

Epipolar Consistency in Transmission Imaging

André Aichert, Martin Berger, Jian Wang, Nicole Maass, Arnd Doerfler, Joachim Hornegger, Andreas K. Maier
Pattern Recognition Lab, Friedrich-Alexander Universität Erlangen-Nürnberg, Germany

Abstract—This paper presents the derivation of the Epipolar Consistency Conditions (ECC) between two X-ray images from the Beer-Lambert law of X-ray attenuation and the Epipolar Geometry of two pinhole cameras, using Grangeat’s theorem. We motivate the use of oriented projective geometry to express redundant line integrals in projection images and define a consistency metric, which can be used, for instance, to estimate patient motion directly from a set of X-ray images. We describe in detail the mathematical tools to implement an algorithm to compute the Epipolar Consistency metric and investigate its properties with detailed random studies on both artificial and real FD-CT data. A set of 6 reference projections of the CT scan of a fish were used to evaluate accuracy and precision of compensating for random disturbances of the ground truth projection matrix using an optimization of the consistency metric. In addition, we use three X-ray images of a pumpkin to prove applicability to real data. We conclude, that the metric might have potential in applications related to the estimation of projection geometry. By expression of redundancy between two arbitrary projection views, we in fact support any device or acquisition trajectory which uses a cone-beam geometry. We discuss certain geometric situations, where the ECC provide the ability to correct 3D motion, without the need for 3D reconstruction.

I. INTRODUCTION

The pinhole camera model applies to the geometry of visible-light cameras for photography and X-ray source and detector alike, albeit the imaging process itself is very different. The analogy opens up a field of established methods in Computer Vision which are ready for application to transmission imaging problems [1], [2]. For example, the term epipolar geometry describes the intrinsic geometry between two pinhole cameras. In Computer Vision, it is used, most notably, to estimate the distance of objects to the observer via stereo disparity. While it is in general impossible to estimate the depth of objects from two X-ray images [3], [4], [5], the theoretical model of physics and geometry of transmission imaging still impose certain constraints on X-ray images. This paper illustrates the epipolar geometry of X-ray images and makes the connection to Grangeat’s theorem [6], establishing constraints on the information along corresponding epipolar lines. This redundant information can be expressed as consistency conditions within pairs of X-ray images, for instance, from a flat-detector computed tomography (FD-CT) acquisition. The three most important sources of errors in FD-CT are noise in X-ray acquisitions, inaccuracies and inconsistencies of the geometry due to motion and over-simplified physical models, for example, ignoring scattered radiation. Our goal is to recover some of the degraded geometric information

by optimizing for data consistency. Both epipolar geometry and Grangeat’s theorem have previously been used for this purpose [7], [3], but their connection has only recently been established [8]. In contrast to Wein et al. [3] or Kyriakou et al. [9], the use of Epipolar Consistency Conditions (ECC) [8], [10], [7] does not require intermediate reconstruction and uses a relatively simple and fast metric on pairs of 2D projections. In this paper, we present the derivation of the ECC exploiting synergies in notation of projective geometry and tomographic reconstruction. In Section II, we review oriented projective geometry and epipolar geometry and introduce the notation used throughout the paper. In Section III, we apply the notation of projective geometry to transmission imaging and derive the equation for a 1D-family of consistency conditions between any two X-ray images of the same object. Further in Section IV, we turn to the geometric consistency between X-ray images, as suggested by Debbeler et al. [7] and extend the algorithm by Aichert et al. [8] for the computation of an Epipolar Consistency metric. We conclude in Section VI with an investigation of the properties and geometry of the Epipolar Consistency Metric with respect to observable directions of motion, the parameters of the algorithm, as well as the effect of resolution, noise and using real input images.

II. METHODOLOGY

A. Oriented Lines in Projective Two- and Three-space

1) *Oriented Projective Space*: Projective geometry is a powerful tool for computational geometry, as it simplifies mathematical statements compared to their “Euclidian” notation. It generalizes over many special cases, notably the intersection of parallel lines and planes. Plücker embeddings allow us to define general join and meet operations, i.e. connecting and intersecting flat objects such as points, lines and planes, simply in determinants of their homogeneous coordinates [11]. The use of homogeneous coordinates in projective spaces with the equivalence classes of scalar multiples is well-established in Computer Vision. A major drawback of “classical” projective geometry is that it is not orientable. In this paper, however, it will be advantageous to use the framework of oriented projective geometry [12]. In its representation as homogeneous coordinates, only **positive** scalar multiples of homogeneous coordinates are equivalent. We summarize our notation in Table I. We define oriented projective n -space over the field \mathbb{R} as an embedding in $n + 1$ -space not containing the zero-vector $\mathbf{0}$ modulo positive scalars $\mathbb{P}^{n+} = \frac{\mathbb{R}^{n+1} \setminus \{\mathbf{0}\}}{\mathbb{R}^+}$. The Euclidian point $(u, v)^\top$ is usually represented as the homogeneous vector $(u, v, w)^\top$ with $w = 1$. We introduce the equality relation $\mathbf{a} \cong \mathbf{b} \iff \exists \lambda > 0 \in \mathbb{R}^+ : \lambda \mathbf{a} - \mathbf{b} = \mathbf{0}$ and find that $\forall \lambda > 0 : \mathbf{x} \cong \lambda \mathbf{x}$. In the case of oriented

projective space, we obtain **(positive) points** with $w > 0$ and their “orientation-reversing” equivalents, **negative points** with $w = -1$. Those negative points allow us, for example, to determine if an oriented line pierces an oriented plane from the “back” or from the “front”. Oriented projective two-space always contains **infinite points** $\mathbf{x} \cong (u, v, 0)^\top$, which can be interpreted as directions.

2) The Depth of Points and the Orientation of a Camera:

Albeit less common, oriented projective geometry has been applied to Computer Vision problems, especially Oriented Epipolar Geometry [13]. Orientation is important for a pinhole camera because only points “in front of the camera” (i.e. on one side of the principal plane) are visible. We describe the geometry of an X-ray source and detector using an oriented pinhole camera model, defined by a single projection matrix

$$\mathbf{P} \cong \begin{pmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{pmatrix} = (\mathbf{M} | \mathbf{p}_4) \in \mathbb{R}^{3 \times 4} \quad (1)$$

with the sub-matrix $\mathbf{M} \in \mathbb{R}^{3 \times 3}$ and a column vector \mathbf{p}_4 . The projection matrix \mathbf{P} maps a world point between source and detector in oriented projective three-space $\mathbf{X} \in \mathbb{P}^{3+}$ to an image point on the detector in the oriented projective plane $\mathbf{x} \cong (u, v, 1)^\top \cong \mathbf{P}\mathbf{X} \in \mathbb{P}^{2+}$. The direction of the principal ray is $\mathbf{m}^3 \cong (P_{31}, P_{32}, P_{33})^\top$ the first three elements of the last row of \mathbf{P} . Further, we denote the location of the X-ray source as the non-negative point $\mathbf{C} \cong \ker(\mathbf{P}) \in \mathbb{P}^{3+}$, where $\ker(\cdot)$ denotes the kernel (null-space) of the matrix. Let $\mathbf{X} = T \cdot (X, Y, Z, 1)^\top$ denote a finite world point in a right-handed coordinate frame and $\mathbf{x} = \mathbf{P}\mathbf{X} = w \cdot (u, v, 1)^\top$ be its finite image in pixels, then

$$\text{depth}(\mathbf{X}; \mathbf{P}) = \frac{\text{sign}(\det(\mathbf{M})) \cdot w}{T \|\mathbf{m}^3\|} \quad (2)$$

is the depth of \mathbf{X} w.r.t the pinhole camera \mathbf{P} . Depth is the signed distance to the principal plane $\mathbf{p}^3 \cong (P_{31}, P_{32}, P_{33}, P_{34})^\top$, which is parallel to the image plane and passes through the source position \mathbf{C} (compare Figure 1, right). For a thorough discussion of pinhole cameras in projective space see Hartley and Zisserman [1]. Negative depth is associated with points behind the camera. We can see that Equation 2 depends on the sign of the determinant of \mathbf{M} , as well as on the sign of the homogeneous coordinates of the world and image points.

3) *Oriented Pinhole Camera:* In this paper, we define an oriented pinhole camera, which maps positive points in oriented projective three-space $T > 0$ to positive points in the oriented projective plane $w > 0$ in front of the camera, thus $\text{depth}(\mathbf{X}; \mathbf{P}) > 0$. Since we are explicitly dealing with angles and lengths, our scene is naturally Euclidian and the plane at infinity is fixed. In addition, we restrict ourselves to right-handed coordinate systems, and choose our image coordinate axes accordingly. Since the norm $\|\mathbf{m}^3\|$ is always positive, we may henceforth consider only pinhole cameras with $\det(\mathbf{M}) > 0$. Note that positive world points in front of the camera are mapped to positive image points. Additionally, negative points in three-space behind the camera will be mapped to positive points in the image. This gives us a

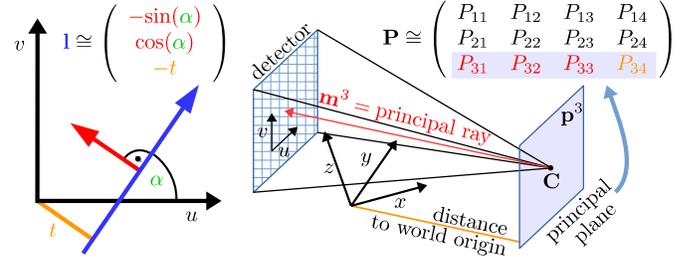


Figure 1. Left: A straight line $\mathbf{l} \in \mathbb{P}^{2+}$ as the set of solutions to the equation $(u, v, 1) \cdot \mathbf{l} = 0$. In this configuration: $t > 0$. Right: The geometry of an oriented pinhole camera with $\|\mathbf{m}^3\| = 1$ and $\det(\mathbf{M}) > 1$ (blue box). Note that the coordinate systems are arbitrary, so long as they are right handed and the image plane faces the source.

natural extension to the concepts in [1] to oriented projective space, where negative points are order-reversing. For details on projective two- and three-space refer to [14], [11], [12]. We refer to tuples (\mathbf{P}_i, I_i) , of projection matrices \mathbf{P}_i and projection images $I_i : \mathbb{R}^2 \rightarrow \mathbb{R}$ with intensity of finite pixels \mathbf{x} (with a slight abuse of notation) $I_i(\mathbf{x}) = I_i(u, v)$ as the i -th view. For convenience, the lower index denotes the view number, for example, $\mathbf{x}_i \cong \mathbf{P}_i \mathbf{X}$ is a point in projective two-space on image I_i . W.l.o.g., we will use the indices 0 and 1 for any two different views.

4) *2D Line Representation:* We can represent a line $\mathbf{l} = (l_0, l_1, l_2)^\top \in \mathbb{P}^{2+}$ in oriented projective two-space for the Euclidian set of solutions to an equation

$$\{(u, v)^\top \in \mathbb{R}^2 : l_0 u + l_1 v + l_2 = 0\} \subset \mathbb{R}^2. \quad (3)$$

The Radon transform on the other hand, is usually parametrized by angle and distance to the origin. The angle α is easily obtained from the direction of the normal $(l_0, l_1)^\top$, while $t = \frac{-l_2}{\sqrt{l_0^2 + l_1^2}}$ is the signed distance to the origin. Note that multiplications by scalars $\lambda < 0$ flip the sign of t and the angle by π . Although the lines \mathbf{l} and $-\mathbf{l}$ contain the same set of points, we would like to differentiate between them, because we need to resolve the ambiguity in orientation α , which is ultimately the reason why we decided to use oriented projective geometry.

$$\mathbf{l} \cong \frac{(l_0, l_1, l_2)^\top}{\sqrt{l_0^2 + l_1^2}} = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \\ -t \end{pmatrix} = \text{line}(\alpha, t) \quad (4)$$

where α is the angle between the x-axis and the line and t is the **signed** distance to the origin (compare Figure 1).

5) *Meet of Two 2D Lines:* The point of intersection $\mathbf{a} = \text{meet}(\mathbf{l}, \mathbf{m}) \in \mathbb{P}^{2+}$ contained in both lines $\mathbf{l}, \mathbf{m}, \in \mathbb{P}^{2+}$ fulfills $\mathbf{a}^\top \mathbf{l} = 0$ and $\mathbf{a}^\top \mathbf{m} = 0$ i.e. the vector representation of \mathbf{a} is orthogonal to those of \mathbf{l} and \mathbf{m} . An orthogonal vector in \mathbb{R}^3 is readily computed as the cross product, which can be written as a multiplication with an anti-symmetric matrix

$$\mathbf{a} \cong \text{meet}(\mathbf{l}, \mathbf{m}) = \mathbf{l} \times \mathbf{m} = [\mathbf{l}]_{\times} \mathbf{m} \quad (5)$$

where $[\cdot]_{\times}$ is an operator which assembles an anti-symmetric matrix out of a three-vector

$$[\mathbf{l}]_{\times} \cong \begin{pmatrix} 0 & -l_2 & l_1 \\ l_2 & 0 & -l_0 \\ -l_1 & l_0 & 0 \end{pmatrix}. \quad (6)$$

6) *Join of Two 2D Points:* Suppose $\mathbf{b} \in \mathbb{P}^{2+}$ is another point on \mathbf{l} , then $\mathbf{a}^\top \mathbf{l} = 0$ and $\mathbf{b}^\top \mathbf{l} = 0$. It follows, that the join operation for connecting two points by a line is also a cross product

$$\mathbf{l} = \text{join}(\mathbf{a}, \mathbf{b}) = \mathbf{a} \times \mathbf{b} = [\mathbf{a}]_{\times} \mathbf{b} \quad (7)$$

and further, it holds that

$$[\mathbf{l}]_{\times} \cong \mathbf{b}\mathbf{a}^\top - \mathbf{a}\mathbf{b}^\top \text{ and } [\mathbf{a}]_{\times} \cong \mathbf{m}\mathbf{l}^\top - \mathbf{l}\mathbf{m}^\top. \quad (8)$$

We say that join and meet operations are dual to each other. While the cross-product is defined only for 3-vectors, the join and meet operations can be extended to arbitrary dimensions using similar concepts.

