E-ISSN 2581 – 7957
P-ISSN 2277 – 3916

CVR Journal of Science and Technology, Volume 22, June 2022
DOI: 10.32377/cvrjst2207

Video Inpainting using self-adaptive GMM with Improved Inpainting Technique

B. Janardhana Rao[1], K. Revathi[2], G. Harish Babu[3]
[1]Assoc. Professor, CVR College of Engineering/ECE Department, Hyderabad, India
Email: janarhdan.bitra@gmail.com
[2]Asst. Professor, CVR College of Engineering/IT Department, Hyderabad, India
Email: revathi.kandula@gmail.com
[3]Asst. Professor, CVR College of Engineering/ECE Department, Hyderabad, India
Email: harish.sidhu12@gmail.com

***Abstract*: Nowadays photography and videography have become part of life. There are many challenging tasks in videography, one of these is video inpainting. Repairing the damaged videos or removing and filling undesired objects in videos is defined as video inpainting. In this paper, an object-splitting video inpainting is proposed. The background subtraction is implemented by self-adaptive Gaussian Mixture Model (SAG). The moving foreground and static background are inpainted by using enhanced inpainting technique with improved patch priority calculation. The structure consistency patch matching is proposed to search for the similar patch in the source region to fill the target region. The video inpainting results are obtained for the own video. The proposed inpainting technique is implemented for state of art videos utilized in other related works. The experimental results show that the proposed technique attained impressible inpainted videos compared to related works.**

***Index Terms*: Video inpainting; SAG, background subtraction; patch priority; Structure consistency**

I. INTRODUCTION

Video inpainting plays a vital role in video processing applications such as video restoration, video stabilization, films post-processing. Video inpainting is used to reconstruct the damaged parts of old videos and to remove the unwanted objects in the video without identifying by human eye visually. Image inpainting is the process of repairing damaged images and restoring the images by removing undesired parts in the image. The traditional image inpainting techniques are Partial Differential Equation (PDE) based [1] and texture synthesis techniques [2, 3]. In PDE based inpainting techniques, the gradient vector is calculated around the unknown region and local structures adjacent to the lines perpendicular to gradient vectors are diffused to the target region. This is completely pixel-based method and failed to fill large target regions. These are well suited for images with more structure information. In texture synthesis algorithms, the texture patches in the source region diffuse to target region and neglects the structured content. Both these methods are combined and formed a new algorithm, exemplar based inpainting [4, 5]. These algorithms preserve both structure and texture information in the inpainted region.

The video inpainting is achieved with image inpainting by dividing the video into sequence of images called frames. Due to this, the temporal coherence between the frames is missing [6]. Bertalmio was proposed video inpainting using Partial Differential Equations (PDE) based image inpainting [7] by preserving the temporal coherence between the frames. But in this, the texture information is not retrieved properly and failed to fill the large portions of target region. Patwardhan [8] consider the motion vector along with image inpainting. This is achieved by taking the preprocessing in video frames by separating the moving foreground from the static background which is called background subtraction. This method has given reasonably good results but not works for videos having nearby moving objects and moving background. Chang [9] proposed the method to remove moving objects in both static and moving background. T. K. Shih [10] given an algorithm which suited for complex camera motions. N. C. tang proposed a new technique for patch searching and patch alignment to get the temporal coherence of the video. C. W. Lin [11] proposed method applicable for videos having dynamic background with a fixed camera or still background with moving camera. In these mosaics of frames are constructed from background-based camera motion. A. Koochari [12], separated the moving objects from the background and constructed the mosaics of moving objects, these are inpainted by using large patches.

The video inpainting problem is achieved in three ways [13]. (i) Patch matching method (ii) Object-background splitting (iii) Structure and texture classification.

In patch matching method, the patches searching to fill the hole in target region is achieved in the spatiotemporal domain. In [14, 15] the color components, spatial and temporal vectors are compared to attain spatiotemporal volume matching. In object-background splitting, background subtraction is used to separate static background from moving foreground. The inpainting will apply individually to both moving objects and static background [8, 16, 17] by maintaining the temporal coherence between the frames. Gaussian Mixture Model

Received on 16-03-2022, Revised on 02-04-2022, Accepted on 23-05-2022.

