Modular Decomposition and Analysis of Registration based Trackers

Abhineet Singh, Ankush Roy, Xi Zhang
and Martin Jagersand

Reading

Abhineet Singh and Martin Jagersand, "**Modular Tracking Framework: A Fast Library for High Precision Tracking**", IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), September 2017 [pdf] [video] [code]

MTF is available at: http://webdocs.cs.ualberta.ca/~vis/mtf/ along with all datasets and papers

UNIVERSITY OF ALBERTA

Registration based Tracking

- Find the optimal warp or geometric transformation that registers each image in a sequence with the template

$$\boxed{\mathbf{p_t} = \underset{\mathbf{p}}{\text{argmax}}\ f(\mathbf{I_0}(\mathbf{x}), \mathbf{I_t}(\mathbf{w}(\mathbf{x}, \mathbf{p})))}$$

$$\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N], \mathbf{x}_k = [x_k, y_k]^T \in \mathbb{R}^2$$

$$\mathbf{I}(\mathbf{x}) = [I(x_1, y_1), I(x_2, y_2), ..., I(x_N, y_N)]^T \in \mathbb{R}^N, I(x, y) : \mathbb{R}^2 \mapsto \mathbb{R}$$

$$\mathbf{p} = [p_1, p_2, ..., p_S], S : \text{DOF of image motion}$$

$$\mathbf{w} : \mathbb{R}^2 \times \mathbb{R}^S \mapsto \mathbb{R}^2$$

$$f : \mathbb{R}^N \times \mathbb{R}^N \mapsto \mathbb{R}$$

Tracking Lucas-Kanade algorithm

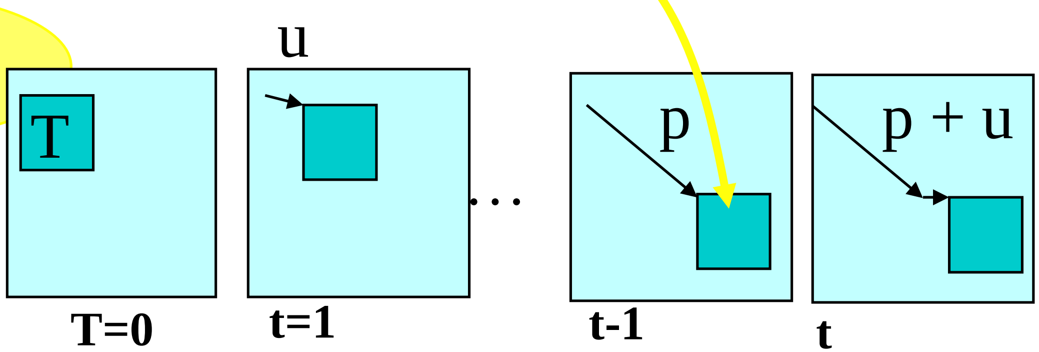
- Create tracking loop, iterate for each new image

Init p=0, Template T

For t = 1…

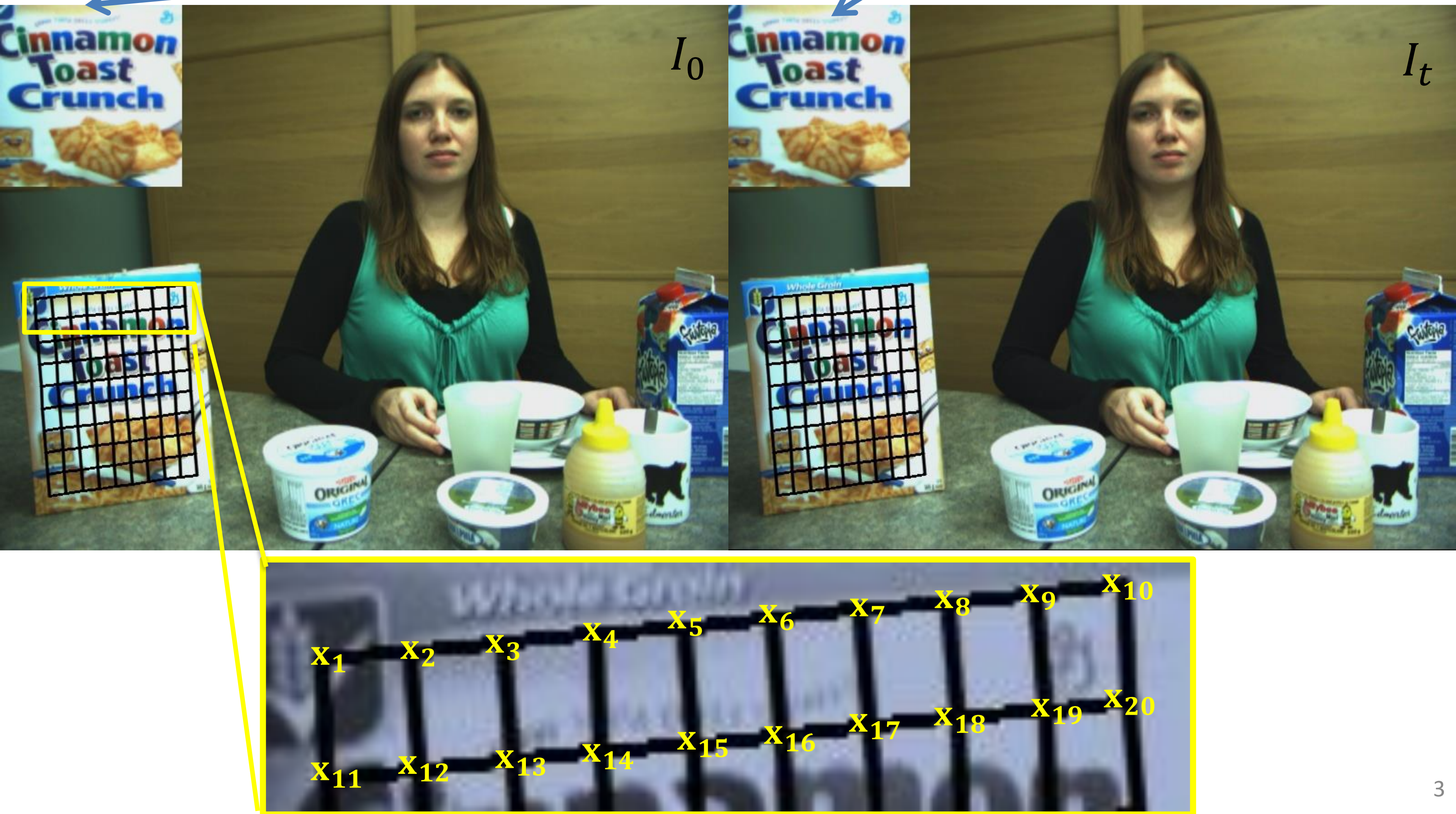
1. Receive I(t+1)
2. Compute dIm = I(t+1, x+p) – T
3. Solve -Im_t = M u
   - Use u = M\Im_t
4. Update p = p + u

Template sourced from pixel window shifted by the state vector p



Registration based Tracking

$$\mathbf{p_t} = \underset{\mathbf{p}}{\operatorname{argmax}}\ f(\mathbf{I_0}(\mathbf{x}), \mathbf{I_t}(\mathbf{w}(\mathbf{x}, \mathbf{p})))$$



Motivation

- Learning/detection based trackers are not suitable for tasks requiring **fast** and **high precision** tracking
  - Visual Servoing
  - Virtual reality
  - SLAM

MTF Usage Example – Multi Target Tracking

UAV Trajectory Estimation

Online Image Mosaicing

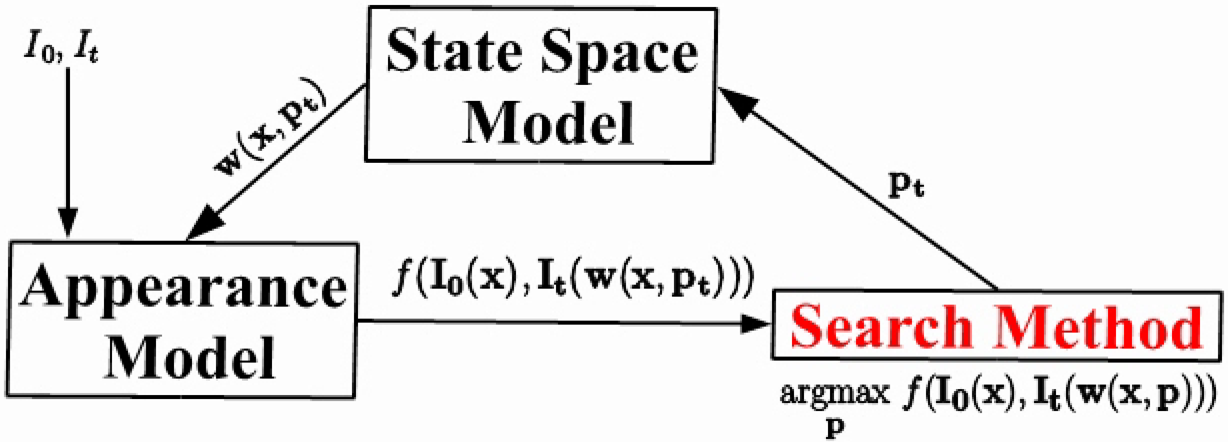
Motivation

- Progress in registration based tracking has become **fragmented** since Lucas Kanade[Lucas81]
  - myriad of contributions that are not well connected
- An intuitive way exists to **relate** these by decomposing the tracking task into three modules
  - most contributions are confined to only one or two of these modules
- Modular Tracking Framework (**MTF**)[Singh16] to easily plug in new methods

B. Lucas, T. Kanade, "An iterative image registration technique with an application to stereo vision", 1981
A. Singh, M. Jagersand, "*Modular Tracking Framework: A Unified Approach to Registration based Tracking*", 2016, available at: http://webdocs.cs.ualberta.ca/~vis/mtf/

Modular Decomposition

$$\mathbf{p_t} = \underset{\mathbf{p}}{\operatorname{argmax}}\; f(\mathbf{I_0}(\mathbf{x}), \mathbf{I_t}(\mathbf{w}(\mathbf{x}, \mathbf{p})))$$



- Appearance Model (**AM**)
  - Measures the **similarity** between a warped patch and the template
- State Space Model (**SSM**)
  - Defines the possible ways to **warp** the object patch
- Search Method (**SM**)
  - **Finds** the warp that maximizes the similarity measure

State Space Model

$$\mathbf{p_t} = \underset{\mathbf{p}}{\operatorname{argmax}}\ f(\mathbf{I_0}(\mathbf{x}), \mathbf{I_t}(\mathbf{w}(\mathbf{x}, \mathbf{p})))$$

- A warping function or geometric transformation that represents the set of allowable image motions of the object
  - embodies any constraints placed on the warp parameter space
    - search efficiency
    - alignment precision
  - includes
    - degrees of freedom (DOF) of allowed motion
    - actual parameterization of the warping function

Registration: from trans $\mathbf{u}$ to warp $\mathbf{w(x,p)}$

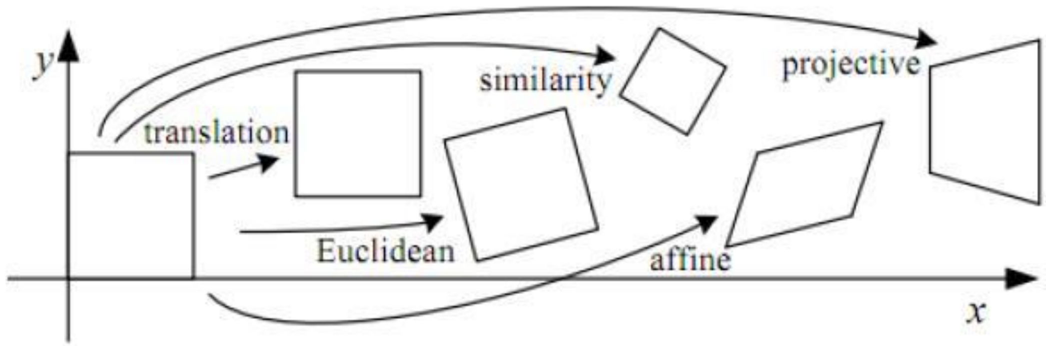
Find parameters of a warping function such that:

$$\mathcal{I}\left(\mathbf{w}\left(\bar{\mathbf{x}};\mathbf{p}_i\right)\right) = \mathcal{I}_T\left(\mathbf{p}_i\right)$$

for all template points $\mathbf{p}_i$

$\mathbf{w}(\bar{\mathbf{x}};\mathbf{p}_i)$

State Space Model – Examples



Translation : $S = 2$

– $\mathbf{w}(\mathbf{x}_k, \mathbf{p}) = \begin{bmatrix} x_k + p_1 \\ y_k + p_2 \end{bmatrix}$

