Wigner Distribution based Motion Tracking of Human Beings using Thermal Imaging

Chandrashekhar N. Padole[1] and Luís A. Alexandre[1,2]
[1] Dept. of Informatics, University of Beira Interior
[2] IT - Instituto de Telecomunicações,
Covilhã, Portugal

Abstract

Motion tracking has become an important technology in various applications like surveillance, non-cooperative biometrics, virtual reality, etc. We propose a novel approach for motion tracking using only thermal images. We exploited two types of data association (spatial and temporal data association) in the proposed approach in order to reduce the false decisions. In addition to these, Wigner distribution was applied to the difference image for reducing the fluctuation in threshold image in terms of false object detections. The results obtained with this algorithm when applied to six image sequences of about 2500 images each, show its robustness for person tracking in thermal images.

1. Introduction

The purpose of this paper is to present an approach for motion tracking of human beings using thermal imaging only. Recently the thermal imaging and applications based on it like motion tracking and face recognition has attracted attention especially from the image processing and vision community [1,2]. There are already some attempts to include thermal imaging in applications like face recognition and motion tracking in addition to the visible cameras, through data fusion. However, if only one of the modalities is used and reliable results are obtained with it, the system would benefit from cost saving and higher processing speed.

Earlier, the availability of infrared cameras was only for military applications. In applications like outdoor surveillance, where the background temperature is largely different from human beings, thermal imaging can play vital role in identifying and tracking persons. This fact has primarily motivated us to use only thermal imaging for motion tracking of human beings. Another advantage of thermal imaging lies in the fact that it cannot sense shadow or light illumination, which is normally the bottleneck in most of the motion tracking based on visible wavelenghts (see figure 1), thus making it more suitable for motion tracking in outdoor environment in day time as well as night time, where shadows and variable illuminations are dominating factors making tracking more difficult. On the contrary, clutters like cool body, variation in temperature across same subjects, blowing winds with different temperature gradients, person overlap while crossing each other, put challenges in thermal imaging and will have to be handled intelligently in order to obtain the efficient performance from motion tracking system using only thermal imaging.

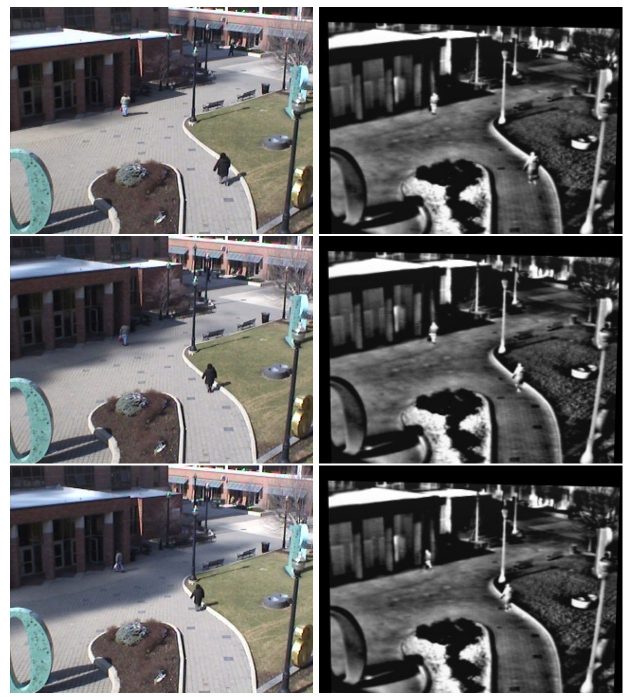

Figure 1. Shadow effects. Left: Visible camera. Right: Thermal Camera.

We propose a novel approach for motion tracking using thermal imaging. This algorithm does not use features of the objects to be tracked which reduces the computational

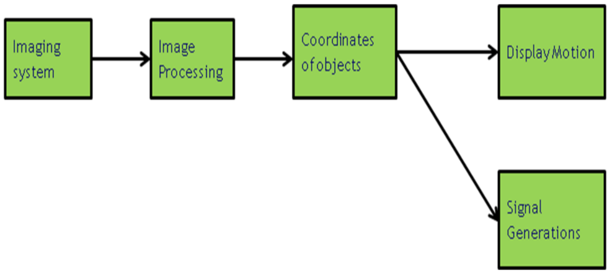

Figure 2. Typical motion tracking system.

complexity of the algorithm.

