Cross-Departmental Post Graduate Course “Optics & Vision”

Department of Medicine

University of Crete

Master Thesis:

Development, realization and assessment of appropriate algorithms for 3D image fusion in Light Sheet Microscopy

Author: Kyparissidis Kokkinidis Ilias

Supervising committee: A. Zacharopoulos, G. Zacharakis and H. Ginis

Evaluation committee: M. Taroudakis, D. Papazoglou, G. Makrakis

Διατμηματικό Μεταπτυχιακό Πρόγραμμα “Οπτική & Όραση”

Τμήμα Ιατρικής

Πανεπιστήμιο Κρήτης

Μεταπτυχιακή Διπλωματική Εργασία

Κατασκευή, υλοποίηση και αξιολόγηση αλγορίθμων, κατάλληλων για τρισδιάστατη σύντηξη εικόνων σε μικροσκοπία φθορισμού πεδίου φωτός.

Κυπαρισσίδης Κοκκινίδης Ηλίας

Επιβλέπον επιτροπή : Α. Ζαχαρόπουλος, Γ. Ζαχαράκης, Χ. Γκίνης

Εξεταστική επιτροπή : Μ. Ταρουδάκης, Δ. Παπάζογλου, Γ.Μακράκης

ΕΙΣΑΓΩΓΗ

Η σύντηξη εικόνων είναι η διαδικασία με την οποία συνδυάζουμε δεδομένα από δύο ή περισσότερες αρχικές εικόνες σε μία, με όλες τις απαραίτητες πληροφορίες των εικόνων προέλευσης. Οι μέθοδοι σύντηξης εικόνων έχουν εφαρμογές σε πολλές διαφορετικά πεδία, όπως τηλεπισκόπιση , κάμερες CCD, ιατρική απεικόνιση και μικροσκοπία. Στην περίπτωσή μας, θα θέλαμε να χρησιμοποιήσουμε μεθόδους σύντηξης εικόνας για να βελτιώσουμε τις εικόνες που αποκτήθηκαν από μικροσκοπία φθορισμού πεδίου φωτός(LSFM) ή μικροσκοπία φωτισμού επιλεκτικών επιπέδων (SPIM). Το SPIM είναι μια τεχνική μικροσκοπίας φθορισμού, στην οποία το δείγμα φωτίζεται από ένα λεπτό πεδίο φωτός λέιζερ, κατά μήκος μιας οπτικής διαδρομής ορθογώνιας προς τον άξονα ανίχνευσης. Δεδομένου ότι μόνο το τμήμα εστίασης του δείγματος εκτίθεται σε φως λέιζερ, επιτρέπει την οπτική τομή του δείγματος σε υψηλό βάθος διείσδυσης και ουσιαστικά μειώνει τη λεύκανση φωτογραφιών. Ωστόσο, όταν ασχολείται με ένα μεγάλο ή εξαιρετικά σκεδαζόμενο δείγμα, η ποιότητα των εικόνων επιδεινώνεται, με αποτέλεσμα την απώλεια πληροφοριών. Για να αυξηθεί το πληροφοριακό περιεχόμενο των εικόνων SPIM και να αντισταθμιστούν οι περιορισμοί του βάθους διείσδυσης, λόγω της απορρόφησης και της σκέδασης φωτός, αποκτάμε σαρώσεις από διαφορετικές προβολές του δείγματος. Στη συνέχεια οι εικόνες αυτές συνδυάζονται με τρόπο τέτοιο που να παρέχει πλήρη εικόνα του δείγματος. Για να γίνει αυτό πρέπει να ευθυγραμμίσουμε με ακρίβεια, και να συγχωνεύσουμε αυτές τις εικόνες. Η ευθυγράμμιση εικόνας είναι η διαδικασία με την οποία μεταβάλλονται διαφορετικά σύνολα δεδομένων σε ένα σύστημα συντεταγμένων. Ορίζουμε ένα μετασχηματισμό T που θα χαρτογραφήσει μία εικόνα σε μια άλλη εικόνα ενός παρόμοιου ή του ίδιου αντικειμένου κατά τρόπο που μεγιστοποιεί την ευθυγράμμισή τους. Έχουμε παρουσιάσει προηγουμένως μια τεχνική που χρησιμοποιεί εικόνες που αποκτήθηκαν με τον τρόπο που αναφέρθηκε παραπάνω, και με άκαμπτη και μη άκαμπτη ευθυγράμμιση εικόνας ευθυγραμμίζει τις εικόνες SPIM. Σε αυτή τη διατριβή παρουσιάζουμε και συγκρίνουμε διαφορετικές τεχνικές που συντήκουν τις ήδη ευθυγραμμισμένες εικόνες για να δημιουργήσουμε μια τελική τρισδιάστατη εικόνα υψηλής ανάλυσης.

Table of Contents

Εισαγωγή 3
1 Abstract 6
2 Introduction 7
2.1 Selective plane illumination microscopy (SPIM) 8
2.2 Image registration 10
2.3 Image Fusion 12
3 Image Fusion Theory 14
3.1 Content based Fusion 14
3.1.1 *Maximum* Intensity 14
3.1.2 Minimum Intensity 14
3.1.3 Average intensity 15
3.1.4 Principal Component Analysis (PCA) based Fusion 15
3.1.5 Focus measure operator based fusion 16
3.2 Transform based fusion 16
3.2.1 Abstract Idea in transform based fusion 16
3.2.2 Wavelet Transform Based Fusion 18
4 Comparison or Image Quality metrics 29
4.1 Histogram, Joint Histogram 29
4.2 Reference Metrics 32
4.3 Non- Reference Metrics 34
5 Experimental work 38
6 Results 39
6.1 Multi-View Fusion 39
6.2 Multi-Spectral Fusion 48
7 Conclusions 51
8 References 53
9 Acknowledgements 55

1 ABSTRACT

Image fusion is the process by which we combine data from two or more source images to a single image with all the necessary information of the source images. Image fusion methods have found applications in many different areas and fields such as remote sensing, CCD cameras, medical imaging and microscopy. In our case we would like to use image fusion methods to improve images acquired from Light Sheet Fluorescence Microscopy (LSFM) or Selective Plane Illumination Microscopy (SPIM). SPIM is a fluorescence microscopy technique in which the specimen is illuminated by a micrometer thin focused laser sheet, along an optical path orthogonal to the detection axis. Since only the in-focus part of the specimen is exposed to laser light, it allows for optical sectioning of the specimen at high penetration depth and substantially reduces photo-bleaching. However, when dealing with a large or highly scattering sample the quality of the images deteriorates, resulting in loss of information. To increase the information content of the SPIM images and compensate for the penetration depth limitations, due to the light absorption and scattering, we acquire scans from different views (projections) of the sample and then combine them in a way that provides a complete view of the sample. To do that we need to accurately register and fuse these images. Image registration is the process by which different sets of data are transformed into one coordinate system; we define a transform T that will map one image onto another image of a similar or same object in a way that maximizes their alignment. We have previously presented a technique that uses images acquired as mentioned above, and by rigid and non-rigid image registration aligns the SPIM images. In this thesis we present and compare different techniques that fuse the already aligned images to create one final high resolution three dimensional image.

2 INTRODUCTION

Optical investigation has always been a fundamental tool of biological research. Since the implementation of optical microscopy the field has being making strides by being able to image biological processes in great detail. However, its ability for *in vivo* tissue imaging has been restricted by light absorption and scattering to superficial investigations. Nowadays, with the use of even more advanced tomographic techniques we are able to visualize biological processes in detail without causing the death of the sample. Nevertheless, even when advanced tomographic methods are used like confocal microscopy, Optical Projection Tomography (OPT) and Selective Plane Illumination Microscopy (SPIM) [1,2,3] light scattering still has a major effect in inhibiting the penetration depth of optical imaging.

The imaging limit of these methods is dictated by the Mean Free Path (MTF) and the Transport Mean Free Path (TMFP) of a photon. The MFP describes the average distance that a photon travels between two consecutive scattering events. The TFMP accounts for the MFP and the average angle by which photons are scattered in each scattering event. Specifically, the TMFP is the distance that photons will travel before losing relation to the original propagating direction, they had before entering the sample, due to multiple scattering events that result in photons becoming diffuse [43]. TMFP can vary a lot depending on the tissue type. Tomographic microscopy and more specifically low cost, in vivo, tomographic approaches such as OPT and SPIM have been shown to be able to image samples up to about 500 μm, that is, at depths where photon scattering occurs[10]. Meaning that as we go deeper into the tissue the quality of the image deteriorates. However, OPT and SPIM have the ability to acquire scans from multiple projections that provide complementary information to each other, to overcome the difficulties of scattering and diffusion in tomographic imaging when visualizing larger samples and organisms. This thesis discusses methods that utilize this ability for multiple projections in SPIM and combines them in an optimal way that provides a clean volumetric image of the samples.

2.1 SELECTIVE PLANE ILLUMINATION MICROSCOPY (SPIM)

Light Sheet Fluorescence Microscopy (LSFM) or Selective Plane Illumination Microscopy (SPIM) is a fluorescence microscopy technique in which the specimen is illuminated from the side by a thin (micrometer-thick) focused laser sheet at excitation wavelength, along a separate optical path orthogonal to the camera detection axis. Since only the in-focus part of the specimen is exposed to laser light, it allows for optical sectioning of the specimen at high penetration depth and substantially reduces light bleaching [8, 9]. Volumetric image formation is then based on stacking together the section images, each collected after translating the sample in relation to the light plane, in a step-wise manner that scans the entire volume of interest. Moreover, the signal emitted from the in-focus section is detected in parallel for the entire field of view, which provides high imaging speeds.

In addition SPIM can be used to acquire multiple scans (volumetric images) from a specimen that offer complementary information to each other. The usual methodology for acquiring the multiple scans requires that after we piece together a volumetric image by translating the sample in relation to the light plane, we can then introduce a rotation of the sample as shown in the Figure below and reacquire a volumetric image of the sample from another perspective. This process can be repeated to acquire any number of volumetric scans (projections) from different angles but most commonly the four (0^0 ,90^0 ,180^0 ,270^0) projections are considered.

In recent years SPIM, has been successfully used to monitor developmental processes for several days, for example, in zebrafish embryos, drosophila melanogaster embryos, and to visualize all muscle in the transgenic Medaka line Arnie [10].

In addition to the multiangle acquisition, SPIM, as a fluorescence technique can also be used for multispectral imaging. For example, as we will see in chapter 6.2, scans of green fluorescence protein (GFP) and dsRED factors (which are proteins that emit at different wavelengths and are used to target distinct parts of the specimen), can be acquired in sequence from the same specimen. The spectral volumetric images are then combined to offer a more complete 4-dimensional view of the animal.

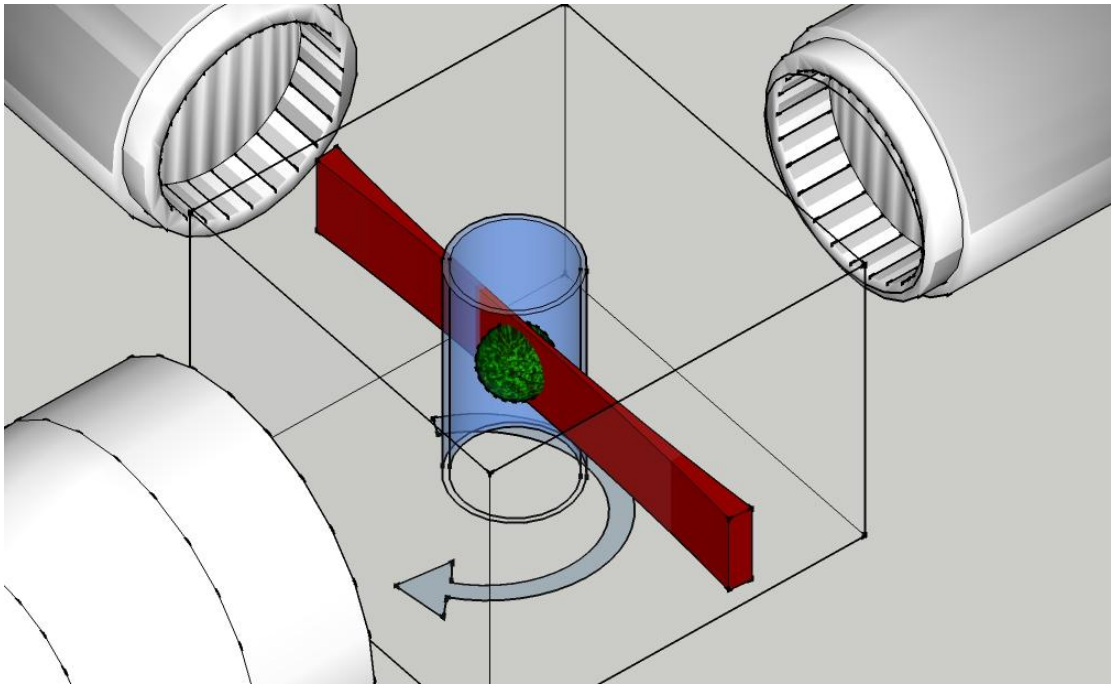


Figure: A thin plane of light is used to illuminate the specimen at a plane perpendicular to the optical detection axis. In addition we can the rotate the sample to acquire multiple views of it.

Even thought, as we have explained above, the creation of the tomographic image for SPIM is not requiring the use of a complex reconstruction algorithm, like OPT, the extraction and combination of the useful and good quality information from the different projections, and possible multispectral volumes, is a significant challenge.

This is where the process of *image fusion* becomes essential. As we can see in the following figure the direct superposition of a low resolution image with a high resolution image would most likely result in a combined lowered resolution image. Therefore, it is not recommended to just co-register and superimpose the different images since that would defeat the whole purpose of the multi-view scans.

