Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT
Image and Vision Computing 22 (2004) 563–570
www.elsevier.com/locate/imavis

A real-time system for monitoring of cyclists and pedestrians

Janne Heikkilä*, Olli Silvén

Machine Vision Group, Infotech Oulu and Department of Electrical and Information Engineering, PO Box 4500, FIN-90014 University of Oulu, Finland

Abstract

Camera based systems are routinely used for monitoring highway traffic, supplementing inductive loops and microwave sensors employed for counting purposes. These techniques achieve very good counting accuracy and are capable of discriminating trucks and cars. However, pedestrians and cyclists are mostly counted manually. In this paper, we describe a new camera based automatic system that utilizes Kalman filtering in tracking and Learning Vector Quantization for classifying the observations to pedestrians and cyclists. Both the requirements for such systems and the algorithms used are described. The tests performed show that the system achieves around 80–90% accuracy in counting and classification.

Keywords: Human tracking; Traffic counting; Target classification

1. Introduction

The analysis of traffic flows has traditionally been based on inductive and microwave sensors. These technologies are not, however, capable of reliable monitoring of pedestrians and cyclists for two reasons. First, the reliability of inductive sensors depend on the materials in the vehicle, and second, the structure of pedestrian traffic is weak and does not focus on easily recognizable lanes. It is necessary to find an alternative technique for pedestrian traffic analysis due to the problems of the traditional approaches.

Traffic flow statistics are mainly used by road administration which tries to resolve the needs for new pavements, cycle paths and tunnels, for instance. Currently, such surveys are performed manually. However, manual counting is tedious and typically continues only for a short period of time which makes the results quite unreliable.

A natural choice for monitoring pedestrians and cyclists automatically is to use visual perception. However, there are still quite few camera based systems developed for people detection and tracking. A well known undertaking in this area was the ESPRIT project PASSWORDS [2]. Another visual surveillance application is W4 [6], which is a real time system for determining interactions between people. Most of the systems presented are based on an assumption of a stationary camera which greatly simplifies the problem of human detection, although solutions for pedestrian detection from moving platforms have also been suggested, for example, in Refs. [5] and [10]. Furthermore, a model based approach for pedestrian detection was proposed in Ref. [11], and a statistical framework for learning and recognizing human movements in Ref. [3].

In human tracking, there are many visual clues that can be utilized. These include color histograms [4], coherent connected regions or 'blobs' [12], and object contours [1]. Kalman filtering is probably the most commonly used algorithm for implementing the tracker, although recently Condensation algorithm [7] and mean shift algorithm [4] have shown to provide certain advantages especially in the presence of significant background clutter. Despite of the methodological advancements, many of the solutions proposed are computationally expensive and they make limiting assumptions about the target appearance and dynamics. Moreover, there is often a long way to go from a controlled laboratory environment to unpredictable outdoor scenes.

When the project for developing an automatic traffic monitoring system was started in 1994, there were no commercial systems available for that purpose. According to our knowledge there still exist no other systems capable of fulfilling even partly the following requirements that were specified by our user group:

* Corresponding author.
E-mail addresses: jth@ee.oulu.fi (J. Heikkilä), olli@ee.oulu.fi (O. Silvén).

0262-8856/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2003.09.010

- counting of the volume of the traffic
- classification of the traffic into separate traffic classes (pedestrians, cyclists,…)
- route tracing
- counting accuracy better than 90%
- classification accuracy better than 85%
- portability and easy installation
- one week stand alone operation time
- transfer of data into a portable field terminal through a telephone network or via a memory card

In addition, the weather conditions set extra requirements on the hardware. The device should work in various temperatures (from −20 to +30 °C), in rain, snowfall, and wind. Also, adaptation to a wide illumination range with sudden changes is typically very difficult for camera based systems.

One of the main problems in developing a stand alone system is the power consumption. The limited capacity of the batteries and long operation time allows us to use only economic hardware solutions at the cost of processing speed. This also forced us to divide the monitoring problem into two parts: real-time tracking and off-line analysis.

