Exemplar-based Video Inpainting for Occluded Objects

Vaishali U. Gaikwad

PG Student

CSE Department,

GECA, Aurangabad, India

P. V. Kulkarni

Assistant Professor

CSE Department

GECA, Aurangabad, India

ABSTRACT

The process of repairing damaged area or reconstructing specific area or removing unwanted objects from video is known as video inpainting. Most of the automatic techniques are available to deal with this problem but most of them are unable to repair large holes. To get sharp inpainted area, in this paper we have proposed an efficient algorithm using exemplar-based technique. Here, considered static camera which gives video having stationery background with moving foreground. To detect the region of moving objects we apply edge detection technique. Once object region detected priority assignment to the patches is applied. A natural image has structures and textures. Structure sparsity was measure to find similarities of the patches. The patch having higher sparseness is then selected and its priority is set which is the highest priority among the patches. This patch is then used for further inpainting.

Keywords
Inpainting, texture synthesis, patch.

1. INTRODUCTION
Inpainting refers to the art of restoring lost parts of images or videos and reconstructing them based on the background information. Basically it is the field of computer vision which aims at restoring damaged parts of video or removing unwanted objects in a video sequence. Numerous image inpainting techniques are available which handles structure of image using only spatial information of image. Nowadays video becomes important medium for communication in the real world. Video inpainting uses spatial information as well as temporal information from image frames of video sequence.

Video inpainting has wide range of applications from the restoration of damaged videos and paintings to the removal or replacement of the selected objects or the objects which occludes important background. Video inpainting is used to remove unwanted or occluded objects or restore missing regions in a video sequence by utilizing spatial and temporal information. The main objective is to generate an inpainted area that is merged seamlessly into the video. This process is carried out so efficiently that visual coherence is maintained throughout and no distortion in the selected area is observable to the human eye when the video is played.

Video consist of sequence of framed images. Normally twenty-five image frames per second are considered as a video. Taking fewer frames than twenty-five does not considered as a video. It is important to distinguish between video and image inpainting from the field of "texture synthesis". The main difference is that in size and characteristics of the region to be considered for inpainting. In texture synthesis the region can be larger with the focus being on the filling in of two-dimensional repeating patterns that have some associated stochasticity i. e. texture. However, inpainting algorithms focuses on the much smaller region which are characterized by linear "structures" such as lines and object contours. Criminisi et al [1] presented a single algorithm to handle both of these cases by giving both the presence of textures ant structures in images.

Over the past few years the research communities work in the field of video inpainting due to various and important applications such as undesired object removal, visual story modification and video restoration. These applications were frequently handled by professionals in a manual fashion, which is very painstaking, slow and also expensive. So automation in this field automatically attracts the commercial organizations (such as broadcasters and film studios) and private individual who wish to edit and improve or to maintain quality of their video collection.

Bertalmio et al [2], [3] designed an algorithm which extends ideas from fluid dynamics. The idea is to propagate isophote lines continuosly from exterior into the region to be inpainted. Author uses only spatial information of images by not considering temporal one. Specifically author designed frame-by-frame PDE based video inpainting algorithm which is mainly focus on edge-continuing methods.

Wexler et al [4] proposed a method by considering both spatial and temporal information. Here authors performed the inpainting by sampling a set of spatial-temporal patches (a set of pixels at frame t) from other frames to fill in the missing data. Here author consider video inpainting as a global optimization problem, which leads to slow running time due to high computational load.

Patwardhan et al [5], presents a basic technique for video inpainting by considering stationary background. A simple spatio-temporal priority scheme is employed by copying patches from undamaged frames to the frames temporally close to the damaged frame, followed by a spatial filling in the step which replaces the damaged region with a best matching patch.

Soon-Yong Park et al [6], designed video inpainting for moving object removal. Also background features are inpainted frame by frame to register background motion. After completing registration process, extraction of moving object as well as its removal takes place in every frame of video.

Video inpainting as a problem of repairing damaged videos is analyzed in et al [7] which involves different techniques by making overall process very complicated.

Zhang et al [8], proposed a motion layer based object removal in video by segmenting video sequences to several layers. Then each separate layer is completed by motion compensation and image completion. All the layers excluding the layer with objects to be removed are then combined to restore the final video.