Isometry/Euclidean : $S = 3$

– $\mathbf{w}(\mathbf{x}_k, \mathbf{p}) =$
$\begin{bmatrix} x_k \cos p_1 - y_k \sin p_1 + p_2 \\ x_k \sin p_1 + y_k \cos p_1 + p_3 \end{bmatrix}$

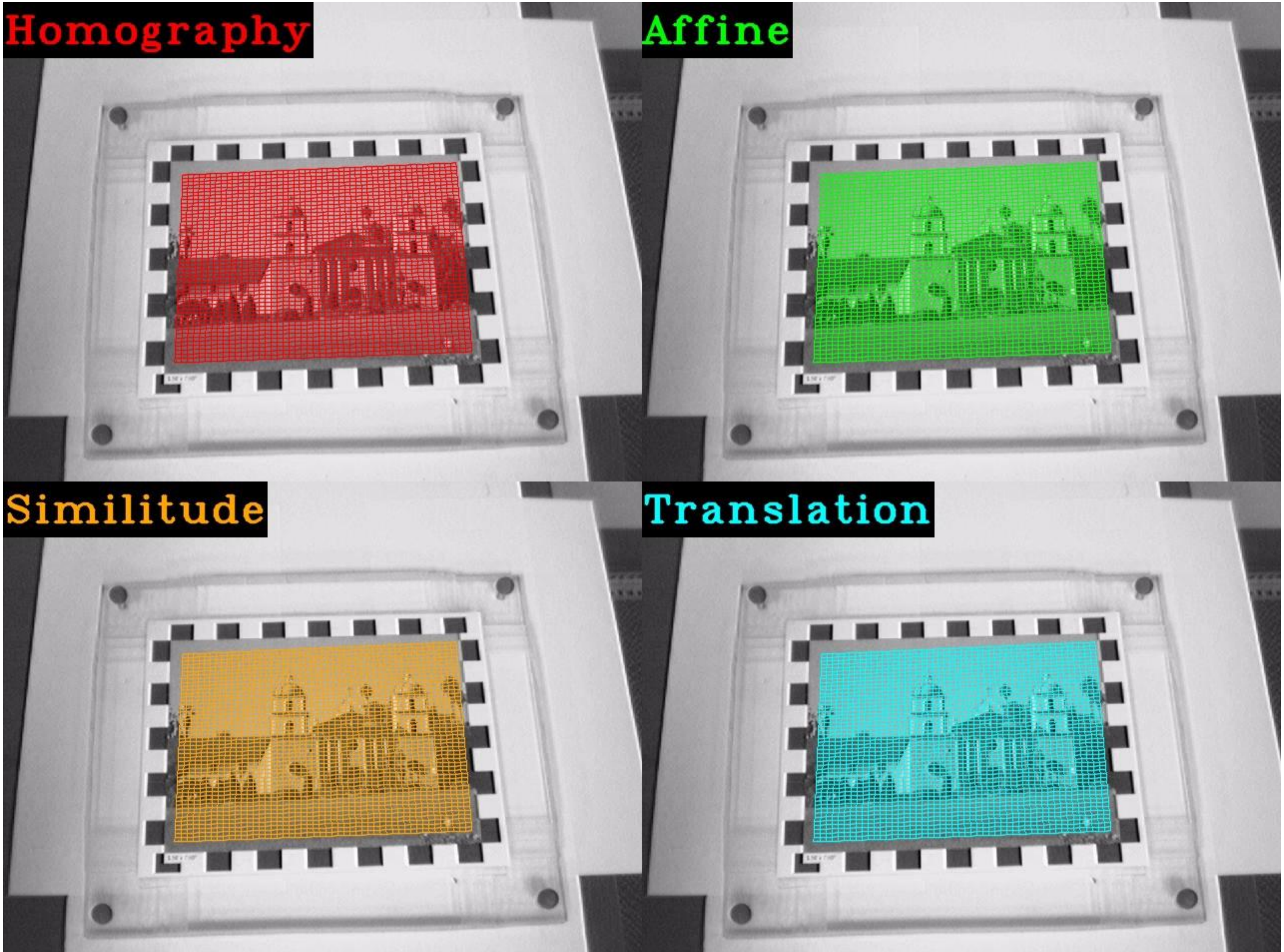
- **Similitude/Similarity**: $S = 4$
  – $\mathbf{w}(\mathbf{x}_k, \mathbf{p}) =$
  $p_4 \begin{bmatrix} x_k \cos p_1 - y_k \sin p_1 + p_2 \\ x_k \sin p_1 + y_k \cos p_1 + p_3 \end{bmatrix}$
- **Affine** : $S = 6$
  – $\mathbf{w}(\mathbf{x}_k, \mathbf{p}) =$
  $\begin{bmatrix} (1 + p_1)x_k - p_2 y_k + p_3 \\ (1 + p_1)y_k + p_2 x_k + p_4 \end{bmatrix}$

State Space Model – Examples (cont'd)

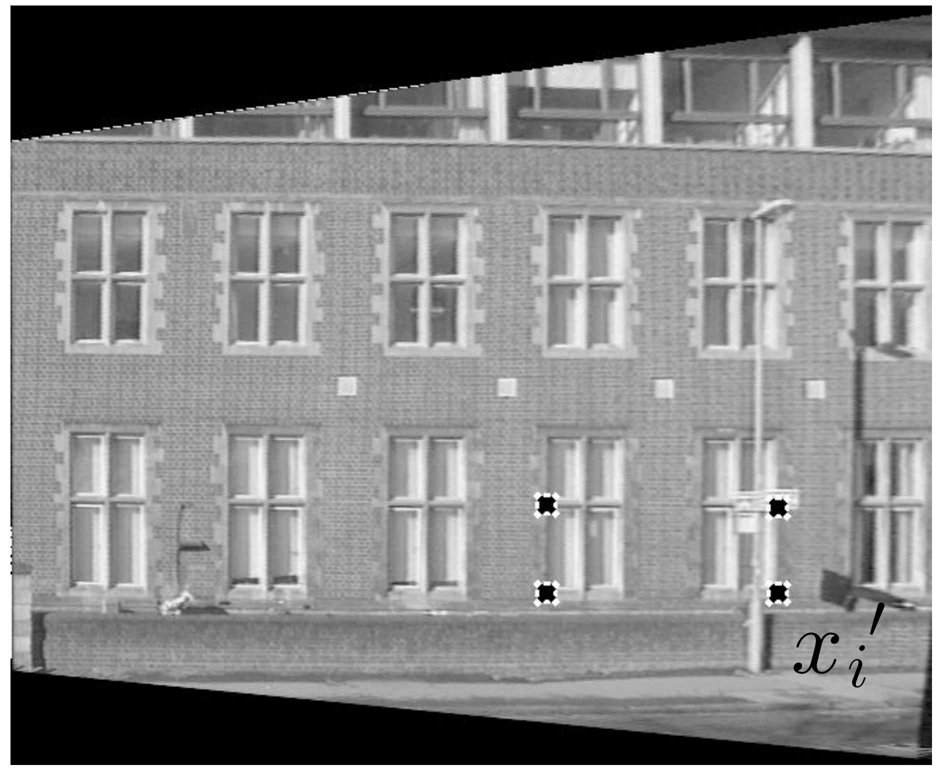
- **Homography** : $S = 8$
  - $\mathbf{w}(\mathbf{x}_k, \mathbf{p}) = \left[\frac{(1+p_1)x_k+p_2y_k+p_3}{(1+p_7)x_k+p_8y_k+1}, \frac{(1+p_4)y_k+p_5x_k+p_6}{(1+p_7)x_k+p_8y_k+1}\right]^T$
- **SL3 Homography**[Benhimane04] : $S = 8$
  - $\mathbf{w}(\mathbf{x}_k, \mathbf{p}) = \boldsymbol{G} \,\hat{*} \begin{bmatrix} x_k \\ y_k \end{bmatrix}$
    - $\boldsymbol{G} = \mathbf{exp}\left(\sum_{i=1}^{8} p_i \boldsymbol{A}_i\right) \in \mathbb{SL}(3)$, $\boldsymbol{A}_i$: $\mathfrak{sl}(3)$ basis
- **Corner Homography** : $S = 8$
  - $\mathbf{w}(\mathbf{x}_k, \mathbf{p}) = \boldsymbol{G} \,\hat{*} \begin{bmatrix} x_k \\ y_k \end{bmatrix}$
    - $\boldsymbol{G} = \underset{\boldsymbol{M}}{\operatorname{argmin}} \sum_{i=1}^{4} \left\| \boldsymbol{M} \,\hat{*} \begin{bmatrix} c_{ix} \\ c_{iy} \end{bmatrix} - \begin{bmatrix} c_{ix} + p_{2i-1} \\ c_{iy} + p_{2i} \end{bmatrix} \right\|^2$
    - $\left\{ c_i = \begin{bmatrix} c_{ix} \\ c_{iy} \end{bmatrix} \middle| 1 \leq i \leq 4 \right\}$: bounding box corners

$\boldsymbol{G} \,\hat{*} \begin{bmatrix} x \\ y \end{bmatrix} = \left[\frac{g_{00}x+g_{01}y+g_{02}}{g_{20}x+g_{21}y+g_{22}}, \frac{g_{10}y+g_{11}x+g_{12}}{g_{20}x+g_{21}y+g_{22}}\right]^T$ with $\boldsymbol{G} = \begin{bmatrix} g_{00} & g_{01} & g_{02} \\ g_{10} & g_{11} & g_{12} \\ g_{20} & g_{21} & g_{22} \end{bmatrix}$

State Space Model – Examples (Demo)



(in 2 weeks) Homography = Planar Projective Warping



A novel view rendered via four points with known structure

$$x_i' = Hx_i$$
$$i = 1 \ldots 4$$

HZ

Results – State Space Models (Demo)

Search Method

$$\mathbf{p_t} = \underset{\mathbf{p}}{\operatorname{argmax}}\ f(\mathbf{I_0}(\mathbf{x}), \mathbf{I_t}(\mathbf{w}(\mathbf{x}, \mathbf{p})))$$

- Optimization method that finds the SSM parameters corresponding to the warped patch that maximizes the AM similarity function.
- Two main categories:
  - Gradient descent
    - Newton or Gauss Newton method
  - Stochastic Search
    - Sampling based

Simple image registration algorithm
SSD error norm

$$E(u,v)=\sum_{x,y}\left(I(x+u,y+v)-T(x,y)\right)^2$$

Exhaustive search:

For each offset *(u, v)*

compute *E(u,v)*;

Choose *(u, v)* which minimizes *E(u,v)*;

(Gauss) Newton optimization:

Solve

u = M\Im_t

$$\begin{pmatrix} \vdots \\ -\frac{\partial \mathrm{Im}}{\partial t} \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots \\ \frac{\partial \mathrm{Im}}{\partial x} & \frac{\partial \mathrm{Im}}{\partial y} \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} u_x \\ u_y \end{pmatrix}$$

Search Method – Examples (Gradient Descent)

- Variants of Lucas Kanade (LK)[Baker01] method
  - Forward Additive (**FALK**)
    - $\mathbf{\Delta p_t} = \underset{\mathbf{\Delta p_t}}{\mathbf{argmax}}\, f\left(\mathbf{I_0}(\mathbf{x}), \mathbf{I_t}\big(\mathbf{w}(\mathbf{x}, \mathbf{p_{t-1}} + \mathbf{\Delta p_t})\big)\right)$
      - $\mathbf{p_t} = \mathbf{p_{t-1}} + \mathbf{\Delta p_t}$
  - Inverse Additive (**IALK**)
    - uses constant approximation of $\mathbf{\nabla I_t}$ computed from $\mathbf{I_0}$
  - Forward Compositional (**FCLK**)
    - $\mathbf{\Delta p_t} = \underset{\mathbf{\Delta p_t}}{\mathbf{argmax}}\, f\left(\mathbf{I_0}(\mathbf{x}), \mathbf{I_t}\big(\mathbf{w}(\mathbf{w}(\mathbf{x}, \mathbf{\Delta p_t}), \mathbf{p_{t-1}})\big)\right)$
      - $\mathbf{p_t} = \mathbf{p_{t-1}} \circ \mathbf{\Delta p_t}$
  - Inverse Compositional (**ICLK**)
    - $\mathbf{\Delta p_t} = \underset{\mathbf{\Delta p_t}}{\mathbf{argmax}}\, f\left(\mathbf{I_0}\big(\mathbf{w}(\mathbf{x}, \mathbf{\Delta p_t})\big), \mathbf{I_t}\big(\mathbf{w}(\mathbf{x}, \mathbf{p_{t-1}})\big)\right)$
      - $\mathbf{p_t} = \mathbf{p_{t-1}} \circ \mathbf{\Delta p_t}^{-1}$
- Efficient Second Order Minimization (**ESM**)[Benhimane04]
  - combines FCLK and ICLK

S. Baker, I. Matthews, "*Equivalence and Efficiency of Image Alignment Algorithms*", 2001
S. Benhimane, E. Malis, "*Real-time image-based tracking of planes using efficient second-order minimization*", 2004

