DEMO : Detection of Medication Intake

Myriam Valin
Biomedical Engineering Institute
Université de Montréal
Montreal, Canada
valinmyr@iro.umontreal.ca

Jean Meunier
Biomedical Engineering Institute and
Département d'informatique et
de recherche opérationnelle
Université de Montréal
Montreal, Canada
meunier@iro.umontreal.ca

***Abstract*— In the context of the growing proportion of seniors in the western world population and the efforts provided in home care services, a computer vision system has been developed for monitoring medication intake. The system can detect automatically medication intake using a single webcam. Person detection and tracking over the video sequence is done using color-based techniques while the recognition of the medication intake activity is performed using a multi-level scenario model. Experimental results in controlled conditions are shown.**

***Keywords*—**Video surveillance, medication intake, tracking, activity recognition, smart homes.

I. Introduction

The system presented here [1] is part of a vaster project of video surveillance for home care services for seniors. It has been developed for monitoring medication intake in the context of a person living alone and whose medications are in bottles. In order to be as accessible as possible, a single low-cost webcam is used. The system has therefore to deal with low-quality digital images.

Recently, efforts have been made on developing computer vision systems for monitoring medication intake [2]. One problem with this approach is that the skin segmentation needs to be very precise to detect effectively occlusions, hand positions and orientations and face regions, which cannot be reached with low-quality images. Because of the expected lack of tracking precision and the complexity of the activity to recognize, we developed a more complex algorithm to improve medication intake activity recognition.

II. Method

A. Overview

In order to perform medication intake monitoring, the system must first detect if the person is taking his medication. The proposed method is inspired by that in [5]. In our method, processing is divided in two parts:

- Low level processing: moving objects (regions) are detected and tracked at every frame.
- High level processing: activity recognition is performed based on moving object characteristics, using our three-level scenario model.

B. Detection and Tracking

In our approach, three types of mobile objects are tracked: the person's head and hands and the medication bottles. Fig. 1a shows these objects. In this experimental setting, three medication bottles are present.

The head tracking algorithm is described in [3]. The head is modeled as an ellipse whose size can vary from one frame to the other. For each image, a local search determines the best fitting ellipse, based on the gradient magnitude around its perimeter and the likelihood of skin color inside it. The gradient magnitude of a pixel corresponds to the rate-of-change of intensities in the gray-scale image over a small local neighborhood. The color histogram used to compute skin color likelihood includes the person's hair color since the head might be turned or leaned.

Hands are positioned based on regions with high skin color likelihood. These regions are extracted from the skin color likelihood map (Fig. 1b) created during the head tracking process. Possible occlusions between the head and the hands or between both hands are also dealt with based on previous positions and a few assumptions.

Since we want to detect which medication is being taken and labels cannot be identified automatically in low resolution video, color bands are affixed on the bottles to better differentiate them.

The detection of medication bottles is done using color models as described in [4]. Pixels classification is performed based on their Mahalanobis distance to the color models, which are created from training samples. Possible bottle regions (objects) are extracted using connected component analysis. Bottle positioning is done considering the size of these objects and bottle previous positions, since bottles can be completely occluded by the hands.

C. Activity Recognition

Medication intake detection is complex because it can be done in many different ways. Thus, activity recognition cannot be done by searching for a fixed sequence. The medication intake activity is then divided in 3 simple scenarios:

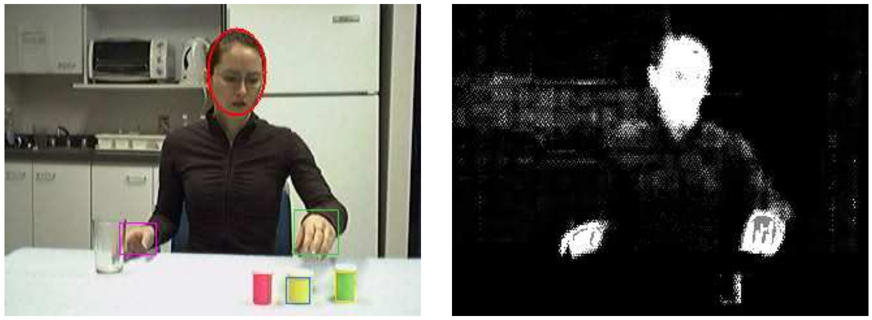
- (S1) the person opens the medication bottle and take the pill(s),
- (S2) the person swallow the pill(s),
- (S3) the person closes the medication bottle.

