THE PUBLISHING HOUSE
OF THE ROMANIAN ACADEMY

PROCEEDINGS OF THE ROMANIAN ACADEMY, Series A,
Volume 13, Number 3/2012, pp. 269–277

NOVEL APPROACH FOR MOVING HUMAN DETECTION AND TRACKING IN STATIC CAMERA VIDEO SEQUENCES

Tudor BARBU

Institute of Computer Science of the Romanian Academy
E-mail: tudbar@iit.tuiasi.ro

An automatic multiple person detection and tracking technique for static camera movies is proposed in this paper. First, a moving human identification method is provided. It detects the video objects by using a novel temporal differencing based algorithm and some morphological processes. Then, our approach decides which moving objects represent walking persons, by computing some human size parameters and extracting some skin segments. A novel human tracking technique using a correlation-based video object matching process is then proposed in this paper.

Key words: video object detection, video tracking, moving person, temporal differencing, morphological operations, skin recognition, normalized correlation, human matching.

1. INTRODUCTION

Human detection and tracking in video sequences has represented a high-interest computer vision domain in recent years [1, 2]. Obviously, it represents the most important sub-class of the video object detection and tracking research area.

Moving object detection and tracking is a challenging computer vision task consisting of two closely related video analysis processes. The first one, object detection, involves locating an image object in the frames of a video sequence, while video tracking performs the monitoring of the video object spatial and temporal changes during sequence, including its presence, position, size, shape, etc. [1]. Any object tracking technique must solve the temporal correspondence problem that is the task of matching the target object in successive frames. Usually, the tracking process starts with detecting the initial instance of the moving object, then identifying that image object repeatedly in subsequent frame sequence. Video tracking is often a difficult process, due to factors such as abrupt object motion, object occlusions and camera motion [2].

Person detection represents the process of identifying the presence of humans in video streams and differentiating them from non-human video objects. Human tracking identifies the instances of each detected person in the frames of the analyzed movie. Detecting humans in images and videos constitutes a challenging task, being complicated by numerous factors, such as: variable people appearance, camera position, wide range of poses adopted by persons, variations in illumination, brightness, contrast levels or backgrounds, and person occlusions [2,3]. Over the last decade the problem of detecting and tracking humans has received a considerable interest. Significant research has been devoted to detecting, locating and tracking people in digital images and videos, since numerous applications involve humans' locations and movements [2]. Many person identification and tracking techniques have been developed recently [3]. Thus, human detection has been approached using methods based on background subtraction [4], frame differences [1, 5], Histograms of Oriented Gradients (HOG) [6], Partial Least Squares Analysis [7], Haar Wavelets with SVM [8], SIFT descriptors [9], Hough transforms [10] and Active Contour models [11]. Also, various human motion tracking techniques have been proposed in this period. Most popular approaches use Kalman filtering [12], Mean-Shift algorithms [13], Optical Flow [14], HMM [15], contour tracking [15] and human matching [16].

Obviously, person detection and tracking has a wide variety of computer vision application fields. Its most important application areas are: video surveillance and security systems, law enforcement, biometrics, human-computer interaction, video indexing and retrieval, medical imaging, augmented reality and robotics.

Some methods for human detection and recognition in digital images were provided in our previous works. Thus, we proposed several face detection [17], face recognition [18] and human skin identification techniques in those papers [17]. Also, we have approached the video object detection and tracking domain in the recent years [19]. So, we consider human body detection and tracking in video sequences as the next important computer vision task to be solved and also a unification of the mentioned research areas. Face and skin features are used by the motion analysis algorithms in our person detection and tracking research.

In this article we propose an automatic multi moving human detection and tracking system for static camera videos. The developed computer vision system only identifies the moving people, not all humans from video frames. It is mainly intended for pedestrian detection and tracking in fixed camera movies, therefore being very useful for video surveillance applications. Our novel detection and tracking approach consists of the following steps. A video pre-processing step, consisting mainly of image denoising and restoration operations, is performed first. Then, a moving object identification technique is applied on the filtered video sequence. The multiple video object detection that is described in the next section uses an improved temporal differencing method and several morphological operations, representing an important contribution of this paper. Then, each identified video objects is classified either as human or non-human, using a method, based on skin detection and some human parameters, that is also provided in the second section and constitutes another original contribution of this article. Next, a human tracking process is performed on the detected persons. The proposed tracking technique that is presented in the third section uses a correlation-based moving object matching algorithm. Some human detection and tracking experiments are discussed in the fourth section. This article ends with a section of conclusions and a list of references.

