A Reliability Measure for Merging Data from Multiple Cameras in Optical Motion Correction

Jakob Ehrl[1,2], Julian Maclaren[1], Murat Aksoy[1], Andreas Maier[2], Roland Bammer[1]
[1]Stanford University, Dept. of Radiology; [2]FAU Erlangen-Nürnberg, Pattern Recognition Lab

Introduction: Optical tracking of head motion is a promising technique for prospective motion correction in MRI. A camera system computes the position and orientation (pose) of the patient's head with respect to the scanner coordinate system and the RF and gradient waveforms are adjusted to compensate for any motion.

Purpose: The overall aim of this work is to improve tracking precision, accuracy and range by filtering pose estimates and merging data from multiple cameras, as illustrated in Figure 1. A quality metric for pose estimates from each camera is required in order to calculate the weights (w_1, w_2, ..., w_N) used for pose combination. The estimate obtained by weighted combination will be more accurate than any of the individual pose estimates. In this work, we develop a reliability measure used to calculate the weights.

Methods: *Hardware* – A tracking unit for prospective motion correction consisting of an in-bore camera and a 3D-marker was used [1]. The marker was placed in front of the camera at a distance of approximately 10 cm (Fig. 2).

Data processing – The discrete linear transform (DLT) based pose detection algorithm presented in [2] was modified. For each camera frame a covariance matrix for the pose is computed using matrix perturbation theory. This is similar to the idea of error analysis in [3,4]. As a heuristic approach, the quality metric for each pose estimate is calculated as $w = \log(\lambda_2)/\log(\lambda_1)$ where λ_1 and λ_2 are the first and second largest eigenvalues of the covariance matrix. Pose estimates from each tracking unit are then combined in a weighted fashion according to their reliabilities using the computed metric. The optimal estimate for translation is found by a linear interpolation between the translations. The optimal estimate for rotation is found by quaternion interpolation of the respective rotations.

Data collection – Video data were collected while the marker was either partly occluded or completely visible to the camera. Occluding part of the marker simulated an ill-posed 3D-pose estimation since only a small fraction of its feature points could be detected. Two such situations were simulated: 1) part of the marker occluded in vertical (Fig. 2,a) and 2) horizontal direction (Fig. 2, b). For each case the occlusion was removed to simulate a well-conditioned pose estimate. The quality metrics for the occluded and the non-occluded cases were determined.

Results: Figure 3 shows the eigenvalues of the computed covariance matrix and the quality metric before and after the horizontal occlusion (Fig. 2b) was removed from the marker. The quality metric, w, has a low value for the ill-posed case and a high value for the well-conditioned case. The result was similar in both experimental cases (Fig. 2). The current implementation requires approximately 1 ms per estimate to calculate the quality measure.

Discussion: Our quality metric serves as a reasonable measure for the reliability of a pose estimated by the DLT algorithm. Based on matrix perturbation theory, it tells us how sensitive an estimate is to changes in the input data (position of feature points). Hence, compared to other metrics such as the reprojection error [1,2], it is capable of accounting for the sources of error (e.g. view angle, noise, number of point correspondences). The small computational overhead makes our quality metric suitable for real-time tracking applications. Given that computing the covariance matrix actually repeats steps made in the DLT algorithm, there is room for further optimization.

The quality metric is applicable to both multi-camera and single-camera systems. It can be used for pose interpolation such that either data from multiple cameras are fused or a single data stream can be filtered. Thereby each pose estimate will be assigned a weight respective to its quality measure (w_i in Fig. 1) determining its impact on the final result. Choosing the best camera placement to optimize through-plane accuracy and enlarge the field of view will be subject of future work.

Conclusion: A measure for quality and reliability of pose estimates generated by the DLT algorithm was defined. This measure constitutes a promising means for pose filtering and merging data from multiple cameras in order to improve accuracy and reduce the sensitivity to errors in real-time tracking applications.

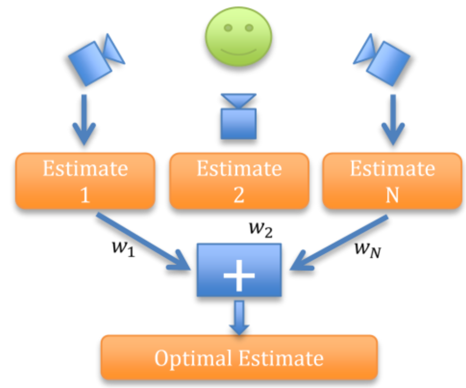

Fig. 1: Computing an optimal pose estimate by a weighted combination of data from multiple single-camera systems

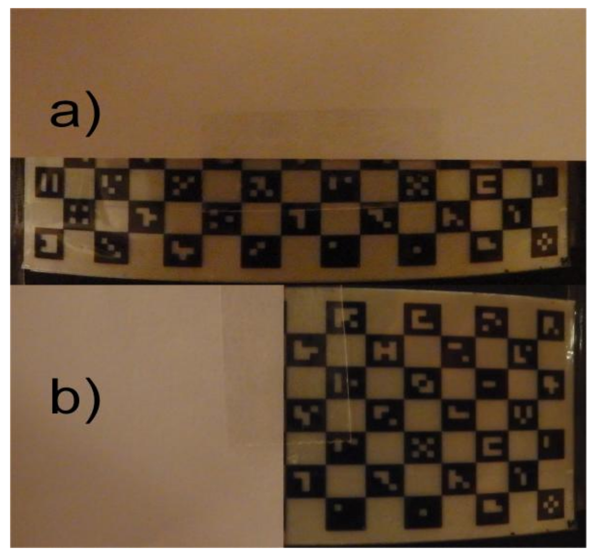

Fig. 2: Camera field of view on circular 3D marker; (a) partly occluded vertically. (b) partly occluded horizontally

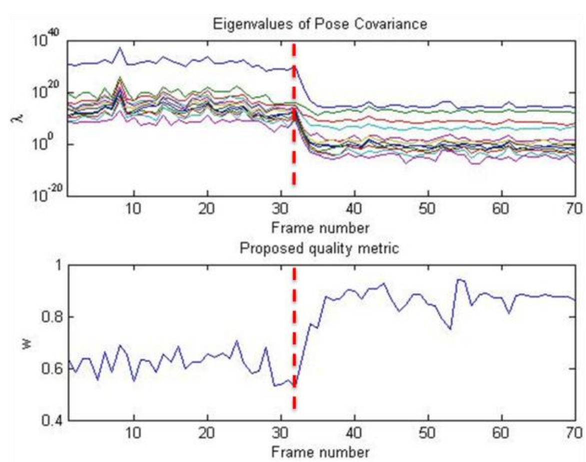

Fig. 3: Eigenvalues of computed covariance matrix (top) and respective quality metric (bottom) while marker is partly occluded (left of dashed line) and when occlusion is removed (right of dashed line).

Acknowledgements: NIH (2R01 EB002711 , 5R01 EB008706, 5R01 EB011654), the Center of Advanced MR Technology at Stanford (P41 RR009784), Lucas Foundation. **References:** [1] Aksoy, et al., MRM 2011;66:366-78. [2] Forman, et al. Medical Image Analysis, vol. 15, no. 5, pp. 708-719, 2011. [3] Criminisi et al., Image and Vision Computing. 17(8):625-634, 1999. [4] Weng et al., IEEE Trans. on Patt. Analysis and Machine Intelligence, 1989;11(5):451-476.