

**Best
Available
Copy**

AFIT/GCS/ENG/PTG 94J-02

AD-A280 617



THESIS

Presented to the Faculty of the School of Engineering
of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Computer Science

GLOBAL PICTURE ARCHIVING AND
COMMUNICATION SYSTEMS (GPACS):
AN OVERVIEW

Russell C. Ellis
First Lieutenant, USAF
AFIT/GCS/ENG/94J-02

DTIC
ELECTE
JUN 23 1994
S B D

April, 1994

Approved for public release; distribution unlimited

 94-19301



1768

94 6 23 129

ACKNOWLEDGMENTS

I would like to thank Lt Col Martin Stytz for his support throughout this thesis effort. He motivated me to continue even when things seemed unattainable.

I would also like to thank my wife, Laurie, for her kindness and understanding throughout my time at AFIT.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS.....	ii
LIST OF ILLUSTRATIONS.....	v
LIST OF TABLES.....	vii
ABSTRACT.....	viii
CHAPTER 1. GPACS "AN OVERVIEW".....	9
PART I. BACKGROUND AND STATEMENT OF THE PROBLEM.....	9
PART II. METHODOLOGY AND LITERATURE REVIEW.....	10
PART III. INTRODUCTION.....	11
PART IV. PACS BACKGROUND.....	13
CHAPTER 2. INTER-OPERABILITY ISSUES: MAKING A CONNECTION..	24
CHAPTER 3. NETWORK TECHNOLOGY AND PACS.....	29
PART I. NETWORK TOPOLOGIES.....	34
PART II. NETWORK CLASSIFICATION.....	42
CHAPTER 4. THE PACS DATABASE: THE BACKBONE OF A PACS.....	51
PART I. STORAGE HIERARCHY LEVELS.....	52
PART II. DATABASE ISSUES.....	56
PART III. RELATIONAL DATABASE.....	58
PART IV. OBJECT-ORIENTED DATABASE.....	63
PART V. THE DATABASE CONTROVERSY.....	68
PART VI. RDB ADVANTAGES.....	70
PART VII. RDB DISADVANTAGES.....	72
PART VIII. OODB ADVANTAGES.....	75
PART IX. OODB DISADVANTAGES.....	79

PART X. A HYBRID SOLUTION.....	81
PART XI. CURRENT APPLICATIONS OF OODBS.....	83
PART XII. MIDB.....	84
PART XIII. DISTRIBUTED & CENTRALIZED DATABASES.....	85
PART XIV. DATABASE SECURITY.....	93
PART XV. THE PACS FUTURE.....	94
CHAPTER 5. IMAGE AND DATA COMPRESSION IN THE PACS DOMAIN..	96
PART I. HUFFMAN CODING.....	108
PART II. LZW COMPRESSION.....	110
PART III. ARITHMETIC CODING.....	116
PART IV. PCM CODING.....	120
PART V. KLT TRANSFORM.....	125
PART VI. DCT TRANSFORM.....	127
PART VII. FRACTAL COMPRESSION.....	133
PART VIII. DWT COMPRESSION.....	137
PART IX. COMPRESSION SUMMARY.....	139
CHAPTER 6. CONCLUSION AND RECOMMENDATIONS.....	140
APPENDIX A: Definition Tables.....	147
BIBLIOGRAPHY.....	155

LIST OF ILLUSTRATIONS

Figure	Page
Figure 1. Medical Database Structure.....	17
Figure 2. The Composition of GPACS.....	20
Figure 3. PACS Relationships.....	20
Figure 4. Network Communication Types.....	33
Figure 5. Fully Connected Network.....	35
Figure 6. Partially Connected Network.....	36
Figure 7. Tree Structured Network.....	37
Figure 8. Star Network.....	38
Figure 9. Ring Network.....	39
Figure 10. Bus Network.....	40
Figure 11. Example of a Wide Area Network (WAN).....	44
Figure 12. ISO Seven Layered Reference Model.....	47
Figure 13. OSI Address Formats.....	50
Figure 14. Time In Storage Triangle.....	54
Figure 15. PACS Storage Hierarchy.....	55
Figure 16. RDB Basic Operations.....	59
Figure 17. Examples of Object Classes.....	67
Figure 18. Example Object Relationships.....	67
Figure 19. One-to-One and One-to-Many Relationships...	73
Figure 20. "Where We Are Today!".....	80
Figure 21. Homogeneous Distributed DBMS Architecture..	88
Figure 22. Address Book Dependence.....	89
Figure 23. Tightly-Coupled Multi Database System.....	92

Figure 24. Loosly-Couple Multi-Database System.....	92
Figure 25. Textual Substitution.....	100
Figure 26. Color-Symbol Substitution.....	106
Figure 27. Example of Huffman Coding.....	109
Figure 28. Example of LZW Coding.....	112
Figure 29. Analog Signal being Digitally Sampled.....	121
Figure 30. Reference Table for PCM.....	121
Figure 31. DCT Encoding.....	128
Figure 32. DCT Decoding.....	128
Figure 33. Zig Zag Sequencing.....	131
Figure 34. Fractal Compression Algorithm.....	134
Figure 35. Fractal Decompression Algorithm.....	135

LIST OF TABLES

Table	Page
Table 1. Inter-Operability Definitions.....	148
Table 2. Network Definitions.....	149
Table 3. Topology comparison.....	40
Table 4. Database Definitions.....	150
Table 5. Three-Level Hierarchy.....	55
Table 6. Modern Three-Level Hierarchy (MDIS at Wright-Patterson Air Force Base, Ohio).....	55
Table 7. RDB Definitions.....	151
Table 8. Relational Database Table (Patient Table)....	60
Table 9. Relational Database Table (Radiology Patient Table).....	61
Table 10. Result of Restricting New Table to Bob Smith's Information Only.....	61
Table 11. Result of a Join on Name of Table 8 and Table 9.....	61
Table 12. OODB Definitions.....	152
Table 13(a). Compression Definitions.....	153
Table 13(b). Compression Definitions.....	154
Table 14(a). LZW Implementation.....	113
Table 14(b). LZW Implementation.....	114

ABSTRACT

The world is currently experiencing a computer revolution. With the expansion and development of new computer technology, the age of video is upon us. The medical world is also currently undergoing a dramatic change. With the development and perfection of non-invasive techniques of probing the body, large amounts of computer data is collected and much of it is used to generate images. The organization and management of these images is the topic of this thesis. The discussion of storage schemes, databases, communications and data compression techniques gives the reader a background and some insight to the current state of the art in Picture Archiving and Communication Systems (PACS) and how this schema will some day be expanded to a global scale called Global PACS (GPACS).

GLOBAL PICTURE ARCHIVING AND COMMUNICATION SYSTEMS

"AN OVERVIEW"

CHAPTER 1

PART I

BACKGROUND AND STATEMENT OF THE PROBLEM

PACS is a new technology that is emerging in the medical field. There is extensive literature on PACS and its place in a hospital system. Although there is an abundance of information available, there is still no one source that identifies the specific areas of a PACS and the current state of the art. Also the idea of moving to a GPACS has not specifically been addressed while relating to a local PACS. The lack of gathered information and insight is what inspired this thesis.

The goal of this thesis is to identify first what a PACS is and what it is used for in a real world setting. Then, while using the information about a PACS, break a PACS down into its sub-parts and briefly describe the different parts. Several of the sections are expanded upon because of the significant role they play in creating and sustaining a PACS. The whole discussion of a PACS is overtone with references to a GPACS being the next step and how we can use what we have today to get to GPACS realization.

PART II METHODOLOGY AND LITERATURE REVIEW

The methodology and literature review are combined into one section because of the nature of this particular thesis. This thesis, rather than following a more traditional engineering thesis format, is a review of PACS. The review of PACS is accomplished by reading over two hundred articles and books to collect information about the current state of the art of PACS and where the development is leading. The articles and books covered all the areas associated with PACS even though some of the topics were not a concern when developing a GPACS. Once the readings had been accomplished, the author used the knowledge gain to identify the different areas of PACS. From these areas, several areas were expanded upon based on the applicability of the area when PACS moves to the next level of GPACS.

PART III INTRODUCTION

A Global Picture Archiving and Communication System (GPACS) is a system that will have the ability to get any type of medical information, including text, sound, images and video, from anywhere in the world, at any time. For the purpose of this thesis, GPACS refers to an all encompassing global computer network system while a Picture Archiving and Communication System (PACS) refers to local hospital systems involving medical image management. PACS is also the most common name for a medical image management system, but several other names such as Image Management And Communications System (IMACS) and Medical Diagnostic Imaging Support (MDIS) are also used.

There are many issues involved in developing a GPACS with several of the basic building blocks currently available. The critical building block, a hospital-wide PACS, has yet to be standardized; therefore many different types of PACS architectures exist. Almost every hospital that deals with medical images has its own unique PACS, with its own collection of equipment. This diversity leads to several of the problems that stand in the way of attaining a GPACS. Some further issues to be resolved when discussing PACS and GPACS include: What type of database should be used to handle multimedia information? Should the system have a centralized database or a distributed database? What kind of communication network should be used? Who will set the

standards and what will the standards be for all the issues in GPACS including compression, transmission, file formats, database formats, software, hardware and security issues?

This thesis presents the current status of work and developments occurring in this field along with some insight into where GPACS development is heading. The first section discusses the inter-operability problems involved not only in a GPACS, but also a PACS. The next section gives a background in the area of networks, describing basic structures and topologies available and in use today. The last two sections deal with types of databases and the types of compression that are available to a PACS. These two areas, databases and compression, are singled out because of the vital impact they have on all aspects of a PACS and GPACS. Although the inter-operability and network issues mentioned in the first two sections are just as important to making a PACS work, these topics are generally restricted to the global community, while the database and compression algorithms involve decisions made on a individual PACS level. A summary of each topic occurs at the end of each individual section with thoughts about where the trends and technology are leading.

PART IV PACS BACKGROUND

An explanation of all the computer systems associated with a hospital is necessary before discussing a PACS. Most hospitals have two main information systems, the Hospital Information System (HIS) and the Radiology Information System (RIS). Smaller systems such as a pharmacy inventory computer and very specific computer products also exist. The HIS is responsible for keeping track of basic patient information, generating bills, paying employees, ordering supplies, reducing excess paperwork in the hospital. The RIS, on the other hand, directly relates to the radiology department and includes management of the film library, scheduling, and storing radiologists' interpretations. The HIS and RIS are generally disjointed and have little, if any communication, allowing redundancy to exist throughout these hospital information systems. Although these systems manage the huge amounts of data necessary to run a hospital efficiently, they lack the ability to acquire, store, retrieve, and view medical images. These operations are currently performed manually.

Traditional medical image management operates by putting the images onto film. The film is then stored in a film library, a large room in the hospital containing folders with patients' X-ray film. This room, depending on the size of the room and the hospital's patient load, will usually accommodate up to two years of images. Once the two

years has elapsed, the hospital staff moves the images to a warehouse or some other storage facility to make room for new images. The images are usually maintained in the warehouse from five to fifteen years depending on state law, hospital policies, and insurance coverage. The problems associated with a film library include lost film, damaged film, slow film retrieval, limited storage space available within the hospital environment, and the overhead cost associated with film chemicals and paper. Protecting the only piece of film in existence for each X-ray or scan remains one of the biggest problems associated with film libraries. The manpower required to organize, retrieve, maintain and track the film is also a problem. Hand delivering of film can delay up to 25 percent of diagnoses because the film is being used elsewhere in the hospital (Busse,1993:2). If the situation is not an emergency, a slow retrieval time can waste the physician's time while waiting for a particular image. If the film is lost or damaged, the physician may prescribe more X-rays. With the current structure of film libraries physicians have difficulty getting a patient's image in a quick, efficient, manner, if at all. Although they were an asset in the past, film libraries are quickly becoming a burden to efficient hospital management.

The solution to the film library problem is a PACS. A PACS' main purpose is to keep track of images in a digital

form. The generic PACS is designed to solve or decrease the afore mentioned problems associated with a film library. One of the main justifications for a PACS, especially today, is the potential cost savings that are associated with a PACS, however, information on whether or not PACS actually saves a hospital money is severely lacking. A cost modeling of PACS software, called CAPACITY, became available (van Gennip et al. ,1992:266) in July 1989. Typically this software shows short term high cost, with large cost savings in the future when installing a PACS. A United States Air Force study determined that the military's \$350 million investment in several PACS (collectively under the title MDIS) would pay off in 26 to 32 months (Taft ,1993:36). Other hospital automation, such as automated billing and computerized scheduling, has already saved some hospitals up to 13% per patient (Busse,1993:1). The unsubstantiated cost savings of a PACS has caused some people to argue the benefits of a filmless hospital. This argument closely resembles the paperless office issue in the business world. The three basic views are: (1) adopt no changes and keep doing only film management, (2) have a PACS with film as a back-up, (3) and have a complete PACS by slowly phasing out the old film managing system. Proponents of each view believe their way is the best way to acquire, store, and manage the medical images in their hospital. Despite these differing views, a change towards digital image manipulation

is occurring in hospitals because of today's high cost of health care.

Two main architectures, Inter-Hospital and Intra-Hospital, encompass the whole database arena for a PACS system. (A stand alone database would not be considered part of a PACS until it is connected in the system.) Figure 1 illustrates the structure available to an individual hospital's database configuration. If a hospital chooses a centralized database, the database can be either relational or object-oriented. On the other hand, if the hospital chooses a distributed database, any number and combination of relational and object-oriented databases can be constructed. The same holds true for a GPACS or inter-hospital network, where several local PACS work together. The GPACS can use either a central repository to store all of the medical images or a distributed architecture allowing the storing and sharing of images by individual PACS. This issue has not yet been seriously discussed because the concept of a GPACS is still in its infantile stage.

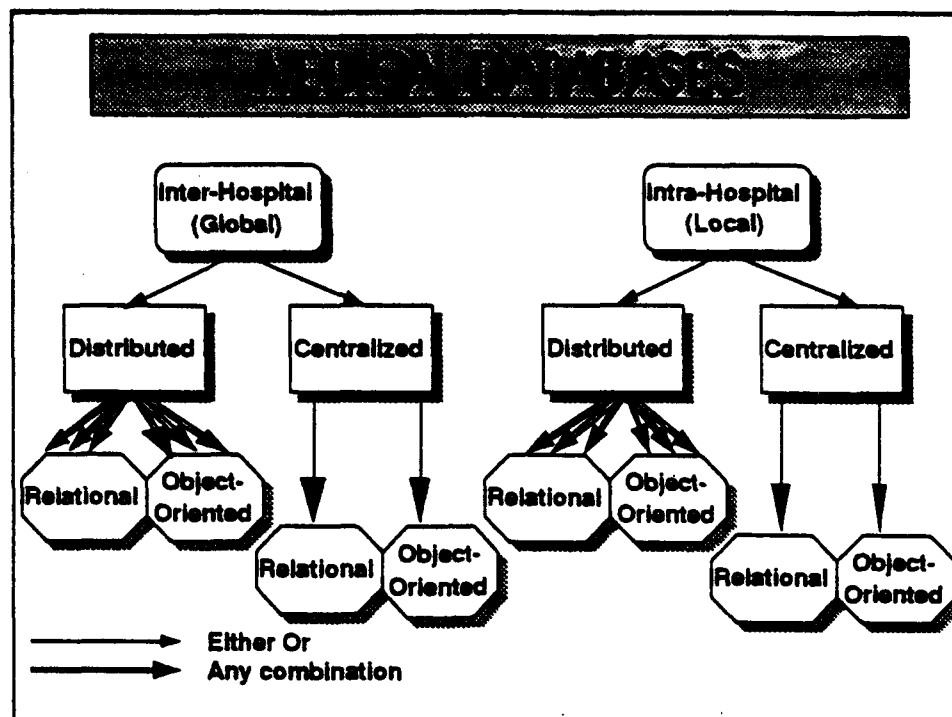


Figure 1. Medical Database Structure

Regardless of the structure of a PACS, getting the images and information to the doctor when the doctor wants them, is the ultimate goal. The following typical process for a first encounter routine exam better illustrates the situation that exists in a hospital that has an HIS, RIS and PACS (separate).

1. Patient calls and makes an appointment.
2. Patient's personal data is entered into the computer.
3. Patient's name is placed in a particular time slot.
4. Patient shows up at the prescribed time and personal information is retrieved.
5. Patient's records are brought up on the computer and sent to the doctor's office.
6. Doctor conducts the exam and inputs the results into

the computer.

The above 6 steps can be accomplished using an HIS. If the physician orders a Magnetic Resonance Imaging scan, the following steps occur:

7. Patient makes another appointment (with the physician's office) to have an MRI scan taken.
8. Patient shows up at the radiology department where they retrieve his records on the computer. Contained within the patient's record are the instructions for the MRI scan.
9. Scan is taken and another appointment is made to meet with the physician.
10. Physician is notified (by e-mail or phone) that the patient has completed the MRI scan.
11. Physician opens the patient's record and views the image. When reviewing the image, manipulation tools are used to rotate, zoom, enhance or whatever else the physician needs to do to make a diagnosis (including comparing the image to other images).

Steps 7-11 are usually carried out independently by the RIS. When the physician requires a second opinion by another physician located 500 miles away he can do the following.

12. Primary physician sends a message to the secondary physician, giving him the patient's name and social

security number (or another identifying characteristic).

15. Secondary physician requests the information at his computer terminal where the information is immediately available.

16. Having established a link, the primary physician communicates his concern to the secondary physician by voice and by using a pointer on both screens.

17. A diagnosis is made.

This is a scenario of the basic functions that are required of a PACS and GPACS. A visual representation of the structure is presented in Figures 2 and 3. GPACS is all encompassing, including different hospitals with their own PACS, which in-turn are comprised of all the systems in the hospital. This has the potential to be a very large system involving many unresolved issues. In order to develop such a system, standardized communication between systems or even standardization across all the subsystems is needed. These issues are discussed in the inter-operability section of the thesis.

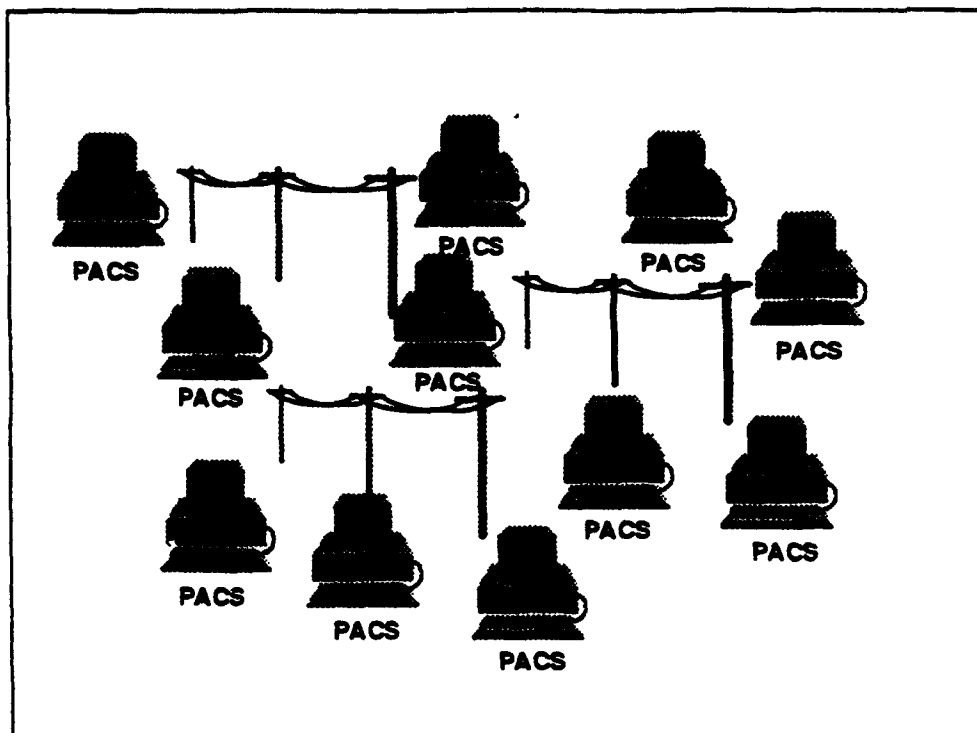


Figure 2. The Composition of GPACS

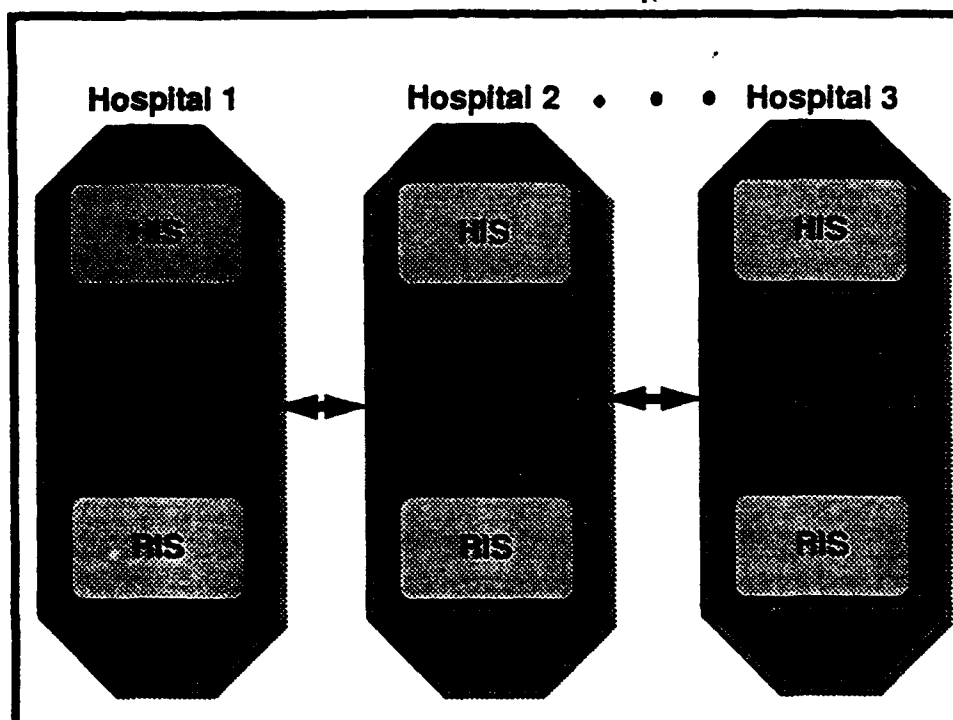


Figure 3. PACS Relationships

As demonstrated above, a PACS can and should be integrated with the other hospital systems such as the HIS

and RIS. Currently, though, it is a separate entity in most hospitals because of the integration problems that exist. Eventually these systems should be connected so they appear to the hospital staff as one complete system. The integration of the PACS, HIS, and RIS is an active area of research. This thesis examines some of the issues involved in accomplishing this integration.

The above information concerns existing systems but what about a hospital looking to attain a PACS capability? The hospital basically has two options, a full PACS implementation or a partial PACS implementation. The full implementation refers to obtaining a PACS that is used throughout the hospital and is installed all at once. This does NOT include an RIS and HIS as these are usually already present. Every department receives the same equipment, with the same software. An all encompassing PACS for a single hospital runs from \$7 to \$9 million dollars (Busse, 1993:2, Quillin, 1992 and Goeringer, 1992:6).

