Chapter-1

INTRODUCTION

Medical imaging provides a vital contribution to the overall diagnosis of a patient. It includes various radiological imaging techniques such as X-Ray, Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET) etc. Imaging modalities provide better anatomical visualization and analysis of the patient body and hence produces improved survival rates. Within the current clinical setting, medical imaging is a vital component in many applications [1]. Medical imaging tests are non-invasive test procedures which help the doctors to diagnose diseases or injuries for treatment planning. Since the information (info) retrieved from couple of images adopted is generally of a free nature, it is frequently demanded that efficient merging of the utile data from the images is necessary. The basic stage in this procedure of merging is to sum the modes necessitated into spatial alignment, which is cited as ‘Registration’. Later, a procedure of Fusion is needed for the unified perception of the demanded data.

In the designing of radiotherapy diagnosis [2], an instance of the utilization of registering distinctive modes can be discovered, whereas at present CT is employed almost entirely. The usage of compounding of both CT and Magnetic Resonance (MR) would provide enhancing results as the earlier is more beneficially accommodated for characterization of tumour tissue (and is in natural more beneficial to gentle tissue counterpoint), while the second is necessitated for exact dose of radiation reckoning.
1.1 IMAGE REGISTRATION

The motive of registering an image is to discover the optimum translation, which properly aligns the construction of pursuit in the source images [3]. It is a significant level of investigating an image wherein, desired information is assumed to be present in more than a single image i.e., images retrieved at distinct durations, from different outlooks or by distinct sensors. Hence, precise consolidation (or Fusion) of the practicable info from one or more images is extremely significant. A lot of research that has been implemented so far for examining the medical images was dedicated to Registration [4].

Registration of images in the field of medical applications admit fusion of images related to anatomical data from CT or MR or operational images from PET, Single Photon Emission Computed Tomography (SPECT) or operational MRI. Registration helps in the design of treatment and intervention, computer assisted diagnosis and adopting disease, atlas construction and contrast, surgery simulation, aided or guided surgery, segmentation of anatomy, radio therapy, computational model structures and deduction of image from contrast upgraded images. Registration is also practicable for the images of SPECT and PET for modifying spread out fading.

The goal of image registration is to determine an unknown geometric transformation that maps one image into another (to a certain degree of accuracy). In other words, after registration problem is solved, for each pixel in the first image the corresponding pixel in the second image is deduced. This assumes that the images are similar in the sense that both images contain the same (or similar) object, which may be rotated, translated, or elastically deformed.
Image registration plays a vital role in the field of processing of medical images and in the process of surgery which is based on image-guidance due the fact that in these processes, vital and useful data are pulled out from the multiple source images. Hence, the registration process gained more attention from the users due to the acquisition of source images from different sensors in different directions with different time frames.

1.2 IMAGE FUSION

Human beings possess wonderful sense to appreciate visuals. Eye plays a key role in supporting various human activities. An image capture of a visual scene always conveys much more information than any other description adhered to the scene. Human beings have five sensing capabilities (systems). They are eyes, ears, nose, tongue and skin. These sensors can acquire independent information. Eyes can visualize a scene. Ears can sense the data by listening to the sounds. Nose can smell the odour of an object. The tongue can sense the object’s taste. Skin can sense the texture and size of the object. All the five sensing systems act as sensors and human brain collects data from these individual sensors and fuses or combines it for compact representation or better description about a scenario. This compact data is useful for decision making and task execution.

Data fusion is a process of combining information from several sources for optimal or compact representation of a huge data supporting better description and decision making. Human brain is a best example of a data fusion system [5].

Even an eye can derive many useful details of a scene by looking at the picture more than once. Brain will integrate the visuals and give the details hidden in a single
view. Multiple views will always improve the decisions. Whenever a snap shot of a scene is taken from a digital camera, a single image captured from this does not provide satisfying results. Hence more snap shots of the same scene would be taken to get more clarity and information. It is not rare to find that none of the images contain all the required qualities. It is common to feel that the positive aspects of these are to be combined to get the desired image. This motivates for the fusion of the images, for a desired output.

