

Wavelet-Based Image Registration

Christopher Paulson, Soundararajan Ezekiel*, and Dapeng Wu

Department of Electrical and Computer Engineering,

University of Florida, Gainesville, FL 32611

* Department of Computer Science,

Indiana University of Pennsylvania, Indiana, PA 15705

Correspondence author: Prof. Dapeng Wu, wu@ece.ufl.edu, <http://www.wu.ece.ufl.edu>

ABSTRACT

Image registration is a fundamental enabling technology in computer vision. Developing an accurate image registration algorithm will significantly improve the techniques for computer vision problems such as tracking, fusion, change detection, autonomous navigation. In this paper, our goal is to develop an algorithm that is robust, automatic, can perform multi-modality registration, reduces the Root Mean Square Error (RMSE) below 4, increases the Peak Signal to Noise Ratio (PSNR) above 34, and uses the wavelet transformation. The preliminary results show that the algorithm is able to achieve a PSNR of approximately 36.7 and RMSE of approximately 3.7. This paper provides a comprehensive discussion of wavelet-based registration algorithm for Remote Sensing applications.

Keywords: Wavelets, image registration, Remote Sensing, multi-modality

1. INTRODUCTION

To introduce the subject of image registration, it is important to understand that it is the fundamental enabling technology in computer vision that aligns two or more images together taken at different times (multitemporal analysis), viewpoints (multiview analysis), and/or sensors (multimodal analysis).¹ Developing an accurate image registration algorithm will significantly improve the techniques for computer vision problems such as tracking, fusion, change detection, autonomous navigation. There has been a significant amount of research that has been conducted in developing image registration algorithms and some of the algorithms have been compiled in a survey by Brown in 1992² and Zitova and Flusser in 2003.¹ However, image registration has not been solved yet because the algorithms are not robust since the algorithm is only able to register particular types of images due to the parameter settings and not able to register other images.³ Another problem within image registration is the parallax problem caused by the high rise buildings because the buildings appear to be swaying; therefore, causing other computer vision algorithms to have inadequate performances. Also registration algorithm needs to perform faster than real time, so other algorithms such as a tracking can track an object in real time which will be vital in defense applications. However, image registration does not have algorithm that is able to perform at such a speed due to high computational complexity.

Image registration typically consists of the following steps:

1. **Preprocessing:** modifies both the sensed (input) and reference (base) image in order to improve the performance of the feature selection and feature correspondence of image registration because some images may be blurry or have significant amount of noise which will dramatically affect the outcome of the algorithm.³ Some techniques alleviating the noise (Image Smoothing) are median filters, mean filters, gaussian filters, etc.³ Also, the techniques for deblurring (Image Sharpening) are the Laplacian, high boost filtering, gradient, etc.⁴
2. **Feature Extraction:** selects the key features such as corners, lines, edges, contours, templates, regions, etc. which will be used to do feature correspondence.³ Some examples of feature selection are Harris corner detector, gradient, Hough transform, etc.³
3. **Feature Correspondence:** matches the key features selected in the reference image and the sensed image to see which points in the reference image matches with the points in the sensed image.³ Cross correlation, mutual information, template matching, Chamfer, etc. are a few examples of feature correspondence.³

4. **Transformation Function** aligns the sensed image to the reference image by the mapping function.³ A few examples of transformation functions are affine, projective, piecewise linear, thin-plate spline, etc.³
5. **Resampling** takes the coordinate points location of the discrete points and transforms them into a new coordinate system because the sensed image is a uniformly spaced sample of a continuous image.^{3,5} Some examples are nearest neighbor, bilinear, cubic spline, etc.³

Most importantly, wavelets is a mathematical method to decompose signals into approximate and detailed coefficients which allows the signal to be described in several levels from the coarse level (lowest resolution) to the finest level (highest resolution).⁶ Some examples of wavelets are Haar, Daubechies, Coiflets, spline, etc.⁷ In general, the functionality of wavelets in 2-D is that the columns of the original image is passed through a high-pass and low-pass filter.⁴ Then the rows of the filtered image are passed through the high-pass and low-pass filter.⁴ If the image is transformed by another level, then the approximate coefficients will be used to transform the image.⁴ Each pass through the filter decrease both the row and column by a multiple of two.⁴ The algorithm keeps repeating these steps until the algorithm has reached n levels which is specified by the user (in our algorithm $n = 4$).⁴ Refer to Fig. 1 to see the process of the wavelet decomposition.