The partial implementation, for most hospitals, is the only way to practically obtain a PACS. The radiology department is usually the department in charge of the digital image acquisition devices, such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), etc. (Stytz, 1990)¹. Once they

¹The discussion of the different modalities associated with a PACS is beyond the scope of this paper.

have acquired a machine that converts or acquires images in a digital format, the management of those images can begin. Now that a source of digital images is available, each individual department begins to acquire hardware and software to store and manipulate the images they currently use. The difficulty with a partial implementation is that if a master plan is not developed in advance, each department within the hospital can end up with different, potentially incompatible, types of hardware and software based on their needs and cash flow. If this occurs, several new problems develop in the areas of communication, sharing images, and the politics involved in trying to determine a standard within the hospital. The easy solution is to convert other departments data formats into the radiology department's format. Partial PACS creation, correctly done over time, will result in a full PACS.

In any given PACS, there is the problem of communication between the different types of computers, each developed by a different vendor. The ideal situation would be a single vendor who could design all PACS and supply all the equipment associated with it, including the modalities (imaging machines), the main database computer, the different storage facilities, and the display systems. This situation is not currently available due to the size of a local PACS and the number of hospitals requiring such a system. A single vendor could help solve many problems such

as standardization of protocols, file formats, databases, processing packages (image display software), reliability, and communications. Even with a single vendor, the interoperability issues associated with a GPACS would still exist. In the GPACS scenario, standardization issues surface again, along with a new set of problems associated with long-haul communications, including protocols, file formats, processing packages, bandwidth restrictions, database, and security issues. The world is still far from attaining a fully operational GPACS because so many issues remain to be resolved. The only practical course of action for this thesis is to address some of the steps leading to the development of a GPACS, concentrating on a PACS and demonstrating the problems on a smaller, easier to understand, scale. The thesis then relates the issues of the PACS to a GPACS. The conclusion sums up the contents of the thesis and offers "big picture view" of what needs to happen before a GPACS becomes a reality.

CHAPTER 2

INTER-OPERABILITY ISSUES: MAKING A CONNECTION

While the background section established the need and basic functions of a PACS, this section concentrates on the components comprising a PACS and their interaction issues. Terms related to inter-operability of PACS are found in Table 1 in Appendix A.

Several pieces of hardware are required to consider a computer system a PACS. The first part of the PACS hardware needed is an image acquiring source. Generally these include different modalities as described by in (Stytz,1990) and there are also other sources of images such as Computed Radiography(CR) and film digitizers. CR consists of the exam terminal, the phosphor plates, the plate reader, and the image processing unit. The CR reusable phosphor plate cassettes replace conventional film cassettes. Once an X-ray is taken, the CR cassette is placed in the CR cassette reader and the information is stored in a digital format. Film digitizers take previously developed film and produce a digital image based on the actual film. More information on CR and digitizers can be found in (Donnelly,1991) with assessments and experiences found in (Hillman and Fajardo,1989) and in (Dietrich et al.,1989).

Once the digital versions of the desired images are acquired, they need to be stored. Typical types of storage include optical disks or tape (for slow long term storage), magnetic disk drives (for semi-fast medium/short storage) and workstation resident memory or Random Access Memory (RAM) (for fast short term storage). The general rule for storage facilities is the faster the memory the more expensive it will be.

The viewing terminals or workstations are an extremely important part of the PACS. Great emphasis and funds should be placed on the workstation because this is the place where the doctor sits down, makes a request and then views the results of his request. The doctor cares that his time is not wasted, the data must be retrieved in a quick and efficient manner, and the quality of the referenced image must be high enough to display on the monitor.

The last piece of hardware to be discussed is the main database server, the machine that holds the database and controls all the traffic throughout the PACS. This is where most of the magnetic disk drive storage resides. In a centralized PACS, this computer has the majority of the control. All of these issues are important in deciding the type of PACS that a hospital is going to develop.

Another important consideration is the environment for this hardware. Some of the requirements can be found in

(Gelish,1991) and are organized in the following three areas:

1. **Physical Environment** : This includes Heating Ventilation and Air Conditioning, Potential hazards, Water Suppression Alternatives, Hazardous Building Materials, Raised Floors, Fire Protection/Suppression, Structural Penetrations, and "Turn Key" Installation.
2. **Ergonomics** : This includes Design and Lighting.
3. **Human Skills** : This contains areas such as User Types, Instructional Systems Development, and Computer Aided Instruction.

The next issue is the software to be used on the hardware. Depending on the manufacturer, the software can either come with the hardware as a "turnkey" system or it might have to be purchased separately. For example, the software cost to run a large database can range from a few thousand dollars to tens of thousands of dollars based on the number of users.

In order for all of the pieces contained within a PACS to communicate with each other, some sort of link has to be established. Two of the popular media are Ethernet cable because it is economical and fiber-optic cable because of its greater bandwidth, reliability, and speed. Several distributed networks are taking advantage of extra bandwidth found in fiber-optic cables. For example, the features that Novel claim for their future distributed heterogeneous

computing systems include items such as supporting flexible network topologies (via virtual mapping), several high-speed communication channels co-existing between distributed processing nodes, low bit-error rates resulting from pure optical networks, fast network routing, and direct network support for application dependent computations and communications (Vetter and Du,1993:17). The benefits of these high-performance communications include not only high-performance computing and support for distributed multimedia information systems, but also the quick, efficient and reliable transfer of medical images.

Long-haul communications will be the link to the future. Along with the increased capacity of optical fiber, other characteristics that are extremely important to efficient image data transfer include the protocol being used between the two transferring entities. Image data is currently transferred using general purpose bulk transfer protocols such as Versatile Message Transaction Protocol (VMTP), and File Transfer Protocol (FTP) (Turner and Peterson,1992:258). However, some have speculated that for different types of transfers, there should be different types of communication pipelines. This idea builds on the fact that data comes in two types, analog and digital, and can represent several different multimedia types of information (text, images, sound, video, etc.). These different types of data can be transmitted and compressed

differently. In the network section of this thesis, the differences between digital and analog data, signals and transmission are discussed. These differences might be better exploited by having separate lines for digital and analog signals, each tailoring itself to a specific type of data. From the compression perspective, discussed in more detail later, information can be compressed in different manners depending upon the type of data and the transmission medium. Restricting one pipeline to only one type of information and compression could increase efficiency.

Once the hardware, software and communication media are available, the issue of standardization becomes apparent not only for a GPACS but also for a local PACS. The problem with standardization is that there is no consensus on what the standards should be. In the United States, the American National Standards Institute (ANSI) is the representative to the International Standards Organization (ISO).

CHAPTER 3

NETWORK TECHNOLOGY AND PACS

The first step in discussing a PACS is to describe how information (images) is shared between its different components. Managing images using a PACS requires the use of communication between several different computers. The task of transmitting information from one location to another is accomplished using a network. Networks are generally geared toward certain applications based on the transmission speed required and the type of users. This section gives a brief description of the different types of network topologies and the transmission media upon what they are based. Refer to Table 2 in Appendix A for a list of terms in this field along with their definitions.

The three areas concerning communication from one computer to another are 1) Data 2) Signaling and 3) Transmissions. Data refers to how the data is stored, signal refers to how signals propagated from one computer to another (on the actual media such as wire, cable and electromagnetic waves), and transmission refers to the communication of data by propagation and the processing of signals (Stallings, 1989:148).

Communication is an essential part of a computer network. Computers communicate in one of two forms, either analog or digital data. Analog data are represented by continuous values on a specified interval, and include voice and video data and data collected by sensors, such as pressure and temperature. Digital data takes on discrete values. Examples of digital data are text and integers.

Once data are obtained, the next step is to transmit the information to another location. This propagation of data can occur by either an analog or digital signal. An analog signal is a continuously-varying electromagnetic signal while the digital signal is a collection of discrete electromagnetic pulses. The analog and digital signals can be propagated on a variety of media including wire (twisted pair or coaxial cable), fiber optic cable, and radio waves (satellite communications). The main benefits of digital transmission over analog transmission are that it has a cheaper cost and it is less susceptible to outside interference commonly referred to as noise. The major disadvantage is that a digital signal suffers from attenuation of its signal that limits the distance of propagation before information is lost. Analog data can be translated into digital signals using a coder-decoder (codec), while digital data can be represented by an analog signal using a modulator-demodulator (modem).

Transmission is how the data and signal are used together. Analog transmission, is the process of transmitting analog signals without regard to their content (analog vs digital data). Analog transmission only deals with analog signals. The signals are transmitted using amplifiers to boost the signal over long distances. These amplifiers not only boost the signal, but they also boost the amount of noise associated with the signal. For analog data, the increased noise does not effect the data. In the digital data case, the increased noise introduces errors into the data.

Digital transmission can be used with either analog or digital signals. If an analog signal is used, digital transmission assumes the data are digital. Using this assumption, during transmission, the digital data, in the analog signal, are received and a new analog signal is generated from this data. This re-transmission of the analog signal generates a new clean analog signal where the noise associated with the signal remains constant. When digital transmission is sending digital signals, repeaters are used to relay the data. A repeater recovers the pattern of 0s and 1s from a digital signal, then it re-transmits the new signal. This process is the same for either digital or analog data when using a digital signal.

Digital signals are not practical for long-haul communications because they cannot be transmitted by

satellite, microwave, and optical fiber systems. To combat this problem, digital data is converted to an analog signal and sent by way of a digital transmission. This gives, the best of both worlds in the sense that we can send digital data over long distances using analog signal media (optical fibers, etc.) while avoiding the introduction of noise by using digital transmissions. The advantages of digital transmission have caused most major long-haul communication systems to convert to digital transmission for all types of data. Analog data suffers some loss when it is sent during a digital transmission only because the analog signal is converted into a digital data and then back again for retransmission purposes. This digital encoding is accomplished using pulse-code modulation (PCM) that is described further in the compression section of this article (Stallings, 1989:151).

All of this information is summed up in *Figure 4* which shows the different ways information moves from one location to another. Converting digital data into an analog signal and then sending it using a digital transmission takes advantage of the high speed digital transmission and the higher capacity transfer media used by an analog signal while eliminating the amplification of noise (errors in transmissions) that occurs with a analog transmission (This process is shown as number 5 in *Figure 4*).

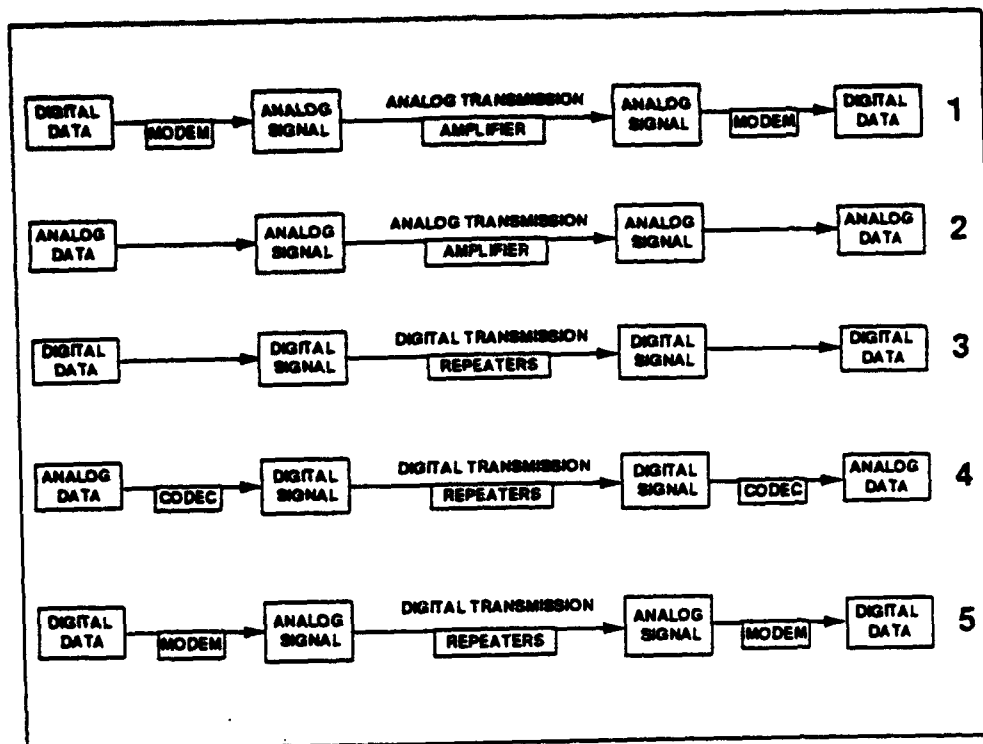


Figure 4. Network Communication Types

PART I NETWORK TOPOLOGIES

There are several ways to connect the sites in a network. These different ways of connecting sites is called topology. The advantages and disadvantages of each topology will be discussed in reference to the basic cost in terms how many links are added per site, communication cost, and reliability.

The first topology is a fully connected network where each site is directly linked with every other site in the network (See Figure 5). Basic cost of this type of system is high because of the number of communication lines. The basic cost grows with the square of the number of sites. Communication cost is very low since only one link is needed to travel between sites. The reliability is very high since several links must fail before the system becomes partitioned into parts that cannot communicate with one another (Silberschatz et al. ,1990:440).

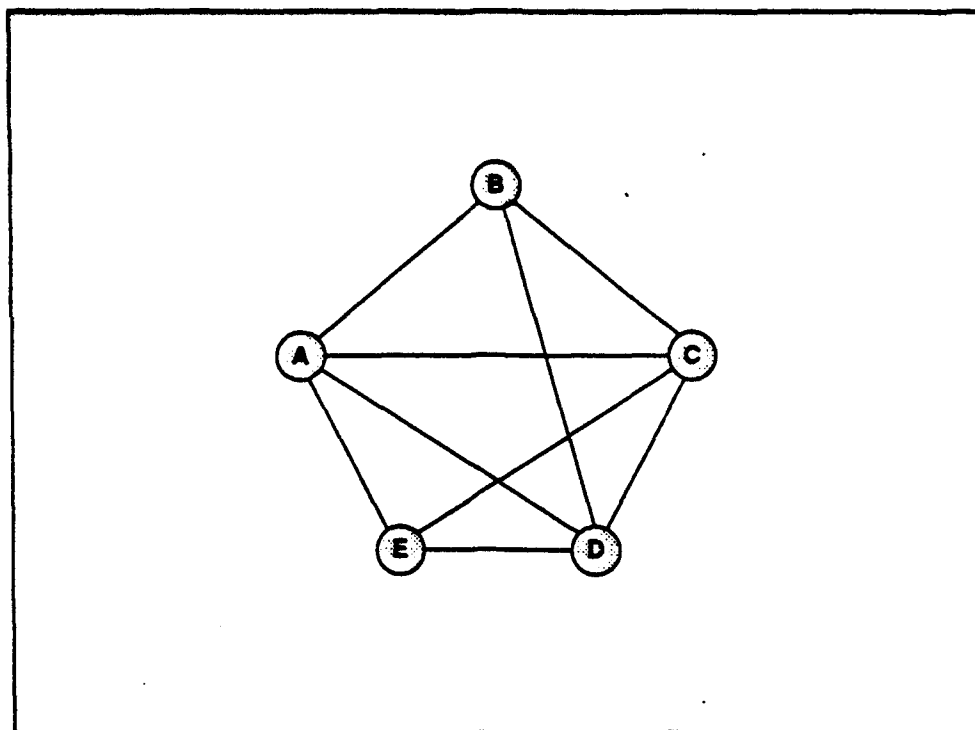


Figure 5. Fully Connected Network

In a partially connected network, direct links exist between some sites but not all (See Figure 6). The basic cost is less than that of a fully connected network. The communication cost is higher than a fully connected network since a message may have to be sent through several intermediate sites before arriving at the desired site. The reliability of the system is reduced since the failure of just one link could force the network to become partitioned. To prevent this, additional links must be added (Silberschatz et al., 1990:440).

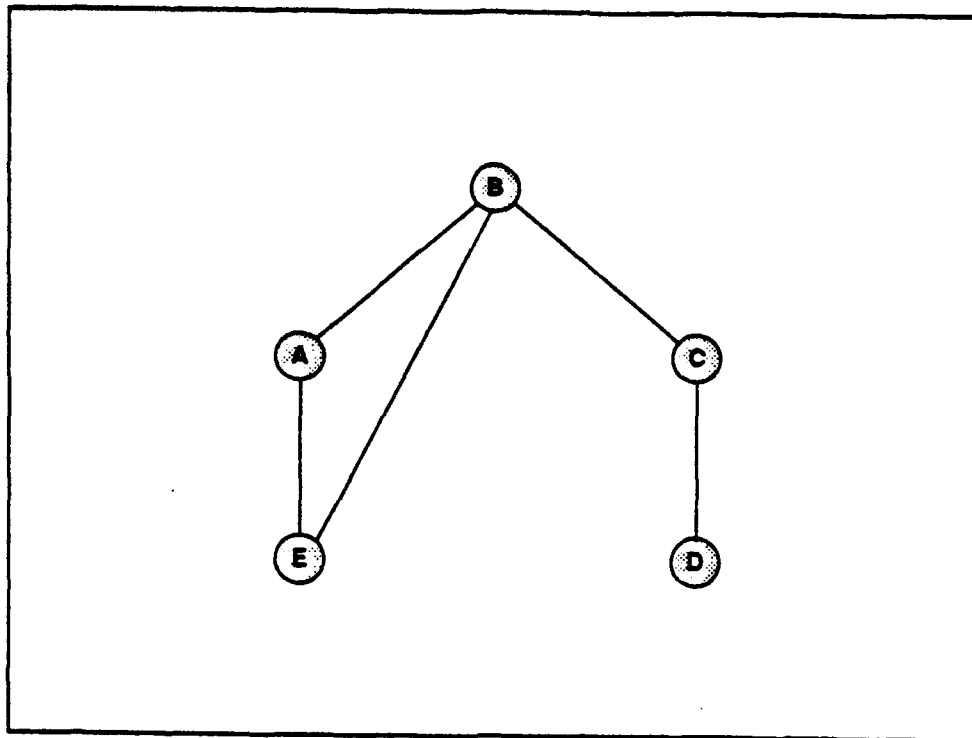


Figure 6. Partially Connected Network

A hierarchical network consists of a network designed as a tree (See Figure 7). This network has one root site that is connected to all other sites known as children. Each of the children must communicate by going through the nearest common ancestor. The basic cost is generally less than the partially connected scheme. Communication cost is restricted to going through a common parent or grandparent etc. The reliability of this scheme is fairly low since any broken link partitions the network into several disjoint trees (Silberschatz et al., 1990:441).

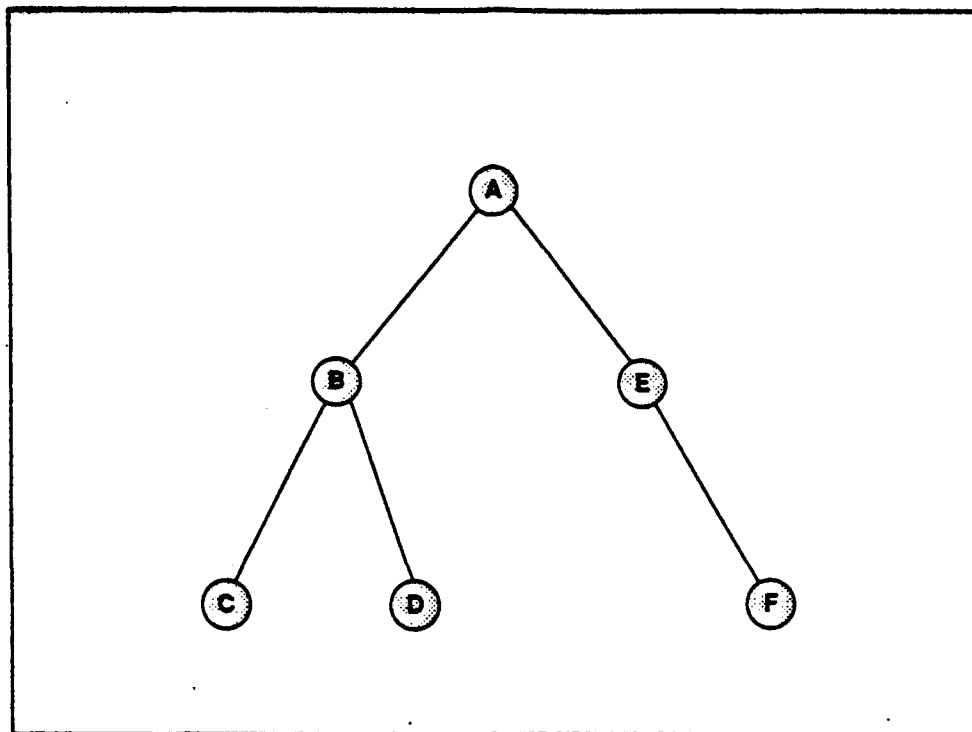


Figure 7. Tree Structured Network

A star network has one central site with all the other sites connected to the it (See Figure 8). The basic cost of this system is linear in the number of sites which is better than the first two types of topologies. The communication cost is also fairly low because it only requires two links to allow two sites to communicate; however, the amount of time to get a message can become long when a lot of messages are being passed and a bottleneck occurs. The reliability of the system is excellent unless the main site fails, then all the sites are completely partitioned (Silberschatz et al., 1990:442).

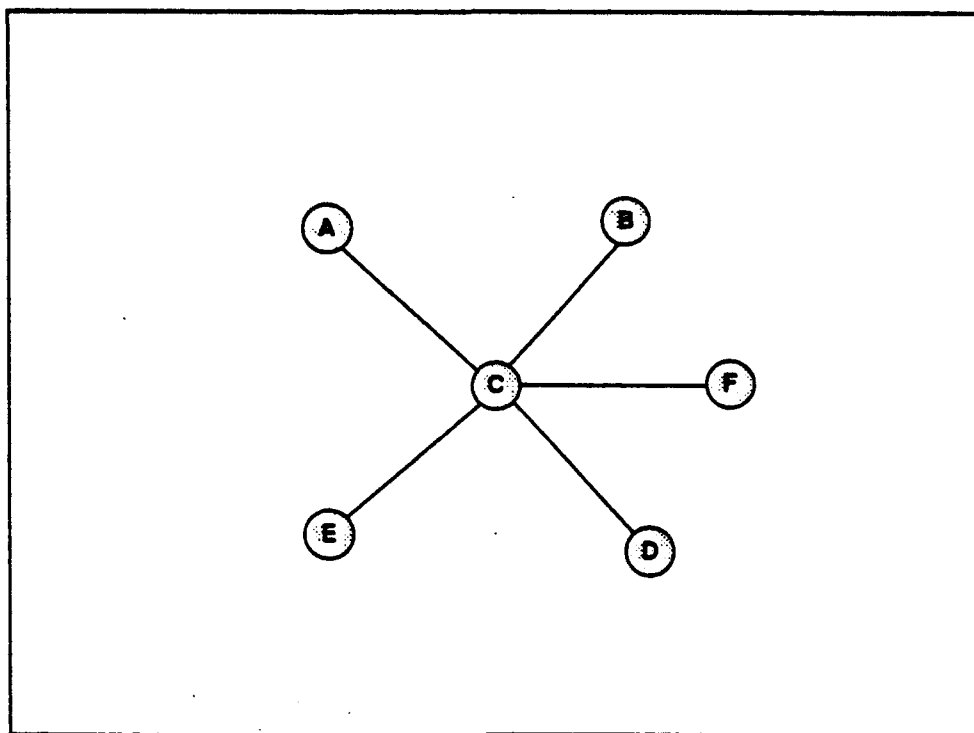


Figure 8. Star Network

A ring network consist of sites that are connected exactly to two other sites (See Figure 9). These connections can be uni-directional or bi-directional. In a uni-directional ring each site can only pass messages to one of its neighbors, while bi-directional architecture a site can transmit information in either direction. The basic cost of this scheme is linear to the number of sites which is the same as a star topology. The communication cost is high. In a unidirectional ring the maximum number of transfers is $n-1$ while a bi-directional ring has a maximum of $n/2$ transfers. A unidirectional architecture requires only one broken link while a bi-directional requires two

broken links to partition the network (Silberschatz et al.,1990:443).

