Spring 2005

Planar visual fusion scintigraphy

Jacob Puthiamadathil
New Jersey Institute of Technology

ABSTRACT

PLANAR VISUAL FUSION SCINTIGRAPHY

by
Jacob Puthiamadathil

Planar scintigraphy, while providing useful information about the distribution of the radiopharmaceutical being used, often does not provide adequate information about the surrounding anatomical structures. A proposed method to solve this problem is by registering and fusing a scintigraphic image with a digital visual image. Fusing planar scintigraphic images with visual photographic images to supply an anatomic correlate to regions of radiopharmaceutical accumulation has been explored in this study. The digital visual image will provide a context for the relevant structures in the scintigraphic image. The validity of using one set of affine transformation parameters for registering all patient data has been confirmed with preliminary data. Results using patient data confirm registration accuracy to within .95 cm on average and 2.35 cm as the greatest amount of error. Error was approximated for an optimal imaging radial distance from the optical axis as being 1.70 cm. Initial clinical imaging has subjectively been very useful in low background applications such as lymphoscintigraphy, whole body I-131 NaI imaging for thyroid cancer and In-111 WBC infection imaging.

PLANAR VISUAL FUSION SCINTIGRAPHY

by
Jacob Puthiamadathil

A Thesis
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Computer Engineering

Department of Electrical and Computer Engineering

May 2005

APPROVAL PAGE

PLANAR VISUAL FUSION SCINTIGRAPHY

Jacob Puthiamadathil

Dr. Atam P. Dhawan, Dissertation Advisor — Date
Professor of Electrical and Computer Engineering, NJIT

Dr. Yun-Qing Shi, Committee Member — Date
Professor of Electrical and Computer Engineering, NJIT

Dr. Lionel S. Zuckier, Committee Member — Date
Professor of Radiology, UMDNJ

BIOGRAPHICAL SKETCH

Author: Jacob Puthiamadathil

Degree: Master of Science

Date: May 2005

Undergraduate and Graduate Education:

- Master of Science in Computer Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2005

- Bachelor of Science in Electrical Engineering,
  Stevens Institute of Technology, Hoboken, NJ, 2002

Major: Computer Engineering

Presentations and Publications:

Zuckier LS, Puthiamadathil JM and Stabin MG. Web-based calculator for counseling patients on potential triggering of homeland-security radiation detectors. Accepted for oral presentation, 52nd Annual Society of Nuclear Medicine Meeting, Toronto, Canada, June 20, 2005

Zuckier LS, Puthiamadathil JM, Dhawan AP, Ghesani NV and Vazirani T. Scintigraphic/visual fusion imaging for improved localization of thyroid abnormalities: method, proof of principle and preliminary evaluation. Accepted for poster presentation, 52nd Annual Society of Nuclear Medicine Meeting, Toronto, Canada, June 20, 2005.

Zuckier LS, Puthiamadathil JM, Vazirani T and Dhawan AP. Scintigraphic/visual fusion imaging for improved localization of abnormalities on planar scanning. 18th Annual Northeast Regional Scientific Meeting of the Society of Nuclear Medicine, October 22 - 24, 2004, Stamford, Connecticut.

Puthiamadathil JM, Zuckier LS and Dhawan AP. Planar visual fusion scintigraphy (poster). 26th Annual International Conference IEEE Engineering in Medicine and Biology Society (EMBS). September 3, 2004.

To my parents and my brother

ACKNOWLEDGMENT

I would like to express my deepest gratitude to Dr. Atam Dhawan and Dr. Lionel Zuckier, who served as my research advisors. Special thanks are given to Dr. Yun-Qing Shi for providing help.

My fellow graduate students in the Intelligent Systems and Image Analysis Laboratory are deserving of recognition for their support and encouragement.

TABLE OF CONTENTS

Chapter	Page
1 INTRODUCTION	1
2 METHODOLOGY	4
2.1 Overview of Methodology	4
2.2 Image Acquisition and Preparation	8
2.3 Affine Transformation	9
2.4 Image Fusion	11
3 BACKGROUND	15
3.1 Photographic Imaging System	15
3.2 Overview of Nuclear Medicine Imaging	17
3.3 Anger Camera and Scintigraphic Images	18
3.4 Single Photon Emission Computed Tomography	22
3.5 Introduction to Multi-Modal Registration	24
3.6 Weighted Geometrical Features Based Registration	26
3.7 Elastic Deformation Based Registration	27
3.8 Overview of Image Fusion	29
4 RESULTS	32
4.1 Registered and Fused Examples	32
4.2 Measurement of Error.	34
4.3 Error of a Single Point	41
5 DISCUSSION AND CONCLUSIONS	44

TABLE OF CONTENTS
(Continued)

Chapter **Page**

6 REFERENCES..... 46

LIST OF TABLES

Table		Page
4.1	The x and y Position of the Centroid of the Points at -27.5 cm	35
4.2	The x and y Position of the Centroid of the Points at -10 cm	35
4.3	The x and y Position of the Centroid of the Points at -27.5 cm	35
4.4	The Pythagorean Distance Between Points at -27.5 cm	36
4.5	The Pythagorean Distance Between Points at -10 cm	36
4.6	The Pythagorean Distance Between Points at 7.5 cm	36
4.7	Ratios of the Lengths of the Identical Line Segments at -10 cm and -27.5 cm	37
4.8	Ratios of the Lengths of the Identical Line Segments at 7.5 cm and -27.5 cm	37
4.9	Ratios of the Lengths of the Identical Line Segments at 7.5 cm and -10 cm	37
4.10	Affine Transformation Parameters for "Generalized" and Patient-Dependent Transformations	39
4.11	Comparison of Generalized Transformation Parameters and Patient Dependent Transformation	39
4.12	Mean and Standard Deviation of Error For Affine Transformation Parameters	40
4.13	Comparison of Error for Generalized Affine Transformation	41
4.14	Distances from Midpoint to Four Outer Points at Different z Distances	42
4.15	Error as a Function of z Displacement	42

LIST OF FIGURES

Figure		Page
1.1	Annual frequency of PubMed citations retrievable based on the keywords 'image fusion' and 'radionuclide imaging'.....	1
1.2	Left Panel: Whole body ^{131}I-NaI studies in a patient with papillary thyroid carcinoma in the region of the thyroid bed (arrow). Because of minimal background activity, anatomic landmarks are sparse. Right panel: ^{111}In-WBC localization in osteomyelistis of the calcaneus (white arrow). The technologist has placed a hot marker over the calcaneus to add an anatomic reference (gray arrow).....	2
2.1	Schematic of a gamma camera. The gantry, bolted to the floor, supports two detectors that collect images from the patient. A moveable pallet translates the patient in and out of the camera.....	4
2.2	Visual camera mounted above gamma camera. Gamma camera head is in horizontal plane for scintigraphic view (left panel) and is rotated to the vertical plane for visual camera acquisition (right panel). Field of view of the visual camera (arrows) is centered on the scintigraphic field of view.....	5
2.3	Images of the calibration phantom by scintigraphic (upper left) and visual (upper right) cameras. On the phantom, points of activity are located at spatial coordinates (0", 0"), (0", 12"), (12",0") and (10",10"). Based on the affine transformation derived from the first 3 points, the registered visual image appears in the lower left panel while the fused image appears on the lower right. The 4th point (arrows) can be used as an internal measure of accuracy. Note the excellent registration of all four points	7
3.1	Weak Perspective Projection.....	15
3.2	Top is spherical aberration. Bottom is chromatic aberration.....	17
3.3	Schematic of Anger Camera.....	19
3.4	Setup of photomultiplier tubes.....	21
4.1	This is a planar scintigraphic image, originally saved in the DICOM format, of In-111 labeled white blood cells localizing in a painful left ankle. The regions that have the highest levels of radioactivity appear whiter in color.....	32

LIST OF FIGURES
(Continued)

Figure		Page
4.2	Digital visual image of the identical view.....	32
4.3	Digital visual image registered using the affine transform and then aligned with the scintigraphic image.....	33
4.4	Fused Image. The hue in this image contains the scintigraphic intensity. The regions with the least radioactivity are red in hue and increased levels of radioactivity are coded as yellow, green, blue, and magenta, respectively.....	33
4.5	The x and y positions of the 5-point phantom images at distances of –27.5 cm (furthest from the photographic camera), -10 cm and 7.5 cm (closest to the photographic camera).....	34

CHAPTER 1

INTRODUCTION

Scintigraphic images produced by a gamma camera contain valuable and unique information about the distribution of the radiopharmaceutical being imaged. In certain circumstances, the activity levels in surrounding normal structures are low, and thus abnormalities appear as "hotspots" without context. For this reason, in the realm of tomographic imaging, successful efforts have been made in the past few years to combine tomographic scintigraphic information (PET and SPECT) with CT volumetric data (1-4). The idea of merging anatomic with molecular image information has been shown to be intuitively correct and clinically meaningful (5), resulting in increasing references in the medical literature (see Figure 1.1).

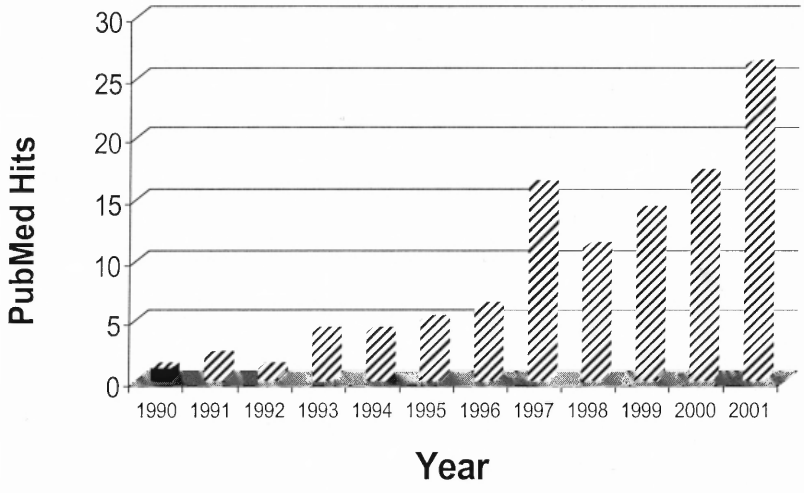

Figure 1.1 Annual frequency of PubMed citations retrievable based on keywords 'image fusion' and 'radionuclide imaging'.

One area in nuclear medicine imaging where fusion with anatomic modalities has not penetrated is planar scintigraphy. Currently available methods to obtain anatomic references for scintigraphic images (such as shadowgrams made by placing a sheet of activity behind the patient, or hot markers placed on anatomic landmarks) are time-consuming and crude (See Figure 1.2). A method of registering and fusing planar scintigraphic images with visual images has been implemented and tested. Combined 2-D visual-scintigraphic imaging would have several advantages over SPECT-CT imaging of the region in question. It would be substantially more available than SPECT-CT imaging, more cost effective, and result in less radiation to the patient.

