NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Faculty of Computer Science
Bachelor’s Programme “Applied Mathematics and Informatics”

UDC XXXXXXX

Research Project Report on the Topic:

Heart Digital Twin and Predictive Analytics

Fulfilled by:

Student of the group БПАД211
Sergazin Iskander

(signature) (date)

Assessed by the Project Supervisor:

Lazarev Mikhail Vladimirovich
Research Fellow
Faculty of Computer Science, HSE University

(signature) (date)

Moscow 2023

Contents

Annotation

The treatment of cardiac illnesses remains a relevant problem in the field of applied medicine. Accurately visualizing the artery network is pivotal in developing a computer-based method that would assist the surgeon in treating heart conditions. This project uses 2D angiography videos, to test a method of automatic 3D reconstruction, drawing conclusions about the applicability of the algorithm for visualizing the cardiac vessel network.

Keywords

Computer Vision, Projective geometry, 3D Reconstruction, Coronary heart disease, X-ray angiography, image segmentation, COLMAP

1 Introduction

1.1 Subject area

Coronary heart disease (CAD) is a major challenge in the healthcare industry and is the leading cause of death globally. Often characterised by plaque buildup in the coronary arteries, CAD leads to an accummulation of plague within the arteries and a narrowing of the heart vessels. Narrow heart vessels reduce blood flow to the heart, which could lead to cardiac ischemia and later to a heart attack. Modern treatment of coronary heart disease involves an operation on the patients heart and a positioning of stents, which expand in the occluded regions restoring normal blood flow. X-ray angiography is a widely used non-invasive approach for visualizing the artery network. The 2D images produced by the imaging device are used by the cardiologist to subjectively decide which regions of the heart require stenting. Due to interobserver variability, it is essential develop a computer-based 3D reconstruction of the artery network, that would assist the surgeon in locating regions of occlusion and assessing stenosis. Such a method could improve clinical outcomes for patients with CAD.

Computer vision is a vastly growing sub-field of modern computer science, which focuses on developing algorithms and methods that would enable the computer to perceive visual information from the external environment. Modern techniques of computer vision have been successfully applied in a wide range of industries, including the healthcare industry, where it was used for guiding and supporting surgeries [7], assessing bone age [12] and creating digital twins [13]. Hypothetically, computer vision, in combination with methods of projective geometry could be used to create an accurate 3D reconstruction of the artery network. This project will use 2D projections taken from x-ray angiography imaging to explore whether an algorithm of automatic 3D reconstruction could be used for visualizing the artery network.

The project begins by reviewing methods of heart reconstruction; techniques for preprocessing angiography images; and algorithms of 3D reconstruction. The project then presents a robust methodology for selecting appropriate frames in angiography videos and a technique for accurately segmenting these frames to isolate the artery network. Thereafter, the project assesses the applicability COLMAP for reconstructing the artery network, formulating relevant hypothesis for why the algorithm fails. The final part of this project tests these hypothesis by following a general Structure-from-Motion framework to manually reconstruct 3D points in two neighbouring images.

2 Literature Review

2.1 Heart reconstruction techniques

Accurately reconstructing the cardiac vessel network is a relevant health problem, which could transform cardiac healthcare and improve clinical outcomes for patients with ischemic heart disease. Different approaches have been tried, for instance in [6], Iyer et al proposed a deep-learning based technique. Initially, the method generates synthetic vessels which are passed as input into a neural network, The NN has two output heads, one for estimating the radii and the other for computing the centerlines of each branch in the coronary tree, both MLP heads optimize a unique loss function, but share a convolutional Resnet backbone. The output of the network is a matrix, which encodes 3D coordinates of the centerline point and the corresponding radius. Notably, the article mentions AngionNet segmentation algorithm for automatically segmenting heart vessels, this algorithm could potentially be used in this project.

Alternatively, Franceska et al [4] proposed a multi-step approach utilizing NURBS for reconstruction. Once the vessel images have undergone preprocessing, 3D reconstruction begins by first defining a global reference system based on the C-arm device. Then initial geometrical estimates are extracted from the radiographic image header file, these geometrical estimates are optimized by reconstructing 3D landmark points (vessel nodes) and re-projecting these points back onto the image plane. A transformation (R,t) is then obtained by minimizing the sum of distance errors between the re-projected points against the original. The next step of the reconstruction involves the calculation of the 3D centreline, this is achieved by taking pairs of 2D centreline correspondences and forming a projective surface from each centreline. The 3D centreline is estimated by the intersection of these surfaces. The centreline then acts as a spine along which the lumen cross-sections are generated at equidistant points. NURBS are then used to functionally encode the contour of the lumen, using NURBS allows the modelling of complex curves and smooth shapes (which is necessary to show vessel deformations due to plague buildups). The continuous 3D luminal surface is then generated by using lofting. Overall the technique presented in this paper produces a really accurate algorithm which calculates vessel diameter measurements that are similar to OCT.

