AN ALIGNMENT SCHEME FOR ATOMIC FORCE TOMOGRAMS OF CONDUCTIVE FILAMENTS IN RESISTANCE SWITCHING DEVICES

An Undergraduate Research Scholars Thesis

by

JOSEPH PIERRE ANDERSON

Submitted to the Undergraduate Research Scholars program at
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by Research Advisor: Dr. Patrick J Shamberger

May 2018

Major: Physics

TABLE OF CONTENTS

ABSTRACT

An Alignment Scheme for Atomic Force Tomograms of Conductive Filaments in Resistance Switching Devices

Joseph Pierre Anderson
Department of Physics and Astronomy
Texas A&M University

Research Advisor: Dr. Patrick J. Shamberger
Department of Materials Science and Engineering
Texas A&M University

The search for a memristive device for use in neuromorphic circuits has brought great interest to research regarding resistance switching devices (also known RRAM devices). These devices, which grow metallic filaments through an insulating material, presently have one major drawback: the electrical characteristics (namely on-state resistance) have a high level of variability from device to device which renders them unusable. Much of this variability can conceivably be attributed to the structure of the metallic filaments. This research presents the development of a reliable means of spatially characterizing the filaments using a tomographic technique based on atomic force microscopy, especially focusing on a post-processing software to analyze the microscopy data.

CHAPTER I

INTRODUCTION

For all the benefits they promise, resistance switching devices have managed to remain just outside of the reach of commercial viability for some time. These devices, which form conductive bridges (filaments) through an insulator under an applied voltage, have the potential to open up entirely new computer architectures; but variability of electronic characteristics in identically prepared devices is often too large to implement them on a large scale. A significant part of this variability may be attributed to variations in the structure of the individual filaments. The purpose of this project is to characterize the shape of these filaments using atomic force microscopy (AFM) and establish the relationship between filament shape and device-level electronic characteristics, as well as to examine a factor by which filament shape may be controlled: the current at which devices are formed.

Resistance Switching Devices

Metal-oxide-metal resistance switching devices (RRAM) consist of three layers, conducting/insulating/ conducting. As the oxide layer approaches the nanometer scale, the device still maintains a high resistance in this layered configuration. Under the influence of an electric field, however, the metal from the positive contact may ionize and be driven into the insulating layer. These ions eventually form a conductive filament, changing the resistance state of the device. This process is also semi-reversible: upon a reverse bias, these filaments can be disrupted, and the device resumes a resistive state. This effect has the potential to be used as a component in neuromorphic computing called a memristor [1]. Similarly to the mechanism by which a neuron connection is strengthened according to the amount of current that has flowed through it, the resistance state of an RRAM device is dependent on the electronic history of the device, and could be used to imitate this connection strengthening effect in a neuromorphic computer.

Precise control of the electrical characteristics of these has proved difficult, limiting large scale usage. Many different factors (e.g. dielectric crystallinity, contact interface roughness, capacitance across the device) can confound forming characteristics we control. As to what aspects of the filament we can control, our understanding is limited. At the moment, while there have been studies demonstrating location and contact area of the filaments [2-3], and even some claiming to control filament size [4], attempts to obtain an idea of the shape of the entire filament have been lacking. Filament shape should, however, have a profound effect on the conduction dynamics through a filament, similar to shape effects seen in nanowires [5].

Objective

We seek to establish a means of obtaining an accurate picture of the filament structure embedded within the dielectric layer. To accomplish this, atomic force microscopy (AFM) will be used to obtain an image stack, essentially a tomogram. However, the relative placement of subsequent images in this tomogram are not credibly obtained from the experimental procedure. Thus, a post-processing method is devised herein to align these images.

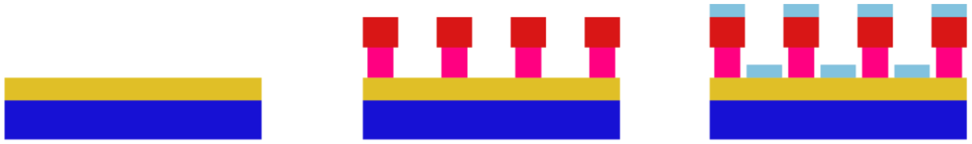
The main goal of this project is to leave behind a post-processing script which can be readily applied to any future AFM tomograms obtained in this manner so that experimental datasets, those which would examine the relationship between structure, forming parameters, and formed properties, can be analyzed.

