Scale and Object Aware Image Retargeting for Thumbnail Browsing

Jin Sun   Haibin Ling

Center for Data Analytics & Biomedical Informatics, Dept. of Computer & Information Sciences

Temple University, Philadelphia, PA, 19122

{jin.sun, hbling} @temple.edu

Abstract

Many image retargeting algorithms, despite aesthetically carving images smaller, pay limited attention to image browsing tasks where tiny thumbnails are presented. When applying traditional retargeting methods for generating thumbnails, several important issues frequently arise, including thumbnail scales, object completeness and local structure smoothness. To address these issues, we propose a novel image retargeting algorithm, Scale and Object Aware Retargeting (SOAR)*, which has four components: (1) a scale dependent saliency map to integrate size information of thumbnails, (2) objectness (Alexe et al. 2010) for preserving object completeness, (3) a* cyclic seam carving *algorithm to guide continuous retarget warping, and (4) a thin-plate-spline (TPS) retarget warping algorithm that champions local structure smoothness. The effectiveness of the proposed algorithm is evaluated both quantitatively and qualitatively. The quantitative evaluation is conducted through an image browsing user study to measure the effectiveness of different thumbnail generating algorithms, followed by the ANOVA analysis. The qualitative study is performed on the RetargetMe benchmark dataset. In both studies, SOAR generates very promising performance, in comparison with state-of-the-art retargeting algorithms.*

1. Introduction

The increasing popularity of digital display devices imposes the need of effective ways for presenting image sets. In this context, browsing a large image set, such as personal photo albums or scientific image collections, becomes an important tool in many applications. Such a tool usually requires casting images into a display device with a fixed small size. The traditional method, scaling, which shrinks original images directly into thumbnails, often brings difficulties in searching and recognition [21], especially for tiny thumbnails shown on small size screens, such as cell phones or PDAs.

Automatic image retargeting, by reducing image size while preserving important content, is an important visual summarization tool and has been attracting a large amount of research attention recently [17]. Many image retargeting algorithms (Sec. 2) generate aesthetically impressive results when the target size is comparable to that of the original image. However, insufficient attention has been paid explicitly to image browsing tasks where tiny thumbnails are widely used. To design effective image retargeting algorithms for thumbnail browsing, several important issues need to be considered:

Thumbnail scales. In image browsing, thumbnails usually have much smaller scales/sizes than do the original images. Studies have shown that scales can have significant effects on human visual perception [12, 14, 9]. To the best of our knowledge, none of existing retargeting algorithms explicitly takes into account such scale information.

Object completeness. To keep objects in the image as complete as possible is important in image retargeting. One challenge lies mainly in the difficulty of explicit definition of object completeness. Consequently, many retargeting methods implicitly handle this problem by keeping low-level gradient-based information as much as possible. The object-level completeness, in contrast, is seldom considered.

Structure smoothness. The contamination in structure smoothness caused by pixel removal methods, e.g. seam carving, usually results in little visual defect when target images have reasonably large sizes. This is unfortunately not always true for targets as tiny as thumbnails, since intense scale change can create serious image structure discontinuities that humans are not comfortable with.

In this paper, we propose a new image retargeting algorithm, *Scale and Object Aware Retargeting* (SOAR), to address the issues discussed above. SOAR is a continuous retargeting algorithm but uses discrete retargeting to guide the warping transformation. In particular, SOAR contains four key components (Fig. 1): (1) We propose a new *scale-dependent saliency* that takes into account the scale information of the target thumbnails. Inspired by the study in visual acuity [12, 14], we compute the saliency response on
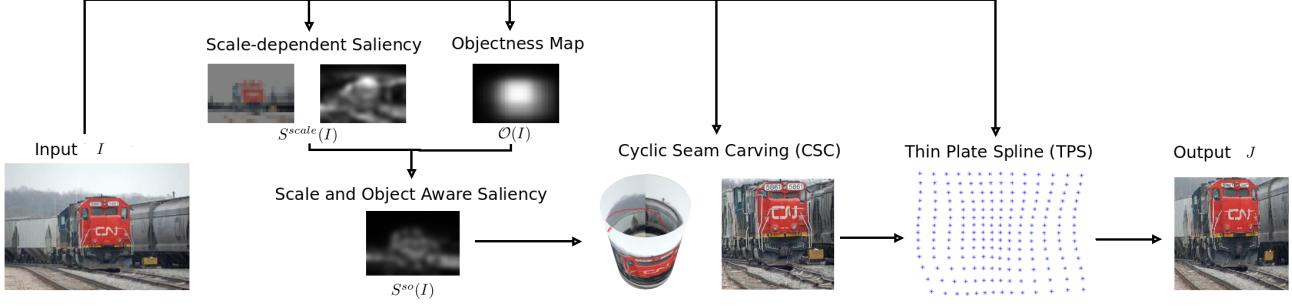
Figure 1. Flow chart of the proposed method.