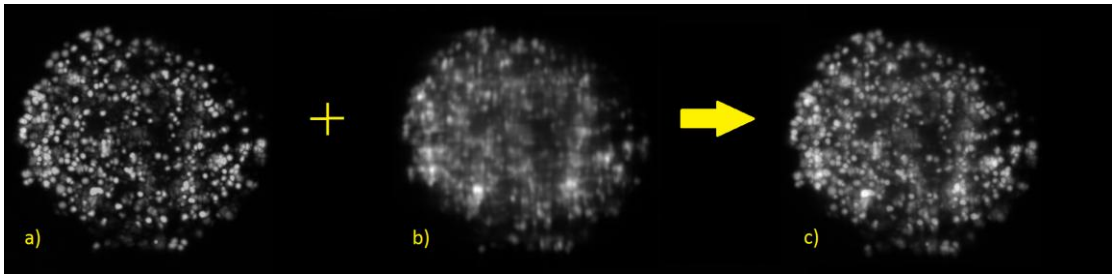


Figure: (a) A high resolution image and (b) a low resolution image and(c) their combination without using fusion methods, where we can see the degradation of the original high resolution

The general idea behind the fusion of multi-view scans from SPIM, is to separate the high resolution content from the low resolution content within the different images and then choose to combine the high resolution content from both the images into a new high resolution image of the whole sample. We should point out that in order for

a successful image fusion to achieve its goal the original images are required to be aligned. This alignment of the original volumetric images is the subject of volumetric image registration.

2.2 IMAGE REGISTRATION

Image registration is the process by which any misalignment issues between images coming from different coordinate systems are resolved by transforming different sets of data into one common coordinate system. Examples, include images acquired from different angles or at different time points or even by different modalities that would display different sources of contrast. In medical imaging registration has contributed to compensate for deformations due to breathing circles or studies of anatomical changes, such as in comparative with atlas images representing a healthy subject [11,12,13,14]. In the following, we will discuss the registration process of misaligned images derived from different perspectives (same fluorescent sources acquired from different views) and from different contrast sources, such as the images in multispectral SPIM (different fluorescence targets in the same subject, excited by different wavelengths). Then, we will describe the different transformation actions that are involved in the registration process.

For the co-registration of the SPIM example images, usually a rigid and (or) an affine transformation process is sufficient, since our data exhibits no local dissimilarities. Rigid transformations are defined as the geometrical transformations that preserve all distances, the straightness of lines, the planarity of surfaces and all non-zero angles between straight lines. These transformations allow only for the translation and the

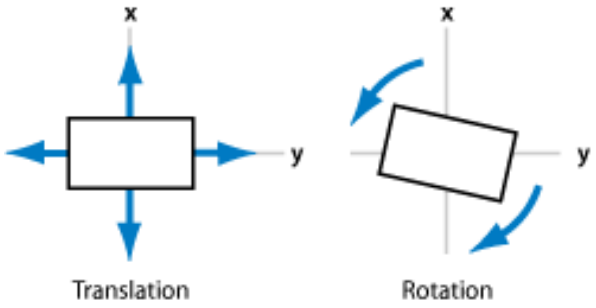
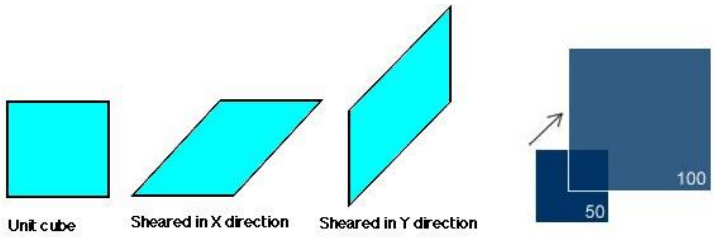


Figure: Rigid transformations

rotation of the image.

Similarly, affine transformations preserve the straightness of lines, and hence, the planarity of surfaces, and also parallelism, but they allow angles between lines to change. Therefore, affine transformations include all the rigid transformations with the addition of the Shear and Scale transformations.



Affine tranformations

Registration is achieved by minimization of a cost functional. There are two main types of cost functions, that we used in our application depending on if we have to register multiangle images (mono-modal) or multispectral images (multi-modal). For the mono-modal, multiangle scans the cost function is based on a similarity metric such as the squared difference of voxel intensities. For the multi-modal, multispectral images, the cost function is chosen to be normalized mutual information (NMI), a metric which is not sensitive to intensity changes. NMI is ideal for images of different spectral information originating from different targets inside the specimens [13, 14]. To transform the images a rigid or affine transformation model applied globally (i.e., to the whole image) is used.

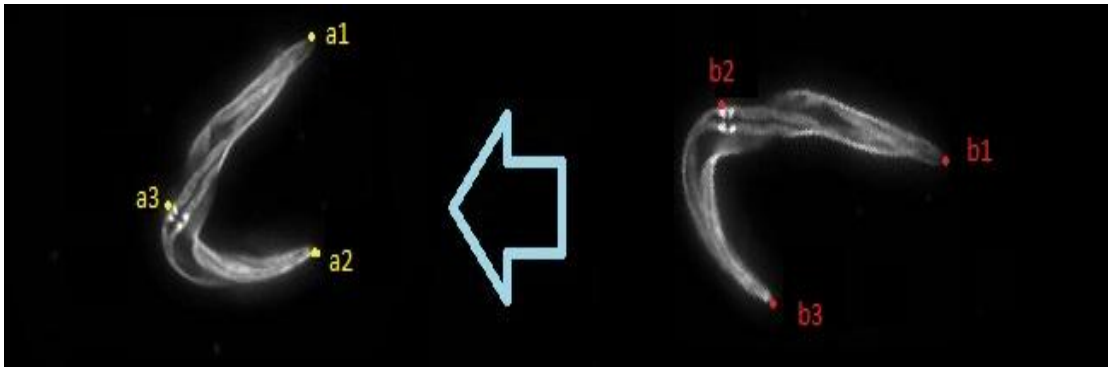


Figure: Multiangle registration, matches different projections of a c-elegans.

Following, a successful image registration the fusion can be performed.

2.3 IMAGE FUSION

The term image fusion refers to a method that allows us to extract information acquired in several images and combine that information into a single image. The resulting fused image will have more information than all the original images separately. Also, the fused image will be more suitable for human and machine

information extraction and further image-processing tasks such as segmentation, feature extraction and object recognition etc. Image fusion methods have found applications in many different areas and fields such as remote sensing, CCD cameras, medical imaging and microscopy. In CCD cameras for example, it is usually not possible for all objects imaged, to be in focus due to the lens limited depth of field. This means that in order to have everything in an image in focus, one should acquire multiple images, each with a different focus distance and then apply a fusion process to combine all those images [15]. Another example where image fusion becomes very useful is in medical imaging where it can provide a bridge between different modalities and combine information from a CT scan, for example, with a MRI scan acquired from the same sample in a single image that would provide the information of both. In our case, the reason we need image fusion is that the SPIM imaging system has a dramatic difference between axial and lateral resolution ie. Resolution drops dramatically in the third dimension. For example a typical SPIM dataset in the x-y dimensions will have the resolution of the camera used ie. 1000x1000 pixels whereas in the z direction will have a resolution determined by the number of slices acquired ie. 20-100 slices. That being the case we will study several image fusion algorithms and compare their results to what suits our imaging requirements and system best.

3 IMAGE FUSION THEORY

Image fusion methods are comprised of two major categories. Those that are content based and those that are transform based.

3.1 CONTENT BASED FUSION

These are methods that deal directly with the content of an image that is the intensity values in the pixels or voxels of an image. These methods include for example simple intensity averaging, maximum intensity, and any selection algorithm based on a neighborhood operation.

3.1.1 *Maximum* Intensity

In this method, we compare the intensities of corresponding pixels in the original source images. The resulting fused image is obtained by selecting the maximum intensity of each of these pixels.

$$Imf(i,j,k) = \max\{I_1(i,j,k)I_2(i,j,k) \dots \dots \dots\} \quad (1)$$

$I_1(i,j,k), I_2(i,j,k)$ are the original images and $Imf(i,j,k)$ is the fused image.

Figure: Demonstration of Maximum Intensity image fusion method

3.1.2 Minimum Intensity

Here we have the same process with maximum intensity but instead, in the selection, we choose the minimum intensity when comparing the corresponding pixels or voxels.

$$Imf(i,j,k) = \max\{I_1(i,j,k)I_2(i,j,k) \dots \dots \dots\} \quad (2)$$

Again $I_1(i,j,k), I_2(i,j,k)$ are the original images and $Imf(i,j,k)$ is the fused image.

Figure: Demonstration of Minimum Intensity image fusion method

3.1.3 Average intensity

Here as the name suggests the fused image will be comprised by the average intensity of the corresponding pixels of all the input images.

$$Imf(i,j,k) = (I_1(i,j,k) + I_2(i,j,k) + \cdots + I_n(i,j,k))/_n \quad (3)$$

Figure: Demonstration of Average Intensity image fusion method

3.1.4 Principal Component Analysis (PCA) based Fusion

As an image fusion method, Principal Component Analysis (PCA) determines weights for each of the original source images using the eigenvector corresponding to the largest eigenvalue of the covariance matrix of the original source images [16, 17]. We will follow with an example of a PCA image fusion algorithm scheme.

Step 1: We insert the images that are to be fused.

Step 2: Compute the Covariance matrix using:

$Cov(I_1, I_2) = \frac{\sum_i^n (I_1(i)-\bar{I_1})(I_2(i)-\overline{I_2})}{n-1}$, n=size of our data and

$$C = \begin{bmatrix} Cov(I_1,I_1) & Cov(I_1,I_2) \\ Cov(I_2,I_1) & Cov(I_2,I_2) \end{bmatrix}$$

Step 3: Compute the Eigenvalues and Eigenvectors of the covariance matrix

Step 4: In this step we order the eigenvectors by eigenvalue from highest to lowest. In this way we get a new matrix of the eigenvectors in order (feature vector).

Step 5: We fuse the original images weighted by the respective values of the feature vector.

3.1.5 Focus measure operator based fusion

Usually focus measure operators are used for autofocusing algorithms in cameras [40].

However, it makes sense to use the focus measure operator as a rank for an image fusion rule. We will make use of the first-order Gaussian derivative to measure the focus score. The σ of the Gauss filter determines the scale of prominent features. The focus function becomes:

$$F(\sigma) = \frac{1}{n}\sum_{xyz}[I(x,y,z)G_x(x,y,z,\sigma)]^2 + [I(x,y,z)G_y(x,y,z,\sigma)]^2 + [I(x,y,z)G_z(x,y,z,\sigma)]^2$$

Where $I(x,y,z)$ is the image intensity values for the respective (x,y,z) pixel , and $G_x(x,y,\sigma), G_y(x,y,\sigma), G_z(x,y,\sigma)$ *are the first-order Gaussian derivatives in the* x , y and z direction at scale σ, *n again denotes* the total amount of pixels in the image.

3.2 TRANSFORM BASED FUSION

In this type of fusion schemes we apply a certain discrete transformation on the images then implement a fusion in the transform domain and obtain the final fused image through an inverse transformation to the image domain[21, 22].

3.2.1 Abstract Idea in transform based fusion

If a signal is in the area spanned by certain basis functions we can decompose the signal into a linear combination of said basis.

$$f(x) = \sum_k a_k \varphi_k(x), \quad f(x) \in L^2(R) \quad (4)$$

$L^2(R)$ is the space of all functions f(x) with a well-defined integral of the square of the modulus of the function.

k: index of the finite or infinite sum
a_k: Expansion coefficients
$\varphi_0(x)$: Expansion functions or basis
If we choose the basis appropriately, there exists another set of basis $\{\tilde{\varphi}_k(x)\}$ such that $\{\varphi_k(x)\}$ & $\{\tilde{\varphi}_k(x)\}$ are orthonormal.
Meaning that their inner product is:

$$\langle \varphi_i(x), \tilde{\varphi}_j(x) \rangle = \int \varphi_i(x), \tilde{\varphi}_j(x)\, dx = \delta_{ij} \quad (5)$$

And $\{\tilde{\varphi}_k(x)\}$ is called the dual of $\{\varphi_k(x)\}$.

Since they are orthonormal, we can find the coefficients by:

$$\langle f(x), \tilde{\varphi}_j(x) \rangle = \int f(x)\, \widetilde{\varphi^*}_k(x)\, dx = \int \left(\sum_{k'} a_{k'}\, \varphi_{k'}(x) \right) \widetilde{\varphi^*}_k(x)\, dx =$$

$$\sum_{k'} a_{k'} \left(\int \varphi_{k'}(x)\, \widetilde{\varphi^*}_k(x)\, dx \right) = \sum_k a_k\, \delta_{kk'} = a_k$$

So, we can calculate the expansion coefficients as:

$$a_k = \langle f(x), \tilde{\varphi}_j(x) \rangle = \int f(x)\, \widetilde{\varphi^*}_k(x)\, dx \quad (6)$$

Note that for the well-known Fourier expansion we have $\varphi_k(x) = \exp\left(j2\pi kt/T\right)$

and $\tilde{\varphi}_k(x) = (1/T) exp\left(j2\pi kt/T\right)$

3.2.2 Wavelet Transform Based Fusion

Introduction to the Wavelet transform in image processing

The most common tool for a transform based signal analysis is the Fourier transform. However, the Fourier transform has the drawback of dealing with just the frequency components of a signal without providing any temporal details. So, in image fusion we mostly use a method based on the Wavelet Transform. Wavelets are functions that are concentrated in time as well as frequency around a certain point. The basic idea is to represent any signal in a series, by using a wavelet function or "Mother Wavelet" and a scaling function or "Father Wavelet" for a basis to expand our signal. The basis functions vary both in the frequency range and in the spatial range, making them ideal for analyzing images. The wavelet transform is designed in such a way that we get good frequency resolution for the low frequency components, which in our case is the average intensity values of an image, and high temporal resolution for high frequency components which are in our case are the edges of an image [41]. The Wavelet transform decomposes the image into four parts. One is the approximation part which shows the general trends for the pixel values of the image and the other three are the horizontal, vertical and diagonal details of the image (Figure1).