After the wavelet decomposition has been completed, the image will be divided into four subimages which are the approximate, horizontal, vertical, and diagonal. In order to obtain the approximate coefficients, the rows and columns are passed through the low-pass filter which resembles the original image, but at a smaller resolution.⁴ Next the horizontal coefficients are obtained by passing the rows through the low-pass filter and the columns through the high-pass filter which will emphasize the horizontal edges.⁴ Also the vertical coefficients obtained by passing the columns through the low-pass filter and the rows through the high-pass filter that will stress the vertical edges.⁴ Lastly, when both the columns and rows are passed through the high-pass filter, this will produce the diagonal coefficients which accents the diagonal edges.⁴ Refer to Fig. 2 to see the result of the decomposition. Now some of the benefits of using wavelets decomposition are that the important features of original image are preserved in the approximate coefficients, decrease computational speed, emphasize strong image features, and can be implemented on a parallel computer.⁸ One disadvantage with wavelet decomposition is that wavelets are not shift invariant; therefore, wavelets are not able to change with the translation operator.⁸

In this paper, we propose a wavelet-based image registration algorithm that uses the approximate coefficients to perform image registration. The data that we are using is the LAIR data of the CLIF2007 data set provided by Wright-Patterson Air Force Base. First the algorithm reads in the base (reference) image and input (sensed) image and removes the noise by applying the Gaussian filter. Then the filtered images are decomposed by using Daubechies wavelets and the approximate coefficients are extracted from the images. Once that is completed, the algorithm performs the gradient in both the x and y direction on the approximate coefficients to find the edges, only keeps the maximum gradient of each row in the x and y direction, and then combines all the maximum gradient of x and y onto one image which these points are the feature points in the coarse level. Next, the algorithm takes these feature points and reconstruct them onto the original size image, but now one point in the coarse level equals sixteen points in the finest level; therefore, the algorithm only keeps the points that lie on the edges of the object. Then feature correspondence is done by template matching between the input and base image by using correlation. After the correspondence is found, then RANSAC is used to eliminate the outliers to allow for better results for registration. Then the registration is performed by using projective transformation function and the resampling technique used was bicubic interpolation. Lastly the algorithm compute the Root Mean Square Error (RMSE) and the Peak Signal to Noise Ratio (PSNR). Some goals that we are trying to meet with this algorithm is to have the algorithm be fast, robust, multi-modal, automatic, RMSE below 4 intensity values, and PSNR above 35.

The remainder of this paper is organized as follows. Section 2 shows the previous work that has been conducted with using wavelets in image registration. In Section 3, describes our wavelet-based image registration algorithm. Section 4 shows preliminary results of the algorithm. Lastly, we conclude the paper with Section 5 which discusses the conclusion of the paper.

2. PREVIOUS WORK DONE WITH WAVELETS IN IMAGE REGISTRATION

Now the paper is going to discuss how wavelets has been used in other image registration algorithms, but some of the key differences from our algorithm compared to the other algorithms are that the wavelets are used only to detect features, use different wavelet transforms, similarity metrics, transformation techniques, and/or resampling methods. For example, Le

Moigne, et al.⁸ uses wavelet decomposition for feature selection process by computing a histogram of the horizontal (HL) and vertical (LH) coefficients for all the levels of the wavelet decomposition and saves only the points that are 13% to 15% above the maxima of the wavelet coefficients. Another image registration algorithm was created by Fonseca, et al.⁹ which selects features by using the local modulus maxima of the wavelet transform and thresholding is applied on features to eliminate insignificant feature points. In order to find the correspondence this algorithm uses the maximum correlation coefficients of the approximate (LL) coefficients and utilizes the affine transformation as the transformation function. Next, Zheng, et al.¹⁰ uses Gabor wavelet decomposition, the algorithm does feature extraction by finding the local maxima of the energy measure, uses affine as the transformation function, bilinear interpolation as the resampling technique, and feature correspondence by mutual correlation coefficients. The next algorithm is created by Li, et al.¹¹ which performs feature extraction by extracting a contour using a wavelet-based scheme. After feature extraction, the algorithm performs a voting algorithm on each contour point based off the intensity value which the algorithm keeps the highest score from the voting algorithm. Then the algorithm uses the normalized correlation as the similarity measure and the transform parameters are computed using the matched points which a consistency test was used to filter out mismatched points. Another example of an image registration algorithm was designed by Corvi, et al.¹² which used the residue images of the discrete wavelet transform (DWT) and clustering technique to obtain the initial transformation parameters. Also this algorithm used both the maxima and minima of the DWT coefficients to allow for more points for the feature correspondence and least mean square estimation. Next Unser, et al.¹³ computed the B-Spline for the images and used the gradient-based least squares optimization in conjunction with the coarse-to-fine iteration strategy. Another image registration algorithm that used wavelets was produced by Djamdji, et al.,¹⁴ this algorithm computed the wavelets by using the algorithm à trous, feature points that are local maxima are the only points kept, compares the position of the detected feature points with that of the reference image, and then repeats the steps in the next level until the algorithm reaches the finest level. Then Quddus, et al.¹⁵ is another example of image registration algorithm which used dyadic wavelet transform as edge detection, and used mutual information in multiscale decomposition. Now Wu, et al.¹⁶ designed an image registration algorithm that used the standard DWT due to the simplicity of the transform and in order to improve the robustness of the algorithm, the algorithm used the approximate (LL) coefficients to register images using sum of absolute differences at the lower resolution and mutual information at higher resolution. Also Wong, et al.¹⁷ created an image registration algorithm that uses complex wavelet phase coherence moment estimation such as Gabor and dual-tree complex wavelets for feature point detection. In order to do feature correspondence, Wong's algorithm uses normalized cross correlation between maximum complex phase coherence moments. Then the maximum distance sample consensus is used to get rid of erroneous control points and for the remaining control points, the location is adjusted iteratively to maximize the normalized cross correlation. Next algorithm created by Xishan, et al.¹⁸ used a feature based approach to do registration by using integrated matching criteria of invariant moments and orientation of contours. In order to extract features from the images a wavelet based edge detection was used by transforming the edge strength into fuzzy field to extract well defined matchable contours. Then this algorithm performs feature correspondence by combining the invariant moment and orientation function to determine the correspondence between the contours in the images. Also Xishan chooses to use the affine transformation as the transformation function. In Bejar's, et al.¹⁹ algorithm, the wavelets were used to extract feature points which were the edge points. Then the algorithm used normalized cross correlation to perform the feature correspondence. Next, Bejar used the moment of inertia in order to estimate the rigid transformation parameters and applied a consistency test in order to eliminate false control points. Another image registration algorithm was designed by Li, et al.²⁰ which decomposes the images by using the discrete wavelet frame transform which is shift invariant compared to the dyadic wavelet transform. Then Li's algorithm computed the energy map from the detail coefficients and used the genetic algorithm to obtain the minimum sum of absolute differences between the two energy maps.

