An Intelligent Video Repairing Approach Using Object Inpainting: A Review

Mr. Vinod R. Thakare ¹, Dr. Nitiket N. Mhala ²
¹PG Student M. Tech. Electronics, B.D.C.O.E. Wardha
²Professor Electronics Department, B.D.C.O.E. Wardha

Abstract — Video inpainting deals with the ability to fix/restore damaged videos and provide flexibility to edit home videos. It also ensures visual privacy in security. Video inpainting is used in large number of applications to inpaint missing objects from a video sequence, remove black spots or holes from video sequence, as a signature remover, a ghost shadow remover. For video repairing the basic approaches has been developed in recent years. Lots of research is done in image inpainting as compared to video inpainting. The video inpainting is broadly categorized as Patch based method & Object based method. There are some critical issues in video completion regarding to keep the spatial-temporal consistencies, and the faithful inference of pixels. Previous work is done in the area of video inpainting and most of the techniques try to ensure to maintain either spatial consistency or temporal consistency between the frames. But no one try to ensure both of them in the same technique with good quality. The research can be done to increase the computational efficiency & to make video inpainting an intelligent. In this survey paper the classification is done in between Patch-based inpainting methods and object-based inpainting methods. This paper represents a brief review on various approaches which is used in video inpainting and it helps to make an intelligent object based inpainting approach for video repairing.

Keywords - Video Inpainting, patch based inpainting, Texture synthesis, Object based Inpainting, pattern recognition, Image restoration.

I. INTRODUCTION

Inpainting is also known as furnishing process through which we can fill the damaged portions of an image. It helps to create an ability of person to visualize the degraded part in the image. Inpainting process is not reserved for the single image. Rather than this inpainting can be used for fixing video frame sequences. Video inpainting is basically based on image inpainting so video inpainting and image inpainting are useful for camera image restoration, occlusion removal, adaptation in image and videos, accumulation in effects and much more. Video inpainting is a major area of research in the field of image processing. Pattern recognition is used in video inpainting process.

As compared to video inpainting more research is to be done in image inpainting because, when we consider video inpainting we need to cover considerably more number of pixels and required large search area. Rather than the spatial consistencies contained by a frame, video inpainting must make sure that temporal consistencies among the frames are well maintained.

In past years, there are many image inpainting algorithms are prescribed. Image inpainting algorithms are utilized in video inpainting. But some of the temporal aspects are not so good as to complete the expectations. Video Inpainting causes inconsistencies and some unwanted artifacts in results. So video inpainting is often considered as much more challenging problem as compare to image inpainting. But still in many cases image inpainting algorithms are successfully applied in video inpainting algorithms directly, but by considering some extensions such as Object based inpainting segmentation, tracking and fragment merging approaches for reducing the video inpainting problems.

Here analysis of inpainting algorithms are depends on how the occluded images got fix. Categorization is based on methods used such as texture-synthesis based methods, partial difference equations based and patch based techniques, and Object based inpainting techniques. Basically texture synthesis method chooses input texture which is replaced into the spoiled or lost objects in frame sequence. In PDE based method data propagation approach is from boundary to centre of the occlude region. The patch based image inpainting was derived to maintain the structural information of image but it caused blurring artifacts. As per results the patch based image inpainting technique is successful to extend this concept for video inpainting. So video inpainting is much more challenging for many methods that have been proposed in Video inpainting. Video inpainting methods basically can be classified into two types such as Patch based inpainting method and Object based inpainting method.

II. PATCH BASED METHODS

Patch based inpainting methods for video inpainting have been impressed from the various patch based inpainting methods which are generally use for image inpainting. Up till now many patch based video inpainting algorithms proposed. Patch based algorithms efficiently works only when it is function for image inpainting. But the patch based methods are having complication when handling spatial and temporal continuity when it relate with video inpainting.