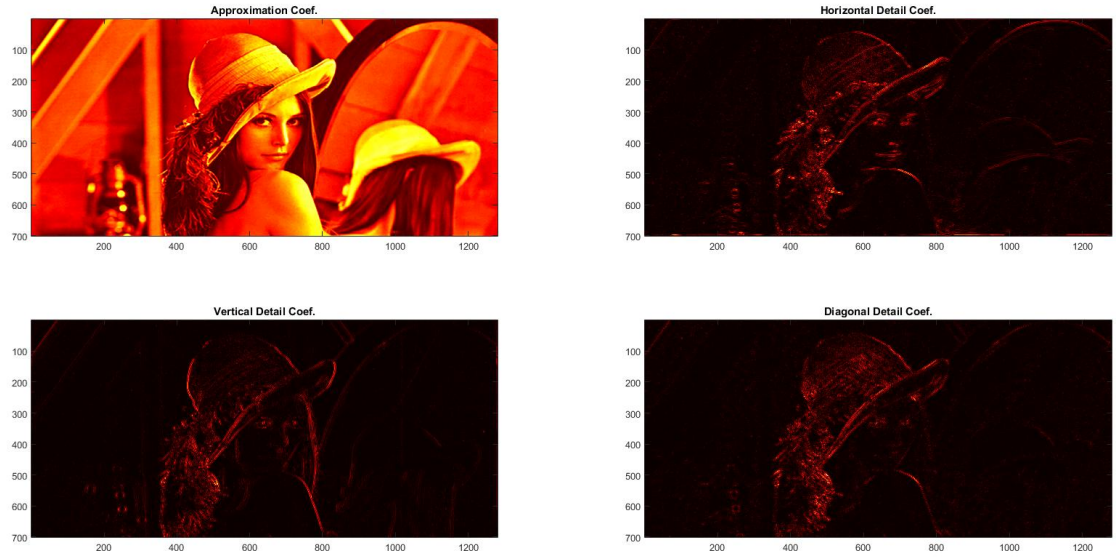


Figure 1: Illustration of the approximation and detail images

The images that go through the wavelet transformation can be fused with a number of decomposition levels and fusion rules. For the wavelet transformation of images, many readily available wavelet processing libraries can be taken (e.g. MATLAB©).

In each sub-band, individual pixels of the two images are compared, based on a fusion rule, which serves as a measure of quality at that particular scale and space. A fused Wavelet transformed image is created by taking pixels from that wavelet transformed image that shows greater quality at that region level. The inverse wavelet transform gives us a new fused image with the highest possible resolution everywhere. For example, since wavelet coefficients that have large values contain the information about the salient features of the images, such as edges and lines, a good fusion rule is to take the *maximum intensity* of the corresponding detail coefficients. In the next few sections we will follow with a mathematical explanation of the Wavelet Transform.

Introduction to the scaling and wavelet function with an example [Haar]: The scaling and the wavelet function

Consider the unit-height, unit-width scaling function $\varphi(x)$ such as [22]:

$$\varphi(x) = \begin{cases} 1 & 0 \leq x < 1 \\ 0 & otherwise \end{cases}$$

It is a rectangular function centered at x = 1/2 with width 1. The scaled version is defined as $\varphi_n(x) = \varphi(2^n x)$. We can see, in the figure 2, that as m=2^n gets higher, the support of this scaled basis will become narrower. As the support of the function becomes narrower, the higher the frequency it represents becomes.

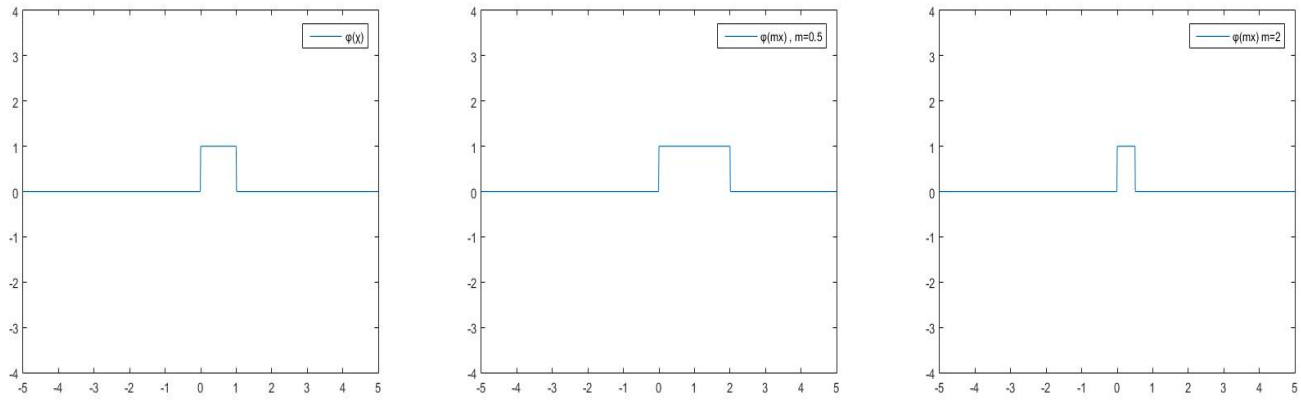


Figure 2a: Haar scaling function

However, the $\{\varphi_n(x)\}$ set of functions are not orthogonal and therefore can not be used for basis construction. Nevertheless, we can use the Gram-Schmidt process to obtain a set of orthonormal basis from existing $\{\varphi_n(x)\}$ like so.

$$\varphi_0'(x) = \varphi_0(x) = \varphi(\chi)$$

$$\varphi_1'(x) = \varphi_1(x) - \frac{\langle \varphi_1(x), \varphi_0(x) \rangle}{\langle \varphi_0(x), \varphi_0(x) \rangle} \varphi_0(x) = \begin{cases} 1/2\ , & 0 \le x < 1/2 \\ -1/2\ , & 1/2 \le x < 1 \\ 0\ , & otherwise \end{cases} = \psi(\chi)/2$$

We can continuously apply this process to extend the basis.

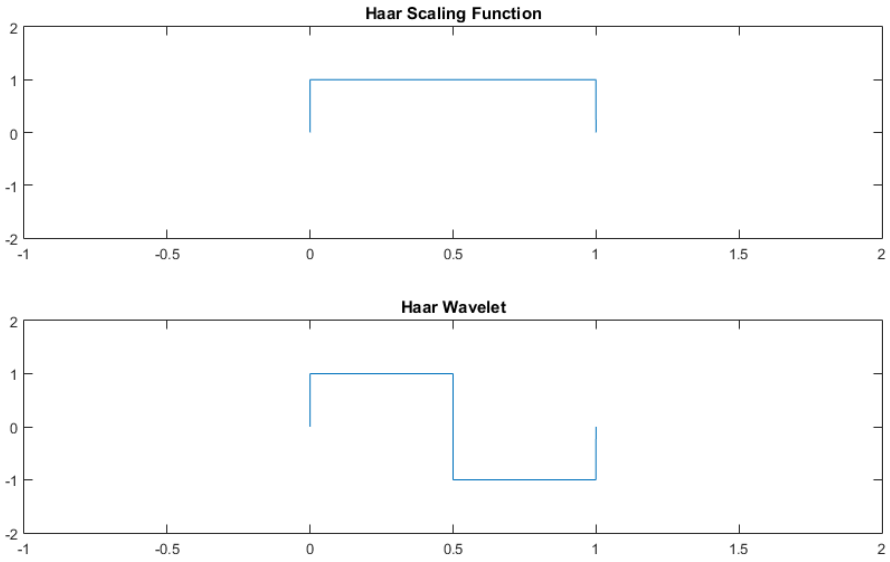


Figure 2b: Haar scaling and wavelet functions

Now, we take a look at the first two basis functions and their Fourier transform. By inspection, $\varphi(\chi)$ has a nonzero mean, while $\psi(\chi)$ is zero in average. Now $\varphi(\chi)$ has two jumps at x = 0, 1 while $\psi(\chi)$ jumps at x = 0, 1/2, 1. Therefore, $\varphi(\chi)$ is more compact at low frequencies while $\psi(\chi)$ concentrates at relatively high frequencies. We call $\varphi(\chi)$ a scaling function for approximations and $\psi(\chi)$ a wavelet function to find the details.

The Scaling Function

Now we can more generally define the scaling function in order to use the idea of multi-resolution, then we can define the wavelet function in terms of it. The scaling function is formally defined as:

$$\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k) \quad (7)$$

Where j is the parameter about the visibility in frequency and k a translation factor for the position. In practice, we may want to see the whole data with "desired" resolution, i.e. for some resolution j. The scaling function is orthogonal to its integer translates. The integer translates of the scaling function do not overlap with the original scaling function. We define the subspace spanned by these functions as:

$$V_j = Span\{\varphi_{j,k}(x)\} \quad (8)$$

This means that for any j, k $\varphi_{j,k}(x) \in V_j$. The subspaces spanned by the scaling function at lower scales are encompassed within those spanned at higher scales: From figure 2, it is clear to show that $\varphi_{-1}(\chi) = \varphi_0(\chi) + \varphi_0(\chi - 1)$.

In terms of subspaces as defined above its: $V_j \in V_{j+1}$ for all $j \in \boldsymbol{Z}$

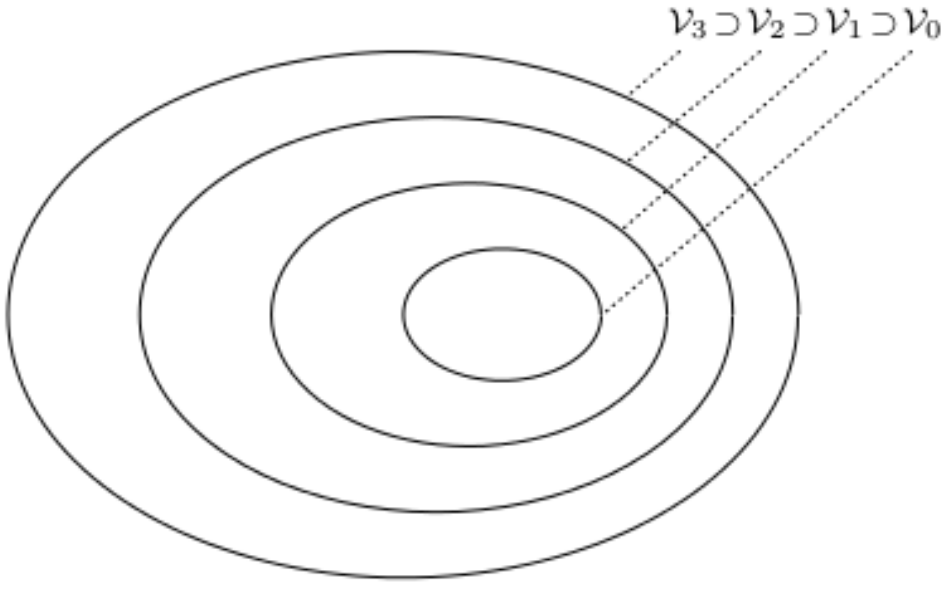


Figure 3 : Subspaces spanned by the scaling functions

Like in the previous example we will use the Haar scaling function, and using the above definition we will get :

$$\varphi(\chi) = \varphi_{0,0}(\chi) = \frac{1}{\sqrt{2}}\varphi_{1,0}(\chi) + \frac{1}{\sqrt{2}}\varphi_{1,1}(\chi)$$

Now, we can expand for $\varphi_{j,k}(x)$ and we have:

$$\varphi(\chi) = \frac{1}{\sqrt{2}}(\sqrt{2}\varphi(2\chi)) + \frac{1}{\sqrt{2}}(\sqrt{2}\varphi(2\chi - 1))$$

In general this is called a refinement equation and it is:

$$\varphi(\chi) = \sum_n h_\varphi[n]\sqrt{2}\varphi(2\chi - n) \qquad (9)$$

$h_\varphi[n]$ are a sequence of real numbers called the scaling function coefficients.

For the Haar scaling functions we have $h_\varphi[n] = \{\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\}$.

The wavelet function

With the scaling function we already have a low resolution approximation of a signal or image. However, instead of using $\varphi_{j,k}(x)$ and increasing j to increase the size of the subspace spanned by the scaling functions to describe the details of an image instead we define a slightly different set of functions that span the *differences* between the spaces spanned by the various scales of the scaling function. These functions are the wavelets. The wavelet function is formally defined as:

$$\psi_{j,k}(x) = 2^{j/2}\psi\left(2^j x - k\right) \quad (10)$$

It is required that scaling functions and wavelets are orthogonal. Also, since the subspaces spanned by the wavelets are the *differences* between the spaces spanned by the various scales of the scaling function and they must also be orthogonal, we have:

$$W_j = Span\{\psi_{j,k}(x)\} \quad (11)$$

So the wavelet spanned subspace W_0will be such that $V_1 = V_0 \oplus W_0$ which can be extended to $V_2 = V_0 \oplus W_0 \oplus W_1$etc.

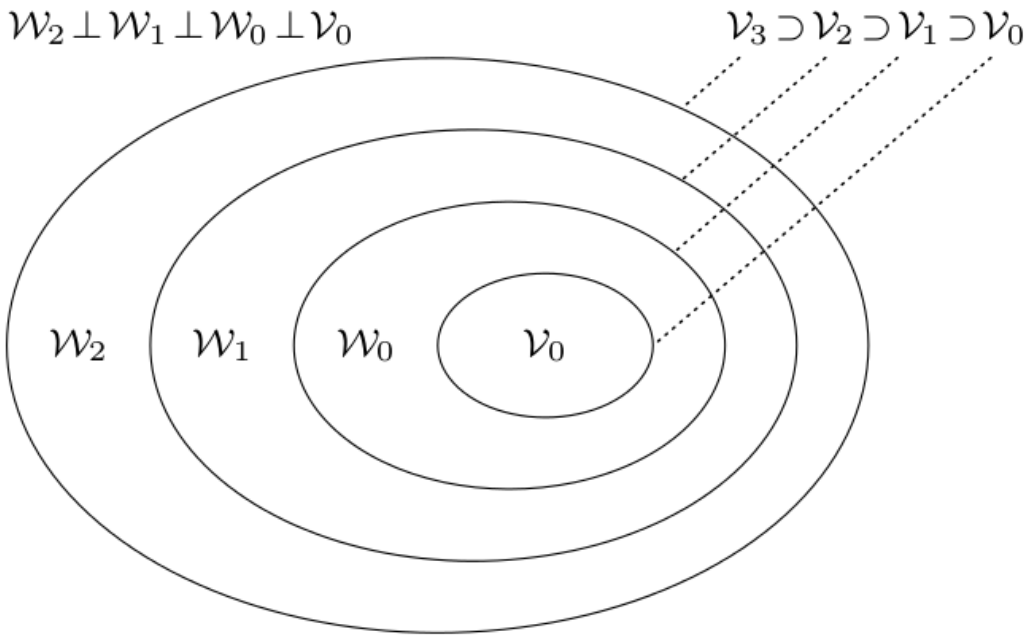


Figure 4: Scaling and Wavelet function subspaces

Since these wavelets reside in the space spanned by the next narrower scaling function ($W_0 \in V_1$), they can be represented by a weighted sum of shifted scaling functions $\varphi_{j,k}(x)$ defined as:

$$\psi(\chi) = \sum_n h_\psi[n]\sqrt{2}\varphi(2\chi - n) \quad (12)$$

From the requirement that the wavelets span the “difference" or orthogonaly complement spaces, we get :

$$h_\psi[n] = -1^n h_\varphi[1-n] \quad (13)$$

Now that we have constructed the $\psi_{j,k}(x)$ and $\varphi_{j,k}(x)$ for our basis, they could span all of $L^2(R)$, so according to (4) any function $f(x) \in L^2(R)$ can be expanded as:

$$f(x) = \sum_{k=-\infty}^{\infty} c(k)\,\varphi_k(x) + \sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty} d(j,k)\,\psi_{j,k}(x) \quad (14)$$

Which is a series expansion in terms of scaling functions and wavelets.