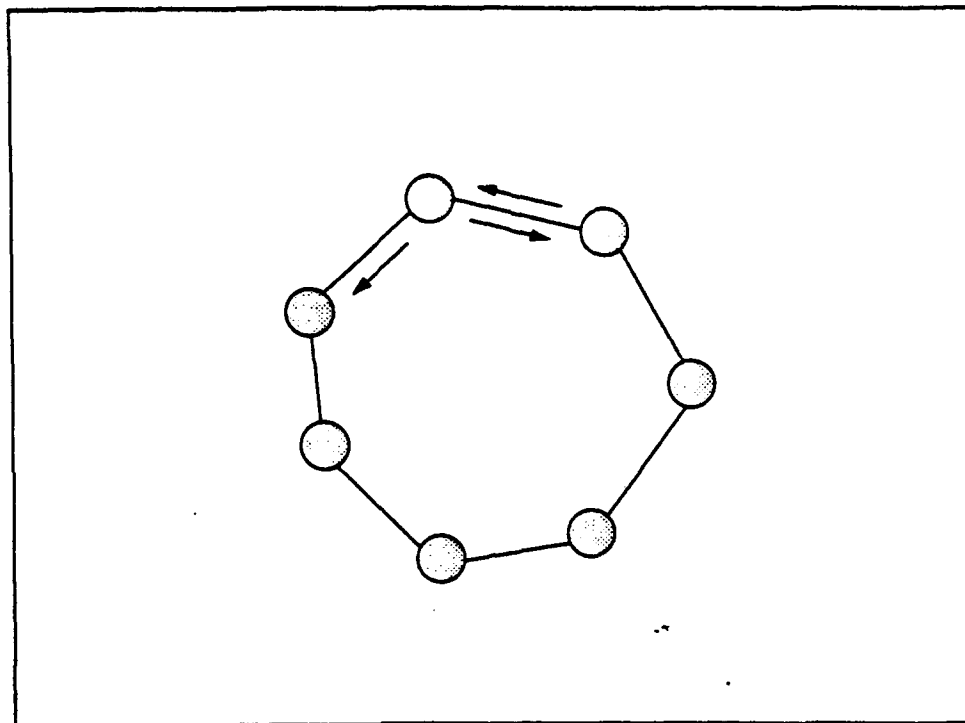


Figure 9. Ring Network

Another network is the multi-access bus network where all sites are connected to a bus that is a single shared link. Two possible configurations can be seen in Figure 10. The basic cost of this network increases linearly with the number of sites, the same as a star topology. The communication cost is quite low unless the bus-link becomes overcrowded. This configuration is similar to a star network with a dedicated central site. Reliability is fairly high unless the bus fails, then the network becomes completely partitioned (Silberschatz et al.,1990:444).

Refer to Table 3 for a summary comparing these different topologies.

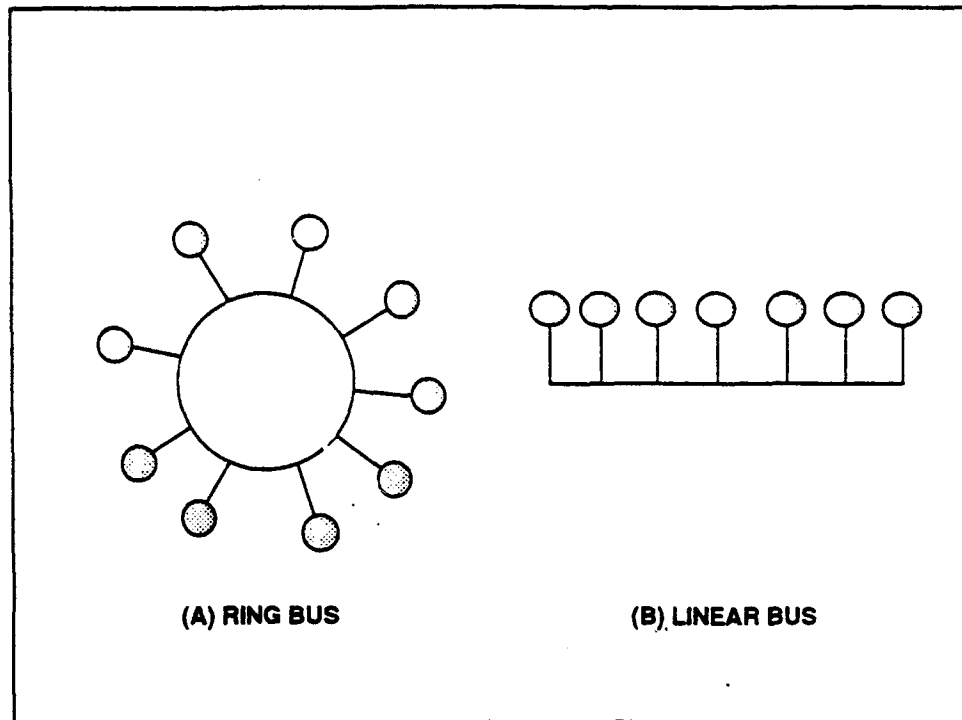


Figure 10. Bus Network

Table 3. Topology comparison

Topology Type	Basic Cost	Communication Cost	Reliability
Fully Connected	High(10) $n*n$	Very Fast (1)	Very Reliable(10)
Partially Connected	Medium(7) $>n$	Fast (3)	Reliable(7)
Hierarchy	Medium(6) $>n$	Slow(5)	Reliable(5)
Star	Medium(5) n	Very Fast (2)	Reliable(4)
Ring	Medium (5) n	Very Slow (8)	Reliable (8)
Multi-access Bus	Medium (5) n	Very Fast (2)	Very Reliable (9)

These topologies establish the basic physical structures available in networks today. Other issues that are involved with networks include routing strategies,

connection strategies and contention resolution. These topics are beyond the scope of this thesis but can be studied further in general network overviews by (Kahn,1972), (Silberschatz et al.,1991), (Doll,1974), (Crowther et al.,1975) and (Tanenbaum,1988).

PART II NETWORK CLASSIFICATION

Networks are usually classified into two categories, local area networks (LANs) and wide area networks (WANs). The difference between these two categories is the distance between sites and the speed of their transmissions (Silberschatz et al., 1990:449).

LANs emerged in the 1970s as a substitute for large mainframe computer systems. It was more economical to have more small independent computers hooked together in a network rather than having a single large system. Generally LANs are restricted to connecting sites that are located within 1 km of each other in order to maintain high transmission speeds. These networks basically are used by computers (also referred to as sites, nodes or hosts) that are located in one building or in several buildings within close proximity. Since the sites are close to one another, the communication links have a higher speed and lower error rates when compared to WANs. Typical bit rates range from 1 megabyte per second on a conventional twisted pair wire to fiber optical LANs having possible transfer rates around 1 gigabyte per second, with 10 megabytes per second being the norm. A LAN's communication link media can include twisted pair wire, baseband coaxial cable, broadband coaxial cable, and fiber optic cable. Most hospital networks are LANs, with the newest being built with fiber optic cable that have

the wide bandwidth required for medical images (Silberschatz et al., 1990:449).

A WAN is usually used to connect computers when the sites are greater than 10 km apart from each other with LANs connecting sites less than 10 km. Typically a WAN connects LANs, with the LAN to WAN connection computer being the host in the WAN topology. In other words, a WAN connects LANs to form one big network. An example of a WAN is the Internet WAN. This network is broken up into regional networks that are connected together using routers to form a worldwide network (See Figure 11). A WAN typically has a transfer rate from 1200 bytes per second to more than one megabyte per second and is generally slower than a LAN. One of the reasons for this slow transfer rate is the media being used to send the data. A "fast" WAN can attain its speed by having a dedicated phone line, known as a T1 line, provided by the telephone system with a transfer rate of 1.544 megabits per second. Slower WANs use standard telephone lines which have a slower transfer rate than the T1 line.

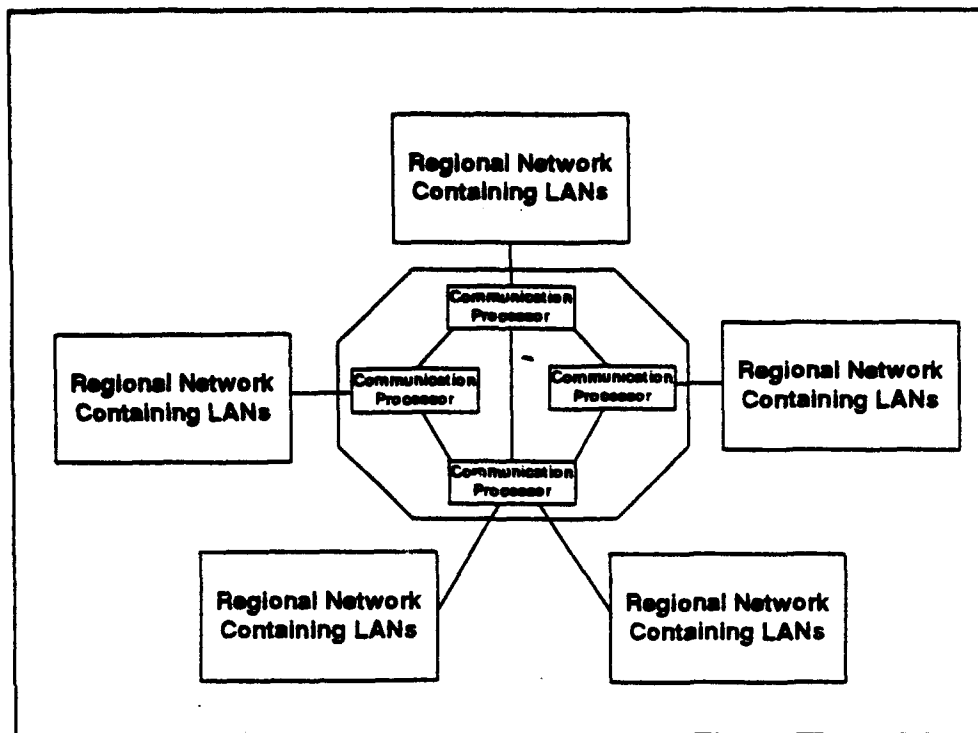


Figure 11. Example of a Wide Area Network (WAN)

To combat the slow transmission rate experienced by WANs, an Integrated Services Digital Network (ISDN) was developed to bring digital channels, recently adopted by the telephone companies, all the way to the customer. The ISDN has a special provision of channels, for tasks such as transmitting images referred to as broadband ISDN or B-ISDN. Specific channels on the B-ISDN have been identified for different transfer rates from 384 kilobits per second to 135 megabits per second, a significant increase over previous T1 lines. This increase in transfer rates will help to make a PACS and GPACS a useful tool. As the capabilities and amount of data increases, the network transmission rates are quickly becoming the bottleneck in the information pipeline,

and the ISDN is a possible solution. For more information on the ISDN, refer to (Stallings,1989 and Griffiths,1992).

The ISO has defined a set of standards that apply to network communications. This standard was established in order to handle slow, error-prone, asynchronous environments that exist on any given network. This standard divides the communication process into seven layers. Each of the seven layers are defined below:

1. **Physical Layer** - This is the layer that handles the mechanical and electrical issues of physically transferring a bit stream.

2. **Data-link Layer** - This is the layer that handles the sending and receiving of frames, fixed length parts of packets, in the bit stream. Also error detection and recovery of data corrupted in the physical layer occurs here.

3. **Network Layer** - This layer is responsible for providing connections and routing of packets in the communication network. This includes handling outgoing and incoming packet addresses while also maintaining routing information in order to be able to respond to changing load levels.

4. **Transport Layer** - This layer is responsible for low-level access to the network and allows the transfer of messages between clients. This task includes partitioning messages into packets, maintaining the order of packets,

controlling the flow of packets and creating the physical addresses.

5. Session Layer - This layer implements sessions otherwise known as process-to-process communication protocols where the two computers are using the same format for file transfers, mail transfers and remote logins.

6. Presentation Layer - This layer resolves the differences in formats among various sites in the network including character conversions and either half-duplex or full-duplex communications.

7. Application Layer - This layer deals with interacting directly with the users. This interface can include file transfers, remote-login protocols and electronic mail.

These layers play a very important part of digital communications because they give a model to follow when communicating with other sources. A pictorial representation of two computer communicating using these layers is shown in Figure 12. The American College of Radiology and the National Electrical Manufacturers Association (ACR-NEMA) has developed a standard within the confines of the ISO layers to promote the interconnection of different imaging modalities and different manufactures (McNeill et al., 1992). The ACR-NEMA standard is being examined as an international standard but, because of some of its shortcomings, it is under scrutiny. This image file standard is the first of its kind to be introduced and although ACR-NEMA specifically

states no claims to being a PACS standard (Maydell and MacGregor,1989), it is still being considered as an international standard because of its potential. The examination and critical review of the ACR-NEMA standard can be found in (Maydell and MacGregor,1989). The following information is taken from that article.

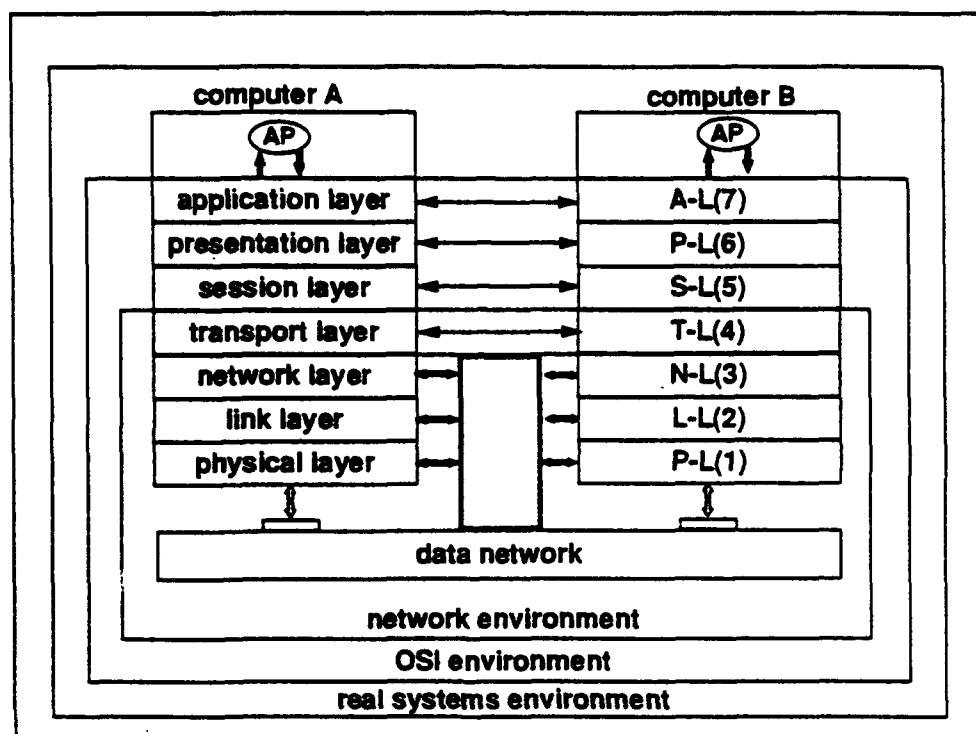


Figure 12. ISO Seven Layered Reference Model

The ACR-NEMA protocol defined in the terms of the ISO seven-layered reference model is as follows:

1. **Physical Layer** - This is a 50 line interconnect, for interfacing two devices. There are 16 data circuits (plus parity and ground), and 6 control circuits and all circuits have two wires per circuit (differential). This could be

characterized as an asynchronous bus, with the control lines used to determine which device is the controller of the bus or bus master.

2. **Data-link Layer** - Flow control is stop-and-wait, with a 1 packet window while there is error checking with no recovery and it has collision arbitration for bus access.

3. **Network Layer** - Packetizing is provided along with sequencing and virtual channels. This layer sets a packet size limit of 32 Kbits and it does not have a routing function.

4. **Transport Layer** - This is the same as the network layer.

5. **Session Layer** - This layer provides direct interface to the user in order to start and end connections with some device. It does not agree with the ISO reference model since it is interacting directly with the users.

6. **Presentation Layer** - This layer builds a message out of ACR/NEMA groups and is the agent for standardizing formats.

7. **Application Layer** - This layer provides image capture

Network addressing problems forced consideration of standardizing PACS. A PACS, by its very nature, cannot truly meet its potential without being in a networked environment. For networking to occur, some standard has to be established, otherwise communication is haphazard. The basic address model, as stated before, has already been

determined by the ISO. Since a PACS deals mainly with medical image data, the current ISO standards may not be appropriate. At the present time, network transmission protocols are designed around the Open Systems Interconnect (OSI) model of transmission adopted by ISO. The ACR-NEMA standard could be incorporated into the already existing Government Services Agency OSI address format or the ANSI OSI format. The break down of the addresses can be seen in Figure 13. The total address space for these addresses is 2^{160} . In each case the Initial Domain Part (IDP in the figure) uses 2^{32} address leaving 2^{128} for the Domain Specific Part (DSP) of the total address. An organization like ACR-NEMA is assigned an administrative authority number in the GSA case or a numeric organization name in the ANSI address format, that then leaves 2^{104} addresses for PACS. In either case, the possibility exists for the ACR-NEMA standard to be utilized.

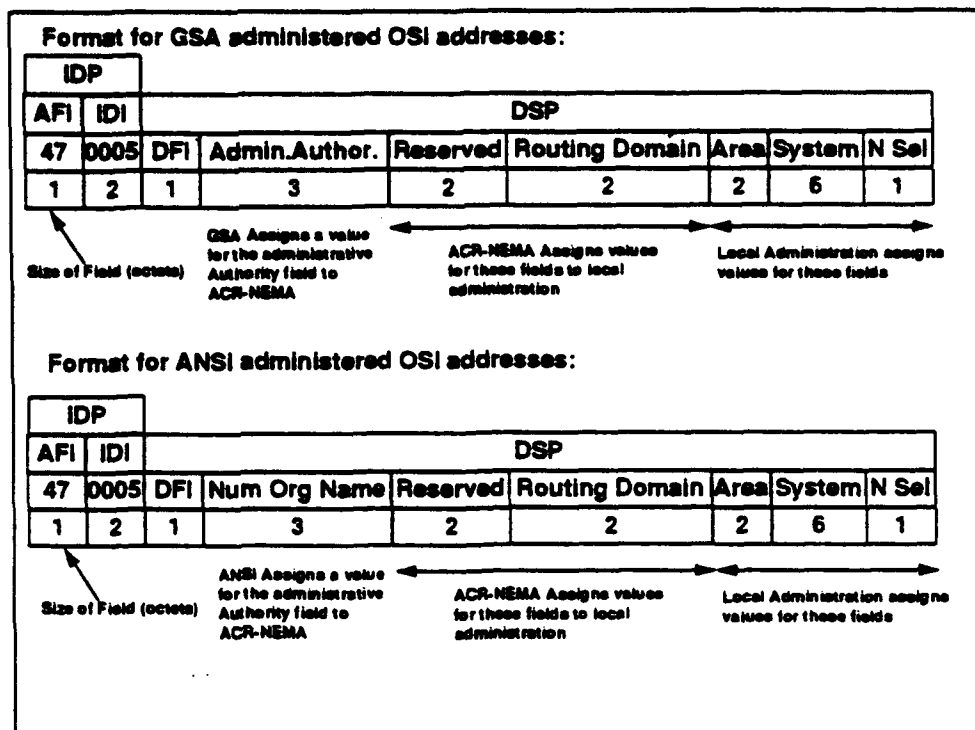


Figure 13. OSI Address Formats

The use of networks is crucial to any large computer system, especially a PACS. A network is the basic underlying structure that allows communication and transmission of data including medical images. At present no addressing schema has been adopted for PACS because of the lack of standardization for transferring image files. The continuing development of networks and their transmission media, along with the establishment of a PACS addressing scheme, will be the determining factors in the development of a real-time GPACS.

CHAPTER 4

THE PACS DATABASE: THE BACKBONE OF A PACS

A discussion of the problems of interacting with other computers and other hospitals leads to an examination of some of the underlying software components of a PACS. The single biggest software component, because of the storage requirements, is the PACS' database. This is the software that stores and manages the medical images of the hospital. This section discusses the storage utilities used along with the different possible types of databases.

PART I STORAGE HIERARCHY LEVELS

A database system is defined by C. J. Date (Date,1990:5) as a computerized record-keeping system that maintains data, making it available on demand. In the context of a PACS, this data can be text, images, sound or video. The actual database consists of a collection of persistent data that is used by an application system of some entity, in this case a hospital. Since a hospital is so large, it can have several databases. The advantages that a database system has over traditional paper-based and film-based methods of record keeping include compactness, speed, elimination of tedious mechanical tasks, and currency of data. The overriding advantage of a database system is that it provides the hospital with centralized control over its data. Terms often used when discussing database systems can be found in Table 4 in Appendix A.

In the PACS environment, a database forms the foundation for storing and retrieving medical image data. Still images typically range in size from 1 megabyte(MB) to 25 MB (Turner and Peterson,1992:258), and medical images are usually 10 MB (Allen and Frieder,1992:42). As the resolution and accuracy of imaging equipment increases, so will the amount of data being collected. Increased data requires a large capacity storage facility. A PACS database, can be broken down into three levels of memory storage(Figure 14). The first level consists of short term

storage. These are the images that have been recently acquired or accessed. An image will remain in short-term storage for approximately 1 to 7 days. Magnetic storage media is used for short-term storage because of its ability to read and write quickly. Once an image has resided in short-term storage for its allotted time, it migrates to the next level of storage, referred to as intermediate-term storage. Once the images have resided in intermediate storage without being accessed for another period of time, typically 3 to 31 days, they are moved to the long-term storage media. The intermediate and long-term storage media consist of optical disks that are very similar except that the intermediate level has an automatic optical disk juke box, whereas the long-term storage has a manually loaded optical disks. The storage required at the intermediate level is around 31 GB (gigabytes(GB) = 10^9 bytes) while the long term storage requirement is around 3.2 TB (terrabytes (TB) = 10^{12} bytes)(Table 5 (Allen and Frieder,1992:43)). With newer technology, the intermediate/long-term storage is combined in an automatic optical disk juke box which contains 100, 10.2 GB optical platters for more than a 1 TB worth of data (Quillin,1992). These platters can be removed, placed on shelf and then replaced with a brand new platter (See Table 6, Figure 15). Figure 15 illustrates how the storage components of a PACS are connected. In either case, the database stores the information based on usage and

available memory. The current trend and practice is to acquire the image, immediately store it onto a Write Once Read Many (WORM) optical disk, then send the digital image out to the requesting department where it is stored in the workstations resident memory (Valentino,1993). This ensures that a copy of any image is always available even if magnetic memory is destroyed. Another common practice is to have complete back ups for the optical disk jukeboxes and hard drives in case either has a catastrophic failure.

