Spatio-temporal Registration of Multiple Gestures for Skills Analysis and Automation in Robotic Surgery

Nicolas Padoy

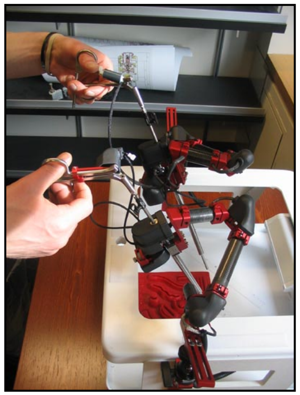
University of Strasbourg, France

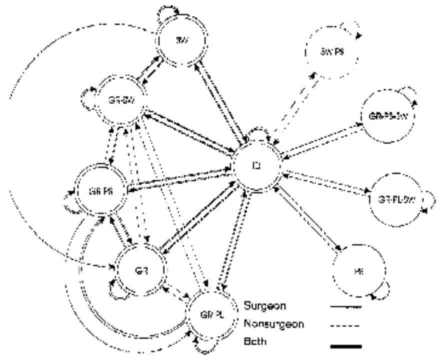
LABORATORY FOR Computational Sensing + Robotics THE JOHNS HOPKINS UNIVERSITY

March 13th, 2013

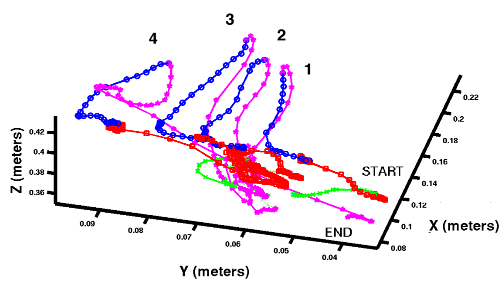
iCUBE

Introduction





Rosen, Brown, Chang, Sinanan, Hannaford. **Generalized approach for modeling minimally invasive surgery as a stochastic process using a discrete markov model**. IEEE Transactions in Biomedical Engineering, 53(3):399–413, 2006.

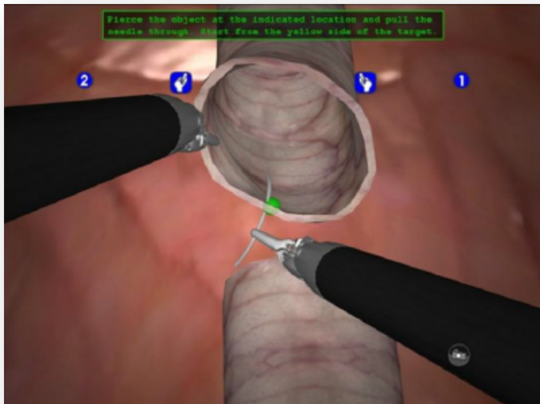


Lin, Shafran, Yuh, Hager. **Towards automatic skill evaluation: Detection and segmentation of robot-assisted surgical motions**. Computer Aided Surgery, 11(5):220–230, 2006

→ Few experiments using such techniques on real data

→ Other applications can benefit from gesture modeling

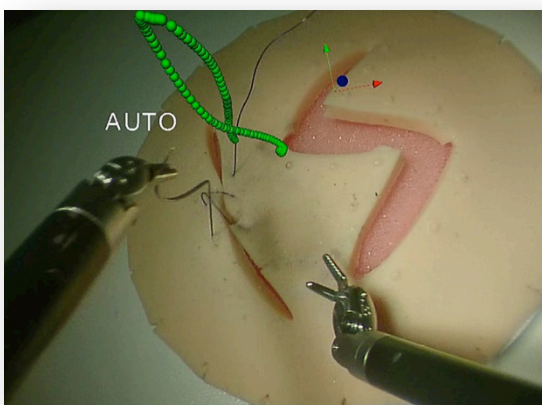
Motivation for Gesture Analysis (1)



Mimic da Vinci simulator

Skills analysis

- Trainee evaluation
- Automated feedback
- **Skills understanding**



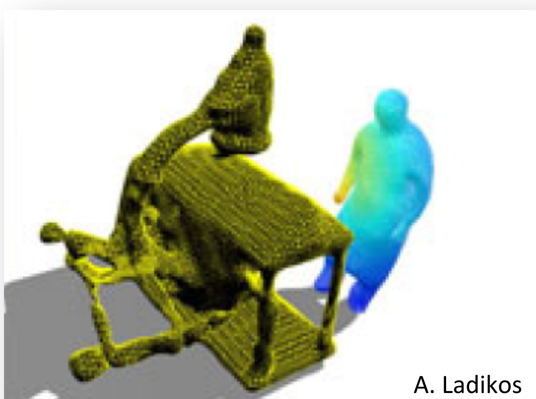
Robotic automation

- Motion replay
- Virtual fixtures
- **Shared control**

Motivation for Gesture Analysis (2)

Context recognition

- Context-aware UI
- OR synchronization
- Reports



Safety monitoring

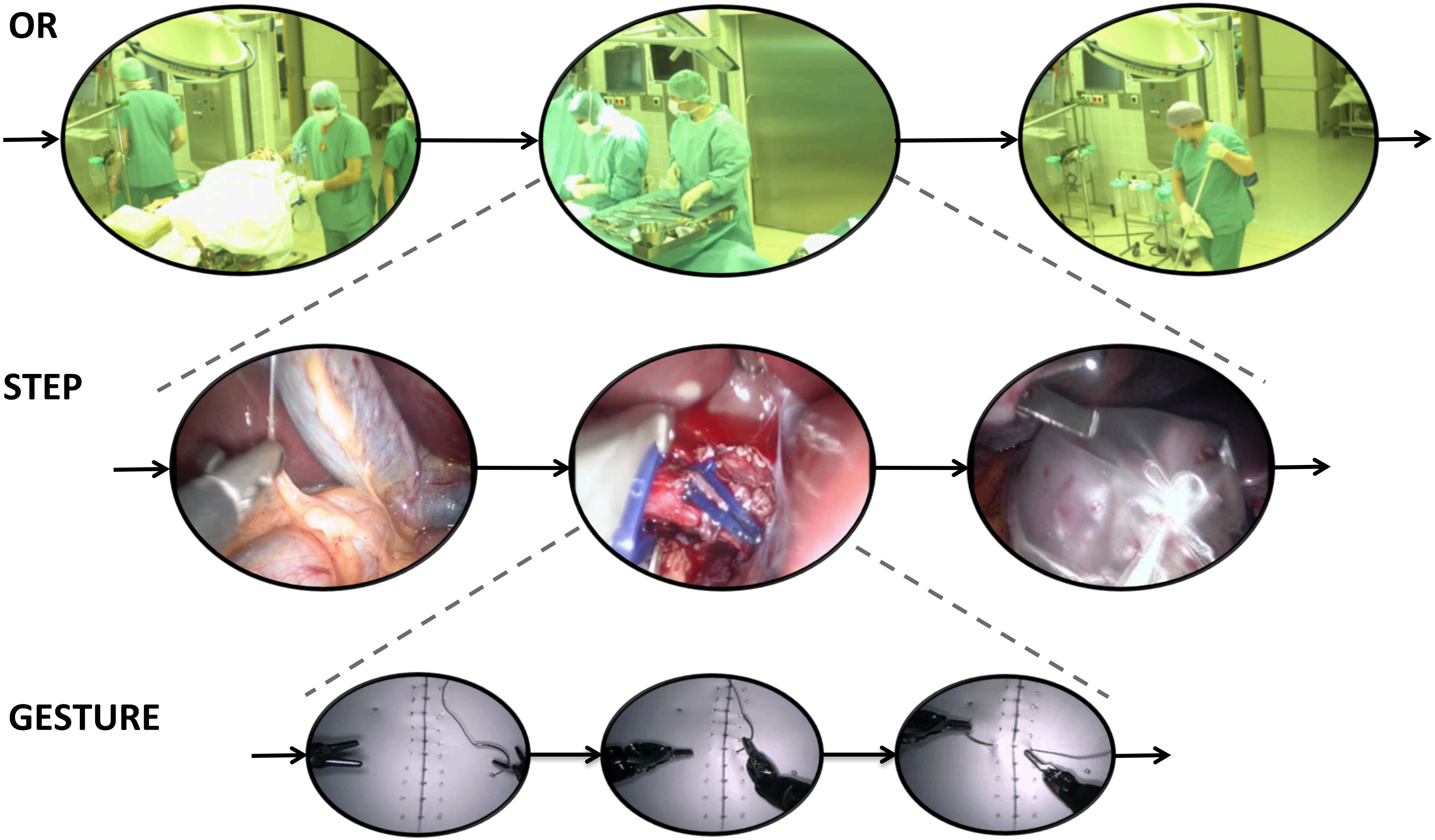
- Tool-tissue interactions (patient)
- Ergonomy analysis (staff)
- Radiation monitoring (staff)

Need for a language of surgical activities…

OR

STEP

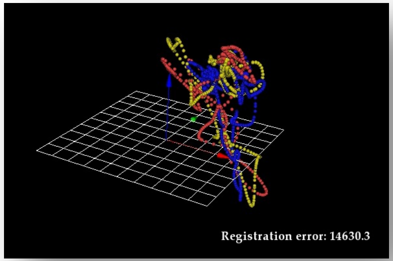
GESTURE



…and for an OR perception and recognition system

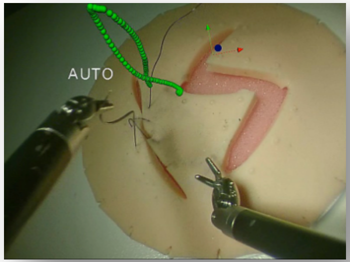
Outline

Gesture Modeling and Human-machine collaboration using the da Vinci robot



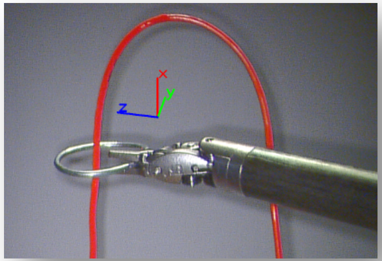
Surgical gesture modeling
(MICCAI'2011)