(GMM) [18, 19] is used for background subtraction. In structure and texture classification, the video frames are divided into structure and structure regions, and these are inpainted using different inpainting techniques by maintaining region priorities. Recently, the novel video inpainting technique is implemented with hybridization of cuckoo search algorithm and multi verse optimization [21] for optimizing the patch matching and Recurrent Neural Network (RNN) for categorizing the patch as smooth or structured. This method produced the optimal video inpainting results compared to available methods. Another video inpainting technique [22] proposed by using enhanced priority computation method and optimal patch selected for inpainting the target region with Grey Wolf Optimization (GWO). This method outperformed the existing techniques of video inpainting in terms of metrics PSNR, SSIM and Edge similarity.

In this paper, an object-background splitting method is implemented. The GMM based background subtraction has its limitations to separate moving foreground and background. The GMM method failed to handle the sudden changes in illumination and irregular background motion. Due to its parametric modelling, it is time-consuming and complex hence didn't produce proper results in real time applications [20]. Proposed a new background subtraction technique, self-adaptive GMM (SAG) has adaptive learning rate and withstand for sudden illumination changes and adaptively select the number of Gaussians to model the pixels in the image. An improved inpainting technique with enhanced priority function is used to identify the priorities of the patches on the boundary of the target region. Structure consistent patch matching method is utilized to search for the best matching patch in the source region.

This paper is organized as, section II explains the proposed video inpainting algorithm, and in section III experimental results are discussed. Section IV is conclusion.

II. PROPOSED VIDEO INPAINTING ALGORITHM

The proposed architecture for video inpainting is shown in figure 1. The input video applied to the background subtraction or foreground detection method which divides the static background and foreground in each frame. In the moving foreground, the object which is removed is identified by tracking using a Kalman filtering technique. This produces the past, present and future steps of the moving object. This is target identification. After this, the target identified is removed by manual method, freehand tool in MATLAB. Then the enhanced inpainting technique is applied to background and foreground frames separately. Finally, all the frames are combined to form the in painted video.

A. Background Subtraction Method

To overcome the disadvantages in GMM, in this paper proposed an advanced background subtraction technique called self-adaptive GMM (SAG). This model uses a dynamic adaptive learning rate to take advantage of fast illumination changes. The flowchart for self-adaptive GMM is shown in figure 2 [20]. This flowchart contains two modules, the background learning module, and a foreground extraction module.

In the background learning module, spatiotemporal filter is used to remove the noise and smoothen the image this increases the robustness of the model. The global illumination changing factor (G) is obtained by using MofQ from the smoothened image. The global illumination changing factor is obtained by equation (1).

$$g = median_{i \in I}\left(\frac{i_{p,i}}{i_{r,i}}\right) \qquad (1)$$

The G value calculated for all pixels of i in the image I between the present image i_p and reference image i_r. The background is learned using SAG, the learning factor (g) calculated between the learned background and the present image, and it keeps on track the global illumination changes. The counter value (c_k) will increase if the parameters of Gaussian model changes. Then the Gaussian is reassigned and set the counter to 1.

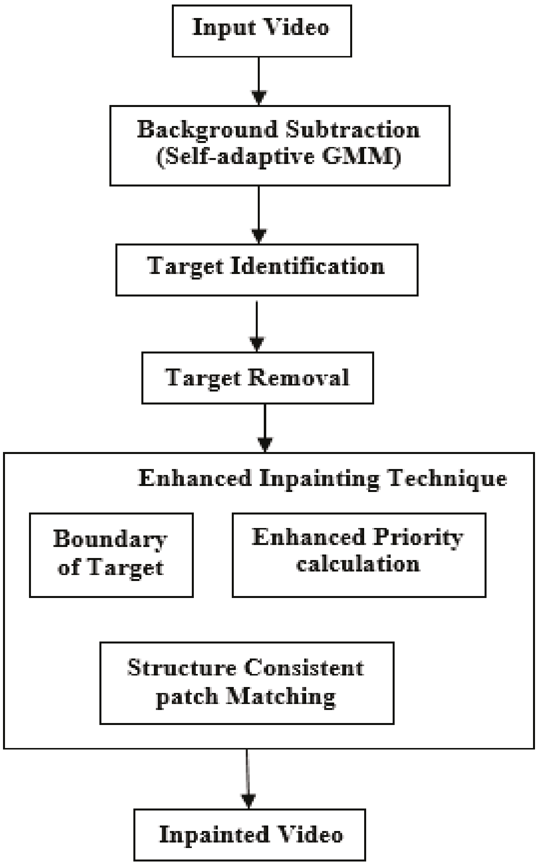


Figure 1. The architecture of proposed video inpainting.

