

TimeViewer, a Tool for Visualizing the Problems of the Background Subtraction

Alejandro Sánchez Rodríguez¹, Juan Carlos González Castolo¹, and
Óscar Déniz Suárez²

¹ Centro Universitario de Ciencias Económico Administrativas, Universidad de
Guadalajara, México

² Grupo VISILAB, Universidad de Castilla La-Mancha, España

Abstract. This paper is about the TimeViewer tool that facilitates understanding of the most common problems in Background Subtraction. The tool displays patterns of each frame, and through the historical values of the pixels allows for visual identification of changes in a sequence of pixels. The paper demonstrates the usefulness of TimeViewer by showing how it visually presents the most common Background Subtraction problems.

Keywords: TimeViewer, Background Subtraction, Change Detection, Motion Detection.

Introduction

A human being generates most of his knowledge through his eyes. It is something so natural and instinctive that, at the beginning of Computer Vision, it was underestimated and considered a trivial operation. Computer vision consists of extracting information from a huge amount of data.

Background Subtraction (BS) consists of minimizing this amount of information, by distinguishing between foreground movement and static background. BS eliminates background data which is not significant in the scene where a movement is of interest. The extracted data is thus prepared and then passed to some other visual process. BS thus helps accelerate the treatment of the relevant information, as well as assign this information the necessary resources in an optimal manner.

Movement is represented by a series of static images which, when played one after another, generate the *illusion* of movement. Hence, the amount of data flow to be analyzed is considerable and lacks the initial significance. BS aims to find patterns of movement from series of frames by determining which pixels should and which should not be considered for further analysis.

This paper considers the most common problems in the Background Subtraction and shows how the *TimeViewer* tool helps in visualizing the patterns unique to a scene corresponding to each problem. The tool uses the historic values of the pixels, from the past frames, and allows to visually differentiate what is happening with

a pixel sequence. After all, the best way to identify possible patterns is by directly observing the images produced by the tool, according to [1].

TimeViewer can also be used as a teaching tool. Researchers can visualize and more easily comprehend a particular BS problem by observing the patterns generated by changes in pixel values through time. Thus, it is easier to analyze the problem represented by a particular scenario.

Section 1 of this paper, provides an overview of the BS. In the section 2 the architecture, as well as performance, of the *TimeViewer* tool is presented. Section 3 presents examples of the implementation of the tool in different problems in BS. Finally, the section 4, provides a brief discussion of the utility of the tool.

1 Background Subtraction Problems

Background Subtraction (BS) is a method for simplification or reduction of data [2]. It is used on videos where the flow of information is abundant and repetitive. BS processes a sequence of images in order to leave a binary mask that represents the objects of interest or *Foreground* as active, and the rest of the scene considered as *Background* is discarded [3]. This is accomplished by comparing the values of color and brightness of each pixel, from image to image.

BS starts out from the fact that there is some *movement* in the scene. It is important to emphasize this supposition, because the humans do not need movement for identification of foreground. Even in a unique image humans identify what is the background and which is the objects of interest. In the field of computational sciences, this is called **Object Recognition**. [4].

BS starts out from the illusion of movement created by a sequence of images with differences between them [5]. Because of this, the method is only applicable to image sequences or video.

According to [6], BS is a primary process in tasks such as **Surveillance, Traffic Monitoring, Tracking and Detection People** and **Gesture Recognition**, because it reduces the amount of information for further treatments.

The *Background* is composed of everything that is uninteresting, like static elements of the scene (walls, floor, furniture, etc) or objects with cyclic or insignificant movements (trees, flags water fountains, ventilators, etc). BS has two main stages: the *initialization* and the *maintenance* of the background model [7].

The **Background** concept used in this paper is taken from [8] and it is defined as: **Everything that remains static or cyclic in a video or image sequence during a predetermined amount of time.**

According to [2], there are four possible types of movement in an image sequence or video:

1. Fixed camera, moving object and **static background**.
2. Fixed camera, several moving objects and **static background**.
3. Mobile camera, relatively still scene.
4. Mobile camera, scene with several moving objects.