Homogenous coordinates: How to translate a 2D point:

- Old way: x' = x+dx

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

- New way: x'=M*dx = M o dx

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \Delta x \\ 0 & 1 & \Delta y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

- Can chain many transf: x'=M1*M2*dx
- Euclidean transform SE2:

$$p' = \begin{pmatrix} & R & & T \\ 0 & 0 & 0 & 1 \end{pmatrix} p$$

Search Method – Examples (Stochastic)

- Nearest Neighbor Search (**NN**) [Dick13]
  - generate samples by warping $\mathbf{I_0}(\mathbf{x})$
  - find the nearest neighbor to $\mathbf{I_t}(\mathbf{w}(\mathbf{x}, \mathbf{p_{t-1}}))$ and update $\mathbf{p_{t-1}}$ with the *inverse* of the corresponding $\mathbf{\Delta p_t}$
  - combined with ICLK for stability (**NNIC**)
- Particle Filter (**PF**) [Kwon14]
  - generate samples by warping $\mathbf{I_t}(\mathbf{w}(\mathbf{x}, \mathbf{p_{t-1}}))$
  - compute weight for each and estimate $\mathbf{\Delta p_t}$ as weighted average of samples

T. Dick et. al, "*Realtime Registration-Based Tracking via Approximate Nearest Neighbor Search*", 2013
J. Kwon, H. S. Lee, F. C. Park, K. M. Lee, "*A Geometric Particle Filter for Template-Based Visual Tracking*", 2014

Appearance Model

$$\mathbf{p_t} = \underset{\mathbf{p}}{\operatorname{argmax}} \; f(\mathbf{I_0}(\mathbf{x}), \mathbf{I_t}(\mathbf{w}(\mathbf{x}, \mathbf{p})))$$

- A similarity measure between two image patches:
  - candidate warped patch from the current image
  - template extracted from the initial image
- Two main categories:
  - SSD like
  - Robust[Richa12]

R. Richa, R. Sznitman, G. Hager, "*Robust Similarity Measures for Gradient-based Direct Visual Tracking*", 2012

Appearance Model – Examples

- Sum of Squared Differences (**SSD**)[Baker01]
  - $f(\mathbf{I_0}, \mathbf{I_t}) = -\frac{1}{2} \parallel \mathbf{I_0} - \mathbf{I_t} \parallel^2$
- Sum of Conditional Variance (**SCV**)[Richa11]
  - $f(\mathbf{I_0}, \mathbf{I_t}) = -\frac{1}{2} \parallel E\left[\mathbf{I_t} \middle| \mathbf{I_0}\right] - \mathbf{I_t} \parallel^2$
  - Using several joint distributions computed from corresponding sub regions of $\mathbf{I_t}$ and $\mathbf{I_0}$ gives a variant called **LSCV**[Richa14]
- Reversed Sum of Conditional Variance (**RSCV**)[Dick13]
  - $f(\mathbf{I_0}, \mathbf{I_t}) = -\frac{1}{2} \parallel \mathbf{I_0} - E\left[\mathbf{I_0} \middle| \mathbf{I_t}\right] \parallel^2$
- Zero mean Normalized Cross Correlation (**ZNCC**)[Ruthotto10]
  - $f(\mathbf{I_0}, \mathbf{I_t}) = -\frac{1}{2} \parallel \frac{\mathbf{I_0} - \mu_0}{\sigma_0} - \frac{\mathbf{I_t} - \mu_t}{\sigma_t} \parallel^2$

R. Richa, R. Sznitman, R. Taylor, G. Hager, "*Visual Tracking Using the Sum of Conditional Variance*", 2011
R. Richa, et. al, "*Direct visual tracking under extreme illumination variations using the sum of conditional variance*", 2014
L. Ruthotto, "*Mass-preserving registration of medical images*", 2010

Appearance Model – Examples (cont'd)

- Mutual Information (**MI**)[Dame10]

$$-f(\mathbf{I_0}, \mathbf{I_t}) = \sum_{ij} \boldsymbol{P}_{\boldsymbol{I_t I_0}}(i,j) \mathbf{log}\left(\frac{\boldsymbol{P}_{\boldsymbol{I_t I_0}}(i,j)}{\boldsymbol{P}_{\boldsymbol{I_t}}(i)\boldsymbol{P}_{\boldsymbol{I_0}}(j)}\right)$$

- Cross Cumulative Residual Entropy (**CCRE**)[Richa12]

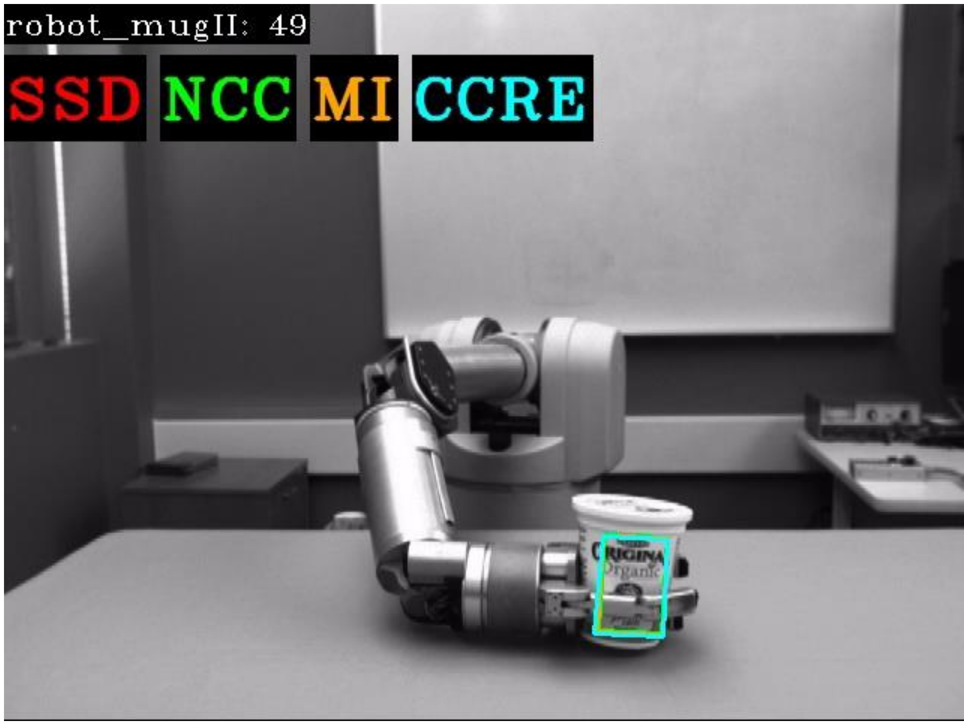
$$-f(\mathbf{I_0}, \mathbf{I_t}) = \sum_{ij} \boldsymbol{P}^*_{\boldsymbol{I_t I_0}}(i,j) \mathbf{log}\left(\frac{\boldsymbol{P}^*_{\boldsymbol{I_t I_0}}(i,j)}{\boldsymbol{P}^*_{\boldsymbol{I_t}}(i)\boldsymbol{P}_{\boldsymbol{I_0}}(j)}\right)$$

- Normalized Cross Correlation (**NCC**)[Scandaroli12]

$$-f(\mathbf{I_0}, \mathbf{I_t}) = \frac{\mathbf{I_0} - \mu_0}{\sigma_0} . \frac{\mathbf{I_t} - \mu_t}{\sigma_t}$$

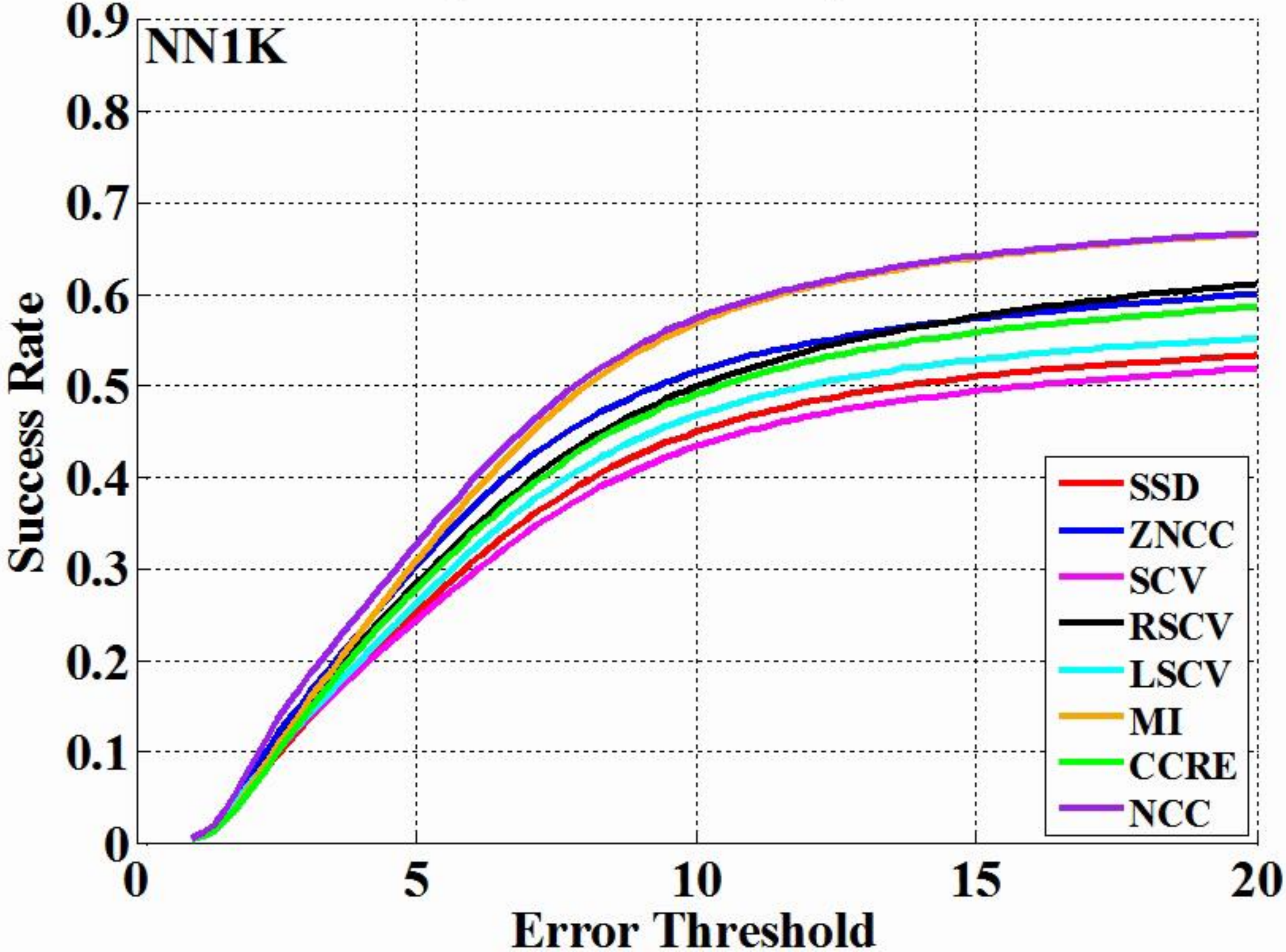
A. Dame, E. Marchand, "Accurate Real-time Tracking Using Mutual Information", 2010
G. G. Scandaroli, M. Meilland, R. Richa, "Improving NCC-Based Direct Visual Tracking", 2012

Results – Appearance Models (Demo)

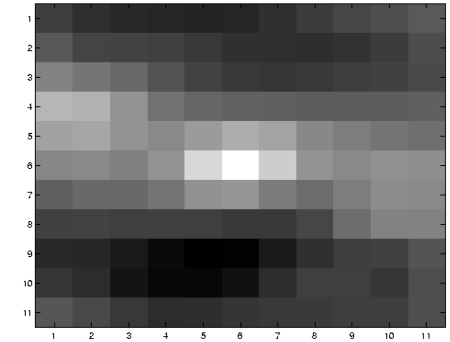


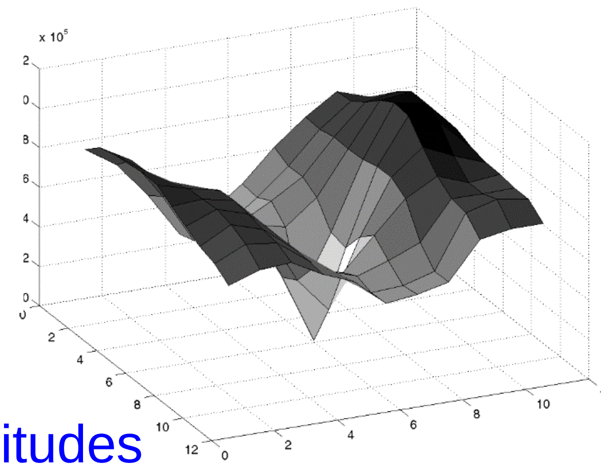
FCLK with **Homography**

Results – Appearance Models (Stochastic)