the image that is actually perceived by human vision system. This image is estimated according to the visual acuity theory, which suggests the visibility of local structures. (2) We improve object-level object completeness preservation by integrating the *objectness* measurement recently proposed by Alexe et al. [1]. The measurement is combined with the scale-dependent saliency to tune down the saliency values for regions with less objectness, e.g., backgrounds. The combination results in *scale and object aware saliency*. (3) We extend seam carving [2] by introducing *cyclic seams* to allow cross-boundary pixels removal. This Cyclic Seam Carving (CSC) is combined with the scale and object aware saliency to conduct a discrete image retargeting. (4) We use the thin-plate-spline (TPS) as warping function in SOAR. Specifically, we use CSC to generate landmark points between input and output image spaces and then fit the TPS model for image warping. This combination inherits the effectiveness of seam carving while benefits the smoothness from TPS.

We evaluate the proposed SOAR algorithm both quantitatively and qualitatively. The quantitative study is very important since human's perception of thumbnail quality and thumbnail browsing can be very subjective. For the purpose, we have conducted an image browsing user study, in which each user is asked to search a target in a screen of thumbnails. The time cost and accuracy of the search procedure are recorded and analyzed statistically (ANOVA and Tukey's significance test). The results show the superiority of our approach over several other ones. Regarding the qualitative study, we apply SOAR to the RetargetMe dataset [17] on which results from many existing algorithms have been collected by [17]. The study shows that, our method, despite emphasizing on the thumbnail browsing effectiveness rather than aesthetic effects, generates images that are visually as good as previously reported results.

2. Related work

Image retargeting has been actively studied recently. A comparative study can be found in [17]. Roughly speaking, existing methods can be divided into two categories:

discrete methods that remove unimportant pixels and continuous methods that reorganize image content with a continuous warping function.

A typical example of discrete image retargeting methods is the well known *Seam Carving* (SC) proposed by Avidan and Shamir [2]. The idea is to selectively and iteratively remove continuous pixel segments while still preserve the image structures as much as possible. Such segments, termed *seams*, usually contain pixels that are less important than others according to certain importance measure. SC attracts a lot of research attention due to its elegancy and flexibility. Several notable works that extend and improve the original SC can be found in [6, 13, 18]. There are other discrete retargeting methods such as [20], which uses the bi-directional similarity to guide the retargeting process.

Continuous warping has been used in many image retargeting algorithms [4, 7, 10, 11, 24]. The basic idea is to deform the original image to the target image through a continuous transformation, which is estimated by image content. The framework of warping is very general because a large variety of transformation models and content models can be used to address different challenges and to adjust to different applications. A hybrid approach [19] has also been proposed that smartly combines several retargeting algorithms for further improvement.

Our method is different than previous studies mainly in two aspects. First, our task is to generate target images for thumbnail browsing, which has not been studied in most previous retargeting methods. This task motivates us to take the scale information into account while determining pixels importance. Second, our method combines both continuous and discrete schemes by using TPS for warping model and cyclic seam carving for tracing landmark points.

The scale dependent saliency in our method is motivated by the studies of human visual acuity [12, 14, 15], which describes how well human can distinguish visual stimuli at various frequencies. If the frequency is too high/low, human can hardly distinguish different stimulus. Given the image size and an observation distance, the theory can be used to decide the minimum scale at which an image structure is perceivable. Another important component used in our

saliency computation is the objectness measure proposed in [1]. Accordingly, the probability that a given bounding window contains an object is predicted by a model learned from training sets. Our method also shares some philosophy with the work on perceptual scale space [23].

3. Overview

Task Formulation. Let I be an input image of size $m_0 \times n_0$ and $S(I)$ be the saliency map of I, such that $S(I)$ is a matrix of $m_0 \times n_0$. We formulate the retargeting problem as to find a retargeting function $\mathcal{F}$:

$$J = \mathcal{F}(I, S(I)) , \tag{1}$$

where J is the result *retargeted image* of *target size* $m_1 \times n_1$. In this formulation, a retargeting method is characterized by retargeting function $\mathcal{F}(.)$ and saliency computation $S(.)$[1].

Note that the function $\mathcal{F}(.)$ can be either continuous or discrete. For example, the *seam carving* (SC) algorithm [2], when carving a vertical seam $seam(i)$, can be defined as

$$\mathcal{F}^{sc}_{i,j} = \begin{cases} I(i,j), & \text{if } j < seam(i) \\ I(i,j+1), & \text{if } j \geq seam(i) \end{cases} , \tag{2}$$

where $i \in [1, m_0]$ and $j \in [1, n_0]$ are row and column indices respectively; $seam(i)$ indicates the seam position at row i calculated by the seam searching strategy to minimize the carving energy defined over saliency $S(I)$. Horizontal seam removal is defined similarly. SC employs a discrete assignment because the pixels along the path of seams, i.e. $\{I(i, seam(i))\}_{i=0}^{m_0}$, are eliminated and the information they carried is discarded.

