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

NUS National University of Singapore

A*STAR I²R

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

NUS National University of Singapore

A*STAR I²R

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;

NUS National University of Singapore

A*STAR I²R

Representation model (Elgammal, Duraiswami & Davis et al 2003)

Kernel-based representation model

$$\Omega = \{\mathbf{x}_i, \mathbf{u}_i\}$$

$$R(\Omega) = p(\mathbf{x}, \mathbf{u} \mid \Omega) = \frac{1}{N} \sum_{i=1}^{N} k_x(\mathbf{x} - \mathbf{x}_i) k_u(\mathrm{u} - \mathrm{u}_i)$$

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

The idea: Use the functional R(Ω) to represent the image region Ω .

Advantage

Accurate characterization of appearance features.

Challenge

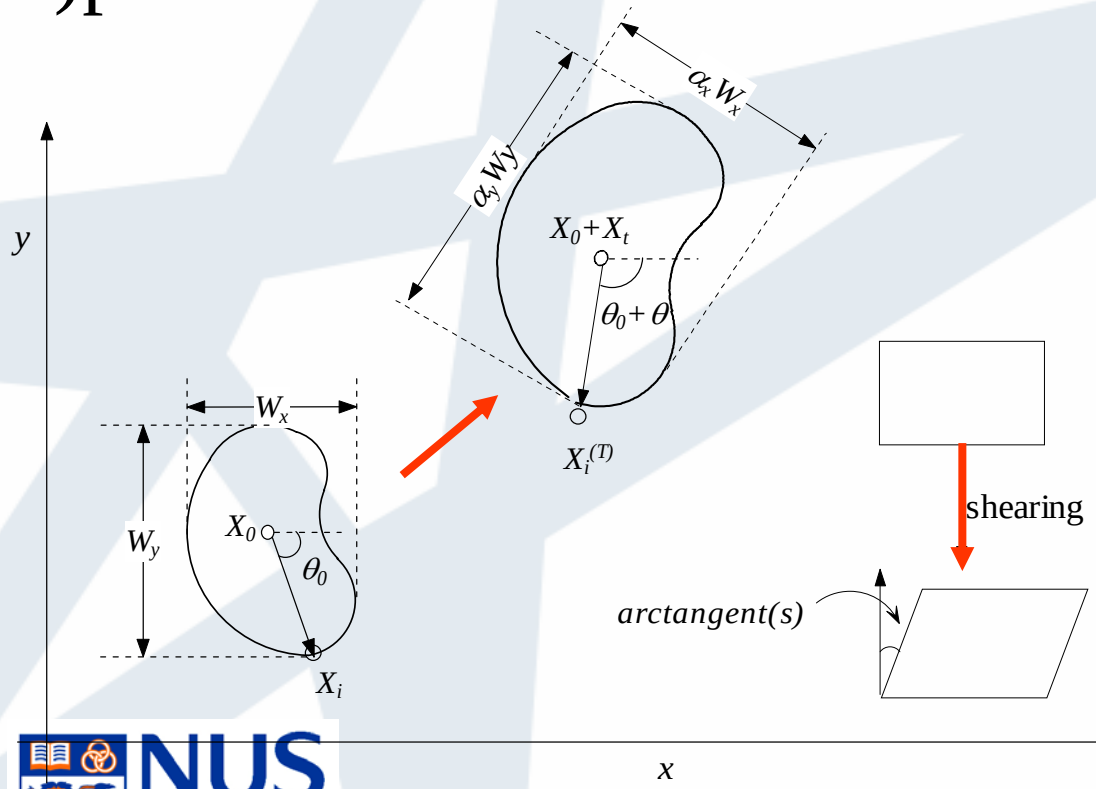
Sensitive to image transformations.

