Geng, Xin, Li, Gang, Ye, Yangdong, Tu, Yiqing and Dai, Honghua 2006, Abnormal behavior detection for early warning of terrorist attack, *Lecture notes in computer science*, vol. 4304, pp. 1002-1009.
.

Abnormal Behavior Detection for Early Warning of Terrorist Attack

Xin Geng[1], Gang Li[1,⋆], Yangdong Ye[2], Yiqing Tu[1], and Honghua Dai[1]

[1] School of Engineering and Information Technology,
Deakin University, VIC 3125, Australia
{xge, gang.li, ytu, hdai}@deakin.edu.au
[2] School of Information Engineering,
Zhengzhou University, Zhengzhou , China
yeyd@zzu.edu.cn

Abstract. Many terrorist attacks are accomplished by bringing explosive devices hidden in ordinary-looking objects to public places. In such case, it is almost impossible to distinguish a terrorist from ordinary people just from the isolated appearance. However, valuable clues might be discovered through analyzing a series of actions of the same person. Abnormal behaviors of object fetching, deposit, or exchange in public places might indicate potential attacks. Based on the widely equipped CCTV surveillance systems at the entrance of many public places, this paper proposes an algorithm to detect such abnormal behaviors for early warning of terrorist attack.

1 Introduction

After September 11, the global "War on Terrorism" has become one of the main challenges of our time. Although great efforts have been taken all over the world to protect innocent people from terrorist attacks, it still appears to be, as the US President George W. Bush appropriately labeled, "a long war against a determined enemy". Besides the political and military actions, new technologies in various areas are urgently demanded to ensure the victory in this war.

One of the most common terrorist attack patterns is to bring explosive devices hidden in ordinary-looking objects to public places. Recent examples are the London bombings (7 July 2005) and the Bali bombings (1 October 2005), both resulting in massive casualty. The early warning of such attacks is extremely difficult since a terrorist carrying a camouflaged bomb is not evidently different in appearance from an ordinary people. Thus it is almost impossible to detect in advance the potential danger of such attacks through conventional CCTV surveillance systems mounted in public places.

Although the isolated appearance is insufficient to distinguish terrorists from ordinary people, clues of abnormal behaviors could be discovered by considering the combination of a sequence of actions. During the procedure of explosive device preparation, delivery and the final attack, the members of terrorist group

⋆ Corresponding Author.

A. Sattar and B.H. Kang (Eds.): AI 2006, LNAI 4304, pp. 1002–1009, 2006.

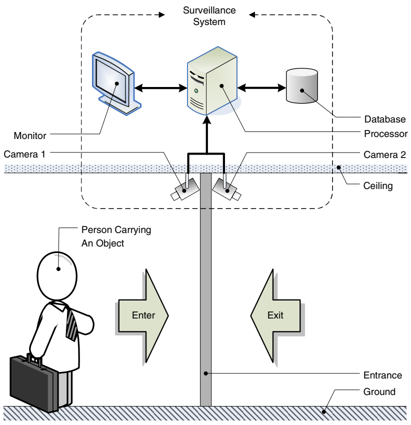


Fig. 1. Application Environment

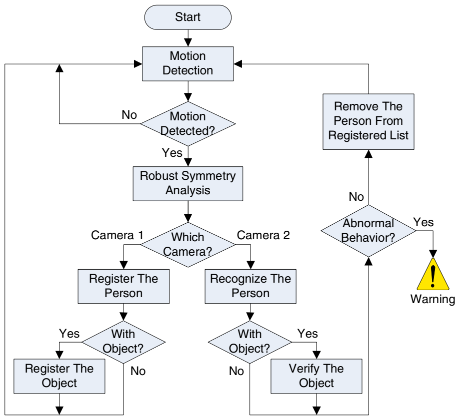


Fig. 2. Flow Chart of ROSY

Table 1. Combination of the Enter/Exit Actions and the Indicated Behaviors

Enter	Exit	Behavior	Code
Without Object	Without Object	Normal	N1
With Object	With the Same Object	Normal	N2
Without Object	With Object	Abnormal (Fetch)	A1
With Object	Without Object	Abnormal (Deposit)	A2
With Object	With the Different Object	Abnormal (Exchange)	A3

need to fetch, exchange and deposit objects of certain size. These behaviors are very rare in many public places, but are not easy to be discovered by security officers because it is almost impossible for human officers to remember every passing individual and the object he/she is carrying. In this paper, a surveillance algorithm named ROSY (RObust SYmmetry) is proposed to automatically detect such abnormal behaviors and warn people before potential terrorist attacks.