The remaining part of this paper is organized as follows. Section 2 presents an overview on motion tracking; section 3 discusses the Wiegner distribution; the following section presents the proposed methodology; section 5 contains the experiments and discussion and the final section presents the conclusions and future research directions.

2. Overview on Motion Tracking

The goal of motion tracking is to determine the position of objects across the frames from image sequences or video. Based on the position of imaging instrument, two types of motion tracking are a possible: static camera tracking or active motion tracking. Based on motion of objects, it is also classified in two categories: with markers (used in 2D and 3D animations) and marker-less (our interest). Depending on the methodology by which motion tracking can be initiated, there are, typically, two classes of motion tracking algorithms: recognition based tracking and motion based tracking. Depending on the type of motion to be tracked in human motion tracking, motion tracking applications can be placed in two categories: articulated motion and moving motion.

Multiple object tracking is required in most of the applications including surveillance. In literature, various motion tracking technologies are described, such as Kalman filter (KF), extended KF, particle filter, unscented KF, hidden Markov model, affine transform and Gabor transform.

The typical motion tracking system is shown in figure 2, which is normally composed of an imaging system, an image processing algorithm to give location of objects and a display device to show the object path and/or signal generation logic to detect the event.

The basic algorithmic approach is depicted in figure 3.

Object detection can be achieved with the following approach: image subtraction between successive frames; if required, morphological operators can be used; segmentation may be required.

Motion tracking using motion estimation involves the matching of a moving object in the surrounding of its location on the next frame. There are the following approaches

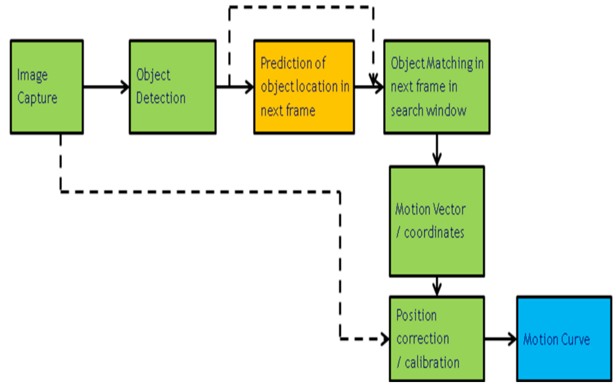

Figure 3. Algorithmic approach.

for finding the best matching unit [3,4,5]: pixel based (computationally complex), block based, region based and mesh based (triangle, hexagonal, content based).

The process of best matching unit to estimate the motion vector includes search based in similarity criterion across the pixels in close vicinity in next frame. This search window is typically 15x15 pixels hence it needs high computational power. In order to reduce this computational complexity there are various optimization techniques in the literature: 2-D logarithm search [6], TSS [7], NTSS [8], FSS [9], two step multiple local winners based [10] and conjugate direction search [11].

Kalman filter [12] and particle filter [13] are the two popular technologies in the field of motion tracking. These methods are widely used in the visible camera based motion tracking.

3. Wigner Distribution

Wigner distribution is a generalized time-frequency representation. The characteristic of the Wigner distribution, to be a function of both time and frequency, is remarkable. The Fourier transform, on the other hand, is strictly a function of frequency. Wigner [14], in 1932, proposed this function for the study of quantum mechanics. Ville [15] proposed it again in 1948. However, researchers did not pay much attention to the method until the 1980s, when researchers in the speech processing area, extensively used this concept.

The Wigner distribution of two signals $f(t)$ and $g(t)$, is defined as [16, 17]

$$WS_{f,g}(t,\omega) = \int_{-\infty}^{\infty} e^{-j\omega k} f(t+k/2) g^*(t-k/2) dk \quad (1)$$

where ω is the frequency, t is time and g^* is the complex conjugate of the function $g(t)$.

