Image Fusion Techniques for Wireless Sensor Networks: Survey

B.Roopa, SunilKumar.S.Manvi
Dept. of Electronics and Communication Engineering
REVA Institute of Technology and Management Bengaluru, India
Email: roopabayareddy@gmail.com, sunil.manvi@revainstitution.org

Abstract— In a Wireless Multi-media Sensor Network (WMSN), each of the camera enabled node can capture from a limited physical area of the environment. The captured image could be noisy, incomplete, and redundant and may be of no practical use. Transmission of individually sensed images to the sink/base station which can further process them could be the intuitive solution. But WMSNs are energy-constrained, have limited bandwidth and are subjected to hostile conditions. Greater portion of node energy is consumed for communication between the sensor nodes. Flooding of sensed images would drain the network to death in short time. To accomplish the desired task (surveillance, tracking) images captured from different sensors have to be fused at different levels of network hierarchy. Fused image free of redundancy, brings out complementary features and aids in further analysis. Thus the application of distributive image fusion techniques in WMSN can prolong the network lifetime. This paper brings out the survey on state-of-the-art image fusion techniques, the fusion techniques that are widely used and that can be possibly used in wireless sensor networks.

Keywords-- WMSN, image fusion, SVD, DWT

I. INTRODUCTION

Wireless networks in combination with image sensors have given scope for numerous sensing applications. Recent trends in wireless communications and sensor technology has en-abled the development of low-cost Wireless Sensor Networks (WSNs). A WSN is a network consisting of thousands of sensors that span a large geographical region. These sensors are able to communicate with each other to collaboratively detect objects, collect information, and transmit messages. Sensor networks have become an important technology specially for environmental monitoring, military applications, disaster management, etc.[1], [2]. A typical WSN application is shown in Fig.1. The availability of low-cost hardware such as CMOS cameras and microphones has fostered the development of Wireless Multimedia Sensor Networks (WMSNs), i.e., net-works of wirelessly interconnected devices that are able to ubiquitously retrieve multimedia content such as video and audio streams, still images, and scalar sensor data from the environment [2]. However, as sensors are usually small in size, they have many physical limitations. For example, due to its limited size, a sensor does not have a very powerful CPU and is limited in computational power and memory. This limitation in energy puts extra constraints in the operations of sensors. As recharging is difficult, sensors should smartly utilize their limited energy in collecting, processing, and transmitting in-formation. In wireless camera-based sensor networks, energy of nodes is used to image transmission, processing and energy has to be spent more for transmission. Thus smart ways of image processing and transmission are essential. Image fusion technique is proven to be boon for accomplishing the desired task and prolonging network lifetime.

There exist numerous image fusion techniques ranging from simple averaging to contourlet transforms in the field of image processing. Application of image fusion include improving geometric correction, enhancing certain features not visible in either of the single data alone, change detection using temporal data sets and enhancing to provide a complete information for diagnosis. Algorithms such as the intensity, hue and saturation (IHS) algorithm and the wavelet fusion algorithm have proved to be successful in satellite and medical image fusion. Image fusion methods can be broadly classified into two categories: spatial domain
fusion and transform domain fusion. Averaging, Brovey method, Principal Component Analysis (PCA), based methods are spatial domain methods. But spatial domain methods produce spatial distortion in the fused image. This problem can be solved by transform domain algorithms such as wavelets. The images to be fused should already be registered. The rest of this article is organized as follows: A brief introduction on the image fusion concept is given in Section II. Section III gives note of multimedia wireless sensor networks. Section IV presents image fusion techniques being used in WMSNs and the possible candidate techniques. Section V contains the conclusions drawn on the fusion algorithms followed by scope for research in section VI.

II. THE CONCEPT OF IMAGE FUSION

Image fusion is the process of merging two images of the same scene to form a single image that is more intelligent and helps for further analysis and processing [3]. Fused image example is shown in Fig.2 in which left hand image has big clock out of focus and right side image has small clock out of focus. The fused image has both object in the focus.

