

1995

Automated Analysis Of Nuclear Medicine Images: Towards Artificial Intelligence Systems

Piotr Jan Slomka

Follow this and additional works at: <https://ir.lib.uwo.ca/digitizedtheses>

Recommended Citation

Slomka, Piotr Jan, "Automated Analysis Of Nuclear Medicine Images: Towards Artificial Intelligence Systems" (1995). *Digitized Theses*. 2522.

<https://ir.lib.uwo.ca/digitizedtheses/2522>

This Dissertation is brought to you for free and open access by the Digitized Special Collections at Scholarship@Western. It has been accepted for inclusion in Digitized Theses by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca, wlsadmin@uwo.ca.

**AUTOMATED ANALYSIS OF NUCLEAR MEDICINE IMAGES:
TOWARDS ARTIFICIAL INTELLIGENCE SYSTEMS**

by

Piotr Slomka

Department of Medical Biophysics

**Submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy**

**Faculty of Graduate Studies
The University of Western Ontario
London, Ontario
November 1994**

© Copyright by Piotr Slomka 1995



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

**THE AUTHOR HAS GRANTED AN
IRREVOCABLE NON-EXCLUSIVE
LICENCE ALLOWING THE NATIONAL
LIBRARY OF CANADA TO
REPRODUCE, LOAN, DISTRIBUTE OR
SELL COPIES OF HIS/HER THESIS BY
ANY MEANS AND IN ANY FORM OR
FORMAT, MAKING THIS THESIS
AVAILABLE TO INTERESTED
PERSONS.**

**L'AUTEUR A ACCORDE UNE LICENCE
IRREVOCABLE ET NON EXCLUSIVE
PERMETTANT A LA BIBLIOTHEQUE
NATIONALE DU CANADA DE
REPRODUIRE, PRETER, DISTRIBUER
OU VENDRE DES COPIES DE SA
THESE DE QUELQUE MANIERE ET
SOUS QUELQUE FORME QUE CE SOIT
POUR METTRE DES EXEMPLAIRES DE
CETTE THESE A LA DISPOSITION DES
PERSONNE INTERESSEES.**

**THE AUTHOR RETAINS OWNERSHIP
OF THE COPYRIGHT IN HIS/HER
THESIS. NEITHER THE THESIS NOR
SUBSTANTIAL EXTRACTS FROM IT
MAY BE PRINTED OR OTHERWISE
REPRODUCED WITHOUT HIS/HER
PERMISSION.**

**L'AUTEUR CONSERVE LA PROPRIETE
DU DROIT D'AUTEUR QUI PROTEGE
SA THESE. NI LA THESE NI DES
EXTRAITS SUBSTANTIELS DE CELLE-
CI NE DOIVENT ETRE IMPRIMES OU
AUTREMENT REPRODUITS SANS SON
AUTORISATION.**

ISBN 0-315-99298-0

Canada

ABSTRACT

Automated methods for the analysis of nuclear medicine images could provide an objective diagnosis, and means to transfer sophisticated expertise to less experienced centres. The goal of this study was to develop software methods for the automated analysis of (a) Quality Control (QC) images, and (b) myocardial perfusion tomography images.

The system for the automated analysis of QC images was based on feature extraction algorithms, which provided input to a higher level diagnostic expert system. Several features characterizing QC images were defined. Rule-based and object-oriented expert systems were created to guide personnel in QC procedures, detect gamma camera faults, and suggest corrective actions. An object-oriented representation of knowledge allowed a natural representation and classification of image features, artefacts, and other concepts used in this knowledge domain. The feature extraction algorithms combined with a prototype expert system could perform diagnosis of gamma camera faults and QC procedure errors on a limited set of examples.

Computer-aided analysis of myocardial perfusion images was accomplished by creating three-dimensional (3-D) reference templates, to which patient's images could be automatically aligned using image registration algorithms. The templates included a normal distribution of activity and perfusion maps corresponding to specific coronary arteries. The quantification was done by a 3-D region-growing procedure that outlined perfusion defects in test-patients based on differences from the normal templates. Alignment and quantification methods of myocardial perfusion images were successfully tested on a group of 168 angiographically correlated patients. Perfusion defects were characterized in terms of numeric parameters, thus avoiding subjective visual assessment. The location of defects relative to the expected hypoperfusion sites was also established.

Analytical and artificial intelligence software methods can be used for automated interpretation of QC and cardiac images. Object-oriented methods are suitable for

encoding the knowledge required for computer-aided analysis of QC images. A comprehensive and fully automated analysis of cardiac perfusion images is possible by comparison of patient data to 3-D reference models.

ACKNOWLEDGEMENTS

To begin, I would like to thank Trevor Craddock for his guidance and advice, as a colleague and supervisor in my graduate studies. His experience and knowledge in the field of nuclear medicine was of considerable help to me during my work and studies at Victoria Hospital. He is partly responsible for my transformation from a callow "computer hacker" to a more mature scientific researcher.

This work would be extremely difficult without fruitful discussions with two other members of my advisory committee: Bob Mercer and Frank Prato. Bob spent many hours with me discussing the concepts of artificial intelligence, and Frank offered his advice in the field of biomedical scientific research. In addition, the enthusiasm of Gil Hurwitz for my work and his dedication in providing research material as well as important, clinically-oriented perspective in many inspiring discussions was invaluable. This multidisciplinary combination was very important during my research. Also, I would like to express my gratitude to all the other staff members at Victoria Hospital who over the years encouraged me, and contributed in many ways to my professional and scientific experience. Soraya Ali deserves a special mention for her patience in proof-reading numerous versions of chapters from this thesis.

Finally, I give special thanks to my friends and family, who provided psychological support; the long hours, when the research results were not quite right, would have been a nightmare without them. In particular, I would like to mention Janice for her belief in my work and words of encouragement.

TABLE OF CONTENTS

	Page
CERTIFICATE OF EXAMINATION	ii
ABSTRACT	iii
ACKNOWLEDGEMENTS	v
TABLE OF CONTENTS	vi
LIST OF TABLES	x
LIST OF FIGURES	xi
NOMENCLATURE	xiv
CHAPTER 1: INTRODUCTION	1
1.1 Computer-aided Medical Image Analysis	1
1.2 Computer-aided Analysis of Nuclear Medicine Images	3
1.3 Thesis Problem	5
1.4 Thesis Outline	8
References:	10
PART I: AUTOMATED ANALYSIS OF QUALITY CONTROL IMAGES	13
CHAPTER 2: SCINTILLATION CAMERA: PRINCIPLE OF OPERATION AND QUALITY CONTROL	13
2.1 Introduction	13
2.2 Scintillation Camera	14
2.3 Assessment of Camera Performance	16
References:	24

CHAPTER 3: AUTOMATED FEATURE EXTRACTION AND QUANTIFICATION OF QUALITY CONTROL IMAGES: TOWARDS A KNOWLEDGE BASED SYSTEM	
.....	26
3.1 Introduction	26
3.2 System Design	27
3.3 Results	33
3.4 Discussion	38
Acknowledgements	41
References:	42
CHAPTER 4: A PROTOTYPE RULE-BASED EXPERT SYSTEM FOR QUALITY CONTROL OF A GAMMA CAMERA	45
4.1 Introduction	45
4.2 System Design and Prototype Implementation	47
4.3 Discussion and Conclusions	53
4.4 Summary	55
References:	56
CHAPTER 5: KNOWLEDGE ACQUISITION, REPRESENTATION AND REASONING IN A GAMMA CAMERA QUALITY CONTROL EXPERT SYSTEM	57
5.1 Introduction	57
5.2 Creating the QC Image Database	59
5.3 Expert System Design	61
5.4 An Example	65
5.5 Preliminary Evaluation of the Prototype	69
5.6 Discussion	69
5.7 Conclusions	71
Acknowledgments	72
References:	73

PART II: AUTOMATED ANALYSIS OF MYOCARDIAL SPECT IMAGES	75
CHAPTER 6: MYOCARDIAL PERFUSION SPECT IMAGING	75
6.1 Introduction	75
6.2 Imaging Techniques	76
6.3 Limitations of Myocardial SPECT	79
6.4 Visualization and Quantification Methods	81
References:	85
CHAPTER 7: AUTOMATED ALIGNMENT AND SIZING OF MYOCARDIAL STRESS AND REST SCANS TO THREE-DIMENSIONAL NORMAL TEMPLATES USING AN IMAGE REGISTRATION ALGORITHM: A METHOD FOR REPRODUCIBLE QUANTIFICATION	91
7.1 Introduction	91
7.2 Methods	93
7.3 Results	105
7.4 Discussion	112
7.5 Summary	118
Acknowledgements	118
References:	119
CHAPTER 8: AUTOMATED CHARACTERIZATION OF MYOCARDIAL PERFUSION DEFECTS BY COMPARISON TO REFERENCE TEMPLATES DEMARCATED WITH HYPOPERFUSION TERRITORIES	122
8.1 Introduction	122
8.2 Methods	123
8.3 Results	129
8.4 Discussion	136
8.5 Conclusions	140
8.6 Summary	140
Acknowledgements	141
References:	142

CHAPTER 9: SUMMARY AND FUTURE WORK	144
9.1 Summary	144
9.2 Future Work	149
References:	153
APPENDIX A: ALGORITHMS FOR EXTRACTION OF FEATURES FROM QUALITY CONTROL IMAGES	155
A.1 Edge-finding Algorithm	155
A.2 Shape-fitting Algorithms	158
A.3 NEMA Calculations	162
A.4 Cold and Hot Spot Quantification	164
A.5 Edge Distortion	164
A.6 Asymmetry	169
References:	172
VITA	173

LIST OF TABLES

	Page
TABLE 3.1: List of simulated and real faults in the collection of floods.	36
TABLE 3.2: Use of distortion coefficients in detecting problems with faulty photomultipliers located on the edge of the field-of-view, and improperly positioned floods.	38
TABLE 5.1: Categories of collected gamma camera faults	61
TABLE 7.1: Simulated-defect experiment. Registration errors (mean \pm s.d.) for three different convergence criteria.	107
TABLE 7.2: Reproducibility of registration after arbitrary misalignments. Mean errors \pm s.d. of transformation parameters for the count-difference method.	108
TABLE 7.3: Effect of alignment errors on segmental count distribution in 2 patients (one normal, one abnormal) with highest registration errors in the reproducibility test.	109
TABLE 8.1: Number of patients with defects in designated territories.	135
TABLE 8.2: Average characteristics of defects.	135

LIST OF FIGURES

	Page
Figure 2.1: Diagram of a scintillation camera.	14
Figure 2.2: QC images of a flood-source (left) and four-quadrant bar phantom (right).	19
Figure 2.3: Decision tree suggested for execution, evaluation and follow-up of a routine quality control test.	23
Figure 3.1: Overview of automated QC system	28
Figure 3.2: Dimension features derived from flood images.	30
Figure 3.3: Three definitions of distortion.	32
Figure 3.4: Sample Interfile keys in QC report, describing numeric features derived from images.	34
Figure 3.5: Various faulty flood images used to assess feature extraction algorithms.	35
Figure 3.6: Generating simulated rectangular shape from circular flood image. . .	37
Figure 3.7: Examples of Prolog predicates corresponding to image features. . . .	40
Figure 4.1: An example of a low level rule.	49
Figure 4.2: Example of a troubleshooting rule.	50
Figure 4.3: Inference strategy in the system.	52
Figure 5.1: Fragment of hierarchical classification of image artefacts.	63
Figure 5.2: Fragment of hierarchical classification of features derived from QC images.	63
Figure 5.3: Logical modules and flow of information in <i>QCMAN</i>	65
Figure 5.4: Example of a quality control image.	66
Figure 5.5: <i>QCMAN</i> reasoning path in an example of floor contamination.	67
Figure 6.1: 180° acquisition of cardiac SPECT images.	78
Figure 6.2: Standard orientation of planes in reconstructed cardiac SPECT datasets.	79
Figure 6.3: Polar map (or bull's eye) representation of myocardial SPECT	82
Figure 7.1: Masking of external activity.	94

Figure 7.2:	Nine linear transformation parameters used by the image registration technique to reorient and resize hearts.	96
Figure 7.3:	Accumulating patient data in the normal model.	97
Figure 7.4:	Male (top) and female (bottom) normal composite stress templates.	98
Figure 7.5:	Three-dimensional surface shaded views of normal male stress (left) and rest templates (right).	99
Figure 7.6:	Three-dimensional surface shaded display of abnormal (LAD) male stress (left) and rest templates (right)	99
Figure 7.7:	Scheme for template segmentation.	100
Figure 7.8:	Mid-ventricular horizontal long-axis, short-axis, and vertical long-axis slices of the encoded segment template.	101
Figure 7.9:	<i>Roving window</i> technique for visual assessment of image registration	102
Figure 7.10:	Example of a simulated defect.	103
Figure 7.11:	Range of scaling parameters in three different directions (X, Y, Z) in normal groups (23 males and 15 females).	105
Figure 7.12:	Mid-ventricular slices of patient data with severe LAD defect (color) overlaid on the model data (greyscale).	106
Figure 7.13:	Standard deviations of relative count distribution in each segment for stress, rest and the difference between stress and rest in the normal male group (n=23).	111
Figure 7.14:	Abnormal segments in the LAD group of 10 male patients.	111
Figure 7.15:	Subtraction of the abnormal (LAD) stress (top) and rest (bottom) patient images (left) from the normal model (middle).	112
Figure 8.1:	Horizontal long-axis (top), short-axis (middle) ,and vertical long-axis (bottom) of the variation template.	130
Figure 8.2:	Small LCX defect in test-patient (top) and results of region-growing procedure (bottom).	131
Figure 8.3:	Average LCX (top) and RCA (bottom) territories (green) overlaid on the respective disease templates.	132
Figure 8.4:	Average PLAD territory (green) overlaid on PLAD template, accumulated from 26 patients.	133

Figure A.1:	8-neighbourhood and 4-neighbourhood definition.	155
Figure A.2:	Morphological operator for edge smoothing.	156
Figure A.3:	Backtracking operation.	157
Figure A.4:	Flowchart of edge-finding algorithm.	158
Figure A.5:	First approximation of fitted shapes for polygons and circles.	159
Figure A.6:	Bevelled-rectangular field-of-view.	161
Figure A.7:	Useful (UFOV) and central (CFOV) field-of-view definitions.	162
Figure A.8:	Polygon contour filling technique.	163
Figure A.9:	Length of the contour.	164
Figure A.10:	Definition of asymmetry for polygons and circles.	170

NOMENCLATURE

ADC - analog to digital converter

AI - artificial intelligence

CAD - coronary artery disease

Co-57 - Cobalt-57

DLAD - distal left anterior descending

FWHM - full width at half maximum

LAD - left anterior descending

LCX - left circumflex

NEMA - National Electrical Manufacturer's Association

QC - quality control

PHA - pulse height analyzer

PLAD - proximal left anterior descending

PM - photomultiplier

RCA - right coronary artery

ROI - region of interest

SPECT - single photon emission computed tomography

Tc-99m - Technetium-99m

3-D - three-dimensional

2-D - two-dimensional

The author of this thesis has granted The University of Western Ontario a non-exclusive license to reproduce and distribute copies of this thesis to users of Western Libraries. Copyright remains with the author.

Electronic theses and dissertations available in The University of Western Ontario's institutional repository (Scholarship@Western) are solely for the purpose of private study and research. They may not be copied or reproduced, except as permitted by copyright laws, without written authority of the copyright owner. Any commercial use or publication is strictly prohibited.

The original copyright license attesting to these terms and signed by the author of this thesis may be found in the original print version of the thesis, held by Western Libraries.

The thesis approval page signed by the examining committee may also be found in the original print version of the thesis held in Western Libraries.

Please contact Western Libraries for further information:

E-mail: libadmin@uwo.ca

Telephone: (519) 661-2111 Ext. 84796

Web site: <http://www.lib.uwo.ca/>

CHAPTER 1: INTRODUCTION

1.1 COMPUTER-AIDED MEDICAL IMAGE ANALYSIS

Currently, most medical imaging techniques involve visual detection and verbal description of various image patterns. These patterns represent a variety of physical properties, which can be registered by several imaging modalities. Often, a mental perception of normal and abnormal images is used implicitly by the interpreting specialist. Medical images can be multidimensional, involving spatial, temporal or other types of measures. In common practice, the task of interpreting images is accomplished visually with the help of various data presentation techniques. These manual methods of analysis are influenced by human skill, knowledge and experience. Subtle patterns and differences in images frequently cause errors and discrepancies in the interpretation. Inter-observer and intra-observer variability are of serious concern in medical imaging disciplines (1-3). The final diagnosis involves a combination of visual perception and specific knowledge about the imaging technology, human anatomy and physiology. The mechanisms involved in this flow of information, from the images to the final conclusions, are fundamental to the whole process of medical imaging.

Automated computer methods offer the promise to streamline and simplify this complex diagnostic process. As computers become more powerful and more affordable, their use in medical image diagnosis becomes more viable. If reliable and reproducible tools for image interpretation are developed, medical imaging knowledge can be transferred and shared between many centres, providing a consistent level of diagnosis as well as objective results. For example, programs developed in experienced academic centres, encoding the expertise of highly trained professionals could be used in smaller community hospitals. Computer-aided diagnosis would be free from miscellaneous psychological factors, for example fatigue or stress. This may be especially important under an intensive workload or in emergencies, which are common in medical practice.

Furthermore, automated computer methods are usually based on quantitative data analysis, which has the potential to reveal subtle differences or visually indistinct patterns (4). As a result, the quality of patient care could be improved and the costs of procedures reduced. Moreover, the development of such methods opens new insights into the understanding of the visual and diagnostic processes.

Many studies have attempted to utilize various computer techniques in the analysis of medical images. Both image processing and artificial intelligence methods have been applied to automate various aspects of medical image diagnosis in radiology, nuclear medicine, and other modalities. Computer techniques have been developed for pattern detection, image interpretation and logical diagnostic steps. Traditionally, algorithmic image processing techniques have been applied to medical images. More recently, artificial intelligence tools such as expert systems (5) and neural networks (6) have become increasingly utilized in the interpretation of image data. Expert systems imitate human reasoning by deriving logical conclusions from a series of rules and facts. The rules and facts represent a knowledge base of a particular domain. On the other hand, artificial neural networks do not use explicit logical dependencies but adapt themselves during the learning process to associate a series of inputs with a series of outputs. In the training phase, input examples with known outcomes are provided. After such training the artificial neural network can predict the outcome of previously unseen examples. In contrast to expert systems, the exact logical path from the input to the output of the artificial neural network cannot be explained.

Expert systems (7) and artificial neural networks (8) have been used in conjunction with other low level image processing techniques to deliver tools for knowledge based image interpretation. The image preprocessing steps are applied to extract the necessary diagnostic data and to reduce the amount of insignificant information. The preprocessing steps involve, for example, normalization of images with respect to position and orientation, or division into several predefined regions (segmentation). Currently, most

implementations of these techniques involve significant user interaction. For example, they often require expertise in placing regions of interests (ROI) or defining edges and orientations of various image elements. The need for skilled user interaction to prepare images for artificial intelligence interpretation introduces subjectivity to the diagnostic process. Although several computer techniques for edge detection, fitting, and image segmentation have been developed, no universal algorithm for these operations exists. To my knowledge, a fully automated medical image diagnostic system has not yet been developed. Potentially, neural networks could be applied directly to raw image data to yield diagnostic answers (9). Nevertheless, the large number of image elements (pixels, voxels) and diagnostically unimportant features, prohibit the practical implementation of neural networks and other machine learning methods to analyze unprocessed medical images at the present time (10).

1.2 COMPUTER-AIDED ANALYSIS OF NUCLEAR MEDICINE IMAGES

The nuclear medicine imaging modality has been one of the most advanced areas of computer-aided medical image analysis for several reasons. Computer-based techniques to simplify or replace visual interpretation are especially needed in this modality and at the same time, they are quite feasible. Unlike in radiology, nuclear medicine images are commonly acquired in a digital form and can be used directly as an input to various software programs. Low resolution of images results in small matrix sizes of stored data, which facilitates fast image processing. Due to low resolution and the high level of noise, the images have a blurred appearance and lack definite anatomical details. Although such data usually convey less visual information than high resolution images, they can be difficult to interpret. These factors introduce a high level of subjectivity in interpretation, often higher than in other modalities. For example, it has been reported that the interpretation of Tl-201 heart scans by different observers showed significant differences in 20% and small disagreements in 33% of cases (11). On the other hand, the number of possible diagnostic outcomes in nuclear medicine image interpretation is usually small;

often only a few final findings are considered. For example, only a negative, positive, or undetermined diagnosis of a particular disease may be derived.

The representation of medical images has been traditionally two-dimensional (2-D) , for example display panel or film. Many types of nuclear medicine data have a complex, multidimensional character and require dedicated presentation techniques. A large series of images acquired sequentially in time, or three-dimensional (3-D) data pose a challenge during visual interpretation, as it is not always suitable or adequate to depict various patterns and changes in a form of a 2-D image. Although more difficult to interpret visually, 3-D data is easier to manipulate by computer software. By using the full 3-D information, algorithms do not have to consider the effects of projections of the 3-D quantity on the 2-D plane (which occurs in planar techniques). For this reason, 3-D scans usually provide a more definite localization of defects.

Image interpretation and pattern recognition in nuclear medicine is often related to image quantification. The use of image quantification to detect abnormalities usually involves a comparison to established normal ranges of count values in predefined image areas. By detecting the areas which deviate from the expected values, a quantitative estimate of abnormality can be made. Such an approach has been pursued by several investigators in the analysis of nuclear medicine scans of the heart (12) brain (13,14), and lungs (15). A multitude of schemes has been devised to divide, or compress images into predefined regions of interests (ROIs), subsequently establishing normal limits in these areas. Computerized analysis of cardiac images represents perhaps the most advanced application; a more detailed description of existing cardiac quantification methods is contained in section 6.3.

The results of quantitative techniques have been used as inputs to artificial intelligence systems, such as expert systems (7), or artificial neural networks (8). Most of these systems are designed to further process the quantitative data obtained from images, and

to provide some diagnostic answers. In particular, neural networks have been applied to the interpretation of brain (16), lung (17,18) and cardiac (19) images. Expert systems have been built to analyze quantitative results of lung (15) and heart images (20). Often, these systems reach beyond the information contained in the image and integrate other available clinical data to provide the final answers.

1.3 THESIS PROBLEM

1.3.1 Research goal and rationale

The goal of my research was to develop automated computer methods for the analysis and interpretation of nuclear medicine images in two areas of focus: nuclear medicine quality control (QC) images, and single photon emission computed tomography (SPECT) images of myocardial perfusion. This choice was dictated by the particular usefulness of computer-aided systems in both applications as well as the resources available at our institution. Although QC images and myocardial SPECT images differ in their origin and in the diagnostic information that they convey, similar methods of image analysis can be applied. QC images contain data pertaining to the performance of the imaging equipment and can be analyzed by comparisons to established technological norms. Myocardial perfusion images contain information related to the patient's physiology; the extraction of specific diagnostic answers can be accomplished by analyzing variations from the expected normal image appearance. In both cases, computer-aided image interpretation can simplify the diagnostic process by providing objective analysis.

The general need for the development of computer-aided image interpretation methods was stated in section 1.1. A more specific motivation to develop the automated diagnostic system for a gamma camera was the practical need for such an application in many nuclear medicine departments. Clinically oriented staff are often inexperienced in the complex technical procedures required to test the instrumentation and the proper implementation of QC protocols often poses a problem. Although in many hospitals the gamma camera manufacturer's technical personnel is responsible for the maintenance

of the equipment, the day-to-day operational checks remain the responsibility of the department. In many third world countries technical service is not readily available (21) and an automated diagnostic system would be especially beneficial. Technical support service could also use such tools to establish an initial diagnosis of equipment faults. Ultimately, automated QC analysis and assistance could be embedded in the manufacturer's software providing auto-diagnostic tests and simplifying the use of the equipment. Although there are many possible findings and diagnostic outcomes relating to the equipment function, QC images are relatively better defined than clinical images. Therefore, the goal of developing a computer-aided QC system seemed feasible and practical.

The motivation to develop software for an automated interpretation of myocardial perfusion SPECT images was also dictated by practical considerations. Myocardial perfusion imaging usually represents a large portion of the workload in a nuclear medicine department. The images are difficult to interpret by visual examination due to the complex character of the data (two separate 3-D datasets which have to be analyzed simultaneously), significant variations in the normal appearance of images, disadvantageous technical imaging factors, and the subtle character of clinically important perfusion defects. Thus, the analysis of myocardial perfusion images requires a high level of expertise and the operator's subjectivity is a particularly serious problem in this technique (11). On the other hand, only a few diagnostic outcomes are possible, which simplifies the implementation of the automated interpretation system. Another practical reason for choosing this objective was the fact that a large set of myocardial SPECT patient data necessary for this project was available at Victoria Hospital.

In both applications, a significant problem was image feature extraction and quantification. A large variety of gamma camera faults required the definition of many image features, which were essential for further logical analysis of image information. Although a variety of quantification methods have been developed for nuclear medicine applications, most

require interaction with an experienced operator. Methods for the elimination of these steps are needed in a completely automated diagnostic system. In particular, the interpretation of myocardial perfusion images relies on correct data alignment; therefore a method of reliable, automated image normalization to common coordinates was fundamental in this application. The application of these image processing techniques is twofold: they can be used independently as tools to describe images in a reproducible and objective manner, and the quantitative results can provide an input to the logical system.

1.3.2 Theoretical framework

To accomplish the goal of computer-aided image interpretation, many software methods and techniques have been utilized in this research; an important aspect was to integrate these various technologies. In the preprocessing phase, algorithms for image reorientation, shape matching, or image registration (22-24) were applied to match the actual image with its expected shape. Although similar methods have been previously developed for other applications, they had to be modified and optimized due to the specific character of nuclear medicine images (lack of well-defined edges, noise). Once the relationship between the image and its expected shape was established, the central mechanism in the formulation of a diagnosis was based on comparison to the reference data. Methods from the field of image segmentation and pattern analysis (25) were applied at this stage to assess image variations. In the case of QC images, these comparisons required additional sources of knowledge that could not be inferred from the images. The knowledge domain had inexact, qualitative character and the number of possible diagnostic findings was large. Therefore, artificial intelligence techniques were employed to encode further diagnostic processes. Expert system tools (5), and object-oriented knowledge representation (26) were utilized to build a full diagnostic system and to construct a framework of well-defined concepts. The theoretical problems in this application extended beyond image interpretation, into the domain of knowledge acquisition and knowledge representation (27,28).

1.4 THESIS OUTLINE

This thesis is organized into two parts, based on the two major objectives of my research. The first part describes methods for the analysis of gamma camera quality control images and the creation of an expert system for its application. The second part is focused on the analysis of myocardial perfusion images. Finally, chapter 9 contains the summary and description of future work. Various portions of this thesis have previously been published or have been submitted for publication. These parts are reproduced with permission of the appropriate publishers. The details of the publications and the description of the author's contributions are indicated in footnotes at the beginning of each chapter.

1.4.1 Automated analysis of quality control images

Chapter 2 of this thesis is the introduction to the first part. It describes the operation of a scintillation camera and the quality control techniques for this instrument. The first step in the creation of a diagnostic system that can analyze gamma camera QC images, was to develop a set of image processing algorithms (chapter 3). These algorithms derive a set of numeric parameters describing the QC images and can be used as an expert system input. A more detailed description of image processing algorithms and definitions of features is contained in appendix A. In the first attempt to develop a gamma camera QC expert system, a traditional rule-based system was created (chapter 4). This system, however, suffered from limitations due to an insufficient number of test cases and the complex structure of rules. To overcome these difficulties, a second study was undertaken (chapter 5). A large set of images representing a variety of faults was collected. Subsequently, a more refined prototype of the expert system was developed. The second design used an object-oriented approach, which resulted in an improved representation of knowledge and enhanced reasoning techniques. These features allowed rigid formulation and organization of various concepts used in the knowledge domain.

1.4.2 Automated analysis of myocardial perfusion images

The second objective of my research was to develop a computer-aided system for the analysis of myocardial perfusion SPECT images. As an introduction, chapter 6 reviews the technique of nuclear medicine myocardial perfusion imaging and discusses existing quantitative techniques. Chapter 7 describes a new method for automated alignment and sizing of 3-D heart images and the creation of normal reference templates. The normal templates provided a reference during relative quantification of the cardiac images. Subsequently, the automated fitting technique was employed to correlate spatially images of a larger group of patients representing specific sites of defects (chapter 8). Three-dimensional maps of expected defect sites were derived, based on correlating angiographic results. These maps, together with the normal templates were used in a refined regional quantification technique. Methods described in chapters 7 and 8 allow fully automatic detection and localization of coronary artery disease in cardiac SPECT scans.

REFERENCES:

1. Franken EA, Berbaum KS. Perceptual aspects of cardiac imaging. In: Marcus ML, Schelbert HR, Skorton DJ, Wolf GL, eds. *Cardiac Imaging*. Philadelphia: WB Saunders; 1991:87-92.
2. Smith MC. *Error and variation in diagnostic radiology*. Springfield, IL: Charles C Thomas; 1967.
3. Spodick DH. On experts and expertise: the effects of variability in observer performance. *Am J Cardiol* 1975; 36:529-535.
4. Links JM. Detection and comparison of patterns in images [Editorial]. *J Nucl Med* 1994; 35:16-17.
5. Kline PJ, Dolins SB. *Designing Expert Systems* New York; Wiley; 1989.
6. Hertz J, Krogh A, Palmer RG. *Introduction to the theory of neural computation*. Redwood City, CA. Addison-Wesley; 1991
7. Fox J, Walker N. Knowledge based interpretation of medical images. In: Viergever MA, Todd-Pokropek A, eds. *Mathematics and Computer Science in Medical Imaging* NATO ASI Series F Vol 39. Berlin: Springer-Verlag; 1988:241-265.
8. Scott R. Artificial intelligence: Its use in medical diagnosis. *J Nucl Med* 1993; 34:510-514.
9. Floyd CE, Tourassi GD. An artificial neural network for lesion detection on single-photon emission computed tomographic images. *Invest Radiol* 1992; 27:667-672
10. Maren A, Harston C, Pap R. *Handbook of neural computing applications*. San Diego, CA: Academic Press; 1990:224.
11. Trobaugh GB, Wackers FJT, Sokole EB, et al. Tl-201 myocardial imaging: an interinstitutional study of observer variability. *J Nucl Med* 1978; 19:359-363.
12. Garcia EV, Cooke CD, Van Train KF, et al. Technical aspects of myocardial SPECT imaging with technetium-99m sestamibi. *Am J Cardiol* 1990; 66:23E-31E.
13. Hooper HR, McEwan AJ, Lentle BC, Kotchon TL, Hooper PM. Interactive three dimensional region of interest analysis of HMPAO SPECT brain studies. *J Nucl Med* 1990; 31:2046-2051.

14. Houston AS, Kemp PM, Macleod MA. A method for assessing the significance of abnormalities in HMPAO brain SPECT images. *J Nucl Med* 1994; 35:239-244.
15. Gabor FV, Datz FL, Christian PE. Image analysis and categorization of ventilation-perfusion scans for the diagnosis of pulmonary embolism using an expert system. *J Nucl Med* 1994; 35:797-802.
16. Kippenhan JS, Barker WW, Nagel J, Grady C, Duara R. Neural-network classification of normal and Alzheimer's disease subjects using high-resolution and low-resolution PET cameras. *J Nucl Med* 1994; 35:7-15.
17. Scott JA, Palmer EL. Neural network analysis of ventilation-perfusion lung scans. *Radiology* 1993; 186:661-664.
18. Banish M, Datz FL, Clark R, et al. Application of neural networks to ventilation-perfusion imaging for diagnosing pulmonary embolism [Abstract]. *J Nucl Med* 1993; 34:176P.
19. Fujita H, Katafuchi T, Uehara T, Nishimura T. Application of artificial neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull's-eye images. *J Nucl Med* 1992; 33:372-276.
20. Herbst MD, Garcia EV, Cooke CD, Ezquerra NF, Folks RD, De Puey EG. Myocardial ischemia detection by expert system interpretation of thallium-201 tomograms. In: Reiber JHC, van der Wall EE, eds. *Cardiovascular nuclear medicine and MRI*. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1992:77-78.
21. Use and maintenance of nuclear medicine instruments in Southeast Asia. IAEA TECDOC-281. Vienna: International Atomic Energy Agency; 1983.
22. Krishnapuram R, Casasent D. Determination of three-dimensional object location and orientation from range images. *IEEE Trans Pattern Anal Machine Intell* 1989; 11:1158-1167.
23. Bhanu B, Faugeras OD. Shape matching of two-dimensional objects. *IEEE Trans Pattern Anal Machine Intell* 1984; 6:137-156.
24. Besl PJ, McKay ND. A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Machine Intell* 1987; 14:239-256.
25. Gonzales RC, Woods RE. *Digital image processing*. Reading, MA: Addison-Wesley Publishing Company 1992; 458-461.

26. Martin J. *Principles of object oriented design*. Englewood Cliffs, NJ: PTR Prentice Hall; 1993.
27. Kahn G, Nowlan S, McDermott J. Strategies for knowledge acquisition *IEEE Trans Pat Anal Mach Intel* 1985; 7:511-522.
28. Forsythe DE, Buchanan BG. Knowledge acquisition for expert systems: some pitfalls and suggestions. In: Buchanan BG, Wilkins DC, eds. *Readings in knowledge and learning*. San Mateo, CA: Morgan Kaufman Publishers; 1993:117-124.

PART I: AUTOMATED ANALYSIS OF QUALITY CONTROL IMAGES

CHAPTER 2: SCINTILLATION CAMERA: PRINCIPLE OF OPERATION AND QUALITY CONTROL

2.1 INTRODUCTION

The single crystal scintillation camera (or gamma camera, Anger camera) is a complex and delicate instrument, which detects and spatially registers gamma rays. The first scintillation camera was built in 1956 by Anger at the University of California (1). Clinically, a gamma camera is used as a diagnostic imaging device to depict the distribution of the radiopharmaceuticals introduced into a patient's body. The invention of the gamma camera revolutionized nuclear medicine technology. Since then, the design of the scintillation camera has evolved significantly due to improvements in components and integration of analog circuitry with digital systems.

Current scintillation cameras are similar in design, but there are many commercial models available. The operating characteristics of a gamma camera are susceptible to various fluctuations and drifts, which may degrade the overall performance; improper or suboptimal function may result in unacceptable clinical images. To ensure constant operation of the device within appropriate limits, frequent monitoring of the imaging capability is needed. These operational requirements necessitate the implementation of a comprehensive quality control (QC) program in nuclear medicine departments.

2.2 SCINTILLATION CAMERA

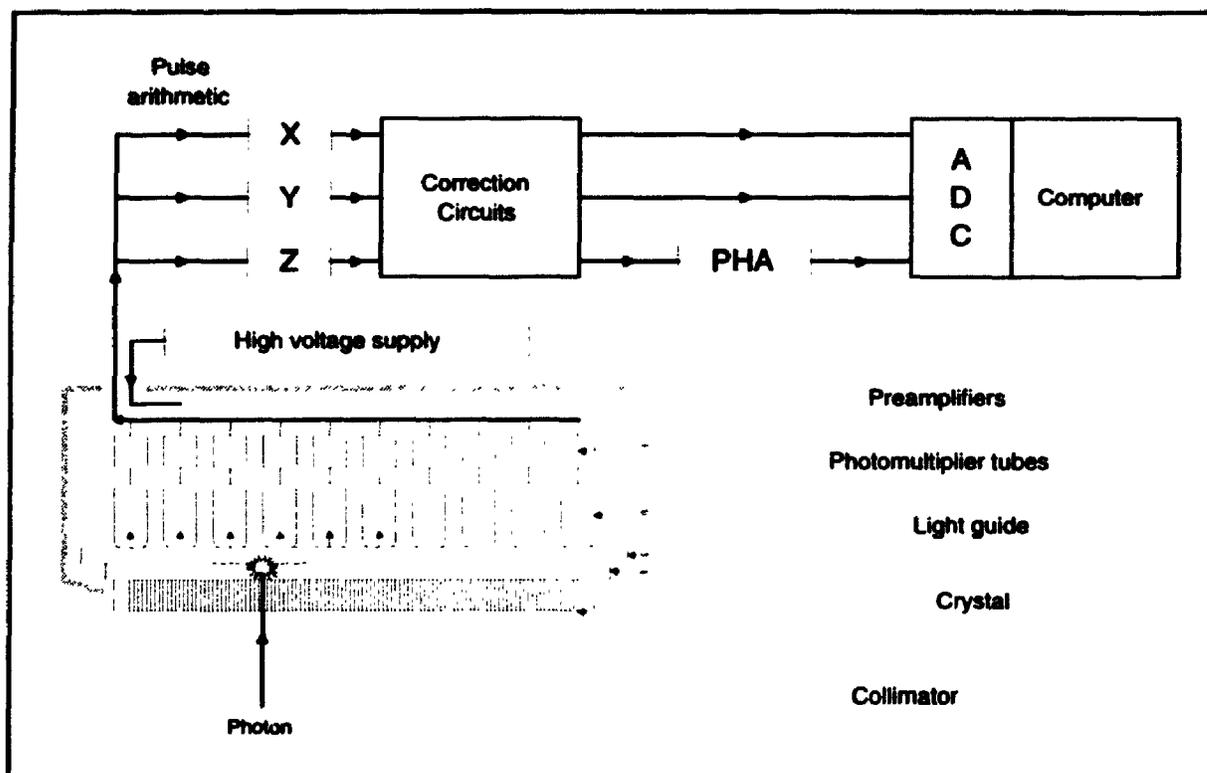


Figure 2.1: Diagram of a scintillation camera.

2.2.1 Basic principles

The principle of gamma camera operation is illustrated in Figure 2.1. A lead collimator is positioned between the radionuclide, which emits photons (gamma rays), and a scintillation crystal. The collimator is needed to relate the sites of photon absorption in the crystal with their origin in the patient. Gamma rays that are parallel to the collimator holes pass through the collimator and are absorbed in the crystal causing a scintillation event; other photons are absorbed in the collimator's lead septa. The most common type is the parallel collimator, which is composed of small parallel cylinders separated by lead or tungsten septa. Low-energy collimators designed for use with low energy photons have thin septa. Thicker septa, and therefore heavier collimators, are required for imaging radionuclides emitting high energy photons.

A large (up to 60 cm in diameter) NaI(Tl) crystal absorbs photons that passed through the collimator. The scintillation crystal is usually the most expensive and delicate part of the gamma camera. To prevent physical damage and to keep it free of moisture, the crystal is sealed in a protective enclosure. When a photon is absorbed in the crystal, a scintillation event occurs and a light impulse is emitted at the interaction site. This flash of light varies in brightness with the amount of energy deposited in the crystal by the photon.

The light signal is passed through a light guide that transmits the light to an array of photomultiplier (PM) tubes. Each scintillation event is registered by several photomultipliers. The photomultipliers that are closest to the origin of the event receive more light. Arrangements of PM tubes in the array vary, depending on the shape of the field-of-view; usually between 37 and 91 PM tubes are used in circular, rectangular or hexagonal configurations. The PM tube converts the light signal to the electrical signal, proportional to the amount of incident light. Each PM tube transmits its output electric signal to a separate preamplifier. The preamplifiers are used to tune the output PM signals, so all of them provide the same response to similar light impulses. In some cameras the gain of the PM signal is monitored and adjusted on-line by computer (2).

The amplified electrical signals enter the pulse arithmetic circuits, which encode the response of the photomultiplier array as X, Y position signals. A separate energy signal, Z, contains information about the energy of the incoming gamma ray. The energy correction circuit normalizes the energy signal with respect to spatial position. In many cameras, separate correction circuits for energy, linearity and uniformity adjust the levels of X, Y and Z signals (3). The energy signal Z is passed to the pulse-height analyzer (PHA), which classifies the signals according to preset energy limits (energy window). The output signal from the PHA circuit is used to exclude events generated by non-primary photons, such as scattered photons. Most scintillation cameras provide 3 or 4 independent energy windows.

If images are formed in a computer system, then the X and Y signals are passed through analog-to-digital converters (ADC) and then to the computer. Finally, the software increments the count value at the appropriate image matrix position. The number of events (counts) in the image varies, depending on the type of scan. Clinical images usually consist of 10,000 to 1 million counts. To avoid statistical fluctuations, quality control (QC) images may contain up to 100 million counts. The images can also be formed in an analog fashion on a radiographic film or a persistence scope.

2.2.2 SPECT systems

Single photon emission computed tomography (SPECT) systems collect radionuclide images from many angles. These multiple views are subsequently reconstructed using mathematical algorithms and a true three-dimensional (3-D) distribution of photon emission sources is found. A rotating detector provides a means to obtain a set of sequential images at various positions. The camera head usually rotates on a circular orbit, but elliptical or other paths are also possible (4). Images are normally acquired over a 360° or 180° head rotation, in a series of steps or in a continuous motion. The number of the angular positions influences the image quality and has to be optimized with respect to the total time of the procedure and count statistics (5). Typically, between 32 and 120 views are collected in a single SPECT acquisition. Dual (6) or triple (7) head configurations, collecting image data simultaneously, have been devised to increase the efficiency of image collection and improve image quality.

2.3 ASSESSMENT OF CAMERA PERFORMANCE

The assessment of camera performance can be divided into two categories: (a) reference and acceptance testing, and (b) routine operational checks. The purpose of reference testing is to establish a measure against which the device can be tested in the future or compared to other cameras. Extensive acceptance testing is necessary to evaluate new systems purchased by the user (8). Often, special devices and phantoms are required to carry out a full set of acceptance procedures. A comprehensive

evaluation of camera characteristics is useful at this stage, because it ensures that the device meets the quoted specifications. A set of standard protocols was designed by the National Electrical Manufacturers' Association (NEMA) in North America (9) and other organizations elsewhere (10,11). Most manufacturers' specifications of gamma cameras are based on NEMA tests. Therefore, NEMA tests are commonly used during acceptance testing of new equipment.