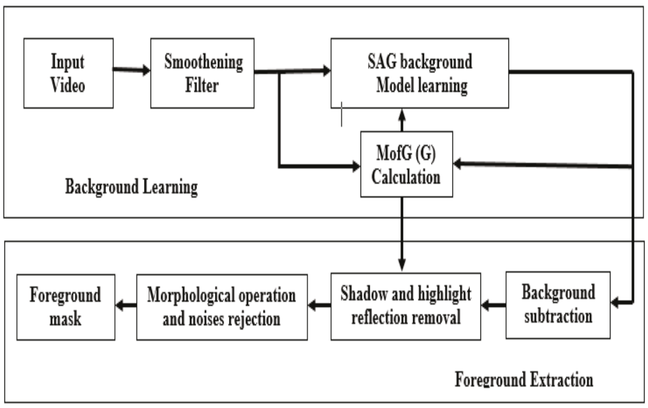


Figure 2. Flowchart of self-adaptive GMM.

When the parameters of Gaussian changed, the new learning rate β_k is calculating from equation (2).

$$\omega_{k,t} = (1-\alpha)\omega_{k,t-1} + \alpha(M_{k,t} + C_T) \quad (2)$$

$$\beta_{k,t} = \alpha(l + c_k)/c_k \quad (3)$$

$$\mu_{k,t} = \mu_{k,t-1} + M_{k,t}(\beta_{k,t}/\omega_{k,t})\delta_{k,t} \quad (4)$$

$$\sigma_{k,t}^2 = \sigma_{k,t-1}^2 + M_{k,t}(\beta_{k,t}/\omega_{k,t})(\delta_{k,t}^T\delta_{k,t} - \delta_{k,t-1}^T) \quad (5)$$

$$c_k = c_k + 1 \quad (6)$$

where l is a constant, $\alpha = 1/T$ is a constant defined to limit the effects of old data and exponentially decreased envelope. The components that are not utilized in the data are represented by negative prior weight C_T. $M_{k,t}$ is taken as 1 for matched components and 0 for other components.

The foreground extraction in self-adaptive GMM is for dividing the foreground mask from the background. It identifies the shadows and highlights spatial, spectral, and temporal features in combination of three domains and with individual features. The distortion effect of highlight and shadow are divided into brightness and chromaticity distortion. The distortion of brightness (B_i) and chromaticity (C_i) are taken as

$$B_i = \frac{g\left(I_{R_i}\mu_{R_i}/\sigma_{R_i}^2 + I_{G_i}\mu_{G_i}/\sigma_{G_i}^2 + I_{B_i}\mu_{B_i}/\sigma_{B_i}^2\right)}{\left(\mu_{R_i}/\sigma_{R_i}\right)^2 + \left(\mu_{G_i}/\sigma_{G_i}\right)^2 + \left(\mu_{B_i}/\sigma_{B_i}\right)^2} \quad (7)$$

$$C_i = \sqrt{\left(\left(gI_{R_i} - B_i\mu_{R_i}\right)/\sigma_{R_i}\right)^2 + \left(\left(gI_{G_i} - B_i\mu_{G_i}\right)/\sigma_{G_i}\right)^2 + \left(\left(gI_{B_i} - B_i\mu_{B_i}\right)/\sigma_{B_i}\right)^2} \quad (8)$$

where, $I_i = \left[I_{R_i}, I_{G_i}, I_{B_i}\right]^T$ is the intensity of the i[th] pixel in RGB space, mean $E_i = \left[\mu_{R_i}, \mu_{G_i}, \mu_{B_i}\right]^T$ and standard deviation $\sigma_i = \left[\sigma_{R_i}, \sigma_{G_i}, \sigma_{B_i}\right]^T$

From the above information the foreground pixels are classified with the following condition,

$$\begin{cases} shadow; & if\ C_i < \gamma_1\ and\ \gamma_2 < B_i < 1 \\ highlight; & if\ C_i < \gamma_1\ and\ B_i < \gamma_3 \end{cases} \quad (9)$$

where, $\gamma_1, \gamma_2, \gamma_3$ are threshold values.

After separating the background from the moving foreground, enhanced inpainting technique is applied to inpaint the frames with enhanced priority calculation and patch matching method, which is discussed in the subsequent sections.

