Vision, Video and Graphics (2005)
E. Trucco, M. Chantler (Editors)

Advances in Shadow Removing for Motion Detection Algorithms

P. Spagnolo & T. D'Orazio & M. Leo & A. Distante

Istituto di Studi sui Sistemi Intelligenti per l'Automazione - C.N.R.
Via Amendola 122/D, 70126 Bari, ITALY
{spagnolo,dorazio,leo,distante}@ba.issia.cnr.it

Abstract

Detecting moving objects is very important in many application contexts such as people detection and recognition, visual surveillance, automatic generation of video effects, and so on. Motion detection algorithms are very sensible to light conditions; in particular they suffer the presence of shadows and sudden changes due to light switches. Here we propose an additional module that can be applied to a generic gray-level motion detection algorithm. The only requirement is the presence of a reference image (background model). The main idea of the proposed approach is that variations in light conditions alter the intensity values of pixels in the image, but the basic structures in the scene remain unchanged. The algorithm we propose is based on the correlation between regions selected from the reference image and the current one. The experiments have been performed on image sequences acquired both in indoor and outdoor environments with natural and artificial lights.

1. Introduction

In the last years, motion detection has attracted great interest from computer vision researchers due to its promising applications in many areas, such as visual surveillance, virtual reality, advanced user interfaces, sport players tracking and recognition, model-based image coding, and so on. These different applicative contexts require as first step to automatically identify people, objects, or events of interest in different kinds of environments.

In literature, motion detection for object segmentation has been treated by several papers. Three classes of conventional approaches can be identified for moving target detection: optical flow [FD98, WH99], temporal difference [ABV85], and background subtraction [HHD00, WADP97, KCL98, FL98]. The last category, background subtraction, is the most used in literature. These methods implement a model of the background and compare the current image with this reference one. In this way the foreground objects present in the scene are detected, and the most reliable shapes of the moving objects are recovered. Many works have been done on motion detection but a general and valid solution has not been pointed out since there are a lot of problems concerning this subject. Some valid approaches focus their attention particularly on the problems of background modelling and updating, providing good solutions in most of the applicative contexts, but they suffer in presence of shadows, and sudden changes in illumination conditions.

There are many differences between outdoor and indoor contexts. Outdoor environments are characterized by the presence of a unique well-defined shadow, caused by the sun, while the presence of sudden light changes is limited to particularly weather conditions, i.e. very fast clouds. On the other side, indoor environments are characterized by the presence of many shadows, caused by the possible presence of a certain number of artificial light sources, and reflective surfaces. Moreover, sudden light changes are very common, due to the presence of light switches.

It's important to note that these two aspect of motion detection are very important, because they heavily alter the segmentation of moving objects, making the global system unreliable. However, the great part of existing systems does not handle these situations, restricting their target exclusively to standard light conditions.

Shadows occur when objects partially or completely occlude direct light from a light source. We can interpret shadows in the image, and the effect they have on the pixels in the scene, as a semi-transparent region in which the scene reflectance undergoes a local attenuation. While in outdoor scenes the shadow regions can be projected on the ground around the object, in indoor environment, due to different light sources and reflection effects, the shadows can be found everywhere in the image producing drastic changes in the background that cannot be afforded by using a simple adaptive model.

Some works have tried to solve the specific problem of shadow detection using similar photometric characteristics; usually they have been applied in outdoor environments and it is not clear what performances could be achieved in indoor scenes. In [EHD99] an interesting approach for shadow elimination, applied on color images, has been tested for the recognition of only one pedestrian's shadow with no dept of field. Authors have improved their work in [HHD98], but they use stereovision for the recognition of shadows of two pedestrians. Similarly, in [CGPP03] the problem is analyzed for color images, using information in HSV space to detect and removing shadows. An interesting approach is proposed in [RE95], but starting from the a-priori assumptions of the authors, it suffers to detect shadows with a small area. The shadow elimination criteria illustrated in [SMO99] is based on a series of assumptions, some of them not applicable in indoor environments, as the high light source intensity, and the geometric characteristics of the background, that must be plane. A good review and comparison of shadow removing algorithms is proposed in [PMTC03].