In proposed method, algorithm is developed for static camera with moving foreground objects. First the input video captured by static camera is extracted into number of frames. Second from that each frame the moving object region is deleted by using Prewitt edge detection technique. Next boundaries of moving object is marked in each image frames. Next, the mask region is inpainted using exemplar based image inpainting using patch priority scheme. Last merging of image frames are displayed as video. The rest of this paper is organized as follows. The section 2 discusses the overview of exemplar based inpainting. Section 3 describes the proposed video inpainting algorithm. Finally, section 4 concludes the paper.

2. EXEMPLAR-BASED INPAINTING

The exemplar based inpainting is an important type of inpainting algorithms. Also it proved its strength by giving very effective output. There are two steps: in the first one priority assigned to the patches and in the second the selection of the best matching patch is done. Exemplar based technique finds the best matching patches from the known region by using its similarity within the patches and pastes into the target region or in missing region. It iteratively synthesis the target region by filling it using most similar patch of source region. This process is called as the patch propagation. Patch propagation is done automatically by the algorithm by inwardly propagating the patches from source region to target region patch-by-patch.
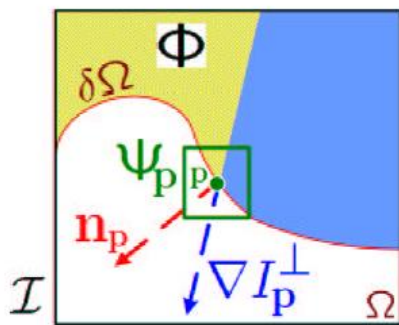

Fig 1: Ψp denotes the patch to be filled. np is the normal to the region contour δΩ of the target region Ω and ∇I_p is the isophote at point p. I denotes the entire image.

Let us first describe the essential terms used in exemplar based inpainting literature.

a) The image to be inpainted is represented as I.
b) The target region i.e. the region to be inpainted is represented as Ω
c) The source region i.e. the region of the image from where the information can be extracted to inpaint the target region is represented as Φ

$$\Phi = I - \Omega$$

d) The boundary of the target region i.e. the pixels that separate the target region from the source region is represented as δΩ

The algorithm first calculates the patch priorities. For calculating patch priority, two properties has to be consider: first one is that the patches which are on the continuations of strong edges are proffered. Second that the patches that are surrounded by high confidence (in terms of their value) pixels are also preferred.

We initialize the confidence value for all the pixels in the source region to be 1 and to the target region as 0. Referring to the above figure, given the patch Ψp, the priority of patch at pixel 'p' is calculated by:

$$P(p) = C(p)\, D(p) \quad\ldots\ldots\ldots\ldots\ldots\ldots..(1)$$

Where $C(p)$ is the confidence term and $D(p)$ is data term. Specifically, the confidence term represents the reliability of the information surrounding the pixel 'p'. Whereas the data term gives priority to the filling in of linear structures first, and this will be done by measuring the strength of the isophotes (contours of equal luminance) arriving the contour δΩ.

After assigning priorities to the patches, the algorithm will proceed to propagate patches from source region having highest priority to the target region. The patch in the source region having the most similarity to the patch in the target region will be selected for copy. Once the copying has been performed, the pixels which are at the intersection of the target region and the region enclosed by the patch are taken to recalculate their confidence value. Now confidence values for these pixels are set to the highest priority. The algorithm then iteratively proceeds like this till target region inpainted completely.

3. VIDEO INPAINTING ALGORITHM

Video can be considered as the display of image frames in sequence. Generally all natural images are composed of the primal sketches like edges; corners etc. which are referred as structures and the regions with feature statistics or homogeneous patterns including the flat patterns are referred to as the texture. Therefore for developing effective algorithm required to consider structure and patterns of natural image. The priority of patch must be assigned in such a way that it should be able to differentiate these two properties of an image. The sparsity of structure in an image is used for assigning priority to the patches to be inpainted. It can be also used for measuring confidence value of a patch which is located at structure. The structures are sparsely distributed at overall image, so pure texture synthesis algorithm are not able to handle the inpainting of missing region.

As shown in flowchart, first capture the video by static camera. This gives the stationary background and moving foreground video. The input video is converted into image frames. For image frame the structure and texture values are calculated. Then the region of moving object which is to be removed is obtained by applying Prewitt edge detection technique.

Next mask the detected object region. Then computing the similarities of patch with its neighboring patches in source region is done. Then the patch priority is calculated by multiplying the structure sparsity with confidence of patch.

