On-line Simultaneous Learning and Tracking of Visual Feature Graphs

Arnaud Declercq    Justus H. Piater

Institut Montefiore, University of Liège
Belgium

June 2007
A Chicken-and-Egg Problem

We don’t have a model neither a tracking result ➤ we do both simultaneously!
Incremental Learning of a Feature Graph

Frame(t) → Feature Tracking → Feature Positions(t) → Relation Learning → Static Graphical Model(t)

Relation Set(t-1) → Feature Tracking
The Current Intermediate Model may be Inconsistent with Future Observations

If the model is built with only the first frame, all the relations between features will be learned as rigid.
The Current Intermediate Model may be Inconsistent with Future Observations

We propose an uncertain model of relations that contributes stability to tracking without exerting overly strong counterproductive bias.

We propose an uncertain model of relations that contributes stability to tracking without exerting overly strong counterproductive bias.
Outline

The Graphical Model and its Tracking

Incremental Learning of the Uncertain Model

Experiments
Relations are represented with a Parametric Potential Function

- online
  - low computational cost

- only interested in specific relations (rigid, articulated, ...)

- here: rigid relations

\[ \psi_{i,j,t}(x_{i,t}, x_{j,t}) = e^{-\frac{(r_t-\mu_t)^2}{2\sigma_t^2}} \]

\[ r_t = x_{i,t} - x_{j,t} \]

\( \mu \)

real position

\( \sigma \)

observation noise

\( x_{i,t} \)

\( x_{j,t} \)
Feature Graph is tracked using 3 Sources of Information

1) The feature position at time $t$ are predicted given their position at time $t - 1$

here: we use particle filters
Feature Graph is tracked using 3 Sources of Information

2) Those positions are adjusted given their image likelihood at time \( t \)

here: we use template matching or color histograms
Feature Graph is tracked using 3 Sources of Information

3) Positions are also adjusted given the potential functions by propagating information between nodes

here: we use belief propagation
Tracking with Potential Function
Learned from Time $t - 1$

blue : connected features
green : unconnected features
The New Observation may be Inconsistent with Previous Observed Relations

The relation is only learned with observations up to time $t - 1$.

$$\psi_{i,j,t-1}$$
Potential Functions must account for Uncertainty on Temporal Coherence

Observations in a video are spatially correlated over time

\[ \psi_{i,j,t}^- = \psi_{i,j,t}^+ \ast N(0, \sigma_\Delta) \]

\[ \psi_{i,j,t}^- = \psi_{i,j,t}^+ \ast N(0, \sigma_\Delta) \]

\[ \psi_{i,j,t}^- = \psi_{i,j,t}^+ \ast N(0, \sigma_\Delta) \]

We convolve with a distribution \( N(0, \sigma_\Delta) \sim p(r_t | r_{t-1}) \)
If Potential Functions account for Uncertainty on Temporal Coherence
Account for this Uncertainty is not Enough!
Is There Other Sources of Uncertainty?

The relation is not rigid anymore
The learned potential function must be more complex
than a simple Gaussian model

$$\psi_{i,j,t}(x_{i,t}, x_{j,t}) = e^{\frac{- (r_t - \mu_t)^2}{2\sigma_t^2}}$$

The relation is not rigid anymore
- The learned potential function must be more complex than a simple Gaussian model
Outline

The Graphical Model and its Tracking

Incremental Learning of the Uncertain Model

Experiments
Two Other Sources of Uncertainty in the Learned Potential Function

- old rigid relation
- new relation
- non 100% rigid relation
- uncertainty in the parametric model
- uncertainty in the parameters
- reliable

The reliable relations should have more influence than other
Learning the Maximum Likelihood Potential Function

\[ \hat{\mu}_t = \frac{\pi_{t-1} \hat{\mu}_t + w_t r_t}{\pi_{t-1} + w_t}, \]

\[ \hat{\sigma}^2_t = \frac{\pi_{t-1} (\hat{\sigma}^2_{t-1} + (\hat{\mu}_t - \hat{\mu}_{t-1})^2) + w_t (r_t - \hat{\mu}_t)^2}{\pi_{t-1} + w_t} \]

\[ \pi_t = \pi_{t-1} + w_t, \]