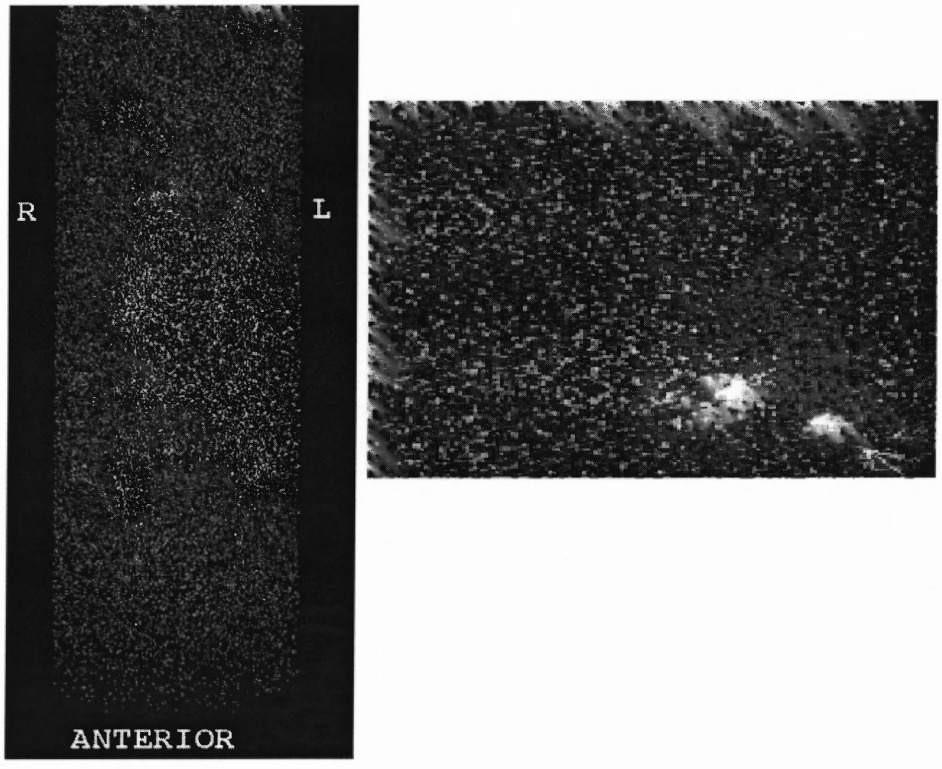

Figure 1.2 Left Panel: Whole body ^{131}I-NaI study in a patient with papillary thyroid carcinoma in the region of the thyroid bed (arrow). Because of minimal background activity, anatomic landmarks are sparse. Right panel: ^{111}In-WBC localization in osteomyelitis of the calcaneus (white arrow). The technologist has placed a hot marker over the calcaneus to add an anatomic reference (gray arrow).

Planar fusion imaging would be especially valuable in "hot-spot" imaging. Hot spot imaging is when certain tissues take up radiopharmaceuticals at a much greater degree than the surrounding tissues. These regions show up with high contrast on scintigraphic images, however anatomic information is often sparse. Examples of hot spot imaging include imaging with ^{111}In-WBC for infection, ^{131}I-NaI for thyroid cancer and ^{99m}Tc-colloid for lymphoscintigraphy.

The goal of this project is therefore to improve the ability of nuclear medicine practitioners to accurately localize sites of abnormal signal by combining digital visual images with planar scintigraphic images.

The proposed method consists of registering planar radionuclide scintigraphic with digital visual (photographic) images obtained from a digital camera located at a fixed geometric location relative to the gamma camera and then fusing these images to provide a composite visualization, which includes both scintigraphic and anatomical information.

CHAPTER 2

METHODOLOGY

2.1 Overview of Methodology

Gamma cameras in clinical use are highly regulated by the FDA, and adding additional equipment to the gamma camera heads without major reengineering is not feasible due to limited attachment areas and clearance. The gamma camera heads are equipped with sensors that stop motion upon contact, and these cannot be easily circumvented. While it might be preferable to attach the digital camera directly to the gamma camera head in a commercial version of the device, in the initial study, a digital camera (2.11 Megapixel Nikon Coolpix 850) was attached 255 cm above the floor and directly above the gamma camera head. See Figure 2.1.

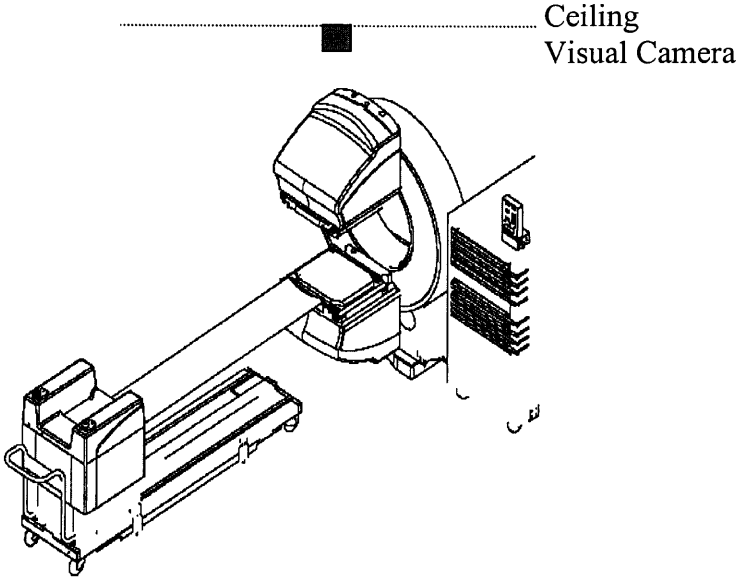

Figure 2.1 Schematic of gamma camera. The gantry, bolted to the floor, supports two detectors that collect images from the patient. A moveable pallet translates the patient in and out of the camera. A visual camera has been suspended from the ceiling and projects directly downwards over the gamma camera field-of-view.

The camera can be remotely controlled by public domain software (www.math.ualberta.ca/imaging/photopc.exe), which allows for control of zoom, resolution, flash, exposure, and image retrieval over a serial cable. These instructions have been simplified by means of scripted and menued batch commands.

All patients have given informed consent to participate in this study, and when a region of interest is identified (based on *a priori* information), a spot view of the area is scintigraphically imaged with the upper gamma camera head oriented parallel to the floor, above the patient lying on the gantry. While the patient remains immobile, the gamma-camera heads are rotated by 90 degrees, and the pallet is raised such that the area of interest is a fixed distance (128 cm) from the ceiling camera, facilitated by a hand-held gauge that is used to estimate offset from the floor. At this distance, the area of interest is photographed using the digital camera. See Figure 2.2.

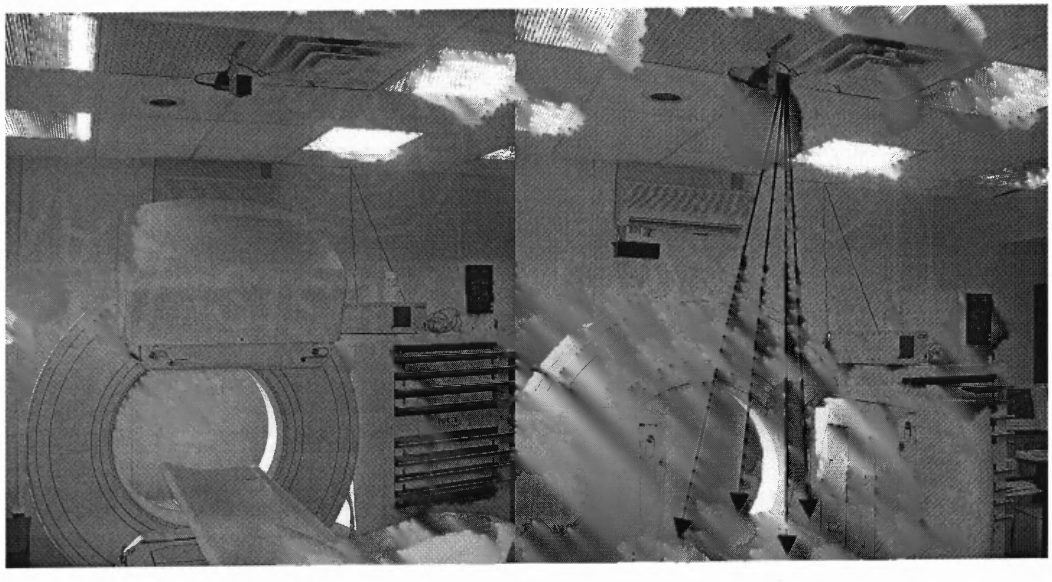

Figure 2.2 Visual Camera mounted above gamma camera. Gamma camera head is in horizontal plane for scintigraphic view (left panel) and is rotated to the vertical plane for visual camera acquisition (right panel). Field of view of the visual camera (arrows) is centered on the scintigraphic field of view.

The photographic images are written to the hard drive in a JPEG format as a 1600x1200 pixel matrix with an 8-bit RGB format that is subsequently converted to grayscale. Scintigraphic images are stored as 256x256 matrices with unsigned 8 bit integers in the DICOM format and exported to PC by ftp.

In the initial implementation and for the data analysis offered in this report, imaging of a 3-point phantom was done using both the gamma camera and the photographic camera. The points consisted of Tc-99m markers that are clearly visualized in both gamma camera and photographic camera images. Matlab (6) was used to find the centroids of each of these points in the photographic and scintigraphic domains based on the following protocol. The images were first smoothed using an averaging filter. A 5x5 filter was used for the photographic images and a 3x3 filter was used for DICOM images. Then the images were thresholded to separate the points from the background, using OTSU threshold levels. 8-connected objects were then selected (the points in question) and the centroids of these objects were obtained. Using these three points in the visual domain and three points in the scintigraphic domain, the parameters of an affine transformation were obtained by mapping the points in the visual domain to the points in the scintigraphic domain (See Figure 2.3). Once the transformation was defined, it may be applied to the visual image as a whole.

