Acta Technica 62 (2017), No. 4A, 667–672 © 2017 Institute of Thermomechanics CAS, v.v.i.

Real-time timestamp removal system for surveillance video

Jun Cheng[2], Tianyin Liu[2], Chuiwei Lu[2]

Abstract. This paper presents a real-time system to locate and remove overlaid video timestamps for surveillance video. The timestamp removal for surveillance video is a challenge research problem because it allows a very short processing time and it requires to recovery the covered objects. And the existing general object localization methods cannot localize the timestamp reliably and the existing object removal methods cannot remove timestamps with high visual scene consistence. For the timestamp localization, this paper customizes the existing method of localizing digital video clocks. For timestamp removal this paper proposes a fast and novel procedure. Its fast speed is because it employs a fast image registration method mainly by fusing the PTZ camera motion estimation and global histogram-based image matching. Then it recoveries the pixels covered by timestamps by using a temporal-neighbor pixel recovery method. The extensive experimental results show that the proposed system can remove a timestamp within 2 seconds and the recovered frames are visually consistent with the real scene.

Key words. Surveillance video, timestamp removal, image registration, image blending.

1. Introduction

Timestamp removal is an active research problem because it is a necessary function in several kinds of applications such as panorama video surveillance and the repurpose of the sports video and the home video. In panorama video surveillance, each individual input video may have a superimposed timestamp; the merged panorama video will appear multiple skew timestamps at the same time if the timestamps of individual input videos are not removed before merging. In addition, these timestamps may cover some objects and it will harm visual effects without timestamp recovery. Figure 1 shows that the three skew timestamps appear in a merged panorama frame because timestamps of three input frames are not removed.

[1]Acknowledgment - This work was supported in part by the Science Research Project of Hubei Provincial Department of Education (Q20164504) ,the Project of Hubei Province Education Science(2016GB096), Hubei Polytechnic University Scientific Innovative Team (13xtz06) and Hubei Polytechnic University Scientific Research Projects (12xjz29Q).

[2]Workshop 1 - Computer School, Hubei Polytechnic University, Huangshi435000, Hubei, China; e-mail: `juncheng@mails.ccnu.edu.com`

The timestamp removal can be divided into two main tasks: timestamp localization and timestamp removal. Timestamp is a static region but the usual static region detection methods cannot localize timestamps in the video surveillance scenario because the scene objects in video surveillance appear as static regions too.

Fig. 1. A merged panorama frame without removing the timestamps of three input frames

Timestamp removal can be done by using the video inpainting, but this approach cannot achieve a real time performance. In the recent years, various video inpainting approaches have been proposed. Wexler et al in [1] extends the image inpainting approach by belief propagation along temporal dimension to build a spatio-temporal 3D graph model. Instead of optimizing the 3D graph directly, Liu et al in [2] completes the motion field before inpainting the regions. However, these approaches via global optimization are often time consuming. Moreover, most of objects and background in the surveillance videos are static. Hence, these approaches cannot recover the covered static object with high visual scene consistence. Image registration has to done before timestamp removal and it is a fundamental research problem in image processing. Multiple algorithms were proposed in image registration [3-4]. And it was indicated that the SURF-based features perform best among feature descriptors in image registration [5]. A hybrid scheme of fusing area-based and feature-based methods was developed in [6]. A two-stage alignment, which first estimates by a histogram-based matching and then fine-tunes with a linear optimization, was proposed for motion estimation in [7]. The algorithm in [7] is faster, but it still cannot achieve a real-time speed.

PTZ cameras purposely and alternatively take two preset settings to reveal the covered area by timestamps. Thus, the image registration can take this knowledge to form a *fast image registration* method, which mainly fuses the PTZ camera motion estimation and global histogram-based image matching. After the image registration, pixel replacement is done by using a *temporal-neighbor pixel recovery* method. Thanks to the image registration procedure is very fast and accurate the whole timestamp removal can be done in real time and with high quality of visual consistence.

2. Overview of the Proposed System

Figure 2 shows a flow chart of the proposed system. The assumptions of this procedure are that the cameras only do slight tilt motion without changing their intrinsic parameters. The input of the system is a video with timestamps from a camera and the output is a removed timestamp video. The objective of the system is to remove the timestamp from the surveillance video as fast as possible. The proposed procedure includes two main stages: timestamp localization, timestamp removal. Furthermore, timestamp removal consists of two steps: image registration and timestamp recovery.