The auto-Wigner distribution of signal $f(t)$ is given by

$$WS_f(t,\omega) = \int_{-\infty}^{\infty} e^{-j\omega k} f(t+k/2) f^*(t-k/2) dk \quad (2)$$

The auto-Wigner distribution of a real function $f(t)$ is given by

$$WS_f(t,\omega) = \int_{-\infty}^{\infty} e^{-j\omega k} f(t+k/2) f(t-k/2) dk \quad (3)$$

In the discrete domain, the Wiegner function is defined as

$$WD_{f,g}(t,\omega) = 2 \sum_{k=-\infty}^{\infty} e^{-j\omega k} f(t+k/2) g^*(t-k/2) \quad (4)$$

and the respective auto-Wigner distribution is

$$WD_f(t,\omega) = 2 \sum_{k=-\infty}^{\infty} e^{-j\omega k} f(t+k) f^*(t-k) \quad (5)$$

For a real function f this comes as

$$WD_f(t,\omega) = 2 \sum_{k=-\infty}^{\infty} e^{-j\omega k} f(t+k) f(t-k) \quad (6)$$

This last equation is important for the development of the Wigner distribution for image processing.

The inclusion of a window in the Wigner distribution defines the pseudo-Wigner distribution in order to reduce computation. For real and discrete function $f(t)$ with a window w of duration $2d+1$, this distribution is given by

$$PWD_f(t,\omega) = 2 \sum_{k=-d}^{d} cos(2\omega k) w(k) f(t+k) w(-k) f(t-k) \quad (7)$$

To use the Wigner distribution function for image processing, it is extended to two-dimensional space. Such an extension results in a four-dimensional Wigner distribution function. The function has two space-domain variables x and y, and two frequency-domain variables, u and v. The extension to 2D space is then

$$WD_{x,y,u,v}(t,\omega) = \frac{4}{MN} \sum_{l=-N'/2}^{N'/2} \sum_{k=-M'/2}^{M'/2} cos(\theta) f(x+k, y+l) f(x-k, y-l) \quad (8)$$

where the image size is $M \times N$, the window size is $M' \times N'$, f is the gray-scale function and $\theta = 4\pi(uk/M + vl/N)$.

For the pseudo-Wigner distribution, the main Wigner kernel is multiplied by another kernel like the exponential, $e^{-\lambda||(k,l)||}$. Equation (8) becomes

$$WD_{x,y,u,v}(t,\omega) = \frac{4}{MN} \sum_{l=-N'/2}^{N'/2} \sum_{k=-M'/2}^{M'/2} e^{-\lambda||(k,l)||} cos(\theta) f(x+k, y+l) f(x-k, y-l) \quad (9)$$

The main objective behind using this kernel is to give the candidate pixel (x, y) where WD is being calculated the maximum influence on the calculations, whereas as one goes away from the candidate pixel, its influence should rapidly decay.

To enhance the image range, equation (9) can be scaled by a multiplicative gain factor α.

Refer to [19] for the in depth study of wigner distribution for image processing. Equation (9) gives the first order Wigner distribution. The second order distribution is obtained by applying the first order Wigner distribution to the output of the the first order Wigner distribution of the image.

4. Methodology

This work has been inspired from tracking done by natural eyes in thermal imaging. In the absence of features, the eye can do tracking in images captured by thermal camera. For this purpose, we exploited two types of data association: spatial data association and temporal data association. The flow chart of motion tracking algorithm is depicted in figure 4 and several images representing different steps of the algorithm are presented in figure 5.

4.1. Spatial data association

First, we calculate the difference image between the current and previous frame. The Wigner distribution is applied to the difference image. We have observed that motion tracking over the difference image is more unstable (fluctuating) than that over the Wigner image of the difference image. Then the Wigner image is thresholded to get a binary image. The foreground object map is created by giving a different label to each object using 8-neighbourhood connectivity.

However, because of forementioned clutters available in thermal images, many false positive objects appear in the thresholded image. Each object is thresholded for size of the object in pixels. This step outputs the objects which are moving and are of sufficiently large size.

