Motion Models That Only Work Sometimes

Cristina García Cifuentes[1]
http://visual.cs.ucl.ac.uk/pubs/MotionModelPrediction
Marc Sturzel[3]

Frédéric Jurie[2]

Gabriel J. Brostow[1]

[1] University College London, UK
[2] University of Caen, France
[3] EADS Innovation Works, France

It is too often that tracking algorithms lose track of interest points in image sequences. This persistent problem is difficult because the pixels around an interest point change in appearance or move in unpredictable ways. In this paper we explore how classifying videos into categories of camera motion improves the tracking of interest points, by selecting the right specialist motion model for each video. As a proof of concept, we enumerate a small set of simple categories of camera motion and implement their corresponding specialized motion models. We evaluate the strategy of predicting the most appropriate motion model for each test sequence. Within the framework of a standard Bayesian tracking formulation, we compare this strategy to two standard motion models. Our tests on challenging real-world sequences show a significant improvement in tracking robustness, achieved with different kinds of supervision at training time.

We implement specialized dynamic models for four types of camera motion and two standard tracking models (Brownian and constant velocity). After training an SVM, we categorize each new video sequence to its most appropriate dynamic model.

The performance of the six motion models is computed over the whole dataset. No individual model performs extraordinarily overall, but each one does well on its "kind" of videos. The experiments show how predicting the most appropriate motion model leads to significant improvement of tracking robustness. We propose that such predictions can be made using a set of training videos, with shared motion properties, and using a classifier to predict the category of a new test videos. There is good reason to expect that further very specialized motion models, that are not good in general but outstanding under known circumstances, are worth developing.

frame 0 | frame 10 | frame 20

Figure 1: Example sequence with overlaid box showing the output of our specialized "forward" motion model, where the velocity and scale of objects approaching the camera tend to increase. Neither Brownian nor constant-velocity motion models are as successful at tracking interest points here.

	Individual motion models						Ideal predictions			Our method
	Br	CVel	TRight	TLeft	Fwd	Bwd	best{Br, CVel}	best{all}	manual labels	
tracking robustness ($\cdot 10^{-2}$)	42.3	43.2	37.9	37.2	44.7	43.7	49.3	56.1	52.6	51.9
$\pm$ std. dev. random runs	0.4	0.4	0.7	0.5	0.2	0.1	0.6	0.4	0.2	0.1
% best choice (± 2)	21	11	12	20	20	16	32	100	52	50

Table 1: White background: average tracking robustness of each individual motion model over all videos (default parameters). Bottom row: percentage of times *that* motion model (among six) was the best choice. Dark gray: best-case results if model is selected by either a performance-based oracle, or our inspection-based labels. Right column: tracking each video using our classifier's suggested motion model, using inspection-based training data.