2. MULTIPLE MOVING PERSON DETECTION TECHNIQUE FOR STATIC CAMERA VIDEOS

We consider the following computer vision task: given a video sequence recorded with a fixed camera, one has to detect and track all the walking persons from this sequence. Often, the movie sequence must be pre-processed before performing the video analysis on it. The pedestrian tracking has to be performed within each shot of the video stream, therefore a temporal segmentation should be performed first on it. In our previous works we proposed some video shot detection approaches which can be applied here [20]. A key frame extraction can also be performed on the analyzed movie, if it has a very large dimension. Depending on the state of the video sequence, some enhancement operations may also be required. We also provided denoising and restoration techniques [21,22], which facilitate image object identification in the video frames.

The multi moving human detection process consists of two steps. In the first subsection we propose a novel foreground segmentation of the enhanced video sequence. The identification of the moving objects representing humans is approached in the second subsection.

2.1. A novel video object detection technique

Our proposed multiple moving object detection approach is based on a novel temporal differencing algorithm and some morphological operations. The frames of the analyzed video are converted into the grayscale form, resulting the sequence $I_1,..., I_n$. The video motion of this sequence can be estimated using a frame difference technique [1, 2, 5]. Thus, the difference of two consecutive video frames indicates the motion between them, the resulted non-black image zones representing the moving regions. Unfortunately, such a moving region does not represent always an entire moving object of the video stream. That happens because both the image objects and the frame background could be composed of more homogeneous regions, characterized by various intensities. Let us express the difference between two video frames as:

$$FD(i, j) = I_i - I_j, \quad \forall i, j \in [1, n], \quad i \neq j. \tag{1}$$

Any moving object that is present in the frames I_i and I_j, is represented in the grayscale image $FD(i, j)$ by some non-black regions. The high-intensity pixels of the object (those having greater values than background pixels) are displayed in $FD(i, j)$ at the locations occupied in I_i, while its low-intensity pixels are displayed in $FD(i, j)$ at their positions in I_j. It is also possible that some non-zero pixels represent some very small camera movements, and not moving objects.

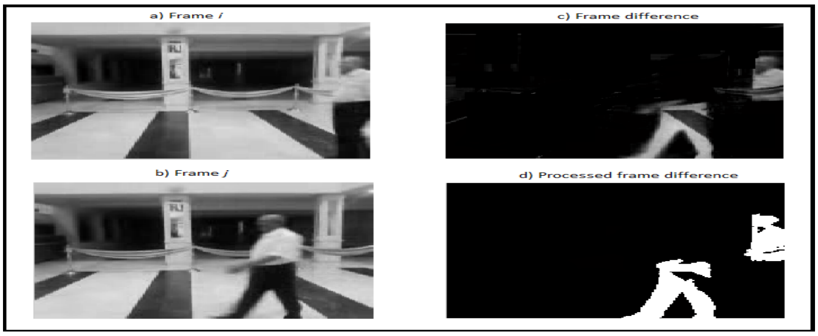

Fig. 1 – Temporal differencing for two video frames.

A video frame difference example is displayed in Fig. 1. In the images (*a*) and (*b*) there are depicted two frames of a video containing a walking person. Their computed frame difference image is displayed in Fig. 1 (c). It contains two major non-black zones, representing the higher intensity region (upper half) and the lower intensity region (bottom half) of the person. Several small area non-black spots represent some errors provided by redundant noise or unintended camera motion. The frame difference computed by (1) can be converted into the binary form by setting to 1 all pixels exceeding a properly selected threshold value *T*, as in:

$$FD_b(i,j) = \begin{cases} 1, & \text{if } FD(i,j) \geq T \\ 0, & \text{if } FD(i,j) < T \end{cases}, \quad \forall\, i,\, j \in [1,n], \quad i \neq j. \tag{2}$$

Some necessary morphological operations are then applied to $FD_b(i,j)$. We consider the dilatation process is quite useful in this case [23]. The dilated binary image is obtained from:

$$FD_S(i,j) = FD_b(i,j) \oplus S = \bigcup_{s \in S} FD_b(i,j)_s \tag{3}$$

where $\oplus$ represents the dilatation morphological operator and *S* is the considered structuring element, a $[l \times l]$ square [23]. The connected components of the dilated binary image $FD_S(i,j)$ are then determined. The components representing errors must be removed, so, our approach identifies and deletes the white regions having inappropriate dimensions. The following categories of connected components are discarded:

- Small white spots: connected components whose area does not exceeds a given threshold
- Connected components whose bounding rectangle has a dimension under a threshold value
- Components characterized by a low solidity (ratio between the region area and its bounding box area)

The result of this binary image processing is noted as $FD_S^p(i,j)$. The processed frame difference image corresponding to the video frames depicted in the figure is displayed in Figure 1 (d). The proposed motion detection algorithm identifies the moving objects from I_1 to I_{n-1}, first. At each step *i*, our approach determines the video objects of I_i, by using the next two frames, I_{i+1} and I_{i+2}. The corresponding morphologically processed frame differences $FD_S^p(i,i+1)$ and $FD_S^p(i,i+2)$ are computed, their intersection being determined next. We consider the *intersection* of two binary images to be the image having the same pixel values at the locations where they coincide and 0 in the locations where they differ:

$$\begin{aligned} &(FD_S^p(i,i+1) \cap FD_S^p(i,i+2))[x,y] = \\ &\begin{cases} FD_S^p(i,i+1)[x,y], & \text{if } FD_S^p(i,i+1)[x,y] = FD_S^p(i,i+2)[x,y] \\ 0, & \text{if } FD_S^p(i,i+1)[x,y] \neq FD_S^p(i,i+2)[x,y] \end{cases}. \end{aligned} \tag{4}$$

Some morphology-based processing can be further applied on the resulted intersection image [23]. The connected components of $FD_S^p(i,i+1) \cap FD_S^p(i,i+2)$ correspond to all high-intensity components of the moving objects of I_i. The low-intensity components of its moving objects are identified similarly. These low-intensity regions correspond to connected components of binary intersection image $FD_S^p(i+1,i) \cap FD_S^p(i+2,i)$. The video objects of I_i are determined by computing sum of these image intersections:

$$Obj(i) = (FD_S^p(i,i+1) \cap FD_S^p(i,i+2)) + (FD_S^p(i+1,i) \cap FD_S^p(i+2,i)). \quad (5)$$

The connected components of the binary image *Obj*(*i*) correspond to moving objects of frame I_i. All moving video objects of the frames I_1, $I_2, \ldots,$ I_{n-1} are detected this way. So, the last stage of our detection approach consists of identification of the final locations in I_n of these objects. This detection is performed through a backward identification process that is expressed as following:

$$Obj(n) = (FD_S^p(n,n-1) \cap FD_S^p(n,n-2)) + (FD_S^p(n-1,n) \cap FD_S^p(n-2,n)). \quad (6)$$

Each binary image *Obj* (*i*) contains the same number of connected components. For each component one determines its bounding rectangle. Sub-images of the frame I_i corresponding to these rectangles represent its detected moving objects (foreground). A moving object detection example is described in Fig. 2:

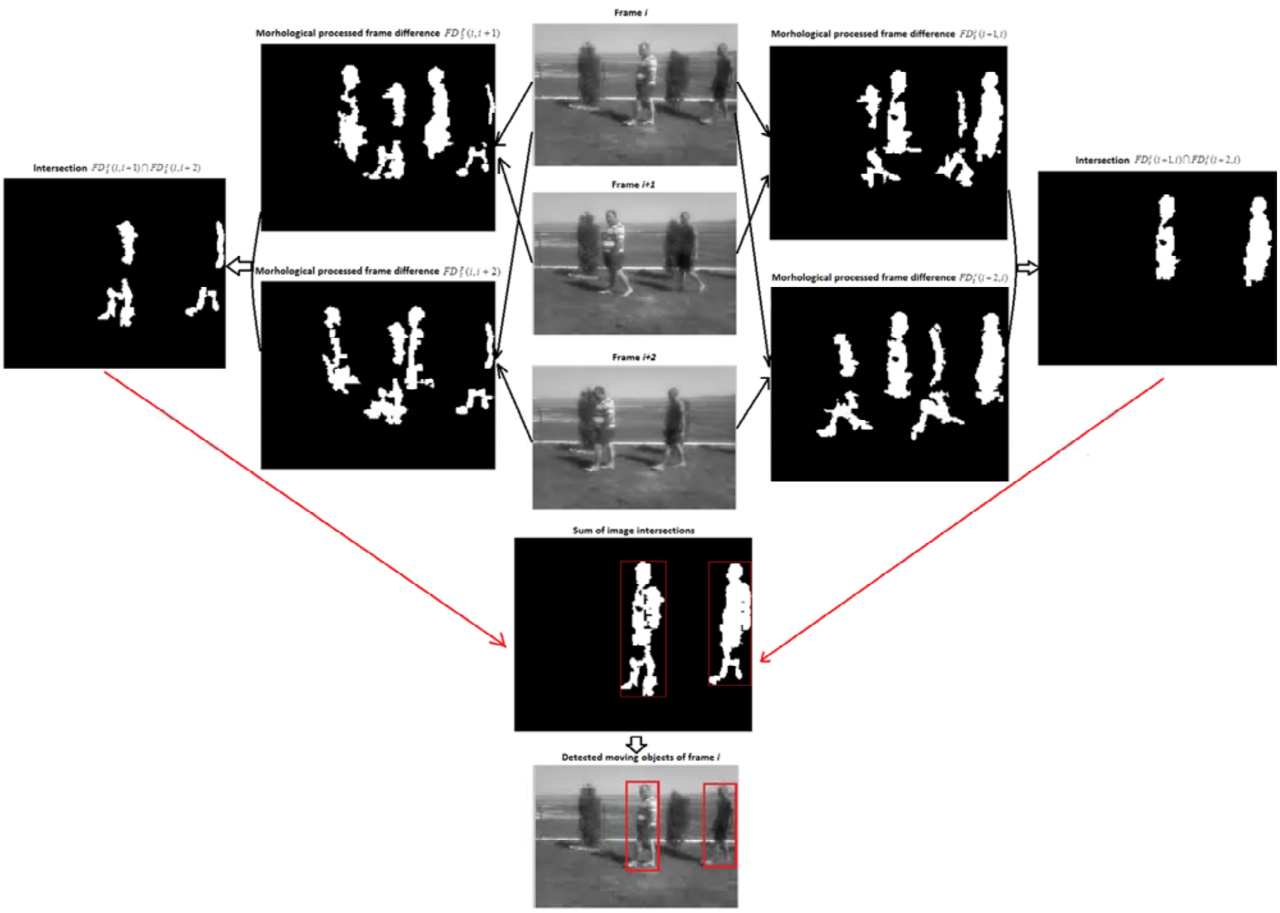

Fig. 2 – Moving object detection scheme.

2.2. Video object analysis for moving person identification

The moving objects of the video sequence being identified, our video analysis must now decide which of them represent human persons. Therefore we consider a set of conditions which have to be satisfied by a video object representing a pedestrian. Our first condition says the height of the bounding box of the image object representing a person candidate has to be at least two times greater than its width. The second condition says that the object solidity must be over 50%. Let $Obj'(i)$ be the binary image obtained after removing from *Obj* (*i*) the components not satisfying the two conditions. The third condition says the moving object has to contain skin regions to be considered a human body. Therefore, we introduce an original skin detection technique for color images. Human skin detection represents the process of identifying skin-colored regions in a digital image. It is usually used as a preprocessing step for diverse application areas, such as face detection, human detection, content filtering and content-aware video compression. In the last decade extensive research has focused on the skin detection task. The most popular skin recognition technique is the detection algorithm proposed by Fleck and Forsyth in 1996 that is based on explicitly defined regions [24]. We also proposed novel skin identification methods using decision rules in our previous works [17]. Thus, our detection approach defines explicitly the human skin regions. The algorithm applies some restrictions to the channels of the color space, each component of a skin pixel being restricted to a specific interval [17]. While it represents one of the most used color spaces for processing and storing of digital image data, RGB is not a favorable choice for skin color analysis, because of the high correlation of its three channels and the mixing of luminance and chrominance data. For this reason our skin segmentation technique uses HSV and YC_rC_b colorspaces. The RGB image is converted into Hue Saturation Value format, by computing the 3 components using the well-known conversion equations. We obtain the components, H, S and V, as 3 matrices whose coefficients belong to [0,1] interval. We are interested mainly in hue value, H.

