Journal of Engineering and Applied Sciences 13 (6): 1382-1385, 2018
ISSN: 1816-949X

Video Inpainting Based on Background Modelling

Israa Hadi Ali and M. Mustafa Ehsan
College of Information Technology, Babylon University, Babylon, Iraq

Abstract: In this study, proposes an efficient method for video inpainting that based on statistical modified background modeling. Deferent challenges and situations can be handled in this algorithm with a good result and less computational cost. There are three major steps in this proposed methodology, in the first step the proposed system makes an estimation to the background of the video to get the background. The second step which is subtracted each frame of input video from the background model in order to get the foreground objects to be deleted. Third step the method detects the trajectory of the target object that selected by the user and shutting off its pixels. Finally, finding the appropriate pixels from the background to be assigned instead of the target object pixels.

Key words: Video inpainting, background model, remove object, video editing, video tracking, computational

INTRODUCTION

Inpainting is the process of removing a specific area from the contact of an image or video (sequence of images) replacing it with background (Panchal and Holia, 2014). Taking into consideration keeping the consistency of the environment around the removed object with pixels that filled instead of original objects pixels. This approach may use to repair damaged images or videos by deleting the scratches or noise (Panchal and Holia, 2014). Or may be used to delete an unwanted object in an image or video file. Inpainting in Video deferent from that's implemented in the image. In image inpainting the depend will be on spatial domain (Criminisi *et al*., 2004) where video inpainting the spatio-temporal domain must be used to fill holes that produce from object deletion (Bhatewara *et al*., 2013). The spatio-temporal dependency achieved by utilize the similarity between video frames (Bhatewara *et al*., 2013; Ebdelli *et al*., 2015). The method will implement in this study assumed the video is captured by a static camera and without any occlusion between objects that moving in the movie by using background model with estimate in first step after reading the video. The background estimation performed by using accumulation histogram that's achieved by four major steps. The first step divides each frame into specific equal-size blocks. The second step is compute the median to each block concentrated in the same place throw the video frames and find the frequency for each median. The third step finds the variance for each block in the blocks which have highest median frequency and find the frequency for each variance. The third step, the block with largest median frequency and largest variance frequency will be chosen as block belongs to background model in the same location.

The inpainting process four major stages after background model estimation. The first stage subtracts each frame from the background model. This step will produce the foreground objects. The second stage tracking each foreground object in order to get the trajectory for each one and save them in a database. This step will help to detect the position of objects pixels in each frame in order to be replaced by background pixels. The third stage after detecting the target object by the user, the algorithm replaces the object pixels by background pixels depending on background model and the target object trajectory that saved in above database. The final stage is performing reaggregation on video frames with keeping the sequence of the video view.

Literature review: Patwardhan *et al*. (2007) produce method for video inpainting implemented on scene with stationary camera with some moving objects. The method inpaint target object and fill its pixels with best matches background pixels by using a priority for each candidate filling pixels (Patwardhan *et al*., 2007). Zarif *et al*. (2013) present a video completion algorithm to remove static object from video scene using local similarity. The proposed algorithm depends on adjacent pixels to the target object pixels and use them to fill the pixels of deleted object.

Ebdelli *et al*. (2015) published a paper discus a new method for video inpainting with free hand camera in

Corresponding Author: Israa Hadi Ali, College of Information Technology, Babylon University, Babylon, Iraq

short-term windows (Bhatewara *et al*., 2013; Yang *et al*., 2016) present a technique of caption removal which is based on the color edge detector "Caption detection and removal from video images with complicated background using intelligent inpainting scheme". Mosleh *et al*. (2013) present an automated framework work in two stage for detect and remove the text that embedded in video and fill-in their remaining regions by appropriate data using unsupervised clustering and stroke width transform.

MATERIALS AND METHODS

The main stages of the proposed method: Figure 1 shows work stages of the proposed method.

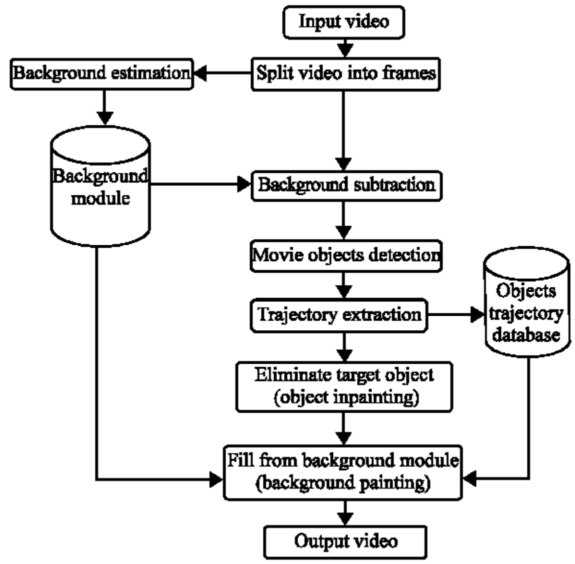

Fig. 1: Block diagram of the proposed method

Input video: In this stage, the video will be input to the proposed method in order to be split into frames those are will use in the next stage.

