A Research on Video Inpainting with Occlusion

Panchal Vishakha M
Department of Electronics & Communication Engineering
Parul Institute of Engineering & Technology, Parul University, Vadodara-391760, Gujarat, India

Mewada Hardik N
Department of Electronics & Communication Engineering
Parul Institute of Engineering & Technology, Parul University, Vadodara-391760, Gujarat, India

Shrivastav Punit S
Department of Electronics & Communication Engineering
Parul Institute of Engineering & Technology, Parul University, Vadodara-391760, Gujarat, India

Abstract

The remarkable demand for universal distribution and consumption of image and video content our various network has pressured the digital consumer electronics industry to launch an almost infinite variety of electronic devices capable of acquiring, processing, editing and storing these very attractive types of content. A computer vision technique is proposed for a video sequence to remove an unwanted object for digital processing. Recently various special effects have been employed in video production. We propose a novel and fast technique of video inpainting which will work on both static and free moving camera videos. We can use this method for error concealment, object removal and background reconstruction. The proposed approach will reduce the algorithm complexity and also provide visually pleasing results.

Keywords: Inpainting, GMM, Exemplar Method, Occlusion Detection

I. INTRODUCTION

“The wide applications of digital camera and the digitization of old photos inpainting has become an automatic process which is operated on digital images. More than scratch removing the inpainting techniques are also applied to object removal, text region and other automatic modification of images and video[19].

Fig. 1: Inpainted result (a) Original image (b) Inpainted image [19]

“Image inpainting is a technique of filling unknown image region with known information from the surrounding of the unknown region in such a way that the result is logically accepted [20].”

Video inpainting is becoming more and more important in the field of digital media in recent years. Video inpainting refers to a field of computer vision that aims to remove objects or restore missing or tainted regions present in the video sequence by utilizing spatial and temporal information from neighboring scenes [15]. The overriding objective is to generate an inpainted area that is merged seamlessly into the video so that visual coherence is maintained throughout and no distortion in the affected area is observable to the human eye when the video is played as a sequence [15]. We must ensure that spatial and temporal consistencies both should be maintained”.

(a) Part of the original video sequence
“In figure 2 here they are employing a video inpainting algorithm to repair a damaged video. The damage consists of the large black region in the middle of the video sequence, possibly caused bit loss due to transmission across an unreliable network. Image (a) depicts the original video sequence - notice here how the damaged region completely occludes the bottom half of the person and the associated background. Image (b) shows the first step of the inpainting algorithm which fills in the missing portion of the person. Image (c) completes the process by filling in the missing background, with Image”[15].”

Shown in Fig. 3, the pre-processing step, we construct object detection Method for background and foreground separately. Then, object detection method is used to distinguish background from foreground. In the step of video inpainting, we firstly fill the missing background as much as possible by copying pixel by pixel information from other frames, and then inpaint the remaining holes in the background by video inpainting technique [8].
Difference between image inpainting and video inpainting is that image inpainting techniques only consider spatial information and does not consider the significant temporal component present in video sequence. The reasons are [15].

1) Image inpainting techniques only consider spatial information whereas for video inpainting temporal information.
2) Image inpainting is valid interpolation. Whereas for video inpainting use extrapolation.
3) Image inpainting considers less number of pixels. Whereas for video inpainting considers huge number of pixels.

“Reasons are only consideration of spatial components that leads to quite severe image artifacts resulting in inpainted video. This is due to assumption that edges is interpolated in smoothly way that is not true for dimension of time. Due to particular pixel containing background in one frame, non-smooth temporal changes are generated. It is possible that inpainted video frames taken separately have no visible artifacts. Video played in a sequence do the artifacts visible themselves to the human eye. Thus video inpainting is more challenging then image inpainting [15].”

In this paper we proposed a fast and efficient video completion method based on occlusion detection and by using exemplar based inpainting technique and will try to solve the occlusion problem.

The remaining of the paper is organized as follows. The section I is introduction about the basic video inpainting. An overview of the difference between image and video process, object detection method and video inpainting techniques. The section II shows overview on object detection methods and video inpainting techniques. The section III gives the comparative analysis of these different techniques. The section IV describe about occlusion detection flow and finally results and conclusion is described in section V and VI.

II. RELATED WORK

Video inpainting is a particularly challenging problem. Difficulties arise due to the possibilities of camera motion, the requirement to handle both static and dynamic regions, as well as the pose and scale variation, lighting and shadows and occlusion that are presently common in video [15].

A. Video Inpainting Techniques [5]

"The literature is replete with a varied array of video inpainting proposals each advocating different approaches to the problem, Bertalmio et al. [3] investigated the Partial Differential Equation (PDE) based technique to arrive for border of the region iterative manner is used. It will smoothly inpaint outside of the border to the inner region. It Produce good results if missed regions are small and target region is non-textured, but take long time if target region is large. Some blurring effect is presented in video sequence [16].”

“Patwardhan et al. [1] suggest a rather simpler method for inpainting stationary background and moving foreground in videos. Here to inpaint Texture Synthesis based inpainting technique is used to fill missed regions and remove damaged area using similar neighbourhoods of the missed pixels. Texture Synthesis based inpainting technique is used to Inpaint small region and can-not handle natural scenes effectively [16].”