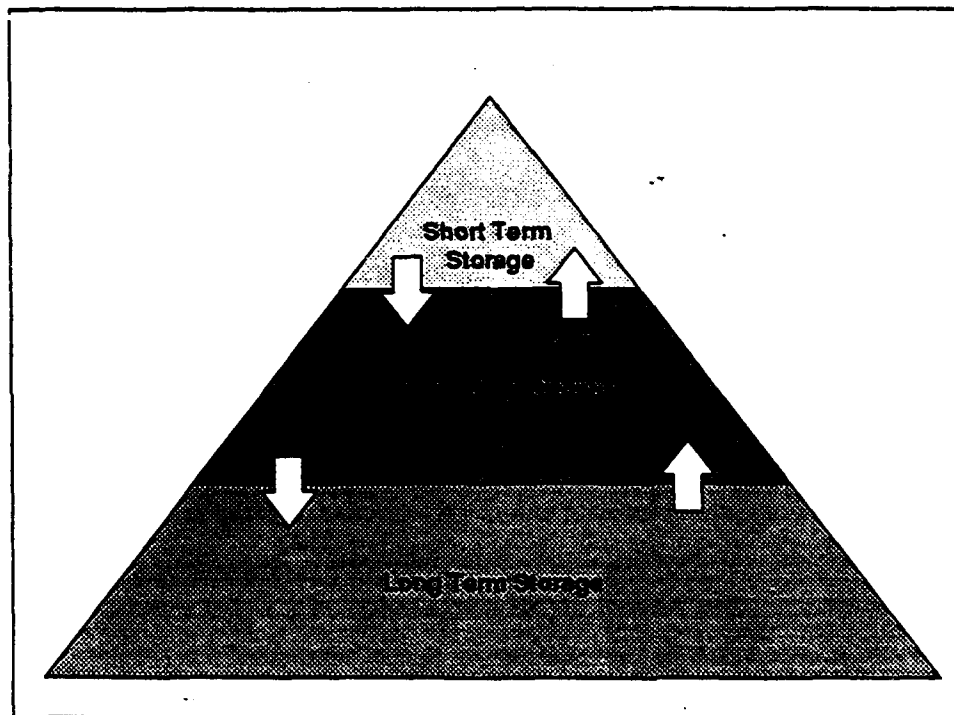


Figure 14. Time In Storage Triangle

Table 5. Three-Level Hierarchy

Level	Retention Period	Typical Storage Media	Approximate Storage Capacity
1	1 to 7 days	Magnetic Disk	21 GB
2	8 to 31 days	Automatic Optical Disk Juke Box	31 GB
3	32 days to 3 years	Manual Optical Disk Juke Box	3.2 TB

Table 6. Modern Three-Level Hierarchy (MDIS at Wright Patterson Air Force Base)

Level	Retention Period	Typical Storage Media	Approximate Storage Capacity
1	48 hours	Magnetic Hard Disk	20 GB (upgraded to 40 MB)
2	3 days to 5 years	Automatic Optical Disk Juke Box	1 TB (100 - 10.2 GB platters, compressed images)
3	Over 5 years	Optical Disk Platters	10.2 GB platters

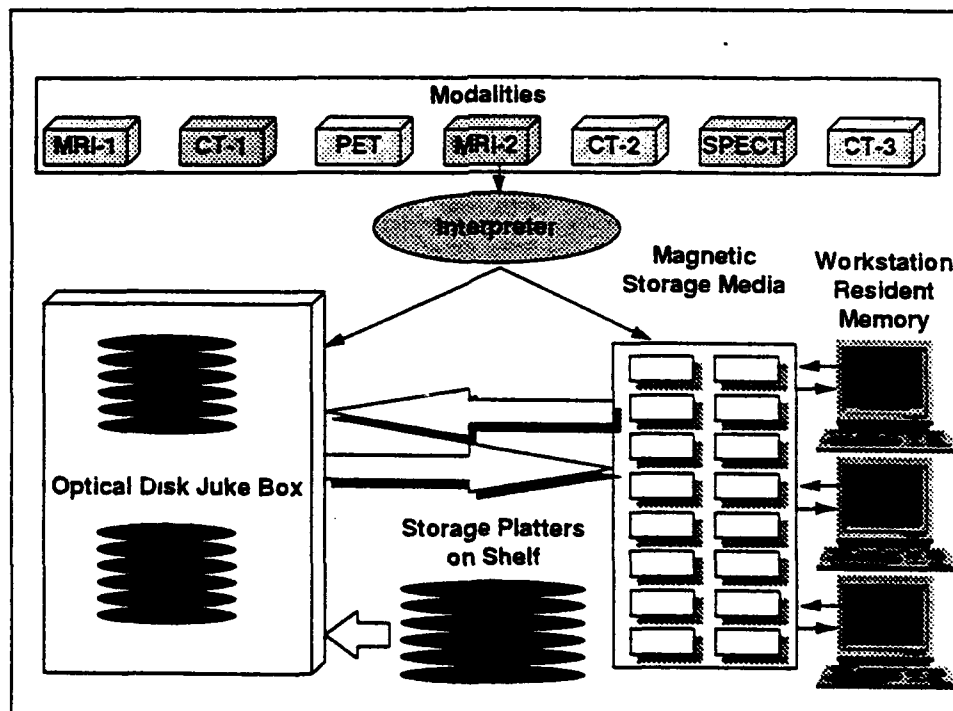


Figure 15. PACS Storage Hierarchy

PART II DATABASE ISSUES

The hierarchical storage media concept solves the problem of storing the medical images, and the next question is how to reference those images in the database. The first databases were flat-file and hierarchical databases. These databases had several problems such as redundancy and lack of data integrity. In 1970, E. F. Codd (Codd,1970) introduced the relational model. This model solved many of the problems associated with the hierarchical and flat-file models (Shindyalov and Bourne,1992:38). The relational model represents data in a tabular structure that allows for a non-redundant description of data in an easily understandable format. Recently, the object-oriented model has emerged to address the shortcomings of the relational model by modeling data closer to real-world objects. The object-oriented model represents data as objects. These objects not only exhibit a state but also a behavior associated with that state. The relational model now has a formidable competitor that challenges its domination of the commercial market as seen by the emergence of object-oriented models in the market.

For clarification purposes, a Database(DB) holds the data location and structure while a Database Management System (DBMS) is the program that interacts and regulates how the data is stored in the DB. The other item that causes confusion is the query language. This is a specific

tool that allows the user to make requests on the DB. DB and DBMS support each other and sometimes these acronyms are used synonymously.

PART III RELATIONAL DATABASE

The following information on relational databases is drawn from (Codd,1970,1979,1990), (Date,1983,1990,1990), (Chen,1983), (Ullman,1989), (Silberschatz et al.,1990), (Frank,1988), (Elmasri and Navathe,1989), (Tsichritzis and Lochovsky,1982) and (Stonebraker,1989).

To better understand the relational database (RDB), a discussion of its data types is necessary. The three data types defined by the RDB are: a table (representing a relation), the row (representing a tuple), and the column (representing an attribute). Some terms that describe these types include number of rows (cardinality), number of columns (degree), the unique identifier (primary key) and allowable values (domain). The four basic properties of a table are: (1) There are no duplicate rows, (2) The rows are unordered (top to bottom), (3) Columns are unordered (left to right), and (4) All column values are atomic (precisely one value, not a list of values at every row-column position). A relational algebra, presented by Codd, that consists of a collection of high-level operators, is used to operate on a table. The operators, that are defined in Table 7 in Appendix A, consist of:

1. RESTRICT
2. PROJECT
3. PRODUCT
4. UNION

5. **INTERSECT**
6. **DIFFERENCE**
7. **JOIN**
8. **DIVIDE**

For a graphical representation of these operations see Figure 16, taken from (Date,1983). This figure illustrates the re-action of the rows and columns when the above operations are applied to the tables containing them. The dark color shows the result of combining two tables with certain operators.

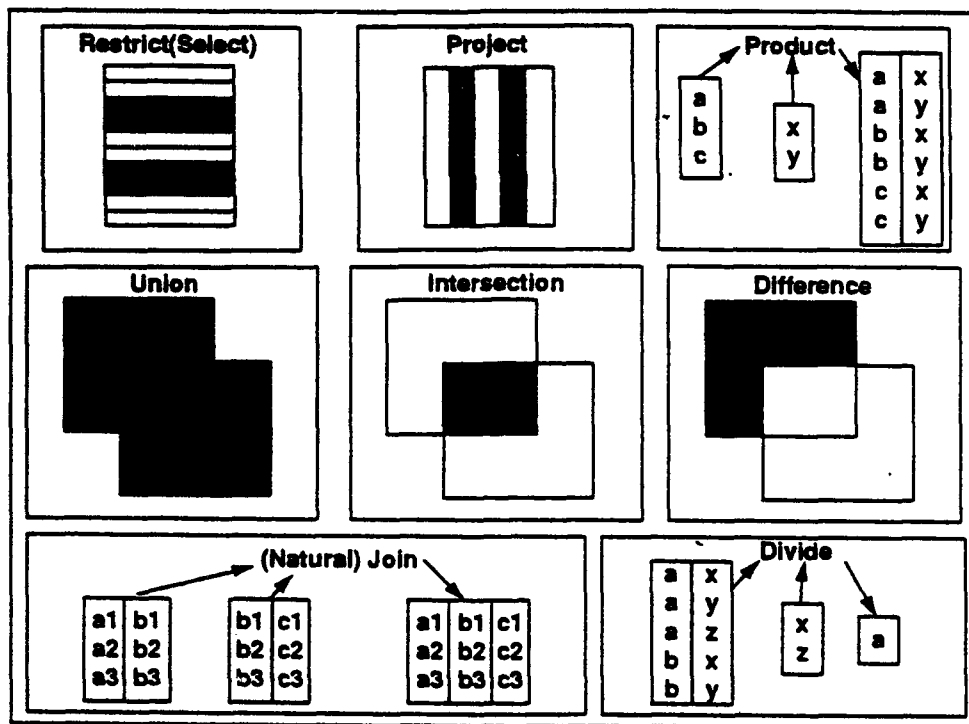


Figure 16. RDB Basic Operations

These basic tools allow someone using the database to access needed information in a easy to understand manner.

Once the information is accessed, it is returned to the user in an understandable format, always a table.

Examples of how a typical RDB would store information about a patient are presented in Tables 8 & 9. Table 8 consists of information concerning the patient's doctor, the patient's next appointment and how many images of the patient are available. Table 9 shows a separate entry for each image acquired including the patient's name, the date the image was received, the type of imaging device used to acquire the image, and the part of the body the image shows. These examples show the results of applying the RESTRICT and JOIN operators on the provided tables. A query requesting all images taken of Bob Smith would use the RESTRICT operator. The resulting table can be seen in Table 10. To combine all of the information in Tables 8 & 9 a JOIN is used on the "patient's name" column. The result is depicted in Table 11. The other basic operators are used to manipulate the tables further as described in Table 7 in Appendix A.

Table 8. Relational Database Table (Patient Table)

Patient's Name	Doctor	Number of X-rays Available	Date of Next Appointment
Bob Smith	Dr. Jones	2	12 July 1993
Sue Adams	Dr. Miller	4	12 July 1993
Mark Brown	Dr. Jones	0	13 July 1993

Table 9. Relational Database Table (Radiology Patient Table)

Patient's Name	Type of Image	Image Pointer	Date Image Acquired	Acquisition Method
Bob Smith	Chest	123456	10 June 1993	CT
Bob Smith	Chest	456734	12 June 1993	MRI
Sue Adams	Head	123456	5 Nov 1991	MRI
Sue Adams	Head	453456	7 Nov 1991	CT
Sue Adams	Liver	143456	21 May 1993	Ultrasound
Sue Adams	Liver	983456	22 May 1993	CT

Table 10. Result of Restricting New Table to Bob Smith's Information Only

Patient's Name	Type of Image	Image Pointer	Date Image Acquired	Acquisition Method
Bob Smith	Chest	123456	10 June 1993	CT
Bob Smith	Chest	456734	12 June 1993	MRI

Table 11. Result of a Join on Name of Table 8 and Table 9

Patient's Name	Doctor	Number of X-rays	Date Next Aptmt	Type of Image	Image Pointer	Date Image Acquired	Acquisition Method
Bob Smith	Dr. Jones	2	12 July 1993	Chest	123456	10 June 1993	CT
Bob Smith	Dr. Jones	2	12 July 1993	Chest	456734	12 June 1993	MRI
Sue Adams	Dr. Miller	4	12 July 1993	Head	123456	5 Nov 1991	MRI
Sue Adams	Dr. Miller	4	12 July 1993	Head	453456	7 Nov 1991	CT
Sue Adams	Dr. Miller	4	12 July 1993	Liver	143456	21 May 1993	Ultrasound
Sue Adams	Dr. Miller	4	12 July 1993	Liver	983456	22 May 1993	CT

To manipulate, that is manage and extract information from, relational databases, several languages have been developed. The most popular of these languages is

Structured Query Language (SQL). However all of these languages are based on the basic operators discussed above. Each language may have a different way of implementing queries (requests of the database), they all serve as a tool to get information from a database and choice of query language is usually decided by the brand of database purchased.

How does a relational database, in a PACS context, manages images? To reference a medical image using a relational database, a pointer to the location of the image data is created(Quillin,1992 and Valentino,1993). When the image is selected, the format of the image needs to be identified to view the image otherwise the computer has no way of extracting the data. In most PACS, when the doctor indicates that he wants to view an image, the software has to identify that the information being referenced is a pointer and that the viewing software has to be activated with the pointer as the place where the data can start to be interpreted. This is not necessarily the most efficient approach to viewing images, especially when there are many image formats to deal with, but it is simple enough to be implemented using a RDB. The future of increasingly complex data types gives rise to new challenges that the RDB is not necessarily prepared to handle. A database strategy that seems to be well suited to meeting these future challenges is the object-oriented database.

PART IV OBJECT-ORIENTED DATABASE

The following discussion on OODBMS is drawn from (English, 1992), (Gupta and Horowitz, 1991), (Kim and Lochovsky, 1989), (Hughes, 1991), (Gray et al., 1992), (Zaniolo et al., 1986), (Tsichritzis and Lochovsky, 1982), (Micallef, 1988), (Frank, 1988), (Booch, 1991), (Silberschatz et al., 1990), (Varma, 1993), (Shindyalov and Bourne, 1992), (Shah et al., 1993), (Martin and Odell, 1992), (Aubry et al., 1991), (Routhier, 1992) and (Hurson et al., 1993). The June 1993 issue of the Journal of Object-Oriented Programming has reviews of current OO books (Bilow, 1993). For further reading and reference please refer to these books and papers.

Object-Oriented Data Base Management Systems (OODBMS) differ from Relational Database Management Systems (RDBMS) in that the Object-Oriented Database (OODB) deals with the concept of objects rather than relationships. The basis of this Database Management System (DBMS), the object, can be described as an abstraction of a real-world entity that exhibits states and behaviors. The object-oriented paradigm incorporates a new concept, that of data encapsulation. The state of an object refers to the values contained within an object such as a name of a patient, the patient's doctor, the date of the next appointment, etc. The behavior of an object is expressed as a set of methods, or operations, that operate on its attributes. The concept of having a state

and behavior is a completely different paradigm from the relational paradigm. With this paradigm, the object cannot operate without having a behavior associated with it because the behavior is contained within the object. This attribute makes an object-oriented design unique. This uniqueness comes from the fact that in order to reference different objects, the type of data contained in the object is no longer an issue to the user requesting it. Therefore, a request no longer needs to be tailored to a specific piece of data, such as text, image, or sound, rather it is just treated as an object within the database. Booch, 1991 describes four major and three minor elements and asserts that the major elements must be present if the system is to be called object-oriented. The major elements, described in Table 12 in Appendix A, are:

1. **Abstraction**
2. **Encapsulation**
3. **Modularity**
4. **Hierarchy**

The minor elements of the Booch object model are:

1. **Typing**
2. **Concurrency**
3. **Persistence**

Although Booch talks about the elements of the object model, no formal object model has been accepted. In an effort to address this issue, a consortium of ODBMS vendors

has formed the Object Database Management Group (ODMG). The current voting members include :

1. Object Design
2. Objectivity
3. Ontos
4. O₂ Technology (O₂)
5. Versant Object Technology (Versant)

The results of their efforts has yielded the ODMG-93 standard. This standard defines an object model and programming language closely related to the current commercial products. This follows considering the members of the ODMG currently have over 90% of the current object database management systems market. Developers claim that this standard will do the same thing for object-oriented databases as the relational model did for the relational databases. For further information on this topic see (Atwood,1993).

To better illustrate the idea of an object in a object-oriented database, an example based on the patient information presented in the relational database example is shown in Figures 17 & 18. The OODB firsts sets up an object with the basic attributes of any person. This object description, because it is generic in nature, can be used for all people in the hospital, therefore, this object is referred to as a class. Notice that each object has its own attributes with operations on these attributes. The objects

also contain relationships with other classes known as associations. A doctor is a person and so is a patient, therefore, they are sub classes of the person class. Each sub-class has its own traits plus the traits of all of its parent classes. If a patient is further classified as a radiology patient, then a new object can be declared under the class of patient, in this case a radiology patient. Although the actual specifications of a radiology patient only include number of X-rays taken (medical images), the patient attributes are inherited and consequently observed as part of the radiology patient. Furthermore, when a person is termed a radiology patient, he inherits the patient and person attributes. Images can also be associated with the radiology patient. Figure 18 shows a graphical representation of the inheritance and associations.

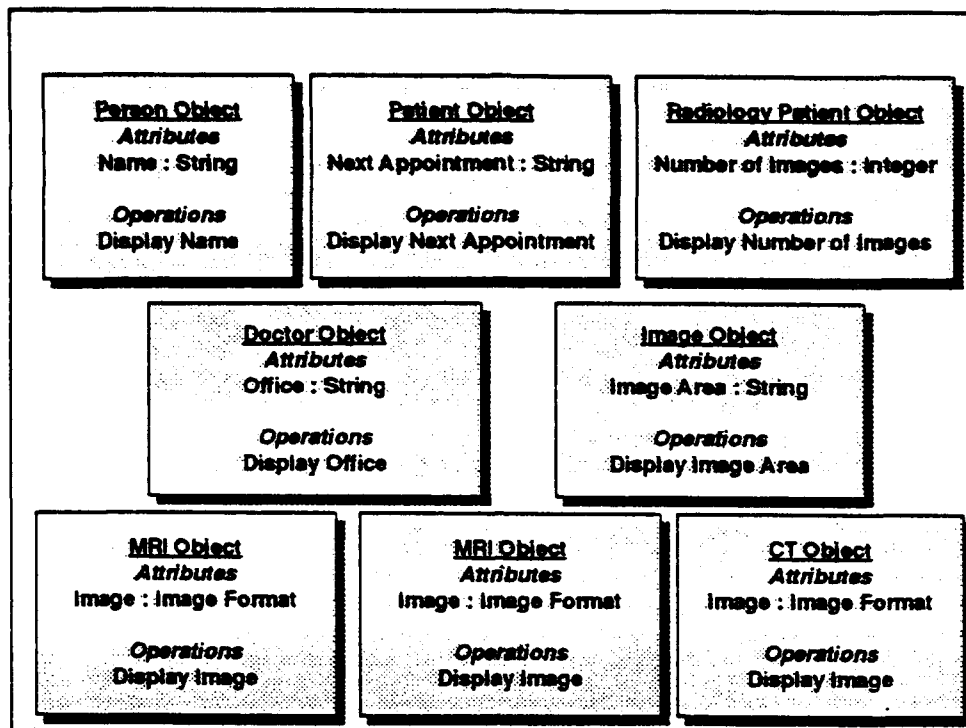


Figure 17. Examples of Object Classes

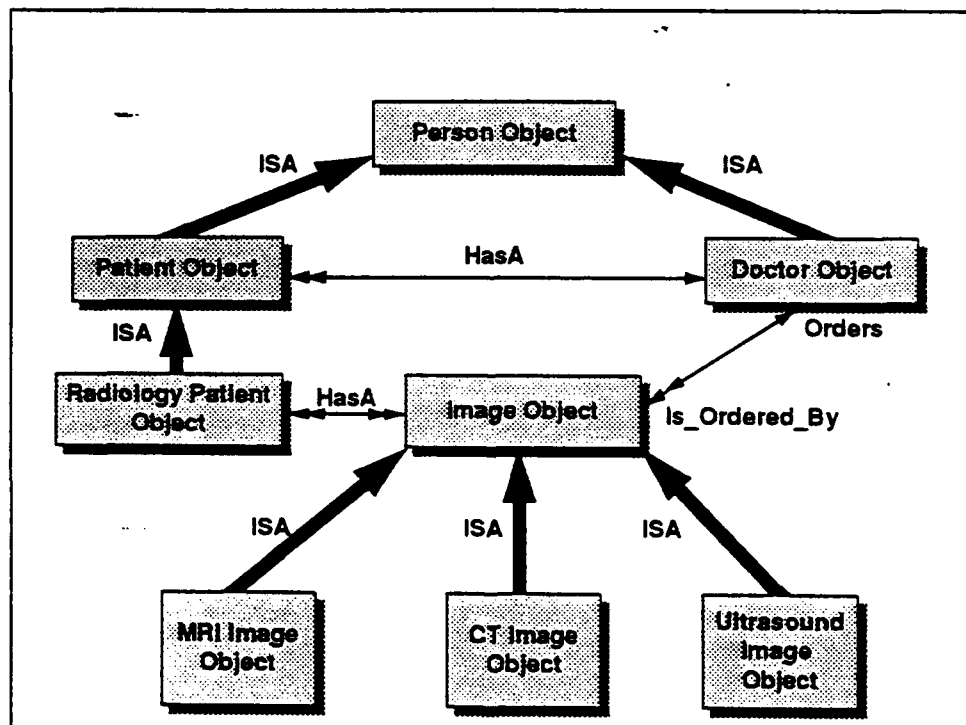


Figure 18. Example Object Relationships

PART V THE DATABASE CONTROVERSY

The controversy of OODBMS versus the RDBMS is a fairly new development in the hospital setting. The controversy stems from the way hospitals operate today compared to the way they will operate tomorrow. RDBMS are currently being used for PACS. At the time PACS was developed, OODBs were not commercially available. That being the case, RDBs were used to meet the data management needs of PACS. There are several reasons PACS continue to use RDBs to reference medical images stored in memory and on disks. The first is the simplicity of having a pointer that points to an area of memory that identifies the image to be viewed. This is an advantage because pointers allow the user to easily identify the location of the desired data without having to go through a complicated path. Also the pointers take up very little space within the actual database. The second reason RDBs are used is because of the significant cost of re-configuring a database. Lastly, the relational model and implementations of this model have been demonstrated to be effective with real database systems.