7) *Join of Two 3D Points as an Anti-Symmetric Matrix:*

In this paper, we extend the operator $[\cdot]_{\times}$ to 4×4 matrices to represent the line through two points in \mathbb{P}^{3+} as follows

$$[\mathbf{L}]_{\times} \cong \mathbf{B}\mathbf{A}^\top - \mathbf{A}\mathbf{B}^\top = \begin{pmatrix} 0 & -L_{01} & -L_{02} & -L_{03} \\ L_{01} & 0 & -L_{12} & -L_{13} \\ L_{02} & L_{12} & 0 & -L_{23} \\ L_{03} & L_{13} & L_{23} & 0 \end{pmatrix} \quad (9)$$

where \mathbf{L} is a six-vector, as defined in the next paragraph and only in this subsection $\mathbf{A}, \mathbf{B} \in \mathbb{P}^{3+}$ denote two 3D points.

8) *3D Line Representation: Plücker Coordinates:* In Equation 9, the matrix $[\mathbf{L}]_{\times}$ is defined by six values of the vector \mathbf{L} , called the Plücker coordinates of the line.

$$\text{join}(\mathbf{A}, \mathbf{B}) = \mathbf{L} \cong \begin{pmatrix} L_{01} \\ L_{02} \\ L_{03} \\ L_{12} \\ L_{13} \\ L_{23} \end{pmatrix} = \begin{pmatrix} A_0B_1 - A_1B_0 \\ A_0B_2 - A_2B_0 \\ A_0B_3 - A_3B_0 \\ A_1B_2 - A_2B_1 \\ A_1B_3 - A_3B_1 \\ A_2B_3 - A_3B_2 \end{pmatrix} \quad (10)$$

Note that Plücker coordinates have only four degrees of freedom, because they are up to scale and there exists an identity going back to the Grassmann-Plücker relations $L_{01}L_{23} - L_{02}L_{13} + L_{03}L_{12} = 0$, which Plücker-Coordinates fulfill [14].

9) *Meet of a 3D Line and a Plane:* Then we get the point of intersection by multiplication $\mathbf{X} = [\mathbf{L}]_{\times} \mathbf{E}$, where \mathbf{E} is the plane in Hessian Normal Form, up to positive scale (analogous to 2D lines from Section II-A4). It is easy to show that

$$[\mathbf{L}]_{\times} \mathbf{E} \cong \mathbf{A}\mathbf{B}^\top \mathbf{E} - \mathbf{B}\mathbf{A}^\top \mathbf{E} = \mathbf{A}\alpha + \mathbf{B}\beta \quad (11)$$

is on the line because it is a linear combination of \mathbf{A} and \mathbf{B} . Additionally, it is a property of anti-symmetric matrices that $\forall \mathbf{E} \in \mathbb{R}^4 : \mathbf{E}^\top [\mathbf{L}]_{\times} \mathbf{E} = 0$, so \mathbf{X} is also contained in the plane \mathbf{E} and must in fact be the point of intersection.

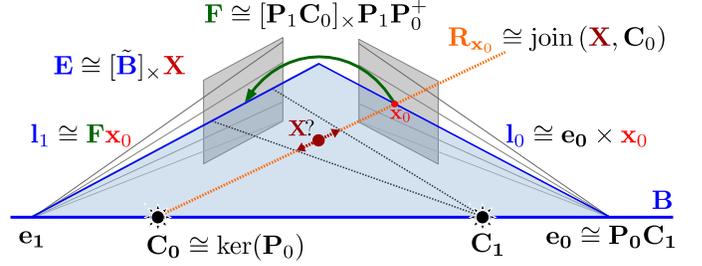


Figure 2. The fundamental matrix \mathbf{F} encodes the relative geometry between two X-ray images for different source and detector positions. It maps an image point \mathbf{x}_0 (bright red) from one detector (gray, right) to a line $\mathbf{l}_1 = \mathbf{F}\mathbf{x}_0$ (left diagonal, blue) on the other detector (left, gray). Suppose \mathbf{x}_0 is the image of a particular feature in 3D space \mathbf{X} (dark red) of unknown depth (for example, the tip of a bone or a metal implant etc.), then the corresponding point $\mathbf{x}_1 = \mathbf{P}_1\mathbf{X}$ which shows the same feature on the other detector is located on that line. In fact, the line is the backprojection ray $\mathbf{R}_{\mathbf{x}_0}$ (dashed, orange) seen from the other source position \mathbf{C}_1 . The baseline \mathbf{B} (solid blue, horizontal) defines a plane through \mathbf{X} , which contains both \mathbf{l}_1 and its corresponding line \mathbf{l}_0 . There is a pencil of such planes, each defining a pair of epipolar lines, for example any of the bright, gray-ish lines, which all intersect in the epipoles \mathbf{e}_1 and \mathbf{e}_0 respectively.

10) *Join of a 3D Line and a 3D Point: Contra-variant Plücker Coordinates:* Without further detail on duality in projective three-space, we can express the dual representation of \mathbf{L} with $[\tilde{\mathbf{L}}]_{\times} [\mathbf{L}]_{\times} = \mathbf{0}$ as an anti-symmetric matrix with

$$\tilde{\mathbf{L}} = (L_{23}, -L_{13}, L_{12}, L_{03}, -L_{02}, L_{01})^\top \quad (12)$$

The plane which contains both the line \mathbf{L} and the point \mathbf{X} is then $\mathbf{E} \cong [\tilde{\mathbf{L}}]_{\times} \mathbf{X}$ [15]. In other words, it represents the pencil of planes around the line \mathbf{L} . We call $[\tilde{\mathbf{L}}]_{\times}$ contra-variant, because it is a representation of a line as the intersection of two planes, and it therefore transforms as such [14].

B. The Fundamental Matrix for Oriented Pinhole Cameras

Epipolar Geometry describes the relationship between two pinhole cameras \mathbf{P}_0 and \mathbf{P}_1 . It is well-known, that it can be fully described by a single 3×3 matrix \mathbf{F} , which maps any point \mathbf{x}_0 on one detector to a line \mathbf{l}_1 on the other detector, compare Figure 2. We present a brief derivation, which takes our special definition of oriented pinhole cameras into account. The projection matrix \mathbf{P}_0 maps a positive world point in front of the camera from oriented projective three-space $\mathbf{X} \in \mathbb{P}^{3+}$ to a positive image point on the detector in the oriented projective plane $\mathbf{x} \cong (u, v, 1)^\top \cong \mathbf{P}\mathbf{X} \in \mathbb{P}^{2+}$. Suppose that the same world point is seen by two cameras as $\mathbf{x}_0 \cong \mathbf{P}_0\mathbf{X}$ and $\mathbf{x}_1 \cong \mathbf{P}_1\mathbf{X}$. The intensity of a pixel \mathbf{x}_0 on the detector goes back to a ray $\mathbf{R}_{\mathbf{x}_0} \cong \text{join}(\mathbf{P}_0^+\mathbf{x}_0, \mathbf{C}_0)$, called the backprojection of \mathbf{x}_0 , where \cdot^+ denotes the pseudo inverse. The epipolar line \mathbf{l}_1 is the backprojection ray projected to the other detector. A line is defined by two distinct points that lie on it. Since every such ray passes through the positive source position \mathbf{C}_0 , we know that $\mathbf{e}_1 \cong \mathbf{P}_1\mathbf{C}_0$ is one point on that line. We call this point epipole and it is positive if and only if \mathbf{C}_0 is in front of \mathbf{P}_1 . We verify that the point $\mathbf{X} \cong \mathbf{P}_0^+\mathbf{x}_0$ is some 3D point, which also has to be on the backprojection ray, because its image is exactly $\mathbf{x}_0 \cong \mathbf{P}_0\mathbf{P}_0^+\mathbf{x}_0$. From our definition of the pinhole camera model, it follows directly that a positive \mathbf{X} is

in front of the camera if and only if \mathbf{x}_0 is positive. In any case, $\mathbf{R}_{\mathbf{x}_0}$ consistently points towards the detector. For a discussion of orientation and epipoles not limited to the oriented pinhole camera used in this paper see the paper by Chum et al. [16].

We have just found two oriented points on the backprojection ray, namely \mathbf{C}_0 and $\mathbf{P}_0^+ \mathbf{x}_0$. We call the join of their projections the oriented epipolar line

$$\mathbf{l}_1 \cong \mathbf{P}_1 \mathbf{C}_0 \times \mathbf{P}_1 \mathbf{P}_0^+ \mathbf{x}_0 = ([\mathbf{P}_1 \mathbf{C}_0]_{\times} \cdot \mathbf{P}_1 \mathbf{P}_0^+) \mathbf{x}_0 \quad (13)$$

for the fundamental matrix for the projections $(\mathbf{P}_0, \mathbf{P}_1)$

$$\mathbf{F}_0^1 \cong [\mathbf{P}_1 \mathbf{C}_0]_{\times} \mathbf{P}_1 \mathbf{P}_0^+ \quad (14)$$

Likewise, the corresponding epipolar line is

$$\mathbf{l}_0 \cong ([\mathbf{P}_0 \mathbf{C}_1]_{\times} \mathbf{P}_0 \mathbf{P}_1^+) \mathbf{x}_1 \quad (15)$$

An oriented epipolar plane \mathbf{E} is readily computed via the backprojection of the corresponding epipolar lines \mathbf{l}_0 or \mathbf{l}_1 respectively as

$$\mathbf{E} \cong \mathbf{P}_0^{\top} \mathbf{l}_0 \cong \mathbf{P}_1^{\top} \mathbf{l}_1 \quad (16)$$

$$\iff \forall \mathbf{X} \in \mathbb{P}^{3+} : \mathbf{P}_0^{\top} [\mathbf{P}_0 \mathbf{C}_1]_{\times} \mathbf{P}_0 \mathbf{X} \cong \mathbf{P}_1^{\top} [\mathbf{P}_1 \mathbf{C}_0]_{\times} \mathbf{P}_1 \mathbf{X} \quad (17)$$

Using the epipolar plane for any world point \mathbf{X} , we can express the relationship between the epipoles

$$\begin{aligned} \mathbf{P}_0^{\top} [\mathbf{P}_0 \mathbf{C}_1]_{\times} \mathbf{P}_0 &\cong \mathbf{P}_1^{\top} [\mathbf{P}_1 \mathbf{C}_0]_{\times} \mathbf{P}_1 \\ \mathbf{P}_1^{+\top} \mathbf{P}_0^{\top} [\mathbf{P}_0 \mathbf{C}_1]_{\times} \mathbf{P}_0 \mathbf{P}_1^+ &\cong [\mathbf{P}_1 \mathbf{C}_0]_{\times} \end{aligned}$$

Substitution yields

$$\begin{aligned} \mathbf{F}_0^1 &\cong [\mathbf{P}_1 \mathbf{C}_0]_{\times} \mathbf{P}_1 \mathbf{P}_0^+ \\ &\cong \mathbf{P}_1^{+\top} \mathbf{P}_0^{\top} [\mathbf{P}_0 \mathbf{C}_1]_{\times} \mathbf{P}_0 \mathbf{P}_1^+ \mathbf{P}_1 \mathbf{P}_0^+ \\ &\cong \mathbf{P}_1^{+\top} \mathbf{P}_0^{\top} [\mathbf{P}_0 \mathbf{C}_1]_{\times} = -\mathbf{F}_1^{0\top} \end{aligned} \quad (18)$$

As a result and in accordance to Laveau et al. [13], we see that the oriented fundamental matrix for the cameras $(\mathbf{P}_0, \mathbf{P}_1)$ is the negative transpose of the fundamental matrix for the cameras $(\mathbf{P}_1, \mathbf{P}_0)$.

C. Plücker Coordinates of the Epipolar Line Bundle

The orientation of epipolar lines determines the orientation of the epipolar plane [16]. Via the fundamental matrix we get for \mathbf{l}_0 and \mathbf{l}_1 consistently

$$\begin{aligned} \mathbf{E} &\cong \mathbf{P}_1^{\top} \mathbf{l}_1 \cong \mathbf{P}_1^{\top} \mathbf{F}_0^1 \mathbf{x}_0 \cong \mathbf{P}_1^{\top} \mathbf{F}_0^1 \mathbf{P}_0 \mathbf{X} \\ &\cong \mathbf{P}_0^{\top} \mathbf{l}_0 \cong -\mathbf{P}_0^{\top} \mathbf{F}_0^{1\top} \mathbf{x}_1 \cong -\mathbf{P}_0^{\top} \mathbf{F}_0^{1\top} \mathbf{P}_1 \mathbf{X} \\ &\cong [\tilde{\mathbf{B}}]_{\times} \mathbf{X} \end{aligned} \quad (19)$$

where the anti-symmetric 4×4 matrix $[\tilde{\mathbf{B}}]_{\times}$ maps points in \mathbb{P}^3 to planes in \mathbb{P}^3 . It is a Plücker representation of a contravariant line through the camera centers, also called the baseline in Stereo Vision. It can be represented as a 6-vector of Plücker coordinates \mathbf{B} with $[\tilde{\mathbf{B}}]_{\times} \cong \mathbf{P}_1^{\top} \mathbf{F}_0^1 \mathbf{P}_0 \cong -\mathbf{P}_0^{\top} \mathbf{F}_0^{1\top} \mathbf{P}_1$ which is the dual to $[\mathbf{B}]_{\times} \cong \mathbf{C}_1 \mathbf{C}_0^{\top} - \mathbf{C}_0 \mathbf{C}_1^{\top}$, compare Equation 9.

