Nuclear Engineering and Technology 54 (2022) 1153–1160

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

journal homepage: www.elsevier.com/locate/net

Original Article

A review of missing video frame estimation techniques for their suitability analysis in NPP

Mrityunjay Chaubey [a], Lalit Kumar Singh [b], Manjari Gupta [a, *]

[a] Computer Science, Centre for Interdisciplinary Mathematical Sciences, Institute of Science, Banaras Hindu University Varanasi, 221005, India
[b] Department of Computer Science & Engineering, IIT (BHU), Varanasi, India

ARTICLE INFO

Article history:
Received 13 June 2021
Received in revised form
19 September 2021
Accepted 10 October 2021
Available online 14 October 2021

Keywords:
Crack
Frame
Nuclear power plants (NPP)
Reactor
Safety
Video processing

ABSTRACT

The application of video processing techniques are useful for the safety of nuclear power plants by tracking the people online on video to estimate the dose received by staff during work in nuclear plants. Nuclear reactors remotely visually controlled to evaluate the plant's condition using video processing techniques. Internal reactor components should be frequently inspected but in current scenario however involves human technicians, who review inspection videos and identify the costly, time-consuming and subjective cracks on metallic surfaces of underwater components. In case, if any frame of the inspection video degraded/corrupted/missed due to noise or any other factor, then it may cause serious safety issue. The problem of missing/degraded/corrupted video frame estimation is a challenging problem till date. In this paper a systematic literature review on video processing techniques is carried out, to perform their suitability analysis for NPP applications. The limitation of existing approaches are also identified along with a roadmap to overcome these limitations.

1. Introduction

Seven of the UK's eight nuclear power facilities are Advanced Gas-cooled Reactors (AGRs), which supply enough electricity to cover roughly 20% of annual demand. The core of the AGR plant is made up of hundreds of channels made out of cylindrical graphite bricks that house the fuel and allow control rods to be inserted to control the reaction. Specialized equipment and tools with various sensors, including a video camera, are used to inspect specific fuel channels. The camera is used for Remote Visual Inspection (RVI) and records footage of the inner wall of the fuel channel in six separate but overlapping orientations. In the inspection video, if any frame got degraded or missed due to any reason then it will create serious problem in the inspection process. For that, we need a reliable and robust technique to estimate the missing/degraded frame.The need for the recovery of missing pixels in any image falls under the category of image restoration, which plays a significant role in many real-world applications such as medical imaging, defense, etc. Some of the well-established methods to restore the images are Median filter, Adaptive filter, Linear filter, Iterative blind deconvolution, Non-negative and support constraints recursive inverse filtering, Super-Resolution restoration algorithms based on gradient adaptive interpolation, Deconvolution using sparse prior, Block matching, Weiner filter, deconvolutionizing regularized filter(DRF), Lucy-Rechardson algorithm techniques, neural network approach [1]. These filters are used in image restoration based on the need and efficiency of the filter. Images and videos have some similarities and differences, such as the fact that the images are 2D/3D spatial distribution of intensity with constant time, while videos are a 3D/4D Spatio-temporal intensity pattern. The missing frames in any video can be estimated, similar to the case of images. Missing frames in any video can be restored based on the available information in the locality (both spatial and temporal) of the missing data. Motion vector recovery techniques attempt to recover lost motion information in compressed video streams and because of its simplicity and reliability, block-matching motion estimation (BMME) is the most widely used method to find the best-matched block; the BMME algorithms use search patterns. A simple technique for error concealment is bi-linear interpolation in which each pixel in the lost area is interpolated from intact neighboring pixels. B. Yan et al. [2] proposed a hybrid motion vector extrapolation algorithm to provide a more precise estimate of the motion vectors of the missing frames, while a well-known frame rate up-conversion technique may improve the visual quality of low frame rate images

* Corresponding author.
E-mail address: manjari@bhu.ac.in (M. Gupta).

https://doi.org/10.1016/j.net.2021.10.012

[3]. A content-adaptive vertical temporal filtering method is used for deinterlacing which effectively uses correlations between adjacent frames [4],23], in which motion estimation is not used and hence provides low complexity.

We attempt to provide a rigorous review of existing techniques on the estimation of missing/degraded video frames to identify the advantages and limitations of these techniques along with suitability analysis of techniques and a roadmap for future work. The organization of the paper is as follows. In Section 2 research methodology is given and Section 3 discusses the existing techniques proposed by different researchers. Section 4 gives a comparative study of these existing techniques along with suitability analysis. The significant findings are brought out in section 5 and section 6 concludes the paper and provides a roadmap for future work.

