

Chapter 6

Frame localisation optical projection tomography

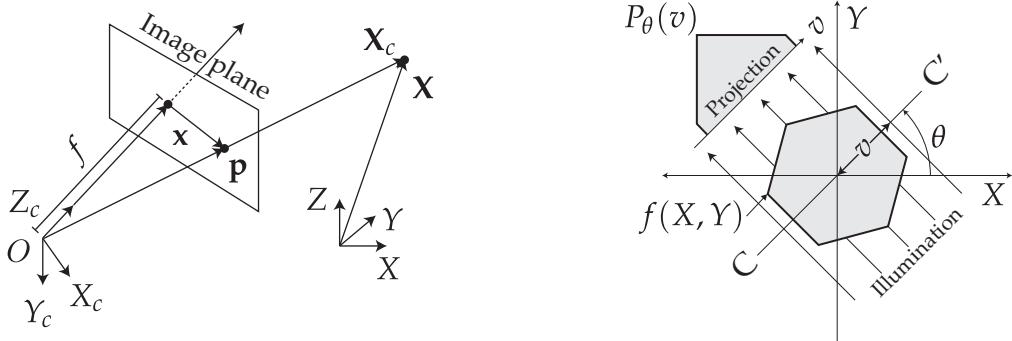
Pascale Sauvage

— Pablo Vallejo

In the previous chapters, volumetric imaging was achieved using wide-field imaging and a relative scanning motion between the system's focal plane and the sample. Volumetric imaging can also be achieved by rotating a sample and tomographically reconstructing the Three Dimensional (3D) distribution of bright signal (scattering, absorption of luminosity) within the specimen. Accurate tomographic reconstruction relies heavily on precision movement and rotation. Here an algorithm will be presented that relies exclusively on multiple (5+) tracked fiducial beads and reconstructs even with systematic mechanical drift.

The projective mathematics introduced in Chapter 5 is a small part of a field of mathematics used in computer vision to localise (in 3D) points in space as projected onto multiple view points. The algorithm presented will use an extension of this projective mathematics to reconstruction Optical Projection Tomography (OPT) volumes from each rotational projection. Considerations were made to combine OPT and light-sheet imaging within the system constructed in Chapter 4, though this was never realised during this thesis.

It should be noted that, due to conventions within the field of computer vision, all lab frame or world coordinates will be represented as capital letters (ex. \mathbf{X} , \mathbf{X}_c) in this chapter as in Fig. 6.1a.



(a) Coordinate system used in this chapter describing a camera with an associated image plane one focal distance f away, imaging an object at point \mathbf{X} .

(b) From an angle θ , an object $f(X, Y)$ and its projection $P_\theta(t)$ are known.

Fig. 6.1 $\mathbf{X}_c = (X_c, Y_c, Z_c)$: camera-centered coordinate point in 3D space. $\mathbf{X} = (X, Y, Z)$: world coordinate point in 3D space. $\mathbf{p} = (x, y, f)$: ray vector to point of image plane. $\mathbf{x} = (x, y)$: image plane coordinates. $\mathbf{w} = (u, v)$: pixel coordinates coordinates (not shown). The optical axis travels along the Z_c axis through the image plane.

6.1 Rotational computer tomography in microscopy

Sharpe *et.al* proposed OPT in 2002 [sharpe_optical_2002] using visible light to image transparent or translucent mesoscopic samples, with micro resolution. OPT addresses a scale gap between the photographic techniques (samples larger than 10 mm), and light microscopy techniques (samples smaller than 1 mm) to image biological samples in the 1 mm to 10 mm regime. OPT is based on computerised tomography techniques [[17]] in which a set of projections of a specimen are imaged as the specimen travels through a full rotation. Algorithms exist to then transform this set of images into a 3D image stack in Cartesian XYZ coordinates. OPT is non-invasive optically but may require specialist invasive preparation for its samples. There are two imaging modalities for OPT, emission Optical Projection Tomography (eOPT) and transmission Optical Projection Tomography (tOPT). In eOPT, a fluorescent sample is excited using an illumination source off axis to the detection, similar to light-sheet but the entire depth of field of the detection of objective is illuminated. Scattered illumination photons are rejected at the detector using an appropriate filter. In tOPT, a white-light source with a diffuser and a collimator is placed along the optical axis to provide near-collimated, uniform illumination onto the sample for transmission to a detector opposite (see Fig. 6.1b). Each photosite at the detector corresponds to a ray that has passed through the sample and been attenuated by the sample. The

eOPT and tOPT modes can work in unison to provide contextual information, with the transmission images indicating overall structure (optical density of absorption or scattering) which can be supplemented by the fluorescent signal from a label of interest.

6.1.1 Reconstruction

As the sample is rotated each pixel collects an intensity $I(\theta) = I_n e^{-k}$ at discrete (n) angles through a full rotation of the sample; where I_n is the unattenuated radiation intensity from the source to the detector, k is the attenuation caused by the sample along a detected ray and $I(n)$ is the measured intensity, see Fig. 6.1b. Rays from the sample to the detector approximate straight lines, and so the rays reaching the detector with a line integrals. A projection is then the resulting intensity profile at the detector for a rotation angle, and the integral transform that results in $P_\theta(v)$ is the Radon transform.

