DESY Summer Student Report

Darie Picu, University of Edinburgh, United Kingdom

September 4, 2013

Abstract

This report is about two projects. The first projects task is to read the metadata of nexus files provided and make this available via a WSDL end point. The task of the second project is analysing the tomograms of magnesium tubes that are usually implanted in human hearts. The magnesium tubes slowly decay until they are completely absorbed by the body. The analysis is required for understanding how the decay takes place. Main steps of analysis involve filtering the images with different noise removal filters, finding the edges of the regions of interest, repainting images for better visualisation and providing means of visualisation for the time evolution of the decay.

Contents

1 Acknowledgements

We would like to thank our supervisors Steve Alpin, Felix Beckmann, Friederike Nowak and Jürgen Starek for their guidance and support throughout our projects.

2 Programming language choice

There was a debate whether to choose a language we (me and Oscar Byrne) were most comfortable with: Java and C++, or choose something new. Our supervisor, Jurgen Starek, suggested that we pick a new language, since one of the main goals of the summer student programme is for us to learn as much as possible. After some research, we decided to choose the Python programming language. The main reasons are: the language is popular, easy to learn and we found useful libraries for both the projects. As for a version control, we decided to use GIT.

3 Nexus to WSDL project

3.1 Introduction

This project was done under the supervision of Jürgen Starek. Nexus files store data taken from experiments in a format similar to how XML files store data. This projects first task was to read the metadata of these files: important information of interest such as, experiment identifier, instrument used and others. The second part of the project was to design a WSDL end point. This server would provide metadata about any Nexus file that was read.

3.2 Reading Nexus metadata

Most of the work went intro installing the right libraries in order to deal with the Nexus files: Nexus, HDF5, Distribute, Numpy and others. Once this was done, reading the metadata of a given Nexus file was easily implemented. Oscar Byrne continued working on this part, more details in his report. The outcome of this is a Nexus reader that reads all the Nexus files at a given path and stores the metadata into a database file.

3.3 WSDL

We encountered library problems on this part too. Some libraries were not maintained anymore and others had too little documentation. In the end we chose the Spyne library. I worked on this part of the project and implemented a server that would use the Nexus reader. By specifying the experiment identifier as a key, together with the required metadata field, one can access the metadata of the Nexus files read by the Nexus reader.

4 Magnesium decay analysis

4.1 Introduction

This project was done under the supervision of Felix Beckmann.
Magnesium tubes are used as implants for various purposes. They are better than conventional titanium ones because they do not require a second surgery for removal, as they slowly decay in roughly half a year and get absorbed in the body with no bad side effects.
These magnesium tubes are scanned at PETRA III. They are submerged in a fluid that mimics blood so that the decay is as similar as possible to the decay that would happen to the implants. Tomograms are taken at different times, from the moment they are submerged in the fluid to the moment when they are in a full state of decay. These tomograms need to be processed, in order to find different parameters of the magnesium decay.
Figure 1 shows a slice from the original tomogram. Each pixel location represents the absorption rate at that point in space. Absorption rate is directly proportional to density. Black represents low density and white represents high density.
Three regions can be distinguished by eye

- Black region - outside the tube which corresponds to the fluid in which the tube is submerged and to the vacuum inside the tube
- Dark grey region - outer regions of the tube. This corresponds to magnesium that is about to decay
- Light grey region - the inner regions of the tube. This corresponds to pure magnesium, unnaffected by decay

Note that there are some white dots all over the magnesium that look like noise. They are highly absorbant particles. The magnesium is not completely pure, it has some other alloy in it.