Once put into use, the performance of the gamma camera must be constantly monitored to ensure that any deterioration is immediately detected. Routine checks are usually based on a modified subset of NEMA protocols (12). The sophisticated instrumentation required for the acceptance testing is usually unavailable during the routine assessment of camera performance in a clinical setting. Routine testing needs to be simple enough to be done frequently, in a short time, and with minimal equipment (13). Operational QC tests are normally performed with technetium-99m (Tc-99m), as it is the most commonly used radionuclide during clinical procedures. Due to the relatively short half-life of Tc-99m (6 hours), sometimes, a cobalt-57 (Co-57) source (half-life 9 months) is used as a practical substitute. The most common QC test is flood-field uniformity. Other tests, which are done less frequently, include linearity and resolution tests, sensitivity and count-rate performance.

2.3.1 Flood-field uniformity

Flood-field uniformity is the most widely used test of scintillation camera performance (Fig 2.2). When imaging a uniform flux of radiation, areas of increased or decreased count density reflect camera malfunction. Such non-uniformities can be caused by a variety of gamma camera faults involving most of the camera components. Main reasons for degradation of the uniformity are: imbalanced PM tubes, failure of PM tubes, spatial non-linearities, crystal defects and deterioration, light guide separation from the crystal, incorrect setting of PHA window, defects in the collimator, ADC faults, and drifts in gains on the line amplifier or ADCs (10). The uniformity test is a simple and sensitive indicator

of gamma camera performance and it is recommended to be performed daily (10,12,13). This test may not be specific, however, and other procedures may be necessary to localize the fault, when non-uniformities are observed.

Uniformity can be measured with the collimator in place (extrinsic uniformity). A uniform flux of radiation is provided in such case by either a plastic Co-57 source with activity embedded in a flat shape, or a solution of radionuclide in a plastic container (phantom). Such a source is positioned directly on the collimator. The uniformity measured in this manner includes the contribution from irregularities in the collimator. If the uniformity is tested without the collimator (intrinsic uniformity), then a small radioactive source (point source), located at a distance from the crystal, is sufficient to provide a uniform photon flux to the detector.

The uniformity can be tested under varying imaging conditions. The uniformity image obtained with a modified PHA window may reveal subtle PM tube imbalance (14). The flood field uniformity, with an observed count rate of 75,000 counts per second, is recommended to assess high count performance (8). As a coincidental result, the sensitivity of the camera can be obtained by dividing the count rate by the activity of the flood source (10). The loss of sensitivity may reflect the degradation of energy resolution of the system or incorrect PHA energy calibration. The function of correction modules can be verified by acquiring flood images with and without the correction enabled (10).

Care must be taken during preparation of the liquid-filled phantom to ensure uniform distribution of the radionuclide and to avoid spillage on the camera face or the exterior of the flood source. During intrinsic uniformity testing the camera without the collimator is sensitive to external sources of radiation; these sources may corrupt the intrinsic flood image and should be removed from the area of the test. Often an image without any test source is acquired (background check) to identify the presence of external radiation.

Camera uniformity can be assessed visually or quantitatively. Visual assessment may fail to detect subtle non-uniformities; the quantitative approach is less subjective but there are several indices of uniformity established, and they may differ in sensitivity (15,16). Quantitative analysis requires computerized acquisition of the data. One advantage of quantitative analysis is that steady deterioration of the camera performance may be detected by analyzing a log of stored uniformity values (17). Several quantitative uniformity indices are described in section 3.2.1.

2.3.2 Spatial resolution

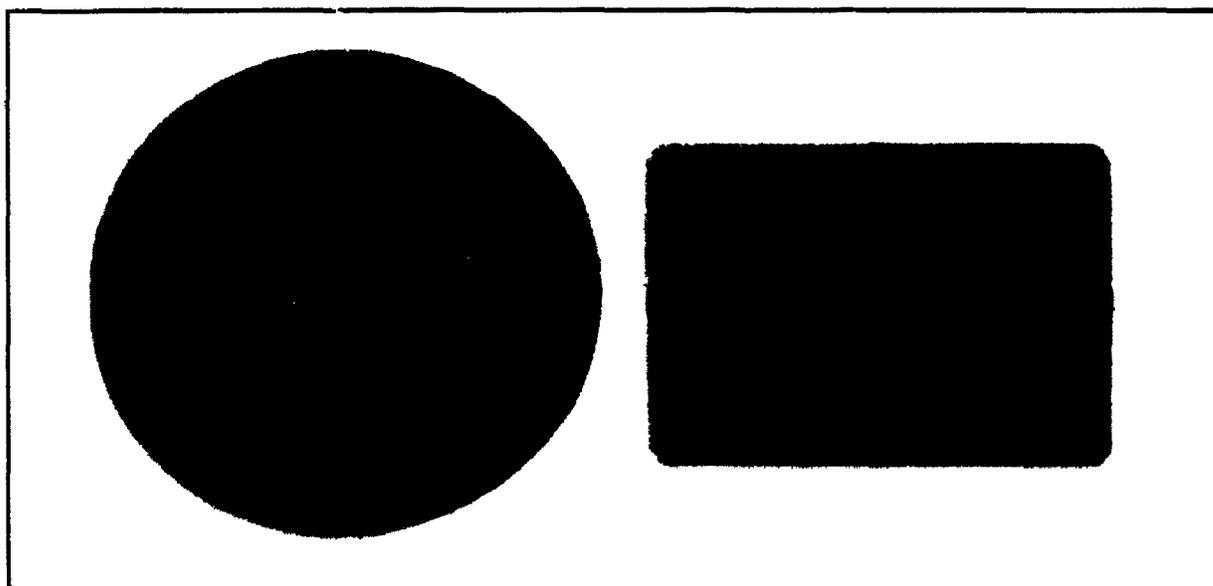


Figure 2.2: QC images of a flood-source (left) and four-quadrant bar phantom (right).

Changes in spatial resolution do not occur as frequently as changes in camera uniformity. Moreover, deterioration of spatial resolution is usually indicated by changes in uniformity (12). Therefore, spatial resolution is evaluated less often than uniformity, usually on a weekly or monthly basis. There are several types of phantoms available for spatial resolution tests. The NEMA resolution and linearity phantom (8) allows the quantification of the full width at half-maximum (FWHM) of the gamma camera line spread function. Other types of phantom, such as four-quadrant bar (Fig 2.2) are used to visually assess

the quality of the image. Resolution of the camera can be tested with and without collimator.

2.3.3 Spatial linearity

Linearity is commonly assessed by imaging similar types of phantoms, as in spatial resolution tests. Quantitative linearity parameters can be derived from the NEMA linearity and resolution phantom using specialized software; however, linearity is usually assessed by subjective visual interpretation of images. Degradation of spatial linearity is usually coupled with an increase in non-uniformity (18). The spatial linearity test is often combined with the resolution test, when using phantoms such as the four-quadrant bar phantom.

2.3.4 Other operational tests

Simple tests, such as checking the background count rate, energy calibration, or film processing equipment are commonly included in departmental QC protocols. Other tests are recommended by NEMA as routine checks, but they are rarely performed on a regular basis (13). A test of multiple-window spatial registration detects shifts in spatial registration of counts for different energies. The energy resolution test measures the accuracy of energy registration and consequently the ability to discriminate between scattered and unscattered radiation. The effect of an increasing count rate on system sensitivity is assessed by the count-rate test. In many institutions these tests have been substituted with more practical procedures for particular settings, which estimate equivalent camera parameters.

2.3.5 Performance of SPECT cameras

SPECT camera testing includes all the standard procedures as well as a few additional checks. Performance of SPECT cameras is more critical because non-uniformities are amplified during tomographic reconstruction. Thus, even subtle irregularities may result

in visible artefacts on clinical images (19). If a gamma camera is used for SPECT studies, the flood-field uniformity should be monitored carefully. Furthermore, the uniformity of the camera may depend on the angular position of the detector, and a check of uniformity at various angles is sometimes necessary (20).

Rotation of the camera head during SPECT acquisition must proceed along a predefined orbit. Any deviation from such a path may cause artefacts on reconstructed images. The mechanical precision of the rotation can be monitored by a test of centre-of-rotation offset (10). It is accomplished by acquiring tomographic images of a small source (point-source) positioned in a fixed location inside the camera orbit. Deviations of the point-source image positions from the expected geometrical path are then calculated. These data can be used to offset the position of acquired images electronically, or via software correction, thus cancelling the mechanical and electronic error. The rotational offset test is usually scheduled on a daily or weekly basis in most institutions.

Another common test, which measures the total performance of the SPECT system, is tomographic imaging of 3-D phantoms, for example the Jaszczak phantom (Data Spectrum Corp., Hillsborough, NC). The usual method of analysis is by visual inspection of the tomographic slice for artefacts and for overall image quality. Quantification of contrast and uniformity on the reconstructed image is sometimes utilized (10). Such tests, similar to the uniformity check, are sensitive to a variety of faults. Total performance phantoms are also useful in the assessment of reconstruction software, which contributes to the image quality.

Multi-head SPECT camera systems are tested in a fashion similar to the single-head systems, but the performance of each head has to be assessed separately. Moreover, the performance of the whole system and the alignment of images from multiple heads must be evaluated (21). Sometimes positioning of the radioactive test source is complicated by the detector geometry. Multi-head systems often require specialized

holders for the test-sources and standard protocols are modified to accommodate these configurations.

2.3.6 QC decision making

In a practical setting, QC procedures require decision-making, based on the test analysis, rather than on a fixed program (13). Apart from scheduled operational checks, new tests may be initiated to localize certain faults. Standard decision trees for the assessment of faults are desirable, but difficult to establish, due to the variety of available resources and specific departmental setups. A general flow chart of decisions in the QC testing process (Fig. 2.3) was proposed by Busemann Sokole (22). Encoding decision-making processes in QC procedures as expert system software is presented in chapters 4 and 5.

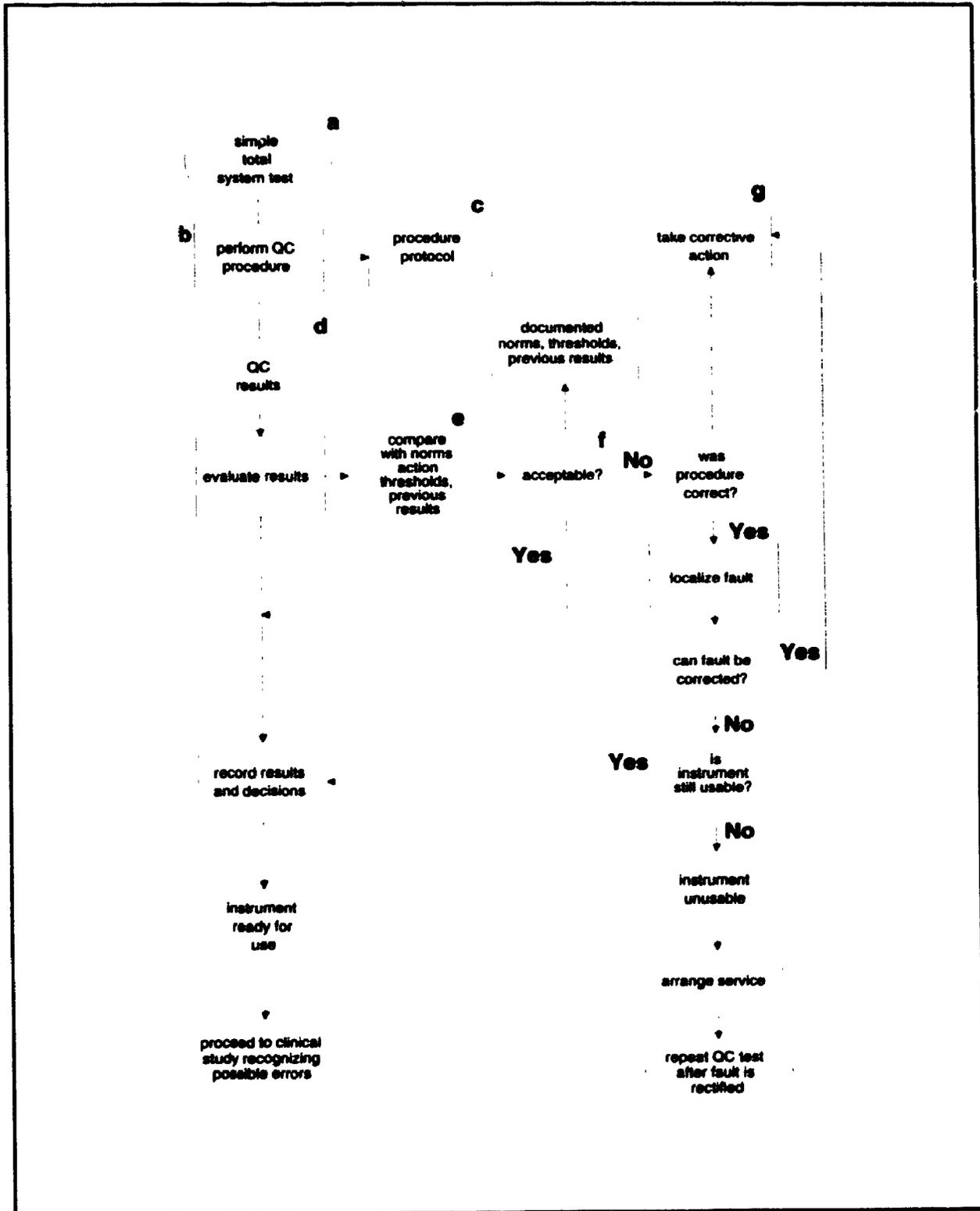


Figure 2.3: Decision tree suggested for execution, evaluation and follow-up of a routine quality control test. The symbols indicate: a, start or end of test; b, process to be performed, c, protocol; d, intermediate results; e, checks required; f, decision to be made; g, action command. Reproduced with permission from Reference 22.

REFERENCES:

1. Anger H. Scintillation camera. *Rev Sci Instrum* 1958;29:27.
2. Graham SL. Automatic tuning of scintillation cameras: a review. *J Nucl Med Tech* 1986; 14:105-110.
3. Simmons GH. On-line correction for factors that affect uniformity and linearity. *J Nucl Med Tech* 1988; 16:82-89.
4. Gottschalk SC, Salem D, Lim CB, Wake RH. SPECT resolution improvements by noncircular orbit. *J Nucl Med* 1983; 24:822-828.
5. Lim CB, Han KS, Hawman EG, et al. Image noise, resolution and lesion detectability in single photon emission CT. *IEEE Trans Nuc Sci* 1982; 29:500-505.
6. Jaszczak RJ, Chang LT, Stein NA, et al. Whole-body single-photon emission tomography using dual, large field-of-view scintillation cameras. *Phys Med Biol* 1979; 24:1123-1143.
7. Lim CB, Gottschalk S, Walker R, et al. Triangular SPECT system for 3-D total imaging: design concept and preliminary imaging results. *IEEE Trans Nuc Sci* 1985; 32:741-747.
8. Murphy PH. Acceptance testing and quality control of gamma cameras, including SPECT. *J Nucl Med* 1987; 28:1221-1227.
9. Performance measurements of scintillation cameras. Standards Publication No. NU1-1986. Washington, D.C.: National Electrical Manufacturers Association; 1986.
10. Quality control of nuclear medicine instruments. IAEA TECDOC-602. Vienna: International Atomic Energy Agency; 1991.
11. Quality control of gamma cameras and associated computer systems. Report No. 66. York, UK: The Institute of Physical Sciences in Medicine; 1992.
12. Raff U, Spitzer VM, Hendee WR. Practicality of NEMA performance specification measurements for user-based acceptance testing and routine quality assurance. *J Nucl Med* 1984; 25:679-687.
13. Craddock TD, Busemann Sokole E. Use of NEMA protocols for routine quality assurance [Letter]. *J Nucl Med* 1986; 26:95-96.

14. Graham SL, LaFontaine RL, Stein MA. Effects of asymmetric photopeak windows on flood field uniformity and spatial resolution of scintillation cameras. *J Nucl Med* 1986; 27:706-713.
15. Sharp P, Marshall I. The usefulness of indices measuring gamma camera non-uniformity. *Phys Med Biol* 1981; 26:149-153.
16. Hughes A, Sharp PF. The sensitivity of objective indices to changes in gamma camera non-uniformity. *Phys Med Biol* 1989; 34:885-893.
17. Raff U, Spitzer VM. Computerized monitoring of scintillation camera uniformity. *J Nucl Med Tech* 1988; 16:57-60.
18. Wicks R, Blau M. Effects of spatial distortion on Anger camera field uniformity correction [Concise communication]. *J Nucl Med* 1979; 20:252-254.
19. Todd-Pokropek A, Zurowski S, Soussaline F. Non-uniformity and artefact creation in emission tomography [Abstract] . *J Nucl Med* 1980; 21:P38.
20. Greer K, Jaszczak R, Harris C, Coleman RE. Quality control in SPECT. *J Nucl Med Tech* 1985; 13:76-85.
21. Kouris K, Clarke GA, Jarrit PH, Townsend CE, Thomas SN. Physical performance of the Toshiba GCA-9300A triple-headed system. *J Nucl Med* 1993; 34:1778-1789.
22. Busemann Sokole E. Quality assurance in nuclear medicine imaging: hardware and software aspects. PhD dissertation. Amsterdam: Rodopi; 1990.

CHAPTER 3: AUTOMATED FEATURE EXTRACTION AND QUANTIFICATION OF QUALITY CONTROL IMAGES: TOWARDS A KNOWLEDGE BASED SYSTEM*

3.1 INTRODUCTION

A highly automated, standardized protocol for testing the gamma camera performance is an important objective for several reasons. Automated Quality Control (QC) procedures seem well adapted for busy or inexperienced nuclear medicine centres. Often, the nuclear medicine staff has a medical, rather than technical, background. For this reason, the QC procedures, even if relatively uncomplicated, can present difficulties during the execution and interpretation of the results. A computer-driven QC procedure could enforce the use of strict test standards and protocols (1-3). It could also maintain a full record of QC results which can be valuable to service engineers in analyzing past performance of the equipment. If computer-aided diagnosis were incorporated into such software, it would provide troubleshooting assistance when the gamma camera malfunctions (4).

The first, and crucial, step in the QC procedures that requires significant human interaction and expert knowledge is the visual interpretation of the images. Thus, to accomplish automated image interpretation, the QC package must be able to describe quantitatively various features of the images. Based on the results of image analysis, and a comparison to established normal ranges, the troubleshooting and diagnostic decision algorithms may be further developed.

*The content of this chapter has been published as an abstract (Slomka P, Todd-Pokropek AE. A general purpose micro-computer based quality assurance package designed to be used under an expert system shell. *Eur J Nucl Med* 1990; 16:S83.) Andrew Todd-Pokropek inspired me to undertake this study and assisted me during the design of new image features. I wrote the software and conducted all experiments.

Typically, quantitative assessment of imaging equipment is done using some proprietary software packages specific to each camera manufacturer. It may be difficult to reliably compare the image quality of cameras from different manufacturers, due to certain ambiguities in the implementation of standard protocols (5). This could be a particularly important problem during the acceptance testing of a camera (6). If the automated QC protocol is independent of the camera and computer system, it would provide a means for the direct comparison of the equipment from different manufacturers.

With the ultimate goal of developing a fully automated, knowledge based system for QC procedures, we designed and implemented a software package which quantitatively analyses QC images. The package accepts flood images as input and estimates a comprehensive set of image features. To provide the widest coverage of existing imaging systems, all possible combinations of detectors and computers were considered. In particular, quantification algorithms were developed for a variety of gamma camera crystal shapes. Several new image parameters were defined in order to extract all the needed information from images, and to eliminate the need for a subjective visual assessment. These image parameters can be used to describe various gamma camera faults and are used in the QC expert system software, which interprets images at a higher, logical level. The design of this artificial intelligence system integrated with the image processing package is presented in chapters 4 and 5.

3.2 SYSTEM DESIGN

The overall system design is presented in Figure 3.1. This chapter discusses the derivation of image features (shaded box in Figure 3.1). The image processing algorithms and definitions of parameters are described in appendix A. The image processing routines are implemented in the C programming language on an IBM PC compatible computer. The package creates the report, which contains several numeric parameters describing the flood image. An important design criterion was to make the processing modules independent of the camera to be analyzed, and to have a common

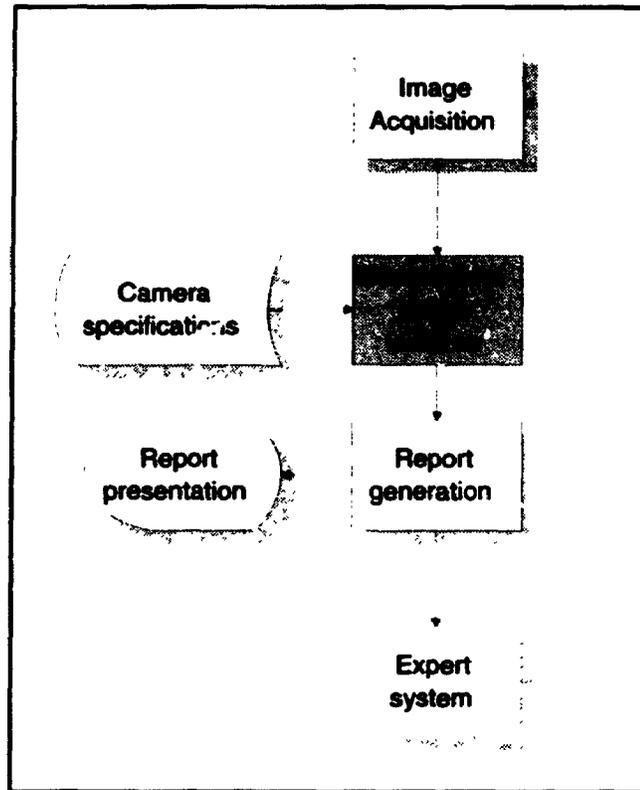


Figure 3.1: Overview of automated QC system.

format for reporting of the results. Some numbers, specifically the uniformity in central and useful field-of-view, are standard NEMA (National Electrical Manufacturer's Association) parameters (1). The system can analyze flood images from a variety of gamma camera field-of-view shapes such as circular, square, bevelled-rectangular (rectangle with cut corners) and hexagonal.

Several flood image features were defined and evaluated. The coefficients of variation, measured and theoretical (7,8), were calculated. The NEMA definitions for central and useful field-of-view were extended for the polygonal shapes. The linear dimensions in different directions were compared. The asymmetry of the shape, defined as a deviation from the regular predefined figure, was assessed and returned as a numeric parameter. Three types of edge distortion were defined. Hot and cold spot parameters were estimated. All these results form an array of numbers, which condense the essential

features describing the image and thus allow image interpretation without visual inspection by an expert. Derived image features are compared with a predefined configuration, which is kept as a separate database of camera-related information. The camera database contains the information for various types of cameras, such as the original dimensions, pixel sizes, tolerance limits, mode of operation etc. The software uses this information during evaluation of the QC images. In addition, a log of all previous results for all the examined cameras is stored together with the specification data.

3.2.1 Image Features

Edge detection. The first operation applied to the flood image is the edge detection algorithm. The result of this procedure is an ordered and chained numeric list (9), which contains all the point coordinates positioned on the edge of the flood. The chained representation of the edge is necessary for further manipulations. Edges in flood images are usually well defined and the background can be easily removed by a thresholding method. The edge detection algorithm is described in appendix A.1.

Shape fitting. After finding the ordered edge list, the appropriate geometrical figures corresponding to the camera field-of-view are shifted, rotated and sized to fit the edge of the flood. Similar techniques and algorithms used for various purposes have been reported (10-13). In this application, the fitting task is constrained to just a few shapes. These shapes, however, can deviate significantly from their normal appearance due to gamma camera faults. The current version of the program supports the following four different shapes of camera head: circular, rectangular or square, bevelled-rectangular (rectangle with cut corners), and hexagonal. The type of shape is known before the analysis because the camera information is recorded in the data header. This information is used by the shape-fitting routines simplifying the algorithms and increasing their robustness. The algorithms for shape fitting are described in appendix A.2

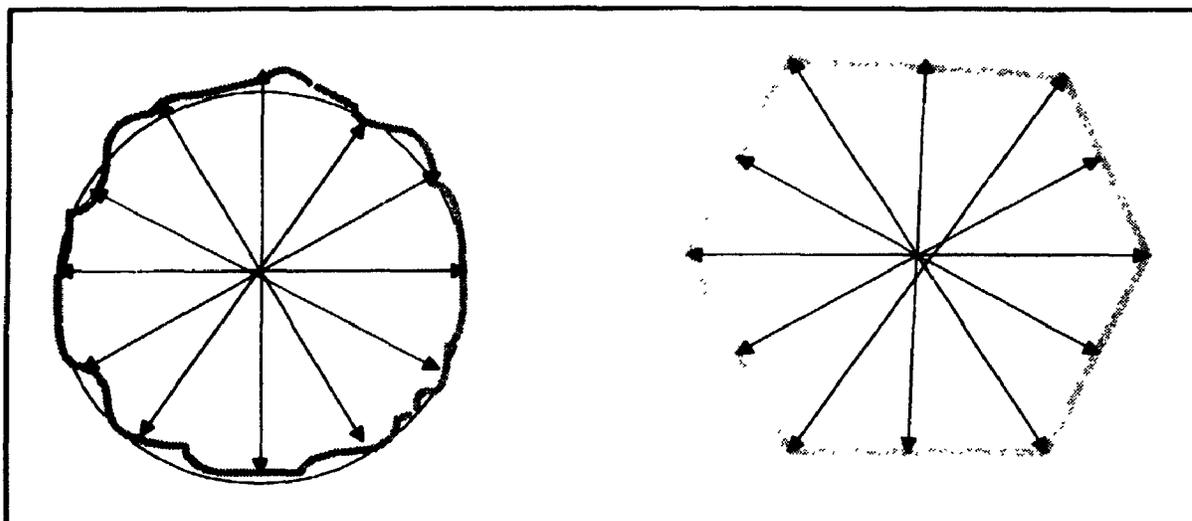


Figure 3.2: Dimension features derived from flood images. The circular field-of-view is defined by six samples of the diameter. The hexagonal field-of-view is specified by three longer diagonals and three distances between the opposite sides of the hexagon. The rectangular and bevelled-rectangular field-of-view is specified by two side dimensions and two diagonals. As additional parameters for polygons, all side dimensions are calculated.

Flood image dimensions. Flood dimensions are defined separately for different field-of-view shapes (Fig. 3.2). These dimensions can be used to monitor any changes in the image shape, X/Y zooms, and X/Y pixel sizes of the camera. Both the camera specification and the image header data contain the image pixel size information. These values are compared to each other; if there is a difference, a warning flag is set in the results indicating an incorrect calibration, and the image pixel size is used for further calculations. Subsequently, all the side-to-side and vertex-to-vertex absolute dimensions (or circle diameters) and relative sizes based on the camera specification are calculated.

NEMA results. The NEMA parameters are obtained for central and useful field-of-view. Standard NEMA definitions provide only the description of the circular and rectangular field-of-view shapes (1). Thus, we modified the definitions of central and useful field-of-view to include the hexagonal and bevelled-rectangular shapes. These field-of-view types were defined as sets of new vertices. Obtaining the useful or central field-of-view

parameters was equivalent to the geometric minification of an appropriate shape (appendix A.3).

Statistics Calculations. The statistics report included several parameters of the flood image, describing its noise and count distribution properties. The Poisson variance and measured variance were calculated (7,8). The Poisson variance (V_p) is defined as the mean count in the image

$$V_p = \sum_{x,y \in FOV} \frac{J(x,y)}{N} \quad [3.1]$$

where N is the total number of pixels in the image.

The measured variance (V_m) is given by

$$V_m = \sum_{x,y \in FOV} \frac{(J(x,y) - V_p)^2}{N} \quad [3.2]$$

where $J(x,y)$ is the image data and FOV is the field-of-view.

Two types of coefficients of variation, Poisson (C_p) and measured (C_m) and are calculated according to the following definitions (7,8).

$$\begin{aligned} C_p &= 100 * \frac{1}{\sqrt{V_p}} \\ C_m &= 100 * \frac{\sqrt{V_m}}{V_p} \end{aligned} \quad [3.3]$$

In total, four coefficients of variation are calculated: two for the useful field-of-view, and two for the central field-of-view. In addition, the mean values of counts and positions for

maximum and minimum count in the useful and central fields-of-view are evaluated. The other statistics related results are the total counts (T_c) of the flood and the area of the field-of-view.

Description of cold and hot spots. A recursive "region-growing" algorithm, similar to the algorithm described by Rosenfeld and Kak (14) was chosen to estimate hot and cold areas on the image. A description of this algorithm is contained in appendix A.4. Subsequently, the sizes, centroid positions and mean counts of such spots were calculated. The flood image, with marked spots, was displayed during the process for visual verification.

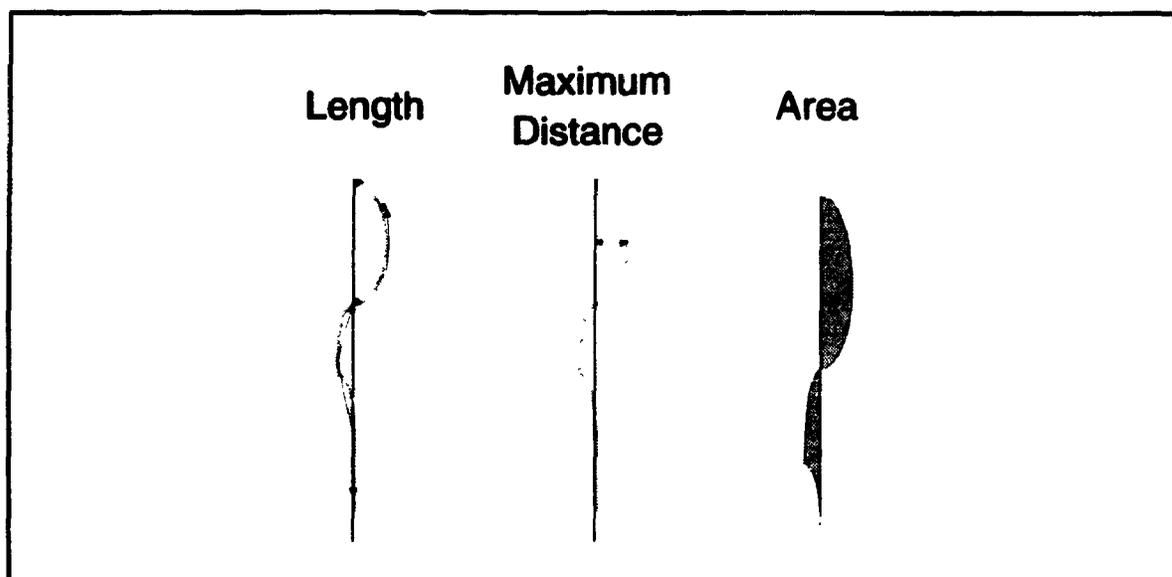


Figure 3.3: Three definitions of distortion. The first definition is based on the actual length of the contour as compared to the geometric length of ideal shapes (left). The second definition of distortion is based on the maximum distance of the fitted edge points from the ideal shape (middle). The third method uses the area of the difference between the ideal shape and the actual shape (right).

Distortion of edges. The distortion was defined as a measure of the difference between the edge contour and the polygon or circle that approximates its shape; distortion characterizes the irregularity of the edge. We developed three definitions of distortion and included all of them in the image description (Fig. 3.3). The distortion for rectangular or

hexagonal cameras is determined separately for each edge. A complete description of distortion parameters is contained in appendix A.5.

Asymmetry of shape. The asymmetry feature describes the difference between the image and its geometric specification. For example, an image zoomed in one direction should yield an increased asymmetry coefficient. The asymmetry coefficient was defined as a measure of dissimilarity between the *ideal* and the *closest-to-ideal* shape (appendix A.6).

3.3 RESULTS

3.3.1 Report Generation

All the parameters described in section 3.2.1 were used to describe the QC image. After the image features were evaluated, the software generated the report (Fig. 3.1). The flood image parameters, as determined by the image processing component, were compared with the predefined camera data retrieved from the camera-types database. The comparison was done for the dimensions and NEMA tolerance limits. The full report contained the acquisition parameters, all the results returned by the data analysis software, and the results of the comparisons defining the conformance to the predefined limits. The most important information was presented as a screen report. The results were also saved in an ASCII file, using a format similar to the Interfile (15) definition (Fig. 3.4).

```
number of pixels on edge := 0
coefficient poisson variation := 2.45
coefficient measured variation := .20
mean count:= 1677
uniformity integral := 4.78
uniformity differential := 5.23
central point coordinates := 32.17 33.50
central point offset := .17 1.50
coordinates of minimum value := 13 53
minimum count := 993
maximum count := 3209
size of cold spot := 1218.00
mean count in hot spot := 2279
size of hot spot:= 850.00
total distortion := 4.45
total counts := 5078376
asymmetry := 2.39
```

Figure 3.4: Sample Interfile keys in QC report, describing numeric features derived from images.

3.3.2 Flood and fault simulation

To assess the robustness and the usefulness of defined image features in the detection of certain faults, several types of flood images were tested. Hexagonal, circular, rectangular and bevelled-rectangular shapes were processed (Fig. 3.5). In total, 43 various flood images were processed. These flood images included several types of faults, which are listed in Table 3.1.

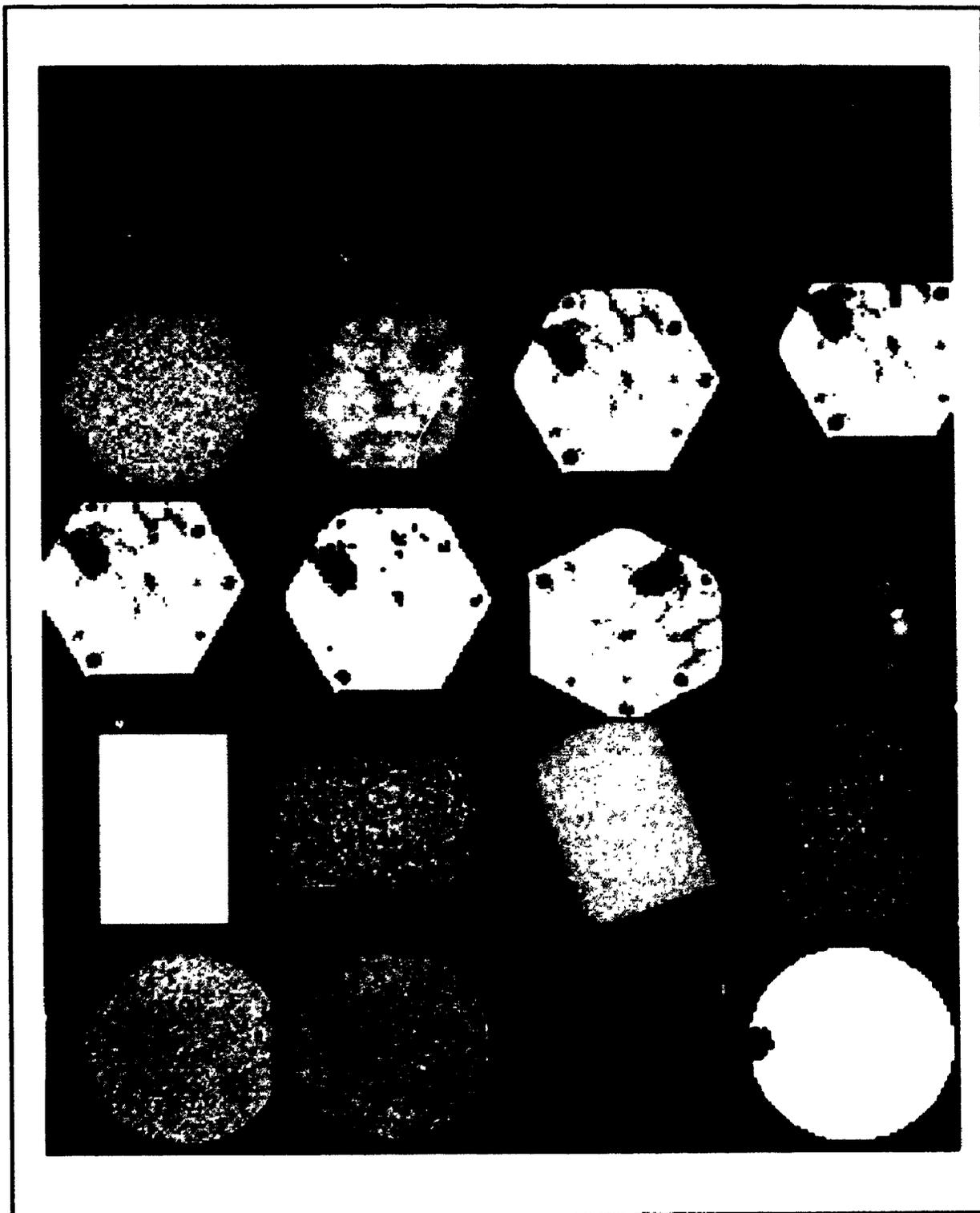


Figure 3.5: Various faulty flood images used to assess feature extraction algorithms. For example, flood image with deliberately switched off photomultiplier tube positioned on the edge (2nd column, 1st row), semicircular shape due to improper position of the flood image in gamma camera field-of-view (1st column, 2nd row), shifted image (4th column, 2nd row).

TABLE 3.1: List of simulated and real faults in the collection of floods.

photomultiplier faults in the centre
photomultiplier faults on the edge
detuned photomultiplier tubes
shifted position of the image (X, Y)
improperly positioned flood
low count images
disconnected X, Y cables
collimator contamination
external source contamination
wrong energy window

To obtain test-images the original acquisition data were used from the Scintronix gamma camera (Edinburgh, UK), Elscint SP-1 camera (Haifa, Israel), and Siemens (Hoffman Estates, IL USA) ZLC 7500 camera. The floods were acquired in 512x512 or 256x256 matrices, and then compressed to 64x64 for processing. As neither a rectangular nor a bevelled-rectangular field-of-view camera was available, these shapes were simulated (Fig 3.6). The original images were then shifted and rotated, to create test-images with parts of the contour positioned on the edge of the matrix. Additionally, the rotation of images was introduced for polygonal field-of-view types, to assess how the algorithms can cope with different orientations of the horizontal and vertical sides of the flood image.

Some examples of faulty images with semicircular and semi-polygonal shapes were tested (Fig. 3.5). This type of artefact can be caused by a smaller than field-of-view flood source, or the wrong position of the phantom with respect to the camera field-of-view. For

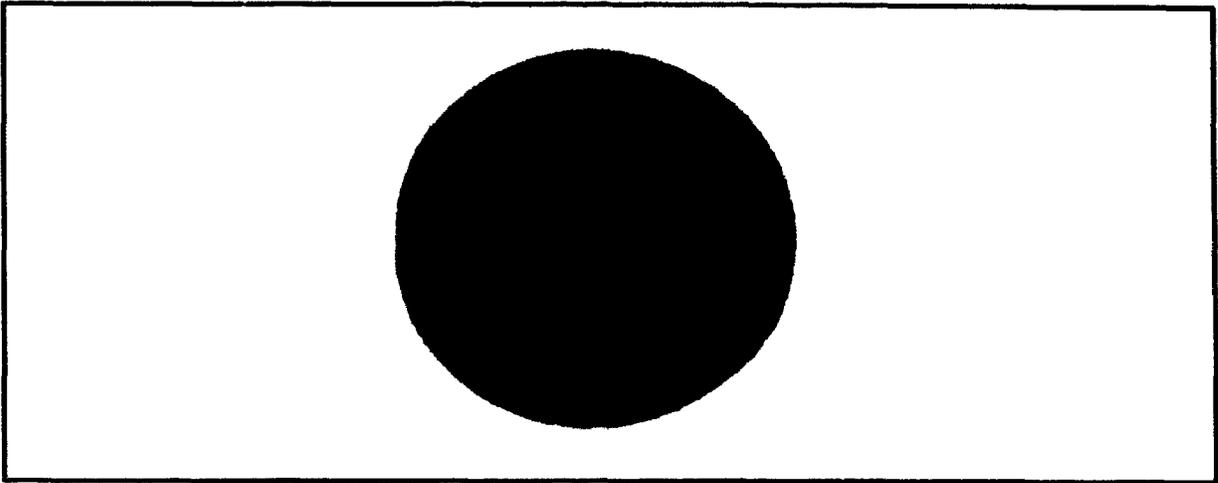


Figure 3.6: Generating simulated rectangular shape from circular flood image. The flood texture from the circular field-of-view image was magnified and clipped (white outline) to the rectangular or bevelled-rectangular outline.

the hexagonal camera it was possible to simulate the photomultiplier fault by disconnecting several tubes in different positions (Fig. 3.5). Low count images were acquired to assess the ability of the edge tracking algorithm. It was found that more than 300 K counts were required to ensure robust determination of the edge. The images with total counts below 300 K caused inaccuracies in the edge tracking program, as assessed by subjective visual interpretation. Typical QC images, however, contain more than 1 million counts. In all cases, the image parameters returned by the software reflected the anticipated image faults. When images were shifted to various positions, the centre-offset parameter changed. If the photomultipliers were disconnected, then the position and size of the cold spot correctly indicated the fault. The defects involving photomultiplier faults on the edge of the field-of-view were identified using the distortion coefficients.