Due to non-uniform temperature within the same object, the same person may exist in more than one blob. To connect these blobs, we use vertical and horizontal threshold of distances between two nearby blobs. Due to vertical shape of person moving, vertical threshold was kept larger than horizontal one. This also helps in separating two persons close by moving in same direction. This step outputs the same label to closeby blobs indicating close blobs are of the same person.

4.2. Temporal data association

We used temporal buffer to apply the temporal data association on the object yielded by previous step. Each object

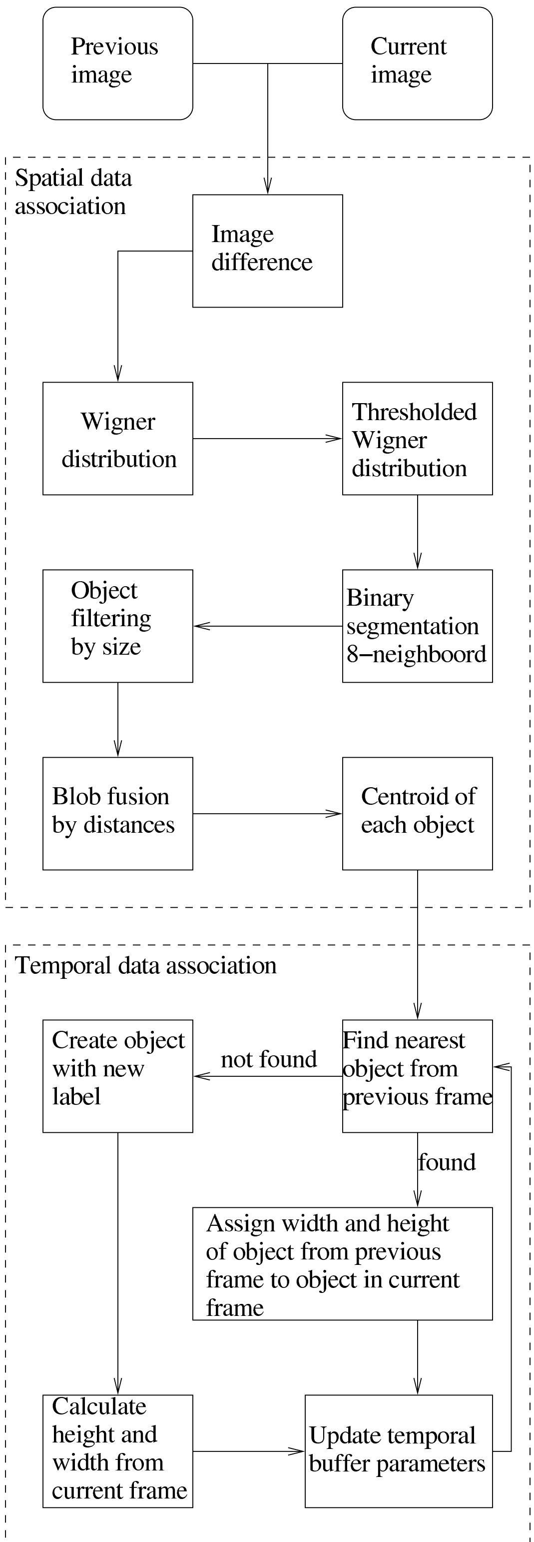

Figure 4. Flow chart for the proposed method.

is evaluated for its centroid. This centroid is tested with existing objects in previous frames for their closeness.

If an object from the previous frame with distance less than threshold is found, its label is assigned to the object from the current frame indicating that the object from the previous frame is now at a new position given by its centroid. The object width and height from the previous frame are also assigned to the object in the current frame. The temporal buffer is updated with the new object location.

In case the nearest object has a distance of more than a pre-defined threshold then an object with a new label is created indicating the entry of new person. Its width and height are calculated from the current frame. The temporal buffer is updated with this new object parameters. Then, all the objects in current frame are stored in the temporal buffer.