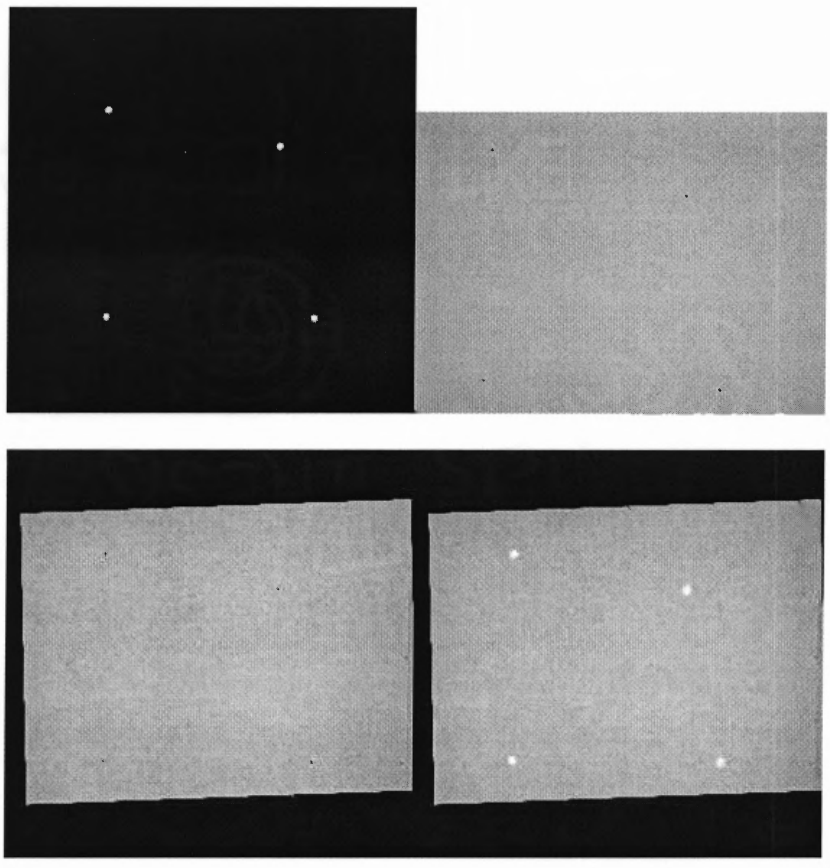

Figure 2.3 Images of the calibration phantom by scintigraphic (upper left) and visual (upper right) cameras. On the phantom, points of activity are located at spatial coordinates (0",0"), (0",12"), (12",0") and (10",10"). Based on the affine transformation derived from the first 3 points, the registered visual image appears in the lower left panel while the fused image appears on the lower right. The 4th point (arrows) can be used as an internal measure of accuracy. Note the excellent registration of all four points.

In subsequent implementations, ^{57}Co markers, which are easier to manipulate and use, have been employed. With maturation of the technique, other changes that were introduced include the usage of the open source Java program ImageJ (7) and two plugins, TurboReg (8) and the Image Layering Toolbox (9) that were used to perform image registration and fusion. Centroid locations were obtained through visual inspection for reasons of speed and convenience.

Testing of robustness of registration was performed in several ways. The affine transformation defined for a certain plane using three markers was compared to a 4^{th} point also on that plane. The transformation that was defined by the three points was performed on the entire photographic image, which contained an image of a 4^{th} point as well and this was compared to the actual location of the 4^{th} point in the gamma camera image.

Testing of the imaging system was performed to ensure that it was linear and spatially invariant. The phantom was centered on the patient bed and the bed height was changed so that the phantom was at three different distances from the photographic camera. The phantom was photographed at these distances. These measurements were used to predict how much the affine transformation parameters could be expected to change due to imprecision in the distance between the camera and the patient bed.

The imaging system was also tested under actual clinical conditions. During a period of several months, a generalized affine transformation which was defined on a certain day, was compared to patient specific affine transformations defined for certain patients. The affine registration parameters for the generalized affine transformation were compared to the specific patient-based affine transformations.

2.2 Image Acquisition and Preparation

Images of the phantom from the photographic camera were acquired as 1600x1200 8 bit RBG JPEGs. Images of the same phantom from the gamma camera were saved as 256x256 8 bit grayscale images in the DICOM format. OTSU threshold levels were determined to separate image of the points from the background image in both

photographic and gamma camera images. The threshold categorized the images into 0 and 255 pixel intensities. The 8-connected objects were chosen which corresponded to the images of the points in question. Centroid calculations were performed on these connected objects having the centroid be the center of mass of these objects, calculated:

$$C_x = 1/N \sum_{1}^{N} X_i \text{ and } C_y = 1/N \sum_{1}^{N} Y_i$$

where Cx and Cy are the x and y coordinates of the centroid of the selected object. N is the number of pixels found, and Xi and Yi are the coordinates of the pixel edges (10).

2.3 Affine Transformation

The affine transformation is a special case of the rigid body transformation that includes translation, rotation and scaling operations (11). Fiducial markers were used to define planes in both the photographic image domain and the gamma camera image domain. Three points were used to define the respective planes. The affine transformation can be written as **x'**=A**x.**

In expanded form the transformation can be written as:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & p \\ 0 & 1 & q \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

The equations can also be written in the form:

$$x' = ax + by + c$$

$$y' = dx + ey + f$$

where x,y were the coordinates of the fiducial marker in the input image domain and x',y' are the coordinates of the same fiducial marker in the output image domain, since the affine transform parameters are unknown.

For the study, the photographic domain was transformed to the gamma camera domain. Since there are six unknowns (a,b,c,d,e, and f), six equations or 3 points were needed to find the parameters of the transformation. In matrix form, the relationship can be written as:

$$\begin{pmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \\ x_3' \\ y_3' \end{pmatrix} = \begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_2 & y_2 & 1 \\ x_3 & y_3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_3 & y_3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix}$$

It can be shown that the 6x6 matrix containing the initial x,y positions of the object to be transformed is full rank, or rank of 6 since the three points selected in both image domains were carefully selected so as to not be collinear. In this case, the inverse of the 6x6 matrix exists. The column vector with the affine transform parameters can be solved for in this manner:

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_2 & y_2 & 1 \\ x_3 & y_3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_3 & y_3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1' \\ y_1' \\ x_2' \\ y_2' \\ x_3' \\ y_3' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix}$$

Once the parameters were obtained, the affine transformation was applied to the entire photographic image which had a resolution of 1600x1200 and was previously converted into a grayscale image using the formula, I=(R+B+G)/3. The intensity of the

pixels in the transformed photographic image was interpolated using bilinear interpolation. Bilinear interpolation is an extension of linear interpolation used for interpolating functions of two variables (12). Since the original pixels values will tend to lie on fractional x and y values in the transformed image, bilinear interpolation uses a weighting of the four nearest neighbors to come up with a value for the integer grid values in the transformed image.

2.4 Image Fusion

In the study, the final implementation was such that image fusion was performed within the ImageJ environment. The Image Layering Toolbox (ILT) plugin was used to fuse the gamma camera image and the registered photographic image. There are two different layering options available within the ILT plugin: contour and colorburn. The colorburn mode is the default mode, but the contour mode was used to modify the gamma camera image.

For the colorburn mode, a colormapper was defined by an array of object Colors with a size of 256. The array was called mLUT. The array mLUT[g], for g=0...255 was populated using the equation:

mLUT[g] = new Color (mC1.getred() + mag*mCDiff[0], mC1.getgreen() + mag*mCDiff[1], mC1.getblue() + mag*mcDiff[2])

The variable mC1 was the first Color selected by the user, by default white. mC1.getred() refers to the integer value of the red component for mC1 and likewise for

green and blue, and mCDiff[0] refers to the integer difference in red between mC1 and mC2 (which was by default black but could also be modified), mCDiff[1] refers to the difference in green between mC1 and mC2, and mCDiff[2] refers to the difference in blue between mC1 and mC2. mag is defined as:

$$\sqrt{cx.getred()^2 * cx.getgreen()^2 * cx.getblue()^2} / \sqrt{3 * 255^2}$$

The Color cx was defined as cx=new Color (g,g,g) for g=0...255. Thus mag ranged in value from 0 to 1. The array mLUT basically takes an input in the form of mag and maps it to a line with y-intercept of mC1.getred, mC1.getgreen, or mC1.getblue and with the slope of Δred, $\Delta green$, or $\Delta blue$, respectively. If the initial defaults for mC1 and mC2, which are white and black, respectively, this colormapper can be used to convert an RGB image into an 8 bit grayscale image.

Mapping into the colormapper is done by reading the color at a particular pixel and converting it into the Color class. This new Color object called cx is then normalized using Equation above and then inserted into the mLUT array as follows mLUT[mag*255].

The contour mode's colormapper is defined as:

$$\text{mLUT[i-1]}= ((huval - ((huval - hlval) * (i / depth)),1,1)$$

The array mLUT is once again an array of Color objects with a depth of 256. The color, this time, is defined in the HSB domain. The variable huval refers to the upper hue value,

which is set to .7. The lower hue value, hlval is set to 0. The variable i runs from 1 to 256. mLUT[0] is defined as Color (0,0,0), which is transparent, but for all other indices of the array, the saturation and brightness were set to 1 and 1 respectively on a 0 to 1 scale. This is another linear mapping of the input hue into a different set of hues. The y-intercept is hovel and the slope equals (huval-hlval)/depth.

Mapping into the contour mode colormapper is performed by first reading the pixel value in the image to be mapped (which is the variable f) and then normalizing it based on this formula:

$$x=(f-mMin)/(mMax-mMin)$$

The value of x varies from 0 to 1. Then this fraction was multiplied by 256 to access the appropriate index in the mLUT array.

The registered photographic image was selected as the bottom layer and it was automatically converted into a colorburn image. The gamma camera image, after it was converted into a contour mode image was added to the layer below it. The addition of these images was done with the Porter-Duff equations.

In the trivial case where:

```
s_rgba_arr[0]==0 && s_rgba_arr[1]==0 && s_rgba_arr[2]==0,
        new_rgba_arr[0]=destination_rgba_arr[0]
        new_rgba_arr[1]=destination_rgba_arr[1]
        new_rgba_arr[2]=destination_rgba_arr[2]
```

otherwise:

new red=destination red + source alpha * (source red-destination red)

new green= destination green + source alpha * (sourcegreen-destinationgreen)

new blue = destinationblue + source_alpha * (sourceblue-destinationblue)

The new image consists of a red, green, and blue channel. This is the new_rgba_arr array. To obtain the new image, the destination image (the image that is already in place), is added with the source image multiplied by the source image's alpha value. This variable is source_alpha. By default the alpha value is .7 on a floating point scale from 0 to 1.

Because there are extreme ranges of activity in hotspot imaging which are best portrayed after log transformation, an additional fusion mode was created. The natural log of the gamma image's pixel intensities were taken before the image was color coded with the existing contour colormapper and added to the previous layers with the same equations.

Essentially the new equation is:

x=(log(f)-log(mMin))/(log(mMax)-log(mMin))

The value of x is a float but when it is multiplied by 256 to get the value in the mLUT array, it is converted to an integer, thereby a binarization process is occurring.

CHAPTER 3

BACKGROUND

3.1 Photographic Imaging System

A photographic camera can be approximated by a pinhole imaging system (13). Assuming that all the points of interest lie in a plane equidistant from the camera, the relationships between a point in the object domain (x,y) and the image of that point in the imaging domain (x',y'), are governed by x'=-mx, and y'=-my, where m= -(f'/z0). f' is the distance from the optical center to the image plane along the k direction. –z0 is the distance along that same direction from the optical center to the object domain (9).