This study analyzes the cases one and two, where the background remains *constant* and the camera is fixed.

Currently, the evaluation of the different BS algorithms is performed by using *Datasets* manufactured for this purpose. These datasets consist of three elements: a) a group of training images b) images for the algorithm's execution and c) images used to check the results. See [7], [9], [10], [6].

In BS many unique and varied situations appear. This makes it necessary to test each BS algorithm in the various scenarios for which it is intended. In some of the scenarios errors in the background models may appear. According to [7] the main difficulties are these:

- Moved Objects. Objects that are considered part of the Background model and for some reason they change their location in the scene. This movement integrates such objects to the foreground, while they should still be considered part of the background.
- Time of Day. The global illumination of the scene changes gradually throughout the day in open environments, which causes different background models having to be considered.
- Light Switch. Global and sudden changes of illumination provoke that the Background model has to be discarded and a new one applied. Unlike the previous problem, the global sudden change of all Background can be regarded as a trigger for the change of Background model.
- Waving Trees. This refers to periodic and sporadic movements that are brought about by objects that should belong to the Background. The common example is the waving of the leaves in the trees (hence the name). Such movement should be ignored though it is frequently detected as Foreground.
- Camouflage. It is a problem that occurs when the pixels pertaining to Foreground objects have similar values to the pixels belonging to the Background model. This causes that even when there is a movement in the Foreground, changes are not perceived, because they are considered as part of the Background model.
- Bootstrapping. Some scenes present situations where it is not possible to obtain a full Background model, because of the constant appearance or movement of objects.
- Foreground Aperture. When an object of a specific homogeneous color in the foreground spans a large part of the vision field and stays there long enough, even though it is moving, some parts of its interior will be considered as part of the background model due to its movement being undetectable.
- Sleeping Person. This problem occurs when an object is initially identified as part of the Foreground and goes into a state of repose for a period of time long enough for the algorithm to include it in the Background model.
- Walking Person. When an object initially considered part of the Background moves, the region becomes Foreground. Here, the problem lies in that the newly discovered region of the Background should be considered as part of the Background immediately and not considered as another Foreground object.

- Shadows. Refers to the change in tonality or brightness of Background regions (shadows) that cause a changes in pixel values, but should not be considered as part of the Foreground.

The authors in [11] propose two more scenarios:

- Local Light Switch. Situation which arises when only some regions of Background have a sudden change of illumination and should continue to belong to the Background model.
- Reflections. If an object and its reflection are assumed as part of the Foreground, the size and shape of the object in question are different than the expected one. For example, a person and his reflection come to be regarded as a single object, which may hinder its location due to oversize.

2 Architecture

In Background Subtraction, one of the most efficient ways of checking the results of an algorithm, is **visually**, just as recommended by the authors in [1]. Even the images utilized for checking the results in [7] y [9], were manually segmented. In [1], the authors mention that *quantitative* forms of evaluation, are inferior and cannot be compared to the *qualitative* ones that are done visually by the investigators themselves.

Certainly, humans naturally and unconsciously keep the history of what we see. For the analysis of the behavior of patterns it was necessary to develop the TimeViewer tool, that allows observation of patterns of behavior of the pixels through time. The tool allows to see (not only register) the values of the pixels, and that way bring more *natural* view that allows understanding of what is *seen*. The algorithm is inspired by the work of [12] and [5].

2.1 Features

The *TimeViewer* tool was designed and coded to allow a view of the history of changes in a sequence of images, under the next features:

1. Read the images from a video or directory
2. Transform the data to make pixel history observable
3. Visualize the original sequence and the results of the data transformation
4. Allow the user to stop the sequence at any time
5. Restart the reproduction after being stopped
6. Save the images shown in the windows at any time
7. The lecture, transformation process and visualization must be done in real time, without delays
8. Optional: Obtain the images from a video-camera, live
9. Optional: Save the sequence of results images either one by one or in video format

The TimeViewer architecture is divided into three layers as shown in Figure 1:

- User Interface. Allows the user to stop viewing, display resume and save images from the screen to the working directory.
- TimeViewer Functions. Here the functions of image reading, image accumulation, transforming matrices of accumulated data and displaying the processed information for qualitative analysis are performed.
- TimeViewer Interface. The reading of video or images is done directly from a video camera or a from directory. Also, the resulting images can be stored in the working directory.

