Personenverfolgung

Projektleiterin
Prof. Dr. Gabriele Peters

Zeitraum
2007

Kontakt
Prof. Dr.
Gabriele Peters
Fachbereich Informatik
Fachhochschule
Dortmund
Emil-Figge-Straße 42
44227 Dortmund
Tel.: (0231) 755-6796
E-Mail: gabriele.peters
@fh-dortmund.de

Abstract
In this project we develop an object tracking system which is capable to handle difficult situations in a dynamically changing environment. We evaluate the concepts of the proposed system by means of the task of person tracking in crowded seminar rooms. The contributions are threefold. We propose an expansion of the condensation algorithm which results in a more stable tracking in difficult situations such as sudden camera movements. Secondly, a new way of particle diffusion is introduced which allows for the adaption of the tracking module to movements of the camera as, e.g., induced by the user. These two contributions apply not only to person tracking but to object tracking in general. The third contribution, which applies to person tracking only, consists in a flexible way how to represent a moving person. The proposed tracking module meets demands such as real-time tracking of a target person who is allowed to move freely within a group of other persons and thus can be occluded. An interactive selection of new target persons is possible and also the challenge of a dynamically changing background could be coped with.

1 Introduction
Tracking objects in a dynamically changing environment belongs to one of the most difficult problems in computer vision. Solutions of this problem are crucial not only for person tracking, which is the application on which we evaluate our proposed tracking method in this project, but also, for example, for dynamic object acquisition where sequences of objects viewed from different view points are analysed. The appearance of objects can change dramatically from frame to frame in a video sequence due to changes in the environment. Intelligent systems should be able to react dynamically to variations in the object's appearance. These variations can have many causes. If, for example, a camera moves around an object the lighting conditions can change or parts of the object can appear or disappear. In addition, the background can change which increases the degree of difficulty to assign corresponding landmarks from frame to frame, i.e., track the object. On the other hand, not only the camera can move but the object can move by itself or, even worse, can change its shape, which holds true for walking persons. The most difficult task is on hand, when also the background varies dynamically. Thus, robust object tracking in dynamic environments should be by itself adaptive to changes in the environment.

In this project we try our hand at a tracking system which reacts dynamically to recognized changes in the environment (i.e., changes in the camera parameters and interaction by the user). As a result the acquired data depend on the environmental dynamics. These concepts of active, visual object acquisition are applyed to moving persons who interact in seminar rooms. The demands of such a person tracking system are manifold and some of them are listed in the following. The person tracking system should be capable of:

1) tracking a selected target person in real-time,

2) tracking a target person within other interacting persons and in front of a dynamic background,

3) adapting to changing camera parameters such as orientation and focal length, where changes are induced either by an active camera or are initiated by a user,

4) bridging occlusions of the target object,

5) allowing for an interactive selection of a new target person during runtime.

The tracking system we present in this article is part of a larger project called Virtual Camera Assistant. The Virtual Camera Assistant is a prototype of a semi-automatic system, which supports documentary video recordings of indoor events such as seminars. It combines an automatic person tracking system with a controllable pan-tilt-zoom camera. Depending on the output of the tracking module the camera parameters (i.e., orientation and focal length) are determined automatically in such a way that the observer gets a natural impression of the recorded video (e.g., without jerky movements) and the whole movie appears pleasant to the eyes. In addition, (that is the „semi" in „semi-automatic") the user of the Virtual Camera Assistant is able, on the one hand, to control the camera parameters manually as well and, on the other hand, to select the target person to be tracked interactively at any time during recording.

A distinctive feature in comparison to other systems for automatic video production is the possible scenario of person tracking in highly cluttered scenes and within groups of interacting persons. In this article we omit a detailed technical description. Rather, we concentrate here on the presentation of some results.

2 Test Sequences

We carried out our experiments on three sequences with interacting persons in a seminar scene. The target person is selected interactively by the user and tracked autonomously. The test sequences (TS) are characterized as follows with increasing degree of complexity:

(TS1) Mainly one slightly moving person only at a time, relatively static camera, (typical seminar scene with moving instructor and relatively static audience)

(TS2) Mainly one moving person only at a time, but additionally a zooming and panning camera following the person's movements (and thus a dynamically changing background)

(TS3) Many simultaneously but non-uniformly moving persons with many crossing roads and mutual occlusions; in addition, we have interaction by the user, who selects different target persons

A subsequence of TS1 is shown in figure 1, subsequences of TS2 are displayed in figures 2, 3, and 4, and figures 5, 6, and 7 show subsequences of TS3.