Then the patch which is having highest priority is considered for filling in target region. The algorithm iterates this process for filling all image frames until target region is filled completely. Finally displaying all the individual inpainted image frames in sequence gives the inpainted video as output.
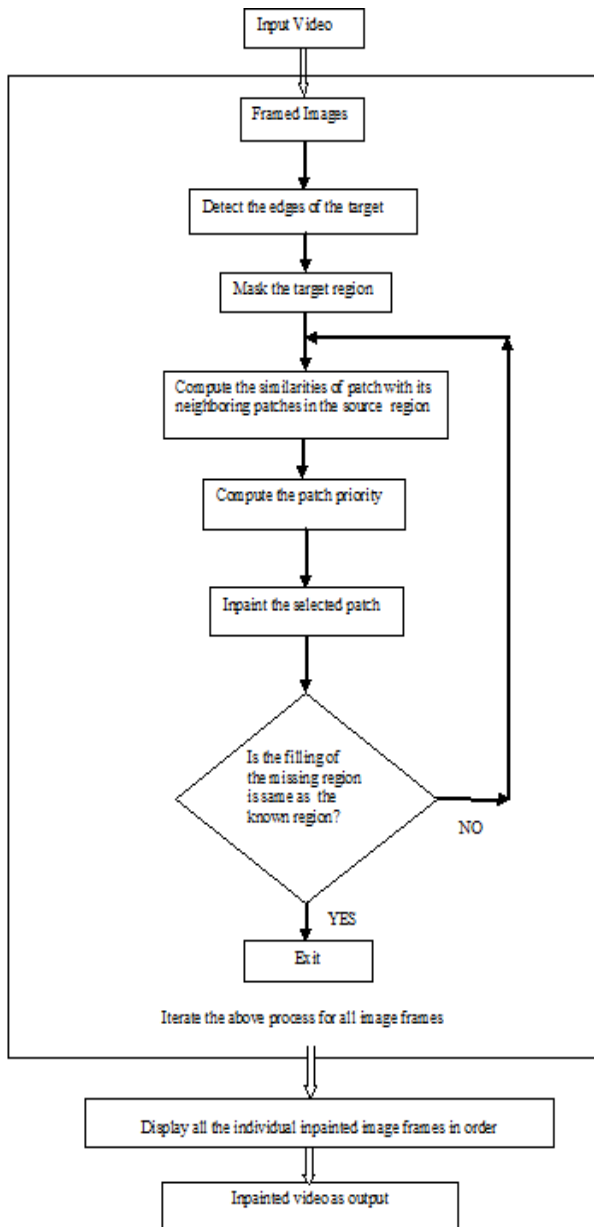
4. CONCLUSIONS

This paper proposed an exemplar based video inpainting algorithm for occluded object. Here considered static camera which gives video having stationery background and moving foreground. The object will be detected by edge detection method which gives sharp boundaries. The proposed algorithm uses exemplar base inpainting which has already proven its best result.

5. REFERENCES

[1] Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by examplar-based image inpainting," IEEE Trans. Image Process.,vol. 13, pp. 1200–1212, 2004.

[2] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in Proc. SIGGRAPH, 2000, pp. 417–424.

[3] M. Bertalmio, A. L. Bertozzi, and G. Sapiro, "Navier–Strokes, fluid dynamics, and image and video inpainting," in Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, pp. 417–424, 2001.

[4] Y. Wexler, E. Shechtman, and M. Irani, "Space-time video completion," Proceedings. 2004 IEEE Computer Society Conference on Computer Vision and Pattern *Recognition*, vol. 1, 2004

[5] K.A. Patwardhan, G. Sapiro, and M. Bertalmio. "Video inpainting of occluding and occluded objects", *in Proc. ICIP 2005.*, Vol. II, pp. 69-72.

[6] Soon-Yong-Park, Chang-Joon-Park, and Inho Lee, "Moving Object Removal and Background Completion in a Video Sequence"

[7] J. Jia, T. Wu, Y. Tai, and C. Tang. "Video repairing: Inference of foreground and background under severe occlusion", *Proceedings. 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. I, pp. 364-371, 2004.

[8] Y. Zhang, J. Xiao, and M. Shah, "Motion layer based object removal in videos," 2005 Workshop on Applications of Computer Vision, 2005.

Fig 2: Proposed Algorithm