Next the paper is going to discuss the pitfalls of these algorithms. To start off with the algorithms that uses mutual information would significantly decrease computational speed due to mutual information being computationally expensive; therefore, this will increase the amount of time that it takes the algorithm to do registration.²¹ Since one of the goals of image registration is to do image registration faster than real time, so other algorithm such as tracking can be done in real time, but the first step of tracking is to do image registration. Therefore, if image registration is not faster than real time, then tracking cannot be performed in real time, so we need to find a better way to do feature correspondence. When the algorithms uses DWT, the algorithm will not be shift invariant; therefore, these algorithms will not be robust.²² Some wavelets that are able to be shift invariant are algorithm à trous and Gabor wavelets; however, these algorithms are computationally expensive which is undesirable since the algorithm needs to be faster than real time.^{22, 23} According to Fauqueur, et al,²³ we can use the Dual Tree Complex Wavelet Transform (DTCWT) to solve the shift invariant problem

and has limited redundancy so this method will be less computationally expensive. Also the algorithm that uses affine transformation as the transformation function is not suitable for all images which can be seen from the results obtained from Table 2; therefore, the algorithm is not robust.

Furthermore, our algorithm is different from the other wavelet algorithms by the method in which our algorithm uses the wavelet coefficients for the feature extraction and feature correspondence, the kind of wavelet used, different similarity measures were used, and/or transformation function. Since we use DWT our algorithm will not be shift invariant; however in future work we will try other wavelets such as DTCWT to see if this improves the robustness of the algorithm. But the novelty of the algorithm will come from the future work with the ability of the algorithm to analyze when the image is registered well and when it is not. Also having the algorithm being able to determine where poor registration has occurred within the image and the potential reason for the failure which will be beneficial for defense application because it will allow other algorithm to know how much leniency to give images and know when the algorithm will not perform as well.

3. WAVELET-BASED IMAGE REGISTRATION

Now we are going to discuss our wavelet-based image registration algorithm and explain each step of the image registration process.

3.1. Preprocessing

First step of our algorithm is to use a Gaussian filter was used in the preprocessing step to eliminate the noise in the image to improve the feature extraction. Without the Gaussian filter, the final results were approximately 5.6 for the RMSE and 33.1 for the PSNR, but with the Gaussian filter the final results were significantly improved to the RMSE being approximately 3.7 and the PSNR being approximately 36.7. The reason why there is a significant improvement in the results is because the Gaussian filter smoothed out some of the noise in the image; therefore, allowing the feature selection to select better points in both the sensed and reference image.

Now the equation for the Gaussian filter is Eq. (1). For the images we are using we set $\sigma = 1$ and the mask size of the Gaussian filter is 101×101 . In order to filter the image with the Gaussian filter, the image must be convoluted with the Gaussian filter Eq. (2). Where $I(x,y)$ is the image, $a = \frac{m-1}{2}$, and $b = \frac{n-1}{2}$.⁴ Now m and n are the size of the mask and in our case $m = 101$ and $n = 101$.

$$G(x, y) = \frac{1}{2\pi\sigma^2} * \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (1)$$

$$G(x, y) \star I(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b G(s, t)I(x - s, y - t) \quad (2)$$

In the future what we would like to do for the preprocessing phase is to try to filter the image using the wavelet filter. This would allow us to get rid of the Gaussian filter to save computation time since the algorithm already calculates the wavelets and the algorithm can use the information to filter the image.

