Multi-Operator Content Aware Image Retargeting on Natural Images

M Abhayadev1* and T Santha2
1-2Department of Computer Science, Dr G R Damodaran College of Science, Coimbatore, Tamil Nadu – 641014, India

Received 22 June 2018; revised 13 December 2018; accepted 04 March 2019

The image retargeting can be done in two methods, i.e. the traditional image retargeting methods (such as cropping and scaling) and content aware image retargeting methods. This research paper proposes a Multi-Operator Content Retargeting Image resizing Technique (CRIST) for image retargeting. The proposed retargeting operator is a combination of the visual saliency based scaling operator and the seam carving. The efficiency of the proposed research is experimented on a natural image dataset. CRIST works well on all natural images based on subjective and objective quality metrics.

Keywords: Target Prediction, CRIST, Filters, DoG, Erosion, Peak-Signal-to-Noise Ratio

Introduction

Image media retargeting has been retained as a very important research topic in the digital image processing research. In order to protect certain visual saliency in important areas, so many resizing methods are introduced by Zhang Luming and Yan Bo, using the significance map and the importance map which are based on local low levels and high level features. The table no.’1’ represents different types of content aware image retargeting operators in the image resizing research. Authors introduced an efficient retargeting of shadow images by using Multi-Operator improved CRIST. Another article of the same authors also proposed a content aware image seam carving technique for object resizing which could preserve the salient objects in images.

Materials and methods

Image retargeting is generally considered as one of the best content aware image resizing technologies in the area of digital image processing research. The proposed technique will work better on the natural images and the proposed method efficiency could be compared to the existing state-of-the-art image retargeting operators. The proposed Multi-Operator Content Retargeting Image reSizing Technique (CRIST) contains three stages. They are the following

• Image object boundary identification by using enhanced high pass frequency filtering algorithm and morphological erosion structuring element.
• Location based visual saliency map generation.
• Proposed natural image retargeting based on hybridized Multi-Operator.

Figure 1 presents the framework of the proposed technique for content aware image retargeting. This framework is explained in the following sub-sections:

Image object boundary identification by using enhanced high pass frequency filtering algorithm and Morphological Erosion Structuring Element

The first stage in content aware image retargeting is the process of identifying the image objects boundary. This particular method uses a high pass frequency filtering algorithm followed by morphological erosion structuring element to identify object boundary. The Difference of Gaussian (DoG) high pass filter is applied for identifying the object boundary. The high pass filter works efficiently in finding the high energy values, gradient amplitude value of pixels in input images. Image thresholding will segment the image into different parts according to the region of significance and less significance. Finally the important object boundary will be normalised by using morphological erosion structuring element.

Location based visual saliency map generation

The proposed visual saliency map works on the location based saliency map technique. The important image’s object features are identified by using both top-down and bottom-up saliency computation methods. The Location Based Visual Saliency
Computation Map (LBVSM) is more suitable for natural image saliency map generation because natural image contains different low and high level objects. The proposed Spatial Domain Location Based Saliency Map identifies the local and global features along with Chromaticity and Structural SIMilarity (SSIM) index measure of the natural input image objects. The combined result of the top-down and the bottom-up Saliency computation will improve the result of visual saliency map.

Proposed natural image retargeting based on hybridized multi-operator

The proposed Multi-Operator is a combination of two sub operators; Visual Saliency based Scaling and Seam Carving. Initially, the natural image object boundary identification is done by using high pass frequency filtering algorithm and morphological erosion structuring element. Visual Saliency Map of the input image object is to be calculated in the next stage. The Saliency Map helps to identify the important objects present in an input image and the content aware image Scaling and Seam Carving will be carried out on the input natural images. The proposed combined hybridized image resizing operation is known as Content Retargeting Image resizing Technique (CRIST). The Subjective Analysis and Objective evaluation experiment being done on the natural image datasets show that the proposed method gives good results. The proposed Multi-Operator content aware image retargeting technique outperforms other traditional approaches on the natural images.