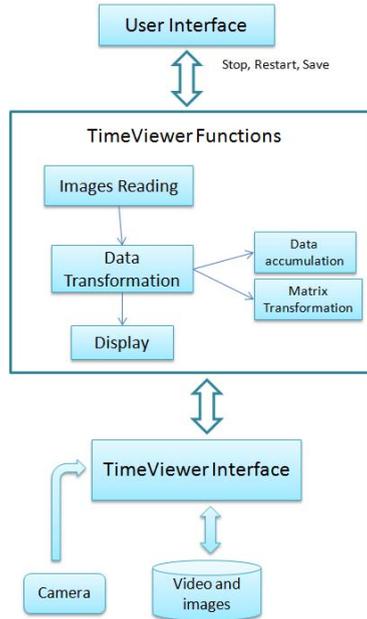


Fig. 1. TimeViewer Architecture

2.2 Implementation

TimeViewer was developed with the OpenCV library, which contains functions that allow the reading of image archives and videos. This library temporarily stores the images or frames in dynamic arrangements. Those arrangements can be manipulated by performing operations directly in the video RAM memory. Besides, the library possesses functions necessary for visualization and interactive control, that allow manipulation and the reproduction in real time. It also features storage functions for the images and video in different formats.

The main functions of TimeViewer are:

Image reading: The reading of the images was done in RAW uncompressed format.

Data Transformation: The transformation of data is done in two steps:

1. Reading and accumulation.
2. Matrix Transformation, as it can be seen in the Figure 2, and according to the equation:

$$Origin[x, y, z] = Destination[x, z, y]$$

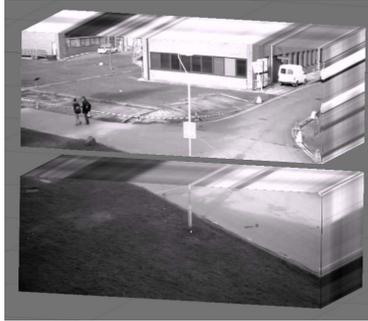


Fig. 2. Image accumulation and transformation

Visualize: The algorithm presents a screen with the original images and the image of the resulting transformation in another one.

Currently, the tool has diverse variations depending on the type of history that is desired. The first variation works with videos in grey scale, generating a window of pixel history at the time. Other variations split images in channels (depending in the color model) generating a reproduction and pixel history window for each channel.

3 Results

The *TimeViewer* tool allows the user to play a video and, at the same time, check the pixel history of certain line of the original image, to stop the reproduction and to save images. It works in real time and it can show the pixel history of any line of the original image.

3.1 Moved Object

Objects that are considered part of the background and that for some reason change their location in the scene, are integrated into the foreground, while they should be kept as part of the background. Figure 3(a) shows initial pattern, in Figure 3(b), person moved an object, then Figure 3(c) shows the final pattern.

Scene in Figure 4(a) shows the initial position of objects, then in Figure 4(b) a person moved the chair, and left it in different location, Figure 4(c).

