SOLUTIONS FOR ENHANCING TARGET RECOGNITION

Lieut. eng. Marius Tita

Military Equipment and Technologies Research Agency

Prof. Emil Cretu, Ph.D.

Titu Maiorescu University

Abstract:*In time the night vision devices were developed from the functionality principle point of view in two main directions: image intensifiers and thermal cameras. In this paper we try an approach with a more sensible domain of developing this two kind of sensors in one single device and fusing the two images obtained in one with a superior quality.*

The work by trying to obtain a better image, with more significant details then the two initial fused images, one thermal and other intensified fits with the symposium topic "Research and Technology" and refers to enhancing the target recognition.

1. MILITARY NEEDS

Modern warfare cannot be imagined without night vision devices. The study of these systems is in accordance with actual research directions for making possible the military actions during night time or in bad weather or poor illumination conditions. The time for finding and identifying the targets must be very short for night vision devices to provide clear images for better visualization distances so the own troops have an advantage upon the enemy. Night vision devices are very important in modern battlefield and their existence is vital.

There are two main categories of night vision devices: night vision devices with image intensifiers and thermal cameras. The main purpose of the two kinds of devices is to make possible the visualization during night time or in poor illumination conditions. The night vision devices with image intensifiers amplify the low level light radiation to a value which allows human eye to observe the scene. Thermal cameras use the thermal contrast of the objects (targets in viewed scene) in regard to medium were those are placed (background), when have the temperature or the emissivity different from the medium (always valuable).

These modern systems are not used only by the commanders on the ground that needs up-to-date imagery information to get an accurate appreciation of the battlefield as well as to assist in the planning and conducting of ground operations. Often this information can be obtained through various agencies and units. Imagery intelligence services are gathering military intelligence via the use of satellite or aerial imagery. In recent conventional warfare, this has been translated to the use of unmanned aerial vehicles (UAV) or other forms of manned and unmanned means for the collection of timely visual reports on the enemy. Information from imagery is useful and simplifies the complexity of having ground troops updating via sighting reports or verbal reports via voice communications, which frequently is inadequate to provide information that would have impacts on the outcome of operations.

There are major differences between the two viewing systems. The night vision systems with image intensifiers provide high quality images for short observation distances and the possibility of reading maps, documents and inscriptions. Also these systems have low energetic consumption, long function autonomy, low price, short time of being ready to use and have a relative simple fabrication technology. Although the advantages, these systems have a limited range of observation, high dependence of illumination conditions, are very influenced by the atmospherical parameters and a poor ability to detect camouflaged targets.

The thermal cameras provide good quality images for long observation distances (high ability to detect targets) even in poor atmosferical condition, have no dependence on the level of illumination and can be used any time on day or night. These systems being technologically advanced have possibilities of capture, storage and easy transfer of images. All the advantages come together with large dimensions of the devices, high energy consumption (especially thermal cameras with cooling systems), complex fabrication technology, short lifetime and raised prices.

2. Image fusion and limitations

In recent times, multi-spectral imagery systems have been used extensively for area surveillance and intelligence collection. Multi-spectral imagery systems help to provide additional information and allow for the capture of imagery under imperfect conditions, such as low light and foggy conditions that may not favor conventional imagery systems which capture imagery in the visible light spectrum. They also range from sophisticated systems installed in a satellite for large area imagery intelligence to hand-portable devices used by soldiers in the battlefield for situational awareness, and its applications are highly versatile. Conventional visible light images are taken with three colors, whereas multi-spectral images explore other frequencies in the electromagnetic (EM) spectrum from visible light as well as frequencies close to visible light such as infrared(IR). A multi-spectral image allows the extraction of additional information that a conventional visible light image fails to capture. Recent developments include researches to explore the ultra-violet regions of the electromagnetic spectrum. However, most imaging systems still only have the capability to detect wavelengths from a specific region of the EM spectrum. Often the information provided by a particular sensor is complementary to another if both images are viewed as one and any drawbacks in one sensor can be filled in by another. So by combining imagery from multiple sensors, an improved imagery can be conceived to help in providing completeness to the information required.

3. Image fusion methods

3.1 Wavelet transformation fusion

Wavelet transform fusion involves decomposition of the original images into wavelet coefficients and combining them into a single image using a fusion scheme. The level of wavelet decomposition considered is selective and can be application-oriented. However this approach requires that the fusion step be done only at the same level of decomposition of each of the images. In the following sub-sections, the essentials of wavelet transform fusion such as wavelet transform, application of discrete wavelet transform (DWT) to image fusion, fusion at the decomposition level and the fusion schemes are examined.

Figure 1 explains the overall wavelet transformation fusion process.

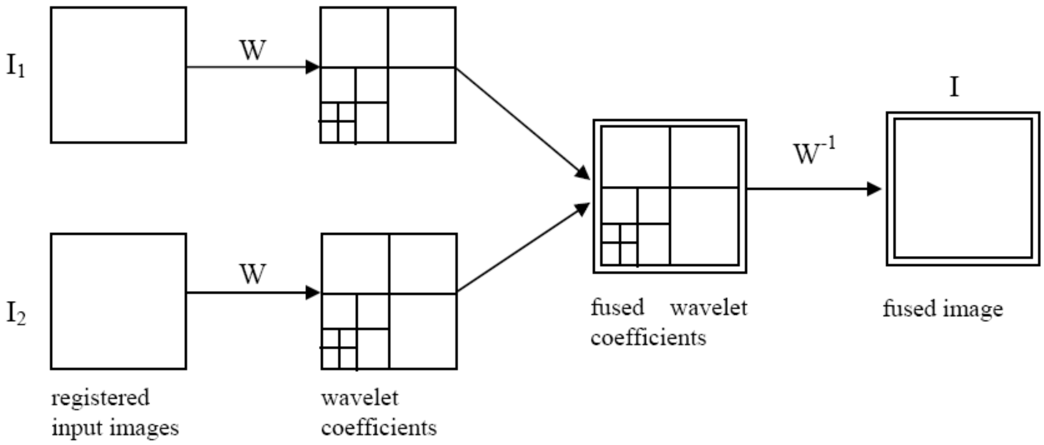

Figure 1 - Fusion of the Wavelet Transforms of Two Images
(From: Nikolov et al., 2002)