The next method by K.W.H Maas et al evaluates the use of NeRF [9] for 3D reconstruction. The method eliminates common limitations of traditional 3D reconstructions by eliminating the need for 2D segmentation. The article begins by identifying the main challenges of 3D reconstructing angiography images, these include: sparse views, limited number of angles, vessel sparsity, vessel overlap and poor vessel visibility. The authors then discuss modelling the scene as afunction which depends only on the desnity of the material through which the x-0ray passes the have adapted NeRF for x-ray reconstructions. The absorption equation had to also be modified and discretised, for the specific use-case. The final modification suggested by the authors involves the removal of positional encoding from the NeRF, since it impacts performance, while having limited impact on the quality of reconstruction. The results of the paper show that image reconstructions can be performed from more than 4 projections with an angle range exceeding 30 degrees. However, the main drawback of the method is that the reconstructions are time-intensive and this limits its application in applied medicine.

2.2 Data preparation methodologies

To accurately reconstruct an object in 3D, the data must be collected and preprocessed accordingly. In our case, images of the same heartrate phase need to be collected. Furthermore, the vessels in these images have to be saturated with medical contrast. To meet the aforementioned criteria different methods have been tried [1],[3],[8],[2]. These methods can be divided into two groups: ECG-based and non-ECG-based. For example, in [2] the authors present several ECG-based classification algorithms, while in [3] the authors use ECG information to train a neural network that can then process and classify images without relying on ECG. On the other hand, non-ECG methods focus on algorithms for cardiac segmentation. In [1], the authors use a top-hat operator and calculate the motion with integral scaling. Whereas [8] presents a semiautomatic detection method, based on selecting key points and then tracking the trajectories of these points, because in our study we could not retrieve ECG data an algorithmic approach similar to that of Blondel et al was used.

2.3 Algorithms for 3D reconstruction

NeRF works by optimizing a neural network whose input are 3D coordinates of a point in space and 2 rotations angles while the output is a volume density at this point and an RGB prediction (emitted radiance). The voulume density is the probability of a ray terminating at a particle in location x. In the MLP discrete points along a ray are sampled and used to estimate the expected colour. The estimated expected colour proposed by the authors is a discretized (using quadrature), differentiable version of expected colour in radiance field theory. Additionally, the authors propose two modifications which further improve their approach: Positional encoding and hierarchial volume sampling. The former applies a function to the input features, which in turn maps them to a higher dimensional space, which improves MLP fitting by capturing high frequency variation. At the same time the latter works by simultaneously optimizing two networks, a coarse network samples points using stratified sampling and then the "fine" network using the input from the "coarse" network determines which locations are relevant and samples rigorously from these points. This approach samples regions in the scene with visible content. Overall, this is a SOTA approach, which has already been tried for angiography reconstruction, in this project we will therefore attempt to use a different reconstruction method.

Incremental structure-from-motion (SfM) is a sequential methodology for reconstructing a 3D object using multiple camera views. The process begins by extracting features from each input image, these features are then matched among the different images. Geometric correspondence is established through geometric verification and outliers are removed using RANSAC, this ensures robustness of the matched points. The reconstruction then begins by initializing an image pair and then new images are added to build a graph of the scene. To add an image to the graph, a PnP problem is solved to recover the pose the image, and new scene points are trinagulated. Lastly, a bundle adjustment algorithm is used to refine the point parameters and camera parameters by minimizing reprojection error. COLMAP [11] is an SfM algorithm which improves the general pipeline by In this project an attempt will be made to evaluate COLMAP [11] for 3D reconstruction. COLMAP is an incremental structure from motion (SfM) algorithm, that introduces scene graph augmentation, next best view selection, robust and efficient triangulation, new bundle adjustment and redundant view mining to improve the 3D reconstruction quality. In this project we will evaluate COLMAP as a means of reconstructing 3D angiography images.