2. Review methodology

The primary objective of this study is to conduct a systematic literature review in order to identify the most advanced techniques for estimating missing video frames and determining their suitability for usage in NPP inspection videos. Our systematic literature review is conducted by searching for research papers published between 1998 and 2020, on IEEE and Web of Science databases. The databases were screened for titles, abstracts, keywords containing phrases “Missing video frame estimation” AND (“Motion estimation” OR “Video restoration” OR “Frame interpolation”). The search was limited to full-length journal articles written in English language only. As a result, we got 51 articles related to the keywords. After removing duplicates and conference papers, we read titles and abstracts of publications. References that did not investigate missing video frame estimation were excluded from further screening. In the following step, we filtered out studies that does not contain the missing video frame estimation techniques. We included 20 articles published in various reputed journals for our literature review shown in Table 1.

We defined some criteria for the suitability analysis of existing missing video frame estimating techniques in NPP inspection videos, and based on those criteria, we can conclude if a methodology is relevant in the NPP inspection video or not.

The following are the suitability criterion for using the existing techniques in the NPP inspection video:

i) By comparing the video dataset's characteristics to the NPP inspection video.
ii) If the characteristics of the used video dataset and the NPP inspection video match, then the given technique will be equally applicable to the NPP inspection video.
iii) If the characteristics of the used video dataset and the NPP inspection video are not similar, we will customise the NPP inspection video and apply the existing technique based on the advantages and limitations of the existing technique.

3. Related work

To overcome the issues of missing video frames,a hybrid frame concealment algorithm for H.264/AVC is proposed by B.Yan et al. [1], in which the limitations of pixel-based motion vector estimation are addressed.The pixel in the missing frame can be divided into two categories.

i) For a pixel that is covered by at least one extrapolated macroblock (MB); In this case, the motion vector(MV) is estimated by averaging the MVs of all overlapped MBs.
ii) For a pixel that is not covered by any of the extrapolated MBs; In this case, MV is estimated by duplicating the MVs of the same pixel in the previous frame.

If the estimated motion vector is:
$MV = (MV_x, MV_y)$, each missing pixel $p_m(x, y)$ can be recovered as

$$p_m(x, y) = p_r(x + MVx, y + MVy)$$

where $p_r(x, y)$ refers to the pixels in the previous frame. The hybrid motion vector extrapolation approach [2] gives a more accurate estimate of the motion vector of the missing frame at low frame rates than standard methods. If the NPP inspection video has the same characteristics as the used dataset video, then the same technique can be applied to estimate the missing frames if the block matching criterion threshold is properly set.

Another algorithm was proposed by Ci Wang et al. [3], in which frame rate up-conversion, using trilateral filtering, is used.

Let, in the unidirectional motion, f_{t-1} andf_{t-2} are two consecutive frames, and $f_t$is the intermediate frame to be interpolated.

Let the motion be uniform, then f_t can be modeled as

$$f_t(x, y) = \frac{1}{2}\left(f_{t-1}\left(x - \frac{1}{2}\Delta x, y - \frac{1}{2}\Delta y\right) + f_{t+1}\left(x + \frac{1}{2}\Delta x, y + \frac{1}{2}\Delta y\right)\right) + n_t(x, y)$$

where $(x,y)$and $(\Delta x, \Delta y)$ are the pixel location in the unidirectional motion vector between f_{t-1} and f_{t+1} respectively and $n_t(x, y)$is the noise.

The frame rate up conversion algorithm comprises three components.

i) Motion Estimation
ii) Initial estimation
iii) Trilateral filtering

This method reduces the cost of refining motion vectors while also reducing interpolation noise and miss registration errors. This technique can be applied in NPP inspection videos by selecting the appropriate search strategy, block matching criterion, and block size.

Nick C. Tang et al. [6] proposed a video inpainting algorithm to repair damaged contents in vintage digitized film, which concentrates on maintaining good Spatio-temporal continuity.In this method, a video inpainting algorithm is used to repair damaged contents in vintage digitized film, which concentrates on maintaining good spatio-temporal continuity.

There are two key techniques used in this algorithm.

i) Motion completion recovers missing motion information in the damaged area to maintain good temporal continuity.
ii) Frame completion repairs damaged frames to produce a visually pleasing video with good spatial continuity and stabilized luminance.

Table 1
Number of articles selected from different databases.

Article Database	No. of Articles
IEEE Transactions	16
Elsevier	2
SPIE Digital Library	1
Springer	1

For videos with unstable motion and luminance, motion completion and frame completion algorithms are utilised, and they work better if the damaged content does not span a broad area. We can utilise this approach to estimate missing or degraded frames in NPP inspection videos if the aforementioned conditions are met.

Multiple hypothesis Bayesian frame rate up-conversion by adaptive fusion of motion-compensated interpolations is proposed by Hongbin Lu et al. [7]. In the proposed work, they have given a multiple hypothesis Bayesian frame rate up-conversion scheme for estimating the intermediate frame with maximum aposteriori probability. By considering the time complexity of the reconstructed frame, this technique improves the objective and subjective quality of the reconstructed frame. This method can also be used to approximate the missing frames in NPP's inspection video.

Let f_t be the intermediate frame to be estimated, and f_{t-1} and f_{t+1} are the previous and following neighboring frames of f_t, respectively. Here the go al is to find the most probable pixel of f_t based on f_{t-1} and f_{t+1}.