The standard NEMA results were almost identical (< 0.5 % maximum difference) to the results obtained using the software on the *Elscint* system. Small differences arise from the fact that radii of the field-of-view were slightly different. The results for other shapes could not be compared with the manufacturer's calculations because the polygonal field-of-view was not included in the *Elscint* QC software. All three types of distortion parameters were lower than 10 % for normal images. Higher distortion coefficients were

related to gamma camera faults (Table 3.2). The area-based and maximum distance-based definitions provided better classification than the length-based approach.

TABLE 3.2: Use of distortion coefficients in detecting problems with faulty photomultipliers located on the edge of the field-of-view, and improperly positioned floods.

Type of flood image	Type of distortion coefficient		
	area	maximum distance	length
Normal (n = 10)	< 8%	< 5%	< 10%
Faulty photomultiplier on edge (n = 7)	> 15%	> 17%	> 11%
Misplaced flood image (n = 7)	> 12 %	> 10%	> 15 %

3.4 DISCUSSION

3.4.1 Rationale for the calculation of features

Conventionally, the quantitative evaluation of flood images is limited to the calculation of NEMA uniformity indices (2). Most of the other results are reported by subjective visual inspection of the flood images, or are obtained via some additional tests. It is possible, however, to estimate several other flood image parameters and use them to assess various aspects of the gamma camera performance. For instance, the initial assessment of X-Y gains (chapter 2) can be deduced from the flood dimensions, or asymmetry coefficients, for some types of cameras. The initial assessment of camera sensitivity can be accomplished using flood image acquisition parameters. It is possible to detect automatically the locations of faulty photomultipliers using cold spot information or distortion coefficients if the photomultipliers are positioned on the edge. The hot-spot description may be useful in identifying contamination problems. The analysis of noise properties can be accomplished by comparing various uniformity indices (8). The position parameters and the distortion coefficients can be used to detect improper acquisition geometry (Table 3.2). The uniformity check is performed on a daily basis and is one of

the simplest QC procedures. Moreover, it has been shown that the uniformity of the image is sensitive to several types of gamma camera faults (16-20). Therefore, maximizing the diagnostic information from this test is important, especially if it can be obtained automatically, without user intervention.

Since our goal was to avoid visual interpretation, some features evident to the observer had to be described numerically. An example of such a visually obvious feature is the flood-field image positioned partially out of the field-of-view. The interpretation of such an image is obvious to a human observer, but if meaningful results are to be derived without any user verification, the image processing component must be robust enough to cope with any possible shape deviations.

3.4.2 Quantification of other QC parameters

The image parameters were defined and implemented for the extrinsic and intrinsic flood tests, but their application is potentially wider. One possibility is to assess the flood images under varying conditions such as, for example, modified energy window of the pulse-height analyzer; the observed changes in image parameters could identify certain faults (21). Similar image parameters could also be derived for the transverse cross-sections of the uniform three-dimensional (3-D) phantom (22), to evaluate quantitatively the SPECT performance. Although a comprehensive number of parameters was designed, it may be necessary to further extend the set of image features, in order to describe other types of faults. For example, it has been reported that the Fourier analysis of QC images can provide the quantification of the analog-to-digital converter faults (23).

The feature extraction methods can be extended to the linearity and resolution tests using the orthogonal-hole transmission phantom (2), and in the case of SPECT to 3-D phantoms, such as the Jaszczak phantom (Data Spectrum Corp., Hillsborough, NC) and centre of rotation datasets (22). The expected distribution of activity is known in these tests and the actual data can be compared to predefined shapes. The appropriate image

features could be generated and then used in the interpretation of the results. One problem with these supplementary tests is that several types of specialized phantoms are available, and different sets of image features would have to be defined in each case. Moreover, some standard phantoms (1) are not suitable for the rectangular field-of-view cameras. Perhaps, the design of more standardized linearity-and-resolution phantoms would be necessary to establish common image features and image processing routines.

Quantifying the images of 3-D phantoms could be accomplished using techniques for fitting shapes in 3-D (24,25) to align and compare the actual phantom images to the expected radioisotope distribution. Perhaps the method described in section 7.2 could be adopted. Specifically, it is conceivable to use the synthetically generated predefined shapes of the Jaszczak phantom (26) and then apply a 3-D shape fitting technique to correlate this shape with the acquired phantom data. Then, using the known geometry of the synthetic data, it would be possible to derive a series of numeric estimates, such as uniformity and contrast indices at given locations. The use of similar techniques for clinical data interpretation is presented in chapter 7.

3.4.3 Use of image features with artificial intelligence software

The results of the quantitative image assessment are suitable as input data to higher level artificial intelligence (AI) software (27). For example, the report file can be used directly to create a knowledge frame (28) in the Prolog logical programming language, by converting the fields of the report (Fig. 3.4) into Prolog predicates (Fig. 3.7).

```
integral_nema_cfov (4.78)
central_point_coordinates (32.17, 33.50)
central_point_offset(.17, 1.50)
```

Figure 3.7: Examples of Prolog predicates corresponding to image features.

These predicates would form a knowledge frame, containing the QC image features important in the fault diagnosis. It would be possible to call the image analysis routines directly from the logical level. The expert system may also run autonomously and access previously generated image reports. In this way, the AI component would be free from the image processing tasks, for which currently available AI tools are not well suited (27). A similar design is described by Fox, et al. (29) and referred to as "image as a database." This technique can be applied to the QC phantom interpretation, because the number of parameters that can be derived from the images is relatively small. The design and implementation of the expert system, based on derived image features is presented in chapters 4 and 5.

ACKNOWLEDGEMENTS

This work has been supported by the International Atomic Energy Agency Vienna under Research Contract 4381/TC.

REFERENCES:

1. Performance measurements of scintillation cameras. Standards Publication No. NU1-1980 Washington, D.C.: National Electrical Manufacturers Association; 1986.
2. Quality control of nuclear medicine instruments. IAEA TECDOC-602 Vienna: International Atomic Energy Agency; 1991.
3. Quality control of gamma cameras and associated computer systems. Report No. 66 York, UK: The Institute of Physical Sciences in Medicine; 1992.
4. Todd-Pokropek AE. The design of a quality assurance package integrated into an expert system. In: Hofer R and Bergmann H, eds. *Radioactive isotopes in clinical medicine and research*. Stuttgart, New York: Schattauer; 1988:469-478.
5. Craddock TD, Busemann Sokole E. Use of NEMA protocols for routine quality assurance [Letter]. *J Nucl Med* 1981; 26:95-97.
6. Murphy PH. Acceptance and quality control of gamma cameras, including SPECT. *J Nucl Med* 1987; 28:1221-1227.
7. Sharp P, Marshall I. The usefulness of indices measuring gamma camera non-uniformity. *Phys Med Biol* 1981; 26:149-153.
8. Grossman LW, Anderson MP, Jennings RJ, Kruger JB, Lukes SJ, Wagner RF, Warr CP. Noise analysis of scintillation camera images: stochastic and non-stochastic effects. *Phys Med Biol* 1986; 31:941-953.
9. Gonzales RC, Woods RE. *Digital image processing*. Reading, MA: Addison-Wesley Publishing Company 1992; 458-461.
10. Sklansky J, Chazin RL, Hansen BJ. Minimum-perimeter polygons of digitized silhouettes. *IEEE Trans Comput* 1972; 21:260-268.
11. Pavlidis T. *Structural pattern recognition*. New York: Springer-Verlag; 1977.
12. Bengtson A, Eklundh JO. Shape representation by multiscale contour approximations. *IEEE Trans Pattern Anal Machine Intell* 1991; 13:85-93.
13. Feng HF, Pavlidis T. The generation of polygonal outlines of objects from grey level pictures. *IEEE Trans Circuits Sys* 1975; 22:427-439.
14. Rosenfeld A, Kak AC. *Digital Picture Processing* New York: Academic Press; 1982.

15. Maguire GQ Jr, Noz ME. Image formats: Five years after the AAPM standard for digital image interchange. *Med Phys* 1989; 16:818-823.
16. Simmons GH. On-line corrections for factors that affect uniformity and linearity. *J Nucl Med Tech* 1988; 16:82-88.
17. Muehlechner G, Colsher JG, Stoub EW. Correction of field nonuniformity in scintigraphic cameras through removal of spatial distortion. *J Nucl Med* 1980; 21:771-776.
18. Wicks R, Blau M. Effect of spatial distortion on Anger camera field-uniformity correction: concise communication. *J Nucl Med* 1979; 20:252-254.
19. Graham LS. Automatic tuning of scintillation cameras: a review. *J Nucl Med Tech* 1986; 14:105-110.
20. Todd-Pokropek AE, Erbsman F, Soussaline F. The nonuniformity of imaging devices and its impact in quantitative studies. In: *Medical Radionuclide imaging*, Vol II. Vienna: International Atomic Energy Agency; 1977:67-84.
21. Graham SL, LaFontaine RL, Stein MA. Effects of asymmetric photopeak windows on flood field uniformity and spatial resolution of scintillation cameras. *J Nucl Med* 1986; 27:706-713.
22. English RJ, Zimmerman RE. Performance and acceptance testing of scintillation cameras for SPECT. *J Nucl Med Tech* 1986; 16:132-138.
23. Borm JJJ, Busemann Sokole E. The use of Fourier techniques for the evaluation of a/d and d/a converter performance of scintillation cameras [Abstract]. *Eur J Nucl Med* 1990; 16:442.
24. Krishnapuram R, Casasent D. Determination of three-dimensional object location and orientation from range images. *IEEE Trans Pattern Anal Machine Intell* 1989; 11:1158-1167.
25. Bhanu B, Faugeras OD. Shape matching of two-dimensional objects. *IEEE Trans Pattern Anal Machine Intell* 1984; 6:137-156.
26. Todd-Pokropek A, Margetic N. The generation of synthetic data for testing SPECT computer systems [Abstract]. *Eur J Nucl Med* 1990; 16:442.
27. Luger GF, Stubblefield WA. *Artificial intelligence and the design of expert systems*. Redwood City, CA: The Benjamin/Cummings Publishing Company Inc; 1990.

28. Clocksin WF, Mellish CS. *Programming in Prolog*. 3rd edition Berlin, New York: Springer-Verlag; 1987.

29. Fox J, Walker N. Knowledge based interpretation of medical images. In: Viergever MA, Todd-Pokropek A, eds. *Mathematics and Computer Science in Medical Imaging* NATO ASI Series F Vol 39. Berlin: Springer-Verlag; 1988:241-265.

CHAPTER 4: A PROTOTYPE RULE-BASED EXPERT SYSTEM FOR QUALITY CONTROL OF A GAMMA CAMERA^b

4.1 INTRODUCTION

Quality control (QC) of a gamma camera in a nuclear medicine department involves several complex tasks that require experience. These tasks include acquiring test data, interpreting results, and following troubleshooting guidelines when test results are unsatisfactory (chapter 2). Improper quality control of the camera can result in unacceptable clinical images. Typically, a person with extensive experience in instrumentation analyzes the data, decides whether the device is working properly, and assesses the need for additional tests to diagnose occurring problems. Data acquisition protocols are relatively complex (1) and often present problems when routinely implemented. Finally, the total performance of the camera is assessed and decisions are made about the clinical use of the instrument; the clinical studies that would result in unacceptable artefacts should not be permitted. A flow chart of actions and decisions for execution, evaluation, and follow-up of a routine QC test was shown in chapter 2 (Fig. 2.3). Specifications of standard required tests are included in several technical documents (2-4). Guidance in fault diagnosis is scattered across various papers, books, and proprietary technical documentation (chapter 2). Some practical diagnostic techniques are not documented at all. Often, the procedures are empirical and are applied only with limited confidence, or are limited to the specific types of equipment.

^bContents of this chapter have been published (Slomka PJ, Craddock TD. An expert system for quality control. In: Höfer R, Bergmann H, Sinzinger H, eds; *Radioactive isotopes in clinical medicine and research. 20th International Symposium*, Stuttgart, New York: Schattauer;1993:91-95). It has been reproduced with the permission of the publishers. Trevor Craddock provided expert knowledge in the field of quality control. I designed and implemented the expert system software.

Hence, it is desirable to create an automated system to guide staff in acquiring and analyzing the necessary QC data (5). Such a system could streamline the QC process improving its overall efficiency and reproducibility.

Expert systems are generally defined as computer software that in some way replaces or aids a human expert by automated interpretation of data. A distinct feature of expert system software is that the information used to solve problems has a dynamic character and the decision path chosen by the software is not predetermined or "hard coded" into the body of the program. Such information, usually referred to as the *knowledge base*, is typically represented as a collection of rules and facts about a given knowledge domain (6). Quality control of scintillation cameras appears to lend itself to this kind of analysis. It would be difficult to use a conventional program in the overall analysis of the QC data because conditions and assumptions often change and a well defined, universal troubleshooting algorithm does not exist. Moreover, a considerable amount of information is specific to a particular manufacturer configuration, or even a departmental setup.

Previously, we developed the software package for automated feature extraction from QC flood-field images (chapter 3). In this study we designed a rule-based expert system that uses derived features as input parameters. During the system design, we attempted to automate and objectify qualitative image interpretation and formulation of expert judgements. The application of such a system could reduce the subjectivity of the QC test, reduce human errors, and standardize the QC process. The goal of the project was to determine the feasibility of such an approach to QC procedures in nuclear medicine.

4.2 SYSTEM DESIGN AND PROTOTYPE IMPLEMENTATION

To implement the expert system, we used a general purpose expert system toolkit *NEXPERT OBJECT* (Neuron Data, Palo Alto, CA), running under X-windows on a SUN SPARC 2 (Sun Microsystems, Mountain View, CA) computer. The use of a high level toolkit facilitated management of the user interface and accelerated the prototyping stage. A separate set of image processing routines (chapter 3) was integrated with the system, to derive quantitative description of QC images. During development of the expert system we used the simulated data that were created to test the image processing modules (chapter 3). We also conducted several interviews with the personnel responsible for QC tests, and with technical personnel, who provided service support for the equipment. Based on standard guidelines (2), we created a set of general rules, which defined some procedures in gamma camera QC. In the prototype design we considered the following input data: extrinsic and intrinsic flood-field images (chapter 2), and simulated numeric results of linearity and resolution tests (without image analysis).

4.2.1 General considerations and system requirements

The main objective of the system is to provide dual functionality: a routine guidance to the staff in quality control procedures on a daily basis, and troubleshooting capabilities. The routine task of the expert system is to determine whether the camera is usable for clinical studies. When a problem is detected, there are several possible scenarios. For example, the camera performance may be unsatisfactory and needs correction, but the system can still be used for certain clinical studies; or it might be totally unacceptable to proceed with any clinical studies and service should be called. If there is a problem, the expert system should try to localize the fault and decide whether it can be corrected by the user. If it can, the expert system should also be able to suggest the corrective action. After a user attempt to correct the fault, the expert system should be used again to confirm that the problem was indeed corrected. The QC expert system should be used on a daily basis to guide the user in operational checks, rather than specifically for fault diagnosis, when something goes wrong.

Conventionally, QC tests are assessed visually, and a human "expert" formulates a verbal opinion about the image. Most troubleshooting principles have this imprecise, qualitative character as well. Since an integral part of the QC process is image analysis, which requires professional knowledge, the expert system should be able to aid humans in this task. The most common gamma camera test, used to diagnose various faults, is the assessment of detector response to a uniform flux of radiation (flood-field uniformity). The description of the test and its diagnostic value is contained in chapter 2. Although several standard quantitative parameters can be derived from flood images (1), image artefacts are usually characterized by an informal and subjective description.

Poor QC image quality can result from a hardware fault. The knowledge about camera function deduced solely from QC images is usually incomplete. Often, supplementary information is needed to identify the source of a problem (for example voltage levels on photomultiplier tubes). This information can usually be acquired by additional equipment and requires service engineers to check the electronic hardware. Such testing cannot be done by clinical staff in the hospital for two reasons: (a) lack of technical qualification, and (b) lack of necessary equipment. Furthermore, the procedures for hardware testing are specific to each camera manufacturer and therefore cannot be included in a general purpose design. In these instances an initial identification of cause should be made by the expert system and the rest should be left for service.

Frequently, unacceptable QC results are the consequence of human errors during procedure preparation and setup. The expert system could provide full explanation and solution to problems in such situations (there is no hardware fault) and the error can be corrected by the user. Sometimes, the diagnosis is simple; for example, the problem can be directly deduced from the QC image header which records inappropriate settings and other relevant acquisition information. In other cases, problems are more camouflaged and image analysis is necessary, for example, when the radioactive test source is not prepared properly.

4.2.2 Rule development and implementation

The rule system was structured as follows. At the lowest level, all parameters relating to a particular type of test are considered for the most basic comparisons between test data and standard reference values for that test. The conclusions (hypotheses) drawn from the results of these comparisons reflect whether the test data are within acceptable limits, or how much they deviate from the standard values. Various elementary features are rated on a scale of *OK*, *bad*, or *alarming*. (Fig 4.1).

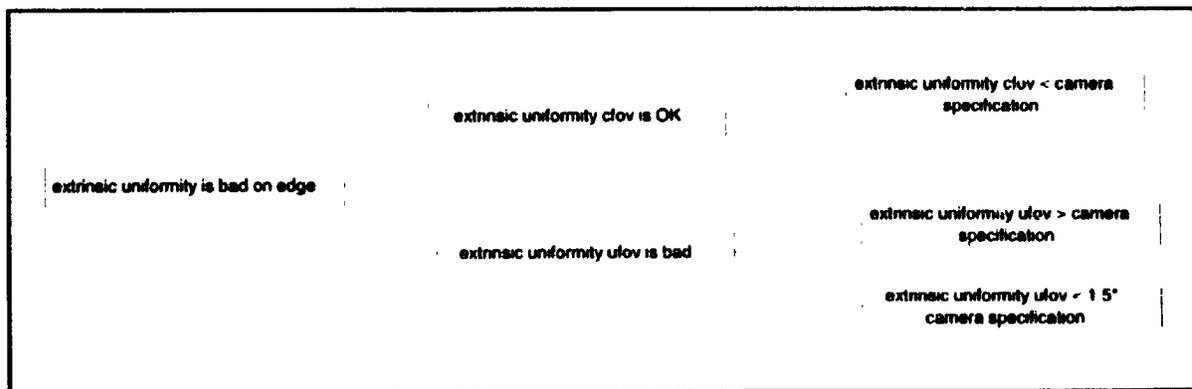


Figure 4.1: An example of a low level rule. The intermediate concept: *uniformity bad on edge* is established, based on the comparison of the test results to the camera specification values. The direction of information flow can be also reversed: the question about uniformity may evoke comparisons to the specification values.

At the next level, these basic features are used to formulate more complex features of the image, referred to as *intermediate concepts* (Fig 4.1). This process continues, limiting the possibilities under consideration, until the number of parameters has been reduced sufficiently and general categories such as *uniformity* or *spot type* can be evaluated. To determine whether the camera is usable, the top level rules use hypotheses drawn from evaluating these general categories. These hypotheses, as well as more specific ones, are used in the troubleshooting rules.

The rules are divided into five categories based on their function in the expert system. *Feature-interpretation rules* are low level rules interpreting the most basic features of the images. The request to derive conclusions of the *feature-interpretation* rules, leads to evaluation of image features using image processing routines written in the C

programming language (chapter 3). The results of these routines are shared with the expert system at the file level. *Acquisition-check* rules verify the information in the acquisition data header; these parameters could indicate discrepancies in acquisition procedures. *Scheduling rules* confirm a timetable for various tests. *Summary rules* are used to derive general conclusions about the usability of the camera. The rules containing general QC guidelines (summary rules, and scheduling rules) encoded standard QC protocols and procedures (3). Finally, *troubleshooting rules* were developed to determine the cause of poor image quality, and to suggest a corrective action. An example of a troubleshooting rule is shown in Figure 4.2.

Troubleshooting and acquisition-check rules were developed during interviews with technologists and service personnel. We also used previously acquired simulated-faults data (chapter 3) to design some of these rules. In total, 247 rules in all categories were developed; however, some feature classification rules had a repetitive character, assigning certain linguistic interpretations to obtained image parameters. In the troubleshooting category, 37 rules were developed.

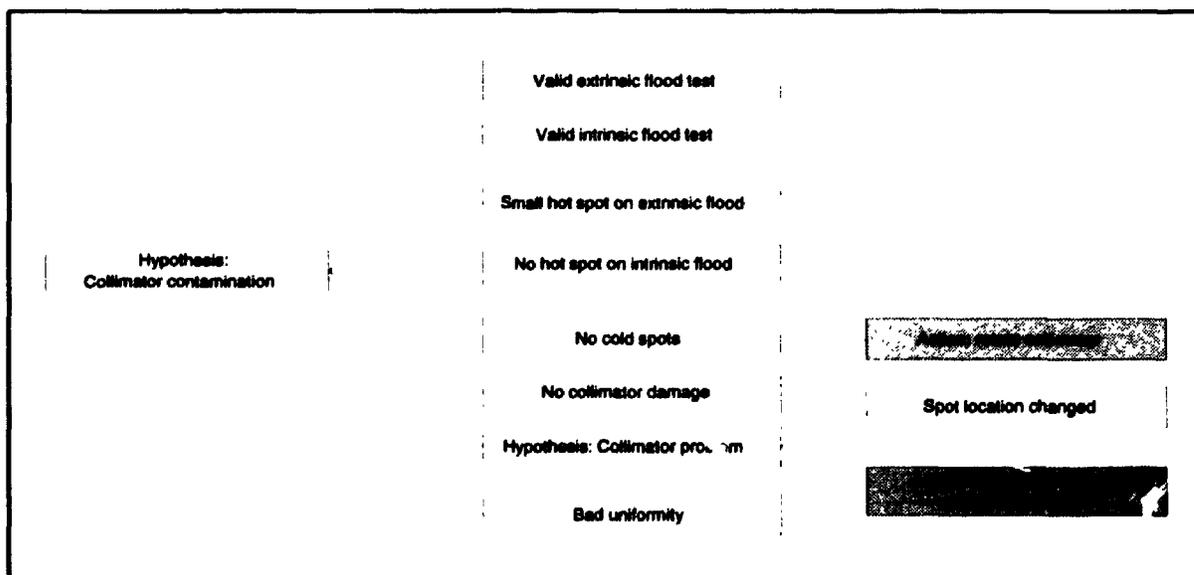


Figure 4.2: Example of a troubleshooting rule. The rule with the hypothesis *collimator contamination* has 8 premises. These premises are hypotheses of other rules. For example premise *collimator problem* is the hypothesis of another troubleshooting rule. Another condition *small hot spot in useful-field-of-view* on the extrinsic flood is an intermediate concept established with low level rules. Sometimes an action is necessary (gray boxes) to satisfy the hypothesis.

During interviews with technologists, we encountered difficulties in the proper designation of problem categories, and troubleshooting procedures, as the description had an intuitive character. Sometimes diagnosis involved a significant amount of common sense knowledge. Often, the fault was deduced based on past experience with the same, or similar faults. Faults that occurred most frequently on a particular camera were first on the probable hypothesis list. For example, one camera was very sensitive to high voltage fluctuations, and the question of power stability was usually raised when it malfunctioned. Faults often have an intermittent character, and many repetitions of tests are needed to reproduce and diagnose the fault. We did not include such unclear troubleshooting methods in this prototype; they are, however, commonly used in QC troubleshooting practice.

4.2.3 Inference mechanism

The system uses a mixed forward and backward chaining strategy (6) to implement the inference mechanism. A top level question about the usability of the camera has to be answered routinely by the expert system and it leads to a backward chain of inference. During this process subconditions and intermediate concepts are established, for example: *acceptable uniformity* or *acceptable linearity*. Several hypotheses about the state of the camera are evaluated. The user is then informed about the condition of the device. If the expert system detects unacceptable results, it attempts to determine the source of the problem. To verify a particular hypothesis, it may be necessary to repeat the test with changed parameters. For example, the flood source can be rotated by 90°, in order to check whether the spot on the image changed location. In such cases, the results of previous tests are compared to the new results. Troubleshooting reasoning is implemented in a forward fashion as opposed to the *"is it usable?"* backward chaining. If the system cannot verify camera usability, the subconditions and intermediate concepts evaluated in the backward chain are used as premises to the troubleshooting rules. During the forward chaining, all possible inference paths leading to all faults are traversed using the NEXPERT *forward-when-true* inference strategy (7). If a new test has been

4.3 DISCUSSION AND CONCLUSIONS

Troubleshooting rules are difficult to formulate for several reasons. They contain a considerable amount of common sense knowledge about gamma cameras and other general assumptions. In many cases, it is rather unclear which answers should be considered first. The human "expert" has some concept of the priority and likelihood of faults under certain circumstances. For example, if a service engineer attempted to fix a particular photomultiplier tube, and the image showed an artefact after the repair, the most likely explanation is that the artefact is caused by the same photomultiplier. Such intuitive short-cuts were not addressed in this prototype. During discussions with experienced staff regarding the fault identification, it was sometimes difficult to formulate rules in the *if-then* form. Instead, loose dependencies were expressed, for example: "usually small hot spots are caused by analog-to-digital converter faults but it depends what they look like." In addition, there are serious problems in deciding which basic image parameters can define complex features of the image, such as: *distinct straight lines, very hot small spots, horse-shoe artifacts*. These image features are routinely used to subjectively describe the QC flood-field images.

To provide the full assessment of camera function, additional quality control procedures such as count rate tests and sensitivity tests (chapter 2) have to be included. The possibility of a request for auxiliary tests during the expert system reasoning (when the test is a premise of a rule) requires the assessment of test feasibility. Therefore, the mechanism for finding the easiest, rather than the first, path to the solution can be implemented. Moreover, the probability of various conditions may be considered.

The developed system, despite covering only a subset of possible faults became very complex, containing about 250 rules. It became evident that in order for the expert system to be fully functional, a large number of rules must be designed. The conventional rule-based approach did not provide an organized structure of rules; such a monolithic representation caused difficulties in expanding the rule base and analyzing the system

performance. Thus, some other scheme of organizing the knowledge base is necessary to ensure expandability and to reduce the complexity of such a system. In chapter 5 we present a refined approach to the knowledge organization problem. On the other hand, the logical dependencies are relatively simple; therefore, the search and inference strategies used in this prototype are sufficient to describe the QC process.

It seems that the *a priori* approach to formulating rules is not sufficient. It should be feasible to build a database of cases representing several well-documented faults. Such a database can be used in "case-testing" of the system (8). Analyses of expert system failures to recognize certain faults should help in the formulation of new rules. It is conceivable that the rule database can be expanded and modified in a learning fashion. A comprehensive set of camera fault cases could also provide validation material for the expert system software. To acquire such a set of faults, a multicentre trial would have to be established since the collection of faulty QC images at one hospital is usually limited, due to the uniform type of equipment and small number of cameras.

Serious implementation problems exist in streamlining the data acquisition process and integrating it with the expert system. In most nuclear medicine camera-computer systems it is not easy to automate the acquisition procedure, and solutions to this problem would have to be system-specific. The expert system designed to run on a general purpose computer may not be able to control the acquisition on a proprietary nuclear medicine system. In the present design, the acquisition components require most human intervention. Even if the data are acquired, transferred to a generic computer, and then analyzed by the expert system, it may be necessary to initiate additional acquisitions in certain cases. These problems may be solved in the future by an "open system" approach to data-acquisition modules (9). Nevertheless, the ability to operate the system on a general purpose computer, even if the data acquisition is done on a separate unit is important, especially if the system is to be used in third world countries or in small centres.

4.4 SUMMARY

A prototype of a rule-based expert system for the quality control of gamma cameras was developed using a general purpose expert system toolkit. The system is designed to guide staff in quality control procedures, to interpret results, and to perform troubleshooting when a camera is malfunctioning. It is hoped that by using such artificial intelligence tools, problems can be solved and eliminated in many cases by an inexperienced user without the help of technical support. The main input data for the expert system are quality control test data such as uniformity images or linearity and resolution tests. By applying a series of image processing algorithms, various features and indices of images are extracted. These features are then compared to manufacturer specifications or expected results. The main question to be answered by the expert system is: "Can the camera be used for clinical studies?". If the results of the tests do not meet standards, the expert system initiates the troubleshooting reasoning. The limiting factors at this stage were the intuitive character of QC knowledge, the complexity of the rule-based system, and the small number of test-cases.

REFERENCES:

1. Raff U, Spitzer VM, Hendee WR. Practicality of NEMA performance specification measurements for user-based acceptance testing and routine quality assurance. *J Nucl Med* 1984; 25:679-687.
2. Performance measurements of scintillation cameras. Standards Publication No. NU1-1980 Washington, D.C: National Electrical Manufacturers Association; 1986.
3. Quality control of nuclear medicine instruments. IAEA TECDOC-601 Vienna: International Atomic Energy Agency; 1991.
4. Quality control of gamma cameras and associated computer systems. Report No. 66 York, UK: The Institute of Physical Sciences in Medicine; 1992.
5. Todd-Pokropek AE. The design of a quality assurance package integrated into an expert system. In: Hofer R, Bergmann H, eds. *Radioactive isotopes in clinical medicine and research*. Stuttgart, New York: Schattauer; 1988:469-478.
6. Kline PJ, Dolins SB. *Designing expert systems*. New York, Toronto: Wiley; 1989.
7. Turner SR. Nexpert Object. *IEEE Expert* 1991; 12:72-77.
8. Kolodner JL. Improving human decision making through case-based decision aiding. *AI Magazine* 1991; 12:52-68.
9. Slomka P: A review of nuclear medicine computer systems and cameras. *Nucl Geneeskd Bull* 1992; 141:16-21.

CHAPTER 5: KNOWLEDGE ACQUISITION, REPRESENTATION AND REASONING IN A GAMMA CAMERA QUALITY CONTROL EXPERT SYSTEM^c

5.1 INTRODUCTION

Gamma camera performance must be assessed frequently to ensure the acquisition of diagnostically reliable images (1-4). This is done by several quality control (QC) procedures (chapter 2). Usually, the QC tests involve acquisition and analysis of test images. The QC procedures are complex and require significant technical expertise, often not available in the clinical environment. Therefore, there is a need to build an expert system which would automate and standardize these tasks (5).

Previously, we developed a set of algorithms for the derivation of image features from QC images (chapter 3) and created a rule-based prototype of the expert system for automated QC assistance and troubleshooting (chapter 4). The goal of this study was to develop a refined, object-oriented design of such an expert system, overcoming the difficulties encountered in the rule-based approach. To categorize various QC problems and scenarios, we compiled a large database of faulty QC images and associated case histories. Thus, we created a database of cases consisting of QC images and associated

^cPortions of this chapter have been published (Slomka PJ, Chudziak JA, Mercer RE. Knowledge elicitation and processing in a gamma camera quality control expert system. In: Liebowitz J, eds. *Moving toward expert systems globally in the 21st century*. Elmsford, NY: Cognizant Communication Corp. 1994:66-71) and accepted for publication (Slomka PJ, Chudziak JA, Craddock TD. Knowledge acquisition, representation and reasoning in a gamma camera quality control expert system. *Radiation Protection and Dosimetry*. Ashford, UK: Nuclear Technology Publishing; 1994). These sections are reproduced with permission of the publishers. Jarek Chudziak assisted me in designing the reasoning modules for the expert system. Bob Mercer provided advice in knowledge representation design. Trevor Craddock offered his expert knowledge in many aspects of gamma camera quality control. I designed and implemented the expert system and performed all experiments.

problem descriptions. These cases were used in the process of formulating models for knowledge representation and reasoning used by the expert system.

The main obstacles in creating the expert system for this application were: the lack of precise documentation for troubleshooting methods, the large variety of possible gamma camera faults, and the informal character of image interpretation. To overcome these problems, we approached the initial process of gathering the information in an organized fashion. First, we created hierarchical generalization-specialization trees of various concepts such as image features, image artefacts, corrective actions, gamma camera problems, and QC test procedures (higher levels corresponding to more general concepts). This categorization was useful in building the knowledge representation model. Using these knowledge structures we built a prototype object-oriented expert system called *QCMAN*. The system was implemented using the *NEXPERT OBJECT* (Neuron Data, Palo Alto CA) expert system toolkit. The improved knowledge representation schemes allowed us to reduce the number of troubleshooting rules and to achieve high modularity of the system, thus eliminating some of the limitations of the previous rule-based approach.

The outline of this chapter is as follows: In section 5.2, we describe the process of creating the database of QC image faults needed to build the system. Section 5.3 contains the design of the expert system. A detailed example of a QC problem representation and diagnosis is presented in section 5.4. In section 5.5 the results of the preliminary testing of system performance are presented. Section 5.6 discusses the findings, and finally section 5.7 contains the conclusions of this study.

5.2 CREATING THE QC IMAGE DATABASE

We acquired a large database of gamma camera faults and associated case histories retrieved from old service logs and archived images. In order to compile an extensive and representative database, many hospitals and camera systems from different manufacturers were surveyed. When possible, an original computerized study was used and converted from proprietary to standard *Interfile* format (6). Most of the time, however, it was only possible to obtain a hardcopy of the images representing gamma camera faults. Such hardcopy film images were digitized using a high resolution video camera (*Video hi-8*, Sony) and frame grabber software (*Videopix*, Sun Microsystems, Mountain View, CA). The film images were digitized as 640x480, 8-bit greyscale images and subsequently converted to a 64x64 matrix size, and nuclear medicine *Interfile* format. In addition, for purposes of creating an electronic visual teaching file, the images were stored in the popular *Tif* image display format, maintaining the original image resolution.

Over a period of one year, 217 quality control flood images representing 154 cases of 53 different gamma camera faults and user errors were collected from nine institutions. The main categories of accumulated faults are listed in Table 5.1. Each fault case included: an informal description of the problem and relevant QC images with various artefacts. Specifically, the associated QC information included: data pertaining to the past performance of the faulty gamma camera, frequency of the problem (if it happened more than once), type of equipment, and the type of action taken to remedy the problem. Computer-acquired images were preferable to film images but in the majority of cases they were not available because many hospitals did not archive their QC images in the digital form. In particular, faulty images were rarely saved in the computerized format and it was only possible to obtain 32 digital image examples. A number of collected film images (n=35) represented problems related to the film processing or display equipment. We did not consider these faults during the creation of the expert system, which was designed to interpret digital data.

We encountered significant problems with the quality of the digitized data. Although it was possible to analyze quantitatively the shape characteristics of such images (for example, asymmetry, distortion, and dimensions), the statistical analysis of count variation was unreliable. The analysis of image histograms under varying scanning conditions revealed large variations (up to 50%) of relative pixel intensity due to external lighting and settings of scanning parameters. Quantitative analysis of digitized scans of uniform (as assessed visually) film images resulted in central field-of-view NEMA parameters of up to 20%, which normally indicates unacceptable image quality (2). Moreover, these digitized images did not contain the relevant acquisition parameters usually found in the data header of computer-acquired images; the description of cases contained only selected parameters. Therefore, we did not use digitized images during the quantitative analysis of image features. These images, however, provided a valuable reference during the development of the system. Perhaps more sophisticated scanning devices would permit a more reliable estimation of the count variation, but ideally these data should be available in the original digital format. A detailed analysis of quantification errors during digitization of images was reported by Boone et al. (7).

TABLE 5.1: Categories of collected gamma camera faults

Collimator defects
Signal electronics circuits
Crystal defects
Photomultiplier tubes
X, Y, Z signal cables
High voltage circuits
Radioactive contamination
Radioactive source
Incorrect acquisition settings
Digital data corruption
Display artefacts
Film processing artefacts

5.3 EXPERT SYSTEM DESIGN

5.3.1 System requirements

System requirements were described in section 4.2.1. Briefly, the purpose of the software is to help personnel with QC procedures, interpret the results and perform troubleshooting if necessary, thus acting as an intelligent QC assistant. The expert system has to be capable of automated quality control image interpretation, avoiding subjective visual inspection of the images. The system should be used on a daily basis during routine check-up rather than only for fault diagnosis.

5.3.2 Image processing component

Raw image data are not suitable for symbolic analysis by the logical system; therefore the images have to be preprocessed (8). The preprocessing phase estimates various image features. These features form a symbolic description of the image necessary for the analysis by the expert system. The image processing algorithms for feature derivation are described in chapter 3 and appendix A. When the values of features are

required by the reasoning module, the appropriate algorithms are executed. Image processing routines are written in the C programming language.

5.3.3 Knowledge representation

One of the most important tasks in any expert system design is thorough and systematic organization of knowledge (9). We used the accumulated database of QC images and associated case histories to derive an ordered taxonomy of QC concepts. This structured and precise classification of faults, features, and artefacts rendered the data in a form suitable for implementation in the object-oriented environment (10). Problems were categorized based on their location (i.e. part of the camera), and type (e.g. user error, hardware defect). Specific types of problems are represented as subproblems of more general concepts. For example a *collimator problem* and *crystal problem* were classified as a specialization of the statement *camera problem*.

Similarly, we organized a hierarchy of all possible image artefacts. We discovered that the interpretation of subjective "expert" terminology is imprecise; for example, images were referred to as *blotchy* or *gridlike*. Therefore, it was necessary to create a more ordered set of definitions. Sometimes, certain types of artefacts were linked invariably to certain problems and it was difficult to separate these two concepts. For instance, an image with the size of the cold area equal to the size of the photomultiplier tube immediately suggests a defect of this element. In other cases, however, artefacts could be symptoms of several problems and such immediate inference was not possible. Figure 5.1 illustrates a fragment of the artefact classification tree.

Image features were organized in a separate hierarchical format. The difference between image features and artefacts is subtle but very important. Features are numeric values derived from images; they describe the image without using high level concepts. Estimation of features does not require knowledge of the function of the camera. The concept of the artefact includes some information relating to the gamma camera function.

Therefore, artefacts are high level interpretations of image features. Furthermore, certain features may be linked to certain artefacts but there is no *one-to-one* relationship. A fragment of the feature classification tree is shown in Figure 5.2.

The NEXPERT *OBJECT* expert system shell provides the concept of class hierarchy that was used to implement our data structures (11). This tool allows the representation of hierarchical trees as classes and subclasses. Subclasses inherit their properties from the classes to which they belong, and they represent further specialization of various concepts. We built hierarchies of problems, artefacts, and features based on the models discussed above. The inheritance was used to propagate various properties and their values. Such a paradigm allowed reasoning on various levels of abstraction that corresponded to the hierarchical levels. In total, 37 features, 35 artefacts and 53 problems derived in the knowledge acquisition stage were encoded as NEXPERT *OBJECT* classes.

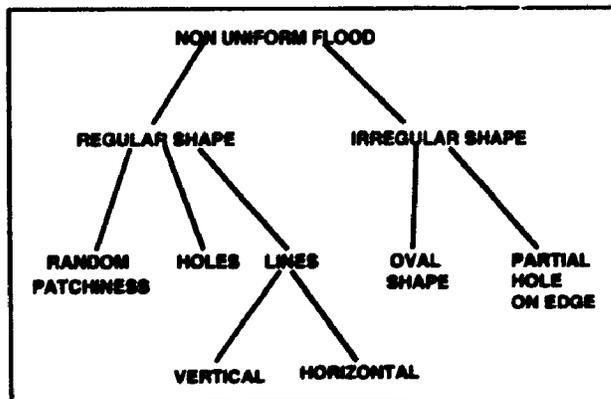


Figure 5.1: Fragment of hierarchical classification of image artefacts.

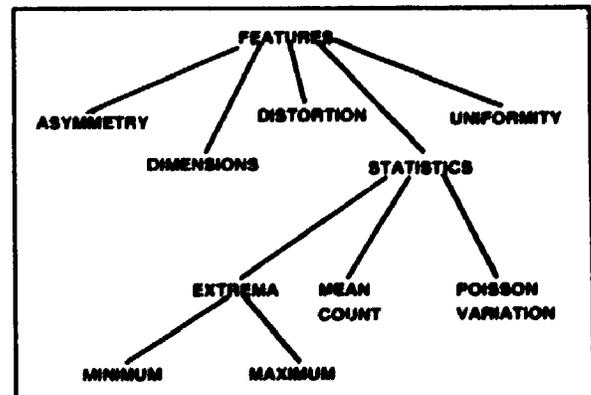


Figure 5.2: Fragment of hierarchical classification of features derived from QC images.

5.3.4 Reasoning strategy

The design of reasoning methods was strongly influenced by the capabilities of the NEXPERT *OBJECT* expert system shell (11). We tried to efficiently combine rule-based inference technology with modular object-oriented inheritance techniques in a fashion similar to that previously described by Torasso et al. (12). There are several mechanisms in NEXPERT *OBJECT* that can be used in the implementation of reasoning. These are: (a) backward and forward chaining methods of rule execution; (b) context links that allow loose association of groups of rules, otherwise not connected by premises or conclusions; (c) pattern-matching between objects; and (d) inheritance of properties. NEXPERT *OBJECT* provides control of the current reasoning strategy and allows temporary changes to strategies, which can be most appropriate at any given moment. We used several of these techniques to implement different aspects of the information flow in the knowledge structures. For instance, we created self-evaluating trees to compute current image feature values using inheritance combined with meta-slot procedures (procedures executed upon the request of the particular slot value).

The reasoning process starts from the elicitation of image features. If features are abnormal, the artefact tree is evaluated. Through pattern-matching governed by general control meta-rules (rules which define pattern-matching mechanism), the most probable preliminary hypotheses are established. To confirm these initial findings, it might be necessary to initiate some additional actions that would lead to the refinement of knowledge and consequently to improvement of the diagnosis. During this process the gamma camera history log, available resources, and feasibility of various actions are evaluated. The action may require interaction with the technologist. If, as a result of an action, additional images are acquired, a new feature tree is instantiated and the existence of artefacts is verified. Then a new set of possible hypotheses is chosen taking into account previous results. This cyclic process stops when one of the hypotheses is confirmed (Fig 5.3).

