PREDICTION BASED SEAM CARVING FOR VIDEO RETARGETING

A DISSERTATION

Submitted in partial fulfilment of the requirements for the award of degree of

MASTER OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING

By

HARPREET KAUR

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY

ROORKEE – 247 667 (INDIA)

MAY – 2016

Declaration

I declare that the work presented in this dissertation with title, "**Prediction Based Seam Carving for Video Retargeting**", towards the fulfilment of the requirements for award of the degree of **Master of Technology** in **Computer Science & Engineering**, submitted to the **Department of Computer Science and Engineering**, **Indian Institute of Technology-Roorkee**, India, is an authentic record of my own work carried out during the period from **June 2015 to May 2016** under the guidance of **Dr. Debashis Sen,** Assistant Professor, Department of Electronics and Electrical Communication, Indian Institute of Technology, Kharagpur and **Dr. Vaskar Raychoudhury**, Assistant Professor, Department of Computer Science and Engineering, Indian Institute of Technology, Roorkee.

The matter presented in this dissertation has not been submitted by me for the award of any other degree of this or any other institute.

Date:

Place: Roorkee **(Harpreet Kaur)**

Certificate

This is to certify that the statement made by the candidate in the declaration is correct to the best of my knowledge and belief.

Date:

Place: Roorkee

Dr. Vaskar Raychoudhury

Assistant Professor

Department of Computer Science and Engineering

Indian Institute of Technology, Roorkee

Dr. Debashis Sen

Assistant Professor

Department of Computer Science and Engineering

Indian Institute of Technology, Kharagpur

Publication

1. Harpreet Kaur, Swarnjeet Kour, Debashis Sen, "**Prediction Based Seam Carving Method for Video Retargeting**," submitted in **International Conference on Pattern Recongnition, Cancun, Mexico, 2016.**

ABSTRACT

Image Retargeting is content aware image resizing which takes into account the resizing of image according to the resolution of display screen but without distorting salient content of image and without the loss of important information. This is challenging task as important areas of image must be preserved to maintain its aesthetics while resizing it.Image resizing techniques like scaling and cropping does not produce adequate results as scaling scales the image equally from all directions distorting the fineness of objects and cropping may remove important information from image.

Similar to image retargeting is the video retargeting where video is transformed to fit in an arbitrary display while considering the saliency of frames in video. As the video is sequence of frames inducing motion and human eye is more sensitive to movements. So in addition to saliency, retargeting should consider the temporal coherency as well where pixels are removed from almost similar position in subsequent frames so that similar kind of data is removed from each frame to maintain smoothness in videos. Temporal coherency is essential in video so that frames are aligned in similar way as if they were in original video. Hence video retargeting aims to have utmost balance between spatial coherency and temporal coherency.

Various algorithms are proposed on warping based techniques, patch based

techniques, seam based technique to retarget videos. Each has its own advantages and disadvantages. We presenta prediction based spatio-temporal seam carving scheme for video retargeting. Itresizes the video maintaining appropriate balance between spatial and temporal coherence. In a video frame, the proposed approach finds a‘temporal’ seam by using Kalman filter estimation and then modifies it with the help of ‘spatial’ seam considering both spatial and

temporal coherency. Unlike image retargeting, it is of utmost importance in retargeting a video frame to consider temporal coherency along with spatial coherency to remove or replicate unimportant background portion. This will ensure that insignificant amount of motion artifacts are introduced during resizing. The proposed Kalman filter based approach not only predicts a spatio-temporal seam to mark a portion of the frame where there is more possibility of having spatially and temporally coherent seam, but also has low time complexity. The proposed approach outperforms other state-of-the-art video retargeting methods which are illustrated by experimental results.

Acknowledgements

I would never have been able to complete my dissertation without the guidance of my supervisor, help from friends, and support from my family and loved ones.

I would like to express my deepest gratitude to my co-supervisor, **Dr. Debashis Sen**, for his excellent guidance, meaningful insights and moral support. He has been supportive since the day I began working on this dissertation and gave me the freedom I needed to explore this area of research on my own, while pointing me in the right direction in the times of need. His comprehensive knowledge in the area of Video Retargeting and hard working nature has been a constant source of inspiration.

I am highly grateful to my supervisor, **Dr. Vaskar Raychoudhury,** for his moral support, continuous encouragement in doing research, and gave me the freedom to explore in this area.

I am also grateful to the **Dept. of Computer Science, IIT-Roorkee** for providing valuable resources to aid my research.

I would like to thank **Nimita Mangal, Swarnjeet Kour, Sartaz Kanwar and Sweta Arya** and my other friends who supported me, were always willing to help and give me their best suggestions.

I would also like to thank my lab attendant **Raj Khati** who motivated me throughout the course.

Finally, hearty thanks to **my parents and siblings**, who encouraged me in good times, and motivated me in the bad times, without which this dissertation would not have been possible.

Dedication

To my parents, for giving me the best education they could

Table of Contents

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

With the growth in digital technology, digital videos have become an important component of media. A wide variety of displays are used to view them, like television, mobile, tablet, monitor, which have different resolutions, sizes, format, etc. Hence, digital media have to experience changes in size and aspect ratio to adapt to different resolution screens. Video retargeting is resizing the video to the screen resolution while preserving its salient content. Figure1.1 illustrates the process of video retargeting where original frame on left side of Figure 1.1 is converted to retargeted frame as shown on right side of Figure 1.1, preserving all salient content of image. It is gaining popularity these days and has wide range of applications.

Figure 1.1 Video Retargeting (source [2])

1.1. Video Retargeting Problem

A video is a sequence of frames placed at regular interval of time. It can be treated as 3D cube having frames in space-time and frame is 2D space having rows and columns. A frame of size m x n constitutes of m rows and n columns. Each pixel in frame has a value depending on type of frame. For black & white video, pixels are given intensity value while coloured video, pixels are given [r, g, b] value that is, red, green, blue values respectively. The video retargeting problem can be stated as resizing a video v of resolution m × n to a new video v' of resolution m' ×n' such that video v' is a good representative of the video v where m' and n' may be smaller or greater than m and n respectively. Video v' is a good representative of the Video v if it satisfies following constraints:

- Retargeted video v' should contain important content of original video v
- Retargeted video v' should maintain structure of original video v
- Retargeted video v' should not contain any distortions
- Retargeted video v' should remove similar kind of pixels from all frames to maintain smoothness of videos.

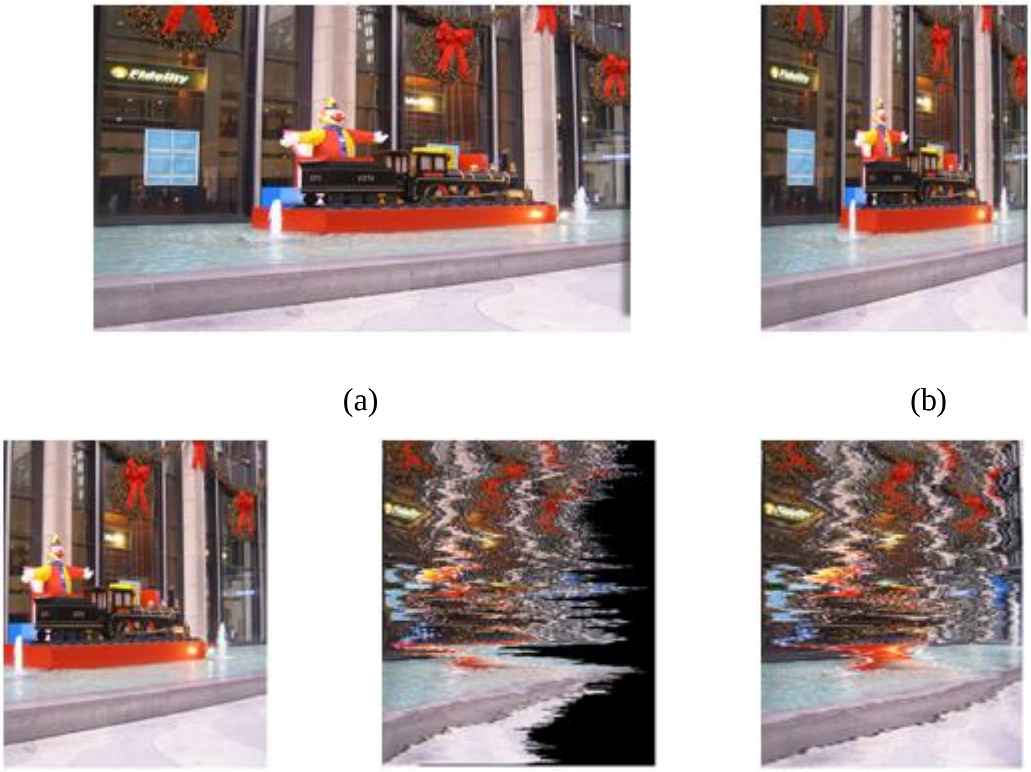
Scaling and cropping cannot provide desired results as scaling scales the frame equally from all directions which may distort the salient content in video while cropping will crop them from the boundary which may have cause important information loss.

1.2. Resizing methods for Video Retargeting

- Scaling**:** It is linear transformation that enlarges or shrinks frame by some scale factor which is same in all directions. Scaling do not consider content of image and it scales image uniformly which cannot give satisfactory results as shown in Figure 1.2(b). Figure 1.2(a) shows original frame.

- Cropping**:** Cropping removes the pixels from the periphery of frames irrespective of content of image. It preserves both shape and visual coherence but it removes the boundary pixels which can contain useful information. It is not maintaining quality of frames. Figure 1.2(c) explains this.

- Find the pixels having lowest energy from frame in increasing order of their energies and then remove them. It is optimal method but it can destroy rectangular shape as well as visual coherence of frame as shown in Figure 1.2(d).
- Remove the pixels from every row that has low energy. It will also make the frame optimal but it creates zig-zag effect in frame as shown in Figure 1.2(e).



(a) (b)

(c) (d) (e)

Figure 1.2 Resizing Methods (source[1])

1.3. Seam Carving for Image Retargeting

A content aware resizing can be represented as retargeting which give a way for various retargeting methods. Seam carving is one such method for retargeting images given by avidan et al[1]. It removes the 8-connected path of pixels running from left to right for horizontal seams and from top to bottom for vertical seam. It takes into account an energy function to form the low energy seams using dynamic programming. These low energy seams are removed or added in images to reduce or increase the resolution of images. However, this image seam carving method, if directly employed on videos, can create motion artifacts. Both spatial and temporal coherency of seams needs to be considered in video seam carving while treating frames in space-time as 3D cubes. Seams thus formed in video are taken as 2D surface as explained in Figure 1.3.