CHAPTER II

METHODS

Device Fabrication

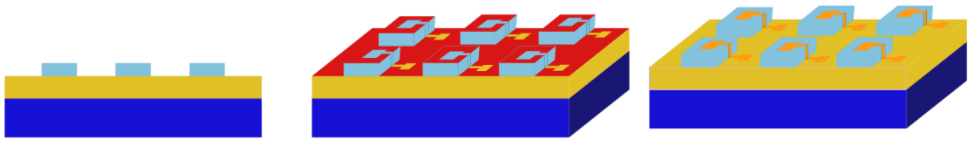
Two device substrates are used, p+Si and TiN. On top of these substrates, HfO_2 is deposited using atomic layer deposition, then annealed at 400 or 600 ^{o}C for 1 hour to induce crystallization, resulting in 3 different crystallization treatments, where an amorphous sample has been left untreated (not annealed). Contact isolation stages are at this point sputtered on, and copper device contacts and connections are then deposited using e-beam and photolithography. The end result is a device which is defined by the small top contact (Figure 1 e-f) on a large plane of HfO_2 and substrate material. This top contact is then connected to a larger contact isolated from the HfO_2 layer.



(a) As received samples of HfO_2 on p+Si substrates

(b) Dual photoresist patterning for lift-off

(c) SiO_2 deposition by RF sputtering



(d) Lift-off of photoresist and SiO_2 in acetone/stripper

(e) Patterning of photoresist for copper devices and bondpads

(f) Cu deposition and lift-off in acetone or photoresist stripper

Figure 1: *Cross sections of sample at various stages in device fabrication*

Electrical Characterization

Once the devices are fabricated, the virgin devices underwent the initial filament forming process. To perform this process, we used a Keithley 4200-SCS semiconductor parameter analyzer. A constant bias stress ranging from 3-15 V was applied to the devices for some time to form filaments through the material. Upon formation, there is a current spike, which the Keithley is able to limit to a compliance current (100 µA in our case). However, there is a lag between the current spike and the compliance circuitry, resulting in an unknown current overshoot to flow through the device for ~100 µs.

Atomic Force Microscopy

Prior to microscopy, the top electrode is removed. A solution of 35 wt% $FeCl_3$ and 6 wt% HCl was diluted to .1 wt% $FeCl_3$, and the devices were immersed in the solution for 20 s (long enough for the top contact not to be optically visible). This process removes the copper by oxidizing the copper and diffusing the resulting compound into the solution. A simplified reaction is given below:

$$2FeCl_{3(aq)} + 3Cu_{(s)} \rightarrow 3CuCl_{2(aq)} + 2Fe^{2+}(aq)$$

Once the top contacts are etched off, the devices are examined using a Bruker Icon atomic force microscope. A conducting tip (doped diamond coated Bruker DDESP-V2, f_o = 400 kHz, k = 80 N/m) was used to initially scan the entire surface which once was under the top contact. A sample bias of -5.0 mV was used to obtain conductivity maps, which is low enough to avoid disturbing the filament structures. The Bruker has a sensitivity range of 1-12 nA, so this low bias was sufficient to obtain a signal. After a preliminary scan is performed, any conductive spots are singled out for closer examination. To these areas a scratch-through method is used.

A 3D scalpel technique was used: this involves using a hard, conductive tip (the doped diamond used previously), to scan the surface at a high setpoint (this corresponds to a large force exerted by the tip). When the tip scans the surface in this manner, a small layer of the sample is actually eroded away. This process is repeated to obtain a series of scans which are gradually deeper and deeper into the sample. In this case the technique was used to scratch through the

HfO_2 layer until the bottom substrate was reached. The rate at which the tip eroded the film was determined by at various points stopping the scan and performing a non-scratching scan of a larger region, giving the depth of the scratched trench. Etch rate is assumed to be linear between these reference scans. Taking multiple of these reference scans is necessary to compensate for erosion of the microscope tip itself, resulting in decreased etch rate as the scan progresses.