The equation of a set of parallel rays from a source passing through the specimen to a point v along the detector is:

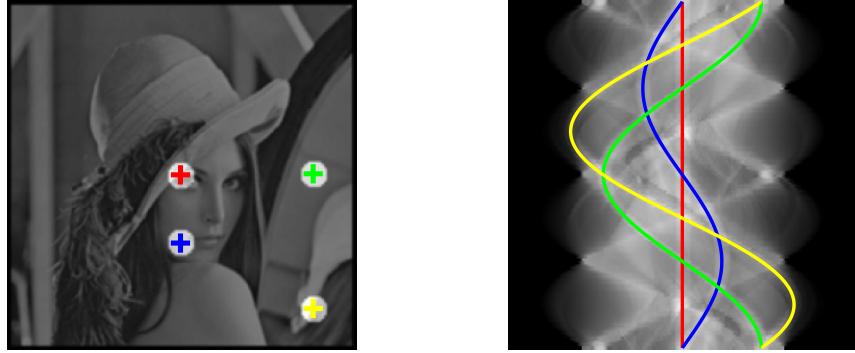
$$X \cos(\theta) + Y \sin(\theta) - v = 0 \quad (6.1)$$

Projecting many such rays through a sample with structure $f(X, Y)$ gives:

$$P_\theta(v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(X, Y) \delta(x \cos(\theta) + y \sin(\theta) - v) dX dY \quad (6.2)$$

Where $P_\theta(v)$ is the Radon transform of $f(X, Y)$ which represents the contrast image of Two Dimensional (2D) slice of the specimen. The Radon transform of an image produces a sinogram as in Fig. 6.2

An inverse Radon transform is used to recover the original object from the projection data; which is achieved by taking the Fourier transform of each projection measurement, then reordering the information from the sample into the respective position in Fourier space. This is valid due to the Fourier Slice theorem (see Appendix ?? for a derivation)[**Fourier slice theorem**], which states that the Fourier transform



(a) Raw input for OPT simulations, Lena ($f(x, y)$)

(b) Image of Lena (Fig. 6.2a) after rotation and projection in 2D, giving the sinogram ($P_\theta(v)$).

Fig. 6.2 Reference images for OPT reconstruction.

of a parallel projection equivalent to a 2D slice of the Fourier transform original sample.

$$f_{\text{fpb}}(X, Y) = \int_0^\pi Q_\theta(X \cos(\theta) + Y \sin(\theta), \theta) dX dY \quad (6.3)$$

Where Q_θ is the filtered projection data, and $f_{\text{fpb}}(X, Y)$ is the back-projected image. A spatial filtering step is applied during back-projection to avoid spatial frequency oversampling during the object's rotation (see Fig. 6.5b) a high pass filter is commonly used to compensate for the perceived blurring. The blurring arises as Q_θ is back-projected (smeared) across the image plane for each angle of reconstruction; which means that not only does the back-projection contribute at the line it is intended to (along line C in Fig. 6.1a), but all other points along the back-projecting ray.

Aim

The Radon transform relies heavily on the assumption of circular motion with constant angular steps about a vertical axis. This chapter presents an improved reconstruction method which explains, some of the techniques used in stereoscopic imaging to register back projections rather than relying on line integrals. The algorithm proposed here is therefore robust to mechanical drifts across acquisitions as well as inconsistent angular steps

6.2 Stereoscopic Imaging

Imaging scenes in stereo allows for the triangulation of individual features in three dimensional space (known as world points) when the features or fiducials in one detector are uniquely identifiable (such as fluorescent beads).

Triangulation requires each feature to be detected in both images of a stereo imaging system, and for these detections to be correctly associated with one another: this is known as the correspondence problem. Many methods exist to ensure that features are detected from image data and accurately associated between two cameras or views. Properties of scale independent features and their surrounding pixel environment in one image can be matched to a similar feature in the second image.

Now, suppose we know the relative positions of the two cameras and their intrinsic parameters, such as magnification and pixel offset. Given the camera parameters, we can translate pixel coordinates $\mathbf{w} = (u, v)$ into image plane coordinates $\mathbf{x} = (x, y)$:

$$u = u_0 + k_u x \quad (6.4)$$

$$v = v_0 + k_v y \quad (6.5)$$

Knowing the focal length (f) of the imaging system, image plane coordinates may be translated into a ray in 3D. The ray can be defined by a point \mathbf{p} in camera-centred coordinates, where it crosses the image plane as:

$$\mathbf{p} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (6.6)$$

From the definition of a world point, as observed through an image, we can construct a dual-view model of world points in space as in Fig. 6.3. Using a model of a system with two views allows for the triangulation of rays based on image correspondences, this is an important part of stereo-vision. The most important matching constraint which can be used is the *epipolar constraint*, and follows directly from the fact that the rays must intersect in 3D space. Epipolar constraints facilitate the search for correspondences, they constrain the search to a 1D line in each image.

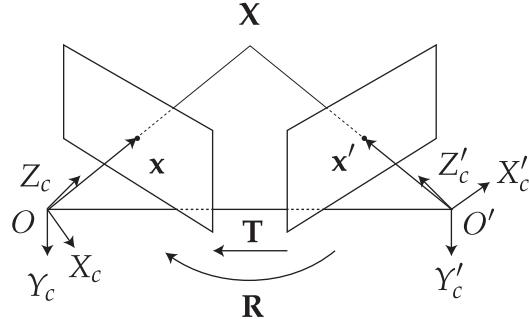


Fig. 6.3 Epi-polar geometry described for two adjacent views (or cameras of a scene). Coordinates as expressed in Fig. 6.1a with prime notation ($'$) denoting the additional right camera view. Transforming from right to left camera-centered coordinates (X'_c to X_c) requires a rotation (R) and a translation (T).

To derive general epipolar constraints, one should consider the epipolar geometry of two cameras as seen in Fig. 6.3

The **baseline** is defined as the line joining the optical centres. An **epipole** is the point of intersection of the baseline with the image plan and there are two epipoles per feature, one for each camera. An **epipolar line** is a line of intersection of the epipolar plane with an image plane. It is the image in one camera of the ray from the other camera's optical centre to the world point (X). For different world points, the epipolar plane rotates about the baseline. All epipolar lines intersect the epipole.