NUS National University of Singapore

A*STAR I²R

Adaptation to Transformations

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



Moving pixels

$$\mathbf{x}_T = T(\mathbf{x}) = \mathbf{M}\mathbf{x} + \mathbf{x}_t$$

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

Consistent colors

$$\{\mathbf{x}_i, \mathbf{u}_i\} \rightarrow \{T(\mathbf{x}_i), \mathbf{u}_i\}$$

NUS National University of Singapore

A*STAR I²R

Kernel based affine matching

Searching Best T on Similarity Surface

- Our Mathematical Solution

$$\arg\max_{T} S(T;\Omega_1,\Omega_2) \qquad \frac{\partial S(T;\Omega_1,\Omega_2)}{\partial T} = 0$$

  - Computing translation

$$\nabla_{\mathbf{x}_t} S = 0 \longrightarrow \mathbf{x}_t^{(m+1)} = f_x(\mathbf{x}_t^{(m)})$$

  - Computing rotation angle

$$\nabla_{\theta} S = 0 \longrightarrow \theta^{(m+1)} = f_\theta(\theta^{(m)})$$

  - Computing shearing factor

$$\nabla_{s} S = 0 \longrightarrow s^{(m+1)} = f_s(s^{(m)})$$

  - Computing scaling factors

$$\nabla_{\mathbf{x}_t} S = 0 \longrightarrow \boldsymbol{\alpha}^{(m+1)} = f_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}^{(m)})$$

