Sensing Human Behavior in the Built Environment

Research on Behavior Visualization Based on Thermal-Infrared Photography and Image Processing Technique

Zhe GUO[1], Xiang WANG[2], Philip F. YUAN[3]

[1] Tongji University
1732143@tongji.edu.cn
[2] Tongji University
18310021@tongji.edu.cn
[3] Tongji University
philipyuan007@tongji.edu.cn

Abstract. This paper shows a new application of infrared photography technique in human behavior sensing situated in outdoor built environment. By building a system integrated behavior thermal-infrared images acquisition and processing, the characteristic of city pattern and human behavior related to that certain environment can be captured by the infrared camera equipped auto-control unmanned aerial vehicles (UAVs) and be displayed in a processed visualization interface. By exploring a more efficient, smart and accurate method of collecting high spatial and temporal resolution data, a situated and context-aware behavioral visualization workflow can be developed which inform the behavior related environmental literacy in different culture of the architects and urban designers in order to reveal hidden patterns in the cities.

Keywords: Behavioral visualization; Thermal-Infrared Photography; Application of Auto-control UAVs; Image Processing Technique

1 Introduction

1.1 Traditional method in behavior research

There is a huge correlation between human behavior and environment. In a large scale of landscape design, especially some open public space design in city, designing is intertwined by many factors, among which user's behavior plays a important role. According to Altman I [1], certain environment lead to a specific behavior, meanwhile by analysing crowd behavior characteristics the internal connection between environment and behavior can also be revealed [2]. Although some empirical-oriented traditional guidelines have been proposed for designing urban public space, several usability problems can be identified only by studying the behavior of real users in virtual environments. The study of the behavior pattern and its characteristics would also help architects to truly understand the specific demand of the behavior to the space.

The term "behavioral sciences" was officially named in an interdisciplinary conference in Chicago, USA in 1949 [3]. At the same time, building related environmental behavior research also emerged. Some psychologists in that period explored the real built environment layout and space arrangement had a significant impact on people's behavior. After 1960s, architects gradually joined this field of research. The most representative of these researches is Kevin Lynch's book *The Image of City* [4] which describes the cognitive style of environmental sites and behavior and bounds human psychology and emotion with certain built environment. This part was later summarized by Gary T. Moore [5] and extended to an interdisciplinary research branch. Moore pointed out that there is a large deviation between the behavioral cognition and the real situation of built environments. Before the development of digital technology, Space Syntax [6] and Behavior Mapping [7] are the most famous and wide-used method studying human behavior in traditional research, which pave the way for the subsequent quantitative research of behavior to a certain extent. Nevertheless, the traditional behavior research methods are restricted by many factors, such as sample size and data accuracy.

1.2 Behavioral sciences research in the digital age

Some advanced technologies are applied in dealing with collecting and processing behavior characteristic in this digital age. According to the different measurement scenarios, data collection methods are varied. In some researches of indoor behavior visualization, some positioning technique such as RFID, Wi-Fi, Zigbee and UWB [8] are commonly used and suitable for high precision and small scope condition. In the research of outdoor large-scale behavior data collection, mobile GPS positioning devices are chosen, such as mobile phones with GPS function [9]. Besides, due to the Location Based Service (GSM network) provided by the communication company the behavior condition could also be tracked and recorded [10]. However, these frequently-used outdoor positioning tools have their unavoidable limitations. Participants need to wear the corresponding GPS positioning module, which not only causes the effect of the subject, but also requires the time-consuming preparation of the equipment in the early stage. In addition, it is more difficult to obtain real-time cellular data which need to be supplied by third-party providers.

Based on the above discussion, this study aims to propose a new behavior data acquisition and analysis method by applying special photographic techniques. Behavior detection (or visualization) methods based on video or images are not novelty in fields of video surveillance domain in which the gray scale [11,12] and color camera [13,14] are most widespread applied. However, using the visible-light information is problematic when facing sudden changes in lighting or illumination problems, also can not quickly identify pedestrians' behavioral characteristics. In comparison thermal-infrared images have a number of distinctive features compared to frames acquired by a visible-light spectrum camera [15-18].

The rest of the article is organized as follows: Section 2 introduce tools used in behavior detection. In Section 3 several algorithms aiming analyze thermal-infrared image commonly applied in image processing are described under different situation depending on what behavioral characteristic is supposed to excavate. This way it is possible to reveal some hidden pattern of outdoor human behavior in certain built environment. Finally, some conclusions are provided in Section 4.