If there are two points P and Q and their images P' and Q' the vectors PQ and P'Q' are parallel and we also have |PQ| = m|PQ|, See Figure 3.1. If the object plane is moved a further distance from the optical center, all of the line segments in the image plane will be magnified by a factor m. It can be assumed that the image has been scaled uniformly.

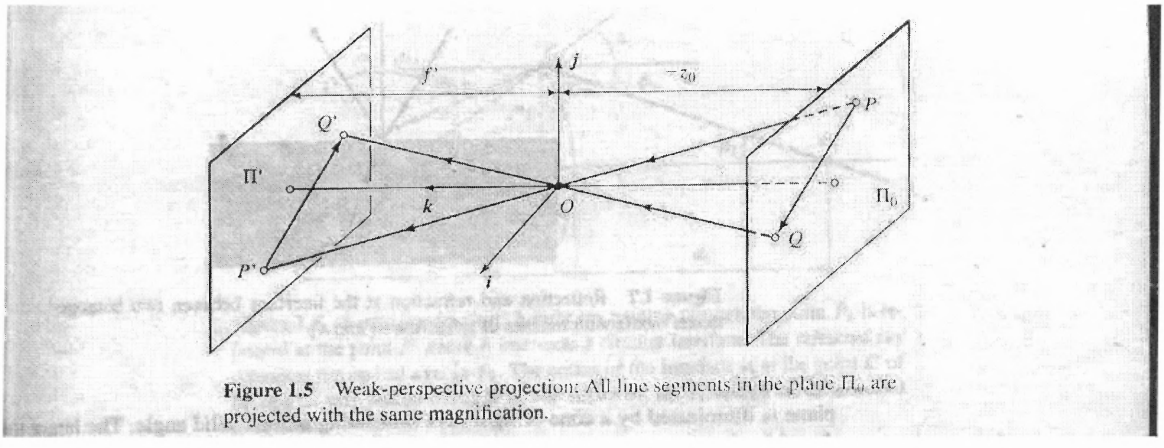

Figure 1.5 Weak-perspective projection: All line segments in the plane Π_0 are projected with the same magnification.

Figure 3.1 Weak Perspective Projection (13).

The thin lens equation can also be used to model the imaging process of a camera. The equation governing the thin lens model is $1/z' - 1/z = 1/f$. Where -z is the distance from the lens to the object and z' is the distance from the lens to the image and f is the focal length of the lens. An affine transformation would not be able to perfectly describe an image where the object plane is at an angle to the lens. In this situation, perspective error would be introduced as the magnifications would be variable. For most applications, it is not a given that the object or region of interest will lie on a plane parallel to the lens in the camera.

Even if the object to be imaged were to be located in one plane parallel to the lens in the camera, spherical and chromatic aberrations would cause errors in the registration process. Spherical aberration is defined as a point in the object domain, P, would appear as a circle of confusion centered at P' (where P' is the paraxial image of P) in the image. Spherical aberration that lies off the optical axis of the camera is called coma. Coma would cause errors in the registration process. Chromatic aberration could also potentially be a problem. Chromatic aberration occurs because different wavelengths of light have different focal lengths, thus making locating the centroid of points inaccurate. This is a fiducial location error. (See Figure 3.2).

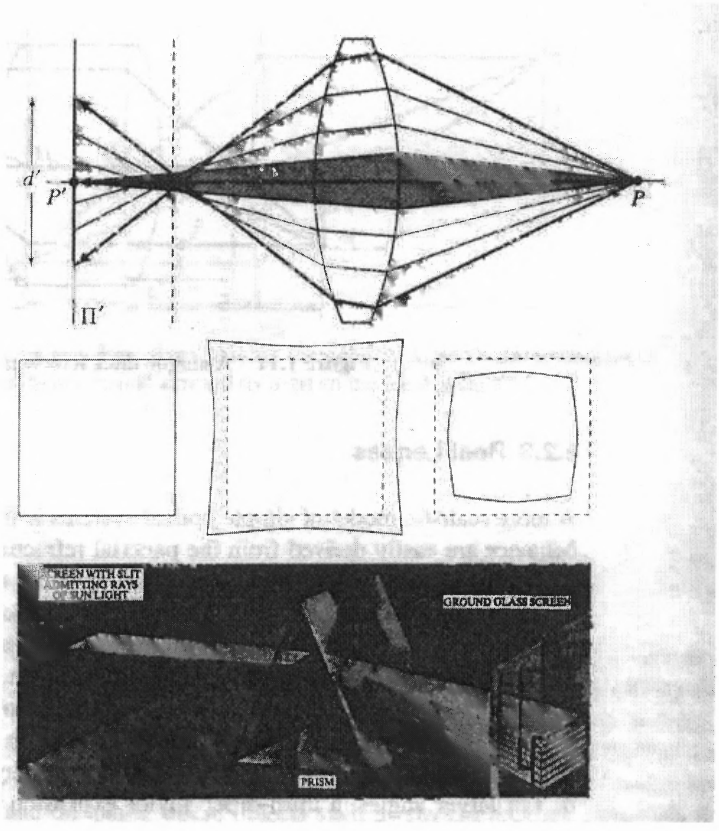

Figure 3.2 Top is spherical aberration. Bottom is chromatic aberration (13).

3.2 Overview of Nuclear Medicine Imaging

Anatomical imaging modalities such as planar X-ray and Computed Tomography are limited to structural information about the body and generally do not provide metabolic or functional information about an organ or tissue (14). This information is important because it provides valuable insight into physiological functioning. Radionuclide imaging methods directly involve an organ and associated tissue in the body in a way that the organ itself becomes a source of radiation that is used in the imaging process. These methods are called emission imaging methods and primarily utilize radioactive decay. In the process of radioactive decay, an unstable nucleus decays into a stable nucleus by the

release of nuclear energy and by emitting photons and/or specific particles such as positrons, electrons and alpha particles.

Radioactive decay can be described as an exponential process with respect to time:

$$N(t) = N(0)e^{-\eta t}$$

N(0) is the initial number of radioactive atoms, N(t) is the number at time t and η is the radioactive decay constant.

The half-life of a radionuclide, T_{half} is defined by the time required for half of the radionuclides to transform, and can be expressed as:

$$T_{half} = .693/\eta$$

3.3 Anger Camera and Scintigraphic Images

In the Anger camera, a γ-ray image is projected by the collimator onto the NaI(Tl) detector crystal, which creates a pattern of scintillations in the crystal that reflects the distribution of the radioactivity in the front of the collimator (15). Photomultiplier tubes that are attached to the collimator amplify the light signal while electronic position logic circuits help determine the location of each scintillation event as it occurs in the crystal. Individual scintillation events are analyzed for energy by pulse-height analyzer circuits. If an event falls within the selected energy window, indicating a non-scattered event, the matrix position X-Y is incremented by one, which corresponds to the location in the crystal where the event occurred (See Figure 3.3).

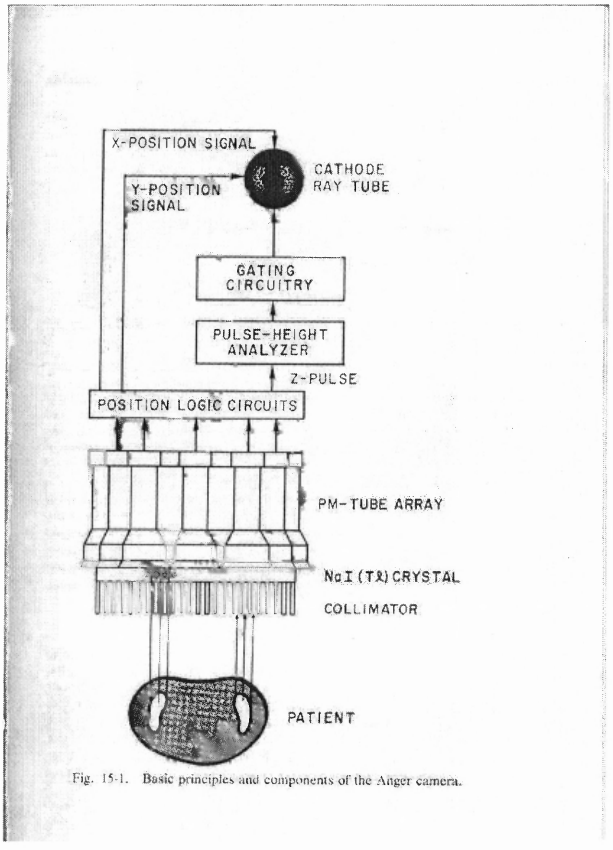

Fig. 15-1. Basic principles and components of the Anger camera.

Figure 3.3 Schematic of Anger Camera (15).

Gamma rays cannot be focused, so a lens principle similar to photography cannot be used. Most γ-ray imaging systems employ the principle of absorptive collimation for image formation. There are four basic types of collimators: pinhole, parallel-hole, diverging, and converging. A parallel-hole collimator was used for this application. The parallel-hole collimator is the most commonly used collimator in most imaging laboratories. Parallel holes are drilled or cast in lead or shaped from lead foils. The lead walls between the holes are called septa. Septa thickness is designed so that γ-rays do not cross from one hole into another. The parallel-hole collimator projects an image of the same size as the source distribution onto the detector. A fundamental relationship for

the collimator is that for a given septal thickness, collimator resolution is improved only at the expense of decreased collimator efficiency.

The Anger camera uses a single, large-area, NaI(Tl) detector crystal, usually 1.25 cm thick x 30-50 cm in diameter. Some low-energy radionuclide cameras, such as those used solely for ^{99m}Tc, have crystal thicknesses of only 6-8 mm. Relatively thin detector crystals are preferred for the Anger camera, even at the expense of less efficient radiation detection, because of their ability to provide better intrinsic resolution and some improvement in detail.

An array of photomultiplier (PM) tubes is coupled optically to the back face of the crystal with a silicone fluid or grease (See Figure 3.4). Most modern cameras use 37, 61, 75, or 91 tubes arranged in a hexagonal pattern. The PM tubes are split into X+, X-, Y+ and Y- regions, and depending on the strength of these signals, X and Y position signals are obtained which are used to position a recorded event in the proper matrix location. The equation for the X position signal is: $X = k(X^+ - X^-)/Z$ and similarly for the Y position signal, $Y = k(Y^+ - Y^-)/Z$. K is a scale factor adjusted for the size of the matrix. Z is a signal proportional in amplitude to the total amount of light produced by a scintillation event in the crystal. These equations do not describe a true mapping of source position because the PM tubes are not "point" detectors.