The definition of image fusion is as follows, “Image Fusion is the process of merging or combining or integrating useful or complementary information of several source images so that the resultant image provides more accurate description about the scene than any one of the individual source images” [5].

Image is a two-dimensional quantity. It can be viewed as the combination of illumination and reflectance. Illumination stands for light from the source falling on the object and reflectance corresponds to the amount of light that is reflected back from the same object. The visual information present in a scene can be captured as a digital image $f(x, y)$ using a sensor array. All the elements in the sensor array will be of same modality. Hence image capture using a sensor array is simply referred to as single sensor image capture. One may be interested in the details of a scene using multiple sensor arrays, each operating in a different wavelength range. This is simply termed as multi sensor image capture. In the following discussion, it can be noted that, the term sensor is used simply in place of a sensor array.

Acquiring pictures by employing single sensor may not extract entire info every time regarding a directed scene. In certain cases, one or more images would be
required for enhanced visual understanding of the scene, which is acquired by enforcing a single sensor of similar mode or by employing numerous sensors of distinctive modes, based on the application. These pictures render appropriate complementary or optical discrete info. Effective or complementary info of these pictures should be unified into a single image to render precise description regarding the scene, than any one of the separated input images.

Two different problems related to capturing multiple images and combining the required data is discussed below:

One is Single Sensor Imaging [6], in which multiple images of the same scene are captured using a single sensor to extract more details of a targeted scene, other is Multi-Sensor Imaging [7], which requires multiple images using different sensors for the same propose.

1) **Problem 1: Single Sensor Imaging**

In digital photography, objects of a scene at different distances can’t be focused at the same time. If the lens of a camera focuses on an object at a certain distance, then other objects appear blurred. Image formation model of a sensor or a camera system is displayed in Figure 1.1.

If a point $G_1$ (green dot) on an object is focused, then a dot $g_1$ corresponding to that point will be generated on the sensor plane. Therefore, all the points which are at the same distance as that of $G_1$ from lens will appear sharp. The region of acceptable sharpness of the object is referred to as Depth of Field (DoF).
Consider another point $R_1$ (red dot) on the object behind the point $G_1$. Since this $R_1$ point is out of DoF, it will generate a dot $r_1$ somewhere before the sensor plane. As the distance of the lens from the $G_1$ increases, the object will appear more blurred. As shown in Figure 1.1., $B_1$ (dark blue dot) is in front of $G_1$ on the object plane. Since $B_1$ fall out of the DoF, it produces a dot $b_1$ behind the sensor plane, resulting in an unsharp dot on the image plane (sensor plane). For better visual quality, images should have all objects in focus. One of the best approaches to do this is to capture images with different focusing conditions and combine them to generate an all-in-one focus image.

2) **Problem 2: Multi Sensor Imaging**

Visual information present in a scene can be captured as an image using a Charge Coupled Device (CCD). The wavelength of the Visible Light that can be captured by CCD sensor ranges from 400 nm to 700 nm. However, in most of the computer vision applications, CCD image alone is not enough to provide all the details of the scene. To extract more details, complementary images of the same scene should be
captured by using multiple sensors of different modalities. This can be done by capturing images in wavelengths other than the visible light band of the electromagnetic spectrum.

As shown in Figure 1.2, medical imaging is considered to justify the need of combining information of various source images of the same scene.

![Figure 1.2: Problem specification via different medical imaging (a) dataset1 (b) dataset 2](image)

In digital photography, more than one object cannot be focused at the same time from a scene due to inherent system limitations. If the focusing is done on a single object, the information related to other objects may be lost and vice versa.

In medical imaging, different modalities like PET, SPECT, CT and MRI are employed to snap complementary info [8]. These individual image captures do not provide all the required details. Therefore, information from different captures has to be incorporated into a single image. Figure 1.2 (a) and Figure 1.2 (b) shows different sets of CT and MR images of a human brain. As shown in these figures, CT can
capture bone structure or hard tissue information. Whereas, MRI can capture soft tissue information present in the brain. For a radiologist, fused image obtained from these two images will be helpful in computer assisted surgery and radio surgery for better diagnosis and treatment.