Table I
LIST OF IMPORTANT SYMBOLS CONCERNING PROJECTIVE GEOMETRY.

Symbol	Interpretation
\cong	Equality up to positive scale.
$\mathbf{X} \in \mathbb{P}^{3+}$	World point in oriented projective three-space.
$\mathbf{X} = T \cdot (X, Y, Z, 1)^{\top}$	Coordinates of a finite world point ($T \neq 0$).
$\mathbf{P} \cong (\mathbf{M} \mathbf{p}_4) \in \mathbb{R}^{3 \times 4}$	Projection matrix.
$\mathbf{x}_i \cong \mathbf{P}_i \mathbf{X}_i \in \mathbb{P}^{2+}$	The image of \mathbf{X} in the oriented projective plane of the i -th projection image I_i .
$\mathbf{x}_i = w \cdot (u, v, 1)^{\top}$	Coordinates of a finite image point ($w \neq 0$).
$\mathbf{M} \in \mathbb{R}^{3 \times 3}$	Left 3×3 sub-matrix of \mathbf{P} . $\det(\mathbf{M}) > 0$.
$\mathbf{m}^3 \in \mathbb{R}^3$	The last row of \mathbf{M} points in direction of the principal ray.
$\mathbf{p}^3 = (\mathbf{m}^3, P_{34}) \in \mathbb{P}^{3+}$	The last row of projection matrix \mathbf{P} encodes the principal plane, which is parallel to image plane and contains center of projection \mathbf{C} .
$\mathbf{C} \cong \ker(\mathbf{P}) \in \mathbb{P}^{3+}$	Center of projection ("source position"). Null-space of \mathbf{P} and per definition a positive or infinite world point.
$\mathbf{P}^+ \in \mathbb{R}^{4 \times 3}$	Pseudo-inverse of \mathbf{P} (for back-projection).
$\mathbf{B} = \text{join}(\mathbf{C}_0, \mathbf{C}_1) \in \mathbb{R}^6$	Plücker coordinates of the baseline i.e. the line through the source positions \mathbf{C}_0 and \mathbf{C}_1 .
$[\mathbf{B}]_{\times}, [\tilde{\mathbf{B}}]_{\times} \in \mathbb{R}^{4 \times 4}$	Anti-symmetric matrices representing the join and meet operations with the baseline \mathbf{B} . It holds $[\mathbf{B}]_{\times} [\tilde{\mathbf{B}}]_{\times}^+ = \mathbf{0}$.
$\mathbf{E} \cong [\tilde{\mathbf{B}}]_{\times} \mathbf{X} \in \mathbb{P}^{3+}$	An epipolar plane containing the baseline \mathbf{B} and some world point \mathbf{X} .
$\mathbf{l}_1 = \mathbf{P}_1^{+\top} \mathbf{E} \in \mathbb{P}^{2+}$	An epipolar line in the oriented projective plane of the projection image I_1 .
$\mathbf{l}_0 = \mathbf{P}_0^{+\top} \mathbf{P}_1^{\top} \mathbf{l}_1$	The corresponding epipolar line to \mathbf{l}_1 .

We have not been able to find this derivation in literature. It gives us directly the relationship between the two epipolar line bundles in the images and the pencil of planes in three-space. We will exploit this in section IV-A. Due to the number of variables, we decided to summarize our notation in Table I.

III. DERIVATION OF EPIPOLAR CONSISTENCY CONDITIONS

A. Intuition

In the following, we will show that there is a relationship between the sum of intensities along each two epipolar lines in X-ray images. In fact, for parallel projections of small objects, the integral over the projection image along each of these lines would be exactly the integral over the epipolar plane \mathbf{E} through the object. First, we will express the integral over a line in the projection image and a plane through the object and second, we will show that they are related through Grangeat's theorem [6]. Since any epipolar plane corresponds to a pair of epipolar lines \mathbf{l}_0 and \mathbf{l}_1 , this gives us two redundant ways of computing the same integral through the object, one for each line. The result of this section is, that there exists a 1D family of ECC between any two projection images, because there is

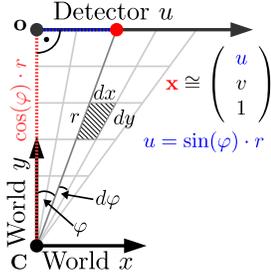


Figure 3. Coordinate system for a specific epipolar plane.

a pencil of such epipolar planes around the line joining the source positions, namely the baseline (compare Figure 2).

B. The Radon Transform of Epipolar Lines and Planes

1) *3D Radon Transform on Epipolar Planes:* We will now establish the relationship between the observed intensities along pairs of epipolar lines and the corresponding plane integral of the object $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $(x, y, z)^\top \mapsto f(x, y, z)$. We denote the backprojection ray $\mathbf{R}_{\mathbf{x}_0} = \text{join}(\mathbf{P}_0^+ \mathbf{x}_0, \mathbf{C}_0)$ which is contained in an epipolar plane $\mathbf{E} \cong \mathbf{P}_1^\top \mathbf{l}_1 \cong \mathbf{P}_0^\top \mathbf{l}_0 \cong (\mathbf{n}^\top, -n)^\top \in \mathbb{P}^3$ with normal \mathbf{n} and signed distance from the origin n .

Let the origin of our coordinate system be the (finite) X-ray source \mathbf{C}_0 with the z -axis pointing in orthogonal direction to an epipolar plane \mathbf{E} and let the y -axis point in direction of the shortest distance to the image plane. Further, let the u -axis of the detector coordinate system coincide with the plane \mathbf{E} and finally, let the x -axis be parallel to the detector. Both world and image coordinates are measured in the same unit. All of this can be achieved by a single rigid transformation of the world and image coordinates and is therefore without loss of generality. In this coordinate system, the plane equation becomes $\mathbf{E} \cong (0, 0, 1, 0)^\top$, hence we need only consider X and Y coordinates in the following, compare Figure 3. We define the 3D Radon transform of the object at a plane \mathbf{E} as the integral

$$\begin{aligned} \rho_f(\mathbf{E}) &= \iiint f(x, y, z) \delta((x, y, z, 1) \cdot \mathbf{E}) \, dx dy dz \\ &= \iint f(x, y, 0) dx dy \end{aligned} \quad (20)$$

using homogeneous coordinates and the Dirac impulse δ .

2) *2D Radon Transform ρ_{I_0} along Epipolar Lines:* Let \mathbf{o} denote the closest point of $\mathbf{l}_0 \cong \mathbf{P}_0^+ \mathbf{E} \cong \text{join}(\mathbf{e}_0, \mathbf{x}_0)$ to \mathbf{C}_0 and d its distance to the origin. We then have a right angle between \mathbf{l}_0 and $\mathbf{R}_{\mathbf{o}} = \text{join}(\mathbf{P}_0^+ \mathbf{o}, \mathbf{C}_0)$. For the points $\mathbf{x}_0 = (u, v, 1)^\top$ on the epipolar line we can thus write

$$F_v(u, r) := f\left(r \cdot (u, d)^\top \cdot \sqrt{u^2 + v^2}^{-1}\right), \quad (21)$$

where r is the distance of \mathbf{x}_0 to $\mathbf{C}_0 \cong (0, 0, 0, 1)^\top$ and the angle $\varphi = \text{atan}(ud^{-1})$ is the ray direction with respect to the closest point \mathbf{o} , hence $u = r \cdot \sin(\varphi)$ and $d = r \cdot \cos(\varphi)$. This fact is visualized in Figure 3. For points on \mathbf{l}_0 , the function $\int F_v(u, r) dr$ is essentially the X-ray transform and samples f along the ray $\mathbf{R}_{\mathbf{x}_0}$. It follows, that for the X-ray intensity

detected in $\mathbf{x}_0 \cong (u, v, 1)^\top$ attenuated by an object f along the ray $\mathbf{R}_{\mathbf{x}_0}$ reads

$$I_0(u, v) = I_{\text{tube}} \cdot \exp\left(-\int F_v(u, r) dr\right) \quad (22)$$

with initial intensity I_{tube} . The X-ray intensity at a single detector pixel is

$$-\ln\left(\frac{I_0(u, v)}{I_{\text{tube}}}\right) = \int_0^\infty F_v(u, r) dr = \int_0^\infty f\left(\begin{array}{c} r \cdot \cos(\varphi) \\ r \cdot \sin(\varphi) \\ 0 \end{array}\right) dr \quad (23)$$

The distance r to the X-ray source defines a specific point on $\mathbf{R}_{\mathbf{x}_0}$. Finally, assuming that the object f is fully visible on the detector (i.e. zero everywhere else or no truncation), the integral over an epipolar line in the polar coordinates of \mathbf{E} is

$$\begin{aligned} \rho_{I_0}(\mathbf{l}_0) &= \iint -\ln\left(\frac{I_0(u, v)}{I_{\text{tube}}}\right) \cdot \delta((u, v, 1) \cdot \mathbf{l}_0) \, dudv \\ &= \int_{-\infty}^{+\infty} \int_0^\infty F_v(u, r) \, dr du \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^\infty \frac{1}{\cos(\varphi)} f\left(\begin{array}{c} r \cdot \cos(\varphi) \\ r \cdot \sin(\varphi) \\ 0 \end{array}\right) \, dr d\varphi \\ &\approx \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^\infty f\left(\begin{array}{c} r \cdot \cos(\varphi) \\ r \cdot \sin(\varphi) \\ 0 \end{array}\right) \, dr d\varphi \end{aligned} \quad (24)$$

This is only an approximation, because the infinitesimal element is regular in u , which thus puts higher weights $\frac{1}{\cos(\varphi)}$ on rays the farther away they are from \mathbf{o} (compare Figure 3). Because φ is bounded by half fan angle, this factor is in practice almost constant and close to one.

C. Cone-beam Geometry

1) *Radial Weights and the 3D Radon Transform ρ_f :* If we write the same plane integral in terms of the polar coordinates, we get the relationship with the integral over the epipolar line \mathbf{l}_0 . The Jacobian determinant of the cartesian-to-polar transformation of the x - y -plane is exactly

$$J_\Phi = r \cdot \cos(\varphi)^2 + r \cdot \sin(\varphi)^2 = r \quad (25)$$

which yields

$$\begin{aligned} \rho_f(\mathbf{E}) &= \iint f(x, y, 0) dx dy \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{+\infty} f(\Phi(\varphi, r)) J_\Phi \, dr d\varphi \\ &\approx \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{+\infty} f\left(\begin{array}{c} r \cdot \cos(\varphi) \\ r \cdot \sin(\varphi) \\ 0 \end{array}\right) r \, dr d\varphi \neq \rho_{I_0}(\mathbf{l}_0) \end{aligned} \quad (26)$$

We observe for cone-beam projections, that the integrals over epipolar lines generally differ by a weighting with the distance to the X-ray source. In the following, we will derive a formulation of a derivative of the epipolar plane integral which happens to cancel out that weighting factor.

2) *Grangeat's Theorem*: The relationship between line integrals on the projection image and plane integrals of the object has been investigated in a different context by Grangeat [6], [17]. For the moment, we restrict ourselves to a single projection image. If we assume w.l.o.g. that the origin of the u - v -plane is located in the principal point \mathbf{p} . We can write the epipolar line $\mathbf{l} \cong (0, 1, -t)^\top$ in terms of a distance t from principal point. As before, the point \mathbf{o} is the orthogonal projection of \mathbf{C} to that line. The 2D Radon transform for \mathbf{l} is

$$\rho_I(\mathbf{l}) = \iint I(u, v) \delta((u, v, 1) \cdot \mathbf{l}) \, dudv = \int I(u, t) \, du \quad (27)$$

Figure 4 (a) reveals the geometric relationships between the 2D Radon transform $\rho_I(\mathbf{l})$ and the 3D Radon transform of the object $\rho_f(\mathbf{E})$. The 3D distance from \mathbf{C} to the line \mathbf{l} is exactly the distance to the image plane within a projection in direction of t . Its orthogonal projection must therefore be again \mathbf{o} . It follows that the lines $\text{join}(\mathbf{p}, \mathbf{o})$ and $\text{join}(\mathbf{C}, \mathbf{o})$ are orthogonal to \mathbf{l} . An arbitrary point \mathbf{x} on \mathbf{l} can be written in terms of the angles κ (between \mathbf{E} and the principal ray) and φ (between \mathbf{o} and \mathbf{x} measured at \mathbf{C}). The distance from \mathbf{C} to \mathbf{l} is then $\cos(\varphi)r$ and the focal distance $\cos(\kappa)\cos(\varphi) \cdot r$ (via triangle $\mathbf{p}, \mathbf{C}, \mathbf{o}$).