Mathematically it can be given as

$$\widehat{f_t} = \underbrace{argmax}_{\widehat{f_t}} P_r(f_t|f_{t-1}, f_{t+1}) \tag{1}$$

where $P_r(.)$ is the probability density function.

Considering the impact of motion, equation (1) can be formulated as

$$\begin{aligned} \widehat{f}_t &= \underbrace{argmax}_{\widehat{f_t}} \int P_r(f_t, m_t|f_{t-1}, f_{t+1}) dm_t \\ &= \underbrace{argmax}_{\widehat{f_t}} \int P_r(f_t|f_{t-1}, f_{t+1}) P_r(m_t|f_{t-1}, f_{t+1}) dm_t \end{aligned} \tag{2}$$

Instead of employing a single uniquely optimal motion, multiple optimal motion trajectories are utilised. To obtain accurate estimation for the pixels in missing intermediate frames, the motion-compensated interpolations generated by all these motion hypotheses are adaptively fused according to the reliability of each hypothesis. To obtain the multiple motion field, a set of block matching sizes is used and the motion fields are estimated by progressively reducing the size of the matching block.

Kwok-Wai Hung et al. [8] proposed an algorithm that discusses a low complexity, computationally scalable, and data-adaptive image interpolation. The image interpolation as a denoising problem was formulated by proposing a new image model to relate the observed low-resolution pixels and missing high-resolution pixels. Applying the maximum-likelihood estimation using the new image model results in an adaptive linear filter, where the filter coefficients depend on the local noise covariance matrix, which is estimated by the local noise sample.

Data adaptive image interpolation and maximum likelihood estimation approaches are utilised because of their simple structure and low complexity, but they cannot be used directly for missing video frame estimation since the interpolation problem must first be turned into a denoising problem. As a result, this technique cannot be used directly to the NPP inspection video.

Haichao Zhang et al. [9] proposed a non-local kernel regression model for various image and video restoration tasks. Their method exploits both the non-local self-similarity and local structural regularity properties in natural images. The non-local self-similarity is based on the observation that image patches that tend to repeat themselves in natural images and videos, and the local structural regularity observes that image patches that have regular structures where the accurate estimation of pixel values is possible via regression. By unifying both the properties explicitly, the non-local kernel regression framework is more robust in image estimation, and the algorithm applies to various image and video restoration tasks. Taking into account the computational complexity, this method produces more accurate results and can be used in NPP inspection videos.

An algorithm proposed by Farhang Vedadi et al. [10] uses non-local cost not only to measure the fitness of the interpolator at the missing pixel position but also to approximate the transition matrix of Markov chain interpolators. The viterbi algorithm was used to find the optimal overall interpolating sequence while keeping the cost function and nearby pixels in hand. It automatically updates the transition matrix for individual frames. Now, maximum aposterior formula was used for estimating the interpolating sequence. Forward-backward algorithm was used to estimate the most successive interpolators in each of the missing pixels.

This method produces better results and may be applied to NPP inspection videos if the properties of the utilised dataset and the NPP inspection video match, but it is computationally more complex due to several assumptions.

Gi-Hyun Na et al. [11] proposed a frame-based recovery of corrupted video files using video codec specifications. Their method addresses how to extract video frames from a portion of the video to be restored as well as how to connect extracted video frames together according to the codec specifications. A video restoration technique for fragmented and partially overwritten video files was proposed and it guarantees the integrity of the restored frames.

If the degraded content does not cover a large area, the frame-based recovery method is used, which recovers the most video data from non-overwritten video segments. If the parameters of the used video dataset match those of the NPP inspection video, this technique can be used to estimate missing video frames in the video.

Ting-Lan Lin et al. [12] proposed an improved interview video error concealment on entire frame loss. An improved depth image-based rendering(DIBR) based whole frame error concealment method for multi-view video with depth was designed. Then, optimal reference view selection was developed. Correctly mapped pixels from another view were used as an initial estimate for one-to-one mapping. By the previous frame, the illumination compensation was estimated across the views of the current frame. Motion vectors for the mapped pixels were taken and a reverse DIBR is applied. Many to one pixel were generated with the better decision using depth information, producing better pixel quality. The hole pixels were efficiently recovered with the estimated motion vectors. It was computed using the neighboring available motion vectors and weighted by the inverse of the spatial distance.

Considering the limitations of this technique, we can utilise this approach to estimate the missing video frames of the NPP inspection video.

Chih-Chung Hsu et al. [13] proposed a method, in which, they addressed the problem of hallucinating the missing high frame resolution details of low-resolution video while maintaining the temporal coherence of the reconstructed high-resolution details using dynamic texture synthesis. Their method divides the input low-resolution video frames into key-frames and non-key frames. The texture synthesis-based super-resolution method was used to upscale each key-frame, followed by a low complexity bidirectional overlapped block motion compensation that reconstructs the high-resolution details of each non-key frame between two successive anchors key-frames.