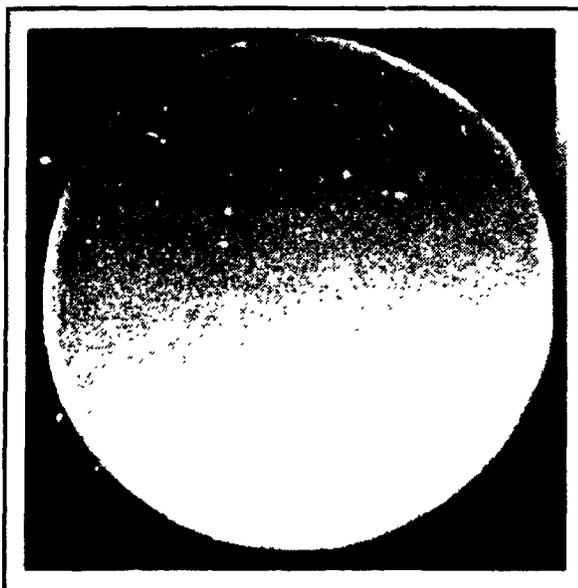


Figure 5.4: Example of a quality control image. The artefact is caused by the radioactive contamination of the floor near the gamma camera

The following is an informal description of the problem and the problem-solving methodology. During a daily QC check, the flood image quality is found unacceptable. Specifically, the NEMA uniformity value is outside the limits. Since a large area of higher intensity is noticed, isotope contamination or some background radiation can be suspected. The next step is to move the gamma camera away and repeat the test. During the second test, there is no artefact visible; therefore, it could be confirmed that indeed it was isotope contamination of the floor. If the same artefact was still observed, it would suggest that the problem is not external to the camera.

associated with a given artefact. A pattern-matching rule, which finds artefacts corresponding to evaluated features, is fired. The evidence of an artefact described as *high count region on side* is confirmed since it has the following features associated with it: (a) dimension is *OK*, (b) uniformity is *alarming*, (c) hot area is *big*, (d) hot area is *definite*, and (e) hot area is positioned *on side*.

There are several hypotheses linked to this artefact, namely: (a) *floor contamination*, (b) *collimator contamination*, (c) *camera face contamination*, (d) *source contamination* and (e) *uniformity correction faulty*. Each hypothesis has some initial evidence and a set of confirming conditions. If these conditions are fulfilled, then the hypothesis property *confirmed* becomes *TRUE*. First, a search is initiated for any confirmed hypothesis; if there are none, then the initial findings have to be verified through some user actions. These actions are listed in the *verify actions* property. Actions have a measure of feasibility. Subsequently, a pattern matching meta-rule finds the hypothesis with the most feasible confirming action. In the case of the floor contamination an associated action *move camera* is initiated. After the user accomplishes the task, a new QC image is acquired and its features are evaluated. In this second pass, the *feature interpretation* property becomes *OK*. The action object has the property *findings* which contains a list of linguistic concepts. In our example, the finding is *spot disappeared* since image features are *OK* now.

When the action is completed, a search for new confirmed hypotheses is initiated. This time, the hypothesis *floor contamination* is established because all its confirming conditions are satisfied. These conditions are as follows: (a) artefact *high count region on side* is present, (b) action *move camera* is done, and (c) action finding is *artefact disappeared*.

5.5 PRELIMINARY EVALUATION OF THE PROTOTYPE

In the current prototype, only a subset of all possible faults was considered. To evaluate the system performance, 30 flood images, which represented 10 different faults, were selected from the collected cases and the simulated images previously developed (chapter 3). The faults included: analog to digital converter gain and offset problems, contamination, photomultiplier defects and detuning, optical decoupling, overfilling the flood phantom, count overflow and wrong pulse-height analyzer settings.

The expert reasoning path and classification were analyzed for each of these images as explained in the example. In 28 cases, the diagnoses or suggested actions were correct. In two cases, the diagnoses were false due to erroneous classification of artefacts (photomultiplier on the edge and incorrect pulse-height analyzer setting). It should be possible to rectify such errors by improving the feature extraction algorithms. Since a small subset of possible faults was analyzed, further refinement and testing of the system would be required to consider all possible cases.

5.6 DISCUSSION

The QC fault database created in this study can be used for various purposes. Structured classification of gamma camera faults has not been previously documented; such an organized scheme could provide a valuable reference in QC procedures. Furthermore, since the images are stored in digital form, they can be used to provide on-line visual examples. This function could be incorporated into the expert system. We are investigating the possibility of providing an electronic distribution of such a database in the form of a teaching file. One limitation of collected film images is the lack of the acquisition parameters in the data header which could provide important information to the expert system. Furthermore, a reliable quantitative analysis and automated diagnosis of digitized film QC images cannot be accomplished.

In comparison to the rule-based approach described in chapter 4, the object-oriented design provided several advantages. The number of rules was significantly reduced as most of the information flow was contained in the data structures. The object-oriented approach provided a natural representation for the classification schemes. The system is easily expandable by introducing new components and relations to the existing data structures. The modular design allowed easy debugging and understanding of the logic. Although the prototype was constrained to the analysis of a limited number of faults, it contains fully developed structures that represent most concepts in this domain; therefore, it provides an expandable framework for a full-scale system.

The main obstacle in creating the expert system for this application was the knowledge acquisition process. With considerable effort, we accumulated a database of cases; some of these cases, however, could not be used directly in testing the system due to their non-digital character. The variety of gamma camera faults and user errors during QC procedures presents a challenge in designing the expert system. These faults can be manifested dissimilarly on different types of imaging equipment; therefore, generalization to all gamma camera types may not be valid in some instances. The collection of a comprehensive set of cases was difficult, because faults occur infrequently and tend to repeat themselves. A typical gamma camera may exhibit a problem only a few times a year at most. This difficulty is compounded by the fact that QC images showing equipment faults are rarely archived by clinically-oriented hospitals. Perhaps, a coordinated effort of gamma camera manufacturers and users would be needed to expand the collected database of faults. Furthermore, the interpretation of faults is subjective, not standardized, and it involves a significant amount of common sense knowledge not directly related to the images. In summary, the construction of the expert system for this application was difficult due to the *knowledge acquisition* bottleneck common in many artificial intelligence applications (13,14).

Routine application of the system presents some problems. The production version of the system would require a significant amount of *knowledge* engineering considering the variety of faults that can be encountered. In addition, it is presently difficult to interface some proprietary acquisition units with general purpose software (section 4.3). There are several new types of cameras available requiring specialized QC procedures (section 2.2.2); to support these cameras, further development of the expert system would be required. Sometimes QC tests that require supplementary phantoms are necessary (2). These phantoms are usually not standard, therefore they were not considered in this study. All these factors can pose practical difficulties in the routine implementation of a general purpose expert system for this application.

The experience in building the QC expert system can provide feedback to the standardization process. One important aspect of the knowledge acquisition process is the need for rigid formulation and systematic organization of various concepts used in the reasoning in the particular domain (13). The knowledge organization was an important step on its own; the exact formulation of concepts simplifies the knowledge domain and reduces the arbitrary character of the interpretation. Thus, the process of knowledge acquisition may provide the necessary background to enhance and objectify standards. For example, the classification criteria derived in this study could be used for such purposes.

5.7 CONCLUSIONS

We built a prototype expert system for gamma camera quality control. We approached knowledge acquisition in an organized fashion and attempted a rigid classification of various concepts, which have been used previously in an informal way. Thus, we created an organized database consisting of quality control images that represented various camera faults. Automated image processing algorithms for deriving features from images were incorporated in the system. The knowledge representation used by the system was object-oriented providing modularization, better control of the reasoning, and natural

representation of knowledge. Although we foresee some practical difficulties, we hope that such an expert system could be used by non-technical personnel in centres without adequate technical support, for example, in third world countries.

ACKNOWLEDGMENTS

We acknowledge the active participation of nuclear medicine technologists Alice Yun Sang and Janice Stephenson in the knowledge acquisition process. This research was supported in part by NSERC Grant 0036853 and KBN Grant 507/032/277/1.

REFERENCES:

1. Performance measurements of scintillation cameras. Standards Publication No. NU1-1980. Washington, D.C: National Electrical Manufacturers Association; 1986.
2. Quality control of nuclear medicine instruments. IAEA-TECDOC-602. Vienna: International Atomic Energy Agency; 1991.
3. Raff U, Spitzer VM, Hendee WR. Practicality of NEMA performance specification measurements for user-based acceptance testing and routine quality assurance. *J Nucl Med* 1984; 25:679-687.
4. Craddock TD, Busemann Sokole E. Use of NEMA protocols for routine quality assurance [Letter]. *J Nucl Med* 1986; 26:95-96.
5. Todd-Pokropek AE. The design of a quality assurance package integrated into an expert system. In: Hofer R, Bergmann H, eds. *Radioactive isotopes in clinical medicine and research*. Stuttgart, New York: Schattauer; 1988:469-478.
6. Maguire GQ Jr, Noz ME. Image formats: Five years after the AAPM standard for digital image interchange. *Med Phys* 1989; 16:818-823.
7. Boone JM, Seibert JA, Barret WA, Blood EA. Analysis and correction of imperfections in the image intensifier-TV-digitizer chain. *Med Phys* 1991; 19:236-242.
8. Fox J, Walker N. Knowledge based interpretation of medical images. In: Viergever MA, Todd-Pokropek AE, eds. *Mathematics and Computer Science in Medical Imaging NATO ASI Series F Vol 39*. Berlin: Springer-Verlag; 1988:241-265.
9. Kahn G, Nowlan S, McDermott J. Strategies for knowledge acquisition *IEEE Trans Pat Anal Mach Intel* 1985; 7:511-522.
10. Martin J. *Principles of object oriented design*. Englewood Cliffs, NJ: PTR Prentice Hall 1993;169-186.
11. Turner SR. Nexpert Object. *IEEE Expert* 1991; 12:72-77.
12. Torasso P, Console L. *Diagnostic problem solving: combining heuristic, approximate and casual reasoning*. New York: Van Nostrand Reinhold; 1989:91-121.
13. Forsythe DE, Buchanan BG. Knowledge acquisition for expert systems: some pitfalls and suggestions. In: Buchanan BG, Wilkins DC, eds. *Readings in knowledge and learning*. San Mateo, CA: Morgan Kaufman Publishers; 1993:117-124.

14. Olson J, Reuter HH. Extracting expertise from experts: Methods for knowledge acquisition. *Expert systems* 1987 4:152-168.

PART II: AUTOMATED ANALYSIS OF MYOCARDIAL SPECT IMAGES

CHAPTER 6: MYOCARDIAL PERFUSION SPECT IMAGING

6.1 INTRODUCTION

Coronary artery disease (CAD) is a major cause of death and morbidity in North America and Europe (1). CAD is caused by narrowing or blockage of the coronary arteries; it may manifest itself in a given patient as myocardial infarction (tissue damage) or transient myocardial ischemia (decreased blood supply). The symptoms of ischemia (exertional chest pain) are frequently atypical or may even be absent. Usually, in the presence of coronary artery stenosis without infarction, blood flow to the myocardium is normal when a patient is at rest. Under stress, however, stenotic vessels do not respond adequately to the increased oxygen demand, which requires increased blood flow. Therefore, under stress, regions of the myocardium involving diseased vessels exhibit a relative decrease in perfusion, which relates to symptoms and adverse effects in patients. Thus, depiction of myocardial perfusion in CAD patients at stress and rest provides valuable diagnostic and prognostic information (2).

A widely used and accepted technique for evaluation of myocardial perfusion is nuclear medicine scintigraphy. Planar or single photon emission computed tomography (SPECT) imaging can be performed. Planar imaging suffers from limitations in ability to localize defects due to superimposition of counts from many locations on one view. SPECT imaging characterizes a three-dimensional (3-D) distribution of the radiotracer and provides better separation of overlapping myocardial regions. Cardiac perfusion images

are acquired during rest and stress conditions. Coronary artery disease can be detected by analyzing defects on stress images, which may relate to either infarction or ischemia. Detection of a significant defect during stress, which is not present or less severe at rest is the hallmark of ischemia. Such a finding provides a presumptive diagnosis of CAD, helps explain the patient's symptoms, and predict future cardiac events (2,3).

In contrast to other tests, such as stress echocardiography or electrocardiography used to detect CAD, myocardial scintigraphy provides a direct measure of perfusion which yields additional diagnostic and prognostic information (4). Myocardial perfusion SPECT may in some cases preclude the need for coronary arteriography (angiography) in the detection of CAD; in other cases these techniques may be complementary. Coronary angiography is a roentgenographic technique that visualizes arteries after an intravascular injection of contrast material and it requires cardiac catheterization. The angiographic test can define vessel morphology accurately and is used as a "gold standard" in the assessment of CAD (5); it also forms a basis for treatment of CAD (6). Nonetheless, angiography cannot measure the effect of the lesion in the artery on the perfusion of the myocardial tissue (7). Furthermore, it is an invasive procedure associated with certain risks and complications (8) and therefore is not warranted when symptoms are not severe; in such cases myocardial perfusion SPECT can provide a means for an objective and noninvasive screening of patients.

6.2 IMAGING TECHNIQUES

6.2.1 Imaging agents

The most established agent for myocardial SPECT perfusion tests is thallium-201 (9). Other radiopharmaceuticals with better imaging characteristics are becoming routinely used. Technetium based agents, Tc-99m-sestamibi (10,11) and Tc-99m-teboroxime (10,12) have been shown to provide advantages over Tl-201; in particular, Tc-99m-sestamibi is used clinically in many centres. Tc-99m radionuclide emits photons of higher energy (140Kev) than Tl-201 (60 to 83 Kev), which is more suitable for a typical

gamma camera. The other benefits of Tc-99m imaging are less scatter and attenuation, and greater activity that can be injected due to the shorter half-life. These features result in improved spatial resolution of images. Moreover, since most other scintigraphic procedures are performed with Tc-99m, the use of agents based on this radionuclide is simpler in clinical practice (10).

6.2.2 Imaging protocols

Imaging protocols differ, depending on the type of agent. In Tl-201 studies stress images are acquired first, 10 minutes after exercise, and rest images are collected 3-4 hours later allowing time for re-distribution of the radionuclide (9). If Tc-99m-sestamibi is used, two separate injections are required, typically with a smaller dose injected at rest and larger dose injected several hours later during stress (11). Some researchers developed dual isotope Tl-201/Tc-99m techniques to acquire stress and rest images simultaneously (13,14), or sequentially (14,15), which has the potential to simplify the study protocol.

Stress studies are often difficult, due to patients' inability to exercise to the appropriate level, or lack of cooperation. Submaximal levels of exercise may result in false-negative tests. In such situations it is necessary to increase myocardial perfusion pharmacologically. Agents used for this purpose are dipyridamole (16), adenosine (17) and dobutamine (18). These agents differ in the type of exercise simulation they provide, their effect on image quality, and also in their potential side effects (16-18).

6.2.3 Acquisition setup

Usually, SPECT acquisitions are performed over a 180° orbit between right anterior oblique and left posterior oblique positions (Fig 6.1). It has been argued that 180° acquisition improves image contrast and resolution by avoiding high scatter and attenuation at posterior positions (19). The differences between 180° and 360° orbits are, however, debatable (20). Standard imaging parameters range from 32 to 64 views, lasting 20 to 30 seconds per view (21). Improved detector designs, such as orthogonal

dual heads, or triple head systems are increasingly utilized in cardiac SPECT to enhance the image quality and shorten the time of the procedure (22).

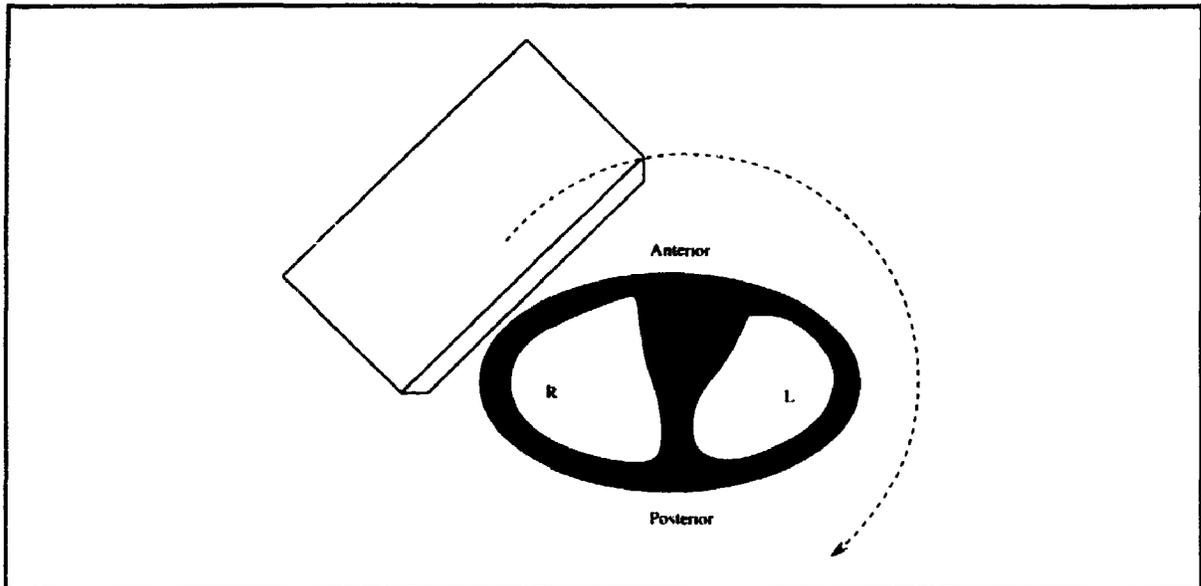


Figure 6.1: 180° acquisition of cardiac SPECT images.

During a standard SPECT acquisition, the images are blurred due to movement of cardiac walls. Gated SPECT technique permits the acquisition of separate 3-D images corresponding to several portions of the cardiac cycle (23). If gated SPECT acquisition is performed, each view is divided into a number of images (usually 8 to 16) containing counts corresponding to particular segments of the average cardiac cycle. To obtain this separation, image acquisition is synchronized with the electrocardiogram (ECG) signal. Although each gated segment contains fewer counts than conventional SPECT images, the effect of heart motion is reduced.

6.2.4 Image reconstruction

Acquired projection images are usually reconstructed by a filtered backprojection technique into a series of transaxial planes. Depending on particular software capabilities, various types of filter and reconstruction options are possible. Individual optimisation of reconstruction parameters for stress and rest studies is recommended because of the

difference in count statistics (21). In general, no attenuation correction is done, unless attenuation maps are available (21). The resulting slices are then reoriented into short-axis, horizontal long-axis and vertical long-axis views (24), to provide a standardized format for the display and comparison of scans (Fig 6.2). Sometimes, 3-D visualization techniques may be utilized during data analysis (25,26).

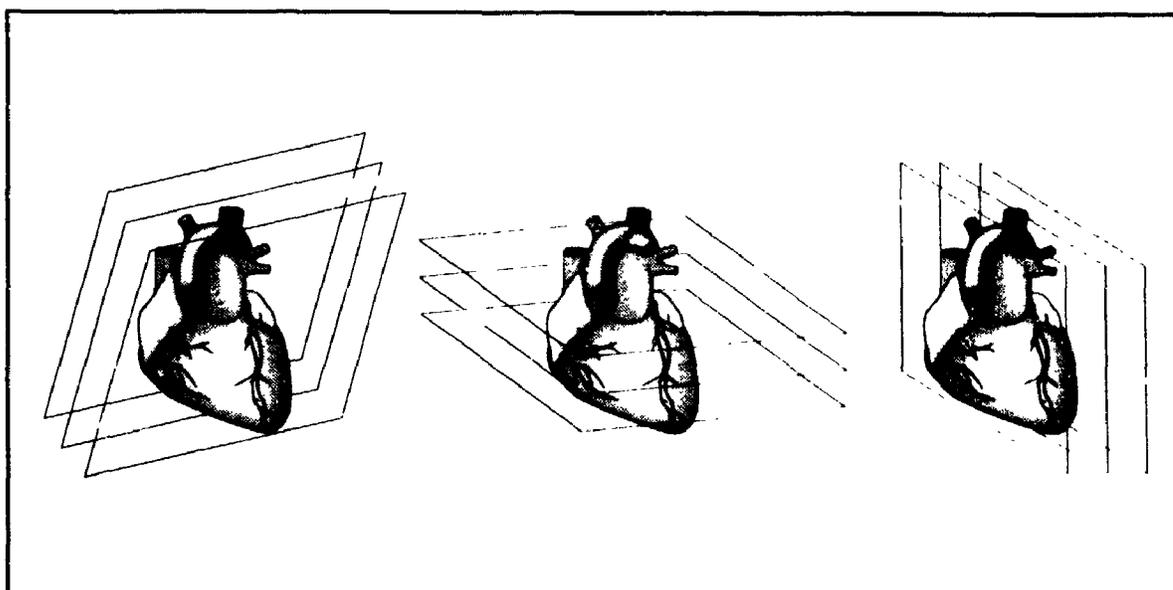


Figure 6.2: Standard orientation of planes in reconstructed cardiac SPECT datasets. Short-axis (left) horizontal long-axis (middle), and vertical long-axis (right). Figure drawn by Frank Panczyk, Visual Services, Victoria Hospital.

6.3 LIMITATIONS OF MYOCARDIAL SPECT

6.3.1 Compton scatter and attenuation artefacts

Due to inhomogeneous attenuation properties of the thorax, the effect of scatter and attenuation is difficult to compensate in myocardial SPECT. The presence of scatter reduces spatial resolution in the images resulting in image blurring. Attenuation of photons causes a relative decrease of the observed activity in regions located deeper in the body. The amount of attenuation depends also on the character of surrounding tissue. Scatter and attenuation contribution increases from the apex to the base of the heart. Thus, underperfused regions in basal areas can be more difficult to detect.

Moreover, depth-dependent spatial resolution of projection images may cause reconstruction artefacts, especially with 180° and non-circular orbits (27). High resolution collimators reduce this effect (21). Since photon absorption is greater for lower energies, the effect of scatter and attenuation is more severe with Tl-201 than with Tc-99m based agents. Attenuation artefacts in women are caused also by breast tissue, which reduces the activity in the anterior region.

Several approaches have been proposed to reduce the degrading effects of attenuation and scatter (28,29). Some of these methods involve simultaneous acquisition of the emission and transmission scans. The transmission scans can be obtained with external sources of radiation of different energies than the radionuclide in the patient. Attenuation maps of the thorax can be obtained by imaging the activity transmitted through the patient's body. The attenuation maps are then used to correct the emission images. Scatter contribution can also be estimated by acquiring images with a lower energy window (30). Specialized software attenuation correction methods have been developed by Tsui et al. (31); in their study, 3-D iterative reconstruction is shown to improve the resolution of myocardial images.

6.2.2 Motion artefacts

SPECT acquisition requires that a patient remain still during the scan; motion results in artefacts on the reconstructed images, which can lead to a false-positive diagnosis. Analysis of raw projection images is recommended to reveal such errors (32,33). Shorter acquisition times and specialized devices providing good arm support for patients reduce chances of patient motion during the procedure. It is possible to detect and compensate for the patient movement after the acquisition, by applying correction algorithms (34). Respiratory motion can be another reason for motion artefacts, especially after exercise. SPECT acquisition, which lasts up to 20 minutes is performed over many cardiac cycles. Heart motion is also a source of artefacts which can be reduced by utilizing the gated SPECT technique (35), (section 6.2.3).

6.3.3 Physiological limitations

Even if the imaging of the radiotracer distribution in the heart was perfect, a false-positive or false-negative diagnosis can be made due to various physiological factors. Patients may not be stressed adequately and therefore the effect on myocardial perfusion may not occur. Superimposed radioactivity in the abdomen may degrade the image quality and elevate counts in the inferior region (36). If such a spurious increase in uptake is observed on the rest study and compared to the normal stress study, it can imitate a relative decrease of uptake under stress. Alternatively, it can mask a defect on either a rest or stress study. Other reasons for abnormal studies are noncoronary diseases of the heart such as left bundle branch block, or myocardial hypertrophy (36). Also an unusual shape or contraction of heart can lead to the misinterpretation of defects on both rest and stress studies (37).

6.4 VISUALIZATION AND QUANTIFICATION METHODS

6.4.1 Bull's-eye quantification and visualization technique

In order to simplify and standardize the analysis of cardiac SPECT scans several quantification methods have been developed. Quantitative methods of image analysis developed for planar myocardial perfusion tomography (38) have been adapted by Vogel et al. for limited angle tomography studies (39). A popular scheme for data visualization and quantification is polar map representation (Fig 6.3). Quantitative techniques that use polar map representation have been developed (40, 41), and extensively validated for thallium-201 (42), and Tc-99m-sestamibi agents (43). In these methods a normal database of gender-matched patients was collected, and average polar map representations were created. Subsequently, polar maps of abnormal patients were compared to the normal distribution, and the differences visualized on the polar map display as color coded images (colors corresponding to the number of standard deviations below the normal) and black-out maps (40). Several types of comparisons to normal ranges were obtained (41). The bull's-eye quantification and visualization scheme was used in the detection and localization of myocardial defects (44,45). These quantitative

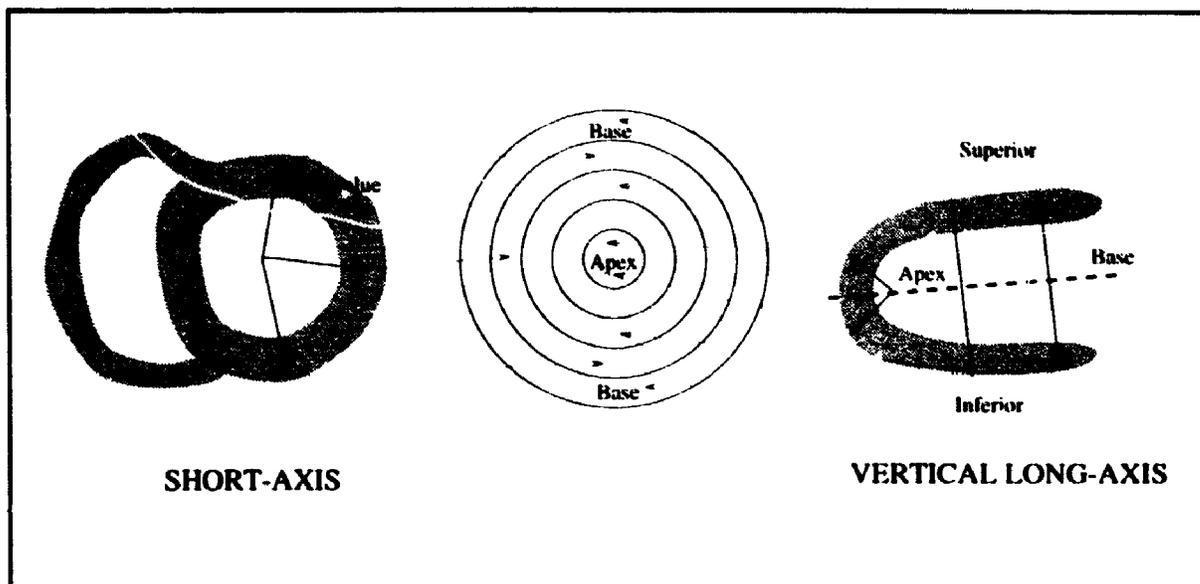


Figure 6.3: Polar map (or bull's eye) representation of myocardial SPECT. Maximum count values in the myocardial wall are found along vectors perpendicular to the long axis of the ventricle. These values are mapped on a two-dimensional circular map. The centre of the polar map corresponds to the apex and the circle positioned at the outside corresponds to the base of myocardium. Thus, each circle (circumferential profile) represents one short-axis slice. Polar map schemes have been adopted in many visualization and quantification approaches (40-45,49,50).

methods were shown to be highly reproducible when used in combination with the visual analysis of images (46).

6.4.2 Other quantification methods

Alternate approaches to quantification of myocardial perfusion tomography have been developed by other investigators. Narahara et al. (47) estimated quantitatively the mass of the ventricle by detecting edges on the short-axis slices. Maublant et al. (20) used a segment map consisting of 36 sectors to determine regional count distribution in myocardium.

Parallel efforts to quantify myocardial perfusion positron emission tomography (PET) images have been undertaken. PET cardiac images are similar to SPECT images in general appearance and noise properties. Thus, analogous quantitative methods can be used in these two modalities (48). Laubenbacher et al. (49) developed a quantitative technique that did not require definition of short-axis slices. Instead, an interactive

definition of the long-axis of the left ventricle was obtained. Subsequently, comparison to normal databases of radial maxima was performed. In another study Hicks et al. (50) divided the polar map into fixed sections and automatically derived a number of quantitative parameters from PET images.

6.4.3 Gated SPECT quantification

Gated SPECT offers the possibility of improving the accuracy of quantitative results (35) and also the ability to estimate the contractile function of the heart (51). Similar methods of quantification can be used in gated and non-gated SPECT to evaluate perfusion defects. Additionally, changes between end-systolic and end-diastolic tomographic images can be quantified and parameters such as ejection fraction can be estimated (52). Nevertheless, such methods are similar in their accuracy to techniques routinely used in planar imaging (53).

6.4.4 Automated methods of analysis

Most of the existing quantification approaches require several interactive steps before comparison to the normal databases can be accomplished. Examples of such operator-dependent steps are adjustment of size and position of the images, identification of correct oblique angles, and elimination of external activity. Incorrect alignment of the myocardium during quantification procedures may introduce discrepancies in the results (33). Although objective criteria for analysis of quantitative results have been developed (45), the final interpretation is usually combined with visual analysis (46), which is subjective in nature. These manual steps introduce inter-observer and intra-observer variability to the results, and prohibit fully automated analysis of the data (54).

In efforts to automate operator-dependent steps, Cauvin et al. (55) developed a geometric algorithm for reorienting the skeleton representation of the left ventricle. They demonstrated no significant difference between manual and automatic alignment. Mullick et al. (56) proposed a method that uses 3-D polygonal representation of the ventricular

surface to compute the orientation parameters. Others have developed semi-automated techniques that require some manual intervention (57-59). Another geometric approach, using principal-axis transformation, was proposed by Faber et al. (60). This method, however, does not perform well in abnormal studies (55). Ezekiel et al. (61) designed an algorithm consisting of a series of filtering and thresholding steps, which eliminates interfering background activity and locates basal and apical slices, centre of the myocardium, and the maximum radius of search.

The expert system approach to the final interpretation of quantitative data was developed by Herbst et al. (62). In their system, polar map data are used as input premises to a set of rules. The expert system can suggest which areas of the myocardium are hypoperfused. In another study, Fujita et al. (63) applied artificial neural network analysis to the bull's-eye quantification results. The values of pixels on the polar map were used as input weights to the network. The results suggested that the neural network system scored better than an average radiology resident but slightly worse than an experienced radiologist. Such techniques have the potential to reduce the variability and subjectivity of the myocardial perfusion test, providing a consistently high level of diagnosis.

Chapters 7 and 8 of this thesis propose three dimensional methods for fully automated image processing and quantification of myocardial scans by comparison to normal reference data. By using image registration techniques, all orientation and sizing parameters are automatically determined in 3-D. This is in contrast to automated methods previously reported (55-60), which estimate only certain subsets of these parameters. The quantification approach presented in chapter 8 is voxel-based and uses a fully 3-D, natural representation of the ventricle, as opposed to geometrical models previously used (20, 38-45,49,50). The results of these quantitative techniques can be naturally visualized and a small set of numeric parameters relating to the defect size, severity and location can be obtained. The use of such methods may ultimately simplify and objectify the interpretation of myocardial perfusion images

REFERENCES:

1. Gotto AM, Farmer JA. Risk factors for coronary artery disease. In: Braunwald E, ed. *Heart disease: A textbook of cardiovascular medicine*. 3rd edition. Philadelphia: W.B. Saunders Company; 1988:1153-1154.
2. Pohost GM, Henzlova MJ. The value of thallium imaging [Editorial]. *N Engl J Med* 1992; 323:190-192.
3. Brown KA. The role of stress redistribution thallium-201 myocardial perfusion imaging in evaluating coronary artery disease and perioperative risk. *J Nucl Med* 1994; 35:703-706.
4. Mettler FA, Guiberteau MJ. *Essentials of nuclear medicine imaging*. Third edition. Philadelphia: W.B. Saunders company; 1991:115.
5. Grossman W, Baim DS. *Cardiac catheterization, angiography and intervention techniques*. Philadelphia: Lea and Febiger; 1991.
6. Baim DS. Interventional catheterization techniques. In: Braunwald E, ed. *Heart disease: A textbook of cardiovascular medicine*. 3rd edition. Philadelphia: W.B. Saunders Company; 1988:1379-1394.
7. Abrams HL, Adams DF. The coronary arteriogram. Structural and functional aspects. *N Engl J Med* 1969; 281:1276-1285.
8. Kennedy JW. Complications associated with cardiac catheterization and angiography. *Cathet Cardiovasc Diagn* 1982; 8:5-11.
9. Iskandrian AS, Heo J, Askenase A, et al. Thallium imaging with single photon emission computed tomography. *Am Heart J* 1987; 114:852-865.
10. Berman DS, Kiat H, Maddahi J. The new 99-m-Tc myocardial perfusion agents: 99-m-Tc-sestamibi and 99-m-Tc-teboroxime. *Circulation* 1991; 84:17-21.
11. Berman DS, Kiat HS, Van Train KF, Germano G, Maddahi J, Friedman JD. Myocardial perfusion imaging with technetium-99m-sestamibi: comparative analysis of available protocols. *J Nucl Med* 1994; 35:681-688.
12. Johnson LL. Myocardial perfusion imaging with technetium-99-m-teboroxime. *J Nucl Med* 1994; 35:689-692.

13. Lowe VJ, Greer KL, Hanson MW, Jaszczak RJ, Coleman RE. Cardiac phantom evaluation of simultaneously acquired dual isotope rest thallium-201/stress technetium-99m dual isotope cardiac SPECT. *J Nucl Med* 1993; 34:2006-2008.
14. Kiat H, German G, Friedman J, Van Train K, Silagan G, et al. Comparative feasibility of separate or simultaneous rest thallium-201/stress technetium-99m-sestamibi dual isotope myocardial perfusion SPECT. *J Nucl Med* 1994; 35:542-548.
15. Heo .I, Wolmer I, Kegel J, Iskandrian AS. Sequential dual-isotope SPECT imaging with thallium-201 and technetium-99m-sestamibi. *J Nucl Med* 1994; 35:549-553.
16. Leppo JA. Dipyridamole thallium imaging: the lazy man's stress test. *J Nucl Med* 1989; 30:281-287.
17. Iskandrian A. Adenosine myocardial perfusion imaging. *J Nucl Med* 1994; 35:734-736.
18. Pennel DJ, Underwood SR, Swanton RH, et al. Dobutamine thallium myocardial perfusion tomography. *J Am Coll Cardiol* 1991; 18:1471-1479.
19. Tamaki N, Mukai T, Ishii Y, et al. Comparative study of thallium emission myocardial tomography with 180° and 360° data collection. *J Nucl Med* 1982; 23:661-666.
20. Maublant JC, Peycelon P, Kwiatkowski F, et al. Comparison between 180° and 360°, data collection in technetium-99m MIBI SPECT of the myocardium. *J Nucl Med* 1989; 30:295-300.
21. Berman DS, Kiat H, Van Train K, Garcia E, Friedman J, Maddahi J. Technetium-99m-sestamibi imaging in the assessment of chronic coronary artery disease. *Semin Nucl Med* 1991; 84:190-212.
22. Galt JR. New instrumentation for cardiovascular nuclear medicine. *J Nucl Med* 1994; 35:20N-22N.
23. Barat JL, Brendel AJ, Colle JP, et al. Quantitative analysis of left ventricular function using gated single photon emission tomography. *J Nucl Med* 1984; 25:1167-1174.
24. ACC/AHA/SNM policy statement: standardization of cardiac tomographic imaging. *J Nucl Med* 1992; 33:1434-1435.
25. Cooke CD, Garcia EV, Folks RD, Peifer JW. Three-dimensional visualization of cardiac single photon emission computed tomography studies. In: Robb RA ed. *Visualization in biomedical computing, proceedings of the 13-16 October 1992 conference SPIE* 1992; 1808:671-675.

26. Loboguerro A, Perault C, Gibold J, et al. Shape preserving three-dimensional display of myocardial scintigraphic data. *Nucl Med Comm* 1994; 15:417-421.
27. Maniawski PJ, Morgan HT, Wackers FJ. Orbit-related variation in spatial resolution as a source of artifactual defects in thallium-201 SPECT. *J Nucl Med* 1991; 30:871-875.
28. Bailey DL, Hutton BF, Walker PJ. Improved SPECT using simultaneous emission and transmission tomography. *J Nucl Med* 1987; 28:844-851.
29. Frey EC, Tsui BMW, Perry JR. Simultaneous acquisition of emission and transmission data for improved thallium-201 cardiac SPECT imaging using a technetium-99m transmission source. *J Nucl Med* 1992; 33:2238-2245.
30. Galt JR, Cullom SJ, Garcia EV. SPECT: a simplified method to compensate for scatter, attenuation and depth-dependent resolution in 180° myocardial SPECT. *J Nucl Med* 1992; 33:2232-2237.
31. Tsui BMW, Gullberg GT, Edgerton ER, Ballard JG, Perry JR. Correction of nonuniform attenuation in cardiac SPECT imaging. *J Nucl Med* 1989; 30:497-507.
32. Freidman J, Berman D, Van Train K, et al. Patient motion in thallium-201 myocardial SPECT imaging: an easily identified frequent source of artefactual defect. *Clin Nucl Med* 1988;13:321-324.
33. De Puey EG, Garcia EV. Optimal specificity of thallium-201 SPECT through recognition of imaging artefacts. *J Nucl Med* 1989; 30:441-449.
34. Cooper JA, Neumann PH, McCandless BK. Detection of patient motion during tomographic myocardial perfusion imaging. *J Nucl Med* 1993;1341-1348.
35. Mannting F, Mannting MGM. Gated SPECT with technetium-99-m-sestamibi for assessment of myocardial perfusion abnormalities. *J Nucl Med* 1993; 34:601-608.
36. Mettler FA, Guiberteau MJ. *Essentials of nuclear medicine imaging*. Third edition. Philadelphia: W.B. Saunders company; 1991:119-121.
37. Eisner RL, Schmarkey LS, Martin SE, Carey D, Worthy MA, et al. Defects on SPECT perfusion images can occur due to abnormal segmental contraction. *J Nucl Med* 1994; 35:638-643.
38. Burow RD, Pond M, Schafer WA, Becker L. Circumferential profiles: a new method for computer analysis of thallium-201 myocardial perfusion scans *J Nucl Med* 1979; 43:771-777.

39. Vogel RA, Kirch DL, LeFree MT, et al. Thallium-201 myocardial perfusion scintigraphy. Results of standard and multipinhole tomographic techniques. *Am J Cardiol* 1979; 43:787-793
40. Garcia E, Van Train K, Maddahi J, et al. Quantification of rotational thallium-201 myocardial tomography. *J Nucl Med* 1985; 26:17-26.
41. Garcia E, Cooke CD, Van Train KF, et al. Technical aspects of myocardial SPECT imaging with technetium-99m-sestamibi. *Am J Cardiol* 1990; 66:23E-31E.
42. Van Train KF, Berman DS, Garcia E, et al. Quantitative analysis of stress Tl-201 myocardial scintigrams: a multicenter trial validation utilizing standard normal limits. *J Nucl Med* 1986; 27:17-25.
43. Van Train KF, Areeda J, Garcia E, et al. Quantitative same-day rest-stress technetium-99m-sestamibi SPECT: definition and validation of stress normal limits and criteria for abnormality. *J Nucl Med* 1993;34:1494-1502.
44. DePasquale EE, Nody AC, De Puey E, et al. Quantitative rotational Tl-201 tomography for identifying and localizing coronary artery disease. *Circulation* 1988; 77:316-327.
45. Maddahi J, Van Train KF, Prigent F, et al. Quantitative single photon emission computerized thallium-201 tomography for the evaluation of the coronary artery disease: optimization and prospective validation of a new technique. *J Am Coll Cardiol* 1989;14:1689-1699.
46. Alazraki NP, Krawczynska EG, DePuey EG, et al. Reproducibility of thallium-201 exercise SPECT studies. *J Nucl Med* 1994; 35:1237-1244.
47. Narahara KA, Thompson CJ, Maublant JC, Criley JM, Mena I. Estimation of left ventricular mass in normal and infarcted canine hearts using thallium-201 SPECT. *J Nucl Med* 1987; 28:1315-1321.
48. Garcia EV, Eisner RL, Patterson RE. What should we expect from cardiac PET? [Editorial]. *J Nucl Med* 1993; 34:978-980.
49. Laubenbacher C, Rothley J, Beanlands R, Sawada S, Sutor R, et al. An automated analysis program for the evaluation of cardiac PET studies: initial results in the detection and localization of coronary artery disease using nitrogen-13-ammonia. *J Nucl Med* 1993; 34:968-978.

