Technical note TN-1

Issued: 07/2008

Improvements to Seam Carving: A Content Aware Image Resizing Technique

Dipti Kapadia

Embedded Systems

Technical University of Eindhoven

Technische Universiteit Eindhoven

Department of Mathematics and Computer Science
and Department of Electrical Engineering

Master of Science Graduation Report

Improvements to Seam Carving:

An Image Resizing technique

By

Dipti Kapadia

Supervisors

Ir. Chris Damkat[1]

Prof. Dr. Ir. Gerard de Haan[1,2]

[1] Philips Research Eindhoven, Video Processing & Analysis

[2] Eindhoven University of Technology

Concerns:

Period of Work: 10-2007 to 06-2008

Notebooks: None

Authors' address Dipti Kapadia dipti.rajkotia@gmail.com

Title: Improvements to Seam Carving: A Content Aware Image Resizing Technique

Author(s): Dipti Kapadia
Embedded Systems
Technical University of Eindhoven

Reviewer(s): Ir. Chris Damkat
Prof. Dr. Ir. Gerard de Haan

Technical Note: TN-1

Additional Numbers:

Subcategory:

Project: ImagEn

Customer:

Keywords: image resizing, retargeting, seam carving, content-aware, visual saliency

Abstract: The recent developments in the field of display technologies have seen great diversity in display sizes. Same content is required to be displayed in different dimensions and aspect ratios for different devices. Standard scaling of an image to match the display size is not proving to be sufficient as it does not consider the image contents. Seam Carving is a *content aware* image resizing technique proposed by Ariel Shamir and Shai Avidan. It *retargets* images to a new size by gracefully carving out or inserting pixels in different parts of the image based on importance of pixels. In this thesis, the seam carving algorithm is analyzed for a variety of content, attention is then drawn to some of the image categories where it gives unacceptable results. A set of improvements by enhancing the so-called *energy model* which drives the seam carving algorithm are proposed. The new energy model reflects the visual saliency of image parts better than the original energy model. The improvements are evaluated in a subjective analysis. Also considering the popularity of wide screen displays, a retargeting algorithm for aspect ratio conversion using seam carving is proposed. Further on an evaluation of the performance of the proposed algorithm for aspect ratio change compared to currently known method of panoramic stretch is carried out.

Conclusions: Analysis of Seam carving, the proposed improvements to it and results of subjective evaluation leads to the following conclusions

- Seam carving in general works best for images with multiple subjects separated by an uninteresting background.
- The results of seam carving can be significantly improved by better modeling of visual saliency in the energy function.
- Higher level cues like face detection greatly improves the results of retargeting.
- Taking into account, the history of previously removed seams, continuous carving from the same region of image can be prevented leading to reduced artifacts.
- Seam carving is a good candidate for retargeting for aspect ratio conversion as it considers the content while achieving target aspect ratio.

Preface

This work was performed at Philips Research in the Video Processing & Analysis department, under supervision of Chris Damkat, and Gerard de Haan. In this department there is much experience with video processing algorithms for television. Besides it the group contributes in image and video analysis for medical care, safety and comfort.

To enjoy digital media content like video and still images on a variety of consumer electronic devices, digital content needs to be effectively adapted. Hence *seam carving* algorithm as a technique for effective resizing of images is studied and analyzed. With the goal to see its applicability to videos, this algorithm is improved by better modeling the visual saliency. Based on seam carving, an algorithm for aspect ratio conversion for adapting an image to a display of different aspect ratio is suggested. Subjective analysis of the improved algorithm helps to visualize the potential of seam carving as a resizing technique. Also the results from the algorithm of aspect ratio conversion using seam carving and its paired comparison with panaromic stretch reflects on its possible use for aspect ratio conversion before retargeting for target display.

The author would like to thank Chris Damkat, Paul Hofmann and Gerard de Haan for their support and supervision. Furthermore thanks go to the department and fellow students for a wonderful time at Philips Research.

Dipti Kapadia,

July 2008.

Contents

Section 1

Introduction

The recent developments in the field of display technologies have seen great diversity in display sizes. Displays vary from low resolution hand-held devices to high definition wide-screen TVs. Increasingly; our computing and communications infrastructure is evolving to support images and video into these ever expanding set of potential displays. Visual content is becoming more important for sharing, expressing, and exchanging information on devices such as, cell phones and hand-held PCs [Liu et al.2003], PDAs with video capabilities, home-networked media appliances, and heads up informational displays in automobiles and helmets. Same content is required to be displayed in different dimensions and aspect ratio for different devices. Standard image processing methods of scaling and cropping are not proving to be sufficient.

With the popularity of wide-screen TVs, better solutions which could effectively display videos on displays other than originally intended is needed. In Philips TV-sets, we see the technique of *Panoramic Stretch*, where the boundaries of the 4:3 aspect ratio videos are stretched to take up the wider screen. However such an anisotropic stretch is not the optimal solution as it does not adapt to the content. Better methods are desired enabling effective resizing for variety of displays especially wide screen.

1.1 Retargeting

All this points to the need of algorithms that could adapt images to displays different than originally intended for. Naive image resizing techniques are linear scaling or cropping. Scaling, uniformly resizes the image giving same importance to all regions of an image. When downscaling images, the size of important features is greatly reduced, this leads to degradation of image quality due to loss of details. For aspect ratio change, using a different scaling factor for each dimension would result into an anisotropic squish or stretch. Cropping on the other hand, discards the region outside the cropping window disturbing the image composition. Hence effective adaption of images considering the image content is needed. Such an intelligent adaption is called *Image Retargeting* or *Video Retargeting* if we are considering video. Figure 1.1 shows difference between resizing and retargeting. Scaling looses the details in regions of interest (birds) whereas retargeting preserves the birds in their original form. Consequently, it is essential for any retargeting technique to identify important features in an image and then use some kind of warping to adapt the image to target size. Computer vision techniques can be used to abstract image information and extract image features. However categorizing the image content automatically is challenging and hence the problem of retargeting is an interesting research topic, which is gaining attention.

Figure 1.1: *On the top is source image (500 x 333), bottom-left shows scaling to (320 x 213) and bottom-right shows retargeting to target size (320 x 213)*

1.2 Retargeting literature

Recently the subject of retargeting is gaining attention and there have been a number of contributions to this field. A few researchers have explored automating image retargeting that seeks to change the size of images while keeping the important features intact. Different retargeting methods identify important image features either based on low level visual saliency or high level image understanding through tools like a face detector, thereafter they use some form of cropping or scaling to resize the image to the target size.

Firstly some techniques based on cropping. Suh *et al.* [13]proposed automatic cropping of an image to a thumbnail, based on visual saliency and face detection. Their evaluation confirms the superiority of automatic cropping over naive scaling for thumbnail generation. However their method cannot handle images with multiple important features. Chen *et al.* [3] had a similar approach to adapt images to mobile devices. Extending on the same line, Liu *et al.* [9] catered specifically to the problem of retargeting images with multiple important regions to mobile device. Their solution was to trade time for space, displaying the multiple objects serially one after the other. The user scrolls between pages to view them. However the user cannot see multiple objects at the same time on the screen. Santella *et al.* [11] proposed gaze-based interaction for semiautomatic photo cropping. With a little user input their method is based on recording the eye movement to identify important image content, an offline learning phase and then automatically generate crops of any size or aspect ratio.

Feng Liu and Michael Gleicher [7] proposed data dependent scaling. Their method called Fisheye-view warping relies on identifying important aspects of source images using low-level saliency and high level object recognition and trying to preserve them at the cost of less important parts of image. Retargeting the image preserves the region of interest and produces a warped image of less important areas. Their technique results into distorted unimportant areas, making the viewer aware of resizing compared to cropping where it would be discarded altogether.

Setlur *et al.* [12] proposed non photorealistic algorithm for automatically retargeting images. Their system is based on segmenting source image into regions and then using an importance map to select regions of importance. Important regions from the image are removed and holes are filled up using inpainting (image interpolation), which forms the background of the image. Next, this background image is resized to the target size and the foreground objects are reinserted at maximum scale possible. The method relies heavily on their capability to solve the separation of the foreground from the background.

Ariel Shamir and Shai Avidan propose *Seam Carving* algorithm for image retargeting [2]. The algorithm is based on identifying regions of interest based on image edge energies. Unlike other image retargeting methods, it is not based on scaling or cropping but instead uses seams (continuous paths of pixels) for retargeting. Moreover their method also works for image enlargement. As seam-based editing has no affect on the rest of image, effects of retargeting are localized compared to other retargeting methods based on scaling.

Next to images, there is also growing interest in adapting videos for different displays. The problem of video retargeting is more challenging than image retargeting. Due to motion and camera movement, determining important aspects of video is difficult. Moreover, maintaining temporal consistency (temporally adjacent neighbors appear visually similar) when regions of interest change dynamically is demanding.

In answer to the challenge of video retargeting, Feng Liu and Michael Gleicher proposed Automating pan and scan method [8] wherein they combine cropping with scaling. Their method finds the best cropping window based on image and motion saliency and scales it to fit target display. The cropping window can be moved during a shot to introduce virtual pans and cuts. Not long ago Wolf *et al.* [6] proposed non-homogeneous content driven video retargeting which uses non uniform global warping. It consists of two stages. First, the frame is analyzed to detect the importance of each region in the frame based on local saliency, motion detection and object detectors. Then, a transformation that respects the analysis shrinks less important regions more than important ones. Recently Shamir and Avidan extended the seam carving algorithm for video retargeting [10]. Instead of a 1D seam they define a 2D seam surface in a 3D video space-time cube. The intersection of surface with each frame defines one seam. The manifold seam surface allows the seam to change adaptively over time, maintaining temporal coherence. They propose a graph-cut based approach to find 2D seam surface from the 3D video space-time cube.

Moving back to the TV application, where standard 4:3 broadcast video needs to be adapted for wider screens, Philips has introduced *Panoramic Stretch*. Panoramic Stretch is basically a non linear scaling depending on the horizontal position of the pixel. It provides a subtle stretch towards the sides filling up the wide screen, but the region in the center, which most of the time is the region of importance, remains undistorted. This method can also be considered a form of retargeting, however with region of importance always being the center.

1.3 Outline of the Thesis

Looking at the growing diversity in display dimensions, there is a great need for retargeting technique for video and still images. Especially we are interested in the opportunities that retargeting might give for conversion of a video to a wider screen. The unique approach of Avidan and Shamir [2], the remarkable results of retargeting still images shown by them and its popularity on the web, led us to further analyze seam carving. In this report firstly the seam carving algorithm is narrated briefly (Section 2). With the idea to extend the seam carving algorithm for video, we started with evaluating it for variety of image categories. Although seam carving can give remarkable results on images, there are some image categories where it fails, giving undesired results. We suggest a set of improvements (Section 3) by redefining the so-called energy function to better reflect visual saliency and human perception. Also in order to support a fully automatic carving without compromising on retargeting results, in Section 4 we investigate the stopping criteria guiding us when to stop carving. Coming back to problem of catering to variety of display dimensions we propose a new technique for aspect ratio change using seam carving (Section 5). Finally an evaluation is done in Section 6 comparing the improved algorithm with the seam carving algorithm originally proposed[2]. Due to lack of time we could not extend this algorithm for video but considering the interest of Philips for wide screen upscaling, we do evaluate seam carving for aspect ratio conversion on still images, comparing it with panoramic stretch (currently used in Philips TVs).

Section 2

Seam Carving

A very novel approach to content aware resizing called Seam Carving [2] has been proposed by Shai Avidan and Ariel Shamir. Here retargeting is achieved by reducing the width (or the height) of the image by one pixel at a time. The image dimension is reduced or enlarged by gracefully carving out or inserting pixels along the height or width of the image. Seam carving is based on finding least noticeable connected seams of pixels. While reducing image height or width, removing such a continuous seam would preserve the visual perception by not distorting regions of interest. Similarly while enlarging, duplicating those less noticeable seams will not much change the regions of interest, maintaining the original view of the image.

