1. Personal Data

Name
Maarten Liebens

E-mail address
maartenliebens@hotmail.com

Home Institute
Catholic University of Leuven
Department of mechanical engineering
Division Production Engineering, Machine Design and Automation
Celestijnenlaan 300B
B-3001 Heverlee (Leuven), BELGIUM
Supervisors: Prof. Dr. Ir. Dominiek Reynaerts
Prof. Dr. Ir. Hendrik Van Brussel

Host Institute
Osaka University
Graduate School of Engineering
Department of Computer-Controlled Mechanical Systems
Active Intelligent Systems Laboratory
2-1 Yamadaoka, Suita, Osaka 565-0871, JAPAN
Supervisor: Dr. Eng. Jun Miura

2. Executive Summary

The research I performed is situated in the world of the camera surveillance. It presents a new approach to the problem of image-based tracking of multiple persons in a heavy occluded space using a single camera.

The presence of heavy occlusions results in uncertain measurement data. Examples of heavy occlusions are objects which impede the observation of a person, the overlap of multiple persons, etc. This measurement uncertainty can be partially by-passed if the process knows more about a person's expected size, i.e. person model. This way the observed measurement can be improved using the introduced person model. Also the uncertainty of the measurements will be calculated with this person model. Subsequently, the improved measurement is used to estimate the person's state (i.e. position and velocity) in the Kalman filter resulting in a more robust tracking.

Next, tracking multiple persons jointly implies the need for a data association technique. The technique used in the research is the Joint Probabilistic Data Association (JPDA) filter which calculates the a posteriori probabilities of the measurements having probably originated from the tracked persons. More information about this topic in section 4.

Finally, the approach has been implemented and tested on a single static camera bearing in mind that it will be applied on a mobile camera or robot. Former mobile approaches make mostly use of laser-range scans wherefrom 3D data is gathered. The approach presented here will verify whether the use of a single camera, wherefrom only 2D image-based data is gathered, can deliver satisfactory tracking results using the Kalman and JPDA filter.

But besides the performed research, there was some time for extracurricular activities during the weekends. Travelling, sight-seeing and leisure. To put it briefly, my participation to DeMaMech was a success and a once-in-a-life-time challenge.

3. Travel Schedule

2005-09-18 Zaventem – Helsinki
Helsinki – Osaka

2006-02-23 Osaka – Helsinki
Helsinki – Zaventem

4. Research

4.1. Introduction

The research is situated in the world of the camera surveillance. The purpose of such a surveillance system is to detect and to prove the presence of a person in the observed space. Another important feature is to follow (track) the persons. Possibilities during this tracking are gathering information about this person, for instance the size, the color of skin, the dynamics and motion, etc.

More specific, this research presents a new approach to the problem of tracking multiple persons jointly using a single camera and assuming a rather continuous motion of the persons. The assignment was to design and implement an algorithm what was able to perform this specifications. Another specification is the future appliance of the algorithm on a moving camera. Former approaches make mostly use of laser-range scans wherefrom 3D data is gathered. However, 3D laser sensors are expensive and take more time to get the data. Other 2D laser sensors only provide range data on a scanning plane. Stereo vision can be used, but needs two or more cameras and extra processing. The Multiple Person Tracker (MPT) described in this research performs image-based tracking using a single camera.

4.2. Person detection

Before tracking a person, the MPT needs to detect one. Based on the color of skin, a person's face can be easily extracted. However, the area of a person's face is only a small prove of its presence and not sufficient to investigate the person's motion and dynamics. For this reason, the MPT builds a color model from the clothing of the upper part of a person's body and initializes its tracking. Charged with this color model, the MPT can extract the regions corresponding to the color model. Using the center of gravity of the extracted regions, the MPT tracks a specific person by calculating its state, namely the position and velocity. However, due to occlusion and lack of information about a person's true size and motion over time, the system will collect measurements perturbed with noise. Subsequently, the algorithm will estimate a person's state using these noisy measurements. In some cases, the noisy measurements can be corrected. One of the most well-known and often-used tools for tracking is the Kalman filter.