→



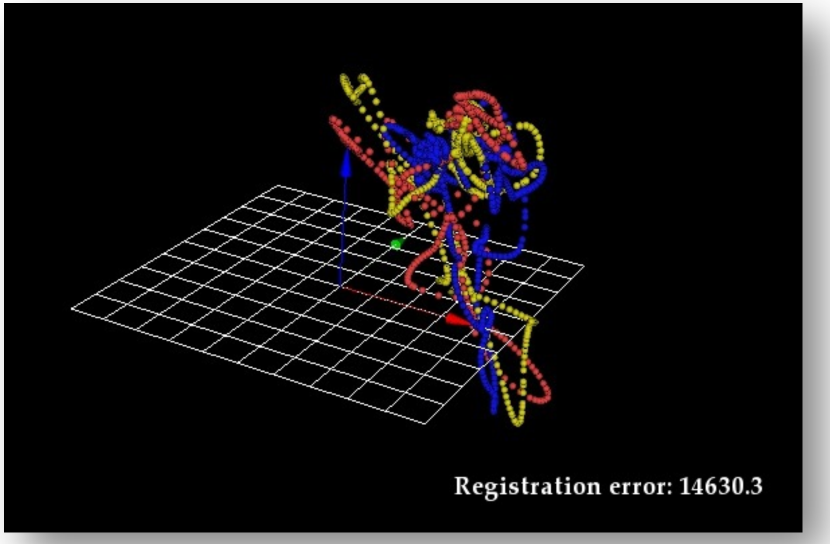
Shared Control
(ICRA'2011)

→

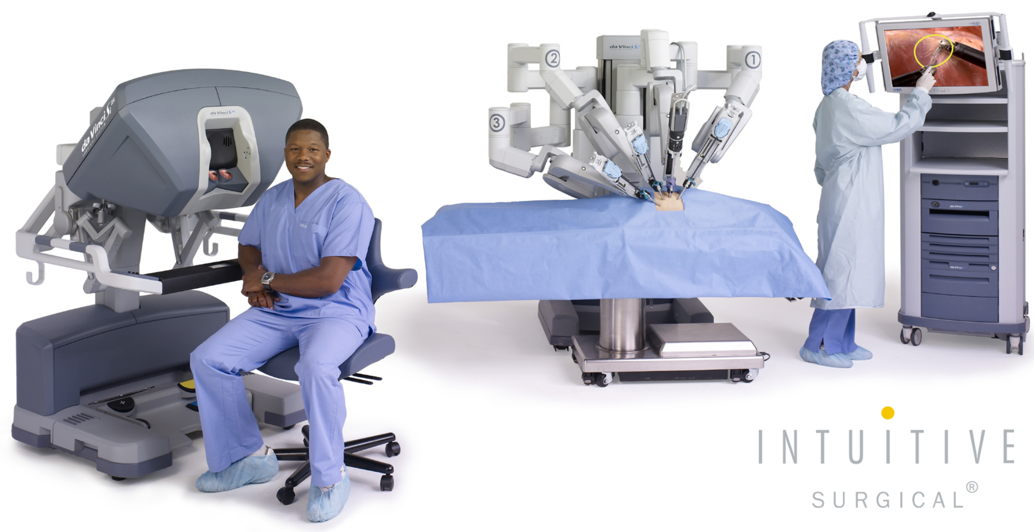


Active fixtures
(RLIHT'2012)

Gesture Modeling



The da Vinci Robot



Surgeon

↕

Master manipulators

↕

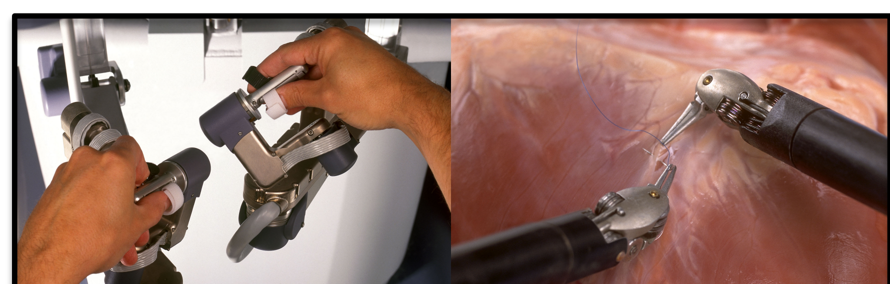
Patient-side manipulators

↕

Patient

Features

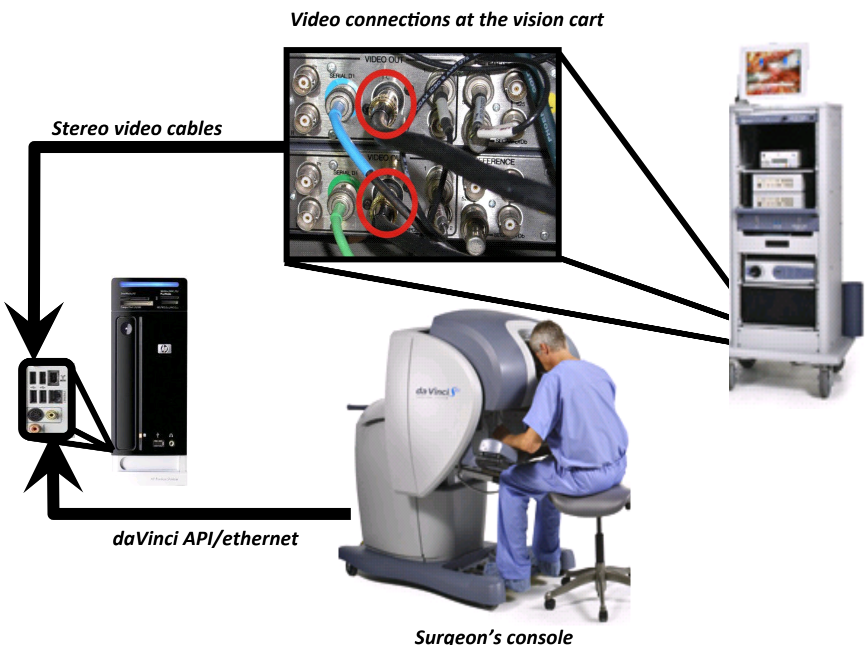
- 4 robotic arms
- HD stereo view
- 7 DOFs instruments
- Motion scaling
- Tremor filtering



Why is Robotic Surgery Well Suited?

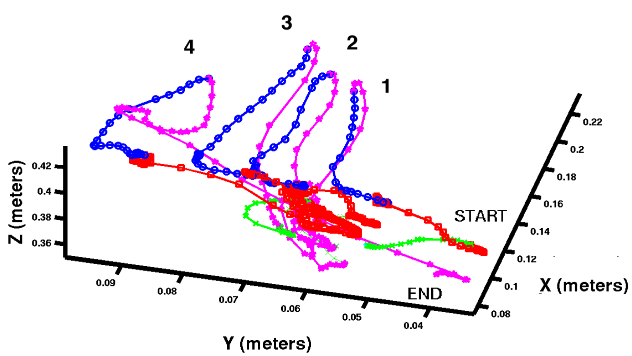
Mock-OR, JHU

Language of Surgery Project @ JHU

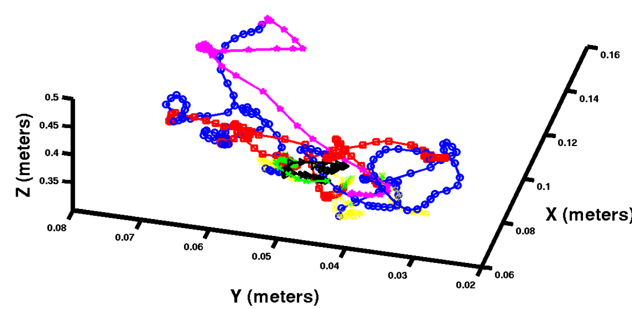


(Lin et al., MICCAI 2006)

- How can surgical skills be evaluated objectively ?