In this section we start with explaining the basic seam carving algorithm, exhibit and discuss some retargeting results and point out some of its limitations.

2.1 Algorithm

Assume we want to reduce the width of the image. The basic idea of seam carving is to remove unnoticeable pixels which blend with their surroundings. Hence the algorithm defines an energy function for each pixel which basically reflects how visually important the pixel is. It appears that the algorithm depends on luminance channel to define the energy model. So let I be a grey-scale conversion of $n \times m$ image with n rows and m columns, and $e(x, y)$ denote the energy of pixel at horizontal position x and vertical position y. Then seam carving defines the energy function as in equation 2.1, which is the sum of absolute gradient in horizontal and vertical direction (L_1-*norm*).

$$e\left(x,y\right) = \left|\partial I\left(x,y\right)/\partial x\right| + \left|\partial I\left(x,y\right)/\partial y\right| \tag{2.1}$$

Figure 2.1 visualizes this energy function. Having defined the energy model (Figure 2.1), clearly the low energy pixels are the darker pixels or the pixels which can be removed without being noticed. However randomly removing pixels from the image could distort its shape. Thus equal number of pixels are required to be removed from every row to preserve the rectangular shape. Taking all this in account, a seam is defined as connected path of pixels on a single image from top to bottom (or left to right).

More formally a vertical seam is defined as:

$$s^x = \{s_i^x\}_{i=1}^n = \{(x(i), i)\}_{i=1}^n \, s.t. \forall i, |x(i) - x(i-1)| \leq 1 \tag{2.2}$$

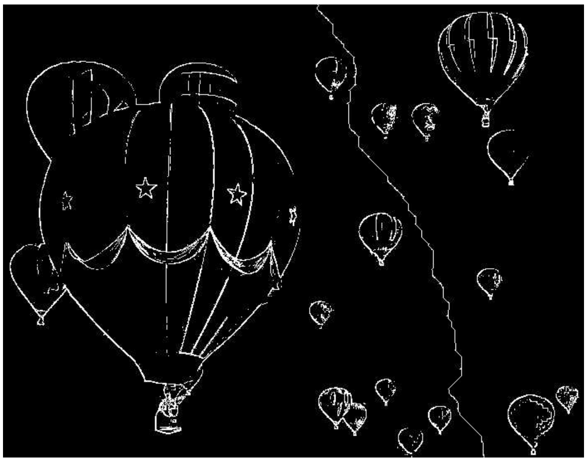

Figure 2.1: *Energy map of 'Balloon' image based on energy function of equation* 2.1.

Figure 2.2: *Vertical seams in Balloon image.*

where x is a mapping $x : [1, ..., n] \longrightarrow [1, ..., m]$ from rows to columns. 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. Figure 2.2 shows the vertical seams.

Similarly, if y is a mapping $y : [1, .., m] \longrightarrow [1, .., n]$, then a horizontal seam is defined as :

$$s^y = \left\{ s_j^y \right\}_{j=1}^{m} = \{(j, y(j))\}_{j=1}^{m} \, s.t. \forall j, |y(j) - y(j-1)| \leq 1 \tag{2.3}$$

With the goal of removing pixels with low energy, we are interested in seams with low energy. For that purpose, we first define cost of seam or energy of seam.

Given an energy function e, cost of seam can be defined as:

$$E(s) = \sum_{i=1}^{n} e\left(I\left(s_i\right)\right) \tag{2.4}$$

Optimal seam s^* is then the seam which minimizes the seam cost. Mathematically s^* is defined in equation 2.5.

$$s^* = min_s E(s) = min_s \sum_{i=1}^{n} e(I(s_i)) \tag{2.5}$$

Dynamic programming with bottom-up approach is used to find optimal seam. Energy along the height of the image from top to bottom is integrated. Since the seam has to be connected we find the cumulative minimum energy at each pixel combining it with one of the three neighbors in previous row. In first step, the image is scanned from second to last row and for every pixel position (i, j), cumulative minimum energy M is computed, using the equation 2.6, where $e(i, j)$ is the energy of pixel at (i, j).

$$M(i,j) = e(i,j) + min(M(i-1,j-1), M(i-1,j), M(i-1,j+1)) \tag{2.6}$$

At the end of this process, the minimum value of the last row in M will indicate the end of the minimal connected vertical seam. From this minimum entry, the path of the optimal seam can be traced backwards. Figure 2.3 shows the cumulative minimum energy map for the energy map in Figure 2.1.

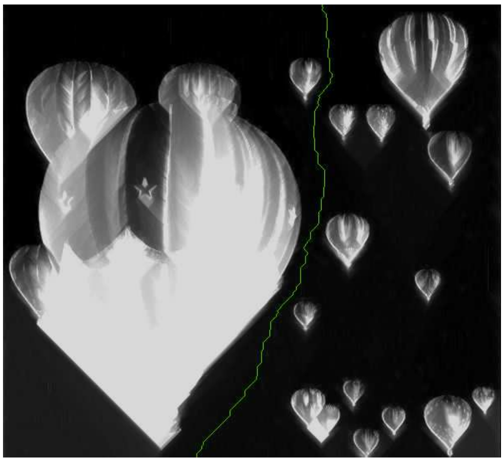

Figure 2.3: *Cumulative minimum energy map showing how energy is integrated over the height from top to bottom. Also an optimal seam consisting of low energy pixels is visible in green.*

The definition for M for horizontal seams is the same with columns substituted as rows:

$$M(i,j) = e(i,j) + min(M(i-1,j-1), M(i,j-1), M(i+1,j-1)) \tag{2.7}$$

Now that the optimal vertical seam is found, the $n \times m$ image is retargeted to a new size of $n \times (m-1)$ by removing (carving) the optimal vertical seam. By repeatedly removing seams in this manner a new size can be targeted. To enlarge an image, the optimal seam is duplicated by averaging the seam pixels with their left and right neighbors. However, when enlarging by more than one column, the same seam will be selected again and again, which will cause a stretching effect. Hence for enlarging a set of seams are analyzed before the process of duplicating them.

2.2 Results of Seam Carving

Seam carving detects low energy pixels, parts of the image where resizing is least noticeable and resizes in that area. While downsizing it removes least noticeable seams of pixels and while

Figure 2.4: *Retargeted 'Balloon' image. On the left the width is reduced by removing low energy seams. The right image shows the same image enlarged in width by duplicating low energy seams*

enlarging it duplicates the seams of unnoticeable pixels. Similar to the effect of removing a row or a column from an image, removing a seam only has a local effect on the image as all the pixels are shifted left or up to compensate for the missing seam. Consequently, the retargeted image differs only along the path of the seam. In this process, the energy function plays a key role as it indicates the areas to be preserved. Therefore having the right energy model identifying the image features is crucial to seam carving. As shown in figure 2.4, when multiple objects are scattered in the image, retargeting is aimed by bringing those objects nearer in case of downsizing. While enlarging, seam carving operator guards the region of interest from being stretched and seams are duplicated in regions identified as unnoticeable or less important which is often the background.

Moreover, other than for retargeting, seam carving can be used for specific image processing like object removal and content amplification. For object removal, the objects to be removed are marked as low energy and then retargeted to a smaller size. On the other hand content amplification can be achieved by retargeting an image to a smaller size or removing uninteresting parts in the image, and then cascading these results with linear upscaling.

2.3 Limitations of Seam Carving

Retargeting is based on categorizing an image, so as to preserve important areas while relinquishing other areas of image. However when a subject occupies an entire frame, there is no less important area to be sacrificed. Consequently any retargeting technique fails. Similarly, if the image is too condensed, all parts of image are equally important, retargeting cannot give satisfactory results. In such cases, it is best to switch to resizing.

Seam carving as proposed originally, works best for images with multiple subjects separated by low energy background. However there are some image categories where it fails to identify interesting subjects giving undesired results. When multiple regions of interest are separated by high frequency textured area, seam carving misinterprets the texture as region of interest and seams are wrongly selected in regions which should be preserved. Moreover, we noticed that, as a result of seam carving, objects in the retargeted image suffer uneven distortion. For example one leg of an animal would appear fatter than another. The seams change the structure and geometry of objects. Especially in symmetric objects like cars or an animal figures, this leads to an unpleasant distortion. As human beings tend to see objects referencing from their memory, structural distortions in very commonly known objects results into unrealistic image.

The key component contributing to the performance of seam carving is the energy model used. The energy model decides the important features of an image. This clearly points out that the above limitations are due to the limited modeling of saliency by the energy function. In Section 3, we bring forward such cases where the original energy function failed, reason why it failed and then suggest our improvements to it.

As the algorithm chooses to remove the seams with the least amount of energy from the image, it ignores the energy that is inserted into the retargeted image. The inserted energy is due to the new edges created by previously non adjacent pixels that become neighbors once the seam is removed. Also repeatedly removing seams in the same place introduces higher energy, creating new edges. In order to reduce these disturbing artifacts, we suggest a technique to control seam traversal in the Section 3.2. Interestingly, in parallel with our work, the recent paper about seam carving for video [10] suggests the forward energy concept which attempts to minimize the introduction of energy due to seam removal instead of minimizing the energy which is removed.

Section 3

Improvements to Seam Carving

A key step in automatic retargeting is to identify the important aspects of the image so that these can be emphasized. Given no other information about the meaning of an image or the needs of the viewer, seam carving is based on *visual saliency* as an approximation of important features. Visual saliency is the distinct subjective perceptual quality which makes some items in the world stand out from their neighbors and immediately grab our attention [4]. The motivation to use visual saliency is that the parts of the image that stand out are likely to be noticed by the low-level human visual system and hence are most likely to be important to the meaning of the overall image. Therefore seam carving uses gradient magnitude of an image (edge detector), a coarse approximation of the visual saliency, as the energy model.

In the section 3.1, we discuss the evolution of the improved energy function to better reflect visual saliency. In section 3.2, we show how seam traversal can be controlled to reduce some of the artifacts introduced by carving. Finally section 3.3 redefines the improved energy function showing how all the different feature extraction techniques (saliency maps) are combined.

3.1 Redefining the energy function

The energy function guides the optimal seam selection. As seam carving is based on removing unnoticeable pixels, the original algorithm defines the sum of absolute gradient in horizontal and vertical direction (also called $L_1 - norm$) as the energy function:

$$e_{fine}\left(x,y\right) = \left|\partial I\left(x,y\right)/\partial x\right| + \left|\partial I\left(x,y\right)/\partial y\right| \tag{3.1}$$

The paper on seam carving [2] talks about testing several other energy functions like the L_1 and $L_2 - norm$ of the gradient, Harris corner detectors, entropy and Histogram of oriented gradients and concluded that no single energy function performed well across all the images. All the energy functions suggested in the paper, perform a very fine scale analysis of the image and hence give similar results. With the goal to overcome some of the limitations of seam carving and considering the key role of the energy function, we took a different approach in defining the energy model of an image. We use a bottom-up visual saliency model to compute a saliency map which reflects the eye-catching features of an image. In the following section we see how the energy function is improved by analyzing the image at different scales, how the color in the image can be used to model human perception and how high level object detection can add valuable cues to energy model.