Our goal is to design an image retargeting algorithm in the context of image browsing, where tiny image thumbnails are the final results. A naive solution is to first set the target size as thumbnail size and then apply retargeting algorithms directly. This is however too aggressive since thumbnails are usually much smaller than the original images. Instead, through out this study, we use retargeting methods to first get retargeted images to a target size that is comparable to the original image size, and then shrink the retargeted images to get the final thumbnails.

Framework Overview. For the goal mentioned above, we propose a new image retargeting algorithm, which we call *scale and object aware image retargeting* (SOAR), denoted as $\mathcal{F}^{so}$. To reduce the contamination of semantic structures in input images, SOAR uses the *thin-plate-spline* (TPS) model as the warping function, which is well known for modeling real world smooth deformations. Fitting a TPS model requires a set of landmark points, which are obtained using an improved seam carving algorithm.

---

[1]We use saliency to indicate the importance measurement used in general, which is not limited to the visual attention-based saliency.
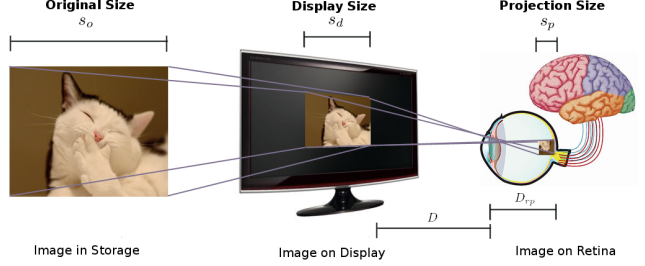

Figure 2. Demonstration of image sizes in different stages.

An important issue we address is the thumbnail scale. Inspired by the study of visual acuity in human vision, we propose using scale dependent saliency, denoted as $S^{scale}(I)$, to reflect the relative scale variation from the original image to the image thumbnail. The saliency is further augmented by combining with object awareness, measured by objectness map $\mathcal{O}(I)$, which is derived by the recently proposed *objectness* measurement [1]. The resulting *scale and object aware saliency* $S^{so}(I)$ is then integrated in the carving energy and fed into our improved seam carving algorithm, named *cyclic seam carving* (CSC). The result of CSC is used to extract landmark points for estimating the TPS warping. In short, for the input image I, SOAR creates the target image J through the following process

$$J = \mathcal{F}^{so}(I, S^{so}(I)) . \tag{3}$$

A summary of the algorithm flow is given in Fig. 1.

4. Scale and Object Aware Image Retargeting

4.1. Scale-dependent Saliency

When a subject is observing an image thumbnail on a screen, sizes of three different images are involved: the original image size s_o in pixels, the display size s_d in inches, and projection size s_p on the retina. These sizes are related by two distances: distance D between the observer's eye to the display device and distance D_{rp} between human retina and pupil. Fig. 2 illustrates the relations between the these variables, which is summarized below:

$$s_d = \frac{s_o}{\text{DPI}}, \quad s_p = \frac{s_d \cdot D_{rp}}{D} , \tag{4}$$

where DPI (Dots Per Inch) is the screen resolution.

Since a thumbnail is eventually presented to and perceived by human visual system, it is critical to explore how well such a system preserves the image information. In particular, we want to make the foreground object/content/theme of the image as "clear" as possible. This has been studied in psychology and ophthalmology in terms of *visual acuity* [12, 14, 15]. According to the study, not all patterns in an image are recognizable by human. In fact, under the above configuration, the perceived image, denoted as I_p, can be derived as

$$I_p(I(i,j)) = \begin{cases} I(i,j), & \text{if } s_d/D \in [\kappa, \rho] , \\ I_N(i,j), & \text{otherwise} , \end{cases} \qquad (5)$$

where $I_N(i,j)$ is the mean value of N-connected neighbors of pixel $I(i,j)$; N is determined by display device specifications; κ and ρ are the lower and upper bounds (in cycles per degree) respectively of human visual acuity. They define the limits at which the visual stimuli frequency becomes too low or too high to be recognized by human. We use $\kappa = 50$ and $\rho = 1$ according to [14].

According to Eqn.5, for a small display size (thumbnail size in our case), a pixel may become indistinguishable from its neighbors to a human observer. Consequently, an image patch that was salient in the original image may not appear salient to a human observer when the patch is displayed as a small thumbnail. Inspired by this observation, we propose using scale dependent saliency to encode the scale information of the final thumbnails.

