This is a preliminary version of an article published by
Johannes Kiess, Benjamin Guthier, Stephan Kopf, and Wolfgang Effelsberg.
SeamCrop for Image Retargeting.
Proc. of IS&T/SPIE Electronic Imaging (EI) on Multimedia on Mobile Devices, Vol. 8304, 2012.

SeamCrop for Image Retargeting

Johannes Kiess, Benjamin Guthier, Stephan Kopf and Wolfgang Effelsberg
Department of Computer Science IV
University of Mannheim, Germany

Figure 1: Original image and examples of four image retargeting operators including our approach. The crop example was done manually.

ABSTRACT

In this paper, we present a novel approach for the adaptation of large images to small display sizes. As a recent study suggests, most viewers prefer the loss of content over the insertion of deformations in the retargeting process.[1] Therefore, we combine the two image retargeting operators *seam carving* and *cropping* in order to resize an image without manipulating the important objects in an image at all. First, seams are removed carefully until a dynamic energy threshold is reached to prevent the creation of visible artifacts. Then, a cropping window is selected in the image that has the smallest possible window size without having the removed energy rise above a second dynamic threshold. As the number of removed seams and the size of the cropping window are not fix, the process is repeated iteratively until the target size is reached. Our results show that by using this method, more important content of an image can be included in the cropping window than in normal cropping. The "squeezing" of objects which might occur in approaches based on warping or scaling is also prevented.

Keywords: image retargeting, image resizing, seam carving, cropping

1. INTRODUCTION

With the increasing popularity of smartphones and other small handheld multimedia devices, the retargeting of media to different display sizes and aspect ratios is currently a popular research topic. Most state-of-the-art image resizing techniques try to show as much content as possible from the original image by modifying the shape of the objects in the image. A detailed, comparative study in this field has indicated that most viewers prefer the loss of content over the insertion of deformation to an image.[1] This can be seen from the fact that manual cropping has ranked among the top three retargeting methods in nearly every category of the study. Therefore, the authors suggest that the automatic search for a cropping window is still a highly relevant research topic.

Further author information: (Send correspondence to Johannes Kiess.)
Johannes Kiess: E-mail: kiess@pi4.informatik.uni-mannheim.de, Telephone: +49 621/181-2609
Benjamin Guthier: E-mail: guthier@pi4.informatik.uni-mannheim.de, Telephone: +49 621/181-2605
Stephan Kopf: E-mail: kopf@pi4.informatik.uni-mannheim.de, Telephone: +49 621/181-2613
Wolfgang Effelsberg: E-mail: effelsberg@pi4.informatik.uni-mannheim.de, Telephone: +49 621/181-2600

Figure 2. Example of seam carving. The red seams are removed to get the adapted image.

Our idea is to combine seam carving with cropping in order to get more image content into the cropping window but at the same time limiting visual artifacts. Non-uniform scaling is not included in our approach because viewers may be bothered by stretching artifacts that can be introduced by this operation (see Figure 1).

In this paper, we present a novel algorithm for image retargeting. Our main contributions are:

- We combine two simple operators into an effective image retargeting algorithm.
- A novel *dynamic threshold* is presented that finds the switching point between seam carving and cropping.
- We evaluate our algorithm in detail and compare it to other state of the art image retargeting operators.

The rest of this paper is structured as follows. Chapter 2 presents related work and the seam carving algorithm. The details of our new algorithm are explained in Chapter 3. An evaluation compares our approach to the current state of the art techniques in Chapter 4. Chapter 5 concludes the paper.

2. RELATED WORK

Image retargeting can be either done in a *discrete* way by the removal of content (e.g. cropping) or in a *continuous* way by merging pixels (e.g. scaling).[2]

Discrete methods try to identify important regions in an image in order to discard pixels from the other regions. This approach can be used to find the optimal position of a cropping window.[3] In previous work, we have developed a media retargeting operator to visualize Web pages and images on handheld mobile devices.[4] Our server-based system identifies and crops border regions which do not contain relevant semantic content. In a second step, attention objects like faces or text regions are identified. We propose a semi-automatic approach which allows users to correct obvious errors in the adapted images. A software tool allows users to validate, add, delete, or modify all automatically extracted data.

Continuous methods retarget an image by rearranging and merging the pixels in the source image to fit into the target size. Current techniques attach a mesh to an image and then deform it to fit to the target size. These meshes are either composed of quads[5] or work on individual pixels.[6] *Hybrid methods* combine discrete and continuous operators in order to overcome their individual limitations. *Rubinstein et al.* presented a combination of cropping, scaling and seam carving where the optimal order of operators is found in a multi-dimensional resizing-space.[7]

Image Resizing based on Seam Carving

Seam Carving introduced by *Shamir and Avidan* is a discrete technique that resizes an image by removing connected paths of pixels called *seams*.[8] Each removed seam reduces the size of an image by one. Vertical seams reduce the width (see Figure 2), horizontal seams the height.

