Stabilized Ultrasound Imaging of a Moving Object Using 2D B-mode Images and a Convolutional Neural Network

Tian Xie, Mahya Shahbazi, Yixuan Wu, Russell H. Taylor, Emad Boctor

\textsuperscript{a} Laboratory for Computational Sensing and Robotics, Johns Hopkins University

\textsuperscript{b} Department of Computer Science, Johns Hopkins University

\textsuperscript{c} Russell H. Morgan Department of Radiology and Radiological Science, Johns Hopkins Medical Institutes

Co-robotic ultrasound + tracking algorithm  \rightarrow  steady imaging

Images are subject to physiological motions.  
Stabilized imaging with the tracking algorithm.
<table>
<thead>
<tr>
<th>Outline</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Challenge and Motivation</strong></td>
<td>In a biopsy, compensate physiological motions and reduce musculoskeletal trauma for sonographers using co-robotic ultrasound</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td>6-DoF motion estimation. A convolutional neural network is embedded to estimate out-of-plane motions. Joint velocities control.</td>
</tr>
<tr>
<td><strong>Evaluation and Results</strong></td>
<td>Feasibility of the tracking algorithm within simulation environments built with scans obtained on biopsy phantoms, pork, etc.</td>
</tr>
<tr>
<td><strong>Discussion and Conclusions</strong></td>
<td>Advantages. Limitations.</td>
</tr>
</tbody>
</table>
Challenge: A Biopsy Procedure

- An ultrasound-guided biopsy

1. Local anesthetic
2. Localize the mass
3. Insert the needle
4. Extract a tissue sample
5. Insert the needle
6. Extract a tissue sample
7. Repeat until diagnosable samples are acquired

Ultrasound-guided liver biopsy on a 47-year-old female patient1.

1 Courtesy of https://www.youtube.com/watch?v=2Si2Q0qJU4

Challenge: Physiological Motions

- Imaging of abdominal organs is subject to physiological motions:
  - Respiratory motion
  - Cardiac-induced motion
  - Patient’s small movement

- Amplitudes of motions:
  - E.g. Shimizu et al.3 investigated liver tumor motion using high-speed MRI
  - 21 mm, 8 mm, 9mm in the SI, AP and lateral directions.

Ultrasound sequence of liver undergoing respiratory and cardiac-induced motion2.

2 Courtesy of https://www.youtube.com/watch?v=KkDTNvClII4

Challenge: Work-related Musculoskeletal Disorders (WRMSD)

Main causes of WRMSD\(^5\)

An accumulation of repeated exposure to physical risk factors:

- Awkward or sustained postures
- Repetition
- Force
- Contact pressures

2,963 participants in Evans et al.\(^4\)

SONOGRAPHERS WITH WRMSD

\(\begin{array}{c|c|c}
\text{WRMSD} & \text{No pain} \\
\hline
10\% & 90\% \\
\hline
\end{array}\)

- 15\% in 6 months
- 45\% in 3 years
- 72\% in 10 years


Motivation

- Robotic arm:
  - force control, accuracy, repeatability
- Co-robotic ultrasound

In a biopsy procedure, physiological motions of the target is another problem to solve.

"Hand-over-hand control"

Respiratory motion
Proposed Method: System Setup

Co-robotic setup

Tracking algorithm for stable imaging

• In-plane motion: $\Delta q_{pr}(t) = q_{pr}(0) \Delta q_{pr}^1(t)$
• Out-of-plane motion: $\Delta q_{tar}(t)$
• 6 degree of freedom (DoF) motion: $\Delta q_{pr} = q_{pr}(0) \left(q_{pr}(t) q_{se(3)}(t)^{-1}\right)$
• Proportional control on SE(3) error

Proposed Method: Significance

A co-robotic ultrasound system with a tracking algorithm

• compensates physiological motions:
  • keep track of the target frame
  • suppress motion artifact and enhance image quality

• reduces musculoskeletal injuries for sonographers:
  • use a robot for long-lasting procedures
  • integrate with force control

• requires only B-mode images:
  • apply to almost all commercial ultrasound machines

Have a robotic system to track the target during the biopsy

Sphere: moving anatomy
Serial link: UR5

Without the tracking algorithm.
Overview: Tracking Algorithm

**CALIBRATION**

*Step 1:* acquire parallel B-scans across the tissue within a small volume.
*Step 2:* fine tune the CNN weights with a small learning rate.