All these works try to propose a solution for the shadow removing problem, but they don't cope with the possible presence of sudden changes in illumination conditions. This problem has been remarked in [XRB02] and [TKBM99]. In literature, most of background subtraction approaches are not able to handle this situations, as indicated in [XRB02]. In [JRAS03] a solution is proposed: many different images of background acquired in different illumination conditions are maintained: when too many pixels are detected as moving pixels, the algorithm checks which background model gives the best result, and select it as current background. Conceptually similar approaches are proposed in [TKBM99] and [HBBZ03]: they are based on the presence of discontinuities in the training set. Unfortunately, this constraint can alter the results of the algorithm when moving object is not much different from the background model. A good treatment of this problem can be found in [SRP*01], where an interesting solution based on HMM is proposed.

The goal of this work is to propose an algorithm for shadow removing able to handle sudden changes in light conditions. The only requirement of this algorithm is the presence of a background model, so it can be easily added to any motion detection algorithm. The basic idea is that the intensity variation of each pixel covered by shadow is proportional to the intensity value of the same pixel in the reference image. So, in order to distinguish shadow regions from foreground objects, an algorithm based on the comparison of the correlation exhibited between regions selected from the reference image and the current one has been implemented. In this way no constraint is imposed on the directions of shadows, and the problem of multiple shadow regions in indoor scenes is solved. In addition, the same algorithm permits to cope with the problems due to light switches: when there are sudden changes, such as turning on/off some lights, the large number of false alarms is eliminated by comparing correlation between background and current regions in the same way of shadows removing.

In the rest of the paper, firstly some information about the background model we have used for our experiment are presented (section 2); then, a subsystem for shadow removing and recovering the correct shape in presence of sudden changes is explained (section 3). Finally, the experimental results obtained on real image sequences acquired in an indoor site are reported (section 4).

2. Background Subtraction

Foreground object segmentation is a primary and fundamental step of motion detection systems, independently of their applicative context: visual surveillance, advanced user interfaces, sport players tracking and recognition, and so on. The results of this step are the inputs for the subsequent processing (object recognition, motion analysis, human activity recognition, data compressionĚ). So it is very important to correctly extract the moving objects. That makes necessary to develop very reliable motion detection algorithms, that should be adaptive to luminance variations and able to reduce the number of false alarms. In literature many background modeling algorithm have been proposed: our opinion is that they work well in standard condition, but have same problems in particular situations, such as in presence of shadows and light switches. For these reasons, we have chosen to implement a standard background subtraction algorithm, improved by our shadow removing algorithm, that is useful to handle the light switches. A statistical model of the background has been implemented: for each pixel value at time t, a running average $B^t(x,y)$ and a form of standard deviation $V^t(x,y)$ are evaluated and maintained. A pixel (x,y) is considered a foreground pixel if its value differs from $B^t(x,y)$ more than 2 times $V^t(x,y)$, as illustrated in [KCL98]. Formally,

$$|I^{t+1}(x,y) - B^t(x,y)| > 2 \cdot V^t(x,y) \qquad (1)$$

A higher-level paradigm, based, for example, on the tracking information, or frame difference, is necessary to validate the results of this step, in order to avoid to detect as moving object any variation in the background objects. Any background subtraction approach is sensitive to variations of the illumination. Each algorithm needs a reliable background model image consistent at each time instant with the current scene luminance condition. In literature, a great number of different updating procedure can be found. Here we have chose to implement the algorithm proposed in [KCL98]. It is based on the use of an exponential filter to adapt spectral properties of each pixel of the background model to the current light conditions. The filter is of the form:

$$F(t) = e^{\frac{t}{\tau}} \tag{2}$$

where τ is a time constant which can be configured to refine the behavior of the system, as remarked in [KCL98]. So, the updating rules for running average and standard deviation are:

$$B^{t+1}(x,y) = \alpha I^{t+1}(x,y) + (1-\alpha)B^t(x,y) \tag{3}$$

$$V^{t+1}(x,y) = \alpha|I^{t+1}(x,y) - B^{t+1}(x,y)| + (1-\alpha)V^t(x,y) \tag{4}$$

where $\alpha = \tau * f$, f is the system frame rate. It is important to note that the value of α affects the quickness of the updating. Small values of α cause a slow updating of the background, while a very fast updating can be obtained with greater values of α.

The background module here explained is the starting point of the shadow removing approach we propose in this work.