3 Data Preparation

3.1 Video frame selection

The objective of the data preprocessing phase was to select frames from the raw DICOM videos for subsequent reconstruction. Videos were obtained from a medical center in Novosibirsk and acquired using x-ray angiography.

For high-quality reconstruction, the frames extracted needed to meet specific criteria. First, the frames had to be in the same cardiac phase to ensure correspondence between videos. Second, the blood vessels in these frames had to be saturated with medical contrast, to ensure vessel visibility. To identify frames in the same heart phase, the pixelwise difference between each frame and a white background was calculated. This generated a time-series that plotted the difference with respect to the frame number. In certain time-series, like Figure 2.1, this showed a periodic behavior possibly reflective of the heart beat. For the extremum points, the corresponding frame number was recorded.

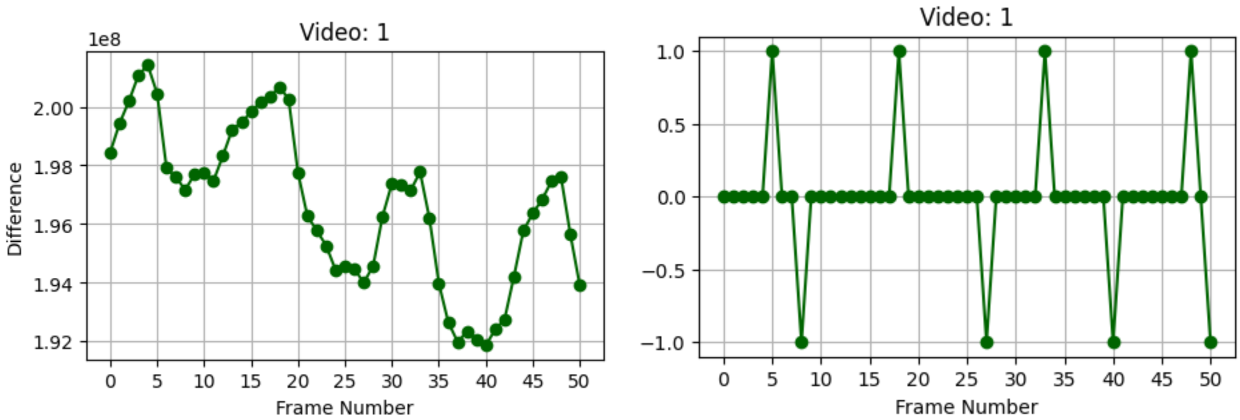

Figure 3.1: Left time-series: The application of the phase-identifying algorithm exhibits periodicity, highlighting the regular intervals of cardiac activity. Right graph: The identified corresponding extrema.

Once frames in the same heart-phase were identified, those that exhibited the highest level of vessel visibility were chosen. A method similar to [10] was employed. Canny edge detection was used to measure edges formed by the visible vessels in the image, images with the most edges were picked. This resulted in a combination of images that are in the same heart-phase and have the highest level of vessel visibility. These attributes make these frames attractive for further 3D reconstruction.

3.2 Image Segmentation

The next step in the data pre-processing pipeline involved the semantic segmentation of frames. The CVAT software was used for this purpose.

In the image, vessels were manually segmented using brush tools provided in the software. Since vessels differ in diameter, differently-sized brushes were used to segment vessel-regions of differing sizes (left image on figure 3.2). Additionally, other objects present in the frame such as the outline of the heart and a metal background pin were also segmented. Potentially, these objects could serve as markers in the reconstruction methods.

Thereafter, a mask of the segmentation was exported (middle image on figure 3.2) and an algorithm was developed which applied the mask onto the corresponding image. This outputted images in which regions of prime focus (vessel networks) were isolated for further reconstruction (right image on figure 3.2).

Figure 3.2: Left: Segmentation within CVAT Middle: Exported mask Right: Mask to image

4 Reconstruction

In this part of the report we evaluate the ability of COLMAP [11] to reconstruct 3D angiography images. COLMAP takes as input images from several views and utilizes feature extraction, matching and bundle adjustment to estimate camera poses and create a dense point cloud, which serves as a 3D model of an object.

Initially, to serve as the baseline reconstruction, images from all available video frames were passed into COLMAP, this yielded anomalous structure (Figure 4.1) which did not resemble the human heart. Possible reasons to this corrupted reconstructions could be attributed to heart expansion/contractions, varying visibility of the medical contrast or additional background structures within the photograph. In the remaining part of this section we address and mitigate these obstacles.