The procedure of computing the scale-dependent saliency goes as follows:

First, the original image is scaled in homogeneous into the final thumbnail size, i.e. the display size, 60×60 pixels in our experiment.

Then, the minimum recognizable pattern, denoted by s'_d, is determined by Eqn.4. Specifically, in our experiment where a monitor of $1680 \times 1050@65hz$ and 120DPI is used, we find the value s'_d is in average 0.009 inches under normal indoor illumination. The size is approximately the distance between two pixel lines on the screen, i.e. $N = 4$ in Eqn.5. As a result, in the final thumbnail four adjacent neighbors of one image pixel patch with value differences in color space within certain threshold will be assigned their mean value, which means those pixels are unable to be distinguished by human. This threshold, intensity differences of 50 in our experiment, is determined according to the gray scale settings in contrast sensitive study [14] and our preliminary experiments.

Finally, the scale dependent saliency $S^{scale}(I)$ is defined as

$$S^{scale}(I) = S(I_p) , \qquad (6)$$

where $S(.)$ is the standard saliency, by which we mean the visual saliency defined in [8]. Some examples of scale dependent saliency are shown in Fig. 3.

4.2. Scale and Object Aware Saliency

Preserving object completeness as much as possible is commonly desired in retargeting algorithms. One bottleneck lies in the difficulty of the explicit definition of such completeness. Recently, Alexe et al. [1] proposed a novel *objectness* measure, which is trained to distinguish object windows from background ones. The objectness measure combines several image cues, such as multi-scale saliency, color contrast, edge density and superpixel straddling, to
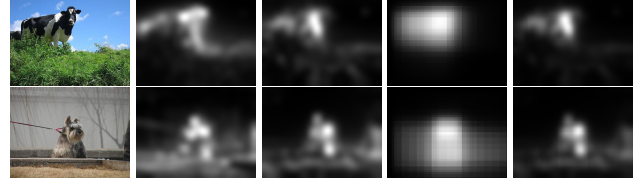

Figure 3. Saliency computations, from left to right: input images, standard saliency [8], scale dependent saliency, objectness map, scale and object aware saliency.

predict the likelihood of a given window containing an object. Specifically, for a rectangular window $\mathbf{w} = (i,j,w,h)$ with a top-left corner at (i,j), width w and height h, its objectness is defined as the probability $p_{obj}(\mathbf{w})$ that $\mathbf{w}$ contains an object.

To get the objectness distribution over pixels, we first sample n_w windows $W = \{\mathbf{w}_i\}_{i=1}^{n_w}$ for an input image and then calculate the *objectness map* $\mathcal{O}$ as the expected objectness response at each pixel,

$$\mathcal{O}(i,j) = \frac{1}{\Gamma} \sum_{\mathbf{w} \in W \wedge (i,j) \in \mathbf{w}} p_{obj}(\mathbf{w}) , \qquad (7)$$

where $\Gamma = \max_{i,j} \mathcal{O}(i,j)$ is used for normalization; $(i,j) \in \mathbf{w}$ means pixel (i,j) falls in $\mathbf{w}$; and $n_w = 10000$ is used in our experiments.

The map $\mathcal{O}(I)$ is then combined with the scale-dependent saliency $S^{scale}(I)$ to inhibit regions with small objectness values:

$$S^{so}(i,j) = S^{scale}(i,j) \cdot \mathcal{O}(i,j) . \qquad (8)$$

We name the augmented saliency S^{so} as *scale and object aware saliency*. Fig. 3 illustrates some examples.

4.3. Cyclic Seam Carving

To guide the TPS warping in our model, we use seam carving (SC) [2] with two extensions, including using cyclic seams and encoding the scale and object aware saliency.

First, we have observed that in many cases a seam has no choice but to cross objects due to the original definition of seam: a continuous polyline from one boundary of the image to its opposite boundary. The results sometimes suffer from damaging the well-structured object in the scene. One solution is to generalize the original seams to discontinuous ones. While this strategy might work in some cases, it introduces new kind of object-structure damage, e.g. pixel shift problem, even when the seam is far away from the object.

To reduce the chance of a seam to trespass objects, we introduce the *cyclic seams* as following: we first warp the image to make it a cylinder shape by sticking its left and right (or top and bottom) margins together. Then a cyclic seam is defined still as the standard continuous seam but on this virtual "cylinder" image. An illustration is shown in

Figure 4. Cyclic seam carving (CSC). Left: the imaginary image in a cylinder shape; right: a cyclic seam on the real image.

Figure 5. From left to right: original image, SC with standard Saliency, CSC with standard Saliency; CSC with scale dependent saliency; CSC with scale and object aware saliency, SOAR result.