![Fig.2 Example of multi-focus image fusion](image)

Image fusion is important in many different image processing fields such as satellite imaging, remote sensing, surveillance and medical imaging. Image fusion reduces uncertainty and minimizes redundancy in the output while maximizing relevant information from two or more images of a scene into a single composite image that is more informative. Fused images would be more suitable for visual perception or processing tasks like medical imaging, remote sensing, concealed weapon detection, weather forecasting, biometrics etc. The human vision mechanism is primarily sensitive to moving light stimuli. If fusion process introduces any moving artifacts, it is highly distracting to the human observer.

The fused images of the image fusion algorithms can be evaluated with respect to reference image using performance metrics. Popularly used metrics are Entropy, Standard Deviation, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Signal-to-Noise Ratio (SNR), Correlation (CORR), Mutual Information (MI), Quality Index (QI), Structural Similarity (SSIM) [11]. A very brief discussion of some well-known image fusion techniques [3],[4],[5],[6],[7],[8],[9],[10] is as follows. Table I, Table II, lists these values obtained in MATLAB for various fixed basis vectors dependent (DCT, DWT, SWT, DTCWT) and data set dependent basis vectors algorithms (PCA, SVD). These results are with respect to 512 X 512 gray scale image shown in Fig.2. Memory of the algorithm files and their execution time is given in Table III. From the tables it can be inferred that depending the application it is wise to choose the particular algorithm for WSN.

A. Simple Averaging and Select maximum

In simple averaging the value of the pixel $P(i;j)$ of each image is taken and added. This sum is then divided by 2 to obtain the average. The average value is assigned to the corresponding pixel of the output image. In select maximum value of the pixel $P(i;j)$ of each image is taken and compared to each other. The greatest pixel value is assigned to the corresponding pixel. Multiplicative method is a simple multiplication of each multispectral band with the panchromatic image. By multiplying the same information into all bands, however, creates spectral bands of a higher correlation which means that it does alter the spectral characteristics of the original image. The Brovey transformation was developed to avoid the disadvantages of the multiplicative method. It is a combination of arithmetic operations and normalizes the spectral bands before they are multiplied with the panchromatic image. The spectral properties, however, are usually not well preserved.

B. IHS and PCA

In IHS three bands of a multispectral image are transformed from the RGB domain into the IHS color space. The panchromatic component is matched to the intensity of the IHS image and replaces the intensity component. After the matching, the panchromatic image replaces the intensity in the original IHS image and the fused image is transformed back into the RGB color space. This method works also well with data from one sensor, but for multi-temporal or multi-sensor fusion the results are in most cases not acceptable.

The PC transform is a statistical technique that transforms a multivariate dataset of correlated variables into a dataset of uncorrelated linear combinations of the original variables. For images, it creates an uncorrelated feature space that can be used for further analysis instead of the original multispectral feature space. The PC is applied to the multispectral bands. The panchromatic image is histogram matched to the first principal component (sometimes to the second). It then replaces the selected component and an inverse PC transform takes the fused dataset back into the original
multispectral feature space. The advantage of the PC fusion is that the number of bands is not restricted (such as for the original IHS or Brovey fusions). It is, however, a statistical procedure which means that it is sensitive to the area to be sharpened. The fusion results may vary depending on the selected image subsets.

C. Pyramids, probability and biological models

An image pyramid consists of a set of lowpass or bandpass copies of an image, each copy representing pattern information of at different scale. Typically, in an image pyramid every level is a factor two smaller as its predecessor, and the higher levels will concentrate on the lower spatial frequencies. An image pyramid does contain all the information needed to reconstruct the original image. When the output of the sensors are non-correlated each image can be represented by a conditional density function

\[ P(\mathbf{x}, \mathbf{y} / I) \]

where \( \mathbf{x} \) represents a particular imaging sensor output. Then the effect of using all of the sensory outputs is equivalent to the use of the total probability function. Biologically inspired fusion uses biological models of color and visible/infrared vision as bases.

