Affine Object Tracking with Kernel-based Spatial-Color Representations

Haihong Zhang, Weimin Huang, Zhiyong Huang* and Liyuan Li

Visual Understanding Lab
Institute for Infocomm Research

*School of Computing
National University of Singapore
Motivation

Simple View based Model to Precise & Robust Representation

• Discerning difference between similar objects
• Robust against noise in images

From Representation to Robust Pose Estimation
Objective

*Simple* View based Model and tracking algorithm for object tracking under affine transform

A generic and robust tracking approach

A robust, iterative optimization procedure for determining the transformation state of target objects *represented by* the kernel-based spatial-color model
The Problem: From Searching to Tracking

• The key: A good similarity surface that leads to the true T

• Difficult Situations we would like to study
  – The target image embedded in background
  – Included background pixels will alter similarity surface S(T): how?
  – Out study. Our findings:
    • Not much with respect to translation;
    • A bit with respect to rotation and shearing;
    • Greatly with respect to scaling;
Representation model (Elgammal, Duraiswami & Davis et al 2003)

Kernel-based representation model

$$\Omega = \{x_i, u_i\}$$

$$R(\Omega) = p(x, u | \Omega) = \frac{1}{N} \sum_{i=1}^{N} k_x(x - x_i)k_u(u - u_i)$$

Kernel density estimate – representation in functional form.
Characterizing spatial-spectral correlation

The idea: Use the functional $R(\Omega)$ to represent the image region $\Omega$.

Advantage

Accurate characterization of appearance features.

Challenge

Sensitive to image transformations.
Adaptation to Transformations

- **Affine Transformations**: Image variations of (near-)planar surfaces due to movements in 3D space

\[
\begin{align*}
X_i(T) &= X_0 + X_t \\
\theta &= \theta_0 + \theta_t \\
\alpha &= \alpha_x \wedge \alpha_y
\end{align*}
\]

- **Translation**
- **Rotation**
- **(non-uniform) Scaling**
- **Shearing**

**Moving pixels**
\[
x_T = T(x) = Mx + x_t
\]
- Translation
- Rotation
- (non-uniform) Scaling
- Shearing

**Consistent colors**
\[
\{x_i, u_i\} \rightarrow \{T(x_i), u_i\}
\]
Kernel based affine matching

Searching Best $T$ on Similarity Surface

- **Our Mathematical Solution**

\[
\text{arg max } S(T; \Omega_1, \Omega_2) \quad \frac{\partial S(T; \Omega_1, \Omega_2)}{\partial T} = 0
\]

- Computing translation

\[
\nabla_{x_t} S = 0 \quad \Rightarrow \quad x_{t}^{(m+1)} = f_x(x_{t}^{(m)})
\]

- Computing rotation angle

\[
\nabla_{\theta} S = 0 \quad \Rightarrow \quad \theta^{(m+1)} = f_{\theta}(\theta^{(m)})
\]

- Computing shearing factor

\[
\nabla_{s} S = 0 \quad \Rightarrow \quad s^{(m+1)} = f_{s}(s^{(m)})
\]

- Computing scaling factors

\[
\nabla_{x_t} S = 0 \quad \Rightarrow \quad \alpha^{(m+1)} = f_{a}(\alpha^{(m)})
\]

**Efficient iterative procedures to seeking the best $T$ in the local**
The Advantage of Using Physical Transformation Models

- The transformation matrix $M$: physical model or plain mathematical model?
- What the similarity surface looks like
- A physical navigation searching scheme

(a) The original object
(b) Transformation with cluttering background

The Profiles of the surface

Complex ridges on the surface

$\rho_{11} = 1.21$
$\rho_1 = 0.44$
$\rho_2 = 0.32$
$\rho_2 = 1.15$
Simulations

- Artificial objects under noise corruption
  ![Images of artificial objects under varying noise levels](image1.png)
- Tracking objects under only translation
  ![Graph showing MSE (x̃t) vs. Noise level (σ)](image2.png)

Simulations

- Artificial objects under noise corruption
  ![Images of artificial objects under varying noise levels](image1.png)
- Tracking objects under only translation
  ![Graph showing MSE (x̃t) vs. Noise level (σ)](image2.png)
Simulations

- Tracking objects under affine transformation

- Robust in tracking rotating and shearing objects
- Less robust in tracking scaling objects
Real World Tasks

- Faces and Hands
- Tanks and Vehicles
One special object tracking demo

Spatial color tracking

Mean shift tracking

Thanks