High textured region



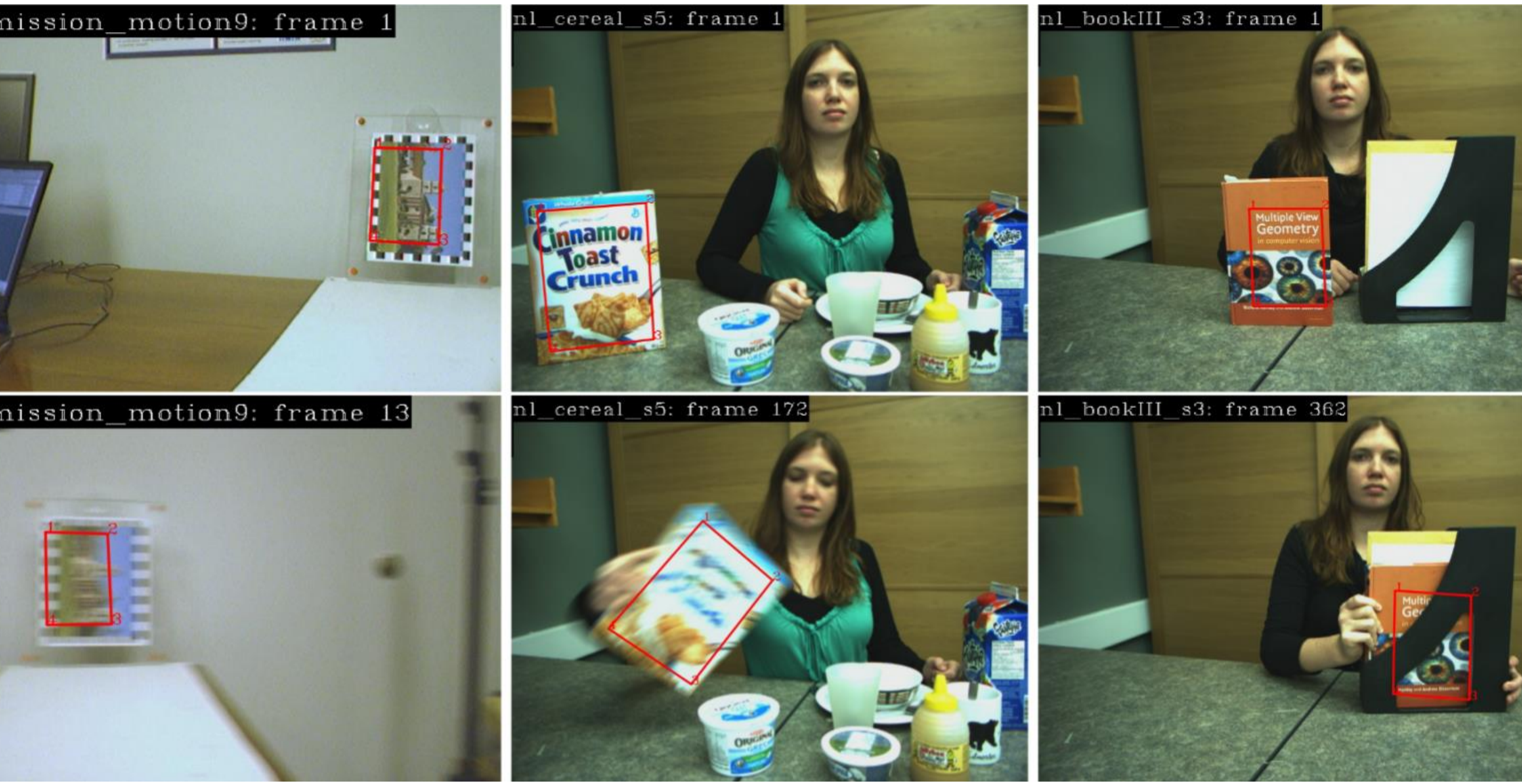


$$\sum \nabla I (\nabla I)^T$$

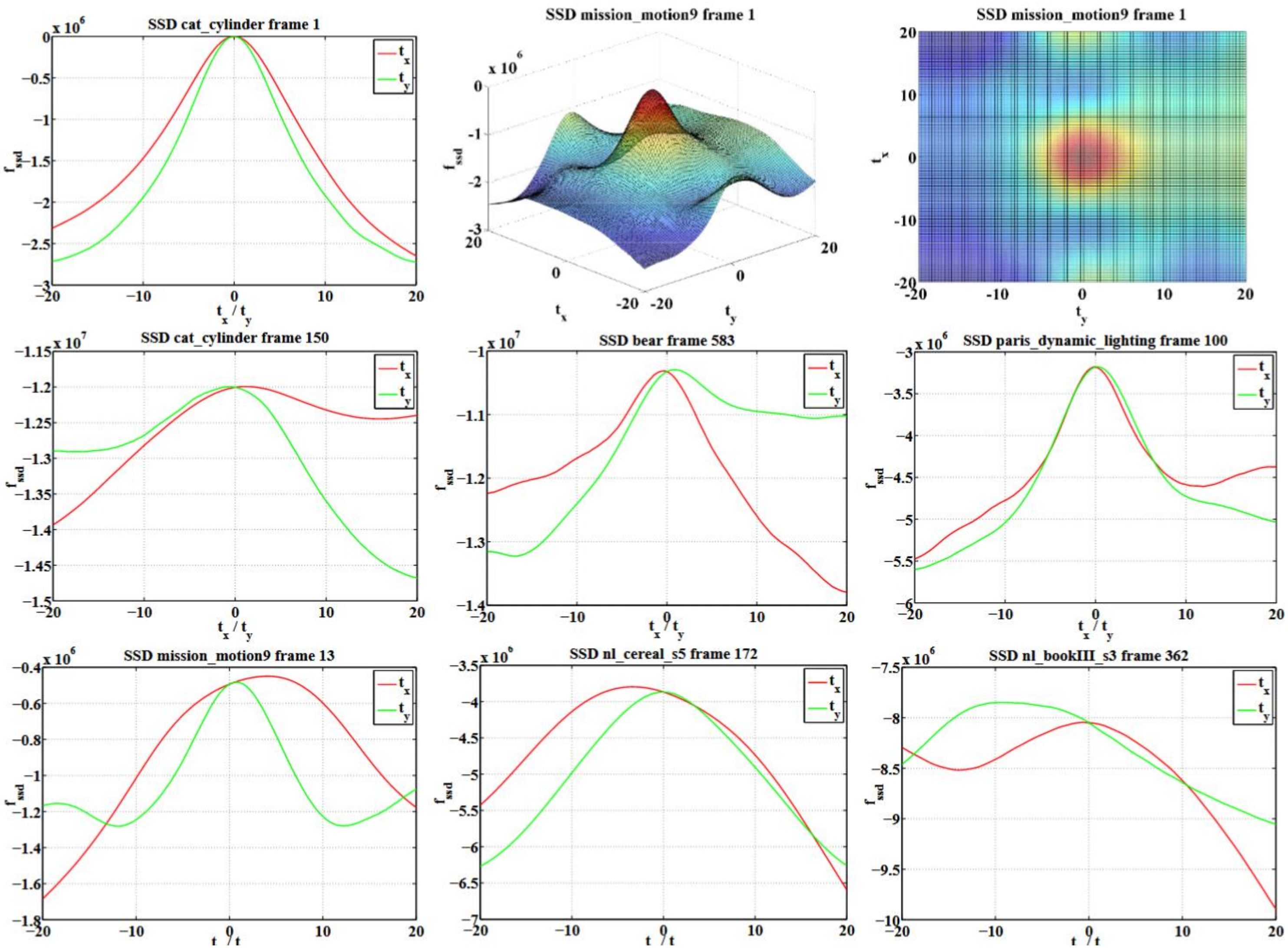
– gradients are different, large magnitudes

– large λ_1, large λ_2

Appearance model
Test images

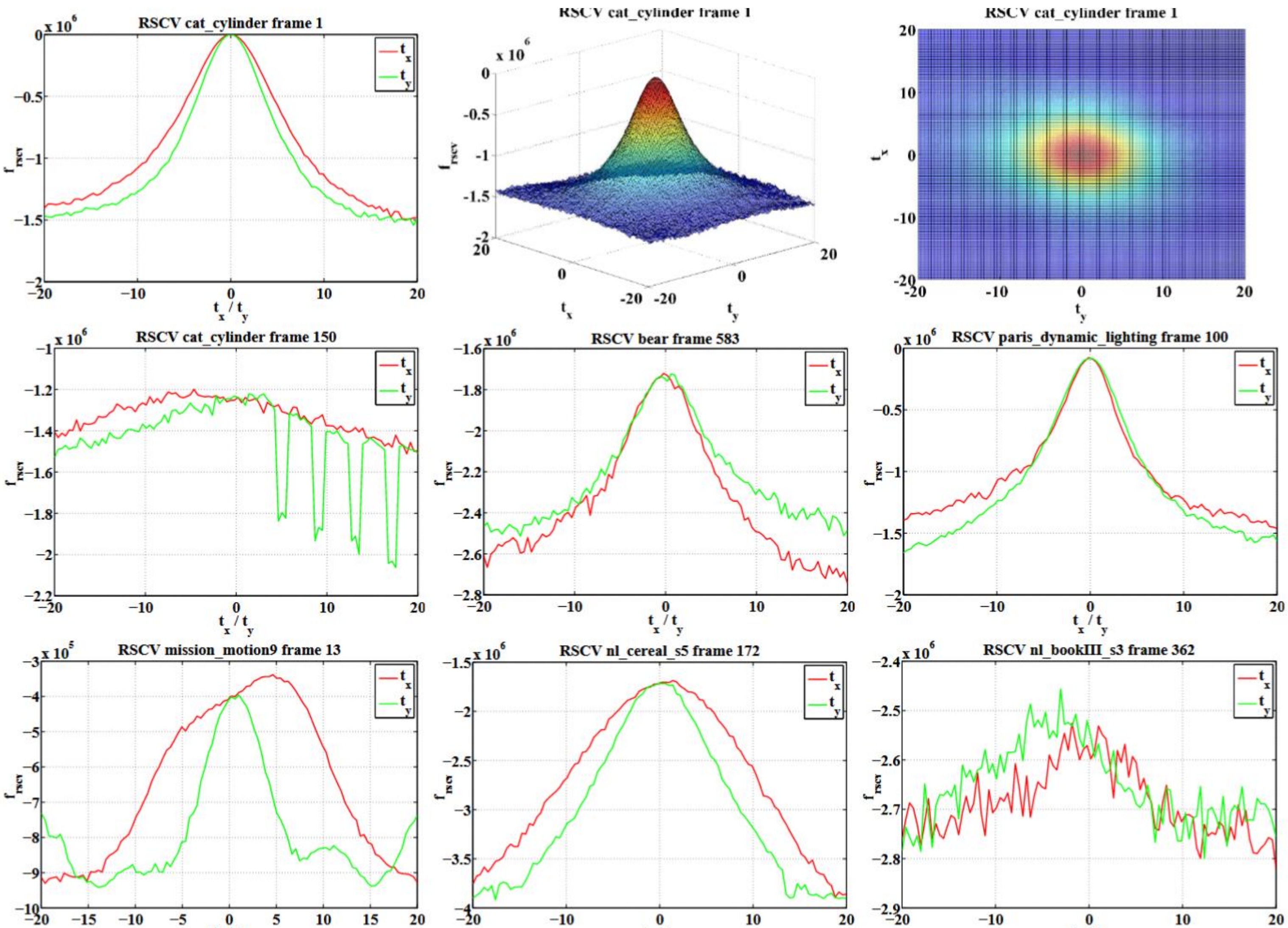


Appearance model
L^2 ||T-I||^2 aka "SSD"