Efficient iterative procedures to seeking the best T in the local

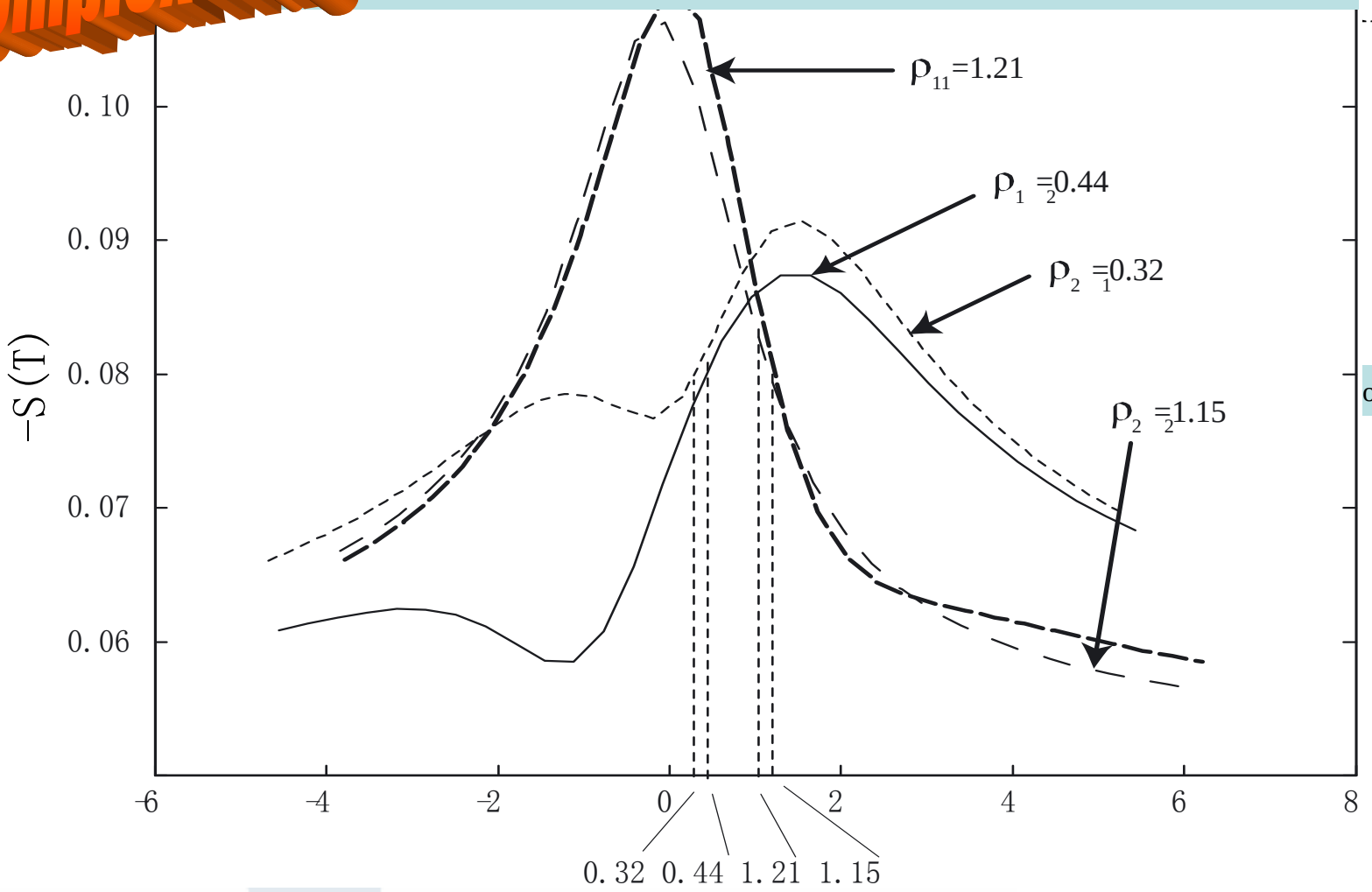
NUS National University of Singapore

A*STAR I²R

The Advantage of Using Physical Transformation

Complex ridges on the surface

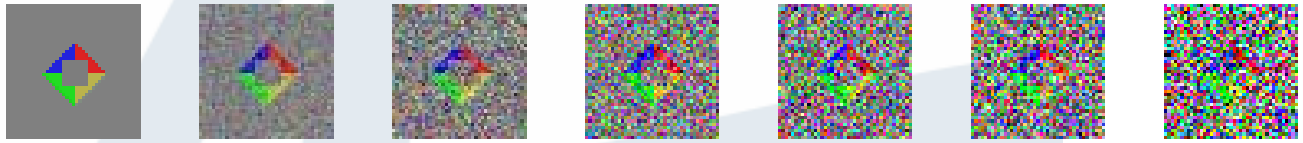
- Complex ridges of the surface
- n
- V
- A