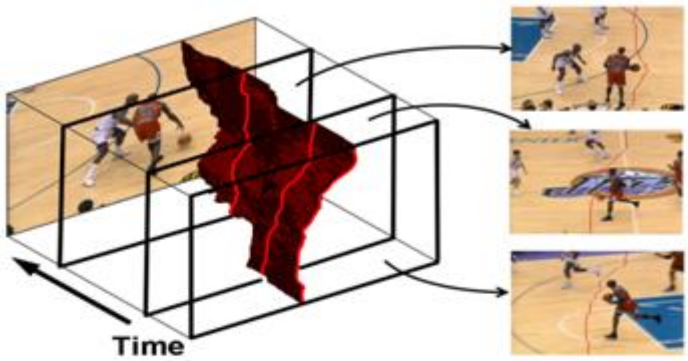


Figure 1.3 Seams as 2D surface in 3D video cube (source [3])

1.4. Motivation

Seam carving technique for image retargeting is a content aware method for resizing images, which considers spatial coherency among pixels. In order to use them in videos, the motion information needs to be considered in such a way that motion distortion like abrupt nonexistent motion changes are not introduced. Maintaining temporal coherency among subsequent frames after the retargeting ensures that motion distortions are not introduced. Seams removed/added from subsequent frames should be such that they are neighbours of containing pixels to maintain smoothness of videos. The state-of-the-art algorithms do not consider temporal coherency directly computed from neighbouring frames. Therefore, there is a scope for improvement by considering both spatial and temporal coherency of seams in video seam carving. Moreover, a large number of computations is required in video retargeting due to the large number of frames in video, and hence, memory requirements and computation speed also need to be optimized. These scopes serves as a motivation to design a video retargeting algorithm that properly balances spatial and temporal coherency in such a way computational complexity encountered is less.

1.5. Video Retargeting Process

Basic flow of video Retargeting is shown in Figure 1.4.

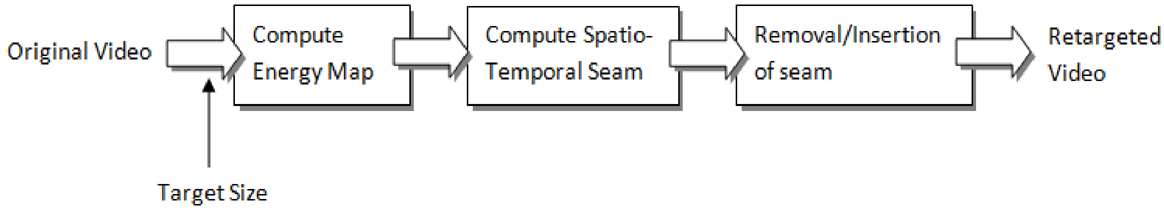


Figure 1.4 Video Retargeting Flow

Video and its targeted size is given as input. Energy map of each frame of input video is computed which quantify the importance of each pixel in the frame. Important regions in the video are given more importance value than unimportant regions so that important regions can be preserved. Important areas in video can be found by object detectors, face detectors, by using gradient energy technique etc. One of the method given by [2] uses four basic principles to find salient features like contrast, colour, suppressing frequently occurring features, high level factors like face. User can also specify constraints on calculating importance map by providing saliency map, structure preservation conditions etc. The basic method for calculating energy map is gradient energy which is given by

$$e = \left|\frac{\partial I}{\partial x}\right| + \left|\frac{\partial I}{\partial y}\right| \tag{1}$$

Apart from gradient energy, other advance techniques can also be used like saliency[7], entropy, histogram of gradients but our main focus is to find seams for video retargeting so most basic energy criteria is taken into account. Next step in this process is computation of spatio-temporal seam. It is actually a combination of spatial and temporal seam where spatial seam is computed by [1] using dynamic programming given as

$$M(a,b) = enr(a,b) + \min((M(a-1,b-1), M(a-1,b), M(a-1,b+1)) \tag{2}$$

Where enr(a, b) is energy at pixel (a, b) of frame calculated from (1). Further, all possible seams are analyzed for each entry (a, b) and cumulative minimum energy, M is computed for each seam taking into account energy for all pixels in a seam. The last row having minimum value of M is having minimal connected vertical seam whose optimal path is obtained by backtracking from this minimum entry. This optimal path act as spatial seam given by

$$X_{spatial} = \min E(s) \quad (3)$$

$E(s)$ is cost of all seams computed from equation (2). Temporal seam is computed considering the neighbors of seams removed/inserted from subsequent frames. These both seams are combined together to find the spatio-temporal seam such that seam have proper balance between spatial coherency and temporal coherency. Spatio-temporal seam is then removed or inserted from frames in order to obtain the retargeted video.

1.6. Applications

- Web pages containing images and videos keep on changing from browser to browser which encourages us to use techniques that can retarget the images and videos accordingly.
- Thumbnails are smaller size genre of images and video. So there is need to imply the retargeting for generating thumbnails.
- Mobile phones are becoming very popular these days but due to their limited resolution, retargeting problem is attaining popularity.
- Retargeting techniques are also useful in photography
- Documents containing embedded images also need to be resized according to the overall layout.

1.7. Thesis Contribution

We present a video seam carving algorithm that estimates a seam in a frame on the basis of current spatial seam and seams computed temporally earlier. Hence, the seam estimated is both spatially and temporally coherent. The proposed seam estimation is performed using Kalman filter. The Kalman filter temporally estimates a seam, and then updates it with respect to a nearby spatial seam having lowest energy. This seam is further combined with spatial seam computed by [1] without considering the temporal past.

Video retargeting is performed using the proposed seam carving algorithm and the performance is compared to the state-of-the-art video retargeting algorithms. Apart from qualitative analysis, the effectiveness of the proposed algorithm is demonstrated in terms of temporal coherence, saliency preservation and time complexity.

1.8. Thesis Organisation

Thesis is organized as follows.

- Chapter 1 introduces video retargeting, problem statement and motivation of doing research in this topic by giving some other resizing methods. We also summarize the basic process of video retargeting, its applications and our contribution made in this dissertation.
- Chapter 2 surveys the literature where research is done in video retargeting under various categories. We also briefly describe state-of-art algorithms implemented by us for comparing our proposed method with them.
- Chapter 3 discusses the proposed method for video retargeting which uses an estimation technique known as Kalman filter to find the spatio-temporal seams. It also explains about why we use this technique for our proposed method.
- Chapter 4 illustrates the results and experiments done by us on multitude of videos.
- Chapter 5 concludes the dissertation.

CHAPTER 2

LITERATURE SURVEY

Various resizing methods were introduced as explained in [1]. Scaling is linear transformation that enlarges or shrinks image by some scale factor which is same in all directions. Scaling do not consider content of image and it scales image uniformly which cannot give satisfactory results. Cropping can be done which removes the pixels from the periphery of image irrespective of content of image. It preserves both shape and visual coherence but it removes the boundary pixels which can contain useful information. These methods are not considering the content of image and do not produce satisfactory results. Hence there is need to have content aware methods for resizing the image so that important objects remain preserved in image and do not get distorted.

2.1. Warping-Based Methods

Warping based methods create nonlinear distortion to obtain the resized videos [7, 18, 19]. Important areas in videos are distorted less while unimportant areas are distorted more. Distortions can be added by either down-sampling or scaling down of unimportant areas to smaller size, which preserves the important contents of videos. This approach for videos is obtained from that for images by introducing additional constraints to maintain smoothness across the adjacent frames. Wolf et al [12] present warping based technique which works in two steps which is illustrated in Figure 2.1. It first find saliency map of video frames and then distorts or removes pixels from non-salient portions of frames. Saliency can be computed using face detectors and motion detectors as well. But it is not always good to remove the homogeneous portion of frames and it may destroy the structure of frames and may also distort the object.

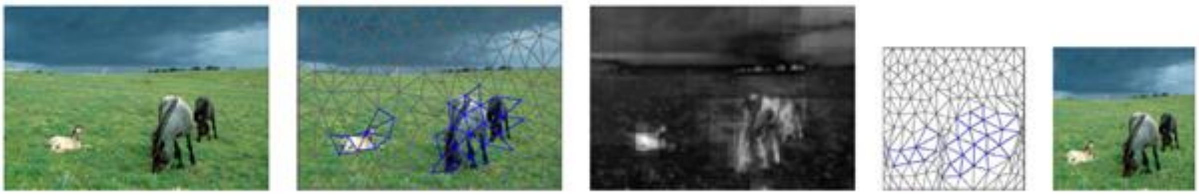


Figure 2.1 Warping based video Retargeting Technique (source [2])

2.2. Multi Operator Methods

These methods combine various operators to retarget the video as each method is having its advantages and disadvantages. So different methods are combined based on the requirements and kind of video to resize the video.

Liu and Gleicher [15] proposed a method which uses the combination of cropping and scaling for video retargeting. Dynamic video shots are represented by virtual pans and cuts in this method. This method optimizes the cropping and scaling by considering saliency of frames, face detection and motion detection. But cropping can remove the important information from frames. Also it treats shots independently which can be a challenge to retarget the videos having multiple shots.

Zhang et al [17] uses shrinkability map to retarget the video. The shrinkability map for each pixel in frames of video is defined by random walk where important pixels are given less shrinkability value than unimportant pixels. Scaling then exploits the target size to assign new position to each pixel in target frame using shrinkability map. It fails to retarget the video as temporal coherency is not handled by this method efficiently.

Tao et al [13] proposed another method where salient areas of frame are extracted by identifying foreground objects that are in motion because it is considering motion as important content in videos. An active window is taken to cover these areas and prevent them from distorting. It cannot produce satisfactory results due to absence of video understanding techniques.

2.3. Seam carving Based Methods

Rubinstein et al [3] proposed a seam carving based video retargeting method which computes the seam using graph cut approach. The method computes seam by finding minimal cut in graph. The cut is monotonic and connected. In addition, this method also considers the energy inserted into image after the removal of seam called as forward energy. Large amount of information about video frames needs to be stored and processed so this method is computationally very intensive and salient content of video frames can be distorted to maintain continuity of seam.

It is believed that no single video retargeting method can alone retarget the video efficiently. Hence, Rubinstein et al [16] proposed another method in 2009 based on seam carving technique but in aggregation with cropping and scaling. This method combines different operators in an optimal manner to maintain spatial and temporal coherency. But this method is very complex as it is difficult to find the number of operators which can have similarity with frame similarity measure.

Chao et al [14] proposed another seam carving method which uses coarse-to-fine temporal optimization to retarget video. This method uses energy functions, motion weight prediction, and pixel based optimizations to extract temporal information. The motion weight prediction reduces the search area of dynamic programming by predicting coarse location of seam in current location. The pixel-based optimization then maintain temporal coherency by exploiting pixel-based optical flow in the reduced search range.

Some of the most recent state-of-art algorithms are given by Matthias et al and Yan et al. These algorithms are using seam carving techniques to compute the spatio-temporal seam for video retargeting. These algorithms are performing better than other algorithms. Hence I have implemented these two methods to compare it with my algorithm. These methods are explained in next subsections.

