A New RMI Framework for Outdoor Objects Recognition

Wong Chuan Ern

Universiti Tunku Abdul Rahman, Malaysia
chuanern@hotmail.com

Ong Teong Joo

Universiti Tunku Abdul Rahman, Malaysia
ongtj@mail.utar.edu.my

Abstract

In this paper, we present an extension to the Recurrent Motion Image (RMI) motion-based object recognition framework for use in development of automated video surveillance systems. We extended the object classes of RMI to include four-legged animals (such as dog and cat) and enhanced the preprocessing and shadow removal algorithms for better object segmentation and recognition. Under the new framework, object blobs obtained from background subtraction of scenes are tracked using region correspondence. In turn, we calculate the RMI signatures based on the silhouettes of the object blobs for proper classification. This new framework is tested on several real world 320 x 240 resolution color image sequences captured with a low-end digital camera, and all of the moving objects in our samples are properly detected, tracked and classified - indicating the applicability of the new framework in similar task environment.

1. Introduction

Moving object recognition has been an active area of research for computer vision and pattern analysis applications. It plays a major role in advanced security systems and video surveillance applications. With the aim to recognize moving objects through video monitoring system, an object recognition algorithm should be able to detect and track moving objects in a surveillance area, and classify objects of interest into various predefined categories.

Enhancement to current security systems and surveillance applications can be realized with this recognition function. For instance, intruder recognition function can be incorporated into a security system to classify intruders in order to reduce nuisance alarm and minimize human errors in manned surveillance system. Another example of its application is in traffic

monitoring system, where it can be used to estimate traffic flow by making vehicle and pedestrian counts.

This paper presents an improved motion-based recognition approach using a specific feature vector called Recurrent Motion Image (RMI) [1] to classify moving objects into predefined categories, namely single person, group of persons, vehicle and four-legged animal. Moving objects detected from image sequences are classified based on their periodic motion patterns captured with the RMI.

In our approach, a new object class has been introduced to the existing RMI framework and various refinements are made on the processing algorithms for improved recognition accuracy and compatibility for use in general outdoor scenes.

2. Background

Extensive research efforts have been dedicated to moving object recognition, where many approaches, such as [1] – [5] have been presented to tackle this problem.

RMI method [1] is one of the few approaches that produce high recognition rate while remaining computationally and space efficient. A specific feature vector called RMI can be used to estimate the repetitive motion behavior of moving objects yielding different RMI signature for different object's motion behavior. Thus, moving objects can be classified as single person, group of persons or vehicle based on their corresponding RMI.

In previous work [1], this approach yields correct classification in about 98 percent of all tested samples. However, the shadow removal algorithm in the original framework suffers an error rate of 30 percent. In addition, the original RMI framework has only been tested on a small set of object classes (human and vehicle).

In light of the shortcomings of the original RMI framework mentioned earlier, we would like to present several refinements we have made to the RMI framework to increase its recognition classes and recognition rate.

3. Methods

Our approach extends the object classes of the existing framework to include four-legged animal and uses different shadow removal and preprocessing algorithms to improve the recognition accuracy.

3.1. Object detection

The new framework extends the preprocessing stages from [1] to include a better shadow removal algorithm and multiple levels of noise filtering for better moving object segmentation, as shown below:

- Background subtraction is carried out by computing an L-inf distance image [6] in the Red-Green-Blue (RGB) color space.
- Foreground points are obtained by applying a low threshold to the L-inf distance image and the points will go through a morphological opening denoise layer [6] to remove first level noise.
- Shadow points in the resultant image are located by transforming the image pixel values to the Hue-Saturation-Value (HSV) color space [7].
- Shadow points are removed from the image, and connected components labeling [8] algorithm is applied to extract the foreground blobs.
- Blob analysis [6] is then performed to filter noise clutters using a blob size threshold.
- Lastly, a high threshold is applied to the L-inf distance image to select points with large difference from the background. Blobs consisting of at least one of these salient points are validated while the others are removed as non-salient blobs.

While the aforementioned procedure may work well for most scenarios, it may fail when objects exhibit similar color to the background. As observed from our earlier experiments [9] that certain objects are sensitive to point-level filters because their results from color segmentation are weak. The resultant scattered foreground points from background subtraction are easily diminished by morphological opening denoise layer that is responsible for the first level noise filtering. The missing information induces problem in proper object segmentation and accurate detection of the corresponding moving objects.

To circumvent this problem, we used connected components labeling algorithm and a blob size filter with low threshold to replace the point-based morphological opening layer at stage 2 instead. The combination allows us to remove the first level noise while retaining the scattered pixels of the moving object candidates.