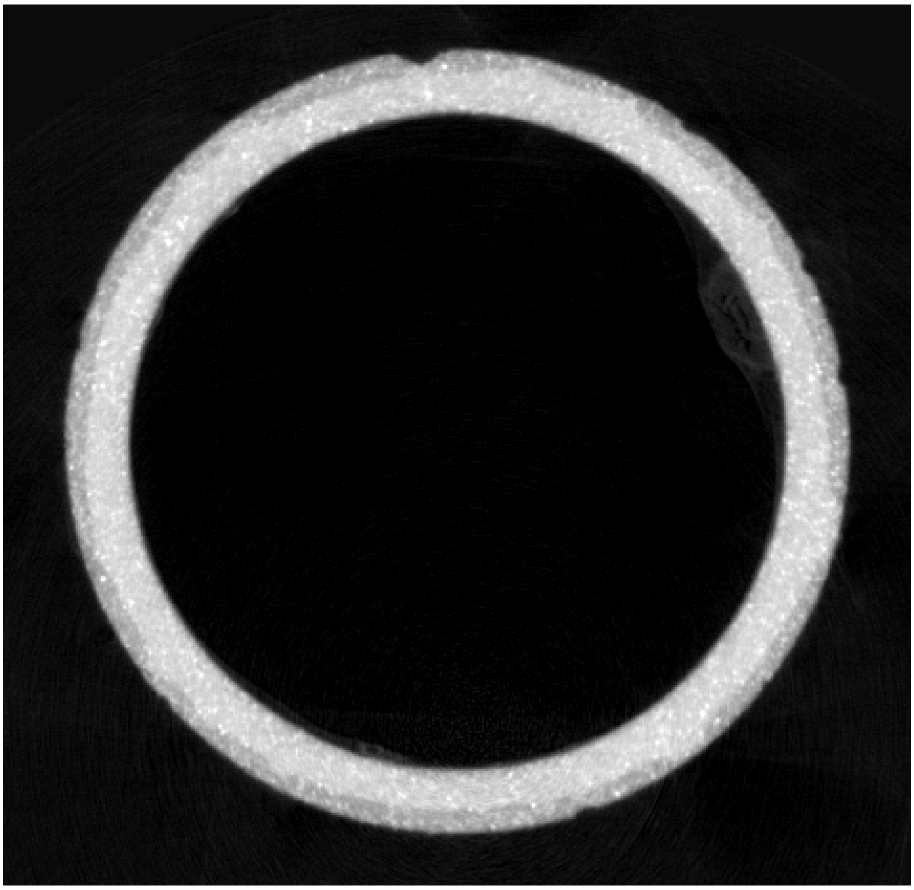

Figure 1: Slice in the original tomogram

4.2 Main steps

- The task is to process a series of tomograms, each tomogram being stored in .sli files or tif. images, one file per slice. Processing should be done in an automatic way, the user should only provide the path to the directory containing all the folders for each time step
- Remove the the highly absorbent particles from the images, as there is no interest in studying these. Due to their random distribution throughout the magnesium, the task is basically noise removal
- Find threshold grayscale values between the 3 regions: fluid, decaying magnesium and pure magnesium need to be computed. These would then be used to differentiate between the three regions
- The filtered images are re colored in high contrast colours for better visualisation of the data

- Provide means of visualising the decay of the magnesium by using two tomograms of the same tube from different time steps

4.3 Image filtering

4.3.1 Median filter

There are many filters that may be applied to remove the noise. One of the most commonly used is the median filter. The most basic median filter does the following to an image: for each pixel in the image, replace its value by the median value of the pixel and its neighbours. Noise pixels are usually stray pixels surrounded by pixels of a different colour. The median filter would remove such pixels. The median footprint decides how many neighbours to consider for each pixel. The basic footprint is 3 by 3 (the pixel and its 8 neighbours), but the footprint may be bigger.
The problem with the median filter is that it blurs the image, most importantly the edges are less sharp. Figure 2 shows the result of applying the median filter with a big footprint to the original image. In this project, analysing the edges between decaying magnesium and pure magnesium is important, therefore using the median filter alone is a poor solution.

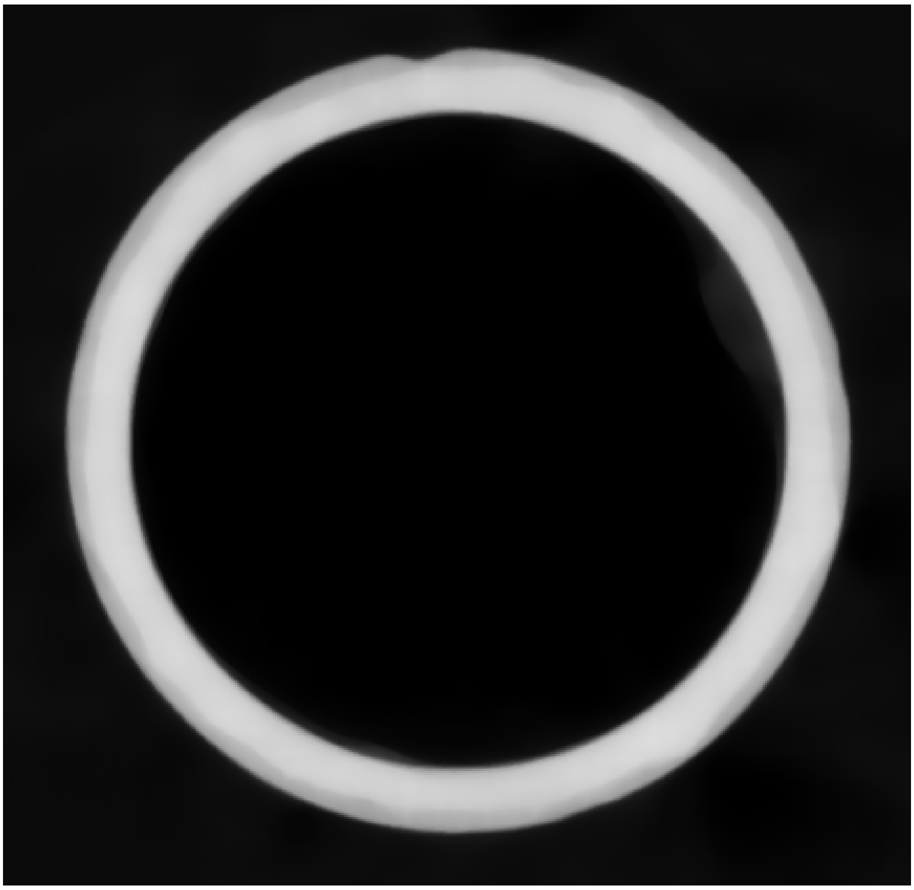

Figure 2: Median filtered image

4.3.2 Anisotropic diffusion filter

Anisotropic diffusion is much more complex that the median filter. By tuning its parameters, one can use it to diffuse (fade out) the highly absorbent particles while preserving the sharp edges. One of the parameters, called Kappa, encodes how important the preservation of edges is. Low Kappa means edges must be preserved, lowest Kappa preserves even the edges of the highly absorbent particles. Figure 3 shows the result of filtering the original image with a low Kappa anisotropic diffusion filter.