B. Enhanced Priority Calculation

In the existing exemplar-based inpainting method [5], the inpainting results are slightly defected due to the unusual calculation of confidence term. The confidence term reduced to a low value for less number of iterations, this is called dropping effect [23]. Due to this the priorities of the patches on the boundary of the target region also decreases drastically. This gives improper result of inpainting. To avoid this an enhanced priority calculation method is introduced with a modified confidence term. In the existing exemplar-based inpainting method, the priority of the patches on the target region is calculated by product of confidence term and data term. The modified priority function [23] is taken as equation (10).

$$prority(p) = x * R_c(p) + y * D(p) \quad (10)$$

where x and y are multiplication factors of confidence term and data term respectively, $x \geq 0$, $y \leq 1$,and $x + y = 1$. The modified data term to avoid dropping effect is shown in equation (11).

$$R_c(p) = (1-\beta)\, C(p) + \beta \quad (11)$$

where β is called the regularization term ranging the values over 0.1 to 0.7.

The confidence term $C(p)$ from existing exemplar inpainting method is given by expression (12).

$$C(P) = \frac{\sum_{t \in \psi_p \cap \Phi} C(t)}{|\psi_p|} \quad (12)$$

where, t denotes the coordinates of pixel p on both source region Φ and patch on the boundary of the target region ψ_p. The data term, $D(p)$ is defined by the equation (13).

$$D(p) = \frac{|\nabla I_P^{\perp} . n_p|}{255} \quad (13)$$

where n_p is the unit vector orthogonal to the boundary of the target region at pixel p and $\nabla I_P^{\perp}$ is the isophote vector. Isophote is a line used to combine similar pixel points.

C. Structure Consistent patch Matching Method

After obtaining the highest priority patch to fill on the boundary of the target region, it is to be filled with the similar matching patch in the source region. The similar matching patch is searched by equation (14).

$$\psi_{q^1} = arg \min_{\psi_q \in \Phi} d\left(\psi_p, \psi_{q^1}\right) \quad (14)$$

where, $\psi_{q^1}$is the best matching patch in the source region.

In the traditional exemplar-based inpainting method, Sum of Squared Difference (SSD) is used for finding $d\left(\psi_p, \psi_{q^1}\right)$. The SSD method calculates only the difference between the intensity values and avoids the structural variations between the two patches. Considering the structure consistency between the source region and target region along with content obtained in SSD calculation improves the quality of the in painted image.

The structure consistency is achieved by utilizing the spatial distribution through standard deviation [25] of patch differences as given by equation (15).

$$d\left(\psi_p, \psi_{q^1}\right) = d_{SSD}\left(\psi_p, \psi_{q^1}\right) * \left\{\left[d_{STD}\left(\psi_p - \psi_{q^1}\right)\right]^{\alpha} + 1\right\} \quad (15)$$

with

$$d_{STD}\left(\psi_p - \psi_{q^1}\right) = \sqrt{\frac{1}{n}\sum_t \left[\psi_p(t)N_p(t) - \psi_{q^1}(t)N_{q^1}(t) - \left(\overline{\psi_p} - \overline{\psi_{q^1}}\right)\right]^2} \quad (16)$$

where $n = \sum_t N_p(t)$; α is a constant adjust, the weight of d_{STD} over the final quality. $\overline{\psi_p}$ and $\overline{\psi_{q^1}}$ are mean values. The value of α set to 0.8. From this, it is observed that if d_{STD} is