Now, we apply Grangeat's theorem and look at the derivative of $\rho(\mathbf{E})$ with respect to the distance to the origin n .

$$\begin{aligned} \frac{d}{dn} \rho_f(\mathbf{E}) &\approx \frac{d}{dn} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{+\infty} F_v(u, r) r \, dr d\varphi \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{+\infty} \frac{d}{dn} F_v(u, r) r \, dr d\varphi \end{aligned}$$

Observe in Figure 4 (b) that there is a relationship $dn = \tan(d\kappa) \cdot \cos(\varphi)r$. Because for small angles $\tan(d\kappa) = d\kappa$ it holds

$$\frac{d\kappa}{dn} = \frac{1}{\cos(\varphi)r} \quad (28)$$

and by chain rule we obtain

$$\frac{d}{dn} F_{\mathbf{x}}(r)r = \frac{d}{d\kappa} \frac{1}{\cos(\varphi)} F_v(u, r) \quad (29)$$

Again, we ignore $\cos(\varphi) \approx 1$ in our computations, because φ is bounded by half fan-angle. We also ignore that the plane normal \mathbf{n} is tilted slightly out of the detector plane, because κ is small. We can compute the derivative w.r.t t instead of n .

$$\begin{aligned} \frac{d}{dn} \rho_f(\mathbf{E}) &\stackrel{d\kappa \approx 0}{\approx} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{+\infty} \frac{d}{d\kappa} \frac{1}{\cos(\varphi)} F_v(u, r) \, dr d\varphi \\ &\stackrel{\varphi \text{ small}}{\approx} \frac{d}{d\kappa} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_0^{+\infty} F_v(u, r) \, dr d\varphi \stackrel{\kappa \text{ small}}{\approx} \frac{d}{dt} \rho_I(\mathbf{l}) \end{aligned} \quad (30)$$

Equation 30 states, that the derivative of a line in t -direction, where t is its distance to the origin, is approximately the same as the derivative of the corresponding plane integral in plane normal direction. If $\mathbf{E} \cong \mathbf{P}_1^\top \mathbf{l}_1 \cong \mathbf{P}_0^\top \mathbf{l}_0$ is an epipolar plane, then the Epipolar Consistency Condition states

$$\frac{d}{dt} \rho_{I_0}(\mathbf{l}_0) \approx \frac{d}{dn} \rho_f(\mathbf{E}) \approx \frac{d}{dt} \rho_{I_1}(\mathbf{l}_1). \quad (31)$$

Table II
LIST OF IMPORTANT SYMBOLS CONCERNING GRANGEAT'S THEOREM.

Symbol	Interpretation
$\delta(\cdot)$	Dirac impulse.
$-\ln \left(\frac{I(u, v)}{I_{tube}} \right) = \int F_v(u, r) \, dr$	X-ray transform in the coordinates of an X-ray fan contained in an epipolar plane, its u -axis lies at the intersection of the detector and the epipolar plane, and r is the distance to the source position.
$\mathbf{E} \cong (\mathbf{n}^\top, n)^\top \in \mathbb{P}^{3+}$	A plane not at infinity is defined by unit-length normal \mathbf{n} and distance to the origin n .
$\rho_f(\mathbf{E}) = \rho_f(\mathbf{n}, n)$	3D Radon transform of the object f is the integral over the plane \mathbf{E} .
$\mathbf{l} = \text{line}(\alpha, t) \in \mathbb{P}^{2+}$	A line not at infinity is defined by angle α and signed distance to the origin t .
$\rho_I(\mathbf{l}) = \rho_I(\alpha, t)$,	2D Radon transform of the image I is the line integral over \mathbf{l} on the detector.
κ	Angle between principal ray and plane \mathbf{E}
\mathbf{o}	Orthogonal projection from principal point \mathbf{p} to line \mathbf{l} .
$\frac{d}{dn} \rho_f$	Derivative of the 3D Radon transform of an object f in n -direction i.e. plane distance to world origin
$\frac{d}{dt} \rho_I$	Derivative of the 2D Radon transform of the image I in t -direction i.e. line distance to image origin
$\frac{d}{dt} \rho_{I_0}(\mathbf{l}_0) \approx \frac{d}{dt} \rho_{I_1}(\mathbf{l}_1)$	Epipolar Consistency Condition for two corresponding epipolar lines \mathbf{l}_0 and \mathbf{l}_1 . For epipolar plane \mathbf{E} see Table I.

Due to the number of variables, we decided to summarize our notation in Table II.

IV. A METRIC FOR GEOMETRIC CONSISTENCY

A. Redundant Information in Two Views

The main result of this paper and the connection between Equations 14 and 30 is, that for any two corresponding epipolar lines \mathbf{l}_0 and \mathbf{l}_1 we find redundant information in two X-ray projections of the same static object, which is fully visible in both views (\mathbf{P}_0, I_0) and (\mathbf{P}_1, I_1)

$$\frac{d}{dt} \rho_{I_0}(\mathbf{l}_0) - \frac{d}{dt} \rho_{I_1}(\mathbf{l}_1) \approx 0. \quad (32)$$

We remind the reader that $\frac{d}{dt} \rho_I(\mathbf{l}) = \frac{d}{dt} \rho_I(\text{line}(\alpha, t))$, or with a slight abuse of notation $\frac{d}{dt} \rho_I(\alpha, t)$, denotes the derivative of the 2D Radon transform of image I in the direction normal to the lines (i.e. distance to the image origin). Equation 32 can also be characterized in other geometrical entities. Let $\mathbf{C} = \ker(\mathbf{P})$ denote camera centers, let $\mathbf{e}_0 = \mathbf{P}_0 \mathbf{C}_1$, $\mathbf{e}_1 = \mathbf{P}_1 \mathbf{C}_0$ be the eipoles. Further let $\mathbf{x}_0 = \mathbf{P}_0 \mathbf{X}$ on I_0 be an image of some world point \mathbf{X} and let $\mathbf{B} = \text{join}(\mathbf{C}_1, \mathbf{C}_0)$ denote the camera baseline, then we have multiple expressions of the epipolar plane \mathbf{E} and its intersections with the images:

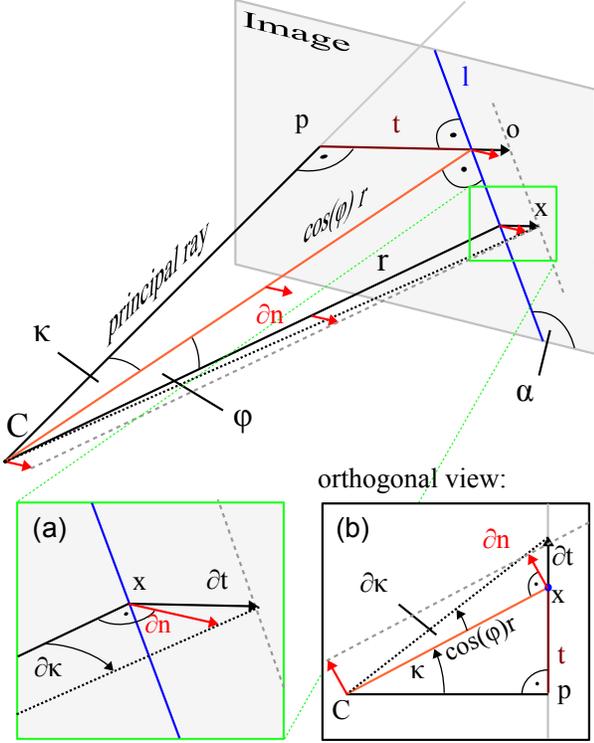


Figure 4. Grangeat's theorem: relationship between angle κ and normal n .

$$\begin{aligned}
0 &\approx \frac{d}{dt}\rho_{I_0}(\mathbf{l}_0) - \frac{d}{dt}\rho_{I_1}(\mathbf{l}_1) \\
&= \frac{d}{dt}\rho_{I_0}(\mathbf{e}_0 \times \mathbf{x}_0) - \frac{d}{dt}\rho_{I_1}(\mathbf{F}_0^1 \mathbf{x}_0) \\
&= \frac{d}{dt}\rho_{I_0}(-\mathbf{F}_0^{1\top} \mathbf{P}_1 \mathbf{X}) - \frac{d}{dt}\rho_{I_1}(\mathbf{F}_0^1 \mathbf{P}_0 \mathbf{X}) \\
&= \frac{d}{dt}\rho_{I_0}(\mathbf{P}_0^{+\top} [\tilde{\mathbf{B}}] \times \mathbf{X}) - \frac{d}{dt}\rho_{I_1}(\mathbf{P}_1^{+\top} [\tilde{\mathbf{B}}] \times \mathbf{X}) \\
&= \frac{d}{dt}\rho_{I_0}(\mathbf{P}_0^{+\top} \mathbf{E}) - \frac{d}{dt}\rho_{I_1}(\mathbf{P}_1^{+\top} \mathbf{E}) \\
&= \frac{d}{dt}\rho_{I_0}(\mathbf{l}_0) - \frac{d}{dt}\rho_{I_1}(\mathbf{P}_1^{+\top} \mathbf{P}_0^\top \mathbf{l}_0)
\end{aligned} \tag{33}$$

Each formulation produces an algorithm for identifying ECCs. Let κ be an angle around the baseline \mathbf{B} , then a specific epipolar plane \mathbf{E}^κ out of the pencil of epipolar planes for two views is identified. Note that we need the other source position \mathbf{C}_1 to define the baseline. We will discuss a complete algorithm to determine \mathbf{E}^κ in Section V. Meanwhile, we can define the following convenient expression for all redundant information found in one particular image I_0 , given its projection matrix \mathbf{P}_0 with respect to the source position \mathbf{C}_1 of another view:

$$\mathcal{R}_0(\mathbf{P}_0, \mathbf{C}_1, \kappa) := \frac{d}{dt}\rho_{I_0}(\mathbf{P}_0^{+\top} \mathbf{E}^\kappa) \tag{34}$$

and \mathcal{R}_1 accordingly. From the second-to-last row of Equation 33 we have an expression for the 1D-family of all epipolar redundancies

$$\forall \kappa \in \left[-\frac{\pi}{2}, +\frac{\pi}{2}\right] : \mathcal{R}_0(\mathbf{P}_0, \mathbf{C}_1, \kappa) \approx \mathcal{R}_1(\mathbf{P}_1, \mathbf{C}_0, \kappa) \tag{35}$$

B. Orientation and Sign

While Debbeler et al. and Maass et al. [7], [10] use a Euclidian framework in world space for their computations, Aichert et al. [8] has provided a formulation in projective space of the projection images. In this paper, we have expressed the same geometry taking orientation into account. From Section II-A4, we immediately see that a projective formulation is up to sign of the intercept and rotation by 180° . The problem with this formulation is the following identity

$$\frac{d}{dt}\rho_I(\mathbf{l}) = -\frac{d}{dt}\rho_I(-\mathbf{l}) \tag{36}$$

The difference of Radon derivatives of two epipolar lines depends on their joint orientation. From Figure 4 we immediately see, that the normals of the epipolar lines have to point towards the **same** half space with respect to the normal of the epipolar plane. This is reflected in Equation 16, which gives us two ways of describing the epipolar plane, one as the backprojection of \mathbf{l}_0 and the other as the backprojection of \mathbf{l}_1 . We can define $\sigma_1^0 = \text{sgn}(\mathbf{l}_0^\top \mathbf{P}_0 \mathbf{P}_1^\top \mathbf{l}_1)$, which is the sign of the cosine of the angle between the two planes. An un-oriented formulation of Equation 32 becomes

$$\frac{d}{dt}\rho_{I_0}(\mathbf{l}_0) - \sigma_1^0 \frac{d}{dt}\rho_{I_1}(\mathbf{l}_1) \approx 0 \tag{37}$$

An alternative is to define $\sigma_1^0 = \text{sgn}(\mathbf{l}_0^\top \mathbf{x}'_0 \cdot \mathbf{l}_1^\top \mathbf{x}'_1)$ by a known point correspondence $0 = \mathbf{x}'_0 \mathbf{F} \mathbf{x}'_1{}^\top$ not on the epipolar lines \mathbf{l}_0 and \mathbf{l}_1 . This is directly related to Chum et al. [16] as it depends on the joint orientation of the epipoles. In experiments we observed, that σ_1^0 is in fact constant for circular trajectories of the x-ray source, as is the case for FD-CT [8]. The sign of σ_1^0 is affected by the orientation of the image planes w.r.t. the baseline \mathbf{B} . This is important when two nearby cameras look in almost the same direction. In this case, the original formulation by Aichert et al. would provide incorrect results, because the sign flips when the cameras are rotated.

C. Metric for Geometric Consistency

1) *In Image Space of I_0* : An alternative approach of the definition of an Epipolar Consistency metric to Equation 35 can be found in Aichert et al. [8]. They parametrize the one-parameter family of epipolar lines \mathbf{l}_0^α using an angle α in image I_0 . In image space, the points $\mathbf{x}_0^\alpha = (e_u + \cos(\alpha), e_v + \sin(\alpha), e_h)^\top$ lie on a circle around a finite epipole \mathbf{e}_0 , no matter the radius $\frac{1}{e_h}$. It is thus possible to define a metric of consistency between two views as an integral over the angle α using the second row of Equation 33:

$$\hat{M}_0^1 = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\frac{d}{dt}\rho_{I_0}(\mathbf{e}_0 \times \mathbf{x}_0^\alpha) - \frac{d}{dt}\rho_{I_1}(\mathbf{F}_0^1 \mathbf{x}_0^\alpha) \right)^2 d\alpha \tag{38}$$

This formulation allows us to discretely sample an arbitrary number of lines based on image points \mathbf{x}_0^α . The downside is the integration happens over an angle α , which is related to the plane angle κ but this formulation is in general not symmetric in the images as we will now show.

