A HIERARCHICAL SUPER RESOLUTION BASED VIDEO INPAINTING TOOL

Zakir Yusuf Khan[1], Prajot Randive[2], Krishna Kale[3], Anish Kumar[4]

[1] Bachelors of Engineering, Department of computer engineering, VIIT, Pune, India
[2] Bachelors of Engineering, Department of computer engineering, VIIT, Pune, India
[3] Bachelors of Engineering, Department of computer engineering, VIIT, Pune, India
[4] Bachelors of Engineering, Department of computer engineering, VIIT, Pune, India

ABSTRACT

This paper provides a framework for creating a video inpainting tool which is capable of removing unwanted objects in a video sequence. This tool is a unique tool to fill in the missing parts in the video. It is also capable of restoring a damaged image frame in a video sequence. Many video inpainting techniques have been proposed by researchers. This paper represents a video inpainting technique which is derived from an extended approach of image inpainting. This paper serves a technique of creating a complete packaged tool capable of performing video inpainting. This tool is basically a desktop application capable of running on different operating platforms. The input will be a video file. The given video sequence will be converted into series of frames based on number of frames per second. Now, the frames which are to be inpainted are selected. In this technique the video is inpainted on frame by frame basis. As a result of which this tool becomes capable of not only providing a flexible framework for video inpainting but also it makes removal of unwanted objects simultaneously in different time intervals. The video is then converted into frames and then inpainting of the image frame is performed accordingly. The frames in which inpainting is to performed are first converted into coarse version, then after that using exemplar based inpainting technique the coarse version of the image will be inpainted and then a hierarchical super resolution algorithm is used to restore the missing areas . It is always easier to inpaint the low resolutioned image as we have to deal with less number of pixels as a result of which we will gain advantage in terms of computational complexity. After inpainting the low resolutioned image by exemplar inpainting method we will then use a super resolution algorithm to enhance the resolution of the coarsed inpainted version of the frame so that we willgain advantage in terms of visual quality as well. The proposed technique improves the existing video inpainting techniques by proposing a framework which involves exemplar based inpainting followed by a super resolution algorithm to reduce the computational complexity and to enhance the visual quality of the inpainted video sequence..

Keywords: *- Inpainting , Super resolution, Exemplar inpainting.*

1. INTRODUCTION

Inpainting is a process of removing unwanted objects while preserving the background as it is [1]. The process of inpainting is basically done on images as well as on videos. In the field of computer vision, video inpainting is one of the toughest problem to deal with. Video inpainting is basically used in film industry for video editing, video processing surveillance, video repairing and removing subtitle or unwanted logos in the video. Many researchers have neglected the high temporal correlation that exists amongst various video frames and considered video inpainting as an extension of image inpainting as a result of which many technical challenges has arrived. Such technical challenges must be resolved. Many researchers have proposed many techniques of video in painting so as to resolve such technical challenges but still many techniques resulted in either high algorithmic complexity or poor vision quality or both[2]. The proposed technique in this paper is an attempt to resolve such technical challenges by serving a framework to perform video inpainting without compromising with computational complexity and vision quality. The proposed technique is capable enough to consider both the spatial information as well as the temporal information. As far as image processing is concerned, there are two types of inpainting: diffusion based and exemplar based inpainting. Diffusion based inpainting is based on partial differential equations and variation methods which concerns diffusion-based approaches that propagates linear structures or level lines through diffusion. Sometimes diffusion based approach tend to introduce some blur while filling the void if the void is very large. Exemplar based inpainting performs sampling and copies best matched texture patches from the kno wn image neighborhood. This method uses texture synthesis technique which is very efficient in case of regular or repeatable textures. This paper serves a technique of creating a complete packaged tool capable of performing video inpainting. This tool is basically a desktop application capable of running on different operating platforms. The input will be a video file. The given video sequence will be converted into series of frames based on number of frames per second. Now, the frames which are to be inpainted are selected. In this technique the video is inpainted on frame by frame basis. As a result of which this tool becomes capable of not only providing a flexible framework for video inpainting but also it makes removal of unwanted objects simultaneously in different time intervals. While inpainting each individual frame, first and foremost the selected frame is initially converted into a coarse version. After that an exemplar based inpainting technique is applied on that coarse version of frame. In order to improve the quality of the lower resolution frame we will apply super resolution algorithm which will enhance the resolution of the inpainted image. The above steps will be performed on each and every frame which is to be inpainted. These inpainted images along with those frames which were unselected will be arranged in proper order so as to convert the frames into a video sequence. In short, the proposed technique improves the existing video inpainting techniques by proposing a framework which involves a combination of multiple inpainting versions followed by a super resolution algorithm to enhance the quality of inpainted frames.