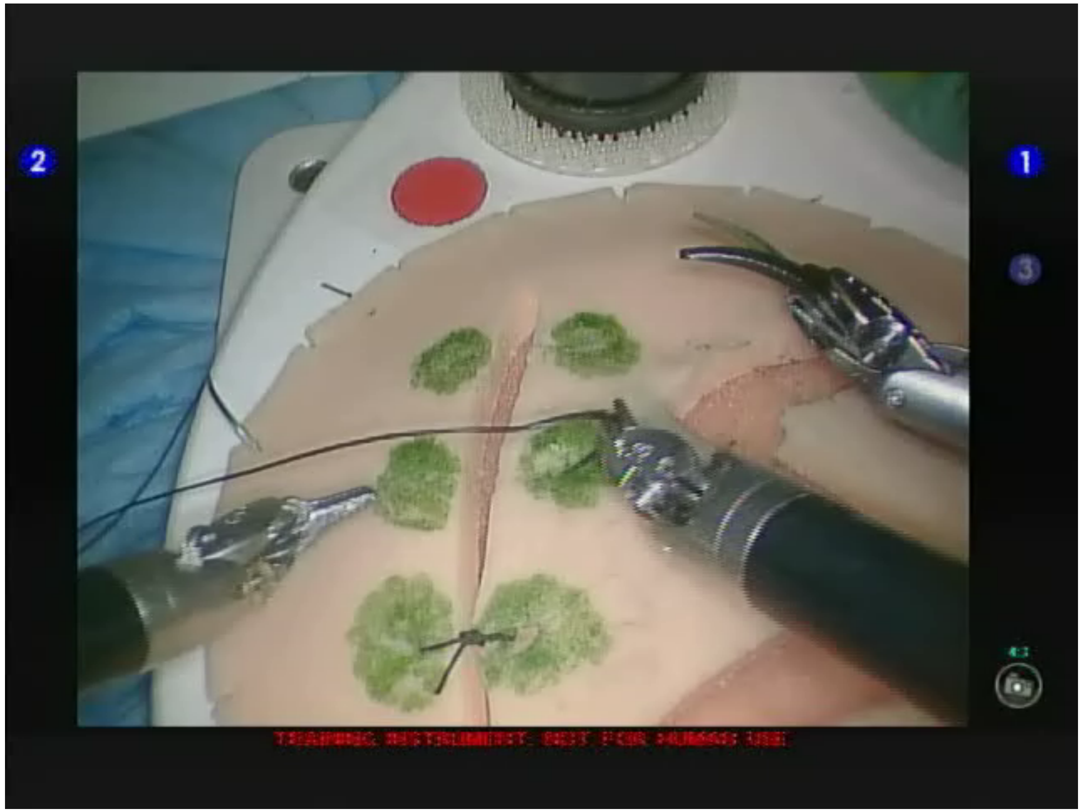


Expert Surgeon - trial 4



Intermediate Surgeon - trial 22

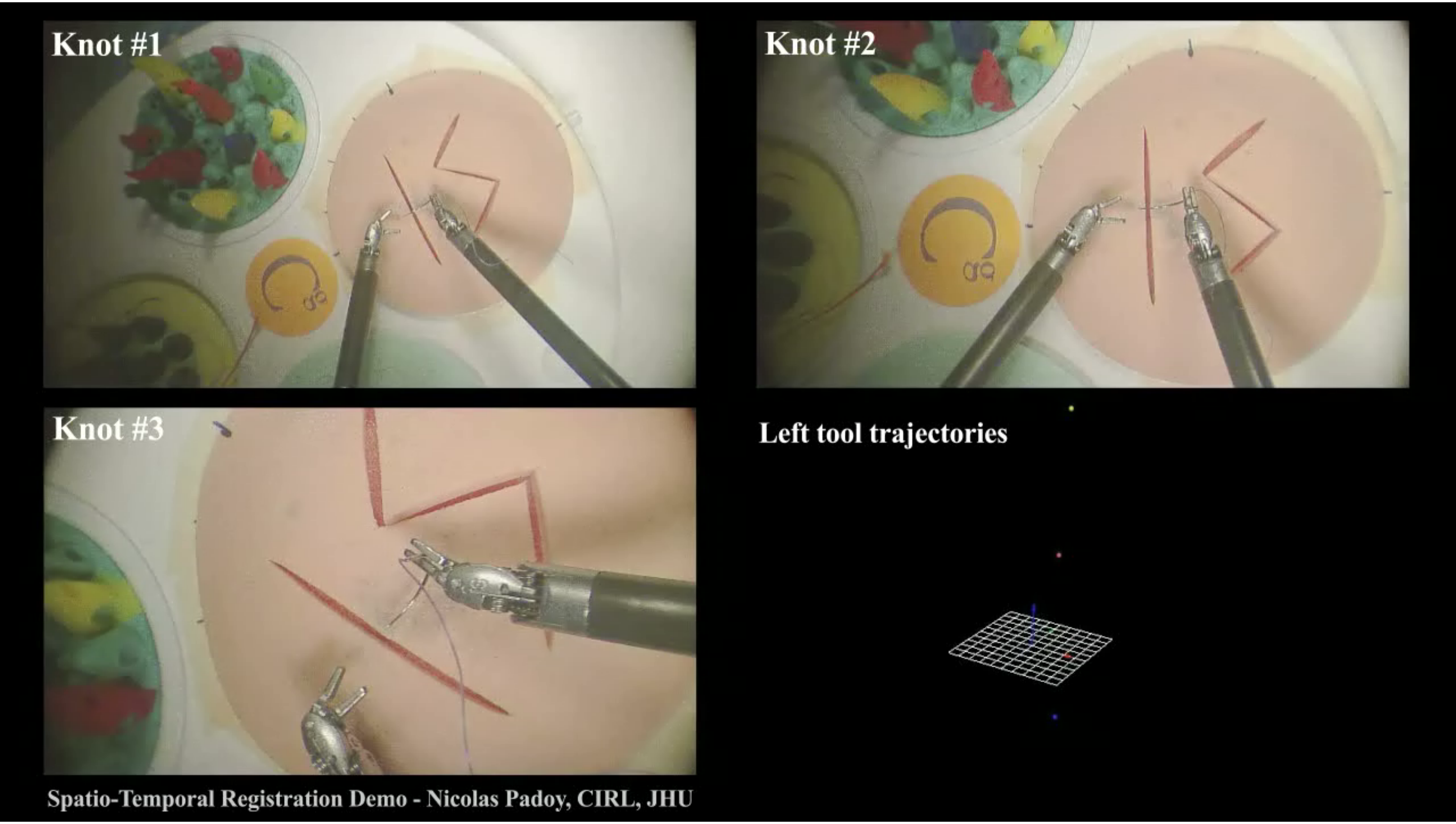
What about this dataset?



Kumar, Hager et al. (speed x5)

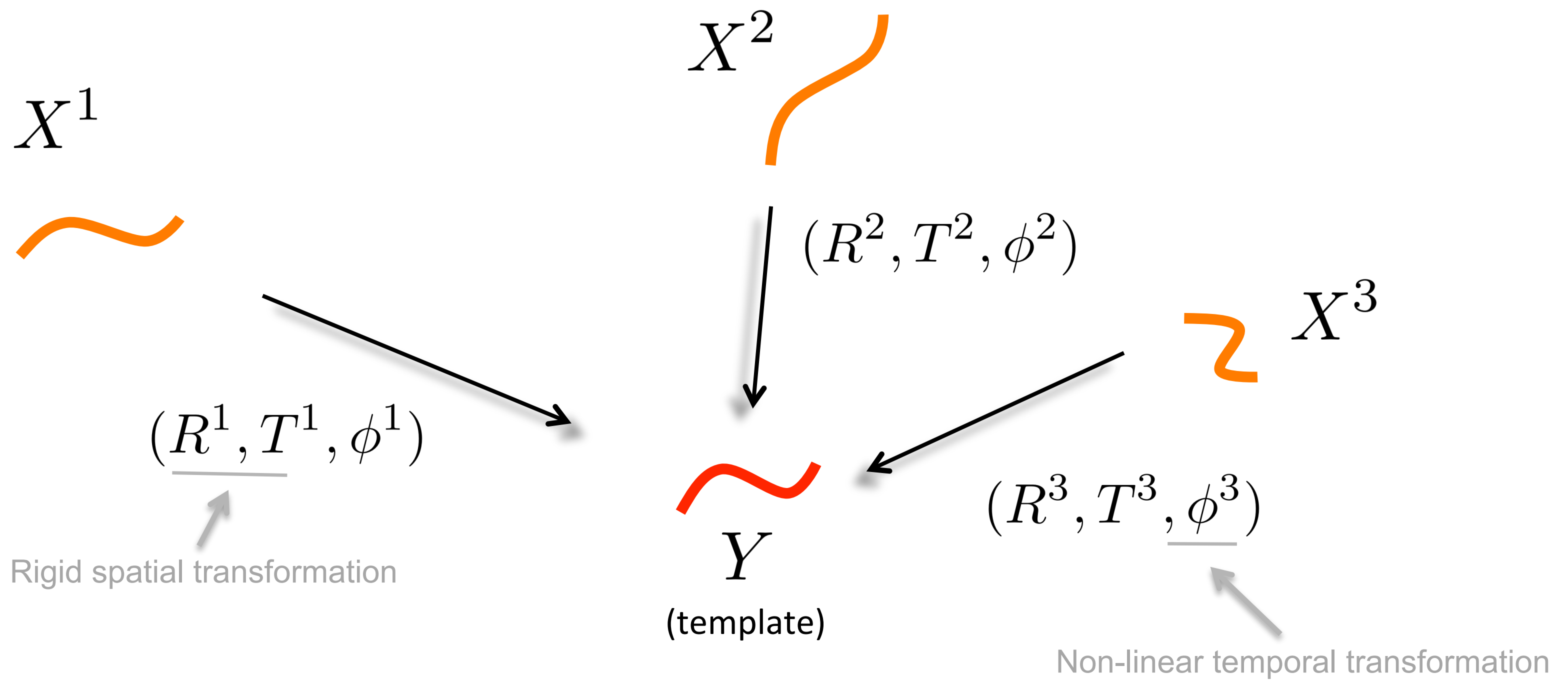
Longitudinal study of resident surgeons in several US hospitals over 4 training tasks

→ large data variability



(raw data: 3 of 6 sequences)

Spatio-temporal Gesture Registration



- Spatial overlap
- Non-linear temporal variations
- Local and global spatial variations

Spatio-temporal Time Warping (STW)