2 Tools for behavioral data sensing

2.1 Method and workflow

The whole workflow designed in studying human outdoor behavior consists of three parts: Environmental point cloud simulation model building; On-site infrared images acquisition through Unmanned Aerial Vehicle (UAV); Visualization of behavior characteristics through image processing algorithm (see Fig.1).

With the gradual maturity of UAV and sensor technology, the collection of high spatio-temporal resolution data is becoming much more feasible. Locating crowds movement information from high altitude by infrared camera carried on UAV is an efficient way to position outdoor human location for its high accessible in space and less limitations in operation. In this research an image record and data transmission system is assembled to a one UAV which can be automatically drove by Pixhwak autopilot hardware through Mission Planner control software. Through the GPS position module connect to Pixhwak, the UAV can achieve better hovering condition in the air to ensure that the infrared camera can obtain stable images data. These acquired thermal-infrared images will be passed back to ground control computer via image transmission equipment and could be processed in near real-time.

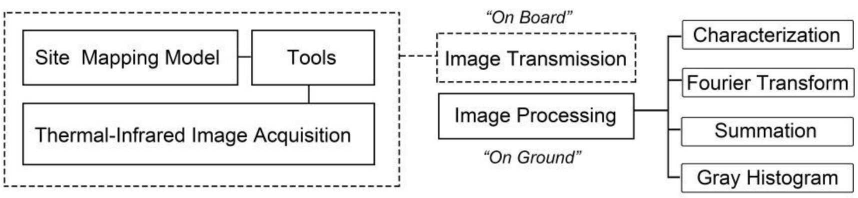

Fig. 1. Workflow

2.2 Usage of UAVs technique

Civilian use of UAVs has become increasingly common in recent years. Improvements in airframe design and electronics, particularly the mass production of comparatively inexpensive miniaturised inertial and positioning sensors, has enabled the application of UAVs to many and varied tasks. One area of growth in the scientific community has been the use of UAVs for Environmental Remote Sensing (ERS) where high spatial and temporal resolution, the ability to fly on-demand, and data collection from multiple sensors offers substantial advantages over traditional techniques. For small regions, Micro-UAVs (MAVs), typically weighing less than 5 kg with flying duration of up to 30 minutes, present an excellent option for collecting the required remotely sensed data for understanding environmental processes that require high resolution (centimeter scale), multi-sensor data.

Four rotors UAV is the most commonly used small autopilot, whose wheelbase is between 350 cm and 550 cm. With the maturity of open source flight control firmware in recent years, this study intends to use self-assembled UAV as the carrier of sensor flight for its further development ability. A complete UAV is consists of two parts, power and control; four brushless motors provide upward thrust through forward and backward steering, Pixhawk based flight control firmware assure adjustment of attitude angle, moving speed and hovering height. FLIR VueTM Pro 640 UAV special camera is applied in this research can be easily disassembled and replaced as a whole that is separated form the main body of UAV(see Fig.2).

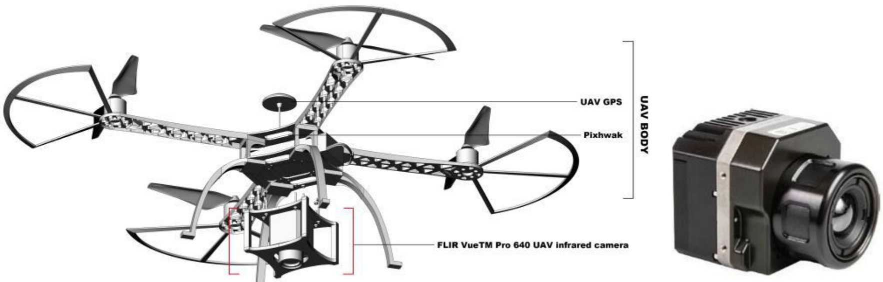

Fig. 2. UAV and infrared camera assembly

2.3 Usage of thermal-infrared photography technique

The operational principle of infrared camera is using infrared thermal imaging technology to detect specific infrared band signal of object thermal radiation, and converts the signal into images and graphics that can be distinguished by human vision. If the surface temperature of an object exceeds absolute zero, it will radiate electromagnetic waves. The electromagnetic waves whose wavelength ranges from 0.75 to 1000 microns are infrared rays, while the visible light of human vision ranges from 0.4 to 0.75 microns. The infrared spectrum has many interesting features which can be exploited for robust human detection. Human body emits constant thermal radiation in natural environment, which is different from non-living body, therefore infrared camera can be used to capture and record human motion sensitively. In previous studies infrared cameras have been used to study the behavioral characteristics of a single human body. S. Iwasawa et al.[19] introduce a new real-time method to estimate the posture of a human from thermal images acquired by an infrared camera. From this branch of usage of infrared camera, bunch of research in the field of human motion analysis are emerged [15-18]. Besides, detection and tracking of moving or stationary targets in FLIR imagery are also considered challenging research topics in computer vision [20].