Background estimation: In this stage, the proposed method will find the ideal background model using accumulation histogram. This achieved by four steps. The first step tacking each frame that produced from the previous stage and divides it into k blocks each block N*M. The second step takes each block in a frame and find the mean for it by apply Eq. 1 and count the frequency for each mean founded for the blocks that occupy the same place. Thirdly take the blocks that have maximum mean frequency and compute variance for each block by using Eq. 2, then count frequency for each variance. Finally, detect the block that belongs to the background by chose it depending two factors maximum mean frequency then maximum variance frequency. As shown in Fig. 1. At the last of these three stages we will get a compilation of highest frequency blocks which represent the background model blocks (Fig. 1):

$$\text{Mean}(b) = \sum_{k=1}^{n*m} \frac{b(k)}{n*m} \quad (1)$$

Where:

b = The block

(k) = The sequence of items on it

$$\text{Variance (b)} = \sum_{k=1}^{n*m} \frac{b((k)\text{-mean}(b))^2}{(n*m)\text{-}1} \quad (2)$$

Figure 2 takes random example blocks as a first block in four successive frames as shown in Fig. 2, the proposed method find the mean for each block and compute the variance frequency for largest frequency mean blocks chose one of the blocks which have highest frequency in the values of variance. The chosen block mark as the first block from background model. Figure 3 shows the obtained background model from a video.

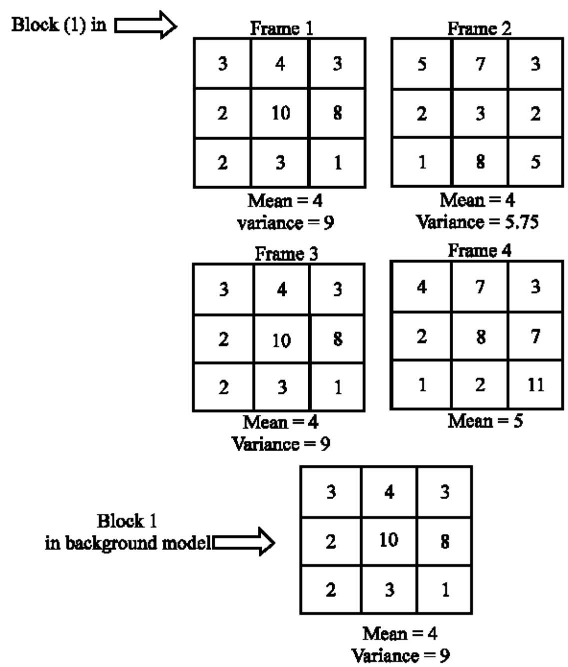

Fig. 2: Accumulation histogram example

Background subtraction: This stage is important to get the foreground objects by subtracting each frame from the background model using Eq. 3 that produced from stage process b:

$$|F_t(y)\text{-}B(y)| > T$$

where, (Ft) represent a specific frame(F) over time (t) for a single pixel (y) in other hand B(y) define the corresponding pixel (y) in the background model B. finally (T) represents a minimum threshold that depended to recognize the deferent between the frame and background as a foreground object. Figure 4 show the result of the subtraction operation of a specific frame from background model.

Movie objects detection: The result from above subtraction is the pixels of moving object. The difference between the background model and each frame represent the foreground object pixels. Figure 4 shows the detecting of two objects result of the previous stage.

Trajectory extraction: Each object trajectory should be saved in a database. This task achieved by detecting the position of each pixel captured by the object centroid from the first frame contains the object to the last frame the object appeared in it. These detected positions saved in the database with a unique label for each object. This database will be invested to detect the object trajectory where the object trajectory will help the method to detect the position of all foreground objects in all frames. It's a useful mean to know which place should be in painted in each frame by detecting the target object trajectory.

Eliminate target object: All previous stages performed in order to prepper all the foreground objects to be deleted by the user. In this stage, the user will select a single target object among the foreground object to be deleted. This selection achieved either by drawing a rectangle around the target object or just click on any pixel inside it by the mouse. The algorithm will detect the trajectory of the target object based on the trajectory database. That will help the algorithm to decide which pixel should be removed and replaced by the corresponding pixel in the background model.

