Target Tracking in FLIR Images Using Mean Shift

Alper YILMAZ
Computer Vision Lab.
University of Central Florida

Characteristics of Targets

- Low contrast with the background.
- Similar gray level distribution to the overall frame gray level distribution.
- Slightly brighter than the background (not always).
- Targets are most of the time 5 to 10 pixels.
- Fast global and ego motion.
- No specific shape information
Mean Shift Vector

Given:
Data points and approximate location of the mean of this data.

Task:
Estimate the exact location of the mean of the data by determining the shift vector from initial mean.

Mean Shift Vector Example

Mean shift vector always points towards the direction of the maximum increase in the density.
Modified Mean Shift

\[ M_h(y_0) = \frac{\sum_{i=1}^{n_x} w_i(y_0)x_i}{\sum_{i=1}^{n_x} w_i(y_0)} - y_0 \]

- \( n_x \): number of points in the kernel
- \( y_0 \): initial mean location
- \( x_i \): data points
- \( h \): kernel radius

Weights are determined using kernels (masks):
Uniform, Gaussian or Epanechnikov

Properties of Mean Shift

- Mean shift vector has the direction of the gradient of the density estimate.
- It is computed iteratively for obtaining the maximum density in the local neighborhood.
Outline

1. Introduction
   1. Kernel density estimate
   2. Possible kernels
   3. Epanechnikov profile
   4. Estimate of density gradient
   5. Mean shift & Epanechnikov Kernel

2. Feature space
   1. Target gray level distribution
   2. Distribution and tracking
      1. Similarity of target & candidate distributions
      2. Distance minimization
      3. Bhattacharya maximization using mean shift
      4. Algorithm
   3. Target standard deviation
   4. Target localization using 2 features
   5. FLIR results
   6. Future work

Kernel Density Estimate

\[
\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K \left( \frac{x - x_i}{h} \right)
\]

- \(n\): number of points in the kernel
- \(h\): window radius
- \(x\): mean vector
- \(d\): number of dimensions
- \(K\): Kernel density function
Possible Kernels

- Uniform kernel

- Epanechnikov kernel (convex, monotonic decreasing)

\[ K_E(x) = \begin{cases} \frac{1}{2} c_d^{-1} (d + 2)(1 - \|x\|^2) & \text{if } \|x\| < 1 \\ 0 & \text{otherwise} \end{cases} \]

\( c_d \): volume of unit d-dim sphere
\( d \): number of dimensions

- Normal kernel (convex, monotonic decreasing)

\[ K_N = (2\pi)^{-d/2} e^{-\|x\|^2/2} \]

\( d \): number of dimensions

Kernel & Profile

- **Kernel function**: defined in terms of vector

- **Profile function**: defined in terms of variable

\[ k_E(\|x\|^2) = K_E(x) \]
Epanechnikov Profile (2D)

Epanechnikov profile yields minimum mean integrated square error

\[ k_E(x) = \frac{1}{2} c_d^{-1} (d + 2)(1 - x) \]

\[ k_{E,d=2}(x) = \frac{2}{\pi} (1 - x) \]

\[ \frac{\partial k_{E,d=2}}{\partial x} = -\frac{2}{\pi} \]

Estimate of Density Gradient

density estimate:

\[ \hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \]

gradient of density estimate:

\[ \nabla \hat{f}(x) \equiv \nabla \hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} \nabla K\left(\frac{x - x_i}{h}\right) \]
Mean Shift Vector in Terms of Epanechnikov Kernel

\[ \hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} \nabla K(\frac{x-x_i}{h}) \]

Using Epanechnikov kernel: \( K_E(x) = \frac{1}{2} c_d^{-1} (d+2)(1-\|x\|^2) \)

\[ \hat{f}(x) = \frac{d+2}{nh^{d+2}c_d} n \left( \frac{1}{n} \sum_{x_i \in \mathcal{S}_x} [x_i - x] \right) = \frac{d+2}{h^{d+2}c_d} M_b(x) \]

\( n \) : number of points in unit d-dimensional sphere

Target Model for Tracking

- Features used for tracking include:
  - Gray level
  - Standard deviation

- Feature probability distribution are calculated by using weighted histograms.

- The weights are derived from Epanechnikov profile.
Target Model for Tracking

\[ p(u|K) = C \sum_{x_i \in S} k(\|x_i\|^2) \delta[S(x_i) - u] \]

\( x_1, x_2, x_3, x_4 \) has the same feature, such as gray level.

