Visual Tracking and Motion Determination using the IMM Algorithm

P. Tissainayagam and D. Suter
[prithi, suter]@zeno.eng.monash.edu.au
Dept. of Electrical and Computer Systems Engineering
Monash University, Clayton Campus, Clayton, Vic. 3168, Australia.

Abstract

In this paper we present a feature tracking system with automatic motion determination of features in an image sequence. The positions of features (corners) extracted in the first frame of a sequence are estimated and predicted in the subsequent frames by using an extension of Bayesian multiple hypothesis technique (MHT [2]) based on different motion models. The tracking of features is based on the Interacting Multiple Model (IMM) [1]. This paper shows how the IMM algorithm combined with a MHT framework can be used in a visual tracking scenario. We considered different order (types) velocity and acceleration models for the IMM algorithm and applied them to two image sequences, the PUMA sequence and Toy car sequence. The study shows that the method proposed can distinguish between different motions depicted in an image sequence with very good tracking results.

1. Introduction

In the recent years there has been an interest in using surveillance tracking techniques for visual tracking applications. One such proposal is outlined in [2] by Cox et al. In this paper we combine the system in [2] with an IMM to track and determine the motion of objects in a long dynamic image sequence. An important reason for considering the MHT algorithm is because the MHT is one of the statistical data association algorithms that integrates all the capabilities such as track initiation, track termination, track continuation, explicit modelling of spurious measurements, and explicit modelling of uniqueness constraints.

In this paper we consider the trajectories of 3 features appearing in the first frame of a sequence and analyse their motion. Our contribution is primarily on determining the motion model appropriate for the feature and introducing the MHT/IMM tracker. In section 2 & 3 we outline the MHT and feature extraction procedure used. Section 4,5 outlines visual tracking in general and the IMM technique implemented, and section 6 provides the results and a brief discussion. Finally section 7 gives the conclusion.

2. Multiple Hypothesis Algorithm

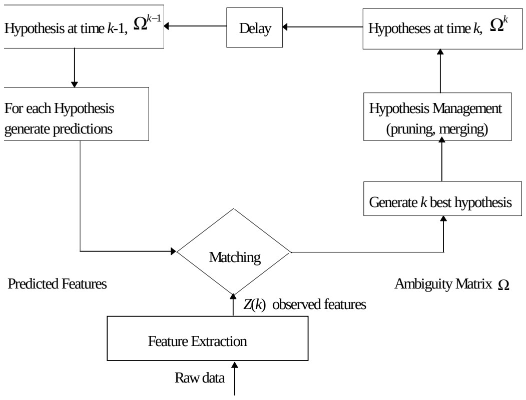

Fig. 1: Outline of the multiple hypothesis algorithm

The Multiple Hypothesis Tracking (MHT) algorithm was originally developed by Reid [7] in the context of multi-target tracking. Cox et. al. later modified the MHT with significant computational efficiency. Fig.1 shows a typical MHT framework. See [7,2] for complete details.

It has been shown in [2] that the predicted next hypothesis (Θ_m^k), given measurements upto k (Z^k) will be given as follows.

$$P\left\{\Theta_m^k | Z^k\right\} = \frac{1}{c}\lambda_N^{\nu}\lambda_F^{\phi}\prod_{i=1}^{m_k}\left[N_{t_i}[\mathbf{z}_i(k)]\right]^{\tau_i}$$
$$\left\{\prod_t (P_D^t)^{\delta_t}(1-P_D^t)^{1-\delta_i}(P_\chi^t)^{\chi_t}(1-P_\chi^t)^{1-\chi_i}\right\}P\left\{\Theta_{l(m)}^{k-1}|Z^{k-1}\right\} \quad (1)$$

Using (1) (with appropriate pruning strategy) combined with a tracking system (IMM) is what we are going to use to track features.

3. Feature Extraction

To use the multiple hypothesis tracking (MHT) technique for visual tracking, it is necessary to extract the features to be tracked in every frame of the image sequence. The positions of features appearing in the first frame are predicted in the subsequent frames (matched /discarded) by the MHT. The MHT uses the *Mahalanobis* distance as the main validation gate, and further, to reduce the search area uses a correlation matching strategy (based on a 5x5 patch size) [2].

For the PUMA sequence, we used the corner detector proposed by Harris [3] to extract the corners. For the Toy car sequence we used a variant of the Lucas and Kanade's corner detector [4]. We maintained the number of corners extracted per frame to around 20-50 for both sequences purely for clarity.

4. Tracking Features

For a visual tracking system to be efficient and reliable, the tracker needs to evolve around a correct motion model. Most visual tracking systems assume a single motion model. This assumption can be wrong if there's a change in motion depicted in the image sequence or there's multiple motions of an object. It is also quite well known that a potential weakness of an estimator based on a single model is that it can lead to under-modelling and/or over-modelling [8].