5. Experimental Results and Discussion

The algorithm was developed in MATLAB and applied on the database available on OTCBVS website [18]. We used the third database, namely, OSU Color and Thermal Database. This database includes six image sequences, each of around 2500 frames (exact size varies from sequence to sequence). The first three sequences are from one location with pedestrians. The remaining sequences are taken from a different location.

The difference image is passed through the Wigner distribution. It can be easily observed from figure 5 that the dynamic range of the diffrence image is improved after the Wigner distribution. This is because of the product term from equation (9) which is reponsable for reducing high frequency noise component and enhancing local contrast of the image, where the object is present [16].

Wigner image was thresholded with threshold value of 0.3. The threshold of size for filtering out false objects was 15 pixels. Depth of temporal buffer was 4. Vertical and horizontal distance thresholds were 20 and 10 respectively. Minimum distance for object to be present in next frame as same object was 5 pixels in any direction. All these threshold parameters were kept constant for all the sequences. All the images are 8-bit gray scale.

The following discussion will be on the experimental results obtained in different sequences. It can easily observed that our algorithm has detected objects even when they are not easily visible with human eye. All images are in figure 6.

We applied our approach to all the six sequences available in dataset. The two images in the first row of figure 6 show the a result obtained on the first frame sequence. In this frame, two persons are moving and third person is stopped, which is very clear from the results obtained with our algorithm. In the second row, all three persons are moving, which has been captured by our algorithm. Similar situation with many people can be verified in the third row. In the forth row, it can be observed that our algorithm is

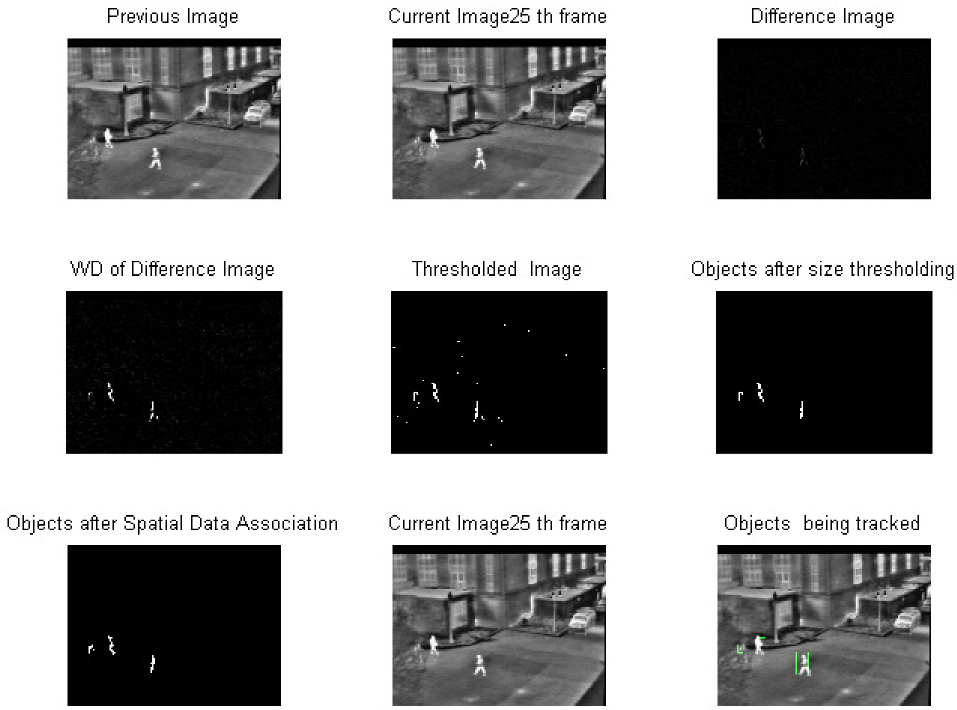

Figure 5. Sample images from several steps of the proposed method.

capable of identifying the movement of multiple objects located at far distance from camera. The fifth row shows the strength of this approach by identifying the movement even behind a tree, which is not visible to humans easily. The six row shows a critical situation: the changing temperature environment. Our approach was able to recognize the true positive object movement. In the last row, though person and vehicle (parked) have the same intensity, our algorithm has detected the moving person.