- \( w_t = p(z_{i,t} | x_{i,t})p(z_{j,t} | x_{j,t}) \) is the weight of the observed relation at time \( t \)

- \( \pi_t \) is the cumulative weight of previous observations

  ➤ less reliable observations will have less influence on the potential function
1) Uncertainty in the Parameters

\[ P(\tilde{\sigma}_t^2 \leq \sigma_t^2) = \alpha \]

\[ \tilde{\sigma}_t^2 = \frac{\pi_t}{\chi^2_{\pi_t-1}(\alpha)} \hat{\sigma}_t^2 \]

\[ \chi^2_{\pi_t-1}(\alpha) : \text{chi-square distribution at } \alpha \]

\[ \psi_{i,j,t}(x_{i,t}, x_{j,t}) = e^{-\frac{(r_t-\hat{\mu}_t)^2}{2\tilde{\sigma}_t^2}} \]
2) Uncertainty in the Parametric Model

\[ \psi_{i,j,t}(x_{i,t}, x_{j,t}) = \lambda_t e^{-\frac{(r_t - \hat{\mu}_t)^2}{2\tilde{\sigma}_t^2}} + (1 - \lambda_t) \frac{1}{2} \]

\[ \text{informative} \quad \text{uninformative} \]

\[ \lambda_t : \text{probability that the relation corresponds to a gaussian model} \]
We estimate $\lambda_t$ with a Method inspired from the Kolmogorov-Smirnov Test

- distance between ML model and observations:
  
  $$D = \frac{1}{|I|} \int_I |\hat{F}(x) - F_n(x)| \, dx$$

- probability of Gaussianity:
  
  $$\lambda_t = e^{-\frac{D^2}{T_D^2}}$$

$T_D$ : allowed deviation of observed relations from Gaussianity.
The Whole Method for a Time \( t \)

\[
\psi_{i,j,t} = \psi_{i,j,t-1} \otimes \mathcal{N}(0, \sigma_\Delta)
\]

\[
\psi^+_{i,j,t}(x_{i,t}, x_{j,t}) = \lambda_t e^{\frac{-(r_t - \hat{\mu}_t)^2}{2\sigma_t^2}} + (1 - \lambda_t)^{\frac{1}{2}}
\]

Sequential Belief Propagation

Frame(\( t \))

Feature Positions(\( t \))

Relation Set(\( t-1 \))

Graphical Model(\( t \))

Tracking

Learning
Exemple of Simultaneous Learning and Tracking
The relations related to the mouth are learned more slowly than the others due to their lower likelihood in the image.
The variance of the non-rigid relations increase once the head start to move separating the rigid relations from the others.
Evolution of the Gaussian Model Probabilities

The probabilities of the non-rigid relations become insignificant once the head moves but do not increase at the end because they were clearly non-rigid during the major part of the sequence.
Conclusions

Unsupervised online learning of feature graphs.

Robust tracking aided by information extracted from previous frames.

Relations based on a parametric model that incurs only a low computational cost.

Uncertain model that contributes stability to tracking without exerting overly strong, counterproductive bias.
Uncertainty in the Parametric Model

\[ \psi_{i,j,t}(x_{i,t}, x_{j,t}) = \lambda_t e^{-\frac{(r_t - \hat{\mu}_t)^2}{2\tilde{\sigma}^2_t}} + (1 - \lambda_t)^{1/2} \]

\[ \lambda_t : \text{probability that the relation corresponds to a gaussian model} \]
The Whole Method for a Time $t$

\[ \psi_{i,j,t} = \psi_{i,j,t-1} \otimes \mathcal{N}(0, \sigma_{\Delta}) \]

\[ \psi^+_{i,j,t}(x_{i,t}, x_{j,t}) = \lambda_t e^{-\frac{(r_t - \hat{\mu}_t)^2}{2\hat{\sigma}_t^2}} + (1 - \lambda_t)^{\frac{1}{2}} \]

Sequential Belief Propagation

Frame($t$)

Feature Positions($t$)

Relation Set($t-1$)

Graphical Model($t$)

Sequential Belief Propagation Tracking Learning

Tracking Learning