3. Shadow Removing

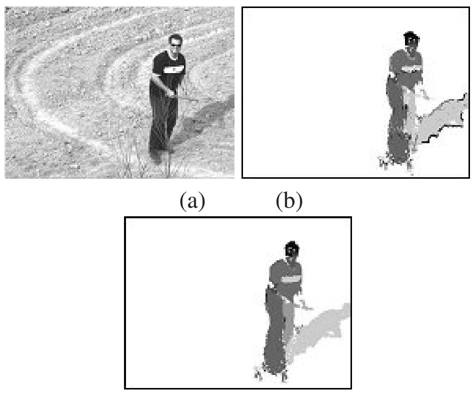
After the background subtraction, in the resulting binary image many small clusters of pixels are still observable: a one-step filter removes blobs whose size is lower than a certain threshold. Finally, an image with only foreground objects is generated, where each object contains also its own shadows. The presence of shadows is a great problem for a motion detection system: independently from the kind of algorithm implemented for foreground regions extraction (background subtraction, frame difference, optical flow), shadows are always detected because they appear as moving objects. This problem is mostly remarked in indoor contexts, where shadows are emphasized by the presence of many reflective objects; in addition shadows can be detected in every direction, on the floor, on the walls but also on the ceiling, so typical shadow removing algorithms, that assume shadows in a plane orthogonal with the human plane, cannot be used.

Shadows drastically change the topological characteristics of the objects in an unpredictable way, causing serious trouble to other subsystem processing on them, such as an object recognition or a behaviors analysis. To prevent this problem, correct shapes of the objects must be extracted: the system needs the implementation of a shadow removing algorithm.

The shadow removing approach here described starts from the assumption that a shadow is an abnormal illumination of a part of an image due to the interposition of an opaque object with respect to a bright point-like illumination source. From this assumption, we can note that shadows move with their own objects but also that they have not a fixed texture, as real objects: they are half-transparent regions which retain the representation of the underlying background surface pattern. Therefore, our aim is to examine the parts of the image that have been detected as moving regions from the previous segmentation step but with a texture substantially unchanged with respect to the corresponding background. A segmentation procedure has been applied to recover large regions characterized by a constant photometric gain; then, for each region previously detected, the correlation between pixels is calculated, and it is compared with the same value calculated in the background image: regions whose correlation is not substantially changed are marked as shadow regions and removed. In the following, each of these steps will be explained in details.

The first step of this algorithm is the segmentation of the image, with the purpose to separate the foreground object from its own shadows. The detected foreground object is segmented in regions $\{F_i\}$ characterized by a constant photometric gain:

$$\Lambda^t(x,y) = \frac{I^{t+1}(x,y)}{B^t(x,y)} \tag{5}$$



(a) (b)

(c)

Figure 1: *: (a) Original grey level image; (b) segmentation of a foreground object into a number of regions characterized by the same photometric gain, (c) the result of the merging process.*

As we can observe in fig.1-b, this segmentation process produces several regions on the foreground object. The actual shadow is given by a large central region and a number of neighboring small patches generally covering the shadow contour. This is due to the different photometric gain exhibited by the shadow edge points with respect to the rest of the shadow. The reason of this phenomenon can be searched in the effects of the finite size illumination source, that reduce the attenuation. For our purposes, each candidate shadow region should grow to include the neighboring smaller regions. A merging step of these regions is performed on the basis of their constant photometric gain similarity. For each region we calculate a running average and a standard deviation on the photometric gains exhibited by all internal points. Then, for each region whose area is less than a fixed threshold, we use the photometric gain running average and standard deviation to establish the more similar neighboring region for merging. After each merge new values are determined for the relative photometric gain statistics. The process is iterated until all the smallest regions have been merged. In Fig. 1-c the output of the merging step is shown. At this point a further step is required to discriminate the real shadow regions, and remove them. All candidate shadow pixels are labelled and the regions with a notable percentage (almost 80%) of these candidate shadow-pixels are removed. Candidate shadow pixels are detected as follows: their photometric gain has to be lower than unit; their correlation values with neighboring points are compared with the corresponding one obtained at the same location on the reference background image; finally if the observed difference is too small, then the pixel is labelled as candidate shadow point. The reliability of this approach improves by increasing the number of neighboring points that are correlated with every pixel. However, a satisfying trade-off must be found with the computational time constraint. For this purpose, we have implemented a simplified version of this algorithm that gives similar good results, but with a lower computational time. In detail, we have proved as the simple correlation between only two adjacent pixels belonging to the same region, i.e. their ratio, is sufficient for an efficient shadow detection. In other words, our shadow elimination algorithm examines all the regions previously detected; for each of them, the ratio between all couples of adjacent pixels is calculated and compared with the corresponding value in the reference image; if these values are similar, these pixels are labelled as candidate shadow points. Formally, for each region F_S:

$$D = \begin{cases} \left|\frac{I(i,j)}{I(i,j+1)} - \frac{B(i,j)}{B(i,j+1)}\right| & j < NumCol \\ \left|\frac{I(i,j)}{I(i+1,j)} - \frac{B(i,j)}{B(i+1,j)}\right| & j = NumCol \end{cases} \tag{6}$$

If D is greater then a threshold experimentally selected, the pixel (i,j) is strictly correlated with (x,y), so they can be considered as shadow points. Otherwise, they probably will be foreground points. All the regions containing a great number of shadow points are removed.

We have verified that the results obtained using only two pixels ratio are similar to those obtained using a more complex correlation, but in a fraction of time.

The proposed method works very well both on indoor and on outdoor sequences. The two pixels ratio is a very fast shadows elimination algorithm, but in theory it could have problems removing not only the shadows, but also some points of people whose texture is similar to the background model. In practice, in our experiments on different situations, these cases have not been encountered.

3.1. Sudden light changes

The approach we propose for the shadow removing starts from the assumption that a shadow region presents about the same texture with respect to the reference image. In other words, the absolute intensity values change, but the relation between them remains the same. This observation can be used to reduce the effects of light switches in indoor context. In this case, traditional motion detection algorithms, as detailed explained in section 1, fail due to the sudden and unpredictable illumination changes. Even the examined background subtraction algorithm is not able to handle these situations. The background updating rules (3) and (4) are able to adapt the background model to standard light changes, but cannot work in presence of a sudden variation of such conditions. So, the resulting images seem totally unusable. However the successive application of the shadow removing algorithm produces surprising results, making the complete motion detection system more robust and reliable compared to similar approaches proposed in literature. Some works have tried the use of supervised approaches including different background images acquired in several light conditions in the reference model. Their main drawback is that in indoor environments there are unpredictable light variations for the simultaneous presence of both natural and artificial light sources. The possibility to cope with these sudden illumination changes in an unsupervised way make the proposed system very general and robust.

4. Experimental Results

The experiments have been performed on real image sequences acquired with a static TV camera Dalsa CA-D6 with 528 X 512 pixels; the frame rate selected is 20Hz. The processing is performed with a Pentium IV, with 1,5 GHz and 128 Mb of RAM. We have chosen to test algorithm both in outdoor and indoor condition; in particular, indoor sequences have been acquired in different light conditions, and in presence of sudden changes, due to light switches. The characteristics of each test sequence are resumed in the table 1.

The results obtained applying the proposed motion detection algorithm are very encouraging. In fig. 2, some images obtained during the elaboration are plotted. We have chosen to report an image for each sequence, so in the first block (2-a) an original grey-level image for each test sequence is illustrated. The results after the background subtraction step are reported in the second block (2-b). Finally, the last section (2-c) shows the results obtained after the second step of shadow removing.

It can be observed that the aforementioned problem of multiple shadows is evident for the images acquired in an indoor environment. Shadow regions are projected on the floor, on the near desks, and also on the wall above the person. The second image has been taken from a sequence after a sudden light change due to a light switched off. The image obtained after the background subtraction step is not significant since a large number of background pixels have not been removed. The final images produced after the shadow removing step become completely clear in every condition.

The test sequences described in table 1 together with the corresponding sequences obtained with the proposed algorithm can be seen at the site "http://www.tnl.it/cnr".

In order to have a quantitative estimation of the error, we have characterized the Detection Rate (*DR*) and the False Alarm Rate (*FAR*), as proposed in [JUS03]:

$$DR = \frac{TP}{TP+FN} \quad FAR = \frac{FP}{TP+FP} \tag{7}$$

where TP (true positive) are the detected regions that correspond to moving objects; FP (false positive) are the detected regions that do not correspond to a moving object; and FN (false negative) are moving objects not detected. For this test we randomly selected 25 frames for each sequence, leading to total of 100 sample frames for evaluation. The "ground truths" of these 100 frames were manually generated. In table 2 we can se the results obtained on the four test sequences. It should be note that only the results obtained after the application of the proposed shadow removing procedure are presented, because the same results obtained without this algorithm could be irrelevant.