50. Hicks K, Ganti G, Mullani N, Gould K. Automated quantitation of three-dimensional cardiac positron emission tomography for routine clinical use. *J Nucl Med* 1989; 30:1787-1797.
51. Faber TL, Akers MS, Peshock RM, Corbett JR. Three-dimensional motion and perfusion quantification in gated single-photon emission computed tomograms. *J Nucl Med* 1991; 32:2311-2317.
52. DePuey EG, Nichols K, Cobrinsky C. Left ventricular ejection fraction assessed from gated technetium-99m-sestamibi SPECT. *J Nucl Med* 1993; 34:1871-1876.
53. Hurwitz GA, Driedger AA, Dain RP, et al. Estimation of left ventricular contractile function on ECG-gated thallium-201 myocardial perfusion images. *Can J Cardiol* 1989; 5:343-351.
54. Garcia EV. Quantitative myocardial perfusion single photon emission computed tomographic imaging: Quo vadis? (Where do we go from here) [Review]. *J Nucl Cardiol* 1994; 1:83-93.
55. Cauvin JC, Boire JY, Maublant JC, et al. Automated detection of the left ventricular myocardium long axis and center in thallium-201 single photon emission computed tomography. *Eur J Nucl Med* 1992; 19:1032-1037.
56. Mullick R, Ezquerra NF, Garcia EV, Cooke CD, Folks RD. 3D visualization of pose determination: application to SPECT imaging. *Visualization Biomed Computing Proc SPIE* 1992; 1808:225-234.
57. Boire JY, Cauvin JC, Maublant J, Veyre A. Automated alignment of thallium-201 myocardial tomographic views. In: Kim Y, Spelman FA, eds. *Proceedings of 11th annual international conference of the IEEE engineering in medicine and biology society*. Seattle, WA; 1989:578-579.
58. Cooke CD, Folks RD, Jones ME, Ezquerra NF, Garcia EV. Automatic program for determining the long axis of the left ventricular myocardium used for thallium-201 tomographic reconstruction [Abstract]. *J Nucl Med* 1989; 30:806.
59. He ZX, Maublant JC, Cauvin JC, Veyre A. Re-orientation of the left ventricular long axis on myocardial transaxial tomograms by a linear fitting method. *J Nucl Med* 1991; 32:1794-1800.
60. Faber TL, Stokely EM. Orientation of 3D structures in medical images. *IEEE Trans Pattern Anal Mach Intell* 1989; 30:626-633.

61. Ezekiel A, Van Train KF, Berman D, Silagan D, Maddahi J, Garcia EV. Automatic determination of quantification parameters from Tc-sestamibi myocardial tomograms. In: *Computers in Cardiology*. New York: IEEE, 1991:237-240.
62. Herbst MD, Garcia EV, Cooke CD, Ezquerra NF, Folks RD, De Puey EG. Myocardial ischemia detection by expert system interpretation of thallium-201 tomograms. In: Reiber JHC, van der Wall EE, eds. *Cardiovascular nuclear medicine and MRI*. Dordrecht, The Netherlands: Kluger Academic Publishers; 1992:77-78.
63. Fujita H, Katafuchi T, Uehara T, Nishimura T. Application of artificial neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull's-eye images. *J Nucl Med* 1992; 33:372-276.

CHAPTER 7: AUTOMATED ALIGNMENT AND SIZING OF MYOCARDIAL STRESS AND REST SCANS TO THREE-DIMENSIONAL NORMAL TEMPLATES USING AN IMAGE REGISTRATION ALGORITHM: A METHOD FOR REPRODUCIBLE QUANTIFICATION^d

7.1 INTRODUCTION

Myocardial perfusion imaging using single photon emission tomography (SPECT) is an important diagnostic procedure for coronary artery disease (1). The traditional method of data analysis is tomographic data reconstruction, followed by visual assessment of perfusion defects on both stress and rest images. One of the most challenging aspects of myocardial SPECT image processing is reliable and accurate quantification (2). Quantitative analysis is used to reduce the variability and subjectivity of the test. Several researchers have been developing methods to improve the accuracy and reproducibility of quantification (3,4). Most existing quantification schemes require user interaction, which can introduce errors in the final results (1). Operator-dependent steps usually include identifying the position of apical and basal slices, delineating the edge of the left ventricle, and adjusting the orientation and translation parameters. Studies on reducing or eliminating certain user-dependent steps have been previously reported (5-8). Full automation of these procedures should result in overall improved reproducibility of the test interpretation.

^dThe content of this chapter has been submitted for publication (Slomka PJ, Hurwitz G, Stephenson JA, Craddock TD. Automated alignment and sizing of myocardial stress and rest scans to three-dimensional normal templates using an image registration algorithm: a method for reproducible quantification *J Nucl Med.* June 1994). Gil Hurwitz selected patient data. Janice Stephenson verified all automatically derived results. I wrote the software, designed, and conducted all experiments.

In this chapter, a completely automated technique for the quantification of tomographic myocardial images is described. The technique is based on the concept of three-dimensional (3-D) image registration. The use of 3-D shape registration has been studied by computer scientists for several applications in computer vision (9-11). In medical imaging, the applications of such algorithms to Magnetic Resonance (MR), Computed Tomography (CT), Nuclear Medicine (NM), and Positron Emission Tomography (PET) images of the brain have been explored extensively (12-14). In most cases, however, these techniques have been applied to register serial studies from the same patient.

In contrast, we propose applying these techniques to register images of different patients, to a common, normalized orientation and size. Compared to the registration of serial data from one patient, the main difference is the need to independently adjust three scaling parameters, which are not known *a priori*. We tested and applied several image registration algorithms to correlate the images of several patients to common 3-D templates. The registration procedure adjusted nine linear image transformation parameters including three angles, three shifts, and three scaling factors. Based on these algorithms, we devised a technique for automatic quantification. To provide a normal reference, 3-D templates were built cumulatively from normal studies registered to the same orientation and size. Subsequently, these templates were divided into predefined regions. A relative quantification of counts in these regions could be performed on the test-images because they were registered to the same orientation as the templates. Normal templates could be also used for a comparative display of patient images, providing a natural 3-D representation of normal count distribution.

A particular problem with myocardial SPECT is the definition of the ventricular surface, which is significantly blurred due to the resolution of the camera and cardiac motion in non-gated studies (section 6.3). Moreover, in the case of abnormal studies, the apparent edges of tracer distribution can be quite different from the actual shape of the ventricle. Therefore, volume-based image registration techniques that do not require definition of

surfaces are more suitable for this application. Such methods are usually computationally more intensive but they avoid potential surface definition problems by utilizing full image data to perform the registration.

7.2 METHODS

7.2.1 Patient selection criteria

Patients were selected from a referral base of approximately 3,000 cases according to correlating angiographic and clinical information. All angiograms as well as scintigraphic studies were ordered on clinical indications. Paired scintigraphic and angiographic studies were considered for inclusion if both procedures were performed within 4 months of each other, and there was no history of Q-wave infarction. Angiograms were interpreted by a single cardiovascular radiologist unaware of the scintigraphic results. Angiographic studies were considered to be normal if there were no coronary stenoses of 50% or larger in diameter. Twenty-three male and fifteen female Tc-99m-sestamibi stress-rest studies were used for the normal model. In the abnormal group, 10 male patients with single-vessel disease who had proximal stenoses of 60% or more in the proximal left anterior descending coronary artery (LAD) were selected.

7.2.2 Data acquisition and reconstruction

Images were acquired on a *Siemens* (Hoffman Estates, IL) ZLC 7500 NCO single head camera with a low-energy all-purpose collimator. Data were collected using 180° rotation, 64x64 matrix, 32 stops, 30 seconds per stop, and circular orbit. The pixel size of the acquisition data was 6.3 mm. Centre of rotation and uniformity corrections were applied to the projection images. The data were reconstructed on the *Nuclear Diagnostics* (Stockholm, Sweden) HERMES workstation. Two-dimensional (2-D) Metz adaptive prefiltering of projection frames was applied. Filtered backprojection reconstruction was performed without applying attenuation correction. Tomograms were reconstructed with 200% zoom in the transaxial plane. Thus, the pixel size of the reconstructed images was 3.15 mm, and slice thickness was 6.3 mm.

In three cases, an extensive area of high activity in the abdomen was manually masked from the reconstructed slices by operator-drawn regions. In other studies, it was possible to automatically eliminate considerable amount of external activity with a simple masking scheme (Fig 7.1).

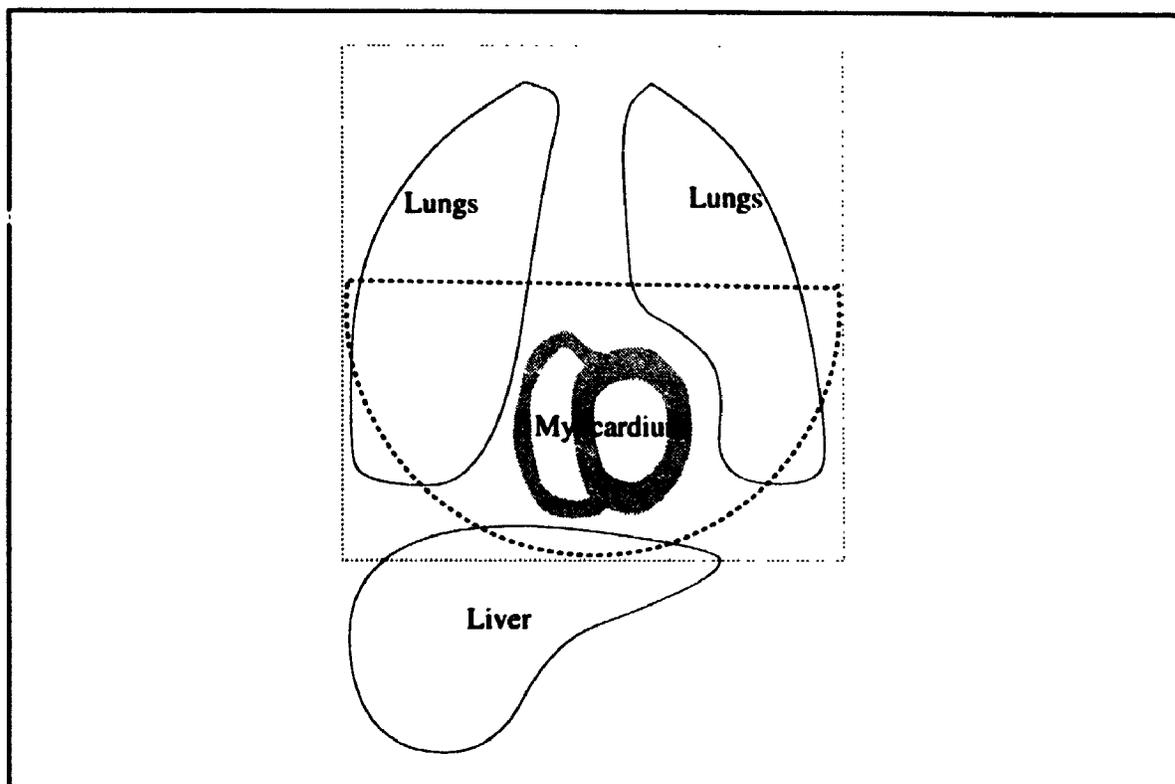


Figure 7.1: Masking of external activity. Voxels outside of semi-circular outlines (dashed line) with the diameter equal to the matrix size and positioned perpendicular to the transverse slices were set to zero. Dotted line shows the original matrix outline.

Although this operation did not usually remove all of the abdominal activity, the registration was considerably improved. The reconstruction and masking process were the only steps requiring user interaction.

7.2.3 Three-dimensional image registration

All registration algorithms were implemented in the C programming language using the Solaris operating environment on a *Sun Microsystems Inc.* (Mountain View, CA) SPARC 10/512 workstation.

Reconstructed transverse datasets were correlated to each other in 3-D by means of an automated image registration algorithm. The algorithm had a hybrid nature. In the first step, an approximate image alignment was accomplished using a technique based on the principal-axes transformation (15,16). It was discovered that the principal-axes technique produced large angle errors in abnormal studies, and could robustly estimate only scaling and translational parameters; therefore, the angle adjustment was disabled in this method. Despite low accuracy, this initial geometric registration was useful as a preprocessing step before a simplex minimization technique (17) was applied to refine the result.

The simplex algorithm finds the minimum of a multidimensional function; it requires only the evaluation of the function values and not the derivatives. The method starts with $N+1$ estimations of the function to be minimized (where N is the dimension of the function), and through a series of steps it converges toward the smallest value. The simplex-downhill minimization algorithm was applied to iteratively improve upon the initial principal-axes image fit by independently adjusting nine transformation parameters: X scale, Y scale, Z scale, X shift, Y shift, Z shift, XY tilt, XZ tilt, and YZ tilt (Fig. 7.2). The X, Y, and Z scaling parameters were included in the search for the best fit between the two datasets to compensate for size variations between the hearts of different subjects. The measure of the difference between 2 image volumes represented the function value. Thus, the problem was equivalent to finding the minimum of a nine-dimensional function. Each evaluation of function value involved a full reorientation of one dataset (32 slices 64x64) using 3-D linear interpolation (18). The reorientation was applied to the raw transverse data, thus avoiding the accumulation of linear interpolation errors. Other minimization algorithms (Powell and Conjugate Gradient) were also tested, but despite general recommendations of such methods (17), it was observed that the simplex method generally required less iterations to converge.

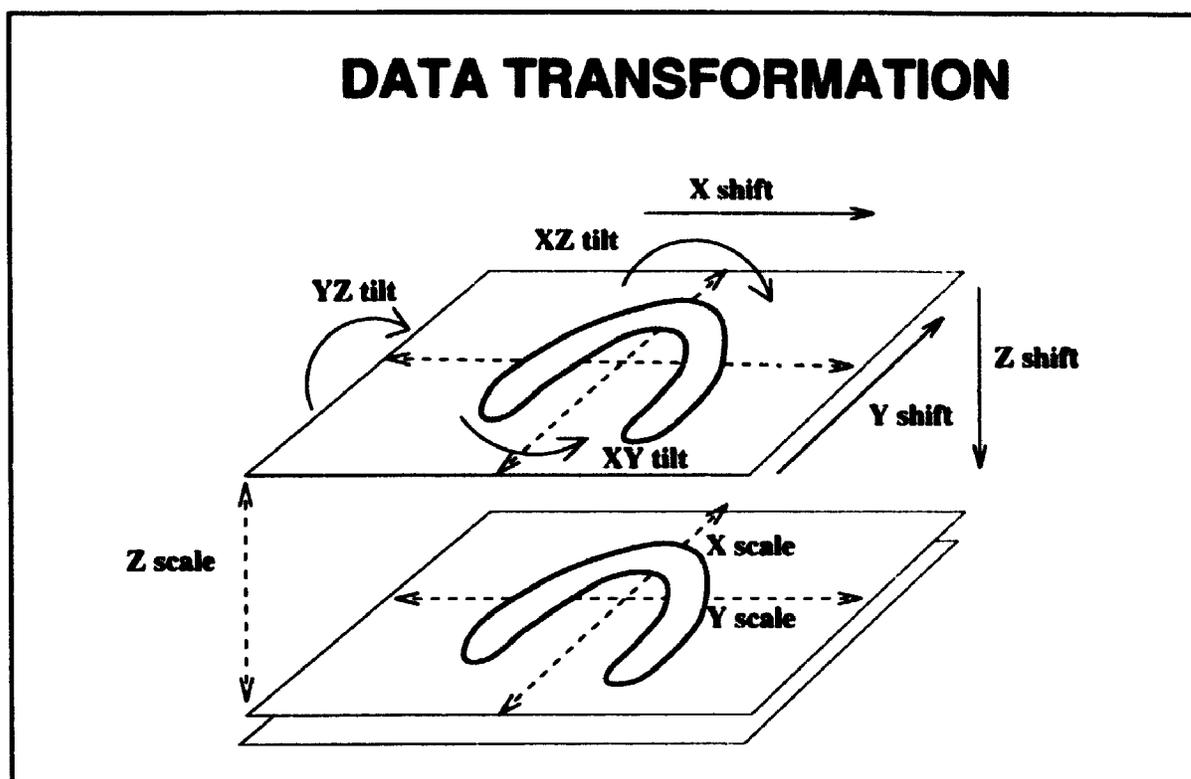


Figure 7.2: Nine linear transformation parameters used by the image registration technique to reorient and resize hearts.

Three different convergence criteria were tested for the minimization procedure: the count-difference (sum of absolute differences) as described by Hoh et al. (19), the Stochastic Sign Change (SSC), as described by Venot et al. (20), and the total variance of counts between two volumes (21). Based on the analysis of alignment errors in the simulated-defect test (section 7.2.8), the count-difference method was chosen for the amalgamation of templates and the registration of individual patient images to the template.

7.2.4 Building composite templates

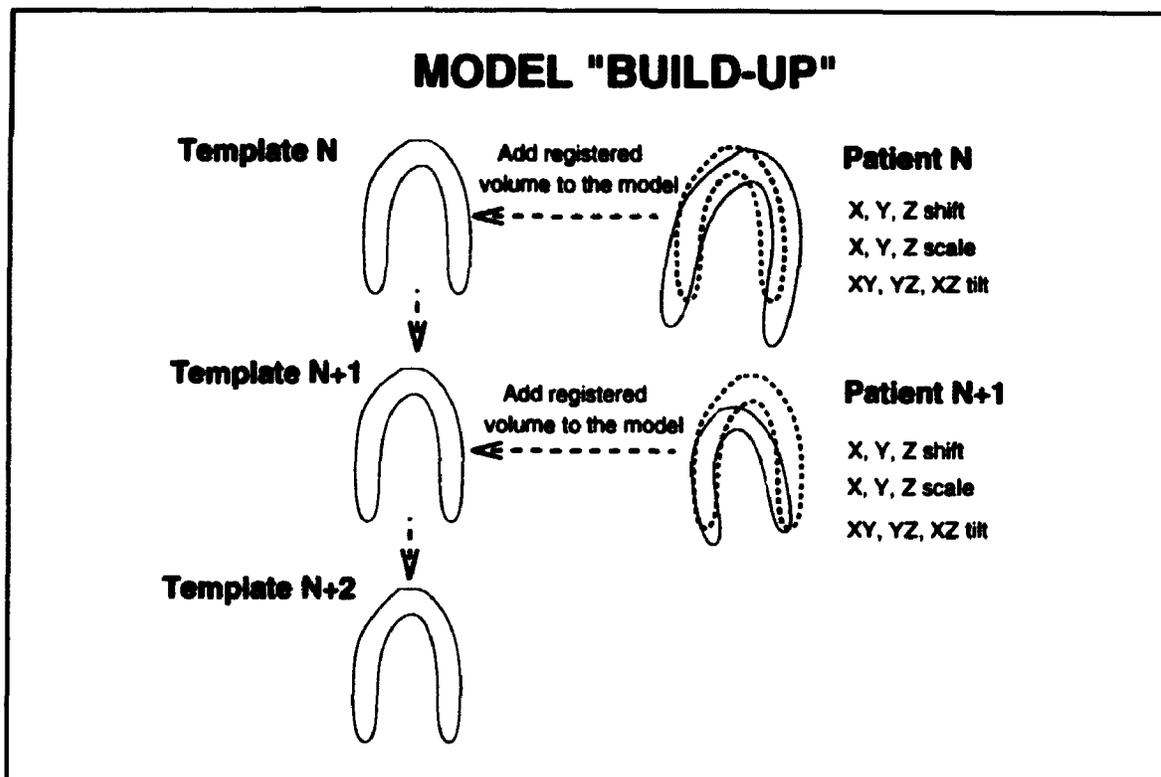


Figure 7.3: Accumulating patient data in the normal model.

Aligned images were sequentially added to each other, creating amalgamated datasets (Fig. 7.3). Stress and rest templates were built in a cumulative fashion by registering individual patient studies with the composite dataset and continuously improving it by adding new registered images. Initially, two patients were selected and their images automatically registered to the same orientation and size. After multiplying the voxel values by the ratio of maximal counts in the $3 \times 3 \times 2$ voxel region in the template and in the patient study, the data from these two registered volumes were added "voxel by voxel" (compare with the normalization scheme in section 8.2.3). Subsequently, other normal patient datasets were registered and added to the composite model. All datasets were registered to the templates using the same registration technique. At the end of this process, we obtained 3-D composite heart templates which could be viewed as tomographic scans (Figs. 7.4 and 7.5).

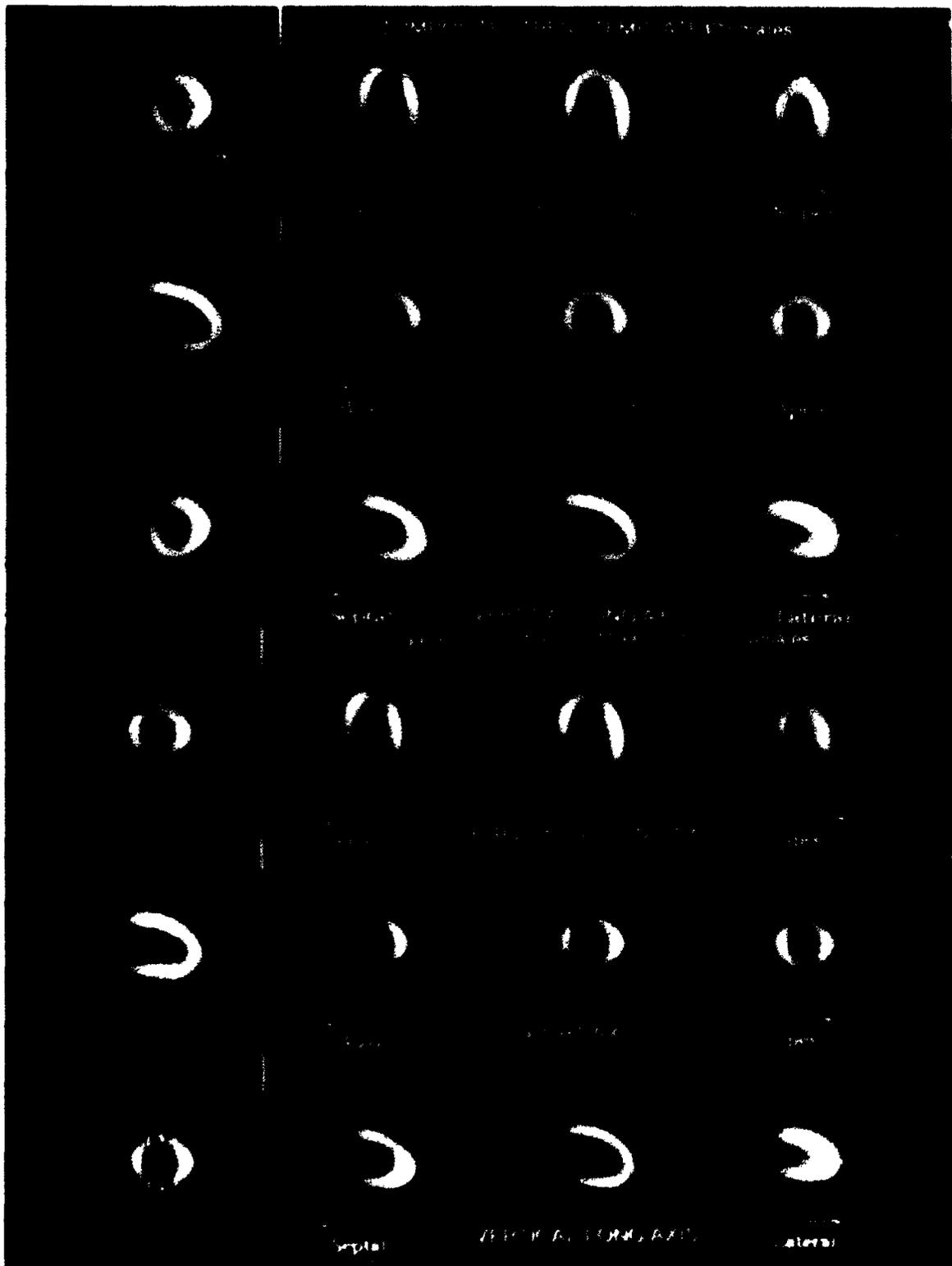


Figure 7.4: Male (top) and female (bottom) normal stress templates. The female template was resized to the male template dimensions.

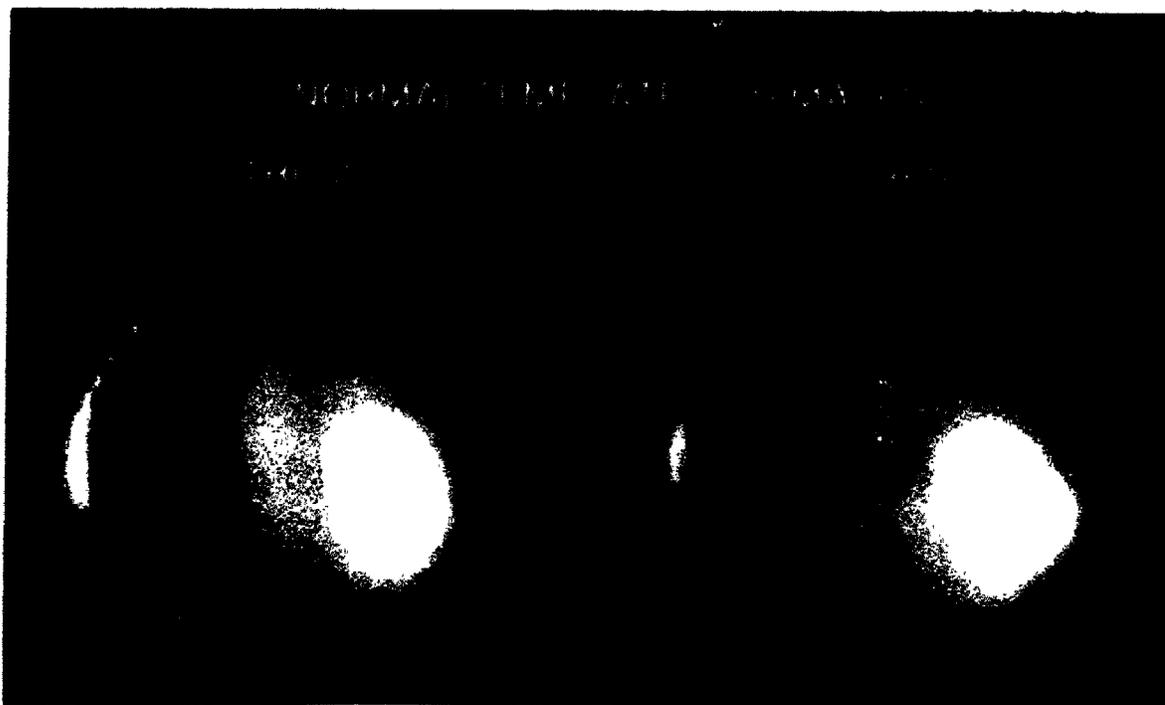


Figure 7.5: Three-dimensional surface shaded views of normal male stress (left) and rest templates (right).

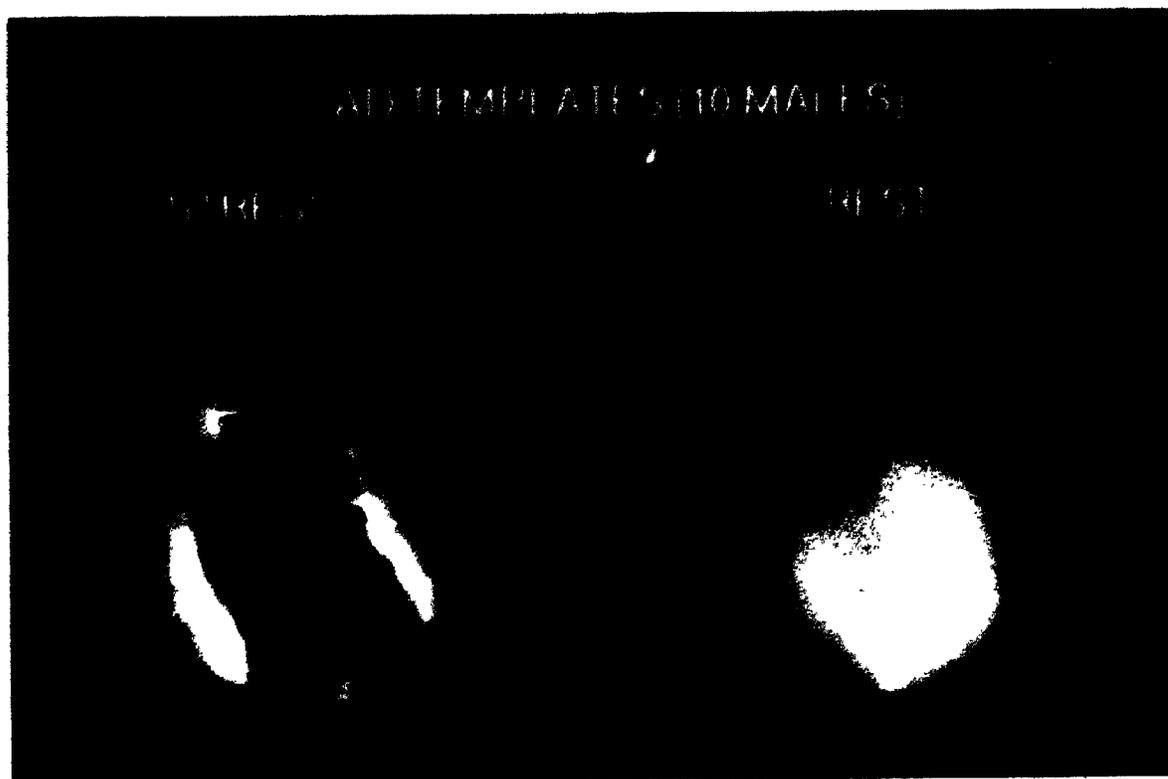


Figure 7.6: Three-dimensional surface shaded display of abnormal (LAD) male stress (left) and rest templates (right).

The composite template represents an average normalized 3-D distribution of counts in the myocardium. Therefore, it can be treated as a normal database that can be used for the comparison of patient images on a voxel level. Corresponding stress and rest templates were then "coregistered" to common spatial coordinates using the same image registration technique; as a result, stress and rest images of a given patient could be directly compared to each other after the registration to the templates. Separate templates for males and females were built since we expected differences due to attenuation artefacts. The attenuation artefact in the inferior wall was more pronounced on the male composite template (Fig. 7.4) Stress and rest templates of 10 male LAD disease patients were also built to visualize the characteristics of this defect (Fig. 7.6).

7.2.5 Segmentation scheme

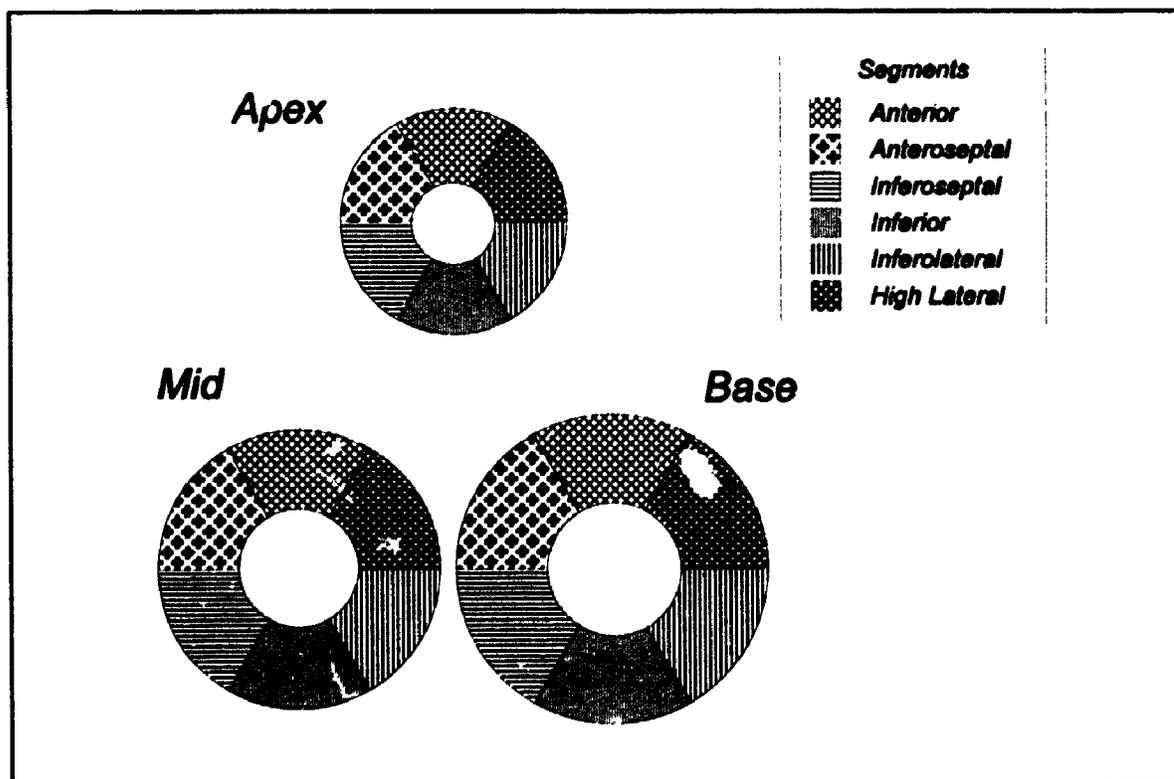


Figure 7.7: Scheme for template segmentation.

In order to estimate the regional distribution of counts, a 3-D segmentation of the templates was performed according to a popular scheme (22). The templates were

reformatted to the short-axis orientation and 16 short-axis slices were derived. Six anterior slices were classified as apical, four middle slices as medial, and six posterior slices as basal. The centre of the ventricle was identified on all 16 short-axis slices; six segments were derived in each of the three zones (Fig 7.7). Thus 18 segments were specified in total. This division provided the most equal distribution of segment volumes because the apical slices were much smaller than the mid ones, and the basal distribution was asymmetrical (Fig 7.4). The result was stored as a 3-D template with segment codes attached to individual voxels, allowing segment identification in the registered individual patients' images (Fig 7.8).



Figure 7.8: Mid-ventricular horizontal long-axis, short-axis, and vertical long-axis slices of the encoded segment template. Colors represent various segments.

7.2.6 Data quantification

To assess the variance of the count distribution in segments, all of the normal images were re-registered to the composite templates and the counts for each patient in the individual segments were calculated. Images of individual patients were normalized to the counts in the template in the same way as during the template accumulation (Fig. 7.3). Finally, the distribution of normalized count values, expressed as a percentage of counts in the templates was derived.

The same method was used to register abnormal patient scans to the stress-rest template pairs. Counts in each segment were calculated and normalized to the maximum in the

normal template (averaged in the 3x3x2 voxel region). In a preliminary assessment of automated detection of perfusion defects, we used this quantification technique in a group of 10 patients with LAD lesions. The results were then compared to the count distribution in normal templates. In addition, the difference between stress and rest counts in individual segments was assessed. The count-difference could be readily obtained by subtracting the stress and rest images since their respective templates were registered to each other. Finally the segments were classified as abnormal if the normalized counts were at least 2 standard deviations below the template counts in corresponding segments.

7.2.7 Visual assessment of registration accuracy



Figure 7.9: *Roving window* technique for visual assessment of image registration. The greyscale image represents the template; the color image represents the patient data. A patient image subwindow can be displayed and interactively moved on the template image (left). A template subwindow can be displayed and moved on the patient image (right).

The result of each automated alignment was carefully assessed visually for discrepancies by an experienced nuclear medicine technologist. Several visual presentation techniques were utilized to verify the registered image accuracy. The template was overlaid with the individual patient images, providing an interactive verification of the registration. Synchronized cursors were used to indicate corresponding

positions on the registered image and the normal template. A *roving window* display technique was also used to assess the quality of the fit (Fig. 7.9). The difference between the template and the registered and normalized patient image was also displayed to reveal the registration discrepancies.

7.2.8 Quantitative assessment of the accuracy and reproducibility of fitting

To assess quantitatively the robustness of the fitting algorithm, we analyzed the effect of simulated defects and the effect of the initial alignment on registration accuracy. The statistical significance of the results was determined using analysis of variance of repeated measures and follow-up comparisons of specific means by Newman-Keuls test (23).

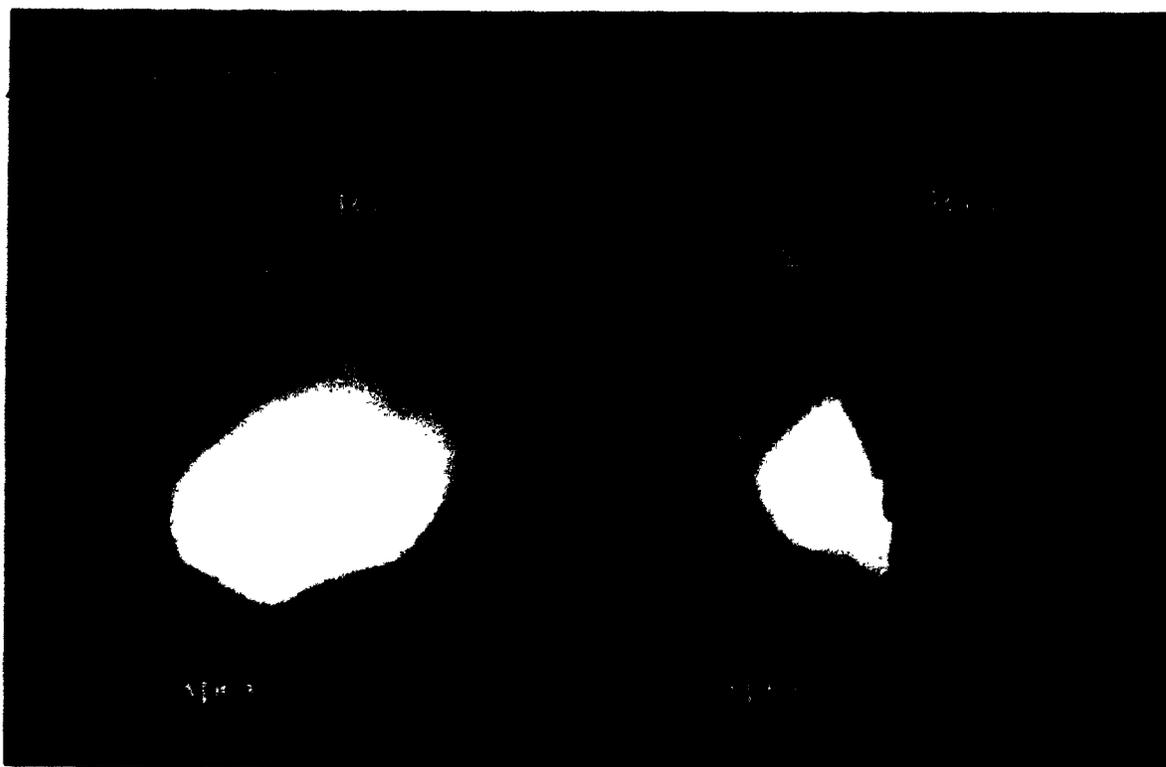


Figure 7.10: Example of a simulated defect. Three-dimensional surface-rendered image of a normal myocardium before (left) and after the introduction of the simulated defect (right).

To estimate the accuracy of the algorithm for abnormal data, 16 simulated defects were generated on one normal stress-rest dataset pair (Fig 7.10). Simulated defects were introduced at various locations in the left ventricle on stress and rest images; their size varied between 20 and 40% of the myocardial volume. Simulated defects did not contain any counts. Images with simulated defects were then arbitrarily misaligned (angles and shifts) and subsequently registered to the normal templates. Registrations of images with simulated defects were performed separately for three different convergence criteria. Registration parameters from the original study without defects provided an absolute reference to which the registrations of images with simulated defects were compared. The registration error for each transformation parameter was defined as the absolute change related to the introduction of the defect.

The reproducibility of the spatial registration of the count-difference method was assessed in a quantitative fashion. Patient studies were misaligned using 20 different combinations of position and tilt parameters and subsequently registered to the normal templates; the angles were changed between -20° and 20° and shifts between -8 and 8 pixels. This task was performed on 10 normal and 10 abnormal stress-rest datasets, resulting in 800 registrations. The reproducibility error for each transformation parameter was defined as the maximum difference from the mean value for 20 misalignments and it was estimated individually for each patient study.

The effect of the worst misalignment errors on the count distribution in 18 segments was also assessed. The two patient studies (one normal, one abnormal) with the highest errors in the registration reproducibility test were selected for this test. Count values were derived from each segment of both stress and rest datasets after the images were registered with 20 different initial orientations. Subsequently, coefficients of variation of the alignment results were evaluated.

7.3 RESULTS

7.3.1 Calculation time

The mean number of iterations required by the simplex-downhill algorithm for each patient fit was 275; each iteration took about 1 second. Thus, the total time for finding the minimum was approximately 4 minutes for each patient. The time involved in the initial principal-axes reorientation was roughly equivalent to one iteration of the simplex algorithm. Testing 800 misaligned cases during the fitting reproducibility test took several days.

7.3.2 Scaling parameters

Scaling parameters calculated by the registration program varied up to 28% within a group of normal patients. Ranges of scaling parameters, for male and female stress and rest groups are given in Figure 7.11.