4.3. State estimation

Mathematically, the state estimation problem for a specific person can be subdivided into and modelled by the person's dynamic system and the measurement system.
The dynamics of a person are modelled by the equation

$$\mathbf{x}_{k+1} = \mathbf{A}\ \mathbf{x}_k + \mathbf{w} \qquad (1)$$

where $\mathbf{x}_{k+1}$ is the 4-dimensional state vector (i.e. position and velocity in two directions) at frame k, $\mathbf{A}$ is the known transition matrix which maps the previous state on the current state based on the previous velocity and $\mathbf{w}$ is the process noise, assumed to be normal distributed with zero mean and variance $\mathbf{Q}$.
The measurement system is modelled as follows. If the measurement originates from the person, then

$$\mathbf{z}_k = \mathbf{H}\ \mathbf{x}_k + \mathbf{v}_k \qquad (2)$$

where $\mathbf{H}$ is the known (2 x 4) matrix which maps the true state on the observed state and $\mathbf{v}_k$ represents the measurement noise, also assumed to be normally distributed with zero mean and variance $\mathbf{R}_k$.
Former approaches using laser sensed data apply a constant variance $\mathbf{R}$ into the Kalman filter. However, the MPT senses the populated space with a camera and extracts measurements having their own size, characterized by height and width. Compared with the expected height and expected width of a specific person, i.e. person model, the (2 x 2) variance of the measurement's noise can be calculated.

The MPT's estimation of the person's state $\mathbf{x}_k$ at time k, given data up to time i, is denoted $\hat{}\mathbf{x}_{k|i}$. The error in the state estimation is represented by its covariance matrix $\mathbf{P}_{k|i}$. In the absence of measurement origin uncertainty, the discrete-time Kalman filter calculates the person's state estimation and its covariance by correcting their predictions as follows

$$\hat{}\mathbf{x}_{k|k} = \hat{}\mathbf{x}_{k|k-1} + \mathbf{W}_k\ .\ (\sim\mathbf{z}_k) \qquad (3)$$
$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{W}_k\mathbf{S}_k\mathbf{W}_k^T \qquad (4)$$

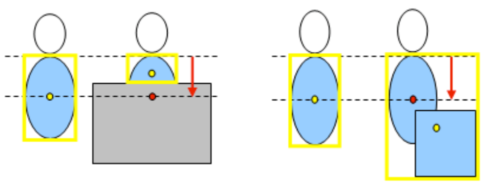
where the innovation vector $\sim\mathbf{z}_k$ is the difference between the observed measurement and the predicted person's state

$$\sim\mathbf{z}_k = \mathbf{z}_k - \mathbf{H}\ .\ (\hat{}\mathbf{x}_{k|k-1}) \qquad (5)$$

where $\mathbf{S}_k$ is the covariance matrix of the innovation $\mathbf{S}_k = \mathbf{A}\ \mathbf{P}_{k|k-1}\ \mathbf{A}^T + \mathbf{R}_k$.

4.4. Measurement Correction Using Person Model

Fortunately, by using another feature presented in this research, measurements can be corrected when the size of a measurement changed gravely compared to the true expected size (i.e. person model) of an associated tracked person. This correction can only happen if there are resembling features shared by previous and current extracted measurements. For instance, the algorithm can determine that only the upper part of a person is visible when the person is partially occluded by an object in the observed space. By this correction the uncertainty or noise variance of the measurement will be reduced.