3 Image processing technique in behavior visualization

3.1 Characterization of behavioral pattern from original infrared image

At the high altitudes infrared cameras works, the human body concentrates on emitting more heat than environment because their head are not covered. In thermal-infrared images, the human body presents a high-brightness spot of nearly circular shape compared to surrounding irregular objects. However, in the actual image acquisition process, some objects with high reflection coefficient in the built environment will reflect sunlight, which may interfere with the identification and segmentation of human infrared images. Distinguishing noise region between human body and environment will be solved in the next stage. Prior to that, the original infrared image needs to be characterized based on some basic image processing method.

The representation method of image is the basis of computational describing and processing algorithm. A two-dimensional image is usually summarized as a two-dimensional array $f(x, y)$ or a two-dimensional matrix $M*N$ (where M is the number of rows of image pixels and N is the number of columns of image pixels). The commonly used image expression methods are RGB image, Index image, Gray image and Binary image (see Fig.3). Among them, the gray image is black with 0 and white with 255, and the numerical value between them indicates the different degree of gray colour. The binary image is divided into matrices represented by 0 and 1 respectively according to the gray value of the pixel by setting threshold. Because the representation structure of pixel data in gray image and binary image is simple and clear, it would be used as the basis of a series of algorithms in subsequent image processing.

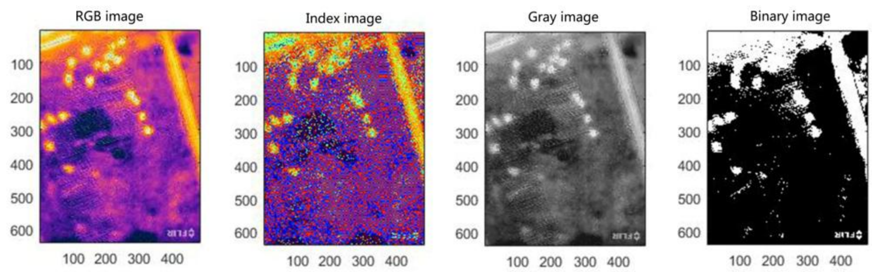

Fig. 3. Four image representation types

According to the bright spackle feature of human body in the infrared image, the shape characteristic can be quickly extracted by converting RGB image into Gray image when processing. Then we need to continue dilation of this Gray image. The function of dilation algorithm is to compute a certain area (line and point features) in the image, which visually enlarges the area from around to outside. Its operation can further highlight the characteristics of human punctuated areas. The operator of dilation algorithm defined as ⊕, dilation of A with B can be recorded as A ⊕ B, Expressed by the following formula (1):

$$\mathrm{A} \oplus \mathrm{B} = \{\mathrm{x} \mid [(\widehat{\mathrm{B}})_x \cap \mathrm{A}] \neq \Phi\} \quad (1)$$

Where $\widehat{\mathrm{B}}$ represents the reflection of set B, $(\widehat{\mathrm{B}})_x$ represents the displacement *x* of the reflection of set B,. It is shown in the above formula that the process of dilation A with B is to map B to its origin set and then translate it into *x*, where the intersection of A and B cannot be an empty set. The processed image see Fig.4. Image 4-(a) is the original gray image, (b) is the dilation translated image and (c) is the binary image based on (b). From this translated result the human behavioral pattern can be separated from the environment accurately.