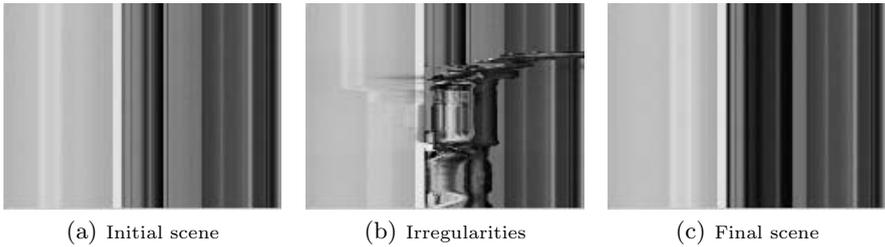


Fig. 3. Moved Object, historic

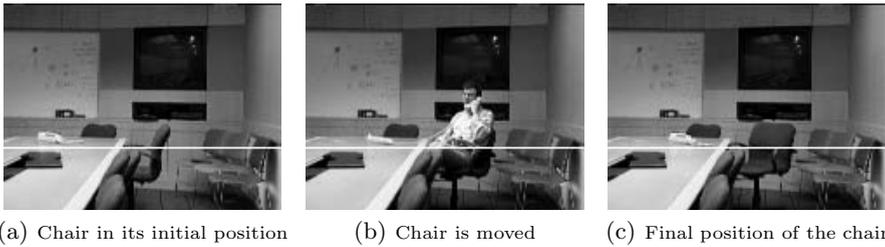


Fig. 4. Moved Object, scene

3.2 Time of Day

The global lightning of the scene changes gradually during the day in outdoor environments, which causes the necessity to consider different background models.

Patterns showed in Figures 5(a),5(b) and 5(c) belong to the same scene, only the illumination changes. Figures 6(a), 6(b) and 6(c) show the scene.

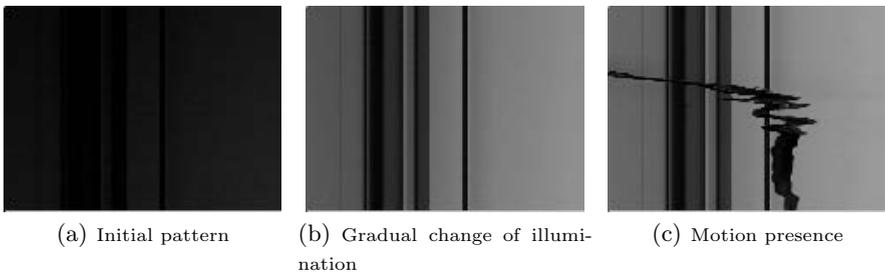
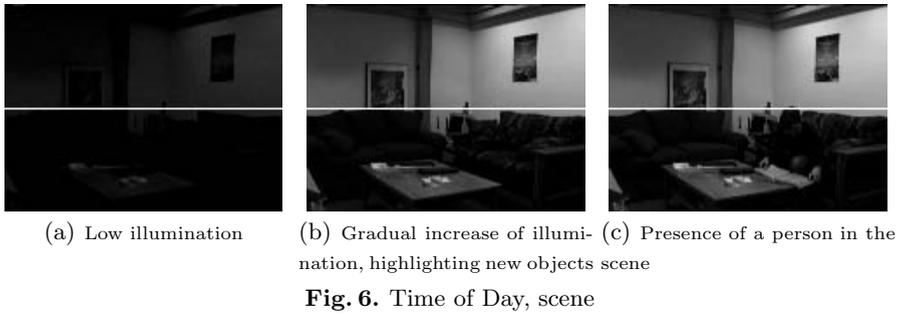


Fig. 5. Time of Day, historic



3.3 Light Switch

Global and sudden changes in lightning cause the background model to be discarded and a new one to be generated. Unlike the last problem, the global and sudden change of all the background model can be considered as a trigger for the change of background model.

Initial pattern (Fig. 7(a)). When the light is turned on the background pattern changes (Fig. 7(b)), Here, a global change can be appreciated. Figure 7(c) shows when the light is turned off and the pattern switches back to its initial model.

Figure 8(a) shows the initial scene. In Figure 8(b), the light is turned off (8(c)).