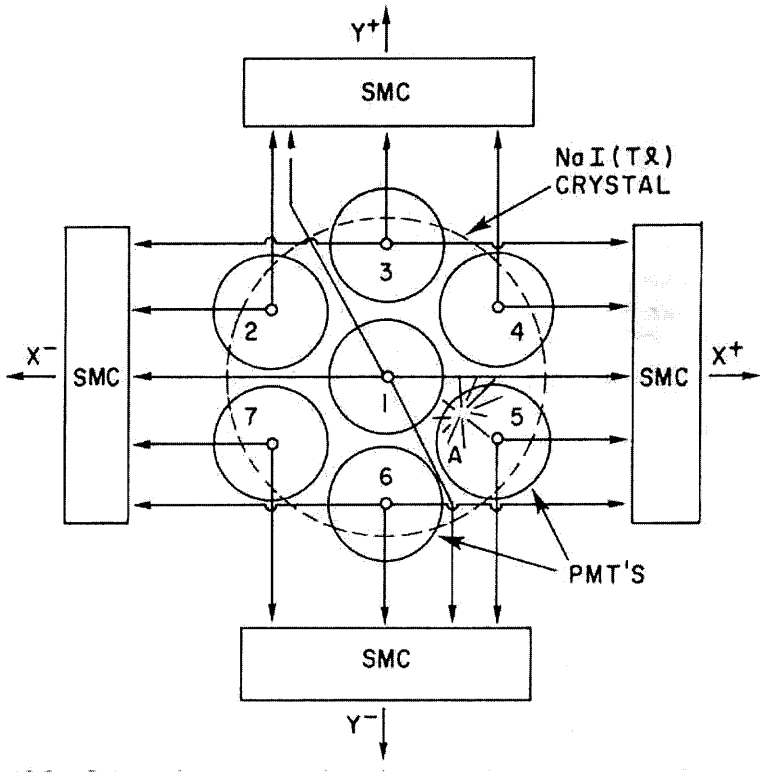

Figure 3.4 Setup of photomultiplier tubes (15).

Two basic types of problems that are found in the Anger camera are nonlinearities and nonuniformities. Nonlinearities occur when X- and Y-position signals do not change linearly with displacement of a radiation source across a PM tube. An example of this is when a source is moved from the edge of a PM tube towards its center, the light collection efficiency of the tube increase more rapidly than the distance the source is moved. The result is a pincushion distortion in areas of an Anger camera directly in front of PM tubes and barrel distortions between them. The other type of problem is nonuniform detection efficiency. This occurs mainly from small differences in the pulse-height spectrum for different PM tubes.

Some cameras incorporate microprocessor-based circuitry to correct for image nonuniformities. In one scheme, the test image is first normalized relative to the image

matrix element having the smallest number of counts (the coldest element matrix in the image). This element is assigned a relative intensity of 100. Other elements are assigned greater values, depending on their relative intensities in the uniform field image. In patient imaging, a certain fraction of the counts recorded in each element are thrown out, depending on the relative intensity of that area in the test image.

Recently, because of the development of economical microprocessors circuits, it has become feasible to address the nonlinearity problem as well. The camera is provided with a set of microprocessors that are used to store correction matrices for regional differences in pulse-height spectra and for position distortions. The correction matrices are programmed into the microprocessor using calibration measurements performed at the factory before installation of the camera. The calibrations involve precise measurements using line source of pulse-height spectra for individual PM tubes and position distortions over the entire region of the detector crystal.

3.4 Single Photon Emission Computed Tomography

In SPECT imaging, radionuclides are injected into the body via radiopharmaceuticals, which distribute into tissue, making the tissue a source of gamma rays (16). The gamma rays from the tissue pass through the body and are captured by the detectors surrounding the body to acquire raw data for defining projections. The projection data are then fed into reconstruction algorithms, to display images with the help of a computer. Commonly used radionuclides include, Technetium ^{99m}Tc, Iodine ^{123}I, and Gallium ^{67}Ga. These radionuclides decay by emitting gamma rays with photon energies ranging from 140 keV to 400 keV.

The attenuation of gamma rays can be expressed as: $I_d = I_0 e^{-\tau x}$, where I_0 is the intensity of the gamma rays at the source, I_d is the intensity at the detector after the gamma rays have passed a distance of x with a linear attenuation coefficient ν that depends on the medium and the energy of the gamma ray photons.

After the tissue or organ to be imaged becomes radioactive, the patient is positioned in a scanner with detectors placed surrounding the patient. Detectors may be arrayed in a ring surrounding the patient covering a full 360-degree range, however more commonly, detectors subtend only a small angle around the patient, but rotate during the scan to provide a complete sampling of angular data. Sodium iodide NaI(Tl) crystal-based scintillation detectors coupled with photomultiplier tubes are most commonly used for detection of gamma rays. Scattering and low photon statistics are major problems with SPECT imaging. Scattering events cause loss of source information, as it is difficult to identify the path of travel for the photon originating from the source. Lead-based collimators are used to reduce the detection of scattered events, while pulse-height energy windowing is used to identify non-scattered photons. These maneuvers improve the signal-to-noise ratio.

The nonlinearities caused by the detectors and the photomultiplier tubes make the SPECT imaging process spatially variant. Images would show degradations and geometric artifacts such as cushion and barrel artifacts because the registered detector count fluctuates from the actual count based on detector location. Intensity nonuniformities are corrected by using phantom data of uniform emission activity in the region of imaging. Pre-processing methods are needed to correct the geometric artifacts. Assessment of point-spread and line spread functions are useful in this regard.

The source of radiation is inside the body so gamma photons interact with the matter inside the body for attenuation and scattering processing, including photoelectric absorption and Compton scattering. Photoelectric absorption leads to the loss of a photon and Compton scattering can cause the direction of a photon to be changed significantly. The distribution of the data collected in the detectors might be different from the actual distribution of emission source activity within the object. Attenuation correction methods incorporate weighting factors that are based on the average regional or pixel-level activity. Scatter correction methods are based on estimation techniques using photo statistics derived from distribution models or experimental measurements.

3.5 Introduction to Multi-Modal Registration

There are three major approaches to multi-modal image registration, both for areas and volumes (17). These include:

1) External Marker And Stereotactic Frames Based Landmark Registration: These include external markers in multiple scans or the use of stereo tactic frames to establish a common coordinate system. Markers, which are visible in different modalities, have been used for image registration. External landmark-based registration methods use coordinate transformation and interpolation computed from visible markers to optimize the mean squared error.

2) Rigid-Body Transformation Based Global Registration: This category treats the volume (usually brain volume) as a rigid body for global registration. Levy et al presented a Principal Axes Transformation for 3-D registration, which does not require user intervention. However this method has its weaknesses for interposal PET-MR registration because PET scans do not cover the entire brain, whereas MR scans usually do. So these volumes will not match. Dhawan et. al developed an Iterative Principal Registration method which overcomes this limitation.

3) Image-Featured-Based Registration: These are algorithms which register images based on features such as edges, contours, surfaces, volumes and internal landmarks. This can be divided into two categories.

a. Boundary and Surface Matching Based Registration: Pelzer et al. presented a surface matching technique, in which 3-D models of the surface to be matched are first produced by outlining contours on serials slices of each scan. 3-D models of the brain surface from both scans are then corrected for possible geometrical errors. A geometric transformation is obtained by minimizing a predefined error function. Once the two surfaces are matched, the registration information is transferred between the scans using the volume of interest solid model created from boundary contours.

b. Image Landmarks and Feature Based Registration: Algorithms that utilize predefined landmarks that can be identified or extracted from images has been used recently. Bayesian probabilistic models have been used to integrate 3-D image information for registration. Landmark-based elastic matching algorithms have been implemented. Also, probabilistic models using maximum likelihood estimation have also been investigated.

If landmarks can be identified a priori, then corresponding points can be used to define a transformation. For non-rigid transformation with isotropic scaling, a transformation can be defined as $\mathbf{x}'=s\mathbf{r}\mathbf{x} + \mathbf{t}$ such that the registration error defined as $E(\mathbf{x}) = T(\mathbf{x}) - \mathbf{y}$ is a minimum. The vectors $\mathbf{x}$ and $\mathbf{y}$ are corresponding points in the source image X and the output image Y.

The error function can also be written in the form $\sum_{i}^{N} w_i^2 \mid s\mathbf{R}\mathbf{x}_i + \mathbf{t} - \mathbf{y}_i \mid^2$. w_{ig} is a weighting factor assigned to the point based on the confidence in the landmark and N is the total number of points used. The algorithm used to calculate the values of $\mathbf{R}$, s and $\mathbf{t}$ is as follows:

1) compute the weighted centroid of the fiducial configuration in space:

$$\bar{\mathbf{x}} = \sum_{i}^{N} w_i^2 \mathbf{x}_i / \sum_{i}^{N} \mathbf{x}_i^2 \quad \text{and}$$

$$\bar{\mathbf{y}} = \sum_{i}^{N} w_i^2 \mathbf{y}_i / \sum_{i}^{N} \mathbf{y}_i^2$$

2) Compute the displacement from each fiducial point to the centroid

$$\overline{\mathbf{x}_i} = \mathbf{x}_i - \overline{\mathbf{x}} \text{ and}$$

$$\overline{\mathbf{y}_i} = \mathbf{y}_i - \overline{\mathbf{y}}$$

3) Calculate the weighted fiducial covariance matrix:

$$\mathbf{H} = \sum_i^N w_i^2 \overline{\mathbf{x}_i} \overline{\mathbf{y}_i}^t$$

4) Perform singular value composition of H:

$$\mathbf{H} = \mathbf{U}\Lambda\mathbf{V}^t$$

where $\mathbf{U}^t\mathbf{U} = \mathbf{V}^t\mathbf{V} = \mathbf{I}$, $\Lambda = diag(\lambda_1, \lambda_2, \lambda_3)$, and $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$.

5) $\mathbf{R} = \mathbf{V}diag(1, 1, \det(\mathbf{VU}))\mathbf{U}^t$

6) $s = \sum_i^N w_i^2 \mathbf{R}\overline{\mathbf{x}_i} \cdot \overline{\mathbf{y}_i} / \sum_i^N w_i^2 \overline{\mathbf{x}_i} \cdot \overline{\mathbf{x}_i}$

7) $\mathbf{t} = \mathbf{y} - s\mathbf{R}\overline{\mathbf{x}}$

For the case where the scaling is anisotropic, a closed form solution does not exist and an iterative solution must be used.