The epipolar line constrains the search for correspondence from a region to a line. If a point feature is observed at x in one image frame, then its location x' in the other image frame must lie on the epipolar line. We can derive an expression for the epipolar line. The two camera-centered coordinate systems X'_c and X_c are related by a rotation, R and translation, T (see in Fig. 6.3) as follows:

$$X'_c = RX_c + T \quad (6.7)$$

Taking the vector product with T , we obtain

$$T \times X'_c = T \times RX_c + T \times T \quad (6.8)$$

$$T \times X'_c = T \times RX_c \quad (6.9)$$

6.2.1 The Essential matrix

Taking the scalar product with \mathbf{X}'_c , we obtain:

$$\mathbf{X}'_c \cdot (\mathbf{T} \times \mathbf{X}_c) = \mathbf{X}'_c \cdot (\mathbf{T} \times \mathbf{R}\mathbf{X}'_c) \quad (6.10)$$

$$\mathbf{X}'_c \cdot (\mathbf{T} \times \mathbf{R}\mathbf{X}_c) = 0 \quad (6.11)$$

A vector product can be expressed as a matrix multiplication:

$$\mathbf{T} \times \mathbf{X}_c = \mathbf{T}_{\times} \mathbf{X}_c \quad (6.12)$$

where

$$\mathbf{T}_{\times} = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix} \quad (6.13)$$

So equation (6.9) can be rewritten as:

$$\mathbf{X}'_c \cdot (\mathbf{T}_{\times} \mathbf{R}\mathbf{X}_c) = 0 \quad (6.14)$$

$$\mathbf{X}'_c \mathbf{T} \mathbf{E} \mathbf{X}_c = 0 \quad (6.15)$$

where

$$\mathbf{E} = \mathbf{T}_{\times} \mathbf{R} \quad (6.16)$$

\mathbf{E} is a 3×3 matrix known as the *essential matrix*. The constraint also holds for rays \mathbf{p} , which are parallel to the camera-centered position vectors \mathbf{X}_c :

$$\mathbf{p}'^T \mathbf{E} \mathbf{p} = 0 \quad (6.17)$$

This is the epipolar constraint. If a point \mathbf{p} is observed in one image, then its position \mathbf{p}' in the other image must lie on the line defined by Equation (6.17). The essential matrix can convert from pixels on the detector to rays \mathbf{p} in the world, assuming a calibrated camera (intrinsic properties are known), and pixel coordinates can then

be converted to image plane coordinates using:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (6.18)$$

We can modify this to derive a relationship between pixel coordinates and rays:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{k_u}{f} & 0 & \frac{u_0}{f} \\ 0 & \frac{k_v}{f} & \frac{v_0}{f} \\ 0 & 0 & \frac{1}{f} \end{bmatrix} \begin{bmatrix} x \\ y \\ f \end{bmatrix} \quad (6.19)$$

$\tilde{\mathbf{K}}$ is defined as follows:

$$\tilde{\mathbf{K}} = \begin{bmatrix} fk_u & 0 & u_0 \\ 0 & fk_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.20)$$

then we can write pixel coordinates in homogenous coordinates:

$$\tilde{\mathbf{w}} = \tilde{\mathbf{K}}\mathbf{p} \quad (6.21)$$

6.2.2 The Fundamental matrix

From (6.17) the epipolar constraint becomes

$$\tilde{\mathbf{w}}'^T \tilde{\mathbf{K}}^{-T} E \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{w}} = 0 \quad (6.22)$$

$$\tilde{\mathbf{w}}'^T F \tilde{\mathbf{w}} = 0 \quad (6.23)$$

The (3×3) matrix \mathbf{F} , is the called the *fundamental matrix*. With intrinsically calibrated cameras, structure can be recovered by triangulation. First, the two projection matrices are obtained via a Singular Value Decomposition (SVD) of the essential matrix. The SVD of the essential matrix is given by:

$$\mathbf{E} = \mathbf{K}'^T \mathbf{F} \mathbf{K} = \mathbf{T}_\times \mathbf{R} = \mathbf{U} \Lambda \mathbf{V}^T \quad (6.24)$$

It can be shown that

$$\hat{\mathbf{T}}_\times = U \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^T \quad (6.25)$$

and

$$\mathbf{R} = U \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} V^T \quad (6.26)$$

Then, aligning the left camera and world coordinate systems gives the projection matrices:

$$\mathbf{P} = \mathbf{K} \left[\mathbf{I} \mid \mathbf{0} \right] \quad (6.27)$$

and

$$\mathbf{P}' = \mathbf{K}' \left[\mathbf{R} \mid \mathbf{T} \right] \quad (6.28)$$

Where $\left[\mathbf{I} \mid \mathbf{0} \right]$ is the identity matrix augmented column-wise with a zero matrix, and the two projection matrices (\mathbf{P} and \mathbf{P}') project from camera pixel coordinates to world coordinates. Given these projection matrices, scene structure can be recovered (only up to scale, since only magnitude of \mathbf{T} ($|\mathbf{T}|$) is unknown) using least squares fitting. Ambiguities in \mathbf{T} and \mathbf{R} are resolved by ensuring that visible points lie in front of the two cameras. As with the essential matrix, the fundamental matrix can be factorised into a skew-symmetric matrix corresponding to translation and a 3×3 non-singular matrix corresponding to rotation.

6.3 The proposed algorithm

The shift of a camera or a camera pair around a scene separated by a transformation matrix ($[R \mid T]$) is analogous to transforming the sample in the fixed view of an imaging detector, as in OPT. During a typical OPT acquisition, a marker will appear to follow an elliptical path in the XY plane. In the following volume reconstruction there will then be a fitting step to recover the path of the fiducial, to then apply a correction before applying a Radon transform. This type of reconstruction not only ignores any mechanical jitter of the sample but also any affine, systematic, mechanical drift ($X, Y, Z, \theta, \phi, \psi$).