3 Results

The results reported in this section are obtained mainly by visual inspection of tracking examples of subsequences of the test sequences TS1, TS2, and TS3. In the image examples the current estimated position and size of the target person are marked by a yellow rectangle with a red cross in its center. As soon as the rectangles are colored red the target person was lost (as, e.g., shown in figure 1). (An exception is shown in figure 4 where the large red rectangle has a different meaning.) If the tracking was successful blue bars in the bottom left of the yellow rectangles visualize the confidence value of the estimation.

3.1 Lost and Found – Functionality of the Particel Filter

Figure 1 illustrates the functionality of the particel filter, which is able to relocate a totally lost target object by employing multiple hypotheses. In particular the third, fourth, and fifth image show the expansion of the particle cloud in the case of a loss of the target person. But the cloud is attracted again by the correct target person as soon the target is recognized again by at least one hypothesis.

Figure 1. Lost and found – functionality of the particle filter. Frames of test sequence TS1.

3.2 Exposure to Changes in Camera Parameters – Pan

Figure 2 shows a scene with a moving person and a continuously panning camera.

The person tracking is stable and precise even if the background is partly cluttered with other persons, although the confidence values in images 6 and 7 of this figure decrease significantly as visualized by the short confidence bars.

Figure 3 shows another example of a sequence with a panning camera. Here the target person has a large velocity itself, and in addition, also the focal length of the camera varies. Even if the camera captures the target object only partly at the edges of a frame the estimations remain on the target.

3.3 Exposure to Changes in Camera Parameters – Focal Length

The other camera parameter that should be variable without affecting the tracking module is the focal length. Figure 4 shows an example where the camera zooms into the scene. The person is tracked robustly across the different focal lengths and with high confidence values.

In figure 5 another example for the exposure of the tracking system to varying focal lengths of the camera is displayed. The camera zooms out of the scene and as soon as the person who disappeared in the previous frames reappears, its position and size are estimated correctly again, although now displayed at a different focal length.

3.4 Occlusions

Occlusions belong to the most difficult challenges to deal with in tracking systems. The number of occlusions in sequence TS3 is quite high. But the target persons are relocated altogether after a short period of time while they were occluded, frequently only a few frames after they reappeared in the captured scene. Figure 6 shows three example scenes from TS3 which are effortlessly bridged by the tracking module. In the third image of the bottom row, for example, one can infer

Figure 2. Exposure to chnages in camera parameters – pan. Depicted are frames of sequence TS2.

Figure 3. Exposure to changes in camera parameters – truncation at frame edges. Frames of sequence TS2.

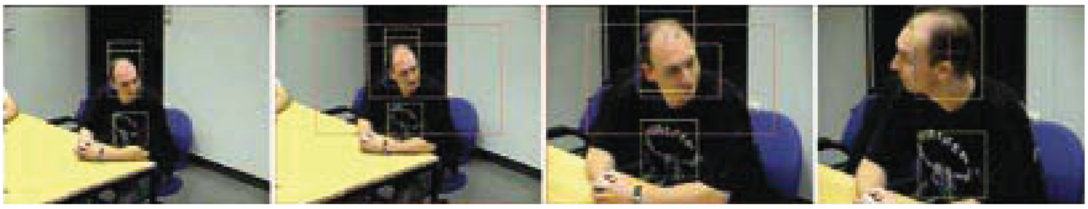

Figure 4. Exposure to changes in camera parameters – focal length. Frames of sequence TS2.

Figure 5. Exposure to chnges in camera parameters – successful relocation after zooming out. Frames of sequence TS3.

Figure 6. Occlusions. Three sample scenes for TS3 with occlusions are shown.