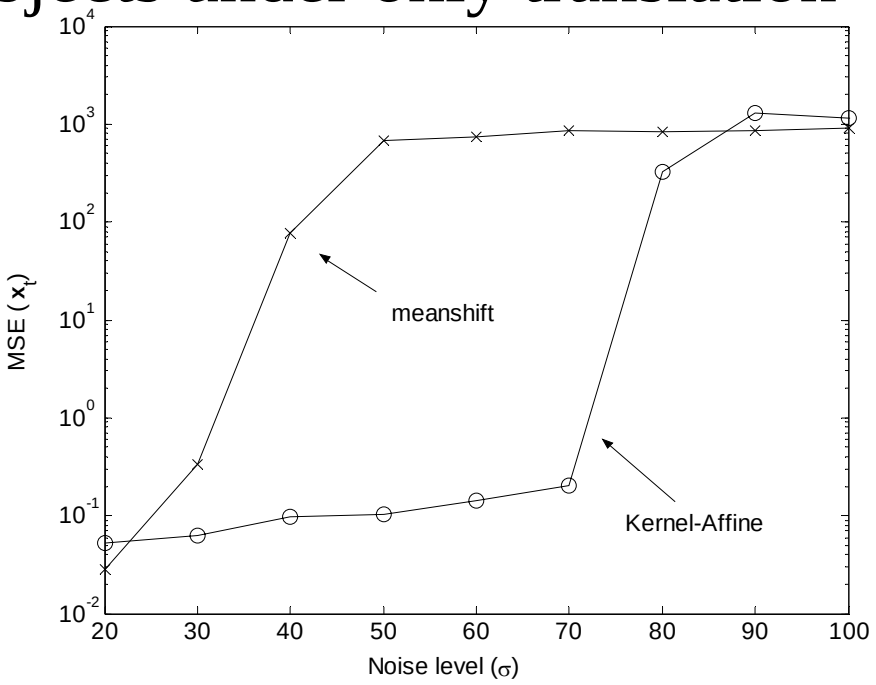
on

Simulations

- Artificial objects under noise corruption



- Tracking objects under only translation

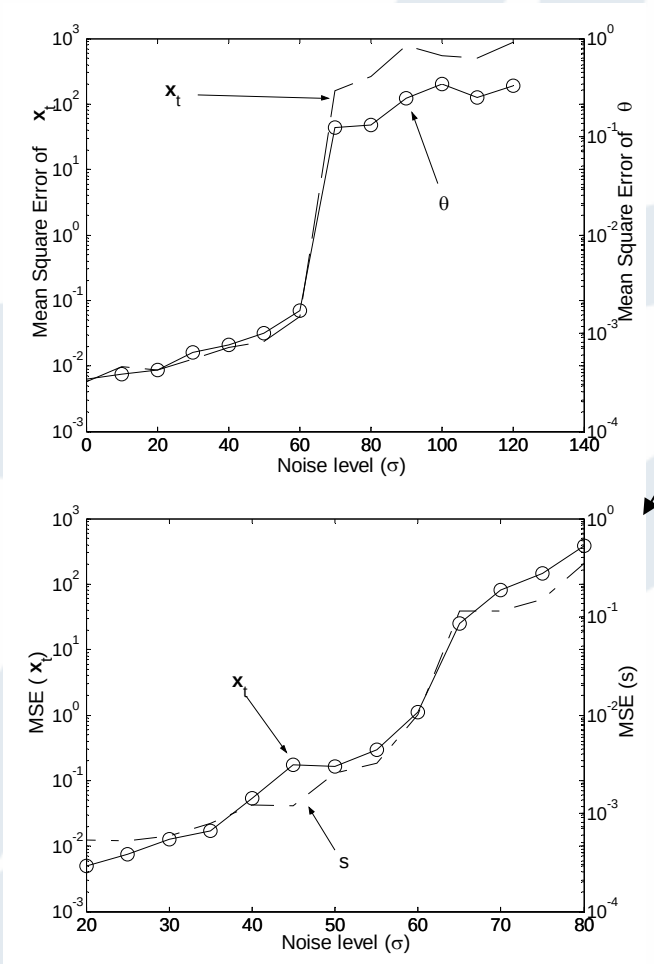


NUS National University of Singapore

A*STAR I²R

Simulations

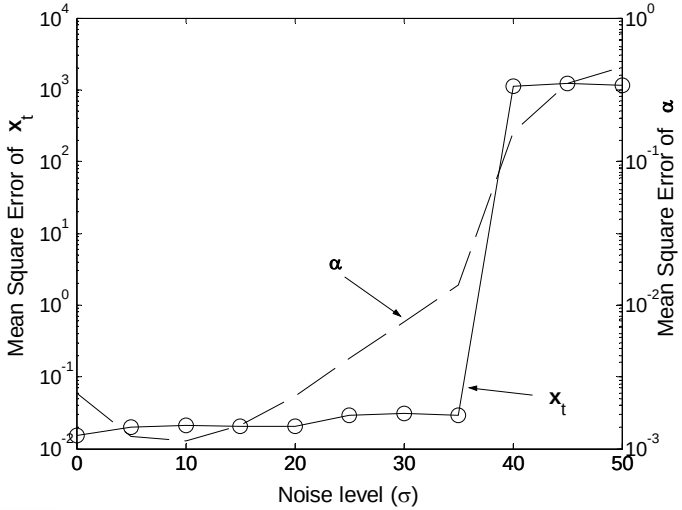
- Tracking objects under affine transformation



Rotation-Trans.