D. Wavelets

Discrete wavelet transform (DWT) has gained popularity as a fusion tool [4]. However, DWT is a shift variant transform. The shift invariant DWT (SIDWT) case is identical to the one in the generic wavelet fusion case: the input images are decomposed into their shift invariant wavelet representation and a composite shift invariant wavelet representation is built by the incorporation of an appropriate selection scheme [5]. Gurpreet Singh et al. [10] have proposed Modified Haar Wavelet Transform which is an enhanced version of Haar Wavelet Transform which can reduce the calculation work and is able to improve the contrast of the image. The main achievement of MHWT is sparse representation and fast transformation. MHWT is relatively efficient because at each level, only half of the original data has to be stored. Two important properties: wavelet symmetry and linear phase of BWT can be exploited for image fusion in a pixel-level image fusion scheme using multi-resolution Bi-orthogonal wavelet transform (BWT) [7] because they are capable to preserve edge information and hence reducing the distortions in the fused image.

The wavelet transform concentrates on representing the image in multi-scale and it is appropriate to represent linear edges. For curved edges, the accuracy of edge localization in the wavelet transform is low. So, there is a need for an alternative approach which has a high accuracy of curve localization such as the curvelet transform [6]. Contourlet transform gives sparse representation of images i.e. most of the contourlet coefficients are close to zero. This property of contourlet transform is useful in denoising. The basic elements of contourlet transform have different aspect ratios and are oriented in various directions. The contourlet coefficients of different decomposition level have different characteristics. Thus use of level dependent threshold gives further improvement in the denoising results. Denoising before fusion gives better results than that of denoising after fusion. Contourlet based image fusion gives better result than SWT and DTCWT in terms of PSNR, entropy, edge strength etc. [8].

E. Singular Value Decomposition

In linear algebra, singular value decomposition is a matrix factorization method. A novel image fusion technique based on multi-resolution singular value decomposition (MSVD) has been presented and evaluated in [12]. It is observed that image fusion by MSVD perform almost similar to that of wavelets. It is computationally very simple and it could be well suited for real time applications. Moreover, MSVD does not have a fixed set of basis vectors like FFT, DCT and wavelets; its basis vectors depend on the data set.

<table>
<thead>
<tr>
<th>Parameter/Algorithm</th>
<th>Averages</th>
<th>Select maximum</th>
<th>SVD</th>
<th>PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>10.9533</td>
<td>17.3132</td>
<td>10.1336</td>
<td>10.9522</td>
</tr>
<tr>
<td>MAE</td>
<td>4.2372</td>
<td>4.2535</td>
<td>3.9407</td>
<td>4.2369</td>
</tr>
<tr>
<td>PFE</td>
<td>4.7906</td>
<td>7.5722</td>
<td>4.4321</td>
<td>4.7901</td>
</tr>
<tr>
<td>CORR</td>
<td>0.9989</td>
<td>0.9971</td>
<td>0.9990</td>
<td>0.9989</td>
</tr>
<tr>
<td>SNR</td>
<td>26.3922</td>
<td>22.4156</td>
<td>27.0678</td>
<td>26.391</td>
</tr>
<tr>
<td>MI</td>
<td>1.3529</td>
<td>1.3915</td>
<td>1.3707</td>
<td>1.3407</td>
</tr>
<tr>
<td>QI</td>
<td>0.5874</td>
<td>0.7818</td>
<td>0.4764</td>
<td>0.4562</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.9575</td>
<td>0.9362</td>
<td>0.9638</td>
<td>0.9575</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter/Algorithm</th>
<th>DCT</th>
<th>DWT</th>
<th>SWT</th>
<th>DTCWT</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>7.1313</td>
<td>9.7342</td>
<td>10.1336</td>
<td>0.9234</td>
</tr>
<tr>
<td>MAE</td>
<td>4.9386</td>
<td>1.5177</td>
<td>3.7968</td>
<td>0.5524</td>
</tr>
<tr>
<td>PFE</td>
<td>3.0593</td>
<td>1.4116</td>
<td>4.2574</td>
<td>0.4039</td>
</tr>
<tr>
<td>CORR</td>
<td>0.9995</td>
<td>0.9999</td>
<td>0.9991</td>
<td>1.0000</td>
</tr>
<tr>
<td>SNR</td>
<td>30.2875</td>
<td>37.0059</td>
<td>27.4171</td>
<td>47.8749</td>
</tr>
<tr>
<td>MI</td>
<td>1.1106</td>
<td>1.3538</td>
<td>1.3806</td>
<td>1.5269</td>
</tr>
<tr>
<td>QI</td>
<td>0.1568</td>
<td>0.5125</td>
<td>0.4953</td>
<td>0.5575</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.8236</td>
<td>0.9817</td>
<td>0.9684</td>
<td>0.9978</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>SAm(deg)</th>
<th>Corr Coeff</th>
<th>ERGAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>BROVEY</td>
<td>0</td>
<td>0.81</td>
<td>23.4</td>
</tr>
<tr>
<td>CS-BROVEY</td>
<td>0</td>
<td>0.71</td>
<td>4.84</td>
</tr>
<tr>
<td>PCA</td>
<td>9.59</td>
<td>0.89</td>
<td>21.4</td>
</tr>
<tr>
<td>CS-PCA</td>
<td>4.85</td>
<td>0.74</td>
<td>11.62</td>
</tr>
</tbody>
</table>
III. WIRELESS MULTIMEDIA SENSOR NETWORK