Hence, for better understanding of the scene there is a need to combine essential visual information present in the source images for obtaining a meaningful image. Image fusion is a phenomenon of integrating or combining or merging visually significant or complementary information of several co-registered source images of a scene into a single image. The resulting image provides more accurate description about a scene than any one of the individual source images.

### 1.3 TYPES OF IMAGE FUSION SYSTEMS

There are two categories in image fusion systems which are Single-Sensor Image (SSI) fusion system and Multi-Sensor Image (MSI) fusion system. In SSI fusion system, image sequences of the same scene are integrated into a single image, which are captured from the single sensor. In noisy environment and in some illumination conditions, human observers may not able to detect the objects of interest which can be easily found from the fused images of that targeted scene. Digital photography applications such as Multi-Focus Imaging and Multi-Exposure Imaging come under SSI fusion system. However, SSI fusion system is not without flaws. They depend on conditions like illumination and dynamic range of the sensors. For example, Visual-Inertial (VI) sensor like digital camera can capture visually good images in high illumination conditions. However, they fail to capture good visual images under low illumination conditions such as night, fog and rain.
Hence, MSI fusion system is introduced to restrict the drawbacks of SSI fusion system by capturing the images in adverse environment conditions.

**Figure 1.3: Single-Sensor Image Fusion System**

**Figure 1.4: Multi-Sensor Image Fusion System**
In MSI fusion system, multiple images of the same scene are captured using various sensors of different modalities. For example, VI sensors are good in high lighting conditions. However, IR sensors can capture images in low lighting conditions. Required and necessary information from these images are combined into a single image by the process of fusion. Applications such as medical imaging, military, navigation and concealed weapon detection fall under this category.

Advantages of MSI fusion systems are:

- **Reliable and accurate information:** This type of fusion systems provides reliable and accurate description of the scene compared to source images.

- **Robust performance:** Even if one sensor of fails, this fusion system generates a composite image by considering the redundant information of all sensors. Hence, it is robust.

- **Compact representation:** Fused image of this system is compact and provides all the necessary information of source images in a single image.

- **Extended operating range:** The range of operation is extended by capturing images at different operating conditions.

- **Reduced uncertainty:** Combined information of different sensors reduces the uncertainty present in the individual captures of the scene.

### 1.4 FUNDAMENTAL STEPS IN IMAGE FUSION

Figure 1.5 shows the steps involved in the process of image fusion system. There are 5 main steps. 1) Pre-processing. 2) Image registration. 3) Image fusion. 4) Post-processing and 5) Fusion performance evaluation.
**Figure 1.5:** Block diagram of the steps involved in Image Fusion system

*Pre-processing:* In this step, mitigation of the noise or artifacts is done which are introduced during the acquisition of image.

*Image registration:* It is the process of aligning or arranging more than one image of a same scene according to a geometrical co-ordinate system. In this process, one of the source images is considered as a reference image and then geometric transformation is applied on the remaining source images. The alignment is done with respect to the reference image.

*Image fusion:* Fusion process can be performed at three levels: Pixel, Feature and Decision. Pixel level fusion is done on each input image pixel by pixel. Fusion procedure that is performed at feature level on the input image features is called Feature level fusion. If it is executed on probabilistic conclusion of local information, it is called Decision Level Fusion, which is successively formulated from the features evoked. The approaches based on feature level are more desirable for fusion in contrast to other methods because of their effectiveness and ease of development.

*Performance analysis:* During the fusion process, some required information of source images may be lost and visually unnecessary information or artifacts may be introduced into the fused image. Hence, fusion algorithms are needed to be assessed
and evaluated for better performance. This performance analysis can be carried out by evaluating them qualitatively by visual inspection and quantitatively by using fusion metrics.

Post-processing: In this stage, obtained fused images are further processed based on the application requirement. This processing may involve Recognition, Segmentation, Classification and Feature Extraction.

In this thesis, new registration approaches and feature level image fusion algorithms are developed for medical images.