2.3.1. *Discontinuous Seam Carving Method*

This method (given by [4]) is appearance based temporal coherence method which means that seams should be such that it maintains the appearance based consistency in temporal domain. It processes the video frame by frame and hence temporally discontinuous seams are computed. Seam removal is done sequentially frame by frame. It takes into account spatial cost (S_c), temporal cost (T_c) and saliency cost (S) which are combined in ratio S_c: T_c : S as 5:1:2 to get one measure M. Minimum cost seam is computed for the frame by minimizing value of M using dynamic programming.

- Measuring Temporal Coherence Cost

  Various steps to measure the temporal coherency cost are:

  - Remove seam S^i from frame F^i to get resulting frame R^i which should be visually close to most temporally coherent frame R^c. R^c is obtained by applying previous seam S^{i-1} to current frame F^i.
  - For each pixel (x,y), sum of squared difference is found which tells how much different is resulting frame from coherent frame.
  - Consider Figure 2.2, "The previous seam S^{i-1} in red is applied to current Frame F^i. If pixel B is removed from row, it results in ACDEF. The optimal temporally coherent seam removes pixel F from row, so that R^c would contain ABCDE. The temporal coherence cost for pixel B is |C-B|+|D-C|+|E-D|+|F-E|" [4].

- Measuring Spatial Coherency Cost

  Three scenarios are considered while removing the pixels in spatial domain.

  - Pixel A and B in Figure 2.3(a) are having color difference, but pixel C is having same color as B. So on removing pixel B no details are lost.
  - In Figure 2.3(b), if we remove B then high frequency will get lost because both pixel A and C are having different color from pixel B. So, It have high spatial error
  - In Figure 2.3(c), there is slight difference in color for pixel B and C. There will not be much change in appearance of image if pixel B will be removed. It have very small spatial error.

Spatial coherence cost $S_c = S_h + S_v$ where S_h is error in horizontal direction and S_v is error in vertical direction by removal of specific pixel. Horizontal direction considers pixels in same row while vertical direction takes pixels from adjacent rows to find gradients. Consider example in Figure 2.4, On removing pixel E and D from Figure 2.4(a) and (b) respectively, we get measure of horizontal gradients as $S_h(E) = |D- E| + |E- F| - |D – F|$, and $S_h(D) = |D – E| - |E – F|$. Similarly vertical gradients considering Figure 2.4(a) are: $S_v(E,B) = 0$, $S_v(E,A) =| |A – D| - |B – D| | + | | B – E | - | B – D| |$, $S_v(E,C) =| |C – F| - |B – F| | + | | B – E| - | B – F| |$.

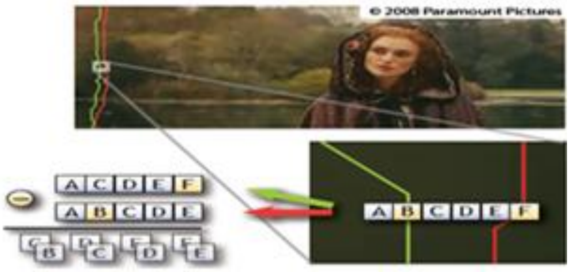


Figure 2.2 Temporal Coherence Cost (source [4])

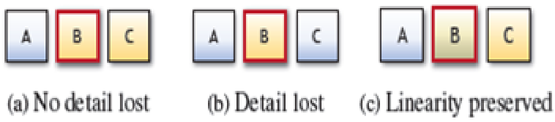


Figure 2.3 Spatial Error cases (source[4])

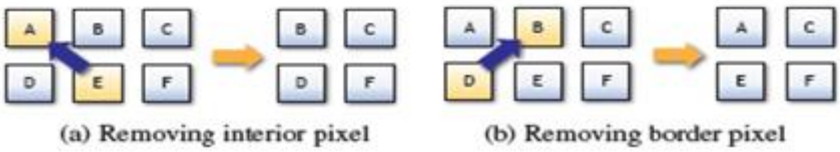


Figure 2.4 Spatial Coherence Cost (source [4])

Saliency cost is measured by techniques like face detection, object detection etc. Example below shows the seams for discontinuous seam carving methods where white colored seam is spatial seam, green colored seam is temporal one. By combining them, we get red colored seam which is final spatio-temporal seam as shown in Figure 2.5.

Figure 2.5 Seams produced by our implementation of algorithm by [4]

2.3.2. Matching Area Based Seam Carving Method

This Method, proposed by [5] is video seam carving method uses the concept of modifying energy map on the basis of matching area. The method divides the pixels of frame into reward/punish regions on the basis of their similarity with the pixels of previous seam. On the basis of this R/P region, energy map will be modified. Further, seam carving method for image retargeting[1] is applied to compute the seam in current frame. Figure 2.6 show algorithm for this method.

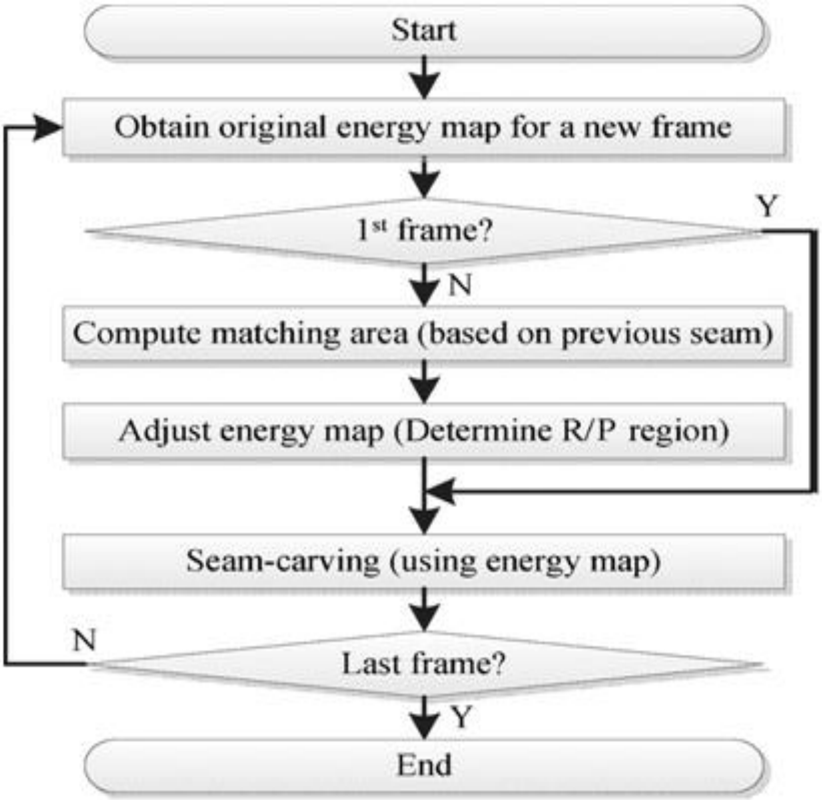


Figure 2.6 Matching-Area-Based Seam carving algorithm (source [5])

Energy map consists of gradients given by (1) which are used to find the seam in first frame of video using dynamic programming as mentioned in (2). Seam is divided into equal regions. The points having the maximum energy value in the particular region are selected as key points(KP). Hence the number of key points(KP) are equal to number of regions. The key points are referred in next frame to find reward/punish regions. In a frame, search areas (SA) are square areas whose center are KPs of previous seam and radius is given by search range(SR). "Match area centered at $P(X_p,Y_p)$ can be defined as $MA_i(P) =\{(x,y)|X_p - MW \leq x \leq X_p + MW , Y_p - MW \leq y \leq Y_p + MW \}$ where MW is match window set to 3" [5] and i are number of regions. Match index (MI) between P in the (i + 1)th frame and KP in the ith frame is computed for every pixel in SA using sum of squared differences. On the basis of match threshold of match index, pixels are divided into rewarded pixels and punished pixels and this information is stored in RPMap. Using this RPmap, energy map is adjusted to new energy map which further is used in seam carving. Figure 2.7 explains this algorithm and figure 2.8 shows example of this algorithm where red portion in second frame is punish region while green is reward region hence energy increased at red portions which lead to seam formed at other location. This example is taken as a result of our own implementation of this algorithm.

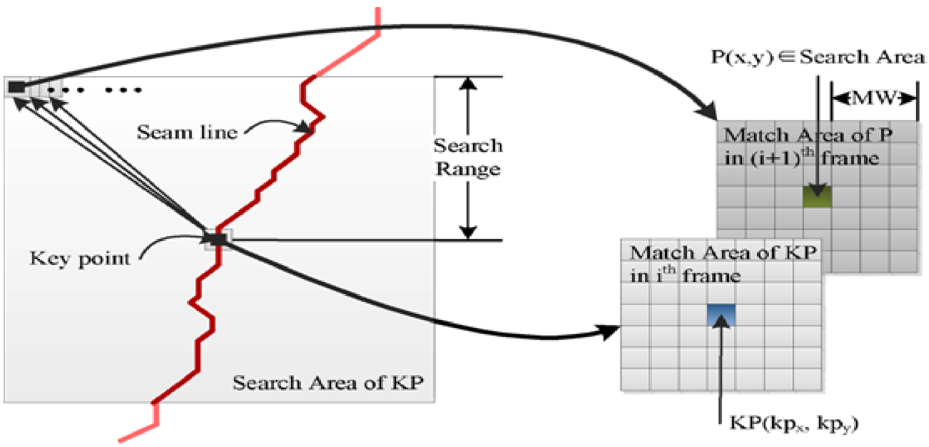


Figure 2.7 Matching-Area-Based seam carving algorithm explanation (source [5])

Figure 2.8 Matching area based seam carving results (own implementation)

Discontinuous seam carving technique for video retargeting takes into account the difference of appearance in retargeted and most temporally coherent frame to find the seam. The method also uses automatic spatio-temporal saliency where per-frame gradient based saliency is not sufficient. Spatio-temporal saliency is obtained by segmentation [20]. Though results are good but it can remove the pixels from different objects if video objects are moving very fast. Also spatio-temporal saliency uses segmentation which limit the video length to 30-40 seconds. Matching area based seam carving algorithm is aiming to adjust the energy by marking pixels in reward/punish regions but it may not always produce temporally coherent results as seam may move far from previous seam due to change in energy of frame.

The little artifacts in video may result in larger distortions as human eye is more sensitive to motion. So, video retargeting aims to have a content aware approach where the balance and maintenance of spatial and temporal coherence is utmost. Without this balance video will contain artifacts. Current research is heavily focused on this aspect of retargeting. Approaches explained in previous sections are computationally very expensive, the objective is to reduce complexity as well by not working on whole image to find seam but to search in area where there is large probability of finding the seam.

A video retargeting technique is envisioned where the balanced spatial and temporal coherence is achieved through prediction of energy in a frame. Prediction is done using kalman filter. Such a technique will result in smoother video retargeting and put less burden of the already computationally expensive seam-carving methods.