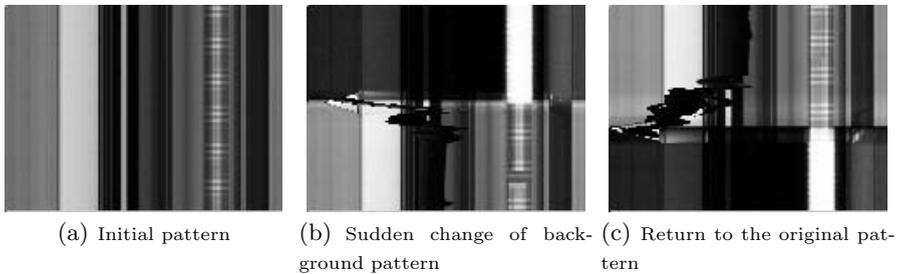


Fig. 7. Light Switch, historic

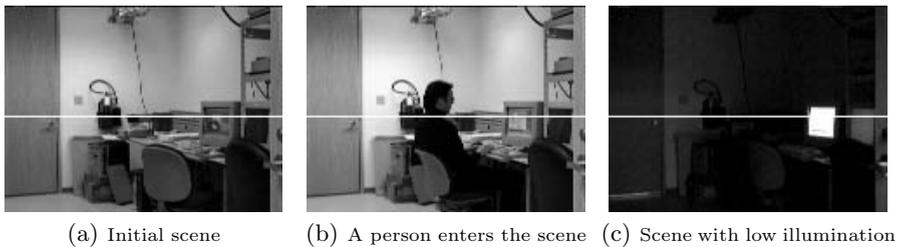


Fig. 8. Light Switch, scene

3.4 Waving Trees

Waving trees refers to movements that do not cease during a period of time, but which are provoked by object that must belong to the background. The common example is the undulation of the leaves in the trees, a movement which must be discarded but is noticeable.

In Figure 9(a), pattern is constantly undulating, yet a foreground region is perfectly noticeable (Figs. 9(b) and 9(c)).

Scene with a tree in constant undulation (Fig. 10(a)), then, a person appears (Fig. 10(b)); back to initial scenario (Fig. 10(c))

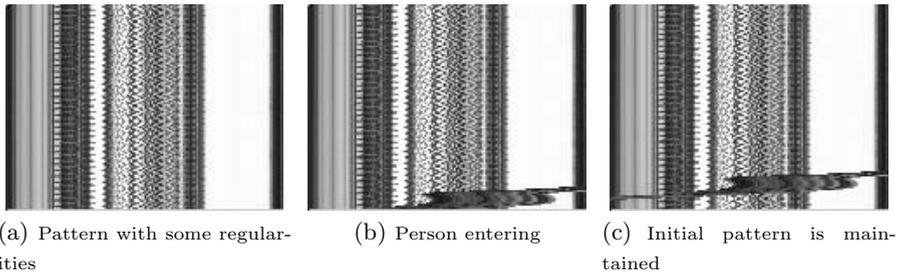


Fig. 9. Waving Trees, historic

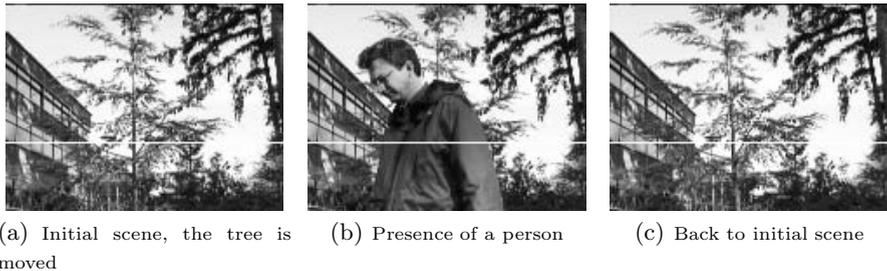


Fig. 10. Waving Trees, scene

3.5 Bootstrapping

Some scenes are situations where a full background model is not obtained due to an object that periodically appears and disappears.

In this problem, an ideal background model can not be constructed, due to constant appearance-disappearance of objects. Figures 11(a), 11(b) and 11(c) shows this situation. Figures 12(a), 12(b) and 12(c) display the same scene at different time.



(a) Some sections have regular patterns (b) New sections with regular patterns (c) New sections with regular patterns

Fig. 11. Bootstrapping, historic



(a) Initial scene (b) Scene after some time (c) Final scene

Fig. 12. Bootstrapping, scene

3.6 Light Switch (local)

This situation only occurs when only some regions of the background suddenly change their illumination but must be kept as part of the background model. Figures 13(a), 13(b) and 13(c) display patterns in which such changes can be appreciated. Scenes in Figures 14(a), 14(b) and 14(c) show different local light changes.



