THE PASSWORDS PROJECT

Marc Bogaert[*] , *Nicolas Chleq*[#] , *Philippe Cornez*[*] ,
Carlo S. Regazzoni[$] , *Andrea Teschioni*[$] , *Monique Thonnat*[#]

[$] Department of Biophysical and Electronic Engineering (DIBE), University of Genova,
Via all'Opera Pia 11A, Genova, Italy

[*] Research Center of the Belgian Metalworking Industry (CRIF),
Avenue Franklin Roosevelt 50, Bruxelles, Belgium

[#] Institut National de Recherche en Informatique et Automatique (INRIA),
Route des Lucioles 2004, Sophia Antipolis, France

ABSTRACT

The objective of the PASSWORDS Project is to design and develop a prototype of an intelligent video image analysis system for video surveillance and security applications, based on concrete needs expressed by potential users.

The goal of the system developed within PASSWORDS is to detect certain dangerous situations in some scenes (e.g. vandalism in a metro station), providing for example a remote operator with an alarm signal.

This paper illustrates the different steps which compose the PASSWORDS system, focusing the attention on the Image Processing Module.

I. INTRODUCTION

Computer-assisted surveillance of complex environments is getting more and more interesting, thanks to recent significant improvements in real-time signal processing. The main role of automatic computation in such systems is to support the human operator to perform many tasks, such as detecting, interpreting and understanding, logging or giving alarms. In this way, possible human failures are expected to be overcome, many functionalities can be added, and better surveillance performances are obtained.

In the surveillance research field applied to public areas, crowding monitoring is very useful but presents particularly complex problems. Recognizing objects and persons and tracking their movements in complex real scenes by using a sequence of images are among the most difficult tasks in computer vision [1][2]. Main difficulties arise both from the high level of shape variability associated with people, and from the complexity of environmental real scenes [3].

II. THE PASSWORDS PROJECT

Within this context, the Passwords Project (EP 8433, Parallel and Advanced Surveillance System With Operator assistance for Revealing Dangerous Situations) intends to demonstrate the feasibility of image sequence analysis on low-cost hardware for visual surveillance and security applications.

The applications cover indoor as well as outdoor scenes, in which people as well as vehicles or other objects are moving around. The goal of the system is to detect certain particular situations in some scenes, which are of interest to the operator, such as persons remaining a too long time in a certain area (e.g. a prowler, waiting on a car park for an occasion to steal something from inside a car, or a member of the right cleaning team too much "interested" by some articles in a shelf of a supermarket or persons with an agitated behaviour, individually as well as in group (e.g. potential vandals in a subway).

The development work is being carried out on parallel digital signal processors, TMS320C40, common to all developing partners of the project. The operating system Virtuoso is used, allowing programming on a "virtual" single processor.

III. THE STRUCTURE OF THE PAPER

This new architecture can be effectively considered an innovation for real-time image processing and also a market

0-7803-3258-X/96/$5.00 © 1996 IEEE

analysis revealed the existence of key interest areas of potential customers of a "passwords" project.

The present paper is mainly addressed to present the main image processing phases which are at the basis of object recognition and tracking:

- illumination change compensation
- shadow compensation
- people matching and tracking
- detection of abnormal behaviours

Thanks to real-time functioning, accuracy and robustness, the method can be used in surveillance systems devoted to assure people safety in public areas.

IV. THE BLOB DETECTION PHASE

The purpose of this phase is to identify the mobile areas (i.e. the blobs) of each image successively acquired by the system.

The acquired color images are first filtered in order to reduce the noise level present in each original images. A background image, i.e. a reference image without any mobile object, is then continuously updated in order to adapt itself to scene changes as for example a new static object

The next step is the application of an *illumination change compensation* algorithm to the filtered images, in order to make the whole processing chain run also in outdoor conditions. This algorithm is based on the determination of the illumination change factors using the log difference between the input image and the background image and on a correction of the change in order to smooth the variation of these factors wrt the time.

