**An Iterative Approach for local 3D-Reconstruction of Non-Rigid Movements of Heart Region From Many Angiographic Views**

erschien in

```
@inproceedings{Bouattour2003AIA,
  crossref = {Ertl2003VMA},
  pages = {183-190},
  author = {Bouattour, Sahla and Heigl, Benno and Paulus, Dietrich
            and Hornegger, Joachim},
  title = {An Iterative Approach for local 3D-Reconstruction of
           Non-Rigid Movements of Heart Region From Many Angiographic Views},
  category = {3D Rekonstruktion, 3D+t Herzrekonstruktion}
}

@proceedings{Ertl2003VMA,
  publisher = {IOS Verlag},
  year = {2003},
  editor = {Ertl, Thomas and Girod, B. and Greiner, G. and Niemann, Heinrich
            and Seidel, H.-P. and Steinbach, E. and Westermann, R.},
  title = {Vision Modeling and Visualization}
}
```

# An Iterative Approach for local 3D-Reconstruction of Non-Rigid Movements of Heart Region From Many Angiographic Views

Sahla Bouattour (1) Benno Heigl(2) Dietrich Paulus (1)
Joachim Hornegger(2)

Computational Visualistics, University Koblenz-Landau, Universitätsstraße 1, 56070 Koblenz, Germany
Medical Engineering Group, Siemens AG Erlangen-Forchheim,Germany
Email:bouattour@uni-koblenz.de

## Abstract

In this paper we examine the reconstruction of a *local* region of a non-rigidly moving heart with unknown motion, using methods of reconstruction of rigid objects and without using (electrocardiogram) ECG-information. The reconstruction procedure chosen is based on the well-known Feldkamp filtered-backprojection algorithm. It reconstructs the volume of a rigid object from a reduced set of projections and their corresponding *known* projection matrices. The basic idea of our approach is to assume *rigidity* for the local region of interest (ROI), and to *fixate* it by compensating its unknown motion by a camera movement accounting for it. This is achieved using an iterative 2D-3D *intensity-based* registration approach.

## 1 Introduction

Heart attacks are one major cause of death for humans today. The early detection and correction of aberrations of coronary vessels is therefore of highest medical importance to avoid occlusions or to minimize the damage of the heart after the heart attack. Image data for subsequent visualization and analysis are usually acquired with cardiac C-arm devices as shown in Figure 1. These systems allow the acquisition of image sequences of the beating heart (Figure 2). Usually 30 frames per second are captured. High-end systems provide even two views from different directions simultaneously. Two views support the treating physician to conclude the real 3D–structure of the vessel tree. It is obvious that the 3D–reconstruction of the vessel or at least the reconstruction of the aberration will lead to an improvement of the workflow and of the treatment.

Figure 1: An AXIOM Artis DFC C-arm

Figure 2: An image of a beating heart showing coronary vessels

Due to the non–rigid motion of the beating heart, the 3D–reconstruction is difficult. No algorithm is known in the literature that allows the computation of the volume of a moving object using images acquired by a C-arm device. The reconstruction comprises usually two main problems: the extraction and tracking of corresponding points and the computation of the structure from these points. The segmentation of coronary vessels, and precisely the extraction of centerlines is addressed by many researchers within different contexts other than heart segmentation such as retina vessel segmentation or road segmentation [7, 1]. However, the reconstruction problem is mostly reduced to the reconstruction from two views of the same ECG-state [2, 1]. The rest of the views acquired by the C-arm device, is not used for reconstruction.

In this paper we introduce a new technique that will at least lead to a local 3D–reconstruction of a stenosis from many angiographic views. Section 2 states the problem. Sections 3 and 4 explain the chosen procedure. Section 5 presents preliminary series of experiments on generic data, aiming to prove the concept. Finally section 6 summarizes the

results and describes future work.

