ONLINE APPEARANCE MODELS FOR VISUAL TRACKING

Bachoo AK., Nicolls F.
abachoo@csir.co.za

Abstract

An ongoing research area in object tracking is the representation and online update of the target appearance. This work examines online appearance models for improved visual tracking. The primary idea is to use both background and foreground models in a statistical framework for improved target tracking. Motivating works are discussed for the general research and then results are presented for a basic tracking system. The tracker uses template matching within a particle filter. The template is updated online and a robust error function detects occlusions and outliers. The results are very promising and the tracker is robust to occlusions, white caps, low contrast and camera motion when tested in a maritime environment.

1. Online Appearance Models

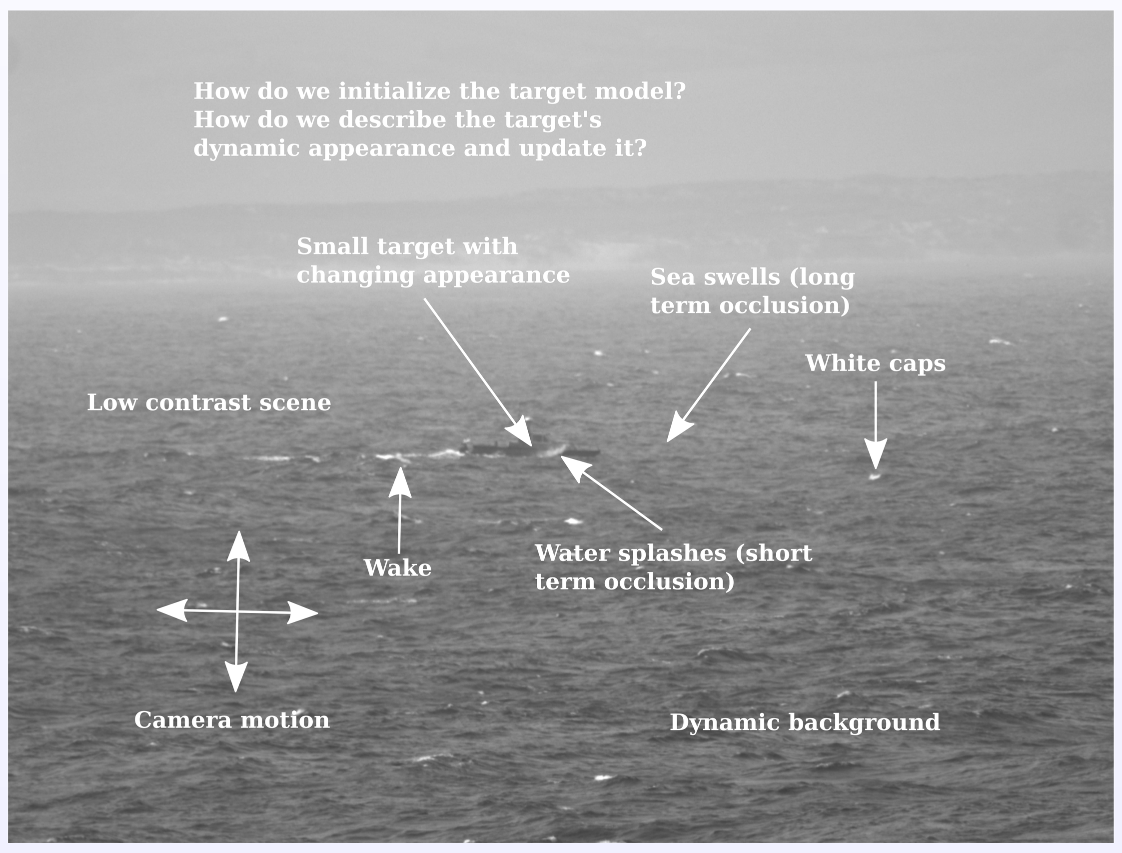
Our motivating ideas for appearance models are drawn from the mixture models of Jepson *et al.* [1] and the view-based appearance models of Morency *et al.* [2]. Our online model will incorporate different views of the target.

2. Segmentation

Segmentation will be considered for separating background and foreground pixels. Mixtures of optical flow [3] and flexible sprites [4] motivate our intention to use a layered or mask approach to the tracking problem.