We have seen that using two adjacent images, of a scene separated by some rotation and translation, world points in 3D space may be triangulated within the scene given the rotational and translational matrices of the respect camera views. The inverse is also possible, given a sufficient number of known fiducial points in a scene the translation and rotation matrices can be recovered. The recovery of a more exact description of the motion of the scene can eliminate any need for a fitting and may recover and correct for drift, as well as eliminate any mechanical jitter. Errors may however then be introduced from fiducials mechanically slipping and localisation errors. Fiducials in this sense refer to an accurately locatable marker common through different views in a sample.

Once a sufficient amount of fidicuials are reliably tracked from the first to the second image, the fundamental, essential and Homography matrices can be computed. Using the factorisation one of these matrixes, between each adjacent view of a rotating scene, the translation and rotational matrices may be recovered. Here we will discuss a reconstruction using F but the same principle applies for E and H .

Here are two ways of reconstructing using the fundamental matrix as described above. The first method involves computing F for two neighbouring images with 5 or more fiducials, having additional beads helps to remove ambiguity and increase confidence in F . Once F is calculated F is then decomposed into R_n and T_n between each view n and $n + 1$. The image at view $n + 1$ is then back projected along the virtual optical axis within a virtual volume where the sample will be reconstructed. The size of this back projection and virtual volume is chosen to be suitably large (so that important data is not lost). Then, all the prior rotation and translation matrices are serially multiplied from $[R_0 \mid T_0]$ until $[R_n \mid T_n]$, this final matrix is inverted and applied to the back projected volume. The matrix inversion step is important as

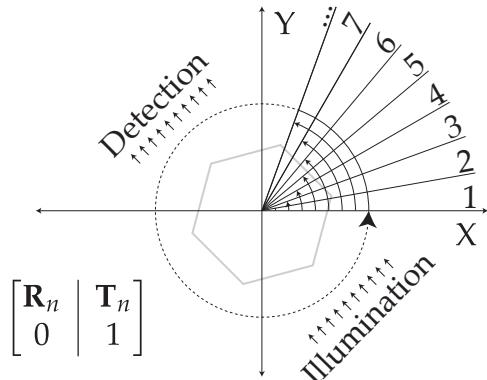


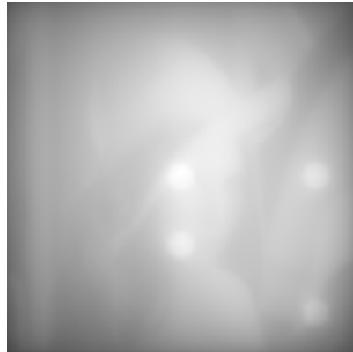
Fig. 6.4 Principles of the proposed algorithm. Each iteration during an OPT acquisition will have an associate \mathbf{R} and \mathbf{T} (shown here in augmented form using homogenous coordinates), these matrices can be recovered from comparing the current iteration (n) to the next iteration ($n + 1$).

it realigns the back projection in the volume to where it originally was compared respective of the first projection. This process is repeated for every angle and the back projected volume from each step is summed with every other step. Finally the remaining volume is filtered using a high-pass filter; here a Ram-Lak filter is used.
¹By producing a series of transformation matrices from adjacent acquisitions, errors compound and the reconstruction of volumes degrades with more projections, see Fig. 6.5.

The second approach is less prone to compound errors but relies on precise identification and tracking of fiducial markers. Instead of calculating \mathbf{F} between neighbouring images \mathbf{F} is calculated between the current projection and the very first projection. \mathbf{F} is then decomposed and the transformation matrix is inverted and applied to the back projected volume. The reoriented back projected volumes are summed and finally filtered to remove the additional spatial frequencies imparted from rotating the sample.

The second approach is robust against compound errors but an additional programmatic step is needed to know which beads in the first image correspond to beads in the n^{th} image. This can be achieved using tracking and momentum particle tracking algorithms, though confounding issues can arise i.e. if a particle orbits too far away or occlusions occur. In both cases a decomposed \mathbf{F} matrix will produce four possible transformation pairs ($\mathbf{R}, \mathbf{T}; \mathbf{R}, -\mathbf{T}; -\mathbf{R}, \mathbf{T}; -\mathbf{R}, -\mathbf{T}$). Once the transformation matrix between the first view and the second view is calculated the

¹Linear real (amplitude) ramp filter in Fourier space



(a) Unfiltered output of the Radon transform i.e. a reconstruction



(b) Ram-lak (Fourier ramp) filter applied to Fig. 6.5a.

Fig. 6.5 The result of a tomographic reconstruction requires Fourier filtering to normalise spatial contrast.

proceeding transformation matrices are then easily chosen by similarity and general direction of motion. An example of this type of selection would be:

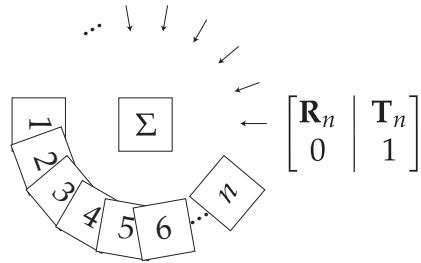
$$\min_{I(n)} \left[I(n) = \left(\left[\mathbf{R}_n \mid \mathbf{T}_n \right] - \left[\mathbf{R}_{n-1} \mid \mathbf{T}_{n-1} \right] \right)^2 \right] \quad (6.29)$$

The first two views are more difficult to choose the correct decomposition for, but it is possible if a suitable ideal matrix is given as a comparison. Such an ideal matrix is composed using *a priori* knowledge of the likely angle of rotation of the system's imaging properties.