CHAPTER III

ANALYSIS

Alignment Procedure

Coming off of the raw experimental collection, a series of topography maps are obtained, as seen in figure 2 (a). The relative depths of these layers are obtained by interpolating between the measured depths of the “well” that has been dug (measured by aforementioned calibration scans). In this raw form, however, the data tells us little about the true structure of the filament in question. To demonstrate this, a preliminary reconstruction of the filament was created by selectively plotting the height data corresponding to conductive regions offset by the estimated well depth. This preliminary reconstruction is seen in figure 2 (b and c). From this reconstruction the following issues are evident: 1) the presence of instrumental drift, creating a slanted filament structure, 2) significant kinks and in some occasion discontinuities in the filament corresponding to the calibration scans (especially seen in figure 2b around $z = -1$ and -8 nm).

To obtain a filament structure which we can assert as a characterization of the *in situ* structure, this data must be aligned, as there is no experimental calibration for relative xy position of successive scans. The only data available for us to use to this end is the height data and conductivity data. An obvious option is to align the scans using the conductivity data, but this would only produce questionably accurate reconstructions, as it would inevitably produce conductive filaments which proceed in a beeline from top to bottom, with little lateral motion. This does not necessarily reflect the true structure of the filament, but it does prevent discontinuities. Thus, this alignment scheme is only to be used as a last resort.

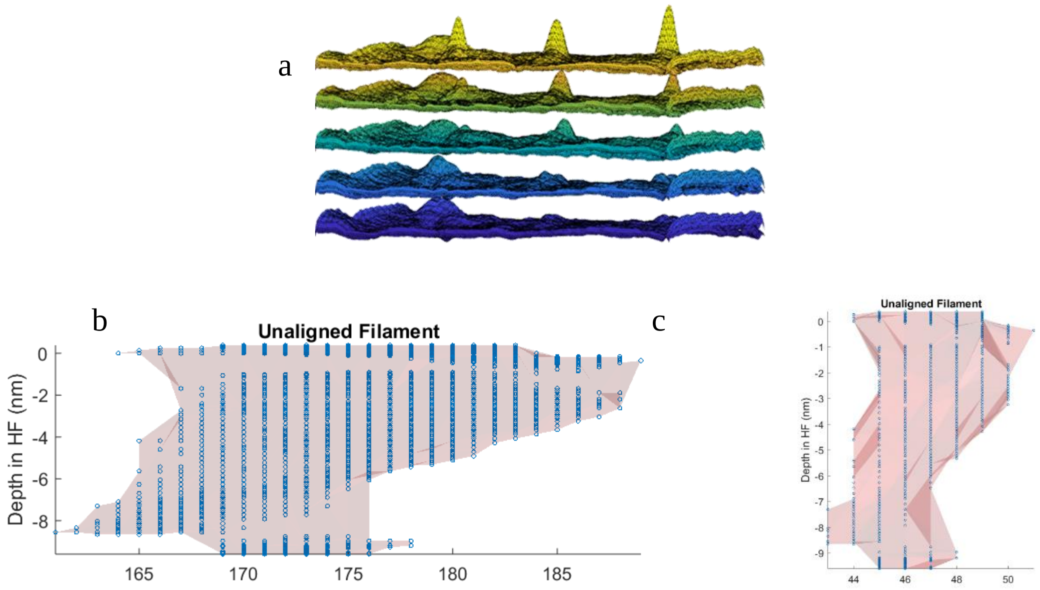


Figure 2: *a) Sample of 5 scans showing the progression of topography change in the sample, b and c) preliminary reconstruction of the filament structure using the unaligned data along the x and y directions.*

The other option is to use the height data to align the scans. This was immediately perceived as the favorable option considering that although the filament may progress laterally between scans, topographical features (on average) should not. This is because our experimental method ideally peels off layers of material from whatever topography is preexisting (with some exceptions allowing for local variations in hardness). Thus, several schemes were devised to align these topographical features.

Correlation Coefficient

The quantity subsequently referred to as the correlation coefficient of two layers (arrays of equal size) is calculated as follows: one array (a) is cropped to some smaller size, and for every possible placement of **a** onto the uncropped array **b**, the two are elementwise multiplied, and the resulting array is summed. This gives a parameter for "how well the two arrays correlate to each other", which we refer to as the correlation coefficient. The interpretation of this coefficient may vary depending on the data being used. The shift of array **a** (dx, dy) which gives a maximum value of the correlation coefficient can later be used to align the array by shifting each image by dx, dy pixels with respect to the previous.