→ Combine DTW with ICP to obtain better point correspondences (STW)

$$C_{warping}(\phi) = \sum_{t=1}^{m} \| X_{\phi_t^x} - Y_{\phi_t^y} \|^2 \qquad \text{(DTW)}$$

trajectory temporal warping

$$C_{rigid}(R,T) = \sum_{l=1}^{p} \| R U_{\psi^u(l)} + T - V_{\psi^v(l)} \|^2 \qquad \text{(ICP)}$$

points correspondences

$$C_{stw}(R,T,\phi) = \sum_{t=1}^{m} \| R X_{\phi_t^x} + T - Y_{\phi_t^y} \|^2 \qquad \text{(STW)}$$

rotations translation

→ Solve using coordinate descent

For multiple sequences, construct a template Y

$$C_{multi}(R^k,T^k,\phi^k) = \sum_{k=1}^{K} \sum_{t=1}^{m_k} \| R^k X^k_{\phi_t^{k,x}} + T^k - Y_{\phi_t^{k,y}} \|^2 \qquad \text{(MDTW)}$$

Algorithm 2. Template generation

input : $\{X^k\}_{1 \leq k \leq K}$, initial template $\tilde{Y}$
output: new template Y and new warpings ϕ^k
for $1 \leq k \leq K$ **do**
    Compute warpings $\tilde{\phi}^k = (\tilde{\phi}^{k,x}, \tilde{\phi}^{k,y}, \tilde{m})$ between X^k and $\tilde{Y}$ using DTW;
Compute $\mu(t) = \sum_{k=1}^{K} \tilde{\phi}_t^{k,x}$, where $1 \leq t \leq \tilde{m}$;
Define $Y_t = \sum_{k=1}^{K} \tilde{Y}_{(\tilde{\phi}^{k,x}(\mu^{-1}(t)))}$, where $1 \leq t \leq \frac{1}{K}\sum_{k=1}^{K} n_k$;
for $1 \leq k \leq K$ **do**
    Compute warpings $\phi^k = (\phi^{k,x}, \phi^{k,y}, m)$ between X^k and Y using DTW;

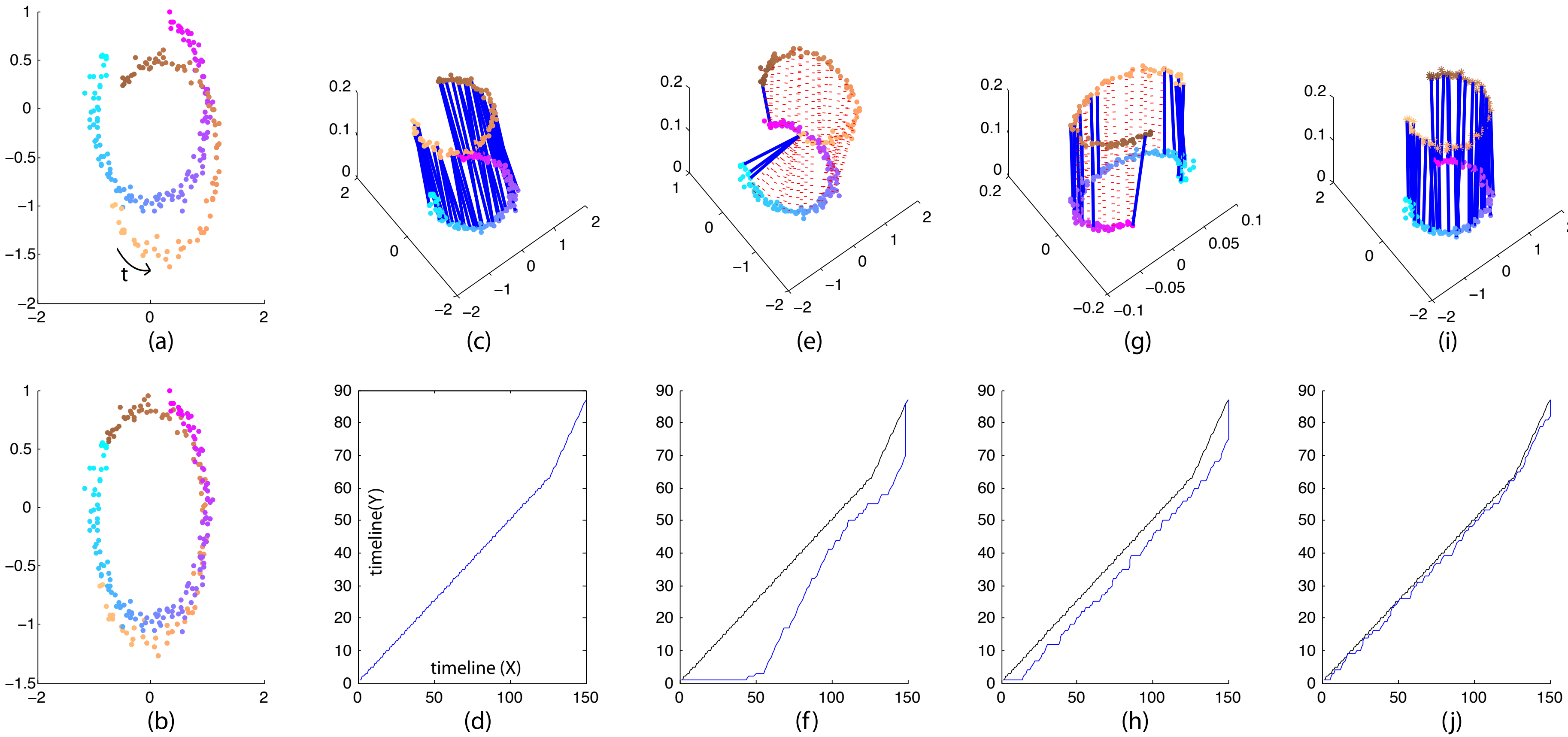
Algorithm 3. Multiple trajectories registration

input : $\{X^k\}_{1 \leq k \leq K}$
output: $\omega = \{R_k, T_k, \phi^k, Y \mid 1 \leq k \leq K\}$
Y is initialized as the X^k with median length;
for $1 \leq k \leq K$ **do**
    $R^k = \mathbf{Id}, T^k = \mathbf{0}$
repeat
    Update Y and $\{\phi^k\}$ using Algo. 2 with input $\{R^k X^k + T^k\}_{1 \leq k \leq K}$ and Y;
    for $1 \leq k \leq K$ **do**
        compute R^k, T^k using rigid registration and correspondence set $\{(X^k_{\phi_t^{k,x}}, Y_{\phi_t^{k,y}}) \mid 1 \leq t \leq m_k\}$;
until C_{multi} *converges* ;

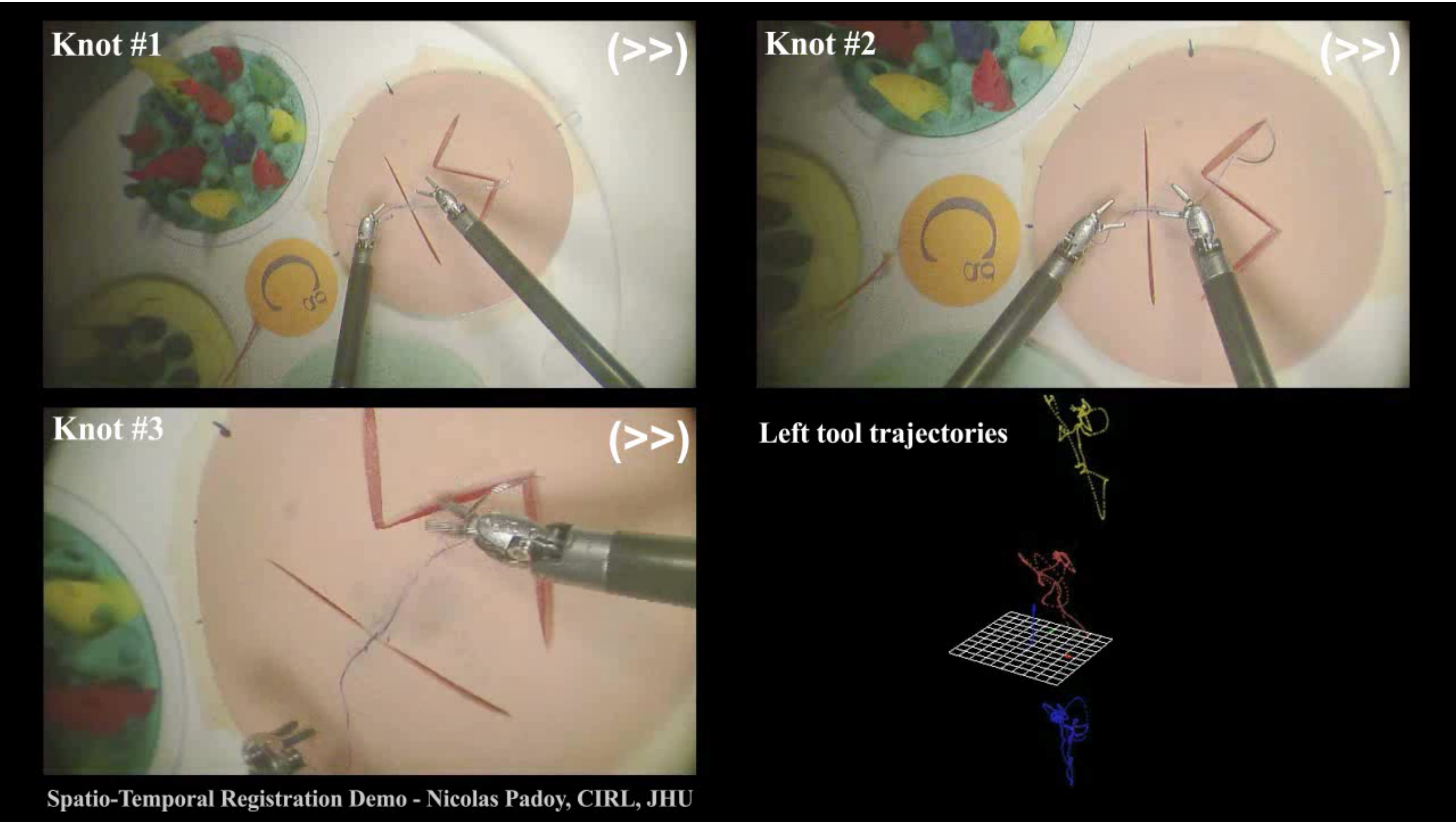
[Padoy, Hager, MICCAI'11]

STW - Synthetic Experiments

Two synthetic sequences. Time is color coded.