The rest of this paper is organized as follows. In Section 2, the ROSY algorithm is explained in detail. Then the experimental results are reported and analyzed in Section 3. Finally conclusions are drawn in Section 4.

2 The ROSY Algorithm

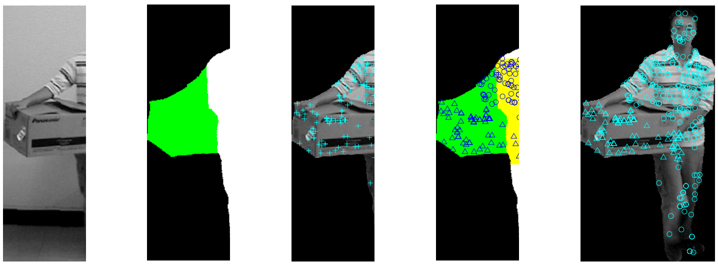
The application environment of ROSY is illustrated in Fig. 1. The algorithm is designed to work with monochromatic stationary video sources, either visible or infrared. Camera 1 is used to image the entering person, and Camera 2 is used to image the exiting person. Images from both cameras are sent to the processor installed with ROSY. Each entering person is registered in a database, and each exiting person is checked with the database. If any abnormal behavior is detected, the system will send a warning message to the security officer and display the suspicious behavior on the monitor.

The possible combinations of the entering and exiting actions, together with the indicated behaviors are tabulated in Table 1. Three of the five situations indicate abnormal behaviors, corresponding to object fetching, deposit and exchange, respectively. The behaviors are coded as N1, N2, A1, A2, and A3.

The flow chart of the ROSY algorithm is show in Fig. 2. The algorithm assumes that the background is relatively steady. Thus motion detection can be achieved through training a Gaussian model for each background pixel over a short period and comparing the background pixel probability to that of an uniform foreground model. The result of motion detection is the silhouette map of the foreground, see Fig. 3(b) as an example. If significant motion is detected, which means a person is passing the entrance, the foreground image is sent to Robust Symmetry Analysis (RSA) to determine whether the person is carrying an object and segment the human body and the object. The details of RSA will be described in Section 2.1. If the person is entering (captured by Camera 1 in Fig. 1), then the person and the object (if any) are registered into the database. If the person is exiting (captured by Camera 2), then the person is recognized and the object (if any) is verified. If any abnormal behavior (A1, A2, A3) is detected, the algorithm will raise a warning message. Otherwise, it will remove the record of the person from the database. The registration and recognition/verification of persons and objects will be further discussed in Section 2.2.

2.1 Robust Symmetry Analysis

The basic idea of ROSY is from the observation that the frontal view of the human body is approximately symmetric, and such symmetry will be violated if people carries an object. The problem is that, in many cases, body motion will also cause some body parts to violate the symmetry. Take Fig. 3(a) as an example, except for the object, the arms and legs are also not symmetric about the middle line of the body. So the algorithm must be able to distinguish whether the asymmetry is caused by an object or body motion. Backpack [1] did this through shape periodicity analysis. Unfortunately there are no sufficient periodic images in the scenarios of this paper. On the one hand, while people pass an entrance, they often do something, such as slow down walking speed, stop to identify themselves, and open the door, all of which will violate their usual walking periodicity. On the other hand, the entrance is often located at relatively narrow places, such like the entrance of a subway train, or a door by the corridor, where no periodicity can be detected since the person appears in the scene for only a short time. Thus Backpack is not suitable for the application here. Instead, ROSY achieves this by Robust Symmetry Analysis (RSA), based on as few as one image. There are mainly two steps in RSA. The first step is to analyze the symmetry of the silhouette, which narrows down the interested region for the second step, appearance symmetry analysis.



(a) (b) (c) (d) (e)

Fig. 3. Robust Symmetry Analysis. (a) Original image;(b) Symmetric axis and nonsymmetric region;(c) SIFT key points;(d) Key points classification;(e) Segmentation.

Silhouette Symmetry Analysis. Fist of all, the symmetric axis of the human body is estimated. Suppose the height of the silhouette is h, the algorithm only consider the highest $h/10$ and the lowest $h/10$. The horizontal middle point of the upper $h/10$ silhouette m_1 is regarded as one end of the symmetric axis. The lower $h/10$ silhouette might consist of several disconnected regions. The horizontal distance from m_1 to the horizontal middle point of each region is calculated, and the mean of those middle points within a certain distance to m_1 is regarded as the other end of the symmetric axis m_2. The estimated symmetric axis of the silhouette is shown as a red line in Fig. 3(b).