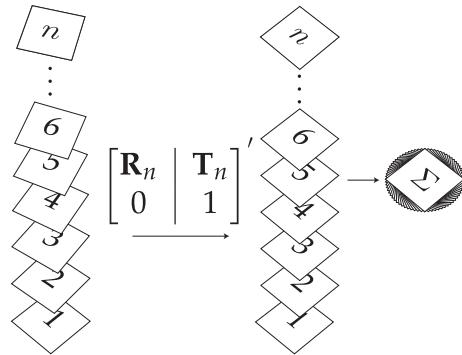
6.4 Verification of the proposed algorithm

To verify that the proposed algorithm successfully reconstructs the specimen, it was applied to simulated data. The image of Lena² is used here as a test image to verify the validity of each reconstruction. Superimposed on Lena are fiducial beads to track the rotation of the image, see Fig. 6.2a. The reference image was then rotated through 128 angles over 2π radians and projected along the Y axis and a slice in XY was taken to create a single line projection, shown three dimensionally in Fig. 6.7. This is repeated for each angle with each line projection stacked to create a sinogram, see Fig. 6.2b.

²Standard reference image data



(a) Forward model



(b) Reconstruction method, solving the inverse problem.

Fig. 6.6 Two dimensional representation of the reconstruction algorithm. (a): For each projection of the object (Σ), at rotation (\mathbf{R}_n) and translation (\mathbf{T}_n), produces n data sets. (b): For each data set the rotational and translation matrices are recovered (\mathbf{R}'_n and \mathbf{T}'_n) and inversely applied to align back project the datasets, the matrices are shown in augmented form using homogenous coordinates. The now realigned back projections are summated to produce an unfiltered back projection.



(a) 3D sample for reconstruction using Cameraman and Lena testcard images.



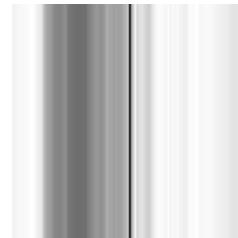
(b) Raw
Frame (n): 0
View: XY



(c) Reconstructed
Frame (n): 0
View: XY



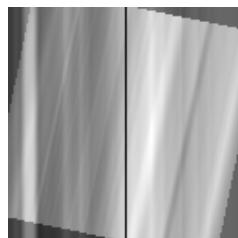
(d) Raw
Frame (n): 0
View: ZY



(e) Reconstructed
Frame (n): 0
View: ZY



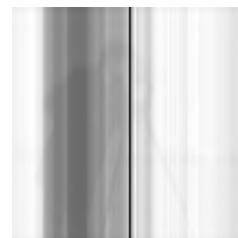
(f) Raw
Frame (n): 1
View: XY



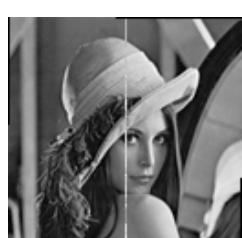
(g) Reconstructed
Frame (n): 1
View: XY



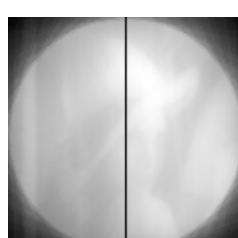
(h) Raw
Frame (n): 1
View: ZY



(i) Reconstructed
Frame (n): 1
View: ZY



(j) Raw
Frame (n): 31
View: XY



(k) Reconstructed
Frame (n): 31
View: XY



(l) Raw
Frame (n): 31
View: ZY



(m) Reconstructed
Frame (n): 31
View: ZY

Fig. 6.7 a 3D test-volume of two orthogonal and different testcard images was used to verify the reconstructive capabilities of the proposed algorithm. The results are unfiltered for clarity of demonstrating the iterative reconstruction.



Fig. 6.8 Filtered reconstruction of a reference image using the new proposed algorithm, (XY)

In the standard approach for OPT reconstruction, the sinogram undergoes the inverse Radon transform, see Fig. 6.5a and then post-filtering, see Fig. 6.5b. This step is substituted for the proposed algorithm; in Fig. 6.9 the two techniques are compared for ideal conditions of smooth, predictable rotation. The proposed algorithm produces a faithful reconstruction on the original image, as shown in Fig. 6.8 Both techniques lose some of the original contrast of the object due to under-sampling of rotations. When taking the histogram of the absolute pixel-wise difference between the original source image to the images produced by the new algorithm and the Radon transform, see Fig. 6.10. The mean square errors (*MSE*, see Equation (6.30)) of the new algorithm and the Radon transform are 15.01 % and 14.84% respectively, see Fig. 6.10 for a histogram of a pixel-wise comparison. This suggests that the new algorithm is producing an accurate reconstruction of the object, similar to the standard Radon transform.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (6.30)$$

The more challenging case of a sample drifting, with a constant velocity, systematically along the X axis was then considered; this produced a helical path of a single fiducial within the sample, see Fig. 6.11b. In Fig. 6.12a, the Radon transform entirely fails to produce a recognisable reproduction of the test image with the addition of a slight helicity to the rotation. The proposed algorithm produces an equivalent result to that of a sample rotating without any systematic drift, see Fig. 6.5b. In Fig. 6.13 the respective images from each algorithm were compared, as before, while the helical shift was incremented. See Fig. 6.11b for a sinogram of a sample whereby