## 2 Problem Statement

In the considered angiographic setting the camera is part of a C-arm which is rotated around the patient with known motion parameters. Let $\boldsymbol{C}_i$ be the center of camera at time $i$. The acquisition geometry is characterized by a projective mapping $\boldsymbol{P}_i$ for the $i^{\text{th}}$ image. The projection matrices are computed during a geometrical calibration step (see [3]). The heart is a *deformable* object and is *moving* with unknown motion model and unknown parameters. For visualization purposes let the geometric point $\boldsymbol{w}_i$ denote a heart position at time $i$. The projection of $\boldsymbol{w}_i$ through the camera $\boldsymbol{C}_i$ is observed on frame $F_i$ and is referred to by $\boldsymbol{q}_i$. The observed point $\boldsymbol{q}_{i+1}$ in frame $F_{i+1}$ arises from the simultaneous movement of the camera from $\boldsymbol{C}_i$ to $\boldsymbol{C}_{i+1}$ and the unknown non-rigid heart movement from $\boldsymbol{w}_i$ to $\boldsymbol{w}_{i+1}$. Figure 3 illustrates the scenario and enhances particularly the movement of the heart from one frame to the next.

As already mentioned the diagnosis of stenosis does not require the reconstruction of the whole heart. The small region containing just the malformation would be sufficient for the physician to extract the needed information such as the degree of the stenosis. In the following we will concentrate on this issue. It is important to note that we will keep the heart locally rigid[1] which means that we will assume that the 3D–motion of the ROI is the same for all points of the ROI. Let $\boldsymbol{H}^{w}_{(i,i+1)}$ be the transformation matrix describing the 3D–motion of the ROI, formulated in homogenous coordinates. The only constraint we have for $\boldsymbol{H}^{w}_{(i,i+1)}$ is that it is a $4 \times 4$ (non-singular) homogenous matrix; we will not assume any concrete model for it. In this case we can write the following equation:

$$\tilde{\boldsymbol{q}}_{i+1} \backsimeq \boldsymbol{P}_{i+1}\boldsymbol{H}^{w}_{(i,i+1)}\tilde{\boldsymbol{w}}_i \ , \qquad (1)$$

where $\tilde{\boldsymbol{q}}_{i+1}$ and $\tilde{\boldsymbol{w}}_i$ refer to the homogenous coordinates of the $(i+1)^{\text{st}}$ frame point and the $i^{\text{th}}$ world point. The projection matrix $\boldsymbol{P}_{i+1}$ is a $3 \times 4$ homogenous matrix specifying the current camera position at time $i+1$. The equality is valid up to a scale factor, due to the projective mapping. Equation 1 establishes the relationship between the 3D–point $\boldsymbol{w}_i$ at time $i$ and its projection after motion at time $i+1$. This relationship remains if we fix the 3D–point in a *reference position* $\boldsymbol{w}_0$ and express the relationship with respect to this reference:

$$\tilde{\boldsymbol{q}}_{i+1} \backsimeq \boldsymbol{P}_{i+1}\boldsymbol{H}^{w}_{(0,i+1)}\tilde{\boldsymbol{w}}_0 \ . \qquad (2)$$

Assuming rigidity for the ROI, equation 2 can be generalized to $m$ points:

$$\tilde{\boldsymbol{q}}^{j}_{i} \backsimeq \boldsymbol{P}_i\boldsymbol{H}^{w}_{(0,i)}\tilde{\boldsymbol{w}}^{j}_0; \ j = 1, \cdots, m, \ i = 1, \cdots, n. \qquad (3)$$

The gain of this formulation is that we are now able to update the projection matrices from each viewpoint ($\boldsymbol{P}_i\boldsymbol{H}^{w}_{(0,i)}$) in such a way that we have a non-moving ROI in 3D. That is we are able to compensate the motion of the ROI by integrating its movement into the projection matrix. The ROI is thus fixed and it is possible to use known reconstruction algorithms based on the updated projection matrices such as the filtered-backprojection algorithm [8] in order to get the volume of the stenosis. To the best of our knowledge the idea of *fixation* of the ROI and the usage of rigid reconstruction algorithm from many angiographic views for local-heart reconstruction has not been yet tackled in the literature.