2) *The Perspectivity on the Line at Infinity*: The fundamental matrix is a rank deficient 3×3 matrix defined up to scale and has $9 - 1 - 1 = 7$ degrees of freedom (DOF). The metric \hat{M}_0^1 is defined over two line bundles in the two image planes. Given one line bundle defined by the first epipole there remains the freedom of selecting a second epipole plus a dual perspectivity from one bundle to the other. Again there is a total of $2 + 2 + 3 = 7$ DOF which define the relationship between two views. One such parametrization for finite epipoles can be found in Luong et al. [18], where an epipolar line is associated with its intersection with the line at infinity $\mathbf{y}_0^\infty \cong \mathbf{I}_0 \times \mathbf{I}^\infty \cong (1, \tau_0, 0)^\top$, then the mapping to the other line bundle $\mathbf{y}_1^\infty \cong (1, \Psi(\tau_0), 0)^\top$ with $\Psi : \tau_0 \mapsto \tau_1 = \frac{a\tau + b}{c\tau + d}$ has 4 parameters up to scale. W.l.o.g. let $e_h = e'_h = 1$, then the fundamental matrix can be written as $\mathbf{F} = f_{ij}$ with $a = f_{12}$, $b = f_{11}$, $c = -f_{22}$ and $d = -f_{21}$, defined up to scale. With $\mathbf{l}_0^\alpha = \mathbf{e}_0 \times \mathbf{x}_0^\alpha$ and $\mathbf{E}^\alpha = \mathbf{P}_0^\top \mathbf{I}_0^\alpha$ Equation 38 can be written

$$\hat{M}_0^1 = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\frac{d}{dt} \rho_{I_0}(\mathbf{e}_0 \times \mathbf{x}_0^\alpha) - \frac{d}{dt} \rho_{I_1}(\mathbf{e}_1 \times \mathbf{y}_1^\infty) \right)^2 d\alpha \quad (39)$$

In consequence $\hat{M}_0^1 \neq \hat{M}_1^0$, because the Jacobian determinant J_Ψ of the perspectivity Ψ is in general not equal to one. In addition, we have gained an expression for the relative geometry between two projection images, which explicitly parametrizes the epipoles and the mapping between the two line bundles.

3) *In World Space*: Apart from notation and Euclidian and projective descriptions of the ECC, the difference between the algorithms proposed in Debbeler et al., Maass et al. [7], [10] and Aichert et al. [8] is the space in which sampling occurs. We have discussed in the previous section the somewhat artificial asymmetry in computing \hat{M}_0^1 and \hat{M}_1^0 and in contrast, we have Equation 35, which uses the angle κ around the baseline \mathbf{B} . We now replace $\mathbf{x}_0 = \mathbf{P}_0 \mathbf{X}^\kappa$ and use the angle κ in world space as in Figure 4. We can define a metric for the consistency between two views as the squared difference of these redundant 1D-signals depending only on κ :

$$M_0^1 = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\mathcal{R}_0(\mathbf{P}_0, \mathbf{C}_1, \kappa) - \mathcal{R}_1(\mathbf{P}_1, \mathbf{C}_0, \kappa))^2 d\kappa \stackrel{!}{=} 0 \quad (40)$$

The formulation is now symmetric in 1 and 0, so we have $M_0^1 = M_1^0$. The practical question remains, how to discretize κ for an appropriate sampling, which we will discuss in Section V.

4) *In Radon Space*: Suppose we have a finite epipole, w.l.o.g. $e_h = 1$, we can parametrize the line bundle by a single angle α as $\mathbf{l}_\alpha = \mathbf{e} \times \mathbf{x}_\alpha$, where $\mathbf{x}_\alpha = (\cos(\alpha) + e_u, \sin(\alpha) + e_v, 1)^\top$ is any point on the unit circle around \mathbf{e} . We have observed that there exists a pencil of epipolar planes, which intersect in the baseline joining the two source positions. Therefore, all epipolar lines form a bundle which has the epipole $\mathbf{e} \cong (e_u, e_v, e_h)^\top$ as the common point. We get

$$\begin{aligned} \mathbf{l}_\alpha &= \begin{pmatrix} e_v - e_h(\sin(\alpha) + e_v) \\ e_h(\cos(\alpha) + e_u) - e_u \\ e_u(\sin(\alpha) + e_v) - e_v(\cos(\alpha) + e_u) \end{pmatrix} \\ &= \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \\ e_u \sin(\alpha) - e_v \cos(\alpha) \end{pmatrix} \\ &= \text{line}(\alpha, e_u \sin(\alpha) - e_v \cos(\alpha)) \end{aligned} \quad (41)$$

The points $\mathbf{l}_\alpha = \text{line}(\alpha, t(\alpha))$ in Radon space describe a sinusoid curve $t(\alpha) = e_u \sin(\alpha) - e_v \cos(\alpha)$. Note that $t(0) = -e_v$ and $t(\frac{\pi}{2}) = t(-\frac{\pi}{2}) = -e_u$ and that $\|\mathbf{e}\|$ becomes large for translations parallel to the image plane with little rotation outside the image plane. In consequence, the curve will be close to linear within the narrow interval of α , for which the lines intersect the image. If the epipole is at infinity, the line bundle is parallel and therefore $t(\alpha)$ is undefined. The curve in Radon space degenerates to a line parallel to the t -axis.

This observation allows us to sample directly in Radon space. A fast implementation of the ECC requires a pre-computation of $\frac{d}{dt} \rho_I(\alpha, t)$, which is a derivative in t -direction of the Radon transform of a projection image. This means, we have to choose a discretization in α and t . Using the last row in Equation 33 we have a similar formulation to Equation 39:

$$M_0^i(\alpha) = \frac{d}{dt} \rho_{I_0}(\mathbf{l}_\alpha) - \frac{d}{dt} \rho_{I_1}(\mathbf{P}_1^{+\top} \mathbf{P}_0^\top \mathbf{l}_\alpha) \quad (42)$$

but we can choose all \mathbf{l}_α on $t(\alpha)$ and take into account the angular distance. Despite this approach allows us to optimally sample the pre-computed information, the authors prefer the more elegant implementation from Section IV-C3.

D. In Multiple Views

All formulations done until now only considered two arbitrary X-ray images. This is an advantage, because we have imposed no constraints whatsoever on potential future applications, which might include motion correction in fluoroscopy, tomosynthesis and FD-CT. All of these deal with many more than two images. We simply sum up the metric over all pairs of views. If we want to optimize over parameters in \mathbf{P}_0 , for example, we need not compute redundancies between (\mathbf{P}_i, I_i) and (\mathbf{P}_j, I_j) for $i \neq 0 \neq j$, because they remain constant if only \mathbf{P}_0 changes. The correct metric for consistency in a fluoroscope sequence would therefore be $M_0 = \sum_i M_0^i$. If however, the geometry of all images should be optimized, for example, within an FD-CT scan, then the global metric for all pairs would be $M = \sum_{\forall i, j: i < j} M_j^i$. This observation also lets us compute a numerical derivative more efficiently.

V. IMPLEMENTATION

A. Parametrization of the Pencil of Epipolar Planes

This section provides a detailed implementation of the algorithm to compute an Epipolar Consistency Metric in world space, compare Section IV-C3. We construct a mapping from

a single angle κ to an epipolar plane \mathbf{E}^κ through both source positions. We do so, by finding an angle-preserving mapping $\mathbf{A} \in \mathbb{R}^{4 \times 3}$ from the points on the unit-circle of two-space

$$\mathbf{x}^\kappa = (\cos(\kappa), \sin(\kappa), 1)^\top \quad (43)$$

to the points on a circle around the baseline

$$\mathbf{X}^\kappa = \mathbf{A}\mathbf{x}^\kappa = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{t} \\ 0 & 0 & 1 \end{pmatrix} \mathbf{x}^\kappa, \quad (44)$$

with orthogonal column vectors $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^3$ and a translation vector $\mathbf{t} \in \mathbb{R}^4$, i.e. to find as a single 4×3 matrix $\mathbf{K} = [\tilde{\mathbf{B}}]_\times \mathbf{A}$, which maps a unit-vector in two-space directly to an epipolar plane $\mathbf{E}^\kappa = \mathbf{K}\mathbf{x}^\kappa = [\tilde{\mathbf{B}}]_\times \mathbf{A}\mathbf{x}^\kappa = [\tilde{\mathbf{B}}]_\times \mathbf{X}^\kappa$. This type of transformation is also called a correlation.

We start by picking some point \mathbf{X}^0 as a reference point, for example the center of the object or the center of rotation of an FD-CT trajectory. We would like the zero-plane for $\kappa = 0$ to pass through that point.

$$\mathbf{E}^0 = [\tilde{\mathbf{B}}]_\times \mathbf{X}^0 \quad (45)$$

We denote the normal direction with the three vector \mathbf{a}_2 of unit length. The direction of the baseline can be represented as the intersection with the plane at infinity $\pi_\infty \cong (0, 0, 0, 1)^\top$

$$\mathbf{N} \cong [\mathbf{B}]_\times \cdot \pi_\infty \cong \begin{pmatrix} \mathbf{a}_3 \\ 0 \end{pmatrix} \quad (46)$$

with a three-vector \mathbf{a}_3 , again of unit length. \mathbf{N} can be interpreted as a plane orthogonal to \mathbf{B} through the origin. We find that its normal is the direction of the line \mathbf{B} and hence the axis of rotation for the angle κ , in other words $\mathbf{N} = \text{null}(\mathbf{A}^\top)$. We complement a set of orthonormal vectors with $\mathbf{a}_1 = \mathbf{a}_2 \times \mathbf{a}_3$, which points in a direction from the line to the points \mathbf{X}^0 . Observe, that the base vectors \mathbf{a}_1 and \mathbf{a}_2 span the plane \mathbf{N} and that the matrix

$$\mathbf{A}' = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (47)$$

maps the points $(\cos(\kappa), \sin(\kappa), 1)^\top$ to a circle around the origin of three-space contained in the plane \mathbf{N} . Finally, we add a translation \mathbf{t} to move the origin to any point on the line \mathbf{B} , for example, one of the source positions. We summarize the aforementioned steps in Algorithm 1.

B. Sampling Redundant Line-Integrals

Given an epipolar plane \mathbf{E}_k , this section finds the corresponding epipolar lines to sample the pre-computed Radon transform. In practice, algorithms which compute the Radon transform usually parametrize lines by angle α and distance to the origin t relative to the center of the image. Different coordinate systems can be accounted for with a projective transformation of two-space $\mathbf{H} \in \mathbb{R}^{3 \times 3}$, under which a line transforms as $l' \cong \mathbf{H}^{-\top} l$ (i.e. contra-variant). The intersection of the epipolar plane \mathbf{E}^κ with the image plane \mathbf{I}_0 provides the epipolar line $[\mathbf{L}_0^\kappa]_\times = \mathbf{E}^\kappa \mathbf{I}_0^\top - \mathbf{I}_0 \mathbf{E}^{\kappa\top}$ in three-space, \mathbf{L}_1 respectively. A Plücker line projects to the image as a 3×3 anti-symmetric matrix

$$[\mathbf{L}_0^\kappa]_\times = \mathbf{P}_0 [\mathbf{L}_0^\kappa]_\times \mathbf{P}_0^\top. \quad (48)$$

Algorithm 1 Parametrization of epipolar planes by angle κ .

- 1) **Input:** Source positions $\mathbf{C}_0, \mathbf{C}_1$, reference point \mathbf{X}^0
- 2) Compute Plücker coordinates of \mathbf{B} via Equation 10.
 - $[\mathbf{B}]_\times \leftarrow \mathbf{C}_1 \mathbf{C}_0^\top - \mathbf{C}_0 \mathbf{C}_1^\top$
- 3) Find direction of \mathbf{B} via Equation 9.
 - $\mathbf{N} \leftarrow [\mathbf{B}]_\times \pi_\infty$
- 4) Compute plane containing \mathbf{B} and \mathbf{X}^0 via Equation 12.
 - $\mathbf{E}^0 \leftarrow \text{join}(\mathbf{B}, \mathbf{X}^0) \cong [\tilde{\mathbf{B}}]_\times \mathbf{X}^0$
- 5) Find orthonormal basis from plane normals.
 - $\mathbf{a}_2 \leftarrow$ first three elements of \mathbf{E}^0
 - $\mathbf{a}_3 \leftarrow$ first three elements of \mathbf{N}
 - $\mathbf{a}_1 \leftarrow \mathbf{a}_2 \times \mathbf{a}_3$
 - Rescale $\mathbf{a}_1, \mathbf{a}_2$ and \mathbf{a}_3 to unit length
- 6) Find any point \mathbf{O} on \mathbf{B} , for example \mathbf{C}_0 or \mathbf{C}_1 , and use it as translation vector.
 - $\mathbf{t} \leftarrow \frac{1}{O_4} \cdot \mathbf{O}$
- 7) Assemble matrix \mathbf{A} .
 - $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{t} \\ 0 & 0 & 1 \end{pmatrix}$
- 8) **Output:** A 4×3 matrix $\mathbf{K} = [\tilde{\mathbf{B}}]_\times \mathbf{A}$, which maps $\mathbf{x}^\kappa = (\cos(\kappa), \sin(\kappa), 1)^\top$ directly to \mathbf{E}^κ .

However, for planes through the center of projection, notably epipolar planes, there is a simpler relationship

$$\mathbf{E}^\kappa = \mathbf{P}_0^\top \mathbf{I}_0^\kappa = \mathbf{P}_1^\top \mathbf{I}_1^\kappa. \quad (49)$$

From this equation, we can readily compute all corresponding epipolar lines

$$\mathbf{I}_0^\kappa = \mathbf{P}_0^{+\top} \mathbf{E}^\kappa = \mathbf{P}_0^{+\top} \mathbf{K} \mathbf{x}^\kappa \quad (50)$$

and \mathbf{I}_1^κ accordingly. To sample the pre-computed Radon transform, which is typically parametrized by angle α and distance to the origin t , we can apply Equation 4 as summarized in Algorithm 2.