From each such corrected image, the mobile areas of the image are extracted using the difference between the corrected image and the background image.

A color segmentation follows, in order to determine the global characteristics of each region within the image, and to make a discrimination between coloured pixels (belonging to mobile objects) and shadows pixels. This *shadows analysis* allows in fact to eliminate from the moving areas (i.e. the blobs) the present shadows in each area and it makes easier the task performed by the next image processing algorithms.

The following stage is the *blob tracking* phase.

V. THE BLOB TRACKING PHASE

The target of this module is to identify with the same label each blob corresponding to the same physical object present in the sequence of images over time (matching phase) and, on the basis of the matching results, to track each blob movement on the image plane over time.

The module labels with the same number each blob corresponding to the same object, object part or group present in the sequence over time. On the basis of the matching result, each blob can be tracked over time, hence providing a further blob characterization by means of cinematic parameters. Matching is performed in two steps:

A. A first rough result is reached by comparing the lists of blob characteristics referring to the current (time step k) and previous (time step k-1) frames. Blob correspondences are organized as a graph: all nodes of each level are the blobs detected in each frame, and the relationship among blobs belonging to different adjacent levels are represented as arcs between the nodes. Arcs are inserted on the basis of the superposition of blob areas on the image plane: if a blob at step (k-1) overlaps a blob at step k, then a link between them is created, so that the blob at step (k-1) is called "father" of the blob at time step k (its "son"). Different events can occur:

1) If a blob has only one "father", its type is set "*one-overlapping*" (type o), and father label is assigned to it.
2) If a blob has more than one "father", its type is set to "*merge*" (type m), and a new label is assigned.
3) If a blob is not the only "son" of its father, its type is set to "*split*" *(type* s*)*, and a new label is assigned.
4) If a blob has no "father", its type is set to "*new*" (type n) and a new label is assigned.

B. Blob matching is refined by substituting, if possible, the new labels with the labels of some blob either belonging to a time step *previous* than (k-1), or belonging to the (k-1) frame and erroneously labelled in phase *A*. This processing phase is based on the comparison between each current blob not labelled with "o" with the set of recent previous blobs, whose label was inherited by no successive blob, collected in a "long-memory" blob graph. This approach is useful for recover some situations of temporary wrong splitting of a blob (corresponding to a single object) into more blobs, because of image noise or static occlusions, or of temporary merging of two overlapping objects. Comparison is performed on the basis of the blob shape characteristics which have been tested to be approximately time/scale-invariant.

Blob matching provides as output the final *graph* of blob correspondences over time, in which matched blobs have the same label over time, and each blob is classified with one of the described types.

VI. THE IMAGE UNDERSTANDING PHASE

The image understanding module takes as inputs image objects (blobs) and features extracted from each image of the temporal sequence and performs the understanding process up to the on-line recognition of ongoing scenario, with respect to a library of scenario models dedicated to a particular application.

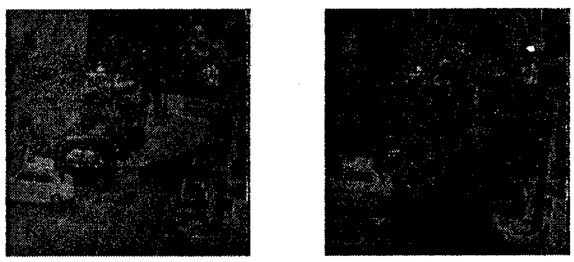
The main feature of this module (see also [4]) is the ability to perform the interpretation in an incremental fashion starting from primitive low-level events and refining the interpretation by relying on a hierarchical recognition of scenario models. Another characteristic is that the image understanding module does not need a high computing power. Currently, the cycle time for one image goes from 10 to 50 milliseconds: provided that the image processing programs can be scheduled at the convenient rate, this allows a real-time interpretation to be performed.

VII. RESULTS

Extensive tests on real image sequences were performed in the context of CEC-ESPRIT Project PASSWORDS: we will present in this section the results coming from the application of two modules developed within the Project itself: the illumination change compensation algorithm and the blob tracking module.