3. Preliminary Work: Application

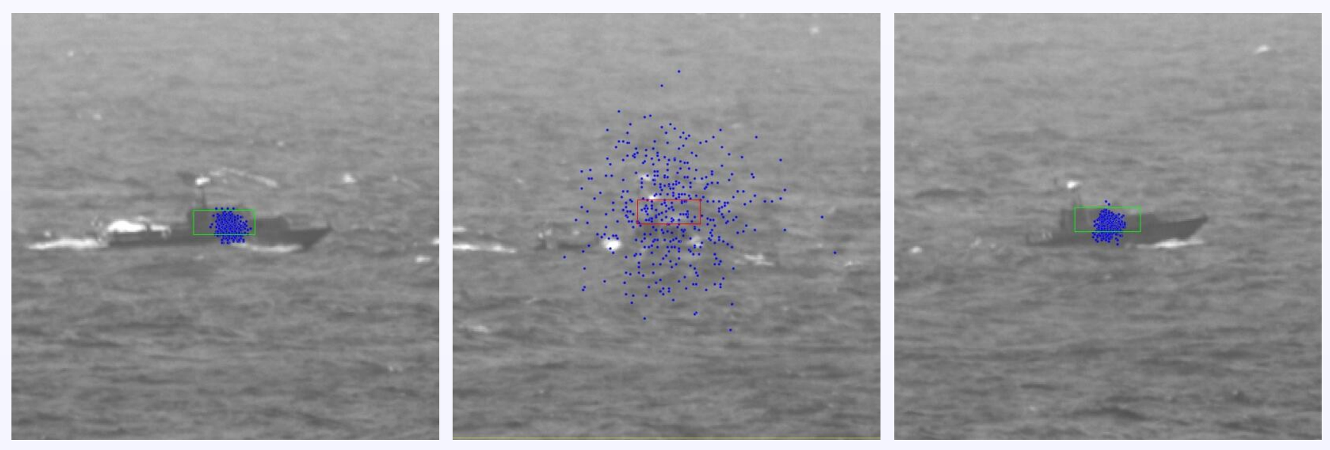
Our initial tracking system was developed for the maritime environment. The tracker uses template matching within a particle filter framework [5]. The template is updated online [6]. Occlusion detection is performed using a robust error function; the target trajectory is predicted using the motion model and estimated state vectors.



4. Preliminary Work: Results

The tracker was evaluated using three test videos of resolution 1360×1024 (shown below). The data was captured in real world conditions with camera motion, low contrast, sea swells, white caps and occlusions. The tracker operates at 15 frames-per-second on a single target with feedback to a pan-tilt unit. The system is also capable of tracking multiple targets. A track is initiated by the user.

The figure below shows the detection of a water splash on a boat, the scattering of particles and then the continuation of tracking. The particle scatter is large due to the unpredictable camera motion.



The tracker is able to track through long occlusions as shown below. A history of state estimations is maintained so as to produce a smooth prediction during occlusions.

5. Preliminary Work: Tracking Performance

Tracking results are presented for the 3 test videos (above). The average absolute pixel errors for the estimated x and y positions are shown; standard deviations are displayed within brackets and are indicative of the camera motion and prediction during occlusions. For example, the x standard deviations are larger than y due to the camera panning. Ground truths were generated manually. Errors are relatively low when tracking the target; they are much larger when outliers are detected.

Video	Global Metrics				Target		No Target	
	n	l (%)	ϕ	r(%)	x_{error}	y_{error}	x_{error}	y_{error}
1	683	83	0.15	57	6.68(7.95)	2.47(1.92)	17.42(20.48)	23.75(21.10)
2	1457	83	0.06	73	5.53(5.00)	2.27(1.60)	14.72(16.35)	2.92(1.89)
3	2278	16	0.57	53	5.82(4.84)	2.58(1.81)	31.39(28.05)	4.62(5.31)

n - number of frames in the video.
$l(\%)$ - longest consecutive track.
ϕ - the number of re-initializations per 100 frames.
$r(\%)$ - how often the tracker believes it is tracking the target (determined using outlier rejection).

6. References

[1] A. Jepson, D. Fleet, and T. El Maraghi, "Robust online appearance models for visual tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1296–1311, 2003.

[2] L.-P. Morency, A. Rahimi, and T. Darrell, "Adaptive view-based appearance models," in *Proceedings of CVPR*, 2003, pp. 803–810.

[3] A. Jepson and M. J. Black, "Mixture models for optical flow computation," in *Proceedings of CVPR*, 1993, pp. 760–761.

[4] N. Jojic and B. J. Frey, "Learning flexible sprites in video layers," in *Proceedings of CVPR*, 2001, pp. 199–206.

[5] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2001.

[6] I. Matthews, T. Ishikawa, and S. Baker, "The template update problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 810 – 815, 2004.