Navier Stokes video inpainting algorithm [1] work under the patch based video inpainting algorithms provides the
application for computational fluid dynamics. It generates Navier-Stokes equation for fluid dynamics. To propagate isothingo andes from the exterior continuously it uses Classical fluid dynamics concepts. Image intensity is a stream function which is used for a 2-D incompressible flow. Degraded information got automatically replaced to the inpainting region which is chosen by user. Different structures and surrounding backgrounds is allowed for simultaneous filling of multiple regions. Navier Stokes video inpainting algorithm considered the spatial aspects only for video inpainting with limited applications. This technique produces good results and there are no colors or motion artifacts occurred. In this technique there is one drawback that it is not suitable for filling large non textured holes.

Wexler’s approach for video inpainting [2] the video completion is the global optimization approach for handling both ordered and unordered dynamic objects automatically. In Wexler’s approach image sequences which do not exist in the dataset are formed. It uses various space time patches selected from different parts of the video sequence. The suitable objective function removes inconsistencies and injuries by using patches and rank with video completion quality. Particulate sizes of cubes with 3-D were used as unit of similarity measure, which measure the Sum of Squared Differences. Here spatial and temporal information are handled simultaneously and multi-scale nature of results causes blurring. The overall method is very slow. This technique can complete frames or portions of frames that never existed in the dataset. In this technique both spatial and temporal information’s are handled simultaneously. Due to Multi-scale nature this technique causes blurs, expensive and high resolution videos are not accepted.

Jia et.al approaches a video completion in [3] using tracking and fragment merging. Texture synthesis is used for holes removal from the image sequence completely. But texture synthesis methods cannot use directly in video inpainting algorithms due to presence of large amount of data in video sequences and source fragment searching in the whole dataset is slow. This is an intellectual process which selects the most appropriate goal at the boundary of the hole. Space-time goal fragment is followed by choosing source fragment which is similar to the goal fragment from the search region using Mean Shift Algorithm (MLA) and finally merging the source and goal fragments which results to reduce the size of the hole. Tracking is used to reduce search space of source fragments size. Temporal consistency ensures consistency of the source fragments. This technique is suitable only when the object motion is periodic in nature and when an image does not change its scale. The region filling process may fails if the object tracking process is interrupted, and it occurs artifacts in the hole region. While comparing Navier Stokes & Wexler’s approach both it is observe that Wexler’s approach is more efficient and quite faster if search space is carefully selected. This technique work fast since holes are completed frame-by-frame and produce visually appealing results. It suffers from flickering artifacts at the hole region. Object tracking is lost when Filling process may get fail and restricted for object motion is periodic and has no scale change.

Cheung et.al introduced video epitome [4] which is patch-based probability models that are learned by compiling together a large number of examples of patches from input images. A video which is useful for video inpainting is containing the basic structure and motion characteristics of video. Probabilistic patch based model is used where patches are used for synthesizing damaged images sequence from videos. Best applicable patches from a part of the image sequence are stitch together to synthesize new faithful image with exact texture and fills missing or occluded regions in the image sequences taken from video. It provides a representation which retains the natural flow of input data and provides computational and statistical advantages than patch based libraries. Epitomes trained corrupted or degraded data when data is repetitive. This method results low resolution and containing over smoothing artifacts.

Shen et.al proposed in video completion for perspective camera under constrained motion [5] an approach to reconstructor the hidden static background and shaking foreground pixels while camera is in motion or stationary. By creating manifolds of space-time volume, search space is reduced from three dimensional to two dimensional. There are two advantages of this system such as good temporal coherence and providing effective periodic motion. It approach based on the motion manifold, which is a good representation of temporal coherence. Basically it is efficient for Pan Tilt zoom cameras which repairs video with some distortion related to varying foreground and global illumination. This technique can synthesize data without structure information. In this paper results are low resolution and containing over-smoothing artifacts. In this technique large area inpainting done properly and static background filled is consistent. Here only stationary cameras used and this technique is Unable to handle completely occluded objects.