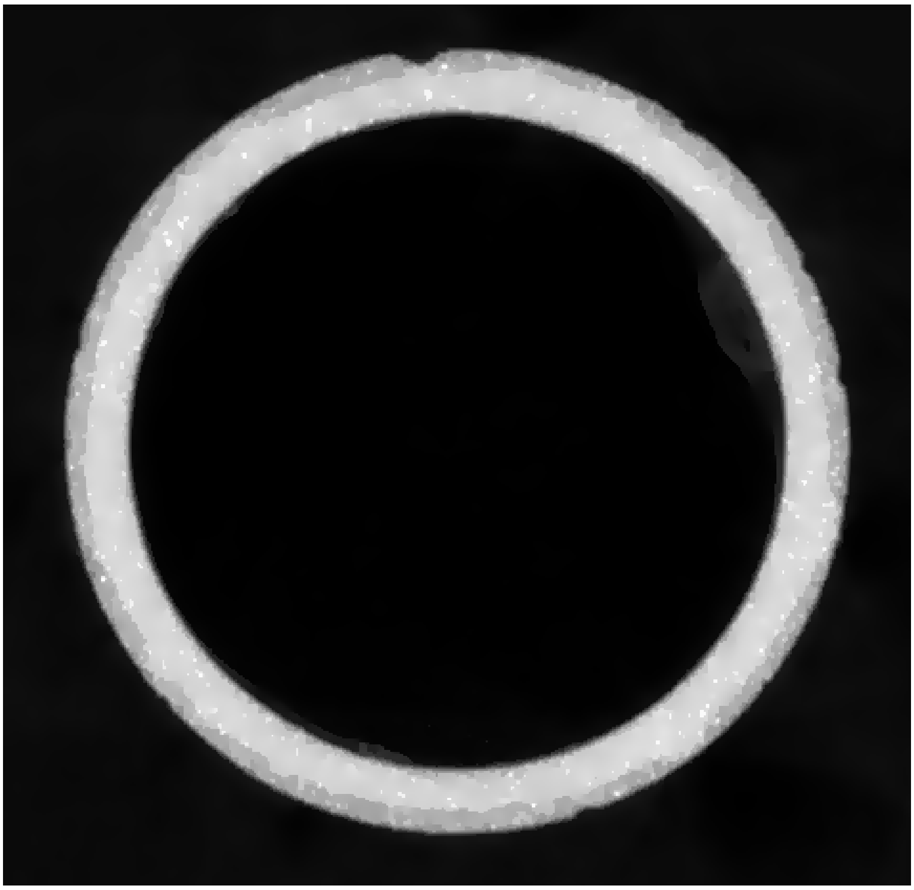

Figure 3: Anisotropic diffusion with low Kappa

High Kappa results in linear diffusion, which blurs out everything, edges included. Figure 4 shows the result of filtering the original image with a high Kappa anisotropic diffusion filter.