1.5 TEXTURE ANALYSIS AND CHARACTERIZATION

Texture is an important feature of images which helps to classify the information by segmenting the features based on the region of interest. Texture analysis provides the details of the spatial distribution of the intensity levels in a neighbourhood. Texture, usually is a repeating pattern of intensity levels according to the local variations in the neighbourhood of pixels. Texture analysis can be done in two steps:

a. Texture Segmentation: This step deals with defining the boundaries between various regions of texture in an image.

b. Texture Classification: This step is concerned with the classification of the segmented textures based on their characteristics. In this, statistical methods like Gray Level Co-Occurrence Matrix (GLCM) are involved.

Different approaches for doing the texture analysis are available. Three methods are listed below:
a. **Structural:** In this method, texture is treated as a set of primitive texture elements sometimes called as texels. The texels would be organized in a regular or repeated form.

b. **Statistical:** Here, texture is considered as a set of feature vectors. These vectors are the quantitative measure of the arrangement of pixel intensities in a region.

c. **Modelling:** In this, constructive models which define the textures are involved.

### 1.6 ORGANIZATION OF THE THESIS

Remaining chapters of the thesis are outlined as follows:

Chapter-2 briefs the classification of image fusion literature. Spatial domain techniques are briefly reviewed. Various multi-scale fusion methods namely Pyramid, Wavelet and Edge Preserving Decomposition methods are discussed. Comments on the merits and demerits of various fusion methods are provided. Motivation for proposing new methods is presented.

Chapter-3 presents a novel medical image fusion algorithm using Non Linear Anisotropic Filtering (NLAF) in Principal Component Analysis (PCA) domain, which preserves the texture information of fused images most effectively. NLAF is utilized to decompose the source images into approximation and detail layers. Final detail and approximation layers are computed with the support of PCA. Finally, fused image is generated from the linear combination of final detail and approximation layers. Qualitative and Quantitative performance of NLAF-PCA algorithm is assessed with
the help of image quality metrics like Peak Signal to Noise Ratio (PSNR), Correlation Coefficient (CC), Entropy (E), Root Mean Square Error (RMSE) and Structural Similarity Index Measurement (SSIM).

Chapter-4 describes the medical image registration and fusion using Active Slope Meagerness (ASM) regularizer with Statistics Based Steered Image Filtration (SSIF). This focuses on simultaneous registration during the fusion procedure. Initially, the MR image is focalized to enhance resolution, so that precise registration of images is obtained. At the same time, fusion of images can be done precisely by extinguishing the misalignment gradually. It iteratively persist these two procedures until the convergence. Further, performance of fusion is assessed both qualitatively and quantitatively. The comparative analysis with the traditional and recent image fusion methods is also included.

In chapter-5, a novel MR and CT image fusion approach is proposed which utilizes the Deep Learning Convolutional Neural Networks (CNNs) with Statistics Based Steered Image Filter (SSIF) is presented, where a deep learning convolutional network is adopted to generate a weight map which integrates the pixel activity information from MR and CT images. The fusion process is conducted via SSIF fusion rule which computes the weights of obtained detail layers using image statistics. Further, weighted average method is utilized to obtain the fused image. Experimental results also demonstrated that this method obtained promising results in terms of both visual quality and objective assessment.

In chapter-6, a texture analysis method for the input image obtained from the output of a Fusion algorithm is proposed. The Nonlinear Anisotropic Filtering in PCA
domain [16] (NLAF-PCA) fusion algorithm is used to generate the fused image. The MR medical images data set consisting of both Tumorous Image (TI) and Non-Tumorous Image (NTI) is considered. The NLAF-PCA method effectively aligns the TI and NTI of the applied input data.

The fused data obtained helps to analyse the tumour information effectively. The features are extracted from the fused images using Discrete Wavelet Transform (DWT). The extracted features are trained using Support Vector Machine (SVM) classifier. Finally, by using the segmentation the tumour is extracted.

Finally, chapter-7 concludes the thesis and gives directions for the future research in the area of Image Registration and Fusion.
REFERENCES


