Efficient Object Tracking Using Level-Sets

We have investigated an approach for efficient object tracking using level sets. In this framework, the target object contour is not modeled explicitly, but is only implicitly defined as the zero-level set of a continuous embedding function. Thus, the object contour can be changed by evolving the embedding function. The advantage of this approach is that the discrete segmentation task is now formulated as a continuous optimization problem, which can be solved by gradient descent methods. Level-set tracking proceeds by iterating a segmentation step, where the current level-set contour is evolved to adapt to the changed image content, and a tracking step, where the warp defining the local reference frame is adapted to best fit the next image. Both steps can be implemented very efficiently, since they operate only on a narrow band around the tracked contour. Our approach is based on the formulation by (C. Bibby & I. Reid, ECCV’08), who report a run-time performance of 85Hz on a regular PC. This makes the approach suitable for applications on mobile phones.

Segmentation step (evolving the level-set contour)

\[
\frac{\partial \Phi(x, t)}{\partial t} = \frac{\delta_{	ext{fg}}(\Phi(x, t))}{\Omega_{\text{fg}}(\Phi(x, t))} \frac{1}{\sigma^2} \nabla^2 \Phi - \text{div}(\nabla \Phi) + \lambda \text{div} \left( \frac{\nabla \Phi}{|\nabla \Phi|} \right)
\]

Tracking step (adapting the warp)

\[
\Delta \mathbf{p} = \sum_{i=1}^{N} \frac{1}{2P(x_i, \mathbf{p}, y_i)} \left( P_i - P_i \Omega_{\text{fg}}(\Phi(x_i)) \right) \nabla \Phi(x_i) + \nabla \Omega_{\text{fg}}(\Phi(x_i))
\]

where \(P(x_i, \mathbf{p}, y_i) = \Omega(x_i) \mathbf{p} + (1 - \Omega(x_i)) \mathbf{p}_b\)

Example Application
Face tracking for mobile user interfaces

Extension to Multi-Object Tracking

We have developed a novel approach that integrates level-set trackers into a tracking-by-detection framework for robust street-level multi-person tracking. Our approach is based on the idea to track individual pedestrians by efficient level-set trackers that are automatically initialized by a pedestrian detector. We propose robust methods for performing this initialization and for integrating additional geometric constraints and consistency checks. The tracked person contours in each video frame are then automatically converted to 3D world locations and are transmitted to a high-level tracker, which integrates those measurements into a robust multi-hypothesis tracking framework using physical motion models. This high-level tracker is responsible for initializing new tracks, correcting the low-level tracker’s predictions when drift occurs, and for tracking person identities through occlusions. An interesting property of our proposed approach is that it does not require the computationally expensive object detector to be executed on every video frame. This is especially relevant for deployment on mobile platforms, where real-time performance is crucial and computational resources are limited.

Our system achieves robust performance under very challenging real-world conditions. Its capabilities are important building blocks for applications in mobile robotics and automotive driver assistance systems.

Applications
Tracking through Crowds  Mobile Robotics  Automotive Driver Assistance Systems