**Overview**

1. Acquire target frame $I_0$, and a reference image $I_{ref}$ along the (+) direction
2. Compute the transformation $g^*$ using template matching via NCC when no large rotation.
3. Compute in-plane transformation and out-of-plane transformation for each patch using CNN.
4. Calculate the overall out-of-plane transformation by minimizing the reconstruction error using all patches.
5. Joint velocity control on SE(3).
6. Current Image $I_c$

**Controller**

- $g(t)$
- $g^*$
**Estimation of Out-of-plane Motion**

(1) Use a CNN to estimate the unsigned distance between two patches

(2) Get the sign using both the target and reference frames

(3) Fit the overall out-of-plane transformation using all patches

\[ g_{ar} = \text{argmin} \sum |\tilde{p}_{r,i} - \tilde{p}_{a,i}| \]

**Speckle Decorrelation**

**Speckle**
- Granular appearance in ultrasound images
- Diffuse scattering in a resolution cell

**Speckle decorrelation**
- Imperfect focus along the elevational direction \( \rightarrow \) neighboring patches are correlated
- Gaussian using RF signals and assuming Gaussian resolution cell

**Conventional speckle decorrelation methods**:  
- RF signals may not be accessible  
- Hard to convert back to RF signals  
- Rely on fully developed speckles

**Our method only requires grayscale B-mode images**:  
- High non-linearity of the CNN  
- “Correlation”

---

CNN to Estimate Elevational Distance

Prevost et al. 7 3D freehand ultrasound estimation using deep learning

- Is used for tomography
- Depends on anatomical features of the specific part of the body (e.g., forearm)

![Frame-to-Frame Motion Estimation](image1)

In a tracking scenario, the target appearance varies. A model invariant to the appearance of the masses.

Training data containing only speckle noises without anatomical features (e.g., vessels).

---

CNN Architecture

- A standard convolutional neural network
- # of parameters: ~ 2.7 million
- Real-time tracking. Less than 50 ms on a PC with an NVIDIA GTX 1050 Ti GPU (use 16 pairs of patches)
- Input: two neighboring B-mode image patches
- Output: elevational distance between two patches

---

**Estimation of Out-of-plane Motion**

(1) Use a CNN to estimate the unsigned distance between two patches

(2) Get the sign using both the target and reference frames

(3) Fit the overall out-of-plane transformation using all patches

\[ g_{ar} = \text{argmin} \sum ||g_{ar} \vec{p}_{r,i} - \vec{p}_{a,i}|| \]
Velocity Control on SE(3) Group

- Implement a proportional (P) controller with double-geodesic feedback on SE(3)\(^8\)

\[
\begin{align*}
\dot{g} & = g_{PP} \\
\dot{v}_{pp}^b & = -K_{wo} \log_{SO(3)} R_{pp} \\
\dot{v}_{pp}^b & = -K_{pp} v_{pp}^b
\end{align*}
\]

Body velocity \(\rightarrow\) joint velocities

Data and Experiments

CNN Training Data
- **Patch size:** 60 x 100 pixels
- **Patches contain only speckle noises**
- Elevational displacement (ground truth) ranges from **0 to 1 mm** (close to the size of a resolution cell)

- **Sonix-Touch Q+ Ultrasound machine** (Ultrasonix Inc., Richmond, BC, CA)
- **Probe:** BK L14-5/38 linear, 10 MHz
- **Phantoms:** CIRS Elasticity QA phantom model 049 and 049A
- **Ground truth:** linear stage + a dial indicator (0.001mm)
- **12,000 pairs:** 80% for training, 20% for validation
- **2,000 patches** collected on the other phantom for testing
**Data and Experiments**

**Data**
- Sonix-Touch Q+ Ultrasound machine
- Probes:
  - (1) BK L14-5/38 linear; (2) C5-20/60 curvilinear
- Phantoms:
  - (a) CAE Blue Phantom tissue biopsy ultrasound training model
  - (b) Pork tenderloin

**Motion Simulation**
- Construct a volume using parallel scans
- Simulate cyclic motions
- Implement P control to track the moving target

\[
x(t) = \frac{A}{2} - Acos^2\left(\frac{\pi t}{T} - \phi\right)
\]
\[
\alpha(t) = \theta \sin^2\left(\frac{\pi t}{T} - \phi\right)
\]

**Outline**

- **Challenge and Motivation**
  In a biopsy, compensate physiological motions and reduce musculoskeletal trauma for sonographers using co-robotic ultrasound

- **Methods**
  6-DoF motion estimation. A convolutional neural network is embedded to estimate out-of-plane motions. Joint velocities control.

- **Evaluation and Results**
  Feasibility of the tracking algorithm within simulation environments built with scans obtained on biopsy phantoms, pork, etc.