This technique improves subjective and objective visual consistency, thus it may be utilised in the NPP inspection video, but it is less successful in capturing motion patterns if the texture includes persistent motions.

Jing Ge et al. [14] proposed a spatio-temporal super-resolution method to enhance both the spatial resolution and the frame rate in a hybrid stereo video system. In this system, a scene was captured by two cameras to form two videos, including a low spatial resolution with high frame rate video and a high resolution, with low frame rate video. For the low spatial resolution video, the low-resolution frames were spatially super-resolved by the high-resolution video via the stereo matching, bilateral overlapped block motion estimation, and the additive overlapped block motion compensation, while for the low frame rate video, these missed frames were interpolated using the high-resolution frames obtained by fusing the disparity compensation and the motion compensation frame rate up-conversion.

This technique provides both subjective and objective properties of a mixed spatiotemporal superresolution approach, and it can be employed in NPP inspection videos if the characteristics of the used dataset video match those of the NPP inspection video.

Xang Xu et al. [15] proposed a method to complete the holes left by removing objects with motion-guided pixels assignment optimization. Initially, the motion field in the holes was estimated by applying a two-step motion propagation method. Then, using the estimated motion field as guidance, the missing parts were completed by performing pixels assignment optimization based on the Markov process. It optimally assigns available pixels from other neighboring video frames to the missing region. Finally, an illumination adjusting approach to eliminate the illumination inconsistency in the completed holes was applied.

This method is used to fill the hole, and it produces better outcomes than other methods. If the characteristics of the used video dataset and the NPP inspection video match, it will be useable in NPP inspection video, taking into mind the computational cost.

Pierre-Henri Conze et al. [16] proposed a multi-reference combinatorial strategy towards longer long-term dense motion estimation. In which, with computer vision applications such as video editing in mind, optical flows were estimated. It was done with various inter-frame distances and combine through multi-step integration and statistical selection.

The combination of the work was two-fold:

i) An exhaustive analysis of available single reference complexity reduction strategy.
ii) A simple and efficient alternative related to multi-reference integration and statistical selection was proposed.

The used method automatically inserts intermediate reference frames.

This technique is used to reduce complexity while enhancing the accuracy of missing frame estimation in low frame rate videos. If the characteristics of the used dataset and the NPP inspection video match, this technique is appropriate for NPP inspection videos.

Werui Hu et al. [17] proposed the twist tensor nuclear norm for video completion. In this, a new low-rank tensor model based on the circulant algebra, namely, Twist Tensor Nuclear Norm(t-TNN) model was tested on a video completion application that aims to fill missing values. Especially it was used while dealing with video recorded by a non-stationary panning camera. The block circulant matricization of the twist tensor can be transformed into a circulant block representation with the nuclear norm invariance. This representation exploits the horizontal translation relationship between the frames in a video after transformation. It endows the t-TNN model with a more powerful ability to reconstruct panning videos.

If the NPP inspection video contains insufficient information regarding the content in the frame, this procedure is used. For panning video, this method produces a more precise estimate, while for fixed surveillance video, it provides less accurate results.

D. M.Motiur Rahman et al. [18] proposed a new view synthesis technique that exploits temporal correlation for hole filling. The views were interpolated from adjacent texture images and their corresponding depth map. In this, several models in the Gaussian mixture separate the background and foreground pixels. The missing pixels were recovered from the adaptive weighted average of the pixel intensities from the corresponding models of the Gaussian mixture model and the wrapped images.

This technique considerably enhances the PSNR, and because of this, it can be employed in the NPP inspection video if the dataset and the NPP inspection video have similar characteristics.

Brian E.Moore et al. [19] proposed a panoramic robust PCA for foreground-background separation on noisy and free motion camera video. The proposed method records the frames of the corrupted video, encodes the varying perspectives caused by camera motion in a global model as missing data. This can generate a panoramic context portion that automatically stitches together corrupted data from partially overlapping frames to recreate the entire field of view.

If the used dataset and the NPP inspection video have similar characteristics, this technique will work similarly in NPP inspection video.

Muhammad Saquib et al. [20] proposed a crowd counting system for low to medium-density crowded images. The framework used a faster-recursive convolutional neural network to detect a pedestrian in crowd video, followed by a motion-guided filter to recover misdetections and thus increase the overall mean detection precision.

This method is used to estimate miss detections and count crowds. This technique may be applied for NPP inspection video as well, based on the similarity between the used dataset video and the NPP inspection video.

Avinash Paliwal et al. [21] proposed a deep slow-motion video reconstruction technique with a hybrid imaging system. The proposed techniques used a hybrid imaging method to present a deep learning approach for video frame interpolation. By coupling the low frame rate video with a high frame rate video with low spatial resolution, the lack of temporal information in the input video was resolved.

The two-stage learning method correctly combines the two videos. First, aligning the high-resolution neighboring frames to the target frame. Second, merging the matched images to recreate the high-quality frame. Finally creating simulated training data and perturbing it to fit the statistics of real data.

