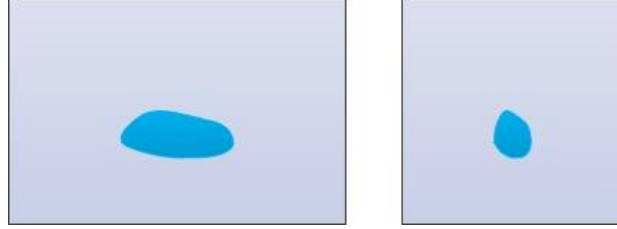
Visual effect similarity mainly involves shape preservation of visual salient regions or contents between original and retargeted images. For measuring visual effect similarity of two images, we have to detect deformation of visual salient regions. We wonder which deformations should be taken into account. According to Rubinstein's work [11], the retargeted image result is rejected and not considered to be similar to the original image owing to five main reasons of deformation: over-squeezing, over-stretching, distortion of proportion, distortion of geometric structure and removal of content. So we should deal with these deformations during evaluating the similarity between original and retargeted images. Actually, in our proposed approach, removal of content has been taken into account at the stage of measuring content similarity of two images. Thus, at this stage of assessing visual effect similarity between original and retargeted images, we mainly cope with the first four deformations. We will detect these deformations of visual salient regions and accordingly determines visual effect similarity values.

### 4.1. Deformation Detection of Visual Salient Regions

This section focuses on elaborating detection of four deformations of visual salient regions between original and retargeted images, which include over-squeezing, over-stretching, distortion of proportion and distortion of geometric structure.

**4.1.1. Detect over-squeezing:** Let  $w_{or}$  and  $h_{or}$  represent width and height of original image respectively, and  $w_{re}$  and  $h_{re}$  be width and height of retargeted image respectively. When the original image is shrunk during the retargeting process, visual salient regions are vulnerable to over-squeezing as illustrated in Figure 4, where the left is original image and the right is retargeted image and the blue block *b1* suffers from over-squeezing. In our proposed

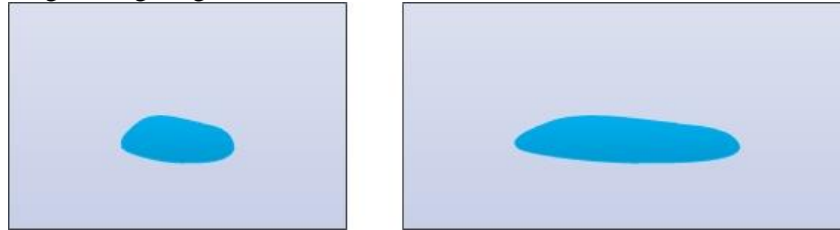
approach, we design and apply formula (6) to determine whether a visual salient region is over-squeezed, where  $w_{or}^{b1}$  and  $h_{or}^{b1}$  mean width and height of the blue block  $b1$  in original image respectively, and  $w_{re}^{b1}$  and  $h_{re}^{b1}$  are width and height of the blue block  $b1$  in retargeted image. If the inequality (6) holds, there is no over-squeezing for this region during retargeting process. Otherwise, over-squeezing of this region is brought into due to image retargeting.



**Figure 4. Detect Over-squeezing**

$$\frac{w_{or}^{b1}/h_{or}^{b1}}{w_{re}^{b1}/h_{re}^{b1}} \leq 0.75 * \frac{W_{or}/H_{or}}{W_{re}/H_{re}}, \text{ (if } w_{or} \geq H_{or} \text{ ) } \text{ or } \frac{h_{or}^{b1}/w_{or}^{b1}}{h_{re}^{b1}/w_{re}^{b1}} \leq 0.75 * \frac{H_{or}/W_{or}}{H_{re}/W_{re}}, \text{ (if } w_{or} < H_{or} \text{ ) } \quad (6)$$