3.1.1 Wavelet Transform

Wavelet transformation is somewhat similar to the Fourier transform where raw signals can be broken down into frequency components to be represented in the frequency spectrum. What is missing in the Fourier transform is the temporal information when the signal is Fourier transformed to the frequency spectrum. Wavelet transformation evolved from the need to preserve the temporal information when representing the signal in the frequency spectrum. Wavelet transforms belong to the class of time-frequency transformations, which allow the user to know when specific frequency components occur in a signal. Different basis functions can be used for time-frequency representation. In wavelet transformations, two specific functions known as the mother wavelet and the scaling function are used to compute the wavelet transform. Two main possible wavelet transform approaches are available to transform a signal into wavelets for analysis, continuous wavelet transform (CWT) and discrete wavelet transform (DWT). The CWT is computationally intensive and more information about this approach can be found at Polikar (1995). In this work we focus on the DWT as this transform has been shown to be significantly more computationally efficient.

3.1.2 Discrete Wavelet Transform: Implementation to Image Fusion

Filters with different cutoff frequencies were used to analyze the signal at different scales in the DWT approach (Polikar, 1995). This operation is accomplished by passing the signal through a series of high pass and low pass filters and filtering out the low frequencies for the analysis of the high frequencies and the high frequencies for the analysis of the low frequencies. According to Nikolov et al., in one dimension the aim of the wavelet transform is to represent the signals as a linear combination of wavelets. Thus, using wavelet decomposition properties, a signal $f(t)$ may be expressed as

$$f(t) = \sum_{m,n} c_{m,n} \psi_{m,n}(t)$$

where, $\psi_{m,n}(t)$ is the dilated and/or translated version of the mother wavelet given by the equation

$$\psi_{m,n}(t) = 2^{-m/2}\psi[2^{-m}t - n]$$

where *m* and *n* are integers.

Every level of decomposition yields approximate coefficients (obtained from low-pass filtering) and detail coefficients (obtained from high-pass filtering). The approximate coefficient of a function at resolution level 2^m is $a_{m,n}$ and at resolution level 2^{m-1} is $a_{m-1,n}$. The approximate coefficient expression is given by

$$a_{m,n} = \sum_k h_{2n-k} a_{m-1,k} \ ,$$

where h_n is the low pass finite impulse response (FIR) and $a_{0,n}$ is the sampled signal. $c_{m,n}$ represents the detail coefficients and is given by

$$c_{m,n} = \sum_k g_{2n-k} a_{m-1,k} \ ,$$

where g_n is the high pass finite impulse response (FIR) and $a_{0,n}$ is the sampled signal.

The one-dimensional wavelet transform can be extended to two dimensions. In such a case, filtering and down-sampling operations need to be carried out in horizontal and vertical dimensions. The resultant is an arrangement of four subbands expressed as horizontal frequency first followed by the vertical frequency second. These subbands are high-high (HH), high-low (HL), low-high (LH) and low-low (LL) images. In this essence, the low-low subband represents the approximate coefficients. So by applying the same decomposition to the LL subband, a multi-resolution decomposition of the original image is feasible. The reverse is carried out for the reconstruction of the fused image, where up-sampling is applied instead of down-sampling and the complement of high pass and low pass FIR filters are used instead. Figure 2 illustrates the multi-resolution decomposition of an image.

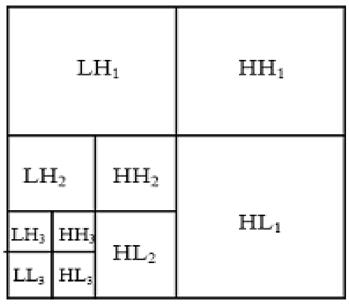

Figure 2 - Subbands Representation in a Multi-resolution Decomposition of an Image

Figure 3 and Figure 4 illustrate a decomposition tree and reconstruction tree, respectively for a thermal image.

The four subband images obtained after the decomposition phase relate closely to the original image. The approximate coefficients LLn+1 represent a coarse replica of the original without the details and the small variations. The coefficients LHn+1 represent the detail horizontal fluctuations from the original image and the coefficients HLn+1 represent the detail vertical fluctuations in the image. Finally, HH_{n+1} coefficients represent the diagonal fluctuations in the image which relate closely to the diagonal features in the image. The ability for subsequent approximate coefficients to be further broken down allows comparison of the two images features in a multi-resolution approach. The multi-resolution also allows the individual coefficients from the images to be combined based on a fusing scheme and to carry out an inverse DWT to recover a fused image. Note the wavelet function or the mother wavelet plays an important role in the decomposition step. Many wavelet families have been developed over the years with different properties.