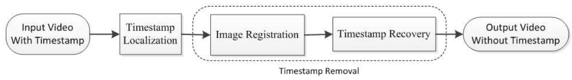

Fig. 2. The flow chart of the timestamp removal

3. Timestamp Localization

S-digit localization: S-digit localization is the first critical step of the proposed algorithm, which finds the bounding box of s-digit. On the pixel value change of s-digit a piece of knowledge is acquired, called pixel secondly-periodicity: a certain number of pixels in s-digit region will change their values when s-digit changes its digit. A set of functions is prepared to capture this knowledge and hence the s-digit localization problem is converted into computing these functions. Two steps are used to localize s-digit place. The first step is to find the s-digit transit frames. Then the second step obtains a set of s-digit pixel candidates; and the main cluster of this set is identified as the approximate region of s-digit. The precise s-digit bounding box is found by a further image analysis in a small area.

Digit color acquisition: Due to s-digit bounding box by s-digit localization method has been known, digit instances of the given video from s-digit can be collected. We also know the rough color of s-digit during s-digit localizing. Then the digit color through analyzing this collection can be found. All the digits of timestamp can be extracted using the acquired digit color and then the bounding boxes of timestamp are obtained.

4. Timestamp Removal

This section develops a procedure to remove the timestamp. The procedure can be divided into two stages: initialization and real-time processing. During the initialization phase, the proposed fast image registration algorithm in this paper first captures two frames in the two preset settings respectively within a short time gap, which are defined as the considering frame and the neighbor frame, and obtains the estimating translation according to the PTZ moving degrees. Then, the final translation can be obtained using histogram-based matching and Least Square method between the considering frame and the neighbor frame. In real-time processing stage, according to the final translation, the procedure can locate the area which covered by timestamps can be seen in the neighbor frame, and the area covered by timestamps is recovered by using a temporal-neighbor pixel recovery method.

4.1. Image Registration

Here presents the proposed fast image registration algorithm. It is based on PTZ motion estimation and global histogram-based image matching with reference to [7]. After the PTZ motion estimation is done, the hierarchical Gaussian pyramid procedure in [7] can be flattened. Hence the proposed procedure can be faster and have the same exactness of registration with the registration procedure presented in [7].

A. PTZ Motion Estimation

First, $\theta_x$and$\theta_y$are the angles of view of the camera lens in the horizontal and vertical directions, respectively. $l \times w$ is the physical size of CCD or CMOS sensor for the camera. Let$f$be the focus of the camera. According to the pinhole camera model, we obtain$\theta_x$and$\theta_y$by equation (1). $S$is the frame captured by PTZ camera. $M \times N$is resolution of $S.k_x$and $k_y$are the number of pixels moving as PTZ rotating 1 degree in the horizontal and vertical directions, respectively. $k_x$can be obtained as dividing$M$byθ_x,and $k_y$can be obtained as dividing Nbyθ_y. Next we begin to capture the two video frames in a sequence. We use the PTZ to move the camera to the preset setting 1 and captured the frameI_1, and get the PTZ degrees θ_{1x}andθ_{1y}in the horizontal and vertical directions. Then, the camera is moved to the preset setting 2, and the frame$I_2$is captured and the PTZ degrees of the horizontal and vertical directions$\theta_{2x}$and$\theta_{2y}$are obtained. At last, we can use the equation (2) to calculate the estimating translation p_x and p_y betweenI_1 and I_2.

$$\left\{ \begin{array}{l} \theta_x = 2arctg(l/2f) \\ \theta_y = 2arctg(w/2f) \end{array} \right. \tag{1}$$

$$\left\{ \begin{array}{l} p_x = k_x \left| \theta_{1x} - \theta_{2x} \right| \\ p_y = k_y \left| \theta_{1y} - \theta_{2y} \right| \end{array} \right. \tag{2}$$

B. Translation Improvement

The translation between I_1 and I_2 is first improved by using histogram-based matching method. This improving task is done as follows.I_1 and I_2 andare pre-processed with gray scaling, smoothing and gray stretching. The overlapping area between I_1andI_2 and is estimated according to the acquired translations p_x and p_y. Let $M^{'}$ and $N^{'}$ be width and height of the overlapping area, respectively.

$$\left\{ \begin{array}{l} M^{'} = M - p_x \\ N^{'} = N - p_y \end{array} \right. \tag{3}$$

Then the histogram in the direction for each row of the overlapping area is calculated as:

$$x_{ijk}\left(R\right) = \left\{ \begin{array}{l} 1, e_{ijk} \in R \\ 0, \quad \text{else} \end{array} \right. \tag{4}$$

$$R = (x_s, \cdots, x_i, k, x_j, \cdots, x_e) \in D \tag{5}$$

$$y_i(R) = \begin{cases} 1, x_i \in R \\ 0, \quad \text{else} \end{cases} \tag{6}$$

A fine-tuning procedure is used to improve the accuracy of the estimated translations further. For the captured frames I_1andI_2, the translation errors in x and $y$directions are minimized separately by using Least Squares method. The time cost in the optimization is fast due to linear regression, and the solution can be achieved to sub-pixel accuracy.