Patwardhan et.al proposes inspecting missing parts of a video recorded with moving or stationary camera [6] for simple pre-processing stages followed by video inpainting. Pre-processing frames are segmented in foreground and background, helps the creation of image mosaics named foreground mosaic, background mosaic and optical flow mosaic. It helps to reduce the search space of algorithm. In foreground inpainting the holes caused by the moving objects in the foreground are filled by information from moving foreground in other frames using a priority based approach. Background inpainting inplanes frames arranged, copy pixels directly, and remaining pixels packed by updating spatial texture synthesis method. This Method assures better visual effect for every frame, finish in improved and disorderly
backgrounds which is unstable for arbitrary camera motions, and the case of complete occlusion does not addressed. In this technique temporal continuity along with the temporal axis is not maintain in well manner. So it adds some flickering artifacts. In this technique there is one condition is that PTZ and stationary cameras are works properly. Main advantage is that Spatial and temporal coherence, as well as periodic motion patterns all are well maintained. This technique is unable to handle spatio-temporal dimensions simultaneously. Computationally this technique is expensive, complex & time consuming.

Amanna Ghanbari, Mohsen Soryani proposes in contour based video inpainting [7] assumes, background is stationary and foreground object is in periodic motion. It separates foreground and background of image by using a threshold value. Squared window is set to boundary of moving object, set ideal windows together and then performs image mosaicing. At last finish work with background inpainting using exemplar based method. Contour based video inpainting overcomes the over smoothing effect of background, consumes less time. But run time quality of video maintenance is a real challenge. It works in rich and cluttered background works properly. In this technique spatial consistency maintained but temporal consistency can’t maintain, hence it produces flickering artifacts.

K. A. Patwardhan proposes implant technique to load missing part of the video [8] which has to be taken by static camera. Main goals are loading the static background keeping as it is with temporal consistency and load moving foreground keeping with globally consistency. Typical values are assigned to all pixels in the frame. Background filling involves temporal as well as spatial filling. The moving object completion is done by one by one frame. It has better performance when inpainting region is large. This method having one drawback is that the algorithm is appropriate for stationary cameras only. Produces few ghost shadows unlike other inpainting algorithms. It deals with different camera motions. Continuous block matching and selecting are major issues which are not handled properly.

T. K. Shih, N. C. Tang, and J.-N. Hwang proposes exemplar based video inpainting technique without ghost shadow artifacts [9]. In this different camera motion segments with differentiate temporal continuity takes different candidate patches. This results to inpaint holes after electing video object after tracking and removing. Basically algorithm generates very less ghost shadows dissimilar most inpainting algorithms which are directly applied on videos. The earlier algorithms also modified to handle complicated camera motions but here multiple objects removed serially and then we have to run the program still getting faithful results. This technique removed over-smoothing effect from the background and it can find best patch quickly. In this paper obtained results are not acceptable when there is non-stationary background and non-periodic motion occurred. And one main challenge is that to reduce the running time while keeping the quality.

B. A. Ahire and N. A. Deshpande propose “Video inpainting of objects using modified Patch based technique” [20]. An object-based video inpainting scheme maintains the spatial as well as temporal motion continuity of an object simultaneously. In this system also handles the problem of deficiency of available postures. The system consist four modules: Object detection and tracking, virtual contour construction, Key posture selection and mapping and synthetic posture generation. In first, object is detected & tracked with the help of background subtraction algorithm. In next module, scheme is proposed that it is able to derive the virtual contour of an occluded object. High resolution videos can handle by this technique. This technique work fast & Provide simple solution to solve over-smoothing problem. One challenge is that the spatio-temporal patches present in result can’t be handled. It is hard to choose that, which patch is more effective in terms of dimensionality and reduction in computational cost.

Waykule J.M. proposes “Modified Image Exemplar-Based Inpainting” [19] a novel algorithm for removing large objects from digital images. The result is an image in which the selected object has been replaced by a visually believable background that mimics the look of the source region. This technique provides preservation of edge sharpness, No dependency on image segmentation, Balanced region filling to avoid over-smoothing artifacts, speed efficiency, Accuracy in the Synthesis of texture, accurate propagation of linear structures. Still this algorithm is too slow.

III. OBJECT BASED INPAINTING METHOD

Video inpainting is can be done by object based method. By using object based inpainting method both spatial and temporal consistencies can covered one by one. In Object based inpainting Method we can improve quality of visual aspects. In object based inpainting we can more concentrate on the occluded object. So as to improve overall damaged video sequence.