Although the RDB is currently entrenched in the commercial hospital community, the object-oriented model's potential is being appreciated and is causing a push to use object-oriented programming and databases. The database community appears to be leaning towards OODBs as the databases of the future. With this thought in mind, many

companies have recently developed either extended RDBs containing features relating to OODBs, or they have developed an OODB. These new products are becoming more abundant everyday, challenging the reign of the relational products. A look at specific advantages and disadvantages of both the RDBMS and the OODBMS outlines the differences between these two types of databases.

PART VI RDB ADVANTAGES

As discussed previously, the RDB has been around since the early 1970's and is based on a mathematical foundation developed by E. F. Codd (Codd,1971). These two facts give the RDB a significant advantage in the commercial market because it has stood the test of time (relatively speaking), and every relation and manipulation can be substantiated using the mathematical foundation. Relational databases have more general advantages based on its concepts. One advantage of the relational database is the simplicity of the data organization. The tables that represent the relationships between the data are easy to understand. If the user inputs a query, the response is usually in the form of a table that is easy to read and comprehend. A second advantage is that RDB products are mature, leading to several user benefits such as increased database capacity and complexity. Several different query languages have been developed to complement them. These query languages, along with the actual databases, have continued to improve making a very user-friendly environment for the people that use this particular type of database. The evolution of the RDB has allowed developers to improve products to meet the needs of users. The streamlining increases the acceptability of the database, consequently more RDB products are purchased. Some of the major RDB vendors are Oracle, Sybase, Informix and Ingres. A list of current relational and some object-

oriented databases can be found in (Garcia and Davis,1993). This list is not complete but can give the reader a relatively recent list of available databases and their prices. The RDBMS used in clinical PACS are meeting the demands of today, making the RDB appear to be the solution for handling the organization of data but whether they it can meet the challenges of tomorrow is an open question.

PART VII

RDB DISADVANTAGES

As mentioned above, an RDB meets the needs of text oriented data systems. The problem with the RDB is the limited flexibility available when handling different forms of data such as images, sound, video, and any other Binary Large Objects (BLOBs). When one of these real world entities does not fit into the relational model directly, artificial decomposition, or breaking data into parts, or using a pointer to where the data is really located becomes necessary (Hurson et al, 1993). The reason these complex data types will not fit into the relational database is because all the fields are normalized, and by definition, normalized relations are non-decomposable, which is untrue for complex entities. The solution currently being used to solve this data handling problem is to follow a pointer to the location of the data and process it appropriately. Since the primary goal of the relational database is to maintain data independence, the obvious benefits of tightly coupling behaviors with the data cannot be exploited. As data becomes more complex, the queries and access methods for the relational database will have to continue to compensate for the database's short-comings.

Another disadvantage of the relational database is that data integrity is in the control of the application programs acting on the RDB. If an application fails to correctly handle constraints such as one-to-one and one-to-many

relationships, then the integrity of the data within the database is lost. The one-to-one and one-to-many relationships being referred to can be seen in Figure 19. This figure shows how data is related to other data in either a one-to-one, or one-to-many relationship. Also, entity identity integrity is not assured in a relational database because if a primary key changes, so does the entity's identity.

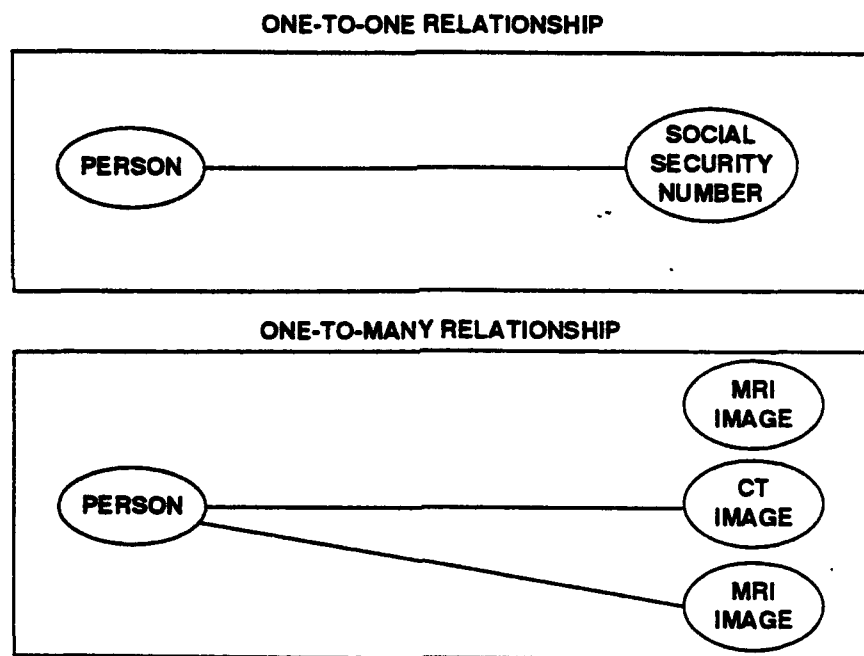


Figure 19. One-to-One and One-to-Many Relationships

Lastly, the relational database lacks the ability to make a structural change to the database without having to modify all the applications associated with it. The

difficulty of evolving with changing requirements makes retaining a relational database future PACS unlikely.

PART VIII

OODB ADVANTAGES

In contrast to RDB, an OODBMS models real-world entities as objects and attaches operations to the data that are specific to the particular data being represented. This allows applications to directly reference data rather than setting up specific applications to interpret the data, which is the case in the relational database. The basic benefits that an OODB would bring to a PACS include the following (Routhier, 1992):

1. **Reusability** : Classes can and must be designed for reuse when designing a system. Once these classes have been created they can be instantiated and used over and over again.
2. **Reliability** : Once a class has been built and tested, then larger systems using these proven classes will tend to be more reliable than systems built from scratch in its entirety.
3. **Integrity** : Data structures can only be used with their specific methods. This feature, by its very nature, also ensures the security of the data, which is an extremely important for legal reasons.
4. **Ease of Maintenance** : When modifying a particular class, the attributes and methods contained in the class only need to be changed once. Once the changes have been made to the class, the results propagate through the rest of

the database using that particular class. This capability is available because an OODB is modular in nature.

5. A Graphical User Interface(GUI) : The Object-Oriented model leads to a graphical representation. Once an object is selected, the user can easily extract the desired information. This method of interacting with an environment, as opposed to typing in specific instructions and operations to extract specified data, is much easier. For example, a physician can point to icons rather than typing the name in of a desired file or image.

6. Images, Video, Speech and Complex Data Types : By using BLOBs to represent these complex data types, methods to interpret the BLOBs can be incorporated directly into the object. This releases the designer from the task of determining what type of application needs to be implemented in order to interpret the requested data. The user is able to view specific types of files even if they are compressed. When the object is activated, the file is uncompressed and displayed on the screen. (More information on specific compression algorithms is the other topic in this thesis).

7. Inter-operability - Different vendors may be able to use each others' classes, which helps to reduce the necessity for new code to be written.

8. Client-Server Computing : Server classes may be used by many different clients. As these classes are accessed, data integrity is assured by the methods associated with

them. For instance, a doctor may be able to access an image and manipulate it, but he is not allowed to modify the database.

9. Massively Distributed Computing : Worldwide networks employ directories of accessible objects, allowing classes in one machine to interact with classes in another. This characteristic permits an OODB to support a GPACS environment so that two doctors can access the same image, and using a pointer (seen on both screens), discuss a possible diagnosis.

10. Machine Performance : OODB has already demonstrated much higher performance than the RDB for applications involving complex data structures. Because PACS main objects are images, this type of database, along with concurrent computing using object-oriented design, promises major leaps in database performance.

11. Posing Queries : The ability to pose queries is made much easier by using a GUI. The GUI helps the user to go through steps to choose the type of information needed. The interface also allows the user to type queries without knowing what type of data, (image, text etc.) is being requested. This makes accessing information much easier by releasing the user from the need to know specific details of data and database structures.

The OODB offers new flexibility not found in the RDB. Ideally the final phase of this technology will be a person

telling the computer what type of program it needs, and the computer will design the model and write the code to implement the user's request. A list of object-oriented consultants that are currently in the business of assisting users with object-oriented technology are listed in (Newling,1993). A more detailed discussion about how some of the existing OODBs manage objects is available in (Hurson et al.,1993).

PART IX OODB DISADVANTAGES

There are several conceptual and implementation drawbacks to OODB management systems. The concept of an OODB can often be difficult to understand. The layout for any particular database can be confusing and more complex than the tabular format. From a programmer and database manager perspective, this complexity makes the job of designing and implementation more difficult. The relationships, instantiations, and inheritance occurring between objects can make it very difficult for a programmer to create and modify an OODB. If the design of the database is not done correctly, the performance can be degraded below other types of databases.

Another disadvantage of an OODB system is the lack of testing in the real world (English,1992). The relatively new OODB approach, has not proven itself to the same extent as the RDB approach. Some of the issues involved in real-world testing include maintainability, reliability, capacity, and efficiency. Figure 20 shows how society is progressing through a timeline where object-oriented databases are part of the future. This figure shows "Where We Are Today" (Martin and Odell,1991) along with the progression of the languages and databases of the past. Predictions of the future include integrated Computer Aided Software Engineering (CASE) and OODBs. The trend apparent

from this figure is that relational technology is fading out as object-oriented ideology takes over.

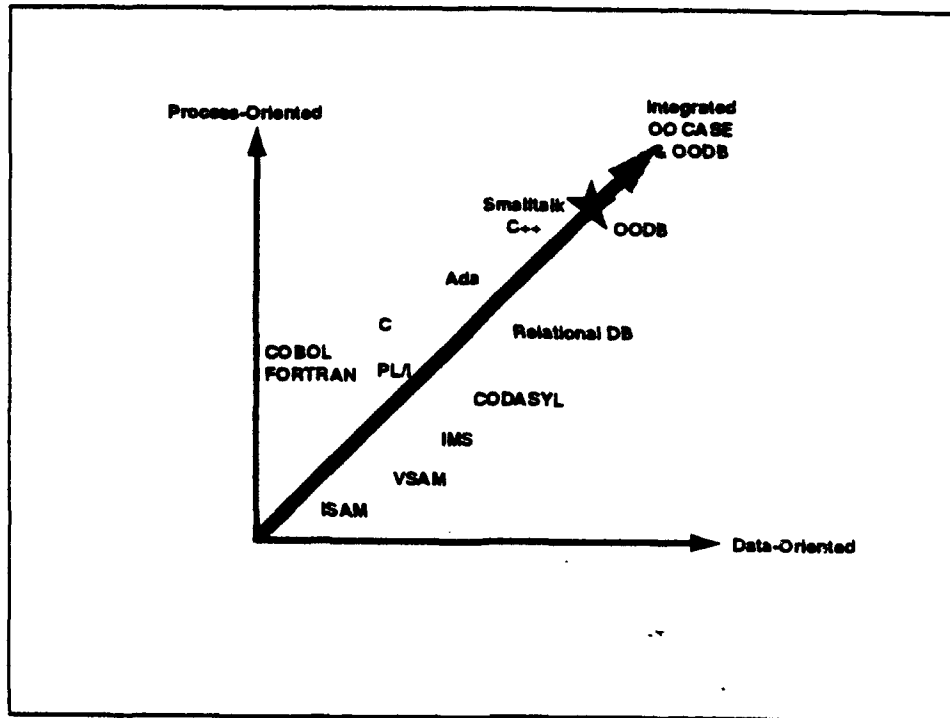


Figure 20. "Where We Are Today!"

PART X A HYBRID SOLUTION

Today's commercial database market is changing focus. It appears that no one is developing a new RDB. Rather, they are looking to integrate the RDB with the OODB. There are three basic types of databases being developed and sold today. The first is a homogeneous ODBMS, the second is a homogeneous RDBMS with objects unbundled and mapped into the database, and the third is a hybrid database system in which an object layer can be accessed from a relational database (English,1992). The homogeneous RDBMS and the homogeneous OODBMS have been discussed above. Developing a hybrid database system is an approach that attempts to retain users and buyers during the transition from an RDB to an OODB. UniSQL is an example of such a database (Finkelstein,1993:48). UniSQL has designed an object-oriented technology into a RDBMS. This feature lets the developers of the database incorporate the object-oriented features into the relational foundation at their own pace. The other benefit of this product is that if a particular database works well with the RDB, but a need to handle images occurs, the object-oriented features can then be incorporated. UniSQL uses nested tables, pointer from table to table, and associates ("registers" in their terms) a method (procedure) with any particular table to handle the complex data structures. The pointer from one table to the next incorporates the idea of inheritance. If one table

points to another, then the table being pointed to is considered to be a sub-class of its parent table. If the parent table is searched for a piece of information, so is its sub-class. To incorporate the idea of data encapsulation, UniSQL allows procedures (methods) to be associated (registered) with a table. Lastly, UniSQL has established a set of system-defined tables, called generalized large objects (GLOs), that are organized in a hierarchical relationship, using inheritance, to specifically support a variety of multimedia data types. All these features in combination provide a relational database with an object-oriented flavor. This approach appears to be an easy transitional tool for users to go from the RDB paradigm to the OODB one (Finkelstein,1993). Eventually, if UniSQL continues to develop, it will evolve into a homogeneous OODB, eliminating the idea of a mix of OODB and RDB approaches. For a more detailed description of UniSQL, (Finkelstein,1993).

PART XI CURRENT APPLICATIONS OF OODBS

Currently, there are several OODB management systems on the market or in development and near commercial release. The main applications of OODB management systems include geographic information systems (GIS), electronic computer-aided design (ECAD), real time process monitoring and control (RTPC), network management (NMD), and computer-integrated manufacturing (CIM) (English,1992).

PART XII

MIDB

A Medical Image Database (MIDB) is the database contained within a PACS. A MIDB can be based on one of the three models specified above. Presently, there are several relational MIDBs in clinical use today. The military system, called Medical Diagnostic Imaging Support (MDIS), and University of California at Los Angeles (UCLA), in the UCLA Medical Imaging Division and Clinical PACS Project, both use relational databases as their MIDB. On the object-oriented side, although sites are non-existent, the NRV-PACS group in France has proposed an object-oriented model for a MIDB. The NRV-PACS group's model has been based on the following requirements: "the server should be able to manage not only flat data, such as numerals, character strings, and dates, but also: 1) structured data such as multidimensional arrays or data sets; 2) semantics such as data set structure, data links into a data set or between data sets... 3) retrieval processes that involves criteria on the image environment and on the data set structure, and leads to data set selection."

PART XIII DISTRIBUTED & CENTRALIZED DATABASES

Other important questions that arise are the implementation of either a centralized or a distributed database, whether or not a homogenous or heterogeneous database should be used with these architectures, and how the databases should be coupled together. To answer these questions, a look at the composition of each type of implementation is discussed.

A centralized database is a database that is located in one machine. This means that it controls all communications and queries, and it holds all the relative database information. Only one type of database can be associated with this system (RDB , OODB or a hybrid). A distributed database refers to either a database that has different parts stored on separate machines, or it can also refer to the idea of linking together different databases (all the same type or a mix) to form one larger hybrid database. Notice that the possibility of having either a centralized or a distributed database exists for both the inter and intra-hospital cases as seen in Figure 1. The intra-hospital side of the figure shows the possible internal construction of a hospital's database. The inter-hospital side of the figure shows the possible structure for a global database where either a centralized repository for all medical images exists, compared to the other side of the tree in which the global system consists of distributed

databases (the intra-hospital systems), where several sites in the system have their own independent databases.

The centralized database is the traditional PACS model, and this approach offers several advantages such as consistency, integrity, and relative simplicity of implementation. These advantages stem from the fact that there is only one database. This database does not have to address consistency issues, such as avoiding communication and update errors, that arise when several machines manage a database. The placement of all the information in one machine simplifies the problem of locating the data. The centralized database works well for smaller database systems. Unfortunately, as the size of the database increases, so do the problems associated with this approach. As the amount of data increases, so does the workload including organization, retrieval time, and the number of queries. Since there is an increased workload, the response time increases. As the database becomes very popular, the centralized database computer eventually becomes a bottleneck in the data request pipeline. Along with the disadvantage of increased response time comes the increased dependency placed on the database machine because of a single point of failure for the system. If for some reason the database gets overwhelmed or breaks down, all information becomes inaccessible. To handle this problem, a back up system is needed. The centralized database for a

GPACS would have to be enormous in order to handle a collection of all the images in the world. Because of these problems and expenses, a GPACS based on a centralized system is impractical. To address the shortcomings of the centralized system, a distributed database must be used.

As discussed previously, there are basically two definitions of a distributed database: a homogeneous system and a hybrid system. The homogeneous system is one in which a database is distributed among several machines that all use the same architecture. Figure 21 shows a typical distributed architecture for a homogeneous Distributed DBMS (Bell and Grimson, 1992). This architecture takes the approach of storing the images across many sites on a network. This is commonly referred to as fragmenting the database. For example, once an image is created, it is stored at the location where it will be utilized the most. This idea of intelligent storing decreases network traffic along with decreasing retrieval time. The location of all the data is managed by the system permitting an image to be requested at any one of the stations (Allen and Frieder, 1992). Several algorithms exist for managing the location of images in memory, the most simple a table of images designated by the system. If this is the case, then when a site needs an image, it references the image location table, called an address book, and finds out its location, then retrieves the desired image.

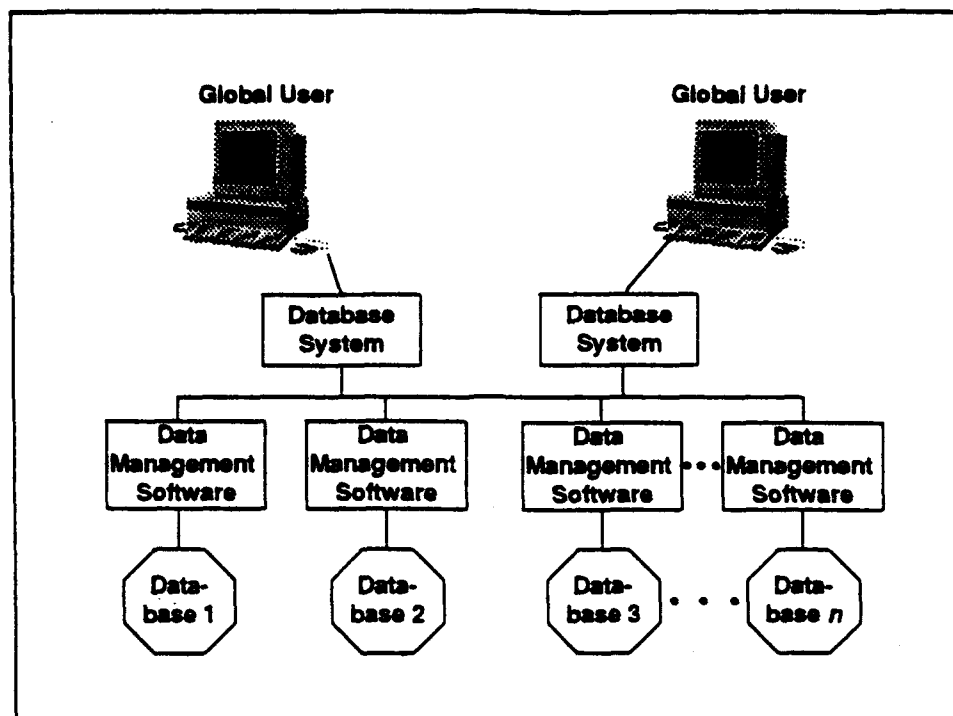


Figure 21. Homogeneous Distributed DBMS Architecture

Another benefit of fragmenting is increased reliability. In the case of software or hardware failure, only the fragment of the database located where the failure occurs will be affected. The other fragments continue to operate normally without being able to retrieve data from the malfunctioning fragment. If the computer containing the address book fails, then only images located locally can be retrieved by the distributed systems. Figure 22 shows how all independently operating parts of the distributed database rely on the address book for image location information. If communication fails between a site and the address book, only images at that site can be referenced since the remote location is unknown. With this in mind, a

certain amount of redundancy is necessary. To insure that images are available in emergency situations, all images that are critical in nature should be stored not only in the requesting departments' database but also in the memory of the unit where the patient is physically located. That way, if the address book fails or the network is down, the image is still available on location.

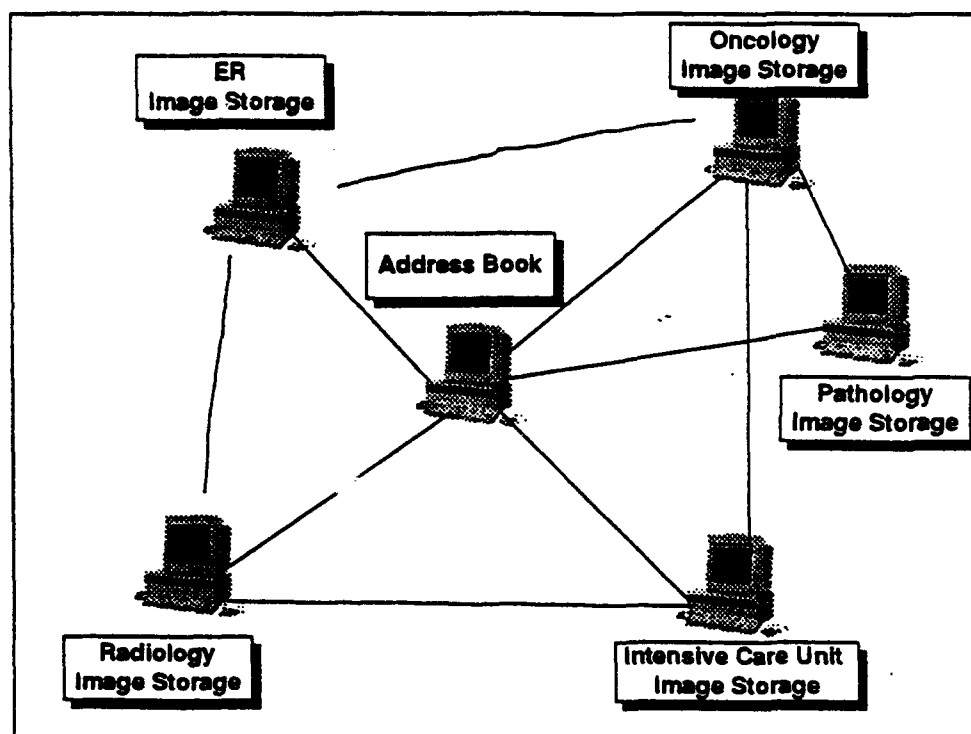


Figure 22. Address Book Dependence

The homogeneous distributed approach using fragmentation is more cost effective than trying to develop the whole system at one time (Allen and Frieder, 1992). This cost savings results from the use of the incremental approach. For example, if a hospital can only buy enough

hardware to outfit radiology and the emergency room, then a distributed system is usually the only option. The distributed system allows the setup of the departments receiving the equipment while other departments wait until they have funds available. Once the other departments attain the required equipment, they should, if planned correctly, be able to connect into the system.