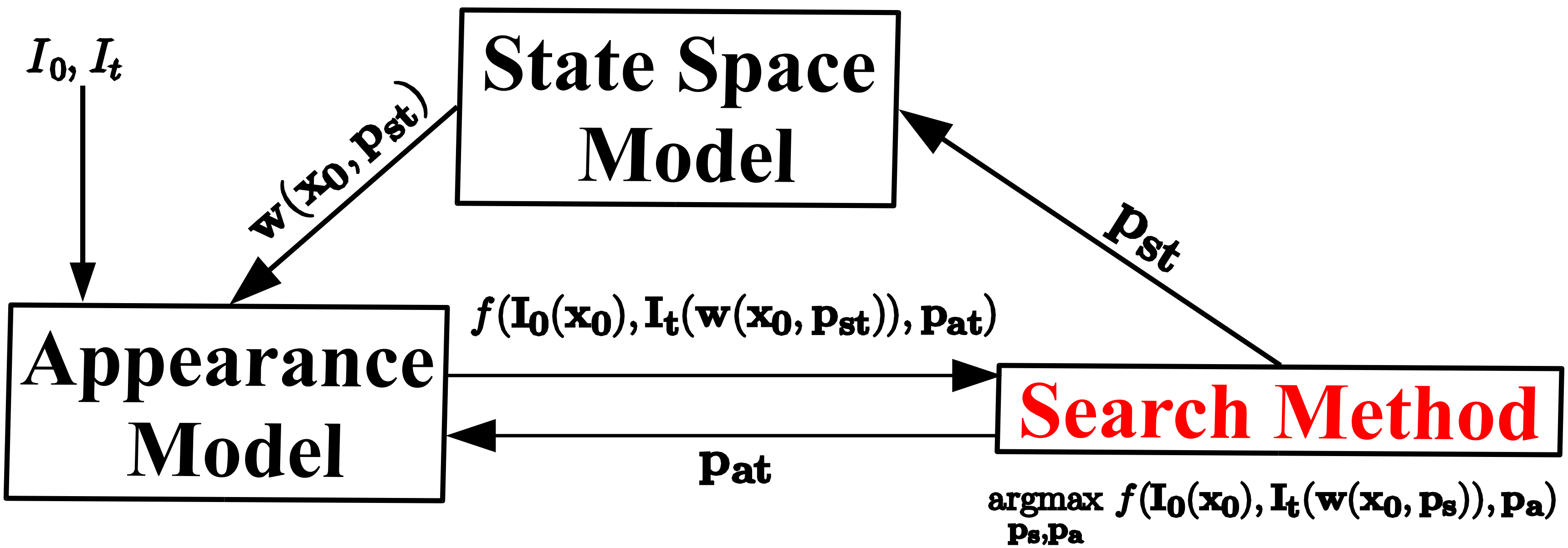
Appearance model

RSCV – Reversed Sum of Conditional Variance



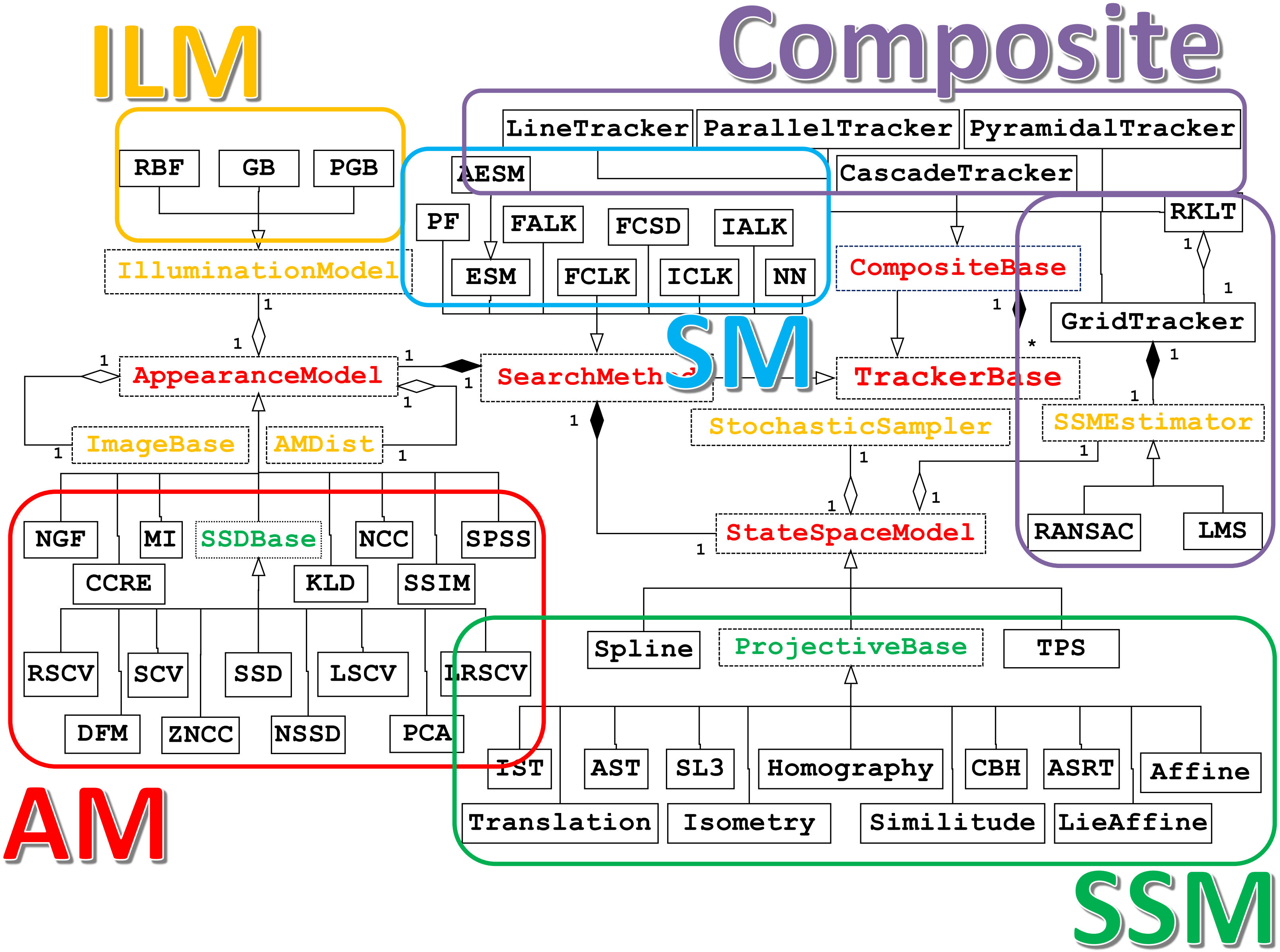
System design

$$\mathbf{p_t} = \underset{\mathbf{p_s},\mathbf{p_a}}{\operatorname{argmax}} f(\mathbf{I_0}(\mathbf{x_0}), \mathbf{I_t}(\mathbf{w}(\mathbf{x_0}, \mathbf{p_s})), \mathbf{p_a})$$



- Search Method (**SM**)
  - **Finds** the warp that maximizes the similarity measure

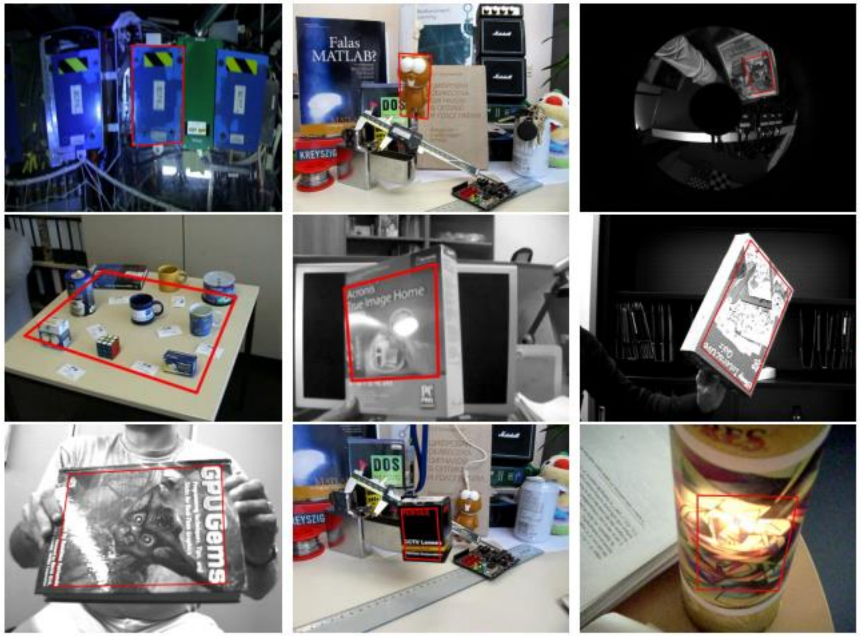
System Design



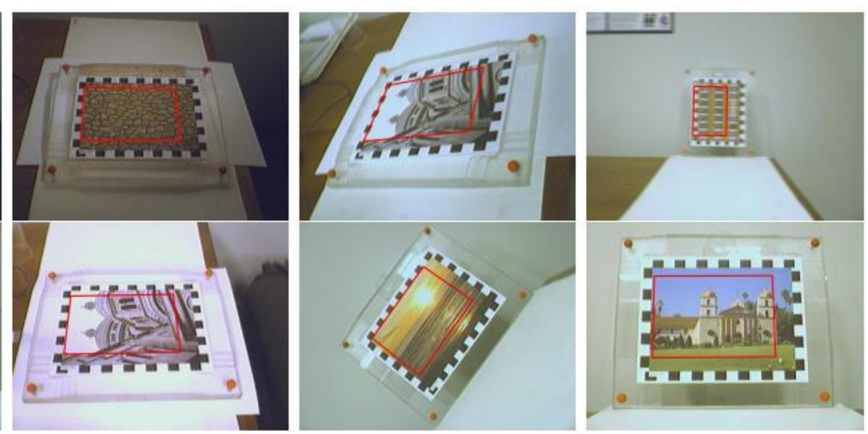
Evaluation Benchmarks

TMT

PAMI



UCSB



LinTrack

Evaluation Methodology - Datasets

- 4 large publicly available datasets with a total of over **100K** frames
  - TMT
  - UCSB
  - LinTrack
  - PAMI

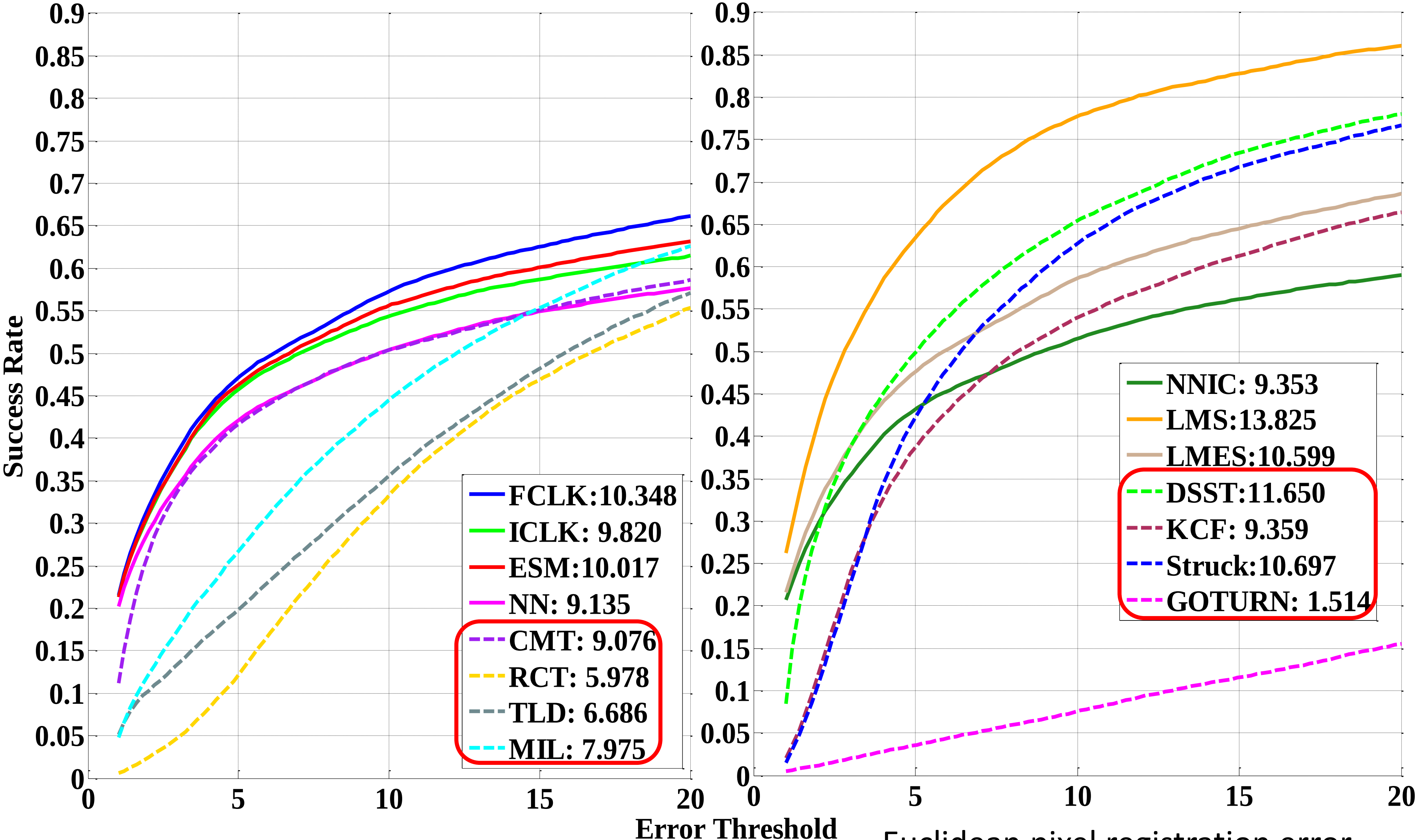
Dataset	Without Subsequences			With Subsequences		
	Sequences	Total Frames	Trackable Frames	Sub-sequences	Total Frames	Trackable Frames
TMT	109	70592	70483	1090	390470	389380
UCSB	96	6889	6793	960	41170	40210
LinTrack	3	12477	12474	30	68700	68670
PAMI	28	16511	16483	280	91400	91120
Total	**236**	**106469**	**106233**	**2360**	**591740**	**589380**