Shearing-Trans.

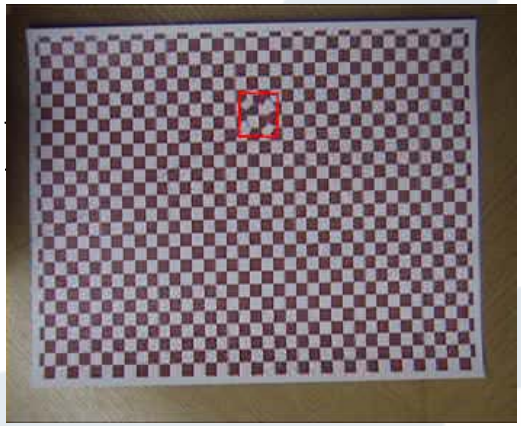
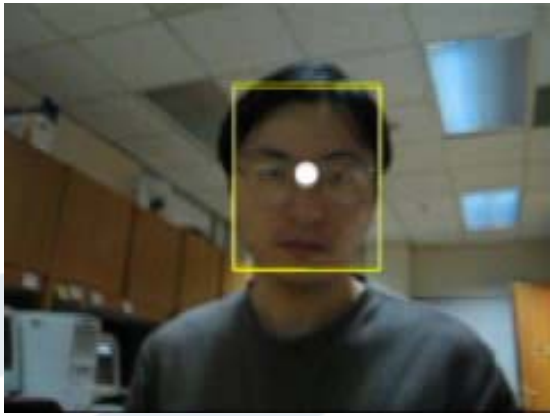
Scaling-Trans.



- Robust in tracking rotating and shearing objects
- Less robust in tracking scaling objects

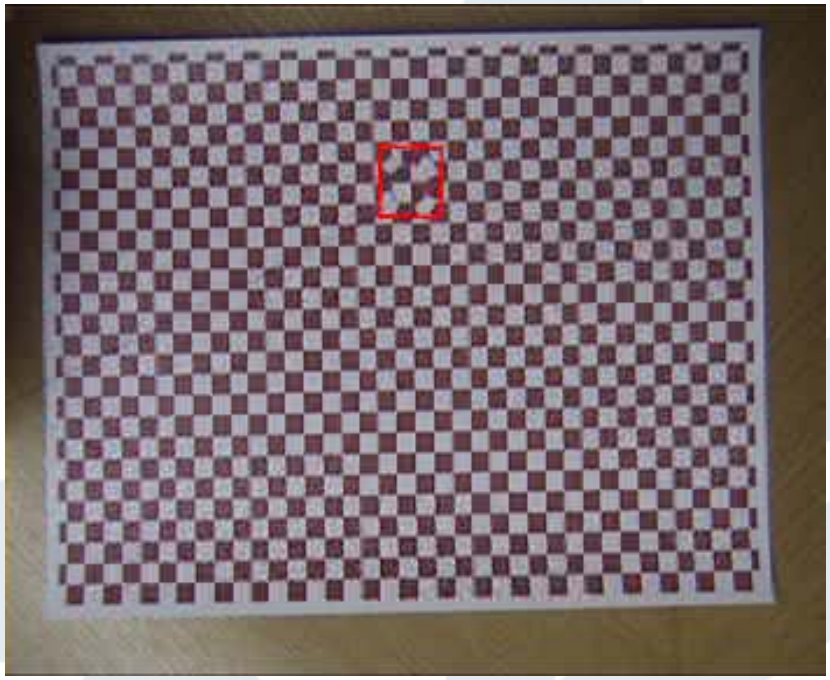
A*STAR I²R

Real World Tasks

- 



- Tanks and Vehicles

One special object tracking demo



Spatial color tracking

Mean shift tracking

NUS National University of Singapore

Thanks

A*STAR I²R