Jia et al. proposed in video repairing under variable illumination using cyclic motions [10] a user can made layer segmentation of complete video which classify the video sequence into color and illumination videos followed by a tensor voting technique which maintains spatial as well as temporal consistency. Here the available pure or ideal objects are synthesized to recover occluded objects. In background textures are considered that’s why synthesized objects do not have a real trajectory. The inpainted video stores pieces of spatial as well as temporal structural consistency and variable illumination. But it has one drawback that is it cannot fix shadows of injured object and there are chances of incidence of mistaken lighting on the fixed move. In this technique large class of camera motions are used. Due to abrupt change of light
& shadow this technique generates artifacts and it has restricted for complex structures.

Zhao et al proposes an efficient object based video inpainting technique for video recorded by stationary cameras. [11]. Mostly background pixels are well-suited to previous one frame which is used for background inpainting and available objects are used for foreground inpainting. A fix size sliding window including set of continuous objects and concludes the similar function which is to be measures similarity between two continuous templates of object. Main ability of this technique is to recover holes and errors, where the occluded objects totally missing from video sequences. This technique is not having sufficient number of postures which produces unsatisfactory test results. This technique does not suggest a systematic way to recognize good filling objects. In this paper variable illumination & spatio-temporal Consistency are well maintained, but this technique is limited for aperiodic motion. And one more drawback is that shadows of damaged moveal are unable to get repair.

Ling et.al proposes in another object based video inpainting technique [12] handles spatial and temporal both aspects one by one was projected. It can also handle the over smoothening artifacts and insufficiency postures sequence available. Virtual contour construction is done at the start and a posture sequence matching is applied on it to retrieve the most similar sequence of postures from the available non-occluded postures. At last key posture selection process is done and posture sequence estimation is done to solve substring search problem. In case there are no good matches among the available postures in the database synthetic key postures are generated. Result generates objects with good subject quality but cannot tackle condition where the object movement is random. In this paper user can address videos from both static and moving cameras and it can handle large holes including cases where the occluded object is completely missing from several frames. For satisfactory results numbers of postures are sufficient. The drawback of this technique is that no systematic way to find a good filling position.

Lin et al in virtual contour guided video object inpainting using posture mapping and retrieval [13] proposed a method which mainly has three main parts such as virtual contour construction, key posture selection, posture mapping and synthetic posture generation. This method considered that the trajectory of the occluded objects can approximate to linear line segmentation during the period of occlusion. Mosaic schemes as well as correspondence maps are used for background inpainting. This method is tried to avoid problems caused due to insufficient number of postures due to data set. But the synthetic posture generation technique which is used is not suitable for complex postures generation and do not deal with the illumination change problem.

Tang et.al proposes another video repairing approach [14] by projected to assume the missing static background and moving foreground due to damage or occlusion in video sequence. Expansion to various video repairing methods used for recovering background pixels, layer segmentation and homographic blending is used to preserve temporal coherence and flickering effects. For repairing moving foreground involves sampling as well as alignment phases. After that reference video mosaic used, so system can process for large classes of camera motions. This method occurs drawback that it is not applicable for more complex dynamics problems. Due to that a few spatial and temporal both artifacts are observed at a time.

Ding et al proposes a rank minimization approach for video inpainting [15]. A manifold based learning method and local linear embedding was used here to map the image to low-dimensional space. The method works in three parts. First is that set of descriptors with the necessary information to reconstruct the frame was found damage. Second is that to find optimum values for descriptors were found. And at last frames were reconstructed based on estimated values. Wexler’s approach to find the best matching patch from the adjacent frames used for filling, difference in the searching technique used. Using Rank Minimization Interpolation used for finding descriptors for the missing area which used for dynamic models thereby reducing the computational complexity. This technique also permits to use video sequences taken by moving cameras. This method is non-iterative and computationally attractive algorithm and not restricted for periodic motion. But this technique having drawback that it cannot handle scaling information and causes blurring and ghost image artifacts where the object’s motion is random due to the RBF inverse mapping.