3.1.1 Multi Scale Approach

On retargeting an image with multiple scattered subjects, the background is identified as low energy and retargeting is aimed by removing the uninteresting background. However when there is texture in the uninteresting background, the energy function as defined in equation 3.1 fails, identifying not so important textured background as the high energy area and preventing the seam to pass through it. As a result the algorithm wrongly tends to select a seam through an important region of the image while protecting not so important textured area. To make it more clear, see the example in Figure 3.1

Figure 3.1: *Original image with texture in the background*

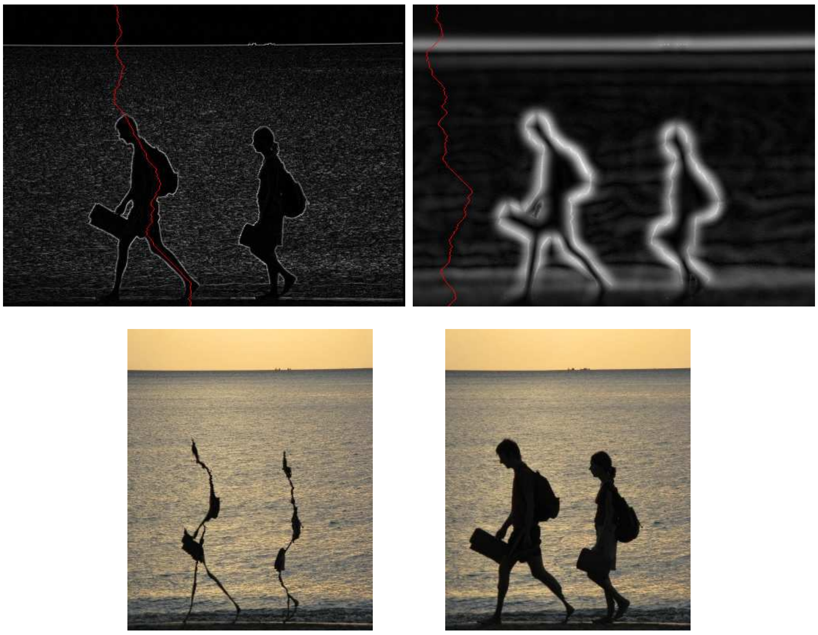

Figure 3.2: *On the top-left is the energy map used by the original seam carving and to the top-right is the energy map suggested by our multi-scale approach. The bottom row shows the corresponding retargeting results.*

In order to achieve effective retargeting, human perception needs to be considered. Our eye tends to perceive structures in images and recognize familiar objects. Our eye is capable of combining information at different scales and also can suppress fine-scale details when needed. For instance, when looking at a table made of wood, we automatically disregard its fine-scale texture in our everyday life. Organization at multiple scales is our way of simplifying the complicated environment before us. Hence the multi-scale feature extraction technique is popular in computer vision. For our problem of identifying interesting regions in an image we could similarly use different scales.

Since we are interested in suppressing finer patterns corresponding to high spatial frequencies, a blurring or a low pass filter is required. It has been established that within the class of convolution transformation, it is best to rely on Gaussian kernel as it does not generate artifacts by smoothing [5]. Therefore, with some heuristics and seeing some results with images of the size 500 x 400 we considered Gaussian smoothing kernel of σ of 5. However, it should be noted that the value for σ need not be strictly 5 and could be tuned keeping in mind the intention of smoothing is to block out the finer details and extract regions of interest. Choosing a smaller value for σ will mean giving importance to fine details and very high values for σ will mean neglecting sizable amount of detail, focusing strictly on bigger objects.

The energy function can be redefined as a combination of fine scale gradient (luminance channel) of image with normalized gradient at a coarser scale, giving higher weightage to the coarser scale and using the finer scale only where coarser scale had nothing to add. The improved energy function now looks like:

$$e_{multiscale}(x, y) = e_{fine}(x, y) + \sigma * (|\partial G(I(x, y), \sigma)/\partial x| + |\partial G(I(x, y), \sigma)/\partial y|) \quad (3.2)$$

where $G(I(x, y), \sigma)$ is convolution of a grey-scale image I with a 2D Gaussian smoothing function with standard deviation of σ. Effectiveness of such an energy function can be seen in the example of figure 3.2.

3.1.2 Color

In general most of the image analysis is done on the luminance channel. However when trying to interpret the contents of an image, color cannot be discarded. Color does play an important role in object identification and recognition in the human visual system. Learning from the importance of color in object segmentation, it is evident that color gradients emphasize the object boundaries. Due to high sensitivity of human perception to the edges and contours, the exact extraction of object boundaries and preserving the same could provide additional visual cues for better retargeting.

Figure 3.3 shows that, in addition to the luminance gradient, including the color gradients in the energy function leads to better highlighting the subjects which in this case are the leaves and the flowers. Also as we want to emphasize object boundaries and are not interested in color features of fine details, we only use color gradients at coarser scale. There exist several choices of the color space for image segmentation. Since we want to separately treat luminance and color channels, we use the YCbCr color space. Equation 3.3 defines the energy function for a coarser scale analysis of image.

$$\begin{aligned} e_{color}(x, y) \quad = \quad & \sigma * (|\partial G(I_{Cr}(x, y), \sigma)/\partial x| + |\partial G(I_{Cr}(x, y), \sigma)/\partial y| + \\ & |\partial G(I_{Cb}(x, y), \sigma)/\partial x| + |\partial G(I_{Cb}(x, y), \sigma)/\partial y|) \end{aligned} \quad (3.3)$$

where G is a 2D Gaussian smoothing function with standard deviation $\sigma = 5$.

3.1.3 Color Histogram

Human scene categorization [15] suggests that humans rely on local, region-based information as much as on global, configural information. In addition, humans seem to integrate both types

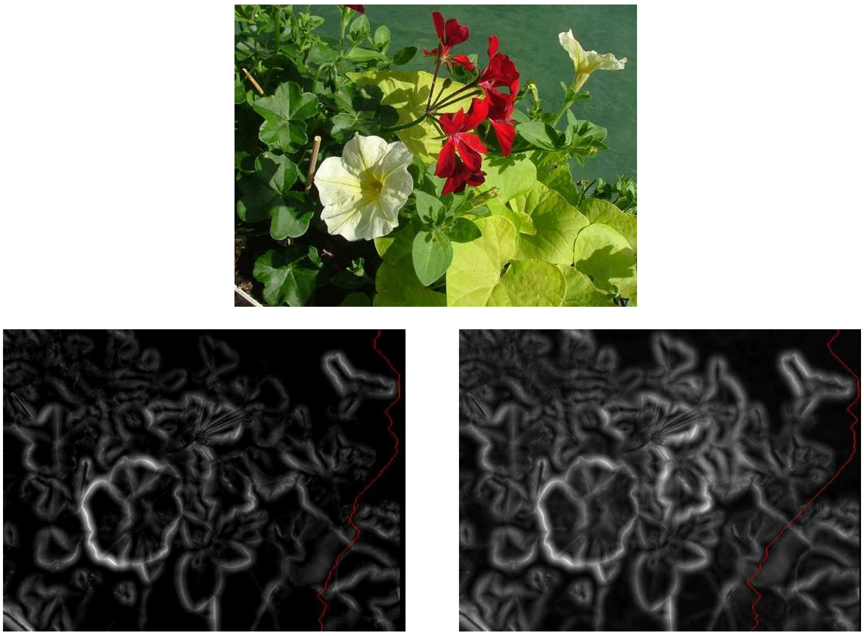

Figure 3.3: *On top is the image being retargeted. On bottom-left is the energy map as combination of coarser and fine scale gradients, on the bottom-right energy map with additional color gradient at coarser scale.*

of information for intact scene categorization. Objects seek attention depending on surrounding context in which they occur. Colors which are rare in the image stand out to attract human gaze.

To preserve the parts of an image with rare colors, we use a color histogram to analyze distribution of colors in an image. For our implementation, we have discretized each of the Cr and Cb channels into 10 bins. A two dimensional (2D) normalized CbCr-histogram of an image is produced counting the number of image pixels in each bin. Pixels in the image are then weighted in inverse proportion to the histogram bin they belong to. To be more specific, if pixel at (x, y) falls in a bin whose histogram value is given by H, then energy due to CbCr-histogram analysis is given by:

$$\begin{aligned} \mathrm{e}_H(\mathrm{x,y}) &= \frac{1}{H(I_{Cb(x,y)}, I_{Cr(x,y)}) * 100} \quad \text{when} \quad H(I_{Cb(x,y)}, I_{Cr(x,y)}) > 0.015 \\ &= \frac{1}{0.015 * 100} \quad \text{when} \quad H(I_{Cb(x,y)}, I_{Cr(x,y)}) \leq 0.015 \end{aligned} \tag{3.4}$$

Note here the energy for each pixel is defined in the range of [0-1]. As we want to emphasize rare colors, but not really distinguish between their exact proportions in an image, we take a maximum cut-off histogram energy of 0.66. All the rare pixels with $H(I_{Cb(x,y)}, I_{Cr(x,y)}) \leq 0.015$ will get a maximum weight of 0.66. Thus we avoid infinitely high e_H values possible due to very small histogram values in the denominator in equation 3.4. Moreover by defining such a cut-off for histogram energy, when combining histogram energy with gradient energy, the edges will still remain more important than the rare color pixels. The example in figure 3.4 shows the retargeting result after using histogram analysis of color channels to extract visually attracting image areas.

Figure 3.4: *Retargeting results on left without color analysis and on the right using color histogram giving more importance to rare colors. Red and white flowers which catch our attention in a fairly green background are better preserved on the right.*

3.1.4 Spatial Position

Seeing the success of Panoramic stretch in displaying 4:3 aspect ratio TV broadcast on a 16:9 wide-screen display, it has become evident that focus of the image or the region of interest is around the center in most images. When defining the energy map of an image, edges and hence the objects around the center could be given more importance. Distortion of objects towards the boundaries of image could be less annoying compared to distortion of objects around the center.

To implement this we bias the energy of every pixel with respect to its spatial position in the image. Let (xc, yc) be the center of an image I with n rows and m columns. Suppose $d_E((x, y), (xc, yc))$ is the euclidean distance between the pixel p at (x, y) and the center (xc, yc). The maximum distance $maxd_E$ will be between one of the corners of the image and its center and can be defined as in equation 3.5. To give greater importance to the visually salient features in the center compared to those towards the boundaries we bias the energy map $w_{spatial}$ as:

$$maxd_E = \sqrt{(n/2)^2 + (m/2)^2} \tag{3.5}$$

$$w_{spatial}(x, y) = 1 - \left(\frac{d_E((x, y), (xc, yc))}{maxd_E}\right)^2 \tag{3.6}$$

3.1.5 High level Cues

Any form of retargeting is bound to lead to some deformation in the retargeted image. Especially when the image has familiar objects with regular well known form, removing a seam through it may cause a noticeable artifact. For instance, distortion to a straight-line or a symmetrical object or a human figure distortion would be very quickly noticeable and also annoying. Hence we thought of combining the energy function depending on visual saliency with some high level object recognition like skin tone detection and face recognition. The original paper [2] also mentions about using a face detector. We show how it can be incorporated with the improved energy function and later on the evaluations in chapter 6 shows its importance in the seam carving algorithm.

Figure 3.5: *Original image on the left. In the middle retargeted image using seam carving without spatial weighting and on the right with spatial weightage decreasing as we move away from the center of image.*

- Skin-tone detection: Many would be interested in receiving videos and images of a live sport or a family or group picture on a hand-held device. In order to prevent any distortion in human figures in such images, we thought of detecting skin-tone to preserve human figure. However we were not very successful here for mainly two reasons: To our knowledge, there is no accurate skin color detector suitable for all lighting conditions and secondly even where skin is accurately detected, it does not solve the problem of detecting the entire human figure. Since a human figure is a combination of face, skin part and some colored areas of human apparel it was difficult to accurately identify all the parts and give them equal importance to evenly preserve it.

- Face recognition: A very simple way to prevent face distortions while seam carving is using face detection to redefine the importance map of image. We use freely available *OpenCV* library which implements a face detector as proposed by Viola and Jones in [14]. It is a very efficient detector for frontal faces at several scales. It outputs bounding boxes of the detected faces. The energy map of an image is updated with the additional energy in areas corresponding to the bounding boxes, thus identifying the faces as energy areas to be preserved.

3.2 Seam Traversal

In addition to the energy function, the seam carving algorithm also depends on seam traversal. We have seen that a seam is defined as a continuous path of pixels from top to bottom or left to right. A single vertical seam beginning at leftmost corner of image could extend up-to the rightmost part of image. However seam traversal could be controlled in couple of ways. For instance, straighter seams can be introduced, making the seams more rigid. The result could be similar to that of cropping, retargeting effects are further localized.