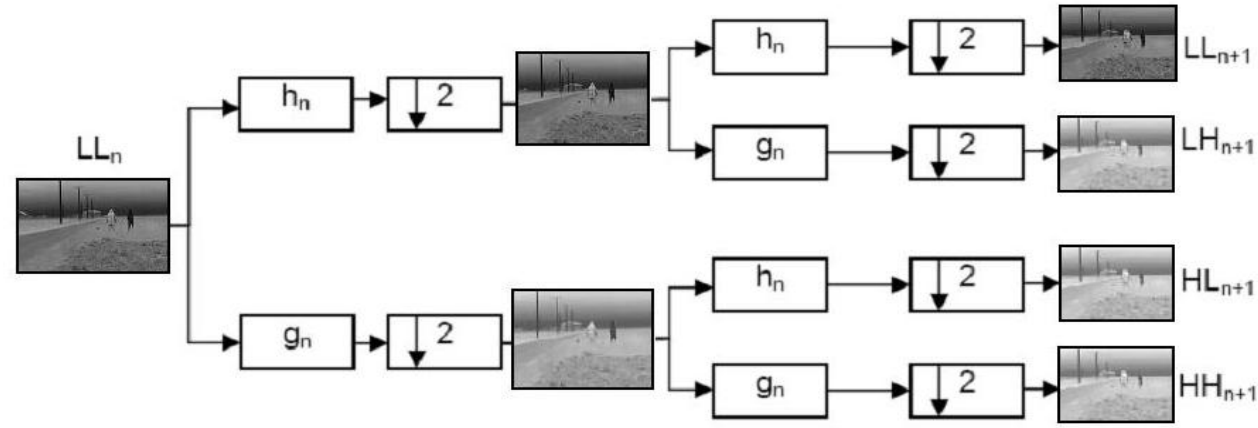

Figure 3 - Decomposition Tree of the Original Images into Four Subband Images. (After: Chow 2004)

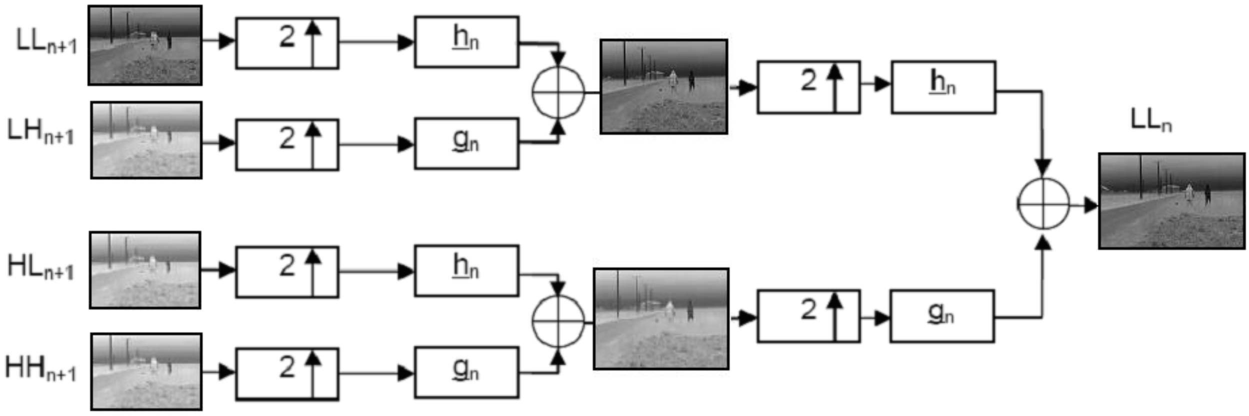

Figure 4 - Reconstruction Tree from 4 Subband Images Back to Recover the Original Image. (After: Chow 2004)

3.1.3 Wavelet Families

Wavelet decompositions are well suited to detect edges, which is especially important when the images considered have low contrast or the features in the images are not distinct. Even though pre-decomposition techniques such as image enhancement could be carried out to improve edge detection, it is important that the wavelet selected can perform to an acceptable standard even when there is prior image enhancement. A study on the comparison of several wavelet families conducted by Chaganti (2005) suggested that Haar wavelets (Haar) have better performance for edge detection applications than Daubechies (db) and Coifman wavelet (Coiflet/Coif) families when dealing with first level decompositions. Chaganti showed that Coiflet wavelets have better edge detection performances than db and Haar wavelet families when applying a second order level decomposition when noise is present in the image. This study also showed that the performance of Db and Coiflet are similar when dealing with third order level decompositions and the images under consideration are noisy, which is typical of night vision images. Chaganti also compared Biorthogonal (Bior), Haar wavelet decompositions, Db and Coiflet decompositions. He found out that Haar performance at first level is better than the rest of the wavelet families. However, Db and Coiflet perform better at higher level, with db more resistant to noise and Coiflets ability to pick up more details.

The focus of this paper is to select a few commonly used wavelet families as bases before expanding to others. Since Db and Coiflet performances over a series of images were shown to be more consistent, we chose to work on the variations of Db and Coiflet. In addition, we also included the symlet (Sym) family as it exhibits similar symmetry properties as Coiflets. A similar experiment as conducted in Chaganti’s study was applied to the selected wavelet families to test their suitability in the application considered. By observing the wavelet decomposition images, we were able to examine the edge detection ability of the wavelet families. Figure 5 shows the wavelet families mentioned.