The hybrid distributed database is similar in many ways to a homogeneous distributed database with the exception that the fragments of this system are autonomous. Both systems have fragmented databases and use the same data storage and retrieval strategies. The hybrid, however, connects different types of databases so that the user is unaware that different systems are being used. A certain fragment of the hybrid database could be an OODB, a RDB, an OODB/RDB combination or even perhaps a homogeneous distributed database. The idea of interconnecting different types of databases supports the concept of the GPACS. The development of the wide area networks (WAN) are forcing developers to address many of the problems associated with the interaction of all the different type of systems that exist today. The standardization issues that are still being resolved will play a key role in making a hybrid distributed database a reality.

The last distributed database design question is whether or not the distributed database architecture should

be loosely coupled or tightly coupled. In Figures 23 and 24, the difference between the two couplings can be seen (Bell and Grimson, 1992). The tightly coupled multi-database system forces the global user to access the databases through a global schema (Figure 23). The global schema is the structure of all the cooperative databases made available to global users. This allows a participating database to control the part available to the global schema. The global user can only access the information available in the global schema. The capability to have information available to the global users while still having areas of the system exclusively for local users all under local control is known as local autonomy. The other way to implement a multi-database system is loosely coupled as in Figure 24. This allows the global user to establish a link directly with each database, allowing the remote users the same privileges as local users. Since global networks are still developing, much of the distributed database potential has yet to be realized.

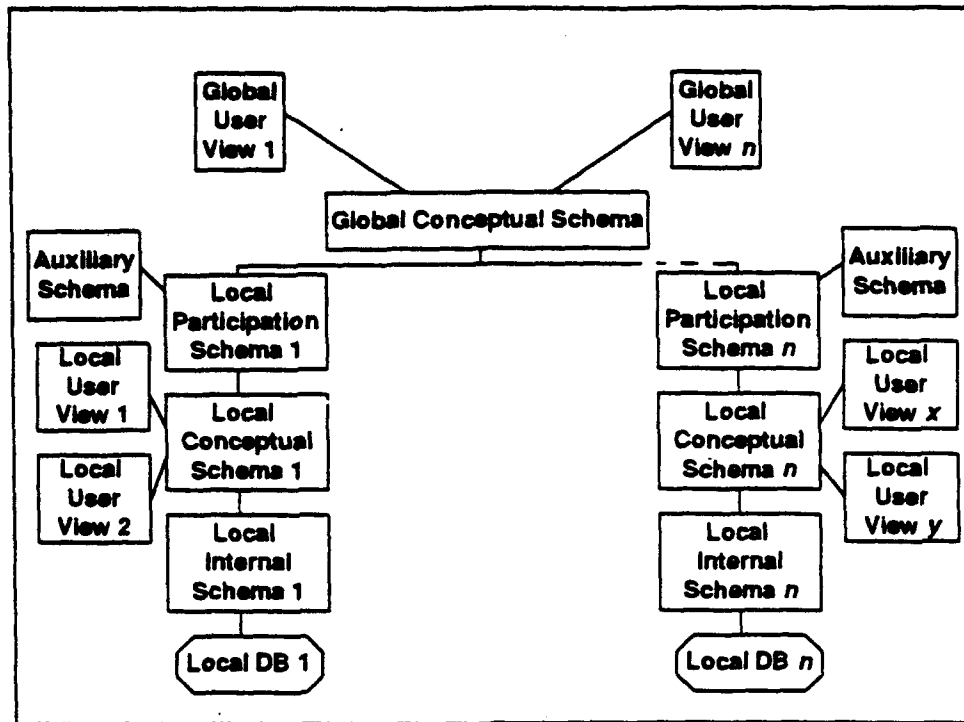


Figure 23. Tightly-Coupled Multi-Database System

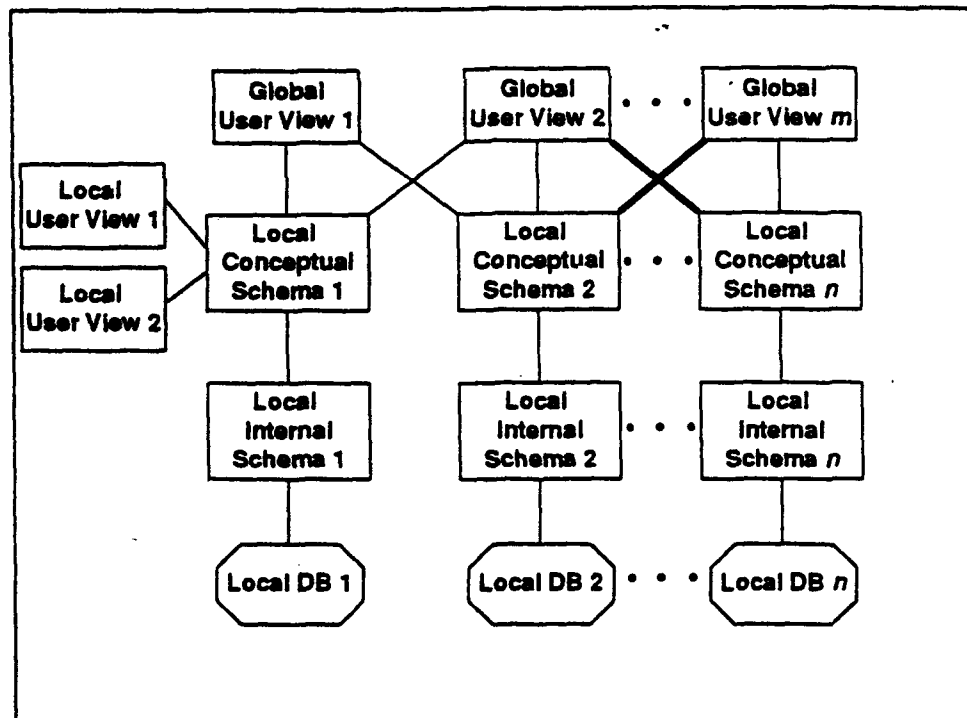


Figure 24. Loosely-Couple Multi-Database System

PART XIV DATABASE SECURITY

In a PACS, as with any other computer that holds sensitive information, security is of great concern. The required local security, that is security for a particular localized database, is usually provided by the local database rather than being implemented by an external source. In this case, when looking for a system, an emphasis should be placed on the capabilities of the database. The main problem occurs when a PACS starts letting outside sources query its information. A doctor may request information about a patient from a hospital he has never interacted with before. The problem with this is whether or not to enforce the security at a local or global level. At a local level, a system needs to be able to determine user authorization. This can be done by allowing the user to request permission, get a password, and assign permission in that fashion. At a global level, the issue becomes more complicated. If a global network is set up with an administrator, then a centralized list of all the possible users would have to be established or a security hierarchy must be set up. The other option is to enforce the security locally and each time a remote inquiry is made, a confirmation by phone would be required. Either way, this issue must be considered when interconnecting PACS. For more information on database security refer to (Lunt and Fernandez, 1991).

PART XV THE PACS FUTURE

Currently, PACS are functional utilizing RDBs. The RDBs are in place today because they met the initial requirements imposed by a PACS. As the idea of a GPACS emerges, users are beginning to realize that the RDBs of the past will not meet the needs of the future. To meet future demands, a new paradigm has appeared, the object-oriented paradigm. This new paradigm offers a new way to organize and retrieve all types of information. OODB management systems are becoming more popular and better accepted, as evidenced by the large number of object-oriented systems appearing on the market. The hybrid databases will make the transition from one model to the next significantly easier, allowing people to experiment with object-oriented features without the cost of a complete redesign. In the medical community, pure and hybrid OODBs will be used extensively in the research areas of hospitals. The properties of identifying objects should appeal to physicians wanting to identify organs of the body as objects. This identification would allow them to isolate a particular organ to determine its problem, thereby offering a better informed diagnosis. The object-oriented paradigm will also be used for image interpretation and manipulation. It is in this area that object-oriented concepts will be initially used the most. Once object-oriented databases mature, the features of

object manipulation and retrieval can be exploited further by using them in distributed systems such as GPACS.

The global database concept is slowly becoming a reality with new distributed databases appearing everyday. These databases will pave the road to global networks and databases. In the next 10 years, the establishment of world-wide, high-speed, high-volume networks will open the door to other global systems such as GPACS. The ability to transfer large amounts of information will allow real-time transfer of images, video, and other high bandwidth applications. Also, the further development of PACS will hopefully lead to standards and lower cost systems.

CHAPTER 5

IMAGE AND DATA COMPRESSION IN THE PACS DOMAIN

A PACS requires the manipulation of large amounts of data in many different forms. The database of choice decides where to store the information and who to send it to, but how to transport and store the data using a reduced capacity is the job of a compression algorithm. This section addresses the compression issues and offers possible solutions.

The manipulation of large amounts of data, especially in the form of medical images, can often decrease the efficiency of a PACS by slowing down transmission time and consuming large amounts of disk storage space. To address the decreased efficiency due to the size of data, compression techniques are employed. Compression of data reduces the size of the data during transmission and storage. When it is time to use the data, an inverse compression algorithm is used to return the data to its original form. The principle behind compression is the basic idea of entropy. Entropy is a measure of randomness. All data has a certain entropy value based on its randomness. The entropy for any particular data is different and basically unknown. The assumption is once you reduce or

compress data until it reaches its entropy, then it can be reduced it no further. Since entropy is an immeasurable value, it is impossible know if the data is ever fully compressed. The notion of entropy assumes the following intuitively reasonable facts: 1) Random data cannot be compressed, 2) Data compressed by an optimal compressor (one that achieves the data's entropy) cannot be compressed further and 3) One cannot guarantee that a data compressor will achieve any given performance on all data (Storer,1988:7). The following section introduces different types of compression along with several algorithms that are used to compress different types of data in a PACS environment. A list of key terms for compression can be found in Table 13.

The following section is based on the following sources: (Abramson and Kransner,1989), (Cockroft and Hourvitz,1991), (Cormack,1985), (Held,1983), (Ho et al.,1993), (Horspool,1991), (Howard and Vitter,1991), (Jackson and Szasz,1992), (Kajiwara,1992), (Kao et al.,1992), (Klien and Carney,1991), (Lin and Lu,1991), (Lynch,1985), (Manduca,1992,1992), (Moll et al.,1992), (Perl et al.,1991), (Turner and Peterson,1992), (Pronios and Yovanof,1991), (Resnikoff,1992), (Sayre et al.,1992), (Seiter,1991,1992), (Suarez,1991), (Zip and Lempel,1977,1978), (Blinn,1993), (Ho et al.,1989), (Chen and Flynn,1992), (Lee et al.,1992), (Leehan et al.,1992), (Lo et

al.,1992), (MacLeod,1991), (McNitt-Gray et al.,1992), (Peterson et al.,1991), (Saito and Kudo,1989), (Stern,1991), (Urbano et al.,1992), (Roth and Van Horn,1993), (Huffman,1952), (Storer,1988), (Lelewer and Hirschberg,1987) and (Sedgewick,1983). For a more general description of compression see (Storer,1988), (Lelewer and Hirschberg,1987), (Lynch,1985), (Sedgewick,1983), (Owen,1982), (Held,1983) and (Cormack,1985).

Image and Data compression is a technology that is beginning to be applied to many areas involving computers. In the PACS domain, image compression is vital because it reduces storage requirements and speeds up data transmission. For example, in a DBMS, one of the problems that affects database performance is the storage of large amounts of data. Compressing that data before it is stored decreases the amount of storage space required, thereby decreasing the time necessary to store the data. The image retrieval time in a PACS can also be reduced significantly by reducing the size of an image using compression. Once the image is compressed, it is sent through the network to its destination where it is uncompressed and used normally. This idea of reducing the size of the data is the cornerstone principle in understanding data compression.

The first advantage associated with compression is the reduction of actual physical storage requirements. In the case of PACS, its job is storing images for short and long

term use. Therefore, if stored images can be compressed, there is a reduction of required storage. This savings results in a PACS that can store a greater number of images.

The second advantage is that compression plays a significant role in the transmission of images. The amount of bandwidth required to send a medical image can be rather large. Because the transmission time increases as the size of the file increases, medical images take a relatively long time to transmit. Although networks are attempting to meet future needs by increasing their bandwidth capacity, they are currently lacking, especially in long-haul communication. Also in the future, as video conferencing is introduced, the network traffic will continue to increase, once again degrading the timely transmission of information. Compression decreases the size of the files. Once the data is compressed, the bandwidth requirement is decreased, making the transmission time considerably less. Also, if the data is compressed using certain compression algorithms, that are discussed later, the data becomes more resistant to transmission errors (Pronios and Yovanof, 1991).

The benefits of compression not only include saving storage space and transmission time but also inherently lead to the security of the data being compressed. If the exact compression method is unknown, it can be very difficult to decode the compressed information without significant analysis by a crypto-analyst or a decoding computer. This

principle is demonstrated by Figure 25, part three which is the text substitution encoded message, "#@'s_image_w&#_same_&#_o#r@s." . Figuring out this message encoded by a simple compression method, although it can be done, is very difficult to do without a key. This difficulty increases as the complexity of the compression algorithm increases. The coded(compressed) information provides security as a desirable side effect. This example uses the method described in (Storer,1988:12).

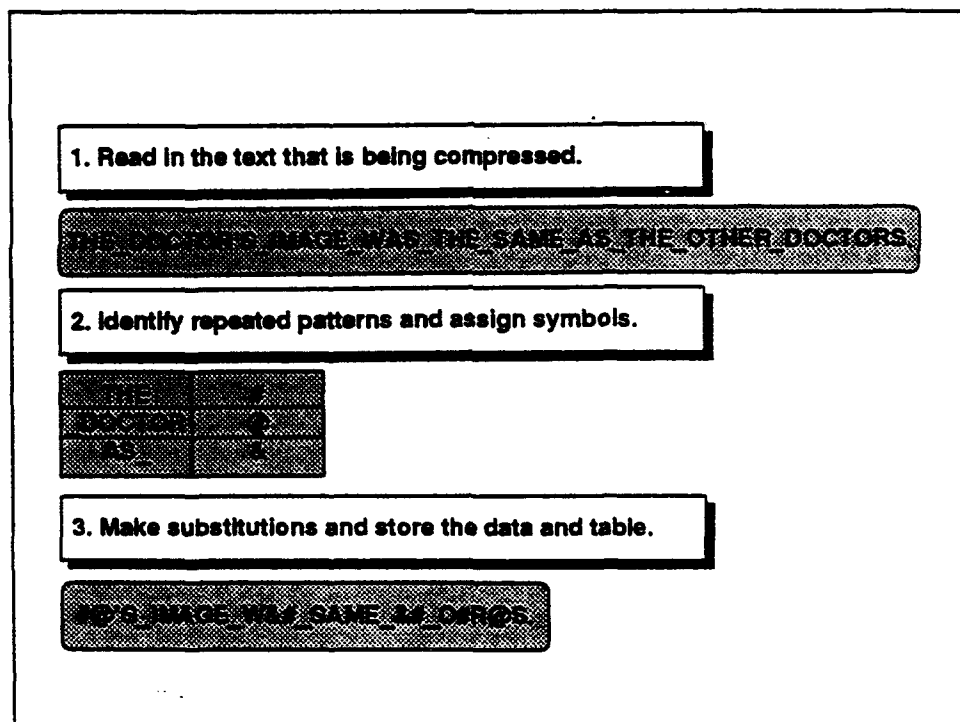


Figure 25. Textual Substitution

Progressive image transfer is a method of quickly getting basic information about an image initially, such as the shape of the image, then, over time, the empty holes in the first rough image, are filled in displaying the full

image in its entirety. In this scheme, the data is extracted evenly from the entire image rather than from a continuous array or block of pixels. This allows the loss of network packets to minimally affect the overall image because the information lost occurs at random points throughout the image. For example, if a 512 x 512 pixel image is sent across a network and a packet is lost, the final image would have missing pixels scattered throughout the image. This method can also quickly display a rough estimate of an image by displaying the packet data as it arrives. While the user is looking at the low resolution image, more of the image information is transferred. As the information is added to the image the image quality increases. This continues until all of the information in the image has been transferred. An example of progressive image transfer is the idea of browsing through low resolution images (less data transferred) until, lets say, a specific body part is identified. By spending more time looking at the identified image, a higher quality image is achieved (more transmissions are arriving as the image is viewed). This process continues until the whole image is transmitted. Progressive image transfer can be used in a WAN as well as a LAN environment to allow external users to quickly browse through images until they find what they are looking for, then as they spend more time looking at any particular image, the quality increases right before their

eyes. This process can be used to prevent unnecessary time waiting for each image to be fully transferred before moving onto the next image. For example, if a doctor has a patient with ten images that say chest x-ray, with the actual one he wants in the sixth position, he can quickly look at the first five possibilities (5 seconds each) and stop on the sixth image that is then transferred fully (2 minutes).

The advantages of compressing also lead to increasing the performance of many applications such as databases, digital phone communications, and have even been utilized in the newest versions of operating systems such as Disk Operating System (DOS).

The term compression ratio(CR) is used to describe the amount of compression that occurs and is defined as :

$$CR = \frac{\text{Data_Size_Before_Compression}}{\text{Data_Size_After_Compression}} \quad (1)$$

The amount of reduction attained by a compression algorithm is given by:

$$\text{Amount of Data Reduction} = (1 - \frac{1}{CR}) \times 100\% \quad (2)$$

These two equations can be used to assess the performance of different compression techniques (Roth and Van Horn, 1993:55).

There are two basic categories of algorithms used to describe image compression, lossless and lossy. The term

lossless describes algorithms that compress and uncompress the image without **any** loss of information. The lossy type of algorithm compresses and uncompresses the image resulting in a new image. The new image may appear to be identical to the original image when viewed by the human eye, but the image has actually changed. Lossless compression is a compression algorithm wherein the data is compressed in a way that allows it to be returned to its original state. Using the following information:

Basis

a = Original Data

C(x) = Compression function

U(x) = Inverse compression function

Lossless compression is represented as:

$$U(C(a)) = a \quad (3)$$

while **lossy** compression algorithms are represented as:

$$U(C(a)) = b \quad (4)$$

where $b \neq a$.

As seen above, the lossless compression takes the original data "a", compresses it, applying the inverse compression function to the compressed data, recovers the data exactly. The lossy compression goes through the same process, but rather than retrieving the exact data, only a close approximation is recovered.

The use of lossy algorithms is a controversial topic in the medical field because of the possibility that important

information might be eliminated. The reason for the continued interest in lossy algorithms is its ability to attain compression ratios up to 50:1 CR with minimal loss, while their counterparts, the lossless compression algorithms, can only attain around a CR of 5:1. The quality of the highly compressed lossy images is reduced significantly. Deciding what is an acceptable amount of compression for lossy algorithms is a continuing area of research while the search for better lossless algorithms moves forward.

Before further discussion of lossless vs lossy compression, the reader should understand that different compression algorithms meet different needs. Most algorithms being developed today target one type of data in order to exploit the attributes of that particular data. An example that highlights the difference in techniques is the comparison of text data versus image data. Text data can be compressed by substituting symbols for frequently used character strings that are larger in size compared to the symbol being used (See Figure 25). In this figure, strings that are frequently repeated in the text are assigned a symbol. These symbols are stored in the symbol lookup table. This table is used to encode and decode the message. For example, in Figure 25, the string "the" is replaced with "#", so that whenever "the" appears in the text a "#" is substituted to generate the compressed data. The result is

a shorter string of characters accompanied by a lookup table. When the data is uncompressed, the process is reversed by putting in the appropriate strings wherever the symbols occur, revealing the original text.

In contrast, an image may get compressed by concentrating on the color of each pixel, once again substituting a symbol for more commonly used colors (See Figure 26). In this example, the image consists of a 6 x 6 block of pixels (36 pixels). Each pixel is stored as a 24 bit value (8 bits for red, 8 bits for green and 8 bits for blue). Since this image only consists of two colors, red and white, we can store each color as a one bit value where red is represented by a "1" and white is represented by a "0". In this very specialized example, the image is reduced from a total of 864 bits to a total of 36 bits and a symbol table. Just looking at the image part of the reduction, excluding the symbol table, we have obtained a CR of 24:1. Even though this is an extreme case, it shows how knowledge of the data being compressed can aid in finding an optimal compression algorithm.

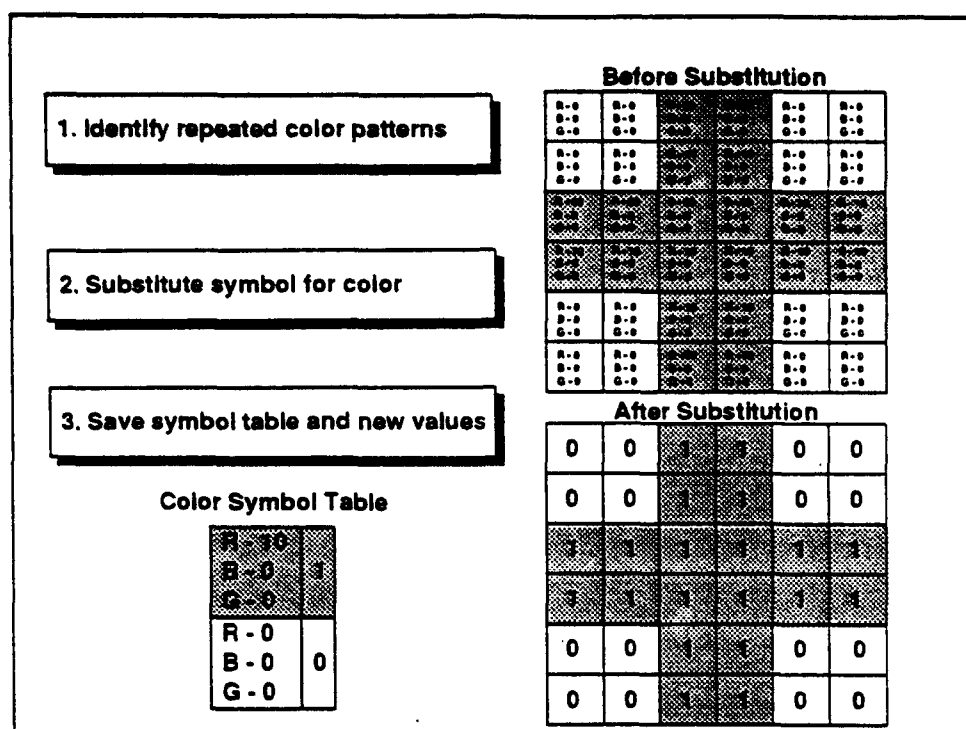


Figure 26. Color-Symbol Substitution

Obviously text and image compression techniques are only two out of many compression techniques specific to different types of data, but it shows how the focus of a particular algorithm changes. Along with these compression algorithms based on data content, there are also generic compression algorithms that can be used on any type of data. These algorithms may be based on the same type of schema as other specialized algorithms except that they have no prior knowledge of the data to be compressed and generally do not result in high CRs.