Fig. 4. We name this extended SC algorithm *Cyclic Seam Carving* (CSC). Intuitively, CSC allows a seam to cross image boundaries to stay away from highly salient regions. On the other hand, a cyclic seam is still continuous in most of its segments. In addition, the pixel shift problem caused by cross-boundary seams can be moderated by adjusting proper penalty values at boundary pixels.

Our second extension to the original SC is to augment the energy function with the proposed scale and object aware saliency. Denote E_{sc} as the original energy used in SC that captures the distribution of histogram of gradients, our scale and object aware energy E_{scale} is defined as

$$E_{scale} = \rho \cdot E_{sc} + (1 - \rho) \cdot S^{so} , \qquad (9)$$

where ρ is the weight and set to 0.3 through out our experiments. The improved energy is then combined with the CSC algorithm to provide landmark point pairs needed for estimating TPS warping (Sec. 4.4).

Fig. 5 shows some results comparing CSC with different saliency definitions. From the figure we can see how different components help improving the retargeting qualities.

4.4. Image Warping Function

As pointed out in previous sections and also observed in [20], many discrete retargeting methods generate excellent results in general but they sometimes create serious artifacts when the target has a size much smaller than the input. In [20] a bidirectional similarity is used to alleviate the problem. We study this problem differently by combining a continuous warping model with a discrete retargeting guidance.

First, we use the thin-plate-spline (TPS) [3] for our continuous warping model. TPS has been widely used in many vision tasks, such as registration and matching, due to its attractive property in balancing local warping accuracy and smoothness. Specifically, given n_l landmark points $P = \{\mathbf{p}_i \in \mathbb{R}^2, i = 1, 2, \cdots, n_l\}$ and $Q = \{\mathbf{q}_i \in \mathbb{R}^2, i =$

$1, 2, \cdots, n_l\}$, where $\mathbf{p}_i$ is mapped to $\mathbf{q}_i$, the TPS transformation $\mathcal{T}$ is defined as the transformation from P to Q that minimizes the regularized bending energy $\mathcal{E}(f)$ defined as

$$\mathcal{E}(f) = \sum_i \| \mathbf{q}_i - f(\mathbf{p}_i) \|^2 + \\ \lambda \iint (\frac{\partial^2 f}{\partial x^2})^2 + 2(\frac{\partial^2 f}{\partial x \partial y})^2 + (\frac{\partial^2 f}{\partial y^2})^2 dx dy , \qquad (10)$$

where x, y are the coordinates and λ is the weight parameter. The TPS warping is then defined as $\mathcal{T} = \arg\min_f \mathcal{E}(f)$. In our image warping problem, f is the displacement mapping from the input image to the target image. Intuitively, f achieves a balance between local warping accuracy and smoothness.

To estimate the TPS warping function, we need to provide the landmark point pairs P and Q. This is derived from the CSC retargeting algorithm. A natural way is to define a point set in the input image and find the corresponding point set in the target image by tracing the CSC process. However, the point set achieved this way can be unstable due to the singularity caused by the seam carving iterations. Instead, we design a two-way solution described in the following: We first sample randomly a landmark set $\widehat{P}$ (Fig. 6(a)) from original image and then trace their shifting in the CSC process. During the process, if a landmark point was eliminated, its direct neighbor will be assigned as a landmark point. After the CSC iterations, we get the corresponding set $\widehat{Q}$ (Fig. 6(b)). Then, point set Q (Fig. 6(c)) is re-sampled uniformly on the target image, which is generated by CSC. Finally, a sample set P (Fig. 6(d)) is generated by mapping Q to the original image using warping estimated by $\widehat{Q}$ and $\widehat{P}$. The landmark sets P and Q are then used to estimate the warping used in the final retargeting.