3.2. Feature Extraction

After the sensed and reference image have been preprocessed then the images are decomposed by Daubechies 1 wavelets to the fourth level. In this algorithm, the Daub1 wavelets were used which is equivalent to the Haar wavelets, but in the future we are going to use a higher DaubJ wavelets such as Daub4. The reason for using a higher order DaubJ is because it will usually result in the signal converging faster to the original signal which means the first few levels of the detail coefficients are negligible so the approximate coefficients at the first few levels are very similar to the original image.⁷ Also DaubJ multiresolution analysis produces a smoother signal than Daub1 because Daubechies wavelets have overlapping windows; therefore, Daubechies wavelets are able to detect the high frequency changes in the high frequency coefficients.^{7, 24} Whereas the Haar wavelet calculates the average and difference of a pair of values, then slides over two time unit, and repeats the calculation until the end of the signal is reached.²⁴ Therefore, causing the wavelet transform not to show the high frequency changes in the high frequency coefficients when a dramatic change occurs between an even value to an odd value.²⁴ However, higher DaubJ wavelets are not always the solution because sometimes a higher DaubJ

will cause the wavelets to have a longer support which means that the detail coefficients are going to contain a significant amount of energy.⁷ Since the detail coefficients contain a significant amount of energy, this will require more values to be used, so the data will not be able to be compressed as much.⁷ One thing to note is that the term longer support refers to all the values in the wavelets that are not zero.⁷

First we must get the equation for the scaling function and separable wavelet functions, so Eq. (3) is the separable scaling function, Eq. (4) - Eq. (6) are the separable wavelets. Next we need to define what the terms mean; therefore, ψ^H refers to the change along the columns which means the horizontal edges, ψ^V is the difference along the row which refers to the vertical edges, ψ^D is the variation along the diagonals, $\varphi(\cdot)$ means the 1D scaling function in the \cdot direction, and $\psi(\cdot)$ stands for the 1D wavelet function in the \cdot direction.⁴

$$\varphi(x, y) = \varphi(x)\varphi(y) \quad (3)$$

$$\psi^H(x, y) = \psi(x)\varphi(y) \quad (4)$$

$$\psi^V(x, y) = \varphi(x)\psi(y) \quad (5)$$

$$\psi^D(x, y) = \psi(x)\psi(y) \quad (6)$$

Now that we have the separable scaling and wavelet functions, we can assign the scaled Eq. (7) and translated basis Eq. (8) - Eq. (10).⁴

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

$$\psi_{j,m,n}^H(x, y) = 2^{j/2}\psi^H(2^j x - m, 2^j y - n) \quad (8)$$

$$\psi_{j,m,n}^V(x, y) = 2^{j/2}\psi^V(2^j x - m, 2^j y - n) \quad (9)$$

$$\psi_{j,m,n}^D(x, y) = 2^{j/2}\psi^D(2^j x - m, 2^j y - n) \quad (10)$$

Once we have the basis function we can now define the discrete wavelet transform for the image which can be found at Eq. (11) - Eq. (14) where $M, N, H, V, D, W_\varphi(j_0, m, n), W_\psi^H(j, m, n), W_\psi^V(j, m, n), W_\psi^D(j, m, n)$ represents the number of columns in the image, number of rows in the the image, horizontal, vertical, diagonal, approximate coefficients at scale j_0 which is usually equal to 0, horizontal, vertical, and diagonal detail coefficients at scale j where $j \geq j_0$ respectively.⁴

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\varphi_{j_0,m,n}(x, y) \quad (11)$$

$$W_\psi^H(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\psi_{j,m,n}^H(x, y) \quad (12)$$

$$W_\psi^V(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\psi_{j,m,n}^V(x, y) \quad (13)$$

$$W_\psi^D(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\psi_{j,m,n}^D(x, y) \quad (14)$$

Since we use the Daub1 which is essentially the Haar wavelets, the scaling function of the Haar wavelet is Eq. (15) and wavelet function is Eq. (16).⁴

$$\varphi(x) = \begin{cases} 1 & 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

$$\psi(x) = \begin{cases} 1 & 0 \leq x < .5 \\ -1 & .5 \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

After the wavelet decomposition of both the base (reference) and input (sensed) image, the algorithm extracts the approximate coefficients from the wavelets. The purpose for decomposing the images is to accelerate the computation speed by making the original images smaller. Once the algorithm has the approximate coefficients, then the gradient in both the x and y direction are calculated in the input and base image, and only the maximum gradient in each row is saved because that is where an edge point should lie since that should be an area of greatest change. In order to see the equation for the gradient, refer to Eq. (17) which is the rate of change in the x and y direction.⁴ Next, the maximum gradients in the x and y direction are combined together for both the base and input image. Then the algorithm create two black images the same size as the coarse image, finds the location of the control points, extracts the corresponding approximate coefficients, places the approximate coefficients into the black images, and the images are reconstructed to the original size. Now the reconstructed image provides the algorithm with a mask to tell where the control points are located on the original image. One problem is that one point in the coarse image is equal to 16×16 points in the finest level, since the algorithm went to the fourth level of the wavelets. To fix this problem, for each 16×16 control point, the algorithm takes the absolute difference of each column of the control point, selects the n greatest difference of each column to become the control points, and $n = 10$ in our algorithm. These points should be the points that lie on the edge of the object because that is where the greatest change would most likely occur.

$$\nabla(f) \equiv grad(f) \equiv \begin{bmatrix} g_x \\ g_y \end{bmatrix} \equiv \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (17)$$

However, there are some issues that need to be addressed in the future work. One example of what we are going to try in the future work is to see if the algorithm will improve by using the Horizontal, Vertical, and Diagonal coefficients rather than using the gradient method to save some computational time. Another issue I see with this feature extractor is the control points could potentially not be spread throughout the image which would have a negative effect on the outcome of the image registration. Some examples of such an issue is that you have a body of water running through the image which will not pick up significant features or the image has a black and white line next to each other in the image, so the greatest change would be along that line which will produce horrible results because the points will be concentrated in a certain region of the image.