Then, a conversion in $YC_r C_b$ color space is performed. In fact, it does not represent an absolute color space, but a way of encoding the RGB information. In this format Y represents the luminance, while C_r and C_b are the blue-difference and red-difference chroma components. These three components of the color space are computed as linear combinations of R, G and B image components. Therefore, the computation formulas of the chroma components, C_r and C_b, have the general form $\alpha \cdot R+\beta \cdot G+\gamma \cdot B+128$, where the coefficients α, β, and $\gamma \in [-0.5, 0.5]$. We consider some proper values for these coefficients and obtain the components:

$$\begin{cases} C_r = 0.15 \cdot R - 0.3 \cdot G + 0.45 \cdot B + 128 \\ C_b = 0.45 \cdot R - 0.35 \cdot G - 0.07 \cdot B + 128. \end{cases} \tag{7}$$

A set of restrictions is then applied on these two components and on the hue, to identify the skin-colored regions. Thus, each pixel of the analyzed $[M \times N]$ image I belongs to a human skin segment if the corresponding values in C_r, C_b, H are situated in certain intervals. One obtains a binary image S_I, identical in size with I, whose white regions correspond to skin segments. So, our skin detection model is expressed as

$$S_I(i, j) = \begin{cases} 1, if \;\; C_r(i, j) \in [150,165], C_b(i, j) \in [145,190], H(i, j) \in [0.02, 0.1] \\ 0, \qquad \qquad \text{otherwise} \end{cases}, \tag{8}$$

where $i \in [1, M]$ and $j \in [1, N]$. The connected components of S_I could represent some human body parts such as faces, hands, arms, legs and feet. Therefore, the detection of skin regions in the image indicates the human presence. Obviously, our skin recognition approach works on the color versions of the frames I_i. A skin detection example is displayed in Fig. 3. The color frame i, whose moving objects have been detected in Fig. 2a, while its identified skin regions are displayed in the binary image of Fig.2b.

Fig. 3 – Skin detection example.

If these skin segments modeled by (8) are identified in the moving object locations, then those video objects should represent human persons. So, for each frame I_i, our approach computes the intersection $Obj'(i) \cap S_{I_i}$ and determines its connected components. The moving objects of $Obj'(i)$ containing these components are considered humans. Let the subimages of I_i determined by the bounding boxes of these *human* objects be noted as $\{P_1^i,..., P_K^i\}$. Obviously, the two moving objects detected in the example displayed in Fig. 2 satisfy the width to height ratio and solidity conditions. Also, the skin segments from Fig. 3b are detected in the locations of these moving objects, therefore these objects represent walking persons.

3. CORRELATION-BASED MATCHING APPROACH FOR MULTIPLE HUMAN TRACKING

In this section we consider a video tracking method for the previously identified moving persons. Now, one knows the moving objects representing humans of each frame, but the instances of any given person in the next frames remain unknown. The proposed tracking approach identifies for each pedestrian from a frame its instance in the next frame, which must represent the most similar moving person.

We consider a correlation-based human matching approach for video tracking. Thus, for each video frame I_i we have the moving person sequence $\{P_1^i,..., P_K^i\}$. For each $j \in [1, K]$, one has to find in the next frame I_{i+1} the human object $P_t^{i+1} \approx P_j^i$, where $t \in [1, K]$ and symbol $\approx$ represents the content similarity between image objects. Let us note $t = match_i(j)$ the index value that must be determined for all i and j. This video object matching task is performed by using a normalized cross-correlation procedure [25].

Our approach computes the 2D cross-correlation coefficients between the current object of the current video frame and all the human objects of the successive frame. Its perfect match is determined as the moving object corresponding to the maximum correlation coefficient value. The same matching process is applied to all human objects and all frames. So, the entire video tracking process is modelled as following:

$$\mathrm{match}_i(j) = \arg\max_{k \in [1,K]} \frac{\sum_x \sum_y \left(P_j^i(x,y) - \mu(P_j^i)\right) \cdot \left(P_k^{i+1}(x,y) - \mu(P_k^{i+1})\right)}{\left(\sum_x \sum_y \left(P_j^i(x,y) - \mu(P_j^i)\right)^2\right) \cdot \left(\sum_x \sum_y \left(P_k^{i+1}(x,y) - \mu(P_k^{i+1})\right)^2\right)}, \tag{9}$$

$$\forall i \in [1, n],\ j \in [1, K].$$

Obviously, any moving person tracked in this grayscale video stream is modeled as an object sequence $\{P_j^1, P_{match_1(j)}^2,..., P_{match_i(j)}^{i+1},..., P_{match_{n-1}(j)}^n\}\Big|_{j\in[1,K]}$. Then, the corresponding instances of this person in the frames of the initial colored movie are easily detected and marked accordingly. Another human matching solution has also been tested. Thus, one can use the edges of the P_j^i image objects instead of these objects in the correlation procedure expressed by (9). We have detected the edges of these sub-images using a Canny filter [26], then successfully tested the obtained edge images in the matching process. Unfortunately, computing the edge information raises substantially the complexity of the video tracking process and its execution time.