Image objects target predictions

The boundary of an image object is also the result of changes in the light-weight, colour, shade and texture. These changes may not resolve the depth, the size orientation and the surface properties of a digital image. Detection of those edges is incredibly troublesome and time intense particularly once an image is corrupted by noise. Implementation stages for proposed boundary detection method is given below.

Step1: A $3 \times 3$ matrix is proposed for masking.
Step2: Using the mask, Difference of Gaussian filter is applied to the input natural image.
Step3: Compute gradient energy of the Difference of Gaussian filtered image $Q(i,j)$ using equation

$$\theta(i,j) = \arctan\left(\frac{f_y(i,j)}{f_x(i,j)}\right)$$

Step4: Calculating the optimal threshold $T^{(t+1)}$ of natural mage using the equation $T^{(t+1)} = \frac{\mu^b + \mu^o}{2}$

Step5: Key feature extraction and boundary tracking using $3 \times 3$ tracking window process.
Step6: The binary image of edge identified natural images taken as an input from morphological erosion structuring element, $A \ominus B$ which is calculated using equation $A \ominus B = \{z \mid (B)_z \subseteq A\}$

Step7: The continuous edge of natural images is extracted and important objects are identified.

The proposed method initially identifies the image object boundary using high pass frequency algorithm and morphological erosion structuring element, and the method’s output result is shown in the figure (figure 2).
Image object’s visual saliency feature identification

Computer vision is an electronically duplicate system similar to the abilities of the human vision system. The visual saliency computation can be described as predicting, locating and mining the salient visual information from an image. The saliency model mainly focuses on the selective characteristic of image attention. The implementation stages for proposed visual saliency method are given below. This phase comes after the image object identification.

Step1: Compute the location based feature selection saliency map of the input image \( I \), to get \( S_C \) the final importance map.

Step2: Compute the image \( I \) intensity value using equation, \( I = (r + g + b)/3 \), where \( r, g, b \in [0,1] \)

Step3: Compute the Image \( I \) normalized red, green and blue channels using equations

\[
\hat{r} = \begin{cases} r/1, & \text{if } I \geq 0.1 \\ 0, & \text{otherwise} \end{cases}, \quad \hat{g} = \begin{cases} g/1, & \text{if } I \geq 0.1 \\ 0, & \text{otherwise} \end{cases}, \quad \hat{b} = \begin{cases} b/1, & \text{if } I \geq 0.1 \\ 0, & \text{otherwise} \end{cases}
\]

Step4: Compute the orientation features from the intensity channels of image \( I \) by convolving it with Gabor filter in four directions using equation,

\[
G_{\psi}(x, y, \theta) = \exp \left( -\frac{x^2 + y^2 \gamma^2}{2\delta^2} \right) \cos \left( 2\pi \frac{x \cdot \lambda}{\lambda} + \psi \right).
\]

Step5: Compute the local irregularly map by adding intensity, colour and orientation maps.

Step6: Compute the high level features of the image \( I \) using Transformation function equation,

\[
f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v)e^{-j2\pi(ux + vy)} \, du \, dv.
\]

Step7: Compute the final normalized Saliency Map \( S_n(x, y) \).

\[
S_C = \alpha S_n + \beta S_g
\]

Here \( S_n(x, y) \) is the combined Top and Bottom Saliency Map. \( \alpha, \beta \) are the scaling factors. \( S_g(x, y) \) is the gradient energy of the input image \( I \).

From the figure (figure 2) top row (from left to right) shows a Palmyra Palm image and the proposed visual saliency of Palmyra Palm image.

