

Epipolar Consistency in Fluoroscopy for Image-Based Tracking

André Aichert¹

andre.aichert@cs.fau.de

Jian Wang¹

jian.wang@cs.fau.de

Roman Schaffert¹

roman.schaffert@fau.de

Arnd Dörfler²

arnd.doerfler@uk-erlangen.de

Joachim Hornegger¹

joachim.hornegger@cs.fau.de

Andreas Maier^{1,3}

andreas.maier@cs.fau.de

¹Pattern Recognition Lab

Friedrich-Alexander Universität

Erlangen-Nürnberg, Germany

²Department of Neuroradiology

Universitätsklinikum Erlangen

Erlangen, Germany

³Graduate School in Advanced Optical Technologies (SAOT),

Erlangen, Germany

Geometry and physics of absorption imaging impose certain constraints on X-ray projections. Various consistency conditions have been derived in the Computed Tomography (CT) community under the assumption of specific scanner trajectories. They can be exploited, for example, in truncation and motion correction. Recently, the Epipolar Consistency Conditions (ECC) [1] have been derived from the epipolar geometry between two arbitrary flat-panel projections. The fact that they make no assumption on trajectory enables us to apply the ECC to patient tracking in fluoroscopy for interventional radiology. For the first time, we perform 3D tracking of an object solely based on the consistency between 2D X-ray shots. The core idea is that before an intervention under fluoroscopy, between two and five reference images of the region of interest are acquired from different angles. We suggest, that an optimization of the ECC of an unseen image with respect to a few reference images enables us to determine a rigid 3D pose of an object or patient.

The ECC are based on the observation, that certain lines in X-ray images contain redundant information. Intuitively, integrating over a line in a projection image is *roughly* the same as integrating over the corresponding plane \mathbf{E} of absorption coefficients through the object. The plane \mathbf{E} is the plane through the source position which intersects the detector in the respective line. For the projection matrices $\mathbf{P}_0, \mathbf{P}_1 \in \mathbb{R}^{3 \times 4}$, the epipolar lines $\mathbf{l}_0, \mathbf{l}_1 \in \mathbb{P}^2$ are special lines in the two images, whose corresponding epipolar plane \mathbf{E} is the same [2]:

$$\mathbf{E} \cong \mathbf{P}_0^\top \mathbf{l}_0 \cong \mathbf{P}_1^\top \mathbf{l}_1. \quad (1)$$

There are two redundant ways to compute the plane integral over \mathbf{E} from either \mathbf{l}_0 or \mathbf{l}_1 , respectively. Epipolar Consistency can thus be quantified by taking several epipolar planes and measuring the difference between these redundant line integrals. Let $\rho_I(\mathbf{l}) = \rho_I(\alpha, t)$ denote the Radon transform of the projection image I at line \mathbf{l} of angle α to the image u -axis and distance t to the origin. The Epipolar Consistency Conditions state

$$\frac{d}{dt} \rho_{I_0}(\mathbf{l}_0) - \frac{d}{dt} \rho_{I_1}(\mathbf{l}_1) \approx 0 \quad (2)$$

where the derivative $\frac{d}{dt}$ in direction of line normals accounts for non-parallel ray geometries [1].

In this work, we assume to have a fluoroscopic video of a patient and a set of reference projections of known projection geometry. Patient movement results in inconsistencies of the fluoroscopic image I_0 with respect to the reference projections. Our goal is to estimate patient motion by minimizing these inconsistencies. A 6-DOF rigid motion is represented by the parameter vector ϕ . For each reference image I_i we define a metric over a set of epipolar planes $\mathcal{E} = \{\mathbf{E}_1, \dots, \mathbf{E}_N\}$

$$M_{0/1} = \frac{1}{N} \sum_{\mathbf{E}_k \in \mathcal{E}} \left(\frac{d}{dt} \rho_{I_0}(\mathbf{P}_0^\top \mathbf{E}_k) - \frac{d}{dt} \rho_{I_i}(\mathbf{P}_i^\top \mathbf{E}_k) \right)^2, \quad (3)$$

which should be minimized over patient motion ϕ .

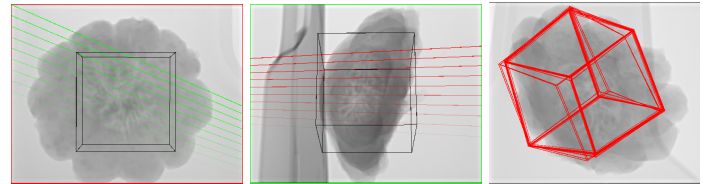


Figure 1: Pumpkin phantom. Left, Center: Two reference X-ray images. Epipolar lines between the images are shown in the color corresponding to the frame. Right: Visualization of 25 iterations of a random study of $\pm 10^\circ$ and ± 25 mm showing two outliers.

There are several assumptions involved in the ECC. Among those, first, an absorption-only model for X-ray physics is assumed when computing the “ray-sums”, which neglects scatter and other non-linear effects of intensity. Second, the method is based on line-integrals in the projection images. Information is therefore obtained only in an orthogonal direction to these lines. It is essential, that reference views are chosen, which ideally produce sets of orthogonal epipolar lines in the input image. Third, and most notably, we present an image-based method. This means, that the object seen by the input and reference views must actually show the same part of the same object in all views.

In our experiments we show, that in spite of these restrictions, it is possible to determine 3D motion based solely on a small set of 2D reference images. We evaluate the proposed method against 2D-3D registration with a previously acquired CT. Our experiments on synthetic data based on a patient CT and pumpkin phantom data from an interventional C-arm (compare Fig. 1) demonstrate that our method is able to compensate for rotations of up to $\pm 10^\circ$ and translations of ± 25 mm between consecutive frames in real time. We successfully track rotations of as much as 45° over 45 images. We show, that small rotations of about 3° in space and translations of about 50 mm can be tracked based on just two reference X-ray images.

The outstanding property of the approach is that no 3D scan is required for tracking a 3D object in space. Since the proposed approach works directly on X-ray images, it exceeds regular 2D-3D registration with a CT in an order of magnitude in computational speed. We conclude that ECC are a simple and effective new tool for pre-alignment and online patient tracking for fluoroscopic sequences.

[1] André Aichert, Martin Berger, Jian Wang, Nicole Maass, Arnd Doerfler, Joachim Hornegger, and Andreas Maier. Epipolar Consistency in Transmission Imaging. *IEEE Trans Med Imaging*, Apr 2015. (<http://dx.doi.org/10.1109/TMI.2015.2426417>).

[2] Richard I. Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.