In this expansion, the first summation in eq. (14) gives a function that is a low resolution or coarse approximation of f(x). For each increasing index j in the second summation, a higher resolution function is added, which in turn increases details. Now since we have an orthonormal basis we can calculate the coefficients of (14) like so:

$$c_k = \langle f(x), \varphi_k(x) \rangle = \int f(x)\, \varphi_k(x)\, dx \qquad (14.a)$$

And

$$d(j,k) = \langle f(x), \psi_{j,k}(x) \rangle = \int f(x)\, \psi_{j,k}(x)\, dx \quad (14.b)$$

However, so far our variable (x) was continuous. If we have a discrete version of the variable (that is x = 0, 1,…, M-1) a discretized version of eq. (14) is needed. And since images are discrete signals this version is the one we will be using.
It is:

$$f(x) = \frac{1}{\sqrt{M}} \sum_k W_\varphi(j_0,k)\, \varphi_{j_0 k}(x) + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_k W_\psi(j,k)\, \psi_{j,k}(x) \qquad (14.c)$$

Where $f(x), \varphi_{j_0 k}(x), \psi_{j,k}(x)$ are discrete functions defined in [0, M-1].
Since the subspaces $\{\varphi_{j_0 k}(x)\}\, k \in \boldsymbol{Z}, \{\psi_{j,k}(x)\}\, k, j \in, \boldsymbol{Z}^2, j \geq j_0$ are orthogonal with each other we have:

$$W_\varphi(j_0,k) = \frac{1}{\sqrt{M}} \sum_k^{M-1} f(x) \varphi_{j_0,k}(x) \qquad (14.d)$$

$$W_\psi(j,k) = \frac{1}{\sqrt{M}} \sum_k^{M-1} f(x) \psi_{j,k}(x) \qquad (14.e)$$

We call (14.d) the approximation coefficients and (14.e) the detailed coefficients.
This is the Discrete Wavelet Transform or DWT.

Fast Wavelet Transform

To reduce the computation time of the DWT another way was developed to produce the approximation and detailed coefficients. From eq. (10) and eq. (12) we have:

$$\psi_{j,k}(x) = 2^{j/2} \psi\left(2^j x - k\right)$$

$$\psi(\chi) = \sum_n h_\psi[n]\sqrt{2}\varphi(2\chi - n)$$

Which can be rewritten as:

$$\psi(2^j x - k) = \sum_n h_\psi[n]\sqrt{2}\varphi(2(2^j x - k) - n)$$

$$= \sum_m h_\psi[m - 2k]\sqrt{2}\varphi((2^{j+1}x - m), n = m - 2k$$

Now from $eq.(14.e)$ we have:

$$W_\psi(j,k) = \frac{1}{\sqrt{M}}\sum_k^{M-1} f(x)\psi_{j,k}(x)$$

$$= \frac{1}{\sqrt{M}}\sum_k^{M-1} f(x)2^{j/2}\psi(2^j x - k)$$

$$= \frac{1}{\sqrt{M}}\sum_k^{M-1} f(x)2^{j/2}\sum_m h_\psi[m - 2k]\sqrt{2}\varphi((2^{j+1}x - m)$$

$$= \sum_m h_\psi[m - 2k]\,(\frac{1}{\sqrt{M}}\sum_k^{M-1} f(x)2^{j+1/2}\varphi((2^{j+1}x - m))$$

$$= \sum_m h_\psi[m - 2k]W_\varphi(j + 1, m)$$

So it is:

$$W_\psi(j,k) = \sum_m h_\psi[m - 2k]W_\varphi(j + 1, m)$$

Similarly for the approximation coefficients it is:

$$W_\varphi(j,k) = \sum_m h_\varphi[m - 2k]W_\varphi(j + 1, m)$$

Discrete Wavelet Transform based fusion

The most common form of a transform based fusion is the Discrete Wavelet Transform or DWT fusion [16, 18, 19, 20]. The basic idea is to represent any signal in a series, by using a scaling function and a wavelet function for a basis. These functions can take many forms like Haar, Daubechies, Symlets, etc. In our previous

example we used Haar. Given a scaling and a wavelet function, [42], the scaled and translated versions can be defined as.

$$\varphi_{j,m,n}(x,y) = 2^j \varphi\left(2^j x - m\,, 2^j y - n\right)$$

(15)

$$\psi^i{}_{j,m,n}(x,y) = 2^j \psi\left(2^j x - m\,, 2^j y - n\right)\; i = \{H, V, D\}$$

$$\psi^H(x,y) = \psi(x)\varphi(y)$$

$$\psi^V(x,y) = \varphi(x)\psi(y)$$

$$\psi^D(x,y) = \psi(x)\psi(y)$$

$\varphi_{j,m,n}(x,y)$: *Is the scaling fuction for low frequency approximations*

$\psi^i{}_{j,m,n}(x,y)$: *Is the wavelet function to find the details along the Horizontal, Vertical and Diagonal plane.*

m: determines the position of (φ,ψ) along the axis x,y

j:determines the width of (φ,ψ) along the axis x,y

Given the basis, the Discrete Wavelet Transform of an image I(x,y) of size M x N is:

$$W_\varphi(j_0,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x,y)\varphi_{j_0,m,n}(x,y) \qquad (16)$$

$$W_\psi^i(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x,y)\psi^i{}_{j_0,m,n}(x,y) \qquad (17)$$

j_0: is an arbitrary starting scale

$W_\varphi(j_0,m,n)$: Are coefficients that define an approximation of I(x,y) at scale j_0

$W_\psi^i(j_0,m,n)$: Are coefficients that add horizontal, vertical and diagonal details for scales j $\geq j_0$.

Generally we set $j_0 = 0$ and select N=M=2^J so that j=0,1,2,3,….,J-1 and m=n=1,2,3,…., 2^j-1.

Given W_φ and W_ψ^i then the image I(x,y) can be obtained by the inverse Discrete Wavelet Transform:

$$I(x,y) = \frac{1}{\sqrt{MN}} \sum_{m} \sum_{n} W_{\varphi}(j_0, m, n)\, \varphi_{j_0,m,n}(x,y) + \frac{1}{\sqrt{MN}} \sum_{i=H,V,D} \sum_{j=j_0}^{J-1} \sum_{m} \sum_{n} W_{\psi}^{i}(j,m,n)\, \psi^{i}{}_{j_0,m,n}(x,y) \quad (18)$$

For the fusion process initially we apply the DWT on the original input images. Following, all respective wavelet coefficients from the input images are combined using a chosen fusion rule. Since wavelet coefficients with large absolute values contain the information about the salient features of the images such as edges and lines, a good fusion rule is to take the *maximum* of the absolute values of the corresponding wavelet coefficients. Then we obtain the final fused image through the Inverse Wavelet Transform (IWT). We will follow with an illustration of the decomposition on a test image.

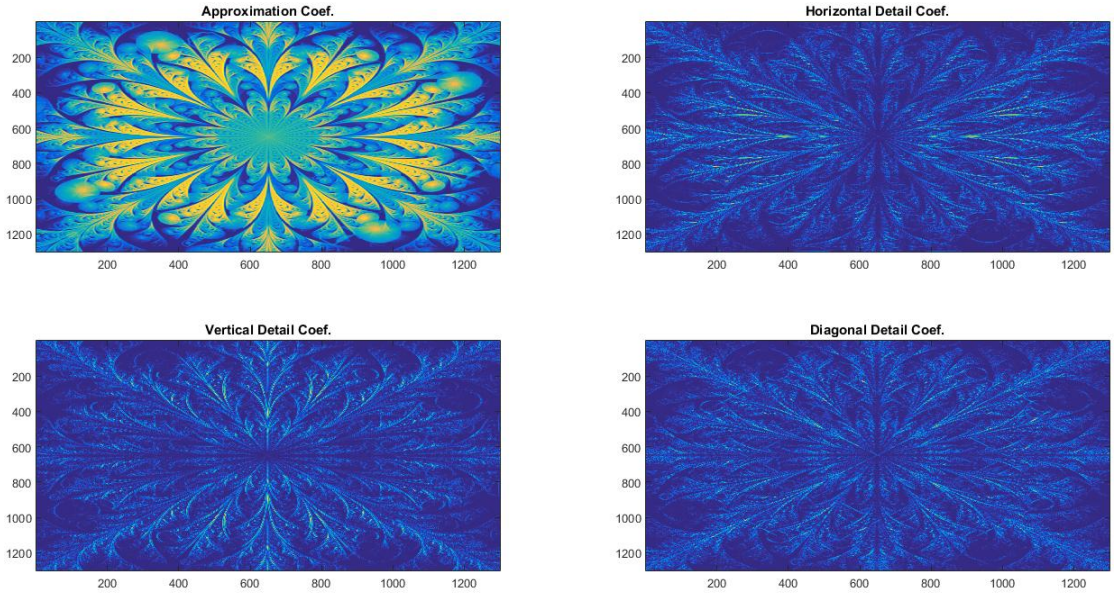


Figure 5: On the left we have the Original image which gets decomposed into Wφ (2nd image upper left side), Wψ: Horizontal (2nd image upper right side), Wψ: Vertical (2nd image lower left side) and Wψ: Diagonal (2nd image lower right side)

The principle of image fusion using wavelets is to merge the wavelet decompositions of the two original images using fusion methods applied to approximations coefficients and details coefficients [33]. After we get the fused decomposition map we use the inverse transform to get the final fused image. The steps of the fusion process are depicted in figure 6.

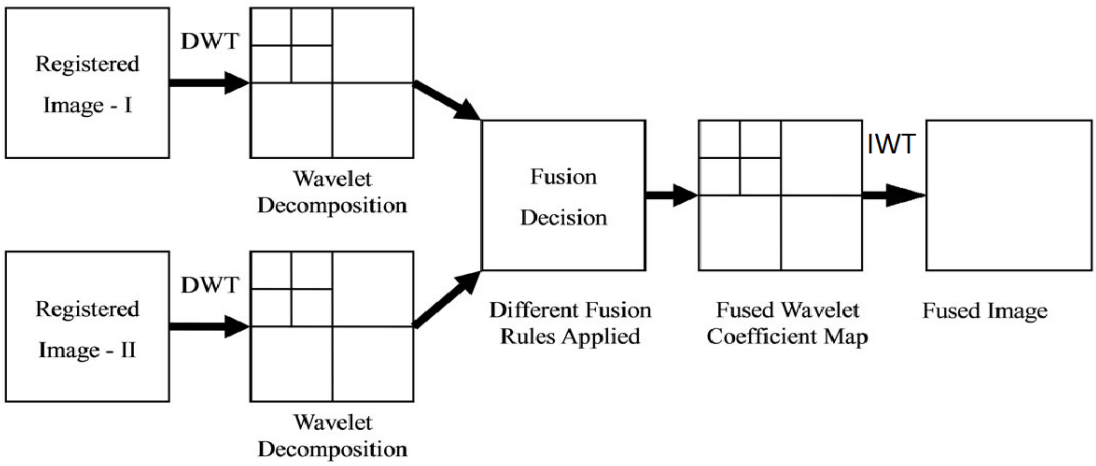


Figure 6: Wavelet Transform based image fusion

4 COMPARISON OR IMAGE QUALITY METRICS

When trying to ascertain which image fusion algorithms are better suited to our imaging system, we need ways to measure the effectiveness of those algorithms. This is where image quality metrics come into play. Many of these metrics are measuring the quality of information in an image. They offer a quantitative way for us to be able to choose between those algorithms.

It is important to note that image quality metrics are not the end all be all for the final assessment. Sometimes other factors, such as computation time of the algorithms for example, can have an affect on the decision. Also a visual inspection from an expert in image analysis or a biologist, on the data can contradict some of the results of these metrics and supersede them (since it's them who will eventually use them). That said, these methods provide a helpful tool in the absence of clear visual qualitative differences between the fused images that the algorithms produce.

Now, image quality metrics can be split into two categories, the reference metrics and the non-reference metrics. As the name suggests reference metrics require a "perfect" ground truth image to use as a reference for comparison with the fused image. On the other hand the non-reference metrics offer no such restrictions as they assess the quality of the fused image alone. The non-reference metrics can be used both with and without a reference image.

4.1 HISTOGRAM, JOINT HISTOGRAM

Histogram

Before we move on to the description of the metrics, we need to know a few things about the image histogram and joint histogram and how we define information in an image.

An image histogram is a graphical representation of the intensity distributions in an image. It plots the number of pixels for each intensity value. The horizontal axis of the graph represents the intensities, while the vertical axis represents the number of

pixels in a particular intensity. Below we can see an image of a C-elegans specimen with its corresponding histogram.

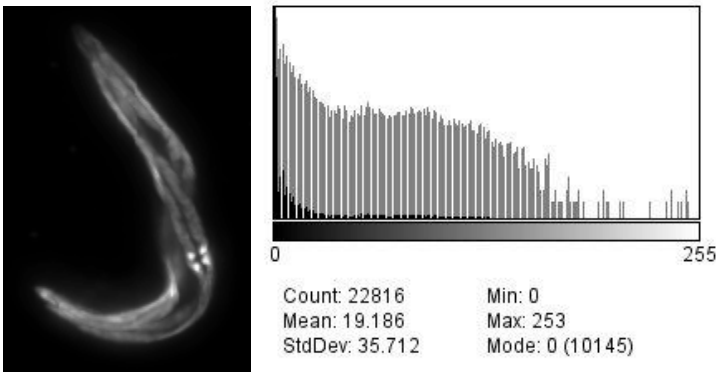


Figure 6 Image histogram of a c-elegans sample image

The probability that a pixel will have a certain value (***a***) can be calculated from the histogram. If N_a is the number of pixels that have the value (***a***) then that probability will be:

$$P(a) = \frac{N_a}{M} \qquad (19)$$

Where M= [(image height) x (image width)] if we have a 2D image or
M= [(image height) x (image width) x (image depth)] for a 3D image.