To conclude, the epipolar lines on a particular image I_0 and projection matrix \mathbf{P}_0 are defined by another source position \mathbf{C}_1 . Using the more intuitive notation from Equation 35 we have

$$\begin{aligned} \mathcal{R}_0(\mathbf{P}_0, \mathbf{C}_1, \kappa) &= \frac{d}{dt} \rho_{I_0}(\mathbf{P}_0^{+\top} \mathbf{K} \mathbf{x}^\kappa). \\ \text{and } \mathcal{R}_1(\mathbf{P}_1, \mathbf{C}_0, \kappa) &= \frac{d}{dt} \rho_{I_1}(\mathbf{P}_1^{+\top} \mathbf{K} \mathbf{x}^\kappa) \end{aligned} \quad (51)$$

with $\mathbf{x}^\kappa = (\cos(\kappa), \sin(\kappa), 1)^\top$ and a $\mathbf{K} \in \mathbb{R}^{4 \times 3}$ as computed using Algorithm 1.

C. Observable Motion in Projection Images

In Aichert et al. [8], it was shown that the metric for two views is reliable in only one spatial direction, except for opposing views. The metric consists of a sum of redundant line integrals as in Equation 32. Think of a line integral as a sum of pixel intensities along the line. When the image is moved in direction of that line, the sum does not change. In consequence, $\frac{d}{dt} \rho_I(\mathbf{l})$ is constant for translations of the

Algorithm 2 Computing a pair of epipolar line integrals $\mathcal{R}_0(\mathbf{P}_0, \mathbf{C}_1, \kappa)$ and $\mathcal{R}_1(\mathbf{P}_1, \mathbf{C}_0, \kappa)$ for an angle κ around the baseline given a point of reference \mathbf{X}^0 .

- 1) **Input:** An angle κ , a matrix $\mathbf{K} = [\tilde{\mathbf{B}}]_{\times} \mathbf{A}$, as computed using Algorithm 1, projection matrices \mathbf{P}_0 and \mathbf{P}_1 and pre-computed Radon transforms, derived by the distance to the origin $\frac{d}{dt}\rho_{I_0}$ and $\frac{d}{dt}\rho_{I_1}$.
 - 2) Compute epipolar plane \mathbf{E}^κ .
 - $\mathbf{x}^\kappa \leftarrow (\cos(\kappa), \sin(\kappa), 1)^\top$
 - $\mathbf{E}^\kappa \leftarrow \mathbf{K}\mathbf{x}^\kappa$
 - 3) Compute corresponding epipolar lines
 - $\mathbf{l}_0^\kappa \leftarrow \mathbf{P}_0^{+\top} \mathbf{E}^\kappa$
 - $\mathbf{l}_1^\kappa \leftarrow \mathbf{P}_1^{+\top} \mathbf{E}^\kappa$
 - 4) In practice: Correct for different coordinate systems with a homography \mathbf{H}
 - $\mathbf{l}_0^{\kappa'} \leftarrow \mathbf{H}^{-\top} \mathbf{l}_0^\kappa$
 - $\mathbf{l}_1^{\kappa'} \leftarrow \mathbf{H}^{-\top} \mathbf{l}_1^\kappa$
 - 5) Compute angle and distance to origin via Equation 4.
 - $(a, b, c) \leftarrow \lambda \cdot \mathbf{l}_0^{\kappa'}$, with $0 < \lambda \in \mathbb{R}$, such that $\sqrt{a^2 + b^2} = 1$.
 - $\alpha_0^\kappa \leftarrow \text{atan2}(-a, b)$
 - $t_0^\kappa \leftarrow -c$
 - α_1^κ and t_1^κ accordingly for $\mathbf{l}_1^{\kappa'}$
 - 6) **Output:** Line integrals for corresponding epipolar lines, $\mathcal{R}_0(\mathbf{P}_0, \mathbf{C}_1, \kappa) = \frac{d}{dt}\rho_{I_0}(\alpha_0^\kappa, t_0^\kappa)$ and $\mathcal{R}_1(\mathbf{P}_1, \mathbf{C}_0, \kappa) = \frac{d}{dt}\rho_{I_1}(\alpha_1^\kappa, t_1^\kappa)$.
-

image in direction of the line \mathbf{l} . We see that the ECC give us information orthogonal to epipolar lines only. In fact, there are parallels between the optimization of the ECC and 1D-3D registration, with all the problems familiar from 2D-3D registration. For example, the effect of a camera being rotated slightly around the y -axis is almost the same as translating the camera slightly along the x -axis, due to the 3D \rightarrow 2D projection. Small translations towards and away from the camera have little impact on the images, other than a minor scaling. In addition, a translation parallel to epipolar lines does not change the integral in that very direction, because integration along lines is in effect a projection from 2D \rightarrow 1D. These ambiguities cannot be resolved using redundancies from just two views.

These observations may appear trivial, but it is one of the most important points of this paper. We believe that by design of the geometry, we can find an optimal set of views in terms of the stability of the ECC. Equally spaced cameras on a sphere around the object and looking directly at its center are one such case. The locations of the epipoles decide if the metric can be used to correct motion in a particular spatial direction. Interventional C-arms generally allow for a lot of flexibility in terms of direction of projection. In an application where a certain motion shall be corrected for, its main direction dictates the geometry of reference projections for a stable result. A recent multi-axis robot system might even automatically acquire a few low-dose shots before any FD-CT short scan, to allow consistency metrics to correct for object

motion or refine calibration.

D. Considerations on Geometric Stability

In Figure 6 the location of the epipole for a circular FD-CT trajectory is shown. Since all source positions are on a plane, we do not get reliable information in the detector u axis, except for opposing views. We would need at least one image, which is not in that plane and greatly improve the condition of the problem. We can summarize these considerations in three conditions for an ideal situation for the sensitivity of our cost function for 3D-parameters:

- 1) There should be two views of the object rotated by an angle of approximately 90° . This makes sure that we can estimate the projection direction (i.e. away from the camera) in either view
- 2) We need to have orthogonal epipolar lines in at least two orthogonal projections. The effect of relative motion on the location of the epipoles has been presented nicely by Hartley [1]. For example, this can be achieved, among other situations, by opposing views or by views translated in parallel to the image planes and orthogonal to the plane through the two camera centers and some point inside the object. Our exploratory experiments have indicated, that the ideal case consists of at least three, preferably four equally spaced X-ray sources around an object in all three dimensions, all seeing the object center. Compare Figure 5.
- 3) The projection images must have variations in intensity in both image dimensions. If the image gradient in the image is mostly orthogonal to a set of epipolar lines belonging to one line bundle, then the metric will not reflect motion in that direction.

VI. EXPERIMENTS AND RESULTS

A. Parametrization and Optimization Algorithm

In Section IV-D we suggested a number of possible applications of the ECC. The exact algorithm depends on the specific problem, i.e. the motion model, the kind of trajectory, the time of acquisition of individual projection images and so forth. In order to study the properties of the metric, we restrict ourselves to the case where we optimize the consistency of one particular projection 0 with a set of reference projections 1...n. First, a parametrization of motion has to be defined. In this paper, we restrict ourselves to the largest set of colinearity-preserving transformations, namely projective transformations. Without going into detail, suppose we have a parametrization of homographies on either the projection image $\mathbf{H}^\phi \in \mathbb{R}^{3 \times 3}$, $\det(\mathbf{H}^\phi) > 0$ or the world $\mathbf{T}^\phi \in \mathbb{R}^{4 \times 4}$, $\det(\mathbf{T}^\phi) > 0$ by some parameter vector ϕ . The projection matrix then transforms as $\mathbf{P}_0^\phi = \mathbf{H}^\phi \mathbf{P}_0 \mathbf{T}^\phi$. Just as with image registration, a higher number of parameters may negatively affect stability. In Section V-C we discussed that not all errors may be observable in the metric. A transformation of the world or the image can also be explained by a transformation of the camera, so we just presented a vast over-parametrization. In reality we have at most $n \cdot 11 + 11 - 15$

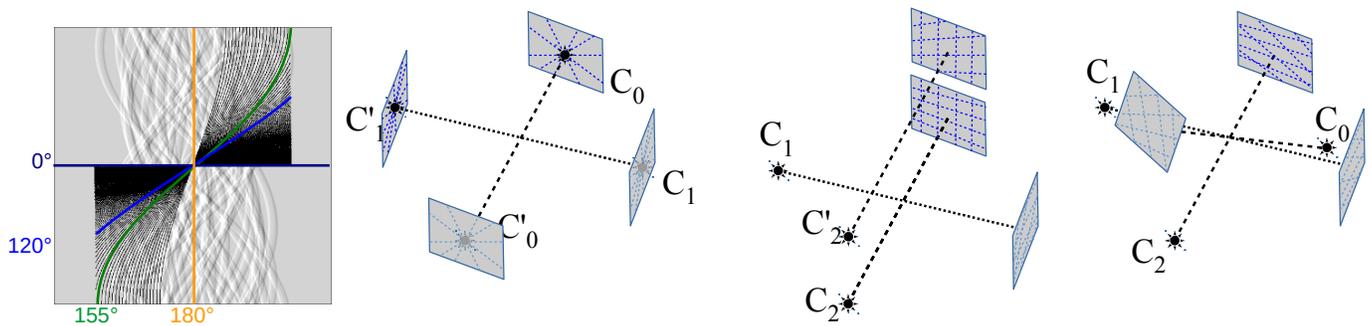


Figure 5. Left: Sinusoid curves in Radon space containing redundant information given an angle between the principal rays of two projections. The color of the lines corresponds to the geometry shown in Figure 6. Others: Three cases which allow us to optimize parameters in all three spacial directions. Left: Opposing views, for all sources on a plane. Only epipolar lines for opposing views are shown. Center: One camera translates out-of-plane. The angle between epipolar lines in image I_2 and I_2 are closer to 90° than in I_1 . Right: Symmetric case when all source positions are roughly equally separated on a half sphere around the object.

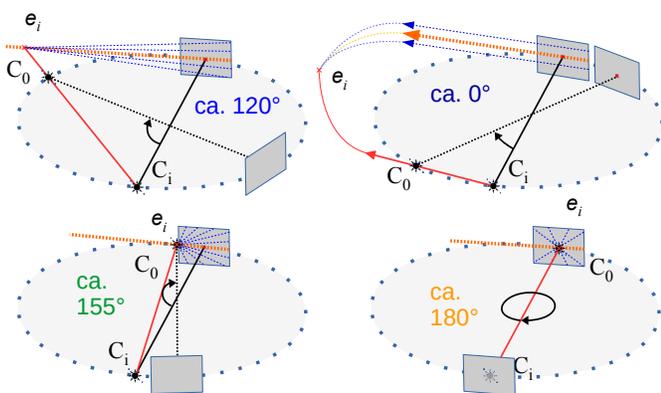


Figure 6. Sketches of the relative geometry for 4 pairs of views out of an FD-CT trajectory. Since there is no out-of-plane motion or rotation in case of a circular trajectory, the epipoles move on a line (dashed orange) in the center of the image. In case of ca. 0° , the detectors are visualized next to each other for better visibility when in reality they would be overlapping: In this case, the epipole is almost at infinity.

independent parameters ($n + 1$ projection matrices up to a transformation of the world). We will thus select only a subset of the aforementioned parameters, for example detector shifts or a rigid motion of the world. Using equation 51 we denote the Epipolar Consistency metric

$$M(\phi) = \sum_{\forall i > 0} \left(\mathcal{R}_0^\phi(\mathbf{P}_0, \mathbf{C}_i, \kappa) - \mathcal{R}_i(\mathbf{P}_i, \mathbf{C}_0, \kappa) \right)^2 \quad (52)$$

and obtain optimal consistency of the 0-th view with the other views for an optimal set of parameters

$$\phi^* = \underset{\phi}{\operatorname{argmin}} (M(\phi)) \quad (53)$$

Second, we need an optimization algorithm. Since the metric is smooth, it seems safe to assume it is in C^2 and we may use one of many local non-linear optimization algorithms and start from an initial guess ϕ_0 . In practice we have an initial guess either from an inaccurate calibration or from previously acquired data, for example in a fluoro sequence. Most non-linear optimization algorithms, which do not assume convexity can be used for this purpose. For example, we can

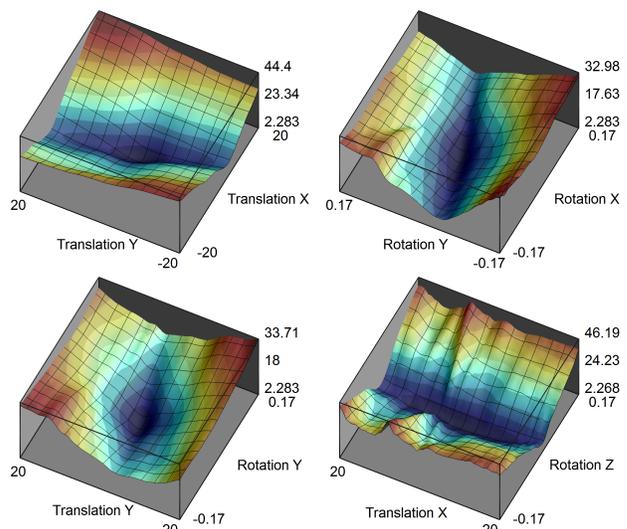


Figure 7. 3D-plots of the Metric of Epipolar Consistency for various combinations of axes of ± 10 and ± 2 cm. . From top left to bottom right: Translation x and y , Rotation x and y , Translation y and Rotation y and Translation x and Rotation z . Note that the same parameter dependencies as in 2D-3D registration apply to this problem: translation in view direction of the camera are more difficult to estimate than the other two spacial directions, because they have a smaller effect on the projection image. Additionally, small rotations about the z -axis have a similar effect as translations along the x -axis (as with y, x swapped), which is also reflected in the valley-shaped plot of the metric on the bottom right.

use a gradient-free optimizer in case of few parameters like Downhill Simplex or, preferably, a gradient-based algorithm and compute a numerical gradient.