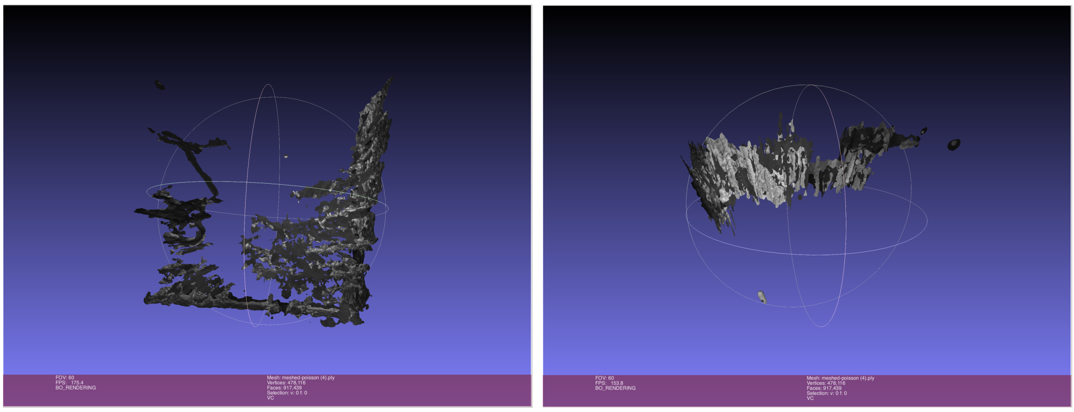

Figure 4.1: Baseline reconstruction

First, to take into account the heart contractions/expansions, images from a single heart phase and without any segmentation were passed into COLMAP, this crashed COLMAP and no reconstruction was produced. Similar result were generated when the aforementioned images under segmentation were passed as input. The results indicate that there may be something wrong with COLMAP and provide room for argument that COLMAP is ineffective at reconstructing angiography images.

Additionally, to check whether it is the background structures that are causing atypical reconstructions in the baseline, it was decided to pass segmented images of all frames in the video. This resulted in projections of the video fragments (Figure 4.2), using a single orientation. These results indicate that there is a strong possibility that background structures ruined our baseline reconstructions. Additionally, since the video fragments that COLMAP has chosen for reconstruction are of objects that are generally stationary within the video fragment (like the aorta in KG05), we can conclude that COLMAP is ineffective at reconstructing non-stationary objects and dynamic videos with heart-expansions/contractions can not be reconstructed using COLMAP