To overcome this limitations, one solution is to use a number of filters based on different motion models (sub-models) covering the range of possible expected observed motions, and to some how combine the estimates from these filters based on the expectation of each model being the correct descriptors of the features' motion. Such a system can be achieved with a multiple model filtering (MMF) based algorithms [6]. As well as improving estimation accuracy, such systems could help in segmenting a scene into independently moving objects. It has been proposed that the segmentation process may be performed by utilising the confidence/ belief measures generated by the individual filters that make up the multiple motion model system. If all objects in a scene are assumed to be rigid, all points on an object will move in an identical fashion, i.e, with the same motion model.

MMF algorithm [6] originally proposed by Maybeck assumes that the system state update is a linear combination of each filter in the filter bank weighted by a probability factor. Such a system is discussed in [6,8,9] and has been implemented for a variety of applications. A potential problem in such a system is that it is assumed that the system obeys one of a finite number of models. No switching from one mode to another occurs during the estimation process. They also have both continuous (noise) uncertainties as well as discrete uncertainties. Because the MMF based filters do not cater for motion (model) switching, these methods can fail or even converge to the wrong motion model [9]. There are ad hoc modifications which are available for MMF algorithms to cope with model switching [9]. In spite of these modifications, the mismatched filter's errors can grow to unacceptable levels. Thus, reinitialization of the filters that are mismatched is in general needed [1,9].

5. IMM Algorithm

The IMM algorithm [1] (fig. 2) is able to cope with mode changes during motion transition and is capable of switching from one mode of motion to another efficiently. In our experiments (for this paper) we have only used one mode of motion with multiple motion models. The main objective of this paper is to determine the motion of the features reliably while maintaining a good tracking performance.

For our analysis, firstly we used the IMM algorithm with 3 second order motion models. These were a constant acceleration model (M1), a constant velocity model (M2) and a constant coordinated turn model (M3) (see [9] for complete description of motion models and detailed results). Secondly we used a third (M4) and a second (M1) order acceleration model for the PUMA sequence and a third order acceleration (M4) and a second order velocity model (M2) for the Toy car sequence. The results of the experiments are graphically presented in figures 3,4.

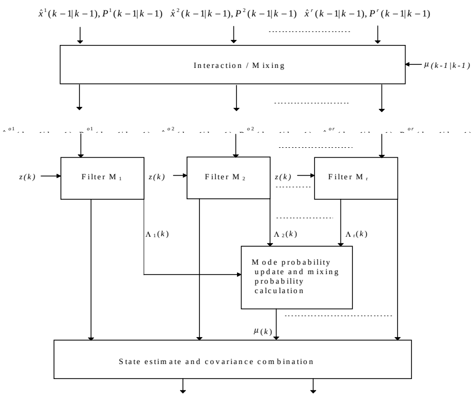

Fig 2: IMM Algorithm (one cycle)

6. Results

Figure 4 shows tracking results obtained by using the MHT algorithm based on a single motion model. From these figures it is quite clear to the naked eye that the constant acceleration model gives the best tracking performance for the PUMA sequence and the constant velocity model gives the best tracking performance for the Toy car sequence. However, we show experimentally that the correct motion can be 'discovered' by the MHT/IMM tracker (fig.3).

For the first experiment, with 3 motion models in the IMM frame work, we initialised the probability of selecting a model to 0.33. That is, at the start all models have an equal chance of getting selected. For the second experiment, with two models, both models were initialised to a selection probability of 0.5. Results show that the MHT/IMM tracker converges to the correct motion model while maintaining good tracking performance.

7. Conclusion

Our study has shown how the Multiple Hypothesis Tracking (MHT) technique combined with an Interacting Multiple Model (IMM) algorithm can discriminate between different motions described by an image sequence. The results have provided evidence of our method being able to identify different motions while maintaining good track result. A drawback of the system is that the features need to be extracted independently of the MHT. A coupled feature detection and tracking mechanism is worth investigating.

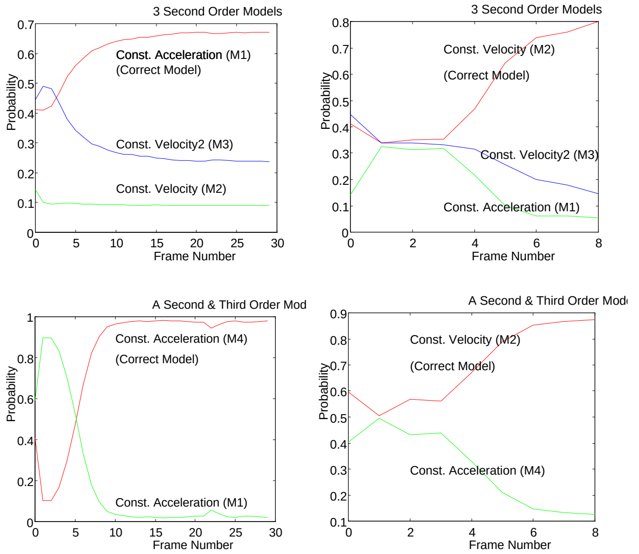

*Fig.3: Model selection probabilities. (a) PUMA Seq., constant acceleration model selected as the **correct** motion model. See corresponding track result in fig. 4. (b) Toy Car Seq., constant velocity model selected as the **correct** motion model. See corresponding track result in fig. 4. (c) 3rd order acceleration model selected over 2nd order acceleration model. (d) 2nd order velocity model selected over 3rd order acceleration model.*