However we have looked at seam traversal from a different perspective of reducing some of the artifacts of carving. It is observed that depending on the image content, the seam carving algorithm may continuously select seams from the same part of the image, sometimes leading to visual artifacts. This is because removing low energy pixels can create higher energy as new

Figure 3.6: *Retargeting results on left without using the knowledge of previously removed seam and on the right considering the history of seams already removed from the image. On the left we see seams selected from same spatial position as result we see extra lines traversal to waves appearing quite annoying. Disallowing the seams to take the same path gives smoother retargeted image on the right.*

neighbors are formed. The artifacts could be seen as a photographic image first ripped apart and then glued together (Figure 3.6). Since the higher energy is created by removal of seams, knowledge of seams already removed could warn us against accumulation of high energy. To discourage additional seams through regions where a seam is earlier removed, we revise the energy map of image to reflect the history of all the seams removed. After a seam removal, the energy map of the image is updated along the seam path, with additional energy proportionate to the energy of the seam pixel being removed.

In our implementation we distribute the energy of the seam pixel carved between the its two neighboring pixel. More specifically when carving a vertical seam, energy of the pixel which is being removed is distributed between its left and right neighbor. This additional energy due to history of seam can be tuned in individual implementation by choosing to add some percentage of the energy of the pixel carved. By choosing a higher percentage we choose to emphasize on the history of seams and by choosing a small percentage we disregard the history, allowing seams to be selected in same regions. In our implementation we have got good results by choosing to distribute 50 percent of the energy of pixel being removed. The overall effect (Figure 3.6) is that of different parts of image are worked on preserving the image composition and giving smoother retargeted results.

3.3 Energy Function Definition

Considering all the different methods to enhance the retargeting results discussed in this section, the improved energy function can be defined as the combination of gradient energy at coarser and finer scales, color gradients at coarser scale and additional energy due to face detection, color composition and spatial weighting. The gradient energy of an image changes after every seam removal. Therefore in the process of retargeting the energy map needs to be recalculated after removing a seam. However the energy due to high level object recognition like a face detector and the energy due to the spatial positioning of a pixel in an image needs to be calculated only once. Hence the improved energy function ($e_{Improved}$) is composed of static *base energy* e_{base} (due to object detection and spatial weighting) which is fixed once for an image and the dynamic *gradient energy* $e_{gradient}$ which varies after every seam removal. After every seam removal, the base energy e_{base} although being static, is adjusted to reflect the seam which is removed and updated to reflect additional energy $e_{seam-history}$ along the seam path.

Block diagram 3.3 shows how we combine different energy functions. Concluding all the image analysis and experiments performed to improve the retargeting results, equations below define the new energy model more formally. i denotes the retargeting iteration and $i = 1$ is the first iteration. The energy functions $e_{multiscale}$, e_{color}, e_H are defined in the equations 3.2, 3.3, 3.4 respectively. Energies e_{face} is the threshold energy added in the region where face is detected or where a seam is removed. The spatial weighting function is defined in equation 3.5.

$$\begin{aligned}
e^i_{gradient}(x,y) &= e^i_{multiscale} + e^i_{color} && (3.7)\\
e^1_{base}(x,y) &= (e^1_{gradient} + e_H + e_{face}) \ * \ w_{spatial}(x,y) && (3.8)\\
e^{i+1}_{base}(x,y) &= 0.25 * E(s) + e^i_{base}(x,y) \ \text{ where } (x,y) \in N(s) && (3.9)\\
&= e^i_{base}(x,y) \ \text{ otherwise} && (3.10)\\
e^i_{Improved}(x,y) &= e^i_{base} + e^i_{gradient}(x,y) && (3.11)
\end{aligned}$$

where $E(s)$ is the cost of seam s being removed as defined in equation 2.4 and $N(s)$ is the neighborhood of seam s. In our implementation we have considered just the immediate neighbors of the seam pixels.

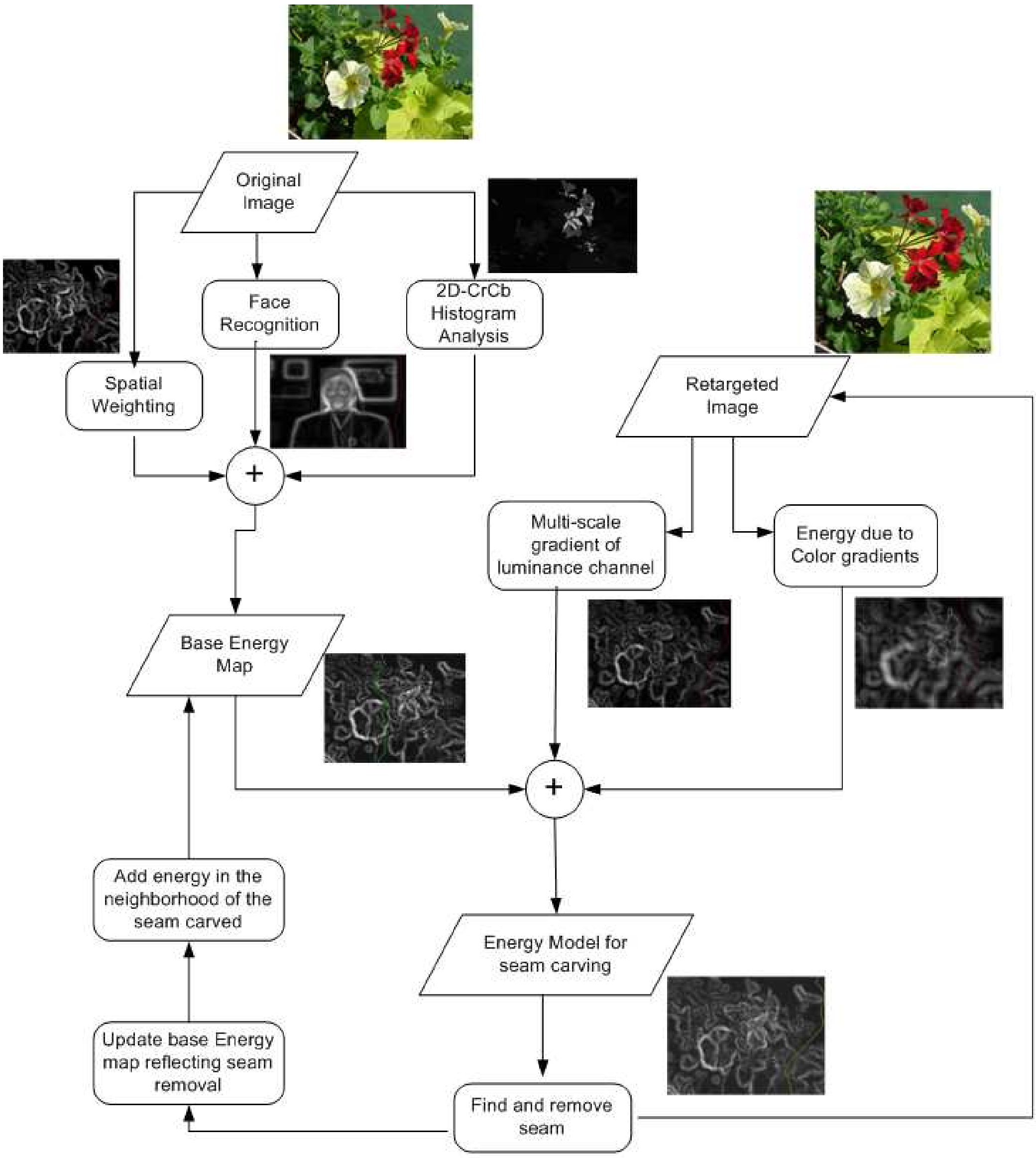

Figure 3.7: *Block Diagram of using the improved energy function for seam carving.*

Section 4

When to stop carving?

Retargeting is content dependent. If the content is too condensed, or it does not contain less important regions, retargeting will not succeed. Stated differently, amount of possible carving, depends on the image content. Hence for a general solution to retargeting, it may be better to combine retargeting with cropping or linear scaling. For that to be possible, we need to automatically detect a stage in every image, which could suggest to stop carving. In this section we put forward our attempt to analyze the process of carving to guide us to stop carving before we loose the visual coherence. However, going through this section, it will be clear that it is hard to find a content independent measure which could detect such a stage.

To monitor the progression of seam carving, we can look at the trends in minimum seam energy, cumulative energy concentration, average energy of an image and the number of important edges crossed as we continue carving. Analyzing some of these trends could lead us to say something about the retargeting stage.

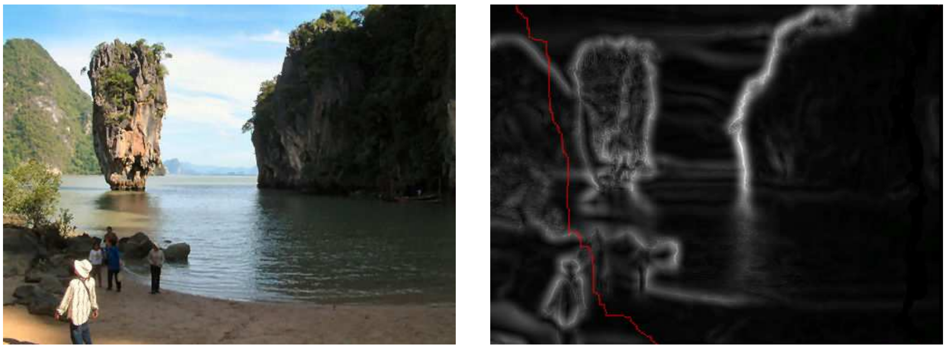

Figure 4.1: *On the right is energy map of 'island' image shown on left, with a seam highlighted in red. Such a seam cutting through a couple of strong edge could be a bad choice.*

A seam traversing through some high energy areas, although still being the seam with minimum cumulative energy will not be preferred as it is possibly distorting the structure of objects of interest. Figure 4.1 shows an energy map of an image with a seam (highlighted in red) crossing some strong edges. It is best to stop carving when seams are selected in regions of interest. A threshold energy can be defined so that energy above it marks region of interest. A seam consisting of quite a few pixels above the threshold energy is an indication to stop carving.

Moreover, in the process of carving, seams with minimum energy are removed with the result that the next seam to be removed though being the least in energy could be equally important.

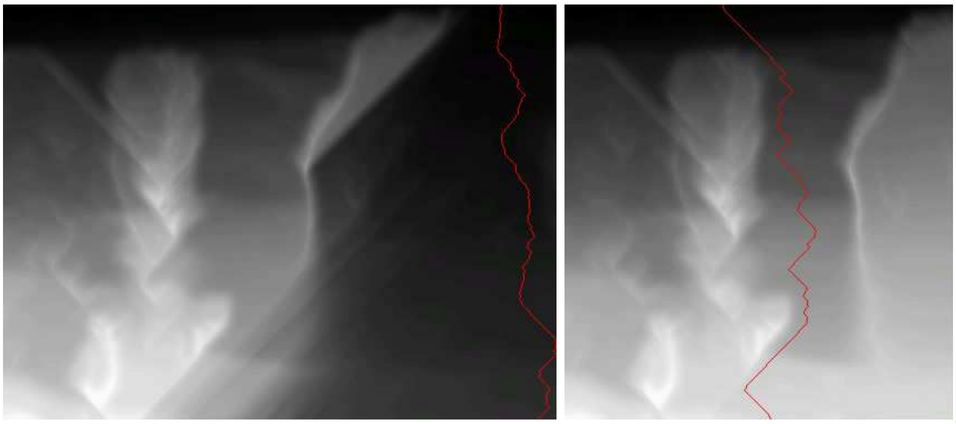

Figure 4.2: *Minimum cumulative energy before and after reducing width of 'island' image showing cumulative seam energy increases.*

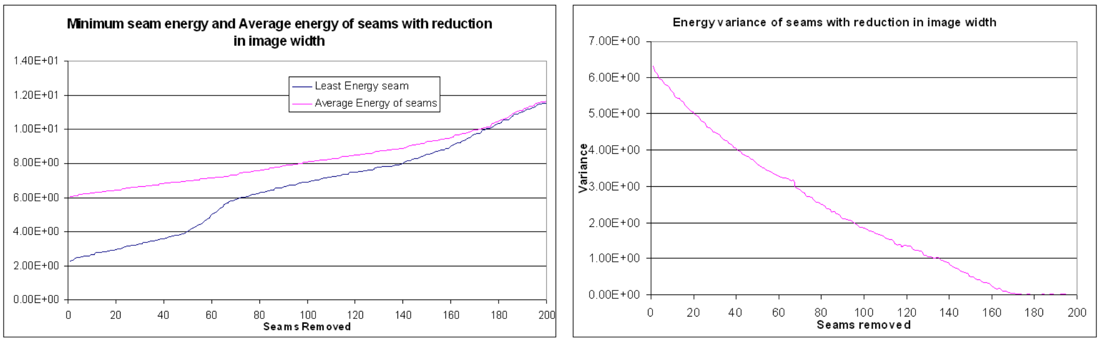

Figure 4.3: *As low energy seams are carved out, energy of seam with least energy approaches average seam energy suggesting no less important area to compromise, variance decreases.*

Stated differently as carving progresses the minimum cumulative seam energy approaches the average cumulative seam energy. Figure 4.2 shows the cumulative energy maps, before and after retargeting. A minimum energy seam highlighted in red, clearly shows that minimum cumulative seam energy is increasing suggesting there is no seam in the retargeted image which is unnoticeable, and hence further carving could spoil visual coherence. Figure 4.3 shows the energy map and the plot of change in average energy while changing the image width. One could also depend on standard deviation as a measure to detect energy variance of seams.