B. Implementation Overview

We present the algorithm in world space using sum of squared difference, to compute the ECC for $k = 1 \dots n$ views.

- 1) Define a parametrization of projection matrices (e.g. $\mathbf{P}_0^\phi = \mathbf{P}_0 \mathbf{T}^\phi$ with one rigid 6 DOF transformation \mathbf{T}^ϕ).
- 2) Input: Several views (I_i, \mathbf{P}_i), where i indexes the view number, an angular distance $d\kappa$ (for example: 0.1°) and the number of subdivisions of α and t for Radon transform $\rho_{I_i}(\alpha, t)$.

- 3) Pre-processing: Compute discrete approximations to the Radon derivatives $\frac{d}{dt}\rho_{I_i}(\alpha, t)$ for all images. This means computing the 2D Radon transform and differentiating numerically in t -direction. The pre-computed Radon transform can be re-used, because the images do not change during the optimization.
- 4) **for each** two-set containing the view which is being optimized for $\{(I_0, \mathbf{P}_0^\phi), (I_i, \mathbf{P}_i)\}$
 - a) **for each** integer j , such that the angle $\kappa = j \cdot d\kappa$ meets $-\frac{\pi}{2} \leq \kappa < \frac{\pi}{2}$ (in parallel)
 - i) Compute a plane \mathbf{E}^κ , see Algorithm 1
 - ii) Sample $\mathcal{R}_0(\mathbf{P}_0, \mathbf{C}_i, \kappa)$ and $\mathcal{R}_i(\mathbf{P}_i, \mathbf{C}_0, \kappa)$, see Algorithm 2
 - iii) $m_0^1 = \sum_{\kappa} (\mathcal{R}_0(\mathbf{P}_0, \mathbf{C}_i, \kappa) - \mathcal{R}_i(\mathbf{P}_i, \mathbf{C}_0, \kappa))^2$
 - b) Sum up $m := m + m_0^1$
- 5) Result: $M(\phi) = m$

A smart algorithm could skip planes, which do not intersect the detectors and/or planes which do not intersect the object. In our implementation, the world origin $\mathbf{X}^0 = (0, 0, 0, 1)^\top$ is in the center of the object. We start with the plane $\mathbf{E}^0 = [\mathbf{B}]_{\times} (0, 0, 0, 1)^\top$ and increment κ until we find a plane, which no longer meets the object. Then, we go back to \mathbf{X}^0 and decrement κ until we find another plane which again no longer meets the object.

C. Random Studies

1) *Artificial Data Set:* We use a multistart local optimizer (NLOpt’s Direct method [19] with a SBPLX [20] local optimizer, but other local optimizers were also successful) in a range of $\pm 10^\circ$ and ± 2 cm consistently in all our experiments. We use attenuation-only forward projections with a resolution of 640×640 px² of a CT of a carp¹ (ca. $200 \times 200 \times 400$ mm³) to test the behavior of the metric without truncation and few symmetries in the data. We rendered digitally reconstructed radiographs using ray-casting based on a simple monochromatic noise-free absorption-only model according to the Beer-Lambert law, which does not model scatter.

Truncation as well as the magnitude and direction of edges plays an important role for the metric, so a small animal seems to be an ideal object to study geometric aspects of the optimization. In order to correct for 3D motion, we need to select reference views from all three sides of the object (e.g. front, side, top). Further, we need enough views, such that we have horizontal and vertical epipolar lines in all of the views. Then, any other view around the object will also have orthogonal epipolar lines for some of these reference views. An example is shown in Figure 8. Note how the location of the epipoles indicates, that we can “explain away” up to three of the six images shown and get stable solutions with as few as three reference images. Observe in Figure 7, that the cost function is smooth and has a clear local minimum for rotations and translations in all spatial directions within the range $\pm 10^\circ$ and ± 2 cm.

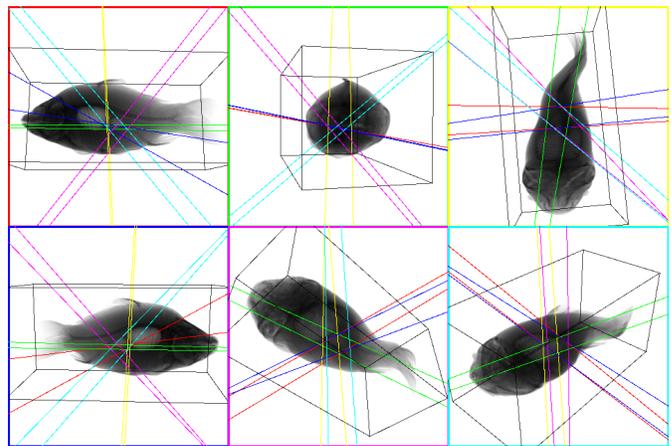


Figure 8. Reference Images of the fish data set. We show two epipolar lines for each other image, in the color of the frame around the other image. In order to optimize for a 3D motion, we need to select views which do not all look in the same or opposing direction and whose respective epipolar lines are almost orthogonal in all images. For example, just by considering the direction of epipolar lines, we can tell, that with respect to the red, green and blue images in the top row, approximately the same information can be derived from the yellow view as from the turquoise and magenta views combined. Another example is, that the magenta, turquoise and any of the top row views together allow for an optimization of 3D object motion .

2) *Error and Accuracy:* Reprojection error (RPE) was calculated from the corner points of the bounding box denoted as the set \mathcal{X} containing 8 points. It is defined as

$$\text{RPE} = \sum_{\mathbf{X} \in \mathcal{X}} \text{distance}(\mathbf{P}_{\text{reference}} \mathbf{X}, \mathbf{P}_{\text{input}} \mathbf{X}) \quad (54)$$

We summed up all RPEs of views that we optimized for. Note that by choosing the bounding box corners for evaluation, we are taking the extreme points of the object, which is an upper bound to the error inside the object. We define the accuracy of the method as the error introduced by the optimization algorithm, when the ground truth is used as a starting point. In a study of 100 different views of the object, we achieved a mean accuracy of about 1.3 px. In most cases, the inaccuracy of the method is not even noticeable close to the center of the object. Accuracy deteriorated when truncation was involved.

3) *Precision:* In order to judge the reliability of our algorithm, we artificially introduced disturbances to the projection matrices. We did so, by applying our parameter model and right-multiplying a rigid transformation of world space to the input projection matrices. The following experiments are “idealized”, since we disturbed only one of the projection matrices, while the reference projection matrices were ground truth. In practice, one would start by optimizing the consistency of the reference views. In all of the following experiments, we chose disturbances uniformly distributed over a range of 10° and 20 mm in all three axes. This corresponds to a quite large mean RPE of about 25 px, or 4% of the image size. If for example, input images were being acquired at a rate of 15Hz, this would correspond to a sudden object (i.e. patient) movement at 30 cm per second and a rotation by 90° in just 0.6 seconds. We define precision as the mean RPE between the most accurate solution and the solution for all randomly disturbed samples. Figure 9 shows the relationship between

¹<http://www9.informatik.uni-erlangen.de/External/vollbib/>

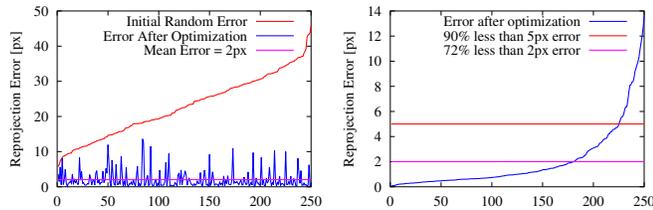


Figure 9. Random Study of 250 random samples with uniform random disturbance of $\pm 10^\circ$ and ± 20 mm sorted by reprojection error before (left) and after (right) optimization. In this case, accuracy was < 0.1 px and we achieved a precision of 2.0 px after optimization, down from ca. 25 px before optimization. In 90% of cases did we achieve an error below 5 px, which is a threshold, beyond which errors are visually noticeable.

random disturbance and optimization result for one of many input frames we tested.

4) *Pre-Computation of Radon Transform*: The algorithm as suggested in this paper has two input parameters: the discretization of the derivative of the Radon transform $\frac{d}{dt}\rho_I(\alpha, t)$ in α and t , as well as the angular distance $d\kappa$ between epipolar planes E^κ . In this section, we study the effect of errors introduced by the pre-computation of the Radon transform. We used the same number of subdivisions in angular α and intercept t direction, i.e. k samples per rotation by π and k samples per translation by image diagonal. We found, that RPE increases rapidly, when the resolution of the Radon derivatives is chosen below 256^2 . It also does not improve when chosen larger than 384^2 , compare Figure 12, left. In all other experiments we constantly used 512^2 . These numbers are related to the resolution of the image, examined in Section VI-C6. Note that we are working with digitally reconstructed radiographs, which due to reconstruction naturally have a lower spatial resolution than a real X-ray image (the carp data set has just $256 \times 256 \times 512$ voxels to begin with).

5) *Sampling the Pencil of Planes*: The other important parameter is the angular difference $d\kappa$, which determines how many samples will be drawn from the Radon derivative. An upper bound is reached, when in both images epipolar lines are less than one pixel apart within image bounds (here: $\approx 0.1^\circ$). In theory, the computational effort depends linearly on this angle. See Figure 10 for the result of random studies for $d\kappa$ from 0.1° to 3° , just like the one in Figure 9. We also show the number of cost function evaluation per second, depending on $d\kappa$. Observe, that at $d\kappa = 1$, we obtained decent results and are able to compute the Epipolar Consistency metric for 7 images about 6 thousand times per second with a CPU implementation on low-end mobile hardware. Interestingly, one can see in Figure 11, that the cost function is less smooth for fewer epipolar planes.

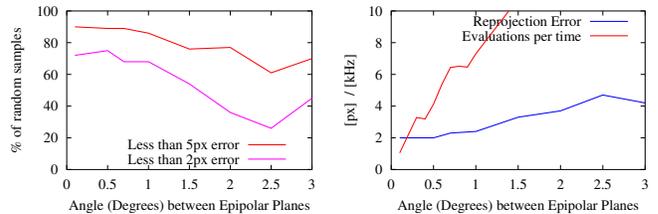


Figure 10. Effect of sampling the Radon transform more densely on quality of optimization and computational speed (mobile Intel i3 CPU).

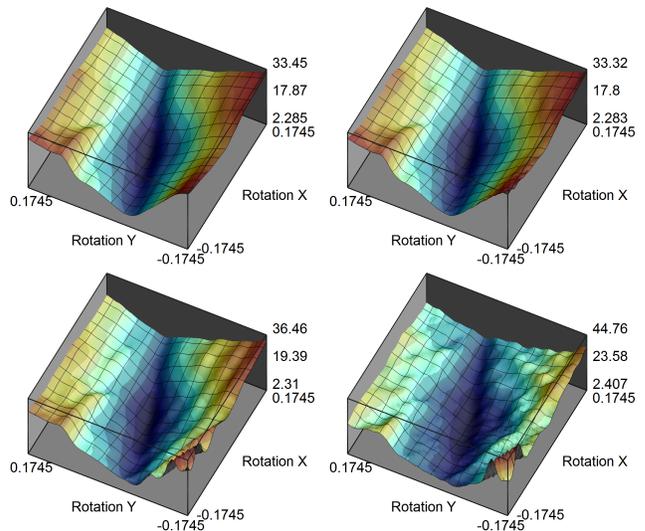


Figure 11. Plots of the cost function for varying angular distance between epipolar planes $d\kappa$. From top left to bottom right: $d\kappa = 0.1, 0.5, 1.0, 2.0$

6) *Image Resolution and Noise*: Find the results of the random study for a Gaussian kernel with standard deviation from 1 – 10 px in Figure 12. Surprisingly, accuracy is merely affected. It may be possible to use a scale-space approach in the optimization. In Figure 13 find results for a simplistic model for both detector and quantum noise. For detector noise, we added Gaussian random values to the line integrals. For quantum noise, we weighted Gaussian random values by intensity. Albeit not a very good approximation to physics, observe in Figure 13, that even for large amounts of noise, we achieve high precision. This is due to the fact that noise averages out during the computation of the line integrals. While this is an expected result, it is nevertheless an important validation, because it implies that we are able to work with very low-dose X-ray images. Interventional C-arms are very flexible in terms of geometry. It would be acceptable, for example, to acquire a few out-of-plane shots before an FD-CT scan, if that allowed us to apply stable motion correction via data redundancies, compare Section V-C. Any robotic C-arm were - in theory - able to acquire two or three low-dose out-of-plane shots before any FD-CT acquisition by default.

7) *Real Data Set*: In addition to the simulation study, we decided to demonstrate applicability on real hardware using projection images from a interventional multi-axial C-arm. To avoid truncation of the object, we use a pumpkin as a phantom, which barely fits the field of view. Note that the

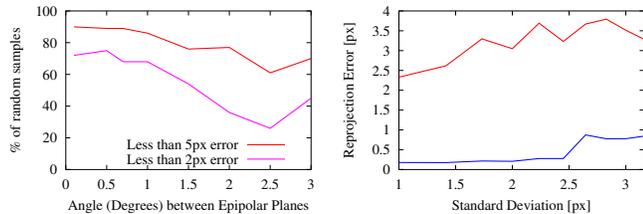


Figure 12. Left: Effect of size of pre-computed Radon derivatives by percentage of random samples with reprojection error less than 2 (magenta) and 5 (red) pixels after optimization. Right: Accuracy (blue) and precision (red) with standard deviations of $\sigma_1 = 0$ and $\sigma_{10} = 3.2$ pixels.