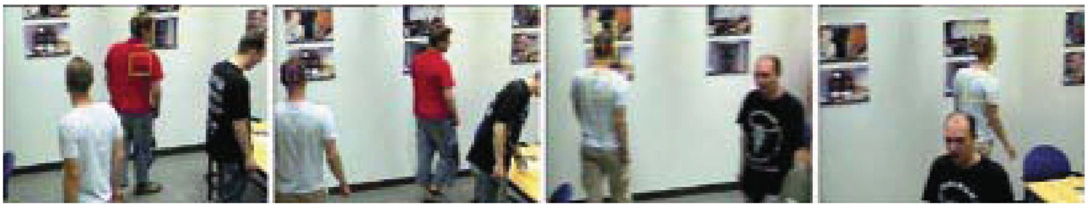

Figure 7. Interactive selection of another target object. Frames of sequence TS3. In the second image the man with the white T-shirt is selected for tracking instead of the previously tracked man with the red T-shirt. He is successfully taken over by the tracking module.

from the length of the confidence bar the decreased weighting of the estimation for the tracked woman. Nevertheless, she is tracked on robustly after her total occlusion.

3.5 Interactive Selection of Another Target Object

The selection of a new target person is done by the user of the system. She has to mark the head of the new target person. In TS3 the user has changed the target person several times. Figure 7 shows an example where the system immediately takes the control and tracks the newly selected person.

4. Weaknesses of the Proposed Tracking System

Some weaknesses of the system striked during the experiments. One problem consists in the fact that the object profile consists of color information only. Although the expansion of the basic profile by a second, dynamic color histogram area below the main rectangle improves the tracking significantly there still exist cases in which the particle cloud expands after an occlusion of the target object, and rearranges at a wrong image area with similar colors. It can happen that, even if the target object reappears in the secene, the particles remain at the wrong position until the similarity drops below the threshold and some particles detect the target again by chance.

Another weakness is the strong dependence of the tracking success on the initial selection of the target object by the user in form of a rectangle. Especially during camera movements the marking of the target can be too unprecise, resulting in the incorporation of wrong color information in the profil. This holds true as well for target persons who display a laterally crooked posture when the extended profil with the dynamically positioned, second rectangle area is employed. As the initial selection always places the bottom rectangle vertically below the head region, in this case it is positioned partly on the background, resulting in wrong color histograms in the object profile.

5 Conclusions

In this project we develop an object tracking system which is capable to handle difficult situations in a dynamically changing environment. We evaluated the concepts of the proposed system (e.g., an improved version of the condensation algorithm or particle diffusion adaptive to variations in the environment) by applying it to the task of person tracking in crowded seminar rooms. The demands made on the system comprise robust real-time tracking of a target person who is allowed to move freely within a group of other persons and thus can be occluded. Furthermore, the background may change from frame to frame, and the tracking method should cope with dynamically varying camera parameters as, for example, induced by a user. In addition, the user of the system should be enabled to interactively select a new target person during tracking of another person.

Our contributions are threefold. First, we proposed an expansion of the condensation algorithm which results in a more stable tracking in difficult situations such as sudden camera movements. Secondly, we introduced a new way of particle diffusion which allows for the adaption of the tracking module to movements of the camera. These two contributions apply not only to person tracking but to object tracking in general. The third contribution consists in a more flexible way how to represent a person than propagated in previous publications. This contribution applies to person tracking only.
Summarizing, the proposed tracking module mostly meets the postulated demands. Real-time tracking and the interactive selection of new target persons are possible. The challenges of a dynamically changing background and multiple occlusions could largely be coped with.

Ongoing research concentrates mainly on an expansion of the current object profiles, which are based on color information only. The incorporation of additional features could disambiguate situations where the current system fails because of similar color distributions in target and background. Moreover, it seems to be promising to synthesize a behavioral model which allows for the prediction of future directions of moving persons. This could represent the basis for an even more advanced and intelligent automatization.

References

- Gabriele Peters and Martin Kluger, Adaptive Object Tracking in Dynamic Environments Applied to Interacting Persons in Seminar Rooms, submitted to: Image And Vision Computing, Special Issue on Computer Vision Methods for Ambient Intelligence, 2008.

- Martin Kluger, Partikelfilterbasierter, virtueller Kameraassistent zur Verfolgung einer Person in einer Gruppe von, diploma thesis, Technical University Dortmund, 2006.