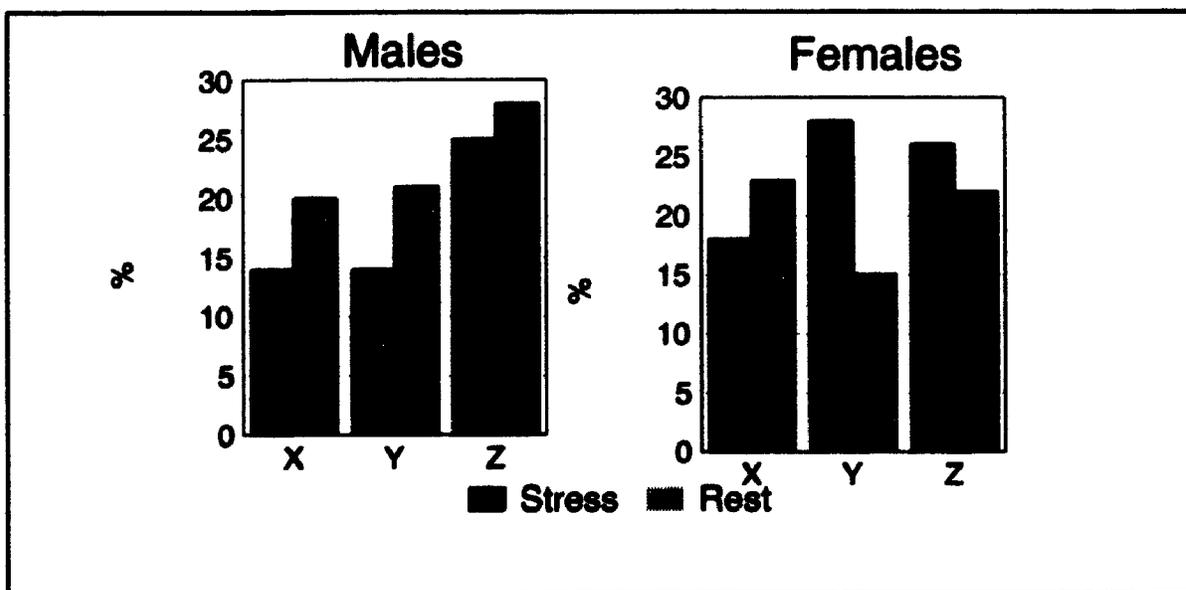


Figure 7.11: Range of scaling parameters in three different directions (X, Y, Z) in normal groups (23 males and 15 females). The values represent the relative difference between the maximum and the minimum scaling parameters

7.3.3 Visual assessment of the registration

A misregistration of size (subjectively assessed as $< 10\%$) was visually detected only in one case. It occurred due to a significant amount of liver activity remaining in the image despite the masking procedure. After removing this external activity, the algorithm properly registered the images. Discrepancies between the patient-data and corresponding templates were observed when the principal-axes registration was performed without further iterative simplex-minimization. The quality of this geometric registration was worse with abnormal scans or in the presence of remaining external activity. When the principal-axes method was combined with the simplex-minimization, errors could not be detected visually. Even in cases of abnormal hearts, the registration algorithm properly reoriented the images (Fig 7.12).



Figure 7.12: Mid-ventricular slices of patient data with severe LAD defect (color) overlaid on the model data (greyscale). Horizontal long-axis is shown on the left, short-axis slice in the middle and vertical long-axis on the right.

7.3.4 Assessment of registration errors in simulated-defect experiment

Registration errors for different convergence criteria due to simulated defects are shown in Table 7.1. Overall, fitting errors were significantly lower for the count-difference method than for the other two methods on stress images ($p \leq 0.0002$) and rest images ($p \leq 0.003$). Specifically, errors for several of the scaling and tilt parameters were significantly lower in comparison to the SSC and variance methods (Table 7.1). Based on these results, the count-difference was chosen for the subsequent test of the fitting

reproducibility and for the final registration of patient images. When comparing fitting parameters to each other for the count-difference method, it was found that Y scaling errors were significantly larger than errors of X and Z scaling ($p \leq 0.0005$) for stress images and significantly larger than Z scaling errors ($p = 0.008$) for rest images. For the count-difference method, the overall differences between stress and rest with respect to errors of 9 fitting parameters were not significant ($p = 0.5$).

TABLE 7.1: Simulated-defect experiment. Registration errors (mean \pm s.d.) for three different convergence criteria.

	Count difference		Stochastic sign change		Variance	
	Stress	Rest	Stress	Rest	Stress	Rest
X shift (mm)	1.2 \pm 1.3	1.2 \pm 1.0	3.1 \pm 3.6	2.0 \pm 2.0	3.2 \pm 2.5	3.6 \pm 2.7
Y shift (mm)	1.5 \pm 1.0	1.4 \pm 0.8	2.7 \pm 2.9	1.6 \pm 1.5	4.4 \pm 3.1	4.4 \pm 3.1
Z shift (mm)	0.8 \pm 1.0	0.9 \pm 1.4	1.4 \pm 1.0	0.9 \pm 1.2	1.7 \pm 2.5	1.9 \pm 2.9
X scale (%)	3.0 \pm 3.1	3.2 \pm 2.3	4.7 \pm 5.9	6.3 \pm 8.5	9.7 \pm 6.5*	11.1 \pm 9.2*
Y scale (%)	5.3 \pm 2.7	3.7 \pm 3.1	4.2 \pm 2.2	2.0 \pm 1.9	11.8 \pm 8.9*	10.5 \pm 6.3*
Z scale (%)	1.5 \pm 3.1	1.9 \pm 3.3	7.1 \pm 8.7*	4.8 \pm 7.0	5.8 \pm 9.4	4.9 \pm 9.4
XY tilt (deg)	1.3 \pm 1.2	1.4 \pm 0.9	5.9 \pm 5.0	2.6 \pm 1.9	5.2 \pm 4.9	3.1 \pm 3.8
XZ tilt (deg)	0.8 \pm 0.8	1.3 \pm 1.4	7.9 \pm 6.5*	4.3 \pm 5.3	2.4 \pm 2.8	3.5 \pm 3.3
YZ tilt (deg)	1.3 \pm 1.1	2.0 \pm 1.0	7.6 \pm 6.8*	6.2 \pm 9.2	1.8 \pm 1.1	2.9 \pm 2.3

s.d. = one standard deviation; n=16 simulated defects at various locations of the myocardium.

* significantly worse than the count difference method $p \leq 0.007$

7.3.5 Assessment of registration reproducibility error

Errors of the transformation parameters for normal and abnormal (LAD defect) patients are shown in Table 7.2. There was no significant difference between fitting of normal and abnormal studies ($p=0.3$), but fitting of stress studies was better than rest studies ($p \leq 0.0007$). Specifically, the fitting of abnormal stress studies was more reproducible than in the other 3 groups ($p \leq 0.002$). For normal stress and rest studies, the errors of the

scaling in Y direction were significantly worse than errors in X and Z direction ($p \leq 0.0002$) and tilt error in XZ direction was significantly worse than XY error ($p = 0.05$). For abnormal stress and rest studies the scaling error in Y direction was significantly worse than in X or Z direction ($p \leq 0.05$).

TABLE 7.2: Reproducibility of registration after arbitrary misalignments. Mean errors \pm s.d. of transformation parameters for the count-difference method.

	Normal studies		Abnormal studies	
	Stress	Rest	Stress	Rest
X shift (mm)	0.2 \pm 0.1	0.4 \pm 0.3	0.4 \pm 0.3	0.3 \pm 0.2
Y shift (mm)	0.7 \pm 0.5	0.8 \pm 0.4	0.7 \pm 0.6	0.6 \pm 0.4
Z shift (mm)	0.3 \pm 0.1	0.5 \pm 0.3	0.2 \pm 0.3	0.4 \pm 0.3
X scale (%)	1.2 \pm 0.4	1.5 \pm 0.6	0.9 \pm 1.0	1.8 \pm 0.9
Y scale (%)	3.8 \pm 1.9	3.6 \pm 2.7	2.0 \pm 2.0	3.0 \pm 1.7
Z scale (%)	2.0 \pm 1.3	2.2 \pm 0.7	1.2 \pm 1.5	2.3 \pm 1.3
XY tilt (deg)	1.1 \pm 0.5	1.1 \pm 0.7	0.7 \pm 0.8	1.1 \pm 0.7
XZ tilt (deg)	2.1 \pm 1.3	2.1 \pm 1.2	0.9 \pm 1.1	2.8 \pm 1.5
YZ tilt (deg)	1.4 \pm 0.8	2.1 \pm 1.6	0.9 \pm 1.0	2.0 \pm 1.0

s.d. = one standard deviation; n=10 datasets (each misaligned 20 times).

The hybrid nature of the algorithm was the reason for good convergence even when the images were grossly misaligned. The application of a fast, but less accurate principal-axes method to estimate the initial position and scaling decreased the number of iterations for the more precise simplex algorithm; it also improved the robustness of the iterative algorithm. We experienced difficulties in simplex algorithm convergence when the images were significantly misaligned and when the initial principal-axes technique was not used.

The results in Table 7.3 describe the influence of the registration error on the final quantification results in individual segments. One normal and one abnormal patient study are presented, for which the largest errors were observed in the registration reproducibility test (Table 7.2.). The variation of counts introduced by the registration errors are much smaller than the variations of counts in the normal studies (Fig 7.13).

TABLE 7.3: Effect of alignment errors on segmental count distribution in 2 patients (one normal, one abnormal) with highest registration errors in the reproducibility test (Table 7.2).

Worst segments	Coefficient of variation*	
	Stress	Rest
Normal patient		
Basal high lateral	3.9 %	3.0 %
Basal anteroseptal	3.0 %	2.3 %
Basal inferolateral	2.6 %	2.0 %
Basal anterior	2.3 %	1.8 %
Others	≤ 2.0 %	≤ 2%
Abnormal patient		
Apical inferoseptal	2.4 %	1.8 %
Apical anteroseptal	2.2 %	1.9 %
Apical inferior	2.2 %	2.8 %
Apical high lateral	2.1 %	1.5 %
Basal high lateral	1.1 %	2.6 %
Basal anterior	0.8 %	2.5 %
Others	≤ 2%	≤ 2%

*Variation of relative count values in segments; n=20 misalignments and registrations.

7.3.6 Analysis of individual patient images

In a preliminary assessment of the potential diagnostic use of our automated method, the count distribution was evaluated for individual segments in normal patients and in patients with angiographically proven LAD disease. The variation of counts in segments in the normal male stress and rest groups is shown in Figure 7.13. The classification results of abnormal patients are presented in Figure 7.14. All LAD patients had abnormal segments based on the 2 standard deviation criteria. In most cases, the abnormality was detected in the antero-septal regions of apical and mid-ventricular zones. Differences between stress and rest provided well defined results. Rest segments had generally higher standard deviations than corresponding stress segments. The stress-rest subtraction images had smaller variation in most segments than the stress and rest images assessed separately, which indicates correlation between stress and rest. The standard deviation was smallest in apical, medial, and basal high-lateral and inferior-lateral segments, and highest in inferior segments.

Difference images created by subtraction from the normal stress and rest templates can illustrate deviations from the normal distribution without using arbitrary segments (Fig. 7.15). Such images can be used to visualize the extent and location of perfusion defects. The differences between defects at stress and rest provide an estimate of ischemia. The stress template built from the group of LAD patients clearly showed the characteristics of this defect (Fig. 7.6). Similar templates can be built for other groups falling into specific diagnostic categories of angiographically defined disease (for example left-circumflex or right coronary artery) ; they could be used to estimate zones of expected hypoperfusion on the templates (chapter 8).

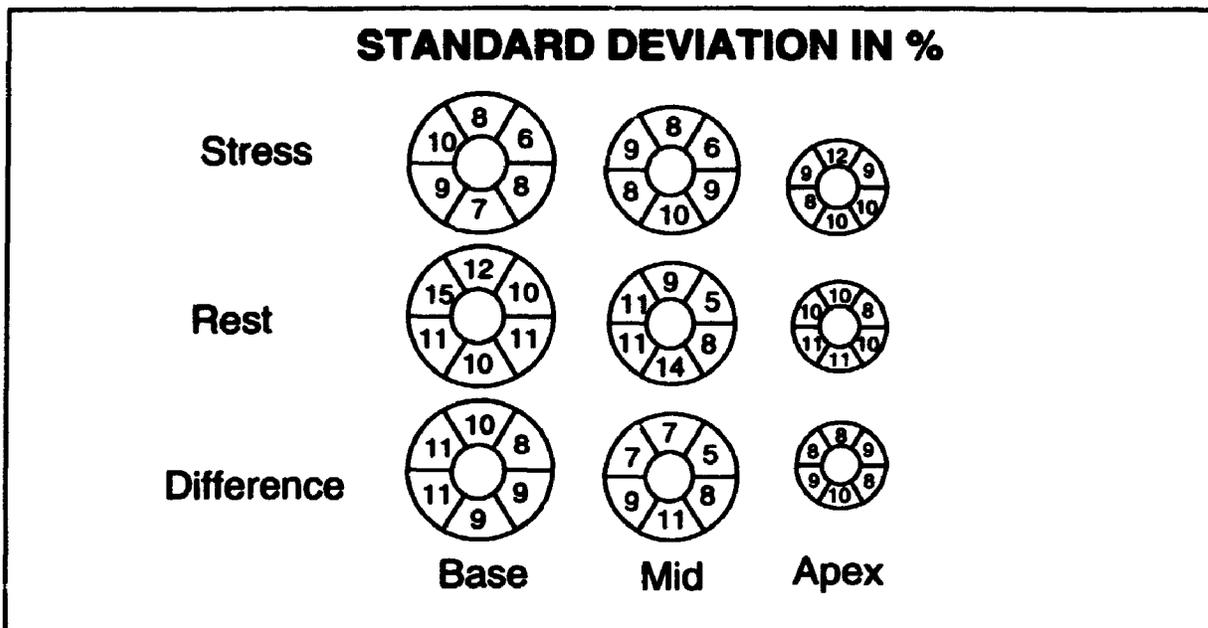


Figure 7.13: Standard deviations of relative count distribution in each segment for stress, rest and the difference between stress and rest in the normal male group (n=23).

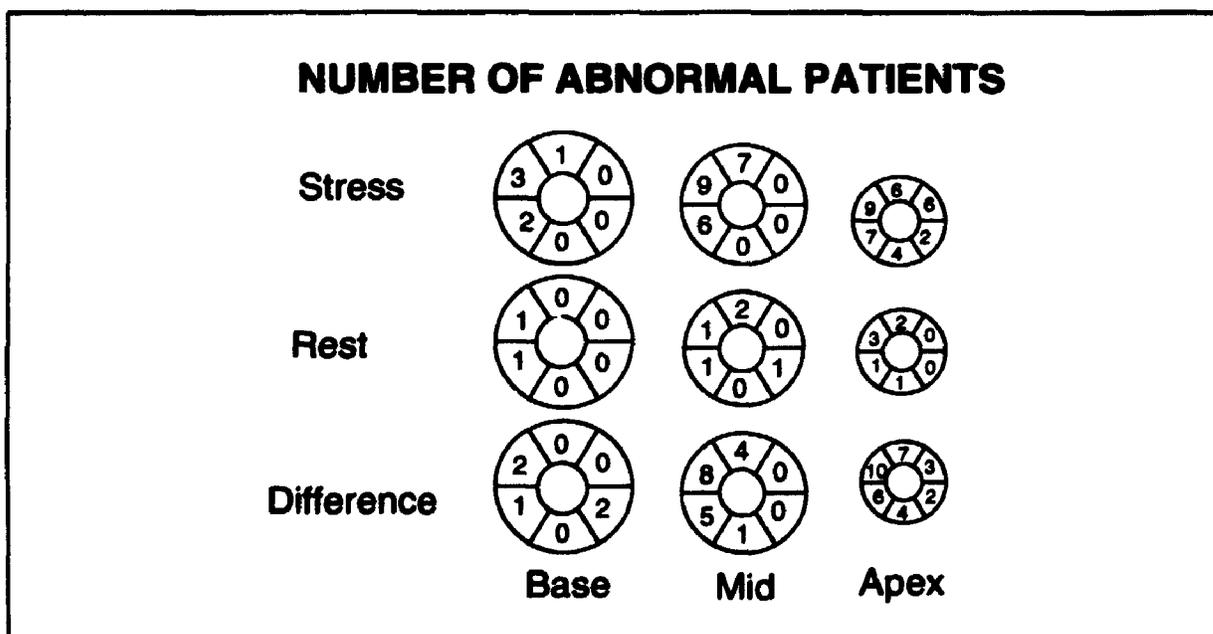


Figure 7.14: Abnormal segments in the LAD group of 10 male patients. The number of patients with abnormal results (below 2 standard deviations) are shown for each segment individually.



Figure 7.15: Subtraction of the abnormal (LAD) stress (top) and rest (bottom) patient images (left) from the normal model (middle). Subtracted images (right) visualize the extent of the defect.

7.4 DISCUSSION

7.4.1 Image registration method

The task of registering myocardial perfusion SPECT studies presents other problems than those experienced in the registration of high resolution images. The edges of the ventricle are not well defined and perfusion defects can significantly modify the shape of tracer distribution. Myocardium can be displaced with respect to the external anatomical landmarks. Cardiac motion significantly reduces the accuracy of the position in non-gated studies. Therefore contour-based methods such as Chen-Pellizari (12) are not suitable for this application. An important distinction from similar techniques applied to the registration of serial scans of the same patient is the incorporation of scaling parameters. When registering images of the same patient, scaling factors are usually known and are based on the pixel sizes of the two image sets. In this study scaling parameters play an

important role because the dimensions of the myocardium vary between patients. Combining three independent scaling parameters (X, Y, Z) into one or two factors would decrease the number of iterations. Nevertheless, because of the individual size variations in each dimension, it was important to include these factors as individual components in the search for the minimum. For example, dilated ventricles (2 cases in the LAD group) had a roughly spherical shape; the algorithm compensated for this effect by adjusting the X scaling parameter, changing the patient images to an approximately ellipsoidal shape of the normal template. Scaling parameters allow estimation of the ventricular volume, which has potential diagnostic value (24, 25).

Registering differently shaped hearts to common templates raises questions of the validity of the linear transformations. It should be possible, for example, to apply non-linear scaling, or "warping" techniques that adjust more than nine parameters to compensate for different shapes (26). Such a method was recently proposed by Houston et al. for the registration of brain studies (27). The disadvantage, however, would be the distortion of relative count distribution in the myocardium; therefore, it would be difficult to compare such techniques with established quantification schemes (3). The use of linear scaling preserves the relative count ratios and defect sizes; therefore, the quantification of data is more straightforward.

Despite amalgamating images of several patients using linear transformations, the composite images retain the essential characteristics of myocardial scans, and there is no significant loss of resolution (Figs. 7.4 and 7.5). The most noticeable characteristics of the composite templates were the low level of noise and background and the slightly smoothed and "blurred" appearance of the images. The smooth appearance could be attributed to a much higher number of counts in the templates than in single-patient images. Blurred edges were most likely caused by non-linear differences in ventricular shapes. Due to the low resolution of the original images, however, these differences do not seem to affect significantly the resolution of the composite images. The relatively good

quality of template images can be explained by the lack of fine details in the original images, which are blurred by the cardiac motion and are generally of low resolution. The attenuation artefact is more pronounced on the male composite template. The myocardial walls on the female template are thicker, probably due to relatively lower resolution of generally smaller hearts. To validate the criteria for normal patient selection we plan to compare the templates derived from studies with a low probability of coronary artery disease and data from multiple centres.

The count-difference convergence criterion provided better results than SSC or the variance method. Hoh et al. (19) did not find any significant difference between the SSC and count-difference method, but they used the algorithm to register serial studies of the same patient and did not search for scaling parameters.

Although the registration algorithm requires a relatively long computing time especially during iterative reorienting of the data, several techniques can be utilized to improve its speed for routine clinical use. Some of the methods for fast image reorientation are described by Hoh et al. (19). The relatively small matrix size and small number of slices in myocardial SPECT studies allowed the implementation of computationally intensive algorithms on a standard, mid-range computer. These algorithms would be currently impractical for larger matrix sizes due to lengthy computation times. The principal-axes technique is fast, due to its analytical character but the accuracy is unsatisfactory for this application. Nevertheless, its use for initial image alignment increases the robustness of the algorithm and reduces the number of subsequent iterations.

7.4.2 Alignment errors

The alignment errors of the automated registration (Table 7.1, 7.2) compared to the pixel size (6.3 mm) and the resolution of the camera (approximately 13 mm) is encouraging. The scaling error of 5% or the angular error of 2° translates approximately to the 3 mm displacement in the apex of an average heart. The smaller errors of scaling in the Y

direction (Table 7.1 and 7.2) and tilting in XZ direction (Table 7.2) may be attributed to the shape of the ventricle; data transformation in these directions corresponds to relatively smaller changes in the count-difference criteria. The registration errors in our automated technique are below the limit of visual assessment. Other methods have been previously designed to automatically adjust selected orientation parameters of the left ventricle (5-8). The advantage of our technique is that all parameters are considered including 3 independent scalings and 3 angles.

An objective measure of the performance of the algorithm with abnormal data was provided by the simulated-defect registration (Table 7.1). The results showed that the alignment was accurate even when a considerable part of the myocardium was missing. The method is robust despite the presence of defects because the algorithm searches for the relative minimum in the count-difference. Although the simulated-defects resulted in overall higher count-difference between patient image and the template, the relative minimum was found for the similar transformation parameters as in the original data registration. Registration errors in the simulated-defect test may have been amplified due to the fact that the introduced lesions did not contain any counts. Real perfusion defects usually contain some counts, which should aid the registration algorithm. In extreme cases, when most of the myocardium is missing, the registration may produce larger errors; in these situations, however, an accurate quantification is usually not critical. The clinical importance of accurate registration may be greatest for mildly-abnormal, borderline cases.

Both normal and abnormal images could be fitted with very good reproducibility. The fitting of abnormal stress studies was more reproducible than the fitting of normal studies probably because of a better-defined minimum in the count-difference between the template and abnormal studies (Table 7.2). Worse reproducibility of rest fitting (Table 7.2) can be attributed to lower counts in rest images. Compared to the traditional manual adjustment used in the bull's eye quantification program (3), the error of 1 slice in the

selection of basal or apical limits would be equivalent to 8% scaling error (assuming 12 short-axis slices). Small errors caused by the misalignments indicate that the technique is relatively insensitive to the initial position and orientation of the reconstructed myocardium. Therefore, the registration can be applied to raw transverse images without prior manual alignment, and true short-axis horizontal long-axis, or vertical long-axis slices can be automatically derived. The effect of the registration reproducibility errors (arbitrary misalignment test) on the quantification results is smaller than the variation in normal count distribution (Table 7.3 and Figure 7.11).

7.4.3 Image Quantification

The division of the templates into segments is one example of the quantification possibilities; the method described could be used with different region definitions. Other arrangements of segments for myocardial perfusion quantification were proposed by Maublant et al. (28) for SPECT and by Kotzerke et al. (29) for PET studies. It is conceivable to combine our method with the bull's-eye quantification scheme (3) by applying the registration algorithm as a preprocessing step. It should also be possible to create a polar map of the average 3-D template which would then serve as a reference database in the bull's eye quantification. Another application of the model could be further image processing on the images subtracted from the normal distribution. For example, it should be possible to estimate the abnormal region location and size by processing the 3-D images subtracted from the template (Fig. 7.15). The subtracted images could be normalized by the standard deviation values for normal studies calculated at each voxel. This approach would eliminate the need for the segmentation scheme. A method for the quantification and localization of defects based on this idea is presented in chapter 8.

The use of 3-D templates has potential advantages in comparison to the popular "bull's-eye" display (3). It can provide a more natural 3-D data visualization and representation, which is not constrained to predefined geometrical shapes. Furthermore, the voxel-based, 3-D character of our approach may facilitate the estimation of defect volume, by

comparing the templates and test-data on a voxel level. In comparison, on the polar map it is possible to estimate only the area, which subsequently can be related to volume (4). Accurate, automated alignment of the data with our registration method could provide more objective results than manual methods of reorienting and repositioning of the ventricle. During 3-D image reorientation the scaling, sizing and tilt parameters are dependent on each other and it is difficult to adjust one parameter accurately without influencing the optimal position of the others; the interactive methods usually rely on changing these parameters sequentially. In contrast, our method optimizes all position, size and orientation alignment parameters simultaneously.

Although the described method has been evaluated for myocardial stress and rest perfusion images, it has more general applications. In particular, the quantification of HMPAO brain scans could be based on a similar scheme. For example, it should be possible to combine pre-segmented Magnetic Resonance scan with normal scintigraphic images to quantify abnormalities in various areas. Work in this direction was reported by Evans et al. (30). They used an adjustable 3-D brain atlas to correlate anatomical areas on PET studies. A similar idea of geometrical 3-D templates for brain quantification was reported by Hooper et al. (31). Our method could be applied to automate these techniques. Another possible application could be the alignment and quantification of ventilation/perfusion SPECT scans.

The relative quantification of counts in segments provides a means to derive essential features and characteristics of the myocardial scan, condensing them into a small set of numeric values. Such values can be used in a variety of ways. One possible application is an artificial intelligence based software that could aid the physician in patient diagnosis (32, 33). Since the alignment procedure is done without user intervention, it would be possible to initiate a fully automated diagnostic system by integrating the results of scintigraphic tests with other clinical data.

7.5 SUMMARY

To optimize the interpretation of myocardial SPECT, we developed an automated method for alignment, sizing, and quantification of images using 3-D reference templates. Stress and rest reference templates were built using a hybrid 3-D image registration scheme based on principal-axes and simplex-minimization techniques. Thirty-eight normal patient studies were correlated to a common orientation, position, and size. Aligned volumes were added to each other to create amalgamated templates. Separate templates were built for normal stress and rest SPECT Tc-99m sestamibi scans of males and females. The same algorithm was used to correlate abnormal patient studies with respective normal templates. The robustness of the fitting algorithm was evaluated by registering data with simulated defects and by repeated registrations after arbitrary misalignment of images. To quantify regional count distribution, eighteen 3-D segments were outlined on the templates, and counts in segments were evaluated for all patients. Our technique provided an accurate and reproducible alignment of the images and compensated for varying dimensions of the myocardium by adjusting scaling parameters. The algorithm successfully registered both normal and abnormal studies. The mean registration errors caused by simulated defects were 1.5 mm for position, 1.3° for tilt, and 5.3% for sizing (stress images), and 1.4 mm, 2.0° and 3.7% (rest images); these errors were below the limits of visual assessment. 3-D myocardial perfusion templates provide a direct reference for comparison of individual patient's images. We conclude that automated 3-D image fitting to normal templates can be used for reproducible quantification of myocardial SPECT, eliminating operator-dependence of the results. The proposed method has potentially wider application to other types of nuclear medicine images.

ACKNOWLEDGEMENTS

The author thanks Dr. Frank Prato for useful discussions, and Ms. Soraya Ali for help in preparing the manuscript.

REFERENCES:

1. Garcia EV. Quantitative myocardial perfusion single-photon emission computed tomographic imaging: Quo vadis? (Where do we go from here?). *J Nucl Cardiol* 1994;1:83-93.
2. Garvin AA, Cullom JS, Garcia EV. Myocardial perfusion imaging using single-photon emission computed tomography. *Am J Cardiac Imag* 1994;8:189-198.
3. Van Train KF, Berman DS, Garcia EV, et al. Quantitative analysis of stress thallium-201 myocardial scintigrams: A multicenter trial. *J Nucl Med* 1986; 27:17-25.
4. Garcia EV, Cooke CD, Van Train KF, et al. Technical aspects of myocardial SPECT imaging with technetium-99m sestamibi. *Am J Cardiol* 1990; 66:23E-31E.
5. Ezekiel A, Van Train KF, Berman D, Silagan D, Maddahi J, Garcia EV. Automatic determination of quantification parameters from Tc-sestamibi myocardial tomograms. In: *Computers in Cardiology*. New York: IEEE, 1991:237-240.
6. Mullick R, Ezquerra NF, Garcia EV, Cooke CD, Folks RD. 3D visualization of pose determination: application to SPECT imaging. *Visualization Biomed Computing Proc SPIE* 1992; 1808:225-234.
7. He Z, Maublant JC, Cauvin JC, Veyre A. Reorientation of the left ventricular long-axis on myocardial transaxial tomograms by a linear fitting method. *J Nucl Med* 1991; 32: 1794-1800.
8. Cauvin JC, Boire JY, Maublant JC, Bonny JM, Zanca M, Veyre A. Automated detection of the left ventricular myocardium long axis and center in thallium-201 single photon emission computed tomography. *Eur J Nucl Med* 1992; 19:1032-1037.
9. Umeyama S. Least squares estimation of transformation parameters between two point patterns. *IEEE Trans Pattern Anal Machine Intell* 1991; 13:376-380.
10. Arun KS, Huang TS, Blostein SD. Least squares fitting of two 3-D point sets. *IEEE Trans Pattern Anal Machine Intell* 1987; 9:698-700.
11. Besl PJ, McKay ND. A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Machine Intell* 1992; 14:239-256.
12. Pelizzari CA, Chen GTY, Spelbring DR, Weichselbaum R, Chen CT. Accurate three-dimensional registration of CT, PET, and/or MR images of the brain. *J Comput Assist Tomogr* 1989;13:20-26.

13. Junck L, Moen JG, Hutchins GD, Brown MB, Kuhl DE. Correlation methods for the centering, rotation and alignment of functional brain images. *J Nucl Med* 1990; 31:1220-1226.
14. Turkington TG, Jaszczak RJ, Pelizzari CA, et al. Accuracy of registration of PET, SPECT and MR images and MR images of a brain phantom. *J Nucl Med* 1993; 34:1587-1594.
15. Faber TL, Stokely EM. Orientation of 3-D structures in medical images. *IEEE Trans Pattern Anal Machine Intell* 1988; 10:376-380.
16. Alpert NM, Bradshaw JF, Kennedy D, Correia JA. The principal axes transformation - a method for image registration. *J Nucl Med* 1990; 31:626-633.
17. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical recipes in C*, 2nd edition. New York NY: Cambridge University Press, 1992;408-412.
18. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical recipes in C*, 2nd edition. New York NY: Cambridge University Press, 1992;123-128.
19. Hoh CK, Dahlbom M, Harris G, et al. Automated iterative three-dimensional registration of positron emission tomography images. *J Nucl Med* 1993; 34: 2009-2018.
20. Venot A, Liehn JC, Lebruchec JF, Roucayrol JC. Automated comparison of scintigraphic Images. *J Nucl Med* 1987; 27:1337-1342.
21. Russ JC. *The image processing handbook*. Boca Raton FL: CRC Press, 1992;240.
22. Berman D, Kiat H, Van Train KF, Garcia EV, Friedman J, Maddahi J. Technetium-99m-sestamibi in the assessment of chronic coronary artery disease. *Semin Nucl Med* 1991; 21:190-212.
23. Myers JL. *Fundamentals of experimental design*. Boston MA: Allyn and Bacon Inc., 1979;162-198.
24. Lamas GA, Pfeffer MA. Left ventricular remodelling after acute myocardial infarction: clinical course and beneficial effects of angiotensin-converting enzyme inhibition. *Am Heart J* 1991; 121:1194-1202.
25. Akinboboye OO, Haines FA, Atkins HL, Oster ZH, Brown EJ. Assessment of left ventricular enlargement from planar thallium-201 images. *Am Heart J* 1994; 127:148-51.

26. Gonzales RC, Woods RE. *Digital image processing*. Reading MA: Addison-Wesley Publishing Company, 1992;298.
27. Houston AS, Kemp PM, Macleod MA. A method for assessing the significance of abnormalities in HMPAO brain SPECT images. *J Nucl Med* 1994; 35:239-244.
28. Maublant JC, Peycelon P, Kwiatkowski F, et al. Comparison between 180° and 360°, data collection in technetium-99m MIBI SPECT of the myocardium. *J Nucl Med* 1989; 30:295-300.
29. Kotzerke J, Hicks RJ, Wolfe E, et al. Three-dimensional assessment of myocardial oxidative metabolism: A new approach for regional determination of PET-derived carbon-11-acetate kinetics. *J Nucl Med* 1990; 31:1876-1893.
30. Evans A, Beil C, Marret C, Thompson C, Hakim A. Anatomical-functional correlation using an adjustable MRI-based ROI atlas with positron emission tomography. *J Cereb Blood Flow Metab* 1988; 8:513-530.
31. Hooper HR, McEwan AJ, Lentle BC, Kotchon TL, Hooper PM. Interactive three-dimensional region of interest analysis of HMPAO SPECT brain studies. *J Nucl Med* 1990; 31:2046-2051.
32. Fujita H, Katafuchi T, Uehara T, Nishimura T. Application of artificial neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull's-eye images. *J Nucl Med* 1992; 33:372-276.
33. Herbst MD, Garcia EV, Cooke CD, Ezquerra NF, Folks RD, De Puey EG. Myocardial ischemia detection by expert system interpretation of thallium-201 tomograms. In: Reiber JHC, van der Wall EE, eds. *Cardiovascular nuclear medicine and MRI*. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1992:77-78.

CHAPTER 8: AUTOMATED CHARACTERIZATION OF MYOCARDIAL PERFUSION DEFECTS BY COMPARISON TO REFERENCE TEMPLATES DEMARCATED WITH HYPOPERFUSION TERRITORIES*

8.1 INTRODUCTION

An important aspect of the myocardial perfusion interpretation is the localization and sizing of diseased portions of the myocardium. The sites of perfusion defects should correspond to anatomical locations of diseased arteries. Therefore, by analyzing defect location it should be possible to optimally predict which of the three arteries is involved in coronary artery disease. Furthermore, false-positive defects positioned outside of the perfusion zones could be excluded. Size and severity of the abnormal areas can be related to the extent of the hypoperfusion. Visual methods of assigning the defect to a given coronary artery are highly subjective because of differences in the coronary artery anatomy and the variation of defect sizes. Thus, an automated method for determining the location and extent of defects could be a valuable step towards computer-aided interpretation of myocardial perfusion images. The goal of the study reported in this chapter was to develop a method for (a) automated assignment of defects to the individual coronary arteries and (b) characterization of defect size and severity.

Previously we developed an algorithm for the automated fitting of patient images to three-dimensional (3-D) normal templates (chapter 7). This study seeks to demarcate the

*The content of this chapter has been submitted for publication. (Slomka PJ, Hurwitz G, St. Clement G, Stephenson JA. Three-dimensional demarcation of myocardial perfusion zones corresponding to specific coronary arteries: towards automated and objective scan interpretation. *J Nucl Med* November 1994). Gil Hurwitz designed patient selection criteria. Grace St. Clement assisted me during the experiments. Janice Stephenson verified automatically derived results. I wrote the software, designed, and conducted all experiments.

normal templates with gender-specific 3-D maps of myocardial perfusion territories corresponding to specific coronary arteries, which would allow automated assignment of defects to perfusion territories. The maps of defect territories in combination with normal templates effectively provide a 3-D reference atlas for an automated interpretation of myocardial perfusion. By fitting and comparing the individual-patient's images to such a model, defects could be characterized in an objective manner. The use of such maps is investigated and a scheme for 3-D defect quantification based on region-growing algorithm (1) is proposed. This new scheme does not require an arbitrary segmentation of the reference models as presented in chapter 7.

8.2 METHODS

8.2.1 Patient selection criteria

One hundred sixty-eight male and female single-vessel disease and normal patients were selected from a population of approximately 400 correlated studies of tomographic perfusion and angiography, which were accumulated over a period of two years. Studies were excluded if: (a) coronary artery bypass surgery had previously been performed, (b) either study (scan or angiogram) was performed within one week of myocardial infarction, or (c) angiographic or other data suggested another cause of the disease (rheumatic valvular disease, idiopathic cardiomyopathy etc.). If angioplasty had been previously done, it was required that a period of three months elapse before the follow-up assessment, and that the angiogram and scan be performed within three months of each other. If there had been no previous angioplasty, an interval of four months between scan and angiogram was allowed. The angiograms and scintigraphic studies were ordered on the usual clinical indications, and were interpreted and reported by standard clinical procedure.

All angiograms were interpreted by a single cardiovascular radiologist blinded to the scintigraphic results and rated as previously reported (2). Angiographically normal cases used for the construction of normal gender-specific templates met the following criteria:

(a) no left mainstem stenosis greater than 30%, (b) no stenosis in the main arteries or branches greater or equal to 50% luminal diameter. Cases with a history of myocardial infarction but no significant stenoses were excluded if (a) Q-waves were present, or if (b) there was a history of a non-Q-wave infarction with an unequivocal resting scan defect, or abnormal contrast ventriculogram. Cases were designated as having single-vessel disease in the LAD, RCA or LCX sites if they met the following criteria: (a) no left mainstem lesion larger than 30%, (b) a lesion of 70% or greater in one of the three main arteries, (c) no lesions in other main arteries or branches of 50% or greater; for the LCX site only, because of the relative paucity of cases, a second site of up to 50% stenosis was allowed. Cases with LAD stenoses were further divided into those with proximal (PLAD) disease (before the first septal perforator artery) and those with more distal disease (DLAD).

8.2.2 Exercise, acquisition and reconstruction protocols

Each patient underwent an individually-tailored stress protocol with intravenous dipyridamole, exercise, or a combined procedure, as previously described (3). The injected dose was dependent on the patient's weight (0.32 mCi per kg in exercise study). Stress images were collected 1 hour after the injection. The acquisition and reconstruction protocols were identical to those described in section 7.2.2. In the majority of studies, most of the external activity was eliminated automatically by masking the reconstructed volumes (section 7.2.2). In 15 cases significant areas of activity in the abdomen were not sufficiently removed by the automatic masking procedure. These areas were manually masked out from the reconstructed slices by interactively drawing regions of interests. The reconstruction and the masking operation were the only steps requiring user interaction. Rest images were not used in this study because the goal was to localize the hypoperfusion zones.

8.2.3 Computer processing algorithms

All computer algorithms were implemented in the C programming language using the Solaris operating environment and a SUN SPARC 10/512 workstation.

Three-dimensional image registration. Reconstructed stress images were aligned to a common 3-D orientation and size by the algorithm previously developed in our laboratory (chapter 7). Briefly, the hybrid algorithm, using principal-axes transformation (4) and iterative minimization (5) of count difference, aligned and sized images in 3-D by independently adjusting nine transformation parameters (three scalings, three translations, and three rotations). Each iteration involved a full reorientation of image volume (64x64x32 voxels).

Creating composite mean and variation templates. The aligned images were used to create three-dimensional reference templates. A refined, two-stage approach was used to build the templates (compare with section 7.2.4). In the first stage, patients' images were sequentially registered and cumulatively added, as described previously as described in section 7.2.4. The male template built in this way was then resized to the mean scaling parameters in three directions (X, Y, Z); the mean scaling parameters were derived from the registration results of the male group. Therefore, the dimensions of the male template reflected average dimensions in this group. The female template was resized to the male template dimensions, which allowed direct comparison of both groups.

In the second stage, the original images of normal patients were re-registered to the resized templates. This time, all individual scans were registered to the same (previously accumulated) templates. The individual studies were normalized using the reciprocal of the ratio of maximum-count (averaged in the 3x3x2 voxel region) in the patient's images, and the count value at the corresponding location in the mean template derived in the first stage. This method of normalization is different than that described in section 7.2.4. which was applied in the previous study. We found that the values of the coefficient of variation

at voxels located around the high-lateral area were approximately 2% higher if normalization was performed in the fashion described in section 7.2.4. Subsequently, lower values of the coefficient of variation resulted occasionally in the creation of spurious abnormal regions in the high lateral area during the quantification stage. This phenomenon may be related to the fact that the maximal voxel in the template were at a different location than the maximal voxels in the patient image, resulting in an artificially low variation of normalized values in certain regions. This however, did not affect the quantification results performed in chapter 7 as the variation values in the segmental method (section 7.2.5) were derived from a large number of voxels. Moreover, the mean templates built using these two normalization schemes looked identical.

A second template was then constructed by calculating the mean of normalized counts at each voxel. Thus, the construction of templates during the second stage was independent of the order in which the patients' images were registered; therefore, true mean voxel values were calculated. Subsequently, the 3-D variation templates were obtained by calculating the coefficient of variation (standard deviation divided by the mean) of the count distribution at each voxel. To test whether there was any residual effect of patient order on the final template, we accumulated the templates using 2 different initial patients and then compared the amalgamated templates; these templates were visually assessed as identical. Quantitatively the maximum voxel difference was 0.5% between such templates.

Finally, all normal and abnormal patients' datasets were registered to gender-specific templates created in the second stage. The process was fully automatic, and the same algorithm and parameters were applied to all patients, without any individual adjustments. An experienced nuclear medicine technologist visually assessed the quality of the automated fitting for each patient, but manual adjustment of registration parameters was not permitted. Abnormal patient datasets were then used to build separate mean templates representing the disease in particular territories (PLAD, DLAD, RCA, LCX). In

total, 4 male (PLAD, DLAD, RCA, LCX) and 3 female (LAD, RCA, LCX) disease templates were constructed. Disease templates were built in the same manner as the normal ones. Normal and disease templates were correlated to the same orientation and size, because all individual studies were registered to the same reference.

Template-based region-growing technique. To estimate size and location of perfusion defects in the patients' images and in the disease templates, a 3-D region-growing scheme was implemented, based on a 2-D version of the algorithm (1). The algorithm had a recursive character and used the variation and mean templates to classify image voxels.

The counts in patients' images were normalized using the reciprocal of the ratio of maximum-count (averaged in the 3x3x2 voxel region) and the count value at the corresponding location in the mean templates. The region-growing was initiated from the point where the difference between voxel values of the individual-patient data and the mean template divided by the coefficient of variation, was maximal. To reduce the effect of noise, this difference was evaluated in 3x3x2 voxel region. The points in the 3-D neighbourhood of the initial point were then considered. New voxels were classified as defect-voxels if their value was smaller than the mean minus a predefined number of standard deviations. This threshold was set arbitrarily to 2.0 standard deviations for all patients. Each voxel satisfying the defect criteria became the new starting point for the recursive algorithm. When the search for abnormal voxels was exhausted, all voxels belonging to the region were marked. The region growing procedure was repeated another two times, avoiding the already marked voxels. Therefore, it was possible to derive up to three separate defects on the patient data. This repetition of the region-growing procedure was necessary to ensure that all defective areas were included. Although in this study single-vessel disease patients were used, in a few cases (n=4) the first detected region was a small spurious defect (usually on the edge) and the true defect was found after the algorithm was re-applied. During the third application of the region-

growing algorithm small defects were found in only 7 out of 168 patients; none of them was related to the primary hypoperfusion site. Nevertheless, in most cases only one region was found, or the secondary regions were negligible in size (< 1% of myocardium); thus, multiple regions were merged for the purpose of further calculations. Regions were saved in the form of 3-D bitmaps.