Binary Map Scheme

The first scheme involves using adaptive thresholding on each layer to produce binary maps of regions that are significantly different from the surrounding area, thus generalizing our height maps to binary maps of “is a feature” or “is not a feature”. As opposed to global thresholding, which would mark anything as a feature which was above a certain global cutoff value, this algorithm assigns a local threshold based on points in the immediate surroundings. A sample feature map can be seen in figure 3. In this scheme the correlation coefficient represents a sum of an **AND** logical operation, thus corresponding to the amount of total overlap of feature regions.

Figure 3: *A sample feature map showing feature regions (white) and featureless regions (black)*

Note the raw data shown in figure 4 (a). It has a distinctive sawtooth characteristic, presumably corresponding to skipping of the microscope tip due to friction. This poses an issue for the adaptive thresholding, as it produces false features over much of the sample. Since the algorithm considers only immediately surrounding data, it may return the peak of each sawtooth as a feature. This is remedied by performing a matrix convolution on the data. This involves a sum of the matrix product of the sample with a smaller matrix (all values equal to 1). The result is in a sense a moving average which preserves more of the underlying structure than a simple moving average. The convoluted surface can be seen in figure 4 (b)

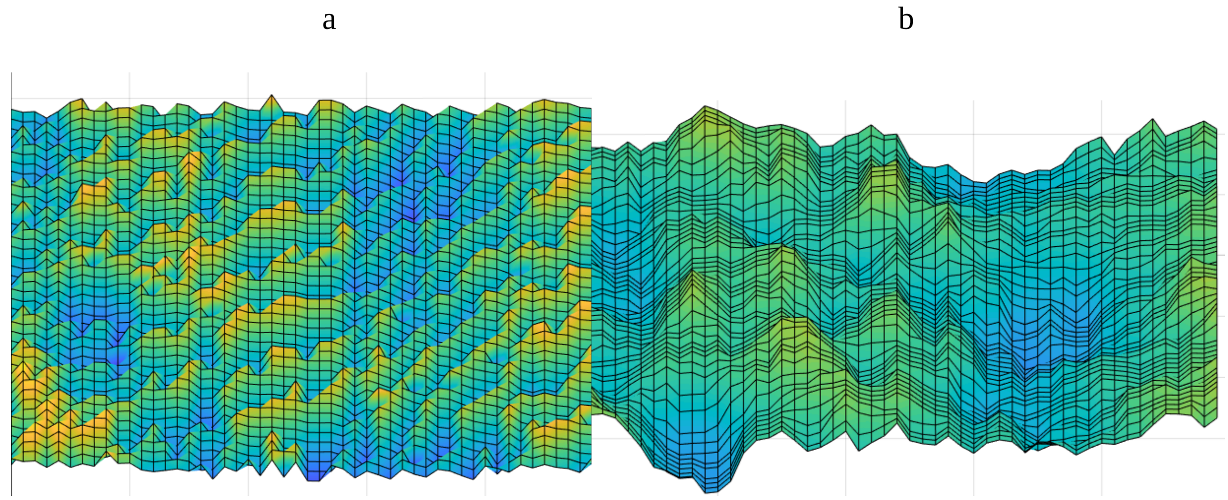


Figure 4:* *a) *Raw height data, showing sawtooth character.* ***b)*** *Convoluted data (1/10 total height range on the convoluted data as compared to the raw data). Note while the surface retains some roughness, it is more continuous with immediate surroundings.*

This having been accomplished, binary maps could now be drawn which were marginally more useful. However, another issue quickly arose. Over the course of the scratch-through process, levels could progress over many scans from a rough to a smooth topography. This leads to total feature area trends as seen in figure 5 (first binarization). To avoid this issue, the following process was used: perform the binarization as before; from the total feature area vector graphed in figure 5 (1st binarization), determine a target feature area ($c\bar{a} \pm \boldsymbol{\sigma}/\mathbf{4}$, where c is some arbitrary sensitivity, two fifths in the case shown, and $\bar{a}$ is the average feature area of all scans);

repeat the binarization on each layer, adjusting the sensitivity of the adaptive threshold algorithm until a feature area in the target region is obtained. The results of this process are also seen in figure 5 (series 2nd binarization). Note that the total feature area remains relatively constant. This is desirable because trying to align a high feature area map with another map will create many false maxima in the correlation coefficient, as most shifts will create a significant overlap of features. Increasing area for low area scans is also desirable as smooth scans (around layer 70) sometimes have no initial features, posing an obvious issue for alignment. Forcing the feature area to a uniform value lessens this risk of false alignments.

