Evaluation of Spatial and Transform Fusion methods for Medical Images using Normalized Non-Reference Quality Metrics

R. Barani
Department of Computer Science,
V.V.Vanniaperumal College for Women,
Virudhunagar, 626001, INDIA.

M. Sumathi
Department of Computer Science,
Sri Meenakshi Govt. Arts College for Women,
Madurai, 625002, INDIA

ABSTRACT

Medical image fusion is a method which enhances the image content by combining the images obtained using different imaging modalities like Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT). The main objective of medical image fusion is to extract and merge the useful information from multi-modality medical images thus highlighting the significant features for improved prediction of the scenario for treatment planning. In this paper, the different image fusion techniques in spatial and transform domain are implemented for MRI/CT and PET/CT images. The resultant fused images are analyzed with non-reference image quality metrics: Entropy (EN), Standard Deviation (SD), Peak Signal to Noise Ratio (PSNR), Spatial Frequency (SF), Average Gradient (AG), Edge Strength (ES), Fusion Factor (FF) and Fusion Symmetry (FS). It is found that the image fusion using Discrete Wavelet Transform (DWT) outperforms all the other spatial and pyramid based fusion methods.

General Terms

Image fusion, quality metrics, medical imaging.

Keywords

medical image fusion, spatial fusion, transforms fusion, non-reference quality metrics.

1. INTRODUCTION

Image fusion is an interesting, up growing and an important technique that combines the essential features from two or more images into a single image with extended information content thus reducing redundant data and artifacts that may exist in the source images. It has its significant importance due to its application in medical science, remote sensing, machine vision, forensic and defense departments etc., The image fusion process is performed for multi-sensor, multi-modality, multi-temporal and multi-focus images [1].

Image fusion can be performed in two different domains of an image namely Spatial Domain and Transform Domain. In Spatial Domain, some sort of mathematical calculations are performed on the intensity value of the pixels in the source images to form the pixel values of the fused image [2]. Whereas in the Transform Domain, the source images are converted into multi-resolution or multi-scale image representation, before the fusion process. Then the fusion operations are performed on multi-resolution or multi-scale representation of an image. The resultant fused image is obtained by taking the inverse transformation on the fused multi-resolution or multi-scale representation. Transform fusion uses pyramid transform or wavelets transform to convert the source image into multi-scale or multi-resolution representation [3].

Image fusion is generally performed at three different levels of information representation of an image: pixel level, feature level and decision level [4]. A pixel level image fusion performs the arithmetic operations (addition / subtraction / average) or selection operations (minimum / maximum) the raw pixel intensity values of an image. Still, these methods smoothens the sharp edges or leave the blurring effects in the fused image [5]. The feature level image fusion on the other hand operates on the various characteristics (features) of an image such as shape, edge, contrast, color, texture etc., These features are extracted first from the source images and then fused using the appropriate fusion rule. Researchers have shown that fusion at feature level give up better fused image than pixel based techniques. But these methods hardly depend on the effectiveness of the algorithms used to segment the features [6]. The highest level of fusion is decision level fusion which deals with the symbolic representation of images and the fusion is performed at the symbol level.

Nowadays, medical imaging becomes an increasingly significant role in human healthcare. Advances in scanning technology, have enabled the radiologists to acquire images of the human body and its internal structures quickly with greater accuracy [7]. In general, medical imaging system is divided into structural and functional imaging systems. Structural imaging shows detailed anatomical informative images inside the body, including bones, organs, tissues and implants with less distortion, but it cannot detect physiological changes. For example, Computed Tomography (CT) uses multiple X-rays to produce cross-sectional layers of hard tissues, whereas Magnetic Resonance imaging (MRI) uses magnetic fields and radio waves to show detailed images of organs, normal and pathological soft tissues, bones, ligaments and cartilage. Functional imaging is another kind of medical imaging technique which focuses on detecting or measuring changes in metabolism, blood flow, regional chemical composition, and absorption. Functional imaging [8] such as Positron Emission Tomography (PET) and Single-Photon Emission Computed Tomography (SPECT) involves injecting, inhaling or swallowing a radioactive 'tracer'. The gamma-rays emitted by this material are used by the scanner to show images of bones and organs and also functional information such as blood flow, blood activity and metabolic activity with low spatial resolution [9]. A single modality cannot provide both anatomical and functional information in a single image. But, for most of the clinical applications, combining these images is necessary.

Hence, Medical Image Fusion plays an important role for clinical applications such as image-guided surgery, image-guided radiotherapy, noninvasive diagnosis, and treatment planning. In all such clinical applications, the complementary information from the images of different modalities is required. For example, combined PET/CT or PET/MRI or SPECT/MRI imaging can concurrently visualize both anatomical and physiological characteristics of the human body. In oncology, this concurrent visualization helps physician to view tumor activity in conjunction with anatomical references to localize and identify the tumor and also to better understand the effects of cancer treatment [10]. So, the major motivation of medical image fusion is to enrich the information content and to improves the medical images by combining two or more anatomical and functional images into a single image that will help the physician during medical diagnosis and treatment planning.