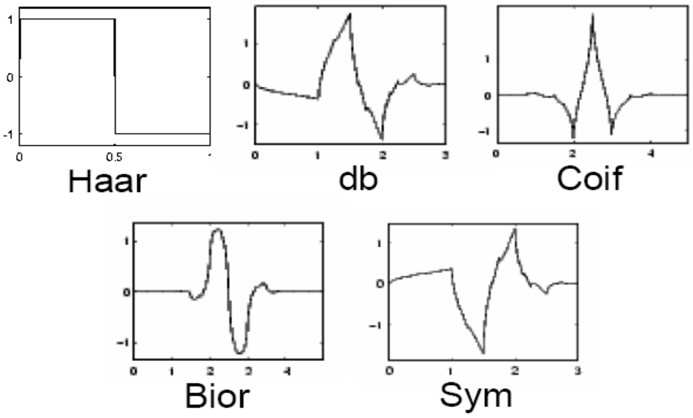

Figure 5 - Wavelet Families (From: Mathworks, 2006)

3.1.4 Image Fusion

Fusion of images can be carried out at different levels of decomposition, depending on the type of methods and schemes adopted. The three levels of information representations are pixel, feature and decision level. Pixel level fusion takes place by comparing individual pixels that are spatially identical in two separate images and fusing them using a fusion scheme. This method assumes that the images are registered prior to fusion, and image resampling and image registration operations are needed prior to fusion when this is not the case.

a. Image Resampling

Images to be fused must have the same size in order for wavelet transform fusion to take place, and images have to be resampled using an interpolation technique when there have different sizes.

b. Image Registration

For a proper fusion of two separate images, both images have to be accurately aligned and registered prior to merging the images. This property is especially important when the images include numerous edges. For this work, the images are assumed to be registered and the image resampling and registration will not be discussed further. Feature and decision level fusion look at larger areas within the image for decision on how the images would be fused. The wavelet transform fusion method is a form of pixel level fusion if surrounding pixels are not considered in the fusion scheme. Wavelet transform fusion may evolve into feature level or decision level fusion when further processing of the image is done, or extraction of information pertinent to the image is done at its decomposition level. Processing and information extraction such as segmentation into regions, characterization into shape, size, contrast texture and intensity from the image means that the fusion method logically develops into the feature level. The fusion method is categorized as decision level fusion when other inputs such as the amount in which individual images contribute to the fused image are taken into account. Extending from the pixel level of wavelet transform fusion is the region-based fusion which incorporates region segmentation using the watershed transform and region based fusion schemes. The region-based fusion approach can be categorized as a feature-level fusion scheme. For the wavelet transform fusion step to be complete, fusion schemes have to be applied before an inverse wavelet transform is carried out. The fusion scheme decides how the coefficients at the decomposition level are to be combined. Fusion schemes can range from a simple summation of coefficients to more complicated methods of sampling the surrounding pixels of the pixel of interest and determining its contribution amount to the fused image. Next is described the fusion schemes considered in this study.

3.1.5 Fusion Schemes

The three fusion schemes described in this section is inspired from the earlier work of Chow (2004) which selected approaches commonly used in the image fusion community. The general framework is shown in Figure 6 using registered images.

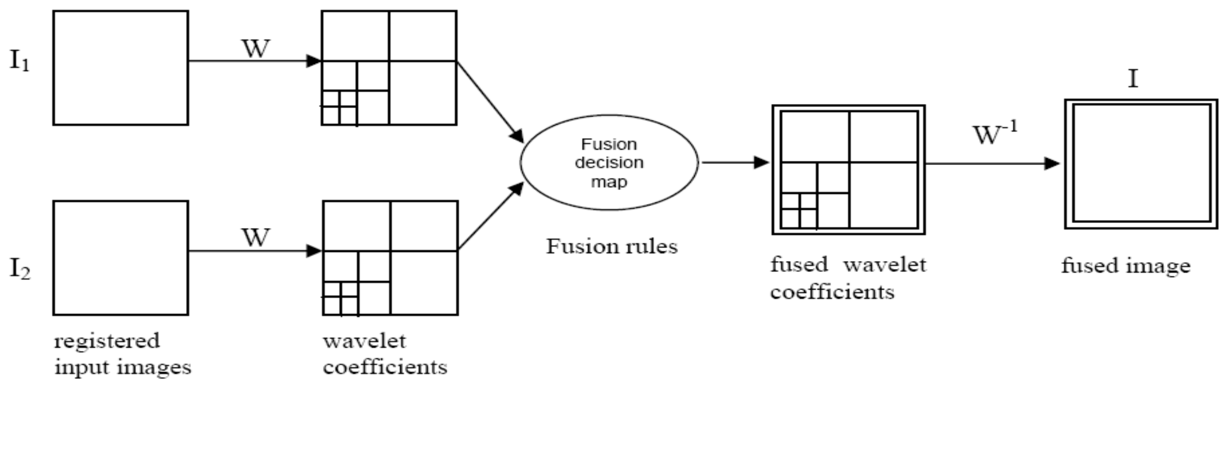

Figure 6 - General framework for wavelet transforms fusion. (After Chow, 2004)

Details coefficients obtained after the wavelet decomposition step represent the image salient features. Next, detail coefficients are compared one pixel at a time to determine their level of contribution to the resulting fused image they are. A fusion rule governs the decision and determines how the resultant detail coefficients highlight the salient features in the fused image. This step is known as the fusion decision map. Approximate coefficients are processed differently. Approximate coefficients obtained from the two images are averaged together to form the new approximate coefficients to the fused image. Figure 7 illustrates the framework for the formation of the fusion decision map.