Properties of the natural image dataset

This research was carried out on a natural image dataset. All 900 images are captured using canon EOS 600D, EOS digital SLR and compact system camera. A fully-labelled natural image dataset provides a unique resource for image retargeting research inquiries and image object analyses in many computational research fields, such as machine vision, machine learning and image resizing. The current research dataset is titled CRIST900, and these images are captured from two localities i.e. Annamalai(Pollachi District, Tamilnadu) and Mannarkkad (Palakkad District, Kerala). These images are captured from January 2015 to December 2016. The entire 900 images are categorized under 9 labels and each label contains 100 images. The label consists of nine different image categories including: Farmers, Mulching, Kālapūṭu (Bullock Race), Vegetables, Flowers and Trees, Farming Machineries and Tools, Germination, Paddy Field and Plough Land. The CRIST900 does have some limitations and the dataset has a variety of different natural objects such as humans, animals, flowers, trees and so on. The created dataset limitations are trivial and could be addressed in future research datasets.

Performance evaluation measures

The image quality assessment metrics used in this research are mainly objective and subjective metrics. One of the techniques in the objective matrices is Peak Signal-to-Noise Ratio (PSNR). PSNR is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of
its representation Hegenbart Sebastian and Andreas Uhl\textsuperscript{(14)}. Peak Signal-to-Noise Ratio is usually expressed in terms of the logarithmic decibel scale.

\[
\text{Mean Square Error (MSE)} \equiv \frac{1}{M \times N} \sum_{(i,j)} |Y(i,j) - \hat{Y}(i,j)| 
\]

\[
\text{Peak Signal-to-Noise Ratio (PSNR in dB)} = 10 \log_{10} \left( \frac{255^2 \times M \times N}{MSE} \right) 
\]

The next evaluation metrics used in this research paper are Receiver Operating Characteristic (ROC), Recall, Precision and Mean Opinion Score (MOS). The Receiver Operating Characteristic is the most popular metrics used to measure the performance of visual saliency model. The receiver operating characteristic takes to perform as a binary classifier and it evaluates the performance of the saliency map. Based on the threshold value, the proposed saliency maps are binarized into forward and backward region and this binarization process is for calculating the true positive, true negative, false positive and false negative values.

**Results and Discussion**

The proposed method generates image resizing result by using the created natural image dataset. And this method uses another image dataset in order to compare the resizing operator's efficiency with the proposed state-of-the-art operator results. Similarly the result of the object boundary identification method is also compared with the state-of-the-art boundary identification methods, and the efficiency is calculated on basis of different quality assessment matrices. The table 2 shows the proposed method objective Peak Signal-to-Noise Ratio (PSNR) quality assessment results of different boundary identification method on the selected experimented natural images.

The second stage of this research paper is to create a visual saliency map for content aware image retargeting. It has proposed Location Based Feature Selection Saliency Map. The location based feature selection visual saliency computation maps more suitable for natural image saliency map generation because the natural image contains different low and high level objects. The Fourier transformation functions are used to detect high level features from the images as the transformation function is an essential tool in the image processing research. The combined result of the top-down and the bottom-up saliency computation will improve the result of the proposed visual saliency map. The saliency map efficiency was calculated using Receiver Operating Characteristic (ROC) quality evaluation metrics. The Area Under the Curve (AUC), Standard Error (SE) and Test Comment are shown in table 3. AUC (Area under the ROC curve) score is an area under the curve calculated in the ROC space. In table 3 the AUC Score is 0.8684; this shows that the performance level of receiver operating characteristic curves is good. The final stage is the Multi-Operator. Single image retargeting operator does not produce finest visual object contents. The content aware scaling was done using the proposed combined location based saliency map and gradient map. In the next stage, the adaptively scaled image goes through vertical or horizontal seam carving, and the Multi-Operator retargeted image visual contents information is preserved by the proposed operator. The retargeted image quality is calculated based on the user study. The user study result shows, that the proposed method visual saliency preservation is good when compared to the other state-of-the-art operators. The original reference image and the retargeted image’s quality accuracy calculation are done by using different subjective and objective quality assessment matrices. It is well evident in the images (figure 2) that the Palmyra Palm image after retargeting, reduces 60% of its width. Here also there is any difference in the height of the input Palmyra Palm image. The proposed image retargeting operator was experimented on flicker benchmark images for comparing its results with other state-of-the-art retargeting operators. The comparison study was conducted with traditional resizing operators such as image scaling and image cropping and also comparing with content aware operators with the same set of inputs. The traditional image scaling and image cropping result was compared with the proposed CRIST. The retargeting is attempted to make the width of the image 40% of the original width without any change in the original image height. From the