3.2. Object tracking

Blobs obtained from the detection section are tracked using region correspondence [10]. Parameters such as centroid, bounding box, size, velocity and change in size of each blob are extracted. Correspondences between regions in previous frame and current frame are established using the minimum cost criteria [1] to update the status of each object over the frames.

Non-corresponded region may possibly involve in object entry, exit or occlusion, so such region from previous frame is examined for object exit. If its predicted position exceeds the frame boundary, it is determined to have exited the surveillance area; otherwise, object occlusion may have happened based on the following reasoning: if its bounding box overlaps the bounding box of another region Q in the current frame then Q is marked as an occluded region. Thus, all of the non-corresponded regions in previous frame overlapping Q are marked as occluding each other while non-corresponded region in the current frame is set to be an object entry.

3.3. Object classification

Each of the moving objects detected and tracked in the image sequences are classified as a single person, a group of persons, a vehicle or a four-legged animal. Recurrent motion which is denoted as repetitive changes in shape of the objects is the main essential feature that differentiates the object classes. RMI will have high values at pixels where motion occurred repetitively and low values at pixels where little or no motion occurred.

RMI is computed with the following equations to determine the areas of moving object's silhouette undergoing repetitive changes:

$$DS_a(x, y, t) = S_a(x, y, t-1) \oplus S_a(x, y, t)$$

$$RMI_a = \sum_{k=0}^T DS_a(x, y, t-k)$$

S_a is a binary silhouette image for object a at frame t, and DS_a is a binary image indicating areas of motion for object a between frame t and t-1. RMI_a is the RMI for object a calculated over T frames.

The RMI is partitioned into N equal-sized blocks in order to compute the average recurrence for each block, as illustrated in Figure 1. Blocks with average recurrence value greater than a threshold τ_{RMI} are set to 1 (white) and vice versa. If there are white blocks at

the middle and bottom sections, the object is classified as a single person or a group of persons. If there are no white blocks, which means no recurrent motion, the object is classified as a vehicle.

Figure 1. RMI for classification [1]

There are two cues to differentiate between single person and group of persons – multiple peak points in a silhouette indicates more than one headcount, therefore representing a group of persons, and, the normalized area of recurrence response at the top section of RMI being greater than that of a single person, due to presence of multiple heads. If either one of the aforementioned criteria is satisfied, the object is classified as a group of persons, whereas if none of them are satisfied, the object is classified as a single person.

To include four-legged animal as an additional object class, we studied the recurrent motion behavior of dogs and cats to derive the motion pattern and criteria for classification. Figure 2 shows the RMI generated for a dog and a cat. Notice that the RMI for dog and cat are similar, whereby their legs and tail demonstrated repetitive motion. This causes white blocks to occur in all top, middle and bottom sections. However, for dogs and cats without a tail, only middle and bottom sections will have white blocks. In addition, dogs or cats which are short in height but long in length may cause white blocks to occur only at the middle section.

Thus, blocks in the middle section are taken into consideration for four-legged animal classification. This is similar to the cue for classifying a moving object as single person or group of persons. Therefore

the classification criteria are modified specifically to differentiate between human and four-legged animal (dog or cat in this case).

Figure 2. RMI of a dog and a cat

The black area within the RMI of an object corresponds to the area where the object demonstrates no recurrent motion. Human and four-legged animals generally do not show any recurrent motion at the main part of the body where the backbone is located. This can be seen in Figure 1 and Figure 2. Notice that the black area within the RMI of a single person and group of persons (Figure 1) has vertical major axis, whereas the black area within the RMI of a dog and a cat (Figure 2) has horizontal major axis. This acts as the cue to differentiate between human and four-legged animal (dog or cat in this case).

When there are white blocks in the middle section of a partitioned RMI, the black area within the respective RMI is observed. If the black area has a vertical major axis, the corresponding object is classified as human, which will then be further categorized as a single person or group of persons. Otherwise, if the black area has a horizontal major axis, the object is classified as a four-legged animal. If the object does not fall under any of the predefined categories (vehicle, single person, group of persons and four-legged animal), it will be classified as other objects.

Note that accurate differentiation between all of the four-legged animals classified by the new algorithm can be carried out using texture, size, color and other relevant information of the object. Our proposed RMI classification method can be used as a filter that classifies moving objects into the proper categories for further processing in a recognition engine.

4. Results

The moving object detection, tracking and classification algorithm is implemented in Matlab and executed on a 1.5GHz Core 2 Duo CPU. To run the experiment, we captured several image sequences with a low-end digital camera (Olympus FE-280) at various housing areas.