Pixel activity level and its salience in the image are measured using a very simple method. A pixel is said to be salient when its wavelet coefficient associated to that location has high value. Thus, the saliency of the coefficients is determined by using the coefficients absolute values as:

$$a_{A_n}(j,k) = \left| c_{A_n}(j,k) \right| \text{ şi } a_{B_n}(j,k) = \left| c_{B_n}(j,k) \right|,$$

where the parameters $c_{A_n}(j,k)$ and $c_{B_n}(j,k)$ represent the n[th] level wavelet coefficients at location (j,k) of input image A and B, respectively (Chow, December 2004). If we expand the area from the pixel of interest to surrounding pixels, one may have a more sophisticated possibility of getting a salient feature from the image. The method known as window-based activity measure (Chow, 2004), allows a selective window size for the operation. A typical 3-by-3 or 5-by-5 window size is chosen because even though having a bigger window size means a more robust fusion system, it also results in heavier computational load. The activity associated with the n[th] level pixel centered in the window at location (j,k) using a window-based approach is determined using

$$a_{W(A,B_n)}(j,k) = \sum_{s \in S, t \in T} c_{A,B_n}(j+s, k+t),$$

where the variable $a_{W(A,B_n)}(j,k)$ is the window-based activity measure and the parameter *S* and *T* are sets of horizontal and vertical indexes that describe the current window (Chow, 2004). Next, a matching criterion is used to have a higher degree of accuracy when comparing the two images. If the two pixels have the same intensity or are closely related to each other, this information is used to decide how the pixel is treated. The matching criteria $m_{F_n}(j,k)$ defined in the range 0 to 1 represents the correlation between the corresponding pixels at location (j,k) for the n[th] level coefficients and is defined as (Chow, 2004)

$$m_{F_n}(j,k) = \frac{2c_{A_n}(j,k)c_{B_n}(j,k)}{\left| c_{A_n}(j,k) \right|^2 + \left| c_{B_n}(j,k) \right|^2}$$

Next, fusion rules are developed to allow the decision on the type of treatment to the individual pixels to be carried out. Three fusion rules adopted again in this work: two pixel-based and one window-based scheme. Further details regarding these schemes may be found in Chow (2004).

a. Fusion Rule 1 – Selection of Dominant Mode

For each pixel in the image, this fusion rule selects the most dominant feature of the two images pixels (the pixel with the highest activity level). Thus, the fusion rule is defined as:

$$c_{F_n}(j,k) = \begin{cases} c_{A_n}(j,k)\_if\ \_\left| a_{A_n}(j,k) \right| > \left| a_{B_n}(j,k) \right| \\ c_{B_n}(j,k)\_if\ \_\left| a_{B_n}(j,k) \right| > \left| a_{A_n}(j,k) \right| \\ \dfrac{c_{A_n}(j,k) + c_{B_n}(j,k)}{2}\_if\ \_\left| a_{A_n}(j,k) \right| = \left| a_{B_n}(j,k) \right| \end{cases},$$

where the variable $c_{F_n}(j,k)$ is the composite coefficient from the two images detail coefficients, $a_{A_n}(j,k)$ and $a_{B_n}(j,k)$ are the absolute values of the coefficients of image A and image B respectively.

b. Fusion Rule 2 – Weighted Average of Modes

This approach is based on a weighted combination of the source images. The matching measure defined previously is used to determine the respective contribution by both images and is given by

$$c_{F_n}(j,k) = \begin{cases} wc_{A_n}(j,k) + (1-w)c_{B_n}(j,k)\_if\ \_\left| a_{A_n}(j,k) \right| > \left| a_{B_n}(j,k) \right| \,\&\, m_{F_n}(j,k) \le T \\ wc_{B_n}(j,k) + (1-w)c_{A_n}(j,k)\_if\ \_\left| a_{B_n}(j,k) \right| > \left| a_{A_n}(j,k) \right| \,\&\, m_{F_n}(j,k) \le T \\ \dfrac{c_{A_n}(j,k) + c_{B_n}(j,k)}{2}\_if\ \_m_{F_n}(j,k) \ge T \end{cases},$$

where *w* is the weighted value that defines the contribution of the selected coefficient, $m_{F_n}(j,k)$ is the matching value and *T* is the pre-defined threshold. A higher weight is given to the pixel with a higher activity level and *T* is set at 0.8, which means an 80% match for the pixel of both images. At the present moment, the threshold value *T* was set at 0.8 by trial and error for the application considered however it could be adjusted if further analysis showed a more suitable value exists.

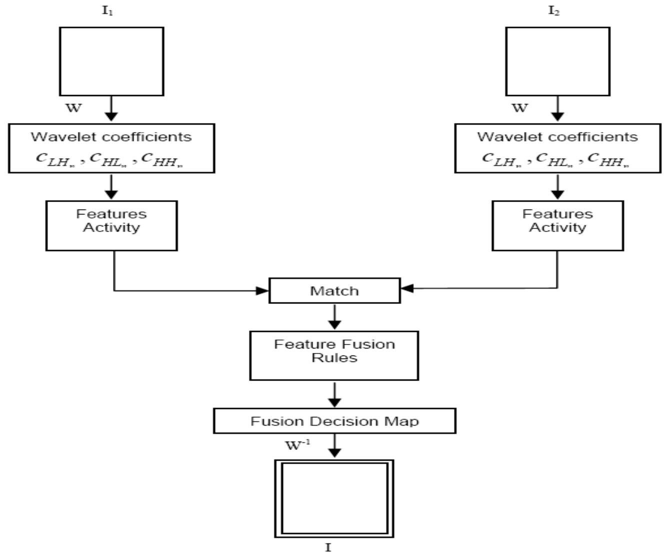

Figure 7 - Framework for the Formation of the Fusion Decision Map.
(After Chow, 2004)