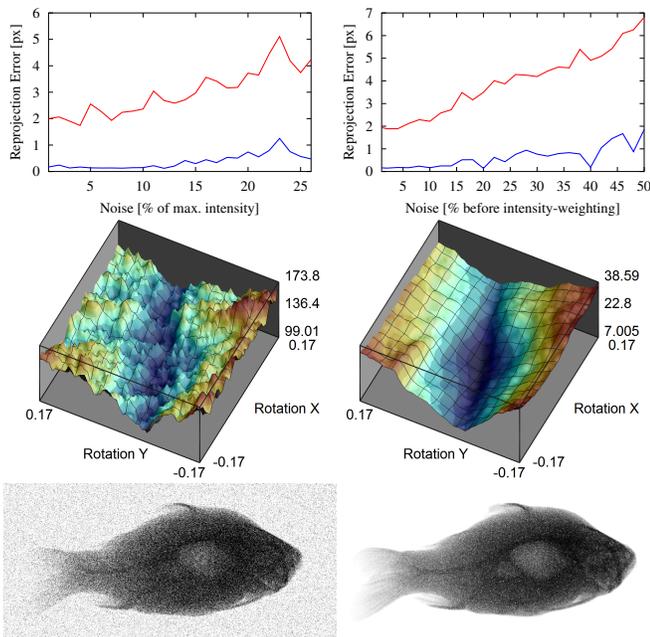


Figure 13. A simplistic model of detector (left) and quantum (right) noise. Top row: a plot of accuracy (blue) and plus precision (red) in a random study with $n = 50$. Center row: A plot of the metric for rotation x,y in case of 25% (left, detector) and 30% (right, quantum) noise, for which the algorithm still works acceptably well. Bottom row: Corresponding noisy digitally reconstructed radiographs. The algorithm proved to be quite resilient to noise.

table is truncated in all projections. We had no control over tube parameters and they vary between projections, which forced us to choose I_{tube} manually per projection. Ideally, the vendor software would extract ray-sums. We present three projection images in Figure 15, along with a visualization of projection geometry in Figure 14. To extract the line-integrals, we divided by a manually chosen initial intensity, which was slightly above the maximum observed intensity. Apart from the bead-phantom presented in Aichert et al. [8], this is the first time motion compensation is applied to real projection images from an interventional C-arm. We conduct two random studies of sample size $n = 500$, for which we randomly disturbed the projection matrix of the image with the black frame in Figure 14 and 15, respectively. The first experiment randomly displaced the object in parallel to the image plane by uniformly distributed offsets in a range of ± 3 cm, which corresponds on average to about 150 px on the detector. Analogously to 2D-3D registration, depth is difficult to optimize in projection

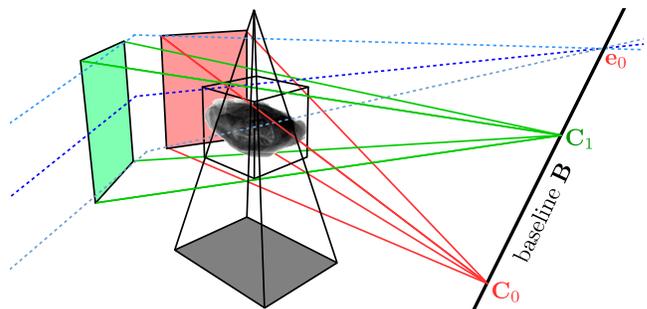


Figure 14. Example for two source positions C_0 and C_1 and detectors (red and green), their joining baseline B (bold black) and three exemplary pairs of epipolar lines (three different shades of blue)

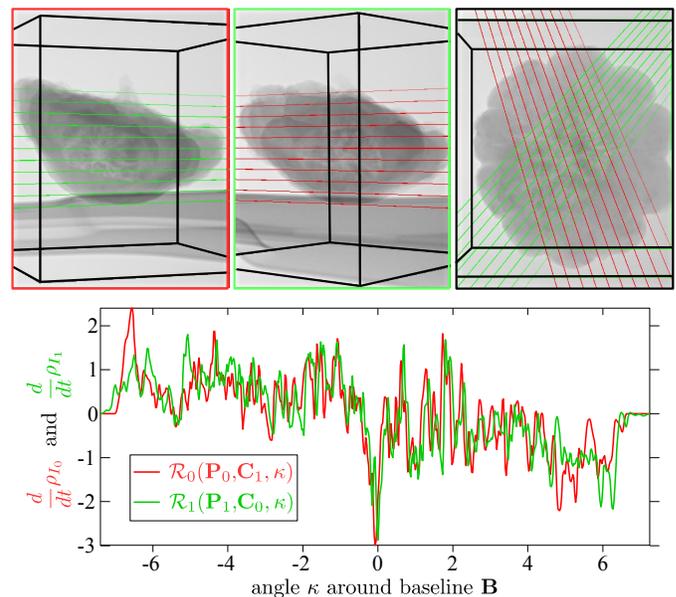


Figure 15. Top: Real projection images (X-ray intensities), which belong to the red and green detectors in Figure 14 and one additional projection (black) which shows a few epipolar lines for both the red and green projections. Ground truth projection matrices were taken from the scanner's odometry. Bottom row: Plot of the signals $\mathcal{R}_0(\mathbf{P}_0, \mathbf{C}_1, \kappa)$ (green) and $\mathcal{R}_1(\mathbf{P}_1, \mathbf{C}_0, \kappa)$ (red) for green epipolar lines in red detector and vice-versa (compare Equation 51).

space, which is why we neglected the direction orthogonal to the image plane. Errors in that direction have relatively little influence on the projection image, hence consistency. The results are shown in Figure 16, right. The correct pose was recovered up to 10 px, except for $< 5\%$ of cases, down from as much as 250 px. This corresponds to a precision of well below three millimeters. The method failed only for extreme offsets of above ca. 300 px. The second random study examined a combination of rotation and translation of the pumpkin. We modeled translations in a range of 10 mm and rotations about all three world axes in a range of $\pm 1^\circ$. The results are shown in Figure 16, left. The correct pose could be recovered up to 15 px down from as much as 100 px, except for two outliers.

VII. CONCLUSION

We present a new formulation for redundancies in transmission images based on the epipolar geometry between any pair

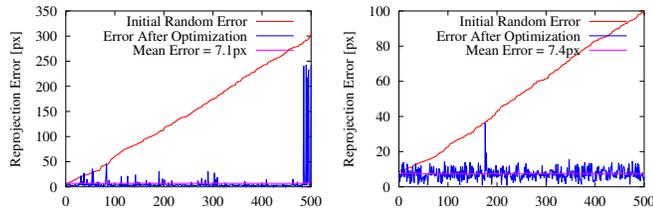


Figure 16. Two random studies of $n = 500$ performed on the rightmost image in Figure 15, sorted by initial error. The left study examines random translations parallel to the image plane in a range of ± 3 cm. The second study included random translations of ± 10 mm and rotations about all three world axes of $\pm 1^\circ$. Accuracy was ~ 2.5 px and mean error ~ 7.5 px in both cases.

of projections of the same object. We use oriented projective geometry to make the connection from Grangeat’s theorem to the epipolar geometry of two X-ray projections, which enables us to identify redundant line integrals in the projection data. The beauty of this paper lies in the combination of three fundamental topics of any course on Medical Imaging and Computer Vision, namely the Beer-Lambert law of X-ray attenuation, the 2D Radon transform and the Epipolar Geometry between two pinhole cameras. We formulate the Epipolar Consistency Conditions (ECC), whose relationship to other consistency conditions, such as Helgason-Ludwig [21], [22] and John’s equation [23] have yet to be established. Possibly, some such formulations of consistency are special cases of ECC. We further derived a metric for geometric consistency, which exploits redundancies of line integrals to optimize the projection geometry. We present a fast algorithm and detailed implementation to compute the metric and observe some of its properties, especially its ability to correct 3D parameters, without the need for 3D reconstruction. By expression of redundancy between two arbitrary projection views, we in fact support any device or acquisition trajectory which uses a cone-beam geometry. We acquired three projections of a pumpkin using a standard clinical C-arm and could show that Epipolar Consistency allows us to optimize geometric parameters. Additionally, we use synthetic data for more detailed and controlled random studies. We argue, that the computation of line-integrals on epipolar lines is essentially a projection from 2D to 1D, which is analogous to the X-ray projection from 3D to 2D. The understanding of the underlying epipolar geometry gives us control over the sampling in Radon space and it helps us to identify geometries for which the metric is reliable. Potential applications might include tracking of the patient or a rigid object in X-ray projections and automatic recalibration of the imaging system for FD-CT reconstruction. Epipolar Consistency could spark more research applications such as truncation or beam-hardening correction. Future work could also attack inconsistencies due to truncation by pre-weighting of intensities or plane-integrals and alternative ways to compare the signals $\mathcal{R}_1(\mathbf{P}_1, \mathbf{C}_0, \kappa)$ and $\mathcal{R}_1(\mathbf{P}_1, \mathbf{C}_0, \kappa)$.

Acknowledgments: This work was supported by Spitzencluster Medical Valley, Verbund Bildgebende Diagnostik: Teilprojekt BD 16 FKZ: 13EX1212 and by the Research Training Group 1774 “Heterogeneous Image Systems” funded by the German Research Foundation (DFG).

REFERENCES

- [1] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.
- [2] O. Faugeras, *Three-dimensional computer vision: a geometric viewpoint*, ser. Artificial intelligence. MIT Press, 1993.
- [3] W. Wein and A. Ladikos, “Detecting patient motion in projection space for cone-beam computed tomography,” in *MICCAI 2011 Proceedings*, ser. Lecture Notes in Computer Science. Springer, Sep. 2011.
- [4] J. Wang, M. Kreiser, L. Wang, N. Navab, and P. Fallavollita, “Augmented Depth Perception Visualization in 2D/3D Image Fusion,” *Computerized Medical Imaging and Graphics*, vol. 2014, 2014.
- [5] M. Wiczelek, A. Aichert, P. Fallavollita, O. Kutter, A. Ahmadi, L. Wang, and N. Navab, “Interactive 3d visualization of a single-view x-ray image,” in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2011*. Springer, 2011, pp. 73–80.
- [6] P. Grangeat, *Mathematical framework of cone beam 3D reconstruction via the first derivative of the radon transform*, ser. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 1991, vol. 1497, ch. chapter 7, pp. 66–97.
- [7] C. Debbeler, N. Maass, M. Elter, F. Dennerlein, and T. M. Buzug, “A new ct rawdata redundancy measure applied to automated misalignment correction,” in *Proceedings of the Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2013, p. 264.
- [8] A. Aichert, N. Maass, Y. Deuerling-Zheng, M. Berger, M. Manhart, J. Hornegger, A. K. Maier, and A. Doerfler, “Redundancies in X-ray images due to the epipolar geometry for transmission imaging,” in *Proceedings of the third international conference on image formation in x-ray computed tomography*, F. Noo, Ed., 2014, pp. 333–337.
- [9] Y. Kyriakou, R. M. Lapp, L. Hillebrand, D. Ertel, and W. A. Kalender, “Simultaneous misalignment correction for approximate circular cone-beam computed tomography,” *Physics in Medicine and Biology*, vol. 53, no. 22, pp. 6267–6289, Nov 2008.
- [10] N. Maass, F. Dennerlein, A. Aichert, and A. Maier, “Geometrical Jitter Correction in Computed Tomography,” in *Proceedings of the third international conference on image formation in x-ray computed tomography*, F. Noo, Ed., 2014, pp. 338–342.
- [11] H. Coxeter, *Projective Geometry*. Springer New York, 2003.
- [12] J. Stolfi, *Oriented projective geometry: a framework for geometric computations*. Boston: Academic Press, 1991.
- [13] S. Laveau and O. Faugeras, “Oriented projective geometry for computer vision,” in *In ECCV96*. Springer-Verlag, 1996, pp. 147–156.
- [14] J. Richter-Gebert. Springer Berlin Heidelberg, 2011.
- [15] J. F. Blinn, “A homogeneous formulation for lines in 3 space,” *ACM SIGGRAPH Computer Graphics*, vol. 11, no. 2, pp. 237–241, Aug 1977.
- [16] O. Chum, T. Werner, and T. Pajdla, “Joint orientation of epipoles,” *British Machine Vision Association*, 2003, pp. 6.1–6.10.
- [17] T. Buzug, *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*. Springer, 2008.
- [18] Q.-T. Luong and O. D. Faugeras, “The fundamental matrix: Theory, algorithms, and stability analysis,” *International Journal of Computer Vision*, vol. 17, no. 1, pp. 43–75, 1996.
- [19] D. R. Jones, C. D. Perttunen, and B. E. Stuckman, “Lipschitzian optimization without the lipschitz constant,” *Journal of Optimization Theory and Applications*, vol. 79, no. 1, pp. 157–181, Oct 1993.
- [20] T. H. Rowan, “Functional stability analysis of numerical algorithms,” Ph.D. dissertation, 1990.
- [21] F. Natterer, *The Mathematics of Computerized Tomography*, ser. Wiley-Teubner series in computer science. Wiley, 1986.
- [22] R. Clackdoyle and L. Desbat, “Full data consistency conditions for cone-beam projections with sources on a plane,” *Physics in Medicine and Biology*, vol. 58, no. 23, pp. 8437–8456, Dec 2013.
- [23] M. S. Levine, E. Y. Sidky, and X. Pan, “Consistency conditions for cone-beam ct data acquired with a straight-line source trajectory,” *Tsinghua Sci Technol.*, vol. 15, no. 1, Feb 2010.
- [24] A. Maier, J. H. Choi, A. Keil, C. Niebler, M. Sarmiento, A. Fieselmann, G. Gold, S. Delp, and R. Fahrig, “Analysis of Vertical and Horizontal Circular C-Arm Trajectories,” in *Proc. SPIE Vol. 7961*, SPIE, Ed., 2011, pp. 7961 231–7961 238.