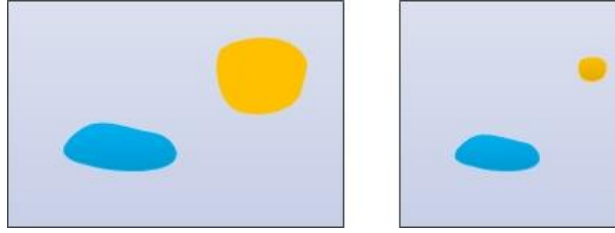
**4.1.2. Detect over-stretching:** When the original image is stretched during the retargeting process, visual salient regions are prone to be over- stretched as illustrated in Figure 5, where the left is original image and the right is retargeted image and the blue block  $b1$  suffers from over-stretching. In our proposed approach, we design and apply formula (7) to judge if a visual salient region is over-stretched. If the inequality (7) holds, there is no over-stretching for this region during retargeting process. Otherwise, over-stretching of this region has been caused by image retargeting.



**Figure 5. Detect Over-stretching**

$$\frac{w_{re}^{b1}/h_{re}^{b1}}{w_{or}^{b1}/h_{or}^{b1}} \leq 0.75 * \frac{W_{re}/H_{re}}{W_{or}/H_{or}}, \text{ (if } w_{or} \geq H_{or} \text{ ) } \text{ or } \frac{h_{re}^{b1}/w_{re}^{b1}}{h_{or}^{b1}/w_{or}^{b1}} \leq 0.75 * \frac{H_{re}/W_{re}}{H_{or}/W_{or}}, \text{ (if } w_{or} < H_{or} \text{ ) } \quad (7)$$

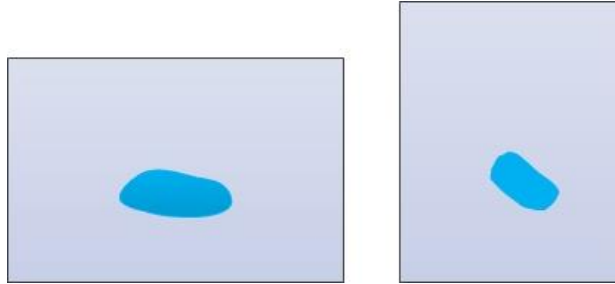
**4.1.3. Detect distortion of proportion:** As shown in Figure 6, there is a distortion of proportion between the blue block  $b1$  and the yellow block  $b2$ , which is derived from retargeting the left original image into the right result image. The yellow block  $b2$  is larger than the blue block  $b1$  in the original image, but notably smaller than  $b1$  in the retargeted image. We can detect the distortion of proportion by virtue of the area of visual salient regions. We design and utilize formula (8) to determine if there is a distortion of proportion between two visual salient regions during retargeting process, where  $s_{or}^{b1}$  and  $s_{or}^{b2}$  represent the areas of the blue block  $b1$  and the yellow block  $b2$  in original image respectively, and  $s_{re}^{b1}$  and  $s_{re}^{b2}$  are the areas of  $b1$  and  $b2$  in retargeted image respectively. If the inequality (8) holds, there is no distortion of proportion between these two regions during retargeting process. Otherwise, distortion of proportion about them is brought into due to image retargeting.



**Figure 6. Detect distortion of proportion**

$$0.75 \leq \frac{S_{re}^{b2}/S_{re}^{b1}}{S_{or}^{b2}/S_{or}^{b1}} \leq 1.25, \text{ (if } s_{or}^{b1} \leq s_{or}^{b2} \text{)} \text{ or } 0.75 \leq \frac{S_{re}^{b1}/S_{re}^{b2}}{S_{or}^{b1}/S_{or}^{b2}} \leq 1.25, \text{ (if } s_{or}^{b1} > s_{or}^{b2} \text{)} \quad (8)$$

**4.1.4. Detect distortion of geometric structure:** Visual effect variation is sensitive to the distortion of geometric structure, so evaluating visual effect similarity has to involve detecting the distortion of visual salient regions. As shown in Figure 7, there is a distortion of the blue block  $b1$  between original and retargeted images. To simplify the process of this deformation detection, we employ the centroid of visual salient region and some special points of region boundary to describe the shape of region, where the detailed steps are included as follows.