However, to apply this idea we have to solve the following three open problems: How to determine the reference pose of the ROI in the space? How to find the transformations $\boldsymbol{H}$ describing heart movement in order to update to projection matrices? How many points $m$ do we need to be able to solve this problem? Section 3 describes the approach for reconstructing the reference pose. Section 4 addresses the problem of updating the projection matrices and the needed number of points.

## 3 Computing Reference Pose

For a certain number of points $m \in \mathbb{N}$, a reference position will be computed from two views using the *optimal triangulation method* described in [5, p. 305], where a set of $m$ corresponding points in 2 views and their corresponding projection matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are given. In order to apply this method the fundamental matrix $\boldsymbol{F}$ has to be computed. It is obvious that $\boldsymbol{F}$ should be computed from the given projection matrices and not using

[1] The validity of this assumption will have to be justified by experiments.

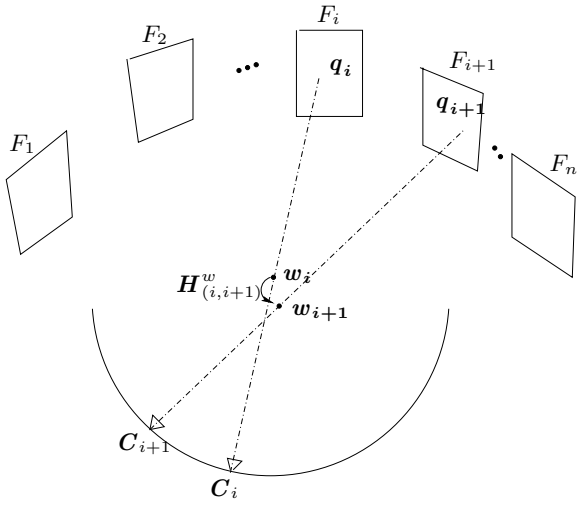


Figure 3: Illustration of the simultaneous heart and camera motion along an angiographic sequence.

the point correspondences (i.e. using the 8-point algorithm). The reason is that the later computation will give a fundamental matrix, for which the given correspondences are ideal, and not necessarily the matrix that reflects the *known* configurations of the cameras given by the projection matrices. The fundamental matrix between two frames is derived algebraically from the following equation

$$\boldsymbol{F} = [\boldsymbol{e}_2]_\times \boldsymbol{P}_2 \boldsymbol{P}_1^+ \ , \tag{4}$$

where the $[\boldsymbol{a}]_\times$ denotes the cross-product matrix with the vector $\boldsymbol{a}$, the epipole $\boldsymbol{e}_2$ in the second view is the projection of $\boldsymbol{C}_1$, the center of the first camera via $\boldsymbol{P}_2$. The advantage of this method in comparison with other linear or iterative solutions that minimize the geometric error is that it is projective–invariant [5]. Due to the motion of the heart this step results in 3D points whose projections do not suit perfectly to the given corresponding image-points, but this is not necessary, because we just need *reference points*, and thus we can chose frames which do not correspond to the same ECG-state, as long as the structure of the ROI in both images is similar.

# 4 Update Projection Matrices

In this section we will review several methods that can be adopted for updating the projection matrices. We will specify the needed number of points to apply each of them, and will show their limitation for our application. In Section 4.4 we propose an iterative 2D-3D intensity-based registration approach for updating the projection matrices at each view while keeping the ROI fixated at the reference position.