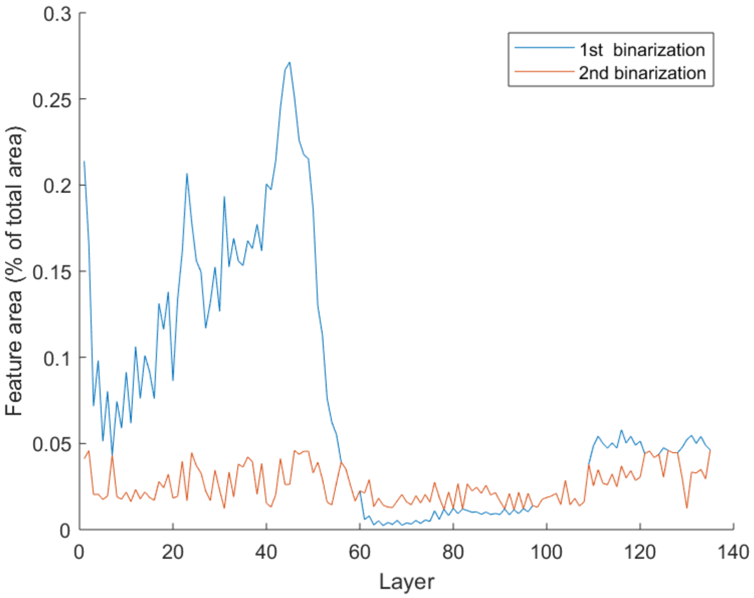


Figure 5: *Shows the total feature area as scans progress. High area corresponds to high roughness, low area corresponds to a smooth and relatively featureless scan.*

One hidden benefit of this scheme is that neither array must be cropped before computing the correlation coefficient. This is allowed because much of the array contributes nothing to the sum (1*0 or 0*0), so the correlation coefficient does not lose any meaning when the two arrays are not completely overlapping. As a result, features close to the edge of the image are able to be used for alignment.

Height Map scheme

The same correlation procedure may be run with the analog height data in lieu of the binarized data. In our case we used the convolution smoothed height map instead of the sawtooth-surfaced raw data. This presents an issue as the interpretation of the correlation coefficient is somewhat unclear. Since it represents an elementwise product of heights, rough scans with a high average height may correlate well to many different scans. However, the correlation algorithm was given a normalization factor to try and account for this effect. This normalization is essentially replacing the height data with distance from the average height value of each scan scaled by the standard deviation.

This scheme allows for false maxima to occur if a mound is aligned to a taller mound, rather than a mound similar in height to itself. With large arrays being aligned this should not be a problem, but it is a risk (especially for smoother layers). This scheme does however carry the benefit of not losing any information in the alignment process: it takes into account the entire continuous range of the initial height data, instead of reducing each height to a logical bit.

Future Directions

There exist several issues with the algorithms in their current state. These issues and possible solutions will now be discussed. These represent directions we intend to pursue before the end of the academic year that are outside the temporal scope of the thesis project.

Large Array Alignment

At the moment, the alignment algorithms are limited to single index shifts, that is they are only capable of aligning the data to increments of 1 pixel. However, the factors that affect the misalignment of the scans (drift, debris buildup, discontinuities in scanning procedure), are inherently analog features. While overshoot should be corrected, this effect imposes an upper limit to how close our alignment could be to the true filament structure. However, this issue can be resolved by enlarging the arrays which are being aligned. This enlargement either be in the form of replication (enlarging individual pixels, allowing truly finer alignment), or by interpolating points on the height map (the analogous treatment for the binary maps would be bilinear or bicubic resizing). Performing either of these, resizing the array by a factor c would allow for a maximum alignment resolution of $1/c$.

