Robust Acquisition of Transportation Information

Jien Kato and Toyohide Watanabe

Department of Systems and Social Informatics
Graduate School of Information Science, Nagoya University
{jien,watanabe}@is.nagoya-u.ac.jp

The main obstacle to robust visual tracking is that distracting features, such as clutter in the background regions, compete for the attention of the tracker and may succeed in pulling the tracker away from foreground (target) objects in a high-level tracking process. To make the tracker reliable, it is a common practice to discriminate the foreground pixels from the background pixels. Earlier researchers have attempted to increase the robustness of the tracker by image differentiation techniques. However, as to applications such as traffic monitoring systems, typical troublesome features are the shadows of vehicles, which are not well handled by traditional techniques.

For enhancing the robustness to different lighting conditions of car trackers, we propose a tracking method based on hidden Markov models (HMMs) that classifies each small region of a traffic monitoring movie into three different categories: foreground (vehicles), background and shadow (including the shadow of moving vehicles), and from the viewpoint of advantages of HMMs at modeling temporal continuity constraints. The temporal continuity here means a pixel belongs to a certain category for a period of time. Once a pixel is inferred to be in, for example, a foreground region, it is expected to be within the same foreground region for some time. Previous work has revealed that the temporal continuity of each category becomes especially significant in a situation where intensity differences among categories are ambiguous. In consideration of this fact, the observations over time for one small region location have been modeled as a single HMM along the time-axis, independently of the neighboring regions. One difficult issue with respect to this method is how to achieve context-dependent classification so that both temporal and spatial dependence among regions is able to be incorporated into a consistent stochastic framework. This is because a scene is understood in not only the temporal but spatial context of objects within it. As an easy-to-understand example, a foreground region is highly unlikely to exist in isolation surrounded by background regions. Such spatially contextual constraints must be entertained in the interpretation of visual information.

It is well-known that Markov random field (MRF) theory provides a convenient and consistent way for modeling context-dependent entities such as image pixels and correlated features. This is achieved through characterizing mutual influences among the entities using conditional MRF distributions. Moreover, MRF used in conjunction with a statistical criterion, e.g., *Maximum a posteriori* (MAP), enable the formulation of objective functions in terms of MAP optimality principle. This MAP-MRF schema is adopted into our HMM-based tracking method so as to realize context-dependent classification in both the temporal and spatial sense. Namely, instead of finding an optimal state for each small region individually, we model the output from the HMMs at all region locations as an MRF, and employs the MAP principle together with the MRF to obtain the optimal configuration of the MRF model. At each time step, the optimal configuration of states is the MAP estimate, which is generated through a stochastic relaxation process. In this way, we achieve the integration of temporal-dependent information obtained from HMMs and spatial-dependent information obtained from the MRF under the Bayes framework.

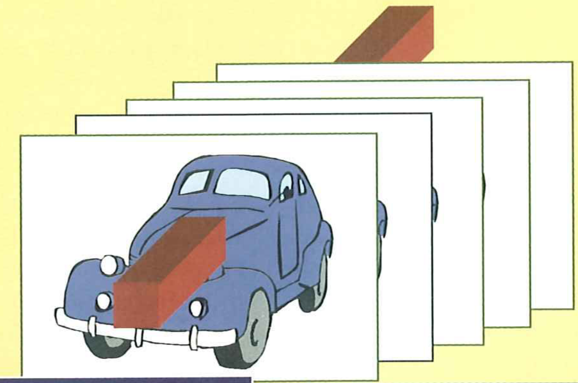
Practical tests on real-world traffic monitoring movies have shown that the proposed approach is able to effectively reduce mis-classification between ambiguous patterns, for example, cars in big shadows, dark cars, or dark regions inside cars. This implies that the approach enhances the robustness of car tracking to different lighting conditions. Because there are no constraints on the shape or movement of foreground objects imposed, the proposed approach is also applicable to other tracking problems or segmentation problems of image sequences.