This method is better for both synthetic and real-world videos, and it is used for high-resolution videos with a very high frame rate. If the auxiliary frames are taken at a low resolution, this technique will not work. This method can be used in NPP inspection videos if the benefits and constraints are considered.

4. Comparative studies of the existing techniques

We performed an appropriateness study of the approaches utilised in the existing literature in this section. On this premise, we have concluded whether or not a certain technique for missing frame estimation can be employed in the NPP inspection video application. Table .2 is showing the study. We have adopted techniques that produce better results while being less computationally expensive.

At low frame rates, the hybrid motion vector extrapolation algorithm [2] produces a more accurate estimate of the missing frame's motion vector than traditional methods. If the block matching criterion threshold is properly selected, this technique will be applicable in NPP inspection videos with low frame rates.

Table 2
Investigation of different techniques used for missing video frame estimation with their advantages, disadvantages and suitability to use in NPP's inspection videos.

Paper	Year of Publication	Technique Used	Advantages	Disadvantages	Suitability to apply the used techniques in NPP inspection video.
B. Yan et al. [2]	2010	Hybrid motion vector extrapolation algorithm.	It gives an accurate estimate of the motion vector of the missing frame than conventional methods.	Computational cost and selection of a threshold for block matching criterion.	This technique is suitable to use in the inspection video of NPP with low frame rate
C. Wang, L. Zhang et al. [3]	2010	Frame rate-up conversion method followed by a trilateral filter.	Reduces the complexity to refine the motion vectors and also suppresses the interpolation noises and misregistration errors.	The performance of unidirectional motion estimation largely depends upon the search strategy, block matching criterion, and block size.	This technique can also be used in the NPP inspection video, because it also uses motion estimation technique with higher frame rate.
K. Lee et al. [4	2010	The modified content-adaptive vertical temporal filter, adaptive weighted vertical temporal filter followed by a field averaging filter.	This method does not use the motion estimation technique and has low complexity.	Spatial and temporal interpolation, which treats the entire frame in the same way regardless of motion in the video, is relatively simple, but the results are lacking when compared to methods that consider motion in the picture to apply corrections.	As this method is not using motion estimation techniques and also has limitations in quality of result, so the use of this method can be avoided.
E. Maani et al. [5]	2010	A cross-layer optimization approach for the H.264/AVC scalable extension.	This method ensures low computational cost.	There is a tradeoff between the size of the optimization window and computational complexity.	This method is applicable, when we try to stream the video through any channel.
N. C. Tang, C. Hsu et al. [6]	2011	Motion completion and Frame completion.	For unstable motion and luminance, it performs better.	If damaged content covers a wide area in each frame then visual defects may appear in the resulting video.	This method is suitable to use in the inspection video of NPP when damaged contaent does not cover wide area.
H. Liu, R. Xiong et al. [7]	2012	Multiple hypotheses Bayesian frame rate-up conversion.	The objective and subjective quality of the reconstructed frame is enhanced.	Require greater computation time.	This method can be used because it provides good quality of reconstructed frame.
Hung, Kwok-Wai et al. [8]	2013	Computationally scalable, data-adaptive image interpolation, maximum likelihood estimation.	A scalable algorithm has the simplest structure and minimal complexity.	The computational complexity is significantly higher than the conventional methods.	This technique can not be used directly because here the image interpolation problem is being converted into denoising problem first.
H. Zhang, J. Yang et al. [9]	2013	Non-local kernel regression.	It is more robust in image estimation.	Due to the patch matching process, the computational cost is more.	This technique can be applied directly in the NPP inspection video and it is robust.
F. Vedadi et al. [10]	2013	Markov-chain of interpolators.	Being competitive on different test sequences, it outperforms the existing de-interlacing algorithm.	Due to many assumptions, computational complexity becomes more.	It is applicable in NPP inspection video but has greater computational complexity.
G. Na, K. Shim et al. [11]	2014	Frame-based recovery technique.	Most video material in non-overwritten segments of video files can be recovered.	Its performance gets degraded if damaged content covers a large area.	This technique is applicable in NPP inspection video if degraded content does not contains large area.
Lin, T.L., Chang et al. [12]	2014	Improved Depth image-based rendering.	The integrated framework of the four proposed methods outperforms the state-of-the-art method	During rendering, unwanted objects (e.g. holes) along depth discontinuities can appear.	It is using motion estimation technique, so this technique is applicable to use in the NPP inspection video.
C. Hsu, L. Kang et al. [13]	2015	Texture-synthesis-based video super-resolution and dynamic texture synthesis.	It improves both subjective and objective visual consistency	In the case of textures having persistent motions, this method	This technique is applicable to reconstruct the high resolution details of the low resolution

(continued on next page)

Table 2 (*continued*)