Joint Histogram

The joint histogram is a concept related to the individual image histograms, where the intensity value of a pixel in one image is plotted against the intensity value for the same pixel location in the second image. A good way to understand how a joint histogram is used is by comparing an image with a different color duplicate of itself and then with transformed versions of the duplicate, as illustrated in Figure 3. To assist in our visual comparisons, we assign different colors (red and green) to the original and duplicate images.

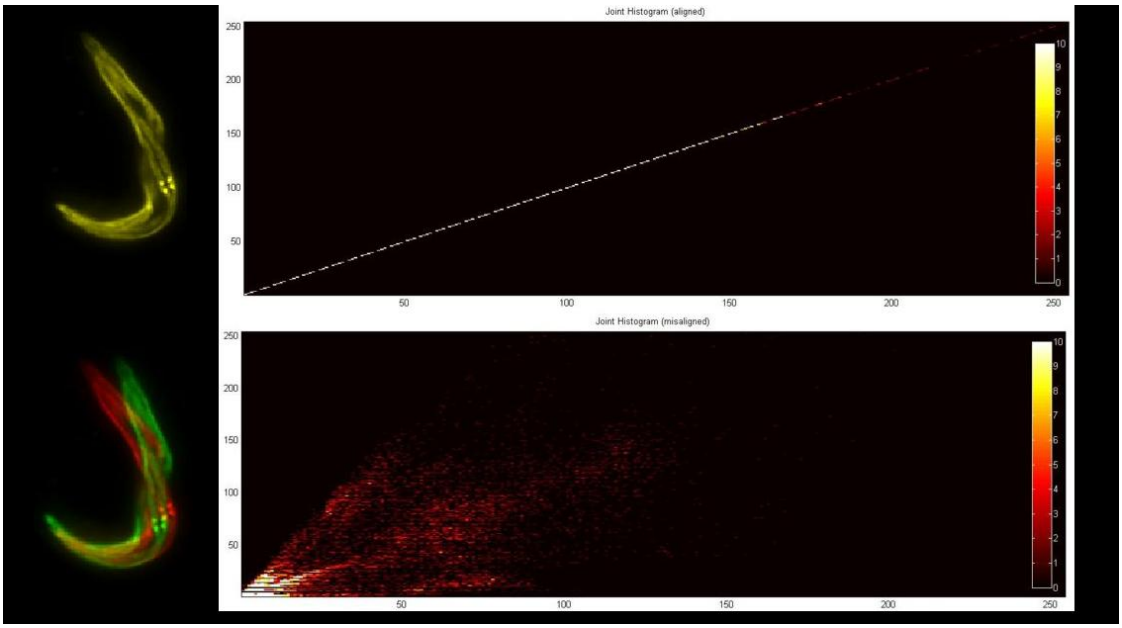


Figure 7 : An example for the use of a joint histogram, The left panel shows correctly aligned (top left) or misaligned (bottom left) green and red images of the worm, while on the right panel their corresponding joint histograms are presented. ***a.)*** *The top-left image represents the case where the two images are correctly aligned and their joint histogram is shown on top-right.* ***b.)****The bottom left picture shows the two images when misaligned and their corresponding joint histogram (bottom-right).*

In figure (7.***a***.) the situation where the two images are aligned is presented. Note the resultant yellow-only color scale. The joint histogram for this situation is a diagonal straight line since the values of all pixel pairs in the two images are identical. Figure (7. ***b***.) shows the effect of a 5 pixels horizontal translation with a 15 degree rotation of the green duplicate image to the right. Notice that regions of mismatch are seen in shades of red for the original image and shades of green for its (transformed now) duplicate, while the overlapping regions are in shades of yellow. The joint histogram is now that of a **scatter plot** because the pixel values in the two images no longer correspond to each other, meaning that a bright pixel in one image might now correspond to a dark one in the other.

It follows that, since we can calculate the probability of occurrence of an intensity value from the marginal histograms we can also calculate the joint probability of an intensity value in a reference image (***a***) and the intensity value (***b***) of an image we want to register. If the joint histogram is normalized, it becomes an estimate of the joint probability distribution function of the intensities in the images. This would be:

$$P(a,b) = \frac{Hist(a,b)}{\sum_{a,b} Hist(a,b)} \qquad (20)$$

4.2 Reference Metrics

As we already stated reference metrics need a ground truth image to work [29, 30, 31, 35, 36]. In most cases were image fusion is used the ground truth image is obtained by manually cutting and pasting the "good" parts of the images we want to fuse. For example, in multi-focus image fusion, were we have two images of the same scenery with the focused part of each of the images being in different areas, we can cut and paste the "in focus" parts of the images in another ground truth image and use that image as a reference. In our case though, where we have three dimensional images acquired from different views the construction of such a reference image is difficult. However, if we can take a two dimensional slice of a certain view (for example the zero degree view) at a depth before any blurring, due to light diffusion, takes place we can use that as a ground truth for that particular section of the image.

Following, we will present some of the most common reference metrics. For the following, let R be the reference (ground truth) image, F be the fused image to be assessed and M x N the size of the images.

Mean Squared Error (MSE) and Root Mean Squared Error (RMSE)

The mean squared error is defined as:

$$MSE = \frac{1}{M}\frac{1}{N}\sum_{j}^{M}\sum_{i}^{N}|R(i,j) - F(i,j)|^2 \quad (26)$$

And the root mean squared error:

$$RMSE = \sqrt{\frac{1}{M}\frac{1}{N}\sum_{j}^{M}\sum_{i}^{N}|R(i,j) - F(i,j)|^2} \quad (27)$$

Peak Signal to Noise Ratio (PNSR)

The peak signal to noise ratio is given by:

$$PSNR = 10\log_{10}\left(\frac{Peak^2}{MSE}\right) \quad (28)$$

Peak is the maximum Possible value NOT the maximum value of the image, it depends on the encoding of the image for example for an 8-bit grayscale image Peak=255.

MSE is the mean squared error defined above.

Correlation coefficient

CC measures the correlation between the ground truth and the fused images. The higher this correlation is, the better the quality of the fused image. The ideal value of correlation coefficient is 1. The formula for calculating the Correlation coefficient is.

$$CC = \frac{\sum_i (\mathrm{R}(i) - \bar{\mathrm{R}})\,(F(i) - \bar{F})}{\sqrt{\sum_i (\mathrm{R}(i) - \bar{\mathrm{R}})^2\,(F(i) - \bar{F})^2}} \quad (29)$$

Structural Similarity Index (SSIM)

The Structural Similarity (SSIM) [35] quality assessment index is based on the computation of three terms, namely the luminance term, the contrast term and the structural term. The overall index is a multiplicative combination of the three terms. It assesses the visual impact of those three characteristics on an image. The measure between two images x and y of same size $N \times N$ is:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{({\mu_x}^2 + {\mu_y}^2 + C_1)({\sigma_x}^2 + {\sigma_y}^2 + C_2)}$$

- μ_x the average of x;
- μ_y the average of y;
- ${\sigma_x}^2$ the variance of x;
- ${\sigma_y}^2$ the variance of y;
- σ_{xy} the covariance of x and y;

- $C_1 = (k_1 L)^2, C_2 = (k_2 L)^2$two variables to stabilize the division with weak denominator;
- L the dynamic range of the pixel-values (typically this is $2^{\#bits\ per\ pixel} - 1$);
- $k_1 = 0.01$ and $k_2 = 0.03$ by default.

4.3 NON- REFERENCE METRICS

When a reference image is unavailable we can use non-reference metrics to measure the quality of the fused image [29, 30, 31, 35, 36].

Information

The definition of information in the context of information theory is based on probabilities. Let us assume an event *a* and the probability of this event occurring *p(a),* then the information contained in that event is defined as [23, 24, 25, 26]:

$$I(a) = \log_b \frac{1}{p(a)}$$

Where b is 2 if we measure information in bits.
Therefore, information is inversely proportional to probability. Events with the least probability of occurring will provide the most information.

Entropy

Now to move on with the actual metrics. Entropy in information theory measures the amount of uncertainty or complexity in the outcome of a situation. For example, the outcome of an experiment. In our case it can be used as a measure for the quality of information in an image. Entropy can be used both as reference and as a non-

reference metric. The most commonly used measure of entropy in image processing is the Shannon-Wiener entropy measure ***H*** which is defined as [23 24 25 26]:

$$H(a_i) = \sum_{1}^{n} p(a_i) \log_2 \frac{1}{p(a_i)} \quad (21)$$

Where n is the number of events.

We can also calculate the joint entropy from the images' joint histogram:

$$H(a_i, b_j) = \sum_{1}^{n} p(a_i, b_j) \log_2 \frac{1}{p(a_i b_j)} \quad (22)$$

Higher values indicate a more complex image and hence a probably better fusion result.

Mutual Information

Mutual information is based on information theory and it expresses the amount of information that an event, or in our case an image, contains about another event or image. Hence, for two images A and B, the mutual information of A and B or the amount of information than A contains about B will be [27 28]:

$$MI(A, B) = H(A) + H(B) - H(A, B) \quad (23)$$

Or, in the normalized form which we will be using:

$$NMI(A, B) = \frac{H(A) + H(B)}{H(A, B)} \quad (24)$$

Where H(A), H(B) are the marginal entropies of A, B and H(A, B) their joint entropy, which can be computed by the marginal and joint histograms of the images. Mutual information can also be used as a non-reference metric with a slight adjustment. Let A and B be the source images and let F be the fused image. When no reference images are available, fusion assessment is performed as follows. The MI-based non-reference measure is defined as [37, 38]:

$$MF(A, B, F) = MI(F, A) + MI(F, B) \quad (25)$$

MF(A,B,F) represents total amount of similarity between fused image F vs. source images A and B. Again, higher values indicate a better fusion algorithm.

Standard Deviation (SD)

This metric is more efficient in the absence of noise. It measures the contrast in the fused image. An image with high contrast would have a high standard deviation. The standard deviation (SD), which is the square root of variance, reflects the spread in the data. Thus, an image with high contrast would have a high standard deviation, and a low contrast image would have a low standard deviation.

$$SD = \sqrt{\frac{1}{M}\frac{1}{N}\sum_{j}^{M}\sum_{i}^{N}|F(i,j) - \bar{F}|^2} \qquad (30)$$

Spatial Frequency (SF)

Spatial Frequency (SF) indicates the overall activity level in the fused image.
Spatial frequency can be calculated by the images row frequency (RF) and column frequency (CF):

$$RF = \sqrt{\frac{1}{M}\frac{1}{N}\sum_{j}^{M}\sum_{i}^{N}|I(i,j) - I(i,j-1)|^2} \qquad (31)$$

$$CF = \sqrt{\frac{1}{M}\frac{1}{N}\sum_{j}^{M}\sum_{i}^{N}|I(i,j) - I(i-1,j)|^2} \qquad (32)$$

And finally the Spatial Frequency:

$$SF = \sqrt{RF^2 + SF^2} \quad (33)$$

5 EXPERIMENTAL WORK

As we have already stated, any optical imaging system would suffer from light scattering and absorption by the tissue. As a result, image quality decreases as the optical path in the sample increases. Tomographic techniques for optical microscopy such as SPIM and OPT device multi-angle acquisition as a countermeasure to mitigate this problem. Multiple 3D data sets of the same object are collected from different directions and combined in a single high visibility data set. The information from each data set can then be fused into a single, superior 3D image. Furthermore, to extract additional information multi-wavelength experiments are possible, where different biomarkers are used to focus and mark different areas of the sample we are studying. In this case two or more scans have to be combined, in each of which, different areas of the sample are illuminated. This means there is a need for a method for the combination of said images in one data set to achieve a more complete view of the object we are studying.

Several, datasets, covering the usual SPIM image requirements have been created and will be used for the purpose of this thesis. The implementation of the image fusion processes will take the following form in general. The source images to be fused will be initially registered to avoid errors due to misalignment. Following, fusion of the original source images will be performed using the methods described above. After the fusion takes place, a comparison of the resulting images from all the fusion methods and in the case of the wavelet transform all of the fusion rules and up to 5th decomposition level will be performed. Since we lack a perfect image due to the nature of the imaging apparatus we will mostly use non-reference comparison metrics. As we stated before, for the reference metrics we will use a “good” 2D slice from one of the source images meaning an image at a depth where we believe that we have minimal diffusion and can be used as reference with the corresponding slice of the fused images.

6 RESULTS

6.1 MULTI-VIEW FUSION

As stated before in multiview SPIM, we acquire several stacks of images-sections of the same sample from different views by rotating the sample. One of the main reasons for this, is to compensate for the massive difference in resolution between the x-y and z dimensions of the volumetric images from a single projection that we have already mentioned. To return the volumetric image to the actual size we apply a scale in the z-axis which results in the deformation of the image.

Experiment 1:

For example, when we scan an approximately spherical object from a single view, and scale it to actual size, we get an object that now will be an ellipsoid. In the following experiment, we scanned an approximately spherical object, a cell from a tumor, from eight different directions, as seen in figure 8. We can see at the individual scans that the object, which should be approximately spherical, show up as an ellipsoid when viewed from the top. Then, if we combine the individual scans by superimposing them on top of one another, we get a star shaped object which is far cry from the afore mentioned sphere we intended. To correct for that, we fuse the images using the wavelet transform. As we already discussed in the wavelet transform we can differentiate between horizontal, vertical and diagonal details between the original images and create fusion rules accordingly to produce the resulting fused image.