Fig -1: Video inpainting

2. EXISTING TECHNIQUES

2.1 Background registration technique[3]

Background registration technique is a regeneration technique for filling the missing parts. In this technique first and foremost the given input video is converted into series of frames based on the number of frames per second. From these set of frames, the first frame is selected and is registered. After that, the process of object removal takes place. In order to start object removal, the edges of the desired object is detected by comparing the edges of the object present in the current frame with that of the registered frame so as to get the boundary of the object to be removed.

This comparison is performed for each and every frame with the registered frame. Now, a masked frame is generated for each time frame. After that the image inpainting process is performed on each and every frame. After image inpainting, the inpainted images are then converted in a video sequence.

2.2 Background subtraction technique[4]

Background registration technique is basically used for detecting the object to be removed and also it is useful for tracking the coordinates of the object. This technique solves the challenging problem of taking temporal and spatial consistency into consideration while dealing with numerous frames. In this technique first and foremost the given input video is converted into a three dimensional artifact. Now, this three dimensional volume is then converted into two dimensional slices in order to maintain the temporal consistency and spatial consistency. Background subtraction technique is now used for object tracking and object detection. Then advanced exemplar based image inpainting technique is applied for removing the object and to fill the missing parts. Now, the retrieval of the most similar postures is performed from the combination of multiple completed slices by forming the sequences of multiple contours. Posture sequence retrieval is a very complex procedure. So, in order to decrease its complexity key posture selection and indexing is used. Synthetic posture generation method is used to increase the number of postures.

2.3 Gradient vector matching technique[5]

The gradient vector matching technique is inspired from fluid dynamics. Various principles of fluid dynamics are used to perform video inpainting which is very innovative method. This is one of the automated method of video inpainting. Principles of classical fluid dynamics are used to propagate isophote lines continuously from the exterior region to the region to be inpainted. Here the image intensity is considered as a streamlined function of two dimensional incompressible flow. The stream function defines a vector field which is used to transport the pixel intensity in the region to be inpainted. The laplacian of image intensity acts as the vorticity of the fluid. This technique is directly based on Navier Stokes equation of fluid dynamics. The technique is designed to continue isophotes while matching gradient vectors at the outline of the region which is to be inpainted.

2.4 Homographic real time video inpainting technique[6]

Homographic real time video inpainting technique provides a platform to achieve a high coherence for translational and rotational camera movements as a result of which inpainting can be performed in real time. Most of the video inpainting techniques have a common restriction of inpainting only those videos which have been captured with stationary cameras. While other inpainting techniques works only when the background is stationary. Such restrictions acts as a hinderance towards the flexibility of the tool as a result of which high quality coherent streams of video are not produces. Therefore such video inpainting algorithms are very computationally expensive without even having significant interactive frame rates. This technique serves a high quality real time video inpainting approach. This approach provides basis for real diminished reality applications to enhance the manipulation of live video streams. This technique is capable of generating coherent video streams even with the presence of heterogeneous backgrounds and non trivial camera movements. In this technique a real time object detection and tracking algorithm is also included. This technique is based only on the previous and current image frame only. So, in order to match the next frame the homographic technique is applied. The overall results showed only a few artifacts as compared to the state of the art methods.