Target Gray Level Feature

image histogram

target 1 distribution

target 2 distribution

non target distribution
Similarity of Target and Candidate Distributions

Target : \( q_u \).
Candidate : \( \hat{p}_u \).

\[
d(y) = \sqrt{1 - \rho(y)}
\]

\[
\rho(y) = \rho[\hat{p}(y), q] = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)q_u}
\]

\( \rho(y) \) : Bhattacharya coefficient.

Distance Minimization

Minimizing the distance corresponds to maximizing Bhattacharya coefficient.

\[
\rho[\hat{p}(y), q] = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)q_u}
\]

Taylor expansion around \( \hat{p}(y_0) \)

\[
\rho[\hat{p}(y), q] \approx \rho[\hat{p}(y_0), q] + \frac{1}{2} \sum_{i=1}^{m} \hat{p}_i(y) \sqrt{\frac{q_u}{\hat{p}_u(y_0)}}
\]

Maximizing Bhattacharya coefficient can be obtained by maximizing the blue term.
Likelihood Maximization

\[ \rho[\hat{p}(y), q] \equiv \rho[\hat{p}(y_0), q] + \frac{1}{2} \sum_{i=1}^{m} \hat{p}_u(y) \sqrt{\frac{q_u}{\hat{p}_u(y_0)}} \]

\[ \frac{C_h}{2} \sum_{i=1}^{m} \left[ \sum_{u=1}^{m} \delta[S(x_i) - u] \sqrt{\frac{q_u}{\hat{p}_u(y_0)}} \right] k\left(\frac{|y - x_i|}{h}\right) \]

h : radius of sphere
C_h : normalization constant
S(x_i) : gray level at x
y : kernel center
m : number of bins

Likelihood maximization depends on maximizing \( w_x \).

Likelihood Maximization Using Mean Shift Vector

\[ M_h(y_0) = \left[ \frac{1}{n_x} \sum_{i=1}^{n_x} (x_i - y_0) \right] \]

\[ y_0 \Rightarrow M_h(y_0) = \frac{\sum_{i=1}^{n_x} x_i}{n_x} = \frac{\sum_{i=1}^{n_x} x_i}{\sum_{i=1}^{n_x} 1} \]

new mean location

y : Spatial coordinates of target
y_0 = [0,0]^T.
Likelihood Maximization Using Mean Shift Vector

Maximization of the likelihood of target and candidate depends on the weights:

\[ w_i(y_o) = \sum_{u=1}^{w_i} \delta[S(x_i) - u] \sqrt{\frac{q_u}{\tilde{P}_u(y_o)}} \quad \text{where} \quad 0 \leq w_i \leq 1 \]

Since \( \sum_{i=1}^{w_i} w_i(y_o) \) is strictly positive, mean shift vector can be written as

\[ M_h(y_0) = \frac{\sum_{i=1}^{w_i} w_i(y_0) x_i - y_0}{\sum_{i=1}^{w_i} w_i(y_0)} \]

Thus, new target center is

\[ \hat{y} = y_0 + M_h(y_0) \]

Algorithm

- Calculate \((q_G)\)
- Initialize estimated center \((y_1 = y_0)\)
- Calculate \((p_G)\)
- Calculate \((w_G)\)
- Estimate new target center \((y_1)\)
- \(d < \epsilon\) ?
  - Update target center \((y_1 = y_1)\)
  - Repeat until end of the sequence
Target Gray Level Distribution (1)

- Target (frame 0)
- Frame 220
- Frame 246
- Frame 288

Target Gray Level Distribution (2)

- Target (frame 173)
- Frame 221
- Frame 382
- Frame 695
Target Std. Deviation Feature

\[ \sigma^2 = \frac{1}{8} \sum_{i=1}^{8} (x_i - x)^2 \]

Modify \( x \) by \( \sqrt{\sigma^2} \)

Gray Level (Reminder)

image histogram

target 1 distribution

target 2 distribution

non target distribution
Target Std. Deviation Feature

Combining Distributions of Gray Level and Std. Deviation

$q_G$: target gray level
$q_S$: target standard deviation
$p_G$: candidate gray level
$p_S$: candidate standard deviation

$$w_G(y_o) = \sum_{u=1}^{n} \delta(S(x_i) - u) \frac{q_G}{\hat{p}_{G_i}(y_o)}$$

$$w_S(y_o) = \sum_{u=1}^{n} \delta(S(x_i) - u) \frac{q_S}{\hat{p}_{S_i}(y_o)}$$
Combining Distributions of Gray Level and Std. Deviation

New target center is determined using

\[ \hat{y}_{k+1} = y_k + \frac{\sum_{i=1}^{n_t} w_{S_i}(y_k) w_{G_i}(y_k) x_i}{\sum_{i=1}^{n_t} w_{S_i}(y_k) w_{G_i}(y_k)} \]