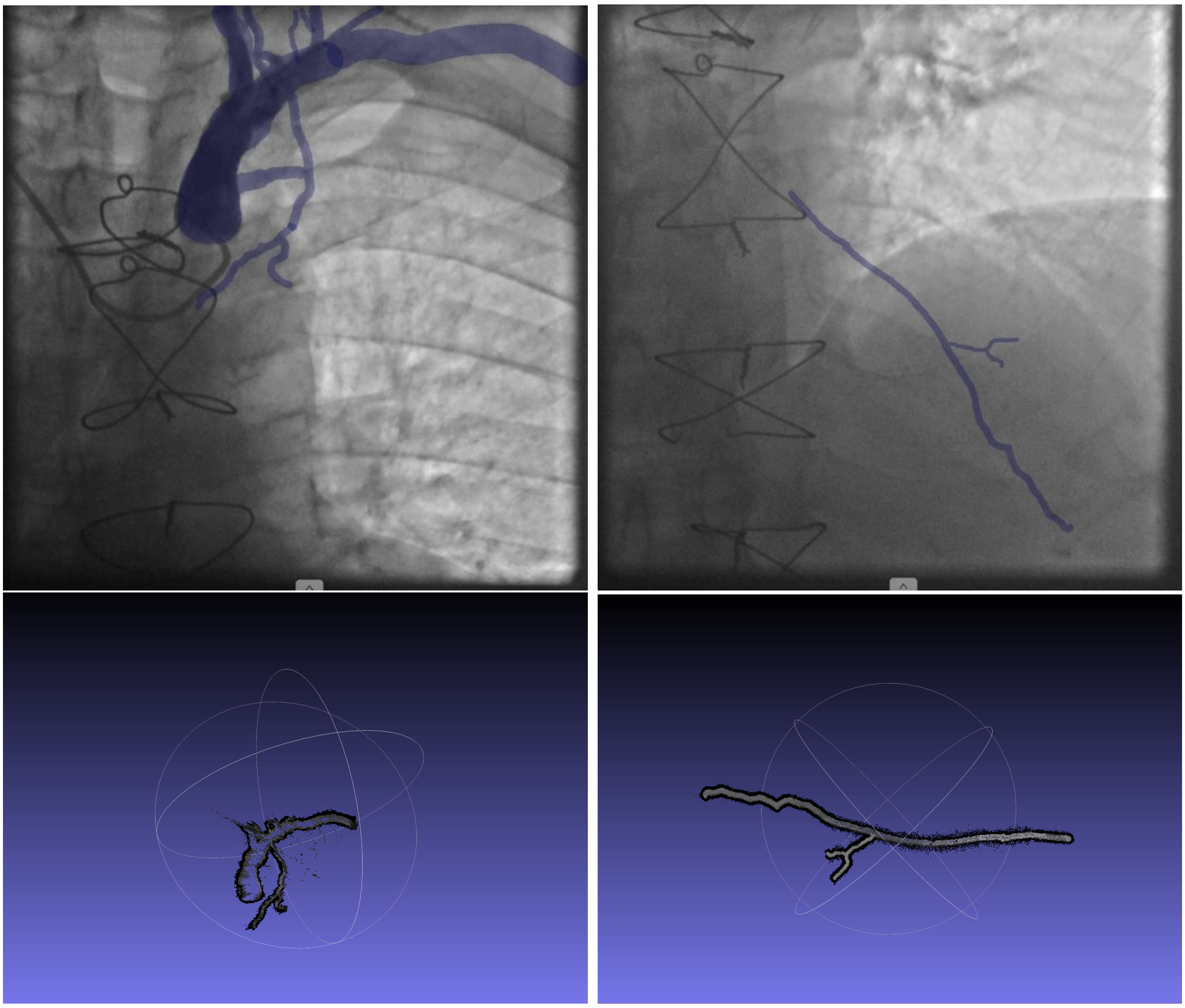

Figure 4.2: Projections of KG05 (left) and KG06 (right)

Overall, the investigation presented in this section evaluated the possibility of using COLMAP for angiography image reconstructions. The results of this investigation provide insights on key issues with using COLMAP for reconstruction of non-stationary objects with a large amount of background structures. The projections produced through COLMAP did not provide the desired result, as COLMAP did not establish correspondences between several camera views. In order to further evaluate observations of this section we will manually reconstruct correspondences and pose estimation in the next section.

5 Manual reconstruction

5.1 Feature Extraction & Feature Matching

In this part of the report multiple-view geometry methods will be used to reconstruct 3D points using images taken from KG01 and KG04. This part of the report heavily relies on methods from [5] and the cv2 python library, generally following a SfM pipeline.

Initially two images were selected using the preprocessing methodology outlined in part 2. Thereafter, in each image features were detected using the ORB feature detector. Then the images were matched using the Brute-Force Matcher, this resulted in pairs of matched features between the two images. Figure 5.1 shows 50 matches identified by the above methodology.