## 4.1 Camera Calibration

Once we have the set of 3D reference points $\tilde{\boldsymbol{w}}_0^j; \quad j = 1, \ldots, m$, and for each frame we have the set of the corresponding 2D points $\tilde{\boldsymbol{q}}_i^j, i = 1, \cdots, n$, we could directly estimate the projection matrices that map the reference points to the measured ones. Different approaches exist that attend to solve this problem. In [5] e.g the projection matrices are estimated directly from the 2D-3D correspondences using linear method followed by a nonlinear iterative optimization procedure. The minimal number $m$ of points which is needed would be *six*. A basic drawback of this method is that it does not exploit a-priori knowledge optimally. Indeed, it assumes no knowledge of the known projection matrices $\boldsymbol{P}_i$ and does not take them into consideration.

## 4.2 Camera Pose Estimation

As we dispose of the intrinsic parameters of the camera from the equations, our problem is actually reduced to a *camera pose determination* problem,

also known as *space resection* in photogrammetry. In [9] e.g. the coordinates of the given world coordinate points are determined in a camera centered framework by solving a set of differential equations. The rigid 3D Motion that best align the 3D points in both coordinate systems is estimated afterwards, using e.g quaternions or SVD. With this non-linear method, the computation of camera pose is uniquely determined by just *four* points.

### 4.3 Direct Estimation of Heart Motion

Another way to update the projection matrices is to directly estimate the heart motion. Equation 6 can be reformulated linearly with respect to the components of $\boldsymbol{H}^{w}_{(0,i)}$ as

$$\boldsymbol{B}_i \boldsymbol{h}_{(0,i)} = \boldsymbol{Q}_i \quad , \tag{5}$$

where $\boldsymbol{h}_{(0,i)}$ is the vector containing the entries of $\boldsymbol{H}^{w}_{(0,i)}$, $\boldsymbol{Q}_i$ is a $3m$ column vector obtained by stacking the columns of $\tilde{\boldsymbol{q}}^j_i$ one above the other and $\boldsymbol{B}_i$ is $3m \times 16$ matrix containing the following three rows for each 3D/2D correspondence:

$$\begin{pmatrix} (\boldsymbol{w}^j_{01}\boldsymbol{P}_i) & (\boldsymbol{w}^j_{02}\boldsymbol{P}_i) & (\boldsymbol{w}^j_{03}\boldsymbol{P}_i) & (\boldsymbol{P}_i) \end{pmatrix}$$

Since the Matrix $\boldsymbol{H}^{w}_{(0,i)}$ has 16 entries, and ignoring scale (15 degrees of freedom), it is necessary to have 15 equations to solve for $\boldsymbol{h}_{(0,i)}$. Each point correspondence leads to two non–redundant equations; so the minimal solution can be computed by *eight* such correspondences.

**Discussion:** The mentioned methods, however, will fail to find a unique and stable solution for certain degenerate configurations. According to [5], the camera cannot be obtained uniquely from images of points, that lie on the union of a plane and a single line containing the camera center. Unfortunately, the centerline of a stenosis can be approximately considered as a planar curve in space; therefore we are handling data which is *close* to a *degenerate configuration* and we expect a poor estimate of the updated matrix, even if we increase $m$.[2]

In addition to that, it is important to notice that the approaches mentioned above focus on the reconstruction problem given known corresponding *points*. It should also be clear that finding correspondences remains in this case one of the *the critical* problems, because the accuracy of the update of projection matrices directly depends on the *number* of matches and the *accuracy* of finding corresponding points. The segmentation and extraction of vessels' centerline is a major subproblem in its own right and is still challenging researchers. Moreover the tracking of centerline points over the C-arm sequence is very awkward, due to the continuity along the ROI and the absence of leading structure elements such as corners. Even if we assume that all these problems are solved in an optimal way, several problems remain to be handled with care such as missing correspondences, outliers or picturing the stenosis along its transversal view.