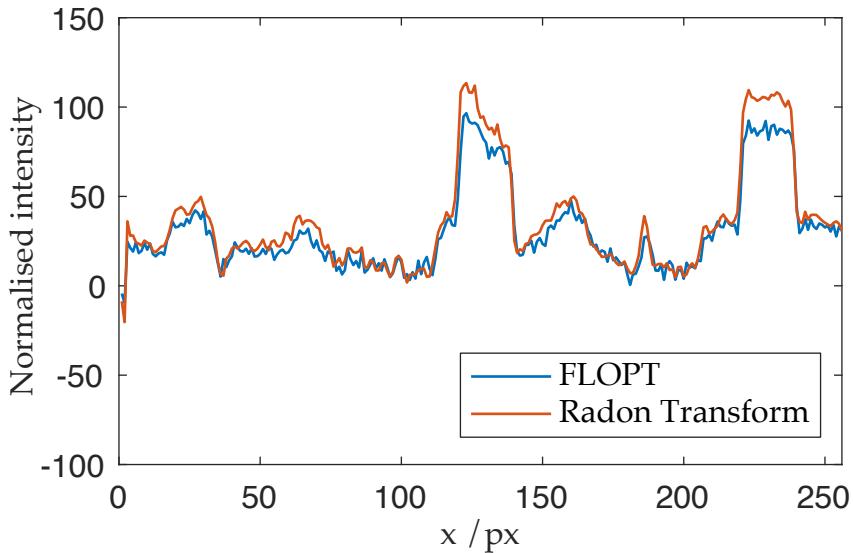


Fig. 6.9 Line profile comparison of the reconstruction of a reference image computationally rotated, projected and reconstructed using the standard Radon transform and the new proposed algorithm.

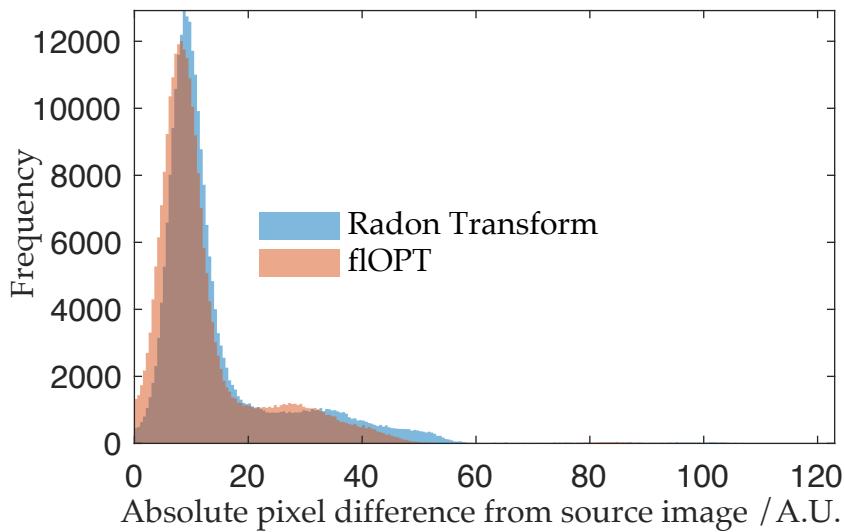
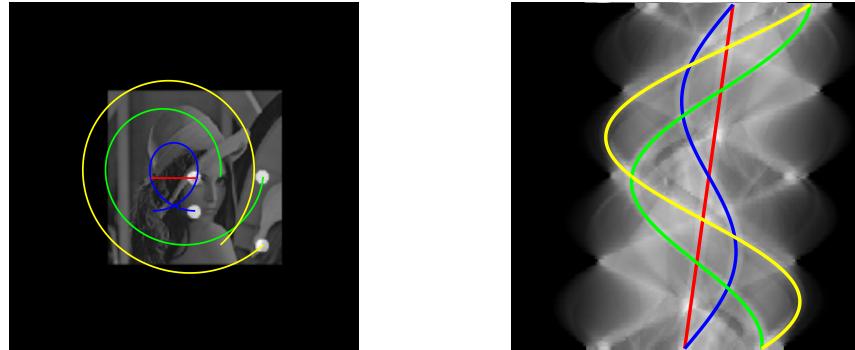


Fig. 6.10 Histogram of of pixel values compared between reconstructions using fLOPT and the Radon transform. The shift of the histogram to towards overall lower deviance from the source image suggests the fLOPT algorithm out performs the Radon transform



(a) Top down views (XY) of the source image with the fiducial paths marked.

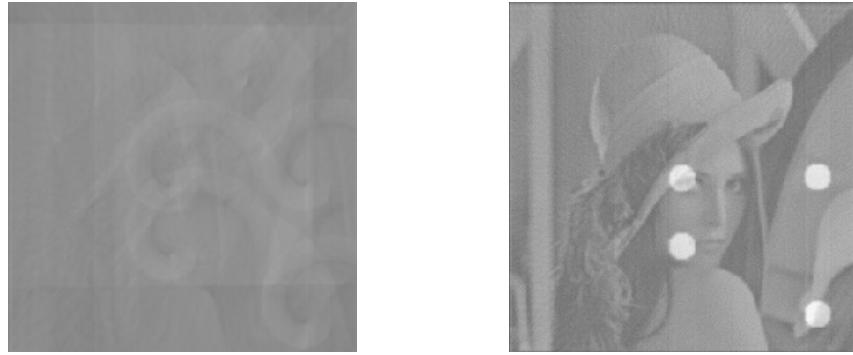
(b) Sinogram (XN) of a sample whose axis of rotation has a systematic drift

Fig. 6.11 Comparison of the two reconstructions under sample imaging with a systematic drift, in 3D though represented here in 2D.

a helical shift has been induced. When using correlation as a metric of reproduction quality, at zero helicity, the new algorithm fairs slightly worse at 94% correlation compared to the Radon transform at 96%. As expected, the Radon transform rapidly deteriorates once a systematic drift is applied; where-as the new algorithm maintains quality of reconstruction, see Fig. 6.13.