**Figure 7. Detect Distortion of Geometric Structure**

(1) We get the centroid of visual salient region from the following formula (9), since the region can be regarded as a homogeneous thin plate with uniform thickness.

$$x_c = \sum A_i * x_i / A, \quad y_c = \sum A_i * y_i / A \quad (9)$$

$A_i$  is the area of the primitive  $(x_i, y_i)$ , and  $A$  is total area of this region. Let  $N$  be the number of points included into this region, formula (9) can be simplified as (10).

$$x_c = \sum x_i / N, \quad y_c = \sum y_i / N \quad (10)$$

(2) We divide the region into four parts by virtue of  $(x_c, y_c)$  and four vectors:  $(x_{max}, y) - (x_c, y_c)$ ,  $(x, y_{max}) - (x_c, y_c)$ ,  $(x_{min}, y) - (x_c, y_c)$  and  $(x, y_{min}) - (x_c, y_c)$  where  $(x_{max}, y)$ ,  $(x, y_{max})$ ,  $(x_{min}, y)$  and  $(x, y_{min})$  represent the points with maximum horizontal coordinate, maximum vertical coordinate, minimum horizontal coordinate and minimum vertical coordinate, respectively. These four parts are numbered as  $p_1$ ,  $p_2$ ,  $p_3$  and  $p_4$ , which are located at the top right, top left, bottom left and bottom right of the region, respectively.

(3) Find the farthest point  $(x, y)_{\rho_{max}}^i$  and the nearest point  $(x, y)_{\rho_{min}}^i$  from the centroid  $(x_c, y_c)$  of the region in every part, where  $i$  is the number of part ( $i=1,2,3,4$ ).

(4) Transform Cartesian coordinates of the vectors  $(x, y)_{\rho_{\max}^i} - (x_c, y_c)$  and  $(x, y)_{\rho_{\min}^i} - (x_c, y_c)$  ( $i=1,2,3,4$ ) into polar coordinates, thus we can get polar radii  $\rho_{\max}^i$ ,  $\rho_{\min}^i$  and polar angles  $\theta_{\max}^i$  and  $\theta_{\min}^i$  respectively.

(5) Sort the polar radii and polar angles. In detail, sort  $\rho_{\max}^i$  in ascending order and record the sequence of corresponding part number as  $s_{\rho_{\max}^i}$ , and sort  $\rho_{\min}^i$  in ascending order and record the sequence of corresponding part number as  $s_{\rho_{\min}^i}$ . Moreover, sort  $\theta_{\max}^i$  and  $\theta_{\min}^i$  in ascending order for every part ( $i=1,2,3,4$ ), and find the polar radius which is exactly corresponding to the larger angle, either  $\rho_{\max}^i$  or  $\rho_{\min}^i$ , and record the label max of  $\rho_{\max}^i$  or the label min of  $\rho_{\min}^i$ . According to ascending order of part number, record the sequence of corresponding label of max or min as  $s_{\max\theta}$ .

(6) Finally, we describe the shape of region with  $s_{\rho_{\max}^i}$ ,  $s_{\rho_{\min}^i}$  and  $s_{\max\theta}$ .

Distortion detection of visual salient region is conducted based on the descriptors of this region in the original and retargeted images. If the visual salient region in original image and the corresponding region in the retargeted image have the same shape descriptor, there is not deformation for this region during the retargeting process. If not, we think the shape distortion of the region has been caused.