### 4.4 Direct Estimation from Intensities

Due to these reasons we propose to use a direct algorithm that reconstructs from the intensity data itself, rather than from computed correspondences. We propose to estimate the motion of the ROI (still assuming rigidity) using a 3D-2D Registration algorithm based on intensity images. Penney designed in [6] such an algorithm to obtain the pose of a CT volume with respect to a single fluoroscopy image. The basic idea is to produce digitally reconstructed radiographs (DRR's) from the CT volume, which are compared to the fluoroscopy image using a similarity measure. The pose of the volume in space consisting of a 3D-rotation (3 angles) and a 3D-translation (3 degree of freedom) is determined by starting an optimization process in a six-dimensional search space. The DRR is built by projecting the volume using the known projection matrix.

In order to apply this algorithm in our case we have first to extend it to apply to $n$ images; as to say we estimate the rigid motion parameters of the volume such that its DRR's along the different camera poses of the C-arm is most similar to the given image sequence. This can be achieved by extending the optimization function to be minimized to the sum of similarity measures over the different $n$ views.

Second we need a volume of the ROI that will be used for registration. As we have a moving object, a 3D-reconstruction using the filtered backprojection [8] approach with the given projection matrices would lead to bleary volume, so registration

[2] Our experimental results on generic data confirmed this assumption.

will not lead to reasonable results. Besides we want to hold somehow the ROI fix in order to be able to get a good reconstruction of it. In order to solve this problem, we propose to segment and track *one* single point ($m = 1$) in the ROI, e.g the centerpoint of the stenosis, reconstruct a reference point using the method described in section 3 and perform a first update for the projection matrices as following:

$$\tilde{\boldsymbol{q}}_i \backsimeq \boldsymbol{h}^w_{(0,i)} \boldsymbol{P}_i \tilde{\boldsymbol{w}}_0;\ i = 1, \cdots, n. \qquad (6)$$

where $\boldsymbol{h}^w_{(0,i)}$ is a $3 \times 3$ homogenous matrix containing the 2D displacement vector from the projection of the reference point wrt. $\boldsymbol{P}_i$ and the detected corresponding point in the considered view. This results in projection matrices $\boldsymbol{P}^0_i = \boldsymbol{h}^w_{(0,i)} \boldsymbol{P}_i$ which hold one point of the ROI fixed in the reference pose. These matrices are used to perform a first volume reconstruction $\boldsymbol{V}^0$.

The volume $\boldsymbol{V}^0$ is then iteratively registered with the given image sequence until stabilization of the estimated rotation and translation. In each iteration $k$ the volume $\boldsymbol{V}^{k-1}$ and its corresponding matrices $\boldsymbol{P}^{k-1}_i$ are used for registration of the local ROI. the estimated Parameters $\boldsymbol{R}^k$ and $\boldsymbol{T}^k$ are used for updating the projection matrices:

$$\boldsymbol{P}^k_i = \boldsymbol{P}^{k-1}_i \begin{pmatrix} \boldsymbol{R}^k & \boldsymbol{T}^k \\ 000 & 1 \end{pmatrix}. \qquad (7)$$

# 5 Experiments

In order to check the effectiveness of the proposition made in section 4, we used a series of 133 images taken by Siemens-C-arm during a rotation of $166^\circ$. The images are gray-level X-ray images of $512 \times 512$ pixels; they represent a phantom experiment, showing a vessel similar structure with a stenosis made by knead. Figure 4 shows some of the gained views.

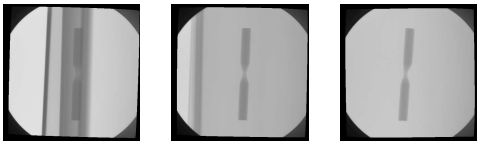

Figure 4: Sample images from the angiographic sequence

## 5.1 Preprocessing

To facilitate the detection and tracking task of the single point, we performed several preprocessing steps. We first subtracted the mask images (images without object) from the X-ray images in order to get rid of the background. The images are then binarized using discriminant threshold analysis. On the binary images we performed a morphological centerline extraction algorithm and filtered, when necessary, the skeleton to delete small artifacts. The extracted centerlines have pixel accuracy. The results of these steps are shown for one image in figure 5. The positions of the middle point of the structure in the phantom stenosis were segmented automatically (by looking along the skeleton for the point having the smallest radius) and corrected manually in an subsequent offline step, where required.