5. Experiments

5.1. Quantitative Experiments

Experiment Setup. We design an image browsing task for user study using a carefully prepared dataset. For each
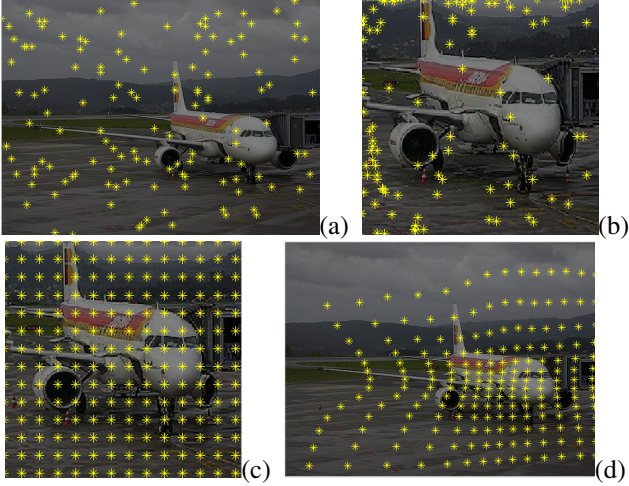
Figure 6. (a) Landmarks in original image; (b) Landmarks traced in seam carved image; (c) New landmarks sampled from seam carved image; (d) Landmarks traced in original image.
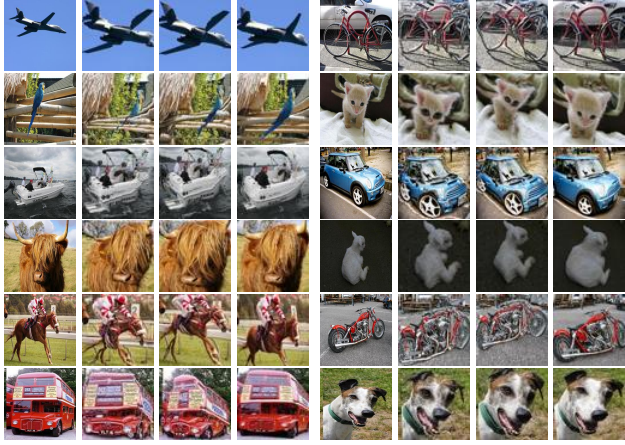

Figure 7. Thumbnails used in the quantitative study. From left to right: SL, SC, ISC, and SOAR (proposed).

Figure 8. The user interface used in the user study.

subject, the task requires him/her to browse and select a target, described verbally, from a set of image thumbnails randomly placed in a page. In each test page, we ask the user to choose one particular image, say "cat", from various types of images: "bike", "train", "bird", and etc. An example page is shown in Fig. 8, where the subject is asked to find a thumbnail of a "motorbike" from the 10×10 thumbnails. There exists only one correct thumbnail in each page to avoid ambiguity.

For images used in the task, we randomly selected 210 images from the PASCAL VOC 2008 database [5]. These images are divided to 14 classes and each class has 15 images. The classes are: aeroplane, bicycle, bird, boat, bus, car, cat, cow, dining table, dog, horse, motorbike, sheep and train. Each subject is requested to browsing in total 210 pages. In each page, the subject will see a word describ-
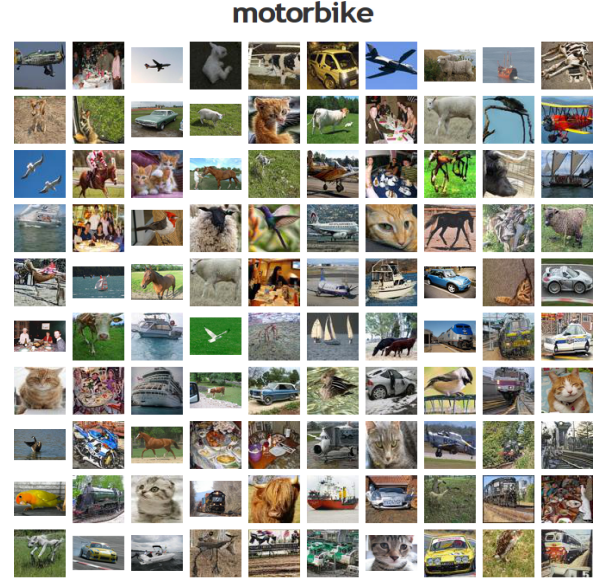
ing the class of the target and 10×10 image thumbnails aligned in 10×10 grid. Only one thumbnail that matches the description will be put randomly in one of the 100 positions. Other 99 "filler" images are chosen randomly from the rest classes. The potential image class conflict is taken into consideration while choosing the filler images: for example, motorbike images and bicycle images should not appear in the same page since they are similar in appearance especially in image thumbnails.

For every image in the dataset, four thumbnails are generated using different methods including scaling(SL), seam carving(SC) [2], improved seam carving(ISC) [18], and the proposed SOAR algorithm. For every thumbnail demonstrated on the screen, one of four thumbnails is randomly picked to be presented in the page. The time a user spend to find the result and the selection are recorded for evaluation. The records of the first 10 pages are discarded to allow the subjects get familiar with the system. Example thumbnails from different methods are shown in Fig. 7.

We recruited twenty college student volunteers to participate the experiments. None of the students has research experience in the topics related to image retargeting. The display device is a monitor with specification of $1680 \times 1050@65$ hz and 120 DPI. Participants sit 0.5 meters away from the monitor with normal indoor illumination. The test takes about one hour per subject.

Results and Analysis. After conducting the experiments, the average time cost per page and recognition accuracy for each retargeting methods are calculated. The results are summarized in Tables 1 and 2.