The image sequences consist of a variety of single persons, groups of persons, vehicles, four-legged animals (dogs and cats), and a few of them are shown in Table 1. In our experiments, the RMI signature of a moving object was generated for one second duration after it has completely entered the frame. All in all, 23 image sequences of various time lengths were taken in parallel to the ground plane. The frames were 320 x 240 pixels in size and sampled at a rate of 8 frames per second. Table 1 shows several instances of the moving object detected and classified using the framework with a threshold (τ_{RMI}) of 2.

Table 1. Examples of moving object detected and classified

Image frame	RMI	Object class
Single person
Group of persons
Vehicle
Four-legged animal

In the early stages of the experiment, we failed to detect or classify some of the moving objects accurately. The problem was caused by the foreground extraction problem due to color similarity between the moving object and its background. Upon further examinations and tests, we found that the point-based morphological opening denoise layer has removed the

scattered pixels of these moving objects, causing incomplete object silhouettes to be fed into our recognition module.

Having replaced the point-based morphological filter with a low-threshold blob size filter which is able to remove small noise blobs (in the size of less than 20 pixels in the example), we are able to preserve the object's weak pixels in the noise filtering layer. The tradeoff is that the silhouette produced by this method is a little coarse at the boundaries. Figure 3 illustrates the results obtained from the use of morphological opening operator and blob size filter at the first level denoising process.

Figure 3. Outputs at preprocessing stages using different first level denoise method

Lastly, table 2 summarizes the result of our experiments where all of the moving objects were

correctly classified - indicating the backward compatibility and successful integration of the new classification list (human, vehicle and four-legged animal) in this new framework.

Table 2. Classification results

Object class	Number of samples tested	Number of samples correctly classified
Single person	5	5
Group of persons	4	4
Vehicle	6	6
Four-legged animal	8	8

5. Conclusion and future work

All moving objects in our experiments were properly detected, tracked and classified into the proper categories: 1) single person; 2) group of persons; 3) vehicle; and, 4) four-legged animal. The problem of weak segmentation for objects with areas that exhibit similar color to the background were solved by replacing the point-based morphological filter with a low-threshold blob size filter at the first denoise layer.

Although the aforementioned weak object segmentation problem caused by color similarity has been solved, the resulting object silhouette is a little rough and requires additional smoothing. The smoothing process may affect segmentation of objects that are relatively small in the image frame, or in noisy image sequences. Further investigations and more experiments are needed to ascertain the effectiveness of the process.

6. References

- [1] O. Javed and M. Shah, "Tracking and Object Classification for Automated Surveillance," *Proc. 7th European Conf. Computer Vision-Part IV (ECCV 02)*, 2002, pp. 343-357.
- [2] C.M. Cyr and B.B. Kimia, "A Similarity-Based Aspect-Graph Approach to 3D Object Recognition," *Int'l J. Computer Vision*, vol. 57, no. 1, 2004, pp. 5-22.
- [3] D. Toth and T. Aach, "Detection and Recognition of Moving Objects Using Statistical Motion Detection and Fourier Descriptors," *Proc. 12th Int'l Conf. Image Analysis and Processing (ICIAP 03)*, 2003, pp. 430-435.
- [4] J.B. Hayfron-Acquah, M.S. Nixon, and J.N. Carter, "Recognising Human and Animal Movement by Symmetry," *Proc. IEEE Int'l Conf. Image Processing (ICIP 01)*, 2001, pp. 290-293.
- [5] Y. Bogomolov, G. Dror, S. Lapchev, E. Rivlin, and M. Rudzsky, "Classification of Moving Targets Based on Motion and Appearance," *Proc. British Machine Vision Conf. (BMVC 03)*, 2003, pp. 429-438.
- [6] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting Moving Objects, Ghosts and Shadows in Video Streams," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, 2003, pp. 1337-1342.
- [7] R. Cucchiara, C. Grana, M. Piccardi, A. Prati, and S. Sirotti, "Improving Shadow Suppression in Moving Object Detection with HSV Color Information," *Proc. IEEE Int'l Conf. Intelligent Transportation Systems (ITSC 01)*, 2001, pp. 334-339.
- [8] R.C. Gonzalez, R.E. Woods, and S.L. Eddins, *Digital Image Processing Using Matlab*, Prentice Hall, Upper Saddle River, NJ, 2004, p. 359.
- [9] C.E. Wong and T.J. Ong, "Moving Object Recognition Using Improved RMI Method," *Proc. 2nd Int'l Conf. Science and Technology (ICSTIE 08)*, to be published.
- [10] B. Jahne and H. HauBecker, *Computer Vision and Applications: A Guide for Students and Practitioners*, Academic Press, San Diego, USA, 2000, p. 379.