The next few paragraphs discuss the basic types of compression algorithms, both lossless and lossy, presenting commonly used algorithms in each area. The last part of

this section concentrates on the different uses of compression in the PACS arena.

The list of lossless compression algorithms include Huffman coding, the Lempel-Ziv or the Lempel-Ziv-Welch (LZW) algorithm, arithmetic coding, abbreviation, null suppression, run-length encoding, pattern substitution, differential compression, and restricted variability codes. The most commonly used lossless compression algorithms are Huffman coding, the LZW algorithm and arithmetic coding. These algorithms are used most often, due to the fact that they work well and are easy to implement. For the following discussions, the terms code words and symbols are used interchangeably because a code word is a binary representation of a symbol.

PART I HUFFMAN CODING

The basic principles used in the text and image compression described above are rooted in Huffman encoding. Both of those examples were very specialized versions of Huffman encoding while the basic principle is better illustrated in Figure 27 (Pronios and Yovanof, 1991). Huffman coding assigns code words of short lengths to the input characters with the highest probability of occurrence and longer code words to lower probability characters. The principle of converting fixed-sized data into variable-length symbols requires the generation of a look-up table to store the values of each symbol used. This overhead is diminished by symbols which reduce the data by a factor that is greater than the size of the table. To decode this type of compressed file, each symbol is replaced by the original data to render the data just as it was before being compressed. Figure 27 shows the actual bit compression techniques inherent to Huffman coding. In the figure, the letter 'a' has the highest probability of occurrence (0.35). Therefore this letter is replaced with a symbol comprised of less than 3 bits (length of all letter representations before compression). Huffman coding (Huffman, 1952) uses a tree structure to represent the code book. To use the code book to determine the code word for any particular letter, follow the path from the top of the tree down to the node containing the desired letter. The code word results by

identifying the value of each path (in this case either a 1 or 0). For more information on Huffman coding see (Huffman,1952) and (Storer,1988).

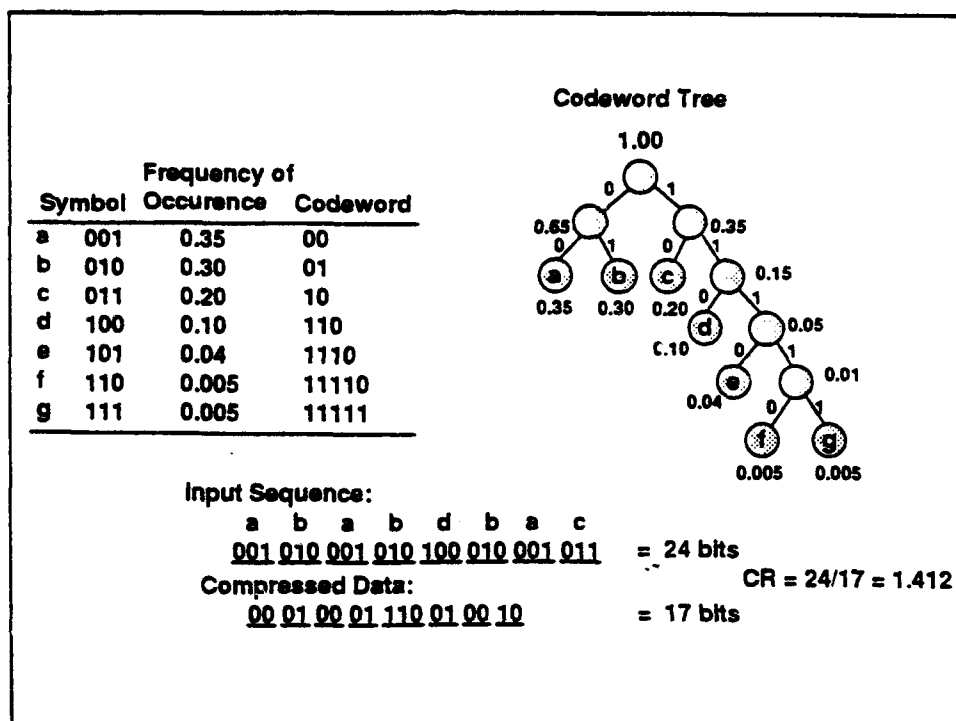


Figure 27. Example of Huffman Coding

PART II LZW COMPRESSION

The LZW compression algorithm (Zip and Lempel, 1977, 1978) (Pronios and Yovanof, 1991), attacks the problem of compressing the data by changing variable length strings into fixed-length symbols. The process creates a unique dictionary (code book) for every different set of information being compressed. The first step in the process is to initialize every element in the alphabet with an associated code and put these assignments into the dictionary. This initialization occurs in both encoding and the decoding. The algorithm then scans the input to the encoder and matches it up with codes already in the dictionary. If no match is discovered, that string is added to the dictionary. The dictionary continues to expand, capturing different size strings and assigning them a fixed-sized number. The following algorithm (Horspool, 1991), shows the process for a basic LZW implementation.

PSEUDO CODE

```
for (i in the range 0 to 3)
    add i as a one character string to the dictionary;
                                                    {This initializes the dictionary}

add the empty string,  $\lambda$ , to the dictionary;

sn = string number of  $\lambda$ ;
                                                    {sn is initially 4 in the example}
while (input remains {
                                                    {Loop until the end of string}
                                                    {(ababdbac) loops 8 times}
    read ( ch );
                                                    {reads in the next character}
    if (<<sn,ch>> is in the dictionary )
                                                    {Checks new string}
        sn = string number of <<sn,ch>>; {If there it identifies the location&}
                                                    {tries adding another character}
    else {
                                                    {if new string not in dictionary}
        write_number (sn);
                                                    {write number of old string}
        if (dictionary is not full)
                                                    {add new string to dictionary}
            add <<sn,ch>> to next position in dictionary;
        sn = string number of <<ch>>; {assign sn to value of last char}
    }
}
write_number (sn);
                                                    {write out last value}
```

Figure 28 (Pronios and Yovanof, 1991), accompanied by Tables 14(a & b), shows an example of a step by step implementation of compressing the string "ababdbac" using the above algorithm. The following notation, <<s,x>>, is used to represent the string formed by appending character x to the string with the number s in the dictionary.

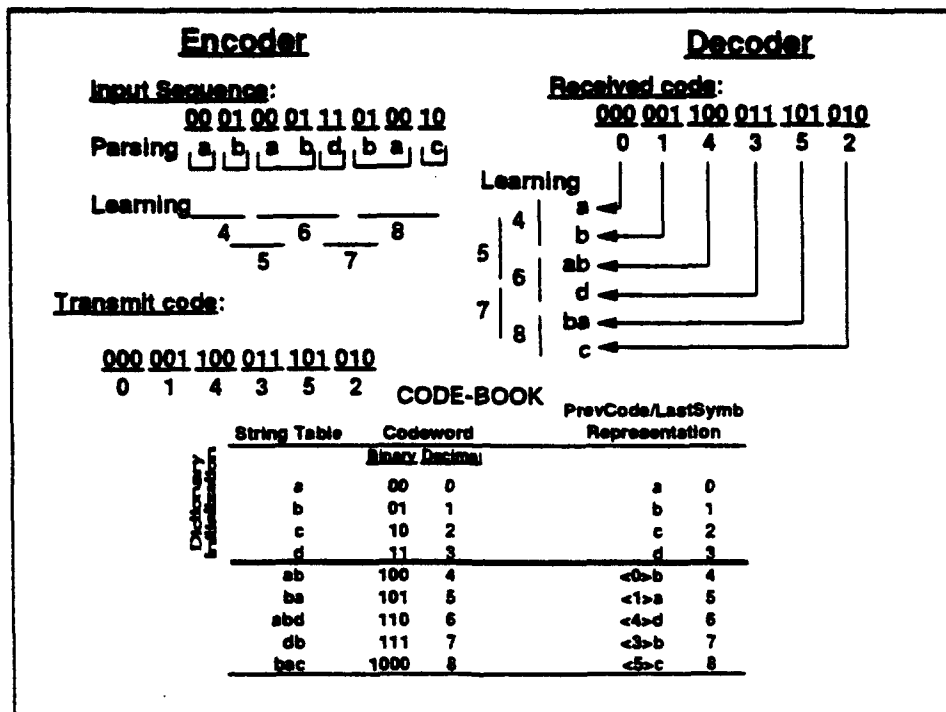


Figure 28. Example of LZW Coding

Table 14(a). LZW Implementation

Read ch	<<sn,ch>>	sn	output	Dictionary
initialize	λ	4	-	000 0 a 001 1 b 010 2 c 011 3 d 100 4 λ
a	a	0	-	000 0 a 001 1 b 010 2 c 011 3 d 100 4 λ
b	ab	1	000	000 0 a 001 1 b 010 2 c 011 3 d 100 4 ab
a	ba	0	001	000 0 a 001 1 b 010 2 c 011 3 d 100 4 ab 101 5 ba
b	ab	4	-	000 0 a 001 1 b 010 2 c 011 3 d 100 4 ab 101 5 ba
d	abd	3	100	000 0 a 001 1 b 010 2 c 011 3 d 100 4 ab 101 5 ba 110 6 abd

Table 14(b). LZW Implementation

Read ch	<<sn,ch>>	sn	output	Dictionary		
b	db	1	101	000	0	a
				001	1	b
				010	2	c
				011	3	d
				100	4	ab
				101	5	ba
				110	6	abd
				111	7	db
a	ba	5	-	000	0	a
				001	1	b
				010	2	c
				011	3	d
				100	4	ab
				101	5	ba
				110	6	abd
				111	7	db
c	bac	2	101	000	0	a
				001	1	b
				010	2	c
				011	3	d
				100	4	ab
				101	5	ba
				110	6	abd
				111	7	db
				1000	8	bac
λ	-	2	010	000	0	a
				001	1	b
				010	2	c
				011	3	d
				100	4	ab
				101	5	ba
				110	6	abd
				111	7	db
				1000	8	bac

The output, shown in Table 14(a & b), is the transmit code. The transmit code is sent alone without the dictionary. The decoder receives the transmitted code and builds its own dictionary as the data is received, reversing the process.

Although this method works well for strings of text, it does not carry over easily to lossless image compression

because images are two-dimensional. This extra dimension makes the data more complex than one dimensional text. Because images are essentially quantitized analog data, exact matches in the data, needed for high-order modeling are fairly rare (Howard and Vitter,1991).

PART III ARITHMETIC CODING

Arithmetic coding is a relatively new compression technique. This algorithm uses mathematical equations to send information. The basic strategy for this technique is to assign a probability of occurrence to each member in an alphabet such that all the probabilities add up to one. In textual data the alphabet would be the set of ASCII characters. For the following example, using the method described by (Roth and Vanhorn,1993:57), the alphabet includes {a,b,c,d,e,f,g} with the probability range defined as :

Element	Probability	Assigned Range
a	0.35	[0.0, 0.35)
b	0.30	[0.35, 0.65)
c	0.20	[0.65, 0.85)
d	0.10	[0.85, 0.95)
e	0.04	[0.95, 0.99)
f	0.005	[0.99,0.995)
g	0.005	[0.995,1.000)

To compress the file containing "ababdbac", begin with the interval [0.0, 1.0]. In order to create an interval that represents our data, the following equations are required (Roth and Vanhorn,1993:57):

$R = \text{Range}, \quad \text{High Value of Range} = \text{HV}, \quad \text{Low Value of Range} = \text{LV},$

$\text{High Value of a particular Element} = \text{EHV},$

Low Value of a particular Element=ELV

Initial HV = 1.0

Initial LV = 0.0

$R = HV - LV$ (5)

$HV = LV + (R \times EHV)$ (6)

$LV = LV + (R \times ELV)$ (7)

These variables and equations are specifically looking at the intervals generated throughout this process. For example the range(R) refers the highest value(HV) in the current interval minus the lowest value(LV) in the current interval. The high value of a particular element (EHV) is a constant based on the assigned range of the element being processed. For the first element "a" in our file of "ababdbac", the EHV is equal to 0.35 and the low value of this element (ELV) is 0.0. Therefore, the information is encoded into the interval. The encoding proceeds as follows:

Start	[0.0, 1.0)
a	[0.0, 0.35)
b	[0.1225, 0.2275)
a	[0.1225, 0.15925)
b	[0.135362, 0.146387)
d	[0.144734, 0.145836)
b	[0.14512, 0.14545)
a	[0.14512, 0.145235)
c	[0.145195, 0.145218)

To decode this interval, knowledge of the ranges of the elements must be known. The decoding equations (Roth and Vanhorn, 1993:57) are :

$$\text{New HV} = \frac{HV - ELV}{EHV - ELV} \quad (8)$$

$$\text{New LV} = \frac{LV - ELV}{EHV - ELV} \quad (9)$$

The initial range sent is (0.145195, 0.145218), the next step is to determine the element whose range surrounds the sent range. The decoding of the interval proceeds as follows:

Current Range	Character Extracted	Range after extracting the character
[0.145195, 0.145218)	a	[0.414842, 0.414909)
[0.414842, 0.414909)	b	[0.216142, 0.216362)
[0.216142, 0.216362)	a	[0.617547, 0.618177)
[0.617547, 0.618177)	b	[0.891825, 0.893925)
[0.891825, 0.893925)	d	[0.41825, 0.43925)
[0.41825, 0.43925)	b	[0.2275, 0.2975)
[0.2275, 0.2975)	a	[0.65, 0.85)
[0.65, 0.85)	c	[0.0, 1.0)

This method could also be applied to other types of data with the alphabet being appropriate to the kind of data being compressed. The interval gets smaller and smaller, requiring additional bits to maintain sufficient precision required to reconstruct the original data. The number of

bits required for this precision however, is not nearly the size of the original data. The drawback associated with this method is the number of floating point operations that are required.

Lossless algorithms are the main focus for many applications, but the main focus for medical images are the lossy algorithms. The lossy algorithms are better suited to medical images for storage and transmission because of their higher compression ratios. The higher CR is necessary because the average size of an image is 10 MB. With this large amount of data, transmission and storage can be difficult and expensive problems. Lossy algorithms fall into basically three categories, Pulse Code Modulation (PCM), Predictive coding and Transform coding (Pronios and Yovanof,1991).

PART IV PCM CODING

PCM is the simplest form of coding an image and dates back to 1939 when it was patented by Sir Alec Reeves (Owen,1982). PCM is a method used to represent analog image data in a discrete space with a discrete amplitude. This is accomplished by sampling the image signal, quantizing each sample, and then binary coding the sample for transmission. This effectively reduces the amount of data in the analog image. The reduction occurs because the digital image data results from only a sample of the analog signal. The lossy effect occurs because not all of the analog signal is captured, therefore slight variations in the image data can be missed. Figure 29 (Owen,1982) depicts an analog signal that will be compressed. The specific data points used (sampled) to create the digital signal are labeled. Figure 30 (Owen,1982) shows the digital encoding table with the bit and signal representation for Figure 29 after applying the PCM compression.

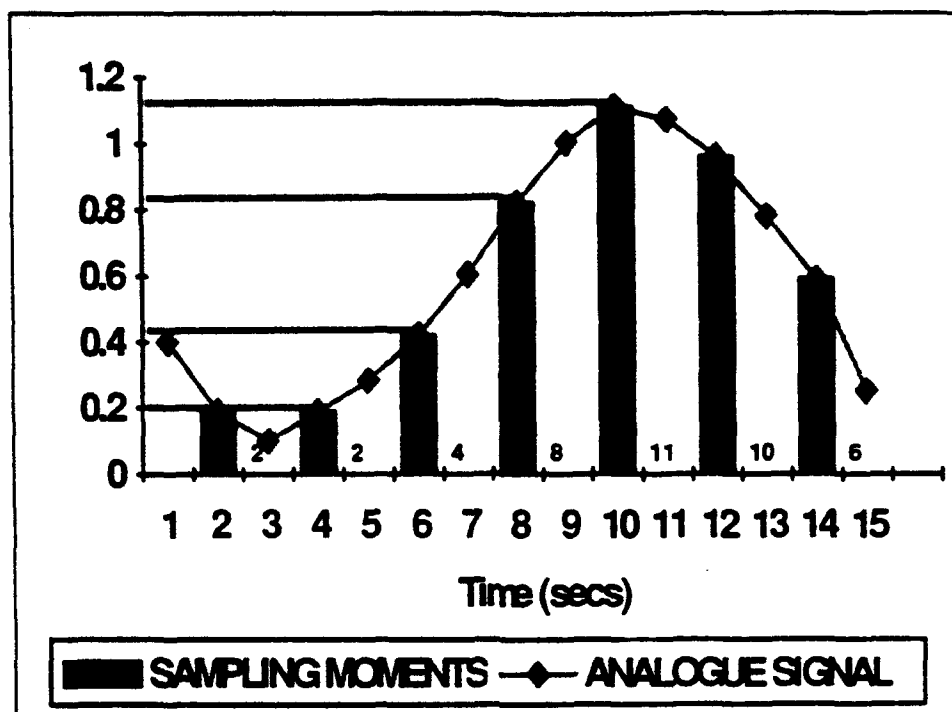


Figure 29. Analog Signal being Digitally Sampled

Amplitude Value	Binary coded equivalent	Pulse code modulated signal
1	0000	
2	0001	
3	0010	
4	0011	
5	0100	
6	0101	
7	0110	
8	0111	
9	1000	
10	1001	
11	1010	
12	1011	
13	1100	
14	1101	
15	1110	
16	1111	

Decimal values : 2 , 2 , 4 , 8 , 11 , 10 , 6
 Binary values : 0001, 0001, 0011, 0111, 1010, 1001, 0101
 PCM signal :

Figure 30. Reference Table for PCM

The PCM process occurs as follows:

Step 1: Sample the analog data at regular intervals. (In the example we sample every 2 seconds) and temporarily store this information. (Figure 29)

Step 2: Use the designated signals contained in Figure 30 and convert the amplitude value into a binary equivalent.

Step 3: Once the binary code is attained, a pulse code modulated signal is associated with each binary value. This signal is a digital approximation of the analog signal.

Step 4: To decode the PCM signal, the binary values are attained by reversing the process in step 3.

Step 5: Using the amplitude values attained from the binary values, an approximation of the actual analog signal is constructed.

The result after decompression is an approximation of the original analog signal. The main use of PCM has been in the digitization scheme for storage and transmission of an image, along with digitizing an image before applying another compression algorithm to the image data. The problem that occurs with just PCM is banding in the decompressed image. Banding is a side affect of PCM because the data is sampled, and the transition from one pixel to the next has to be approximated when uncompressing the image. When the sampling occurs at regular intervals, the missed data appears as bands in the uncompressed image. This effect can be countered adding noise which is random

data that can help smooth the transition from one sample to the next.

Predictive coding takes advantage of the statistical dependencies or redundancies from one image sample to the next. Predictive coding is comprised of two stages. The first is a conversion of input data into a form more amiable to compression, and the second part is a reversible coding process using one of the more popular lossless algorithms.

Differential Pulse Code Modulation (DPCM) is an example of a predictive coding scheme where the decorrelation process involves creation of a differential signal between the actual value of a pixel compared to the estimated value based on previously encoded pixels. The coding stage of DPCM uses a lossless compression algorithm like Huffman coding or the LZW algorithms.

Transform coding compression is performed using an energy preserving transformation of images pixels into another set such that the maximum amount of information is squeezed into a minimum number of samples. Some examples of these types of transforms are the Karhunen-Loéve Transform (KLT), Fast Cosine Transform (FCT), Discrete Cosine Transform (DCT) (with a subset being Adaptive Discrete Cosine Transform (ADCT)), Block Truncation Coding (BTC), Fast Fourier Transform (FFT), Discrete Wavelet Transform (DWT), Fractal-Transform compression and Vector Quantization (VQ). Of these techniques, KLT is the best

linear transformation leading to uncorrelated coefficients but is rarely used because of the computational cost. DCT, although it is a sub-optimal algorithm, is currently the most widely used because it makes a better trade off between computation and optimization (Pronios and Yovanof,1991). The DWT, fractal-transform, and VQ are promising algorithms that are still being researched.

PART V KLT TRANSFORM

The KLT algorithm is an orthogonal transform similar to the Fourier Transform (FT). The FT is based on sines and cosines while the KLT is based on eigenvectors of the covariance matrix. Eigenvalues are the solution for λ in the equation:

$$C_x \bar{x} = \bar{x} \lambda \quad (10)$$

where C_x is an n^2 by n^2 matrix, x is an eigenvector and λ is a scalar called the eigenvalue. The set of non-trivial solutions of eigenvalues are also referred to as the spectrum of C_x . The eigenvalues are found using the following determinant.

$$|C_x - \lambda I| = 0 \quad (11)$$

where I is the identity matrix. The resulting eigenvectors are orthogonal. The result of the KLT is the assurance that it will provide the most efficient and accurate transformation by minimizing the mean square error (MSE) in reconstruction, therefore maximizing the entropy of the representation.

There are five basic steps to calculating the KLT.

- (1) Compute the mean and covariance function for the image.
- (2) Calculate the eigenvalues for a centered covariance matrix, (centered implies that the mean has been subtracted from the image).
- (3) From eigenvalues calculate the respective eigenvectors.
- (4) Arrange these n^2 eigenvectors into descending eigenvalue order.
- (5) Select only the k

largest eigenvalues from the list. The result is a simplification of the representation, and therefore, a compression of the image. The problem went from an n^2 to a k size KLT basis set, which is a significant reduction. The mathematics involved in giving an example are beyond the scope of this thesis but a detailed description including proofs can be found in (Rosenfeld and Kak,1982). For other less detailed description refer to (Suarez,1991) and (Kreysig,1983).

PART VI DCT TRANSFORM

The foundation for most lossy algorithms being used today is the DCT. This technique is widely used due in part to the adaptive DCT (ADCT) adoption by the Joint Photographic Experts Group (JPEG) as the basis for their compression standard. The JPEG's ADCT is a method that uses DCT as its basis while offering lossless methods as an alternative to DCT. Since the JPEG's adoption of DCT, most vendors have implemented this standard in their products (Waltz ,1992). Figures 31 and 32 represent the process that a DCT-based encoder/decoder goes through in order to compress a single-component (grayscale) image. DCT based compression can be thought of as compressing 8x8 blocks of grayscale image samples. To compress color images, the only required change is that the algorithm must process multiple grayscale images, either individually or by interlapping 8 x 8 sample blocks in turn.