## 5.2 Phantom Data

The first experiment is done with a non-moving stenosis object. The goal is to check the accuracy of the segmentation, the reconstruction of the reference position and the estimation of the 2D Displacement if we do not have any motion. The result of this experiment will be taken as a referring point for the amount of error induced by these steps. Figure 6 show the ideal volume reconstructed by the filtered-backprojection algorithm using the projection matrices delivered by the calibrated C-arm, as well as the initial volume $\boldsymbol{V}^0$ reconstructed as described in section 4.4. The artifacts in the second volume are due to the errors in the initial update of the projection matrices, which are due to incorrect segmentation and detection of the point of interest.

In the second series of experiments we introduced generic 2D motion, which could be due to an unknown 3D motion. A simple translational motion is introduced in the 2D images by shifting periodically the pixels in x-direction with the following vectors: 20 30 40 50 50 40 30 20 over the image sequence. The points of interest are also shifted for each view with the corresponding vector. Figure 7 (Top) shows the volume resulting from the introduction of the motion. As expected the reconstruction of the moving phantom led to a bleary volume, in which we can even recognize the introduced generic motion. In the bottom of Figure 7 we observe the initial volume $\boldsymbol{V}^0$ reconstructed by

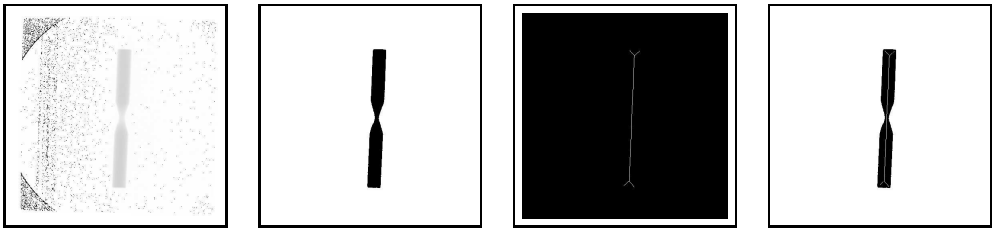

Figure 5: Preprocessing steps: from left to right: subtracted image, binarization, extraction of centerline, combined binary image with the skeleton for visualization purpose.

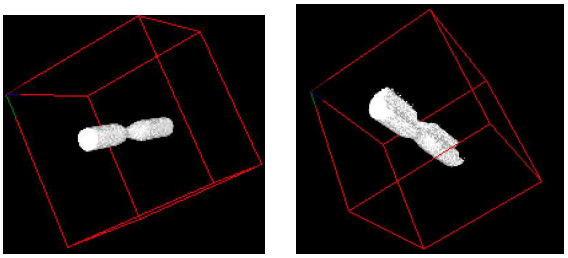

Figure 6: Left: Ideal volume of static-stenosis. Right: Volume $\boldsymbol{V}^0$ of non-moving stenosis

accounting for the 2D displacement of the middle point of the stenosis. Motion artifacts are mostly eliminated and we expect that the reconstructed volume will improve by further iterations. As the goal is to obtain a visually appealing reconstruction and to provide plausible 3D impression that will facilitate medical treatment, judgement of the results solely based on visual impression is acceptable.

We performed third series of experiments in which we reconstructed a really moving object with unknown 3D motion. Figure 8 illustrates the used object consisting in an electrical connector. We performed basically the same preprocessing steps as in section 5.1 for extracting and tracking the upper point of the head of the connector that we considered as the ROI. The results of the reconstruction of the moving object and the initial volume with fixation of the upper point are shown in Figure 9. On the top of the figure, the structure of the connector is lost due to the 3D motion. The fixation of the upper point led to stabilization in the structure ( Figure 9, bottom). Basically the reconstruction is better the nearer we are wrt. the fixated point.