c. Fusion Rule 3 – Weighted Average of Window-based Modes

Expanding on Rule 2, neighboring pixels are taken into account when determining the weight of each corresponding pixel from both images. From the window-based activity measure defined previously, the measure is incorporated into the fusion rule as

$$c_{F_n}(j,k)=\begin{cases} wc_{A_n}(j,k)+(1-w)c_{B_n}(c,j)\_if\ \_\left|a_{W(A_n)}(j,k)\right|>\left|a_{W(B_n)}(j,k)\right|\ \&\ m_{F_n}(j,k)\leq T \\ wc_{B_n}(j,k)+(1-w)c_{A_n}(c,j)\_if\ \_\left|a_{W(B_n)}(j,k)\right|>\left|a_{W(A_n)}(j,k)\right|\ \&\ m_{F_n}(j,k)\leq T \\ \dfrac{c_{A_n}(j,k)+c_{B_n}(j,k)}{2}\_if\ \_m_{F_n}(j,k)\geq T \end{cases},$$

where the variables $a_{w(A_n)}(j,k)$ and $a_{w(B_n)}(j,k)$ are the window-based activity level measures for image A and B.

3.2 REGION- BASED FUSION

An image can usually be segmented according to the objects and features combined within, which also applies to low contrast images such as night vision and thermal images. The application of segmentation is easier on night vision and thermal images as the high intensity regions define either area of light sources or heat sources. The composition of these images will be explored in the next chapter. The region-based fusion may be considered as a more robust method and is able to better interpret the feature intensity in the images than pixel-based schemes are. By including the feature intensity in the decision of the fusion, important features may be better preserved and integrated in the fused image. The purpose of the region fusion is to optimally extract the information from different sources and maximize the “scene content” in the fused image.

3.2.1 Region Segmentation

As mentioned previously, an image may be segmented according to its objects and features. A region-based segmentation scheme is implemented on objects and features by considering the level of homogeneity in these regions, their strong statistical correlation and visual similarities.

3.2.2 Watershed Transform

The basic idea behind the watershed transform can be explained by drawing a parallel with natural features such as ridges and valleys that make up the geographical landscape of the earth. So an image may be segmented into high and low intensity regions by viewing it as a topographical map, where high intensity regions correspond to peaks and low intensity areas as valleys. The ridges forms watershed lines that channel water to form pools and reservoirs known as catchment basins at low lying regions. If this water has nowhere to flow, the water level will continue to rise until it overflows to another region possibility out of the catchment basin. The watershed transformation is a morphological transformation used to segment an image into regions of interest. As previously described if water starts to flood from the watershed into the valley, the water will rise uniformly throughout the low regions. If a dam is erected between the two regions where the flooded regions meet and subsequently for all regions, the image is partitioned according to different regions of regional minima. Once the catchment basin is created, the gradient image is formed using “Dilation” and “Erosion” morphological operations. Note that a direct application of the watershed transform to a gradient image usually produces excessive segmentation. Thus, the so called “maker-based segmentation” approach is usually preferred, which limits the over-segmentation problem encountered in the basic watershed transform approach. The watershed algorithm is applied to the approximate coefficients of the wavelet transformed image rather than to the grayscale image to significantly reduce over-segmentation and remove broken contours, as originally proposed by Jung et. al. (2002). Chow found that generally a good segmented image only needs about forty or fewer regions after numerous experiments. This constraint also helped to ensure that the processing times for subsequent stages were curbed. Figure 8 and 9 illustrate the region segmentation process obtained using the approximate coefficients at various level of decomposition.

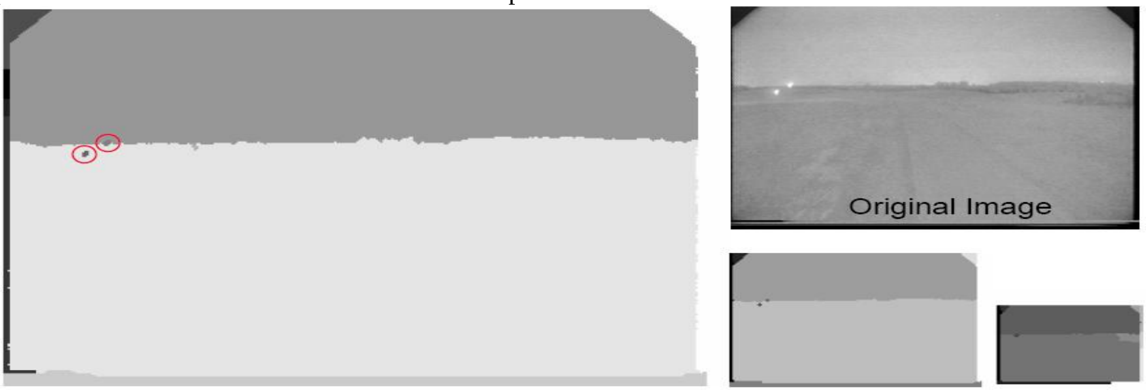

Figure 8 - Region Segmentation of the Approximate Night Vision Intensified Image at Three Levels of Decomposition: Level 1 - left, Level 2 - middle and Level 3 - right (After: Chow, 2004)