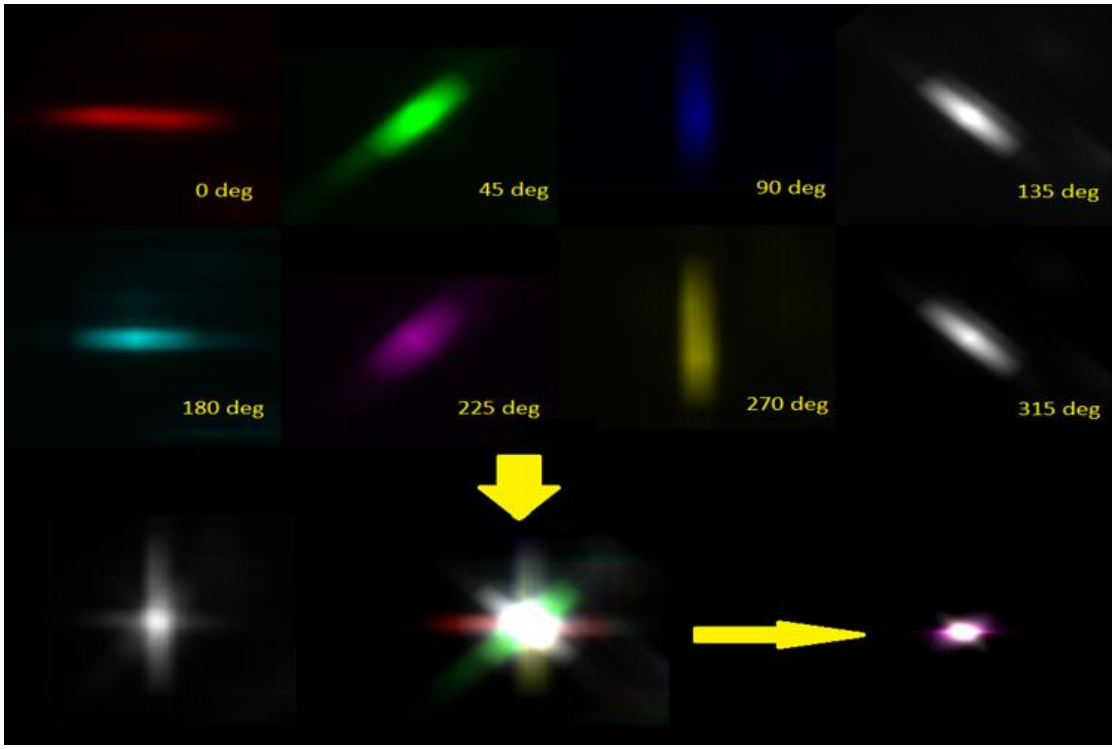


Figure 8: Here we see a single cell scanned from different views and then combined into a single image by PCA fusion (on the bottom left), simple addition (bottom center) and the by wavelet transform based fusion (on the bottom right corner).

Experiment 2

Another problem that is common in SPIM tomographic imaging is the afore mentioned effect of absorption and scattering of light by the samples. Depending on the opacity of the sample, a scan from a single view can result in lost information the deeper we go into the sample. To counteract that a solution of multi-angle scans and eventually fusion, is proposed. To demonstrate this effect we will use two scans of a tumor spheroid, acquired from opposite sides (0^0 and 180^0 angles) that has been engulfed in a transparent medium with scattered fluorescent beads as seen in figure (9.a)

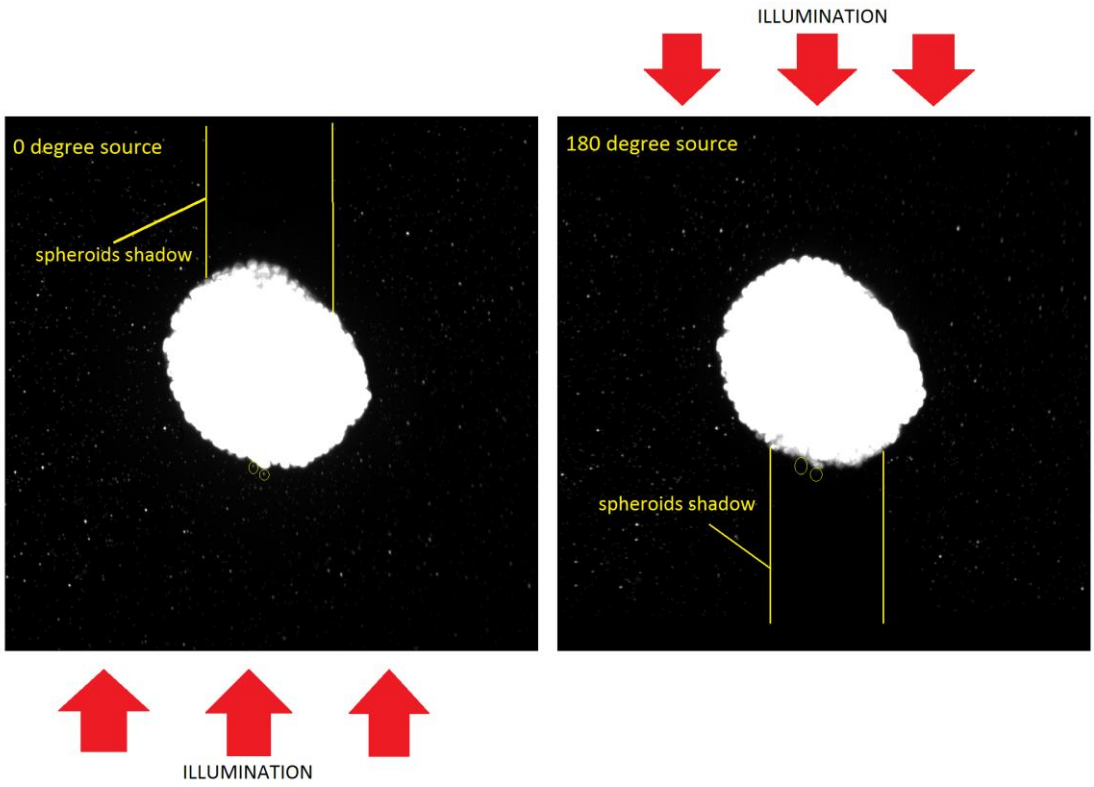


Figure 9a: Cell spheroid with fluorescent beads scanned from 0 and 180 degrees. We can see beads that are visible in both images as well as beads that are visible in only one of the source images due to the spheroids shadow.

In these scans we can see certain beads that are visible only in one of the two source images. This is due to the direction of the illumination in each case and the position of those beads in the shadow created by the spheroid. The goal of the fusion in those images is to combine them so that the unique for each image information (e.g. beads in the shadow) is presented and maximized in the fused image. The original beads and the resulting fusion can be seen in figures (9.b) and (9.c). In figure (9.b) we focus on one of the beads shown in the yellow circles in figure (9.a) below the spheroid that is clearly visible in the 0 degree image and is barely visible in the 180 degree image. We also focus on a bead that is visible in both images (figure (9.c)), since it is not in the shadow in any of the images, in order for it to serve as a control for the fusion. In the fused images (on the right of both figures) we can clearly see that the intensity of the bead that was in the shadow of the 180^0 image has been recovered.

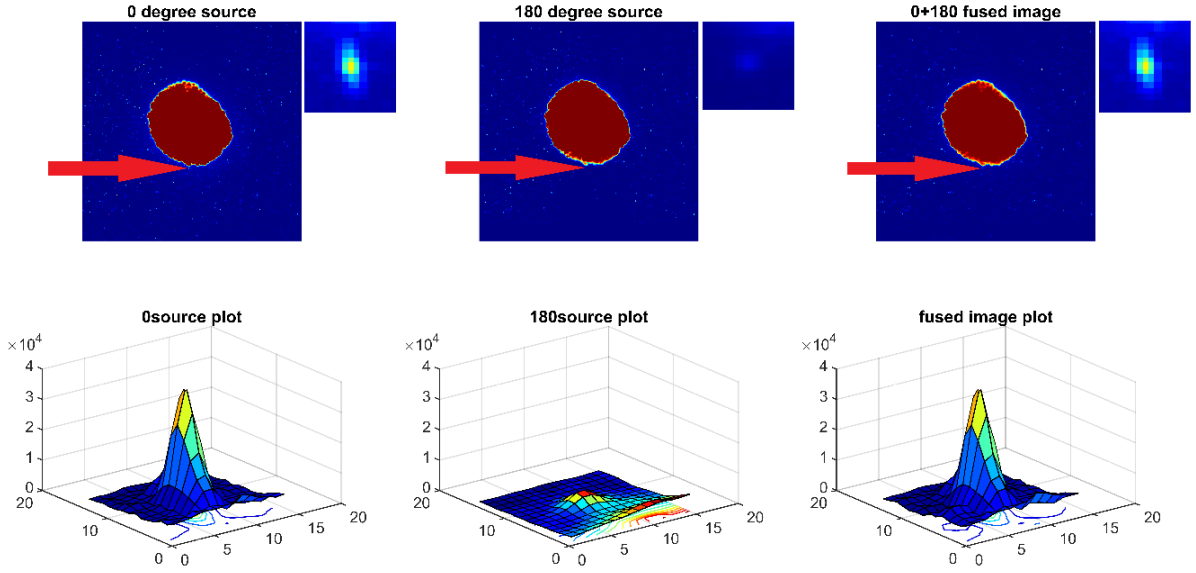


Figure 9.b: In this figure we can most clearly see the fusion of information from two source images. Specifically if we focus on the fluorescent bead just below the main spheroid we can see that it is barely visible at the 180 degree source since it's in the spheroids shadow. However when we fuse the two images we can see intensity returning to the bead in the final image.

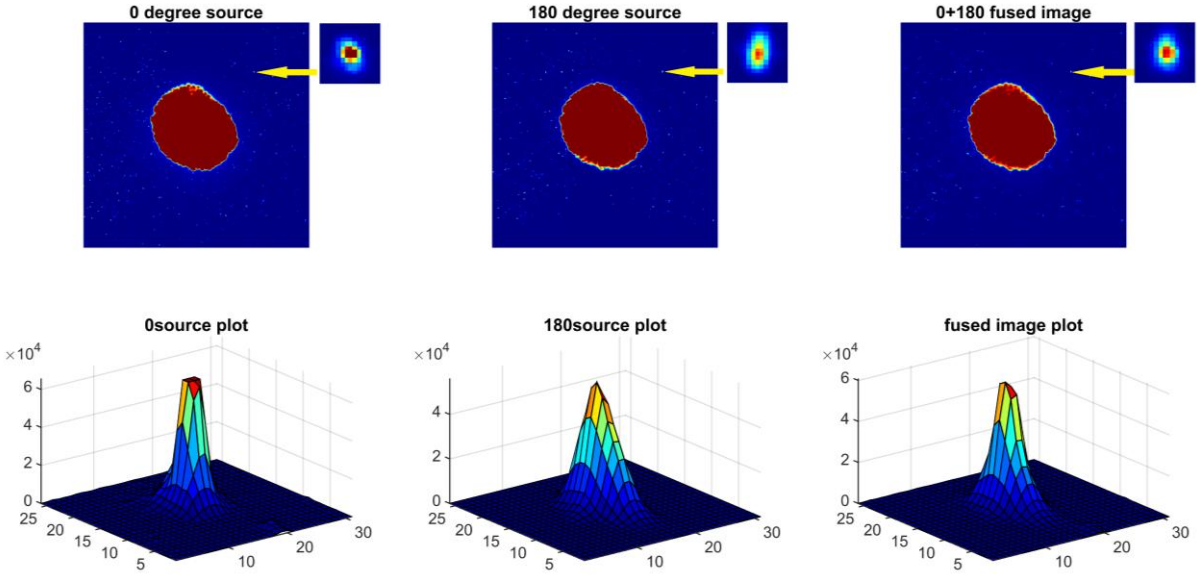


Figure 9.c: Here we focus on a control bead that is visible in both images. In the fused image we see how the fusion process combined the slightly different information from the two scans in the final bead.

Experiment 3

As a final experiment for the effect of fusion on multiangle SPIM imaging, we will use the image fusion methods, described in the previous chapters, to fuse a complete four projections (0^0 , 90^0 , 180^0 , and 270^0) dataset, and compare their results. The datasets that will be used depict a tumor spheroid stained with a far-red fluorescent dye (Draq7) which only stains the nuclei in dead and permeabilized cells. The performance for the three dimensional fusion algorithms is measured using Entropy and Feature Mutual Information for quantification.

Figure 10 depicts maximum intensity projections of the three dimensional volumetric images from the 0 degree angle. More specifically, on the left side we present the 0, 90 and 180 degree original images which were fused using seven different fusion methods. Methods based on the wavelet transform are depicted at the center of figure 10. The fusion rules for the wavelet transform based fusion are mean-max , max-max and mean-mean respectively. Following, on the right side on the top of the figure, we have the average intensity and Gaussian derivative based fusion. Maximum intensity and PCA based fused images are on the bottom. In addition on the bottom two rows of figure 10 we have presented close ups from the respective images. Figure 11 depicts the same results as figure 10 from the ninety degree angle. Finally, the results from the non-reference measures used to assess the fusion quality are shown in the bar-chart of figure 12. We notice that from the entropy measure that the PCA method stands out, while all the wavelet based methods and especially the mean-max over took the simplest methods. At the same time the FMU measure was inconclusive with all the fusion methods scoring very close to each other, with PCA and wavelets mean-mean being bit better.

Since we don't have a ground truth three dimensional image, using the reference measures for the whole images will be difficult. However, we can compare two dimensional slices with the corresponding "ground truth slice" as discussed before. According to the observation of the expert biologist, we have chosen the seventieth slice (Figure 13) of the registered spheroid volumes, as a reference. In this case

reference measures are presented for the different fusion approaches in the bar-chart (figure 14) for Mean Squared Error, Peak Signal to Noise Ratio and Structural Similarity Index and Standard Deviation (figure 12 middle). In general, PCA is the one that stands out from the quality metrics, with the wavelets based fusion methods following more consistently.