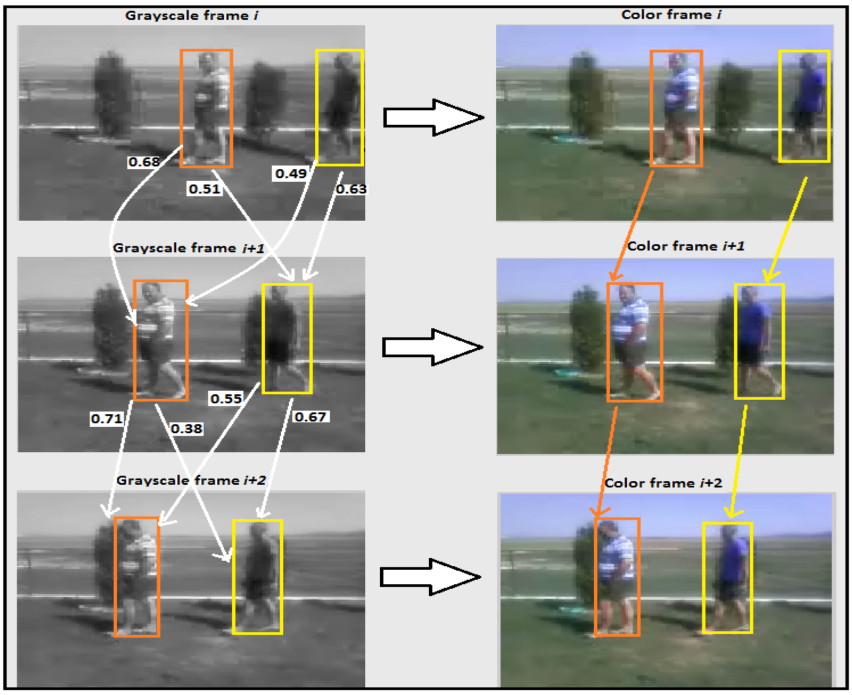

Fig. 4 – Human tracking example.

A human person tracking example is represented in Fig. 4. The proposed object matching algorithm is applied on the same sequence of the three consecutive video frames analyzed for pedestrian detection. On the left column of the figure, one can see the identified human objects, two for each grayscale frame, bounded by colored rectangles. The human matching process is represented by arrows linking objects of different frames and marked by the computed 2D correlation results [25]. The arrows corresponding to the higher correlation values indicate the correct matches. On the right column of Fig. 4, there are displayed the final detection and tracking results: the identified and tracked moving persons in the initial color frames. As one can see in the pictures, instances of the same pedestrian are marked with the same color and linked by same-color arrows.

4. EXPERIMENTS

The pedestrian identification and tracking approaches proposed in this article have been tested on numerous video datasets. One has performed various detection and tracking experiments on video sequences containing people and obtained satisfactory results. The provided techniques produce quite high detection and tracking rates. Thus, the proposed human identification approach achieves a person detection rate of approximately 80%. Our temporal differencing based moving object detection technique has a much higher rate (over 90%), but this rate is lowered by the proposed skin recognition performance.

The described skin segmentation technique produces a robust identification of the skin-colored regions. But because our approach does not perform a semantic analysis, it is possible to identify as skin some regions representing skin-colored image objects, such as parts of clothes. Also, there have been obtained satisfactory values for performance parameters of *Precision*, *Recall* and F_1. Our human detection approach produces few missed hits, meaning undetected moving persons, and also few non-human objects detected as humans. The provided video tracking technique is also characterized by a high person matching rate that exceeds 80 percent. The performance parameters are *Precision* = 0.85 and *Recall* = 0.85, which means very few false positives and false negatives are obtained by the tracking method.