Exemplar based method by Criminisi et al. [4] searches most plausible patch from source region and copies pixels of the found patch directly to the missing region. Exemplar based inpainting technique is considering structure and texture region [4]. This technique consists of two basic steps.

1) Priority is given to all the patches.
2) Best matching patch is selected.
B. Object Detection Techniques [5]

![Object detection flowchart](image)

“In [5] the presence of moving objects is determined by calculating the difference between two consecutive frames. Its calculation is simple and easy to implement. It has a strong adaptability, for a variety of dynamic environments, but it is generally difficult to obtain complete outline of moving object, responsible to appear the empty circumstance, as a result the detection of moving object is not accurate”[16].”

“In [6] Background subtraction is background modelling. It is the core of background subtraction algorithm. Background Modelling must sensitive enough to recognize moving objects. Background Modelling is to yield reference model. "In background subtraction, each video sequence is compared against the reference model to determine possible variation. The difference between current frames and the reference frame in terms of pixels give importance to existence of moving objects. The background subtraction method is to use the difference method of the current image and background image to detect moving objects, target recognitions but very sensitive to the changes in the external environment”[18].”

“In [8] Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are parametric model of the probability distribution of continuous measurements or features in a biometric system. It includes color based tracking of an object in video. It is critical to identify moving objects from a sequence of videos frames [17]. The Gaussian Mixture Model for background subtraction method in that frame pixels are deleted from the required video. GMM is reactive to the various changes such as illumination, starting and stopping of moving objects. Surveillance is the monitoring of the behavior, activities or other changing information usually of people and often in a surreptitious manner [17].”

III. COMPARATIVE ANALYSIS

A comparative analysis of object detection methods and video inpainting techniques are shown in table I and table II respectively.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing time</td>
<td>Low to High</td>
<td>Medium</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Medium</td>
<td>Low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing time</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Handle hole size</td>
<td>Small region</td>
<td>Small region</td>
<td>Large region</td>
</tr>
<tr>
<td>Region</td>
<td>Only used structure</td>
<td>Used both texture and structure</td>
<td>Used both texture and structure</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
</tr>
</tbody>
</table>

IV. PROPOSED WORK

Our proposed method is based on frame local similarity by using exemplar based inpainting technique and by 3D hole filling to complete the damaged background after the removing object. The object can be static or dynamic or we can consider some logos or stamps.
In occlusion detection flow, we have to follow the step accordingly first the video is to be converted in number of frames then by the object detection techniques like frame differencing, background subtraction to separate the foreground and background object and here we are using the background subtraction technique i.e. GMM (Gaussian Mixture Model). Second, find the location of the two objects and then the parameter Euclidean distance is to be used to calculate the distance of the object and if the value is accurate and it is around +15 then the occlusion will be observed.

V. RESULTS

![Fig. 7: Result of background subtraction](image)

![Fig. 8: Result of Gaussian mixture model (GMM)](image)
C. Inpaint Results: Stable Object

Fig. 9: Inpainting result of image using exemplar technique

Fig. 10: Frame No-19 Stable object inpainted

Fig. 11: Frame No-50 Stable object inpainted
D. Occlusion Detection Results:

Fig. 12: At Frame No-91: No Occlusion

Fig. 13: At Frame No-112: Occlusion Start

E. Inpaint Results: Unstable Object

Fig. 14: Frame No-19: Unstable Object Inpainted
VI. ANALYSIS

Table 3: Calculation of Object Detection

<table>
<thead>
<tr>
<th>Method</th>
<th>TP</th>
<th>TN</th>
<th>FP</th>
<th>FN</th>
<th>TG</th>
<th>TF</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approximate Median</td>
<td>90</td>
<td>55</td>
<td>0</td>
<td>55</td>
<td>200</td>
<td>200</td>
<td>72.5%</td>
</tr>
<tr>
<td>Adaptive Gaussian Mixer Model (GMM)</td>
<td>115</td>
<td>50</td>
<td>0</td>
<td>35</td>
<td>200</td>
<td>200</td>
<td>82.5%</td>
</tr>
</tbody>
</table>

Accuracy = TP+TN/TF\(^{[21]}\)

![Graph showing comparison of object detection methods based on parameters](image)

Fig. 16: Comparison of object detection methods based on parameters

![Graph showing comparison of object detection based on processing time](image)

Fig. 17: Comparison object detection based on processing time
VII. CONCLUSION & FUTURE SCOPE

“In this paper, there are three major contributions improvements of exemplar based inpainting, the separation of background and foreground using GMM and occlusion detection. The above mentioned occlusion problem discussed in the paper has been researched extensively by many researchers. The work that has been proposed addresses to the above mentioned problem partially and would be strive to mitigate the occlusion problem as much as possible.”

In future this approach can be improved on video with complex non-static background which has high resolution which is taken by free moving camera.

REFERENCES

[14] Anil K. Jain, Fundamentals of image processing, University of California, Davis
[16] Mr. Mahesh C. Pawaskar, Mr. N. S. Narkhede and Mr. Saurabh S. Athalye “Detection of moving object based on background subtraction”, International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)