6. Conclusion and Future Scope

In this paper we proposed a new method for person tracking using thermal images. The method is based on a Wigner distribution modelling of the interframe difference image and the conjungation of spatial and temporal data associations. The proposed algorithm has small computational complexity and is thus quite fast. We presented results on several images from 6 sequences with around 2500 images each. The results obtained with this algorithm show the efficiency and errorless performance of person tracking in thermal images.

Future work includes the automatic adaptation of the threshold parameters. Also more image sequences will be tested. More robust data association rules will be searched for in order to improve performance in more challenging image sequences.

7. Acknowledgement

We acknowledge the financial support given by ”FCT - Fundação para a Ciência e Tecnologia” and ”FEDER” in the scope of the PTDC/EIA/69106/2006 research project ”BIOREC: Non-Cooperative Biometric Recognition”.

References

[1] Lawrence B. Wolff, Diego A. Socolinsky, and Christopher K. Eveland, Chapter 6, “Face Recognition in the Thermal Infrared”, Computer Vision Beyond the Visible Spectrum, Advances in Pattern Recognition, Springer, 2005

[2] E. Herrero, C. Orrite, A. Alcolea, A. Roy, J.J. Guerrero, C. Sages, “Video-Sensor for Detection and Tracking of Moving Objects” IbPRIA, Pattern Recognition and Image Analysis, LNCS 2652, pages 346-353, 2003.

[3] Van Beek, P.J.L.; Tekalp, A.M.; Puri, A.; “2-D mesh geometry and motion compression for efficient object-based video representation”, Int. Conf. on Image Processing, 1997. Volume: 3, 1997 , Page(s): 440 - 443 vol.3

[4] Altunbasak, Y.; Murat Tekalp, A.; Bozdagi, G.; “Two-dimensional object-based coding using a content-based mesh and affine motion parameterization” Int. Conf. on Image Processing, 1995. Proceedings, Volume: 2 ,1995 , Page(s): 394 - 397 vol.2

[5] Badawy, W.; Bayoumi, M.; “A mesh based motion tracking architecture” The 2001 IEEE Int. Symposium on Circuits and Systems, 2001. ISCAS 2001. Volume: 4, 2001 , Page(s): 262 - 265

[6] Jain, J.; Jain, A.; “Displacement Measurement and Its Application in Interframe Image Coding”, IEEE Trans. on Communications, Volume: 29 , Issue: 12, 1981, Page(s): 1799 - 1808

[7] http://www.ece.cmu.edu/~ee899/project/deepak_mid.htm

[8] Reoxiang Li; Bing Zeng; Liou, M.L.; “A new three-step search algorithm for block motion estimation” IEEE Trans. on Circuits and Systems for Video Technology, Volume: 4 , Issue: 4 1994, Page(s): 438 - 442

[9] Lai-Man Po; Wing-Chung Ma; “A novel four-step search algorithm for fast block motion estimation” IEEE Trans. on Circuits and Systems for Video Technology, Volume: 6 , Issue: 3 , 1996, Page(s): 313 - 317

[10] Hsien-Hsi Hsieh; Yong-Kang Lai; “A novel fast motion estimation algorithm using fixed subsampling pattern and multiple local winners search” The 2001 IEEE Int. Symposium on Circuits and Systems, 2001. ISCAS 2001. Volume: 2 Page(s): 241 - 244 vol. 2

[11] Srinivasan, R.; Rao, K.; “Predictive Coding Based on Efficient Motion Estimation” IEEE Trans. on Communications, Volume: 33 , Issue: 8 ,1985 , Page(s): 888 - 896

[12] C. Stauffer and W. E. L. Grimson. Learning patterns of activity using real-time tracking. IEEE Trans. on Pattern Analysis and Machine Intelligence, 22(8):747-757, 2000.