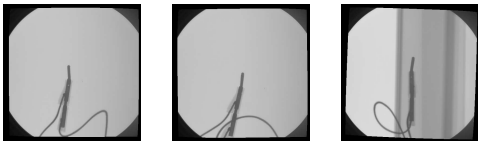

Figure 8: Sample images of the moved connector in 3D

From this initial results we can conclude that segmentation and tracking can have a big influence on the quality of the initial reconstruction. The fixation of a single point leads to an initial volume, that contain sufficient information around the ROI to enable registration. Further experiments will verify that even with the presence of segmentation errors our registration procedure will converge.

### 5.3 Forthcoming Experiments

In the presence of more complex motion, the computation of the initial volume $\boldsymbol{V}^0$ can be improved by a better estimation of the 2D transformation $\boldsymbol{h}^w_{(0,i)}$. This can be achieved by using more than one point. Specifically we need *three* points, in order to be able the compute the gradient of the middle point and thus computing an affine transform for estimating $\boldsymbol{h}^w_{(0,i)}$ using the position of the points and their gradients.

In addition to that future work will concentrate

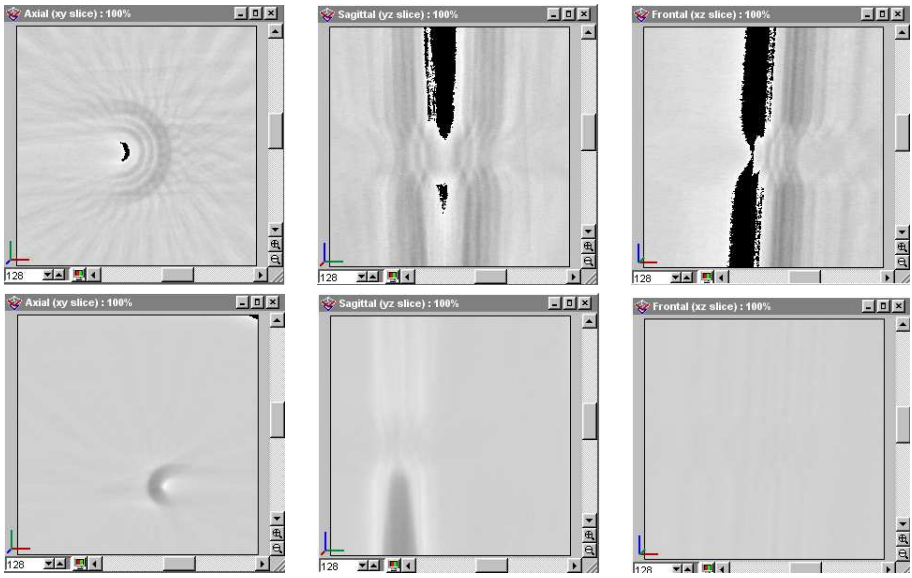


Figure 7: Axial, sagittal and frontal views of the volumes of the phantom stenosis with generic 2D motion. Top: Volume of moving stenosis. Bottom: Volume $\boldsymbol{V}^0$ resulting from fixation of the middle point of the stenosis.

on the experimental study of the convergence of the algorithm by following the initial reconstruction by a number of iterations using the registration algorithm with $n$ views. Besides, the registration algorithm has to be extended to apply for a local region and more experiments on real data have to be performed.

## 6 Conclusion

In this paper we examined the possibility to compensate for the motion of a local non-rigid heart region by integrating it into camera motion, while keeping the region of interest fixed. This has the advantage to use known algorithms of volume reconstruction for fixed objects. We discussed several methods to update the projection matrices and showed their strength and limitation. We proposed an iterative intensity-based registration algorithm for integrating heart motion into camera motion. The initial volume for registration is gained by tracking and fixating a single point of interest. We presented several series of experiments on phantom data that prove the effectiveness of the idea of fixation. Future Work will concentrate on the experimental proof of convergence of the algorithm and on more rigorous experiments using real data.