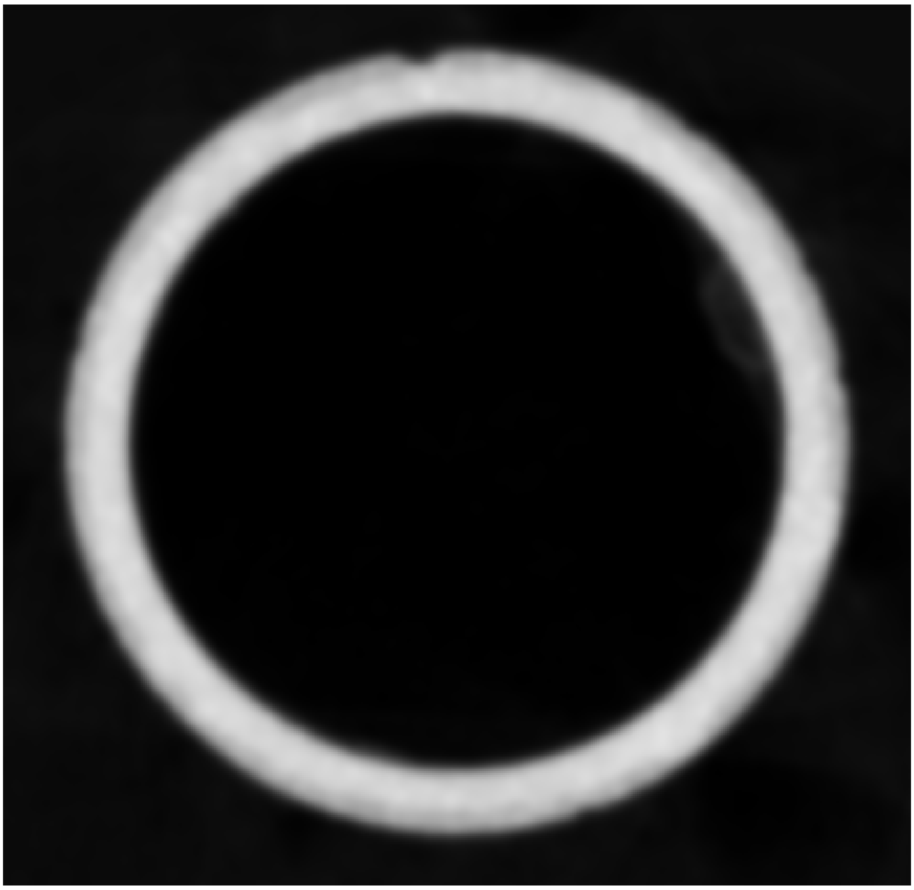

Figure 4: Anisotropic diffusion with high Kappa

4.3.3 Final filter

Many trials were made in an attempt to find parameters that would diffuse all the highly absorbent particles and preserve the edges between decaying magnesium and pure magnesium. Such parameters where not found and was decided that a combination of low Kappa anisotropic diffusion filter coupled with a low footprint median filter would be the best choice. The anisotropic filter removes most of the highly absorbent particles and the median filter removes the last remaining highly absorbent particles. The result of this combination is shown in Figure 5

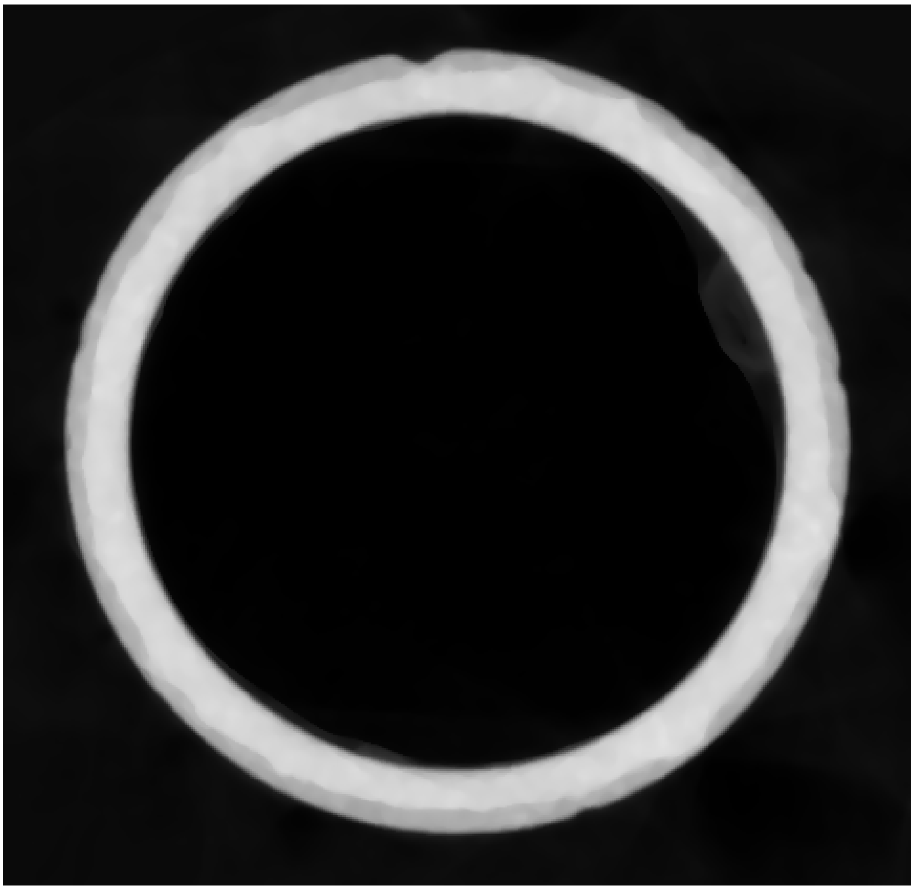

Figure 5: Filtered image

5 Threshholds

5.1 Introduction

Two thresholds are needed in order to differentiate between the 3 main regions: fluid, decaying magnesium and pure magnesium.

- fluid - decaying magnesium
- decaying magnesium - pure magnesium

Note that on the inside of the tube there is vacuum, the threshold between vacuum and magnesium is not of interest since it does not replicate the decay that would happen with an implant.

5.2 Sobel filter and histogram

The Sobel filter determines places in an image where the colour gradient is highest, these places are at the borders between regions. The Sobel filter was used on the images that were already filtered by the anisotropic diffusion filter and the median filter. From now on these will be referred to as simply the filtered images. Figure 6 shows the Sobel filter applied on the filtered original image, the edges between the 3 main regions are visible.