CHAPTER IV

RESULTS

The results of this project fall into several categories which need to be discussed: the experimental data obtained, an assessment of which alignment scheme better aligns topographic features, and whether the alignment of topographic features has any effect on the resulting filament structure

Experimental Results

Shown in table I are all the scratch-through scans which were carried out. The scans highlighted in green are the scans which were performed after my inclusion in this project. The first three were the basis of previous work largely focused on refining our use of the tomography technique itself. The scans which are used for alignment and from which we attempt to reconstruct filament structures are the last few, scans 5-8, which scratched through the entirety of

TABLE I: Tomogram collection

	Thickness (nm)	*Crystallinity*	*Forming*		*Summary*
1	30	P	4 V	1 mA	30 nm hill, partial scratched 2 nm through Hf layer
2	30	P	4 V	1mA	18 nm hill, scratched 25 nm through Hf layer
3	30	P	4 V	100 μA	10 nm hillock, conductive signal ended abruptly
4	10	A	5 V	100 μA	3 small (<10 nm) hills, partial scratch 3.5 nm through
5	10	A	5 V	100 μA	50 nm hill, scratched full through
6	10	P	4 V	100 μA	10 nm hill, scratched full through
7	10	P	4 V	100 μA	5 nm hill, scratched full through
8	10	P	4 V	100 μA	5 nm hill, scratched full through

the thin dielectric layer. In the region where a conductive filament was found on each device (the head of a filament) was also found a large hillock on the dielectric surface. Given that our AFM methods do not allow for compositional characterization, it is impossible to say for sure what these are and when they formed, but they are linked to the filaments. Possibilities include a region of some copper compound that was immune to the etchant used to remove the copper electrode or possibly a preexisting artefact of some stress-relaxation process [6].

In choosing an experiment on which to test our alignment software, we were looking for a dataset which had slight drift, showed no dramatic mechanical deformations, and showed no abrupt loss of conductive signal. In several of these experiments, the root of the filament (near where it connects to the bottom electrode) flaked off. This appeared as, in two successive scans, a small raised region in the first, which then in the second showed as a pit extending well into the bottom electrode. In another experiment, we "lost" the filament partway through: the conductive signal abruptly ended partway through the dielectric layer. Factoring in these requisite criteria, it was settled that we would use experiment 6 (boxed in table I). Successive scans are still being taken, building up a data set of experimental interest rather than to calibrate the analysis and technique, in the hope that will be analyzed with the post-processing discussed. We will now proceed to a discussion of the results of the attempts to align the data obtained in experiment 6.

Alignment Results

To obtain a figure of merit which measures 'how well the alignment algorithm aligns topographic features', we turn once again to the correlation coefficient. In this case, we will use the normalized correlation coefficient. For two images f_1 and f_2, the correlation coefficient (this time comparing two full images, rather than a full image and a subset of another):

$$\boldsymbol{\gamma} = \sum_{x,y} \frac{f_1(x,y) - \overline{f_1}}{\sigma_1} * \frac{f_2(x,y) - \overline{f_2}}{\sigma_2} \tag{1}$$

Note that this still represents a sum of elementwise products, but now the elements translated into distances from the mean value of the image, and each is scaled by the standard deviation of each image. This value now represents how well two images match. Now we may take f_1 to be any layer in the image stack, and f_2 to be the subsequent layer. By comparing the $\boldsymbol{\gamma}$ values for the unaligned stack and the aligned stack, we are now able to compare how much better subsequent images match after the alignment.

To compare between the two alignment schemes we wish to compare the quantity $\frac{\gamma_{aligned}}{\gamma_{unaligned}}$. This represents the percent change in 'matching-ness' caused by the alignment algorithm, which we will refer to as the alignment gain. In figure 6, this gain is graphed for each layer, for both the height alignment scheme and the binary alignment scheme.

Note that the height map alignment scheme performs adequately (a gain greater than one, an increase in alignedness), until it reaches the smooth region around scan 60. At this point, the distinguishing features are too small, and also near the edges of the scan (the height alignment algorithm requires cropping of the layer to be aligned). It quickly dips into a loss in alignment: overlap of distinguishing features decreases. Even in this difficult region, however, the binary scheme still creates a positive gain. Since the algorithm works by maximizing the overlap between the feature maps, gamma never drops below one. In every scan, the overlap of features either stayed constant (no shift) or increased.