Although these trends could be spotted in nearly in all the images we analyzed, the rate at which these variations occurred, greatly varied. For instance, it is very difficult to approximate a particular variance in the seam energy that could be a warning to stop carving. The example image in Figure 4.4 illustrates that the variance in cumulative seam energy is decreasing. However looking at the energy maps 4.5 and the plots of variance in seam energy we see that variance may not approach zero before distorting the image. The minimum seam energy may not reach the average seam energy before distorting the retargeted image.

Additionally, if we look at the trends in average gradient energy of the image, we see that, as we continue carving we are continuously removing low energy pixels, due to which the average gradient energy of the image is increasing. This suggests the content is getting condensed or every part of the image is important. Figure 4.6 shows plot of average gradient energy while

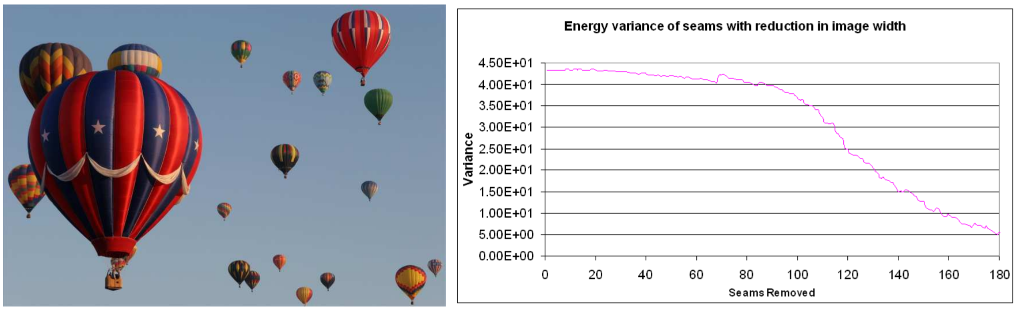

Figure 4.4: *As seams are carved out from the 'balloon' image on the left, variance in cumulative seam energy decreases.*

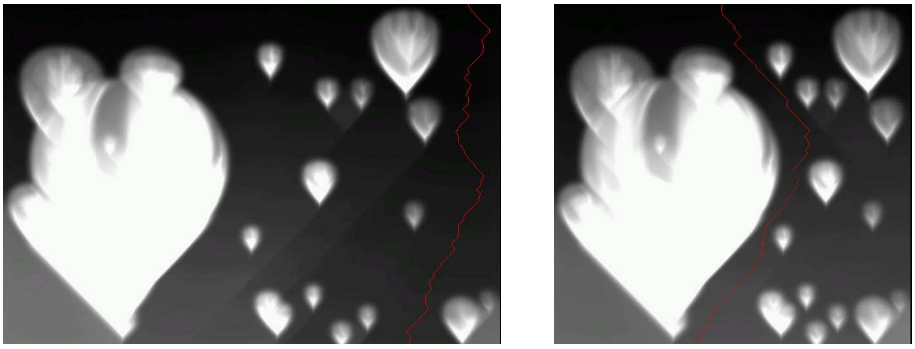

Figure 4.5: *Minimum cumulative energy before and after reducing width of 'balloon' image showing cumulative seam energy increases.*

changing the image width. The measure or the percentage increase in average gradient energy of the retargeted image compared to that of the original image point towards the retageting stage of the image. However on comparing the plots of average gradient energy for various images we found that it is hard to conclude on a stand alone measure that could serve as a warning to stop carving.

Looking at the above mentioned trends on a set of images, it is could be said that there is no one such content independent measure which we could depend upon. However, a combination of all or some such measures could perhaps serve as a good guide. Due to lack of time, we stopped our analysis in this area. Inspite of the complexity of the problem, a rough approximation is presumably possible.

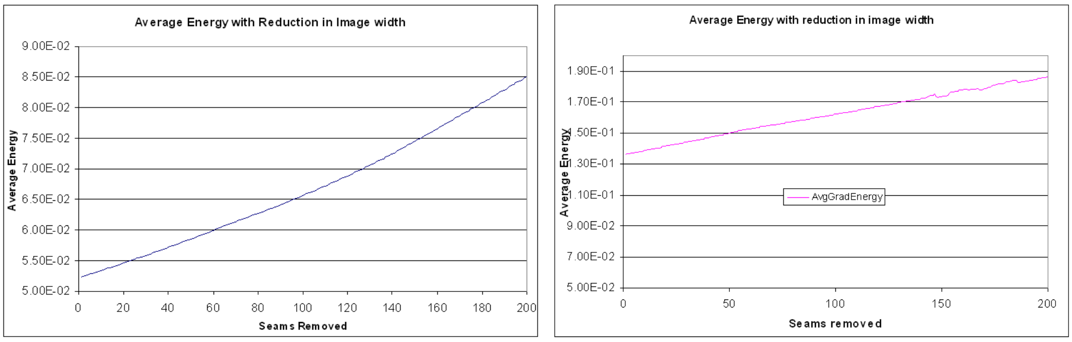

Figure 4.6: *Plots of average gradient energy for island image on the left and balloon image to the right. Average gradient increases as more and more seams are removed, but the rate of increase is different in two images.*

Section 5

Retargeting for aspect ratio change

The aspect ratio is the ratio of the width of an image to its height. These days many TV programs are still broadcasted in 4:3 aspect ratio to support the traditional standard definition TVs. Also a lot of old/legacy material is broadcasted in 4:3 aspect ratio. On the other hand there is more and more video content on DVD or satellite channels at a wider aspect ratio due to the popularity of widescreen TVs. In order to support display other than originally intended, some workarounds or compromises have been sought. Today converting formats of unequal ratios is done by either cropping the original image to the receiving format's aspect ratio, or by adding horizontal mattes (letterboxing) or vertical mattes (pillarboxing) to retain the original format's aspect ratio.

In this section we first look at Panoramic stretch as an retargeting technique for aspect ratio conversion. Next retargeting algorithm for aspect ratio change based on seam carving is suggested.

5.1 Panoramic Stretch

Naturally the traditional 4:3 aspect ratio video shot for standard definition TV, does not fit to the modern 16:9 formatted screen. As a result, viewers are treated with a linearly upscaled image, where everything is replaced with its stretched version. There are no circles left, only ovals. A different stretch technology called the *Panaromic Stretch* has been seen in Philips TV. It uses pixel position as the control function to scaling. It is based on the observations that important parts of the image are in and around the center of the image. Hence the center part retains its normal shape and image boundaries are stretched to fill up the wide screen.

Panaromic stretch follows a nonlinear curve based on the spatial position as shown in the plot in the figure 5.1. The example in figure 5.1 shows that the central parts of the image is preserved, where as regions towards the boundaries are stretched. As we saw, all the methods discussed above have some flaws. Hence we investigate, if seam carving could be an option for aspect ratio conversion.

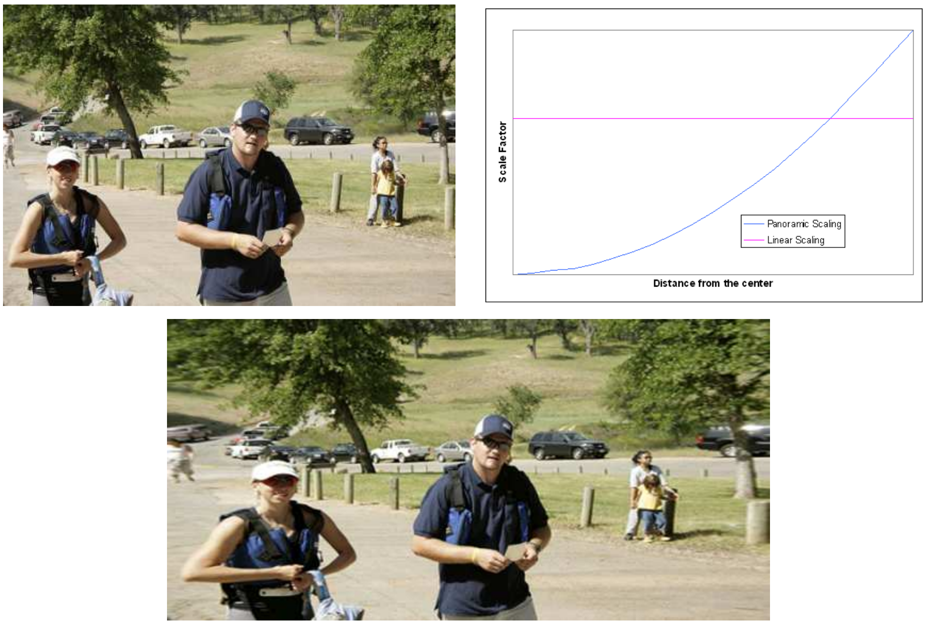

Figure 5.1: *Panoramic Stretch. On the top is the original image and at the bottom is the panoramic stretched image. Notice the stretching of the lady where as the man remains untouched.*

5.2 An Example for aspect ratio conversion using seam carving

Retargeting for aspect ratio conversion can be achieved in couple of ways. We illustrate this with an example. Consider a scenario where an image with resolution *768* x *512* (aspect ratio 1.5) needs to be displayed on a wider screen with resolution *1024* x *578* (aspect ratio 1.77). Using seam carving this can be achieved in following ways.

- Linear scaling (keeping aspect ratio) original image (*768* x *512*) to first get the target width(*1024* x *683*) and then removing horizontal seams to get the target size(1024x578).
- Linear scaling original image(768x512) to first get the target height(867x578) and then enlarging the width using seam carving.

The order of cascading scaling with retargeting could also be reversed to first get the target aspect ratio and then do linear scaling to achieve the target size. When considering the retargeting effects, removing horizontal seams or reducing the height changes the image composition, where as duplicating vertical seams or enlarging width causes stretching. Therefore, considering the image content and the required aspect ratio, depending on either of the two cannot be a reliable solution, as shown in the example of Figure 5.2. To reduce these retargeting effects while changing aspect ratio, we suggest a combination of horizontal and vertical seam carving.