This paper implements the following fusion methods in the spatial domain such as maximum selection (MAX), simple average (AVG), weighted average (WAVG) and Principle Component Analysis (PCA) This paper also implements pyramidal fusion methods that uses Laplacian Pyramid, Gradient Pyramid, Contrast Pyramid, Ratio-Of-Low-Pass (ROLP) Pyramid, Filter Subtract Decimate Pyramid (FSD) and Morphological Pyramid and the fusion methods based on discrete wavelet transform in the transform domain. All these fusion methods are verified for various combinations of analytical and functional medical imaging modalities such as MRI/CT and PET/CT. The resultant images are quantitatively analyzed using the normalized non-reference image quality metrics Entropy (EN), Standard Deviation (SD), Peak Signal to Noise Ratio (PSNR), Spatial Frequency (SF), Average Gradient (AG), Edge Strength (ES), Fusion Factor (FF) and Fusion Symmetry (FS).

2. IMAGE FUSION IN SPATIAL DOMAIN

Spatial image fusion methods perform certain mathematical calculations on the pixel intensity values of source images to create the fused image. Spatial domain fusion methods are classified into primitive methods and sub-space methods [2].

2.1 Primitive Fusion Methods

The Primitive fusion methods perform the fusion process exactly on the source images. These methods perform simple arithmetic operations such as addition, subtraction, division, and multiplication as well as minimum, maximum, median, rank, average and weighted average on the pixel values of input images [11]. It also includes more complicated operations such as Expectation-Maximization, Markov Random Field etc. Unfortunately these methods produce unsatisfactory results such as features that appear in one source image but not in others are rendered in the composite at reduced contrast or superimposed on features from other images. But these methods are good for certain cases where the input images have an overall high brightness and high contrast [12].

2.2 Sub Space Methods

The sub space methods are a collection of statistical techniques which remove the correlation that exists between the input images IK, where, k € {1,2,……..k}. Some important sub space techniques [Mitchell, 2010] are Principal Component Analysis (PCA), Independent Component Analysis (ICA), Non-negative Matrix Factorization (NMF), Canonical Correlation Analysis (CCA), Linear Discriminant

Analysis (LDA) and Nearest Neighbor Discriminant Analysis (NNDA). The sub space methods are used in the applications where computational load and the storage requirements are very important and crucial [2]. One of the major disadvantages of spatial domain fusion methods is that it introduces spatial distortions in the resultant fused image and does not provide any spectral information. These disadvantages were overcome with the use of transform domain image fusion methods [9].

3. IMAGE FUSION IN TRANSFORM DOMAIN

In the transform domain, the image fusion is performed at the multi-resolution or multi-scale image representation of the source images. Normally, the human visual system processes and analyzes image information at different scales. Based on this idea, Burt proposed an efficient method for image fusion that uses multi-scale or multi-resolution representation of an image [13].

A multiresolution analysis [14] decomposes an image into a smoothed version of the source image and a set of detail information at different scales. First, the detail information that distinguishes the sharp edges in the image is removed. The resultant image obtained is slightly blurred. This blurred version of the original image is rendered at a slightly coarser scale. Then the same procedure is repeated recursively. Each time, some detail information and a more and more blurred (or smoothed) version of the original image is obtained. In this process, removal of the detail information corresponds to a bandpass or highpass filtering, and generation of the smoothed image corresponds to a lowpass filtering. Given the decomposition of an image, one can reconstruct the original image.

Transform domain implements a pattern selective approach to image fusion [1,15], so that the composite image is constructed not a pixel at a time, but a feature at a time. Primary features that are identified from each of the source image are copied as a whole to the composite image. Unimportant features that may partially mask the more salient features are discarded. Transform fusion uses pyramid transform or wavelets transform to convert the source image into multi-scale or multi-resolution representation. Image fusion in the transform domain is carried out in four steps [16].

Step 1:	*Convert both source images into multi-scale or multi-resolution representation.*
Step 2:	*Compute match and saliency measures for the source images at each position in the transform representation.*
Step 3:	*Combine the source representations using the combination rule to generate the combined representation.*
Step 4:	*Obtain the fused image from the combined representation using inverse transform.*

Now, consider *A* and *B* are the two source images and *C* be the composite image. D_A, D_B and D_C are the transform representation of the source images. Let in general, $D_I(mnkl)$ represents the transform representation of source image *I, mn* represents sample position, *k* denotes the level of decomposition and *l* specifies the orientation in the transform representation. Simply $D_I(mnkl)$ is denoted as $\vec{m}$.

3.1 Match and Salience Measure

Most of the pattern selective image fusion is guided by two measures: a *match measure* which determines the mode of

combination at each sample position $\vec{m}$ (selective or averaging), and *salience measure* that determines which source pattern is chosen in the selection mode.

3.1.1 Salience Measure

The salience of a particular component pattern is high if that pattern plays a major role in representing important information in the fused image. There can be various measures of salience such as: amplitude, contrast, correlation, edge intensity etc. The choice of salience mainly depends on the application area and the nature of the source images. In this paper, the measure of salience at each sample point $\vec{m}$ is considered to be the local energy or variance within neighborhood p and is defined as follows.

$$S_I(m,n,k,l,) = \sum_{\acute{m},\acute{n}} p(\acute{m},\acute{n}) D_I(m+\acute{m}, n+\acute{n}, k, l)^2 \quad (1)$$

Here, the neighborhood p is small, which includes only sample point itself (point case) or a *3x3 or 5x5* windows of sample points (area case).