3. PROPOSED TECHNIQUE

The technique which is proposed in this paper is a novel framework for providing the prototype of a video inpainting tool which is capable of gaining advantage not only in terms of computational complexity but in terms of vision quality as well. First and foremost, the video which is to be inpainted is taken as and input and is converted into sequence of frames according to the frame rate. Since, the conversion into frames is based on the physical quantity known as frame rate, therefore we can easily estimate the time from which a particular frame is coming into picture. So, it becomes very easy to inpaint non continuos frames of different time intervals as well hence making this tool very flexible. The inpainting will be performed on frame by frame basis. Now after the selection of the frames which are to be inpainted, each and every selected frame will be converted into a low resolution coarsed version of that frame. Such conversion into low resolution is a necessary step because inpainting of a low resolution image is less time consuming as we have to deal with less number of pixels. A low resolution image is mainly represented by its dominant and important structures. Thus it would be very helpful in object recognition. Since, a low resolution image will contain less noise in it so inpainting a low resolution image is much more easier. Therefore the filling order computation will be reduced because of the reduction in local orientation singularities. The coarse version of the selected frames will be displayed to the user and the user will then choose the object to be inpainted. The tool will store the coordinates on the based of mouse click, mouse dragging and mouse click release. Therefore, the object to be removed will be selected based on the highest number of pixels having same intensity present in that selected region. Now an exemplar based inpainting technique will be applied on this coarsed version of the image. Exemplar based inpainting performs sampling and copies best matched texture patches from the known image neighborhood. This method uses texture synthesis technique which is very efficient in case of regular or repeatable textures. While performing exemplar based inpainting, similar patches must be found out. So, in order to find out the similar patches introduction of an a priori rough estimate of the inpainted values using a multi-scale approach is performed. It then results in an iterative approximation of the missing regions from coarse-to-fine levels. The exemplar based inpainting on the coarse version is performed on the basis of two parameters: filling order and patch size. In order to make it more robust, we will introduce a lot of different settings ie; variations in filling order and patch size. Now after identifying the patch sizes, the filling order is determined on the basis of patch priority. In exemplar based inpainting we need to distinguish structure from the textures. So, for distinguishing the structures a priority to the individual patches is assigned on the basis of filling order computation. A high patch priority indicates that the structure is present. Now in order to fill the patches, the patch which is having highest priority will be filled first and likewise the other patches will be filled according to their priorities. To fill in the unknown part of the current patch the most similar patch located in a local neighbourhood centered on the current patch is sought. A similarity metric is used for this purpose. The chosen patch maximizes the similarity between the known pixel values of the current patch to be filled in and co-located pixel values of patches. This is how the inpainting of the coarse image is performed. Now, in order to improve the resolution of the inpainted image we will use super resolution based algorithm. The super resolution algorithm is of two types: single image super resolution and multiple image super resolution. Super resolution process refers to creating a high resolution image from a single or a set of multiple low resolution images. In order to perform super resolution algorithm there must be some prior information. This prior information is known as training set. This training set can be an energy function defined on a class of images. It can also be used as a regularization term with suitable interpolation techniques. After acquiring the multiple inpainted versions of a single frame amongst the selected frame, these multiple low resolution frames are to be converted into high resolution frame by using super resolution algorithm. So, after getting a good quality inpainted image we convert these set of inapinted frames along with other frames into a video sequence. Thus, the inpainted video is produced.

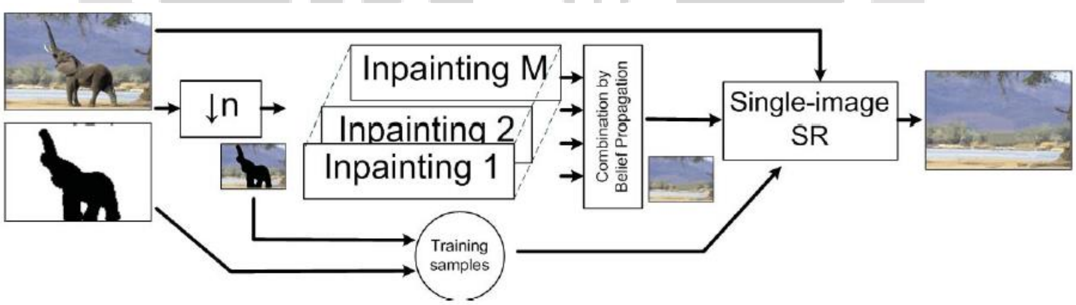

Fig -2: Proposed technique

4. CONCLUSIONS

The experimental results depicted the effectiveness, usefulness and robustness of this proposed technique. This technique was performed on various kinds of videos such as videos consisting of non trivial camera movements and videos consisting of stationary background. According to the experimental results, the implementation of super resolution algorithm took most of the processing time because the template matching was sequential and not parallel. One interesting avenue of the future work would be either to use parallelization techniques while template matching. In fact the template matching technique could be replaced by approximate nearest neighbor search. Other super resolution algorithms can also be used in order to improve the robustness.

5. REFERENCES

[1]. Olivier le meur, Mounira ebdelli, Christine gulliemot "A hierarchical super resolution based inpainting" IEEE Transactions on image processing vol 22, no 10, Oct 2013.
[2]. Nita D Patel, Rashmi prajapati "A survey on video inpainting" IJSRD, Vol 2, Issue 11, 2015, ISSN(online) 2321-0613
[3]. Sanket Sonawane,Sahil Sarangdhar,Siddhi Bhatt,Divya Sharma "Video Inpainting using background registration", International Journal of Research in Engineering and Technology Volume 03 Issue 05 May 2014.
[4]. Sun Xiaoming, Yu Xiaoyang, Yu Shuchun, Guan Yanxia, Meng Xiaoliang, Yu Yang, Liu Yanan, " Video Inpainting Model for Camera Motion Based on Improved Background Subtraction Method", Measurement Information and Control,2013 International Conference on (Volume 01).
[5]. M. Bertalmio **A.** L. Bertozzi **G.** Sapiro "Navier-Stokes,Fluid Dynamics, and Image and Video Inpainting" ,Preceedings of IEEE Conference on Computer Vision and Pattern Recognition,vol.1,pp.355-362,2001.
[6]. Jan Herling, Member, Wolfgang Broll "High- Quality Real-Time Video Inpainting with PixMix" IEEE transactions on visualization and Computer graphics, vol. 20, no. 6, june 2014