## References

[1] C. Blondel, R. Vaillant, F. Devernay, G. Malandain, and N. Ayache. Automatic trinocular 3d reconstruction of coronary artery centerlines from rotational x-ray angiography. In *CARS 2002 Proc.*, Paris, June 2002. Springer Publishers, Heidelberg.

[2] C. Canero, P. Radeva, R. Toledo, J.J. Villanueva, and J. Mauri. 3d curve reconstruction by biplane snakes. *ICPR 00*, 4:4563–4567, September 2000.

[3] O. Faugeras and Q.Luong. *The Geometry of Multiple Images*. MIT Press, 2001.

[4] E. Grossmann and J. Santos-Vistor. Uncertainty analysis of 3d reconstruction from uncalibrated views. *Image and Vision Computing*, 28:685–696, 2000.

[5] R. Hartley and A. Zisserman. *Multiple View Geometry in computer vision*. Cambridge university press, 2000.

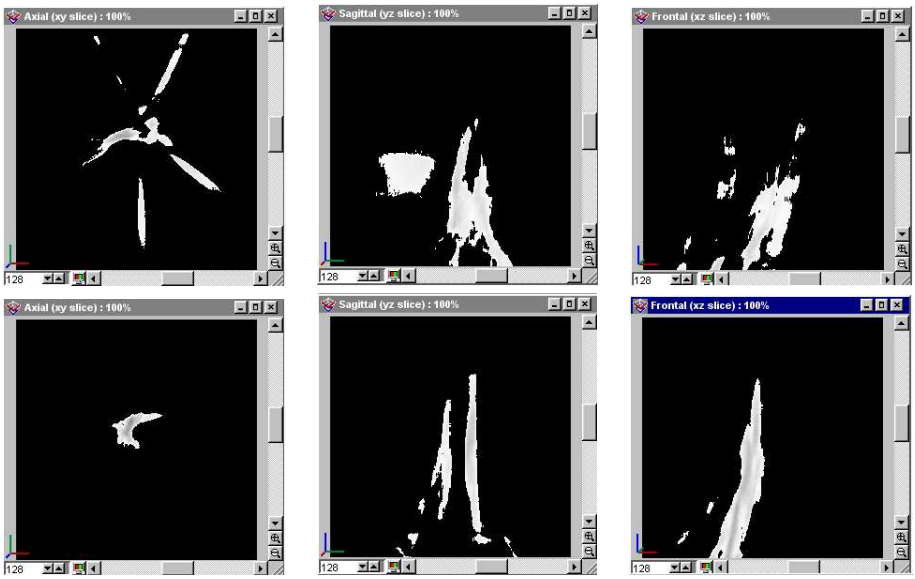


Figure 9: Axial, sagittal and frontal views of the volumes of an electrical connector with unknown 3D motion. Top: Cloudy volume of the connector which result from the direct reconstruction from the original projection matrices. Bottom: Volume $\boldsymbol{V}^0$ resulting from fixation of the upper point of the head of the connector.

[6] G. P. Penney. Registeration of tomographic images to x-ray projections for use in image guided interventions. Technical report, Computational Imaging science Group, Division ofRadiological Sciences and Medical Engineering. King's College London, December 1999.

[7] C. Steger. An unbiased detector of curvilinear structures. *PAMI98*, 1998.

[8] K. Wiesent, K. Barth, N. Navab, P. Durlak, T.Brunner, O. Schuetz, and W. Seissler. Enhanced 3-d-reconstruction algorithm for C-arm systems suitable for interventional procedures. *IEEE Transactions on Medical Imaging*, 19:391–403, May 2000.

[9] L. Zhi and J. Tang. A complete lienar 4-point algorithm for camera pose determination. *MM Reserach Reprints*, 21:239–249, December 2002.