3.1.2 Match Measure

The match measure is used to determine which of the two combination modes, selection or averaging have to be applied at each sample position. This measure can be the relative amplitudes of corresponding patterns in the two source amplitudes. Alternatively, here the match measure at sample point $\vec{m}$ is defined as the local normalized correlation within neighborhood p:

$$M_{AB}(\vec{m}) = \frac{2 \sum_{\acute{m}\acute{n}} p(\acute{m}\acute{n}) D_A(\vec{m}+\acute{m}\acute{n}) D_B(\vec{m}+\acute{m}\acute{n})}{S_A(\vec{m}) + S_B(\vec{m})} \quad (2)$$

Again, the neighborhood p includes the sample point (point case) or a local array of sample points (area case). The value of MAB is 1 for identical patterns, -1 for identical patterns but has opposite signs, and value between -1 and 1 for all other patterns.

3.1.3 Combination Rule

The combination states that how the components patterns of source images are combined together to form the fused image representation. Commonly used combination rule is weighted average and is given by:

$$D_C(\vec{m}) = w_A(\vec{m}) D_A(\vec{m}) + w_B(\vec{m}) D_B(\vec{m}) \quad (3)$$

Where w_A and w_B are the weights assigned to the source images *A* and *B* and depends on the match and saliency measures M_{AB}. At each sample position $\vec{m}$,

$$\left.\begin{array}{l} if \ M_{AB} < \alpha, then \ W_{min} = 0 \ and \ W_{max} = 1, else \\ if \ M_{AB} \geq \alpha, then \ W_{min} = \frac{1}{2} - \frac{1}{2}\left(\frac{1 - M_{AB}}{1 - \alpha}\right) \ and \\ \qquad\qquad max = 1 - W_{min} \end{array}\right\} \quad (4)$$

Where α is the threshold value. The largest weight w_{max} is assigned to the source image with larger salience.

$$\left.\begin{array}{r} if \ S_A > S_B, then \ W_A = W_{max} \ and \ W_B = W_{min} \\ else \ W_A = W_{min} \ and \ W_B = W_{max} \end{array}\right\} \quad (5)$$

Here, the parameters the level of transform *l*, the threshold vale α and the size of the neighborhood *p* may affect the fusion process and can be optimized for better fusion results.

3.1.4 Consistency Check

Consistency verification is required to ensure that a sample point in the fused representation does not come from a different source image from its neighbors i.e. a sample point under verification is from image A whereas most of its neighboring sample points come from image B. Then the sample point under verification is changed to the sample point of image B. Usually this verification is performed with choose max selection rule. Commonly used consistency verification is window-based consistency verification [1]. It uses 3x3 or 5x5 window for considering the neighbors. Li uses a majority filter to implement window based consistency check and is represented in Fig. 1.

Fig 1: 3x3 window-based consistency verification

3.2 Pyramid Transform

The image pyramid is a data structure designed to support efficient scaled image analysis through reduced image representation [17]. It consists of a sequence of low pass or band pass copies of an original image in which both sample density and resolution of an image are decreased in regular steps. An image pyramid does contain all the information needed to reconstruct the original image. The pyramid representations considered here are

3.2.1 Gaussian Pyramid

The Gaussian pyramid [13] consists of low-pass filtered, down sampled images of the previous level in the pyramid with base level is the original image. Let G_k be the kth level of the Gaussian pyramid for image *I*. Then $G_0(i,j) \equiv I(i,j)$ and for $k>0$,

$$G_K \equiv [w * G_{k-1}]_{\downarrow 2} \quad (6)$$

Here w is the weighting function or generating kernel, and the notation $[...]_{\downarrow 2}$ indicates that the image array in brackets is sub sampled by *2*. For example, $w(m,n)$ is the 5-tap filter 1/16 [1 4 6 4 1]. Each element in Gaussian pyramid represents a local average at various scales.

3.2.2 Laplacian Pyramid

The Laplacian pyramid $\tilde{L}_k$ [18,19] is defined as the difference between successive levels of the Gaussian pyramid and is given by

$$\tilde{L}_k = G_k - 4w * [G_{k+1}]_{\uparrow 2} \quad (7)$$

Here $[...]_{\uparrow 2}$ indicates up sampling by *2*. n-1 rows and columns of zero vale are inserted between the original rows and columns. Convolution by *w* has the effect of interpolating the missing samples.