- Each sequence tested from 10 different starting points for an effective total of nearly **600K** frames

Evaluation Methodology – Performance Metric

- **Alignment Error** (E_{AL})
  - $\boldsymbol{E_{AL} = \frac{1}{4}\|C_{track} - C_{gt}\|}$
- **Success Rate** (SR)
  - **x** axis : error threshold $t_p \in [0, 20]$
  - **y** axis : fraction of frames with $E_{AL} < t_p$
  - each sequence tracked from 10 different starting points
  - measures both accuracy and robustness
- **Failure Rate** (FR)
  - reinitialize whenever E_{AL} exceeds 20
  - count the number of such failures
  - additional metric for tracking robustness

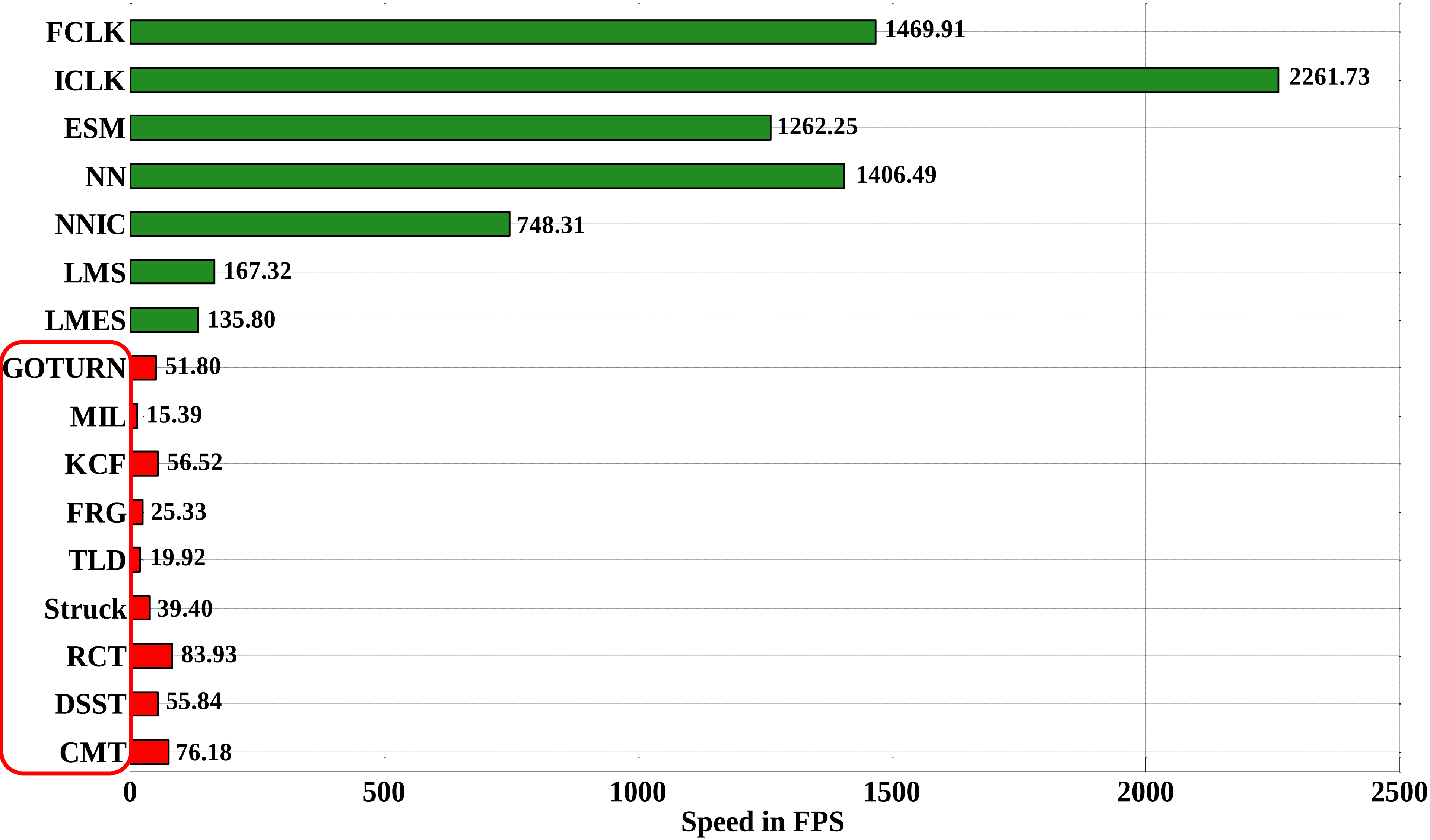
Results: Learning vs. 2DOF Registration Based Trackers (Accuracy)



Euclidean pixel registration error

RBT vs Learn vs GoTurn

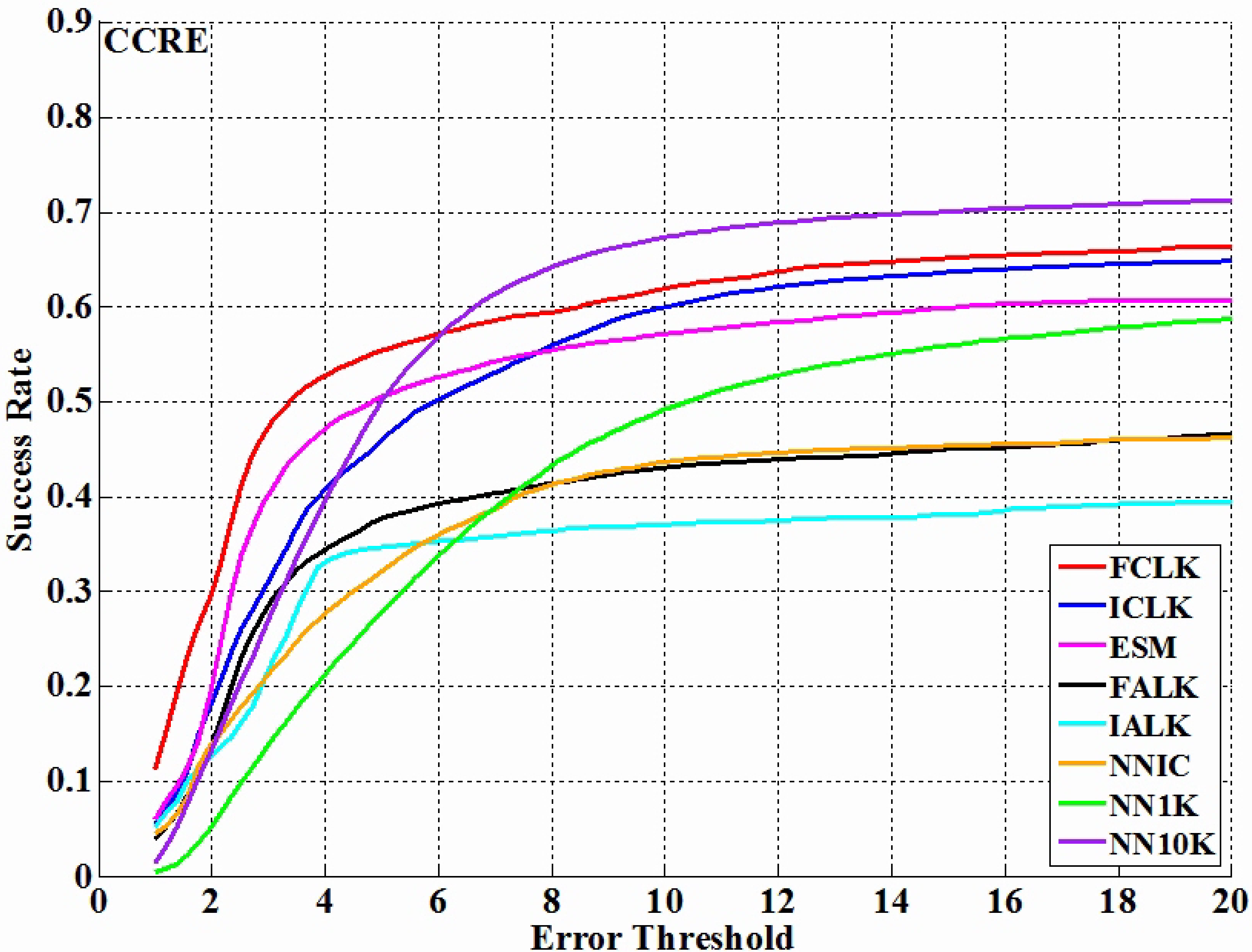
Results: Learning vs. 2DOF Registration Based Trackers (Speed)



Results –Learning Based Trackers (Demo)

ZNCC with **Translation**

Results – Search Methods (Robust)



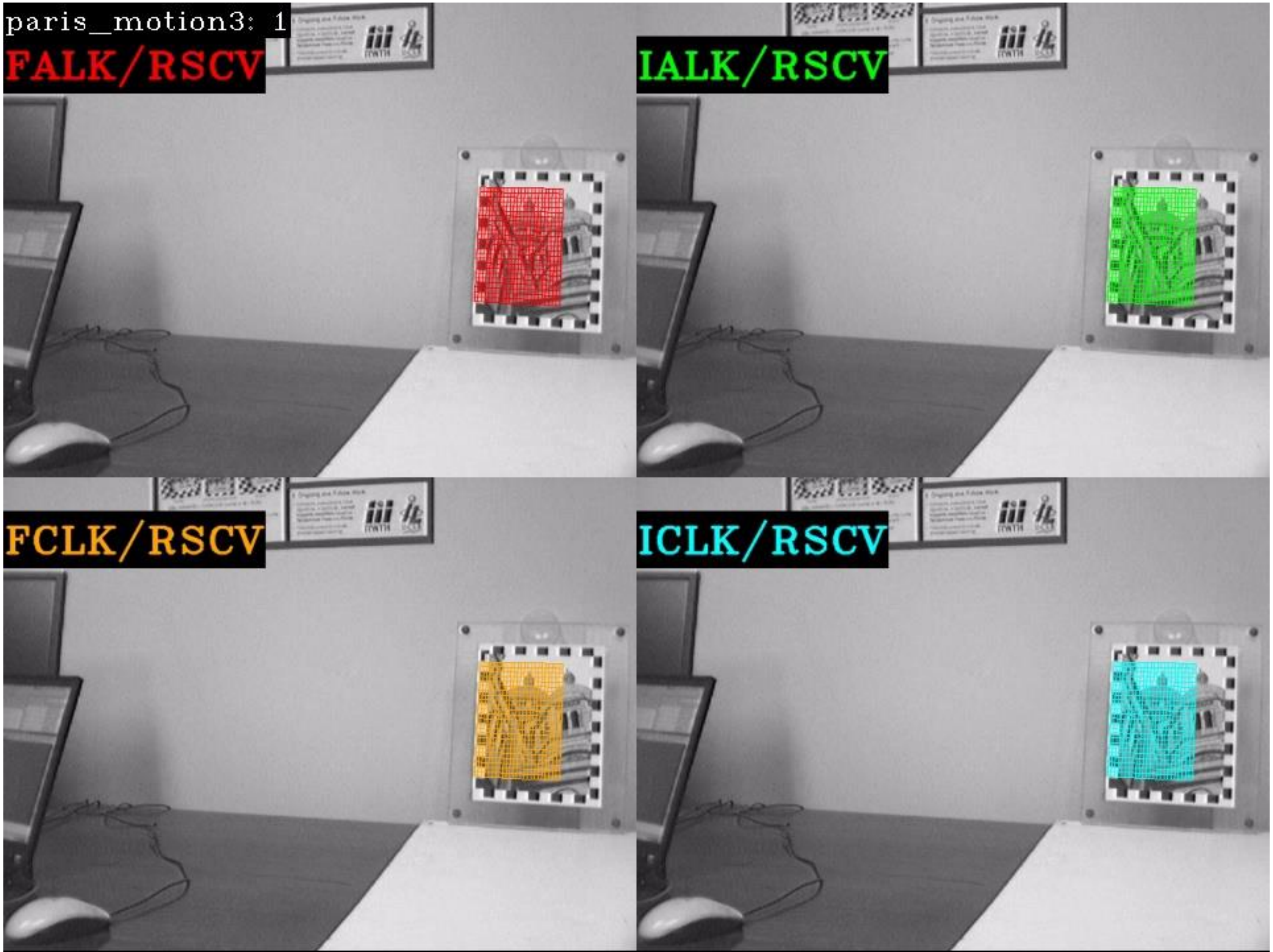
Results – Search Methods (Demo)