3.3. Feature Correspondence

Once the algorithm has extracted the feature points, the algorithm must determine which feature points in the input image matches with the feature points in the base image. First, the points along the edges of the image are discarded because there is a $w \times w$ window (in our algorithm $w = 9$) that scans through the images to get a template to do the correspondence, so if we kept the points that were on the edge of the image then the algorithm would crash. Therefore, the region that will be scanned is $\frac{w}{2}$ through $m - \frac{w}{2}$ and $\frac{w}{2}$ through $n - \frac{w}{2}$, where m and n are the number columns and rows of the image respectively. Next the algorithm must round down $\frac{w}{2}$ because w needs to be an odd number so the control point can be located exactly in the middle of the window which when we divide an odd number by 2 we will have a decimal. Then one control point in the base image is selected, a $w \times w$ window is created around the control point which is considered the middle of the $w \times w$ window, and now there is template of a $w \times w$ window of intensity values of the base image. Once the template is extracted from the base image, then the template is compared to all the control points in a $b \times b$ window (in our algorithm $b = 500$) of the input image and b needs to be much greater than w. Before the comparison is done between the base and input control points, a $w \times w$ window is created around the input control point with the control point being the

middle of the window. Next, the comparison is done by performing the wavelet decomposition on the template, extracting the approximate coefficients, and performing correlation based off of the normalized dot product, refer to Eq. (18) to see the equation. Now the variables are tb , ti , and i which refer to the base template, input template, and the location in the $w \times w$ window, respectively. After the comparison between the two points is done, then the base control point is compared to another input control point within the $b \times b$ window, if the correlation between those two points is higher than the current maximum correlation, then that point is saved as the point with the highest correlation. Then comparing the base feature point to all the input feature points in a $b \times b$ window is done until all the input feature points within the window have been examined. Once all the input feature points have been examined, then input feature point which has the highest correlation and is above a certain threshold (our algorithm the threshold is .995) is considered the corresponding point to the base feature point. Then the algorithm goes to the next base control point and repeats the previous steps to find the corresponding input control point. The algorithm performs the feature correspondence until all the base control points have been examined. To further improve the correspondence of the algorithm, Random Sample Consensus (RANSAC) was used to get rid of the outliers. In order to have a further understanding how RANSAC works, please refer to the paper by Fischler.²⁵

$$corr = \frac{\sum_{i=1}^l tb_i \cdot ti_i}{\sqrt{\sum_{i=1}^l (tb_i \cdot tb_i) \sum_{i=1}^l (ti_i \cdot ti_i)}} \quad (18)$$

One of the biggest issues with this algorithm is that the computation time takes too long, so improvement needs to be done in this area because the computation time takes about an hour for one pair of images. An idea is to do correspondence in the coarse level, then go to the next finest level, perform correspondence just on the points that the algorithm deemed as correspondence points, and do this until the finest level has been reached which has been done in other algorithms. This should speed up computation speed considerably because the majority of the computation will be done in the coarse level which is significantly smaller than the original image. Another idea that we may try is to continue with correspondence in the coarse image, but split the image into section, find a certain number of points in each section, and once number has been reached then continue onto the next region. These are some ideas that we are going to try in our future work.

3.4. Transformation

After Feature Correspondence the algorithm uses the projective transformation because the lens and sensor nonlinearities do not exist and the scene is relatively close to the camera.³ Also we tried the other transformation function such as linear conformal, affine, and polynomial, but projective gave us the best results which was based off the RMSE and PSNR (Refer to Table 2).

The equations for the projective transformation are Eq. (19) and Eq. (20). Now a through h are the eight unknowns that can be solved by at least four non-colinear corresponding points in the images.³ Usually the number of control points that are used are more than the minimum required amount of control points which usually allows for better results at the cost of requiring more computation time to determine feature correspondence of the points.^{1,2}

$$X = \frac{ax + by + c}{dx + ey + 1} \quad (19)$$

$$Y = \frac{fx + gy + h}{dx + ey + 1} \quad (20)$$

3.5. Resampling

Following the transformation step, The algorithm uses the bicubic interpolation to do the resampling of the image. In our algorithm, bicubic interpolation was used because it produces a more accurate result than nearest neighbor or bilinear interpolation but does require more computational time.³ Also refer to Table 3 for actual results obtained by our algorithm.