3.2.3 Filter Subtract Decimate (FSD) Pyramid

FSD Laplacian pyramid [20] is defined as the difference between G_k and the filtered copy of G_k prior to subsampling to form $G_{(k+1)}$. FSD pyramid is generated by using the following recursive rules:

$$\left.\begin{array}{ll} \tilde{G}_{n+1} = w * G_n & Low - pass \ Filter \\ L_n = G_n - \tilde{G}_{n+1} & Subtract \\ G_{n+1} = Decimate \ \tilde{G}_{n+1} & Decimate \end{array}\right\} \quad (8)$$

3.2.4 Gradient Pyramid

A gradient pyramid [21] for image *I* can be obtained by applying a gradient operator to each level of its Gaussian pyramid representation. The image can be represented by a set

of four gradient pyramids one for each derivatives in horizontal, vertical and the two diagonal directions. Let D_{kl} be the kth level and l^{th} orientation gradient pyramid for image I and is obtained from G_k through convolution with four gradient filters d_l.

$$D_{kl} = d_l * [G_k + w * G_k] \qquad (9)$$

3.2.5 Ratio-Of-Low-Pass (ROLP) Pyramid
The ratio-of-low-pass pyramid [22] is defined as the ratio of two successive layers of Gaussian pyramid and is given by

$$\left. \begin{array}{l} R_i = \dfrac{G_i}{\text{expand } [G_{i+1}]}, 0 \leq i \leq N-1 \\ R_N = G_N \end{array} \right\} \qquad (10)$$

ROLP preserves visually important details in the fused image.

3.2.6 Contrast Pyramid
Luminance Contrast [23] is defined as

$$C = (L/L_b) - I \qquad (11)$$

Where L denotes the luminance at a certain location in the image plane, L_b represents the luminance of the background and $I(i, j) = 1$ for all i and j. When the luminance contrast at level i, i.e. Ci is defined as

$$C_i = G_i/expand(G_{i+1}) - I \qquad (12)$$

We have obtained the sequence of contrast levels. Hence the sequence is sometimes known as contrast pyramid.

3.2.7 Morphological Pyramid
Morphological pyramids [24,5] systematically split the input signal into approximation and detail signals by repeatedly applying morphological filters (pyramidal analysis operators) followed by down sampling. The fundamental morphological operators are: *erosion, dilation, opening* and *closing.*

Erosion sets a pixel at (i, j) to the minimum over all pixels in the neighborhood centered at (i, j) and is defined as

$$I\Theta K(x) = \min_{y\epsilon K}\{I(x + y)\} \qquad (13)$$

Where $x \epsilon E^2$, K is a subset of E^2, gray scale function dilated by a structuring element K.

Dilation sets a pixel at (i, j) to the maximum over all pixels in the neighborhood centered at (i, j) and is defined as

$$I\oplus K(x) = \min_{y\epsilon K}\{I(x - y)\} \qquad (14)$$

Morphological Opening on an image is defined as erosion followed by dilation. Also, morphological closing is defined as dilation followed by erosion. They are represented as

$$\left. \begin{array}{l} I \circ K = (I\Theta K)\oplus K \\ I \cdot K = (I\oplus K)\Theta K \end{array} \right\} \qquad (15)$$

3.3 Wavelet Transform
The wavelet transform specifies a multiresolution decomposition, with the wavelet defining the highpass or bandpass filter that determines the detail information. Associated with each wavelet, there is a smoothing function which defines the complementary lowpass or average filter [25,26]. The wavelet function is the mother wavelet which detects the differences and the scaling or smoothing function is the father wavelet that captures the average behavior. A wavelet is represented in terms of scaling and wavelet functions as follows:

$$f(t) = \sum_{k=-\infty}^{\infty} c(k)\phi_k(t) + \sum_{k=-\infty}^{\infty}\sum_{j=0}^{\infty} d_j(k)\psi_{j,k}(t) \qquad (16)$$

Where $\phi_k(t)$ is a scaling function and $\psi_{j,k}(t)$ is a wavelet function.

There are two main groups of wavelet transform, continuous and discrete. Since it is difficult or impractical to apply continuous wavelet transform, DWT is commonly used for image fusion. According wavelet decomposition theory [26,27], the original image is decomposed using 2-dimensional wavelet transform, which results in a low frequency sub band and three high frequency sub-bands. The low frequency sub-band reflects the approximate or smoothed component that is related to image's basic information, while the high frequency sub-bands reflect the detail components along the horizontal, vertical, and diagonal direction which corresponds to edge, line, area, boundary and other information present in the image.

The 2-D wavelet analysis of an image I consists of filtering and down-sampling of each row using the 1-D low-pass filter L and followed by high-pass filter H thus producing the coefficient matrices IL and IH. This step is followed by vertical filtering and down-sampling of column elements using the low-pass and high-pass filters L and H in IL and IH. This first level of decomposition produces four sub images ILL, ILH, IHL and IHH. ILL is a smooth sub image corresponding to a low-frequency band and can be considered as a subsampled version of the original image I containing the basic information. ILH, IHL and IHH are detail sub images corresponding to horizontal, vertical and diagonal directions of the image I containing edge, line and boundary information. Fig. 2 represents the multiresolution wavelet decomposition of an image I after the first level.
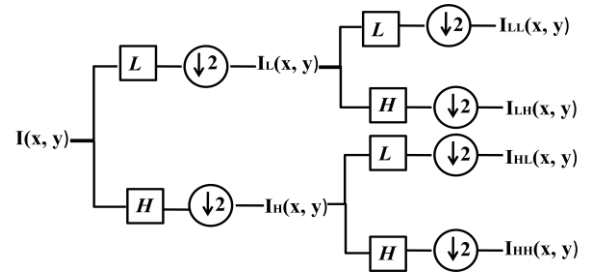
Fig 2: One stage of 2-D DWT multiresolution decomposition