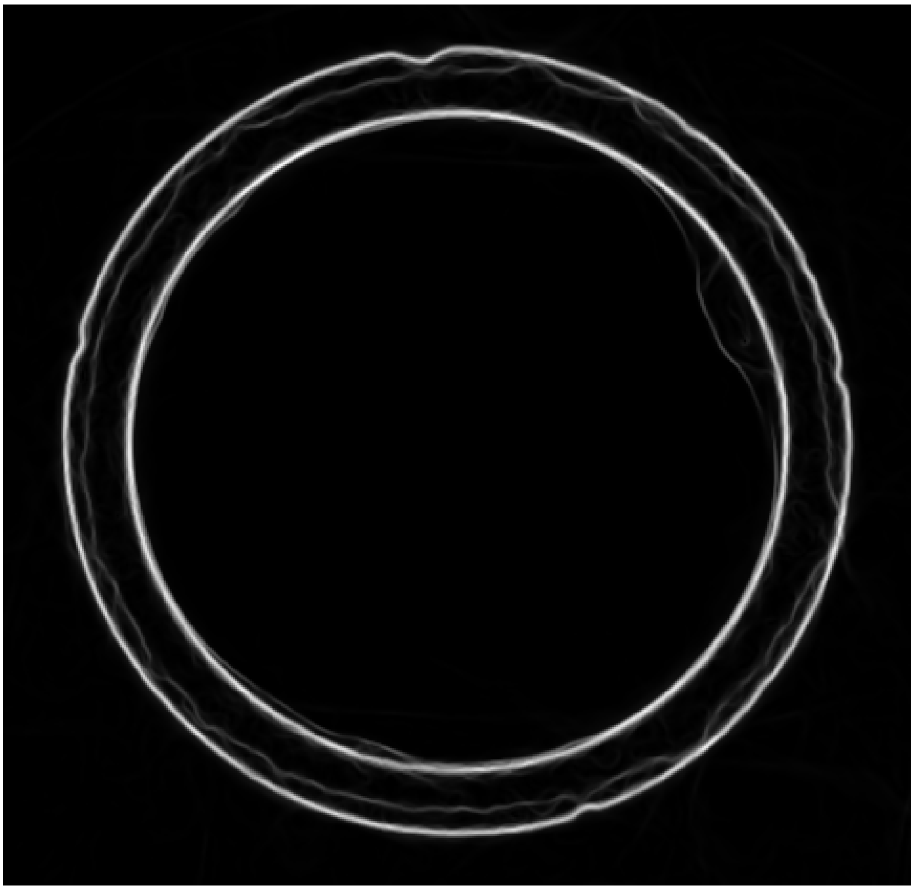

Figure 6: Sobel filtered image

The locations of high gradient edge pixels were recorded, and the grayscale values corresponding to those locations were plotted in a histogram. Figure shows one such histogram of pixel counts versus grayscale value:

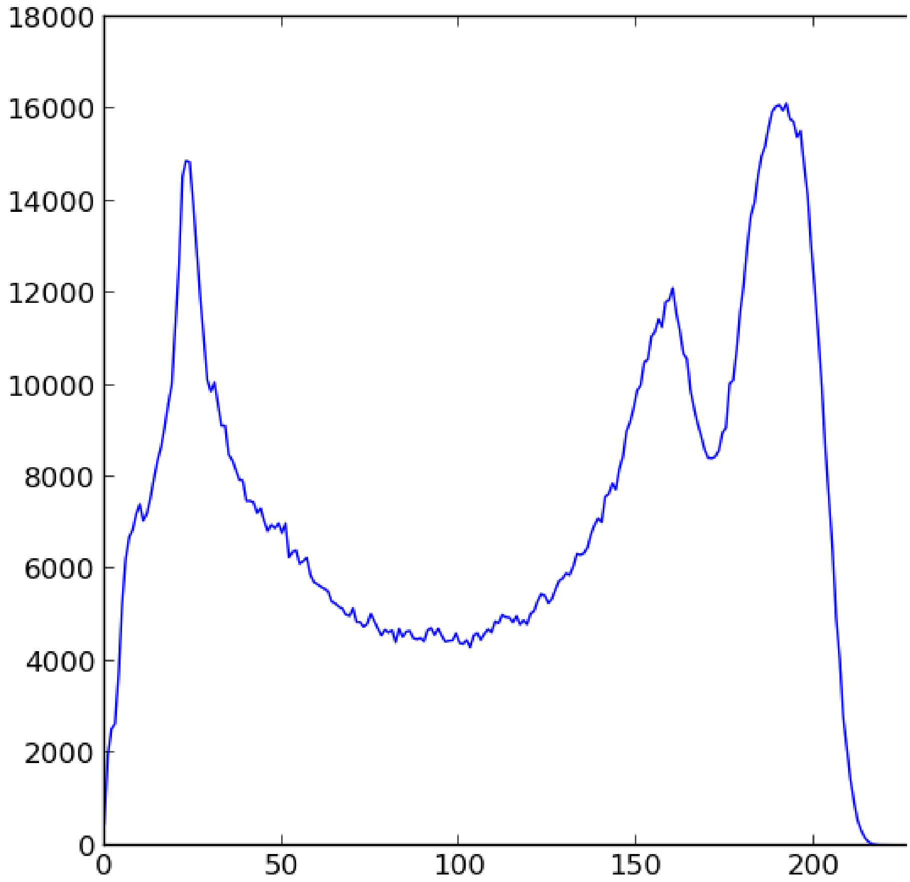

Figure 7: Histogram for edge pixels

The three peaks correspond to: fluid + vacuum, decaying magnesium and pure magnesium. Fitting three Gaussians and finding their intersection points should lead to the required thresholds.