#### 4.2. Assess Visual Effect Similarity

We count the visual salient regions with over-squeezing, over-stretching, distortion of proportion and distortion of geometric structure respectively, and keep them with the variables  $N_{osq}$ ,  $N_{ost}$ ,  $N_{pd}$  and  $N_{sd}$  respectively. Undoubtedly, the more the visual salient regions with deformation after image retargeting, the lower the visual effect similarity between the original and retargeted images. Visual effect similarity is inversely proportional to the number of visual salient regions with deformation. Thus, for the original image  $I_{or}$  and the retargeted image  $I_{re}$ , we measure their visual effect similarity as shown in formula (11), where  $w_{osq}$ ,  $w_{ost}$ ,  $w_{pd}$  and  $w_{sd}$  are the weights of four deformations respectively, which are generally all assigned 0.25.

$$VEM_{(I_{or}, I_{re})} = w_{osq} / N_{osq} + w_{ost} / N_{ost} + w_{pd} / N_{pd} + w_{sd} / N_{sd} \quad (11)$$

### 5. Similarity Measure for Image Retargeting

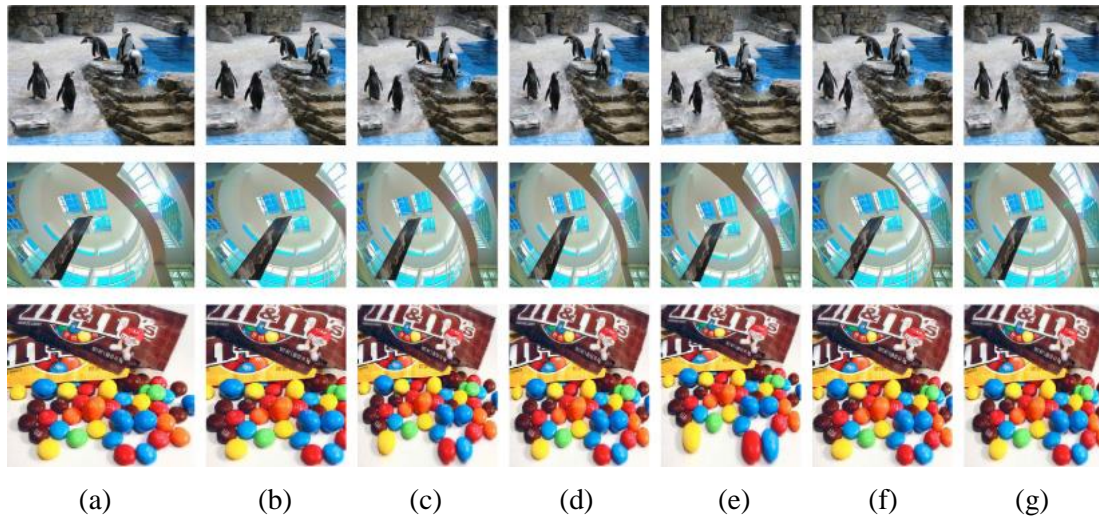
As illustrated by the framework of similarity measure for image retargeting proposed by us in the Section 2, since we have measured content similarity and visual effect similarity between the original image and retargeted image in Section 3 and 4, we can assess the overall similarity between these two images. We design and utilize formula (12) to achieve similarity measure with deformation detection of visual salient regions for image retargeting, which is a weighted combination of content similarity and visual effect similarity, where  $CSM_{(I_{or}, I_{re})}$  comes from formula (5) and  $VEM_{(I_{or}, I_{re})}$  is from formula (11).  $w_{CSM}$  and  $w_{VEM}$  mean the weights of content similarity and visual effect similarity respectively, and generally are all assigned 0.5, and also can be adjusted according to the actual situation of image retargeting.

$$SMDDIR_{(I_{or}, I_{re})} = w_{CSM} * CSM_{(I_{or}, I_{re})} + w_{VEM} * VEM_{(I_{or}, I_{re})} \quad (12)$$