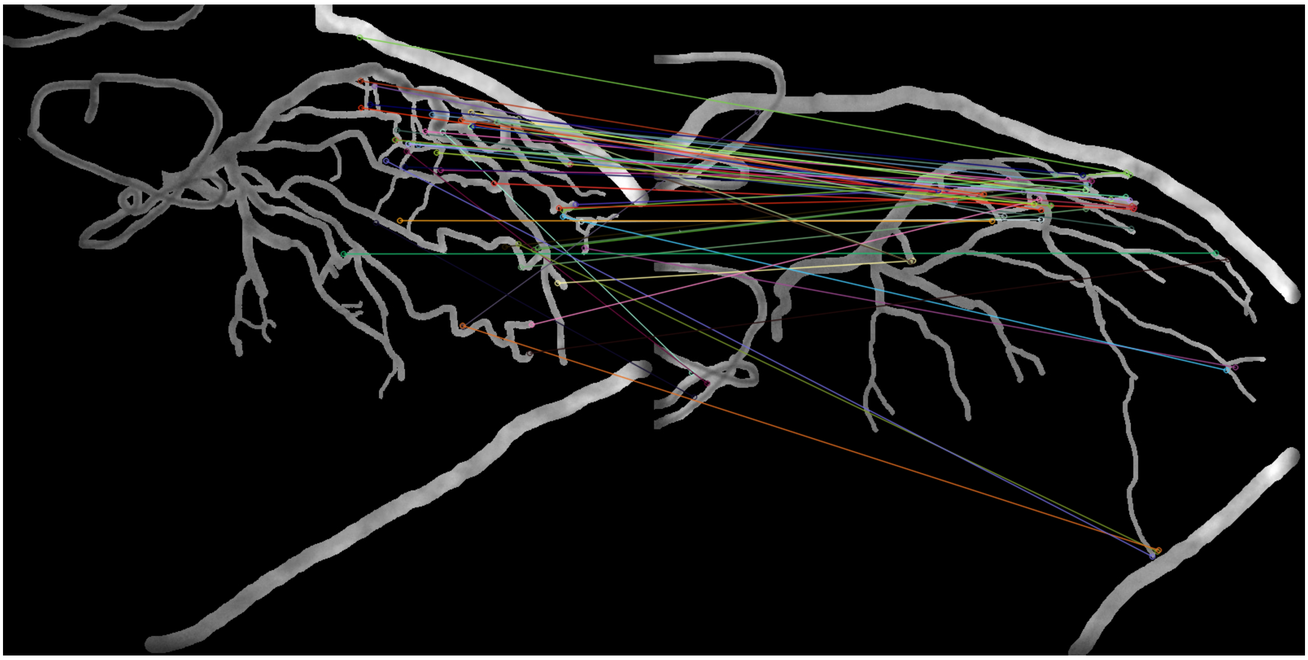

Figure 5.1: Matches between two frames at different camera perspectives

5.2 Deriving the Essential Matrix

To compute the essential matrix we first need to calculate the fundamental matrix. A fundamental matrix is a 3x3 matrix which relates points across views. It is used to map a point from one view onto an epipolar line in the second view. Generally, to find the fundamental matrix F, given points x1 and x2 from two views a system of linear equations will be solved that satisfies equation 1:

$$x1 * F * x2 = 0 \tag{1}$$

The fundamental matrix is then used to calculate the essential matrix E, using equation 2:

$$E = K_2^T * F * K_1 \tag{2}$$

Where K_1 and K_2 are the camera intrinsic matrices for camera 1 and camera 2, but since the same camera was used to capture the video frames $K_1 = K_2$. To find the calibration matrices, metadata was extracted from the .dcm file. The focal length per pixel was calculated by the formula $f_x = f_y = d/ps$, where d is the distance from source to detector and ps is the pixel spacing. Additionally, the center of the image (c_x, c_y) was calculated by dividing the number of rows and columns in the image by two, yielding the calibration matrix:

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

It should again be emphasized that in reality for computation we have used functions from CV2. Moreover, when calculating the essential matrix, the RANSAC algorithm was used to remove outlier matches from the images.

5.3 Camera poses & Triangulation

Having found the essential matrix, it is known that it depends on (t,R) through equation 4:

$$E = [t]_x * R \tag{4}$$

We can thus find t and R through SVD decomposition of E, where the translation vector would correspond to the smallest singular value and therefore the third column of matrix U, since the corresponding singular value would be zero and the translation vector corresponds to the direction where $tE = 0$. On the other hand, R would be found using formula $R_1 = UWV^T$, $R_2 = UWV^T$, where W and W^T are basic rotation matrices and by combining them with the U and V^T from SVD of E, we thereby extract rotation matrices that align with the geometrical properties of the camera motion given by E. Altogether, four possible solutions would be derived and the true solution would correspond to (t,R), such that the 3D points remain in front of the camera (have positive depth).

The next step would involve computing the projection matrices defined by equation 5 & 6:

$$P0 = K[I|0] \tag{5}$$

$$P1 = K[R|t] \tag{6}$$

Having found these projection matrices, 3D points can be reconstructed by solving a system of linear equations given by equation 7 & 8:

$$x0_{pr} = P_0 * X_{3D} \tag{7}$$

$$x1_{pr} = P_1 * X_{3D} \tag{8}$$

For any matched points. Notably, we have three unknowns and four linearly independent equations so we can solve this by algebraically manipulating into $AX = 0$ system and solving it through singular value decomposition. This would yield the reconstructed 3D coordinates of points in space.

5.4 Reprojection and Reconstruction Accuracy

Having derived the 3D points in space, we can reproject them back onto the image and compare with the intially matched points. The results are depicted in figure 5.4 and shows that the algorithm accurately reprojects the points.

This shows that even using simple 3D reconstruction techniques and without complex methodologies such as bundle adjustment, we have been able to accurately reconstruct 3D points in space. These results suggest that the previously produced COLMAP results are questionable and provide room for further investigation, perhaps by implementing further steps of the structure-from-motion algorithm, or manually implementing COLMAP.