The tracking unit is installed beside the road on a pole. It performs several image processing tasks to detect and track moving objects in the scene. The shape and the trajectory of the objects are stored in a removable memory card to wait for further processing. After several days of stand alone operation, the memory card is fetched and the data is transferred to a separate desktop computer equipped with analysis software. The objects are then classified to pedestrians and cyclists based on the features stored during the tracking phase. The architecture of the monitoring system is depicted in Fig. 1.

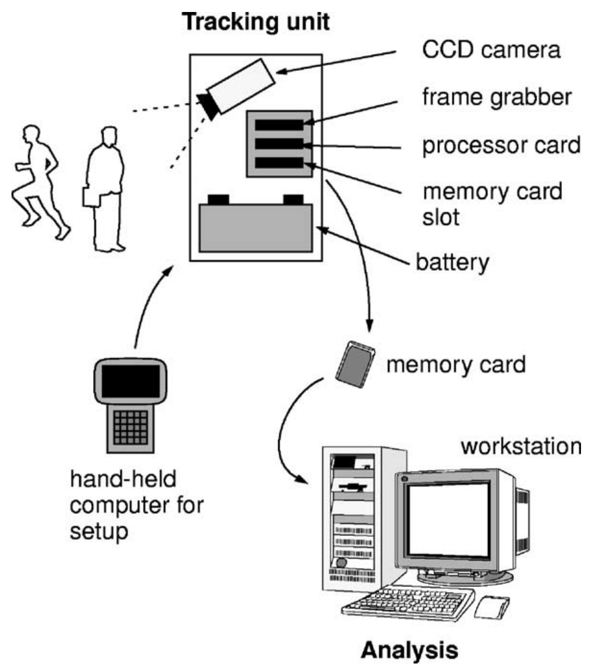

Fig. 1. System architecture.

In Section 2, we will describe the basic techniques used in the tracking unit including the image processing steps for feature extraction and the Kalman filter based tracking scheme. Section 3 introduces the Learning Vector Quantization (LVQ) based analysis for producing the final counting and classification result. The preliminary test results are given in Section 4.

2. The tracking unit

The main modules in the tracking unit are the CCD camera, frame grabber and Ampro 486/50 MHz processor card. The power consumption of the unit was optimized to around 7 W in full operation mode and 1 W in sleep mode. The sleep mode is used at night time when the overall lighting is not adequate for reliable observations.

The weight limit of the tracking device was 5 kg, including the used 20 Ah 6 V sealed lead acid battery. The selection of the battery technology was dictated by the temperature conditions ranging from −20 to +30 °C. The mean power consumption of the tracking unit is approximately 2.5–3 W of which roughly 1.5 W is taken by the camera that is always on, except during the sleep mode. The technology solutions were selected from among several alternatives, mostly based on power control and consumption characteristics.

The purpose of the tracking unit is to detect moving people from the video stream and collect appropriate data of their routes and shapes. Each event is saved and compressed with a time stamp in a removable memory card.

The images read from the frame grabber are decimated to 160 by 120 pixels with only luminance information preserved in order to reduce the computational load and to guarantee adequate frame rate (around 10 fps) for tracking. Each incoming frame goes through four successive image processing stages where the raw intensity data is reduced to a compact set of features which can be used as an input of the classifier. These four stages, motion detection, filtering, feature extraction and tracking, are described next.

2.1. Motion detection

Motion detection is started by computing a pixel based absolute difference between each incoming frame and an adaptive background frame $B(k)$. The pixels are assumed to contain motion if the absolution difference exceeds a predefined threshold level. As a result, a binary image is formed where active pixels are labeled with ‘1’ and non-active ones with ‘0’.

It is necessary to update the background image frequently in order to guarantee reliable motion detection. The basic idea in background adaptation is to integrate