## 6. Experiments and Results

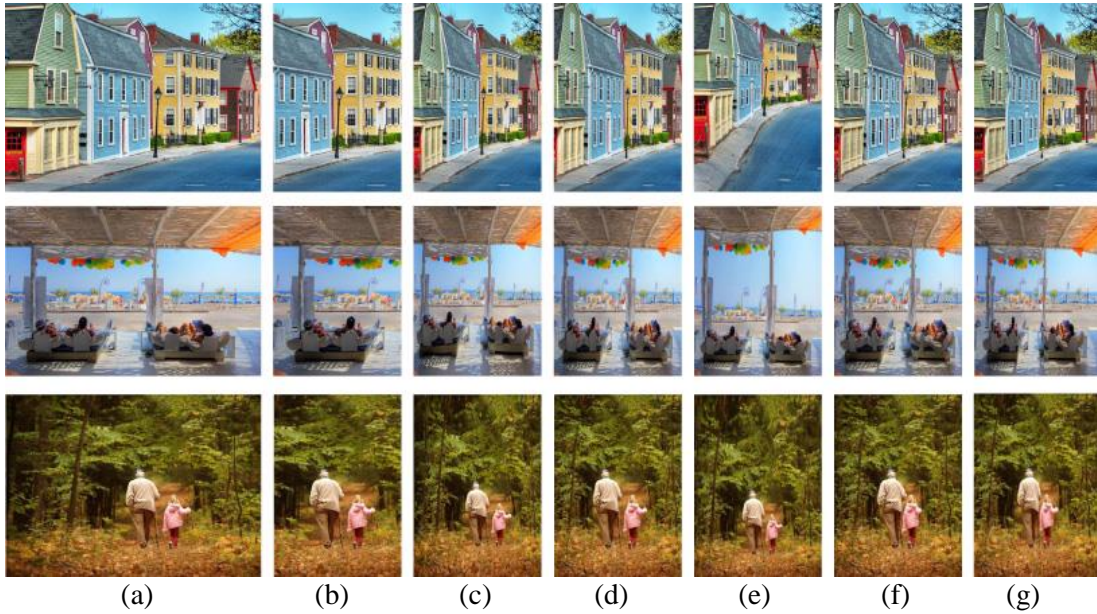
In this section, we took real retargeted images to confront the proposed similarity measure with the real world, so as to validate it. The image dataset in our experiment is picked from the RetargetMe dataset [14] which contains 80 original images with various attributes. Each original image has all or partial of eleven retargeted images generated by manual cropping (CR), streaming video (SV), multi-operator (MULTIOP), scale-and-stretch (SNS), seam carving (SC), non-homogeneous warping (WARP), energy-based deformation (LG), shift-maps (SM), uniform scaling (SCL), quadratic programming (QP) and object size adjusted (OSA), respectively, through shrinking or stretching the original image to different degrees.



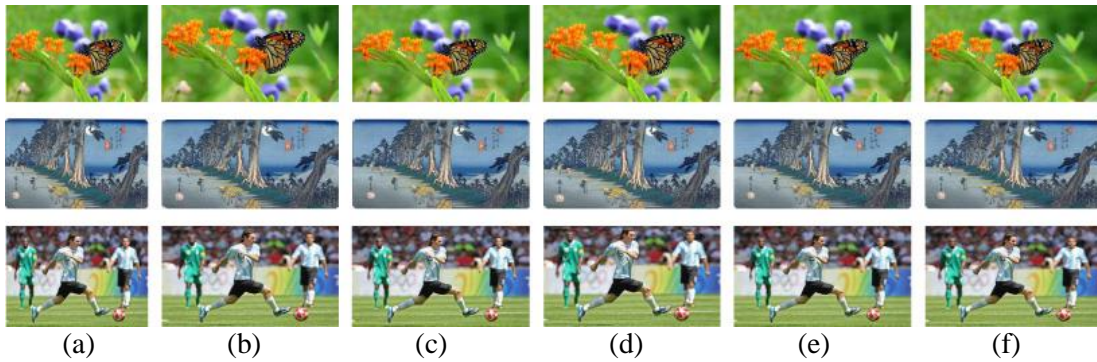
**Figure 8. The First Group of image similarity measure cases where the Original Images have been Shrunk into Retargeted Images with a Ratio of 0.75 in Width. From Top to Bottom, these Cases are “penguins”, “getty” and “mnm”, and from Left to Right, these Images are (a) the Original Image, and the Retargeted Image Generated by (b) CR, (c) SV, (d) MULTIOP, (e) SNS, (f) SC and (g) WARP**