(a) Pattern with local changes (b) Object presence (c) Local lighth changes

Fig. 13. Local Light Switch, historic

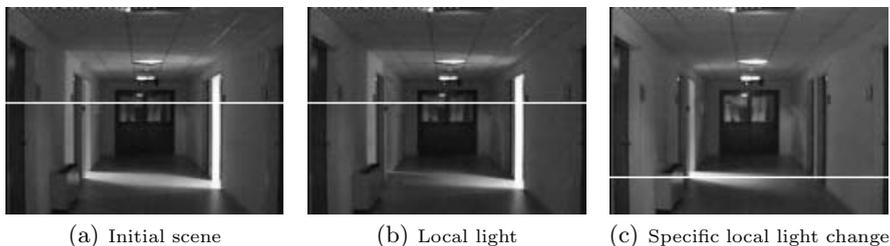


Fig. 14. Local Light Switch, scene

3.7 Sleeping Person

This problem occurs when an object is initially identified as part of the foreground and it then changes to a rest state for a long enough period of time for the algorithm to integrate it to the background model. Figure 15 shows a pattern with some local objects. The entire scene is shown in Figure 16.

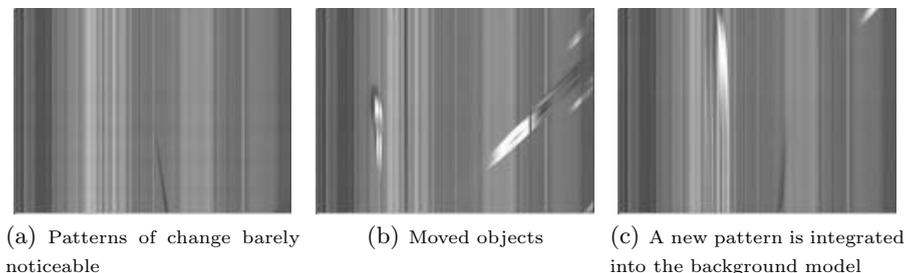


Fig. 15. Sleeping Person, historic

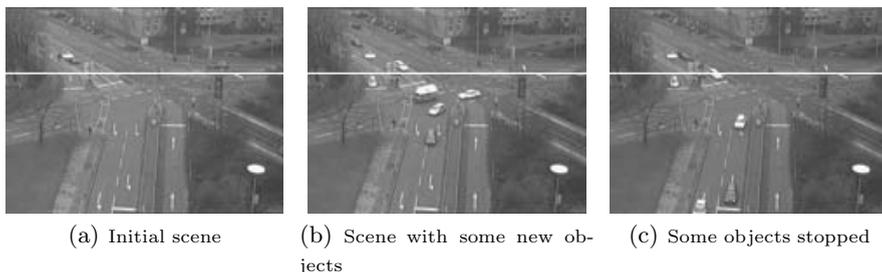


Fig. 16. Sleeping Person, scene

3.8 Shadows

Due to the lighting conditions on the scene, the foreground objects often cast shadows (Fig. 17), which due to the contrast with the background are considered as part of the foreground object. Some algorithms [8] try to resolve this issue through considering variations of color shades.

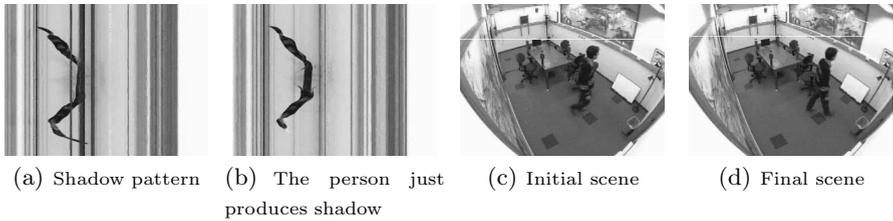


Fig. 17. Shadows: historic and scene

3.9 Reflections

Another phenomenon are the reflections on the object's surfaces. Usually, this effect is considered an special case of shadows, however, one of the main characteristics that makes reflections different from the shadows, is that reflections have well defined borders and can be identical to the foreground object. In Figures 18) the apparent lack of differences between the foreground object and the background can be observed.