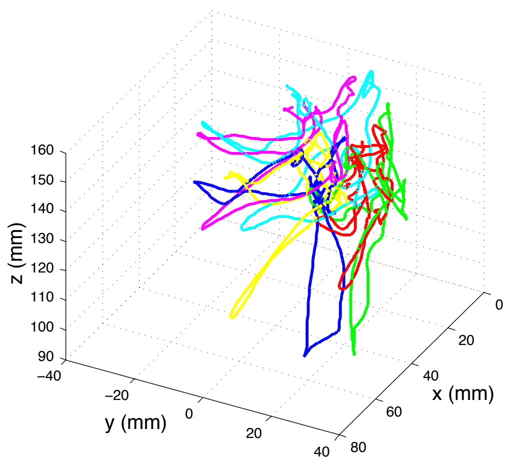


(ICP) (Ground truth) (DTW) (CTW) (STW)

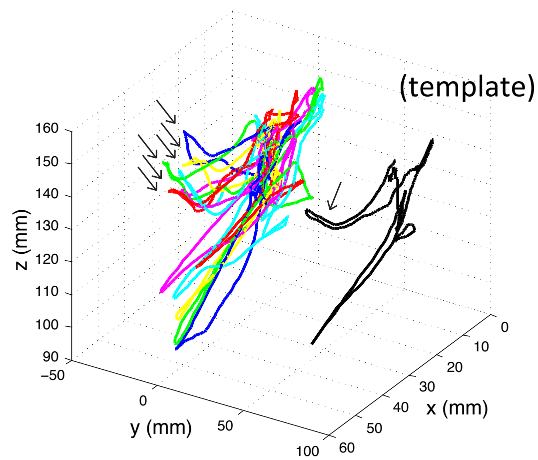


[Padoy, Hager, MICCAI'11]

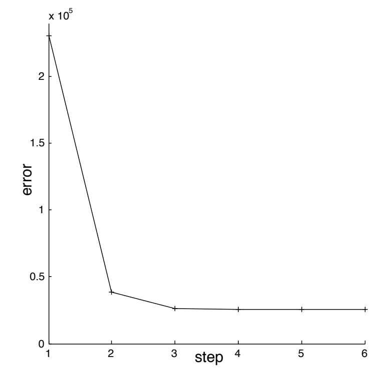
Results – Knot Tying Task



6 raw trajectories

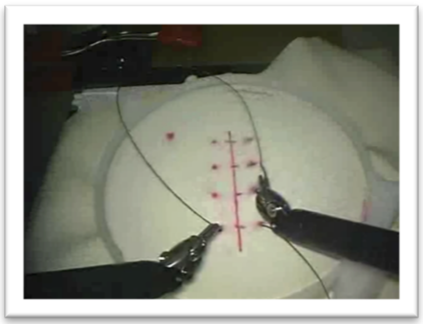


6 registered trajectories

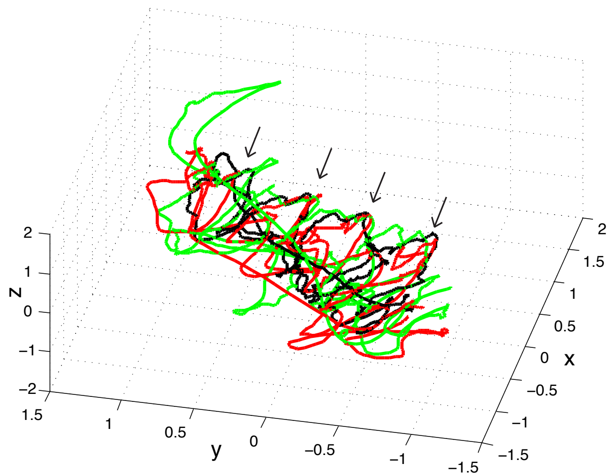


registration error

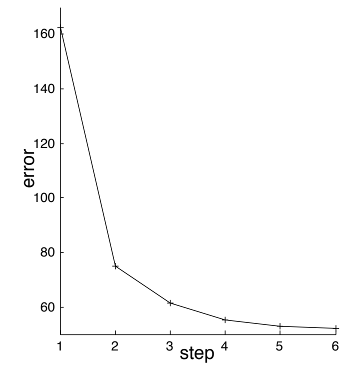
Results – 4-Throw Suturing Task



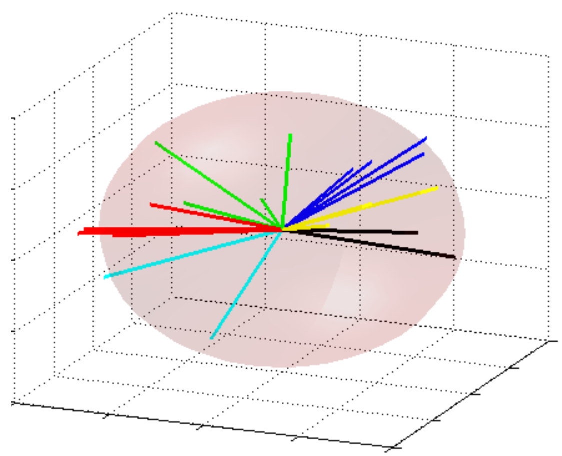
4-throw suturing tasks



3 of 19 registered trajectories

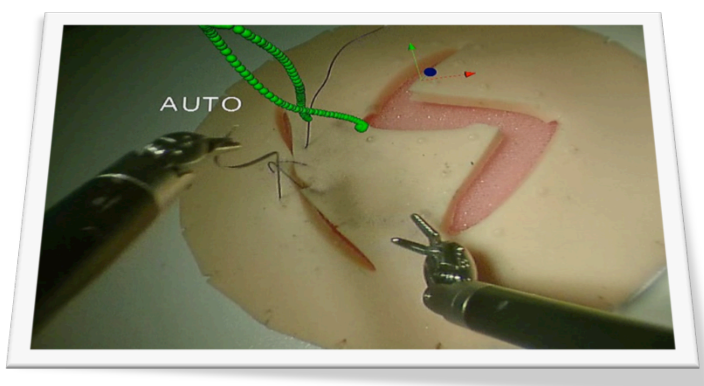


registration error



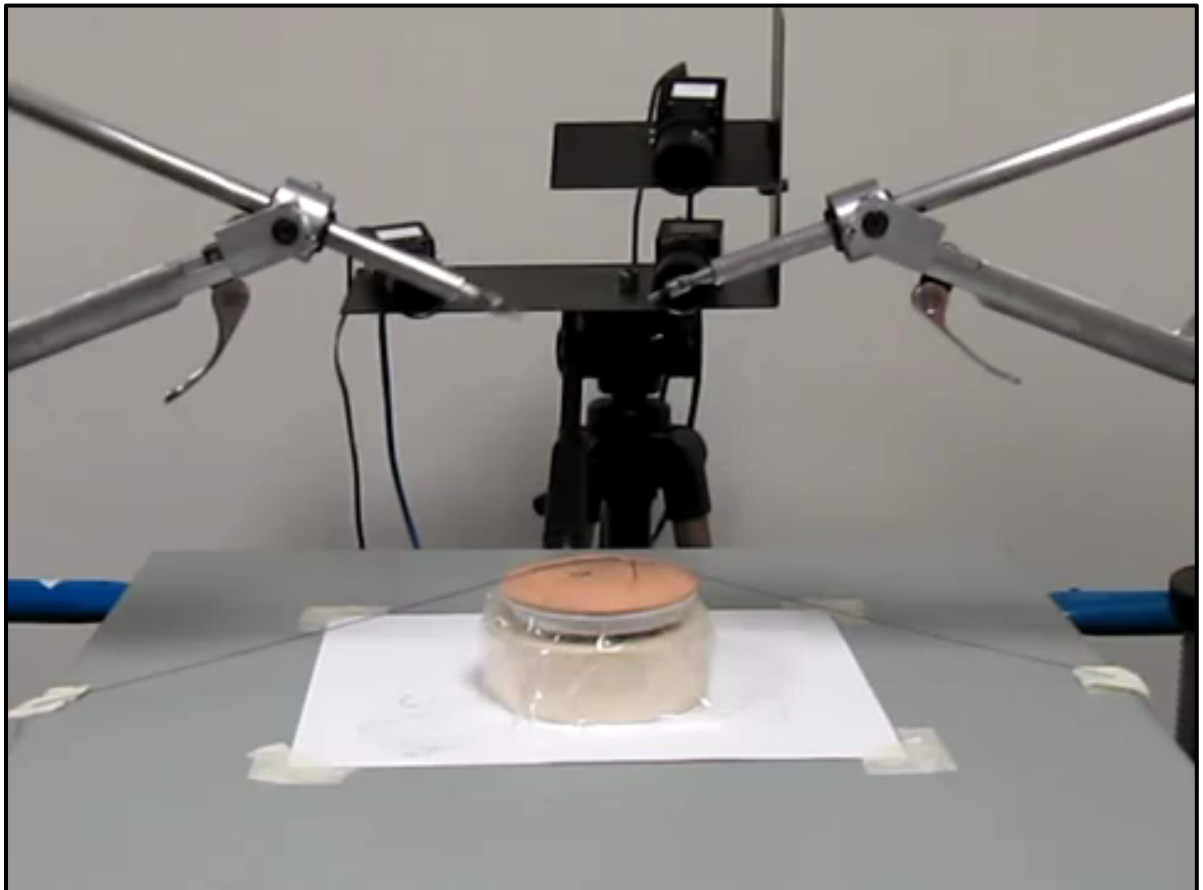
rotation axes colored per surgeon

Human-Machine Collaboration



Collaboration with Intuitive Surgical (Sunnyvale, CA, USA)