<table>
<thead>
<tr>
<th>Images</th>
<th>Proposed Method</th>
<th>Canny</th>
<th>Prewitt</th>
<th>Sobel</th>
<th>LoG &amp; SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tomato</td>
<td>12.4223</td>
<td>11.3130</td>
<td>10.1396</td>
<td>7.3928</td>
<td>11.0019</td>
</tr>
<tr>
<td>Palmyra Palm</td>
<td>10.2722</td>
<td>9.3101</td>
<td>8.4522</td>
<td>7.1208</td>
<td>10.0110</td>
</tr>
<tr>
<td>Farmer 1</td>
<td>5.9362</td>
<td>5.0991</td>
<td>5.0162</td>
<td>4.6002</td>
<td>5.0091</td>
</tr>
<tr>
<td>Chayote fruit</td>
<td>5.8997</td>
<td>4.9012</td>
<td>4.1567</td>
<td>4.9022</td>
<td>4.0091</td>
</tr>
<tr>
<td>Kālapūttu 1</td>
<td>11.9276</td>
<td>11.1453</td>
<td>9.7435</td>
<td>8.8944</td>
<td>10.5093</td>
</tr>
</tbody>
</table>
Figure 3, it is evident that the proposed image resizing operator, CRIST, produces better results than that of other state-of-the-art image resizing techniques. By observing the image titled ‘Sailors at St. Malo’, it is found, that in scaling and cropping operator’s output results, the sailor’s boat shape is not in a good viewing condition. The middle image in the second row of figure 3 shows that the scaling operator’s result does not consider the image object background (i.e. sea area). The third image in the second row, in figure 3, shows the proposed algorithm’s result preserves the image artifacts, and resizes the image efficiently. To prove the performance efficiency of the proposed retargeting approach, the research uses mean opinion score (MOS) techniques. The method is then compared with the existing image retargeting methods. Two images are displayed on the screen; the one is the source image for reference, and the other is the retargeted image to be evaluated. The subjective viewers could judge the perceptual quality of the retargeted image and provide a rank for their operators. In order to evaluate the proposed retargeting operator’s efficiency, the research has taken five different retargeting operator’s results in cropping, scaling and CRIST. Sixty participants 40 male, 20 female and age range 20 years to 40 years participated in this subjective quality assessment study. All the participants belong to the computer science backgrounds. The study uses 60 natural input images as the source image, and all the 60 images are retargeted for evaluation in this experiment. The resolution of the screen for the subjective study is 1920 × 1280 which is sufficient for displaying the images in their original resolution. In this experiment, the computer monitor shows the original natural image from the created dataset and the proposed operator result along with the cropping and scaling competitors. For each set of images, the input original image is separately shown in the first row in the computer monitor, while the different retargeted operator’s results are randomly displayed in two additional rows within the same page. The image quality scales are labelled as ‘Bad’, ‘Poor’, ‘Fair’, ‘Good’, and “Excellent” which has numerical values ranging from “1”, “2”, “3”, “4” and “5”. All the subjective quality assessment participants have given numerical values for each retargeting operator’s image result based on the quality comparison of the reference image and the resized image. Larger the subjective value obtained by user opinion rating of each output image, then better the retargeting operator’s image perceptual visual quality.