Multimedia is an effective tool of communication. A multi-media signal is one that integrates signals from several media sources, such as video, audio, graphics, animation, text in a meaningful way to convey some information. Cyclops and XYZ have developed several prototypes with camera modules directly mounted on wireless sensor platforms. These low-cost camera sensor systems are ideal for quick deployment in unmanageable spaces, such as the battlefield for military applications and the remote areas for habitat studies. Due to severe bandwidth constraints on the low-power radios, however, these devices typically employ lightweight on-board image processing algorithms and do not provide continuous high-resolution images. Challenges involved with multimedia communication are listed below:

(i) Bandwidth limitations of communication channels.
(ii) Real-time processing requirements.
(iii) Inter-media synchronization.
(iv) Intra-media continuity.
(v) End-to-end delays and delay jitters.
(vi) Multimedia indexing and retrieval.

IV. IMAGE FUSION TECHNIQUES AND WIRELESS SENSOR NETWORKS

Though we don’t find exclusive literature on image fusion techniques in WSN, there have been works on energy efficient image transmission which involve compression and image fusion. Manvi et al., [13], proposed a Context Aware agent based Distributed Sensor Network (CADSN) to form an improved infrastructure for multi-sensor image fusion to monitor the militant activities. The proposed work is based on context aware computing which uses software mobile agents for image fusion in WMSN. Instead of each source node sending sensed images to the sink node, images from the different active nodes are fused and sent to sink node by using mobile agent. MinWu et al, used a shape matching method based image fusion [14]. Nasri, et al have adopted distributed image compression taking advantage of JPEG 2000 still image compression which optimizes network life time and memory requirements [15]. Compressive sensing (CS) has received a lot of interest due to its compression capability and lack of complexity on the sensor side. Wan et al [16] have exploited the properties of compressive measurements through different sampling patterns and their potential use in image fusion. CS-based image fusion has a number of perceived advantages in comparison with image fusion in the multi-resolution (MR) domain.

A. Context Aware Image fusion

The scheme comprises of three phases: context gathering, context interpretation, and image fusion. Context gathering - contexts are gathered from the target, i.e., sensed image and time from the target, and stored in the node for a short period until its interpretation is done. Context interpretation-sensed images are compared with previous image and set of critical image features (weapons, explosives, enemy, etc.) stored at the node. If image analysis yields some general or critical object feature existence, information fusion process is invoked. Image fusion-relevant images from active sensor nodes corresponding to object existence are fused to get a clear picture of the object and make some decisions.

Image fusion is classified into two types namely low resolution image fusion and high resolution image fusion. Low resolution image fusion is used for contexts like general object detection, and general image gathering by sink. High resolution image fusion is done for contexts such as critical object detection, image gathering in night time by sink. During nights, it is better to monitor the target periodically since lighting condition is poor and possibility of enemy attack, militant activities, etc., are high. Sink driven image fusion is based on the time of sensing, available network bandwidth and sensor node battery. Static and mobile agents are employed to perform the fusion process. The scheme assumes that an agent platform is available in the nodes of WSNs. However, if an agent platform is unavailable, the agent communicates by traditional message exchange mechanisms such as message passing method. Fusion is carried out using DWT. Agent based context-aware DWT image fusion scheme can be summarized as:

1. Context Agent (CA) of the active sensor nodes gather the context (sensing time, image). Node manager Agent (NMA) interprets the context using its data base.
2. NMA floods the context to the sink.
3. Sink Manager Agent (SMA) interprets the context and creates Fusing Agent (FA) along with fusion code.
4. FA visits the first active node and fuses the image and migrates to the next active node and continues the process till it visits all the active nodes.
5. FA returns to the sink along with fused image

B. Image fusion using image matching and background subtraction

A shape matched method is used to coarsely register images to find out maximal overlap to exploit the spatial correlation between images acquired from neighboring sensors. For a given image sequence, background image is transmitted only once. A lightweight and efficient background subtraction method is employed to detect targets. Only the regions of target and their spatial locations are transmitted to the monitoring center. At sink whole image is reconstructed by fusing the background and the target images as well as their
locations. This indeed reduces energy consumption for image transmission. Consider the network shown in Fig.3. The whole process involves in-network processing and reconstruction of images at the sink as explained below. However this is suitable for applications that involve only one fixed sensor equipped with camera.

1. Transmit the background of the target along the route of sensor 1, sensor 2, sensor 3, and remote sensor and another route of sensor 4, sensor 5, sensor 6, and remote sensor, respectively.

2. At sensors 2, 3, 5, and 6 apply image matching to remove spatial redundancy between images in sensors 1 and 2, sensors 2 and 3, sensors 4 and 5, and sensors 5 and 6, respectively.

3. At each sensor, whenever a target is detected using background subtraction on a new captured image, the extracted target area and its spatial location are transmitted to the remote sensor along the same route.

4. Restore the background image transmitted from each sensor.

5. Reconstruct sensor images by fusing background and target area as well as its spatial location each time after target image and its spatial location are received.

Fig.3. Transmission route

C. Compressive Image Fusion

Compressive sensing (CS) theory states that a signal that is sparse in an appropriate set of basis vectors may be recovered almost exactly from a few samples via optimum minimization if the system (image) matrix satisfies certain conditions. These conditions are satisfied with high probability for Gaussian-like vectors. Since zero-mean image patches satisfy Gaussian statistics, they are suitable for compressive sensing. If the sensed image is sparse or nearly sparse in some basis, then with high probability, the measurements essentially encode the salient information in the signal. Further, the unknown signal can be estimated from these compressive measurements to within a controllable mean-squared error. In this sense, similar fusion schemes that are used in spatial and transform domain can be used in the compressive domain too. DWT, PCA, KLT, Brovey and Non-subsampled Contourlet Transform (NSCT) depicted in Fig.4 can also be used in compressive domain. Comparison table for few algorithms is shown in Table IV which is reprinted from [17].

Fig.4. Compressive fusion using NSCT

D. Multi-modal Image Fusion

Images from different modalities often provide complementary information. For example infra-red sensors can better detect the hot objects properly than the light sensors in case of invisibility (may be due to presence of smoke, fog). Several applications require integration of complementary information for better analysis. DWT and SIDWT can be used for fusing the multi-modal images.

A novel method for adaptive fusion of multimodal surveil-lance images, based on Non-Subsampled Contourlet Transform (NSCT), which has an improved performance over Visual Sensor Networks (VSN) has been proposed in [9]. To reduce the energy and bandwidth used in transmission, the proposed method uses Compressive sensing (CS) which can compress the input data in the sampling process efficiently. Since CS is more efficient for sparse signals, each sensor image is first decomposed into sparse and dense components. Contourlet Transform is used for this decomposition because of its ability to capture and represent smooth boundaries of objects in images, so that the reconstructed images have a better quality. The reconstructed input images are fused using an adaptive algorithm based on NSCT in a centralized server. The improvement in the quality of the fused image is achieved by the use of an image fusion metric and a search algorithm to assign optimum weights to the various regions in the segmented source images.

E. Directional Controlled Fusion in Wireless Sensor Networks

Though data redundancy can be eliminated at aggregation point to reduce the amount of sensory data transmission, it introduces new challenges due to multiple flows competing for the limited bandwidth in the vicinity of the aggregation point. On the other
image fusion flavor in WSN. Since WMSNs are energy constrained ad-hoc networks energy aware distributive nature of these algorithms along with taking advantage of the compression techniques, error resilient multipath routing would better suit the scenario.

REFERENCES