Table 1. Wavelet-Based Image Registration

-
- **Preprocessing:** First, read in both the base and input image and then filter both the base and input image using Gaussian Filter.
 - **Feature Extraction:** After the image has been preprocessed, then apply wavelet transformation on both the base and input images using Daubechies Method (db1 or haar) to the fourth level. Next, extract the approximate coefficients from the fourth level and calculate the gradient of the base and input approximate images in both the x and y direction. After finding the gradient of both images, then find the maximum gradient in both the x and y direction for the base and input images and combine max gradient of the x and y direction onto one image for both images. Then reconstruct the base and input energy images into the original image size. Once energy images have been reconstructed to the original size, the next step is to map the location of the control points on the energy images onto the actual image. Since one point in the coarse level equals 16×16 in the finest level; therefore, the algorithm filters out all the unnecessary point by picking the points that lie on the edges in the finest level.
 - **Feature Correspondence:** After getting the control points, then find correspondence between base and input control points using correlation. Then use RANSAC to eliminate the outliers to further improve the registration results.
 - **Transformation Function:** Once RANSAC is completed the algorithm uses Projective transformation as the transformation function.
 - **Resampling:** The last step of the registration algorithm is to perform the resampling technique by using bicubic interpolation.
-

3.6. Algorithm Summary

In order to see a general overview of how our algorithm functions, refer to Fig. 3. Then look down below in Table 1 which shows step by step how the algorithm performs image registration, but for details for each step look at Section 3.1 through Section 3.5.

4. RESULTS

In order to see how well our algorithm performed we simulated a preliminary test which produced desirable results with the RMSE being 3.7361 and the PSNR being 36.6824 (our goal of RMSE being below 4 and PSNR being above 35 was met) (Fig. 4). One reason for using the RMSE is because the Mean Square Error (MSE) is a multiple of the energy of the difference between the images. Now RMSE is the square root of the MSE which the RMSE will allow us to compare the data as the intensity values whereas MSE would be (intensity value)². Also we used PSNR as a value measure because the PSNR is a logarithmic measure which studies have shown that our brain reacts logarithmically to alteration in light intensity.⁷ Next we explain what the variable mean in the equation below, so P in equation Eq. (21) is the total number of pixels used in the calculation, i is the column location of the pixel, and j is the row location of the pixel. Since this is only tested between one pair of images, there is still much work that needs to be done with testing of the algorithm on the rest of the data set, but the speed of the algorithm needs to improve significantly before we can test it on the other data sets.

$$MSE = \frac{1}{P} \sum_{i,j} (f_{i,j} - g_{i,j})^2 \quad (21)$$

$$RMSE = \sqrt{MSE} \quad (22)$$

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (23)$$

5. CONCLUSION

In this paper, we proposed a wavelet-based image registration algorithm, which uses wavelet coefficients for feature extraction and correspondence instead of using intensity values. Our preliminary results showed that the algorithm was able to achieve a PSNR of 36.6824 and RMSE of 3.7361 which shows that there is potential for this to become a high-performance image registration algorithm. However, there is still a significant amount of work that needs to be performed on the algorithm before we can make an informed decision on how well the algorithm performs; e.g., the algorithm is slow, the robustness of the algorithm is not yet known, and multi-modal registration has not been tested.

Our future work is listed as below:

- Develop a robust image registration algorithm that performs registration faster than real time, so other algorithms that require registration as the first step can still function in real time such as tracking, change detection, fusion, etc. which will be vital for defense application. One way to decrease the time is to utilize the multiresolution of wavelets, but more needs to be done in order for the algorithm to be able to perform faster than real time.
- Test algorithm on more data to determine the robustness of the algorithm and to understand where the algorithm fails. The importance of knowing where the algorithm fails will assist in improving the performance of the algorithm
- Create an analyzer that is able to detect when poor registration occurred, where the misalignment is located, and some possible reason for the failure to produce a well registered image. This will be extremely beneficial for other algorithms, so they know which images there has to be more leniency provided when running the algorithm.

Disclaimers

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of AFRL or the U.S. Government.

Acknowledgement

This material is based on research sponsored by AFRL under agreement number FA8650-06-1-1027. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The authors would like to thank Olga Mendoza-Schrock, James Patrick, Ed Zelnio, and Mark Minardi for all their support and constructive suggestions on how to improve the algorithm.

