Motion planning as model for visual object tracking

L = L₀ − 4πif(z₀) L = L₀ − 2πif(z₀) L = L₀ L = L₀ + 2πif(z₀)

k = −2 k = −1 k = 0 k = 1

∆θ = −3π ∆θ = −1π ∆θ = π ∆θ = 3π

(a) Good plans
zb zg O₁ O₂ γ₁ γ₂ γ₃
(b) Redundant plans
zb zg O₁ O₂ γ₄ γ₅ γ₆ γ₇
(c) Loop plan
zb zg O₁ O₂ γ₈

Example of winding numbers

A path can be written in parametric form: \( x = x(t) \), \( y = y(t) \).

The loop function can be a constant: \( f(z₀) = 1 \).

Then \( \text{L-value} \) can be computed in closed form as

\[ L = \text{Const} + \text{exp} \left[ \int \text{vector}(0) \right] \]

where \( \text{Const} \) is the distance between path \( z₀ \) and tracklet \( F \) and \( S_{\text{Tracklet}}(F, F_i, F_j) \) in the score for picking up the partial occlusions along the gap.

To reduce computation, we prune paths whose costs are higher than the minimal one above a threshold.

Data and Result Comparison

<table>
<thead>
<tr>
<th>Scenario</th>
<th>FPS</th>
<th>Resolution</th>
<th>Data</th>
<th>Result Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>seq #6</td>
<td>1024</td>
<td>700</td>
<td>Object Tracking</td>
<td>Tracking by Planning</td>
</tr>
</tbody>
</table>

Multi-hypothesis Motion Planning for Visual Object Tracking

Halff Gong¹, Jack Sim¹, Maxim Likhachev², Jianbo Shi¹

¹ GRASP Lab, University of Pennsylvania ² Robotics Institute, Carnegie Mellon University

Tracking by Planning

We test our motion model in a batchmode tracking by detection framework.

Tracking a person in the visible state leads to a short trajectory that we call a tracklet.

A conservative threshold is used to terminate the trajectory when the tracking score becomes too low.

After termination, the same person may be picked up again by the detection algorithm, and tracked to produce another tracklet.

After tracklets are obtained, we can link them using both appearance and planning consistency.

Criterias for tracklets linking by planning

- Assume that we have a set of tracklets \( F = \{ F_1, \ldots, F_n \} \).
- Each tracklet is described by 3D point series.
- We then link and extend these tracklets, \( F_i \) to complete trajectories.
- To link tracklets into plausible goal-oriented-obstacle-avoiding walking paths, we design the following criterion for tracking:
  \[ \text{LinkingScore}(F_i, F_j) = S_{\text{Tracklet}}(F_i, F_j, F_k) \]

where \( S_{\text{Tracklet}}(F_i, F_j, F_k) \) is the score for picking up the partial occlusions along the gap.

To reduce computation, we prune paths whose costs are higher than the minimal one above a threshold.

Drawbacks of [Hemantbhattacharya2010]

2. Obstacle marker function must be carefully chosen for numeric stability of \( L \)-values.
3. The representation of state space is an infinite augmented graph.

Tracking with motion planning

- Distance Score
- Partial Occlusion Score

From \( L \)-value to winding numbers

- We propose replacing \( L \)-value with a more informative index, that incorporates the number of loops around obstacles.
- This allows us to screen out any paths with many loops, which are unlikely to be the paths that people actually take.
- The \( L \)-value of a path with respect to a single obstacle \( k = (k_x, k_y) \) is:
  \[ L = \int \text{vector}(0) \]
- \( k \)-values for a single obstacle must be in the discrete set of \( \{ 0, \pm 1, \pm 2 \} \).
- Thus we can use \( k \) (winding number) to distinguish homotopy classes with respect to one obstacle which

Example of winding numbers

- \( k > 0 \) indicates a path to the right of the obstacle that includes \( k \) loops around it.
- \( k < 0 \) indicates a path to the left of the obstacle that includes \( k \) loops around it.
- For a feasible path, the values of \( k \) will likely be 1 or -1, meaning ‘go-right’ or ‘go-left’ around the obstacle.