Figure 10: 0 Degree maximum projection view of the three dimensional source images and fused results and their close ups

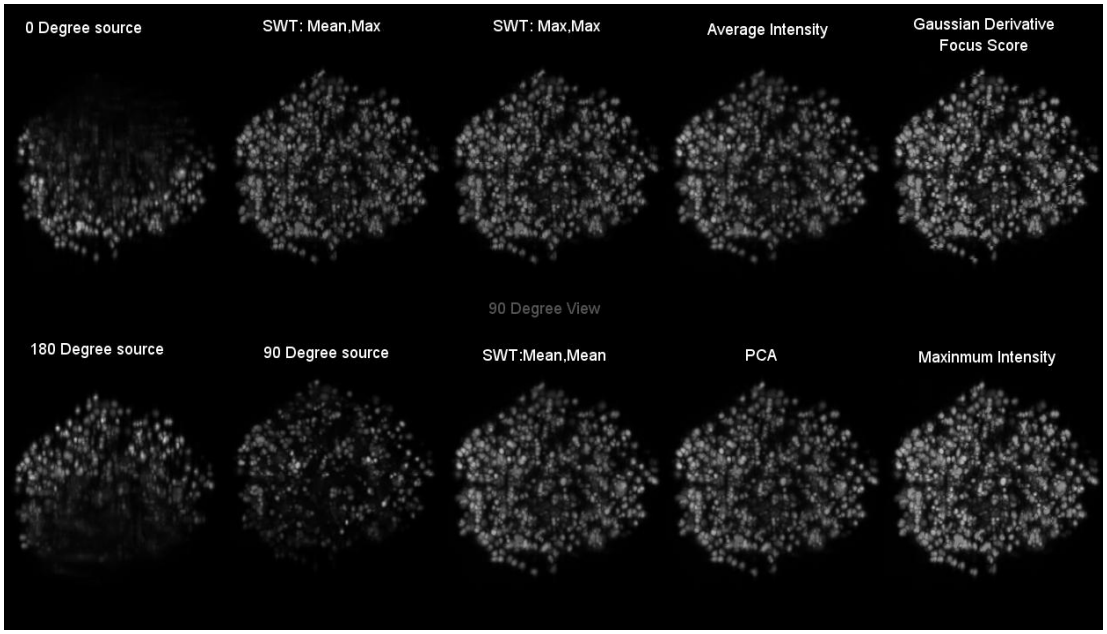


Figure 11: 90 Degree max projection view of the three dimensional source images and their fused results

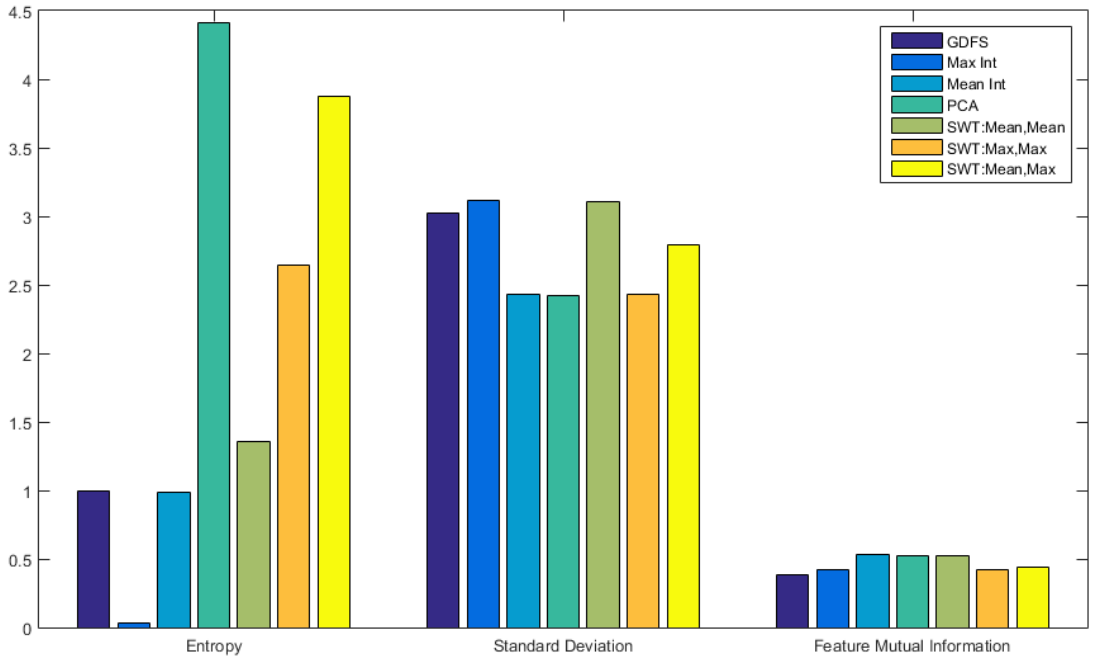


Figure 12: Image Fusion Quality assessment based on: Entropy, Standard Deviation and Feature Mutual Information

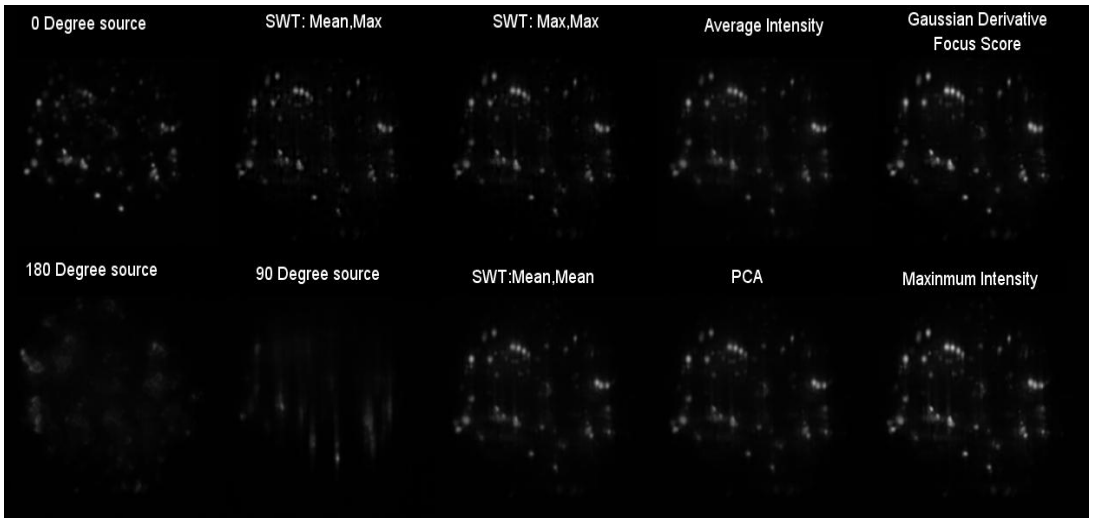


Figure 13: Slice 70 from the three dimensional source images and fused images

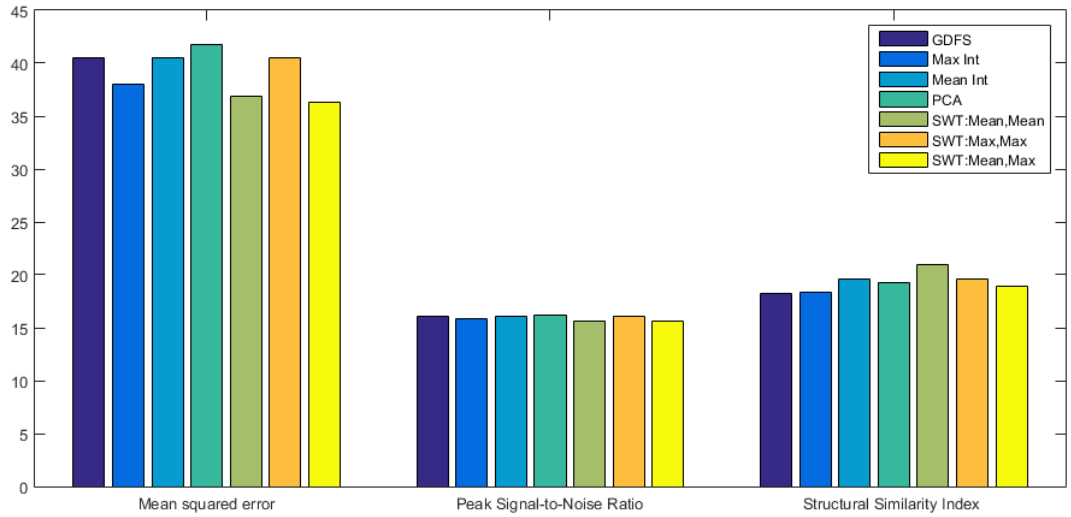


Figure 14: Image Fusion Quality assessment of the figure 13 images using the 0 degree source images as a reference. The Quality assessment is based on Mean Squared Error, Peak Signal to Noise Ratio and Structural Similarity Index

6.2 MULTI-SPECTRAL FUSION

In order to demonstrate the use of fusion for multispectral SPIM imaging a dataset from c-elegans worms marked with green fluorescent protein (GFP) and red fluorescent protein (DsRed) was devised. Two separate image stacks were acquired using different filters for the two different fluorescence agents which depict different areas of the specimen. This presents the opportunity to co-register and eventually fuse the two images and hence combine them in one data set to get a more complete view of the object we are studying. However as we can see from figure 15 when we use the fusion algorithm we fuse the multicolor information into one monochrome image and therefore we lose the effect achieved by a simple composite image, figure 16, where the two images occupy different color channels. So as long as the biomarkers target different areas of the sample, combining the images by setting the different spectra information into multiple color channels is sufficient in getting the complete, useful visual information from the multispectral scan.

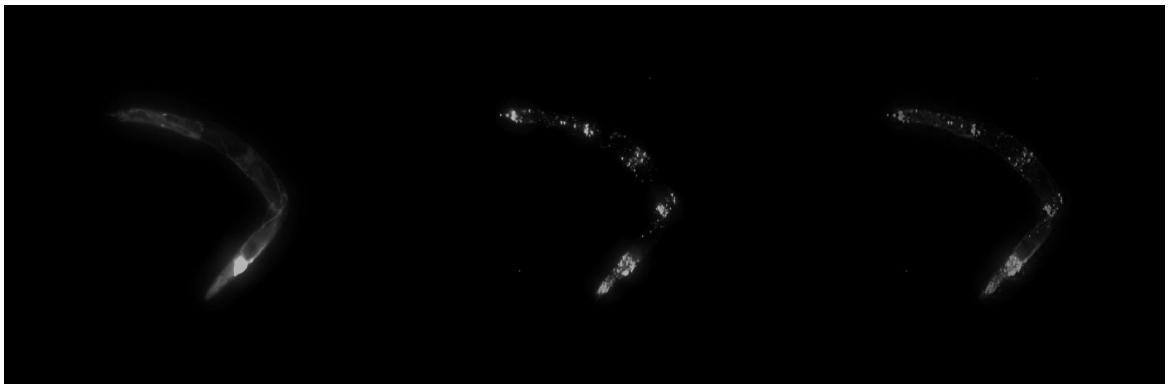


Figure 15 Multi-spectral scans (left, centre) and Their Fused image (right)

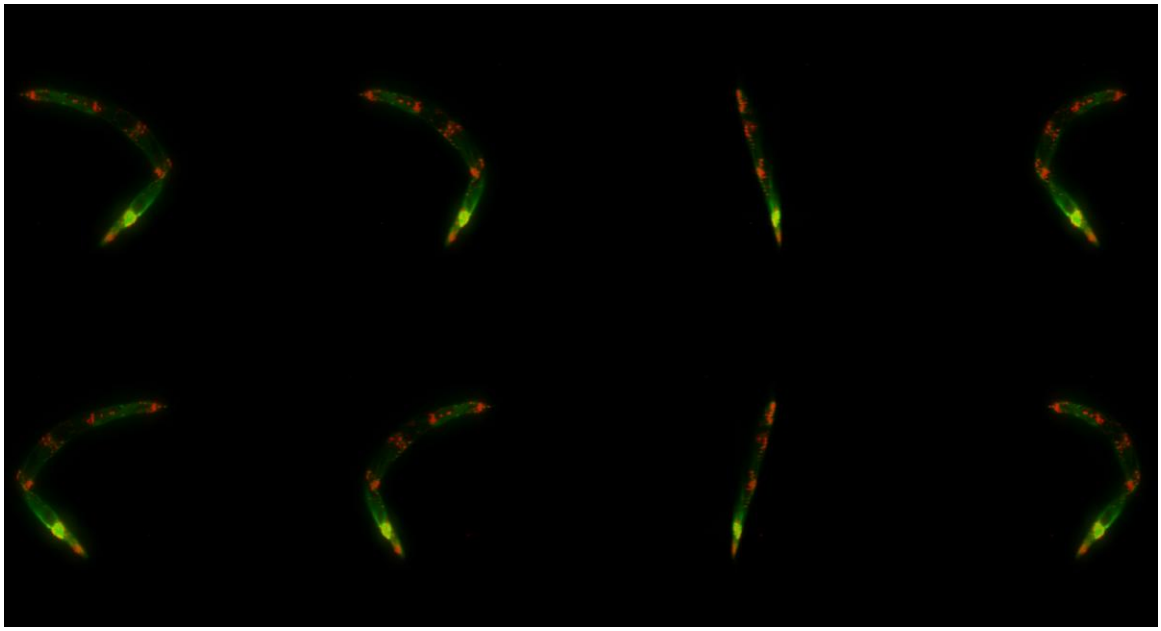


Figure 16: Composite image of the Multi-Spectral scans (360deg rotation)

6.3 COMPUTATION TIME

In this section we measure and compare the computation times of the different image fusion methods. For the measurements we two images of resolution 512x512. The algorithms were run by: Processor:Intel(R) Core(TM) i7-4700MQ CPU @ 2.40GHz, 2401 Mhz, 4 Core(s), 8 Logical Processor(s). All the algoritms were run in a MATLAB environment.

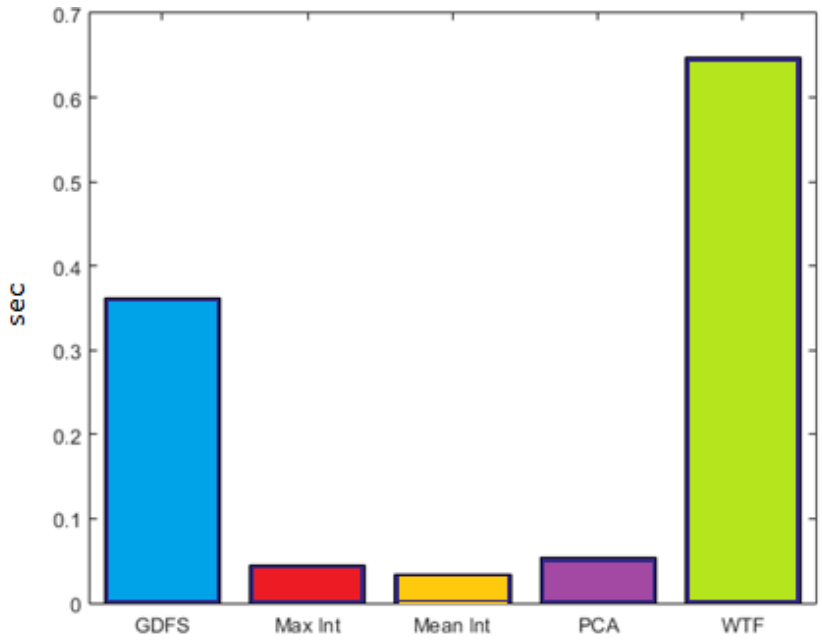


Figure 17: Computation times for a single run of the algorithms.