In the following, only the reduction with vertical seams is explained as the reduction with horizontal seams is similar.

A vertical seam is a path of pixels from top to bottom meeting two constraints: Exactly one pixel is selected in each row, and the seams are vertically connected (8-pixel-neighborhood). An energy function is used to assign each pixel a relevance value. The cost of a seam is defined as the summed up energy of all pixels on the seam. By traversing the image from top to bottom, the cumulated minimum energy values for all possible seam combinations are calculated. In order to get the optimal seam, the algorithm determines the minimum value in the last row of the cumulated energy map and tracks the path back to the top.

A disadvantage of seam carving is the fact that straight lines become curved and cause clearly noticeable errors. In previous work, we have proposed a technique to reduce these effects by modifying the energy in the local neighborhood of the intersection point of a seam and a straight line to prevent other seams from removing adjacent line pixels.[9]

3. THE SEAMCROP ALGORITHM

Our SeamCrop algorithm is designed with the preference of content loss over the insertion of deformation in mind.[1] It combines the two image retargeting operators *seam carving* and *cropping*. Several hybrid techniques use *scaling* instead of cropping or as an additional third operator. We have chosen to not include it in the algorithm because it may cause squeezing artifacts when the aspect ratio is changed. Scaling is used however to uniformly scale down very large images before they are adapted to a different aspect ratio.

In the following, we assume that the width of an image is reduced. The reduction of the height can be done analogously. First, a saliency map of the source image is calculated. Based on this map, vertical seams are taken away from the image until the target size is reached or more than α percent of the energy have been removed. If the target size has not been attained, the algorithm switches to cropping. An optimal cropping window with the smallest possible size is searched in the image until the target size is reached or more than β percent of the energy has been cropped. As the switching point between seam carving and cropping is crucial to the quality of the result, we use thresholds that dynamically adapt to the image content. The algorithm is repeated until the target image size is attained. Figure 3 gives an overview of the workflow.

Energy function

The energy function we use in our approach is a contrast-based saliency map. It is computed on a copy of the original image that was converted into the LUV color space. This color space has the advantage of being perceptually uniform. The perceived distance between two colors can thus simply be calculated using the Euclidean distance between the LUV coordinates. For each image pixel, the color distance to all neighbor pixels in a rectangular window is computed. It is weighted by a Gaussian function to create a value between 0 and 1. The saliency value for the considered pixel is then obtained by averaging the color distance over the entire window. Areas with little color variation therefore have lower energy than those with high contrasts.

Marking seams

For the removal of seams, we mark the paths of the seams in the image and give the included pixels a very high energy value to make them unusable for further seams. Because of this, it may happen that a seam different than the energy optimal one is chosen, since the seam may not be able to cross a previously marked path. In practice, this is a major advantage because the seams are more spread over the image. Our intention is to remove seams very carefully to prevent visible artifacts. Only if the summed energy of the marked seams hits the threshold and fulfills its requirement or if the target size is reached, the seams are actually removed.

Dynamic threshold

The dynamic threshold is composed of two parts: First, seams may only be removed up to a certain percentage α of the total energy of an image. The second part is the requirement that the seams discarding about α percent of the total energy must remove at least the same percentage of image pixels. This is based on the idea that an operator is only effective when it removes a higher percentage of pixels than it removes energy, i.e. relevant content. Only if the requirement is fulfilled, the marked seams are actually removed. If necessary, the algorithm