Robust Acquisition of Transportation Information

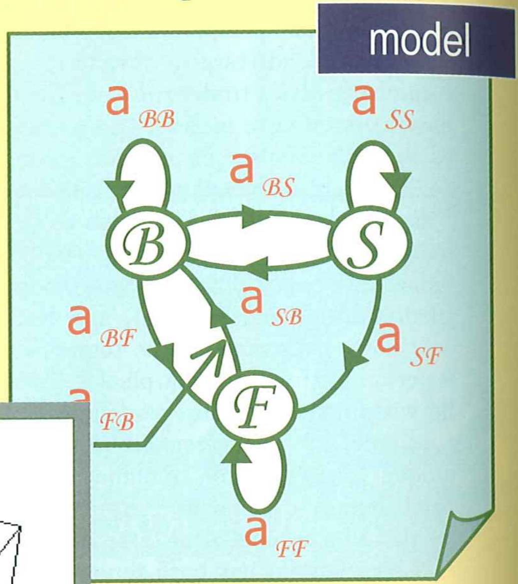
Jien Kato and Toyohide Watanabe

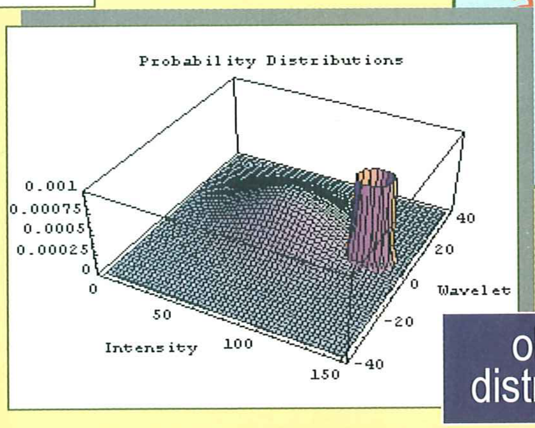
Temporal-dependent modeling by HMMs

-classify each small region into three categories: background ($\mathcal{B}$), shadow ($\mathcal{S}$) and foreground ($\mathcal{F}$)



an HMM along time-axis

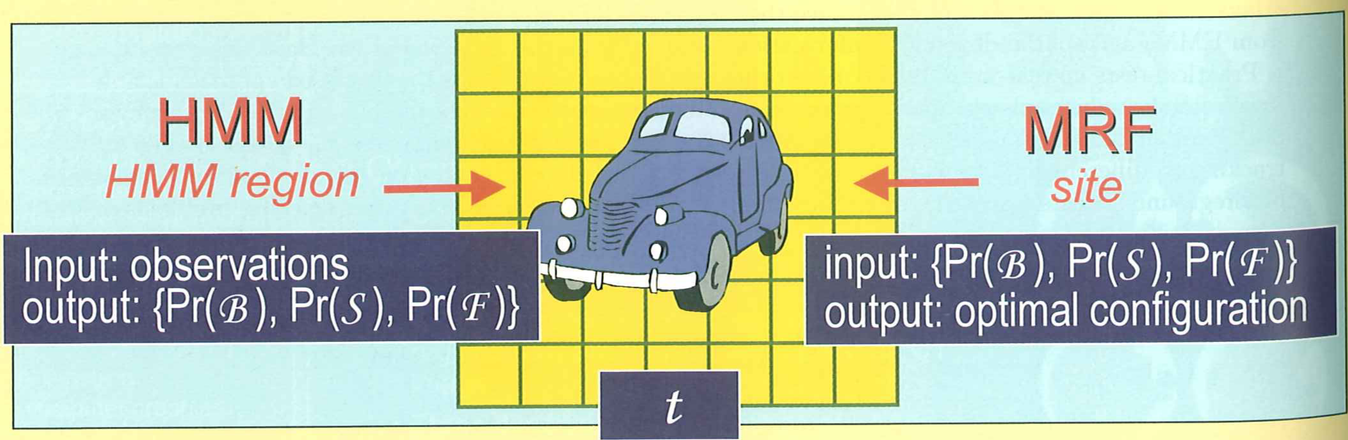




observation probability distributions for $\mathcal{B}$, $\mathcal{S}$ and $\mathcal{F}$