As shown in Figures 8-11, we select ten typical cases of image similarity measure for retargeting in the paper due to space limitation, which involve different scenes including “penguins”, “getty”, “mnm”, “Marblehead\_Mass”, “Perissa\_Santorini”, “family”, “butterfly”, “mochizuki”, “soccer” and “DKNYgirl” for illustrating the extensive application of our method. According to the degree with original image being shrunk or stretched during retargeting process, we divided them into four groups. As illustrated in Figure 8 from top to bottom, the first group involves three cases: “penguins”, “getty” and “mnm”, where the original images have been shrunk into retargeted images with a ratio of 0.75 in width. As shown in Figure 9 from top to bottom, the second group involves three cases: “Marblehead\_Mass”, “Perissa\_Santorini” and “family”, where the original images have been shrunk with a ratio of 0.5 in width, with being narrower. As illustrated in Figure 10 from top to bottom, the third group involves three cases: “butterfly”, “mochizuki” and “soccer”, where the original images have been stretched into retargeted images with a ratio of 1.25 in width. As shown in Figure 11, the fourth group involves the case “DKNYgirl”, where the original image has been stretched with a ratio of 1.5 in width, with being wider. In every case, we measure the similarity between the original image in the leftmost and each of the remaining images using our proposed method. Here the remaining images are generated by manual

cropping (CR), streaming video (SV)[15], multi-operator (MULTIOP) [16], scale-and-stretch (SNS) [17], seam carving (SC) [18] and non-homogeneous warping (WARP)[19] respectively.



**Figure 9. The Second Group of Image Similarity Measure Cases where the Original Images have been Shrunk into Retargeted Images with a Ratio of 0.5 in Width. From Top to Bottom, these Cases are “Marblehead\_Mass”, “Perissa\_Santorini” and “family”, and from Left to Right, these Images are (a) the Original Image, and the Retargeted Image Generated by (b) CR, (c) SV, (d) MULTIOP, (e) SNS, (f) SC and (g) WARP**



**Figure 10. The Third Group of Image Similarity Measure Cases where the Original Images have been Stretched into Retargeted Images with a Ratio of 1.25 in Width. From Top to Bottom, these Cases are “Butterfly”, “mochizuki” and “soccer”, and from Left to Right, these Images are (a) the Original Image, and the Retargeted Image Generated by (b) SV, (c) MULTIOP, (d) SNS, (e) SC and (f) WARP**

More specifically, we apply formula (12) to calculate the image similarity between the original image and every retargeted image for each case. Then the retargeted images are

ranked according to their similarity to the original image and the top ranked image is considered the most similar to the original one.



**Figure 11. The Fourth Group for Image Similarity Measure Involves the Case “DKNYgirl”, where the Original Image has been Stretched into Retargeted Images with a ratio of 1.5 in Width. From Left to Right, these Images are (a) the Original Image, and the Retargeted Image Generated by (b) SV, (c) MULTIOP, (d) SNS, (e) SC and (f) WARP.**

For the first group, the calculated rank results of similarity measure are summarized by Table 1. According to the obtained results in Table 1, the retargeted image generated by CR is almost the most similar to the original image in these three cases, followed by the retargeted image generated by MULTIOP and SV. According to common sense, we also know that CR can achieve good similarity results where the image width or height does not be changed too much during the retargeting process. Besides, in some cases such as “mnm”, WARP also achieves the best rank, which is not surprising since the original image is only reduced by 25 percent in width and a small number of deformations are introduced during retargeting.

**Table 1. The Calculated Rank Results of Similarity Measure for the First Group in Figure 8**

| Similarity Rank | Rank 1  | Rank 2 | Rank 3  | Rank 4 | Rank 5 | Rank 6 |
|-----------------|---------|--------|---------|--------|--------|--------|
| penguins        | CR      | SV     | MULTIOP | SNS    | WARP   | SC     |
| getty           | MULTIOP | CR     | SV      | SNS    | WARP   | SC     |
| mnm             | WARP    | CR     | MULTIOP | SV     | SC     | SNS    |

For the second group, the evaluated rank results of similarity measure are summarized by Table 2. As shown in Table 2, the retargeted image generated by SV is almost the most similar to the original image in these three cases, followed by the retargeted image generated by MULTIOP, SNS and SC. The lowest ranks are achieved by CR and WARP. When the original image is shrunk too much in width or height, for example by 50 percent, CR will remove lots of contents of original image, which will humiliate this operator. Removal of content and distortion of visual salient regions hinder the WARP operator, which might collapse regions in the image during its deformation if not enough areas of homogeneous content are found.

