駅改札シーンにおける人物行動把握のフレームワーク構築

ティティズイン[†] 上條 俊介[†]
東京大学生産技術研究所[†]

Abstract

In recent years, the increasing demand for safety and security in society have led to a growing need for surveillance activities especially in public transport areas such as airports, railway stations, underground and motorways. In surveillance systems, the development of abnormal event detection system that promptly detects events of violations, a dangerous action, and the crime, etc. is an important problem. This paper provides a conceptual framework for analyzing image sequences acquired from a security camera to detect, recognize, and understand human behaviors and activities arising out of the interactions of various objects, as well as their evolution over time to the environments. The emphasis is on establishing general model by using silhouette analysis, movement analysis, geometric and motion correlation representations for human behaviors.

1. Introduction

In modern civilization, the threats of theft, accidents, terrorists' attacks and riots are ever increasing. Due to the high amount of useful information that can be extracted from a video sequence, video surveillance has come up as an effective tool to forestall these security problems. Video surveillance is one of the fastest growing sectors in the security aspects due to its wide range of promising applications, such as intruder detection for shopping mall and important buildings, traffic surveillance in cities and detection of military targets, recognition of violent/dangerous behaviors [1].

Surveillance cameras are already prevalent in commercial establishments, while camera outputs are usually recorded in tapes or stored in video archives. Several intelligent activity/ event detection methods are being proposed but the behavior patterns of real life scenario still remain challenge for the research community. Some examples of human abnormal behaviors at public transportation areas that used in our experiments are shown in Fig. 1. Therefore our emphasis in this paper is to propose a general conceptual framework in the direction of building robust and intelligent surveillance system for human behavior analysis.

Conceptual Framework for Human Behavior Analysis at Station Ticket Zone
† Thi Thi Zin, Shunsuke Kamijo
The Third Department, Institute of Industrial Science, Univ. of Tokyo.

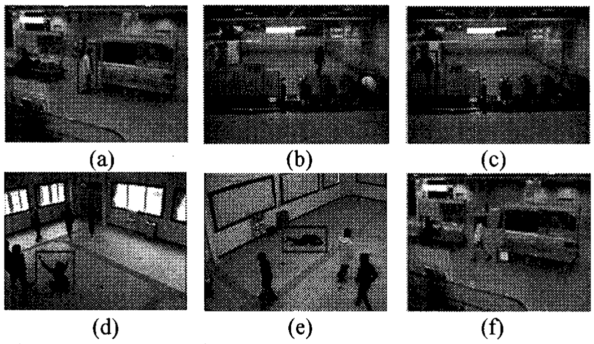


Fig. 1 Examples of human abnormal behaviors at public transportation areas: (a) loitering, (b) turnstile violation: crawling, (c) turnstile violation: hopping, (d) drunker, (e) lying down, and (f) abandoned object.

The rest of this paper is organized as follows. In section 2, we present the overview of proposed framework for human behavior recognition. In section 3, some experimental results carried out over the real image sequences are reported. Finally, in section 4, discussions and concluding remarks are described.

2. Proposed Conceptual Framework

The general framework of an automatic video surveillance system is shown in Fig. 2. Video cameras are connected to a video processing unit to extract low and high level information identified with alert situation. The main video processing stages include background modeling, object segmentation, object tracking, behaviors and activity analysis. Although tracking is now a usable technology, understanding what actions people are taking is still at its infancy. So, human action recognition has been a widely studied topic. Yet, the solutions to the problem are very premature and very specific to dataset at hand [2].

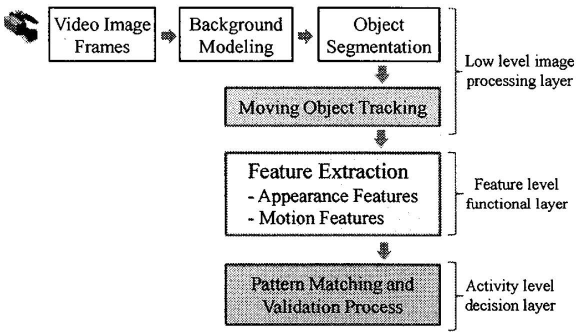


Fig. 2 Overview of the proposed general framework.

Thus the proposed conceptual framework will focus on analysis of human behaviors and activities. In order to do so we assume the existence of a tracking system, able to identify moving blobs in the scene (e.g. through background subtraction, with adaptive background models) and provide information regarding the position of the target over time. Conceptually, human behaviors are classified based on the number of moving objects (here human) and the stationary objects. In this paper we consider three main behavior classes namely: class of one moving object, class of more than one (group) and class of stationary objects involvement.