(a) (b)

4.5. Data Association

4.5.1 Problem formulation

The MPT doesn't only have to deal with the inaccuracy of the measurements which is modelled as additive noise. It also has to deal with additional uncertainty caused by the uncertainty of the origin of the measurements and the correctness of associations of measurements to a person.
When tracking a person in a heavy occluded space and where more persons occur with same colored clothing, there might be several candidate measurements originating from the person in track. The analysis for this appearance requires knowledge of measurement-to-person correspondence, that is, do two measurements taken at different times originate from the same person and what is the possibility that the measurement currently originates from this person? Which measurements are used to correct the predicted state of a person in order to estimate the person's position? This is the formulation of the motion correspondence problem in the world of computer vision where it is referred to as the data association problem.
A first step toward this data association is to estimate the likelihood of a measurement originating from a specific person. One such measure used by the MPT is the Mahalanobis distance which is a statistical validation test. The purpose of this validation technique is to reduce the computational load by avoiding having to evaluate all the measurements in the observed space.
The next step is the actual motion correspondence and has extensively been studied for person tracking and the surveillance community. The MPT applies the Joint Probabilistic Data Association (JPDA) technique. Former approaches to track multiple persons already investigated the use of the JPDA filter. However, most of these approaches used laser sensed data and there are not many works on applying the JPDA theory to image-based tracking.
To explain the JPDA theory, suppose that m measurements are extracted from the frame at time k are denoted $\mathbf{z}_j$ for measurement j and the collection of all measurements up to time k as Z^k. The corresponding individual innovations at time k for person t are:

$$\sim\mathbf{z}_j^t = \mathbf{z}_j - \hat{}\mathbf{z}^t \qquad (6)$$

where $\hat{}\mathbf{z}^t$ is the predicted measurement for person t. The time subscript k will be suppressed from all variables except from **P** and Z for convenience.
The main purpose of the JPDA filter is to combine all these measurements' innovations weighted with their association probability in order to correct the prediction of the state of person t in the Kalman filter:

$$\sim\mathbf{z}_t = \sum_{j=1}^{m} \beta_j^t \, (\sim\mathbf{z}_j^t) \qquad (7)$$

where β_j^t is the association probability that measurement j originated from person t, and $\sim\mathbf{z}_j^t$ is the innovation of measurement j compared to person t.

4.5.2. Joint probabilities

The key to the JPDA algorithm and to the calculation of the association probabilities β_j^t is the evaluation of the feasible joint events χ. The feasible events are those joint events in which only one measurement originates from only one person and vice versa. The probability of the joint event conditioned on the measurements up to time k is denoted as $P(\chi \,|Z^k)$. After obtaining these feasible joint events χ and their probabilities, the association probability β_j^t is the probability of the event where measurement j is associated to person t conditioned on all measurements up to the present time.

The probability of a joint event conditioned on all measurements up to the present time k can be presented as:

$$P(\chi \mid Z^k) = p(\chi \mid \mathbf{z}_1,\ldots,\mathbf{z}_m, Z^{k-1}) = p(\chi \mid \mathbf{z}_1,\ldots,\mathbf{z}_m, X^k) = 1/c\ p(\mathbf{z}_1,\ldots,\mathbf{z}_m \mid \chi, X^k)\ P(\chi \mid X^k) \quad (8)$$

The measurements Z^k in (8) can be split up in the current detected measurements $\mathbf{z}_1,\ldots,\mathbf{z}_m$ and the measurements Z^{k-1} obtained before time k. However, to consider the condition of all measurements obtained before time k is complex and laborious. To overcome this problem, the distributions of the predictions of the persons' states X_k are introduced in equation (8). All previous measurements are indirectly encapsulated in these predictions. On the other hand, the prediction and estimation problem are assumed to be Markovian. The state of a person depends only on the measurements obtained at time k.
Using Bayes' rule, the probability of a joint event χ can be presented as in equation (8) where $P(\chi \mid X^k)$ is the prior probability of an event conditioned only on the distribution of the current predicted states X_k. It is "prior" in the sense that it does not take into account any information about the current measurements $\mathbf{z}_1,\ldots,\mathbf{z}_m$.
The term $p(\mathbf{z}_1,\ldots,\mathbf{z}_m \mid \chi, X^k)$ is the conditional probability density of the current measurements $\mathbf{z}_1,\ldots,\mathbf{z}_m$ given the association event χ and the distribution of the current predicted states of the persons.
The term 1/c is the prior probability of the measurements conditioned on the past data, and acts as a normalizing constant.