As we can see in figure 17 the algorithms based on the Wavelet Transform Fusion (WTF) and the Gaussian Derivative Fusion Score (GDFS) have a fairly longer computation time. So in terms of computation times we see the content based methods outperform the Wavelet based method.

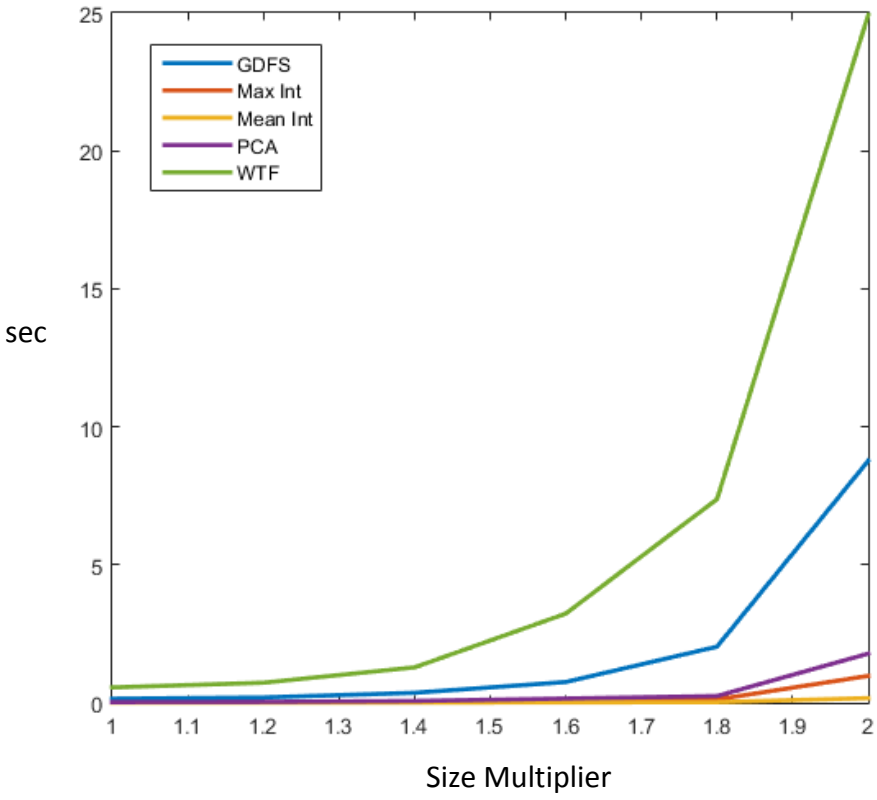


Figure 18: Calculation of computation time with increasing the size of the images

Figure 18 depicts the computation time of all the algorithms as we increase the images size. We can see that for all image fusion methods computation time increases exponentially with image size.

7 CONCLUSIONS

The aim of this thesis was to study the different methods and algorithms for image fusion and compare them in terms of their application to light sheet microscopy. In most cases we modified versions of preexisting algorithms. In some other cases we had to develop new ones based on the methods described to suit our needs. In terms of success, all of the methods described can provide adequate results depending on the dataset. However after taking into account all of the evaluation measurements and under visual inspection, the methods that stand out are the Principal Component Analysis based fusion and the Wavelet Transform based fusion by utilizing the Mean-Max fusion rule (taking the mean for the approximation coefficients and the maximum for the detail coefficients).

On the multispectral level an argument can be made for the composite multicolor image producing more useful results than image fusion overall. Even though image fusion has shown great success at combining multispectral images at remote sensing applications they do not transfer as well to light-sheet microscopy multispectral scans. In remote sensing, [18, 19] it is common to have two sensors with one providing low resolution color information and another with high resolution grayscale information. For light-sheet microscopy instead we work with images acquired by scanning for certain biomarkers emitting at different wavelengths and at different intensities, resulting in the stronger signal overtaking the weaker one in the fusion process if there is a high enough difference. It's good to mention here that the image fusion and its success depends on the type of application that we have in hand. What would be a great fusion for a remote sensing image, for example, would not be the ideal for a microscopy image. It all depends on the type of information that the user is interested in.

In addition, in the SPIM images where there is a massive difference in lateral and axial resolution, the scaling introduced to fix for the actual proportions of the object introduces an elongated image in the axial direction (figure 9). This is why a wavelet type fusion model is particularly useful in light sheet microscopy since it provides access to the horizontal, vertical and diagonal details of an image specifically. By emitting the respective details from the decomposition images, we can dismiss the elongation in the direction that is happening from the appropriate detail image.

Finally, with the application of fusion in SPIM multiangle images, we can get a complete three dimensional high resolution view of the object using the high resolution information from each of the individual projections reducing the effect of scattering and absorption. This results in us being able to mitigate the effect of scattering in penetration depth and being able to go deeper in larger more turbid objects.

As a final point when fusing images one has to be aware of the different needs and shortcomings of the imaging system used, and the required information content determined by each application. Meaning, there is no “one fits all” fusion model that could be adapted for every imaging system and every imaging application.

It should also be noted that for a meaningful fusion to take place a correct registration of the images is essential otherwise there is no point at even attempting image fusion. That being said image fusion can be extremely useful when trying to combine information from different resolution images without the loss of information.

8 REFERENCES

1. J. Huisken, J. Swoger, F. Del Bene, J. Wittbrodt and E. H. K. Stelzer, "Optical Sectioning Deep Inside Live Embryos by Selective Plane Illumination Microscopy," Science, **305**, 1007-1009 (2004)
2. J. Huisken and D. Y. Stainier, "Selective plane illumination microscopy techniques in developmental biology," Development 136: 1963-1975 (2009)
3. M. Rieckher, I. Kyparissidis-Kokkinidis, A. Zacharopoulos, G. Kourmoulakis, N. Tavernarakis, J. Ripoll, G. Zacharakis "A Customized Light Sheet Microscope to Measure Spatio-Temporal Protein Dynamics in Small Model Organisms" PLoS ONE **10(5)**: e0127869 (2015)
4. R. Weissleder, and U. Mahmood, Molecular imaging. Radiology, 2001. 219(2): p. 316-33.
5. Zacharakis, G., Time-Resolved Studies of light Transport Through highly Scattering Media. Applications in: Optical Characterization of Tissue (Optical Mammography) and Light Amplification, in Medicine. 2002, University of Crete: Heraklion.
6. B. Chance, *Optical Methods.* Annual Review of Biophysics and Biophysical Chemistry, 1991. 20: p. 1-28
7. R.Weissleder,. and V. Ntziachristos, Shedding light into live molecular targets. Nat. Med., 2003. 9(1): p. 123-128.
8. J Huisken, Stainier DYR (2009) Selective plane illumination microscopy techniques in developmental biology. Development 136: 1963–1975. doi:10.1242/dev.022426.
9. J Huisken (2012) Slicing embryos gently with laser light sheets. Bioessays 34: 406–411. doi:10.1002/bies.201100120
10. Jan Huisken, Jim Swoger, Filippo Del Bene, Joachim Wittbrodt, Ernst H. K.Stelzer Optical Sectioning Deep Inside Live Embryos by Selective Plane Illumination Microscopy.
11. J. B. Antoine Maintz∗ and Max A. Viergever. A survey of medical image registration
12. B. Zitova´, J. Flusser. Image registration methods: a survey
13. Medha V. Wyawahare, Dr. Pradeep M. Patil, and Hemant K. Abhyankar. Image Registration Techniques: An overview
14. J. B. Antoine Maintz∗ and Max A. Viergever. A survey of medical image registration
15. 2004 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.doi:10.1016/j.patcog.2004.03.010
16. Kusum Rani, Reecha Sharma. Study of Different Image fusion Algorithms (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 5, May 2013)
17. Mohamed R., Metwalli, Ayman H. Nasr, Osama S. Farag Allah, and S. El-Rabaie. "Image fusion based on principal component analysis and high-pass filter", Int. Conf. on Computer Engineering & Systems, ICCES 2009, pp. 63-70. 2009.
18. H. Li, B.S. Manjunath, and S.K. Mitra. Multisensor image fusion using the wavelet transform. Graphical Models and Image Processing, 57:235–245, 1995.
19. J.L. Moigne and R.F. Cromp. The use of wavelets for remote sensing image registration and fusion. Technical Report TR-96-171, NASA, 1996
20. L.J. Chipman, T.M. Orr, and L.N. Lewis. Wavelets and image fusion. IEEE Transactions on Image Processing, 3:248–251, 1995.
21. Chun-Lin, Liu ,A Tutorial of the Wavelet Transform, February 23, 2010
22. Haar, A. [1910]. "Zur Theorie der Orthogonalen Funktionensysteme," Math. Annal., vol. 69, pp. 331–371.
23. Shannon, Claude E. (July–October 1948). "A Mathematical Theory of Communication". Bell System Technical Journal **27** (3): 379–423.doi:10.1002/j.1538-7305.1948.tb01338.x
24. Han, Te Sun & Kobayashi, Kingo (2002). Mathematics of Information and Coding. American Mathematical Society. pp. 19–20.ISBN 978-0-8218-4256-0.
25. Carter, Tom (March 2014). An introduction to information theory and entropy. Santa Fe. Retrieved Aug 2014.
26. T.M. Cover, J.A. Thomas, Elements of Information Theory, John Wiley & Sons, Inc., New York, 1991.
27. Haghighat, M. B. A.; Aghagolzadeh, A.; Seyedarabi, H. (2011). "A non-reference image fusion metric based on mutual information of image features". Computers & Electrical Engineering **37** (5): 744–756. doi:10.1016/j.compeleceng.2011.07.012

28. F. Maes, A. Collignon, D. Vandermeulen, G. Marechal and P. Suetens, "Multimodality image registration by maximization of mutual information," IEEE Trans. Med. Imag., vol. 16, p. 187–198, Mar. 1997.
29. Dr. H.B. Kekre et al. Review on image fusion techniques and performance evaluation parameters / International Journal of Engineering Science and Technology (IJEST) Vol. 5 No.04 p.880-889 April 2013
30. Dr. H. B. Kekre, Archana Athawale, Dipali Sadavarti,"Algorithm to Generate Kekre's Wavelet Transform from Kekre's Transform", International Journal of Engineering Science and Technology,Vol. 2(5), 2010, pp. 756-767.
31. T. Stathaki. Image Fusion: Algorithms and Applications. (1st edition)[On-line].
32. C. Sidney Burrus A multiresolution formulation of Wavelet Systems (online)
33. Zeeuw, P.M. (1998), "Wavelet and image fusion," CWI, Amsterdam, March 1998.
34. K. Kannan and S. Arumuga Perumal: optimal level of decomposition of stationary wavelet transform for region level fusion of multifocused images, ictact journal on image and video processing, november 2010, issue: 02 issn: 0976 – 9102 (online)
35. Wang Zhou, Bovik, Alan C., Sheikh, Hamid R., and Simoncelli, Eero P.Image Qualifty Assessment: From Error Visibility to Structural Similarity. IEEE Transactions on Image Processing, Volume 13, Issue 4, pp. 600–612, April 2004
36. R.C, Gonzalez,., R.E. Woods, S.L. Eddins, Digital Image Processing Using MATLAB, New Jersey, Prentice Hall, 2003, Chapter 11.
37. M.B.A. Haghighat, A. Aghagolzadeh, H. Seyedarabi, "A Non-Reference Image Fusion Metric Based on Mutual Information of Image Features," Computers and Electrical Engineering, vol. 37, no. 5, pp. 744-756, Sept. 2011.
38. M. Haghighat, M.A. Razian, "Fast-FMI: non-reference image fusion metric,"
39. 8th International Conference on Application of Information and Communication Technologies (AICT), pp. 1-3, 2014.
40. Geusebroek, J.-M., Cornelissen, F., Smeulders, A. W.M. and Geerts, H. (2000), Robust autofocusing in microscopy. Cytometry, 39: 1–9.
41. R. Agrawal, "Tthe Theory of Wavelet Transform and its implementation using Matlab",(Video online):site:https://www.youtube.com/watch?v=dSi9mLaa-WE ,accessed: March 2015
42. R.C. Gonzalez, R.E. Woods, Digital Image Processing,Addison-Wesley, Reading, MA, 1993.
43. V. Ntziachristos. Going deeper than microscopy: the optical imaging frontier in biology, Nature methods | VOL.7 NO.8 | AUGUST 2010 | 603-614

9 ACKNOWLEDGEMENTS

Μέσω αυτής της εργασίας απέκτησα πολλές και σπουδαίες γνώσεις καθώς επίσης και πολύτιμες εμπειρίες. Βέβαια η διεξαγωγή και η ολοκλήρωση της δεν θα ήταν δυνατή εάν όλο αυτό τον καιρό δεν υπήρχαν κάποιοι άνθρωποι να με στηρίξουν, να με βοηθήσουν και να με κατευθύνουν κατάλληλα. Για το λόγο αυτό θέλω να ζητήσω ένα μεγάλο ευχαριστώ στους ανθρώπους που με προτρέψανε να κάνω την εργασία αυτή και επιβλέποντες της εργασίας αυτής Δρ. Αθανάσιο Ζαχαρόπουλο, Δρ. Γιάννη Ζαχαράκη για την απεριόριστη βοήθεια και το χρόνο που μου αφιέρωσαν ώστε να πετύχω τους στόχους μου. Ακόμα τους Δρ Βαγγέλη Λιάπη και Στέλιο Ψυχαράκη για την συνεργασία που είχαμε και για την ηθική υποστήριξη όλο αυτό το χρονικό διάστημα που διήρκησε η εργασία.