6 Conclusion

This project has engineered an innovative framework for reconstructing the vessel network within the human heart and tested its robustness with real data. The framework relies on pre-

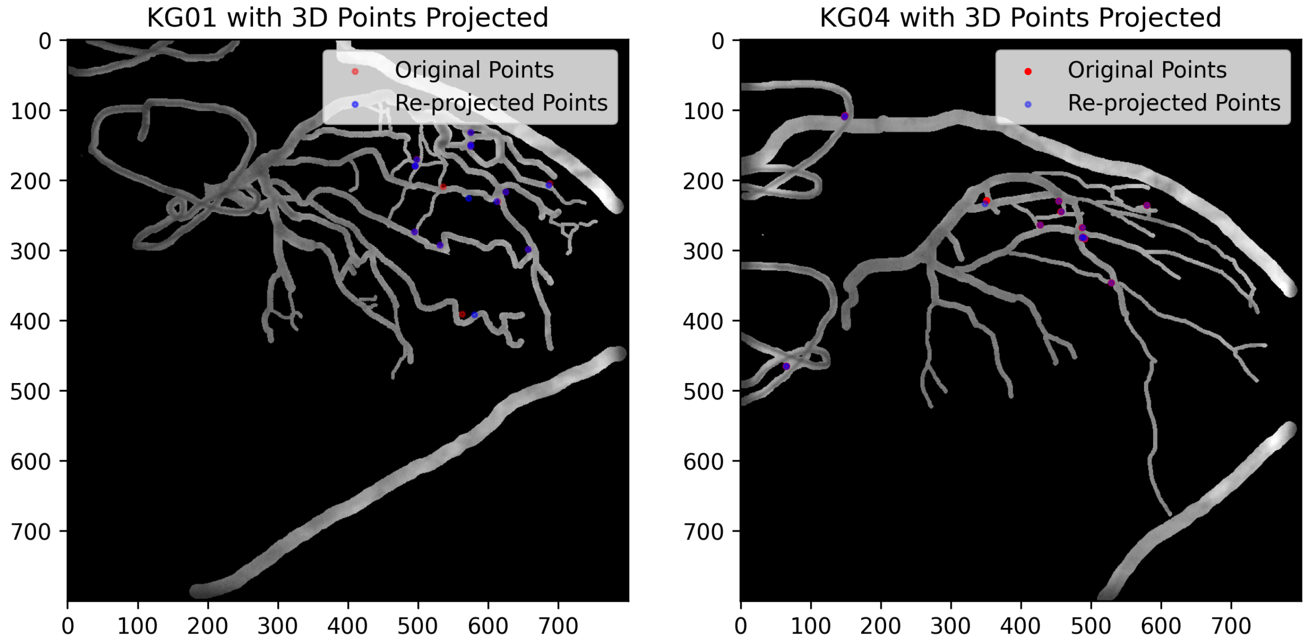

Figure 5.2: Initial and reprojected points

processing steps which segment the videos and extract frames of the heart in a single phase, followed by an automatic reconstruction algorithm which theoretically should be able to produce a 3D reconstruction of the object in question. However, despite the most profound efforts, when implementing the framework we have failed to achieve the desired goal of automatic 3D reconstruction.

Evidence from section 4 points to COLMAPs inability to establish correspondences between the different video frames and capture the changes in camera pose. However, even a simple manual reconstruction was successfull at reconstructing 3D points between neighbouring frames, which raises further questions regarding COLMAP. Several hypothesis can be formulated which suggest areas of further research. First and foremost, it is possible that our image segmentation is inappropriate and more robust techniques need to be tried for automatic segmentation (AngioNet). Second, it is possible that COLMAP failed to work in the specific environment where it was run (an old PC with a CUDA GPU). Lastly, our results could, perhaps, highlight an issue with COLMAP.

To test these hypothesis COLMAP will initially be tested on another PC. Thereafter, image segmentation will be performed using AngioNet and again an attempt at reconstruction will be made. Lastly, more steps of the SfM pipeline will be implemented incorporating additional improvements introduced by COLMAP. A manual reconstruction could highlight where exactly our algorithm crashes, knowing this we could think of how to mitigate these issues in order to produce a reconstruciton of the vessel within the human heart.