We can note that the FAR parameter is always under 6% in the first three test sequences (indoor environments, more sensitive to effects of shadows) and even under 4% in the fourth test sequence (outdoor context, standard shadows).

Finally, we have tested our algorithm on the same sequences proposed in [PMTC03], available at the website http://cvrr.ucsd.edu/aton/shadow; in [PMTC03] a modified version of (7) has been used for testing and comparing different shadow removing algorithms. In table 3 the results we have obtained are reported. It can be note that even on these test sequences, the results we have obtained appear to be acceptable. For an immediate comparison, the results obtained by other authors can be found in [PMTC03].

Table 1: *The characteristics of the test sequences.*

Sequence number	Number of frames	Context
1	7894	Indoor
2	8387	Indoor
3	12058	Indoor
4	12551	outdoor

Table 2: *Rates to measure the confidence.*

Test Sequence	*DR%*	*FAR%*
1	87,46	5,72
2	93,81	4,16
3	89,12	5,83
4	94,31	3,26

5. Conclusion and future works

This work deals with the problem of improvement motion detection in a generic context. The results of the proposed algorithm can be used for object recognition, or human activity recognition, advanced user interfaces and so on. A shadow removing algorithm has been implemented to cope with the problem of incorrect extraction of shapes. The main contribution of this work lies in its unsupervised ability to manage different kinds of situations of indoor and outdoor scenes such as different light sources, small movements in background objects, and sudden variations of light conditions. The proposed approach can be easily added to a generic motion detection system; the only requirement is the presence of a background image.

A large number of experiments have been carried out: the proposed approach gives good results, in particular shadow regions are correctly removed in a great number of frames. The negative effects of sudden changes due to light switches have been easily avoided with the proposed algorithm.

Now we are effecting intensive tests on other test sequences; in particular we are testing our algorithms on the PETS video sets.

Future work will investigate the effective real time implementation of our algorithms in a motion detection system.

Table 3: *Experimental results obtained on the test sequences proposed in [PMTC03].*

Test Sequence	*DR%*	*FAR%*
Highway I	93,65	3,25
Highway II	89,43	4,63
Campus	88,91	6,36
Laboratory	91,77	3,96
Intelligent Room	96,54	3,43

References

[FD98] FEJES S., DAVIS L.S.: What can projections of flow fields tell us about the visual motion, In Proc. Intern. Confer. on Computer Vision ICCV98, 1998, pp. 979-986.

[WH99] WIXSON L., HANSEN M.: Detecting salient motion by accumulating directional-consistent flow, In proc. of Intern. Conf. on Comp. Vis., 1999, vol II, pp 797-804.

[ABV85] ANDERSON C., BURT P., VAN DER WAL G.: Change detection and tracking using pyramid transformation techniques, In Proc. of SPIE - Intell. Robots and Comp. Vision Vol. 579, pp.72-78, 1985.

[HHD00] HARITAOGLU I., HARWOOD D., DAVIS L.S.: A Fast Background Scene Modeling and Maintenance for Outdoor Surveillance, ICPR, pp.179-183, Barcelona,2000.

[WADP97] WREN C., AZARBAYEJANI A., DARRELL T., PENTLAND A.: Pfinder: Real-time tracking of the human body, IEEE Trans. on Patt. An. and Mach. Intell. 19(7): pp.780-785, 1997.

[KCL98] KANADE T., COLLINS T., LIPTON A.: Advances in Cooperative Multi-Sensor Video Surveillance, Darpa Image Underst. Work., Morgan Kaufmann, Nov. 1998, pp. 3-24.

[FL98] FUJIYOSHI H., LIPTON A.: Real-time human motion analysis by image skeletonisation, IEEE WACV, Princeton NJ, October 1998, pp.15-21.

[XRB02] XIE B., RAMESH V., BOULT T.: Sudden Illumination Change Detection Using Order Consistency, Workshop on Statistical Methods in Video Processing (in conjunction with ECCV2002), June 2002.