Spatial-dependent modeling by MRF

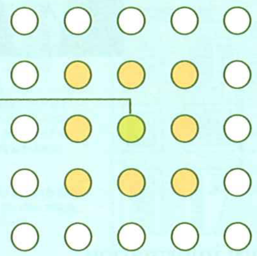
-output of HMMs is modeled as a MRF and an optimal configuration is found with MAP principle



Robust Acquisition of Transportation Information

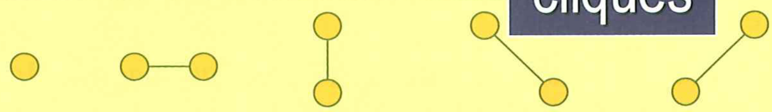
Jien Kato and Toyohide Watanabe

Neighborhood System



= neighbors of site i

cliques



MRF Modeling

$$p(f) = Z^{-1}\exp\{-U(f)/T\}, \quad \text{(in Gibbs form)}$$

$$U(f \mid \Omega) = \alpha \sum_{\{i\}\in\mathcal{C}_1} f_i + \beta \sum_{\{i,i'\}\in\mathcal{C}_2} f_i f_{i'}$$

a priori

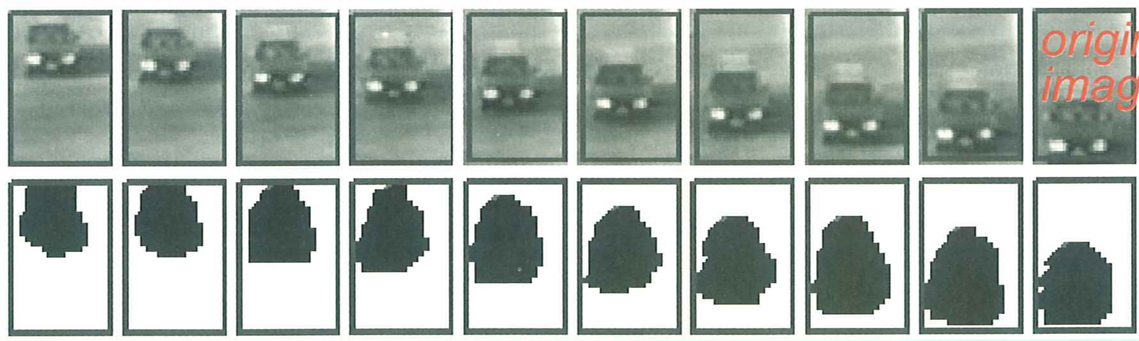
likelihood

$$p(d_i = \vartheta \mid f_i = \vartheta') = \begin{cases} P_r(\mathcal{A}), & \text{if } \vartheta = \vartheta' = \mathcal{A},\ \mathcal{A} \in \{\mathcal{B}, \mathcal{F}\} \\ P_r(\mathcal{F}), & \text{if } \vartheta = \mathcal{B},\ \vartheta' = \mathcal{F} \\ P_r(\mathcal{B}), & \text{if } \vartheta = \mathcal{F},\ \vartheta' = \mathcal{B} \end{cases}$$

Gibbs distribution

$$p(f_{n_t} = \mathcal{A} \mid f_{\mathcal{N}_{n_t}}) = \frac{\exp\{-\alpha f_{n_t} - \beta f_{n_t} \sum_{i'\in\mathcal{N}_{n_t}} f_{i'}\}}{1 + \exp\{-\alpha f_{n_t} - \beta f_{n_t} \sum_{i'\in\mathcal{N}_{n_t}} f_{i'}\}}, \quad \mathcal{A} \in \{\mathcal{F}, \mathcal{B}\}$$

Some Results



original images

segmentation into $\mathcal{B}, \mathcal{S}$ (white) and $\mathcal{F}$ (black)