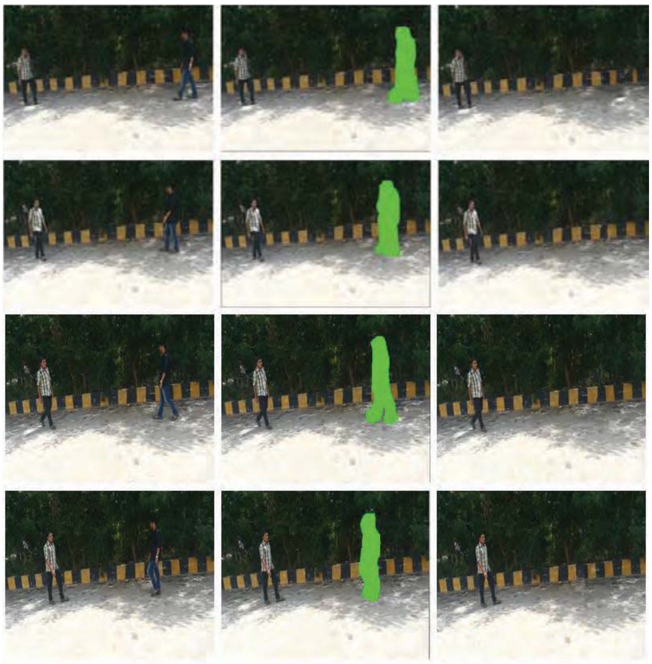
E-ISSN 2581 – 7957
P-ISSN 2277 – 3916

DOI: 10.32377/cvrjst2207

small then the two patches are similar in structure. The structure variations are consistent in both patches if d_{STD} is zero. Under different illumination conditions if the patches have a same structure, then d_{STD} becomes zero, this gives errors in results. To avoid this 1 is added in the equation. In this structure consistent patch matching method due to standard deviation, attain best matching patches considering structure distributions also.

III. Experimental Results

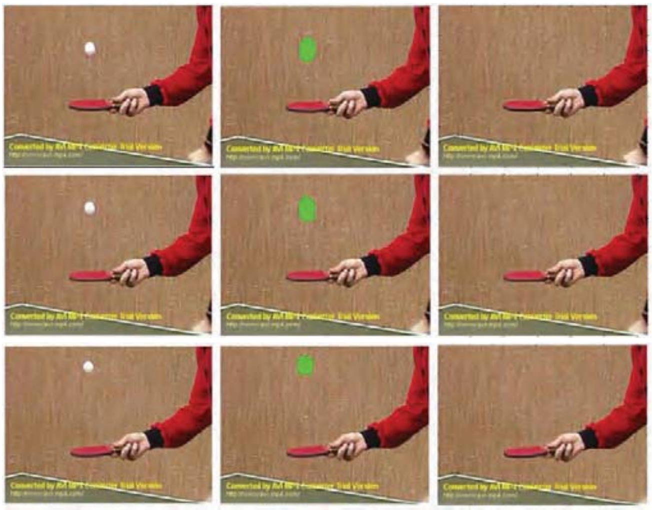
The experimental results are implemented in MATLAB. In enhanced inpainting, the priority of the patches on the boundary of the target is calculated taking the regularization term $\beta = 0.7$, the multiplication factors of confidence term and data terms are $x = 0.8$ and $y = 0.2$ respectively [24]. These values result the best inpainting. The inpainting is applied to my own video. Some of these inpainted frame results are shown in figure 3.



(a) (b) (c)

Figure 3. Own video inpainted Results: a) Input frame b) Object to be removed c) Inpainted frame.

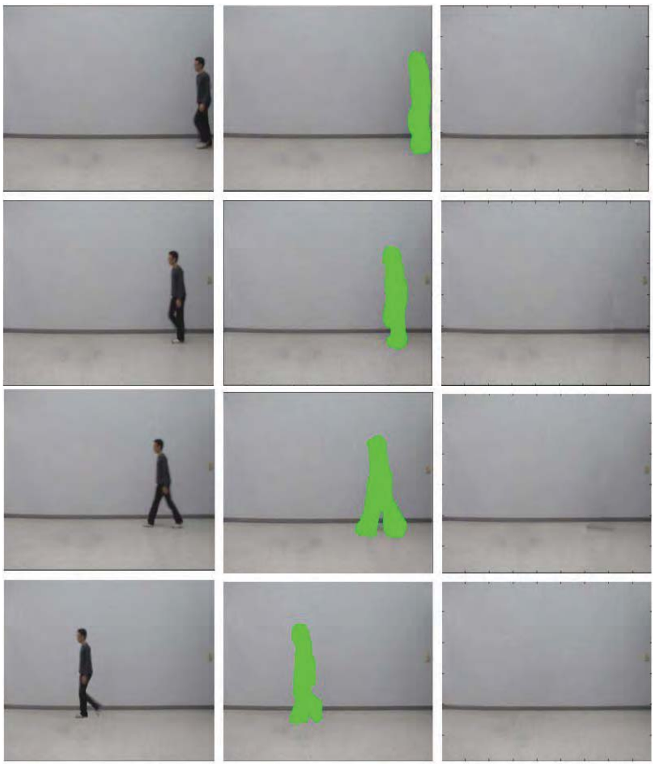
The proposed inpainting is compared to the state of art existing techniques taken their input videos [26]. Applied to the input of table tennis and produces the results as shown in figure 4.