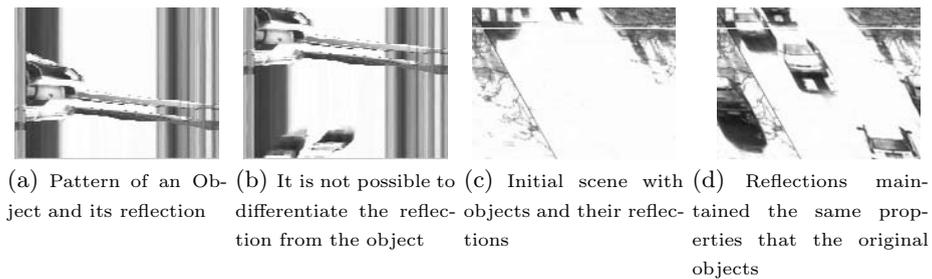


Fig. 18. Reflections, historic and scene

4 Conclusions

As already mentioned, various authors state that there is no better analysis in BS than the qualitative one, which is done visually. TimeViewer is a tool to support such analysis.

TimeViewer has also proved to be an excellent learning tool for those starting their studies and research in the area of **background subtraction**, **motion detection**, **change detection**, etc., because it eases the comprehension of different problems through the visualization of the pixel history in a sequence of images, as well as the storage of said snapshots from said historic for a detailed analysis.

Future work is aimed at using the pixel history for analysis of the movement of objects in a video, detecting patterns and trajectories. There is also potential as a video editing tool that needs to be explored.

Acknowledgements This work has been supported by Consejo Nacional de Ciencia y Tecnología (CONACyT), México, Centro Universitario de Ciencias Económico-Administrativas (CUCEA) of the Universidad de Guadalajara, México, and Project TIN2011-24367 from the Spanish Ministry of Economy and Competitiveness, España.

References

1. Radke, R.J., Andra, S., Al-Kofahi, O., Roysam, B.: Image change detection algorithms: a systematic survey. *IEEE Transactions on Image Processing* 14(3), 294–307 (2005)
2. Shapiro, L.G., Stockman, G.C.: *Computer Vision*. Prentice Hall (2001)
3. Benezeth, Y., Jodoin, P.M., Emile, B., Laurent, H., Rosenberger, C.: Comparative study of background subtraction algorithms. *Journal of Electronic Imaging* 19(3), 033003 (2010)
4. Pinto, N., Cox, D.D., Dicarlo, J.J.: Why is real-world visual object recognition hard? *PLoS Computational Biology* 4(1), e27 (2008)
5. Adelson, E.: Mechanisms for motion perception. *Optics & Photonics News*, 24–30 (August 1991)
6. Cheung, S.C.S.: Robust techniques for background subtraction in urban traffic video. In: *Proceedings of SPIE*, vol. 5308(1), pp. 881–892 (2004)
7. Toyama, K., Krumm, J., Brumitt, B., Meyers, B.: Wallflower: principles and practice of background maintenance. In: *Proceedings of the Seventh IEEE International Conference on Computer Vision* 1(c), vol. 1, pp. 255–261 (1999)
8. Bradski, G., Kaehler, A.: *Learning OpenCV: Computer vision with the OpenCV library*. O’Reilly Media (2008)
9. Tiburzi, F., Escudero, M., Bescos, J., Martinez, J.M.: A ground truth for motion based video-object segmentation. In: *2008 IEEE International Conference on Image Processing, Proceedings of ICIP 2008. Workshop on Multimedia Information Retrieval: New Trends and Challenges*, vol. 1, pp. 17–20. IEEE, San Diego (2008)
10. *Datasets for the Eleventh IEEE International Workshop on Performance Evaluation of Tracking and Surveillance* (2009)
11. Cristani, M., Farenzena, M., Bloisi, D., Murino, V.: Background Subtraction for Automated Multisensor Surveillance: A Comprehensive Review. *EURASIP Journal on Advances in Signal Processing* 2010, 1–24 (2010)
12. Adelson, E., Bergan, J.R.: Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A Optics and Image Science* 2(2), 284–299 (1985)