Paper	Year of Publication	Technique Used	Advantages	Disadvantages	Suitability to apply the used techniques in NPP inspection video.
			compared to existing techniques.	is less successful in capturing the motion patterns.	video. It can be used in NPP inspection video accordingly.
Ge, Jing, et al. [14]	2016	Stereo matching, adaptive overlapped block motion compensation, disparity compensation, and motion compensation.	It improves both subjective and objective qualities of a mixed spatiotemporal super-resolution approach.	The stereo matching technique used has some disadvantages as • Occlusions. • Matching ambiguity.	This technique is applicable for those videos which has been obtained by superresolving low spatial resolution high frame rate video with high spatial resolution low frame rate video to estimate the missing low resolution frames.
Z. Xu, Q. Zhang et al. [15]	2016	Two-step motion propagation method, pixels assignment optimization based on Markov random field, and finally illumination adjusting approach.	It outperforms previous methods in terms of keeping the completed background spatio-temporally coherent, completing video backgrounds with significant depth discontinuity, and maintaining consistent illumination in the completed field.	The key drawbacks of the used methodology are the computational complexity of the Markov random field method and the sensitivity of the results to model parameters.	This technique uses motion estimation technique to fill the hole, so it can be applicable in NPP inspection video
Conze, Pierre-Henriet.al [16].	2016	Optical flow estimation, multi-step integration, and statistical selection.	It reduces the complexity while enhancing the accuracy and outperforms the existing techniques.	Frame differencing accuracy is dependent on object speed and frame rate, and it requires a lot of memory.	This technique uses optical flow estimation and applicable for low frame rate videos. It can be used in NPP inspection video.
W. Hu, D. Tao et al. [17]	2017	Twist-tensor nuclear norm.	It improves the estimate of panning videos.	For stationary surveillance video, this method has weaker advantages than other methods.	This technique uses the concept of sparsity and will be applicable if we have degraded video with less information and can be used in NPP inspection video.
D. M. M. Rahaman et al. [18]	2018	Spatial and temporal correlation of intra/inter-view image, Gaussian mixture model, and adaptive weighted average of the pixels.	Compared to existing approaches the proposed approach gives a significant improvement in PSNR.	Gaussian mixture model has some limitations as Computationally expensive if the number of distributions is large and Converge to local optimal.	This technique for free view point video and multi-view video coding, it can be applicable in NPP inspection video.
B. E. Moore, C. Gao et al. [19]	2019	Principal component analysis(PCA), low-rank matrix estimator, weighted total variation framework.	In terms of foreground and background estimation, the proposed method outperforms the current state-of-the-art method.	The covariance matrix is difficult to be evaluated accurately. Even the simplest invariance could not be captured by the PCA unless the training data explicitly provide this information.	This technique can be used to estimate the missing video frames using PCA and concept of sparsity.
M. Saqib, S. D. Khan et al. [20]	2019	Motion guided filter and Deep convolution neural network.	In terms of crowd counting in low to medium density video, this method outperforms the existing technique.	The guided filter is not specifically applicable for sparse inputs like strokes.	This technique can be applicable for crowd counting and to estimate the miss detections and hence it can be used in NPP inspection video for different purposes.
A. Paliwal et al. [21]	2020	Frame interpolation, Deep learning system, context, and occlusion aware network.	It outperforms prior work on both synthetic and real videos.	If the auxiliary frames are taken at extremely low resolution, this approach will not be able to use the information in them.	This technique can be used for high-resolution video with extremely high frame rate.

A frame rate up conversion approach is utilised for higher frame rate video, followed by a trilateral filter [3], which reduces the computational cost of refining the motion vectors while also suppressing interpolation noise and missregistration errors. By selecting the right search strategy, block matching criterion, and block size, this technique can be used in NPP inspection videos.

Although some algorithms do not use the motion estimation method [4], their results are inferior to systems that do, hence these techniques should be avoided for use in NPP inspection video application.

When attempting to estimate the missing video frame in streaming video over any channel, we use a cross-layer optimization approach for the H.264/AVC scalable extension [5] that ensures low computational complexity, but there is a tradeoff between the size of the optimization window and computational complexity. This technique can be applied in NPP inspection videos by balancing the optimization window and computational cost.

Motion completion and frame completion approaches [6] are used for a video with unstable motion and luminance, and they perform better if the damaged content does not span a large area. This approach can be used to estimate missing/degraded frames in NPP inspection videos.

Multiple hypotheses baysian frame rate-up conversion approach [7] can improve the objective and subjective quality of the reconstructed frame by taking time complexity into consideration. This technique can likewise be applied to NPP's inspection video.

Because of their simple structure and minimal complexity, data adaptive image interpolation and maximum likelihood estimation techniques [8] are used, but this technique cannot be used directly for missing video frame estimation because the interpolation problem must first be converted into a denoising problem. As a result, this technique cannot be directly applied to the video of the NPP inspection.

The missing video frames are estimated using the non-local kernal regression technique [9], which produces a more reliable result. Although this technique can be employed in NPP inspection videos, the main disadvantage is that the computational cost increases owing to the patch matching procedure.