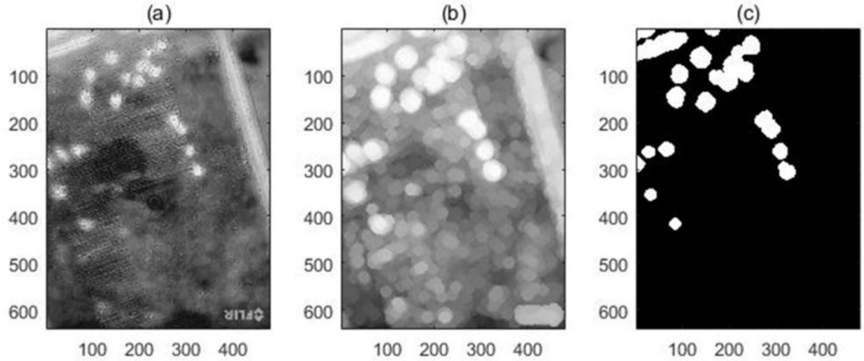

Fig. 4. (a) environment feature gray image; (b) dilation; (c) behavioral pattern segment in the binary image

3.2 Pattern feature location based on Fourier Transform

As can be seen from the Fig.4.(c), although the behavior features can be extracted explicitly and abstractly after image processing, there still be a part of environmental image noise regions that interfere with the accurate judgment of the behavior spackles. In this stage the Fourier transform algorithms are applied to solve this problem.

In the purely mathematical sense, Fourier transform is to transform an image function into a series of periodic functions; in physical sense, Fourier transform aims to transform an image from space-domain to frequency-domain, and then to reverse transform the image from frequency-domain to space-domain according to Fourier inverse operation. In other words, Fourier transform is to transform the gray distribution function of the image into the frequency distribution function of the image. A still digital image can be regarded as a two-dimensional data array, so the digital image processing is mainly two-dimensional data processing. One-dimensional DFT (Discrete Fourier Transform) and FFT (Fast Fourier Transformation) are the basis of two-dimensional discrete signal processing. Then two-dimensional discrete Fourier transform pairs are defined as formula (2-1) and (2-2):

$$F[f(x,y)] = F(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi\left(\frac{ux}{M}+\frac{uy}{N}\right)} \quad (2\text{-}1)$$

$$F^{-1}[(x,y)] = f(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} F(u,v)e^{j2\pi\left(\frac{ux}{M}+\frac{uy}{N}\right)} \quad (2\text{-}2)$$

In these above formulas: u, x = 0, 1, 2,..., M-1; v, y = 0, 1, 2,... N-1; x, y are space-domain variables; u, v are frequency-domain variables.

Fourier transform can be applied to describe image features. The correlation associated with convolution is introduced to locate the template with specific features. By this method, we can separate the human speckle maps near the circle shape from the environmental noise. The convolution of Fourier transform of images can be explained as follow: If the Fourier transform of function $f(x,y)$ is $F(x,y)$, the Fourier transform of function $g(x,y)$ is $G(x,y)$, then the image function $h_1(x,y) = f(x,y) * g(x,y)$, whose Fourier transform $H_1(u,v) = F(u,v) \cdot G(u,v)$; the Fourier transform of function $h_2(x,y) = f(x,y) \cdot g(x,y)$, whose Fourier transform $H_2(u,v) = F(u,v) * G(u,v)$.

Based on the above calculation principle, the human body feature speckle template and the image Fig.4-(c) containing human body characteristics are computed by Fourier transform. Then, the convolution of speckle template and Fig.4-(c) is calculated by fast convolution method, and the peak value of convolution operation is extracted; that is to say, the location result of the corresponding human body feature speckle parts in the image are separated. See Fig.5.

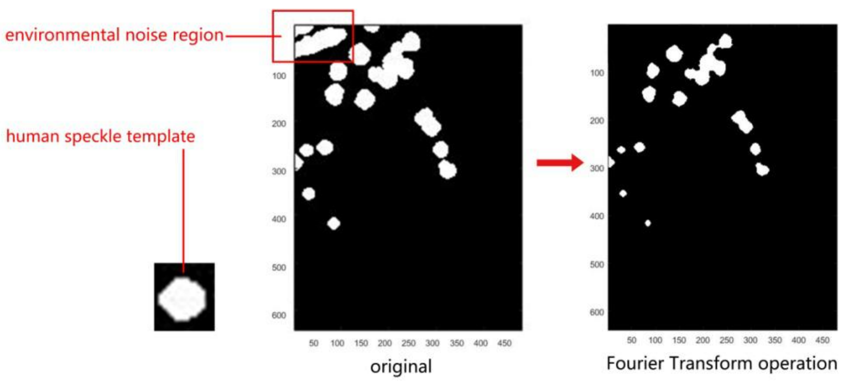

Fig. 5. Pattern feature location based on Fourier Transform

3.3 Summation of frame images

Some on-site experiments have been taken at the central square in public park located at one northern city of China. the images is recorded at the frequency of one frame per second. The final result is a series of static images that are difficult to see its correlation. In the processing stage, all the images are needed be summed together (the equation see formula.3), where *Pn* is an image matrix of equal size. By calculating, the image which is superimposed on each other representing all time nodes could be solved out (Fig.6).