Suppose the estimated symmetric axis is $[m_1, m_2]$, p_l and p_r are a pair of pixels on the silhouette boundary such that the line segment between them perpendicularly intersects with $[m_1, m_2]$ at p_s. Let $d(p_1, p_2)$ denote the distance between p_1 and p_2. Then the symmetry of a pixel p_x on the line segment $[p_l, p_r]$ is determined by

$$sym(p_x) = \begin{cases} false & \text{if } d(p_x, p_s) > min(d(p_l, p_s), d(p_s, p_r)) + \varepsilon \\ true & \text{otherwise,} \end{cases} \tag{1}$$

where $sym(p_x) = false$ means p_x is a nonsymmetric pixel, and ε is a predefined small number. As an example, the nonsymmetric region in Fig. 3(b) is denoted by green. In order to make the algorithm more robust against image noise and the slightly asymmetry of human body, only those nonsymmetric region larger than a certain size (predefined minimum width, height and area) is considered.

Appearance Symmetry Analysis. In order to segment the object and the human body, the appearance of both the symmetric and nonsymmetric regions are compared with the other side of the symmetric axis. ROSY extracts the SIFT features [2] for image matching purpose. The SIFT feature positions of the foreground in Fig. 3(a) are shown in Fig. 3(c). Each SIFT key point is modeled by its position, scale and orientation so that scale and rotation invariance can be achieved. Experimental results [2] have shown that reliable recognition is possible with as few as 3 SIFT features. Thus even when a large portion of the object is occluded by other objects, it is still possible to be recognized. All of these advantages of SIFT make it suitable for body part matching, where scaling, rotation, and partial occlusion are all possible.

The SIFT matching is performed in a batch way, i.e. not only the single feature is matched to each other, but also those groups of features that agree to the same object pose are examined. Thus before matching, the SIFT features in both the symmetric and nonsymmetric regions should first be clustered into groups. In ROSY, the clustering is based on the positions of the SIFT key points and the texture of a small patch around the key point. In detail, each key point is represented by a triplet

$$\mathbf{t} = < x, y, E > \tag{2}$$

where (x, y) is the 2D coordinate of the key point and E is the entropy of the 9-by-9 neighborhood around the key point.

$$E = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i), \tag{3}$$

where z_i is a random variable indicating intensity, $p(z)$ is the histogram of the intensity levels in the 9-by-9 neighborhood, L is the number of possible intensity levels. Then the elements of $\mathbf{t}$ are normalized so that they have zero mean and unity standard deviation. Clustering is done by applying a graph based method called Normalized Cuts [3] on the normalized $\mathbf{t}$. The clusters in both the symmetric and nonsymmetric regions are matched to the other side of the symmetric axis. For each cluster, if at least 3 key points match with the other side, and they agree on the same pose, then it is regarded as part of the human body, otherwise of the object. An example is shown in Fig. 3(d), where the nonsymmetric region is green, the examined symmetric region is yellow, the positions of the key points classified as from body part are marked by circles, and those classified as from object are marked by triangles. It can be seen that most key points are correctly classified. Finally all SIFT key points outside the colored region is classified as from the human body. The result of RSA is shown in Fig. 3(e). If no key point is classified as from the object, then the person is not carrying an object.

2.2 Abnormal Behavior Detection

ROSY maintains a database to remember whoever entered the entrance, and whatever they are carrying with them. Each entry consists of four fields: *img*, *withObject*, *bodySIFT*, and *objectSIFT*, which respectively store the image, whether the person is carrying an object, the SIFT key points from the body, and the SIFT key points from the object.

When Camera 2 in Fig. 1 images somebody, i.e. somebody is exiting the entrance, ROSY first recognizes the person from the records in the database. This is achieved through matching the *bodySIFT* of the current image to those stored in the database. The record with the maximum number of matching key