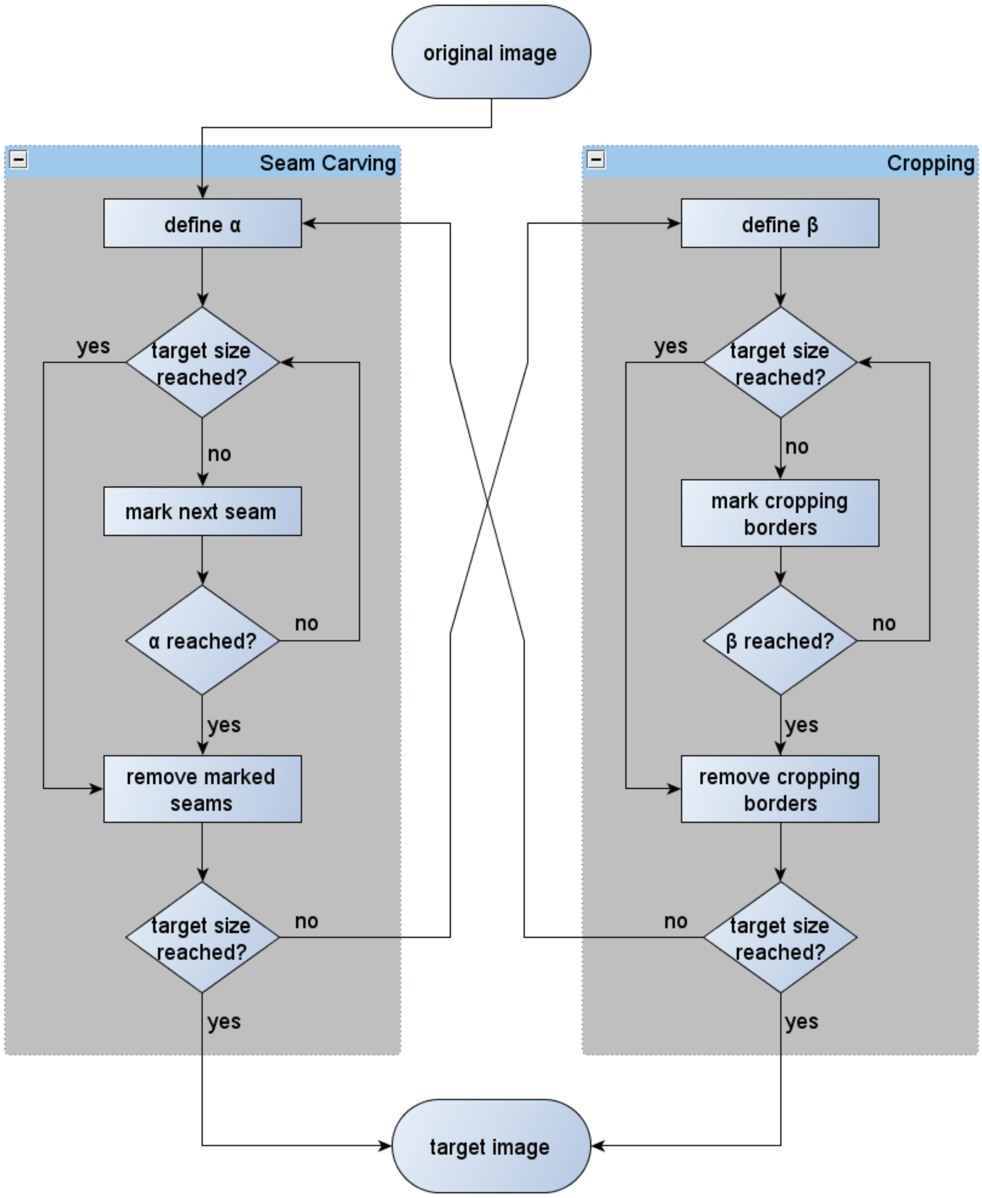

Figure 3. The operator switches between seam carving and cropping based on a dynamic threshold. This is repeated until the target size is reached.

then switches to the cropping operator.

Positioning of the cropping window

As the width of the image has been modified by seam carving, a new saliency map has to be computed first. The saliency map is projected to a one-dimensional energy array by summing up the saliency values for an entire column. After that, we use a brute-force 1D search to find the optimal window position with the smallest possible size that discards less than β percent of the energy of the image. β should be chosen higher than α as the intention is to remove seams only carefully and cropping does not produce artifacts like bend or broken objects.

Cropping starts with the width of the cropping window being one column smaller than the width of the image. The total energy of all possible positions is calculated and the position with the highest energy inside the window is picked. Then, if less than β percent of the energy is discarded, the step is repeated again with the window size reduced by one. This is continued until the removed energy exceeds β percent. The last size and position with discarded energy below the threshold is chosen as the optimal cropping window. Like the marked seams, the cropping window has to fulfill a requirement in order to be executed. The algorithm stops when β percent of the total *energy* of the image would be discarded. At the same time, the crop has to remove at least β percent of the *width*. If the requirement is not met, the crop is skipped. After that, the algorithm switches the operator again if necessary.

It may happen that both operators are skipped repeatedly due to the thresholds. In this case, the requirements are relaxed to allow the removal of lesser width than energy. In practice, this is very unlikely: If high energy regions at the borders would prevent a crop, the energy in the center would be lower allowing the removal of seams. When there is high energy in the middle, a crop is possible.