Chih-hung et al proposed another posture sequence estimation method [16]. Again this is a manifold based learning approach. Here human objects are considered and involves mainly in three main steps such as human posture synthesis, graphical model construction and posture sequence estimation. There are two constraints introduced by this method to manage the motion continuity. First one controls the searching range and second one constraint controls search direction that’s why making the reconstructed motion continuous. In this approach the postures from same video are taken to complete the video, it removes blurring as well as ghost artifacts. But as mentioned in above all algorithms object based algorithms are not handle object segmentation sharply so it will cause disagreeable results. This technique is applicable for both periodic and non-periodic motion and the reconstructed video has continuous motion in nature. Due to inaccurate segmentation it generates artifacts.

Alasdair Newson, Matthieu Fradet, Patrick Pérez, Andrés Almansa, Yann Gousseau, “Towards fast, generic video inpainting” [21], automatic video inpainting with global optimization of a patch-based functional in reasonable execution times is possible, even high resolution videos are
accepted. When extending Patch-Match algorithm in case of spatio-temporal patches, this is able to provide a fast, useable video inpainting algorithm. A fast, simple solution to a problem of over smoothing of video inpainting results which particularly problematic in the case of video textures. This work help to make experimentation with automatic, generic video inpainting easier and therefore make further progress in the domain possible. Even high resolution videos are handled. It Provide simple solution to solve over-smoothing problem new challenge is that, the spatio-temporal patches present in result. It is hard to choose that which patch is more effective in terms of dimensionality reduction and computational cost.

IV. COMPARISON OF VARIOUS ALGORITHMS

The below table shows different video inpainting algorithms used by different authors. Each algorithm has its own merits and demerits. Table (1) describes merits and demerits of the various algorithms which have been discussed below.