Determining the territories of specific coronary arteries. To estimate the average location of defects in each angiographic group, the region-growing procedure was applied to the disease templates. Because the templates, accumulated from many patients, were very smooth in appearance, the thresholds for the region-growing procedure were reduced. The threshold was individually adjusted for each of the disease templates (4 male and 3 female). This was done to obtain the size of the coronary artery territory equal to the average size of the defects derived individually for each patient. The resulting disease regions were coded as 3-D bitmap images, allowing voxels to belong to more than one region (thus permitting overlap between territories). Separate 3-D territory maps were constructed for males and females. The DLAD territory was mostly included in the PLAD region; therefore, these 2 regions were merged for quantification purposes. Thus, three separate hypoperfusion regions (LAD, LCX, RCA) were established in both the male and female groups.

Quantification of the extent, severity, and location of defects on patient data. The region-growing procedure, using the mean and variation normal templates, was applied to each of the test-patient's images. The percentages of the defect belonging to each territory and to undesignated areas were calculated. This was done by counting the overlapping voxels belonging to the defect and a particular perfusion zone. Also, the percentage of each territory covered by the patient's defect was calculated. Several other quantitative parameters were derived for the abnormal regions. The derived parameters included the following: absolute (cm^3) and relative volume of the defect (percentage of myocardial volume), severity of the defect (average number of standard deviations below the mean),

centroid coordinates, and severity product (a multiplication of the severity and defect size). The absolute volume was calculated by dividing the volume of the resized defect by X, Y, and Z scaling parameters. The severity product provided a measure of missing counts in the myocardium.

Visualization of the defects. Two approaches were used to visualize the derived hypoperfusion defects. In the first method, the voxels belonging to the defect were marked by setting the abnormal patient-data voxels to an arbitrary predefined value, different from any data voxel value. In this study the value equal to twice the maximum-count was selected. Then, a color look-up table was constructed, which contained a distinct color corresponding to the defect value. Such modified patient data could be used with any existing display program. The second method was to overlay the bitmap images representing the defect data on the original patient data. The overlay effect was accomplished by interlacing pixels from the defect map and the patient image. Patient data could also be displayed in the original dimensions (before alignment to the template). For this purpose, inverted registration scaling parameters were used to resize the patient images and defect bitmaps.

8.3 RESULTS

The normal templates were similar in appearance to previously constructed templates (chapter 7: Figs 7.4 and 7.5). Tomographic slices of the male variation template are shown in Figure 8.1. An example of the region-growing and the quantification results in a test-patient with the LCX defect is shown in Figure 8.2. The LCX, RCA and PLAD disease templates with superimposed defect territories in males are shown respectively in Figures 8.3 and 8.4. The DLAD territory was mostly included in the PLAD region; therefore these 2 regions were merged for quantification purposes.

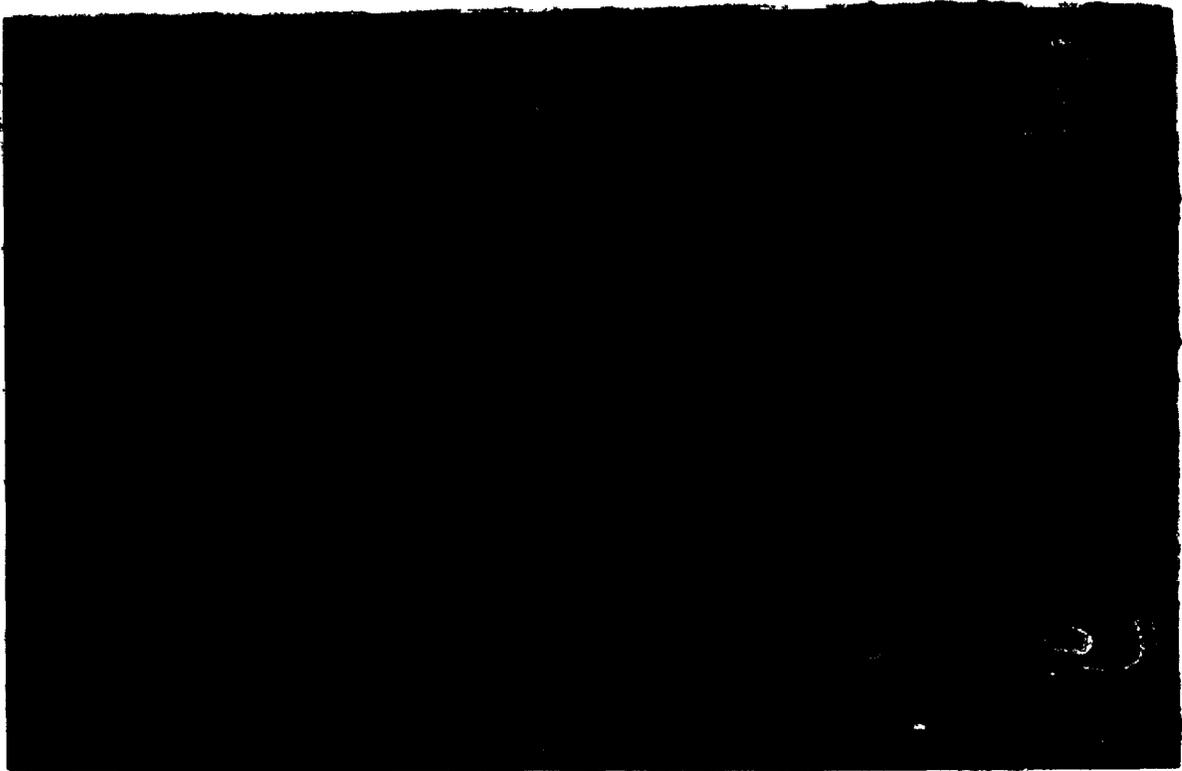


Figure 8.1: Horizontal long-axis (top), short-axis (middle) ,and vertical long-axis (bottom) of the variation template. Higher variation is observed on edges due to shape differences between images included in the template. The lowest values of the coefficient of variation (blue, 7%) are located in the high lateral regions; the highest values (red, 50%) are located on the edges. The voxels outside the mean template outline (defined as less than 15 % of the maximum) are masked (Fig. 7.4).

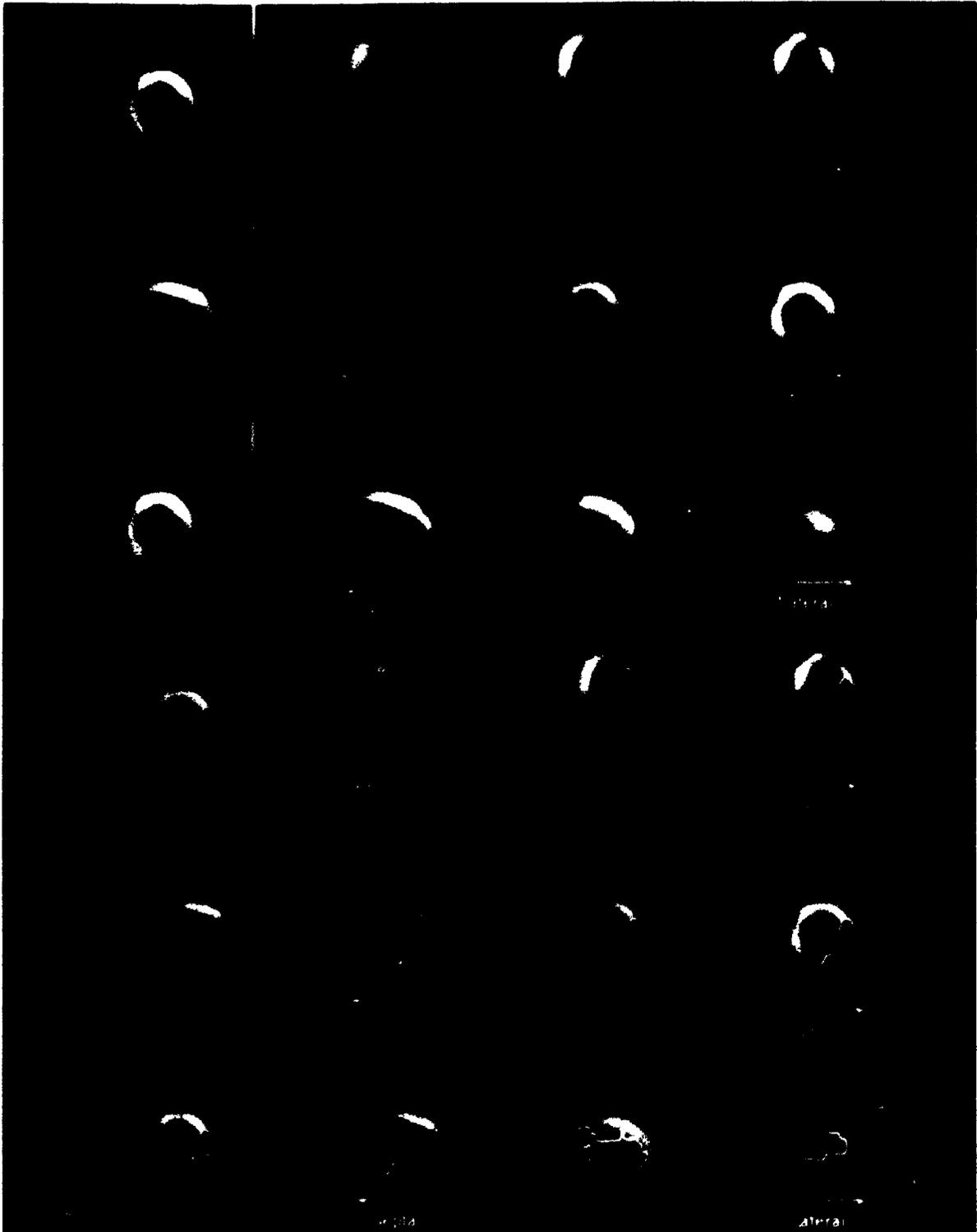


Figure 8.2: Small LCX defect in test-patient (top) and results of region-growing procedure (bottom). The region below 2 s.d. is marked as green. The results of the quantification are as follows: number of voxels (455), absolute volume (30 cm³) relative defect size (11%); defect severity (3.2); severity product (29); percentage of defect contained in territories: LAD (0%), RCA (26%), LCX (89%), undesignated (9%).

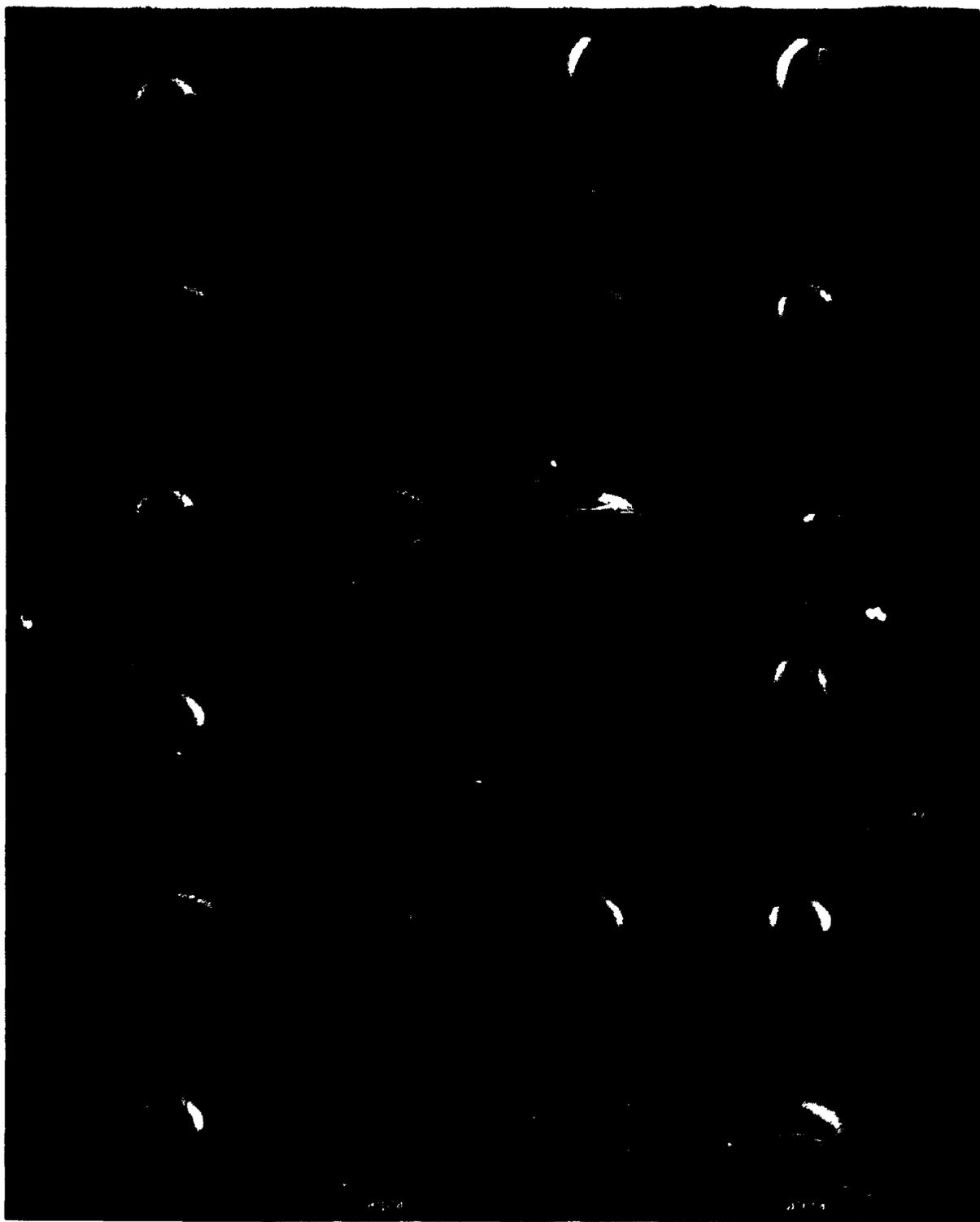


Figure 8.3: Average LCX (top) and RCA (bottom) territories (green) overlaid on the respective disease templates. The territories were derived by region-growing on the disease templates: LCX template included 14 patients, and RCA template 28 patients.

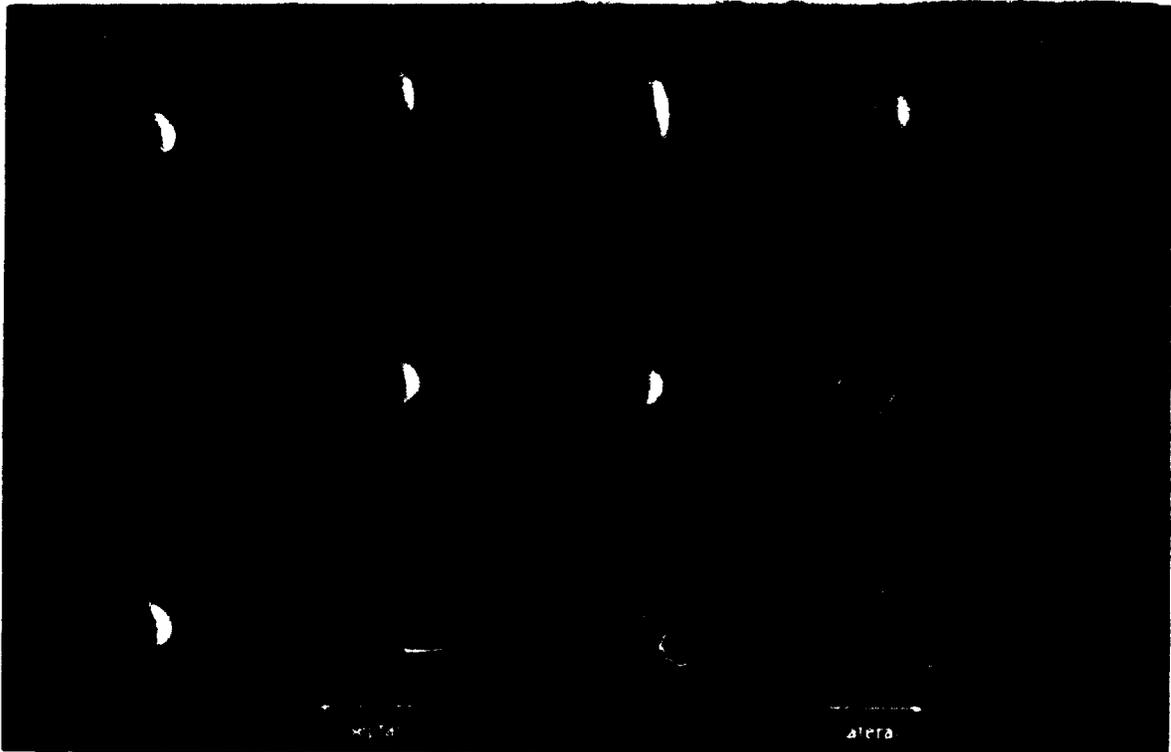


Figure 8.4: Average PLAD territory (green) overlaid on PLAD template, accumulated from 26 patients.

The overall performance of the algorithm is summarized in Table 8.1. Studies with defects $\geq 2\%$ of myocardial volume were considered abnormal. Defects were assigned to the perfusion zone that contained the largest portion of the abnormal region. The defect was considered as unassigned if more than 50% of defect volume was positioned in the undesignated area, and it did not cover more than 10% of any perfusion territory. Based on these arbitrary criteria, defects were identified in 9 out of 49 normal patients, but only 3 of these were in regions of the myocardium designated as perfusion zones. Only 1 of these 3 cases was interpreted as normal by visual analysis. In diseased patients, a unique defect was identified in the majority of the cases (108 out of 119); defects were generally located within the territory of the appropriate artery, given the considerable overlap between RCA and LCX zones. Non-specific defects in the diseased groups were rare (3 out of 119); thus 105 cases in the diseased groups were classified as abnormal and 14 as normal; 9 of these 14 cases were interpreted visually as normal and 5 as borderline or showing mild disease. The average quantification results for all studied

groups are presented in Table 8.2. The RCA defects in the male group were generally smaller than defects in other territories, probably due to the attenuation artefact. Higher severity products generally corresponded to larger defects.

To assess the effect of possible misalignment on the region-growing procedure, we visually examined the quality of fit and region-growing results in all individual cases. We observed small discrepancies of tilt (< 4 deg) and sizing (< 5 %) in 5 cases with a large amount of external activity adjacent to the myocardium, which had not been removed by masking procedures. In 3 cases with severe defects in LAD territory, a "ballooning" of the apex significantly modified the shape of the ventricle. The registration program correctly aligned such data, but part of the myocardial wall was positioned outside of the template and the region-growing slightly overestimated the size of the defects. In 10 other cases the region-growing produced small defects due to shape differences between the registered patient data and the normal template; nevertheless, these defects, were negligible in size ($< 2\%$), or located outside the expected defect territories. Such small differences in shape occurred sometimes at the edge of the myocardium. Presumably the high variation values on the edges (Fig. 8.1) prevented the creation of spurious defects of a larger size.

We observed a significant overlap of territories between RCA and LCX regions in males (Fig. 8.3), which caused an incorrect assignment of the perfusion zone in some cases (Table 8.1). The male LCX group contained 2 cases in which the defect extended significantly toward the high lateral zone, in other cases ($n=10$) it was located in the inferolateral and lateral regions. The average RCA defects were smaller than the LCX defects in males; therefore, in some RCA patients most of the defect was actually located in the LCX zone (Tables 8.1 and 8.2). In females, however, these territories were better separated (Tables 8.1 and 8.2).

TABLE 8.1: Number of patients with defects in designated territories.

Site of stenosis	n	>2%*	Defect assigned to the territory			
			LAD	RCA	LCX	Unassigned
None	20	4	1	0	0	3
LAD†	12	11	11	0	0	0
RCA	8	8	0	8	0	0
LCX	7	5	0	1	3	1
All Females	47					
None	29	5	0	0	2	3
PLAD	26	24	23	0	0	1
DLAD	24	23	22	0	0	1
RCA	28	24	2	15	7	0
LCX	14	13	0	3	10	0
All Males	121					

* size of the defect greater than 2% of the myocardial volume
† in female group the proximal and distal LAD cases were combined in one category

TABLE 8.2: Average characteristics of defects.

Site of stenosis	Size %	Severity product†	% of defect in *			
			LAD	RCA	LCX	Unmarked region ‡
Female						
None	1±3	3±7	-	-	-	-
LAD	21±15	78±65	65	21	6	23
RCA	22±13	78±47	13	72	16	15
LCX	16±16	58±56	6	19	38	34
Male						
None	2±5	4±12	-	-	-	-
PLAD	26±15	104±74	71	2	3	22
DLAD	23±14	90±60	65	7	8	21
RCA	13±11	37±38	18	47	38	17
LCX	23±16	91±67	8	36	55	21

* Part of the defect contained in a given territory. Due to territory overlap the sum of the percentage contained in all territories can be higher than 100%
† Average number of standard deviations below the mean in the defect region multiplied by the defect relative size (in % of myocardium).
‡ average % of the defect positioned outside of all designated territories

8.4 DISCUSSION

8.4.1 Quantification technique

Demarcated perfusion maps were applied to assign lesions to particular arteries and to discard false-negative defects positioned in undesignated areas. They can be used together with normal templates to detect and classify defects in a fully automated fashion. The region-growing technique classified defect voxels into contiguous regions. Such a representation of defects facilitated the calculation of quantitative parameters. For example, the average position or size of individual defects could be estimated. Furthermore, spurious regions of abnormal voxels were not marked because the algorithm searched for only a limited number of defects. Robust performance of the region-growing algorithm, which is reflected in the small number of false-positive results, could be attributed to the incorporation of the 3-D standard deviation map and accurate alignment of test-images. The application of the multiple region-growing technique will have to be further evaluated in the detection and localization of multiple vessel disease. Automated voxel grouping during the region-growing procedure distinguishes our technique from the bull's eye approach (6,7). In the bull's eye method pixels on the polar map are tested against normal criteria and subsequently a clustering algorithm is applied to eliminate isolated abnormal pixels.

The scheme for quantifying and displaying defects employed in this study may offer advantages compared to the methods based on polar map (bull's eye) representation of the left ventricle (6,7). Our 3-D voxel-based technique does not use the concept of maximal-count on circumferential or radial profiles (section 6.4.1); thus, it preserves information about the thickness of the myocardial wall. Consequently, apparent non-transmural portions of defects can be detected, which would not be marked as abnormal on the polar map. Although the resolution of myocardial perfusion scan is not sufficient to depict such lesions accurately (8), a more representative average estimate of the defect may be obtained by including all voxels below normal values. Furthermore, our technique allows one to obtain volumetric, quantitative parameters including the absolute

size of the defect, relative size in comparison to the volume of the myocardium, measure of missing counts, and volume overlap with expected hypoperfusion zones (Table 8.2 and Fig. 8.2).

The natural representation of quantification results in our approach simplifies the interpretation of images. If there are position, alignment, or shape differences between the patient data and reference model, these discrepancies are immediately evident and directly compared with the original data because both the actual patient data and the defect are visible (Fig. 8.2). Integrated images of tomographic slices and defects can be displayed using existing conventional display formats. Therefore, the interpretation of quantitative results as in cases of borderline scans or when motion and attenuation artefacts are present, may be more straightforward. It has been shown that errors in specifying sizing and alignment parameters can produce artefacts on the polar map display (9,10). These errors may produce inaccuracies in quantification and localization, thus an analysis of the original images is always recommended (9,10). Several approaches have previously been developed to visualize the maximal-count profiles in three dimensions (11,12). These techniques plot the values of the circumferential/radial profiles derived from the short-axis slices on the 3-D surface. Such visualization can camouflage positional or distortion errors, in the same manner as the 2-D polar map.

The fully automated fitting of patients' images to the reference templates in our approach eliminated potential operator subjectivity and reduced alignment errors as compared to manual methods. Although the bull's-eye method has been shown to be highly reproducible when combined with the subjective visual interpretation of scans (13), the adjustment of orientation parameters and the interpretation of scans may require experienced readers to achieve good reproducibility.

8.4.2 Overlap of the territories

Previously DePasquale et al. (14) and Maddahi et al. (15) applied the bull's eye quantification scheme (6) to characterize the territories of the three major coronary arteries. The left anterior descending (LAD), right coronary artery (RCA), and left circumflex (LCX) territories were reported in a form reduced to the 2-dimensional (2-D) polar map; on this display, these three territories were represented as non-overlapping geometric areas. In our study, we found a significant overlap between the LCX and RCA perfusion territories in the male group. Our approach, however, differed in several aspects: (a) Tc-99m-sestamibi scans were analyzed, (b) automated data registration was used, (c) the defect maps were created on 3-D templates, (d) somewhat different angiographic criteria were assigned. Specifically, no overlap of the LCX and RCA territories in previous studies (14-15) could have been the result of the different angiographic criteria for dividing the patients into these two groups. For example, some LCX patients were assigned to the RCA group based on the dominance of the arterial system, which is normally not known *a priori*. The overlap between RCA and LCX territories was also reported in echocardiography studies (16).

8.4.3 Image registration technique

The fitting errors of the registration procedure were assessed earlier using simulated-defect and random misalignment tests (sections 7.3.4 and 7.3.5). In this study, small visual discrepancies in 5 severely abnormal patients were found. These observations were based on the subjective interpretation of the images, which cannot be used as the absolute standard. Visual assessment of registration errors varied significantly, depending on viewing parameters such as orientation, color table and display threshold. Automated methods of reducing registration errors caused by severe perfusion defects may be desired in some cases. One conceivable method, for example, is a two-pass registration; in the second pass the template voxels marked as defective would not be considered in the evaluation of count difference.

8.4.4 Limitations of the study

Because the goal of our study was to accurately determine the perfusion territories, the patient population was limited to single-vessel disease cases with well defined defects according to strict angiographic criteria. A prospective evaluation of defect characterization in other groups will be required during the clinical validation of our method (17,18). A full validation of the technique would include multi-vessel disease patients, rest studies and broader angiographic criteria. At this stage, we did not attempt to optimize the criteria for detection of defects or assignment of defects to the perfusion zones, and the threshold for region-growing procedure was arbitrarily set to the 2 standard deviation limit. Moreover, rest studies were not quantified at this time; the comparison between the regional defect quantification of stress and rest images will be necessary to provide an estimate of ischemia.

8.4.5 Other applications

Because our technique does not assume a particular shape or character of the analyzed data, it is possible to use a similar paradigm in the analysis of other types of three-dimensional medical images. Both the normal template concept, and the template-based region-growing method can be used to evaluate deviations from normal count distribution in an automated fashion. In particular, it should be possible to use the same quantification and visualization techniques to analyze brain Positron Emission Tomography (PET) and SPECT studies as well as lung SPECT studies.

8.4.6 Use with automated artificial intelligence systems

Since the developed method of defect localization and quantification does not require user intervention, it would be possible to integrate it with higher level logical software, thus providing a fully automated patient diagnosis. The results of our quantification constitute a relatively small set of parameters, which precisely describes the location, extent, and severity of the defect. These parameters are well suited as inputs to higher level medical diagnostic software. One possibility is to create a partial diagnostic system that only

includes a small number of clinical variables. However, the inclusion of additional information not contained in the images is debatable, since referring physicians use the same information again, together with the results of the nuclear medicine procedure. Perhaps a better solution would be to separate the imaging results from the clinical results. Then, a general-purpose medical decision-making software, involving many variables, could use image interpretation results such as described in this chapter. Thus, from a diagnostic imaging perspective, the development of automated tools for objective image analysis is the most important step towards the computer-aided patient diagnosis.

8.5 CONCLUSIONS

By fitting images of homogeneous groups of diseased patients to normal reference templates, perfusion territories corresponding to the major coronary arteries could be demarcated. An automated method was used to discern, outline and quantify defects in patients with well-characterized single-vessel disease. In test-patients the method provided an appealing, natural display of defect extent and severity. The measure of the defect overlap with demarcated perfusion territories provided criteria for an automated assignment of defects to specific coronary arteries. These techniques may allow for fully automated and standardized interpretation of myocardial tomography and can be applied to other types of nuclear medicine images.

8.6 SUMMARY

Automated localization of myocardial perfusion defects with respect to the arterial supply could result in a more objective image interpretation. In this study, three-dimensional (3-D) maps of specific coronary artery territories were derived and combined with normal-distribution maps to provide a reference for automated characterization of defects including location and size. One hundred sixty-eight normal and single-vessel disease male and female myocardial perfusion scans were selected according to the angiographic data. Five separate groups were established for males and females: normal, proximal left anterior descending (PLAD), distal left anterior descending (DLAD), right coronary artery

(RCA), and left circumflex (LCX). All myocardial perfusion studies were aligned and sized to the same 3-D orientation using the automated image registration technique, previously developed. Mean and variation 3-D templates were constructed from stress images in each group. Normal templates were demarcated with hypoperfusion regions obtained from disease templates. The defects were detected in the individual patient's images by a region-growing algorithm which identified abnormal voxels by comparison to the corresponding voxels in the mean and variation templates. Defects were quantified with respect to volume, location relative to the expected hypoperfusion zones, and a severity index. Abnormal regions could be marked directly on tomographic slices and visualized in various orientations. Single defects greater than 2% of myocardium positioned within demarcated perfusion territories were detected in 105/119 abnormal patients and in 3/49 normal patients. Maps of myocardial perfusion zones created from images of angiographically selected patients provide a reference for the automated localization of myocardial perfusion defects. A template-based region-growing is a robust technique for volumetric quantification and localization of abnormal regions.

ACKNOWLEDGEMENTS

The authors thank Dr. Frank Prato for useful discussions, and Ms. Soraya Ali for help in preparing the manuscript.

REFERENCES:

1. Gonzales RC, Woods RE. *Digital image processing*. Reading MA: Addison-Wesley Publishing Company, 1992;298.
2. Hurwitz GA, MacDonald AC. Stenoses of the left anterior descending artery: predominant role in stress-induced pulmonary uptake of thallium-201. *Can J Cardiol* 1994, In press.
3. Hurwitz GA, Powe JE, Driedger AA, et al. Dipyridamole combined with symptom-limited exercise for myocardial perfusion scintigraphy: image characteristics and clinical role. *Eur J Nucl Med* 1990; 17:61-68.
4. Faber TL, Stokely EM. Orientation of 3-D structures in medical images. *IEEE Trans Pattern Anal Machine Intell* 1988; vol PAMI-10 no.5:376-380.
5. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical recipes in C*, 2nd edition. New York NY: Cambridge University Press, 1992;408-412.
6. Garcia E, Van Train K, Maddahi J, et al. Quantification of rotational thallium-201 myocardial tomography. *J Nucl Med* 1985; 26:17-26.
7. Garcia E, Cooke CD, Van Train KF, et al. Technical aspects of myocardial SPECT imaging with technetium-99m sestamibi. *Am J Cardiol* 1990; 66:23E-31E.
8. Galt JR, Garcia EV, Robins WL. Effects of myocardial wall thickness on SPECT quantification. *IEEE Trans Med Imag* 1990; 9:144-150.
9. De Puey EG, Garcia EV. Optimal specificity of thallium-201 SPECT through recognition of imaging artefacts. *J Nucl Med* 1989; 30:441-449.
10. De Puey EG. How to detect and avoid myocardial perfusion SPECT artifacts. *J Nucl Med* 1994; 35:699-702.
11. Cooke CD, Garcia EV, Folks RD, Peifer JW. Three-dimensional visualization of cardiac single photon emission computed tomography studies. In: Robb RA ed. *Visualization in biomedical computing, proceedings of the 13-16 October 1992 conference SPIE* 1992; 1808:671-675.
12. Loboguerro A, Perault C, Gibold J, et al. Shape preserving three-dimensional display of myocardial scintigraphic data. *Nucl Med Comm* 1994; 15:417-421.

13. Alazraki NP, Krawczynska EG, DePuey EG, et al. Reproducibility of thallium-201 exercise SPECT studies. *J Nucl Med* 1994; 35:1237-1244.
14. DePasquale EE, Nody AC, De Puey E, et al. Quantitative rotational Tl-201 tomography for identifying and localizing coronary artery disease. *Circulation* 1988; 77:316-327.
15. Maddahi J, Van Train KF, Prigent F, et al. Quantitative single photon emission computerized thallium-201 tomography for the evaluation of the coronary artery disease: optimization and prospective validation of a new technique. *J Am Coll Cardiol* 1989;14:1689-1699.
16. Segar DS, Brown SE, Sawada SG, Ryan T, Feigenbaum H. Dobutamine stress echocardiography: a correlation with coronary lesion severity as determined by quantitative angiography. *J Am Coll Cardiol* 1992;19:1197-1202.
17. Van Train KF, Areeda J, Garcia E, et al. Quantitative same-day rest-stress technetium-99m-sestamibi SPECT: definition and validation of stress normal limits and criteria for abnormality. *J Nucl Med* 1993;34:1494-1502.
18. Van Train KF, Berman DS, Garcia E, et al. Quantitative analysis of stress Tl-201 myocardial scintigrams: a multicenter trial validation utilizing standard normal limits. *J Nucl Med* 1986; 27:17-25.

CHAPTER 9: SUMMARY AND FUTURE WORK

9.1 SUMMARY

9.1.1 Principal conclusions

The main objective of this study was to develop methods for the automated analysis and interpretation for two types of nuclear medicine images: gamma camera quality control images, and tomographic myocardial perfusion images. This objective has been accomplished.

Software algorithms for comprehensive feature extraction from QC images were designed, with the purpose of implementing a knowledge based QC system. Several types of cameras (round, hexagonal, rectangular) were considered. Images representing various gamma camera defects were used in the initial evaluation of the software. Image interpretation was based on comparisons of observed image features with predefined specifications. Several image parameters and features were defined to describe patterns related to certain gamma camera faults. These numeric image parameters could indicate various simulated faults. They can be used as input data to a higher level logical diagnostic module.

Results of the QC image quantification were used as input to an expert system. A prototype of a rule-based expert system for the quality control of gamma cameras was developed using a general purpose expert system toolkit. The system was designed to guide staff in quality control procedures, to interpret results, and to perform troubleshooting when a camera is malfunctioning. If the results of the tests did not meet standards, the expert system could initiate troubleshooting reasoning. It was possible to integrate the feature extraction algorithms with the logical component. The system performed an automated diagnosis of gamma camera faults on a small set of simulated images. Limiting factors at this stage were: inexact character of knowledge used to solve

QC problems, additional information necessary to diagnose problems fully, complexity of the system, and the lack of examples of real gamma camera faults.

In an attempt to refine the rule-based expert system, we performed a comprehensive knowledge acquisition and developed an object-oriented scheme for representation of the domain concepts. A large set of QC images representing gamma camera faults was collected; analysis of these cases led to a rigid formulation of the domain knowledge. Object-oriented representation of knowledge domain concepts resulted in natural encoding of dependencies in the form of hierarchical trees. Such an approach simplified the general structure of the system as well as the mechanism of the information flow. The major remaining difficulties were as follows: non-digital character of available QC images, variety of cameras, many possible faults, and non-standard methods of interpretation.

To fulfil the second objective of this thesis, the task of an automated analysis of myocardial perfusion SPECT images was undertaken. In the first step, a novel technique based on image registration was applied to accomplish the automated reorientation and sizing of images of many patients to common coordinates. Subsequently, mean images (templates) of normal patients were built and used as a reference for test cases. Furthermore, this method demonstrated excellent reproducibility. In addition, a preliminary segmentation and quantification scheme was devised based on a division of the normal template into arbitrary segments. Count values in these segments were quantified for both stress and rest studies and results reflected angiographic findings in a small group of abnormal patients.

Subsequently, a study was initiated to demarcate normal myocardial perfusion templates with expected sites of defects. For this purpose, separate single-vessel disease patient groups were established according to the results of angiographic tests. Templates corresponding to various sites of disease were created and the average 3-D territories of

defects were obtained. The demarcation of normal templates with disease territories provided a reference system used for comprehensive regional quantification of individual-patient images. The quantification scheme did not require any operator interaction. A model-based region classification technique localized and quantified defects in a large group of single-vessel disease patients. In contrast to quantification schemes previously devised by other investigators, our technique is fully automated, uses three-dimensional data representation, and characterizes perfusion defects in a clinically relevant format.

9.1.2 Limitations of this study

The primary limitation of this research was that the methods were developed and applied only to specific types of nuclear medicine images. Therefore, these methods cannot claim to be universal, although some aspects of them may be applicable to other types of images. Another limitation was the lack of prospective validation of the developed techniques. The evaluation of the QC expert system was limited by the difficulties in obtaining test data. Due to a variety of equipment and possible faults, the development of a complete diagnostic QC system encompassing all scenarios would require a significant amount of *knowledge engineering* and coordination of resources from many institutions. Such activities are beyond the scope of this scientific project.

The goal of a prospective validation of computer-aided analysis for myocardial perfusion scans is more attainable. Although the alignment methods were comprehensively validated and the quantification methods were tested on a large group of patients, the clinical performance of the system in the evaluation of coronary artery disease was not prospectively evaluated. This limitation was mainly caused by the variability of the disease, and the limited number of angiographically-correlated scintigrams available in a single hospital. The wide variety of stress techniques, and the increasing number of radiotracers create difficulties in accumulation of a large uniform group of test-cases from other institutions. Even in a single hospital the routine procedures may change over the years, resulting in an inconsistent patient database. These limitations may be solved by

organizing appropriate multi-centre studies or by accumulating a larger database in our institution in the future. At the current time, the integration of clinical and scintigraphic data in a uniform diagnostic system has not been attempted, largely for the reasons described above. More specific ideas for future work in these directions are outlined in section 9.2.1.

9.1.3 Implications

There are several potential implications of the research described in this thesis for the professional practice of nuclear medicine. The study of a knowledge based gamma camera QC system revealed the character of data analysis and formulated the knowledge domain. This was an important step in the creation of an artificial intelligence software for this application. If an expert system based on the proposed design is fully implemented and validated, it may simplify and streamline QC procedures and provide technical knowledge in centres lacking experience in instrumentation. Apart from the application in the computer-aided diagnostic systems, the rigorous formulation of the knowledge domain attempted in this study can be beneficial in the revision of standards and guidelines for QC procedures. Perhaps, a coordination of efforts in the improvement of standards and development of computer-aided QC support systems would be beneficial.

The visual evaluation of myocardial perfusion images is presently difficult due to several fundamental factors: (a) gender-specific differences in myocardial count distribution; (b) attenuation artefacts; (c) considerable variations among individuals in size and orientation of the heart; (d) operator difficulty in identical positioning of the patient for the two sets of scans and in realignment during image reconstruction; (e) patient motion artefacts randomly impacting stress and/or rest scans; (f) pronounced variability of the baseline rest scan due to lower counts, and loss of myocardial tissue (myocardial infarction) in a sizable minority of patients; and finally, (g) the subtle character of clinically important defects in tracer distribution. The major implication of my work for the clinical practice is the potential to eliminate the difficulties outlined above by automated detection and

quantification of ischemia, consequently improving the diagnostic value of the myocardial perfusion test. This may be especially beneficial in less experienced centres, where computer-aided methods could provide a consistently high quality of diagnosis. Moreover, the automated quantification software is suitable for integration with more general artificial intelligence systems, which could provide full assessment of clinical information.

9.1.4 Strengths of the study

The major strength of the QC image interpretation system is the broad character of the problem statement covering various aspects of QC procedures and extending beyond the image interpretation task. The consideration of additional knowledge associated with the images is essential in building the diagnostic system for this application. This aspect of the work presented in this thesis demonstrates the complexity of inferring knowledge from images, which is convoluted with mental experience and common sense of the observer. This particular application of a gamma camera QC expert system is unique, and to my knowledge no other similar system has been reported.

The strong points of the research in the quantification of myocardial perfusion images are the complete automation and excellent performance of proposed methods. In particular, accurate and reproducible image alignment and sizing are essential for reliable comparison against the normal distribution as well as between stress and rest images. The automated alignment method was comprehensively validated and tested on a large group of normal and abnormal studies. The extent and severity of defects can be assessed objectively using the three-dimensional standard deviation templates and the region-growing technique. The standard deviation templates summarized the variability of the clinical material and the region-growing provided a clinically acceptable and automated method of outlining abnormal areas on the images. The user-independent nature of the software should facilitate integration with artificial intelligence decision-making systems. The parameters of the quantification results, such as relative size of the

defect, are designed in a natural, clinically-relevant format, which can be easily understood and interpreted by physicians. Moreover, an improved scheme of a true three-dimensional quantification may be more accurate and robust than previously designed methods based on the polar map concept (1-3).

9.2 FUTURE WORK

9.2.1 Implementation of the QC expert system

The schemes for QC knowledge representation and reasoning presented in chapter 5 can be utilized in the development of the comprehensive automated QC software package. The feature, artefact and fault trees can be easily expanded and the same flow of information model can be applied. The important step in such a project would be the further collection of well documented QC cases in a digital form. This may prove to be a practical difficulty and may require collaboration of several centres as well as help from the manufacturers. The quantitative methods described in chapter 3 could be then used for analysis of such images and their performance could be validated on real examples of faults. The acceptance and usefulness of the QC package will be related to the ability of the software to cope with a wide spectrum of problems including trivial procedure errors and miscellaneous vendor-specific aspects of the data acquisition. Specifically, good integration of the system with the acquisition software is essential for the implementation of the cyclic information flow (section 5.3.4), which often requires multiple acquisitions of data.