5.3 Improved algorithm for aspect ratio change

As it is clear from the previous example, depending on the image content, target aspect ratio can be achieved by horizontal seam carving, vertical seam duplication or a combination of it. Assume an image of width w, and height h. The equation to achieve target aspect ratio (AR_T)

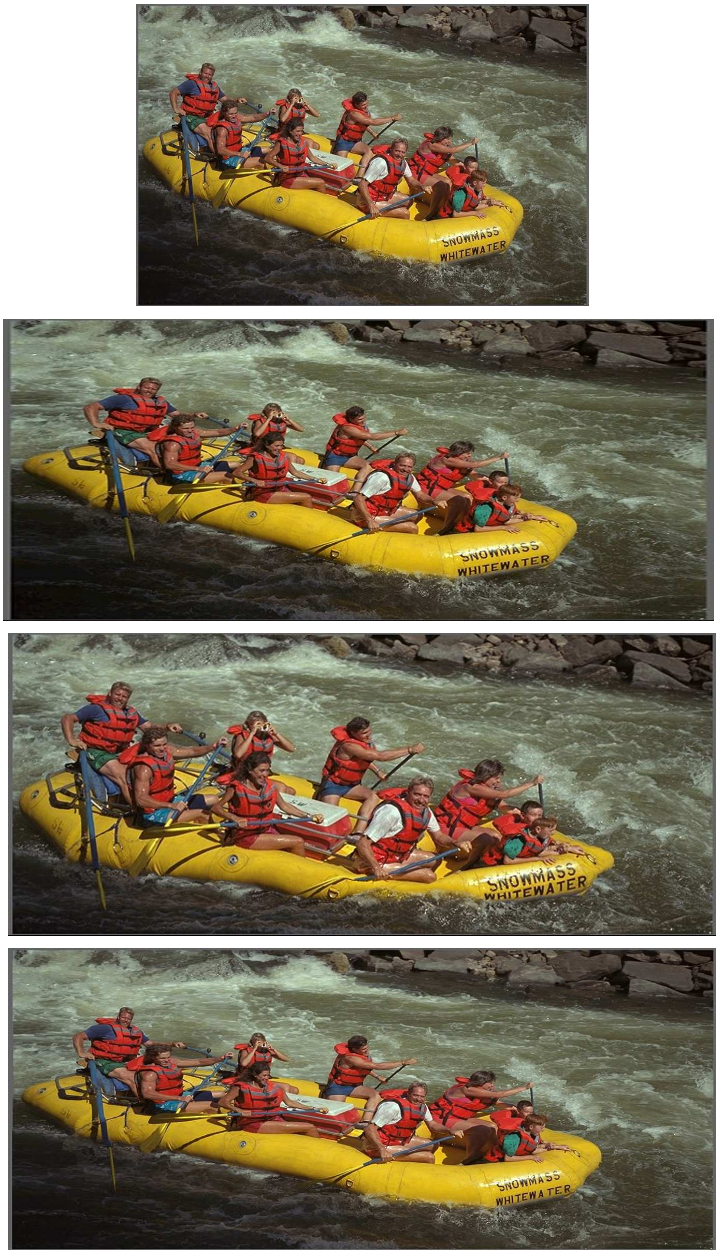

Figure 5.2: *Comparing various methods of retargeting to a wider screen. Top-most is the original image. Next is panoramic scaling showing left corner of cannoe and the person there are both stretched. From the top, third image is retargeting to target aspect ratio by removing horizontal seams (cannoe is distorted and appears sinking). Bottom most is the result of retargeting to target aspect ratio by duplicating vertical seams (notice the stretching of limbs of girl in the center).*

(assuming target aspect ratio is greater than original) from the original image dimensions, is given as:

$$AR_T = \frac{(w+c)}{(h-r)} \tag{5.1}$$

where c vertical seams are added and r horizontal seams are carved. Different combinations of c and r could lead us to the same target aspect ratio. Thus we have several solutions. As energy is related to visual saliency, in the process of retargeting we try to select a seam with minimum energy (cost). Similarly to achieve optimum retargeting, we strive to strike optimal balance between horizontal and vertical carving so that costs of retargeting are minimum. Retargeting costs are the total costs(energy) of all the seams carved. Cost of seam is defined in equation 2.4. We analyze the costs of all the combinations of horizontal and the vertical retargeting that could be possible to achieve the target aspect ratio. Figure 5.3 shows energy plot against different combinations of vertical and horizontal seams. The alliance of vertical and horizontal seams leading to the minimum cost is the optimal retargeting for target aspect ratio.

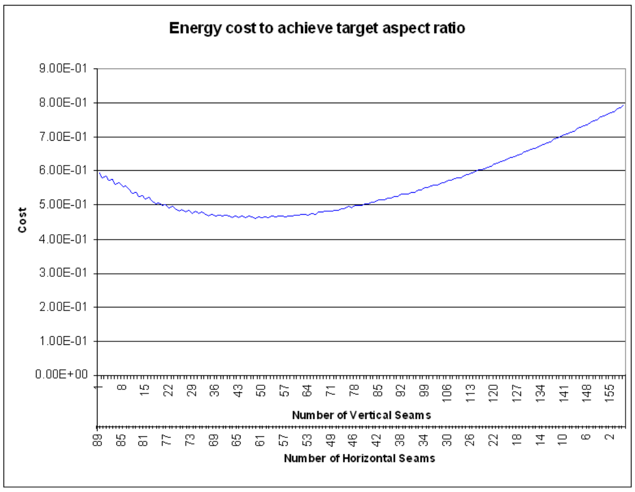

Figure 5.3: *Cost of different combinations of vertical and horizontal seams for 'Balloon' image.*

Algorithm: To find optimal retargeting for aspect ratio change.

- Find maximum number of vertical seams(c_{vmax}) that are required to be duplicated to get the target aspect ratio (AR_T) using :

$$AR_T = \frac{(w+c_{vmax})}{h} \tag{5.2}$$

- For *i*=1 to c_{vmax},

  1. Find cost of duplicating *i* vertical seams.
  2. Find how many horizontal seams(*r*) needed to be removed when *i* vertical seams are added, so that target aspect ratio is achieved, using :

$$AR_T = \frac{(w+i)}{(h-r)} \tag{5.3}$$

3. Find the cost of *r* horizontal seams.
4. Find the total cost as sum of costs in step 1 and step 3

- Find minimum of the total costs found in step 4. Duplicating corresponding number of vertical seam and removing corresponding horizontal seams would give us the best results for the target aspect ratio.

The same algorithm holds when the target aspect ratio is smaller than the original aspect ratio, but instead of duplicating the vertical seams we carve vertical seams and duplicate horizontal seams. Thus seam carving could be extended to serve the needs of any aspect ratio change. Figure and 5.5 shows results of proposed algorithm.

We have seen the effectiveness of this algorithm on a set of images. Moreover in section 6 the results of this algorithm are compared with panoramic scaling. A drawback of this technique is that it is computationally expensive, as it first needs to evaluate the cost of complete range of vertical and horizontal carving and find an optimal solution and then once again carve image as per the optimal retargeting.

Figure 5.4: Comparing panoramic scaling(top) with proposed algorithm(bottom) for aspect ratio change from 1.5 to 2.35. Notice that in the first image, leftmost corner of cannoe is stretched whereas the retargeted image amplifies the cannoe and its people.

Figure 5.5: Comparing panoramic scaling(top) with proposed algorithm(bottom) for aspect ratio change from 1.5 to 2.35. Notice that in the top image, the sheep is stretched and the ladder is distorted.

Section 6

Evaluation

After having analyzed the approach to such an extent, it made sense to evaluate our work against the originally proposed seam carving operator. Judging the results of retargeting is difficult as it mostly depends on human perception. Therefore, it is hard to define an objective quality metric such as the mean square error (MSE) score. On the other hand, subjective methods are known to be the most precise measures of perceptual quality and hence we relied on it. We had set up two goals for the subjective test.

1. How good is the improved seam carving compared to that originally proposed by Ariel Shamir and Shai Avidan?

2. How good is the the algorithm for aspect ratio change using improved seam carving algorithm, when compared to panoramic stretch?

6.1 Test Specifications

- Image Categories: Since retargeting is a fairly new area of study, no standard test database is available for evaluation. Hence we had to come up with our set of test images. Considering the image categories where retargeting is required, we tried to include all range of images in our test. Our image database consists of images with single or multiple scattered objects with textured/non-textured background, nature-landscape, nature-close-ups (flowers, stream etc), group/family pictures, etc.

- Methods to measure quality: Several measurement techniques are described within ITU-R BT.500 standards for subjective picture quality measurement. We use "The Stimulus Comparison Adjectival Categorical Judgment method (the SCACJ method)". The two images are displayed side by side and afterwards the viewer is asked to give his opinion using comparison scale shown in figure 6.2. The viewer is unaware of which image is what and provides a voting that can be represented within a seven-grade measurement (measurement score ranges from -3 to +3) on the quality of both the images.

6.2 Experimental Setup

To answer the two very distinct questions and to show the results of retargeting we designed two very different experimental setups.

- Test Setup 1: Comparing improved seam carving with original seam carving operator. For this we took a mobile or a hand-held device as target display. Original images of different resolutions ranging from 500x333 to 500x480 were retargeted to resolution of 320x240. Also included images of resolution 333x500 to 480x500 which were retargeted to 240x320. The original image was retargeted twice once using the seam carving algorithm as proposed in the paper and second time using our revised seam carving operator. The two retargeted results were shown for pair-wise comparison. The original image is not shown keeping in mind the practical situation where the viewer will only have the retargeted view.

- Test Setup 2: How good is the Seam Carving operator for aspect ratio change? We compare the results of seam carving with panoramic scaling. Panaromic scaling stretches the sides of the images to adjust to a larger aspect ratio. For our test, original image with aspect ratio of 1.5(resolution 768x512) is retargeted to wide screen aspect ratio of 1.777 with resolution 1024x578. We first use seam carving for aspect ratio change on the original image and perform linear scaling to 1024x578. This later result is then compared to panoramic stretch from resolution 768x512 to 1024x578. The two results are shown side by side for comparison.

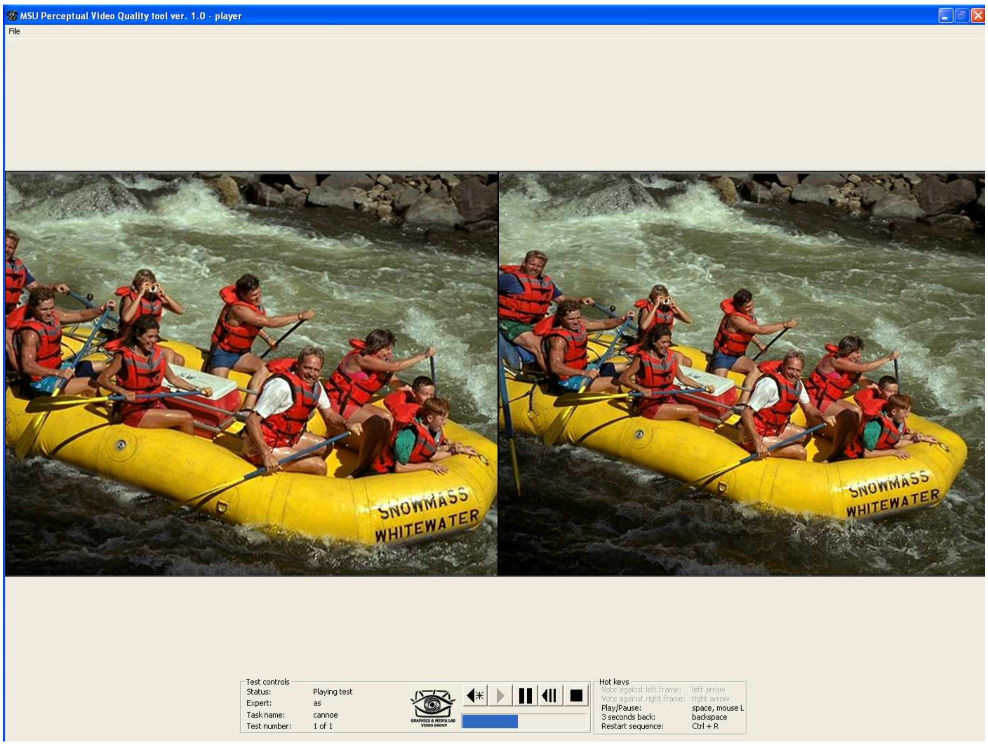

Figure 6.1: *MSU Perception Tool*

For both the setups we took a set of 15 images. A single wide-screen 42" display monitor with resolution 1920 x 1080 is used. However since the two setups are very different, we did provide a guideline for the viewing distance. But the viewer was free to set up his own viewing distance. Also there was no constraint on the viewing duration, sufficient time was given. The pair images were shown for 15 seconds before the viewer was asked to vote. However