REFERENCES

1. B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and vision computing* **21**(11), pp. 977–1000, 2003.
2. L. Brown, "A survey of image registration techniques," *ACM computing surveys (CSUR)* **24**(4), p. 376, 1992.
3. A. Goshtasby, *2-D and 3-D image registration for medical, remote sensing, and industrial applications*, Wiley-Interscience, 2005.
4. R. Gonzalez and R. Woods, "Digital Image Processing. ISBN: 9780131687288," 2008.
5. J. Parker, R. Kenyon, and D. Troxel, "Comparison of interpolating methods for image resampling," *IEEE Trans. Med. Imaging* **2**(1), pp. 31–39, 1983.
6. E. Stollnitz, T. DeRose, and D. Salesin, "Wavelets for computer graphics: A primer, part 1," *IEEE Computer Graphics and Applications* **15**, pp. 75–75, 1995.
7. S. Walker James, "A Primer on WAVELETS and their Scientific Applications," USA, *Universiti of Wisconsin* , p. 300, 2008.
8. J. Le Moigne, W. Campbell, and R. Crompt, "An automated parallel image registration technique based on the correlation of wavelet features," *IEEE Transactions on Geoscience and Remote Sensing* **40**(8), pp. 1849–1864, 2002.
9. L. Fonseca and M. Costa, "Automatic registration of satellite images," in *Proceedings of the Brazilian Symposium on Computer Graphic and Image Processing, Brazil*, pp. 219–226, Citeseer, 1997.
10. Q. Zheng and R. Chellappa, "A computational vision approach to image registration," *IEEE Transactions on Image Processing* **2**(3), pp. 311–326, 1993.

Table 2. Results of using Different Transformation Techniques

	Linear Conformal	Affine	Projective	Polynomial
PSNR	23.4397	30.6937	36.5214	36.1649
RMSE	17.1614	7.4449	3.8060	3.9655

Table 3. Resulting using Different Resampling Technique

	Nearest Neighbor	Bilinear	Bicubic
RMSE	4.2286	3.8654	3.8620
PSNR	35.6068	36.3868	36.3945

11. H. Li and Y. Zhou, "Automatic EO/IR sensor image registration," in *IEEE Int. Conf. on Image Proc.*, volume B, pp. 161–164, 1995.
12. M. Corvi and G. Nicchiotti, "Multiresolution image registration," in *Image Processing, 1995. Proceedings., International Conference on*, 3, 1995.
13. M. Unser, A. Aldroubi, and C. Gerfen, "A multiresolution image registration procedure using spline pyramids," *Proceedings of the SPIE-Mathematical Imaging: Wavelets and Applications in Signal and Image Processing* **2034**, pp. 160–170, 1993.
14. J. Djamdji, A. Bijaoui, and R. Maniere, "Geometrical registration of images: the multiresolution approach," *Photogrammetric Engineering and Remote Sensing* **59**(5), pp. 645–653, 1993.
15. A. Quddus and O. Basir, "Wavelet-Based Medical Image Registration for Retrieval Applications," in *Proceedings of the 2008 International Conference on BioMedical Engineering and Informatics-Volume 02*, pp. 301–305, IEEE Computer Society, 2008.
16. J. Wu and A. Chung, "Multimodal brain image registration based on wavelet transform using SAD and MI," *Lecture Notes in Computer Science*, pp. 270–277, 2004.
17. A. Wong and D. Clausi, "Automatic Registration of Inter-band and Inter-sensor Images using Robust Complex Wavelet Feature Representations," in *Proc. of 5th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS 2008)*, 2008.
18. H. Xishan and C. Zhe, "A wavelet-based multisensor image registration algorithm," in *Signal Processing, 2002 6th International Conference on*, 1, 2002.
19. C. Bejar and D. Megherbi, "A feature-based image registration algorithm using the multi-resolution approach combined with moments of inertia, with application to ladar imaging," in *Proc. SPIE*, **5909**, pp. 506–516, 2005.
20. S. Li, J. Peng, J. Kwok, J. Zhang, and C. Changsha, "Multimodal registration using the discrete wavelet frame transform," in *Proceedings of the 18th International Conference on Pattern Recognition-Volume 03*, p. 880, Citeseer, 2006.
21. A. Wong and P. Fieguth, "Fast phase-based registration of multimodal image data," *Signal Processing* **89**(5), pp. 724–737, 2009.
22. N. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Applied and Computational Harmonic Analysis* **10**(3), pp. 234–253, 2001.
23. J. Fauqueur, N. Kingsbury, and R. Anderson, "Multiscale keypoint detection using the dual-tree complex wavelet transform," in *Proceedings of IEEE International Conference on Image Processing*, Citeseer, 2006.
24. F. Elfouly, M. Mahmoud, and S. Deyab, "Comparison between Haar and Daubechies Wavelet Transformations on FPGA Technology," *International Journal of Computer, Information, and Systems Science, and Engineering* **2**(1), pp. 1047–1061, 2006.
25. M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," 1981.

(This document is approved for public release via 88 ABW-10-1118.)