$$P^{'} = \sum_{i=2}^{n} P_n(x_{n,} y_{n,}) \quad (3)$$

Fig. 5. Static images into summed result

From this operation result, pedestrians' preferences choice and staying time in a certain area of the environment can be clearly seen; besides, human behavior trajectory can also be clearly displayed in the processed images. The traditional experiments with a lot of tracking, positioning and data recording work can be replaced in an efficiently way. By analyzing the characteristics of behavior images collected in different experimental periods and using the visualization results of environmental data as the reference to the outdoor behavior pattern, we can intuitively see the general distribution trend of crowds' preferences for outdoor space area selection and the characteristics of selective staying of crowd in different periods of the day and environment. By comparing these two type of data, it can visually reveal the influence of environmental factors on certain behavioral trends. The data histogram shown in Fig.6. By comparing the gray value segments of three different time periods in a day, we can intuitively extract the trend of outdoor behavior changing and spatial preference with time.

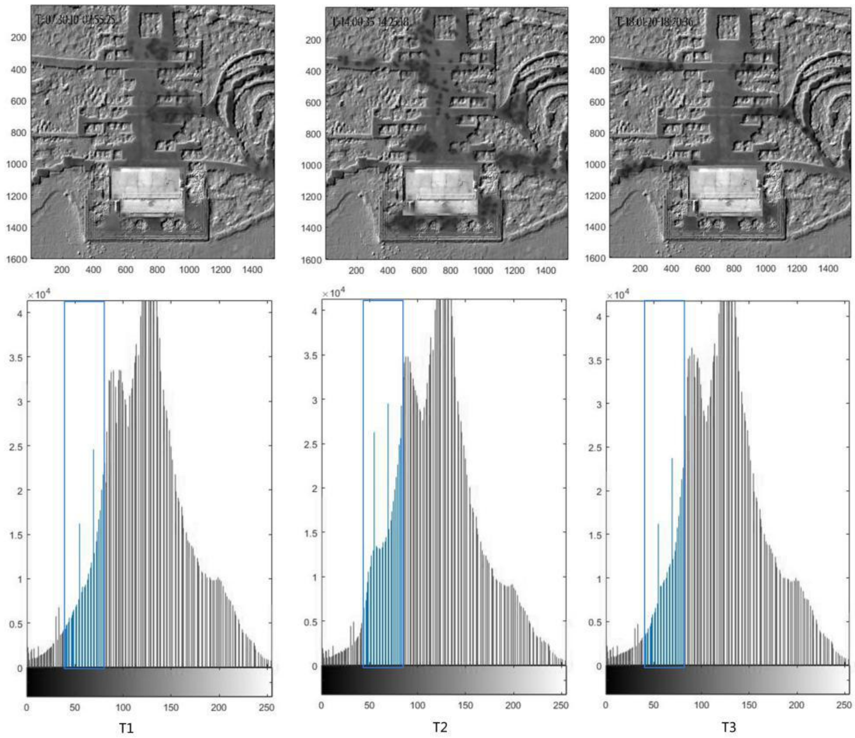

Fig. 6. Gray histogram of behavior changing features in different time

4 Conclusion

Behavior feedback is critical to outdoor built environments. Aided by UAV' high flexibility and controlled ability in urban high altitude sensing, human outdoor behavioral thermal-infrared image data shoot by infrared camera can effectively acquired which could provide big help for designers to discover the characteristics and relevance of spatial behavior and mine hidden factors that can not be obtained through traditional outdoor behavior research methods. Based on a series of image processing algorithms, macroscopic human behavior feature images can be abstracted and separated from environment pattern. These result would be used as the basis for further intervention in machine vision and machine in-depth learning research. After accumulating a large number of behavioral characteristics in different urban cultures and environmental backgrounds, these infrared image data through targeted processing could act as database samples for machine learning, to mine human behavioral preferences and trends in regions with different cultures. In further research, we will apply more efficient image processing algorithms to solve the problem of extracting behavior features from more complex environmental background and manage to find the entry point for docking machine learning.