6.4.1 Recovery of \mathbf{R} and \mathbf{T} using matrix decomposition

To quantitatively verify that the matrix decomposition technique was valid and robust, the accuracy of the reproduction of \mathbf{R} and \mathbf{T} was tested directly. The original \mathbf{R} and \mathbf{T} matrices were computed and compared to \mathbf{R} and \mathbf{T} generated from matrix decomposition, this absolute difference was computed element-wise in each matrix and then an average for each matrix was taken. Overall the worst case scenario produced a percentage error of 2 % see Fig. 6.14 for full statistics. The accuracy of the calculated \mathbf{R} and \mathbf{T} did deteriorate when adding in additional degrees of combined movement, but with no correlation between the degree of helicity and the error produced. Consistently the translational matrix (\mathbf{T}) was more accurately reproduced, this is likely due to there being fewer of degrees of freedom for errors to spread over.



(a) Unfiltered reconstruction using a Radon transform, (XY)

(b) Filtered reconstruction using the new algorithm, (XY)

Fig. 6.12 Comparison of the two reconstructions under sample imaging with a systematic drift, in 3D though represented here in 2D.

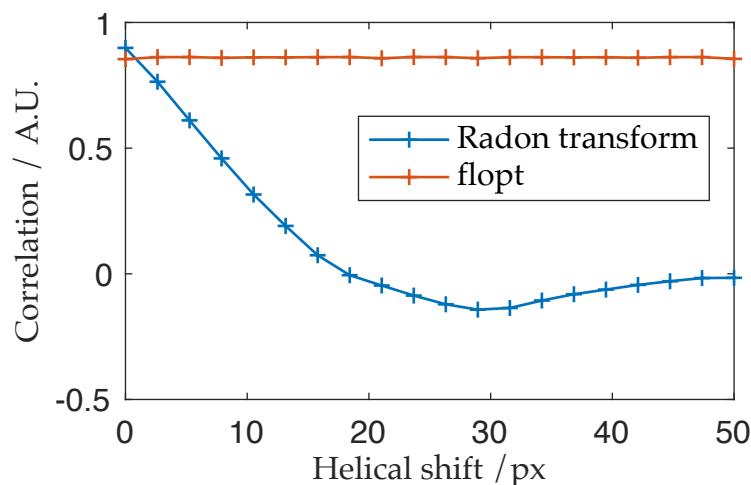


Fig. 6.13 2D correlation of the source image shows that floPT does not degrade under systematic drift compared to Radon transforms.

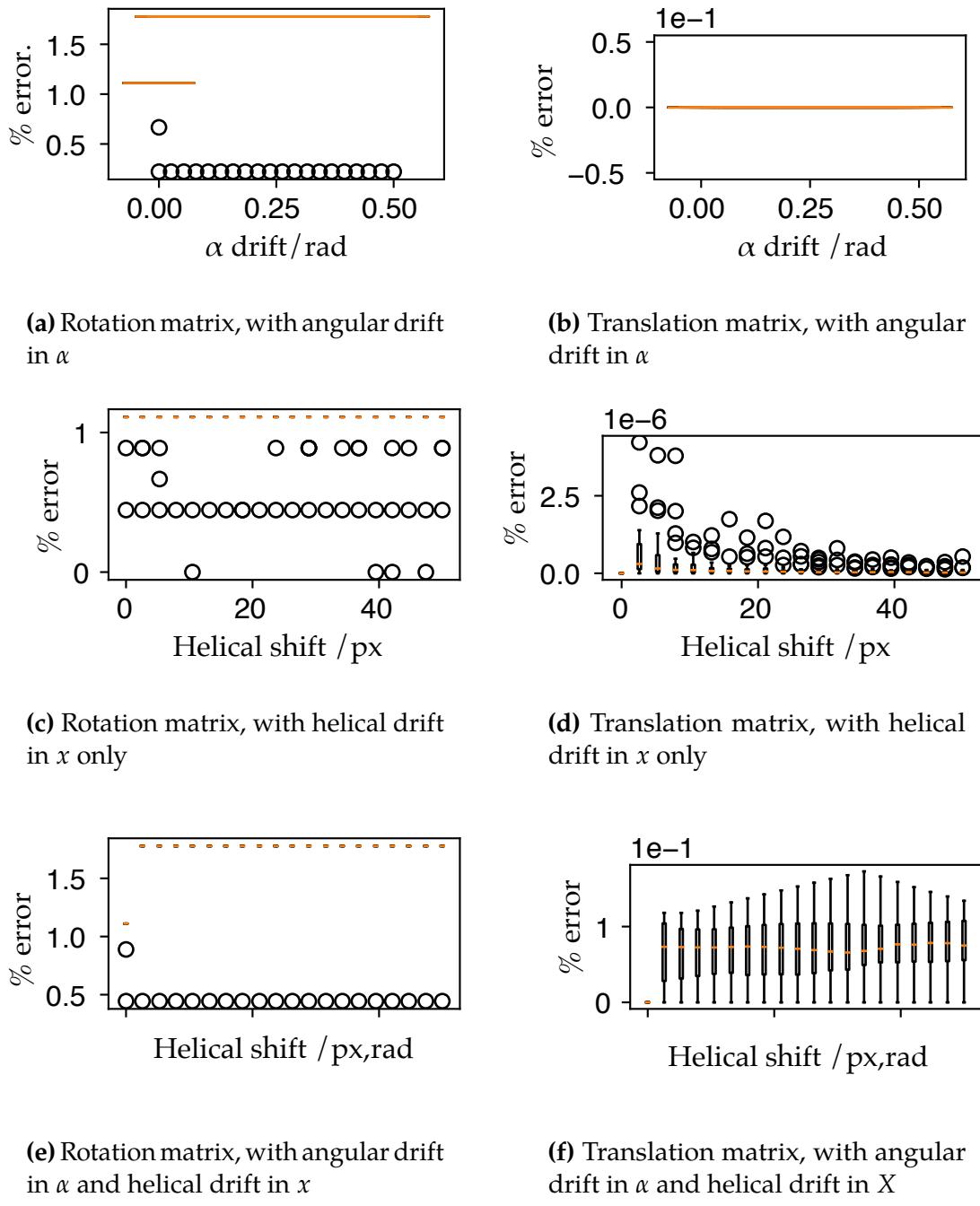


Fig. 6.14 Box plots demonstrating that the rotational and translations matrices can be recovered accurately from fiducial marker positions. Panels 6.14a and 6.14b introduce an angular drift during rotation, to an observer at the detector this would appear as a tip of the sample towards them, causing precession. Panels 6.14c and 6.14d introduce a lateral drift in X causing a helical path to be drawn out. Panels 6.14e and 6.14f combine the two effects. In all cases the percentage error introduced by the the addition of undesirable additional movements was on the order of $< 2\%$.