The 2-D decomposition of an image at the first level can be expressed by the following equations [6]: Let I(x, y) be the original image of size *M x N*, *l(i)* the low-pass coefficients of a specific wavelet basis, where i = 0, 1, 2, ……., N_{l-1} and N_l is the length of the low-pass filter *L*, *h(j)* the high-pass coefficients of a specific wavelet basis, where j = 0, 1, 2, ……., N_{h-1} and N_h is the length of the high-pass filter *H*. Then,

$$I_L(x, y) = \frac{1}{N_l}\sum_{i=0}^{N_l-1} l(i).I((2x + i) \bmod M, y) \qquad (17)$$

$$I_H(x, y) = \frac{1}{N_h}\sum_{j=0}^{N_h-1} h(j).I((2x + j) \bmod M, y) \qquad (18)$$

where $x = 0,1,2,……M/2 -1$ and $y = 0,1,2,……N -1$.

$$I_{LL}(x,y) = \frac{1}{N_l}\sum_{i=0}^{N_l-1} l(i).I_L(x,(2y+i)\,mod\,N) \quad (19)$$

$$I_{LH}(x,y) = \frac{1}{N_h}\sum_{j=0}^{N_h-1} h(j).I_L(x,(2y+j)\,mod\,N) \quad (20)$$

$$I_{HL}(x,y) = \frac{1}{N_l}\sum_{i=0}^{N_l-1} l(i).I_H(x,(2y+i)\,mod\,N) \quad (21)$$

$$I_{HH}(x,y) = \frac{1}{N_h}\sum_{j=0}^{N_h-1} h(j).I_H(x,(2y+j)\,mod\,N) \quad (22)$$

where $x = 0,1,2,\ldots\ldots M/2$ -1 and $y = 0,1,2,\ldots\ldots N/2$ -1.

The 2-D decomposition process can then be iterated on the smooth sub image $I_{LL}(x,\ y)$ to obtain four coefficient matrices in the next decomposition level and so on.

4. RESULTS AND DISCUSSION

The above mentioned methods are tested for MRI/CT and PET/CT images. The effectiveness of the fusion methods and the quality of the resultant fused images can be evaluated and analyzed quantitatively with the image quality metrics.

Image quality metrics are classified into two categories: reference and non-reference metrics. Reference metrics evaluates against the reference image. But in real-time applications, the availability of reference image is not possible. Hence in this paper, the non-reference image quality metrics alone considered for evaluation. The non-reference image quality metrics used in this paper are EN, SD, PSNR, SF, AG, ES, FF and FS [28,9,7]. The EN shows the amount of information in the image. The SD and PSNR indicate the quality of an image. Higher PSNR and standard deviation indicates better quality. The SF shows the overall activity and clarity of an image. ES represents the edge information associated with the fused image which is visually supported by human visual system. Higher the value of ES implies fused image with better edge information. FF shows how much information is derived from source images. Low value of the fusion symmetry indicates the goodness of the fusion algorithm.

Some of these metrics measure the amount of information present in the fused image, whereas the other metrics measure the quality of an image. In medical applications for better diagnosis, an image should have both these qualities. Hence, equal significance should be given to both the quality metrics that measure quality and information content. And also, the values of each of the quality metric may widely vary in their range. In order to bring the metric values in the uniform range and by giving equal importance, the above considered non-reference image quality metrics are statistically normalized. In this paper, a simple normalization which linearly maps a given value from the existing range to the desired range is considered and is expressed as

$$x' = a * x + b \quad (23)$$

Where x is a vale in the existing rang and x' is the new value. a and b are defined as follows:

$$\left.\begin{array}{l} a = (max' - min')/(max - min) \\ b = max' - a * max \text{ (or) } min' - a * min \end{array}\right\} \quad (26)$$

Where *max* and *min* are the maximum and minimum values in the existing range and *max'* and *min'* are maximum and minimum values in the new range.

For simplicity, the transform-based image fusion methods both pyramid transform and wavelet transform assume the decomposing level as 3 and the Combination rule is considered to be the measure and salience with consistency check for detailed coefficient and weighted average for approximate coefficient. The window size for consistency check is considered to be 3x3. The threshold vale is also assumed to be fixed.