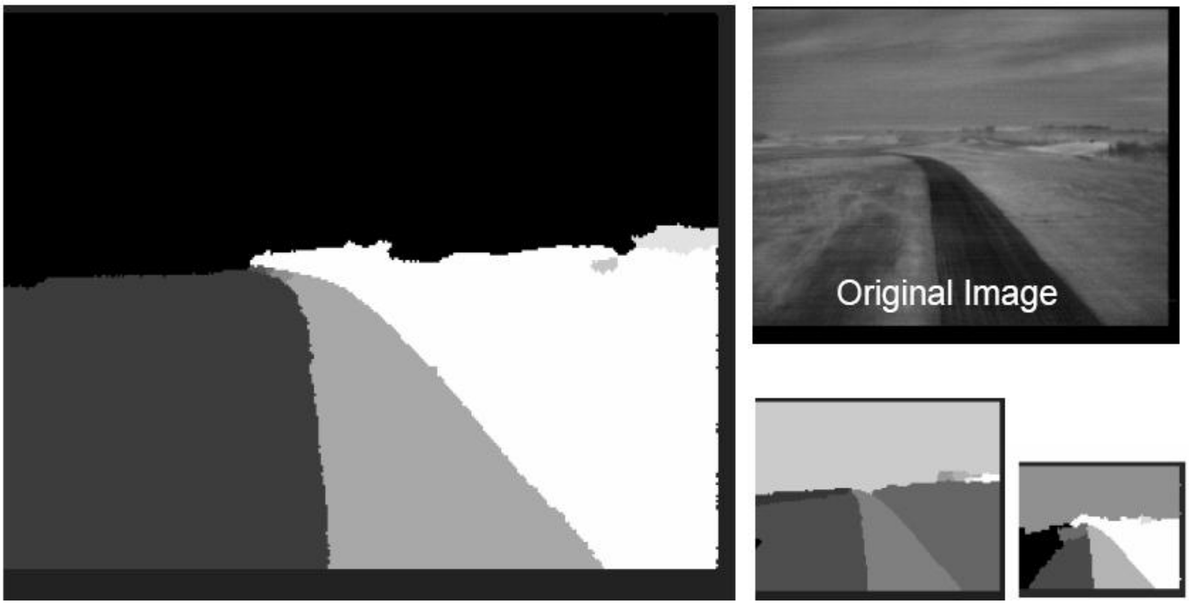

Figure 9 - Region Segmentation of the Approximate Thermal IR Image at Three Levels of Decomposition: Level 1 - left, Level 2 - middle and Level 3 - right (After: Chow, 2004)

3.2.3 Fusion Methodology

After the images to be fused are segmented, regions unique to each image need to be merged at each level of decomposition to direct the fusion of the wavelet coefficients at each level. The fusion methodology is an extension of the wavelet transform framework with an addition of region activity and expands the “Feature Fusion Rules” to “Region and Feature Fusion Rules”. As the name implies, the region information is included in the fusion decision map. The wavelet transform framework with the added features is shown in Figure 11.

a. Segmentation Process and Feature Activity Information

Each image generates a region activity map which comprised the segmentation of regions in the image. Two region activity maps are combined to form the joint region activity map. The joint region activity map is used in the fusion process together with the feature activity information to generate the fusion map. The feature activity information is explained later in this section. Figure 10 shows the construction of the joint region map.

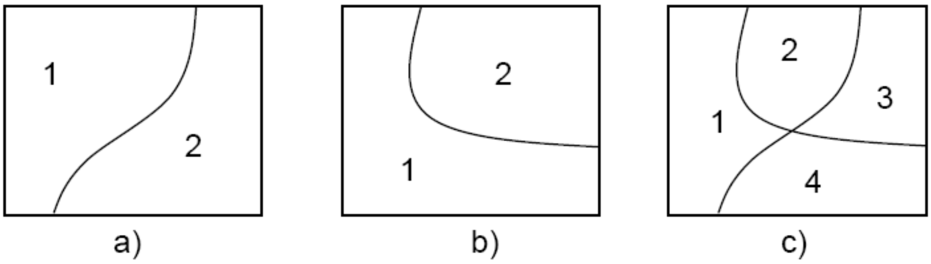

Figure 10 - Region Segmentation: a) Region Representation of Image A; b) Region Representation of Image B and c) Joint Region Map, Indicating the Four Identified Regions. (From: Chow 2004)

The fusion rules are crucial in deciding how the regions in the images are treated. Using the joint region map, the regions are labeled using the following notation

$$R = \{R_n^k\},$$

where the variable R_n^k represents the k segmentation at level n. Following the definition of the joint region map, the sizes of the regions are determined. By overlaying the joint map over the original images, one is able to see the region segmentation superimposed on to the images. The activity levels in each of the regions for image A, known as the feature activity information are computed using

$$A_{A_n} = \frac{1}{S_i}\sum_{S_i} a_{A_n}(j,k),$$

where $a_{A_n}(j,k)$ is the absolute value of the detail coefficient and represent the n^{th} level activity measure at location (j,k) and S_i is the size of the region determined.