3.6 Weighted Geometrical Featured Based Registration

A disparity function can be written which utilizes multiple geometrical features to perform a registration (18):

$$d(T) = \sqrt{\sum_{i=1}^{Ns} \sum_{j=1}^{Nxi} w_{ij}^2 \, || \, T(\mathbf{x}_{ij}) - \mathbf{y}_{ij} \, ||^2}$$

Where $\{X_i\}$ for i=1……N_s is a set of N_s shapes that correspond between **x** and **y** spaces. An iterative approach can be used to minimize the disparity function:

1) Initialization for k=1: $x_{ij}^{0} = x_{ij}$, $x_{ij}^{(1)} = T^{(0)}(x_{ij}^{(0)})$ $x_{ij}^{(1)} = T^{(0)}(x_{ij}^{(0)})$, where $T^{(0)}$ is some initial transformation. K and the superscript parentheses are iteration indices. The algorithm can be implemented with multiple initial transformations to make sure that a global minimum has been obtained.

2) Iteratively apply the following steps, incrementing k after each loop, until convergence within a tolerance ε has been obtained

   a. For each shape X_i, compute the closest points $y_{ij}^{(k)} = C_i(x_{ij}^{(k)}, Y_i)$ for j=1,…., N_{Xi}.

   b. Compute the transformation $T^{(k)}$ between the initial point set, $\{x_{ij}^{(0)}\}$, and the current set, $\{y_{ij}^{(k)}\}$, using the weights $\{w_{ij}\}$.

   c. Apply the transformation to obtain registered points: $x_{ij}^{(k+1)} = T^{(k)}(x_{ij}^{(0)})$.

   d. Terminate the loop when $d(T^{(k)}) - d(T^{(k+1)}) < \varepsilon$.

3.7 Elastic Deformation Based Registration

In the elastic deformation process, one volume which is considered to be made of an elastic material is registered to another volume, which is considered to be rigid (19). The elastic volume is deformed by applying external forces so that it matches the reference model. The matching process starts in a coarse mode in which large differences are corrected first, and then fine detailed adjustments are done later in the registration process.

There are two major constraints that can be applied during the deformation for local matching. These are smoothness and incompressibility. Smoothness guarantees that there will be continuity in the deformed volume while the incompressibility constraint ensures that the total volume remains the same. The forces required to locally match the volumes are calculated with these constraints and can be expressed by the general equation for motion of a deformable body in Lagrangian form as:

$$f(\mathbf{r},t) = \mu\frac{\partial^2\mathbf{r}}{\partial t^3} + \gamma\frac{\partial\mathbf{r}}{\partial t} + \mu\frac{\partial\varepsilon(\mathbf{r})}{\partial r}$$

where f(r,t) is the force acting on a particle at the position **r** at a time t. μ and γ are the mass and damping constant of the deformable body, and $\varepsilon(\mathbf{r})$ is the internal energy of deformation. An image voxel can be treated as a particle for which a movement is calculated for image registration.

Let the elastic volume be represented by V_1 with the coordinate system $\mathbf{x} = (x_1, x_2, x_3)$ and the reference volume can be represented by V_2 with the coordinate system $\mathbf{x}' = (x_1', x_2', x_3')$. A relationship between the two coordinate systems can be expressed as $d\mathbf{r} = \frac{\delta\mathbf{r}}{\delta x_1}dx_1 + \frac{\delta\mathbf{r}}{\delta x_2}dx_2 + \frac{\delta\mathbf{r}}{\delta x_3}dx_3$. The partial differential operators on a position vector **r** of a voxel in the registration space can be represented as $g_i = \frac{\delta\mathbf{r}}{\delta x_i}$ for i=1,2,3. Thus, the above equation can be rewritten as $d\mathbf{r} = g_i dx_i$.

If **u** is an arbitrary vector represented by (u1,u2,u3) with respect to the coordinate system such that $\mathbf{u} = u_1g_1 + u_2g_2 + u_3g_3 = u_ig_i$, it can be shown that $u_i = (u_jg_j)\cdot g_i = g_{ij}u_j$. The quantities g_{ij} are components of the generalized metric tensor G_{ijk} in a 3-D coordinate system. The metric tensor represents a distance measure with respect to a given coordinate system. Any transformation applied to the metric tensor creates a new metric tensor for the new coordinate system.

Considering a position vector **u** with respect to a point $\mathbf{a}(a_1, a_2, a_3)$ in the coordinate system, the metric tensor can be defined as $G_{ijk}\int\frac{\delta\mathbf{u}}{\delta a_i}\frac{\delta\mathbf{u}}{\delta a_j}\frac{\delta\mathbf{u}}{\delta a_k}$. A curvature

tensor $B_{ijk} = \int \frac{\delta^2 \mathbf{u}}{\delta a_i \delta a_j} \frac{\delta^2 \mathbf{u}}{\delta a_j \delta a_k} \frac{\delta^2 \mathbf{u}}{\delta a_k \delta a_i}$ can be defined, which represents the second-order partial derivative.

The potential energy $\varepsilon(x) = k \sum_{ijk} (G_{ijk} - G^0_{ijk})^2$ can be defined in terms of the metric tensor where G^0_{ijk} is the resting metric tensor (no deformation) and k is a constant. The elastic deformation function model can now be represented as $\varepsilon(\mathbf{r}) = \int_{\Omega} \| G_{ijk} - G^0_{ijk} \|^2 \, d\mathbf{u}$ where Ω represents all voxels in the volume for registration.

A model for matching through elastic deformation can be formulated as a minimization problem of the cost function that is based on a similarity measure between the two volumes to be registered. Let S(**x**,**x'**) represent a similarity measure between the local region R centered at the location **x** in V_1 and the region R' centered at **x'** in V_2. A displacement vector **u** is defined as the difference between the two locations. The optimal match for R to the region R' for the displacement vector **u** is the one that maximizes the similarity measure S(**x**,**x'**).

A possible form of the similarity measure can be expressed in terms of metric and curvature tensors as $S(x,x') = \int_{\Omega} (\| G^1_{ijk} - G^2_{ijk} \|^2 + \| B^1_{ijk} - B^2_{ijk} \|^2) da_i da_j da_k$ where the superscripts 1 and 2 represent, the deformable volume V_1 and the reference volume V_2.

3.8 Overview of Image Fusion

Image fusion techniques have been used in 3-D MR and PET brain image data sets (20). Such 3-D data can be considered as a stack of two-dimensional images cross sections that are acquired along a certain axis. In such a case, a slice from one imaging modality can

be overlayed onto a slice from another modality to create a composite visualization. The advantage of this is that boundaries and edges from the anatomical imaging modality can be overlayed onto the functional imaging modality to provide detail that was not available in either image alone.

Direct 3-D visualization methods such as stereo-pairs with color- or polarization-coded depth perception has been investigated in the medical literature. Through the variations of brightness, the observer can visualize or perceive the depth in the volumetric data. With stereoscopic visualization, two views of an object are created with different viewing angles using a binocular vision approximating the same geometry. The difference in viewing angle depends on viewing and interocular distances. Other direct display methods include motion parallax, varifocal mirror systems, rotating light-emitting diode systems and holography. The direct display methods have little significance in diagnostic radiology dealing with tomographic images.

3-D volumetric data can be obtained from a 3-D medical imaging modality such as CT, MRI or PET. In many applications, these 3-D representations will be organized as a stack of serial selections that are 2-D images of a 3-D object along a specific axis. With newer scanners, however, 3-D imaging directly provides 3-D volumetric data. The methods to visualize volumetric data can be broadly classified into two categories: surface rendering or volume rendering methods.

Surface rendering consists of turning 3-D data into a binary volume representation. The 3-D binary volume is searched for the boundary voxels that are interpolated to render a specific geometrical primitive based surface. This method is

useful for understanding the geometrical shape of the object and its variation within a group or with respect to an anatomical atlas.

Volume rendering methods allow visualization of the structural variations within the 3-D volumetric data. A common approach used in volume rendering methods is ray tracing. In the simplest form of ray-tracing, a uniform light source is placed behind the voxels of the volumetric data. For each straight-line based ray from the light source to the pixel in the image plane of the display, the density values and location of the voxels in the path of the ray are used in calculating the intensity value of the pixel in the corresponding image.

CHAPTER 4

RESULTS

4.1 Registered and Fused Examples

Thirty-eight patient images were registered and fused. Here are the results of some of the registration and fusion process:

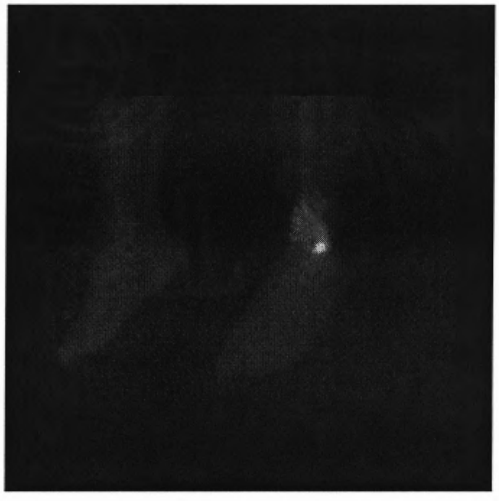

Figure 4.1 This is a planar scintigraphic image, originally saved in the DICOM format, of In-111 labeled white blood cells localizing in a painful left ankle. The regions that have the highest levels of radioactivity appear whiter in color.

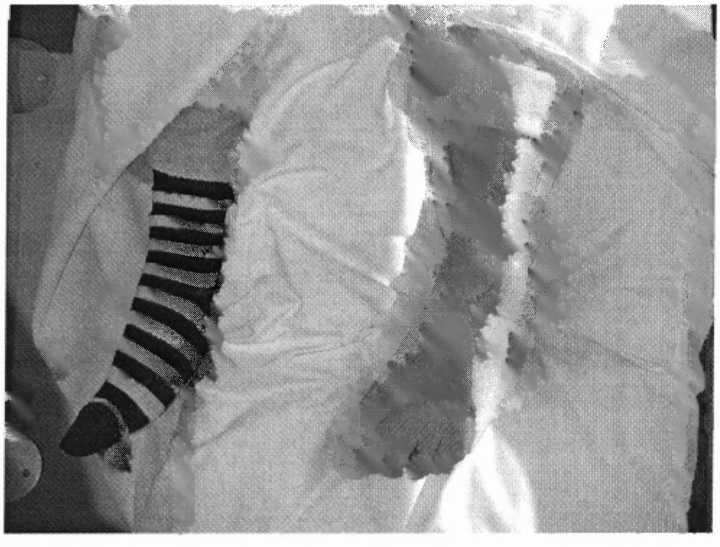

Figure 4.2 Digital visual image of the identical view.

Figure 4.3 Digital visual image registered using the affine transform and then aligned with the scintigraphic image.