CHAPTER 3

KALMAN FILTER FOR SPATIO-TEMPORAL SEAM ESTIMATION (PROPOSED METHOD)

3.1. Kalman Filter in Seam Carving

Kalman filter [8, 9, 21] is also known as linear quadratic estimation (LQE). It is an algorithm that uses a series of measurements observed over time and produces estimates of unknown variables. Kalman filter consists of two phases: prediction and updating. The prediction phase uses a previous frame seam to estimate a current frame seam. In the updating phase, the predicted seam of current frame is combined with an observation in the current frame to refine the seam estimate. The observation is a spatial seam having lowest energy in a pre-defined neighborhood.

Kalman filter is popularly used nowadays in various applications like navigating and controlling vehicles particularly aircraft and spacecraft, planning robotic motion, signal processing etc. We present an algorithm which estimates the seam using Kalman filter on the basis of seams computed so far, hence giving temporally coherent seam. This temporally coherent seam is updated by taking into account Kalman gain and spatial seam having lowest energy and nearby to this seam. Further this updated seam is further combined with actual spatial seam computed by [1] to get spatio-temporal seam. This is iterative process which continues for all frames in video. Kalman filter goes hand in hand with both seams viz. spatial seam and temporal seam. Also it is estimating the position of seam from past history, so we get portion of frame where we can get a seam having supreme balance between spatial and temporal coherency which fastened the algorithm. The block diagram for our proposed work is given in Figure 3.1.

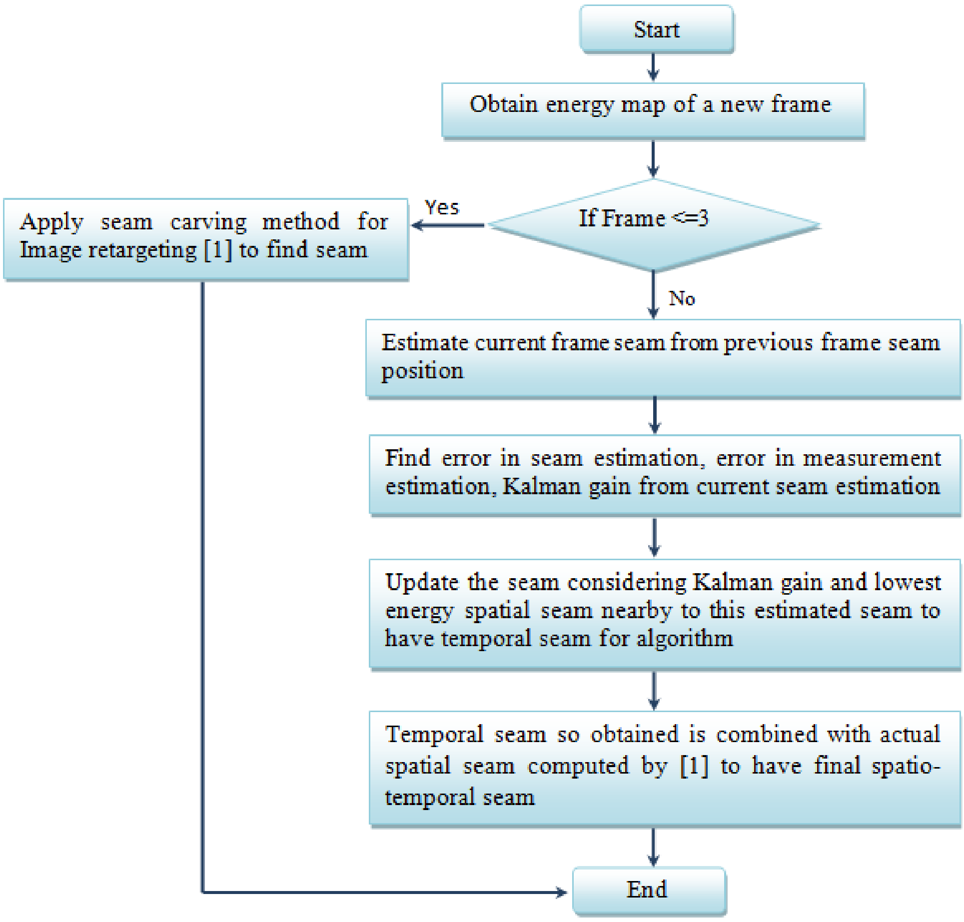


Figure 3.1 Flow Chart for Proposed Method

3.1.1. Seam Prediction

Current frame seam is considered as 2-component state having position and velocity given as.

$$X_t = \begin{pmatrix} Pos_t \\ Vel_t \end{pmatrix} \qquad (4)$$

Where X_t is set of pixels in a seam of frame t, represented as $\{T(1, y_1), T(2, y_2), T(3, y_3), \ldots.. T(i, y_i)\}$ where i is number of rows in frames, y_i is column location in ith row and Pos_t represents of position set of all the pixels. Similarly, Vel_t indicates the velocities for all pixels in seam.

As seam is changing its position from frame to frame, the current frame seam position can be used to find next frame seam position by applying standard equation of motion for position and velocity which are given as

$$Pos_t = Pos_{t-1} + Vel_{t-1} \times t + \frac{1}{2} \times acc_t \times t^2 \tag{5}$$

$$Vel_t = Vel_{t-1} + acc_t \times t \tag{6}$$

These equations of motion can be rendered in our model as

$$\mathrm{X_t} = \begin{pmatrix} \mathrm{Pos_t} \\ \mathrm{Vel_t} \end{pmatrix} = \begin{pmatrix} 1 & \mathrm{t} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{Pos_{t-1}} \\ \mathrm{Vel_{t-1}} \end{pmatrix} + \begin{pmatrix} \frac{\mathrm{t}^2}{2} \\ \mathrm{t} \end{pmatrix} \mathrm{acc_t} \tag{7}$$

Giving generalized equation as

$$\mathrm{X_t} = \mathrm{F_t} \times \mathrm{X_{t-1}} + \mathrm{B_t} \times \mathrm{u_t} \tag{8}$$

Where$\mathrm{X_{t-1}}$is previous frame seam position and $\mathrm{F_t}$, $\mathrm{B_t}$, $u_t$are various control inputs which applies the effect of previous frame seam position to find out current frame seam.

$$F_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, B_t = \begin{pmatrix} \frac{t^2}{2} \\ t \end{pmatrix}, \mathrm{u_t} = \mathrm{acc_t} \tag{9}$$

Let {F1, F2, F3........ Fn} be the frames of video and {S1, S2, S3........ Sn} be the set of pixels for vertical spatial seam (computed using (3)) in {F1, F2, F3........ Fn} respectively. Let a set Si be represented as {Si1, Si2, Si3…… Sim} pixels in frame i, where m is height of frame. Change in position of seam pixels from consecutive frames gives velocity which can be represented as set V= {V2, V3, V4…… Vn} where Vi = (Vi1, Vi2, Vi3…… Vim) and

$$Vi1 = \frac{\mathrm{Si1 - S(i-1)1}}{\mathrm{dt}} \tag{10}$$

Similarly change in velocity gives how much seam pixels accelerate between consecutive frames which can be represented as A={A3, A4, A5……An} where Ai = (Ai1, Ai2, Ai3…… Aim) and

$$Ai1 = \frac{Vi1 - V(i-1)1}{dt} \tag{11}$$

This act as control inputu_t for predicting seam. So seam prediction, and hence temporal seam computation, initiates from the 4th frame in a video.

Initially for a seam prediction, previous frame seam state is given by PS_{t-1} = [PS1, PS2, PS3……….. PSm] where m is height of frame and PS_j is 2x1 matrix represented as

$$\mathrm{PS_j} = \begin{pmatrix} S(t-1)j \\ V(t-1)j \end{pmatrix} \tag{12}$$

The seam measurement prediction, Z_t of current frame seam is function of current frame seam prediction having measurement vector C as 1x2 matrix given by

$$[Pos_t] = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} Pos_t \\ Vel_t \end{pmatrix} \tag{13}$$

which gives generalized equation as

$$Z_t = C \times X_t, C = \begin{pmatrix} 1 & 0 \end{pmatrix} \tag{14}$$

This seam prediction is updated with respect to spatial seam measurement. For this updation to happen, Kalman gain needs to be computed which takes into consideration the variance in seam prediction as well as the variance in measurement prediction. Covariance matrix gives the variance in seam prediction which is given as

$$P_t = F_t P_{t-1} F_t{}' + \mathrm{E_x} \tag{15}$$

Initially $P_{t-1}$is taken as 2x2 matrix which is represented as

$$P_{t-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{16}$$

E_x is error introduced in calculating current frame seam which is variance in position and velocity given as

$$\mathrm{E_x} = \begin{pmatrix} {\sigma_{Pos}}^2 & \sigma_{Pos}\sigma_{Vel} \\ \sigma_{Pos}\sigma_{Vel} & {\sigma_{Vel}}^2 \end{pmatrix} \tag{17}$$

Now from (7), $t^2/2$ is affecting position of seam while ‘t’ is affecting velocity. Hence error in seam prediction can be given as

$$\mathrm{E_x} = \begin{pmatrix} \frac{t^4}{4} & \frac{t^3}{2} \\ \frac{t^3}{2} & t^2 \end{pmatrix} \tag{18}$$

Kalman gain given by

$$K = P_t \times C'(C \times P_t \times C' + E_z)^{-1} \tag{19}$$

.

E_z is measurement error or we can say variance in measurement given by

$$E_z = {\sigma_{Z_t}}^2 \tag{20}$$

This variance is considered on the basis of spatial seam calculated from (3) having minimum energy and closest to the seam measurement predicted by (14), say it as NS. For finding nearest seam, we aretaking window of 100 pixels, 50 pixels on left of seam measurement and 50 pixels on right of it.

3.1.2. Seam Updation

Seam is updated using Kalman gain with formula given as

$$X_{temporal} = X_t + K\,(z_t - Z_t) \tag{21}$$

where$X_{temporal}$ is updated seam which act as temporal seam for our algorithm, z_t is actual observed measurement of seam which is seam having low energy and nearest to predicted seam, NS.If predicted seam measurement is same as actual observed measurement then our estimation is correct and need not to be changed as difference of predicted measurement and actual measurement will be zero. But if predicted measurement is not same as actual measurement then estimated position is having error which is corrected by Kalman gain, K to get more accurate estimation of seam in video frames. Covariance matrix is also updated to be used in next prediction by

$$P_t = (I - K \times C) \times P_t \tag{22}$$

3.1.3. Combining Spatial and Temporal seam

Next step in this algorithm is to combine spatial and temporal seam. For this, area between the spatial seam position and temporal seam position is considered and all possible spatial seams calculated by (3) in this area are considered as candidates for final seam. Let β_t be set of candidate seams in frame t represented as $\{CS_1, CS_2, CS_3, \ldots\ldots CS_r\}$ where r is number of candidate seams.

Two parameters are considered for finding the actual seam. One is the distance of candidate spatial seam from the temporal seam, say $\mathrm{distanceDiff}_t$ for frame t and another is energy difference of candidate spatial seam with the actual spatial seam, say it as $\mathrm{energyDiff}_t$ for frame t. These parameters are calculated for each of candidate seam in set β_t.