- **Discussion and Conclusions**
  Advantages. Limitations.
Evaluation and Results: CNN Estimation

- 10 parallel scans with 0.1 mm separation
- 756 patches in one image (slide by 20 pixels along x, y axes)
- Accurate estimations in range 0.2 to 0.7 mm
- Poorer estimations near 0 or 1 mm (near the size of the resolution cell)

<table>
<thead>
<tr>
<th>Ground truth/mm</th>
<th>Mean abs %</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>10.56</td>
</tr>
<tr>
<td>0.2</td>
<td>1.62</td>
</tr>
<tr>
<td>0.3</td>
<td>0.49</td>
</tr>
<tr>
<td>0.4</td>
<td>0.29</td>
</tr>
<tr>
<td>0.5</td>
<td>0.95</td>
</tr>
<tr>
<td>0.6</td>
<td>2.23</td>
</tr>
<tr>
<td>0.7</td>
<td>0.25</td>
</tr>
<tr>
<td>0.8</td>
<td>5.4</td>
</tr>
<tr>
<td>0.9</td>
<td>9.99</td>
</tr>
</tbody>
</table>

- Label each pixel with the absolute error computed at the nearest patch center
- Dark regions correspond to non-fully developed speckles (*bad* regions for speckle decorrelation)

Evaluation and Results: Biopsy Phantom

- **Target motion**:
  - 4, 11 and 6 mm along axial, lateral and elevational axes
  - 10 and 6 degree about x and y axes (w/o in-plane rotation)
- **Error**:
  - Translations: less than 0.7 mm
  - Rotations: less than 2 degree
- No fine tuning. Same linear probe used for CNN training.
Evaluation and Results: Biopsy Phantom

- **Target motion:**
  - 4, 11 and 6 mm along axial, lateral and elevational axes
  - Less than 10 degree about x and y axes (w/o in-plane rotation)
  - Different random phase angles
- **Low errors compared with the magnitudes of the motions.**

Evaluation and Results: Pork Tenderloin

- **Target motion:**
  - 1-DoF elevational translation (amplitude 10 mm)
- **Error:**
  - Less than 1.05 mm
  - Fine tuning using 400 pairs of patches. 40 epochs within 40 s (GPU)

No tracking

With tracking

Stable imaging
**Outline**

<table>
<thead>
<tr>
<th>Challenge and Motivation</th>
<th>In a biopsy, compensate physiological motions and reduce musculoskeletal trauma for sonographers using co-robotic ultrasound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods</td>
<td>6-DoF motion estimation. A convolutional neural network is embedded to estimate out-of-plane motions. Joint velocities control.</td>
</tr>
<tr>
<td>Evaluation and Results</td>
<td>Feasibility of the tracking algorithm within simulation environments built with scans obtained on biopsy phantoms, pork, etc.</td>
</tr>
<tr>
<td>Discussion and Conclusions</td>
<td>Advantages. Limitations.</td>
</tr>
</tbody>
</table>

**Discussion and Conclusions**

Verified the feasibility of a co-robotic ultrasound system:

- **Real-time Tracking** of a moving object using 2D B-mode images
- **Out-of-plane motion** estimation using CNN based on speckle patterns
- Velocity control on robot joints

**Advantages:**
- Provide stabilized imaging
- Track in real time
- Use only B-mode images
- Reduce musculoskeletal disorder for sonographers

**Limitations:**
- Larger deformation
- Quick and large motion – beyond the size of a resolution cell

**Further development:**
- Algorithms robust to in-plane rotation
- Kalman filter
- Force control
Appendix 1: tracking algorithm

**Tracking algorithm**

**If using different probes**

**CALIBRATION**
- **Step 1**: acquire parallel B-scans across the tissue within a small volume.
- **Step 2**: fine tune the CNN weights with a small learning rate.

**TRACKING**
- **Step 1**: at time $t = 0$, acquire the target image $I_0$.
- **Step 2**: move end effector a small distance to the (+) elevational direction. Acquire a reference image $I_{ref}$.
- **while** tracking:
  - **Step 3**: compute in-plane transformation → NCC
  - **Step 4**: estimate out-of-plane translation for each patch using the CNN
  - **Step 5**: estimate the overall out-of-plane transformation using all patches
  - **Step 6**: compute transformation error and control the robot joint velocities
Appendix 2: P control on SE(3)

• Implement a proportional (P) controller with double-geodesic feedback on SE(3) group \[11\]

\[
\Delta g = g_{p,p} = g_{pr}(0) \left( g_{pr}(t) g_{ar}^{-1}(t) \right)^{-1} \rightarrow (R_{p,p}, t_{p,p})
\]

\[
\begin{cases}
\tilde{g}_{p,p}^b = -K_\omega \log_{SO(3)} R_{p,p} \\
\tilde{v}_{p,p}^b = -R_{p,p}^T K_\omega t_{p,p}
\end{cases}
\]

\[
\tilde{\gamma}_{p,p} = \left[ \begin{array}{c} \tilde{g}_{p,p}^b \\ \tilde{v}_{p,p}^b \end{array} \right]
\]

\[
\tilde{V}_{bp} = Ad_{g_{F_3}^b} \tilde{\gamma}_{p,p} + \tilde{v}_{p,p}
\]

(Positive definite $K_\omega$ and $K_\nu$

(Change coordinate to the base frame)

• Omit object velocity (the first term $\tilde{V}_{bp_0}^b$) on the R.H.S $\rightarrow$ joint velocity

\[
\dot{q} = J_b^T \tilde{V}_{bp}
\]