Some of our detection and tracking experiments are described in the previous figures (Fig. 1 to Fig. 4). We consider the following parameter values in the detection processes of our experiments: $T = 25$ for relation (2), $l = 2$ for structuring element. We have compared the performances of human detection and tracking approaches proposed here with detection and tracking performances of other known techniques, such as those based on Histograms of Oriented Gradients (HOGs) [6,27]. We have found that our detection and tracking system provide comparable good results for non-occluded side view pedestrian detection. Unfortunately, our human identification method performs worse than HOG-based approaches for front view pedestrian detection, rear view pedestrian detection and partially occluded pedestrian detection [27].

5. CONCLUSIONS

A novel automatic human detection and tracking system for fixed camera video sequences containing multiple moving objects have been proposed in this article. Our paper brings important contributions in both the human identification and video tracking research areas. First, we have proposed a new moving person detection technique. The moving objects of the video stream are identified using an improved frame-difference based algorithm and some morphological operations. Then, the detected objects satisfying some certain conditions are classified as humans. One of these conditions is the presence of skin regions in the analyzed image object, so a new color-based skin detection approach has also been provided. The identified people are tracked in the video sequence using an object matching technique. A cross-correlation based human matching approach is proposed in this paper. Numerous experiments have been performed using the proposed approach and satisfactory human detection and tracking results have been obtained.

As we have already mentioned, the person detection technique here does not work optimally for any human detection task. It performs successfully for detection of non-occluded side view pedestrians [28], but provides poorer results for occluded people, front view and rear view moving persons. Also, our approach is not well suited for non-pedestrian human detection and tracking. It can not successfully detect moving persons which are not in the upright position, because of the condition related to width to height ratio. Persons with no visible skin regions, such as masked and gloved humans, are also difficult to identify. They could be detected and tracked if the skin-related condition is removed, but more errors may also appear in this case, like some moving objects detected as humans. Also, the method does not identify the static persons from the movies at all and performs quite unsatisfactory for detection of people moving only parts of their bodies (like head, hands or feet). Because of the temporal differencing procedure used for video motion estimation, the proposed human identification technique is also very sensitive to camera movements. Even the presence of small camera motion could affect quite seriously the entire moving object detection process.

Therefore, our future research in this video analysis domain will focus on improving and extending the proposed techniques in the mentioned directions. So, we intend to make our detection approach work better in the case of partially occluded persons, front view and rear view walking people. Also, the human detection errors caused by unintended camera motions will be better addressed. Non-pedestrian human identification approaches will also be considered. The human detection and tracking results described in this article can be successfully applied in many important computer vision domains. Video indexing and retrieval, robotics, video surveillance and urban traffic monitoring are some of these application areas [29].