The study on recognition accuracy is to ensure that the

Method	SL	SC	ISC	SOAR
Accuracy (%)	94.76±5.6	94.25±4.1	94.95±4.1	96.15±3.8

Table 1. Average searching accuracies. SL, SC, ISC stand for Scaling, Seam Carving, Improved Seam Carving, respectively.

Method	SL	SC	ISC	SOAR
Time Cost (sec.)	13.53±3.0	14.26±3.1	13.05±3.0	10.98±2.7

Table 2. Average time costs per page of the user study. SL, SC, ISC are the same as in Table 1.

thumbnail generated by our method does not bring additional difficulty to image understanding compared with other methods. Intuitively, given enough time, a user can eventually find a given target accurately. Consequently, we are more interested in the time cost, while emphasize less on the recognition accuracy. By checking Table 1, we can see that all methods have achieved high accuracies (94% and up) in the experiment, which confirms our intuition. Furthermore, the proposed SOAR algorithm outperforms other methods by a small margin.

The average time cost reflects the effectiveness of different retargeting methods in the context of thumbnail browsing. The results shown in Table 2 strongly support the proposed SOAR algorithm. To study the significance of the conclusion, we also give the box plot in Fig. 9. For a rigorous evaluation, we have conducted one-way ANOVA analysis on the time cost data of four methods. The F-value is 4.06 and p-value is 0.0101, which indicates that the four methods are significantly different. Furthermore, multiple comparison test using the information from ANOVA has been performed to distinguish if our method is significant in pairwise comparison to other methods. Results are given in Table 3. The 95% confidence intervals for all the differences ([0.58,4.51], [1.31,5.24], [0.13,4.04]) have rejected the hypothesis that the true differences are zero. In other words, the differences between our method and other three methods are significant at 0.05 level.

In summary, the analysis shows clearly the superiority of the proposed SOAR algorithm in the thumbnail browsing task. This promising result shall be attributed to all the ingredients we integrated in the algorithm, especially the scale and object aware saliency, and the combination of the elegant seam carving scheme and the smoothness-preservation TPS warping. In addition, it shows that the other three methods performs similarly.

5.2. Qualitative Experiments

Although quantitative evaluation is critical in our study, where the focus is thumbnail browsing, qualitative study provides important information toward additional understanding of the proposed algorithm. Fig. 7 has some examples generated by different approaches used in our quantita-
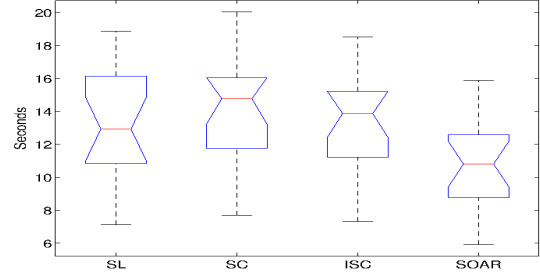

Figure 9. Box plots of time cost (in seconds). From left to right: SL, SC, ISC, SOAR.

Comparison	Lower	Difference in means	Upper
SL and SOAR	0.58	2.55	4.51
SC and SOAR	1.31	3.28	5.24
ISC and SOAR	0.13	2.07	4.04

Table 3. Tukey's least significant difference(LSD) test result for multiple comparison. 95% confidence intervals are [Lower, Upper]. Details are in Sec. 5.1.

tive study, and more examples can be found in the supplementary material. It is hard to draw a firm conclusion from visual inspection. That said, from the results, we can see that in general our method performs the best regarding the (foreground) object size in the thumbnails. Intuitively, this advantage highlights the target object and therefore helps the user in the visual searching task. The quantitative analysis in Section 5.1 has confirmed this intuition.

While we are interested mainly in tiny thumbnails, it is also worth investigating how well the proposed method performs when the size change is less drastic. Recently, Rubinstein et al. [17] released a benchmark dataset, RetargetMe, together with the results from many state-of-the-art image retargeting methods [10, 16, 18, 19, 22, 24]. Taking benefit of their work, we run the proposed SOAR algorithm on the dataset for qualitative inspection. For a fair comparison, our method is tuned to output in the scales used by those methods, which is about 75% of the original image sizes. Some example results are shown in Fig. 10 and more can be found in the supplementary material.

By checking the results, we feel it is hard to visually pick out a method that is better than all others. In particular, the results from our SOAR algorithm look in general similar to the results from other methods. The best and worst results for different input images often come from different retargeting methods. In fact, according to the user study in [17], manually cropping generates results in which users like the most. We conjecture that, as far as the targeting size is comparable to the original image size, (manually) cropping and scaling would be the best choices since they tend to keep the image structure untouched and smooth. However, this may not be true for extremely small targeting sizes, such as

Figure 10. Example results on the RetargetMe dataset. From left to right: input image, multi-operator media retargeting, non-uniform scaling with bi-cubic interpolation, optimized scale-and-stretch, non-homogeneous content-driven video-retargeting, SOAR (proposed). All results except SOAR are from [17].