4.1 MRI/CT Image Data Set

Here, MRI image (Fig. 3, 1A) and CT image (Fig. 3, 1B) of a brain is considered as the source images. The size of the image is 159x159. Both the images are preprocessed and co-registered images. MRI offers better information on normal and pathological soft tissues and CT provides best description about the denser tissues mainly the structures of bones and the useful information like brain boundaries and shape. In most of the clinical applications, the combined images showing clearly the outline, shape, position of both bone and tissues is desirable that can aid in diagnosis and treatment planning for doctors. The images in Fig. 3, 1C through 4C are the resultant fused images of various fusion methods.
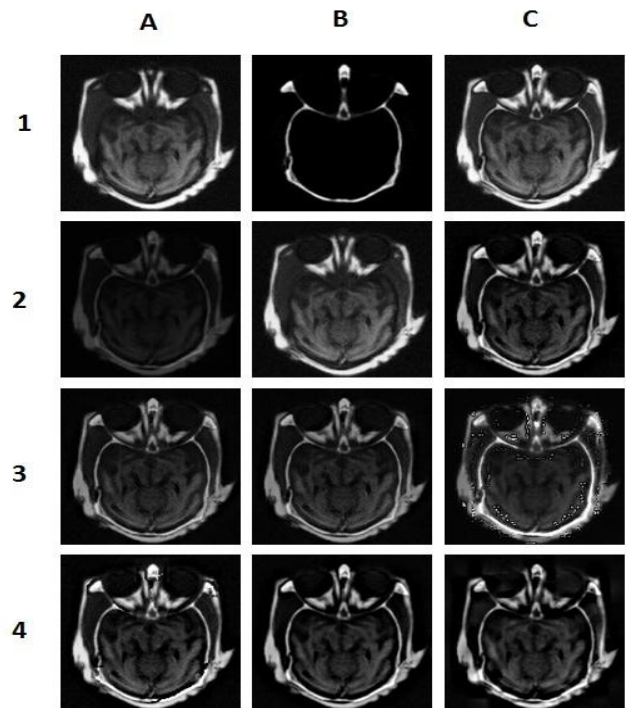

Fig 3: Fusion results of MRI/CT Image Data Set

1A) MRI Image 1B) CT Image 1C) MAX 2A) WAVG

2B) PCA 2C) Lapl. 3A) FSD 3B) Grad. 3C) ROLP

4A) Morph. 4B) Daube. 4C) Haar

The resultant fused images are quantitatively analyzed with the non-reference quality metrics EN, SD, PSNR, AG, ES, FF and FS. All the values are normalized and fit within the range 1 to 100. The Table 1 shows the values of normalized image quality metrics. It shows that the method MAX has the highest EN and FF values and PCA has the highest SD value. The highest values of other quality metrics PSNR, AG, ES and FS are shared by both Haar and Daubechies wavelet

transforms. It shows that the wavelet-based methods better capture the useful patterns from the source images and produces the fused image with more information and good quality.

Table 1. Comparison of various image fusion methods for MRI/CT Images

Method	EN	SD	PSNR	AG	ES	FF	FS	Total
MAX	100.0	95.0	32.8	54.2	65.1	96.1	44.0	487.2
WAVG	27.3	1.0	1.0	1.0	1.0	55.1	39.6	126.0
PCA	1.0	79.3	100.0	38.1	46.4	83.8	56.7	405.3
Lapl.	1.1	81.4	16.5	57.3	68.6	1.0	5.7	231.6
FSD	1.0	49.8	16.6	35.5	41.7	7.8	11.9	164.3
Gradi.	4.2	88.0	3.2	67.0	69.8	7.1	1.0	240.3
ROLP	3.6	50.7	16.6	36.0	42.2	7.5	10.5	167.1
Morph.	41.0	97.1	14.2	67.5	80.3	8.5	12.7	321.3
Daube.	67.3	100.0	45.5	88.6	100.0	81.6	65.1	548.1
Haar	83.4	91.2	52.2	100.0	99.7	100.0	100.0	626.5

The Fig. 4 shows that the wavelet transform methods both Haar and Daubechies have achieved the normalized total nearer to each other showing that those methods can better extract the edges, textures and other required features from the source images.
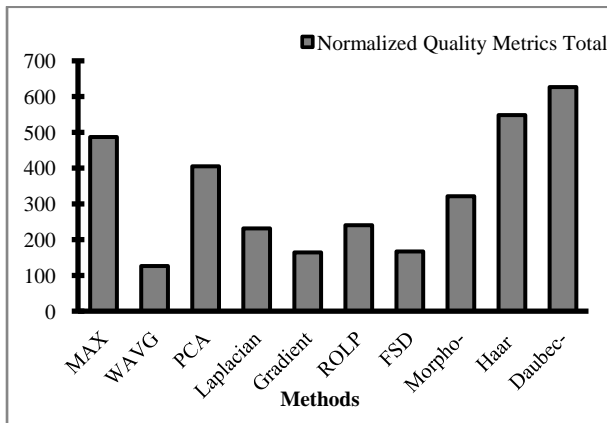
Fig 4: Comparison of various fusion methods based on normalized total of quality metrics for MRI/CT images