4. EVALUATION

We performed an evaluation in order to compare our algorithm to current state-of-the-art image retargeting techniques. The comparative study[1] and the images from the benchmark data set* provided by *Rubinstein et al.* were used as basis for our comparison. Our results of the full benchmark are available on the web†. A manually chosen crop, *multi-operator retargeting*[7] and *streaming video*[6] were the best techniques in nearly all test cases of the study. Therefore, the authors suggest that it is sufficient to demonstrate that a new algorithm outperforms these three.

The evaluation was a no-reference comparison where the original image was not shown to the subjects. This simulates the real-world situation in which a user only gets to see the retargeted result. Nine image sets consisting of four resized images each were evaluated by the participants. Each set consisted of retargeted results of the same source image calculated by the three methods mentioned above and our new algorithm. The nine image sets were randomly chosen by our evaluation software out of the images provided in the benchmark data set. We did not include a reduction in height in the evaluation. This led to 71 possible image sets out of 80.

In an image set, the participants could rank the results by giving 1 to 4 points to each resized image with 1 being the best and 4 being the worst. Additionally, they were asked a number of questions.

A total of 16 subjects (14 male, 2 female) took part in the evaluation. One half were volunteers, the other half were colleagues from our department. A total 144 image sets was evaluated. In each set, the images are reduced to either 75% or 50% of their original width depending on the values used in the benchmark data set. As parameters, we used $\alpha = 1\%$ as the seam carving threshold, and $\beta = 15\%$ for cropping.

Analysis and Discussion

Based on the evaluation results, the mean rank and the standard deviation σ for each method have been calculated. They are presented in Table 1.

	Manual Cropping	Multi-Operator	Streaming Video	Seam Crop
Mean rank	1.7153	2.4525	2.532	2.2116
σ	1.02	1.05	1.04	1.11

Table 1. Mean rank and standard deviation σ of the four evaluated techniques.

Similar to the comparative study,[1] the manually chosen crop was clearly the preferred technique by the participants. The gap between the other techniques is much smaller with our approach being ranked second. As the visual attention analysis of an image is one of the most challenging parts of image retargeting, the advantage of a manually picked cropping window over automatic detection is obvious. The human editor can identify all

*http://people.csail.mit.edu/mrub/retargetme

†http://ls.wim.uni-mannheim.de/de/pi4/research/projekte/retargeting/

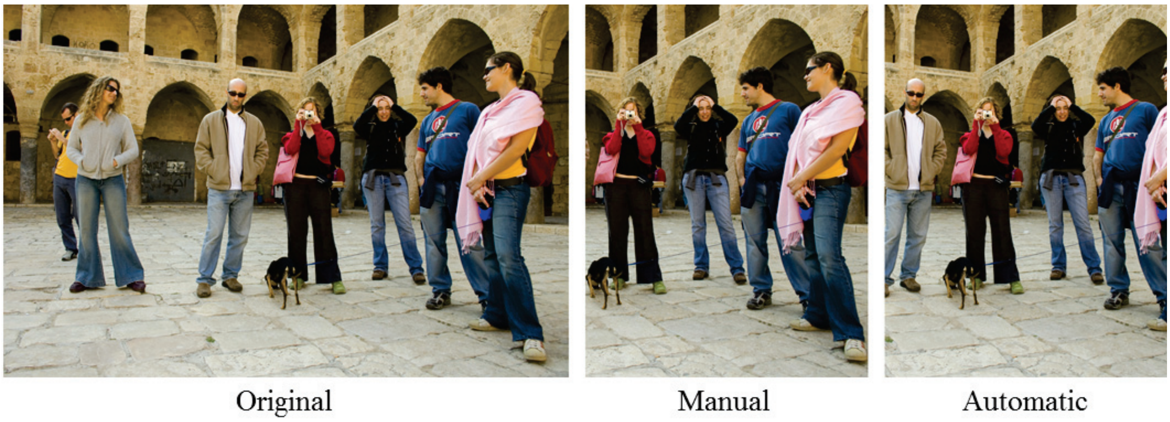

Figure 4. A comparison of manual and automatic cropping.

important objects and pick the perfect position by also taking the composition of an image into account (see Figure 4). Additionally, a crop does not cause any artifacts in the retargeting process. This may be the reasons why manual cropping ranks first in the no-reference comparison. It can be assumed that the quality of the other three techniques would benefit from a manually created saliency map.