Figure 11. Failure examples of the proposed method.

those used for thumbnails.

5.3. Limitations

There are still limitations in our work. In some images, regions in our result images are bend: Fig. 11 shows the examples. One scenario is when the saliency distribution is scattered such that the main theme of the image is not so clear. Another scenario is when the object is too big that seam carving (no matter which version) will definitely break the structures hence affect our final results.

An interesting fact is that the human observer can still correctly recognize and select those images due to the powerful human visual system. It is worth exploring in the future study when such image distortion significantly hurt the recognition rate or searching time costs.

6. Conclusion

In this work, we proposed a scale and object aware image retargeting (SOAR) method. The method aims at generating effective retargeted images for using in the context of thumbnail browsing. For this purpose, we integrate several new techniques, including scale dependent saliency, objectness and cyclic seam carving, into a TPS-based continuous warping model. The proposed method is evaluated both quantitatively and qualitatively. Our method demonstrates promising performances in comparison with state-of-the-art image retargeting algorithms.

References

[1] B. Alexe, T. Deselaers, and V. Ferrari. What is an object? *CVPR*, 2010. 2, 3, 4

[2] S. Avidan and A. Shamir. Seam carving for content-aware image resizing. *SIGGRAPH*, 2007. 2, 3, 4, 6

[3] F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *PAMI*, 1989. 5

[4] Y. Ding, J. Xiao, and J. Yu. Importance Filtering for Image Retargeting. *CVPR*, 2011. 2

[5] M. Everingham, L. Van-Gool, C. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2008. 6

[6] M. Grundmann, V. Kwatra, M. Han, and I. Essa. Discontinuous seam-carving for video retargeting. *CVPR*, 2010. 2

[7] Y. Guo, F. Liu, J. Shi, Z. Zhou, and M. Gleicher. Image retargeting using mesh parametrization. *IEEE Trans. on multimedia*, 2009. 2

[8] L. Itti, C. Koch, E. Niebur. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. *PAMI*, 1998. 4

[9] T. Judd, F. Durand, and T. Torralba Fixations on low-resolution images. *Journal of Vision*, 2011. 1

[10] Z. Karni, D. Freedman, and C. Gotsman. Energy-based image deformation. *C.G.F*, 2009. 2, 7

[11] F. Liu and M. Gleicher. Automatic image retargeting with fisheye-view warping. *UIST*, 2005. 2

[12] J. L. Mannos and D. J. Sakrison. The effects of a visual fidelity criterion of the encoding of images. *IEEE Trans. on Information Theory*, 1974. 1, 2, 3

[13] A. Mansfield, P. Gehler, L. Van Gool, and C. Rother. Scene carving: Scene consistent image retargeting. *ECCV*, 2010. 2

[14] F. L. van Nes and M. A. Bouman. Spatial modulation transfer in the human eye. *JOSA*, 1967. 1, 2, 3, 4

[15] E. Peli. Contrast sensitivity function and image discrimination. *JOSA*, 2001. 2, 3

[16] Y. Pritch, E. Kav-Venaki, and S. Peleg. Shift-map image editing. *ICCV*, 2009. 7

[17] M. Rubinstein, D. Guterrez, O. Sorkine, A. Shamir. A Comparative Study of Image Retargeting. *SIGGRAPH Asia*, 2010. 1, 2, 7, 8

[18] M. Rubinstein, A. Shamir, and S. Avidan. Improved Seam Carving for Video Retargeting. *SIGGRAPH*, 2008. 2, 6, 7

[19] M. Rubinstein, A. Shamir, and S. Avidan. Multi-operator media retargeting. *ACM Trans. on Graphics*, 2009. 2, 7

[20] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani. Summarizing visual data using bidirectional similarity. *CVPR*, 2008. 2, 5

[21] B. Suh, H. Ling, B.B. Bederson, and D.W. Jacobs. Automatic Thumbnail Cropping and Its Effectiveness. *UIST*, 2003. 1

[22] Y.-S. Wang, C.-L. Tai, O. Sorkine, and T.-Y. Lee. Optimized scale-and-stretch for image resizing. *ACM Trans. on Graphics*, 2008. 7

[23] Y.Z. Wang and S.C. Zhu. Perceptual Scale Space and Its Applications. *IJCV*, 80(1):143–165, 2008. 3

[24] L. Wolf, M. Guttmann, and D. Cohen-Or. Non-homogeneous content-driven video-retargeting. *ICCV*, 2007. 2, 7