- The four variants of Lucas Kanade fail at different times
- Sequences from TMT

RSCV with **Homography**

Results – Search Methods (Demo)

- The four variants of Lucas Kanade fail at different times
- Sequence from UCSB

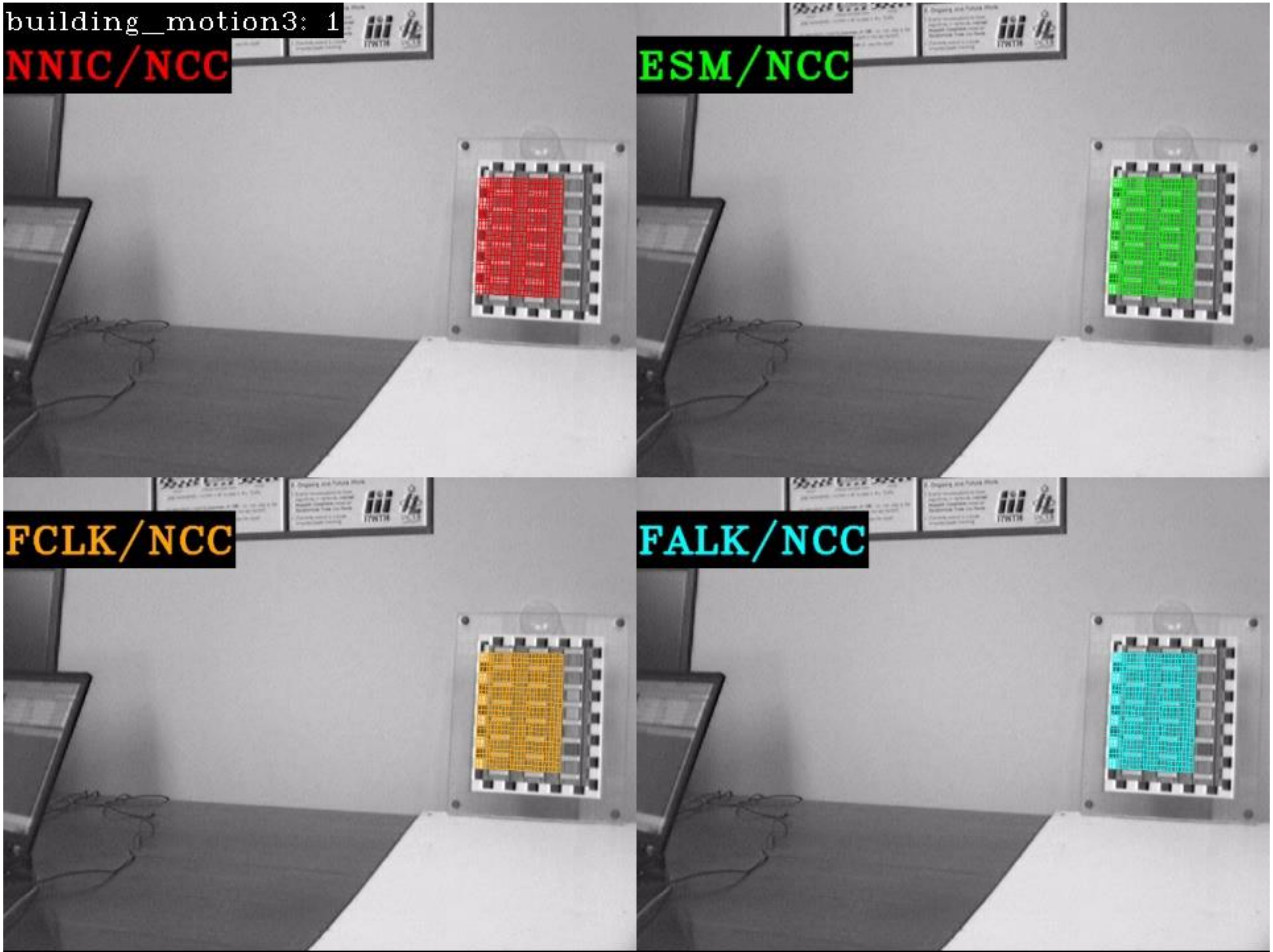


Results – Search Methods (Demo)

- NN has more jitter than LK type SMs
  - decreases with more samples

Results – Search Methods (Demo)

- NNIC is more robust to motion blur
- Sequence from UCSB



Conclusions

Questions?

- Tested **different combinations** of sub modules leading to several interesting observations that were missing in the original papers.
  - used two large datasets with over 77,000 frames in all to ensure statistical significance.
- Compared **robust similarity metrics** with traditional SSD type measures.
- Compared formulations against **online learning based trackers** to validate their usability for precise tracking
- Provided **an open source tracking framework** called **MTF** using which all results can be reproduced
  - can also address practical tracking requirements with its efficient C++ implementation

Conclusions

MTF is available at: http://webdocs.cs.ualberta.ca/~vis/mtf/ along with all datasets and this presentation

Project and Research opportunities in video tracking

- Combine learning and registration tracking search methods
  - Direct deep network methods not precise
  - Predict with deep network, refine with registration
- New appearance models
  - Kullback-Liebler divergence
  - AM based on deep features
- More detailed experimental evaluation
  - Study failure causes in individual frames, solve those
  - Our TMT benchmark data marked up for this (per-fame annotation of blur, motion etc.)

Questions ?

Conclusions

- Tested **different combinations** of sub modules leading to several interesting observations that were missing in the original papers.
  - used two large datasets with over 77,000 frames in all to ensure statistical significance.
- Compared **robust similarity metrics** with traditional SSD type measures.
- Compared formulations against **online learning based trackers** to validate their usability for precise tracking
- Provided **an open source tracking framework** called **MTF** using which all results can be reproduced
  - can also address practical tracking requirements with its efficient C++ implementation

MTF is available at: http://webdocs.cs.ualberta.ca/~vis/mtf/ along with all datasets and this presentation

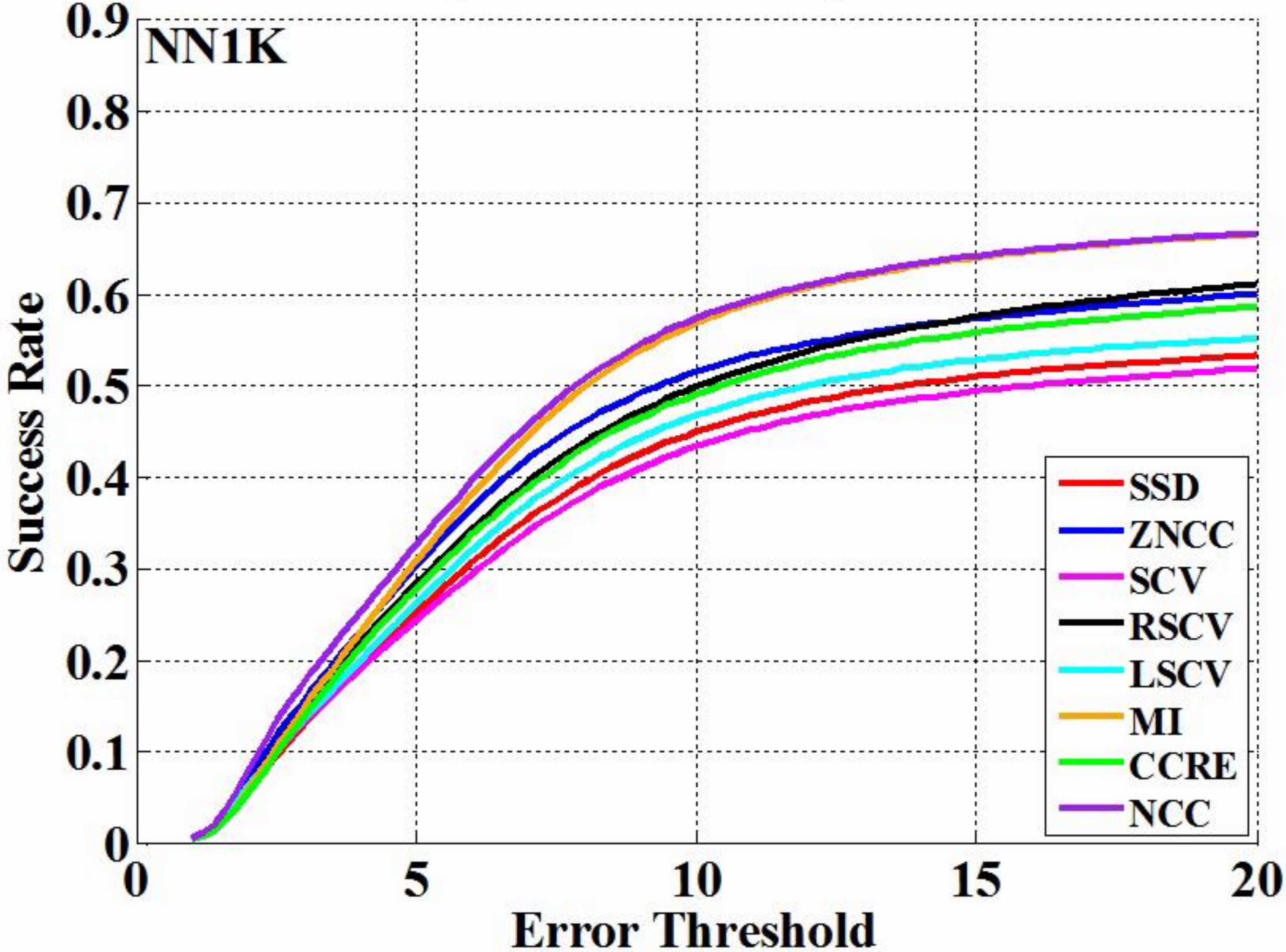
References

- S. Baker and I. Matthews, "*Equivalence and Efficiency of Image Alignment Algorithms*", CVPR 2001
- S. Benhimane and E. Malis, "*Real-time image-based tracking of planes using efficient second-order minimization* ", IROS 2004
- A. Dame and E. Marchand, "*Accurate Real-time Tracking Using Mutual Information*", ISMAR 2010
- T. Dick, C. Perez, A. Shademan and M. Jagersand, "*Realtime Registration-Based Tracking via Approximate Nearest Neighbor Search*", RSS 2013
- S. Gauglitz, T. Hollerer and M. Turk. "*Evaluation of interest point detectors and feature descriptors for visual tracking*", IJCV 2011
- J. Kwon, H. S. Lee, F. C. Park, and K. M. Lee, "*A Geometric Particle Filter for Template-Based Visual Tracking*", TPAMI 2014
- R. Richa, R. Sznitman, R. Taylor and G. Hager, "*Visual Tracking Using the Sum of Conditional Variance*", IROS 2011

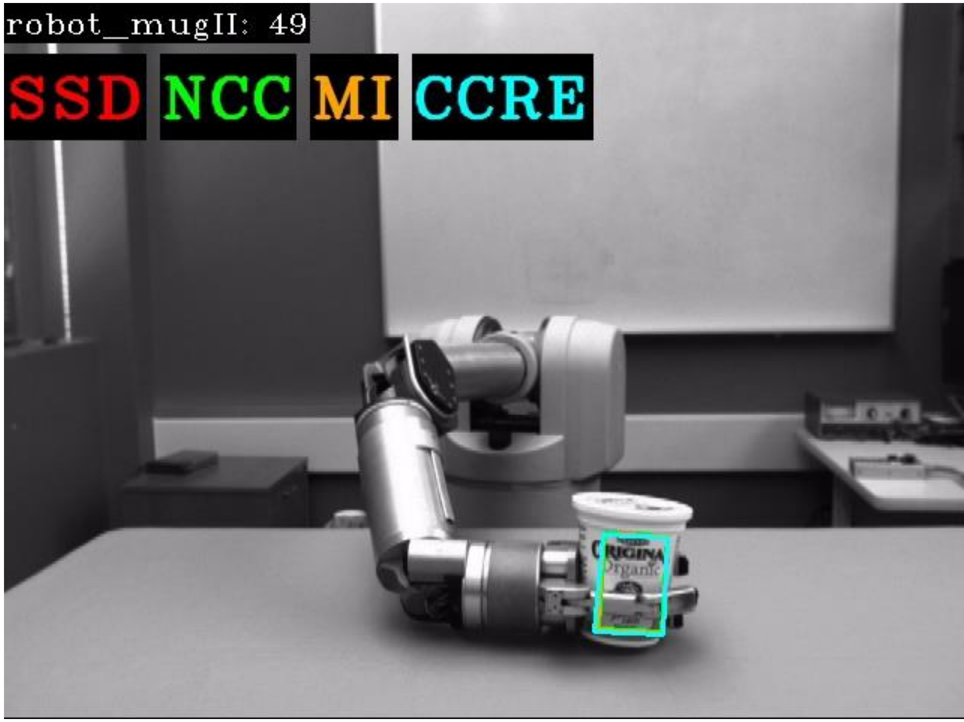
References (cont'd)