9.2.2 Prospective validation of localization of the coronary artery disease

As stated in section 9.1.2, there is a need for the prospective validation of the developed quantification methods. A prospective evaluation would have to include broader angiographic criteria including multi-vessel disease and less severe stenoses (50 to 70% luminal diameter). There are a few specific questions that could be answered by a validation study. The performance of the technique when delineating and locating multiple sites could be assessed and compared to the angiographic data. Then, the ability of

accurate quantification for both ischemia and infarction could be tested. In particular, the detection of ischemia in cases with extensive infarctions should be evaluated. The answers to these questions would require optimization of the classification criteria. Specifically, the sensitivity and specificity of this method should be tested and optimized using receiver-operator-characteristics curves (4) and perhaps artificial neural network tools (5). The optimal cut-off criteria could be derived for the size of the defects and also for the relative location with respect to the expected sites. Furthermore, the variation in normal count distribution and location of hypoperfusion sites due to selection of cases could be tested, for example using a *bootstrap* statistical technique (6). This statistical test could be performed as follows. A series of templates could be constructed using randomly selected subsets of patients from each of the normal and abnormal groups. Subsequently, the variation of counts between these templates would be evaluated. The results of this test would show to what degree the created normal and abnormal models depend on the selection of clinical material.

I do not anticipate a change of methods in the validation phase, as the developed techniques were tested in the current set of patients. I expect that the method may yield better results than the polar map quantification (7,8), due to accurate alignment of images and voxel-based quantification. A direct comparison of accuracy and reproducibility may not be possible, however, because previous techniques have not been validated for pharmacological studies. The stress protocol used in our laboratory is based on dipyridamole, which differs from the exercise test in its background characteristics (9).

At this point, the number of angiographically-correlated scintigrams was not sufficient to undertake the validation study, especially in the female group. The paucity of cases is caused by the fact that both angiographic and scintigraphic studies were required. At our institution, the Tc-99m-sestamibi studies have only been performed in the last 2 years. Assuming the current rate of referrals at Victoria Hospital continues, it will be possible to conduct the validation study in about 12 months. Another possibility would be to

undertake a multicentre validation of the developed technique. I am currently investigating the possibility of collaboration with the European COST B2 group to initiate such a project (10). If other centres participate in this study, it may also be possible to validate the technique for TI-201 tomograms as well as other stressing protocols.

9.2.3 Diagnostic system integrating clinical data

It is conceivable to develop a knowledge based system that combines the results of scintigraphic quantification and other clinical indicators to derive a clinical diagnosis; for example, to formulate a decision about a surgery. Additional data related to the perfusion test such as assessment of patient motion, patient stress level, ventricular dilation, and lung uptake could be included; these variables may alter the probability of CAD in borderline cases. The creation of such a system would require careful analysis of the clinical decision process and would have to be done with close collaboration between both nuclear medicine physicians and cardiologists. The integration of automated image analysis and a decision system could provide the final diagnostic or prognostic outcome in a fully objective manner.

The current diagnostic process in the nuclear medicine department involves limited use of clinical data. Subsequently, the referring cardiologist uses the same set of clinical variables together with the results of the nuclear medicine procedure. Thus, the data are not used independently (11). These aspects of clinical decision-making must be evaluated during the development of the diagnostic system. It may be necessary to integrate the final nuclear medicine report into more complicated artificial intelligence systems reaching far beyond image diagnosis (12). It may also prove that a clear separation between the information inferred from the images and a further medical diagnostic process is more suitable. The automated image quantification software could be then interfaced as a separate component with general purpose decision-support systems.

9.2.4 Application of developed methods to other types of clinical studies

The three-dimensional alignment and model-based quantification methods developed in chapter 7 may be applied to other types of nuclear medicine images. In particular, at our institution there is considerable interest in application of automated fitting, normal templates, and region-growing techniques to HMPAO SPECT brain quantification. To assess the feasibility of such a project, I applied the registration and region-growing procedures to a small set of HMPAO brain scans (n=5) with satisfactory results. An important consideration during such an investigation would be the establishment of criteria for the normal reference model; currently there is no absolute standard for this purpose. A possible method would be to apply statistical (13) or artificial neural network (14) techniques for the optimal classification of the results.

Another potential area for the application of developed algorithms is the analysis of lung ventilation/perfusion scans. Currently, however, these scans are acquired as planar images in most hospitals; therefore, the 3-D methods of alignment cannot be directly used. Although the modification of the alignment methods for 2-D images is straightforward, the accuracy may be much worse. Two options exist: either such a study would be done in the future using SPECT lung scans, or modified methods for 2-D alignment could be applied. After the alignment and defect quantification procedures, established decision criteria could be applied (15) using expert system software. In a preliminary assessment of this project, I developed a simple expert system consisting of 17 rules that implemented Biello criteria (15). Work in this field has been reported by other researchers (16,17). In these studies, however, images were analyzed in an interactive fashion. Thus, the possibility of automated alignment and quantification of lung scans could be perhaps the most important contribution.

REFERENCES:

1. Fujita H, Katafuchi T, Uehara T, Nishimura T. Application of artificial neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull's-eye images. *J Nucl Med* 1992; 33:372-276.
2. Herbst MD, Garcia EV, Cooke CD, Ezquerra NF, Folks RD, De Puey EG. Myocardial ischemia detection by expert system interpretation of thallium-201 tomograms. In: Reiber JHC, van der Wall EE, eds. *Cardiovascular nuclear medicine and MRI*. Dordrecht, The Netherlands: Kluwer Academic Publishers; 1992:77-78.
3. Garcia EV, Cooke CD, Van Train KF, et al. Technical aspects of myocardial SPECT imaging with technetium-99m sestamibi. *Am J Cardiol* 1990; 66:23E-31E.
4. Turner DA. An intuitive approach to receiver operating characteristic curve analysis. *J Nucl Med* 1978; 19:213-220.
5. Scott R. Artificial intelligence: its use in medical diagnosis. *J Nucl Med* 1993; 34:510-514.
6. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical recipes in C*, 2nd edition. New York NY: Cambridge University Press, 1992:691-692.
7. Van Train KF, Berman DS, Garcia E, et al. Quantitative analysis of stress TI-201 myocardial scintigrams: a multicenter trial validation utilizing standard normal limits. *J Nucl Med* 1986; 27:17-25.
8. Van Train KF, Areeda J, Garcia E, et al. Quantitative same-day rest-stress technetium-99m-sestamibi SPECT: definition and validation of stress normal limits and criteria for abnormality. *J Nucl Med* 1993;34:1494-1502.
9. Hurwitz GA, Saddy S, O'Donoghue P et al. The VEX-test (vasodilator plus exercise) for myocardial scintigraphy with TI-201 and sestamibi: effect on abdominal background activity. (submitted to *J Nucl Med*).
10. Bourguignon MH, Busemann Sokole E, Jones B, Van der Wall E. Protocols for selection of cardiac radionuclide studies as a data base of normal studies and typical patterns of diseases. *Eur J Nucl Med* 1993; 20:59-65.
11. Simons M, Parker A, Udelson JE, Gervino EV. The role of clinical data in interpretation of perfusion images [Commentary]. *J Nucl Med* 1994; 35:740-741.

12. Shortliffe EH, Hance E. *Computer based medical consultation: MYCIN* New York: Elsevier; 1976.
13. Houston AS, Kemp PM, Macleod MA. A method for assessing the significance of abnormalities in HMPAO brain SPECT images. *J Nucl Med* 1994; 35:239-244.
14. Kippenhan JS, Barker WW, Nagel J, Grady C, Duara R. Neural-network classification of normal and Alzheimer's disease subjects using high-resolution and low-resolution PET cameras. *J Nucl Med* 1994; 35:7-15.
15. Biello DR, Mattar AG, McKnight RC, et al. Interpretation of ventilation-perfusion studies in patients with suspected pulmonary embolism. *Am J Roentgenol* 1979; 133:1033-1037.
16. Gabor FV, Datz FL, Christian PE. Image analysis and categorization of ventilation-perfusion scans for the diagnosis of pulmonary embolism using an expert system. *J Nucl Med* 1994; 35:797-802.
17. Scott JA, Palmer EL. Neural network analysis of ventilation-perfusion lung scans. *Radiology* 1993; 186:661-664.

APPENDIX A: ALGORITHMS FOR EXTRACTION OF FEATURES FROM QUALITY CONTROL IMAGES

A.1 EDGE-FINDING ALGORITHM

In the first step, the algorithm creates a binary image by assigning the value *high* to pixels with values greater than 50% of the maximum, and the value *low* to the remaining pixels.

The thresholded image can sometimes contain points that do not have any neighbours with the same values. These isolated points could cause the tracking algorithm to fail; therefore, a morphological operator is applied to the thresholded binary image data which tests the 8-neighbourhood region (Fig. A.1) of each pixel.

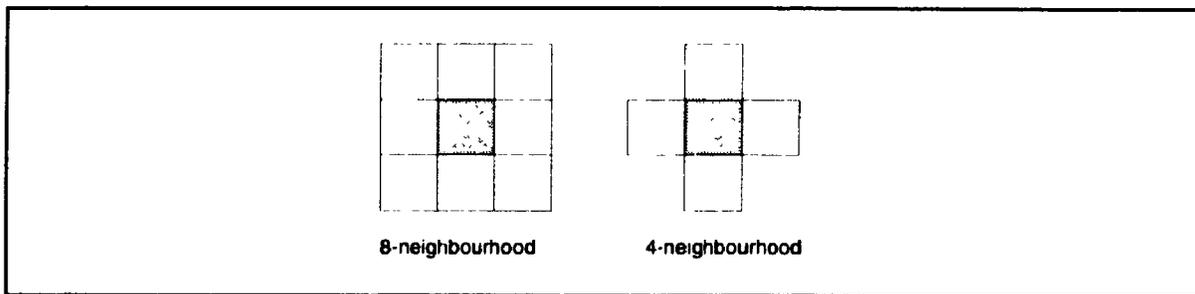


Figure A.1: 8-neighbourhood and 4-neighbourhood definition.

This morphological operator is illustrated on Figure A.2 and is expressed by

$$J(x,y) = \begin{cases} low & \text{if } low\ nbr \geq (8-N) \vee J(x,y)=low \wedge low\ nbr > N \\ high & \text{if } high\ nbr \geq (8-N) \vee J(x,y)=high \wedge high\ nbr > N \end{cases} \quad [A.1]$$

where $J(x,y)$ is the image, nbr - number of neighbours and N - the maximum number of different neighbours allowed. Such an operation is similar in concept to thickening algorithms (1). It smooths the edges and removes the dead lanes that could cause loops in the detected edge.

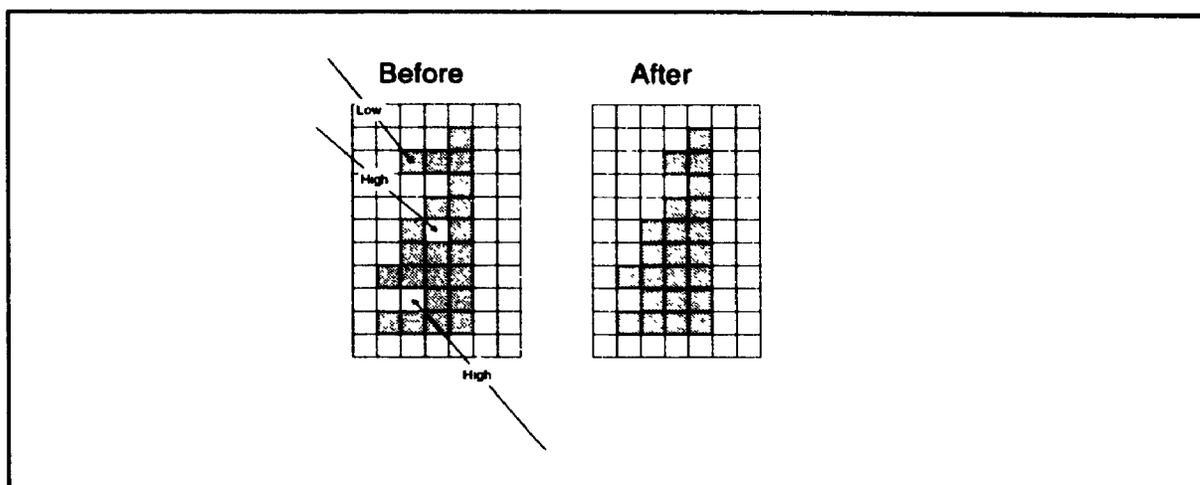


Figure A.2: Morphological operator for edge smoothing. The pixel value remains the same if there are more than 2 neighbours with the same value. For example, if a pixel with value *high* has 4 *low* neighbours and 4 *high* neighbours, its value stays *high*. If the pixel is *low* and has 7 neighbours *high* and one *low*, its value will be converted to *high*.

After these preprocessing steps, the algorithm searches for the first point on the edge from the outside, thus avoiding the possibility of tracking the edge of a cold spot inside the flood. The image matrix is expanded by a 2-pixel margin on each side, allowing the detection of the edge on the matrix border. The pixels on the line between the opposite corners of the image matrix are checked until the first edge-point is found. A point is defined as the edge-point, if in its 4-neighbourhood (Fig. A.1), there is at least one *high* pixel and one *low* pixel.

The edge-tracking algorithm is then applied, starting from the first point found, and sequentially adding pixels that satisfy neighbourhood criteria. Generally, the 4-neighbourhood is more convenient for tracking the edge, because of the speed of the algorithm and a reduced chance of small loops on the distorted edge. Therefore, the edge tracking algorithm is based on the 4-neighbourhood model.

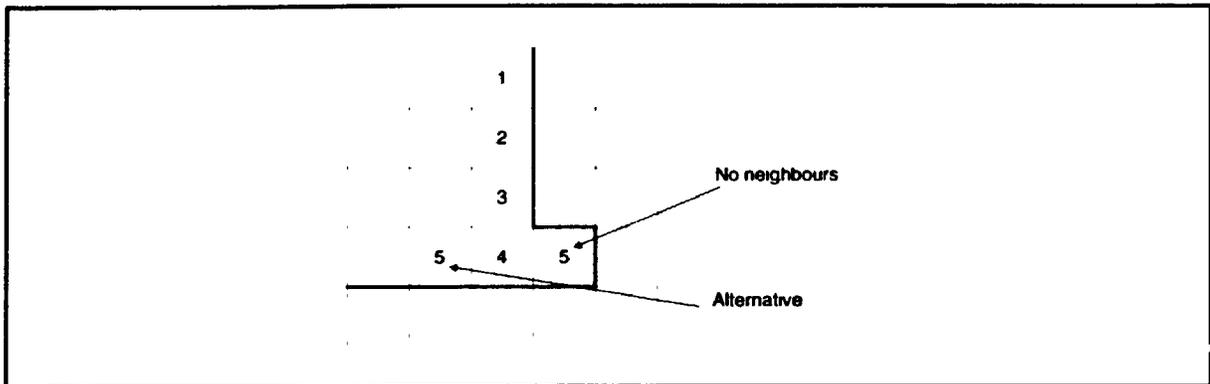


Figure A.3: Backtracking operation. If the last pixel on the edge does not have any neighbours that satisfy the edge condition, the algorithm backtracks recursively, to find an alternative solution for the previous point in the edge list.

When the first point satisfying the edge condition is found, it is pushed on the last position of the scratch list, which contains 10 points. This buffer is used to retreat from dead ends and to prevent the occurrence of small loops in the distorted edge (Fig A.3). Newly found edge pixels are compared with this list; if they are not on it, they are added to the main edge list.

When the distance between the last point on the edge and the first one is smaller than 2 pixels and the edge is already longer than the assumed minimum, the edge is regarded as closed. If the algorithm is not able to close it, an error condition is returned and an interactive edge definition is required. Finally, the list filled with consecutive edge-points is returned. The list of points is fully chained; the predecessor on the list is always a neighbour (2). In some cases, the flood image can be positioned partially on the border of the image matrix and its contour can be different from the geometry of the gamma-camera field-of-view. Therefore, no assumptions are made about the shape of the flood during the edge-finding procedure. A flow chart of the edge-finding algorithm is shown in Figure A.4.

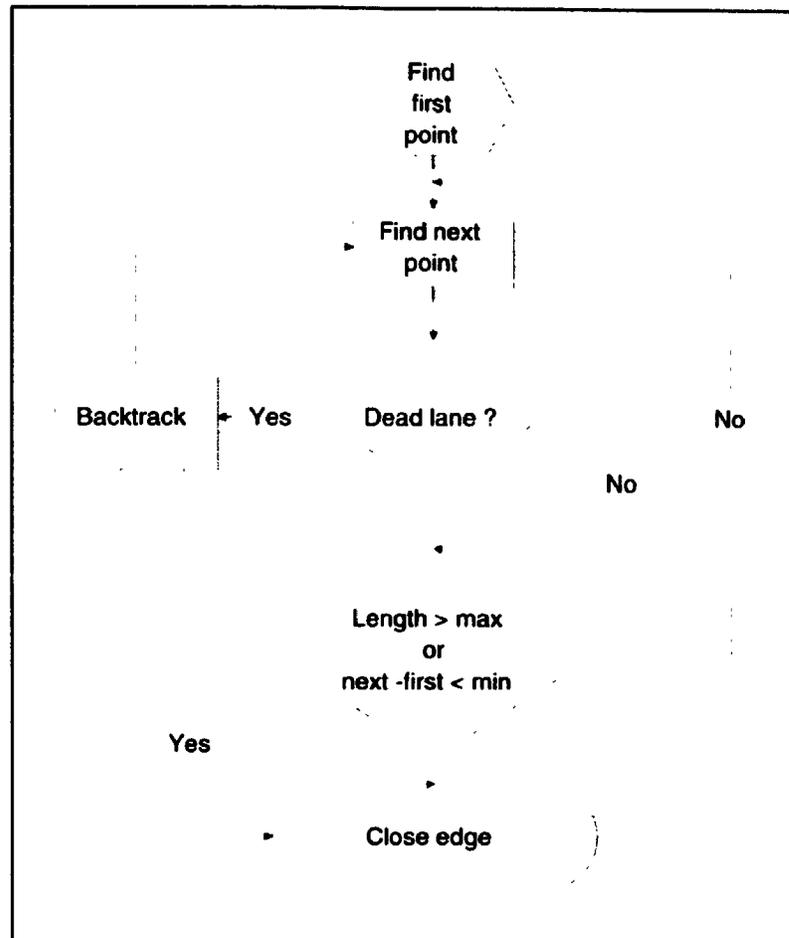


Figure A.4: Flowchart of edge-finding algorithm.

A.2 SHAPE-FITTING ALGORITHMS

A.2.1 Circle-fitting

The circle-fitting algorithm works as follows. The maximum distance between any two points on the edge is found. Next, six equally spaced points from the edge list are chosen, forming an irregular hexagon. Consequently, six hexagon diameters are calculated and six vertices of such a hexagon are determined. The circle is then defined by these vertices. The radius of the circle is assumed to be equal to the maximum diameter of the hexagon. The centre of the circle is calculated as the geometric centre of mass of the hexagon vertices.

A.2.2 Polygon-fitting

The polygon-fitting algorithm is used to fit hexagonal, rectangular and bevelled-rectangular shapes. Two interpretations of the best fitted polygon are used. The first interpretation defines it as a polygon with vertices not necessarily on the edge, but with a preserved geometric shape (rectangle, or ideal hexagon). The other interpretation of the fitted shape is a polygon with all the vertices positioned exactly on the detected edge. In the first case the vertices have to be real numbers; in the second case, the integer coordinates or positions in the edge list are sufficient. The algorithm initially finds the approximate position of the polygon vertices (Fig A.5), and then adjusts it by translation and rotation, to minimize the difference between the actual point list and the shape.

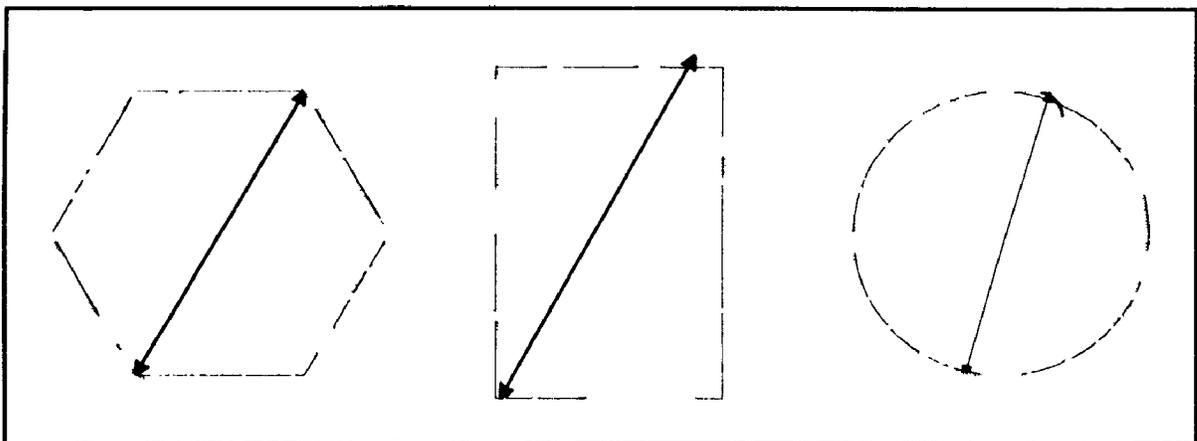


Figure A.5: First approximation of fitted shapes for polygons and circles. Two most distant points from the edge are assumed to be the two opposite vertices and the remaining vertices are found based on the geometric specifications, retrieved from the camera database.

The first approximation of a figure does not fit the contour exactly because it is assumed that the maximum distance points are the correct positions of the vertices. Subsequently, the fitting algorithm finds the parameters of the lines between vertices that define the polygon. The algorithm calculates the polygon sides using the vertices found by the first approximation. These sides correspond to the first fit and have to be corrected. The points from the edge list that are closest to the ideal vertices, are regarded as the ends of the corresponding sides. Next, all the points between two vertices, except for a few points

($n=5$) at the end of each side, are passed to a line-fitting routine. The points near the end are not reliable because they may belong to the adjacent side of the polygon, and therefore could corrupt the calculations.

The line-fitting algorithm uses least squares regression. All line coefficients are represented in polar coordinates; thus each line is represented by its distance ρ from the point (0,0) and by the angle θ between the X axis and the line itself (-90,90). In this manner, exception conditions for vertical and horizontal lines are avoided. If $ax+b$ representation exists, then the parametric representation is given by

$$\begin{aligned}\theta &= \arctan(a) - 90^\circ \\ \rho &= -b/\sqrt{a^2+1}\end{aligned}\quad [A.2]$$

The quality of the line fit found by the regression method is characterised by the correlation coefficient. Subsequently, the angle differences between the first approximation of lines and the lines resulting from the regression analysis are found, and one mean angle of rotation for all sides is calculated. Only the sides with correlation coefficients greater than a certain threshold (value experimentally established: 0.95) were used for calculating the rotation angle. A smaller correlation coefficient indicates that the edge is distorted, therefore such a side cannot properly determine the angle. Therefore, the mean angle of rotation α_{mean} is given by

$$\alpha_{mean} = \sum \frac{\alpha_i}{N} \quad [A.3]$$

where α_i are the angles between the first approximation of lines and the lines resulting from the regression analysis for each side of the polygon, and N is the number of sides used.

The polygon is then rotated by the obtained angle. The translations in X and Y directions (T_x , T_y) are calculated by comparing the new polygon sides after the rotation to the regression results. Again, only the sides with high correlation coefficients ($r > 0.95$) are used. After calculating the mean shifts T_{xmean} and T_{ymean} , all the polygon vertices are shifted in the X and Y directions accordingly. Eventually, the new side coefficients are calculated. Translation values are calculated according to the following formula

$$\begin{aligned} T_x &= \sum \frac{\rho_{1_{xi}} - \rho_{2_{xi}}}{N} \\ T_y &= \sum \frac{\rho_{1_{yi}} - \rho_{2_{yi}}}{N} \end{aligned} \quad [A.4]$$

where $\rho_{1_{xi}}$, $\rho_{1_{yi}}$ are distances from the point (0,0) in X, Y direction of the first approximated line and $\rho_{2_{xi}}$, $\rho_{2_{yi}}$ are similar distances for the regression-fitted line.

In the case of the bevelled rectangle field-of-view (Fig. A.6), its coordinates are calculated after fitting and adjusting the full rectangle. The bevelled-rectangle is defined by 8 coordinates. Since short sides, consisting of approximately 8-10 points, cannot be properly used by regression, the fitting uses initially only four sides (the same as for the ordinary rectangle). The corners are bevelled at a later stage, according to the manufacturer's data.

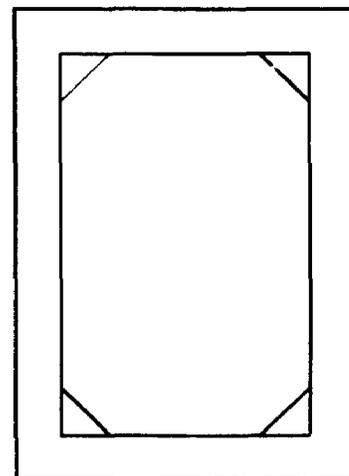


Figure A.6: Bevelled-rectangular field-of-view.

The rotational and translational parameters described above provide the estimation of the fit between the actual flood image and the predefined shape of the flood (symmetric polygon). Finally, the algorithm estimates a polygon, which is the closest-possible to the one fitted, but has all the vertices on the edge (integer coordinates). Thus, there are two fitted shapes calculated: *the ideal polygon*, and *the closest-to-ideal polygon*.

A.3 NEMA CALCULATIONS

Useful and central field-of-view areas are derived by minifying the obtained flood shapes. Each vertex of the polygon has its opposite counterpart. This counterpart is defined as the vertex, whose position in the list of vertices is half of the list length distant from the corresponding coupled pixel. The vertices of the minified polygon are shifted along the line, joining the two opposite pixels, towards the centre. The useful and central field-of-view types are defined as minified polygon shapes. A useful and central field-of-view are defined as minified polygons, with diagonals equal to 95% and 75%, respectively, of the original size. The circular field-of-view was obtained by the minification of the radius. Sometimes, when the flood source of hexagonal shape is not regular on the corners, the useful field-of-view NEMA uniformity coefficients are very high. For this reason, the useful field-of-view for hexagons was defined as 90% of the diagonal, since it is not yet properly standardized for polygons by NEMA. Field-of-view definitions are illustrated in Figure A.7.

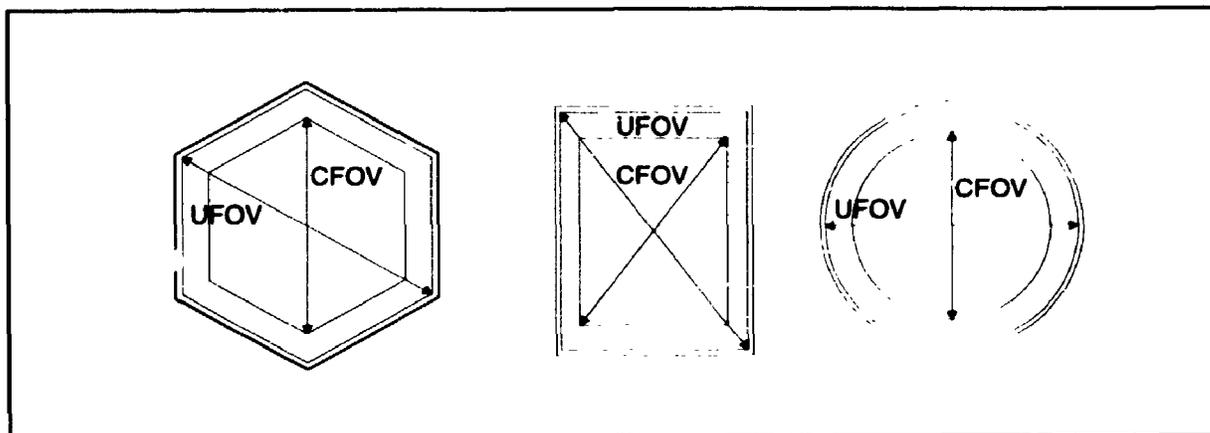


Figure A.7: Useful (UFOV) and central (CFOV) field-of-view definitions.

To perform NEMA calculations, the image has to be filtered by the spatial 3x3 smoothing filter (3). The filtering is applied to the raw data rather than to an image clipped to the particular field-of-view. All the subsequent NEMA calculations are done on the smoothed data.

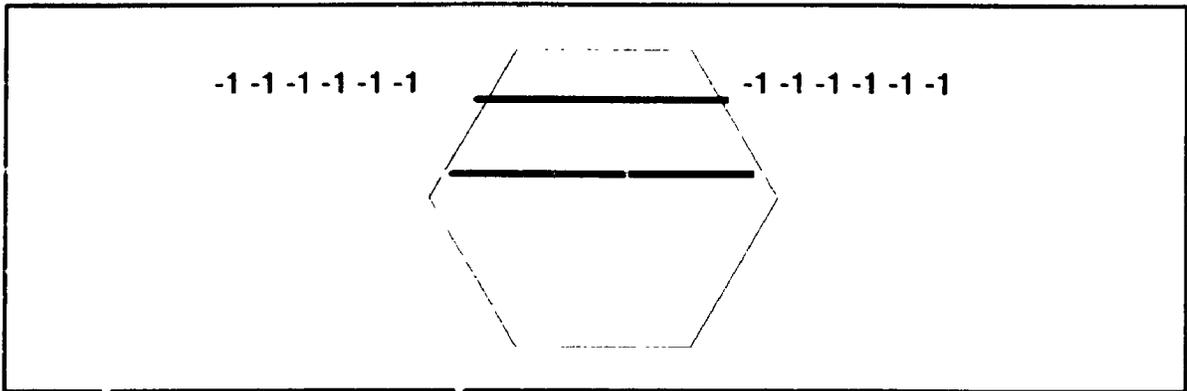


Figure A.8: Polygon contour filling technique. Each horizontal line is scanned for the first and last point of the shape. These points are calculated from the intersection between the current horizontal line and sides determined by pairs of vertices. The left side determines the first point and the right side gives the last point of the shape in the current scan line. All the values that are positioned outside of the polygon or circle are marked by using a negative pixel value. The negative value is used in order to distinguish between 0, which could be the image data in the field-of-view (for example, in the case of a photomultiplier fault) and values regarded as outside of field-of-view.

After defining the field-of-view, its parameters (new radius, or set of vertices) are used to clip the image data to the minified dimensions. In the case of a circle, all the pixels whose distances from the centre are greater than the radius, are masked out. For polygons, a standard contour-filling routine, illustrated in Fig A.8, was used.

The NEMA calculations are performed on clipped images. Points with negative values are ignored in all calculations; this scheme allows implementation of NEMA algorithm for any conceivable shape. Three NEMA coefficients are calculated: NEMA integral non-uniformity, NEMA differential x nonuniformity, and NEMA differential y nonuniformity, according to the NEMA definitions (3). Finally, the algorithm checks whether the calculated NEMA values fall within the specifications and sets the appropriate logical flags. The camera specifications are provided in a separate file and retrieved for comparison.

A.4 COLD AND HOT SPOT QUANTIFICATION

To determine the extent, mean count, and location of the cold and hot spots in the images a region-growing technique can be used (4). The same region-growing algorithm is used for both cold and hot spots. The algorithm is implemented in the following way. The starting point (seed) is the location of the averaged (3x3 region) minimum, or maximum. The 8-neighbourhood of the seed pixel is examined and, if there are new points satisfying the spot condition, they are added to the spot area. It is assumed that points with values that differ more than 2 standard deviations from the average count in the flood image, belong to the spots. Next, the algorithm recursively checks the 8-neighbourhood of each new spot pixel. The pixels that satisfy the spot condition, are added to the set of spot pixels and, subsequently, their neighbourhoods are examined. When no new neighbours have been found, the algorithm stops. Finally, the size, mean count and geometrical centroids of such regions are calculated.

A.5 EDGE DISTORTION

The real length of the contour (RCL) is defined by

$$RCL = \sum_{c_j \in \text{contour}} \|c_{j+1} - c_j\| \quad [\text{A.5}]$$

where c_j are the contour points and double bars represent the vector length. Such a definition calculates the actual length of the contour, which is different than the number of points included.

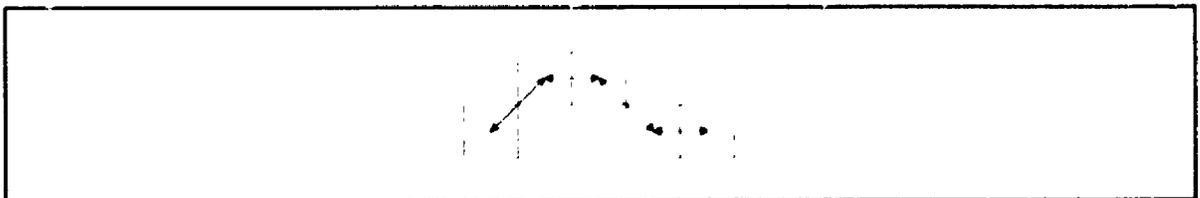


Figure A.9: Length of the contour. When a subsequent pixel is not positioned on the same horizontal or vertical line the length is increased by $\sqrt{2}$ rather than 1.

A.5.1 Length-based distortion - circle

In the case of the circle, the distortion refers to the whole contour. The real length (RCL) is compared to the ideal length (ICL), which is derived from the circle's diameter. Length based distortion coefficient (CDL_{circ}) for a circle, is defined as

$$CDL_{circ} = 100 * \frac{RCL - ICL}{ICL} \quad [A.6]$$

A.5.2 Length-based distortion - polygon

In the case of polygons, in addition to the total distortion, the distortion of each edge is individually assessed. The sides are defined in the edge-fitting algorithm (section A.1). The real length of each side (RCL_i) is calculated according to Eq. A.5 and then the distance between the two opposite points of the side is calculated.

The length-based distortion coefficient for each side CDL_i , is defined as

$$CDL_i = 100 * \frac{RCL_i - \|p_i - p_{i+1}\|}{\|p_i - p_{i+1}\|} \quad [A.7]$$

where i is the side number, and p_i is the first point of the i -th side.

The length-based, total coefficient of distortion for a polygon is calculated using the partial side distortion coefficients according to the following formula

$$CDL_{poly} = 100 * \frac{RCL - \sum_{i=1}^N \|p_i - p_{i+1}\|}{\sum_{i=1}^N \|p_i - p_{i+1}\|} \quad [A.8]$$

A.5.3 Maximum-distance based distortion - circle

The maximum-distance method for circular shapes uses the maximum distance between the actual contour and the ideal geometric shape. This distance ($dmax$) is defined by

$$dmax = \max_i (\|c_i - cO\|) - r \quad [A.9]$$

where c_i is the i th point on the contour, cO is the centre of the circle, r is its radius, and double bars denote the length of the vector. Subsequently, the distortion coefficient maximum-distance based for circles (CDM_{circ}) is defined as

$$CDM_{circ} = 100 * \frac{dmax}{r} \quad [A.10]$$

A.5.4 Maximum-distance based distortion - polygon

The maximum-distance method for polygons uses the maximum distance between the contour and each of the fitted sides. This method can be applied to a single side as well as to the whole contour. The maximum distance ($dmax_n$) is defined for polygons by

$$dmax_n = \max(\text{dist}(p, l_n)) \quad [A.11]$$

where p is the point on the edge, l is the line defining the side of the polygon, i is the edge point index (all points from the given side), and n is the side index.

The distortion coefficient based on the maximum-distance for each side is defined as

$$CDM_n = 100 * \frac{dmax_n}{l_n} \quad [A.12]$$

where l_n is the length of the n -th side of the polygon.

Subsequently, the total distortion coefficient maximum-distance based (CDM_{poly}) for polygons is defined as

$$CDM_{poly} = 100 * \frac{\sum_{n=1}^N dmax_n}{RCL} \quad [A.13]$$

A.5.5 Area-based distortion - circle

The measure of the distortion is defined as the area between the contour line and the ideal shape of the circle. The approximate estimate of such an area is the sum of the distances between each point of the ideal circle, assuming a pixel size of 1. Thus, the measure of the distortion area (DA) is given by

$$DA = \sum_{i=1}^M \|c_i - cO\| - r \quad [A.14]$$

where c_i is the i th point on the contour cO is the centre of the circle, r is its radius, M number of points on the contour and double bars denote the length of the vector.

Consequently, the area-based coefficient of distortion is defined for circles as

$$CDA_{circ} = 100 * \frac{DA}{A_{fov}} \quad [A.15]$$

where A_{fov} is the area of the field-of-view.

A.5.6 Area-based distortion - polygon

For polygons, the distorted area is examined on each side separately. The measure of the distortion is the area between the contour and the straight side line. This area is given by the integral of the absolute value of the contour curve in the direction of the side. The estimate of this area is the sum of the distances between each point and the closest side. Thus, the distortion area for side n (DA_n) is defined by

$$DA_n = \sum_{i=1}^M \text{dist}(p_i, l_n) \quad \text{[A.16]}$$

where p_i are points from the side n , l_n is the line determining n -th side of the polygon, and M is the number of points on one side. To avoid the inclusion of the spurious area due to the pixel aliasing effect, the distance between the contour point and the line is added to the area only when it is bigger than 1 (pixel size). It should be noted that this definition is not exactly equivalent to the area, but is an estimate of such area, which is feasible to compute.

The area-based distortion coefficient for side n is subsequently defined as

$$CDA_n = 100 * \frac{RCL * DA_n}{l_n * A_{fov}} \quad \text{[A.17]}$$

where A_{fov} is the area of the whole field-of-view. Multiplying this coefficient by the contour-side length ratio is necessary, because the distortion area depends on the side length, and, in case of the rectangular shape, the sides are not equal.

The total area-based distortion coefficient is defined as the average of the coefficients for each side and is given by

$$CDA_{poly} = \frac{\sum_{i=1}^N DA_n}{N} \quad [A.18]$$

where N is the number of sides.

A.6 ASYMMETRY

A.6.1 Circle asymmetry

Asymmetry is defined for the circle, using the dimension features. As stated in section A.2, the features of the circle include equally spaced diameters. The asymmetry (AS_{circ}) is calculated as follows

$$AS_{circ} = \sum_i \left| r - \frac{d_i}{2} \right| \quad [A.19]$$

where r is the radius of the ideal circle and d_i is the i -th diameter of the fitted circle (Fig. A.10).

The asymmetry coefficient for circle (CAS_{circ}) is given by the equation:

$$CAS_{circ} = 100 * nf * \frac{AS_{circ}}{r * N} \quad [A.20]$$

where nf is the normalization factor, r is the ideal circle radius and N is the number of diameter samples.

A.6.2 Polygon asymmetry

In the case of the polygon, the ideal shape returned by the edge-fitting algorithm adjusts only the zoom (one scaling parameter) and the position of a previously defined shape. When the condition that vertices must be positioned on the edge is imposed, all the assumptions about the shape are lost; to describe the deviation from the ideal shape, the concept of asymmetry can be used (Fig. A.10)

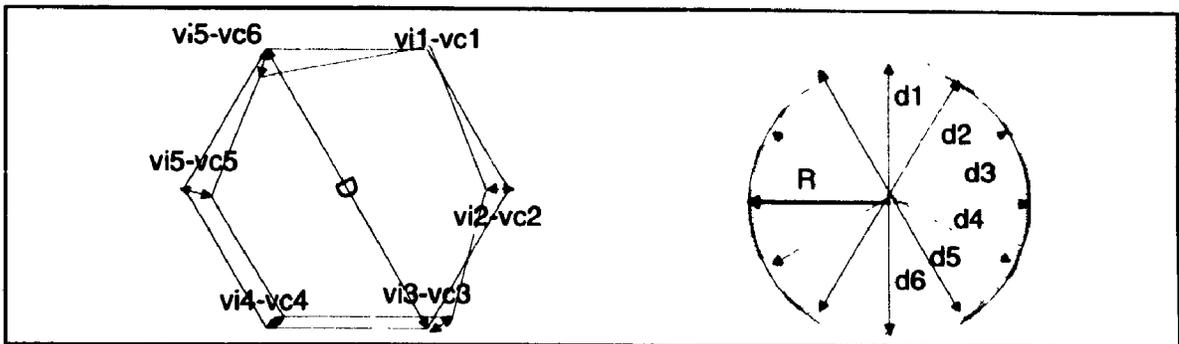


Figure A.10: Definition of asymmetry for polygons and circles. Asymmetry is defined as the sum of all the distances between the vertices of an *ideal* figure, and the *closest-to-ideal*, as defined in section A.2.

The measure of the polygon asymmetry (AS_{poly}) is given by

$$AS_{poly} = \sum_i \|v_i - vc_i\| \quad [A.21]$$

where v_i is i -th *ideal* polygon vertex, vc_i is i -th *closest-to-ideal* polygon vertex and double bars denote the norm (length) of the vector argument.

The asymmetry coefficient for polygons (CAS_{poly}) is defined as follows

$$CAS_{poly} = 100 * nf * \frac{AS_{poly}}{d_{max} * N} \quad [A.22]$$

where d_{max} is the maximum diagonal of the polygon, nf is the normalization factor, and N is the number of sides of the polygon. This factor is important only for the presentation of the data. For example, the value assumed in this program ($nf = 8$) results in CAS values above 10% for significant asymmetry effects.

REFERENCES:

1. Gonzales RC, Woods RE. *Digital image processing*. Reading MA: Addison-Wesley Publishing Company 1992; 538-540.
2. Gonzales RC, Woods RE. *Digital image processing*. Reading MA: Addison-Wesley Publishing Company 1992; 458-461.
3. Quality control of nuclear medicine instruments. IAEA TECDOC-602. Vienna: International Atomic Energy Agency; 1991.
4. Rosenfeld A, Kak AC. *Digital picture processing*. New York: Academic Press; 1976.