Overview: a low-cost and unobtrusive method to register ultrasound scans in 6 degrees of freedom (6 DOF)

A small calibrated camera is rigidly mounted on an ultrasound probe to record skin features when scanning. With each ultrasound scan acquisition, a patch of skin features is recorded at the same time.

Methods: approximate the scan region as a planar structure and estimate camera poses from planar homographies for the camera images

Firstly a square sticker with known dimensions is affixed to the skin surface to obtain the initial world-to-image planar homography and camera pose:

\[
\begin{bmatrix} x_0' \ y_0' \ 1 \end{bmatrix}^T = K \begin{bmatrix} x_0 \ y_0 \ 1 \end{bmatrix}^T, \quad \text{H}_{0i} = K \begin{bmatrix} x_0 \ y_0 \ 1 \end{bmatrix}, \quad K: \text{projection matrix}
\]

The image-to-image correspondences are established by using SIFT (Scale-Invariant Feature Transform) feature descriptors, which are robust to translation/rotation, scale, lighting, and affine distortions.

An image-to-image homography matrix is computed for each consecutive pair of images using their correspondences. The world-to-image homography and thus the camera pose for image \( i \) are then estimated in an iterative manner:

\[
\begin{align*}
H_{0i} &= H_{0,i-1} H_{2i}^T H_{1i}^T, \\
H_{0i} &= K \begin{bmatrix} x_0 \ y_0 \ 1 \end{bmatrix},
\end{align*}
\]

Ultrasound Calibration: convert estimated camera poses to scan poses (i.e. find \( T_{\text{scan}}^\text{camera} \))

The coordinate transformation between the ultrasound scan and world, \( T_{\text{scan}}^\text{world} \), is obtained by a predetermined physical alignment of the probe. The transformation \( T_{\text{world}}^\text{camera} \) is obtained by camera calibration. Combining the two, we have the desired transformation:

\[
T_{\text{scan}}^\text{camera} = T_{\text{scan}}^\text{world} \cdot T_{\text{world}}^\text{camera}
\]

The coordinate transformation between the ultrasound scan and world, \( T_{\text{scan}}^\text{world} \), is obtained by a predetermined physical alignment of the probe. The transformation \( T_{\text{world}}^\text{camera} \) is obtained by camera calibration. Combining the two, we have the desired transformation:

\[
T_{\text{scan}}^\text{camera} = T_{\text{scan}}^\text{world} \cdot T_{\text{world}}^\text{camera}
\]

Linear, tilt, and rotational scanning were tested on a phantom in which a cylinder with known dimensions was embedded.

Experiment Results: the performance of volume imaging was quantitatively evaluated through in-vitro experiments

Volume imaging results and the volume errors: (left to right) linear, tilt, and rotational scanning.