Acknowledgements. This research is funded by National Key Research and Development Program (2016YFC0702104); National Natural Science Foundation Project (51578378); International Cooperation Project of Sino-German Science Center, National Natural Science Foundation of China (GZ1162); Project Topic of Shanghai Science and Technology Commission: Research on 3D printing technology of guideway robot concrete for intelligent building platform (18DZ1205604); Project of Shanghai Science and Technology Commission: Research on Building Digital Construction Technology Based on Intelligent Perception (17DZ1203405).

References

1. Altman I. The Environment and Social Behavior[M]// Environment and social theory /. Routledge, 1-14 (1975)
2. Harold, M., Ittelson, W.H. and Rivlin, L.G.,Environmental psychology: Man and his physical setting, Holt, Rinehart and Winston, New York (1970)
3. Isaac S, Michael W B. Handbook in research and evaluation: A collection of principles, methods, and strategies useful in the planning, design, and evaluation of studies in education and the behavioral sciences[M]. Edits publishers, (1995)
4. Kevin L. The image of the city[J]. Cambridge Massachussettes, (1960)
5. Moore G. T.,Knowing about environmental knowing: The current state of theory and research on environmental cognition[J]. Environment and behavior, 11(1): 33-70 (1979)
6. Hillier B. Space Is The Machine: A Configurational Theory Of Architecture[J]. Journal of Urban Design, (3):333-335 (1996)
7. Cosco N G, Moore R C, Islam M Z. Behavior mapping: a method for linking preschool physical activity and outdoor design[J]. Medicine & Science in Sports & Exercise,42(3): 513-519 (2010)
8. Zhao, Y., Guo, Z., Yin, H., Yao, J.i.a.w.e.i. and Yuan, P.h.i.l.i.p. F.,Behavioral Data Analysis and Visualization System Base on UWB Interior Positioning Technology, Proceedingsofthe23rdInternationalConferenceonComputer-AidedArchitecturalDesign Research in Asia (CAADRIA Conference), Beijing, 217-226.
9. Michael K, McNamee A, Michael M G, et al. Location-based intelligence – Modeling behavior in humans using GPS[J]. (2006)
10. Pei L, Guinness R, Chen R, et al. Human behavior cognition using smartphone sensors[J]. Sensors,13(2): 1402-1424 (2013)
11. Enzweiler, M.; Gavrila, D. Monocular pedestrian detection: Survey and experiments. IEEE Trans. Patt. Anal. Mach. Intell. 31, 2179–2195 (2009)

12. Fernández-Caballero, A., López, M.T., Saiz-Valverde, S. Dynamic stereoscopic selective visual attention (DSSVA): Integrating motion and shape with depth in video segmentation. Expert Syst. Appl.34, 1394–1402 (2008)
13. Fernández-Caballero, A., López, M.T., Castillo, J.M. Maldonado-Bascón, S. Real-time accumulative computation motion detectors. Sensors 10044–10065 (2009)
14. Schwartz, W.; Kembhavi, A.; Harwood, D.; Davis, L. Human detection using partial least squares analysis. In Proceedings of the IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 27 September–4 October. pp. 24–31 (2009)
15. Olmeda, D.; de la Escalera, A.; Armingol, J. Far infrared pedestrian detection and tracking for night driving. Robotica 29, 495–505 (2011)
16. Li, J.; Gong, W.; Li, W.; Liu, X. Robust pedestrian detection in thermal infrared imagery using the wavelet transform. Infrared Phys. Tech. 53, 267–273 (2010)
17. Fernández-Caballero A, Castillo J C, Martínez-Cantos J, et al. Optical flow or image subtraction in human detection from infrared camera on mobile robot[J]. Robotics and Autonomous Systems, 58(12): 1273-1281(2010)
18. Fernández-Caballero A, López M T, Serrano-Cuerda J. Thermal-infrared pedestrian ROI extraction through thermal and motion information fusion[J]. Sensors, 14(4): 6666-6676(2014)
19. S. Iwasawa, K. Ebihara, J. Ohya, S. Morishima, Realtime estimation of human body posture from monocular thermal images, in: Proceedings of the 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 15–20(1997)
20. A. Yilmaz, K. Shafique, M. Shah, Target tracking in airborne forward looking infrared imagery, Image and Vision Computing 21 (7) 623–635(2003)