These two parameters are normalized and combined together as one measure, H_t given as

$$\mathrm{H_t} = (1 + \mathrm{distanceDifft_t})(1 + \mathrm{energyDifft}_{\,\mathrm{t}}) \tag{23}$$

Final seam is one having lowest measure, H_t given as

$$X_{final} = \min_{1 \leq j \leq r} G(CS_j) \tag{24}$$

Where X_{final} is final spatio-temporal seam and G is function that computes H_t for each of candidate seam CS_j. This final seam use to find control inputs, covariance matrix for next frame predictions. So, this is iterative process which continues for all frames in video.

3.2.Why Kalman Filter?

Temporal coherency is crucial for video retargeting and the seam computed considering temporal coherency should not remove or replicate the objects in the frames as well. Kalman filter is one of the methods, which considers the change of seam positions in the previous frames and from it estimates the position of seam in the current frames. Such seam prediction considering the seams in previous frames inherently imparts temporal coherency in our algorithm. Moreover, the update phase of Kalman filter considers spatially coherent seam with a local region as well. So in total, Kalman filter itself imparts both spatial and temporal coherency, which is required for video retargeting. In addition, the spatially coherent seam with minimum energy in the entire frame is also considered along with the Kalman filter estimated seam to get the final spatio-temporal seam.

On the other hand, matching area based seam carving video retargeting method [3] considers temporal coherency by considering the seam at the same position where it was in the previous frame. It then modifies the energy by comparing the portion of area around seam to the previous frame area. While such a technique may avoid removing object parts, there is a good chance that the seams are spatially separated by considerable amount in successive frames, especially if the object motion is large.

The other approach of discontinuous seam carving method for video retargeting [4] combines spatial coherency and temporal coherency to find the seam. But the ratio is predefined and may not be accurate is properly combining these coherencies. There is more possibility of such a scenario, as the ratio is kept the same for all frames of a video.

CHAPTER 4

RESULTS AND EXPERIMENTS

4.1. Experimental Results

We demonstrate our method on different kinds of videos with those having large number of objects, small number of objects, fast paced videos, slow paced videos, indoor and outdoor videos, etc., and in almost all cases it is showing satisfactory results. We compare our method with other seam carving based state-of-art methods proposed by Matthias et al[4] and Yan et al[5]. These methods are most recent contributions in video seam carving. Hence, our algorithm is to further the contributions by these algorithms such that balanced seams having proper mix of spatial and temporal coherency is achieved, which can give more smoothness to videos size change. We have analyzed our method on various criteria of retargeting which are demonstrated below.

4.1.1. Experiment 1: Object Cutting

Figure 4.1 compares the discontinuous video seam carving by Matthias et al[4], matching area based seam carving method by yan et al[5] with our proposed method in context of object cutting. Figure 4.1(a) shows that seam in red color computed by [4] cut through the dog as temporal cost is high at other position where spatial cost is low, even [5] cut through the dog in Figure 4.1(b) though it move the seam towards left to avoid the dog while our method is not cutting object as shown in Figure 4.1(c) as it is considering both spatial and temporal coherency with proper ratio and hence it removes the pixels from extreme left side of frame where there are no objects. Similar results are shown in Figure 4.2 where [4] and [5] cut through the horse while our algorithm avoids the seam through horse. Figure 4.3 is showing a toy in globe having statues inside them. Algorithm by [4] and [5] remove the pixels from them but our algorithm is not doing so.

(a) (b) (c)

Figure 4.1 (a) Seam computed by Discontinuous seam carving method (b)Matching area based seam carving method (c) and proposed method in video of dogs

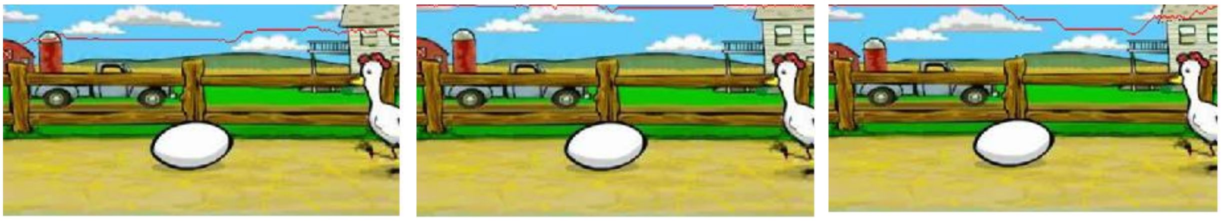
(a) (b) (c)

Figure 4.2 (a) Seam computed by Discontinuous seam carving method (b)Matching area based seam carving method (c) and proposed method in video of horse racing

(a) (b) (c)

Figure 4.3 (a) Seam computed by Discontinuous seam carving method (b)Matching area based seam carving method (c) and proposed method in video of toys

Not only for vertical seams but proposed method shows better results on horizontal seams as well. Figure 4.4 shows the example where algorithm by [4] cut through the house on right side as well as hut and pillar on left side of frame, algorithm by [5] continues to remove clouds while our algorithm is trying its best to preserve all the salient content and removing pixels from background portion only. Similar example shown in Figure 4.5 where buildings at back are distorted by algorithm by [4] and [5] but proposed method is preserving the buildings and removing pixels from portion where there is no salient content.



(a) (b) (c)

Figure 4.4 (a) Seam computed by Discontinuous seam carving method (b)Matching area based seam carving method (c) and proposed method in video of hen

(a) (b) (c)

Figure 4.5 (a) Seam computed by Discontinuous seam carving method (b)Matching area based seam carving method (c) and proposed method in video of boy playing with dog

4.1.2. Experiment 2: Temporal Coherency

Figure 4.6 demonstrates the temporal coherency of algorithm given by [4] and [5] with our proposed method. Discontinuous method for video seam carving takes into account temporal coherency in predefined ratio which cannot always produce temporal coherent results as illustrated in Figure 4.6(a) which shows two consecutive frames of videos having different positions of seam in red color. Figure 4.6(b) represents seam positions in two consecutive frames of video computed by [5] where seam move far away from previous seam position due to modification in energy of frame and then simply using seam carving technique for image on this frame. Our algorithm maintains temporal coherency, as shown in Figure 4.6(c).

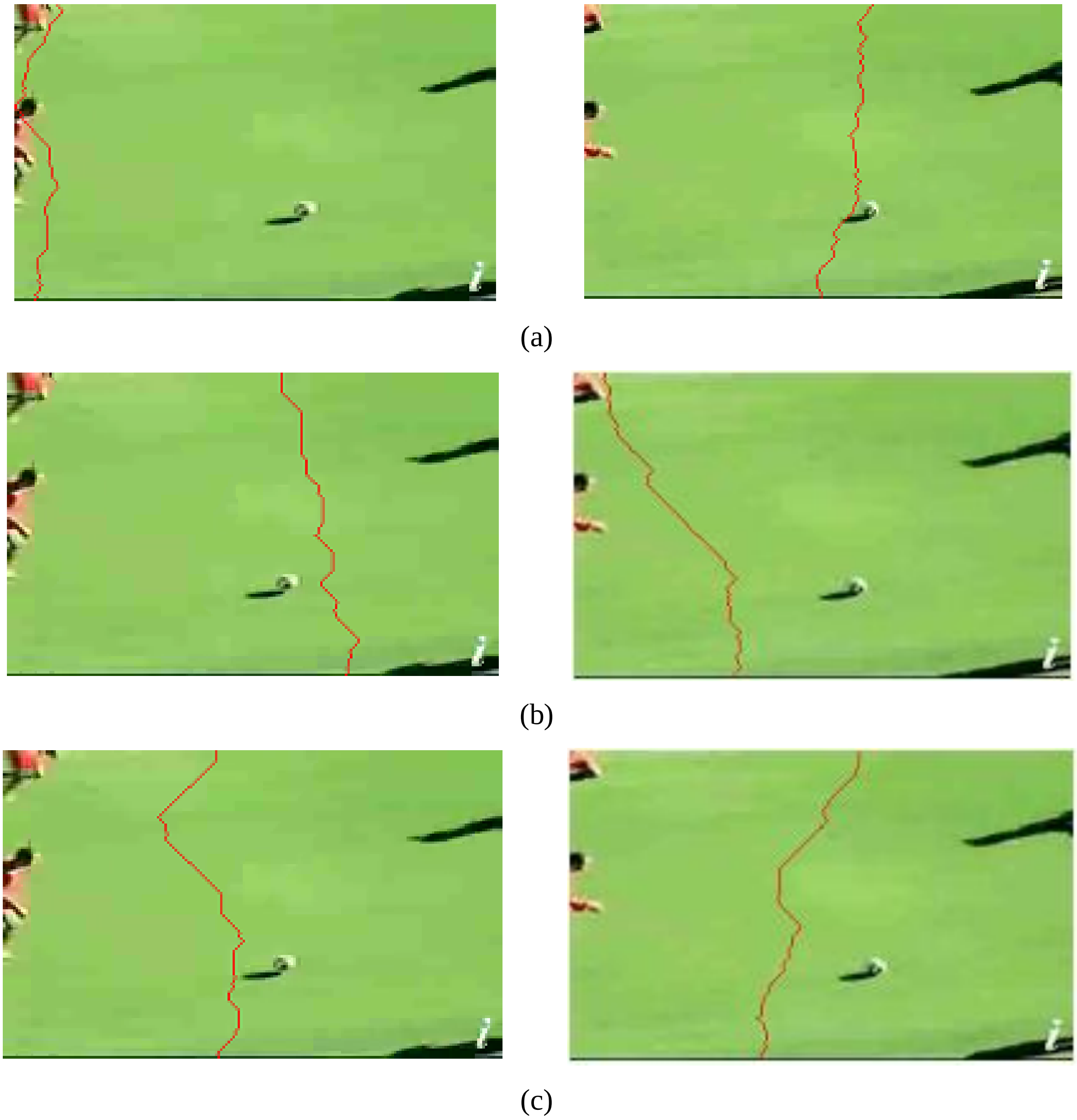
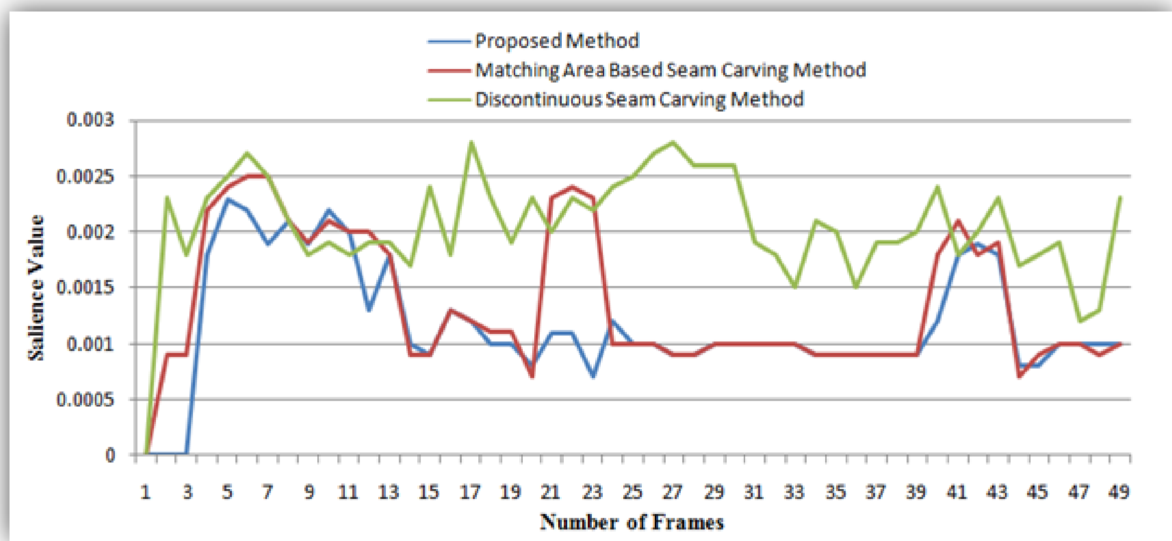