4.2 PET/CT Image Data Set

PET image (Fig. 5, 1A) and CT image (Fig. 5, 1B) of a neck are considered as the source images. The size of the image is 272x336. Both the images are preprocessed registered images. CT provides best description about the denser tissues mainly the structures of bones and the useful information like bone boundaries and shape. Whereas PET images offer functional information such as blood flow, blood activity and metabolic changes with low spatial resolution. In Fig. 5, 1A shows shape and boundary of a bone and other organ structure and 1B shows the metabolic activity (absorption of glucose) indicating the presence of malignant lesion. In most of the clinical applications, particularly in oncology, the combined images showing both the information are desirable for effective diagnosis and for treatment planning. The images in Fig. 5, 1C through 4C are the resultant combined images of various fusion methods.
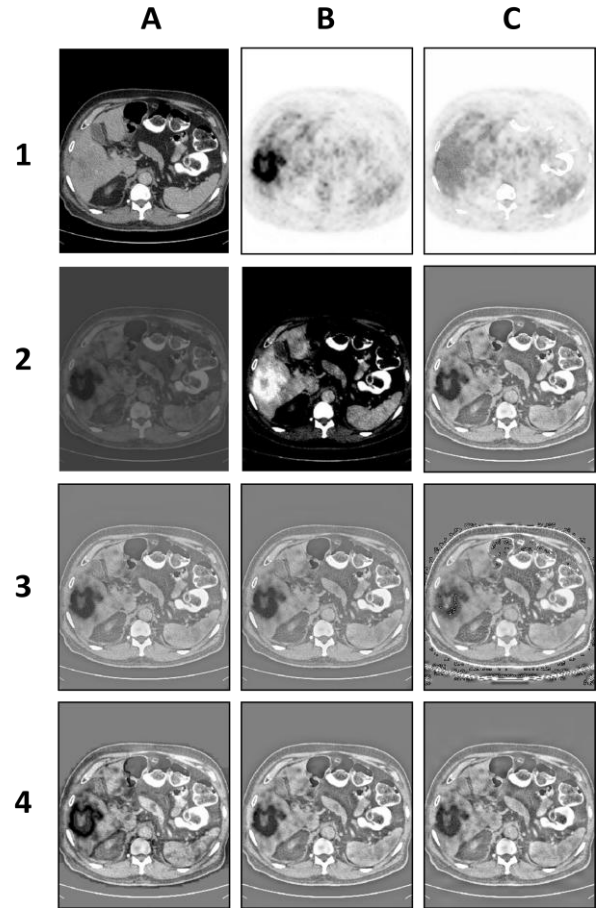

Fig 5: Fusion results of PET/CT Image Data Set

1A) MRI Image 1B) CT Image 1C) MAX 2A) WAVG

2B) PCA 2C) Lapl. 3A) FSD 3B) Grad. 3C) ROLP

4A) Morph. 4B) Daube. 4C) Haar

Table 2. Comparison of various image fusion methods for PET/CT Images

Method	EN	SD	PSNR	AG	ES	FF	FS	Total
MAX	100.0	14.6	1.0	55.0	1.0	85.3	100.0	356.9
WAVG	27.2	1.0	83.2	55.6	2.4	35.0	1.0	205.4
PCA	1.0	100.0	17.0	90.5	68.0	53.0	30.9	360.4
Lapl.	16.9	22.9	60.4	72.0	47.9	15.2	11.1	246.4
FSD	9.7	15.5	59.3	67.3	32.7	19.2	14.6	218.3
Grad.	17.4	15.5	59.3	67.2	32.6	19.5	14.6	226.1
ROLP	3.4	68.3	29.8	86.3	60.9	1.0	31.3	281.0
Morph.	59.0	29.1	58.6	94.2	54.2	18.7	10.3	324.1
Daube.	76.2	70.1	89.0	91.7	100.0	100.0	37.5	564.5
Haar	95.9	95.4	100.0	100.0	91.6	63.9	74.8	621.6

The resultant fused images are quantitatively analyzed with the non-reference quality metrics EN, SD, PSNR, AG, ES, FF and FS. All the values are normalized and fit within the range 1 to 100. The normalized values of image quality metrics are

shown in Table 2. It shows that the highest EN and FS values are scored by the MAX and PCA has the largest SD value. The highest values of other quality metrics SD, AG, ES and FF are shared by both Haar and Daubechies wavelet transforms. It is once again proved that the wavelet-based methods better capture the useful patterns from the source images and produces the fused image with more information and good quality.
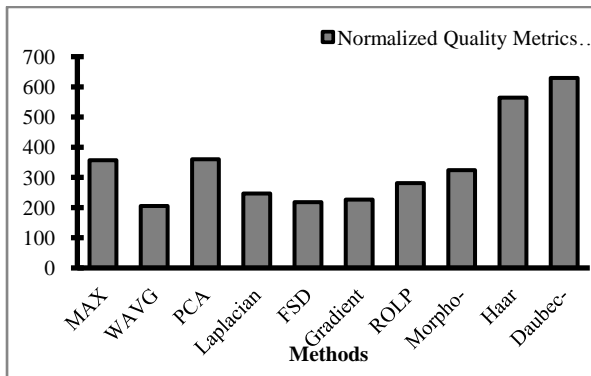
Fig 6: Comparison of various fusion methods based on normalized total of quality metrics for PET/CT images

The Fig. 6 shows that the wavelet transform methods both Haar and Daubechies have achieved the normalized total nearer to each other showing that those methods can better extract the edges, textures and other required features from the source images.

5. CONCLUSIONS

In this paper, some of the fusion methods in the spatial and transform domain are implemented and quantitatively analyzed with non-reference image quality metrics for multi-modality medical images. The analysis says that the wavelet transform results in good quality fused image with extended information than spatial and pyramid transform methods by effectively extracts the useful patterns from the source images. But the effectiveness of the wavelet transform depends on its parameters. Here, for simplicity, all these methods are implemented for fixed decomposition level and fixed threshold value, which may differ for different modality images with varying detailed information. The combination rule is assumed as fixed both for detailed and approximate coefficients. Since the medical images from different modalities differ and are complementary to each other, these parameters may be dynamically chosen that, they well suit to medical image fusion and be optimized.