Beyond motion replay

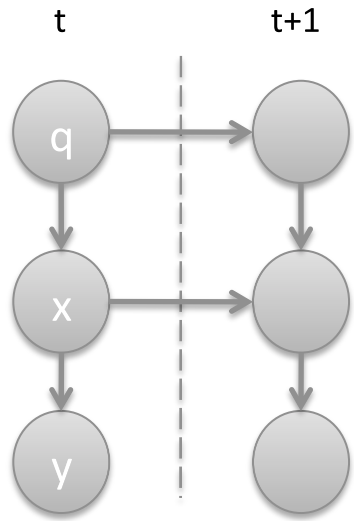


[Superhuman Knot Tying, van den Berg, ICRA'2010]

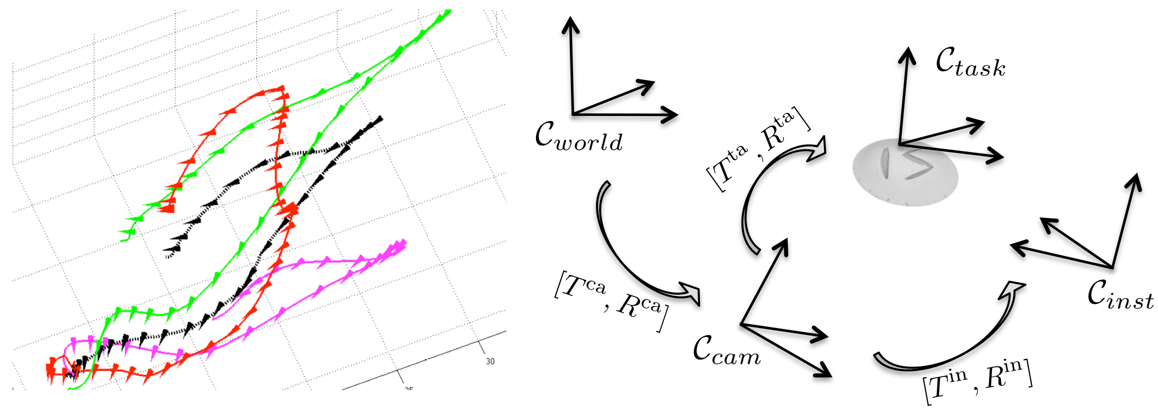
Human Machine Collaboration

Padoy, Hager, ICRA 2011 *(patent pending)*

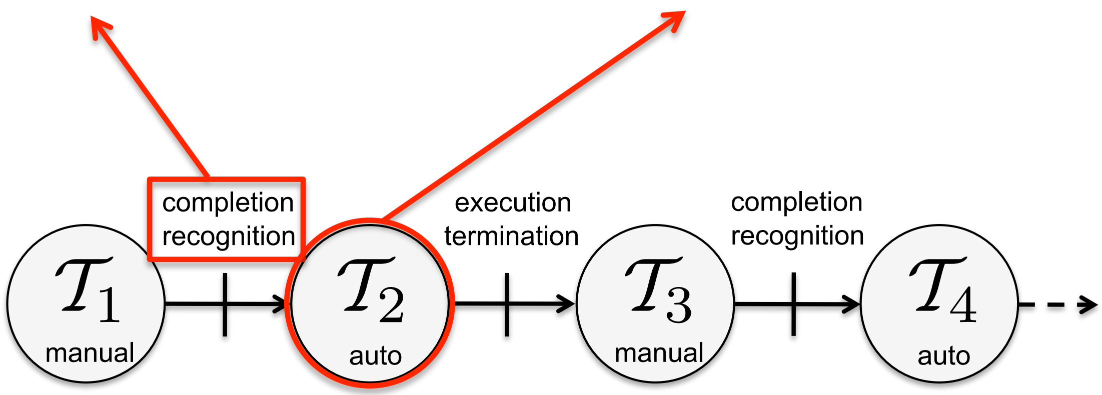
Methods

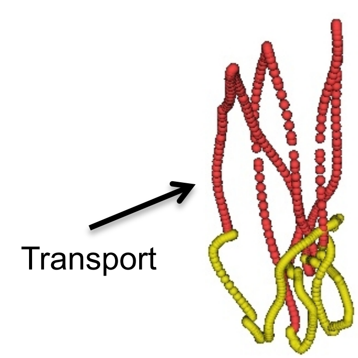


H-HMM



Temporal Motion Averaging + Robot Control

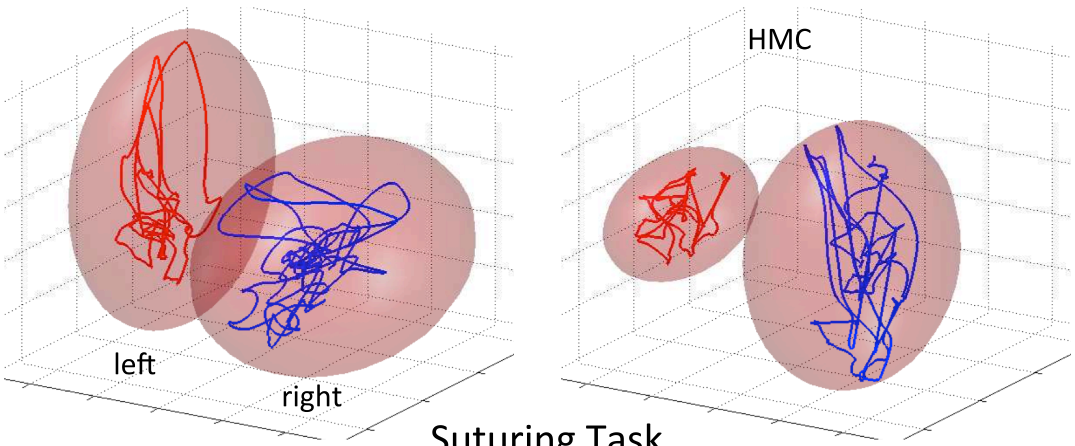




Left tool trajectory

Results

- Reduced master manipulators motions



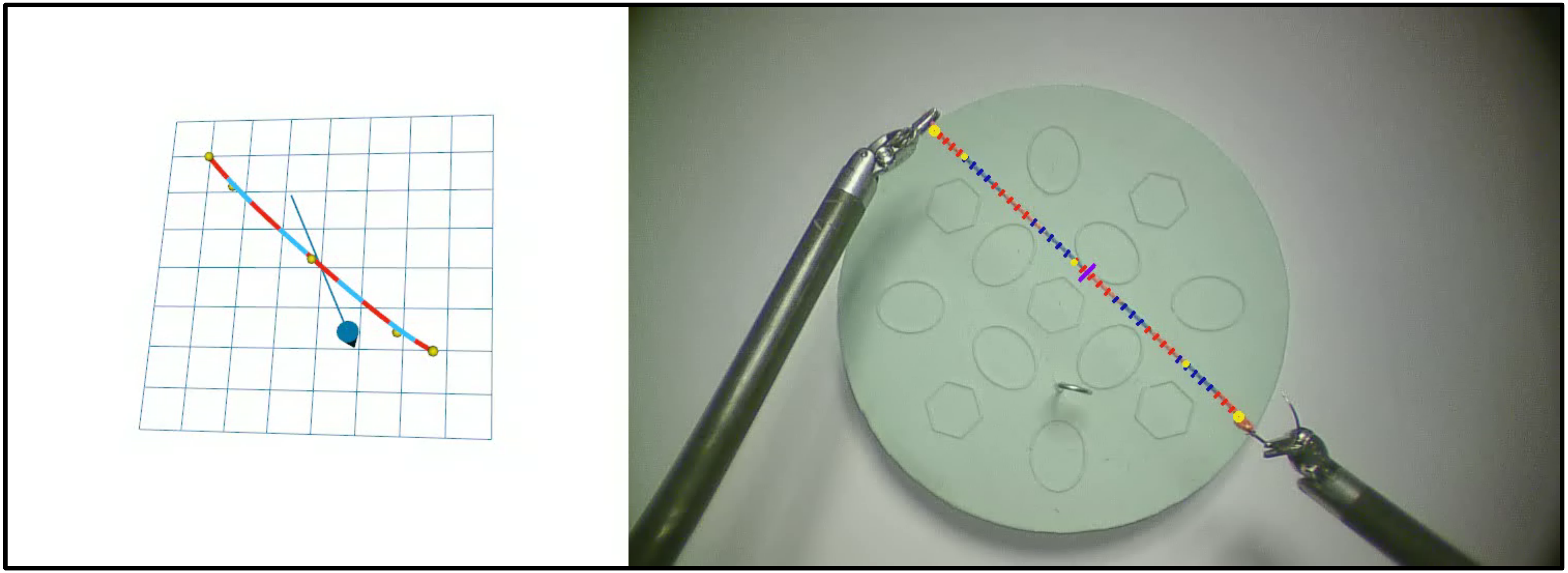
Suturing Task

- Shorter traveled distances

- Smaller standard deviations:

	Pin Task			Sut Task		
	x	y	z	x	y	z
Manual	55.4	50.4	22.1	17.7	21.4	38.3
HMC	10.6	19.9	23.1	12.1	16.7	23.2

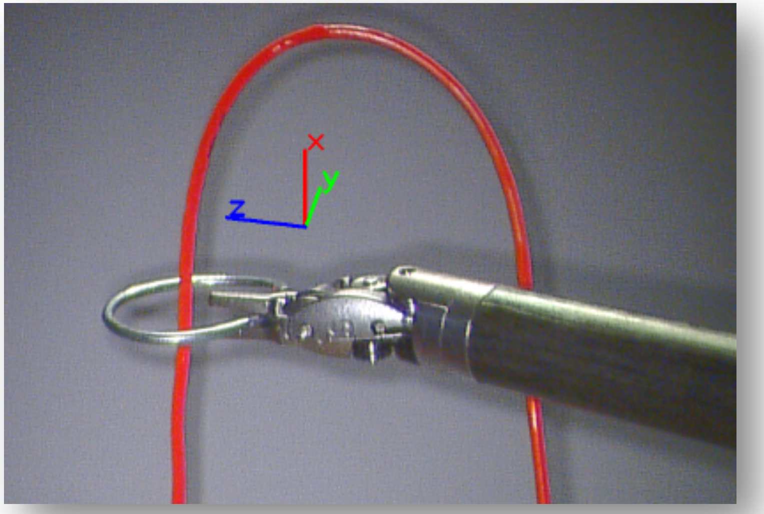
Need for Physical Context Tracking



Padoy, Hager, BMVC 2012

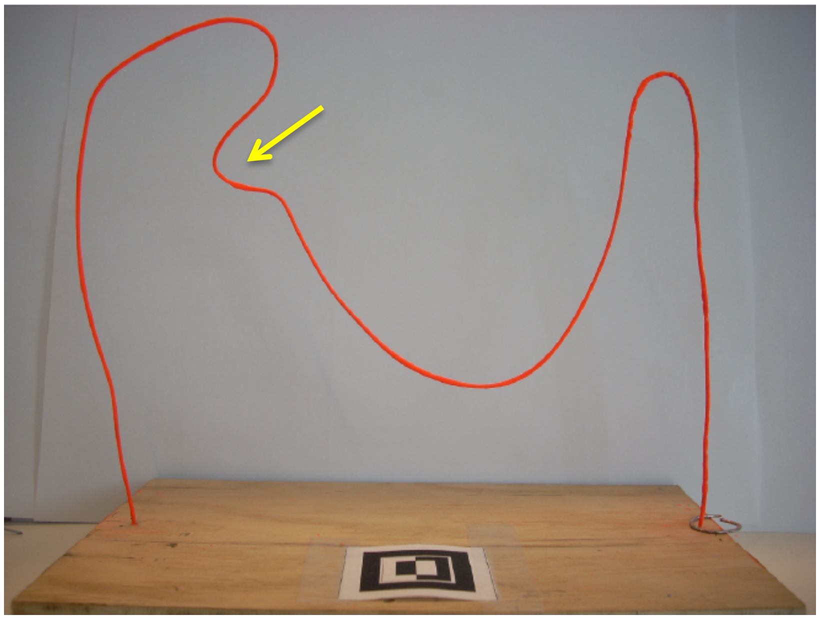
(patent pending)

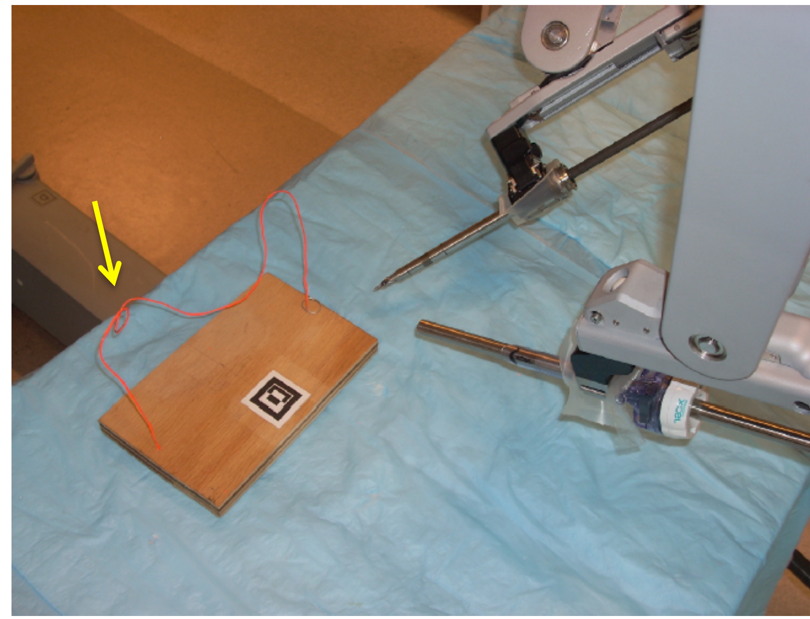
Learning-based Partial Automation (does ergonomy matter?)



(In collaboration with Sebastian Bodenstedt, KIT)

Ring transfer task



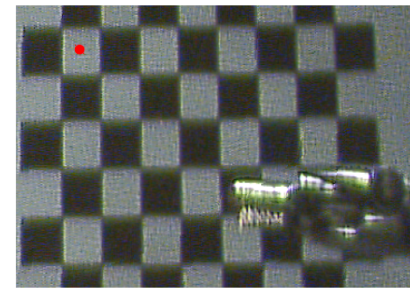


Partial automation of a dexterous task:

- Control position of the instrument through the user
- Control orientation of the tool through the robot

Model of the environment

- Stereo endoscope calibration
  - 3D reconstruction error: 1.58mm
  - Rectification error: 2.43px
- Instrument position calibration
  - Correction with reconstructed chessboard pattern

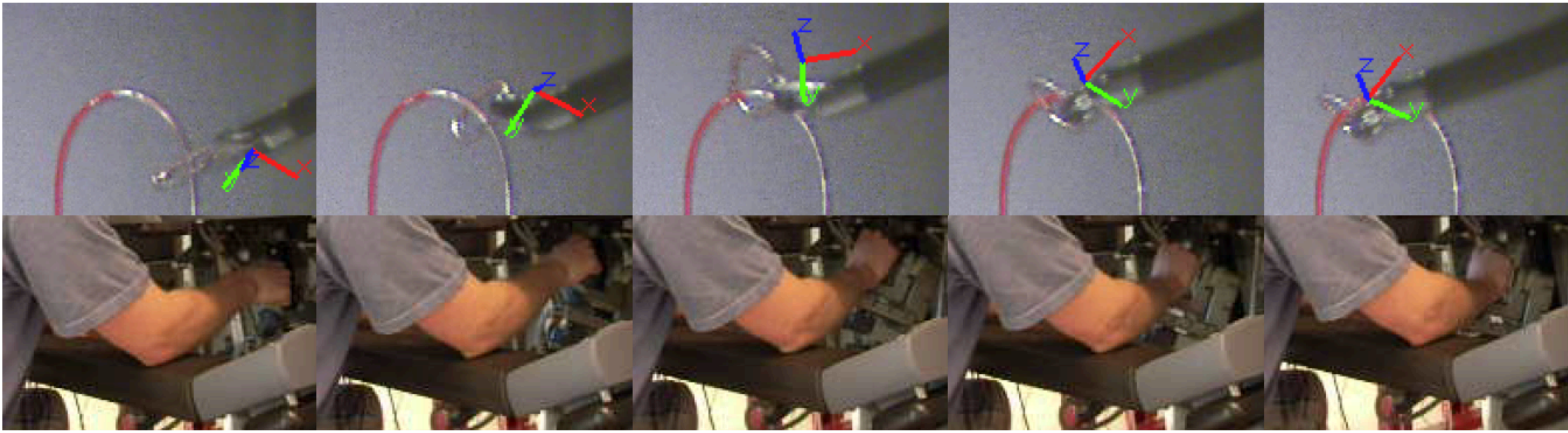


- Model of the environment

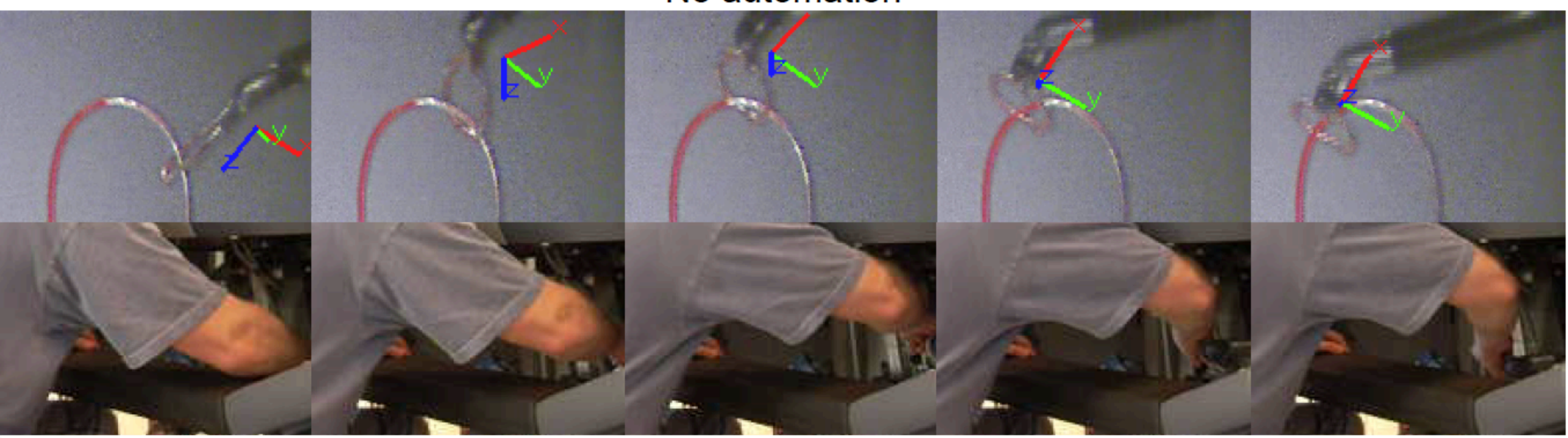
preprocessing + semi-global block matching