**Table 2. The Calculated Rank Results of Similarity Measure for the Second Group in Figure 9**

| Similarity Rank   | Rank 1  | Rank 2  | Rank 3 | Rank 4 | Rank 5 | Rank 6 |
|-------------------|---------|---------|--------|--------|--------|--------|
| Marblehead_Mass   | SV      | MULTIOP | SNS    | SC     | WARP   | CR     |
| Perissa_Santorini | SV      | MULTIOP | SNS    | SC     | WARP   | CR     |
| family            | MULTIOP | SV      | SC     | CR     | SNS    | WARP   |

For the third and fourth groups, the obtained rank results of similarity measure are summarized by Table 3. As shown in Table 3, the retargeted image generated by MULTIOP is almost the most similar to the original image in these four cases, followed by the retargeted



image generated by SV. WARP always gets a poor rank when the original image is expanded in width or height, especially by 50 percent.

**Table 3. The Calculated Rank Results of Similarity Measure for the Third Group in Figure 10 and the Fourth Group in Figure 11**

| Similarity Rank | Rank 1  | Rank 2 | Rank 3  | Rank 4 | Rank 5 |
|-----------------|---------|--------|---------|--------|--------|
| butterfly       | SV      | SNS    | MULTIOP | WARP   | SC     |
| mochizuki       | MULTIOP | SV     | SC      | WARP   | SNS    |
| soccer          | MULTIOP | SV     | SC      | WARP   | SNS    |
| DKNYgirl        | MULTIOP | SV     | SNS     | SC     | WARP   |

From these ten cases, on the whole, the retargeted images generated by SV and MULTIOP generally achieve the best rank about similarity to the original image, and the generated images by WARP and SC always obtain the last rank, according to our proposed approach. It is consistent with the fact that the retargeted image generated by WARP or SC involves some distortions or deformations. For every case, we also can see that the achieved similarity rank results in Figures 8-11 approximate the subjective human choices and are consistent as well as accurate. Thus our proposed method about similarity measure for retargeting has been validated.

## 7. Conclusion

In this paper, we proposed a similarity measure with deformation detection of visual salient regions for image retargeting applications. The proposed method is based on a combination of content similarity and visual effect similarity. Content similarity pays attention to the number and layout of the salient contents in the image and involves the sub-steps of finding seams, dividing image into regions and matching regions. Visual effect similarity places emphasis on shape preservation of visual salient regions and is characterized by deformation detection of visual salient regions between original and retargeted images which mainly focuses on over-squeezing, over-stretching, distortion of proportion and distortion of geometric structure. Experimental results showed that the achieved similarity ranks for image retargeting closely match the subjective human choices, indicating that our proposed objective measure is almost congruent with human perception mechanism. The good results from the experiments illustrated the practicability and effectiveness of our proposed approach.

## Acknowledgements

The work is supported by the National Nature Science Foundation of China (No.61374014), the Natural Science Foundation of Henan Educational Committee (No.12B520068) and the PHD Foundation of Zhengzhou University of Light Industry (No.2011BSJJ002).

## References

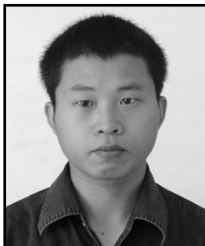
- [1] A. Shamir and O. Sorkine, "Visual media retargeting", In ACM SIGGRAPH ASIA 2009 Courses, (2009) December 16-19, Yokobama, Japan.
- [2] R. L. Gregory, "Eye and brain: The psychology of seeing", Princeton university press, New Jersey (1997).
- [3] A. M. Eskicioglu, "Quality measurement for monochrome compressed images in the past 25 years", IEEE International Conference on Acoustics, Speech, and Signal Processing, (2000) June, Istanbul, Turkey.