After the evaluation, the participants were asked if there were things in the resized images that bothered them. Many subjects noticed squeezing artifacts in several sets. While all of them found the artifacts disturbing when persons were displayed, some found them acceptable for images depicting only buildings or nature without any important objects in the foreground. The participants were also asked if they generally prefer the loss of content or squeezing. All of the subjects stated that they prefer loss of content. However, some specifically added the restriction that prominent foreground objects should not be truncated. This supports our decision not to include scaling in the algorithm due to this effect. An example of an image set from the evaluation is shown in Figure 5.

A comparison between the single used operators and our approach demonstrates that it is effective to combine these two (see Figure 6). Table 2 gives an overview of the average seam carving to cropping ratio that occurred in the benchmark data set images.

	Seam Carving	Cropping	σ
To 75% width	20.23%	79.77%	9.62%
To 50% width	24.96%	75.12%	15.70%

Table 2. Mean ratio and standard deviation σ of the operators used in the benchmark.

When images are reduced do 75% of their width, about 20% or the pixels are removed with seam carving and 80% with cropping. In case of a 50% target width, seam carving is employed a little bit more. As cropping also cuts image content already manipulated by seam carving, this indicates that the seam carving operator is used only carefully like we intended.

5. CONCLUSIONS

We presented an algorithm that combines seam carving with cropping in order to resize images without squeezing artifacts. Our evaluation has shown that most viewers prefer images where content is discarded instead of deformations being introduced into the image. We consider our algorithm to be a promising step in this direction.

In future work, we would like to improve on the energy function as it still is an important quality factor of the algorithm. Further, we want to extend our approach to the retargeting of videos. We will also submit all retargeted results to the authors of the benchmark data set in order to make them available to the research community.

Figure 5. Example image set from the evaluation.

Figure 6. A comparison of the individual operators and SeamCrop.

Acknowledgements

The authors would like to thank Michael Rubinstein, Diego Gutierrez, Olga Sorkine and Ariel Shamir for their comparative study and for sharing their benchmark data set with the resarch community. Several images appearing in this paper were taken from it. The original images were provided by the following flickr.com users via the creative commons license: Wolfgang Staudt (Brasserie_L_Aficion), BurnAway (dancers), Ali Nishan (surfer), yanivba (Umdan) and stuckincustoms (battleship). The example from the evaluation (bicycle2) was published by Vojko Kalan on publicdomainpictures.net. Furthermore, the authors acknowledge the financial support granted by the *Deutsche Forschungsgemeinschaft* (DFG).

REFERENCES

[1] Rubinstein, M., Gutierrez, D., Sorkine, O., and Shamir, A., "A comparative study of image retargeting," in [*ACM SIGGRAPH Asia 2010 papers*], 160:1–160:10, ACM, New York, NY, USA (2010).

[2] Shamir, A. and Sorkine, O., "Visual media retargeting," in [*ACM SIGGRAPH ASIA 2009 Courses*], *SIGGRAPH ASIA '09*, 11:1–11:13, ACM, New York, NY, USA (2009).

[3] Suh, B., Ling, H., Bederson, B., and Jacobs, D., "Automatic thumbnail cropping and its effectiveness," in [*Proceedings of the 16th annual ACM symposium on User interface software and technology*], 95–104 (2003).

[4] Kopf, S., Guthier, B., Lemelson, H., and Effelsberg, W., "Adaptation of web pages and images for mobile applications," in [*Proceedings of IS&T/SPIE Electronic Imaging (EI) on Multimedia on Mobile Devices*], **7256**, 72560C:01 – 72560C:12 (January 2009).

[5] Wang, Y.-S., Tai, C.-L., Sorkine, O., and Lee, T.-Y., "Optimized scale-and-stretch for image resizing," *ACM Transactions on Graphics* **27**(5), 1–8 (2008).

[6] Krähenbühl, P., Lang, M., Hornung, A., and Gross, M., "A system for retargeting of streaming video," in [*ACM SIGGRAPH Asia 2009 papers*], 1–10, ACM, New York, NY, USA (2009).

[7] Rubinstein, M., Shamir, A., and Avidan, S., "Multi-operator media retargeting," *ACM Transactions on Graphics, SIGGRAPH 2009* **28**(3), 1–11 (2009).

[8] Shamir, A. and Avidan, S., "Seam carving for media retargeting," *Commun. ACM* **52**(1), 77–85 (2009).

[9] Kiess, J., Kopf, S., Guthier, B., and Effelsberg, W., "Seam carving with improved edge preservation," in [*Proceedings of IS&T/SPIE Electronic Imaging (EI) on Multimedia on Mobile Devices*], **7542(1)**, 75420G:01 – 75420G:11 (January 2010).