We then start feature extraction process based on two reasonably large sets of features. The first group of features includes the aspect ratio of the Bounding Box (BB) and the second order moments about Gravity Center (GC). The second subset of features includes the instantaneous position and velocity of the tracked subject. The target velocity and speed are obtained through differentiation of the instantaneous position estimate. Hence, we introduce different ways of averaging the velocity and speed estimates over an interval of T frames. For details, we define the following terminologies and notations. Some parameters are described in Fig. 3.

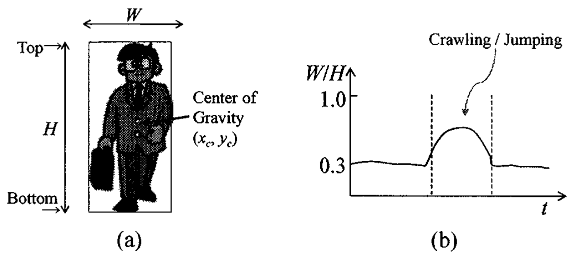


Fig.3 Some parameters used in the proposed algorithm: (a) detected individual region on background subtraction image, (b) illustration of aspect ratio on abnormal behaviors.

Bounding Box (BB) Properties

Position at time t: $p(t) = (x(t), y(t))$,

Width: $w(t)$,

Height: $h(t)$,

Aspect ratio: $A(t) = w(t)/h(t)$.

Moment Properties

Center of Gravity (CG): $CG(t) = (x_c(t), y_c(t))$,

2nd moment about CG along x axis: $m_{20}(t)$,

2nd moment about CG along y axis: $m_{02}(t)$,

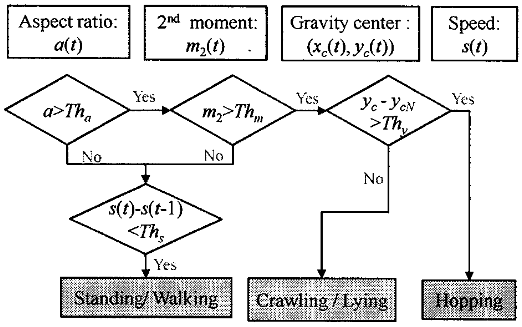
Moment Ratio: $m_2(t) = m_{20}(t) / m_{02}(t)$,

$$m_{20} = \sum_{(x,y)\in BB} (x - x_c)^2 f(x,y)$$

$$m_{02} = \sum_{(x,y)\in BB} (y - y_c)^2 f(x,y)$$

where $f(x,y) = 0,1$.

By using these features, we develop an algorithmic decision process composed of sets of rules to classify various types of human behaviors, actions and interactions. One example of the functional flow diagram for human pose estimation is shown in Fig 4.



y_{cN} = y position of gravity center at normal case

Fig. 4 Flow diagram for human pose estimation.

3. Experimental Results

In our experiments, the proposed method is applied to 6 types of video sequences taken in rail stations. It includes the abnormal behaviors namely: illegal tailgating (crawling, hopping), lying down, loitering, drunk clumsy walking, and abandoned objects. These concerned videos sequences were combined to produce objects tracked frames from which the sets of features are extracted. Then, the rule based algorithms developed in this work are applied for behavior classification process.

4. Discussion and Concluding Remarks

We have presented a simple rule based conceptual framework for monitoring human behaviors over time in video sequences, addressing the surveillance systems for protecting public transport assets. Our system correctly can detect target activities giving very promising results and indicate that there is much to be said in favor of simple methods even when the problem is complex. Although this approach does not cover all problems arising in the behavior analysis context, we suggest that simple classification rules such as those based on motion and appearance feature concepts, should be integrated with other methods such as Hidden Markov Models, possibly using a successive refinement strategy.

References

[1] Teddy Ko, "A survey on behavior analysis in video surveillance for homeland security applications," *37th IEEE Applied Imagery Pattern Recognition Workshop*, Washington, DC, USA, pp.1-8, Oct. 15-17, 2008.

[2] R. Kasturi, et al, "Framework for Performance Evaluation of Face, Text, and Vehicle Detection and Tracking in Video: Data, Metrics, and Protocol," *IEEE Trans. on PAMI*, vol. 31, no. 2, pp. 319-336, Feb. 2009.