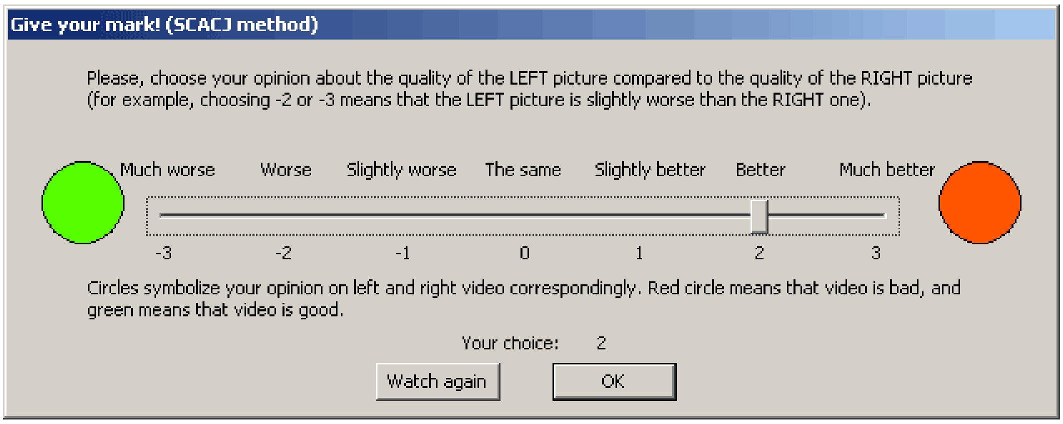

Figure 6.2: *MSU Scoring Dialog*

he still had the option of viewing the images again. We use *MSU Perceptual Video Quality tool* [1] which implements the *SCACJ*(Stimulus Comparison Adjectival Categorical Judgement) method. Figure 6.1 shows the graphical interface of the software with images displayed in pair and Figure 6.2 shows the dialog box where viewers give their voting. 20 viewers participated in the perception test. These viewers are the employees and the students from the Philips Research Video Processing and Analysis group with a good background knowledge in this area.

6.3 Results

The differences between the image pairs are scored on a 7-grade scale +3 to -3 labeled as much better, better, slightly better, the same, slightly worse, worse, and much worse.

Test Setup 1:

Our test database for this setup included 15 images with complex backgrounds and with single or multiple regions of interest. We have considered an image-mix which can be broadly classified as below:

Table 6.1: Image Categories: Test Setup 1

Category	Non-faces	Faces
Single ROI	4	1
Multiple ROI	8	2

Average score plot (Figure 6.3) clearly shows that our improved seam carving operator gives better results in most cases. Out of 15 images we tested, our enhancements showed improved results for 13, comparable results for one image and for only one image the original algorithm performs better. This clearly shows that focusing on higher level content analysis and suppressing the influence of texture gives visually pleasing results. Looking at the improved results for images 3, 4 and 5 it is proved that high level object detection like face detection can achieve better retargeting due to the fact that distortions in objects like faces which are known to have specific form is not found acceptable to the viewers.

Images (like $lilies$ in figure 6.4) where original seam carving worked comparable or slightly better are the images where content occupies more or less the entire image area. Any retargeting

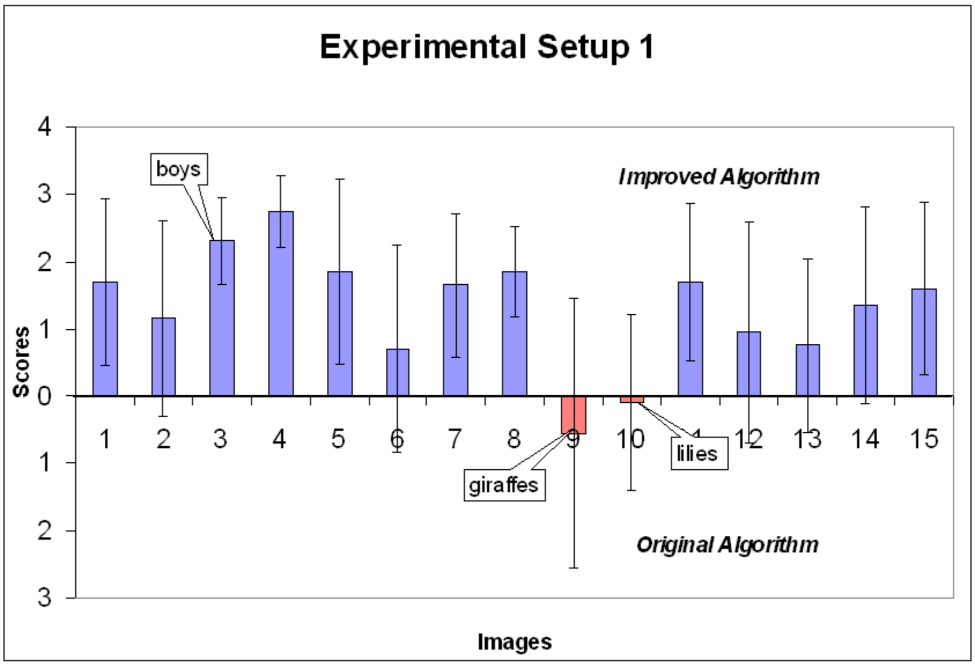

Figure 6.3: *Comparing average scores collected from 20 viewers for 15 images, with improved algorithm scores above the x-axis and the original algorithm below x-axis. Vertical error bars show standard deviation.*

will tend to produce distortion and then it is more matter of personal taste as where and what kind of distortion could be acceptable. The large standard deviation for images where original seam carving works better justifies it.

In the giraffe image (figure 6.4) there is high gradient between the bright sun and green parts of forest. Since our improved algorithm highlights the contrast, it did not allow many seams to pass through it. Where as original algorithm distributes the seam more evenly, as it does not distinguish between object boundaries or texture edges. As a result original algorithm performed better.

Test Setup 2:

Our test database for this setup included 15 images with complex-textured backgrounds. The image set included close-ups as well as nature scenes and be categorized as under:

Table 6.2: Image Categories: Test Setup 2

Category	Non-faces	Faces
Single ROI	3	1
Multiple ROI	9	2

Retargeting scores (Figure 6.5) are higher compared to panoramic scaling for 9 images, for 2 images scores are almost comparable and panoramic stretch performed better in 4 cases. This shows that seam carving is an interesting candidate for aspect ratio conversion. Retargeting images in many cases resulted into content amplification (upscaling of the region of interest), which was appreciated by the viewers compared to the distortion at the image borders produced by panoramic stretch. In our test images, seam carving artifacts which annoyed the viewers giving preference to panoramic stretch can be pointed out as:

1. Distortion in straight lines.

2. New edges created due to removal of low energy seams. This was seen in images with water or sky as the background. New imaginary edges in background were clearly annoying.

3. Disturbing the proportion of objects in the image. Images with human figure tend to preserve the face, while distorting the rest of the human figure. This mismatch between the face and body proportion was disturbing. Symmetric objects like cars, buildings are unevenly preserved resulting into distorted objects.

On the other hand, panoramic scaling caused uniform distortion only towards the image borders. Since in most cases image boundaries are part of image background, the distortion did not bother the viewers much. Inspite of it the plot in figure 6.5 shows that retargeting results outperforms panaromic stretch.

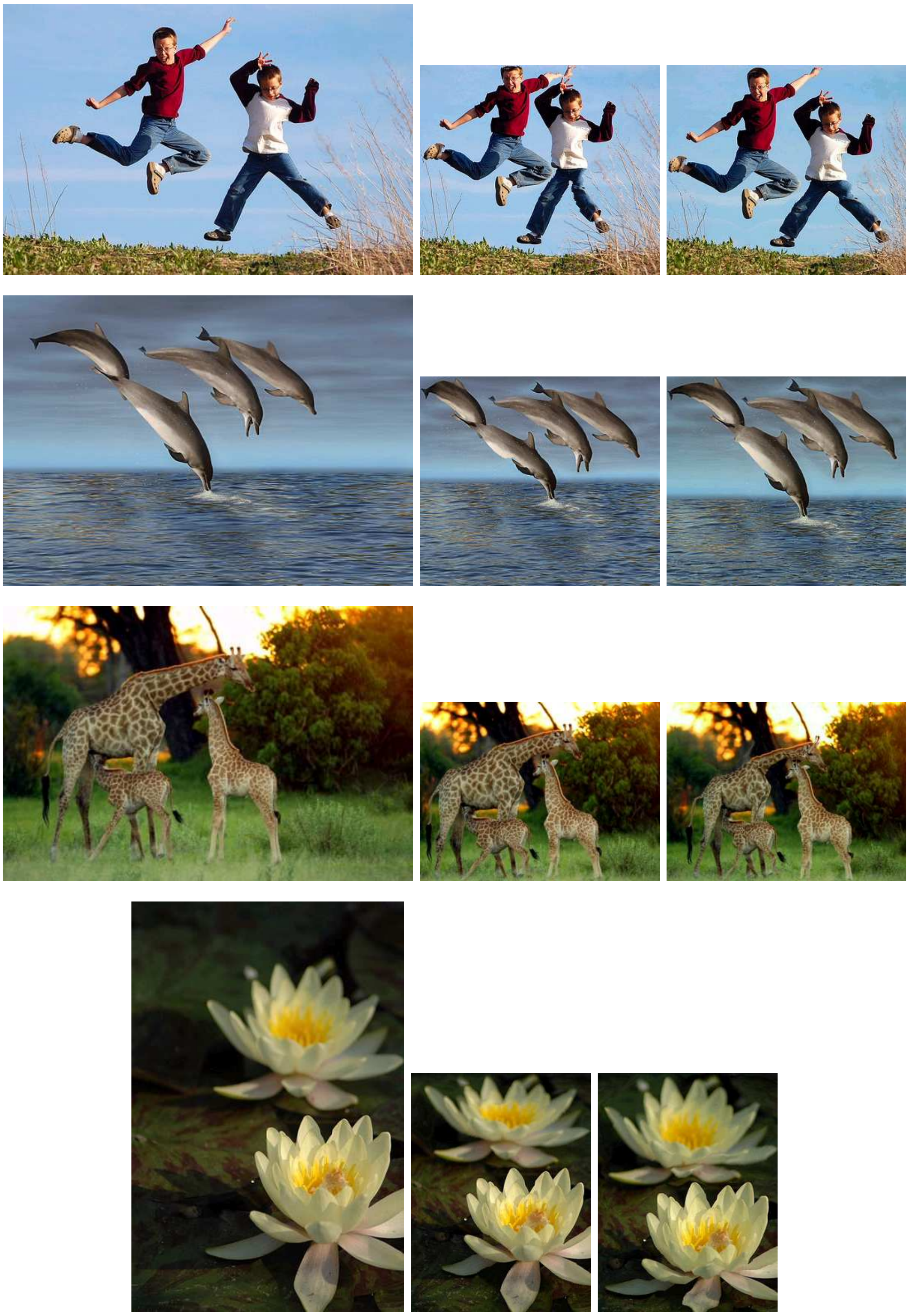

Figure 6.4: *Some images used in experimental setup 1. Original image on the left, retargeting using original seam carving algorithm in middle and right most is the retargeted image using improved seam carving algorithm.*

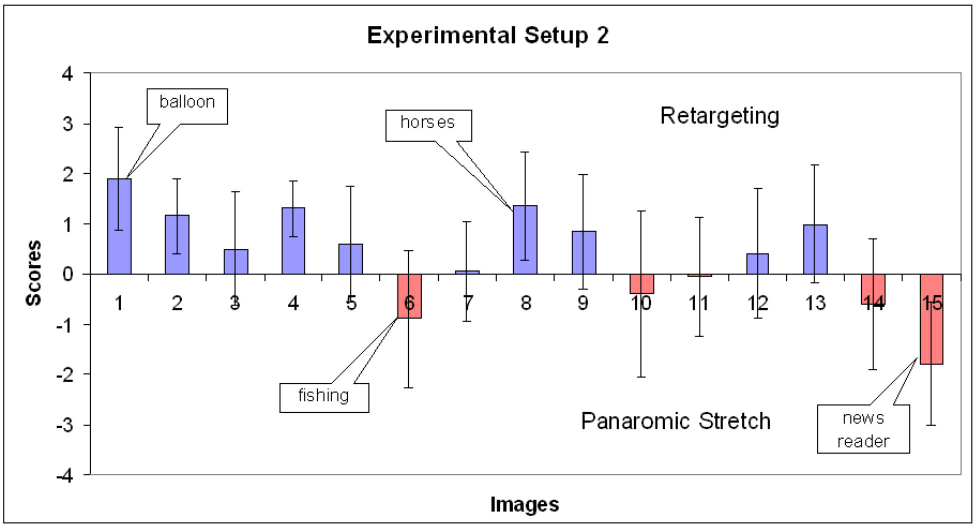

Figure 6.5: Average Scores for 15 images with Retargeting above the x-axis and Panaromic stretch below x-axis. Vertical error bars show standard deviation.