This approach has a problem though: one needs to choose a threshold for the high gradient pixels that discriminates between low and high gradient locations. The shape of the histogram would stay the same, but it would be shifted left or right depending on this threshold. A shifted histogram would lead to different thresholds. It was decided not to use the Sobel filter due to this problem.

5.3 Histogram for all pixels

Instead of considering only the edge pixels, all pixels could be considered. The histogram would look similar to the one in Figure 7, only that the first peak corresponding to air and fluid would dominate the other two peaks, since there is much more air/fluid than there is magnesium in the images. The technique of determining the thresholds is described in detail in the report written by Oscar Byrne.

6 Better Visualisation

6.1 Basic recoloring

The filtered images may be recoloured for better visualisation using the thresholds. Two contrasting colours, dark blue for pure magnesium and light green for decaying magnesium were used. Black was used for fluid and vacuum. Figure 8

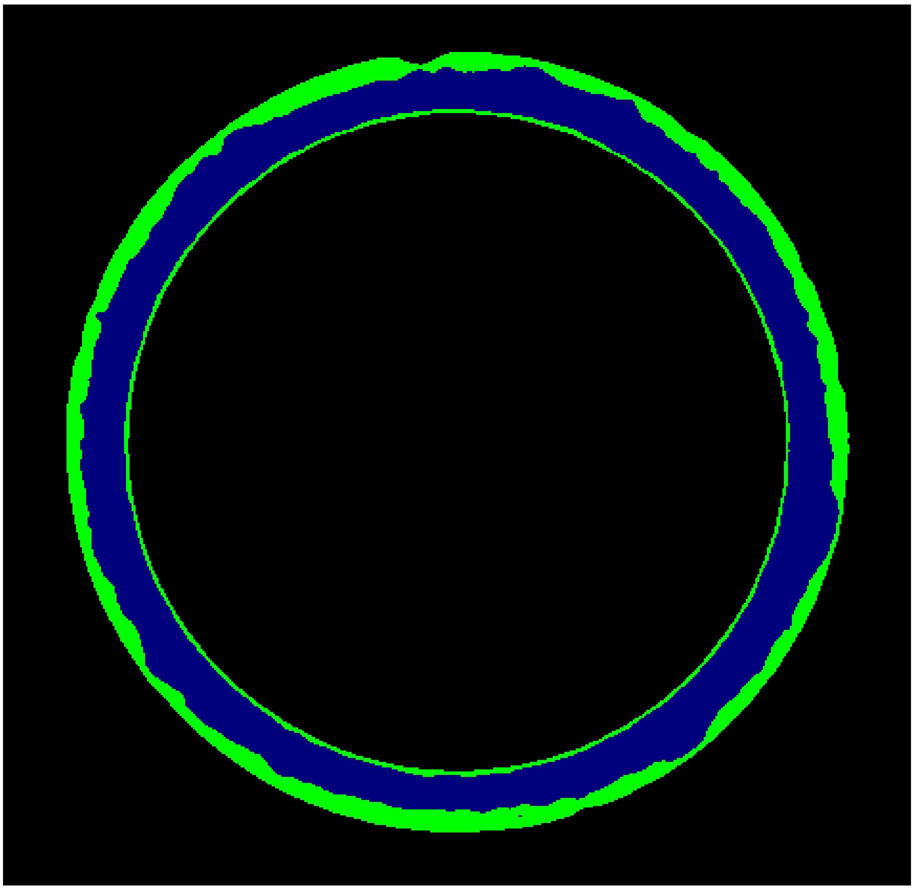

Figure 8: Recolored filtered image

Basic recolouring is used after the two filters have been applied and the thresholds were computed. The resulting images are stored in a folder for each tomogram.

6.2 Advanced recolouring

Using only three colours for recolouring, one misses small details in the structure of the material. For each region, two colours may be defined, one for the minimum and one for the maximum grayscale value. Then for each pixel in the region, linearly interpolate its grayscale value between these two colours. Figure 9 shows one such advanced recoloured image. Pink was used to show the fluid as well. Unfortunately, this type of visualisation is not necessary for our project, therefore advanced recolouring is no longer applied.