The Markov chain interpolation technique [10] is used to be competitive on different test sequences. It provides better results and can be applied to inspection video of NPP, but due to many assumptions, its computational complexity increases.

If the degraded content does not cover a large area, the frame-based recovery method [11] is applied, which recovers the most video material from non-overwritten video portions. This technique can be used to estimate missing video frames in NPP inspection video.

The missing frames can be estimated using improved depth picture based randering [12], which is based on motion estimation. However, undesirable objects might appear along depth discontinuities while randering. This approach can be employed in NPP inspection video considering this limitation.

Texture-synthesis-based video super-resolution and dynamic texture synthesis methods [13] are used to reconstruct the high-resolution details of low-resolution videos. It improves the subjective and objective visual consistency, so this method can also be used in the NPP inspection video, but if the texture has persistent motions, this method is less successful in capturing the motion patterns.

Stereo matching, adaptive overlapped block motion compensation, disparity compensation, and motion compensation techniques [14] are used in those videos which have been obtained by superresolving low spatial resolution, high frame rate video with high spatial resolution, low frame rate video to estimate the missing low-resolution frames. These techniques provide both subjective and objective qualities of a mixed spatiotemporal superresolution approach and can be used in NPP inspection videos.

To fill the hole and provide better results than conventional methods, the two-step motion propagation method, pixel assignment optimization based on Markov random field, and illumination adjusting approaches [15] are used, and will be applicable in NPP inspection video taking into account the computational complexity.

The optical flow estimation, multi-step integration, and statistical selection approaches [16] are utilised for low frame rate videos to reduce complexity while improving the accuracy of missing frame estimation, and this technique is suitable for NPP inspection videos.

The twist-tensor nuclear norm technique [17] is utilised if the NPP inspection video has sparse information about the content in the frame. This method provides a better estimate for panning video, but it provides less accurate results for stationary surveillance video.

For free view point video and multi-view video coding, the Gaussian mixture model and adaptive weighted average of the pixels [18] are employed. It improves the PSNR significantly and can be used in the NPP inspection video.

To estimate the missing video frames, techniques such as principal component analysis (PCA), low-rank matrix estimator, and weighted total variation framework [19] are utilised, and similar techniques can also be applied for NPP inspection video.

Motion guided filter and deep convolution neural network algorithms [20] are utilised for crowd counting and estimating miss detections. These techniques will work for NPP inspection videos as well.

Frame interpolation, deep learning algorithms, and occulusion-aware network approaches [21] are utilised for high-resolution videos with an extremely high frame rate and perform better for both synthetic and real videos. If the auxiliary frames are taken at a very low resolution, these methods will not work. These strategies can be employed in NPP inspection videos, taking into account their benefits and limits.

5. Findings

In this section, we discuss the significant outcomes of the different kinds of literature related to the missing video frame estimation included for the study in this paper. Mainly, in our study, we have observed the different techniques which are used to recover the missing video frames and its suitability to apply these techniques on NPP's inspection video as: The hybrid motion estimation technique gives a better estimate of motion vectors but has more computational complexity than conventional methods and a major challenge is to select the accurate threshold value, and another technique that does not use the motion estimation technique is the modified content-adaptive vertical temporal filtering has low complexity. Some deep learning techniques with frame interpolation are used on the synthetic as well as on the real video which also gives a good result for missing frame interpolation. Motion guided filter with a deep convolution neural network is also used to overcome the limitations of the pedestrian detector in terms of crowd counting in low to medium density videos. Some other technique like principal component analysis with low-rank matrix estimator performs well for foreground and background estimation. Optical flow estimation technique with, multi-step integration and statistical selection is used to reduce the complexity and improve the accuracy. Texture-synthesis-based video super-resolution and dynamic texture synthesis give better results in terms of objective and subjective visual consistency.

Frame-based recovery methods restore the fragmented video files successfully, regardless of the amount of fragmentation. Most video material in non-overwritten segments of video files can be recovered using this technique. A cross-layer optimization approach for the H.264/AVC scalable extension ensures low computational cost since it bypasses highly complex pixel-level motion compensation operations. For such videos containing unstable motion and luminance, the motion completion and frame completion methods are used. Markov random field for pixels assignment optimization is used for two-step motion propagation which is useful in terms of keeping the completed background spatio-temporally coherent, and completing video backgrounds with significant depth discontinuity, and maintaining consistent illumination in the completed field. The non-local kernel regression technique used provides more robust image estimation, and the algorithm is used to restore different images and videos.