6.5 Discussion

A new algorithm for reconstructing OPT data has been demonstrated. The new algorithm uses multiple fiducials to recover the matrix which describes the rotation and translation of the sample. The quality of the reconstructions when compared to a standard Radon transform shows a slight improvement, with a great effect when a systematic drift is introduced. To measure the accuracy of the decomposition of \mathbf{F} into \mathbf{R} and \mathbf{T} they were compared to the ground truth matrices. The element-wise absolute difference $\left(\frac{x-y}{2(x+y)}\right)$ of each matrix was averaged across the matrix for \mathbf{R} and \mathbf{T} . In the worst case scenario a maximum of 2 % average absolute difference was found between ground truth and recovered matrices, suggesting the technique is robust to all forms of drift and general instability. Such an algorithm could be used to help in minimising ghosting effects seen in real samples; particularly in samples where slipping is likely to occur such as in gels or in cheaper OPT systems which tend to be more mechanically unstable and imprecise.

6.6 Future work

6.6.1 Bead Tracking

The work presented here is, so far, is a proof-of-concept demonstrated by simulation and reconstruction from ground-truth testcard objectives, and requires several further steps in order to apply it to real OPT data. Firstly, a bead-tracking algorithm[1] will need to be created to track multiple beads concurrently in an image series. A sensible approach would be to have a user select the fiducials in the image on the first frame and template match in a small window around the selection; this is similar to the algorithm described in Chapter ???. Template matching is robust to occlusions provided the fiducial is not fully eclipsed. If two fiducials occlude each other however, this algorithm may switch their identities or both tracking windows may follow one bead. This is a common problem in particle tracking algorithms, but is solved by using a weighted likelihood based on momentum[2].

The likelihood of a bead occlusion occurring will increase with the introduction of additional beads into the sample. As such occluded beads may need to be omitted. Egregious outliers may be found by tracking a *confidence* estimator as the bead rotates as in Fig. 6.15.

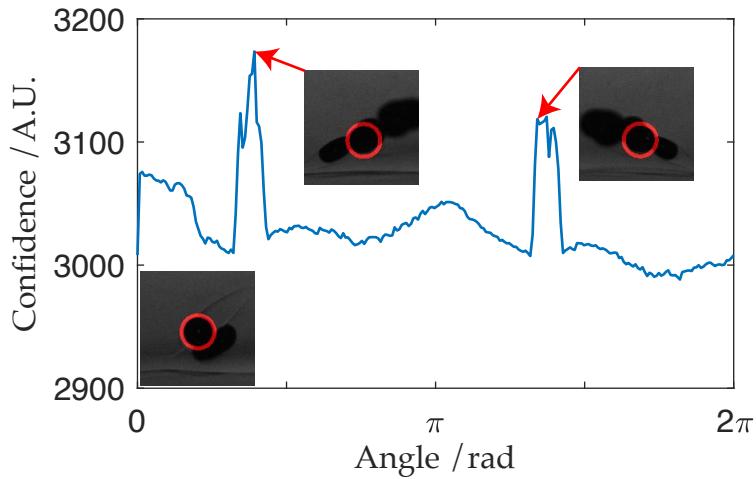


Fig. 6.15 A plot of the confidence value of a single fiducial bead being tracked *in vivo*. Sharp changes in the confidence value occur when the fiducial bead is occluded. The image at the origin shows the fiducial being tracked well in the first frame. (Images courtesy of Pedro Vallejo)

6.6.2 Multiple views tracking

The theory backing the proposed algorithm relies on triangulation between two view points. In this work the two view points refer to the image at frames n and $n + 1$. However, it is possible to use three separate views (frames n , $n + 1$ and $n + 2$) to reconstruct a scene, one such approach being quaternion tensors. Working with tensors is computationally and mathematically more challenging, but a future iteration of the algorithm presented here may benefit from using three views to provide a more accurate transformation matrix. Beyond three views there currently is no mathematical framework at present for four or more views. If such tools did exist, it may be possible to make the algorithm described above as non-iterative and essentially a single shot reconstruction from pixels to voxels.

6.6.3 Fiducial free reconstruction

In computer vision scenes often do not contain known fiducial marks and so such marks are found between views. To find such correspondences, points with similar local texture are found and matched in between each image. This technique is only valid for views with small angles between them, as would be found in OPT. A similar method could be introduced into the algorithm presented here, as each image should have sufficient texture, particularly when using tOPT.

The following chapter will move from improvements in registration using projective matrices and into improves in resolution using on-camera slit-scanning.

References

1. Crocker, J., Crocker, J. & Grier, D. Methods of Digital Video Microscopy for Colloidal Studies. *Journal of Colloid and Interface Science* **179**, 298–310 (1996).
2. Chenouard, N., Bloch, I. & Olivo-Marin, J.-C. Multiple Hypothesis Tracking for Cluttered Biological Image Sequences. *IEEE transactions on pattern analysis and machine intelligence*. pmid: 23689865 (2013).