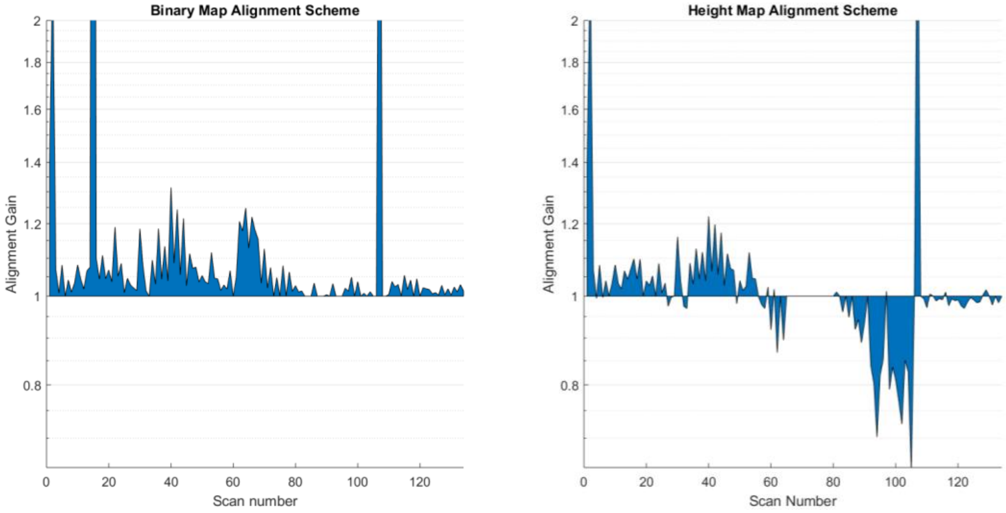


Figure 6: *Shown are the alignment gains between successive scans the feature maps for the data aligned via the binary scheme as well as the height scheme. Values less than one signify a decrease in "aligned-ness", greater than one an increase.*

Impact on Filament Structure

Shown in figure 7 is the filament structure before and after alignment. As was mentioned before, this entails selectively plotting the topographic data which corresponded to a conductive pixel, offset by the estimated depth of each scan. Shown in figure 7 is a boundary surface obtained from a scatter plot of these data. Much of the assessment we can now perform on the result is now qualitative. Even so, there is visible improvement. Notice in the long axis view in figure 7a, there is a waving of the filament seen along the short axis which is removed by the alignment. An amount of what seems to be instrument drift is also corrected along the long axis (figure 7b). Also, the aligned structure shows more clearly what is occurring at the beginning and end of the experiment, namely the redistribution or uncovering of the hillock at positive *z*, as well as some penetration into the (conductive) bottom electrode at $z < -10$ nm. In terms of overall shape, the aligned filament more closely displays the funnel-like structure which was expected.