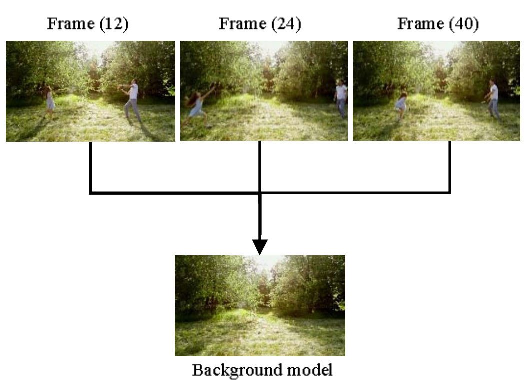

Fig. 3: Background estimation

Fig. 4: Result of the subtraction operation

Background painting: This stage represents the main challenge of how the method can cover deleting a foreground target object from video. The deleting of target object should be not visible to the video viewer. This coverage achieved by filling the background pixels instead of target object pixels. Depending on background model Which obtained in the second stage and the database trajectory Which resulted from the fifth stage. The method replaces each pixel of target object in each frame by corresponding pixel in the background model.

RESULTS AND DISCUSSION

By applying the proposed method on video example (girl play with father). In order to delete a single object from it. By marking the girl as a target object to be deleted we get the results that shown in Fig. 5-9. Figure 7 shows the obtained background model from original video by using accumulation histogram.

Fig. 5: Original video (girl play with father)

Fig. 6: Random frames from the video; a) Frame (12); b) Frame (24) and c) Frame (40)

Fig. 7: Background model

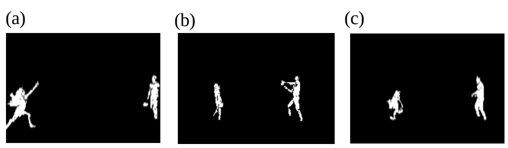

Fig. 8: Binary frames for background subtraction: a) Frame (12); b) Frame (24) and c) Frame (40)

Fig. 9: The result video after inpainting the girl: a) Frame (12); b) Frame (24) and c) Frame (40)

CONCLUSION

This research suggests a new method for video inpainting by using background modeling based on accumulation histogram. The new method applied efficiently on static camera and non-overlapped moving foreground objects. The proposed method results in painted video with no corrupted background while maintaining a coordinated video view. All these results obtained on little processing time and less computational cost from previous methods. The proposed method can produce a good support for editing video application in addition to the ability to apply it productively in the film production.

REFERENCES

Bhatewara, N., P. Kumar and A. Agrawal, 2013. Intelligent video inpainting system for texture reconstruction. Proceedings of the 2013 4th International Conference on Computer and Communication Technology (ICCCT'13), September 20-22, 2013, IEEE, Allahabad, India, ISBN:978-1-4799-1569-9, pp: 60-65.

Criminisi, A., P. Perez and K. Toyama, 2004. Region filling and object removal by exemplar-based image inpainting. IEEE. Trans. Image Process., 13: 1200-1212.

Ebdelli, M., L.O. Meur and C. Guillemot, 2015. Video inpainting with short-term windows: Application to object removal and error concealment. IEEE. Trans. Image Process., 24: 3034-3047.

Mosleh, A., N. Bouguila and A.B. Hamza, 2013. Automatic inpainting scheme for video text detection and removal. IEEE. Trans. Image Process., 22: 4460-4472.

Panchal, K.A. and M. Holia, 2014. A survey: Different techniques of video inpainting. Intl. J. Innovative Res. Comput. Commun. Eng., 2: 2923-2928.

Patwardhan, K.A., G. Sapiro and M. Bertalmio, 2007. Video inpainting under constrained camera motion. IEEE. Trans. Image Process., 16: 545-553.

Yang, S., J. Xue and Y. Zong, 2016. Caption detection and removal from video images with complicated background using intelligent inpainting scheme. Proceedings of the 2016 12th World Congress on Intelligent Control and Automation (WCICA'16), June 12-15, 2016, IEEE, Guilin, China, ISBN:978-1-4673-8415-5, pp: 611-616.

Zarif, S., I. Faye and D. Rohaya, 2013. Static object removal from video scene using local similarity. Proceedings of the 2013 IEEE 9th International Colloquium on Signal Processing and its Applications (CSPA'13), March 8-10, 2013, IEEE, Kuala Lumpur, Malaysia, ISBN:978-1-4673-5608-4, pp: 54-57.