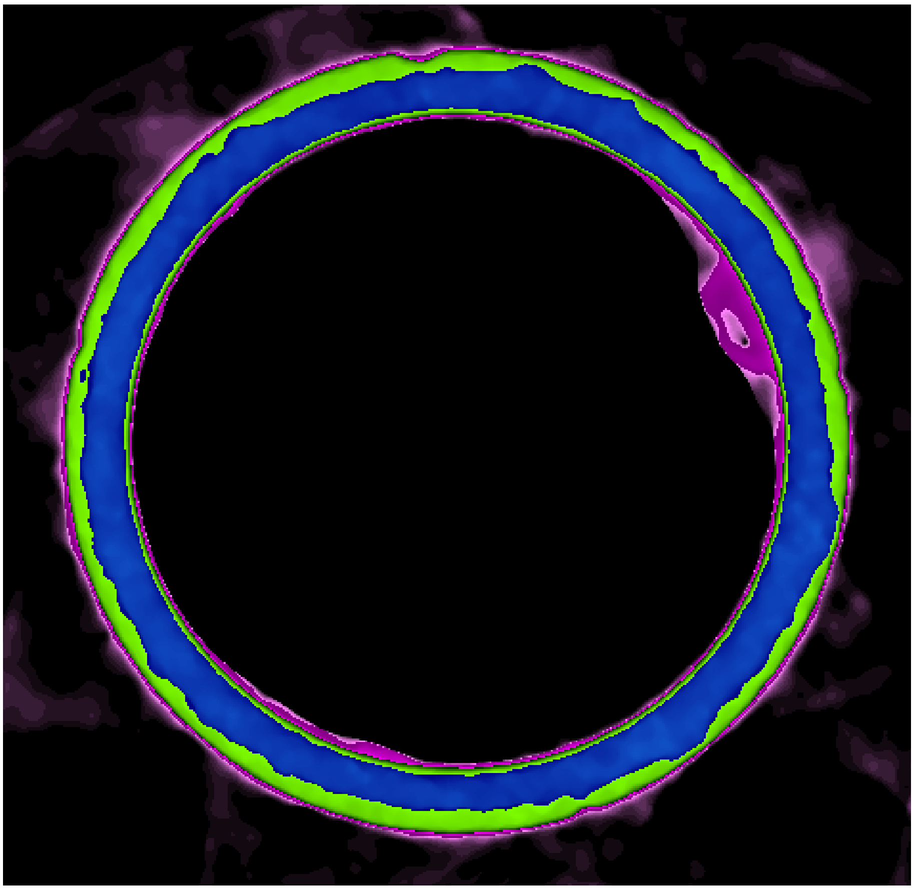

Figure 9: Advanced recolored filtered image

7 Time visualisation

7.1 Position shifts between tomograms

Time visualisation is essential in understanding how the decay takes place. Two tomograms of the same magnesium tube may be used to find regions such as:

- Decaying magnesium which completely decayed
- Pure magnesium which became decaying magnesium

These are the only types of regions that one would expect. However, regions where there was fluid in the first time step that turned into decaying magnesium in the second timestep are possible too if the magnesium tubes shift positions from one scan to the other. Such regions were found to exist, therefore it was concluded that there indeed is a shift in the magnesium tube position between scans. Figure 10 shows these regions in red.

In order to fix this shift problem, the two tomograms have to be aligned. A simple solution was implemented. The two tomograms are stored into 3D bit arrays, 1 stands for any type of magnesium and 0 for fluid and vacuum. The similarity between the two is simply their dot product. Assuming that the tube can only move in the plane of the images, one could check the similarity with the four adjacent positions, call them North, South, East and West. If any position is more similar than the current position, shift the array by one in that direction. This step is repeated until the current position has the highest similarity. This was implemented and proved to work, however our supervisor Felix Beckmann told us that it is better to align the two tomograms by only considering some highly absorbent particles that are in the pure magnesium. Three such particles should be enough for alignment. This solution was not implemented.