b. Fusion Rule

Once the computations of the activity levels in each of the images are completed, the fusion can be carried out using specific fusion rules for region-based fusion. The fusion rule implemented here is a simple weighted average for the detail coefficients using

$$c_{F_n}(j,k) = \begin{cases} wc_{A_n}(j,k) + (1-w)c_{B_n}(j,k)\_if\_\left|A_{A_n}(x)\right| > T \\ wc_{B_n}(j,k) + (1-w)c_{A_n}(j,k)\_if\_\left|A_{B_n}(x)\right| > T \\ \dfrac{c_{A_n}(j,k) + c_{B_n}(j,k)}{2} otherwise \end{cases}$$

where the variable $c_{F_n}(j,k)$ is the composite coefficients from the two images detail coefficients, w is the weighted value that defines the contribution of the selected coefficient, $A_{A_n}(x)$ and $A_{B_n}(x)$ are the activity levels in the region which the coefficients lies in image A and B, and T is the pre-defined threshold. The fusion rule selects the coefficients according to the weight assigned after considering the activity level in the region where the coefficients exist. If however, both coefficients exhibit similar activity levels, the average is taken and no weights are assigned. The ideal weight factor after experimenting with the fusion rule is set at 0.8. The threshold is similarly set at 0.8 to define a region of high activity. Chow noted in his study that this fusion rule retains the most important features to extract from both night vision and thermal images, which should help in better orientation and improved situational awareness. However, our results showed that the region-based image fusion does not necessary produces the best fused image. The segmentation in the region-based fusion may make the image look rather patchy, which may not appeal to most observers. On the other hand, the observer may perceive the image as being easier to interpret when the fused image maintains a certain level of uniformity.

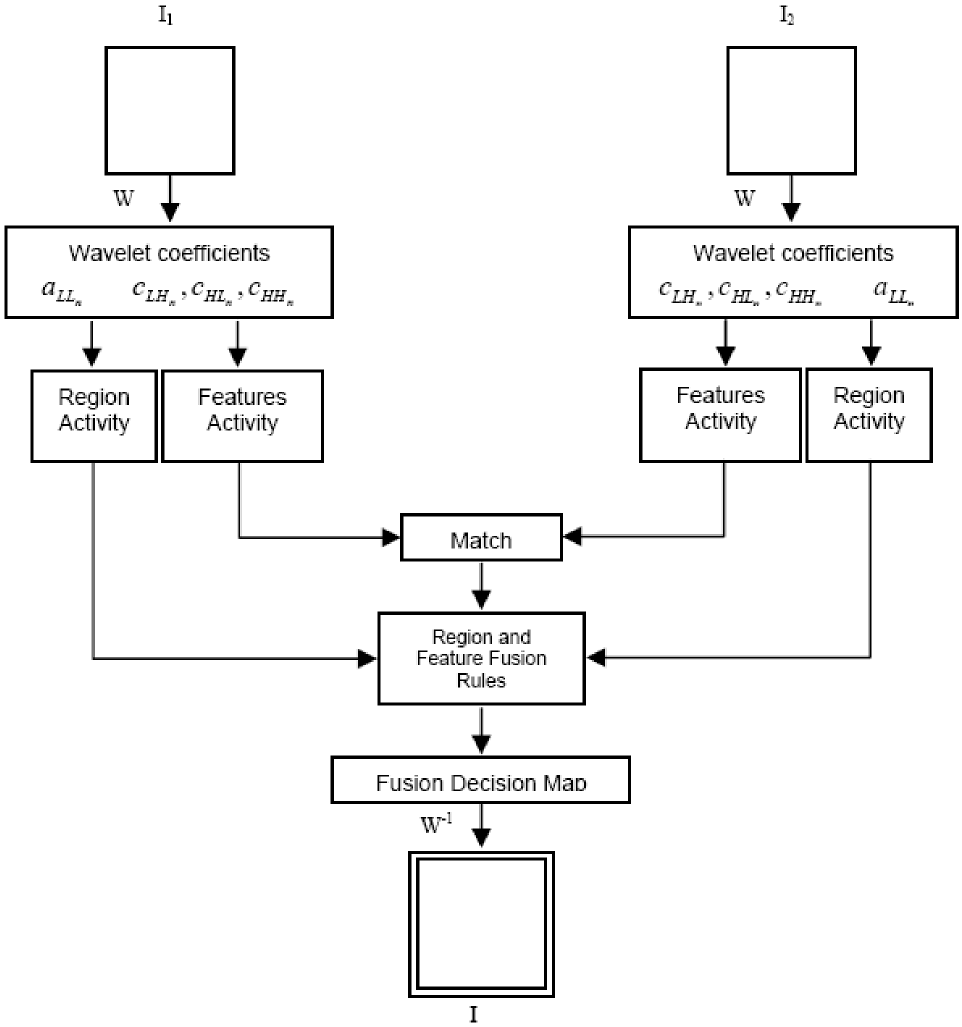

Figure 11 - Framework for the Formation of the Fusion Decision Map for Region-based Fusion. (After: Chow, 2004)

4. Further work

In the future we try to develop a MATHLAB algorithm for fusion the intensifier night vision images and thermal images. Our initial results will be validated by human subject tests mean to verify if the image fusion application has the ability to select the best fusion scheme based on the selected image quality evaluation. Furthermore, additional images should be collected and used in such a testing phase to achieve an accurate evaluation of the fusion application performances.

References

1. *"Fusion Night Vision and Thermal Images"* – Thesis of Tiong Tien Neo, Naval Pstgraduate School, Monterey, California, december 2006.
2. *"Fusion of Images from Dissimilar Sensor Systems"* – Chow K.C., Naval Pstgraduate School, Monterey, California, december 2004.
3. *"Digital Enhancement of Night Vision and Thermal Images"* – Teo C.K., Naval Pstgraduate School, Monterey, California, december 2003.
4. *"Multi-Sensor Image Fusion using the Wavelet Transform*" - Li H., Manjunath B.S. & Mitra S.K., IEEE Computer Society Press, Los Alamitos, C.A., 1994.
5. *"Night Vision Technology"* – internet, www.military.com, nov. 2006.
6. *"Robust Watershed Segmentation using the Wavelet Transform"* – Jung C.R., Scharcanscki J., Brazilian Symposium on Computer graphics and Image Processing, 2002.