(a) (b) (c)

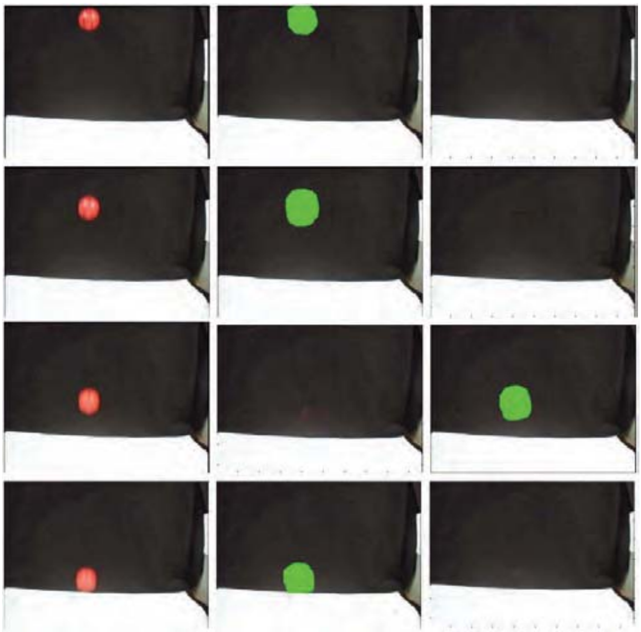
Figure 4. Table tennis inpainted results: a) Input frame b) Object to be removed c) Inpainted frame

The proposed technique is also compared, [27] for videos of walking man and throwing ball. The inpainted results for some frames are shown in figure 5 and 6.



(a) (b) (c)

Figure 5. Walking man inpainted results: a) Input frame b) Object to be removed c) Inpainted frame



(a) (b) (c)

Figure 6. Throwing ball inpainted Results: a) Input frame b) Object to be removed c) Inpainted frame

IV. Conclusions

This paper proposed a new video inpainting technique using self-adaptive GMM (SAG) background subtraction method. The background subtraction method separated the moving foreground and static background any sudden illumination changes in the video. The separated static background and moving foreground are inpainted with improved priority calculation technique and structure consistent patch matching method. The patch priorities on the boundary of the target region have been calculated with multiplication factors of confidence term and data term as $x = 0.8$ and $y = 0.2$ for regularization term $\beta = 0.7$. The experiment results are shown that the proposed method produced the results quite better compared to existing techniques.

References

[1] M. Bertalmio, G. Sapiro, V. Caselles, C. Bellester, Image inpainting, in: Proc. ACM SIGGRAPH, New Orleans, Lousiana, 2000, pp. 417-424.

[2] A .A . Efros , T.K. Leung , Texture synthesis by non-parametric sampling, in: IEEE Conf. on Computer Vision, vol. 2, Sept. 1999, pp. 1033–1038 .

[3] L. Liang , C. Liu , Y.Q. Xu , B. Guo , H.Y. Shum , Real-time texture synthesis by patch-based sampling, ACM Trans. Graph. 20 (3) (Jul. 2001) 127–150 .

[4] A. Criminisi , P. Perez , K. Toyama , Object removal by exemplar-based inpainting, in: IEEE Conf. on Computer Vision and Pattern Recognition, vol. 2, Jun. 2003, pp. 721–728 .

[5] A. Criminisi , P. Perez , K. Toyama , Region filling and object removal by exemplar-based image inpainting, IEEE Trans. Image Process. 13 (9) (Sept. 2004) 1200–1212 .

[6] M. V. Venkatesh, S. S. Cheung, J. Zhao, “Efficient Object Based Video Inpainting”, *Journal of Pattern Recognition Letters*, vol. 30, no. 2, pp. 168-179, 2009.

[7] M. Bertalmio, A. L. Bertozzi, G. Sapiro, "Navier-Strokes, Fluid Dynamics, and Image and Video Inpainting", In Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 355- 362, 2001.

[8] K. A. Patwardhan, G. Sapiro, M. Bertalmio, "Video Inpainting under Constrained Camera Motion", In Proc. IEEE Trans. Image Processing, vol. 16, no. 2, pp. 545-553, 2007.