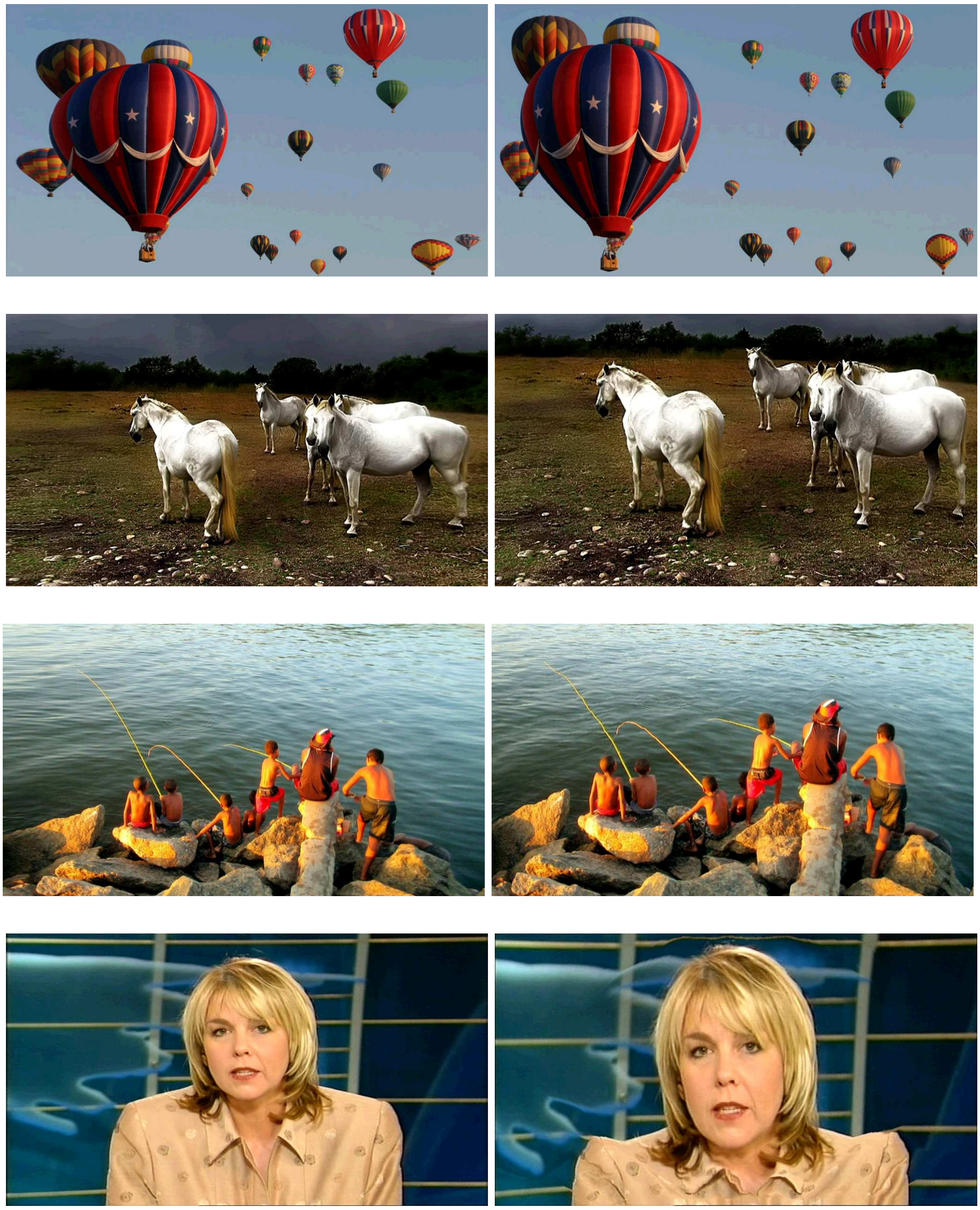

Figure 6.6: Some images used in experimental setup 2 for aspect ratio change. On the left is panoramic stretch, on the right retargeted image. Clearly for the *balloon* image we see panoramic scaling distorts the shape of balloons. Similarly in *horses* image panoramic scaling stretches the horse. In the retargeted *fishing* image we see fishing rods are distorted and in *news-reader* image we see both the methods cannot maintain the body and face proportion. However panoramic distortion is more uniform and hence less noticeable whereas retargeting fails in preserving the straight lines in the background and distorts the shoulders

Section 7

Conclusions

Seam carving gives remarkable results while retargeting images with multiple subjects separated by an uninteresting background. Retargeting is achieved by removing or inserting a seam in the region of image which is less important such as the background. We presented an improved seam carving algorithm which can now handle images with more complex background. Our improved energy function better models visual perception to distinguish between important and unimportant parts of the image. Subjective evaluation confirms the superiority of our improved seam carving operator over the originally proposed operator in most cases. The evaluations also suggest, the coarser scale image analysis as a key contributor in improving the energy model. Although our algorithm might be deforming the original image we can better preserve images' important features in the retargeted image.

However as defined in the algorithm, since the seams have to be continuous, it is possible that a part of seam traverses through regions of interest. Depending on image layout, sometimes the seams cannot bypass some image features. In such cases, if these image features happen to be straight edges or objects known to us having a very specific form and shape, any kind of distortion in them may be annoying. The tests show that, distortions in objects which we tend to recognize from our memory like faces or symmetrical objects like a round ball or a tall building is not acceptable to the viewers. Such distortions are apparently noticeable, giving a non realistic retargeted image. Seam carving needs to take that into consideration for better retargeting. High level object detection could help in preserving such objects.

Besides image retargeting, seam carving can be used for variety of image manipulations like aspect ratio change, content amplification and object removal. We presented an algorithm for aspect ratio conversion using seam carving. Seam carving gives visually pleasant results for the target aspect ratio for most images. At the same time it was observed that, seam carving changes the image composition by removing background, which may not be favored always.

Taking in account the computational complexity of seam carving compared to panoramic stretch, its sometimes unfavorable change in image composition(layout) and also the kind of artifacts which it creates, panoramic stretch could still be an equally favored technique for wide screen conversions due to uniform distortion in all images that too only towards the boundaries where in most cases the content is not so important.

However the experimental results clearly show that seam carving could be an interesting candidate. The reason is while changing aspect ratio, seam carving achieves content amplification which is generally preferred by the viewers. Typically in images with objects towards the image boundaries, retargeting using seam carving is preferred compared to panoramic stretch.

Section 8

Future Work

Depending on image layout, sometimes the seams cannot bypass some image features resulting into a distorted image feature. Moreover, seam carving is oblivious to symmetry of objects, causing uneven resizing in different parts of object, resulting into visually unpleasant distortion. So in order to preserve such objects, seam carving could be combined with higher level object detection.

We have seen how by better understanding visual saliency, the energy model of an image and hence the results of seam carving can be improved. Further understanding of visual saliency could help to acheive better results for image and video retargeting.

The forward energy concept as stated in [10] sees that minimum energy is created when seams are removed. Combining the forward energy concept, with our improved energy function, can result into better edge preserving retargeted images with smoother background.

We showed how retargeting for aspect ratio change can be achieved using seam carving. Considering its computational complexity, we would like to find a better way to combine horizontal and vertical resizing. It is possible to make some rough estimation of horizontal/vertical seam energy required for aspect ratio conversion based on initial energy map. It could be worth to see how good this approximation could be to the optimal solution.

Seam carving could be extended in resizing videos. Recently Avidan *et al* [10] proposed an algorithm for retargeting video. It is based on finding a 2D seam manifolds from a 3D space time volume using graph cuts. We would like to see how this would work with our improved energy function.

Apparently, seam carving results highly depend on the content of image. Hence the extent to which retargeting can be achieved, is also content dependent. We have suggested some ideas to automatically detect a stage in the process of carving to stop retargeting before distorting the image. However further exploration is needed to decide upon a stopping criteria to guide us to combine seam carving with other resizing/retargeting techniques leading us to a general solution for retargeting.

Bibliography

[1] http://compression.ru/video/quality-measure/perceptual-video-quality-tool-en.html.

[2] Shai Avidan and Ariel Shamir. Seam carving for content-aware image resizing. *ACM Transactions on Graphics*, 26(3), 2007.

[3] Li-Qun Chen, Xing Xie, Xin Fan, Wei-Ying Ma, Hong-Jiang Zhang, and He-Qin Zhou. A visual attention model for adapting images on small displays. *Multimedia systems ISSN 0942-4962*, 2003.

[4] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998.

[5] Tony Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2), 1998.

[6] Daniel Cohen-Or Lior Wolf, Moshe Guttmann. Non-homogeneous content driven video retargeting. *Proceedings of the Eleventh IEEE International Conference on Computer Vision*, pages 1–6, 2007.

[7] Feng Liu and Michael Gleicher. Automatic image retargeting with fisheye-view warping. *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 153–162, 2005.

[8] Feng Liu and Michael Gleicher. Video retargeting: Automating pan and scan. *MULTIMEDIA '06: Proceedings of the 14th annual ACM international conference on Multimedia*, pages 241–250, 2006.

[9] Hao Liu, Xing Xie, Wei-Ying Ma, and Hong-Jiang Zhang. Automatic browsing of large pictures on mobile devices. *MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia*, pages 148–155, 2003.

[10] Michael Rubinstein, Shai Avidan, and Ariel Shamir. Improved seam carving for video retargeting. *ACM Transactions on Graphics*, 27(3), 2008.

[11] Anthony Santella, Maneesh Agrawal, Doug DeCarlo, David Salesin, and Micheal Cohen. Gaze-based interaction for semi-automatic photo cropping. *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 771–780, 2006.

[12] Vidya Setlur, Saeko Takagi, Michael Gleicher, Ramesh Raskar, and Bruce Gooch. Automatic image retargeting. *MUM '05: Proceedings of the 4th international conference on Mobile and ubiquitous multimedia*, pages 59–68, 2005.

[13] Bongwon Suh, Haibin Ling, Benjamin B. Bederson, and David W. Jacobs. Automatic thumbnail cropping and its effectiveness. *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 95–104, 2003.

[14] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.

[15] Julia Vogel, Adrian Schwaninger, Christian Wallraven, and Heinrich H. Bulthoff. Categorization of natural scenes: local vs. global information. *APGV '06: Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, pages 33–40, 2006.