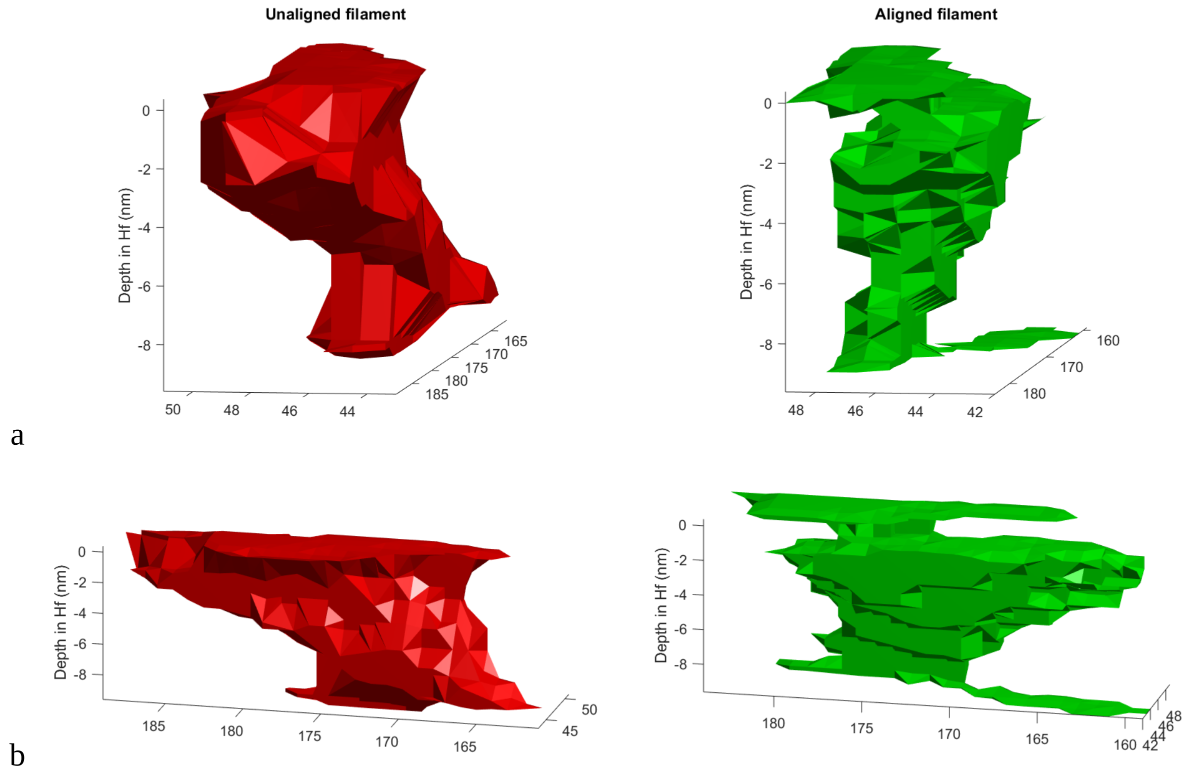


Figure 7: *View of the filament along its long (a) and short (b) axes. Real space depth shown on z axis, where 0 is the top of the Hf layer, -10 is the bottom electrode (visible in the aligned filament). Dimensionless matrix index on x, y axes.*

CHAPTER V

CONCLUSION

Herein we have refined a means of experimentally obtaining a tomogram of conductive filaments in resistance switching devices by introducing a post-processing alignment script. This alignment script provides an increase in similarity of consecutive images in the tomogram stack by shifting them in the in-plane directions. Two schemes to determine this shift were devised, based on aligning the topographic data as an image versus aligning binary feature maps obtained from the topographic data. Of the two, the binary scheme resulted in a more reliable increase in the similarity between consecutive images. This method was shown to produce a gain in similarity in regions where there are few topographic features as well as in rough regions with a preponderance of topographic features. This was accomplished by recursively adapting the binarization sensitivity to obtain maps with a constant feature area.

This approach was then applied to one such tomogram, and the resulting alignment is observed. The shape of the conductive filament is seen to have been corrected for instrument drift and observation discontinuities. The aligned scan also returns a filament shape that confirms the expected filament shape (funnel-like).

Future Directions

Two future directions in which the post-processing itself may be taken have also been discussed, especially padding the data with empty space in order to provide a finer resolution in the shifting process. Any fractional shift should be compensated for eventually by subsequent shifts, but more precision is desirable in this process.

The script in its current form is able to be applied to any future data collected by our atomic force tomography technique, as everything that was done to the data by the script is general to the form of the data obtained by this method. Armed with this post-processing script, the final piece of the process in the search for a means of spatially characterizing embedded conductive filaments, experiments are opened to our reach which would relate device forming parameters to structure, elucidating the effect of these parameters on the final device properties.

REFERENCES

[1] Williams, R. S. (2017). What's Next? Computing in Science & Engineering, 19(2), 7–13. https://doi.org/10.1109/MCSE.2017.31

[2] Bersuker, G., Gilmer, D. C., Veksler, D., Kirsch, P., Vandelli, L., Padovani, A., … Nafría, M. (2011). Metal oxide resistive memory switching mechanism based on conductive filament properties. Journal of Applied Physics, 110(12). https://doi.org/10.1063/1.3671565

[3] Brivio, S., Tallarida, G., Cianci, E., & Spiga, S. (2014). Formation and disruption of conductive filaments in a HfO2/TiN structure. Nanotechnology, 25(38), 385705. https://doi.org/10.1088/0957-4484/25/38/385705

[4] Nardi, F., Ielmini, D., Cagli, C., Spiga, S., Fanciulli, M., Goux, L., & Wouters, D. J. (2011). Control of filament size and reduction of reset current below 10 lA in NiO resistance switching memories. Solid State Electronics, 58, 42–47. https://doi.org/10.1016/j.sse.2010.11.031

[5] Durkan, C., & Welland, M. E. (2000). Size effects in the electrical resistivity of polycrystalline nanowires. Physical Review B, 61(20), 14215–14218. https://doi.org/10.1103/PhysRevB.61.14215

[6] Sarobol, P., Blendell, J. E., & Handwerker, C. A. (2013). Whisker and hillock growth via coupled localized Coble creep, grain boundary sliding, and shear induced grain boundary migration. Acta Materialia, 61(6), 1991–2003. https://doi.org/10.1016/J.ACTAMAT.2012.12.019