6. Conclusion and future research

In this paper, we studied various methods for estimating the missing frames in videos. We conducted an overall study on the available techniques to recover the missing/damaged frames and its suitability to use in NPP's inspection video. Each has its own set of benefits and drawbacks that are identified. This paper gives an overview of different techniques to estimate the missing video frames and its suitability to use in NPP's inspection video, which may arise from different reasons, either it may corrupt while transmitting it over a channel during the streaming or the frames may degrade in any recorded video. The work would be very useful as a benchmark to the future researchers in this field and also to the NPP practitioners. In the future, we shall develop a more robust technique to estimate the missing video frames which can overcome the exiting limitations and has low computational complexity and minimum artifacts.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Mark R. Banham, Aggelos K. Katsaggelos, Digital image restoration, IEEE Signal Process. Mag. 14 (1997) 24–41, 2.

[2] B. Yan, H. Gharavi, A hybrid frame concealment algorithm for H.264/AVC, in: IEEE Transactions on Image Processing, vol. 19, Jan. 2010, pp. 98–107, 1.

[3] C. Wang, L. Zhang, Y. He, Y. Tan, Frame rate up-conversion using trilateral filtering, in: IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, June 2010, pp. 886–893, 6.

[4] K. Lee, C. Lee, High quality deinterlacing using content-adaptive vertical temporal filtering, in: IEEE Transactions on Consumer Electronics, vol. 56, November 2010, pp. 2469–2474, 4.

[5] E. Maani, A.K. Katsaggelos, Unequal error protection for robust streaming of scalable video over packet lossy networks, in: IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, March 2010, pp. 407–416, 3.

[6] N.C. Tang, C. Hsu, C. Su, T.K. Shih, H.M. Liao, Video inpainting on digitized vintage films via maintaining spatiotemporal continuity, in: IEEE Transactions on Multimedia, vol. 13, Aug. 2011, pp. 602–614, 4.

[7] H. Liu, R. Xiong, D. Zhao, S. Ma, W. Gao, Multiple hypotheses bayesian frame rate up-conversion by adaptive fusion of motion-compensated interpolations, in: IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, Aug. 2012, pp. 1188–1198, 8.

[8] Kwok-Wai Hung, Wan-Chi Siu, Computationally scalable adaptive image interpolation algorithm using maximum-likelihood denoising for real-time applications, J. Electron. Imag. 22 (2013), 043006, 4.

[9] H. Zhang, J. Yang, Y. Zhang, T.S. Huang, Image and video restorations via nonlocal kernel regression, in: IEEE Transactions on Cybernetics, vol. 43, June 2013, pp. 1035–1046, 3.

[10] F. Vedadi, S. Shirani, De-interlacing using nonlocal costs and markov-chain-based estimation of interpolation methods, in: IEEE Transactions on Image Processing, vol. 22, April 2013, pp. 1559–1572, 4.

[11] G. Na, K. Shim, K. Moon, S.G. Kong, E. Kim, J. Lee, Frame-based recovery of corrupted video files using video codec specifications, in: IEEE Transactions on Image Processing, vol. 23, Feb. 2014, pp. 517–526, 2.

[12] T.L. Lin, T.E. Chang, G.X. Huang, C.C. Chou, U.S. Thakur, Improved interview video error concealment on whole frame packet loss, J. Vis. Commun. Image Represent. 25 (2014) 1811–1822, 8.

[13] C. Hsu, L. Kang, C. Lin, Temporally coherent superresolution of textured video via dynamic texture synthesis, in: IEEE Transactions on Image Processing, vol. 24, March 2015, pp. 919–931, 3.

[14] Jing Ge, et al., A spatiotemporal super-resolution algorithm for a hybrid stereo video system, Signal, Imag. Video Proc. 10 (2016) 559–566, 3.

[15] Z. Xu, Q. Zhang, Z. Cao, C. Xiao, Video background completion using motion-guided pixel assignment optimization, in: IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, Aug. 2016, pp. 1393–1406, 8.

[16] Pierre-Henri Conze, et al., Multi-reference combinatorial strategy towards longer long-term dense motion estimation, Comput. Vis. Image Understand. 150 (2016) 66–80.

[17] W. Hu, D. Tao, W. Zhang, Y. Xie, Y. Yang, The twist tensor nuclear norm for video completion, in: IEEE Transactions on Neural Networks and Learning Systems, vol. 28, Dec. 2017, pp. 2961–2973, 12.

[18] D.M.M. Rahaman, M. Paul, Virtual view synthesis for free viewpoint video and multiview video compression using Gaussian mixture modelling, in: IEEE Transactions on Image Processing, vol. 27, March 2018, pp. 1190–1201, 3.

[19] B.E. Moore, C. Gao, R.R. Nadakuditi, Panoramic robust PCA for foreground–background separation on noisy, free-motion camera video, in: IEEE Transactions on Computational Imaging, vol. 5, June 2019, pp. 195–211, 2.

[20] M. Saqib, S.D. Khan, N. Sharma, M. Blumenstein, Crowd counting in low-resolution crowded scenes using region-based deep convolutional neural networks, in: IEEE Access, vol. 7, 2019, pp. 35317–35329.

[21] A. Paliwal, N.K. Kalantari, Deep slow motion video reconstruction with hybrid imaging system, in: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, 1 July 2020, pp. 1557–1569, 7.