Homotopy-class planning [Bhattacharya2010]

\( L = L₀ - 4\pi i f(z₀) \)

\( k = (k_x, k_y, k_z) \) with respect to all obstacles as an integer vector (vector of winding numbers, or \( -vectors \))

Theorem Two trajectories \( z₁ \) and \( z₂ \) with \( k \)-vectors \( k₁ \) and \( k₂ \) connecting the same points lie in the same homotopy class if and only if \( k₁ - k₂ \).

Homotopy-class planning [Bhattacharya2010]

- Let \( z \) be a point in the complex plane, \( o \) the start point and \( g \) the goal of an agent (where it is intended to go). A path \( \gamma(t) \) is a complex function of \( at \), with constant end points.
- A complex obstacle marker function is defined as \( f(z) \)
- A path \( \gamma(t) \) is a complex Homotopy function and \( \alpha \) is a point in obstacles.
- Cauchy Integral Theorem Two trajectories \( \gamma₁(t) \) and \( \gamma₂(t) \) connecting the same pair of points \( i \) and \( j \) in the same homotopy class are homotopic if and only if \( f(z) \).

Therefore they use the \( L \)-value, defined as \( f(z) \), \( f(\gamma₁) \), and \( f(\gamma₂) \) to index homotopy classes.

Hemant Bhattacharya 2010

2. Obstacle marker function must be carefully chosen for numeric stability of \( L \)-values.
3. The representation of state space is an infinite augmented graph.

Tracking with multi-hypothesis motion planning

<table>
<thead>
<tr>
<th>k</th>
<th>0</th>
<th>-1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆θ</td>
<td>-3π</td>
<td>-π</td>
<td>π</td>
</tr>
<tr>
<td>L₀</td>
<td>L₀</td>
<td>L₀</td>
<td>L₀</td>
</tr>
</tbody>
</table>

Defining \( \alpha₁ \) be the \( L \)-value associated with the \( i \)-th obstacle, we can homotopy a class with respect to all obstacles as an integer vector (vector of winding numbers, or \( -vectors \))

\[ k = (k_x, k_y, k_z) \]

Vilem Novák, Dejan Milojčić, and Alejandro López

From winding numbers to winding angles

A path \( \gamma(t) \) can be written in parametric form: \( x(t) = x(t) \), \( y(t) = y(t) \).

The loop function can be a constant: \( f(z₀) = 1 \).

Then \( \text{L-value} \) can be computed in closed form as

\[ L = \text{Const} + \text{exp} \left[ \int \text{vector}(0) \right] \]

where \( \text{Const} \) is the distance between path \( z₀ \) and tracklet \( F \) and \( S_{\text{Tracklet}}(F, F_i, F_j) \) is the score for picking up the partial occlusions along the gap.

To reduce computation, we prune paths whose costs are higher than the minimal one above a threshold.

<table>
<thead>
<tr>
<th>Plan</th>
<th>seq #6</th>
<th>seq #4</th>
<th>Plan</th>
<th>seq #6</th>
<th>seq #4</th>
</tr>
</thead>
<tbody>
<tr>
<td>LINEAR</td>
<td>464</td>
<td>512</td>
<td>LINEAR</td>
<td>464</td>
<td>512</td>
</tr>
<tr>
<td>tracking</td>
<td>94</td>
<td>125</td>
<td>tracking</td>
<td>94</td>
<td>125</td>
</tr>
</tbody>
</table>

Data and Result Comparison

Image datasets and bounding boxes over time. Each panel shows the bounding box of a pedestrian in two parts. The top parts show the image patches of ground truth (3rd row), PLAN results (2nd row) and LINEAR results (1st row). The number on each box is the frame number. They are intended on left or right for better visual effects. The bottom parts show side-views frames superimposed with bounding boxes. The magnup bounding boxes are current objects of interests. Yellow bounding boxes are other objects. The bold green lines are the planned routes that the objects follow. The thinner green lines are other planned paths (after passing) that are not followed by the people.

Current objects of interests. Yellow bounding boxes are other objects. The bold green lines are the planned routes that the objects follow. The thinner green lines are other planned paths (after passing) that are not followed by the people.