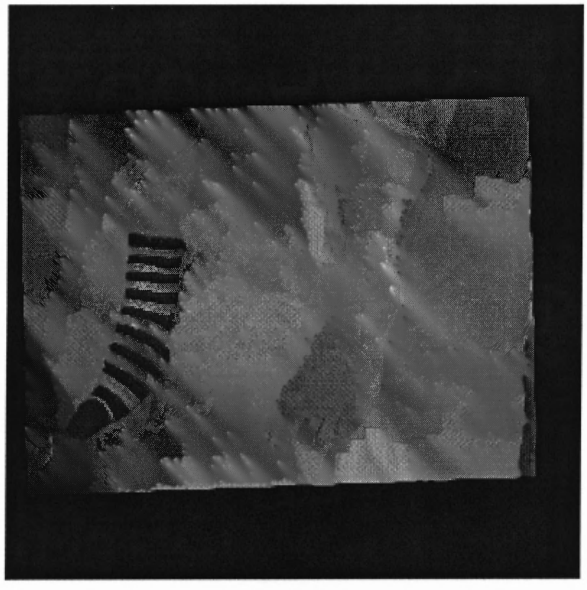

Figure 4.4 Fused image. The hue in this image contains the scintigraphic intensity. The regions with the least radioactivity are red in hue and increased levels of radioactivity are coded as yellow, green, blue and magenta, respectively.

4.2 Measurement of Error

The initial measurement of robustness of registration is error introduced within the image plane. In the above 4-point phantom (Figure 2.3) the centroid of the fourth point was at position (85.5, 168.6) in the scintigraphic image matrix while based on the affine transformation of the visual image, it was predicted to be at a position (85.5, 169.5), indicating an error less than 1 pixel confined to the y axis.

The second test was performed to assess the validity of the affine transformation of the photographic image. The camera position was held stationary and the patient bed was elevated to three separate heights (-27.5 cm, -10 cm, and 7.5 cm) with a phantom located on the bed. At +7.5 cm, the phantom was closest to the camera. The pattern of the phantom consisted of a square with four radionuclide markers at the vertices and an additional marker in the center of the pattern (See Figure 4.5).

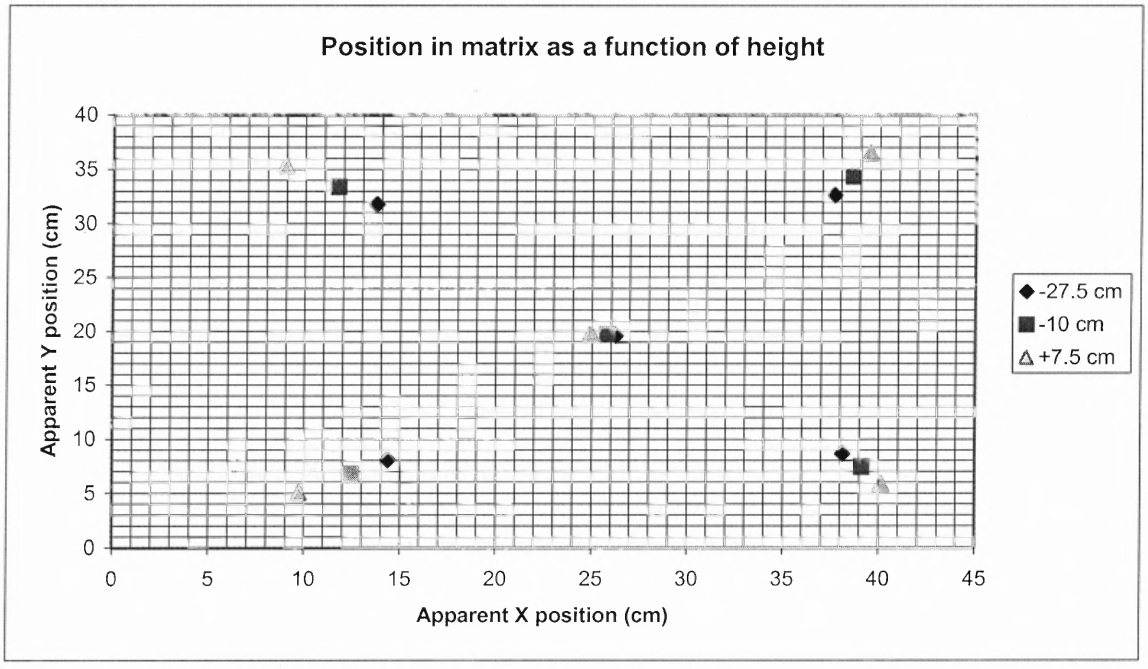

Figure 4.5 The x and y positions of the 5 point phantom imaged at distances of -27.5 cm (furthest from the photographic camera), -10 cm and 7.5 (closest to the photographic camera).

Using the pinhole perspective model of imaging, we see that image plane locations x' and y' of the object located at xylem are related as follows:

$x' = f' (x/z)$ and

$y' = f' (y/z)$

The variable f' is the length along the z direction. It is the distance between the pinhole and the image plane. The variable z is the distance from the pinhole to the object plane. See Tables 4.1-4.3 for a listing of the positions.

Table 4.1 The x and y Position of the Centroid of the Points at –27.5 cm (in cm)

-27.5	X	Y
A	13.74	31.77
B	37.67	32.63
C	14.37	8.05
D	38.14	8.66
E (midpoint)	26.25	19.58

Table 4.2 The x and y Position of the Centroid of the Points at –10 cm (in cm)

-10	X	Y
A	11.74	33.41
B	38.6	34.38
C	12.44	6.77
D	39.14	7.46
E (midpoint)	25.79	19.72

Table 4.3 The x and y Position of the Centroid of the Points at 7.5 cm (in cm)

7.5	X	Y
A	8.95	35.36
B	39.52	36.49
C	9.75	5.05
D	40.16	5.85
E (midpoint)	24.95	19.8

Given any two points (P, Q) in the object plane, the length of the line segment in the image plane will be related as $m|PQ|$, m being the magnification. Since the phantom

was imaged at three separate depths, the distance between any two points P, Q in the images will be related to the original object plane length by $m_1|PQ|$, $m_2|PQ|$, and $m_3|PQ|$. Thus, if the image plane is parallel to the object plane, the line segments that connect P and Q should be related by a common multiple (m_3/m_2 or m_3/m_1 or m_2/m_1). Tables 4.4-4.6 have a listing of the Pythagorean distances between points at a certain distance from the photographic camera. Point A is the upper left point, Point B is the upper right point, Point C is the lower left point, Point D is the lower right point, and Point E is the midpoint.

Table 4.4 The Pythagorean Distance Between Points at –27.5 cm (in cm)

-27.5	A	B	C	D	E
A	X	23.94545	23.72836	33.60702	17.467
B	X	X	33.86837	23.97461	17.34125
C	X	X	X	23.77783	16.55522
D	X	X	X	X	16.14368
E	X	X	X	X	X

Table 4.5 The Pythagorean Distance Between Points at –10 cm (in cm)

-10	A	B	C	D	E
A	X	26.87751	26.6492	37.73808	19.61679
B	X	X	38.03495	26.92542	19.46822
C	X	X	X	26.70891	18.59906
D	X	X	X	X	18.1254
E	X	X	X	X	X

Table 4.6 The Pythagorean Distance Between Points at 7.5 cm (in cm)

7.5	A	B	C	D	E
A	X	30.59088	30.32056	42.95235	22.31846
B	X	X	43.29811	30.64668	22.15493
C	X	X	X	30.42052	21.18024
D	X	X	X	X	20.63847
E	X	X	X	X	X

Overall, the scaling factors between line segments at –10 cm to –27.5 cm were roughly 1.23. The scaling factor between line segments at 7.5 cm to –27.5 cm was 1.27 and the scaling factor between line segments at 7.5 and –10 cm was 1.13. It appears that the scaling is relatively uniform at different depths which means that spherical and chromatic aberrations are not a significant cause of registration error. See Tables 4.7-4.9.

Table 4.7 Ratios of the Lengths of the Identical Line Segments at –10 cm and –27.5 cm

Ratios(-10/-27.5)	A	B	C	D	E
A	X	1.122448	1.123094	1.122922	1.123077
B	X	X	1.123023	1.123081	1.122654
C	X	X	X	1.12327	1.123456
D	X	X	X	X	1.122755
E	X	X	X	X	X

Table 4.8 Ratios of the Lengths of the Identical Line Segments at 7.5 cm and –27.5 cm

Ratios(7.5/-27.5)	A	B	C	D	E
A	X	1.277524	1.277819	1.278076	1.2775
B	X	X	1.278423	1.278298	1.277586
C	X	X	X	1.279365	1.279369
D	X	X	X	X	1.278424
E	X	X	X	X	X

Table 4.9 Ratios of the Lengths of the Identical Line Segments at 7.5 and –10 cm

Ratios(7.5/-10)	A	B	C	D	E
A	X	1.138159	1.137766	1.13817	1.137722
B	X	X	1.138377	1.138207	1.138005
C	X	X	X	1.138965	1.13878
D	X	X	X	X	1.138649
E	X	X	X	X	X

The second evaluation that was performed was to assess error introduced by variation in the image height (Figure 4.5). From Figure 4.5, several factors are apparent.

The center of the phantom aligns with the center of the image in the y dimension, but not in the x axis, where it is displaced approximately 50 pixels to the right. Secondly, as the phantom moves away from the imaging camera, the size of the rectangle boarded by the 4 outer markers decreases by 5.72 pixels per cm height (in the x direction) and 5.75 pixels per cm height (in the y direction) or approximately .62 percent per cm height (in the x direction) and .62 percent per cm in height (in the y direction). This represents an estimate of the registration error based on incorrect camera-subject distance. As the distance between the center point and the photographic camera increases, the center point slightly shifts to the right.

An analysis was done comparing the affine registration parameters of the generalized transformation to the patient specific transformation. The generalized transformation was defined on January 26, 2004. It was compared to patient specific transformations defined on five other dates.

Table 4.10 Affine Transformation Parameters for "Generalized" and Patient-Dependent Transformations

Date	Transformation Parameters
January 26, 2004 "Generalized Transformation"	-.141, .005, 0 -.0044, -.1408, 0 237.2464, 210.7148, 1
June 9, 2004a	-.1456, .0075, 0 .0024, -.1437, 0 234.6083, 209.2214, 1
June 9, 2004b	-.1407, .0045, 0 .021, -.1326, 0 224.1104, 203.888,1
June 11, 2004	-.1406, .0076, 0 -.0017, -.1417, 0 224.5652, 207.5135, 1
June 14, 2004	-.2658, .0041, 0 -.0131, -.2723, 0 333.6204, 290.1921, 1
July 7, 2004	-.1427, .0035, 0 -.0015, -.1472 234.4822, 214.9255, 1

Table 4.11 Comparison of Generalized Transformation Parameters and Average of Patient Dependent Transformations

	Gen. Trans. Param. 1 – Trans. x Param. 1	Gen. Trans. Param. 2 – Trans. x Param. 2	Gen.Trans. Parameter 3 – Transformation x Parameter 3	Gen. Trans. Param. 4– Trans. x Param. 4	Gen. Trans. Param. 5– Trans. x Param. 5	Gen. Trans. Param. 6– Trans x Param. 6
Trans.1	0.0046	0.0025	0.0068	0.0029	2.6381	1.4934
Trans.2	0.0003	0.0005	0.0254	0.0082	13.136	6.8268
Trans.3	0.0004	0.0026	0.0027	0.0009	12.6812	3.2013
Trans.4	0.1248	0.0009	0.0087	0.1315	96.374	79.4773
Trans.5	0.0017	0.0015	0.0029	0.0064	2.7642	4.2107

Table 4.10 refers to a comparison of the affine parameters defined specifically for patients on five dates as compared to the generalized affine parameters. The affine parameter defined on June 14, 2004 appears to be more of an outlier than the transformations defined on the other days.