<table>
<thead>
<tr>
<th>Technique/Methodology used</th>
<th>Approach</th>
<th>Year</th>
<th>Merits</th>
<th>Demerits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patch Based Techniques</td>
<td>Navier-stokes, fluid dynamics, and image and video inpainting.</td>
<td>2001</td>
<td>This method is straightforward, shows good results: sharp, no color &amp; motion artifacts.</td>
<td>Not suitable for filling large non-textured holes.</td>
</tr>
<tr>
<td>Patch Based Techniques</td>
<td>Space-time completion of video.</td>
<td>2004</td>
<td>Can complete frames or portions of frames that never existed in the dataset. Both spatial and temporal information’s are handled simultaneously.</td>
<td>Due to Multi-scale nature blurring is occurred. More expensive. High resolution videos are not accepted.</td>
</tr>
<tr>
<td>Object Based Techniques</td>
<td>Video repairing: Inference of foreground and background under Severe occlusion.</td>
<td></td>
<td>Large class of camera motions is used.</td>
<td>Due to abrupt change of light &amp; shadow generates artifacts. Restricted for complex structures.</td>
</tr>
<tr>
<td>Patch Based Techniques</td>
<td>Video completion using tracking and fragment merging</td>
<td>2005</td>
<td>Work fast since holes are completed frame-by-frame. Produce visually appealing results.</td>
<td>Suffers from flickering artifacts at the hole region. Object tracking is lost when Filling process may fail. Restricted for object motion is periodic and has no scale change.</td>
</tr>
<tr>
<td>Patch Based Techniques</td>
<td>Video epitomes</td>
<td></td>
<td>Can synthesize data without structure information.</td>
<td>Results are low resolution and containing over-smoothing artifacts.</td>
</tr>
<tr>
<td>Patch Based Techniques</td>
<td>Video inpainting of occluding and occluded objects</td>
<td></td>
<td>Large area inpainting done properly. Static background filled-in is consistent.</td>
<td>Only stationary cameras used. Unable to handle completely occluded objects.</td>
</tr>
<tr>
<td>Patch Based Techniques</td>
<td>Video completion for perspective camera under constrained motion.</td>
<td>For PTZ and stationary camera works properly. Spatial and temporal coherence, as well as periodic motion patterns, are well maintained.</td>
<td>Unable to handle spatio-temporal dimensions simultaneously. Computationally expensive, complex &amp; time consuming.</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Object Based Technique</td>
<td>Video repairing under variable Illumination using cyclic motions.</td>
<td>Variable illumination &amp; spatio-temporal Consistency well maintained.</td>
<td>Limited for aperiodic motion. Shadows of damaged moveel are unable to repair.</td>
<td></td>
</tr>
<tr>
<td>Object Based Technique</td>
<td>Efficient object-based video inpainting.</td>
<td>Can address videos from both static and moving cameras. Handle large holes including cases where the occluded object is completely missing from several frames.</td>
<td>For satisfactory results numbers of postures are sufficient. No systematic way to find a good filling position.</td>
<td></td>
</tr>
<tr>
<td>Patch Based Techniques</td>
<td>Video inpainting under constrained Camera motion.</td>
<td>In rich and cluttered background works properly.</td>
<td>Maintain spatial but temporal consistencies can’t maintain. Hence it produces flickering artifacts.</td>
<td></td>
</tr>
<tr>
<td>Object Based Technique</td>
<td>A rank minimization approach to Video inpainting.</td>
<td>Applicable for non-periodic motion. Gives non-iterative and computationally attractive algorithm.</td>
<td>It unable to handle scaling or deformations. Generates blurring and ghost image artifacts if the objects motion is not periodic.</td>
<td></td>
</tr>
<tr>
<td>Patch Based Techniques</td>
<td>Exemplar-based video inpainting Without ghost shadow artifacts by maintaining temporal continuity.</td>
<td>Produces few ghost shadows unlike other inpainting algorithms. It deals with different camera motions.</td>
<td>Block matching and selecting Continuous blocks are major issues which are to be handled.</td>
<td></td>
</tr>
<tr>
<td>Patch Based Techniques</td>
<td>Contour based video inpainting.</td>
<td>Over-smoothing effect is removed from the background. Best patch can find quickly.</td>
<td>Obtained results are not acceptable when there is non-stationary background and non-periodic motion. Main challenge is to reduce the running time while keeping the Quality.</td>
<td></td>
</tr>
<tr>
<td>Object Based Technique</td>
<td>Virtual contour-guided video object inpainting using posture mapping and retrieval.</td>
<td>Avoids over-smoothing artifacts, and compensates insufficient number of postures.</td>
<td>For getting virtual contour, the object must moves linearly during occlusion. The synthetic posture synthesis is restricted for complex postures.</td>
<td></td>
</tr>
<tr>
<td>Object Based Technique</td>
<td>Human object inpainting using manifold learning-based posture sequence estimation.</td>
<td>Applicable for both periodic and non-periodic motion. The reconstructed video has continuous motion in nature.</td>
<td>Due to inaccurate segmentation it generates artifacts.</td>
<td></td>
</tr>
<tr>
<td>Patch Based Techniques</td>
<td>Video inpainting of objects using modified Patch based technique</td>
<td>Helps to remove objects with good subjective quality in terms of the object’s spatio-</td>
<td>If object moves nonlinearly during an occlusion period then the virtual trajectories</td>
<td></td>
</tr>
<tr>
<td>Patch Based Techniques</td>
<td>Modified Image Exemplar-Based Inpainting</td>
<td>2013</td>
<td>This technique provides preservation of edge sharpness, No dependency on image segmentation, Balanced region filling to avoid over-smoothing artifacts, speed efficiency, Accuracy in the Synthesis of texture, accurate propagation of linear structures.</td>
<td>Still algorithm is too slow.</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------------------------</td>
<td>------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Object Based Techniques</td>
<td>Towards fast, generic video inpainting</td>
<td></td>
<td>Even high resolution videos are handled. Technique is fast. It Provide simple solution to solve over-smoothing problem.</td>
<td>New challenge is that the spatio-temporal patches present in result. Hard to choose which is more effective patch in terms of dimensionality reduction and computational cost?</td>
</tr>
</tbody>
</table>
V. CONCLUSION

This survey shows the various patch based and object based video inpainting algorithms works till now. Patch based video inpainting algorithms are usually extended from the existing image inpainting algorithms while object based video inpainting algorithms was introduced. All algorithms have its own merits and demerits on the basis of performance. Most of the patch based algorithms were unable to maintain the spatial and temporal consistency simultaneously and blurring artifacts are added in results. In object based inpainting technique having more advantages than patch based technique. Inpainting technique selection can be done on the basis of requirements.

VI. REFERENCES