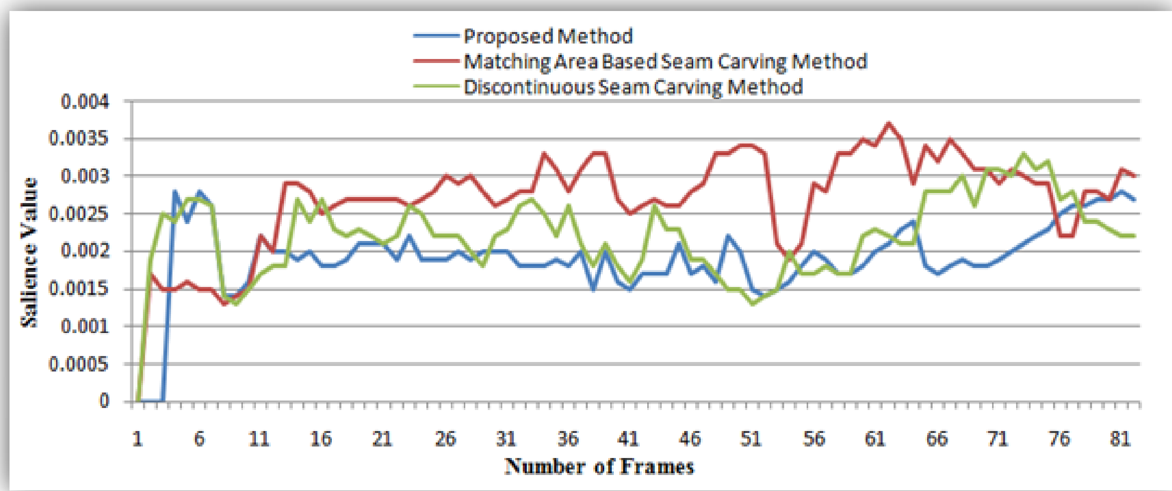
Figure 4.6 (a) Temporal coherency of seams computed by Discontinuous seam carving method (b)Matching area based seam carving method (c) and proposed method in video of football match

4.1.3. Experiment 3: Salience value of seams

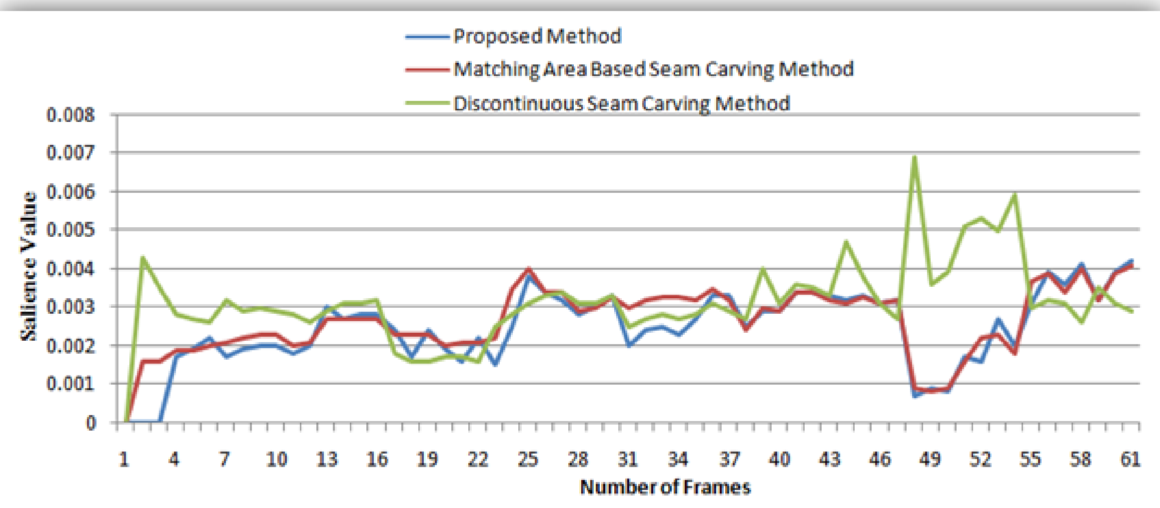
Apart from qualitative demonstrations, we compare our proposed method with other algorithms on the basis of quantitative figures as well. Saliency of a frame is defining the quality by which objects in frame stands out relative to its neighbors. Each pixel in the frame is given some salience value whose value is higher for salient content while lower for other unimportant content in the frame. One of the algorithms to find saliency is proposed by [10, 11] which consider low level features of the frame to statistical structure of frame. We use this method to find the saliency of each frame in video and then average salience value of seams is computed per frame. As the salient content should remain intact while retargeting, so seams computed should have minimum saliency to avoid removal of important areas of frame. Figure 4.7(a) illustrates the graph corresponding to the example given in Figure 4.1 which shows that seam computed by proposed method is having almost lower salience value as compared to other algorithms. Similarly graph in Figure 4.7(b) is against example given in Figure 4.2 and graph in Figure 4.7(c) corresponds to example mentioned in Figure 4.3.



(a)



(b)

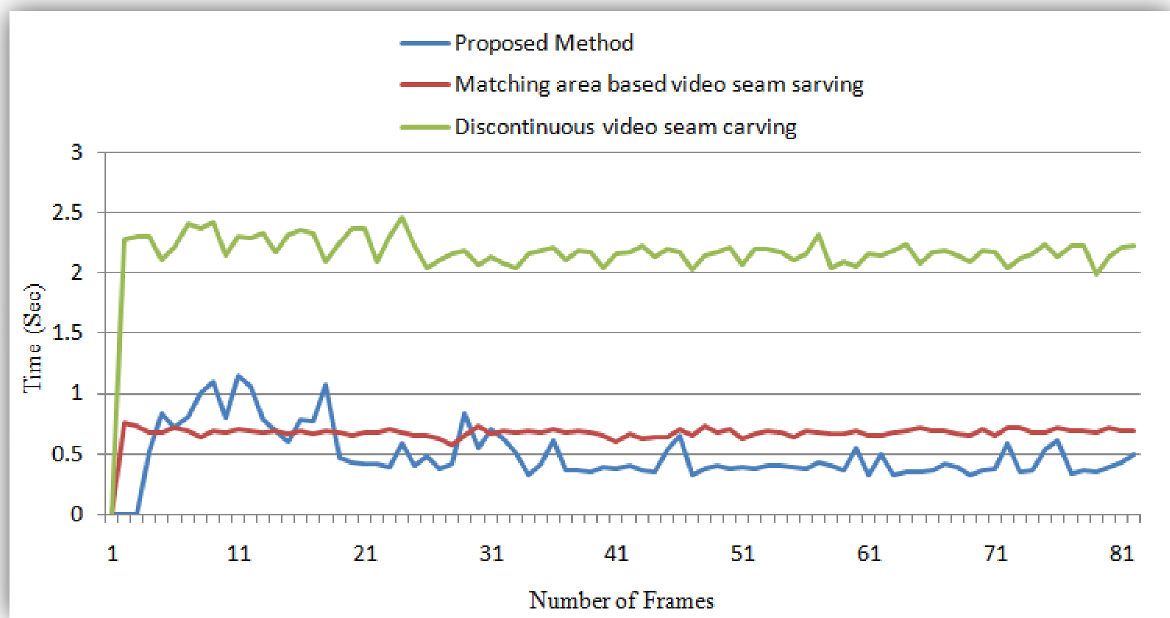


(c)

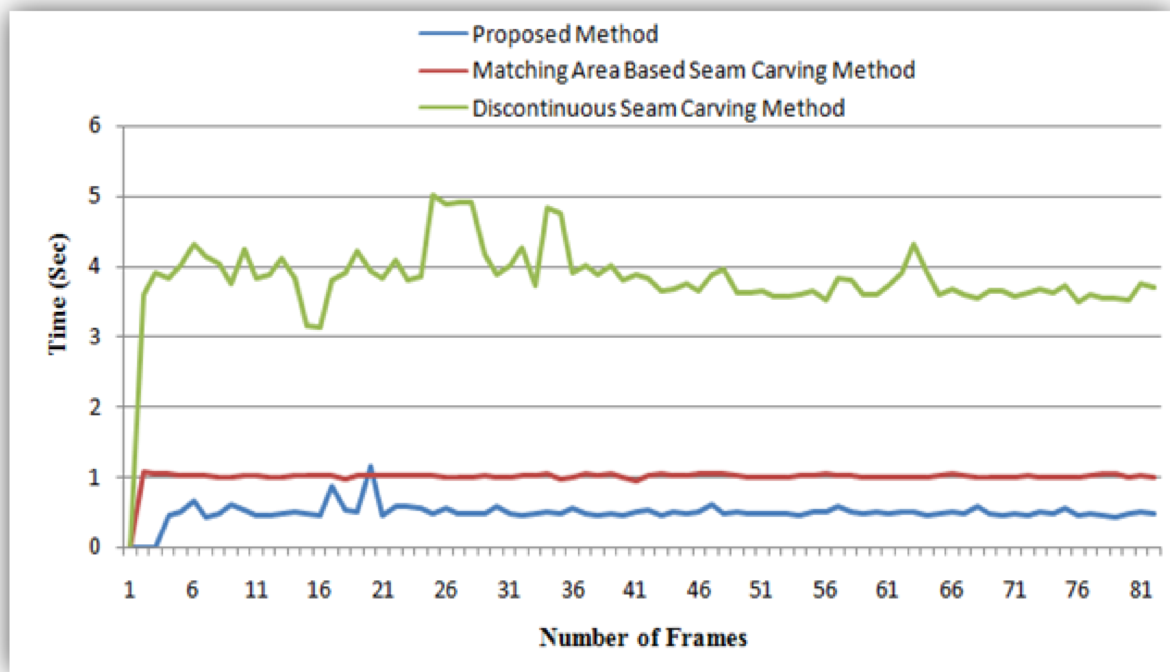
Figure 4.7 Graphs showing salience value per frame for seams produced in different video examples