Table 4.11 is a comparison of the six parameters of the affine transformation parameters between the generalized transformation and the five patient dependent transformations.

Table 4.12 Mean and Standard Deviation of Error for Affine Transformation Parameters

	Mean of Errors	Standard Dev. of Errors	Mean Of Errors without June 14, 2004	Standard Dev. Off Errors without June 14, 2004
Parameter 1	0.02636	0.055056998	.00175	.002
Parameter 2	0.0016	0.000938	.001775	.0009
Parameter 3	0.0093	0.009358686	.00945	.0108
Parameter 4	0.02998	0.056824	.0046	.0033
Parameter 5	25.5187	39.93712349	7.80	5.90
Parameter 6	19.0419	33.83964	3.15	2.23

Table 4.12 has the means of the errors for the six different 2-D affine parameters and the standard deviation of the means. The affine parameters obtained on June 14, 2004 were significantly different from those obtained on the other days. For this reason, the mean of errors and the standard deviation of the errors were recalculated without this data point. The generalized and the patient-specific affine transformations appear to be similar. More data is needed to see if they are statistically similar.

Another measure of error was obtained by using the generalized transformation defined on the data of January 26, 2004 and applying this transformation on patient data acquired on other days. Cobalt markers were placed on the patients and a single cobalt marker was used as a best representative for the error of the transformation. A comparison was done between the predicted locations of the cobalt marker in the transformed image compared to the actual locations in the gamma camera image. Table 4.13 has this information.

Table 4.13 Comparison of Error for Generalized Affine Transformation

	jpg-x	jpg-y	dcm-x	dcm-y	Error x-pixels	Error y-pixels	Error sum pixels	Error sum in cm
40812	90	158	85	157	5	1	5.10	1.196
40709	131	141	127	140	4	1	4.12	.967
407011	187	127	186	126	1	1	1.41	.331
40630	133	137	130	136	3	1	3.16	.741
406091a	154	168	150	167	4	1	4.12	.967
406092a	184	127	187	125	3	2	3.60	.845
40611	105	123	95	121	10	2	10.20	2.34
40614	126	157	125	157	1	0	1	.234

It can be seen from Table 4.13, that in most cases, the error was small. The average error in cm was .95 cm. In one particular case (case number 40611), the error was significant. It is speculated that the height of the patient bed was not properly adjusted for this particular case.

4.3 Error of a Single Point

Error from the digital visual image can be minimized both by adjusting the z displacement from the camera to conform to the height at which the affine transformation was defined and by placing the area of interest as close to the optical axis of the camera as possible. Table 4.14 shows the distances from the midpoint (point E) to each of the four corner points. Table 4.15 shows the error at the four outer points of the five-point phantom, both as absolute distances and as percentages.

Table 4.14 Distances from Midpoint to Four Outer Points at Different z Distances

	Distance from midpoint at 7.5 cm (in cm)	Distance from midpoint at –10.5 cm (in cm)	Distance from midpoint at –27.5 cm (in cm)
Point A	22.3	19.6	17.5
Point B	21.1	19.5	17.3
Point C	21.2	18.6	16.6
Point D	20.6	18.1	16.1

Table 4.15 Error as a Function of z Displacement

Distance	Absolute Dilation for z displacement of 17.5 cm (in cm)	Absolute Dilation for z displacement of 35 cm (in cm)	% Change for z displacement of 17.5 cm	% Change for z displacement of 35 cm
Point A-Point E	2.7	4.8	12.11	21.52
Point B-Point E	1.6	3.8	7.58	18.00
Point C-Point E	2.6	4.6	12.26	21.70
Point D-Point E	2.5	4.5	12.14	21.84

It should be noted that the calculated error in the position of the four peripheral markers represents a worst-case scenario in that in the majority of cases, the area of interest to be imaged lies near the center of the image, and not at the periphery. Because the camera magnification or minification error due to discrepancy in camera-object distance is proportional to the distance between the area images and the center of the axis r, error within the circular field of view of radius 3 inches will only have an error equivalent to $3/\sqrt{(2*6^2)}$ or $3/\sqrt{72}$ of the corner errors. Theses formulae are dependent on the fact that there were 12 inches between any two of the outer points. Assuming that the phantom was centered with the optical axis of the camera (which is not

a bad assumption because it is centered in the y-axis and only displaces by 50 pixels in the x-direction for a z displacement of 35 cm), a reasonable calculation of the error can be obtained by taking the maximum absolute dilation (4.8 cm), converting this into inches and multiplying this by $3/\sqrt{72}$, and converting back into cm. The error obtained is 1.70 cm.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

The goal of this project is to improve the ability of nuclear medicine practitioners to accurately localize sites of abnormal signal by combining digital visual images with planar scintigraphic images based on the hypothesis that fusion of planar scintigraphic images with visual images leads to greater diagnostic confidence and accuracy.

Towards this aim, a working prototype has been constructed and a clinical acquisition protocol for the combined visual and scintigraphic camera (including successful registration and fusion of the visual and scintigraphic images) has been initiated. A total of 38 images have been registered and fused.

Analysis of the imaging system reveals that the imaging system is linear and spatially invariant, as the magnification values for the phantom were uniform at different heights. While spherical and chromatic aberrations were noticed upon visual inspection of the phantoms, they did not contribute to significant error.

The main source of error in the registration between the scintigraphic and photographic camera occur because when the plane of interest is at an angle to the camera, perspective error is introduced into the image. The second source of error occurs by the process of approximating a surface of variable height with a 2-D plane.

One of the original hypotheses was that a generalized affine registration defined for a certain distance from the photographic camera would be of negligible error as compared to the patient dependent affine transformations. This type of study would take into account error due to perspective, variations in the height of the patient bed, fiducial

location errors, etc. It would not take into account the errors that would occur at the variations from the 2-D planar representation of the plane.

Towards this goal, patient specific data has been gathered from approximately 10 patients. A generalized transformation has been compared to the average of the transformations obtained for five patients.

Results from the comparison between the two sets of transformations shows that they are similar, but this must be confirmed with more data.

Further analysis was conducted by comparing the location of one cobalt marker in a registered photographic image and in the accompanying nuclear image. The marker was chosen as a best representative of the location of the lesion. The average error was .95 cm and the maximum error was 2.54 cm.

A theoretical evaluation for the error was conducted by making the assumption that the center point of the 5-point phantom is aligned with the optical axis of the camera. The displacement at the four outer points was measured and then an estimate was obtained for a radius closer to the optical axis of the camera. This region is where patients will ideally be imaged. A point 3 inches from the optical axis in the radial direction would be displaced by 1.70 cm if the z were to be displaced by 35 cm.

REFERENCES

1. Bocher, M., Balan, A., Krausz, Y. et al. (2000). Gamma camera-mounted anatomical X-ray tomography: technology, system characteristics, and first images. Eur. J Nucl. Med., 27, 619-627.

2. Delbeke, D., & Sandler, M.P. (2000). The role of hybrid cameras in oncology. Semin Nucl Med., 30, 268-280.

3. Israel, O., Keidar, Z., Iosilevsky, G., Bettman, L., Sachs, J., Frenkel, A. (2001). The fusion of anatomic and physiologic imaging in the management of patients with cancer. Semin Nucl Med., 31, 91-205.

4. Townsend, D.W., Cherry, S.R. (2001). Combining anatomy and function: the path to true image fusion. Eur Radiol., 11, 968-1974.

5. Czernin, J., Schelbert, H. (2004). PET/CT Imaging: facts, opinions, hopes and questions. J Nucl Med., 45, Suppl. 1:1s-3s.

6. Image Processing Toolbox (3.0) Users Guide, The Math Works, Natick, MA, 1998, pp. 5-13

7. ImageJ. Vers. 1.32j. 26 Feb. 2005 <http://rsb.info.nih.gov/ij>.

8. TurboReg. Vers. 1.0 28 Feb. 2005 <http://bigwww.epfl.ch/thevenaz/turboreg/>.

9. Image Layering Toolbox. Vers. 1.0 19 June 2002 <http://rsb.info.nih.gov/ij/plugins/image-layering.html>.

10. The Centroid. (2003). Retrieved June 12, 2005, from <http://www.maths.abdn.ac.uk/~igc/tch/index/eg1006/notes/node87.html>.

11. Dhawan, A.P. (2003). Medical Image Analysis. Hoboken, NJ: John Wiley & Sons, Inc. (pp. 256-257).

12. Bilinear Interpolation. (1994). Retrieved June 12, 2005, from http://www.cs.ubc.ca/spider/little/505/lectures/subsection3_9_2.html.

13. Forsythe, D.A. & Ponce, J. (2003). Computer Vision, A Modern Approach. Delhi, India: Pearson Education, Inc. (pp. 3-12).

14. Dhawan, A.P. (2003). Medical Image Analysis. Hoboken, NJ: John Wiley & Sons, Inc. (p. 92).

15. Sorenson, J.A. & Phelps, M.E. (1987). Physics in Nuclear Medicine, 2nd edition. Orlando, Florida: Grune & Stratton, Inc. (pp. 298-345).

16. Dhawan, A.P. (2003). Medical Image Analysis. Hoboken, NJ: John Wiley & Sons, Inc. (pp. 93-96).

17. Dhawan, A.P. (2003). Medical Image Analysis. Hoboken, NJ: John Wiley & Sons, Inc. (pp. 266-267).

18. Fitzpatrick, M. J., Hill, D.L.G., Maurer, C.R Jr. Image Registration. In Sonka, Milan & Fitzpatrick, M.J. (Ed.), Handbook of Medical Imaging (pp. 447-513). Bellingham, Washington: SPIE Press.

19. Dhawan, A.P. (2003). Medical Image Analysis. Hoboken, NJ: John Wiley & Sons, Inc. (pp. 268-272).

20. Dhawan, A.P. (2003). Medical Image Analysis. Hoboken, NJ: John Wiley & Sons, Inc. (pp. 277-281).