At a higher scenario level, the system searches for certain sequences of these simple scenarios, which can be separated by unrelated actions such as drinking water or putting the pills on the table, to recognize the whole activity [1].

III. Results

The system uses a Logitech Quickcam personal web camera for notebooks Pro with a digital video capture speed of 30 frames per second and a resolution of 640x480 pixels. However, in practice, the maximum image resolution to avoid lags is 320x240 pixels. Since a frame rate of 15 frames per second is sufficient to perform activity recognition, only one image out of two is considered.

The detection of medication intake has been tested with 48 sequences from 33 videos taken with three different persons. 31 sequences show real medication intake and 17 others show other activities (lures), like eating or manipulating the bottles. Results for all video sequences are summarized in Table I.



a b

Fig. 1. a) Mobile objects detected and tracked and b) corresponding skin color likelihood map.

TABLE I

RECOGNITION RESULTS FOR 48 SEQUENCES.

Activity	Detected	Not Detected
Medication Intake	28	3
Other (lures)	1	16

IV. DEMONSTRATION

The first part of the demonstration will consist of examples of mobile object tracking for different persons and with different camera points of view. The video sequences will show the ability of the system to track hands (separated and in contact) and medication bottles even if they are completely occluded by the hands.

In the second part of the demonstration, examples of activity recognition will be presented. A typical analysis of a video sequence during medication intake involving multiple bottles is shown in Fig. 2.

V. CONCLUSION

Considering that experimentation has been done in controlled conditions, we expect that adaptations will be needed for the medication intake monitoring system to work within real home environments. However, the presented algorithms have shown to be effective and promising.

We have tested the activity recognition process for medication intake, but it should be also applicable to other everyday home activities.

REFERENCES

[1] M. Valin, J. Meunier, A. St-Arnaud and J. Rousseau, "Video Surveillance of Medication Intake," *Int. Conf. of the IEEE Engineering In Medicine and Biology Society*, New York City, USA, Aug. 2006 (4 pages - *submitted*)

[2] D. Batz, M. Batz, N. da Vitoria Lobo and M. Shah, "A computer vision system for monitoring medication intake," in *Proc. IEEE 2nd Canadian Conf. on Computer and Robot Vision*, Victoria, BC, Canada, 2005, pp. 362–369.

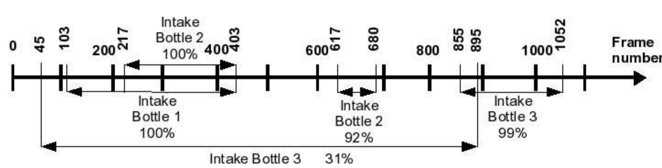


Fig. 2. Example of medication intake scenarios involving multiple bottles.

[3] S. Birchfield, "Elliptical head tracking using intensity gradients and color histograms," in *Proc. IEEE Computer Vision and Pattern Recognition*, Santa Barbara, CA, USA, 1998, pp. 232–237.

[4] N. Habili, C. Lim and A. Moini, "Hand and face segmentation using motion and color cues in digital image sequences," in *Proc. IEEE Int. Conf. on Multimedia and Expo*, Tokyo, Japan, 2001, pp. 377–380.

[5] S. Hongeng, R. Nevatia and F. Bremond, "Video-based event recognition: activity representation and probabilistic recognition methods," *Computer Vision and Image Understanding*, vol.96, no. 2, Nov., pp. 129–162, 2004.