4.1.4. Experiment 4: Time taken to compute seams

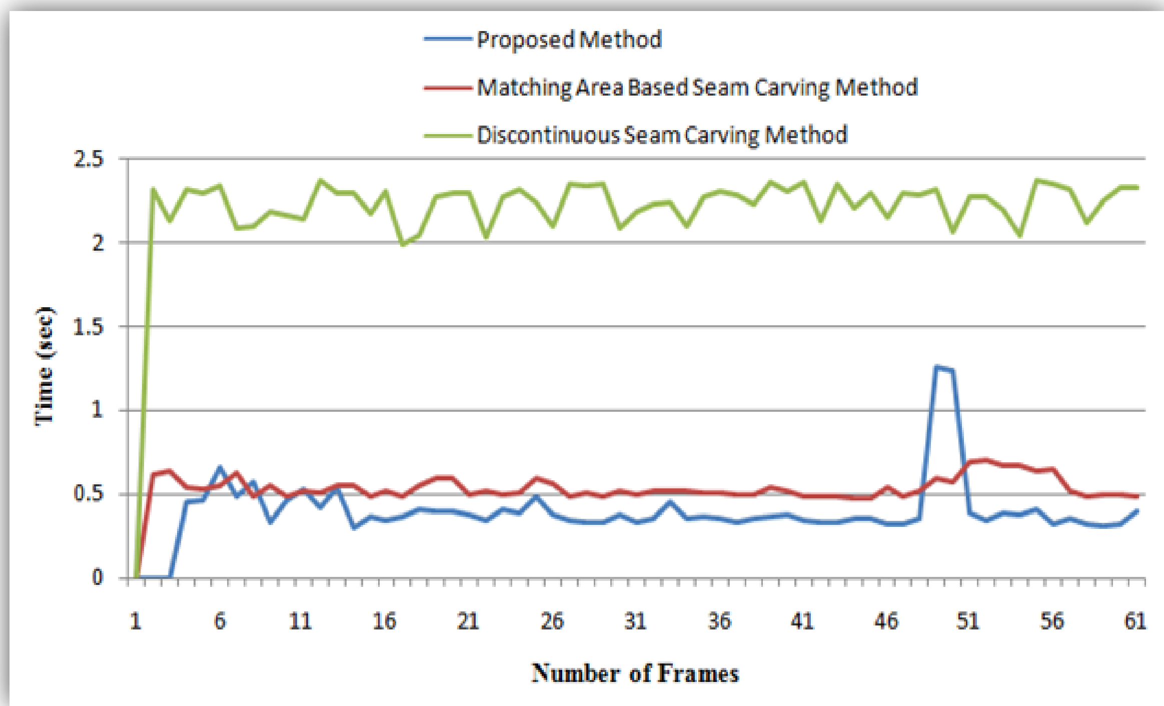
Another experiment done is for time taken by various algorithms to compute the seams. Apart from giving visually satisfactory results, our proposed method is also less complex and is taking less time to compute seams. Our algorithm estimates the seam using Kalman filter which allows us to find the seam in nearby areas of estimated seam only. Hence our algorithm is comparatively faster than other two while discontinuous seam carving method for videos is considering the appearance difference for all pixels per row in frame and the matching area based approach is comparing fixed sized match area in surface area which may take more time. This can be demonstrated by graph given in Figure 4.8(a) corresponding to example shown in Figure 4.6. The graph shows the time taken in seconds to compute seam per frame where our proposed method (in blue color) is taking less time as compared to discontinuous seam carving method in green color and matching area based seam carving method in red color. Another result is shown in Figure 4.8(b) corresponding to the example corresponding to example given in Figure 4.2 and Figure 4.8(c) showing time taken for computing seams in Figure 4.3.


Proposed Method
Matching area based video seam serving
Discontinuous video seam carving
Time (Sec)
3
2.5
2
1.5
1
0.5
0
1
11
21
31
41
51
61
71
81
Number of Frames

(a)


Proposed Method
Matching Area Based Seam Carving Method
Discontinuous Seam Carving Method
Time (Sec)
6
5
4
3
2
1
0
1
6
11
16
21
26
31
36
41
46
51
56
61
66
71
76
81
Number of Frames

(b)

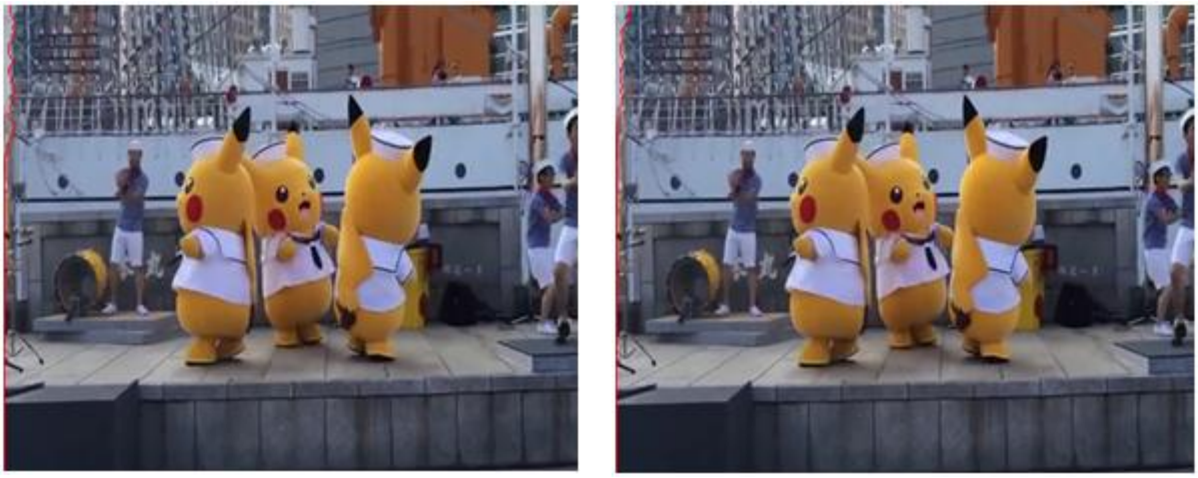


(c)

Figure 4.8 Graphs showing time taken to compute seams per frame produced in different video examples

4.1.5. *Experiment 5: Slow Paced Videos*

Matching area based seam carving method [5] will produce temporally coherent results if the video is having slow pace. It will compare the areas around where the previous frame seam was. Since video is slow so areas will match producing seam at almost similar location maintaining temporal coherency. On the other hand our proposed method will estimate the seam considering the frame rate of video so it will also produce temporally coherent results. So in slow paced videos both the algorithms are working similar as shown in Figure 4.9. Figure 4.9(a) is showing red colored seam created by matching area based algorithm while Figure 4.9(b) is having seam computed by our proposed algorithm. We are taking same energy criteria that is, gradient energy for both algorithms to show comparisons fair and better.



(a) (b)

Figure 4.9 Seams produced by (a) Matching area based seam carving algorithm and (b) proposed method in video of dancing Pikachu

Another example of slow paced video having large number of objects is given in Figure 4.10 where both are having red colored seam at extreme left side of frames.

(a) (b)

Figure 4.10 Seams produced by (a) Matching area based seam carving algorithm and (b) proposed method in video of minions

4.1.6. Experiment 6: Average saliency of seams

We tested our method on multitude of videos. The results shown in previous experiments are on individual videos but to signify our proposed method, we combine the results for all the videos and represent them in one place. We take an average of salience value of seams computed per frame for all videos which is shown in Figure 4.11. As discussed, salience value of seam should be less to maintain salient content in video. It is clearly visible that our proposed method(in blue color) is removing or replicating low salience value pixels as compared to algorithm by [4](in green color) and algorithm by [5](in red color).

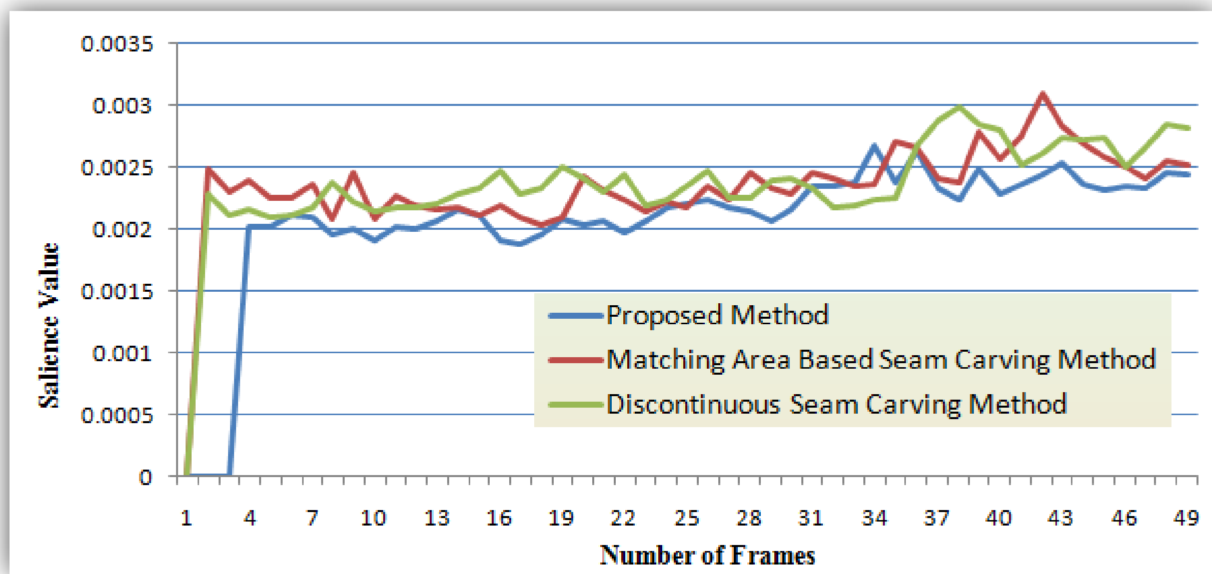


Figure 4.11 Graphs showing average salience value per frame for seams produced in all video examples

4.1.7. Experiment 7: Average time taken to compute seams

Similar to experiment 5, we take average of time taken by state-of-art algorithms for all the videos to illustrate that our proposed method is having low time complexity as compared to [4] and [5]. The graph in Figure 4.12 shows the time taken (in seconds) per frame for all videos by different algorithms. Blue colored curve showing time taken by proposed method is below the green colored curve indicating time taken by [4] and red colored curve indicating time taken by [5].