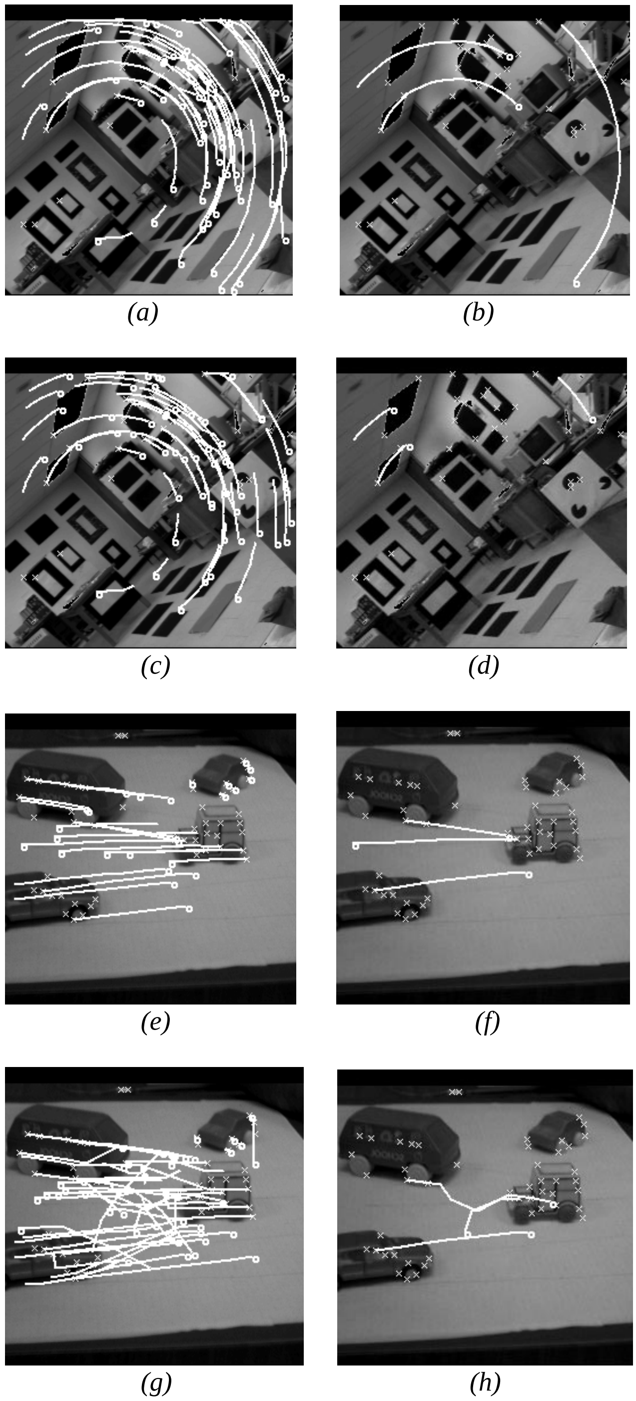

*Fig. 4: Track length of more than 6 are only displayed (frame 1of PUMA & Toy car sequences). For each track the circle indicates the end of track and the 'x' indicates the corners extracted in the first frame. (a) PUMA seq., M1 with all the tracks (**correct** model). (b) The selected 3 tracks for M1 in case (a). (c) M2 with all the tracks (**incorrect** model). (d) The selected 3 tracks for M2 in case (c). (e) Toy car seq., M2 with all the tracks (**correct** model). (f) The selected 3 tracks for M2 in case (e). (g) M1 with all the tracks (**incorrect** model). (h) The selected 3 tracks for M1 in case (g).*

References

[1] Y. Bar-Shalom and X. R. Li, Estimation and Tracking, *Principles, Techniques and Software, Artech House*, Boston, MA (USA), 1993.

[2] I. J. Cox and S. L. Hingorani, "An Efficient Implementation of Reid's Multiple Hypothesis Tracking Algorithm and Its Evaluation for the Purpose of Visual Tracking", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 2, Feb. 1996.

[3] C. G. Harris and M. J. Stephens. "Combined Corner and Edge Detector", In *Proceedings of the Fourth Alvey Vision Conference*, Manchester, Pages 147-151, 1988.

[4] B. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision", *Proc. 7th International Joint Conference on Artificial Intelligence*, pp. 674-679, 1981.

[6] P. S. Maybeck. Stochastic Models, Estimation and Control. Vol 1 & 2, *Academic Press*, NY, 1982.

[7] D. B. Reid, "An Algorithm for Tracking Multiple Targets", *IEEE transactions on Automatic Control.*, vol.24, pp.843-854, Dec. 1979.

[8] J. M. Roberts. Attentive Visual Tracking and Trajectory Estimation for Dynamic Scene Segmentation, *Phd Thesis, Dept. of Elect. and Comp. science*, University of Southampton, UK, 94.

[9] P. Tissainayagam and D. Suter, "Motion Model Selection for Visual Feature Tracking", *Tech. Report. MECS-1997-4, Monash University, Australia.*