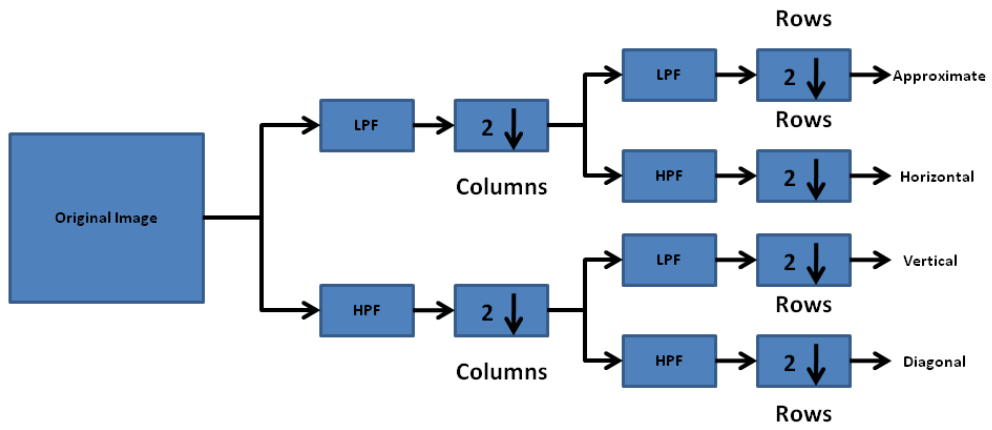


Figure 1. 2-D Wavelet Transform

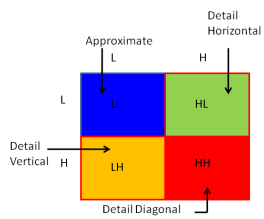


Figure 2. Wavelet Decomposition

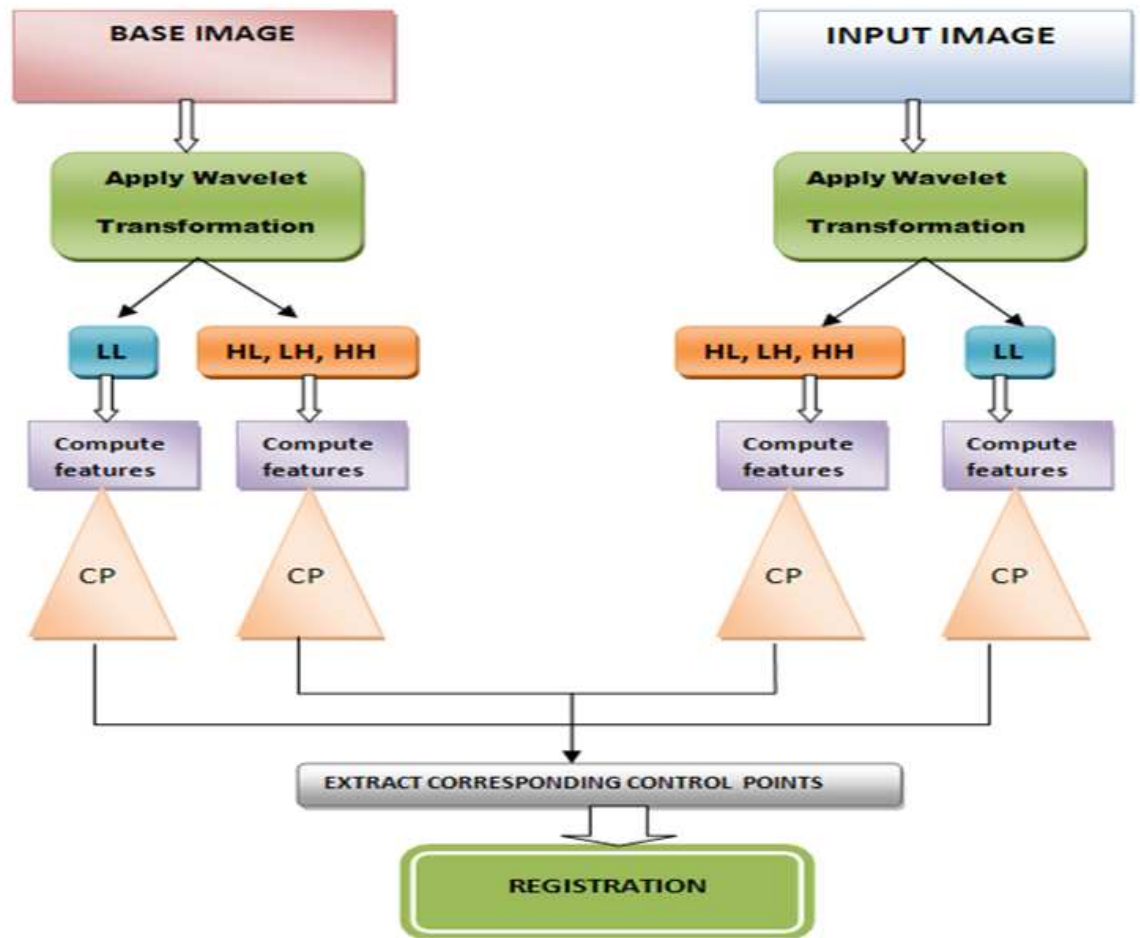


Figure 3. Flow Chart of the Wavelet Based Image Registration Algorithm

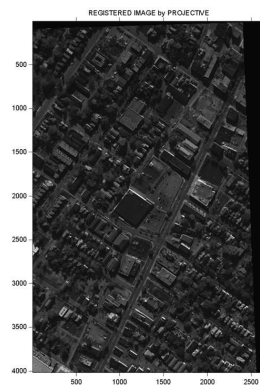


Figure 4. Registered Image