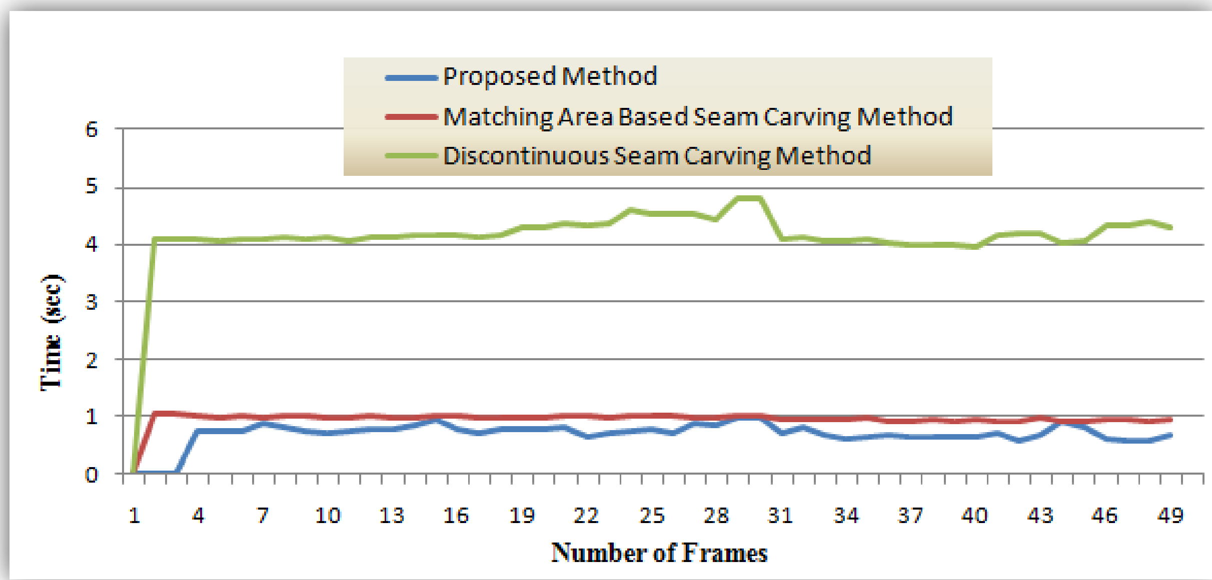


Figure 4.12 Graphs showing average time taken to compute seams per frame for all video examples

4.1.8. Experiment 8: Retargeted video frames

Below are some examples after retargeting the video. Video are retargeted according to target size given by user.

- Figure 4.13(i) showing two babies, out of which, the baby sitting on left is distorted by [4] and [5] while retargeting. Algorithm by [4] distorts the shape of baby while [5] distorts the head of left baby. Our method is showing comparatively good results with respect to original image.
- In Figure 4.13(ii), oval shape of tub in which dog is playing is changed to haphazard curve by [4], though algorithm by [5] does not destroy the shape, but our proposed method removes the pixel from unimportant areas as compared to algorithm by [5] so that salient content remains intact.
- Figure 4.13(iii) representing large number of audience for some event is well preserved by our method while the boy at top left side of frame are distorted by other both algorithms.
- In Figure 4.13(iv) some football match frames are taken and it is clearly visible that two boys at back are getting removed by algorithms [4] and [5], while our algorithm is preserving the actual objects (boys here).

(a)

(b) (c) (d)

(i)

(a)

(b) (c) (d)

(ii)

(a)



(b) (c) (d)

(iii)

(a)



(b) (c) (d)

(iv)

Figure 4.13 Video Retargeting Results (a) Original frames (b) [4]'s results (c) [5]'s results (d) Our Proposed algorithm

4.2. Complexity Comparison

As shown in Table 4.1, Our algorithm is having comparatively less complexity as compared to other two algorithms. Our algorithm is basically working around the number of rows of frame for vertical seam. Seam carving for image retargeting [1] is applied to compute the seam for first frame which consider the energy for all pixels of the frame hence take O(HW) time. Kalman filter runs to initialize variables, do prior estimation of seam, calculate Gaussian errors in prediction phase, update seam state using the pixels that are part of seam in previous frame thus taking time of number of rows per frame. Kalman gain computation requires the seam having low energy and nearest to predicted seam which is estimated in O (HW).

Matching area based seam carving method for video retargeting is also trying to find seam around the previous frame seam to maintain temporally coherency but it takes extra computation time for comparing match areas in surface area with the key point match area. Hence the time of O(K*(M+SM)) per frame is included for finding match area for all key-points and comparing them with all match areas of surface area. Key points are maximum energy pixels on previous frame seam regions found by checking energy values of all pixels of seam taking O(H) computations per frame. After that it takes O(HW) time per frame to update the energy of frame and applying seam carving for image retargeting[1] for finding final seam.

Discontinuous seam carving method for video retargeting is finding spatial, temporal cost of removing all the pixels in addition to finding minimum energy seam in first frame thus taking O(HW+WH) time. It aims to find spatial, temporal and saliency costs. Spatial cost is combination of errors introduced in horizontal and vertical direction by removing pixels from frame which in total takes O(H)+O(WH). Temporal cost is finding sum of squared difference of most temporally coherent frame which is computed in O(HW) time and temporal coherency cost for other pixels in frame taking time $O(WHW+W^2H)$. Saliency cost takes O(WH) as it is finding the salience value for all pixels in frame. All the three cost are added for each pixel in each frame and then minimum seam is found out of them taking O(W) time. So, This method is taking more of time as it is considering all pixels of a frame.

Table 4.1 Complexity comparison for state-of-art algorithms with our proposed method

Seam Carving Methods	Complexity
Proposed Method	N*[O(HW)+4*O(H)+O(HW)]
Matching Area Based Seam Carving	N*[O(HW)+O(H)+O(K(M+SM))+2*O(HW)]
Discontinuous Seam Carving	N*[O(HW)+O(HW+WH)+O(H)+O(WH)+O(HW)+ O(WHW+W^{2}H)+O(WH)+O(W)]
W is frame width, H is frame height, R is number of regions in a frame, S is search area (like 7 x 7), M is matcharea (like 3 x 3), K are number of keypoints and N is number of seams to be removed.	

CHAPTER 5

CONCLUSION AND FUTURE WORK

In this thesis, we presented a seam carving based video retargeting method to resize the video. This method estimates a spatio-temporal seam using Kalman filter and the well-known spatial seam carving approach. The use of Kalman filter ensures temporal coherency and spatial coherency, whereas the spatial seam carving ensures spatial coherency. This temporal seam from Kalman filter and spatial seam from spatial seam carving are combined in such a way that seam has low energy and is temporally coherent. The Kalman filter works only on the pixel location of the seams of pervious frames and hence it is computationally fast.

Through experimental results it has been found that the proposed technique is far better in maintaining temporal coherency reducing motion artifacts than some state-of-the-art video retargeting techniques. It has also found that the proposed approach is faster and does better or as well as existing approaches in seam removal or replication for video retargeting.

We use Kalman filter as an estimation algorithm. This research can be extended to other prediction algorithms that can work better on complex systems as well. The Extended Kalman Filter can be applied to have better prediction of temporal seam.

BIBLIOGRAPHY

[1] Avidan, S. and Shamir, A, “Seam carving for content-aware image resizing,” ACM Trans. on Graphics (Proc. of SIGGRAPH), Vol. 26(3), October 2007.

[2] D. Vaquero , M. Turk , K.Pulli , M. Tico and N. Gelfand, "A survey of image retargeting techniques," in Proc. SPIE, 2010, pp. 779814-1 -779814-15.

[3] Rubinstein, M., Shamir, A., &Avidan, S, “Improved seam carving for video retargeting,” In ACM transactions on graphics (TOG), Vol 27(3), p. 16, August 2008.

[4] M. Grundmann, V. Kwatra, M. Han, and I. Essa, “Discontinuous seamcarving for video retargeting,” in Proc. IEEE Conf. Comput.Vision Pattern Recognit.(CVPR), pp. 569–576, 2010.

[5] B. Yan , K. Sun and L. Liu,“Matching area based seam carving for video retargeting,” IEEE Trans.Circuits Syst. Video Technol, Vol23(2), pp.302 -310, February 2013.

[6] Guo, Y., Liu, F., Shi, J., Zhou, Z.-H., and Gleicher, M, (2009, Aug.)”Image retargeting using mesh parametrization,” IEEE Trans. on Multimedia. Vol 11(5), pp. 856–867, August 2009.

[7] T. Liu, J. Sun, N.-N.Zheng, X. Tang, and H.-Y.Shum., “Learning to detect a salient object”, IEEE CVPR, 2007.

[8] Sen, D., Swamy, M.N.S., Ahmad, M.O,“Computationally fast techniques to reduce AWGN and speckle in videos,” IET Image Process, Vol1(4), pp. 319–334, 2007.

[9] R. Faragher,“Understanding the basis of the Kalman filter via a simple and intuitive derivation,” IEEE Signal Process, Vol29(5), pp.128 -132,September 2012.

[10] Garcia-Diaz, A. Fdez-Vidal, X. R Pardo, X. M and Dosil, R, “Saliency from hierarchical adaptation through decorrelation and variance normalization,” Image and Vision Computing, Vol 30 (1), pp.51-64, January 2012.

[11] Garcia-Diaz, A. Lebor·n, V. Fdez-Vidal, X. R. and Pardo, X. M, “On the relationship between optical variability, visual saliency, and eye fixations: A computational approach,” Journal of Vision, Vol 12 (6), pp. 1-22, June, 2012.

[12] L.Wolf, M. Guttmann, and D. Cohen-Or, “Non-homogeneouscontent-driven video-retargeting,” in IEEE ICCV, 2007, pp. 1-6.

[13] Tao, C., Jia, J., and Sun, H, “Active window oriented dynamic video retargeting,” in Proceedings of the Workshop on Dynamical Vision, ICCV, 2007.

[14]W.-L. Chao, H.-H. Su, S.-Y. Chien, W. Hsu, and J.-J. Ding, “Coarse-to-fine temporal optimization for video retargeting based on seam carving,” in Proc. IEEE Int. Conf. Multimedia Expo (ICME), 2011, pp. 1–6.

[15] Liu, F., and Gleicher, M., “Video retargeting: automating pan and scan,” In MULTIMEDIA Proceedings of the 14th annual ACM international conference on Multimedia, ACM, 2006, p. 241– 250.

[16] Rubinstein, M., Shamir, A., and Avidan S, ”Multioperator media retargeting,” ACM Trans Graph, Vol 28(3), August, 2009.

[17] Zhang, Y. F., Hu, S. M., & Martin, R. R, “Shrinkability Maps for Content-Aware Video Resizing,” In *Computer Graphics Forum*, Vol 27(7), pp. 1797-1804, October, 2008.

[18] Gal, R., Sorkine, O., and Cohen-Or, D., “Feature-aware texturing,” in Proc. of Eurographics Symposium on Rendering, pp. 297–303, june 2006.

[19] Zhang, G.-X., Cheng, M.-M., Hu, S.-M., and Martin, R. R., “A shape preserving approach to image resizing,” Computer Graphics Forum Proc. of Pacific Graphics, Vol 28(7), 2009.

[20] M. Grundmann, V. Kwatra, M. Han, and I. Essa, “Efficient hierarchical graph-based video segmentation,” In IEEE CVPR, pp. 2141-2148, June, 2010.

[21] http://studentdavestutorials.weebly.com/