The Mean Opinion Score (MOS) = \( \frac{1}{N} \sum_{i=1}^{N} Y_i \) ... (3)

In Table 3 Cropping, Scaling and Proposed operator Mean Opinion Score results are shown. Based on the statistics, the proposed operator outperforms all competitors in general. The subjective experiment did not provide a time constraints for the decision time, but the participants were instructed to finish the tests within 20 minutes. The Table 2 shows that the proposed technique’s subjective Mean Opinion Score is high compared to the other content oblivious retargeting

<table>
<thead>
<tr>
<th>Methods</th>
<th>Subjective Value (MOS)</th>
<th>Memory Requirement(KB)</th>
<th>Research Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed CRIST</td>
<td>280</td>
<td>250 - 400</td>
<td>MatlabR2015a</td>
</tr>
<tr>
<td>Cropping</td>
<td>200</td>
<td>150-300</td>
<td>MatlabR2015a</td>
</tr>
<tr>
<td>Scaling</td>
<td>210</td>
<td>150-400</td>
<td>MatlabR2015a</td>
</tr>
</tbody>
</table>

<p>| Table 3 The Statistics of User Study Result and Image Storage Comparison for Different Retargeting Operators |
|---------------------------------------------------------------|---------------------------------------------------------------|---------------------------------------------------------------|</p>
<table>
<thead>
<tr>
<th>Methods</th>
<th>Subjective Value (MOS)</th>
<th>Memory Requirement(KB)</th>
<th>Research Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed CRIST</td>
<td>280</td>
<td>250 - 400</td>
<td>MatlabR2015a</td>
</tr>
<tr>
<td>Cropping</td>
<td>200</td>
<td>150-300</td>
<td>MatlabR2015a</td>
</tr>
<tr>
<td>Scaling</td>
<td>210</td>
<td>150-400</td>
<td>MatlabR2015a</td>
</tr>
<tr>
<td>The Proposed Saliency Map Receiver Operating Characteristics (ROC) Curve Analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>0.86838</td>
<td>0.02801</td>
<td>Good test</td>
</tr>
</tbody>
</table>

Figure 3 - The State-of-the-art Retargeting Operator’s Comparison Result 1
The figure 3 is the image named ‘Sailors at St. Malo’. The first row Sailors at St.Malo image and second row (from left to right) is the Sailors at St.Malo image retargeted using Scaling, Sailors at St.Malo image retargeted using Cropping and Sailors at St.Malo image retargeted proposed CRIST.
operators. It shows the proposed image retargeting operator’s visual saliency has important object feature preservation ability in this content aware image resizing research field. The table (table 3) also gives the retargeted image storage comparison results of various content aware image resizing operators. Those are Cropping, Scaling and Proposed operators retargeted image storage space results are shown in this table. The proposed image retargeting method resize images into 250 to 400 kb memory required size without losing the image object content information with the comparison on remaining methods.

Conclusion
Image retargeting is an imperative research area because it demands much more knowledge reproduction and there is scope for research in the proposed Multi-Operator Content Retargeting Image reSizing Technique (CRIST). The proposed retargeting operator is a combination of Visual Saliency based Scaling operator and Seam Carving and the efficiency of the proposed research is experimented on natural image dataset. It works good on all natural images based on Subjective and Objective Quality Metrics. In future, research can be done on many complex agricultural and natural images. Also, the researchers can focus on video resizing by using the proposed method.

Acknowledgements
We express our respect and gratitude for the great help and contributions of Mr. Aneesh T K (Senior Video Editor Amrita TV, Calicut, Kerala), Mr. Arun Krishnan P (Field Sales Officer in Hegde&Hegde Pharmaceutical LLP, Calicut, Kerala) and Mr. Sakthi Shiva Kumar (Agricultural Scientist and Researcher, Coimbatore, Tamilnadu) in collecting the natural image dataset CRIST900 and for giving many valuable suggestions in completing this research.

References
4. Abhayadev M & Santha T, Content-aware image seam carving technique for object resizing, Saba J of Infot Tech and Network, 5 (2017b) 24-29