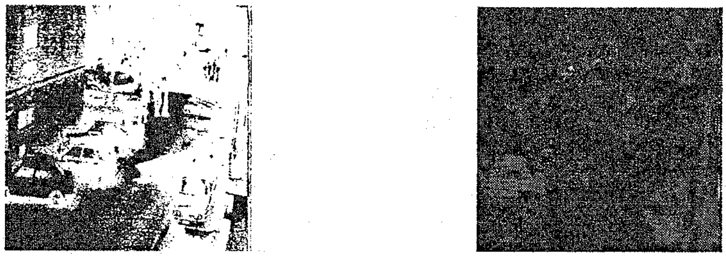
Figure 2 shows an example of results coming from the application to a test image sequence of the illumination change compensation algorithm in outside conditions (cloud occluding the sun). The algorithm carries a good improvement of the detection of mobiles: a lot of blobs due to illumination changes are filtered out, while most of the mobile parts remain.

Image 1 Image 2 Corrected image 2



Logarithmic difference between images 1 & 2 Logarithmic difference between image 1 and corrected image

On these images, the larger the absolute difference, the whiter the pixel



Thresholded logarithimic diff.
between images 1 and 2

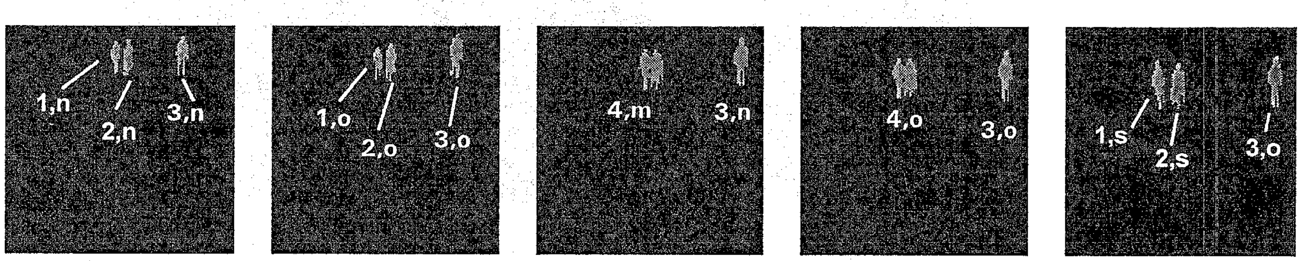
Thresholded logarithmic diff.
between image 1 and corrected image 2

Black parts are considered as mobile parts on these thresholded images

Figure 2. Test images of the illumination change compensation algorithm in outside conditions

Figure 3 shows the result of the blob matching algorithm on a test sub-sequence: each image contains the detected blobs (resulting from blob detection) with their numerical label and type (obtained from blob matching). This example points out in particular the capability of the module to:

1. assign the same label to two blobs before and after their temporary overlapping on the image plane (hence to consider them as the same mobile object) even after several frames (see frames 2 and 5);
2. assign the correct label of a blob which was erroneously classified as new during the matching phase *A* (see frames 2 and 3).



frame 1 frame 2 frame 3 frame 4 frame 5

Figure 3. A sequence of images showing critical cases of blob splitting, merging and displacement.

References

[1] W. Kinzel, and Dickmanns ED, "*Moving humans recognition using spatio-temporal models*", XVII-th Congress Int. Soc. Photogrammetry and Remote Sensing, 1992.

[2] F. Cravino, M. Delucca and A. Tesei, "*DEKF system for crowding estimation by a multiple-model approach*", Electronics Letters, no. 30, vol. 5, 1994, pp. 390-391.

[3] Y. Bar-Shalom, and T.E. Fortmann, "*Tracking and Data Association*", Academic Press, New York, 1988.

[4] N. Chleq, and M. Thonnat, *"Real-time Image Sequence Interpretation for Video-Surveillance Applications"*, accepted for ICIP '96, Lausanne(Suisse)