Iteratively calculate the new target center until the distance is minimized.

\[ d(y) = \sqrt{1 - \rho(y)} \]

where

\[ \rho(y) = \tau \sqrt{p_S(y)^T q_S} + (1 - \tau) \sqrt{p_G(y)^T q_G} \]

Gray Level Distribution at Each Iteration

Sequence rng14_15, frames 87 and 88

![Graph showing probability distribution at each iteration](UCF Computer Vision Lab. 30)
Distance Between Consecutive Frames

For Different Sequences

Experiments

Features used:
- Intensity
- Standard deviation
- Gradient magnitude
- Intensity & standard deviation
- Intensity & gradient magnitude

Mutual probabilities are combined using
- Geometric mean
- Weighting (described in slide 20)

We filtered the frames using 2D-Regularization filter.
Test Set

Our test set is composed of 21 FLIR sequences:

rng14_15, rng15_20, rng15_NS, rng16_04,
rng16_07, rng16_08, rng16_18, rng17_01,
rng17_02, rng17_20, rng18_03, rng18_05,
rng18_07, rng18_12, rng18_13, rng18_16,
rng18_18, rng19_01, rng19_06, rng19_07,
rng19_11

We manually initialize one target in the first frame
and track the target in the sequence.

We have visually confirmed the results
(not from the ground truth)

Comparison of Results

<table>
<thead>
<tr>
<th>Variance</th>
<th>Color</th>
<th>Variance and Color together</th>
<th>Geom. mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Variance</td>
<td>Color</td>
</tr>
<tr>
<td>14_15</td>
<td>Fine</td>
<td>Fine</td>
<td>Fine</td>
</tr>
<tr>
<td>15_20</td>
<td>Fine</td>
<td>Bad</td>
<td>Fine</td>
</tr>
<tr>
<td>15_NS</td>
<td>Fine</td>
<td>Bad</td>
<td>Ok</td>
</tr>
<tr>
<td>16_04</td>
<td>Bad</td>
<td>Fine</td>
<td>Ok</td>
</tr>
<tr>
<td>16_07</td>
<td>Bad</td>
<td>Bad</td>
<td>Fine</td>
</tr>
<tr>
<td>16_08</td>
<td>Ok</td>
<td>Fine</td>
<td>Fine</td>
</tr>
<tr>
<td>16_18</td>
<td>Fine</td>
<td>Fine</td>
<td>Fine</td>
</tr>
<tr>
<td>17_01</td>
<td>Fine</td>
<td>Fine</td>
<td>Fine</td>
</tr>
<tr>
<td>17_20</td>
<td>Bad</td>
<td>Bad</td>
<td>Fine</td>
</tr>
<tr>
<td>18_03</td>
<td>Ok</td>
<td>Bad</td>
<td>Ok</td>
</tr>
<tr>
<td>18_05</td>
<td>Ok</td>
<td>Bad</td>
<td>Fine</td>
</tr>
<tr>
<td>18_07</td>
<td>Fine</td>
<td>Ok-fine</td>
<td>Fine</td>
</tr>
<tr>
<td>18-13</td>
<td>Fine</td>
<td>Bad-ok</td>
<td>Fine</td>
</tr>
<tr>
<td>18-18</td>
<td>Fine</td>
<td>Fine</td>
<td>Fine</td>
</tr>
<tr>
<td>19-01</td>
<td>Fine</td>
<td>Fine</td>
<td>Fine</td>
</tr>
<tr>
<td>19-06</td>
<td>Fine</td>
<td>Fine</td>
<td>Fine</td>
</tr>
<tr>
<td>19-07</td>
<td>Fine</td>
<td>Fine</td>
<td>Fine</td>
</tr>
<tr>
<td>19-11</td>
<td>Fine</td>
<td>Fine</td>
<td>Fine</td>
</tr>
</tbody>
</table>
### Comparison of Results

Using "Geometric Mean" and
- Regularization pre-filtering
- Gaussian pre-filtering
- Median filtering
- Without filtering

|                  | 14_15 | 15_20 | 15_NS | 16_04 | 16_07 | 16_08 | 17_01 | 17_02 | 18_01 | 18_03 | 18_05 | 18_07 | 18_08 | 18_10 | 18_11 | 18_12 | 18_13 | 18_14 | 18_15 | 18_16 | 18_17 |
|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|

### FLIR Tracking Results

- 14_15
- 15_20
- 15_NS
- 17_01
- 17_02
- 17_03
- 18_05
- 19_07
- 19_11
Future Work

- Resolve the drawback of the system for large global motion
- Enhance the sequence to have more distinctive features of the target
- Obtain target model using ellipsoidal region instead of circular region.
- Update initial model periodically.