4.2. Timestamp Recovery

This section presents a temporal-neighbor pixel recovery method that recovers the portion covered by timestamps. It first localizes the ROI, which is the region in the neighbor frame corresponding to the region covered by the timestamp. Then it uses the localized region to replace the timestamp and does a blending to improve the visual consistence. Figure 3 gives an illustration of our method for timestamp removal.The replaced region may have a color difference with respect to the original whole frame. We keep the color values on the seam as they were in the and modify colors of pixels in the replaced region so that no visible boundary remains.

5. Experimental Results

In this section, the proposed method is evaluated with a variety of surveillance frames. We first experiment on accuracy and computing time of the proposed method, and compared with the SURF method in [5]. Then the performances of the proposed method are compared with the Criminisi method in [10].

5.1. Experiments on Image Registration

This paper adopts the method presented in section3.1 to compute the translation parameters for image registration. So, we can extract the translation parameters from the homography matrix. Thus this method can accurately remove the timestamp in a very low cost of computing. We conduct separately experiments one hundred times to evaluate the translation parameters of the two adjacent frames and computing time of timestamp removal. The results are given in Table 1 and Table 2 separately. In Table 1 and Table 2 μ and σ are the means and the standard deviations; and are the translation parameters from the homography matrix by the proposed method;andare the real translation parameters; and are the absolute value of the difference between the translation parameters by the proposed method and the real displacement parameters. From Table 1, we can conclude that our method can achieve a low deviation no more than 1.3 pix for the image recovery by timestamp coverd for 704*576 and 1280*720 videos. From Table 2, we can conclude that our method can achieve a very low computing cost for timestamp removal for 704*576 videos in 291 milliseconds and 1280*720 videos in 504 milliseconds.

5.2. Experiments on the Timestamp Recovery

The proposed temporal-neighbor pixel recovery method is compared with the Criminisi's method in [7]. The Criminisi's method in [7] is proposed for classical algorithms for image inpainting which can produce fairly good results in still images. The results are presented in Figure 4. It is seen that Criminisi's method leads to some misunderstandings of the timestamp area, while our results are comparatively rational in image context. The performance using our method is better than that presented in [7] due to replacing the timestamp with a real image rather than visually plausible backgrounds or the approximate pixels.

6. CONCLUTIONS AND FUTURE WORK

This paper has presented a timestamp removal system for panorama video surveillance. This system contributes two techniques. The first is a fast image registration procedure and it achieves a real time speed because it uses the PTZ camera motion estimation to obtain the rough image registration. The second is a temporal-neighbor pixel recovery method, which is based on the assumption that a pixel has a very low possibility to change within a small gap of time. The experimental results demonstrate that our approach can accurately remove timestamp in real-time. The performance of timestamp removal is authentic and consistent.

References

[1] Yonatan Wexlerr, Eli Shechtman, Michal Irani: *Space-time completion of video.* IEEE Transactions on Pattern Analysis and Machine Intelligence *29* (2007), No. 3, 463–476.

[2] B. Zitova, J. Flusser: *Image registration methods: a survey.*Image and vision computing *21* (2003), 977–1000.

[3] M. Gevrekci, B. K. Gunturk: *On geometric and photometric registration of images.* IEEE Int'l Conference on Acoustics, Speech, and Signal Processing *1* (2007), 1261–1264.

[4] H. Bay, T. Tuytelaars,L. V. Gool: *SURF: Speeded up Robust Features* European Conference on Computer Vision (2006), 404–417.

[5] M. Tico, K. Pulli: *Robust image registration for multi-frame mobile applications.* Asilomar Conference on Signals, Systems and Computers (2010), 860–864.

[6] X. Yu, W. Ding, Z. Zeng: *Reading Digital Video Clocks.* International Journal of Pattern Recognition and Artificial Intelligence *29* (2014), No. 4.

[7] Z. Farbman, G. Hoffer, Y. Lipman, D. Cohen-Or, D. Lischinski: *Coordinates for instant image cloning.* ACM Trans. Graph. *28* (2009), No. 3, 1–9.

Received November 16, 2016