- R. Richa, R. Sznitman and G. Hager, "*Robust Similarity Measures for Gradient-based Direct Visual Tracking*", CIRL Technical Report 2012
- R. Richa, M. Souza, G. Scandaroli, E. Comunello and A. Wangenheim, "*Direct visual tracking under extreme illumination variations using the sum of conditional variance*", ICIP 2014
- A. Roy, X. Zhang, N. Wolleb, C. P. Quintero and M. Jagersand, "*Tracking Benchmark and Evaluation for Manipulation Tasks* ", ICRA 2015
- L. Ruthotto, "*Mass-preserving registration of medical images*", Thesis 2010
- G. G. Scandaroli, M. Meilland and R. Richa, "*Improving NCC-Based Direct Visual Tracking*", ECCV 2012
- A. Singh and M. Jagersand, "*Modular Tracking Framework: A Unified Approach to Registration based Tracking*", arXiv:1602.09130, 2016

Results – Appearance Models (Stochastic)

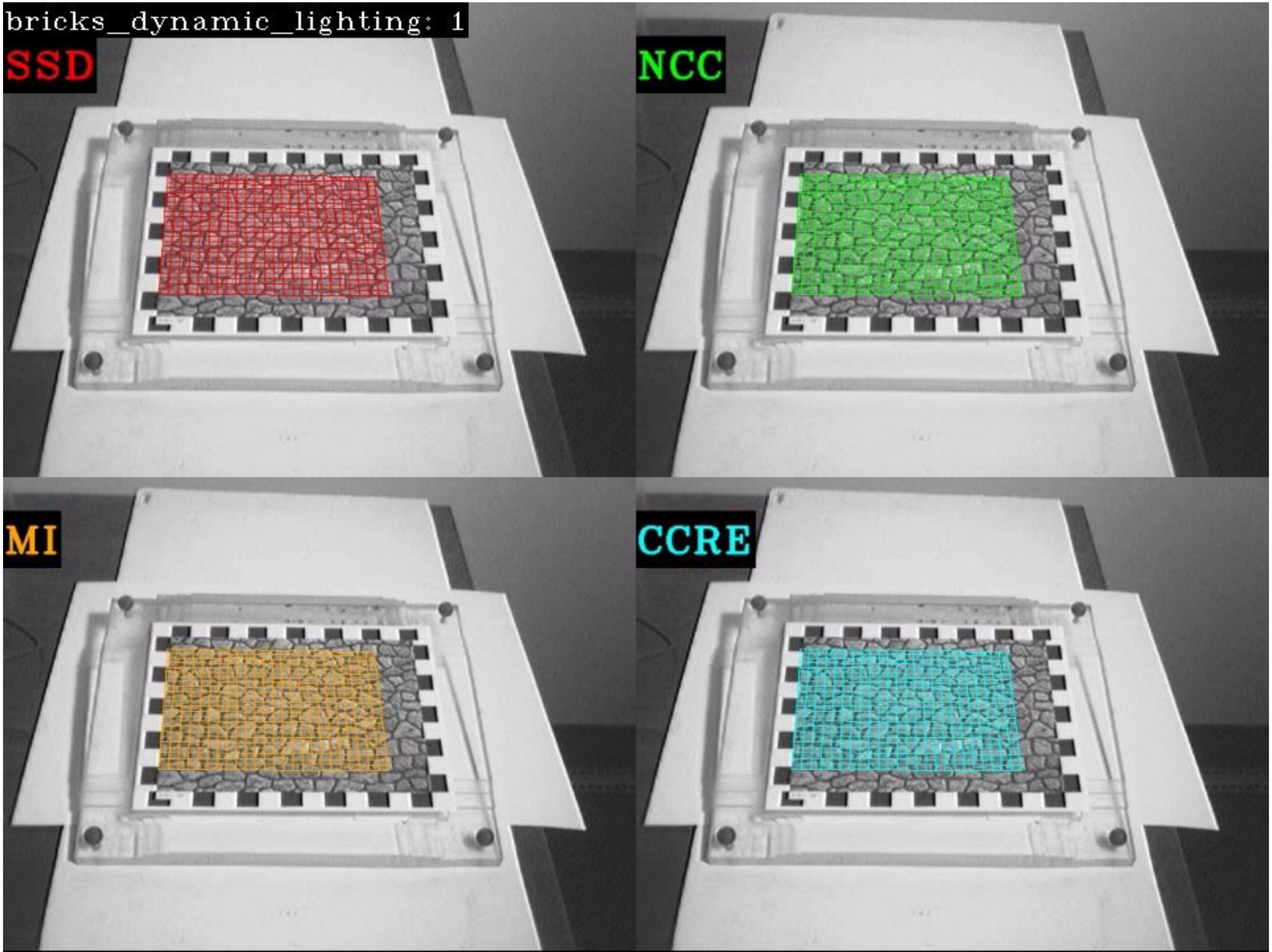


Results – Appearance Models (Demo)



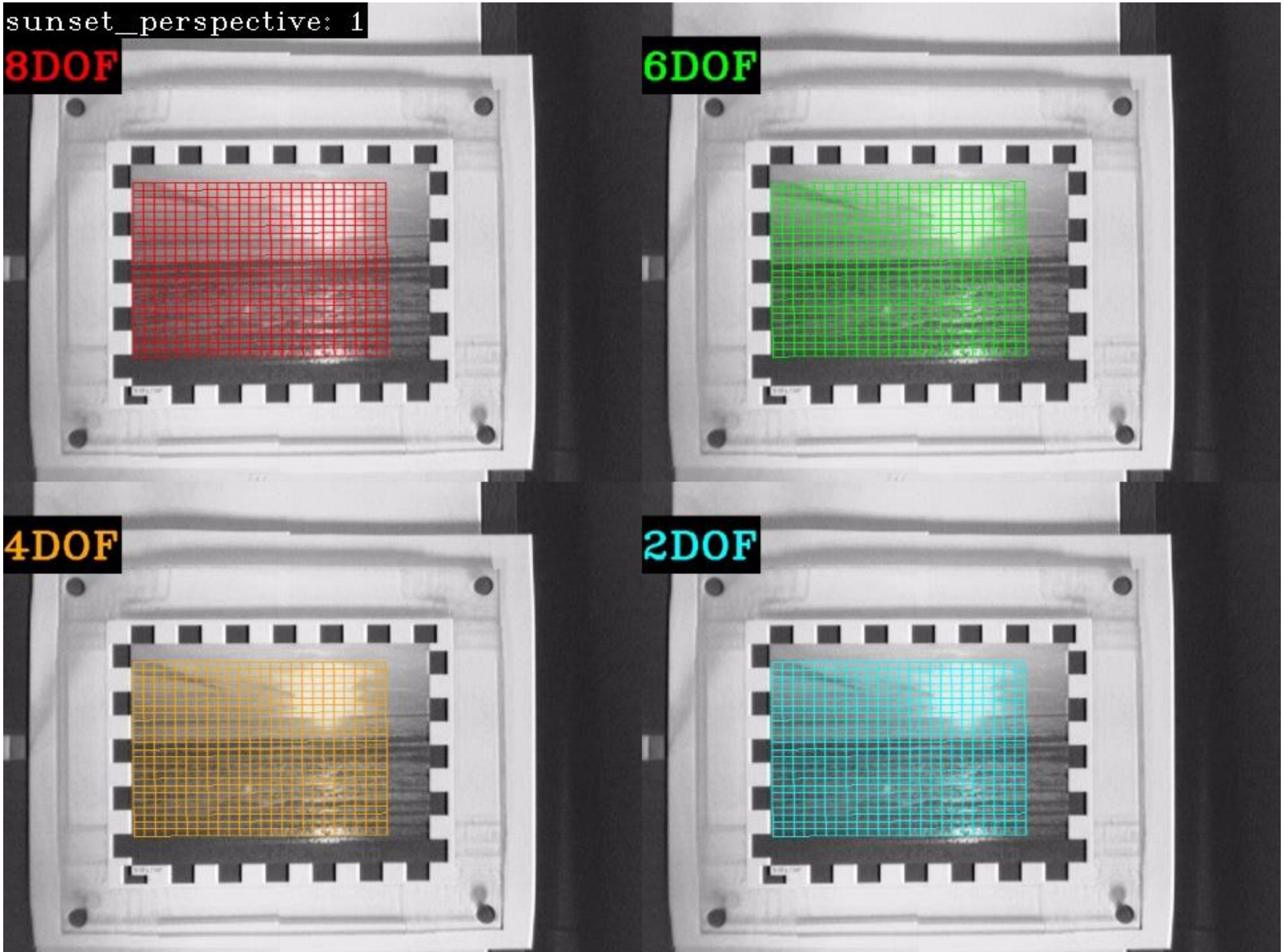
FCLK with **Homography**

Results – Appearance Models (Demo)

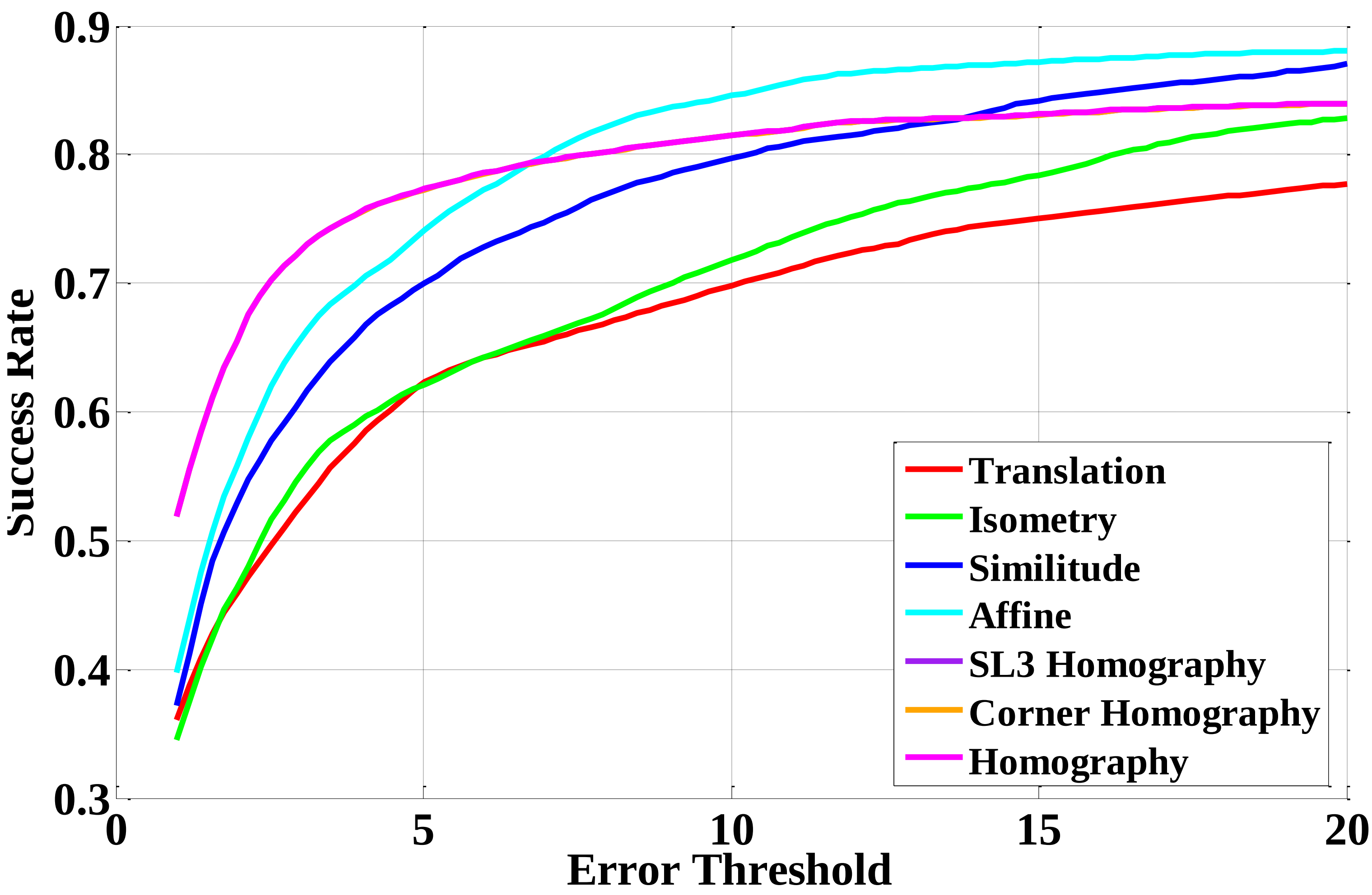


Results – State Space Models (Demo)

Results – State Space Models (Demo)



Results – State Space Models



ESM with **ZNCC**