spline approximation

Comparison no automation / automated rotation



No automation

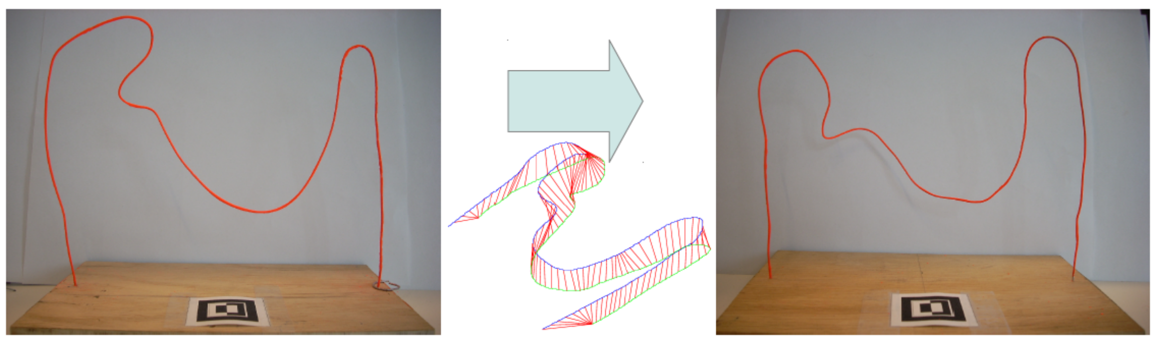


Automated rotation

Solution

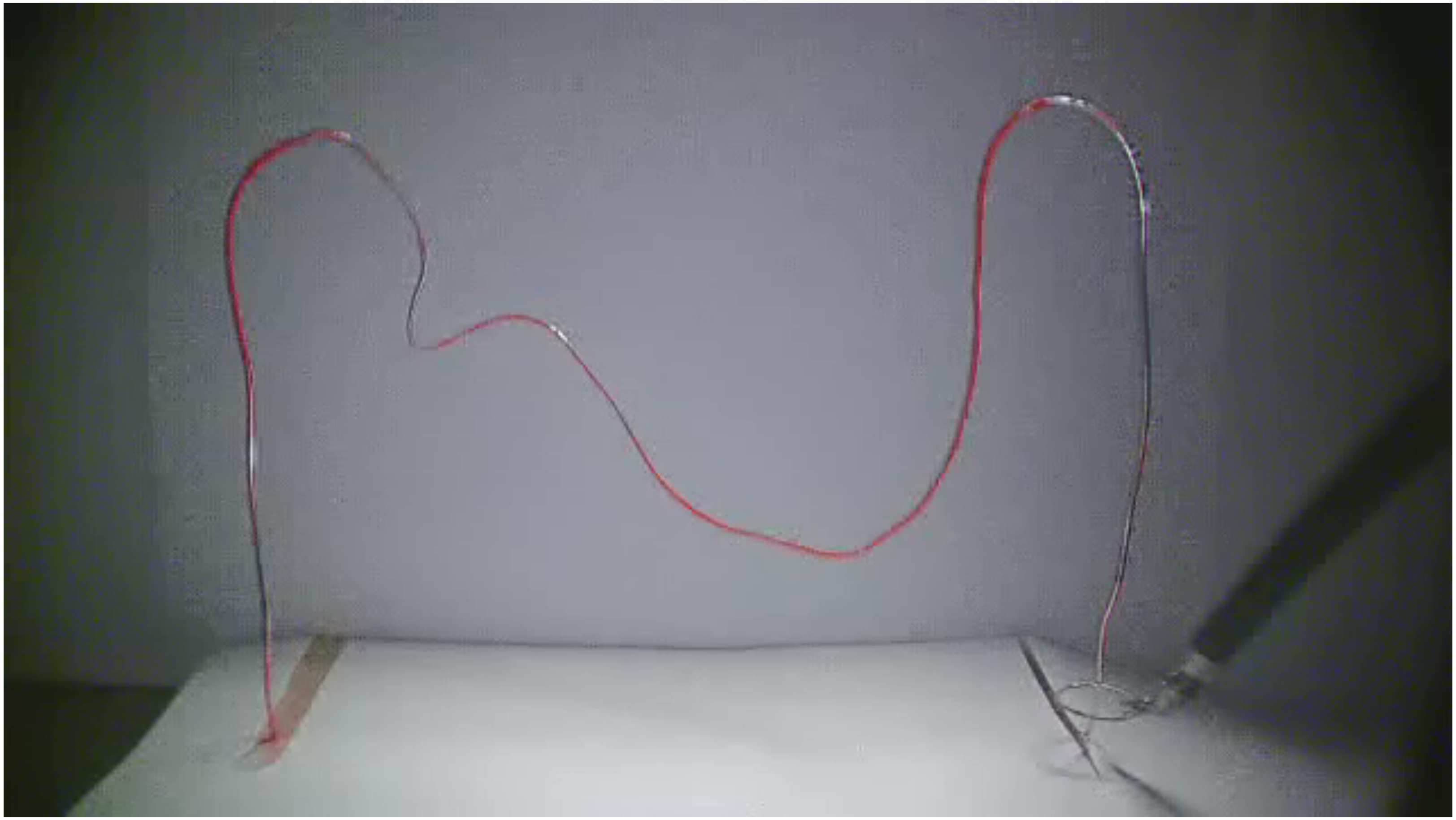
- Use Gaussian mixture regression (GMR) to learn 'ergonomy' from expert demonstration on a reference task
  - Learn conditional PDF P(R|t)

- Map current task to reference task using DTW



- Combine GMR, current position and 3D spline to infer automated rotation

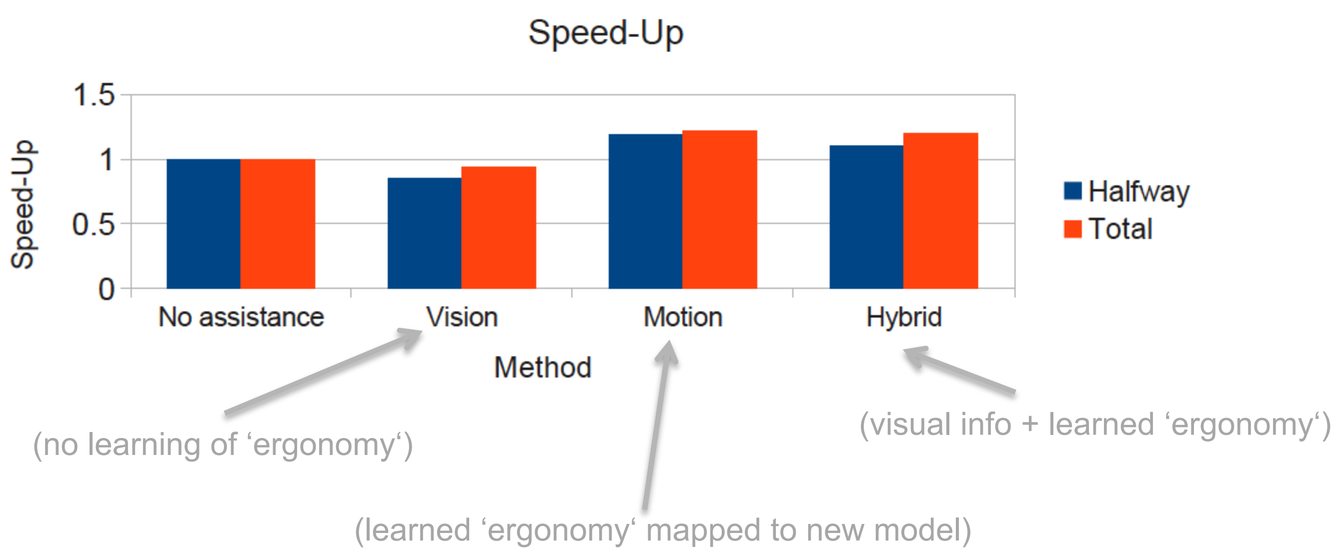
Video



Bodenstedt, Padoy, Hager, RLIHT 2012

User Study

- 10 participants
- 4 tasks per user (randomized order)



- Larger speed-up on beginners
- Hybrid method ranked best, vision method ranked worst

Conclusion

→ Trajectories are not sufficient

→ Models of tool/tissue and operator/tool interactions needed

MIX-Surg

Thank you for your attention !

Spatio-temporal Registration of Multiple Gestures for Skills Analysis and Automation in Robotic Surgery

Nicolas Padoy

University of Strasbourg, France

LABORATORY FOR
Computational Sensing + Robotics
THE JOHNS HOPKINS UNIVERSITY

March 13th, 2013

iCUBE