[9] R. C. Chang, N. C. Tang, C. C. Chao, "Application of Inpainting Technology to Video Restoration", In Proc. 1st IEEE Int. Conf. Ubi-Media Computating, pp. 359-364, 2008.

[10] T. K. Shih, N. C. Tang, J. N. Hwang, "Exemplar-Based Video Inpainting Without Ghost Shadow Artifacts by Maintaining Temporal Continuity", In Proc. IEEE Trans. Circuits and Systems for Video, vol. 13, no. 3, pp. 347-360, 2009.

[11] C. W. Lin, N. C. Cheng, "Video Background Inpainting Using Dynamic Texture Synthesis", In Proc. IEEE Int. Symposium on Circuits and Systems, pp. 1559-1562, 2010.

[12] A. Koochari, M. Soryani, "Exemplar-Based Video Inpainting with Large Patches", Journal of Zhejiang University, vol. 11, no. 4, pp. 270-277, 2010.

[13] Kuo, Tien-Ying, Po-Chyi Su, and Yun-Ping Kuan. "SIFT-guided multi-resolution video inpainting with innovative scheduling mechanism and irregular patch matching." *Information Sciences* 373 (2016): 95-109.

[14] Y. Wexler , E. Shechtman , M. Irani , Space-time video completion, in: IEEE Conf. on Computer Vision and Pattern Recognition, vol. 1, Jul. 2004, pp. 120–127 .

[15] Y. Wexler , E. Shechtman , M. Irani , Space-time completion of video, IEEE Trans. Pattern Anal. Mach. Intell. 29 (3) (Mar. 2007) 463–476 .

[16] J. Jia , Y.W. Tai , T.P. Wu , C.K. Tang , Video repairing under variable illumination using cyclic motions, IEEE Trans. Pattern Anal. Mach. Intell. 28 (5) (May 2006) 832–839 .

[17] J. Jia , T.P. Wu , Y.W. Tai , C.K. Tang , Video repairing: inference of foreground and background under severe occlusion, in: IEEE Conf. on Computer Vision and Pattern Recognition, vol. 1, Jul. 2004, pp. 364–371.

[18] W.H. Press , S.A. Teukolsky , W.T. Vetterling , B.P. Flannery ,in: Numerical Recipes: The Art of Scientific Computing, Third Edition, Cambridge University Press, 2007, pp. 842–848 .

[19] C.H. Yeh , C.Y. Lin , K. Muchtar , L.W. Kang ,Real-time background modeling based on a multi-level texture description, Inf. Sci. 269 (June 2014) 106–127 .

[20] Goyal, Kalpana, and Jyoti Singhai. "Review of background subtraction methods using Gaussian mixture model for video surveillance systems." *Artificial Intelligence Review* 50.2 (2018): 241-259.

[21] B Janardhana Rao, Y Chakrapani, S Srinivas Kumar, Hybridized Cuckoo Search with Multi-Verse Optimization-Based Patch Matching and Deep Learning Concept for Enhancing Video Inpainting, *The Computer Journal*, 2021; bxab067, https://doi.org/10.1093/comjnl/bxab067.

[22] B Janardhana Rao, Y Chakrapani, S Srinivas Kumar, An Enhanced Video Inpainting Technique with Grey Wolf Optimization for Object Removal Application, Journal of Mobile Multimedia (2022), Vol. 18, Issue 3, pp. 561-582.

[23] J. Wang, K. Lu, D. Pan, N He, B. Bao, Robust object removal with an exemplar-based image inpainting approach, Neurocomputing 123 (2014) 150-155.

[24] Janardhana Rao, B., Y. Chakrapani, and S. Srinivas Kumar. "Image Inpainting Method with Improved Patch Priority and Patch Selection." *IETE Journal of Education* 59.1 (2018): 26-34.

[25] H. Wang et al., Exemplar-based image inpainting using structure consistent patch matching, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2016.08.149.

[26] Hung, Kuo-Lung, and Shih-Che Lai. "Exemplar-based video inpainting approach using temporal relationship of consecutive frames." *Awareness Science and Technology (iCAST), 2017 IEEE 8th International Conference on*. IEEE, 2017.

[27] Dusane, Meenal R., and G. S. Sable. "Video Inpainting for Moving Object Removal Using Patch Sparsity." *International journal of computer engineering & technology* (2013): 2277-9655.