REFERENCES

1. C. WREN, A. AZARBAYEJANI, T. DARELL, A. PENTLAND, *Pfinder: Real-time tracking of the human body*, IEEE Trans. Pattern Analysis and Machine Intelligence, **19**, pp. 780–785, 1997.
2. N. A. OGALE, *A survey of techniques for human detection from video*, Department of Computer Science, University of Maryland, College Park; http://www.cs.umd.edu/Grad/scholarlypapers/papers/neetiPaper.pdf
3. W. ABD-ALMAGEED, M. HUSSEIN, M. ABDELKADER, L. DAVIS, *Real-Time Human Detection and Tracking from Mobile Vehicles*, Proc of 2007 IEEE Intel. Transportation Systems Conf., Seattle, USA, Sept.30–Oct.3, pp. 149–157, 2007.
4. Y. LIU, A. HAIZHO, X. GUANGYOU, *Moving object detection and tracking based on background subtraction*, Proceeding of Society of Photo-Optical Instrument Engineers (SPIE), **4554**, pp. 62–66, 2001.
5. R. JAIN, H. NAGEL, On the analysis of accumulative difference pictures from image sequences of real world scenes, IEEE Trans. Patt. Analy. Mach. Intell. **1**, *2*, pp. 206–214, 1979.
6. N. Dalal, N. Triggs, *Histograms of Oriented Gradients for Human Detection*, Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), **1**, pp. 886–893, 2005.
7. W. SCHWARTZ, A. KEMBHAVI, D. HARWOOD, L. Davis, *Human detection using partial least squares analysis*, Proceedings of IEEE 12th International Conference on Computer Vision, 29 Sept. – 2 Oct. 2009, pp. 24–31.
8. C. P. PAPAGEORGIOU, M. OREN, T. POGGIO, *A general framework for object detection*, Proceedings of the 6th International Conference on Computer Vision, 1998.
9. X. HU, Y. TANG, Z. ZHANG, *Video Object Matching based on SIFT Algorithm*, Proceedings of International Conference on Neural Networks and Signal Processing, 2008, pp. 412–415.
10. D. H. BALLARD, *Generalizing the Hough transform to detect arbitrary shapes*, Pattern Recognition, **13**, *2*, pp. 111–122, 1981.
11. M. KASS, A. WITKIN, D. TERZOPOULOS, *Snakes: Active Contour Models*, International Journal of Computer Vision, pp. 321–331, 1988.
12. N. PETERFREUND, *Robust Tracking of Position and Velocity With Kalman Snakes*, IEEE Transactions on Pattern Analysis and Machine Intelligence, **21**, *6*, 1999.
13. C. BELEZNAI, B. FRUHSTUCK, H. BISCHOF, *Human Tracking by Fast Mean Shift Mode Seeking*, Journal of MultiMedia, **1**, *1*, pp. 1–8, 2006.
14. L. WIXON, *Detecting Salient Motion by Accumulating Directionally-Consistent Flow*, IEEE transactions on pattern analysis and machine intelligence, **22**, *8*, 2000.
15. Y. CHEN, Y. RUI, T. HUANG, *Multicue hmm-ukf for real-time contour tracking*, IEEE Transactions on Pattern Analysis and Machine Intelligence, **28**, pp. 1525–1529, 2006.
16. L. XIAOWEI, K. QING-JIE, L. YUNCAI, *A Feature Fusion Algorithm for Human Matching between Non-Overlapping Cameras*, Chinese Conference on Pattern Recognition CCPR '08, 2008, pp. 1–6.
17. T. BARBU, *An Automatic Face Detection System for RGB Images*, International Journal of Computers, Communications & Control, **6**, *1*, pp. 21–32, 2011.
18. T. BARBU, *Gabor filter-based face recognition technique*, Proceedings of the Romanian Academy (Series A), **11**, *3*, pp. 277–283, 2010.
19. M. COSTIN, T. BARBU, M. ZBANCIOC, G. CONSTANTINESCU, *Techniques for static visual object detection within a video scene*, Polytechnic Institute Bulletin, Iași, Automatics and Computers, Tom LI (LV), Fasc.1–4, pp. 75–85, 2005.
20. T. BARBU, *Novel automatic video cut detection technique using Gabor filtering*, Computer and Electrical Engineering, **35**, *5*, pp. 712–721, 2009.
21. T. BARBU, V. BARBU, V. BIGA, D. COCA, *A PDE variational approach to image denoising and restoration*, Nonlinear Analysis: Real World Applications, **10**, *3*, pp. 1351–1361, 2009.
22. T. BARBU, *A PDE based Model for Sonar Image and Video Denoising*, Analele Ştiinţifice ale Universităţii "Ovidius", Constanta, Seria Matematica, **19**, *3*, pp. 51–58, 2011.
23. J. SERRA, P. SOILLE, *Mathematical Morphology and Its Applications to Image Processing*, Proceedings of the 2nd International Symposium on Mathematical Morphology (ISMM'94), 1994.

24. M. M. FLECK, D. A. FORSYTH, C. BERGLER, *Finding naked people*, Proceedings of European Conference of Computer Vision, (ECCV), 1996, pp. 593–602.
25. J. P. LEWIS, *Fast Normalized Cross-Correlation*, Vision Interface, 1995.
26. J. CANNY, *A Computational Approach to Edge Detection*, IEEE Trans. Pattern Analysis and Machine Intelligence, **8**, *6*, pp. 679–698, 1986.
27. X. WANG, T. X. HAN, S. YAN, *An HOG-LBP human detector with partial occlusion handling*, Proc. of ICCV, 2009.
28. C. PAPAGEORGIOU, T. POGGIO, *A Trainable Pedestrian Detection system*, International Journal of Computer Vision (IJCV), pp. 15–33, 2000.
29. P. VANNOORENBERGHE, C. MOTAMED, J. M. BLOSEVILLE, J. G. POSTAIRE, *Monitoring pedestrians in a uncontrolled urban environment by matching low-level features*, Proc. IEEE Intl. Conf. on Systems, Man, and Cybernetics, **3**, pp. 2259–2264, Oct. 1996.

Received August 9, 2011