4.5.3 Association probabilities

The probability β_j^t that measurement j, from the collection of all m validated measurements, belongs to person t, may now be obtained over all feasible events χ for which this condition is true ($\omega_{jt}(\chi)$ will be one only if measurement j originates from person t):

$$\beta_j^t = \sum_{\chi} P(\chi \mid Z^k)\omega_{jt}(\chi) \qquad \text{for all measurements } j = 1,\ldots,m \text{ and}$$

for all tracked persons t = 1,…,T.

These probabilities are used to calculate the combined innovation (7) for every person which will be used in the Kalman filter to calculate the estimation for a person's state.

5. Exchange student life

5.1. Introduction

I spent five months in Osaka and did my research at Osaka University in the laboratory of active intelligent systems. In these five months I also had the opportunity to visit and enjoy the culture of Japan together with my DeMaMech and Japanese friends.

5.2. Accommodation

During my stay in Japan, I lived in an international student house about six kilometres from the university. Although international, the staff couldn't speak English at all. Kitchen and lavatories were in quite poor condition. As a result I never cooked at the dormitory and I could take only a seven minute shower for 50 cent.
Fortunately I could also stay overnight in the laboratory in a special installed bedroom, the "holy place". At the university I could even take a shower and buy breakfast. But most of the times I used to sleep at the dormitory.

5.3. A normal day at the campus

A working day started for me around 8h30. I could arrange my working hours by myself on one condition that I finished my research after five months. On my way to campus, in the beginning by bike and at the end by train, I bought my breakfast. After breakfast I started working from 9h30. Every Thursday I had a weekly meeting with my professor. My work consisted mainly on gathering papers and books from internet or library and learning the theory which I had to implement in an algorithm. With my questions, I could always go to my professor or a PhD student. The communication was in English and most of my lab members could speak English quite well. My presence in the lab was also a way to help them in practicing their English. I started with taking Japanese class but stopped after a while because it was time consuming and my stay of five months is too short to learn the Japanese language.
For lunch and dinner I went to the student restaurant with my lab members or the DeMaMech students. The food was very nice and is one of the things I miss very much. After 20h00 I had the choice to go to gym with some DeMaMech friends, to go home or simply to stay, as already mentioned before. Although the rule of the laboratory was to be present on Saturdays for meetings, I wasn't obliged to this rule. But after some weeks I also started to go to campus on Saturdays.

5.4. Leisure time

In the weekends I had the opportunity to enjoy the Japanese culture. Osaka is a very good location to start an excursion to all this beauty. Kyoto, Nara and Kobe are quite near and accessible in about an hour. Hiroshima and Tokyo are easily accessible by a high speed train or night bus. For New Year's Eve, we had a DeMaMech reunion in Tokyo. I also had the opportunity to go skiing with my laboratory for two days. Impossible without the help of my Japanese fellow students because sometimes the people in the street can't speak English that well. Or sometimes they are just shy.
Other free time I spent with the DeMaMech and other exchange program students by going down town Osaka. There is a big choice in restaurants or bars to visit during the night. And in general, the people in Osaka are very friendly, even a noticeable difference compared to people from Tokyo towards foreigners.

6. Summary

I stayed five months at Osaka University performing my master thesis in order to finish my studies. But staying five months in Japan isn't long enough to discover the full beauty of its country, to learn the Japanese language, etc. But it is certainly long enough to taste its beauty, to build up the fundamentals of a life in a totally different culture, to make friends, to have the time of your life, to perform research or even simply to get to know yourself a little bit better. All of this was a success during my stay and a profit for my future (career) life.

Therefore, I would like to thank my Japanese professor Dr.Eng. Miura who gave me the opportunity to join his laboratory of active intelligent systems, who supported me during my research and discussed with me about relevant topics. I also would like to thank Prof.Dr. Tomiyama and the other founders of the DeMaMech exchange program. And last but not least, thanks to Prof.Dr.Ir. Van Brussel and Prof.Dr.Ir. Reynaerts, the KU Leuven participates also in the DeMaMech exchange program.