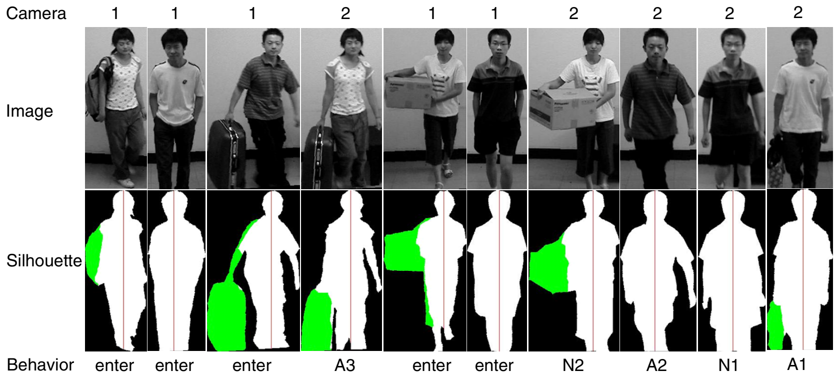
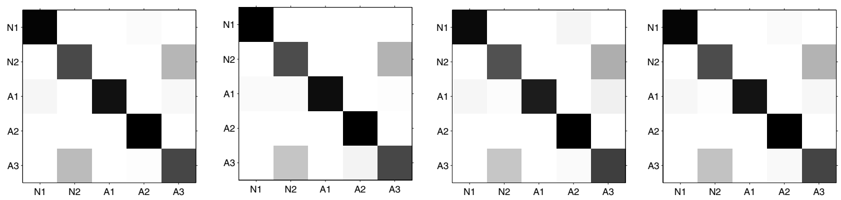


Fig. 4. Behavior detection in part of the random sequence



(a) (b) (c) (d)

Fig. 5. Confusion matrices. (a) Trial 1; (b) Trial 2; (c) Trial 3; (d) Average

points that agree on the same pose is regarded as from the same person. Then the entering and exiting profiles of the same person are compared with each other. Based on the field *withObject*, three behaviors (N1, A1, A2) in Table 1 is decidable. If the person entered and is exiting both with an object, then the object is further verified to see whether it is the same one that the person brought in. This is achieved by matching the *objectSIFT* of the two images, if there are more than 3 matching key points agree on the same pose, then the objects are the same (behavior N2), otherwise they are different (behavior A3).

3 Experiment

We have collected video clips from 20 different persons. Each person entered and exited a room in 6 different states respectively, among which 4 states are with an object (each time a different one), and 2 states are without object (one is walking normally, the other is walking with some variation). There are totally 4 different objects: a backpack, a handbag, a luggage, and a box. The real scenario at the entrance of public place is simulated by creating a random sequence of the video clips. Part of the sequence is shown in the second line of Fig. 4 (labeled

Table 2. Abnormal Behavior Detection Rate (DR) and False Alarm Rate (FAR)

Trial	1	2	3	avg.	std.
DR (%)	88.16	90.06	89.26	89.16	0.95
FAR (%)	5.30	3.65	4.90	4.62	0.86

with ‘image’). Three such random video sequences of length 1000 is generated to test the ROSY algorithm.

Part of the behavior detection results are shown in Fig. 4. The silhouettes of the corresponding images with non-symmetric regions marked by green are show in the line labeled ‘silhouette’. Note that only when the image is captured by camera 2 does the algorithm detect a behavior listed in Table 1. Otherwise the behavior is ‘enter’.

The performance of ROSY is illustrated as confusion matrices shown in Fig. 5, where rows represent the real behavior labels and columns represent the labels classified by ROSY. The intensity of each square indicates the probability of the behavior corresponding to the row being classified as the behavior corresponding to the column. The darker the square, the higher the probability. It can be seen that most behaviors in the three random sequences are correctly classified. The confusion mainly occurs between N2 and A3. This is because that there might be large difference in the view angle from the camera to the object and the object might be remarkably deformed. In such cases, automatic verification of the object is extremely hard. One way to solve this problem is to display any amphibolous behaviors between N2 and A3 on the monitor and let the security officer make the final decision.

As for abnormal behavior detection (A1, A2, A3), the detection rate (DR) and false alarm rate (FAR) of the 3 trials are tabulated in Table 2. It reveals that high detection rate and low false alarm rate can be achieved by ROSY. The average detection rate over all three trials is 89.16%, while the average false alarm rate is 4.62%. Moreover, the small standard deviation indicates the relatively steady performance of ROSY in different situations.

4 Conclusions

This paper proposes an abnormal behavior detection algorithm named ROSY for early warning of potential terrorist attack. Unlike previous work on abnormal behavior detection, ROSY aims to discover the abnormality concealed in a series of sub-behaviors, each of which alone is normal. This endows ROSY with the ability to detect behaviors like object fetching, deposit, and exchange, which is very rare in many public places and indicates potential danger of terrorist attack. Moreover, by using a novel technique called Robust Symmetry Analysis, ROSY can work on as few as one image, which makes it suitable to apply to the widely equipped CCTV systems at the entrance of many public places.

References

1. Haritaoglu, I., Cutler, R., Harwood, D., Davis, L.S.: Backpack: Detection of people carrying objects using silhouettes. In: ICCV, Kerkyra, Greece (1999) 102–107
2. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60**(2) (2004) 91–110
3. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. **22**(8) (2000) 888–905