[13] M. Isard and A. Blake,“CONDENSATION-Conditional Density Propagation for Visual Tracking”, Int. Journal of Computer Vision 29(1), 5-28 (1998)

[14] E. Wigner, “On the quantum correction of thermodynamic equibilibrium”, Phys. Rev 40, 1932,749-759.

[15] J. Ville, “Theorie et applications de la notion de signal analytique”, Cables et Transmission 2eA 1, 1948, 61-74.

[16] Padole, C.N.; Vaidya, V.G.; “Image restoration using Wigner distribution for night vision system” 9th Int. Conf. on Signal Processing, 2008. ICSP 2008. Page(s): 844 - 848

[17] Vaidya, V.G.; Padole, C.N.; “Night vision enhancement using Wigner Distribution”, 3rd Int. Symposium on Communications, Control and Signal Processing, 2008. ISCCSP 2008. Page(s): 1268 - 1272

[18] Davis, J., Sharma, V.: “Background-subtraction using contour-based fusion of thermal and visible imagery”. Computer Vision and Image Understanding 3(2-3) (2007) 162-182

[19] Vaidya, V.G.; ”The use of generalized space frequency representation for motion estimation from noisy image sequences”, PhD Thesis, Electrical Engineering Dept., University of Washington, Seattle (1992) 162-182

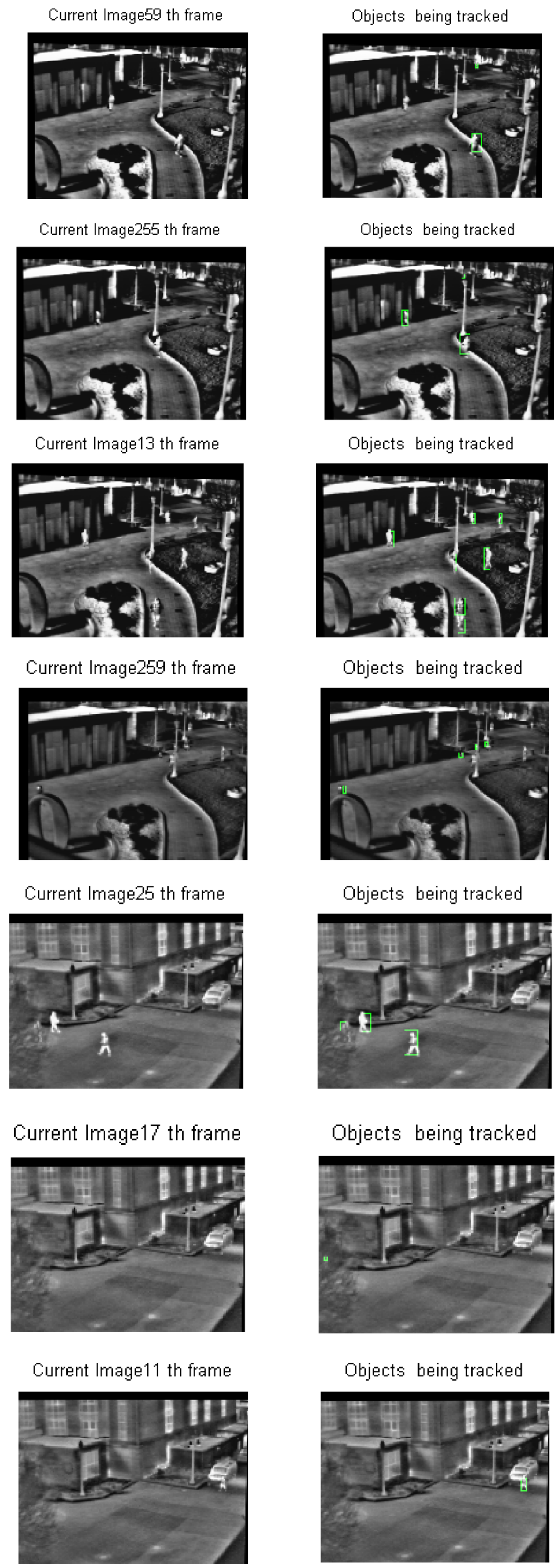

Figure 6. Samples from dataset sequences. Please see text for details