6. REFERENCES

[1] Li H., Manjunath B.S., and Mitra S.K., (1995), Multi-sensor image fusion using the wavelet transform, Graphical Models and Image Processing, Volume(57(3)), 235 – 245.

[2] Mitchell H.B., (2010), Image Fusion: Theories Techniques and Applications, Springer, ISBN 978-3-642-11215-7.

[3] Gemma Piella, (2002), A General Framework for Multiresolution Image Fusion: from Pixels to Regions, Probability, Networks and Algorithms (PNA).

[4] John J. Lewis, Robert J. O'Callaghan, Stavri G. Nikolov, David R. Bull and Nishan Canagarajah, (2005), Pixel-and region-based image fusion with complex wavelets, Information Fusion, Volume(8), 119-130.

[5] John J. Lewis, Robert J. O'Callaghan, Stavri G. Nikolov, David R. Bull and Nishan Canagarajah, (2005), Pixel-and region-based image fusion with complex wavelets, Information Fusion, Volume(8), 119-130.

[6] Pajares G. and de la Cruz J.M., (2004), A wavelet-based image fusion tutorial, Pattern Recognition, Volume(37(9)), 1855 – 1872.

[7] Yang L., Guo B.L., and Ni W., (2008), Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform, Neurocomputing, Volume(72), 203–211.

[8] Sabalan Daneshvar and Hassan Ghassemian, (2010), MRI and PET image fusion by combining IHS and retina-inspired models, Information Fusion, Volume (11), 114-123.

[9] Rajiv Singh and Ashish Khare, (2014), Fusion of multimodal medical images using Daubechies complex wavelet transform – A multiresolution approach, Information Fusion, Volme(19), 49–60.

[10] Sushmita Mitra and B. Uma Shankar, (2015), Medical image analysis for cancer management in natural computing framework, Information Sciences, Volume (306), 111–131.

[11] Xiaoli Zhang, Xiongfei Li, Yuncong Feng, Haoyu Zhao and Zhaojun Liu, (2015), Image fusion with Internal Generative Mechanism, Expert Systems with Applications, Volume(42), 2382–2391.

[12] Vince D. Callhoun and Tulay Adali, (2009), Feature - Based Fusion of Medical Imaging Data, IEEE Trans. On Inf. Tech. in Biomedicine, Volume (13(5)), 711 – 720.

[13] Burt P.J. and Kolczynski R.J., (1993), Enhanced image capture through fusion, Proc. 4th Intl. Conf. on Computer Vision, 173-182.

[14] Zang Z. and Blum R.S., (1999), A categorization of multi-scale decomposition-based image fusion schemes with a performance study for a digital camera application, Proc. of the IEEE, Volume(87(8)), 1315 – 1326.

[15] Liu X. and Yang W., (2000), Enhanced visualization of images through fusion, Proc. of SPIE, Volume (4231), 340–345.

[16] Rick S. Blum, Zhiyn Xue and Zhong Zhang, 2005, An Overview of Image Fusion, Multi-Sensor Image Fusion and its Applications.

[17] Burt P.J., Adelson E.H., (1985), Merging Images Through Pattern Decomposition, Proc. of SPIE, volume(575), 173-181.

[18] Burt P.J. and Adelson E. H., (1983), The Laplacian Pyramid as a Compact Image Code, IEEE Transactions on Communications, Volume (31(4)), 532 – 540.

[19] Adelson E. H., Anderson C. H., Bergen J. R., Burt P. J., and Ogden J. M., (1984), Pyramid methods in image processing, RCA Engineer, Volume(29(6)).

[20] Burt P.J., (1984), The Pyramid as a Structure for Efficient Computation, Multiresolution Image Processing and Analysis, Springer.

[21] Qu X. J., Zhang F. and Zhang Y., (2013), Feature-Level Fusion of Dual-Band Infrared Images Based on Gradient

Pyramid Decomposition, Applied Mechanics and Materials, Volume (347-350), 2380-2384.

[22] Toet A., (1989a), Image fusion by a ratio of low-pass pyramid, Pattern Recognition Letters, Volume (9), 245 – 253.

[23] Toet A., Van Ruyven L. J. and Valeton J. M., (1989b), Merging Thermal And Visual Images By A Contrast Pyramid, Optical Engineering, Volume (28(7)), 789-792.

[24] Neha Uniyal and S.K. Verma, (2014), Image Fusion using Morphological Pyramid Consistency Method, International Journal of Computer Applications, Volume(95).

[25] Lee A. Barford, R. Shane Fazzio and David R. Smith, (1992), An Introduction to Wavelets, HPL, 92-124.

[26] Rioul O. and Veterli M., (1991), Wavelets and Signal Processing, IEEE Signal Processing Magazine.

[27] Mallat S., (1989), A Theory for Multiresolution Signal Decomposition: The Wavelet Representation, IEEE Trans. On Pattern Analysis and Machine Intelligence, Volume (11(7)), 674-693.

[28] Toet A., (1990), Hierarchical Image Fusion, Machine Vision and Applications, volume (3), 1-11.

[29] Changtao He, Quanxi Liu, Hongliang Li and Haixu Wang, (2010), Multimodal medical image fusion based on IHS and PCA, Procedia Engineering, Volume(7), 280–285.