References

[1] C. Blondel, G. Malandain, R. Vaillant, and N. Ayache. "Reconstruction of coronary arteries from a single rotational X-ray projection sequence". In: *IEEE Transactions on Medical Imaging* 25.5 (2006), pp. 653–663. DOI: 10.1109/TMI.2006.873224.

[2] Serkan Çimen, Ali Gooya, Michael Grass, and Alejandro F Frangi. "Reconstruction of coronary arteries from X-ray angiography: A review". en. In: *Med Image Anal* 32 (Mar. 2016), pp. 46–68.

[3] Costin Ciusdel, Alexandru Turcea, Andrei Puiu, Lucian Itu, Lucian Calmac, Emma Weiss, Cornelia Margineanu, Elisabeta Badila, Martin Berger, Thomas Redel, Tiziano Passerini, Mehmet Gulsun, and Puneet Sharma. "Deep neural networks for ECG-free cardiac phase and end-diastolic frame detection on coronary angiographies". In: *Computerized Medical Imaging and Graphics* 84 (2020), p. 101749. ISSN: 0895-6111. DOI: https://doi.org/10.1016/j.compmedimag.2020.101749. URL: https://www.sciencedirect.com/science/article/pii/S0895611120300525.

[4] Francesca Galassi, Mohammad Alkhalil, Regent Lee, Philip Martindale, Rajesh K Kharbanda, Keith M Channon, Vicente Grau, and Robin P Choudhury. "3D reconstruction of coronary arteries from 2D angiographic projections using non-uniform rational basis splines (NURBS) for accurate modelling of coronary stenoses". en. In: *PLoS One* 13.1 (Jan. 2018), e0190650.

[5] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. 2nd ed. New York, NY, USA: Cambridge University Press, 2003. ISBN: 0521540518.

[6] Kritika Iyer, Brahmajee K. Nallamothu, C. Alberto Figueroa, and Raj R. Nadakuditi. "A multi-stage neural network approach for coronary 3D reconstruction from uncalibrated X-ray angiography images". In: *Scientific Reports* 13.1 (Oct. 2023), p. 17603. ISSN: 2045-2322. DOI: 10.1038/s41598-023-44633-2. URL: https://doi.org/10.1038/s41598-023-44633-2.

[7] Pietro Mascagni, Deepak Alapatt, Luca Sestini, Maria S. Altieri, Amin Madani, Yusuke Watanabe, Adnan Alseidi, Jay A. Redan, Sergio Alfieri, Guido Costamagna, Ivo Boškoski, Nicolas Padoy, and Daniel A. Hashimoto. "Computer vision in surgery: from potential to clinical value". In: *npj Digital Medicine* 5.1 (Oct. 2022), p. 163. ISSN: 2398-6352. DOI: 10.1038/s41746-022-00707-5. URL: https://doi.org/10.1038/s41746-022-00707-5.

[8] Yinghui Meng, Minghao Dong, Xumin Dai, Haipeng Tang, Chen Zhao, Jingfeng Jiang, Shun Xu, Ying Zhou, Fubao Zhu, Zhihui Xu, and Weihua Zhou. "Automatic identification of end-diastolic and end-systolic cardiac frames from invasive coronary angiography videos". en. In: *Technol Health Care* 30.5 (2022), pp. 1107–1116.

[9] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. "NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis". In: *CoRR* abs/2003.08934 (2020). arXiv: 2003.08934. URL: https://arxiv.org/abs/2003.08934.

[10] Maxim Popov, Temirgali Aimyshev, Eldar Ismailov, Ablay Bulegenov, and Siamac Fazli. *A Review of Modern Approaches for Coronary Angiography Imaging Analysis*. 2022. arXiv: 2209.13997 [eess.IV].

[11] Johannes Lutz Schönberger and Jan-Michael Frahm. "Structure-from-Motion Revisited". In: *Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.

[12] C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, and R. Leonardi. "Deep learning for automated skeletal bone age assessment in X-ray images". In: *Medical Image Analysis* 36 (2017), pp. 41–51. ISSN: 1361-8415. DOI: https://doi.org/10.1016/j.media.2016.10.010. URL: https://www.sciencedirect.com/science/article/pii/S1361841516301840.

[13] Francesco Viola, Giulio Del Corso, Ruggero De Paulis, and Roberto Verzicco. "GPU accelerated digital twins of the human heart open new routes for cardiovascular research". In: *Scientific Reports* 13.1 (May 2023), p. 8230. DOI: 10.1038/s41598-023-34098-8. URL: https://doi.org/10.1038/s41598-023-34098-8.