- [4] I. T. U. R. B. T. Recommendation, "500-11, Methodology for the subjective assessment of the quality of television pictures", International Telecommunication Union, Geneva, Switzerland (2002).
- [5] H. Marmolin, "Subjective MSE measures", IEEE Transactions on Systems, Man and Cybernetics, vol. 16, no. 3, (1986), pp. 486-489.
- [6] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", IEEE Transactions on Image Processing, vol. 13, no. 4, (2004), pp. 600-612.
- [7] C. Liu, J. Yuen and A. Torralba, "SIFT Flow: Dense Correspondence across Scenes and Its Applications", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 5, (2011), pp. 978-994.
- [8] O. Pele and M. Werman, "Fast and robust earth mover's distances", In: IEEE 12th international conference on Computer vision, (2009) September 27 - October 4, Kyoto, Japan.
- [9] D. Simakov, Y. Caspi, E. Shechtman and M. Irani, "Summarizing visual data using bidirectional similarity", In: IEEE Conference on Computer Vision and Pattern Recognition, (2008) June 24-26, Anchorage, Alaska, USA.
- [10] M. Rubinstein, A. Shamir and S. Avidan, "Multi-operator Media Retargeting". ACM Transactions on Graphics, vol. 28, no. 3, (2009), pp. 23.
- [11] M. Rubinstein, D. Gutierrez, O. Sorkine and A. Shamir, "A Comparative Study of Image Retargeting", ACM Transactions on Graphics, vol. 29, no. 6, (2010), pp. 160.
- [12] L. Chen, "Topological structure in visual perception", Science, vol. 218, (1982), pp. 699-700.
- [13] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing". ACM Transactions on Graphics, vol. 26, no. 3, (2007), pp. 10.
- [14] RetargetMe, <http://people.csail.mit.edu/mrub/retargetme/>
- [15] P. Krähenbühl, M. Lang, A. Hornung and M. Gross, "A system for retargeting of streaming video", ACM Transactions on Graphics, vol. 28, no. 5, (2009), pp. 126.
- [16] M. Rubinstein, A. Shamir and S. Avidan, "Multi-operator media retargeting", ACM Transactions on Graphics, vol. 28, no. 3, (2009), pp. 23.
- [17] Y. S. Wang, C. L. Tai, O. Sorkine and T. Y. Lee, "Optimized Scale-and-Stretch for Image Resizing", ACM Transactions on Graphics, vol. 27, no. 5, (2008), pp. 118.
- [18] M. Rubinstein, A. Shamir and S. Avidan, "Improved seam carving for video retargeting", ACM Transactions on Graphics, vol. 27, no. 3, (2008), pp. 16.
- [19] L. Wolf, M. Guttman, D. Cohen-Or, "Non-homogeneous content-driven video-retargeting", In: IEEE 11th International Conference on Computer Vision, (2007) October 14-20, Rio de Janeiro, Brazil.

## Authors



**Li Canlin**, he received PhD degree in computer science from Shanghai Jiaotong University in 2010. Since 2010 he has been in the School of Computer and Communication Engineering at Zhengzhou University of Light Industry. His research interests include image processing, multimedia, digital entertainment and software engineering. Dr. Li is a member of IEEE as well as ACM.



**Zhu Fubao**, he received PhD degree in computer science from Wuhan University in 2008. He is an associate professor of School of Computer and Communication Engineering at Zhengzhou University of Light Industry. His research interests include image processing, Geographic Information System and computer software.



**Su Rijian**, he received PhD degree in control science and engineering from Huazhong University of Science and Technology in 2010. He is an associate professor of School of Computer and Communication Engineering at Zhengzhou University of Light Industry. His research interests include control engineering, image and signal processing and computer application.