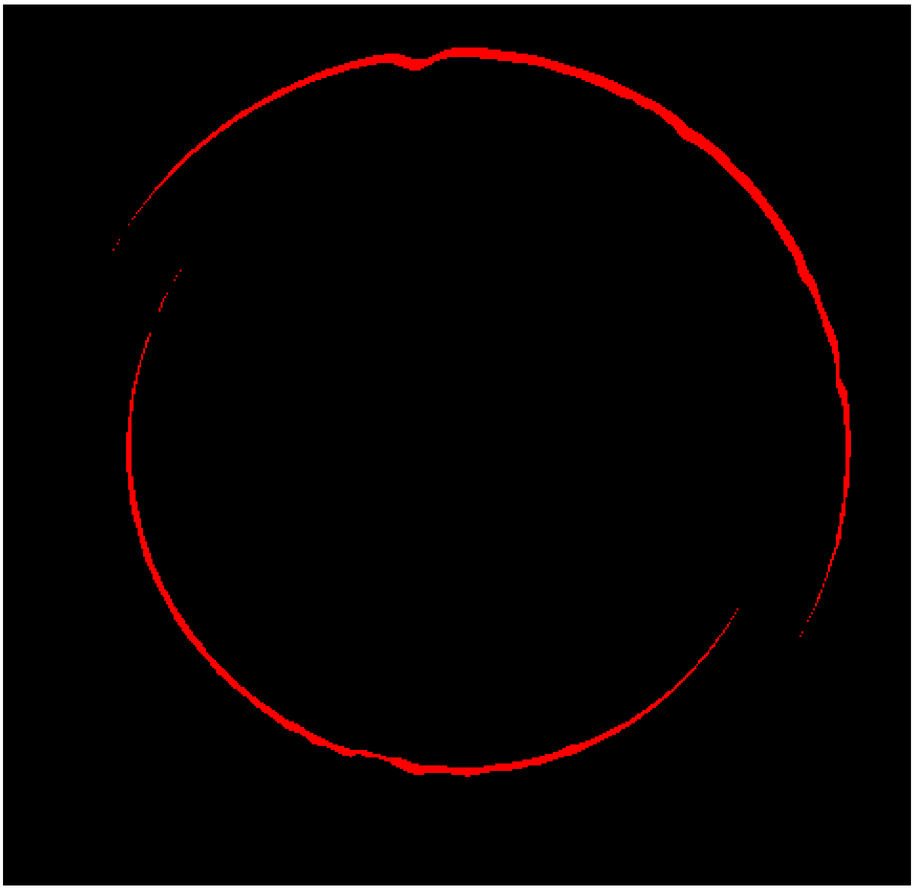

Figure 10: Shift in the magnesium tube

7.2 Dilation of decaying magnesium

One would expect the volume of the magnesium to decrease as decay takes place. While this is true for big time steps, as eventually all the magnesium will have decayed completely, for small time steps it was found that the volume increases. The increase rates range from 0.5% to 0.9%.This makes the analysis described in Section 7.1 harder.

7.3 Visualisation

A simple visualisation method was implemented, Figure 11 shows the time evolution between two time steps.

- Light grey: decaying magnesium in both time steps
- Dark grey: pure magnesium in both time steps
- Green: pure magnesium which decayed

- Blue: dilation of the decaying magnesium
- Red: misalignment

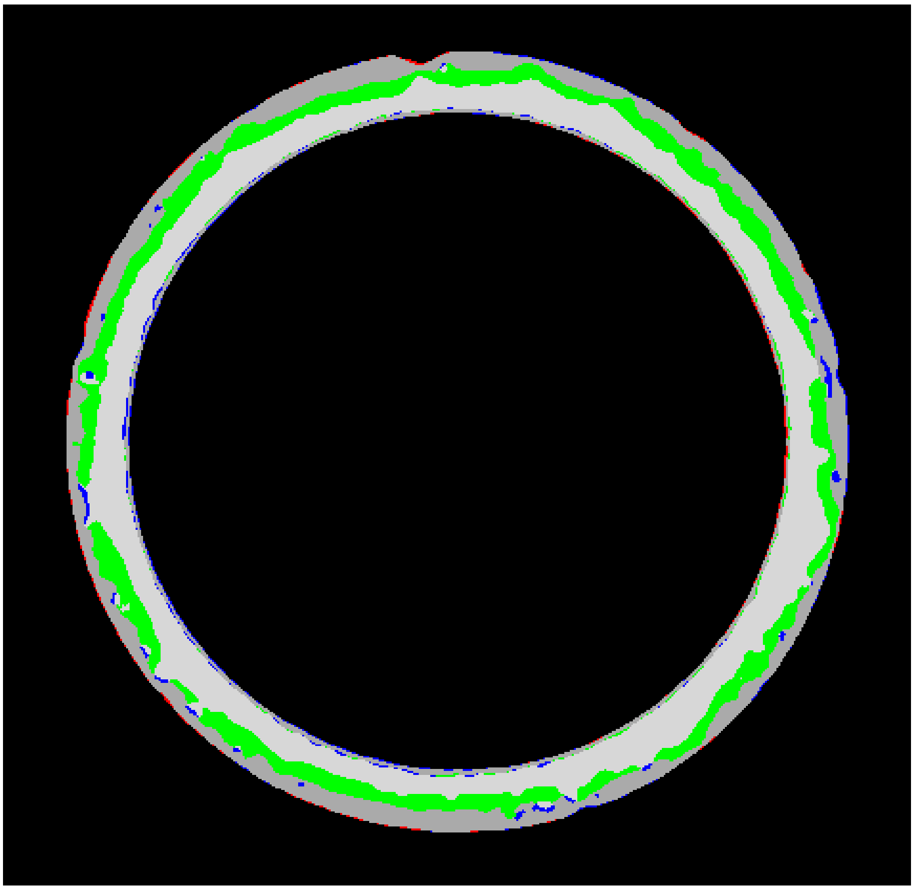

Figure 11: Time evolution

Dilation coupled with slight misalignment makes this visualisation method not very informative. However one could still draw some qualitative conclusions about how the decay advances.

7.4 Overview

The project does the following:

- Receives a path where the tomogram data is. Each folder at this path is expected to be a tomogram at a different timestep. Ordered alphabetically, these tomograms should be in chronological order
- Each tomogram is treated independently. The data is read into a 3D array, vertical slice in this array representing one of the slices presented so far in the report

- The 3D array is filtered with anisotropic diffusion and the median filter. Filtered images are saved into a folder called Filtered
- Using the thresholds, a region 3D array is created which has 0, 1 or 2 for fluid, decaying magnesium and pure magnesium. Recolored images are saved into folder Recolored. The 3D array is saved into folder Masked.

Code for basic alignment of two tomograms and for visualisation of two timesteps is provided, but not used.

References

[1] Nonlinear scale-space by anisotropic diffusion *James Fishbaugh*

[2] Anisotropic diffusion in image processing *Joachim Weickert*