[TKBM99] TOYAMA K., KRUMM J., BRUMITT B., MEYERS B.: Wallflower: Principles and Practice of Background Maintenance, International Conference on Computer Vision, September 1999, Corfu, Greece.

[JRAS03] JAVED O., RASHEED Z., ALATAS O., SHAH M.: KNIGHTM:A Real Time Surveillance System for Multiple Overlapping and Non-Overlapping Cameras, The fourth International Conference on Multimedia and Expo (ICME 2003), Baltimore, Maryland, 2003.

[HBBZ03] HAMID R., BALOCH A., BILAL A., ZAFFAR N.: Object Segmentation Using Feature Based Conditional Morphology- IEEE International Conference on Image Analysis and Processing 2003, Mantova, Italy.

[SRP*01] STENGER B., RAMESH V., PARAGIOS N., COETZEE F., BOUHMAN J.: "Topology free hidden markov models: Application to background modeling," in Proc. IEEE Int. Conf. Computer Vision, 2001, pp. 294–301.

[EHD99] ELGAMMAL A., HARWOOD D., DAVIS L.S.: Non-parametric model for background subtraction, Proc. ICCV'99 Frame-Rate Workshop, 1999.

[CGPP03] CUCCHIARA R., GRANA C., PICCARDI M., PRATI A.: Detecting Moving Objects, Ghosts, and Shadows in Video Streams, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (10) (2003) 1337-1342.

[SMO99] STAUDER J., MECH R. OSTERMANN J.: Detection of moving cast shadows for object segmentation, IEEE Trans. On Multimedia, Vol. 1, N.1, pp. 65-76, 1999.

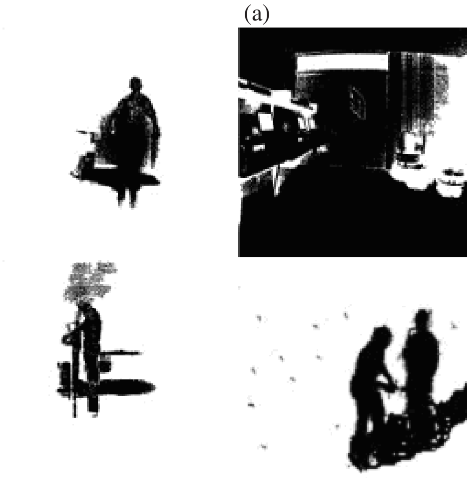
[PMTC03] PRATI A., MIKIC I., TRIVEDI M.M., CUCCHIARA R.: Detecting Moving Shadows: Algorithms and Evaluation, in IEEE Transaction on PAMI, vol. 25, n.7, pp. 918-923, July, 2003

[JUS03] JARABA E.H., URUNUELA C.O., SENAR J.: Detected motion classification with a double-background and a Neighborhood-based difference, Patt. Recogn. Letter, Elsevier, Location, pp. 2079-82, 2003(24).

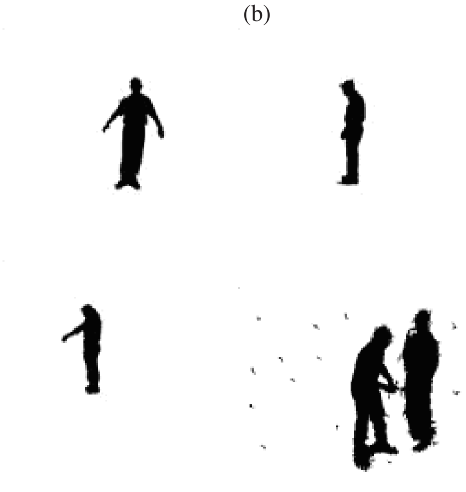
[HHD98] HARITAOGLU I., HARWOOD D., DAVIS L.S.: W4S: A Real Time System for Detecting and Tracking People in 2.5 D, Proceedings of the 5th European Conference on Computer Vision, Freiburg, Germany, June, Vol. 1, pp. 877-892, 1998.

[RE95] ROSIN P.L., ELLIS T.: Image difference threshold strategies and shadow detection, Proceedings of the 6th British Machine Vision Conference, Birmingham, UK, September, pp. 347-356, 1995.

(a)



(b)



(c)

Figure 2: *: (a) original grey level images; (b) results after background subtraction; (c) results after background subtraction and shadow removing*