Unusual human behavior recognition using evolutionary technique

Lin Lin [a,*], YongSu Seo [b], Mitsuo Gen [a], Runwei Cheng [c]

[a] Graduate School of Information, Production and Systems, Waseda University, 2-7 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0135, Japan
[b] Dongseo University, San69-1, Churye-Dong, Sasang-Ku, Busan 617-716, Republic of Korea
[c] JANA Solutions Inc., 1-15-13 Shiba, Minato-ku, Tokyo 105-0014, Japan

Abstract

The vision for surveillance is an important task in many computer vision applications. The monitoring system concerns the tracking and recognition of people, and more generally, the understanding of human behaviors, from image sequences involving humans. Several methods for human tracking and human behavior recognition have been proposed by various researchers. But most of those do not have versatility and flexibility. In this paper, we propose an efficient and robust object tracking algorithm which use the color features, the distance features and count feature based on an evolutionary techniques to measure the observation similarity. And then we will track each person and classify their behavior properties by analyzing their trajectory pattern. We propose multi-layer perceptron based on hybrid genetic algorithm using Gaussian synapse make the recognition algorithm very efficient and robust for classify human behavior by trajectory pattern.
© 2008 Elsevier Ltd. All rights reserved.

Keywords: People tracking; Unusual behavior; Evolutionary technique; Hybrid genetic algorithm; Gaussian synapse

1. Introduction

In recent years, the increasing availability of video sensors and high performance video processing hardware opens up exciting possibilities for tackling many computer vision problems. It is important to develop robust real-time monitoring systems which can process the large amounts of data attainable. And the vision for surveillance is a very important application in computer vision. The monitoring system concerns the tracking and recognition of people, and more generally, the understanding of human behaviors, from image sequences involving humans.

Several methods for human tracking and human behavior recognition have been proposed by various researchers. But most of those do not have versatility and flexibility, and spend a lot of time and cost on building the tracking human behavior recognition system.

---

* Corresponding author. Tel.: +81 90 9565 2964.
  E-mail address: lin@ruri.waseda.jp (L. Lin).

Systems for tracking people have usually employed some form of background subtraction: Haritaoglu, Harwood, and Davis (1998), Haritaoglu, Harwood, Davis, and Hydra (1999) used grey level subtraction, Wren, Azarbayejani, Darrell, and Pentland (1995) modeled pixel color variation using multivariate Gaussians, and Gaussian mixtures have also been used in a similar manner (McKenna, Raja, & Gong, 1999; Stauffer & Grimson, 1999). The car tracking system by Koller, Weber, and Malik (1994) used an adaptive background model based on monochromatic images filtered with Gaussian and Gaussian derivative (vertical and horizontal) kernels. Background 'subtraction' using these filter outputs yielded results superior to the use of raw grey levels.

A number of tracking algorithms for different applications have been presented. While these works present a wide range of different methods, they can be classified into two main categories of increasing complexity as follows:

- Methods using full 3D modeling of human.
- Methods using region 2D appearance model of human.

1.1. Methods using full 3D modeling of human

Most of the work on tracking for visual surveillance is based on 3D modeling of change detection (Elgammal, Duraiswami, & Davis, 2003; Haritaoglu, Harwood, & Davis, 1998; Hongeng & Nevatia, 2001; Rosales & Sclaroff, 1999; Stauffer & Grimson, 2000) or frame differencing if the camera is stationary. Additional stabilization is required if the camera is mobile (Tao, Sawhney, & Kumar, 2002). Some researches combine compact object shape, motion, and appearance in a Bayesian framework. However it does not explicitly handle occlusion of multiple objects since it was designed mainly for airborne video. Much work has been done on estimating human body postures in the context of video motion capture (a recent review is available in Moeslund and Granum (2001)). This problem is difficult, especially from a single view because 3D pose may be under constrained from one viewpoint. Most successful systems (e.g., Deutscher, Davison, & Reid, 2001) employ multiple viewpoints, good image resolution, and heavy computation, which is not always feasible for applications such as video surveillance. Use of constrained motion models can reduce the search space, but it only works on the type of motion defined in the model. Rohr (1994) describes pioneering work on motion recognition using motion captured data. In each frame, the joint angle values are searched for on the motion curves of a walking cycle. Results are shown only on an isolated human walking parallel to the image plane. Motion subspace is used in Sidenbladh, Black, and Fleet (2000) to track human walking using a particle filter. Recently, an approach has been proposed by Efros, Berg, Mori, and Malik (2003) to recognize actions. It is also based on flow-based motion description and temporal integration.

1.2. Methods using region 2D appearance model of human

A variety of 2D appearance models tracking algorithms have been proposed and implemented. Tracking algorithms using 2D appearance models can be roughly classified into two categories: deterministic methods and stochastic methods.

Deterministic methods typically track by performing an iterative search for the local maxima of a similarity cost function between the template image and the current image. The cost function widely used is the sum of squared differences (SSD) between the template and the current image such as in Baker & Matthews, 2004; Hager & Belhumeur (1998). More robust similarity measures have been applied and the mean-shift algorithm or other optimization techniques have been utilized to find the optimal solution (Collins, 2003; Elgammal et al., 2003). Appearance-based approaches apply recognition algorithms to learn the objects either in some basis such as the eigenspace formed from observations or in kernel space (Black & Jepson, 1998; Williams, Blake, & Cipolla, 2003).

On the other hand, the stochastic methods use the state space to model the underlying dynamics of the tracking system. In a linear-Gaussian model with linear measurement, there is always only one mode in the posterior probability density function (pdf), the Kalman filter propagates and updates the mean and