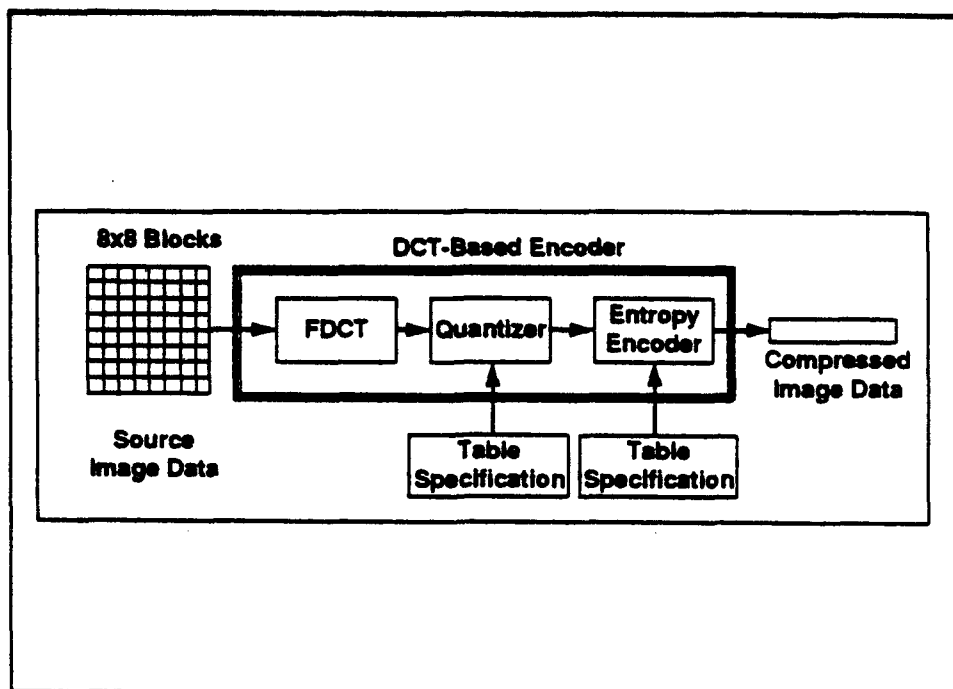


Figure 31. DCT Encoding

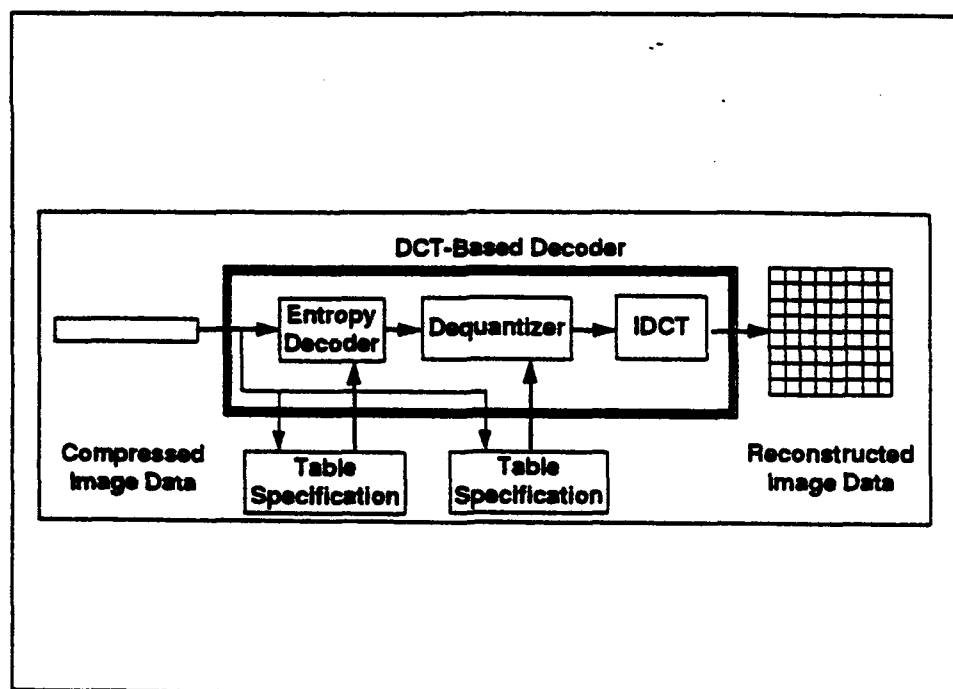


Figure 32. DCT Decoding

Compression using the DCT is a several step process (Figures 31 and 32). The first step is to break an image into the 8 x 8 blocks used by the DCT-based encoder. These samples are shifted from unsigned integers with range ,0..2^P - 1 to signed integers with range , -2^{P-1}..2^{P-1} -1. The forward DCT (FDCT) takes the blocks and evaluates them and generates 64 DCT coefficients. The following equations are idealized mathematical definitions of the 8 x 8 FDCT and the 8 x 8 Inverse DCT (IDCT) (Wallace,1991:32):

$$F(u,v) = \frac{1}{4} C(u)C(v) \left[\sum_{x=0}^7 \sum_{y=0}^7 f(x,y) * \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \right] \quad (12)$$

$$f(x,y) = \frac{1}{4} \left[\sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v) F(u,v) * \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \right] \quad (13)$$

where: $C(u), C(v) = \frac{1}{\sqrt{2}}$ for $u, v = 0$; $C(u), C(v) = 1$ otherwise.

These DCT coefficients (64 basis-signal amplitudes) are the result of decomposing the 64-point discrete signal into 64 orthogonal basis signals, each containing unique 2D spatial frequencies. Once these DCT coefficients have been created, they are then quantized based on a 64-element quantization table supplied by the user. The quantization achieves high compression by only using enough precision to achieve the desired image quality. This step introduces the lossiness in DCT-based encoders and also controls the amount

of compression. The equation used for quantization (Wallace, 1991:34) is:

$$FQ(u,v) = \text{Integer Round} \left(\frac{F(u,v)}{Q(u,v)} \right) \quad (14)$$

the dequantization (Wallace, 1991:34) is the inverse function:

$$FQ'(u,v) = FQ(u,v) * Q(u,v) \quad (15)$$

The third step has two parts. The overall approach in this step is to use entropy coding to put the data into the final compressed format. The first part involves placing the coefficients into a "zigzag" sequence to facilitate the second step of entropy encoding by placing low-frequency coefficients before the high frequency ones (See Figure 33). Once this is accomplished, one of the lossless techniques described above, such as Huffman coding or arithmetic coding, is applied to the data. To uncompress the image, a DCT-based decoder reverses the process.

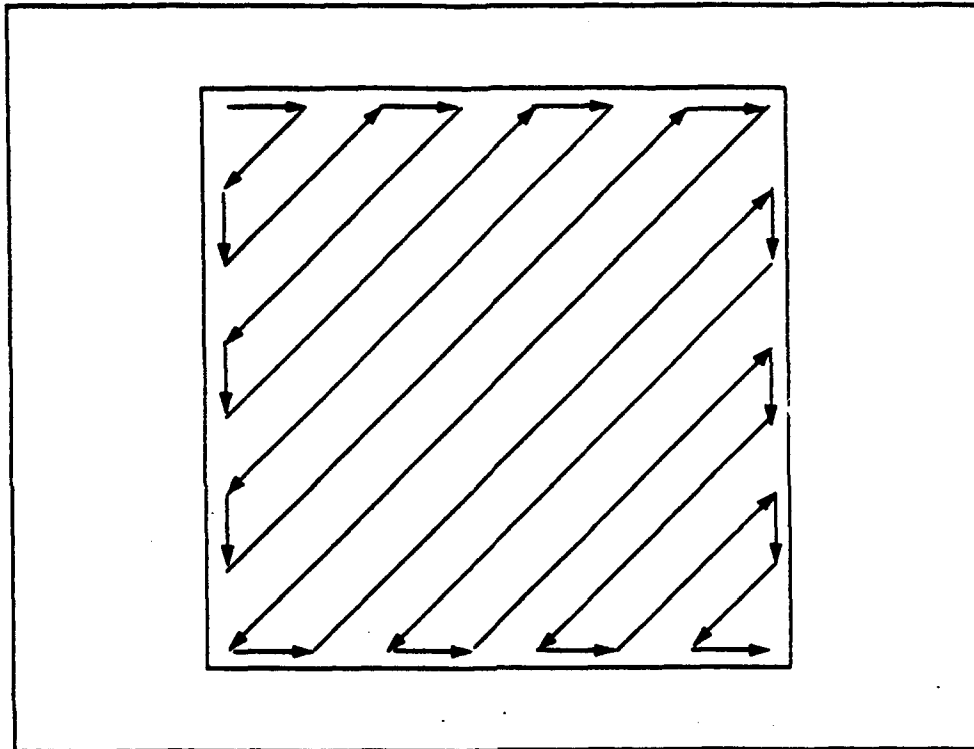


Figure 33. Zig Zag Sequencing

The JPEG standard uses the DCT as the basis for its compression standard, although the standard continues to evolve to keep pace with current technology. Further explanation of the DCT process, formulas, and the JPEG standard can be found in (Wallace,1991), (Blinn,1993), (Kaijiwara,1992), and the current JPEG standard.

Another lossy compression algorithm that is starting to appear in commercial products is called fractal compression. Before discussing fractal compression, it is important that the reader understands the concept of fractals. A fractal is a modern invention that describes objects in nature that cannot be described using Euclidean geometry (spheres cubes etc.) such as clouds, mountains forests etc. The differences

between Euclidean geometry and fractal geometry include the fact that Euclidean shapes have a size and length where fractals possess none of those characteristics. Fractals are described as "self-similar and independent of scale" (Peitgen and Saupe, 1988:25) meaning that no matter how much a viewer "zooms in", the image consists of similar objects as seen in the smaller object. Fractals consider the self-similarities of objects. This means that at different magnifications of an object, such as a leaf, the object consists of the combination of smaller, overlapping objects, of the same shape. The principle is related to the idea similar to the planets revolving around the sun just as electrons revolve around a nucleus. Fractals are usually generated using recursive algorithms that generate overlapping versions of itself, where Euclidean shapes use a specific formula. So how do fractals relate to compression? The answer is that images, like nature, have a lot of redundancy that can be represented by fractals that take up much less space. For example, a leaf could be represented by 4 smaller, overlapping versions of the larger image. These smaller objects along with their positions can be represented using four linear equations. These equations require just a few bytes compared to the image color for each pixel. (Sirota, 1993) The following discussion describes how an image is compressed using fractal compression.

PART VII FRACTAL COMPRESSION

Fractal compression is based on affine transformations. These are the transformations used in a function to scale, rotate, skew and translate points in any number of dimensions. The affine transformation is said to be contractive when the resulting image is smaller than the original. An example of a simple 2D affine transformation is (Anson,1993:196):

$$W(x,y)=(ax+by+e,cx+dy+f) \quad (16)$$

where a,b,c,d define the scale, rotation and skew, e and f determine the translation and x and y are the initial 2D point.

Using the basic concept of a natural order, described in (Mandelbrot, 1982), every picture that exists can be represented by a set of affine transformations. Taking a finite set of N contractive affine transformations W_i every image S is approximated by (Anson,1993:198):

$$S = W_1(S) \cup W_2(S) \cup \dots \cup W_N(S) \quad (17)$$

Michael F. Barnsley (Barnsley,1988) made the observation that real-world images are full of affine redundancy. This observation plus his previous work allowed him to develop the first fractal-transform process that

would automatically compress images. The algorithms to determine the affine transformations and compress an image and the algorithm to decompress a fractal-transformed image are located in Figures 34 and 35. . A more detailed discussion of this transformation of images into affine equations and vise-versa can be found in (Peitgen and Saupe,1988:219). In this reference Barnsley briefly discusses the process. In (Barnsley and Hurd,1993) the mathematics behind fractal compression are discussed extensively. Any further discussion is beyond the scope of this thesis.

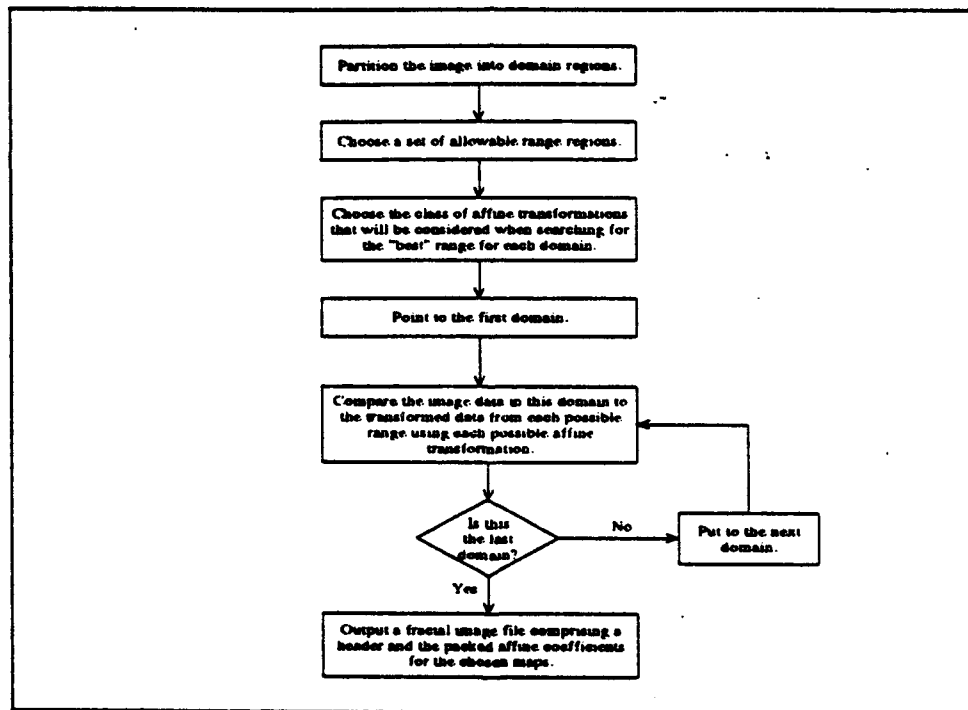


Figure 34. Fractal Compression Algorithm

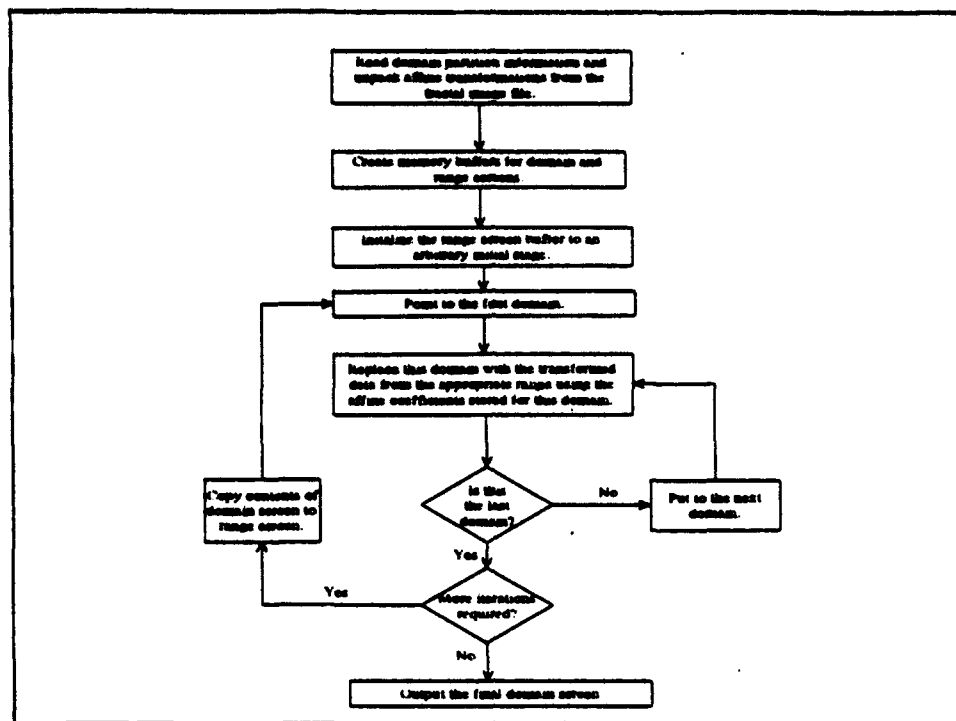


Figure 35. Fractal Decompression Algorithm

This compression algorithm offers some solutions to the problems encountered in lossy compression algorithms such as JPEG's DCT (Anson,1993:195). The shortcomings of DCT that fractal compression addresses include a blocky effect at higher compression ratios, an effect called Gibb's phenomenon where an images sharp edges have ripples that spread out from them, caused by eliminating the higher frequencies used in DCT, and lastly DCT compression is resolution dependent. Resolution dependent means that an image compressed at one resolution will not appear to be the same when uncompressed and displayed at a different resolution. The result of going from a low resolution compression to a high resolution display using DCT is a blockiness effect . As graphics cards and printers continue

to increase their resolution, DCT compression will be less appealing. Fractal compression promises to be a presence in the future (Sirota, 1993).

PART VIII DWT COMPRESSION

The next type of lossy compression that will be discussed is DWT. DWT can be described as a fast linear operation that transforms a data vector into a numerically different vector of the same length. This reversible and orthogonal process is based on a hierarchical set of basis functions that are all scaling and translations of a single basic "wavelet function." Wavelet theory is a new mathematical theory that establishes a class of relationships between discrete functions, such as digitally sampled signals and functions defined on the real line $\mathbb{R}$ (Resikciff,1992). DWT differs from DCT because it treats an image as one block compared to dividing up the image into smaller blocks as done by DCT. Images, when expressed using wavelets, represent most of the information in only a few of the wavelet coefficients. This characteristic allows coefficients with little information to be thrown away without a large effect on the image's appearance. Analogies exist between wavelet transforms and the way a visual cortex of more advanced visual systems (such as a humans) process incoming visual data (Manduca,1992:1225) , (Computer Letter,1993).

The disadvantages of compressing data include processing overhead, disrupting data properties, portability of compressed data between different software and machines, susceptibility of lossless compression to data errors, and

the space required by decoding tables. The most significant disadvantage is the processing overhead. If data is needed quickly, compressing the data could cause problems in utilizing the data in an efficient manner. This problem may be diminished in the future with faster hardware components and software algorithms. When data is compressed, it loses some of its distinguishable attributes (such as color, order etc.). Disrupting the data attributes is one problem that does not currently have a solution. Instead the data must be decompressed before operations such as sorting can take place. The other significant problem that exists in compressing data is the lack of standards. Usually a vendor has his own proprietary "compression technique," making it difficult to transfer compressed data without also sending the software that accompanies it. Standardization should solve these compatibility problems.

PART IX COMPRESSION SUMMARY

Image compression is a crucial part of a PACS system. The desired rates of retrieval would be virtually impossible if the images were not first compressed. Compression can also be applied to different areas of PACS from image compression, to the compression of patient information. One thing to realize, is that for each type of image, whether it is 3-D, 2-D, color or monochrome, and data, text, sound etc., a specific type of compression will yield the best results. That means there is not one compression algorithm that works best for all data, rather several compression algorithms work equally well on different types of data.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

The idea of being able to get all types of medical information about a person, anywhere in the world at anytime with just the touch of a button, is the underlying goal of a global PACS. This thesis has looked at the current state of the art in PACS. The first thing discussed was the background and the emergence of the PACS requirement and issues were addressed such as how PACS is connected, some of the current limitations, the basic components, and how a PACS works now and in the future. The thesis touches on several topics briefly in order to make the reader understand the complexity of a PACS.

This thesis's main concentration included the unresolved issues of database types and compression algorithms. Relational database currently dominates the PACS field, offering enough flexibility to get the job done. The future of PACS databases is a highly controversial topic, and many authors believe that object-oriented databases are the databases of the future. Although object-oriented databases are not proven to the extent of the relational model, they do offer a closer organization and representation to "real world" objects. A distributed database architecture offers more flexibility and resistance to failure for bigger databases than the centralized

architecture. As new databases are developed, the concentration seems to be on distributing the information among users.

The increasing amount of network traffic and software size, as well as a need to reduce data, makes compression necessary when transmitting or storing information. The thesis looks at the two types of compression, lossy and lossless. There is no one correct compression technique for all data. Rather, every type of data has its unique best compression scheme. In a PACS, the best compression technique to use is a combination of several techniques based on the type of data being compressed. Although it comes with some overhead, compression is a necessary part of PACS today and in the future.

In order for a GPACS to work, the underlying structure of a PACS has to work first. Most of the issues that exist on the smaller scale also exist on the global scale and can become more of a problem when different countries and standards are introduced. Cooperation and communication is needed to develop a system that will work in a global market.

After reviewing several hundred articles related to PACS, I suggest that the following actions are needed in order to come to GPACS realization.

GPACS Development

This first section outlines the national effort that needs to occur to offer an effective GPACS. One note here is that since worldwide cooperation is already difficult, the first nation that establishes a system that works, will probably be the leader of GPACS. The rest of the world would then follow in their footsteps. Currently the United States has the largest number of working PACS, which makes the United States a prime candidate for the position of a GPACS leader.

1. A preliminary national committee needs to be established that will initially act in advisory roll and then later in a supervisory roll to all PACS users. (This committee will need full time dedicated members.)

2. This committee should establish a generic PACS template that offers enough flexibility to accommodate hospitals that want to build their PACS piece by piece and those hospitals that want to install their PACS all at once.

3. The committee should offer guidance on different areas of PACS including preferred database capabilities, preferred compression algorithms, preferred image formats along with communication recommendations.

4. The committee should be a collection point for all PACS information from new technology to what hospitals currently have in the way of PACS. This library of

information should be made available to everyone in order to avoid re-inventing existing concepts and applications.

5. The committee should not give specific product recommendations (This allows vendors to develop products that have the desired capabilities outlined by the committee while not limiting users to a specific product).

6. The committee should look at the eventual connection of all PACS and the issues associated with this task. Some of these issues include file formats, communication formats, compression algorithms and database types.

7. The committee should offer a forum for an exchange of ideas and the development of agreed upon standards.

PACS Development

While a national committee is being formed, individual hospitals need to organize themselves internally. The following steps will allow the hospital PACS to become one cohesive PACS that will be able to interact efficiently in a national effort.

1. When a hospital first begins or decides that they will be acquiring a PACS or any part of a PACS, a PACS committee should be established within the hospital. This committee should include a representative from each department, regardless of their interest in PACS (in the future, everyone in the hospital will have a vested interest in the PACS).

2. This hospital PACS committee should reference the GPACS committee referred to earlier (If this GPACS committee is non-existent at the time, the local PACS committee should be required to do thorough research in order to determine the best way to address the PACS development issues).

3. Once a PACS committee is formed, local hospital standards must be set that everyone agrees upon. When setting these standards, a national effort should be kept in mind in order to insure compatibility with other hospital's PACS.

4. Acquiring different hardware and software should be taken into special consideration, making sure all the pieces of the PACS are compatible.

5. Lastly the committee should become a point of contact for inquiries from internal and potential external users of their PACS.

The key things to remember about a PACS is that the development effort is a very large, complicated task consisting of many parts. This task demands many resources including man-hours and money. Along with the financial aspect, is the problem of everyone communicating throughout the PACS life cycle. There is no quick fix to the problems that currently exist, rather a slow tedious process of getting everything to work. Already several million dollars have been spent on PACS development and acquisition (spent by various hospitals). Since these system have been costly, they are unlikely to be significantly modified. These

already existing PACS will try and adapt to the changing technology but it will be the hospitals acquiring the most recent PACS that will provide the stepping stones to the future.

From my readings I conclude that the GPACS trends will include the following:

1. Object-oriented databases are the databases of the future, offering greater flexibility.

2. Distributed database architecture will be the way that a GPACS is implemented, with a common query language being the link between different types of databases.

3. Communication of information will become mainly digital using fiber-optic cables and satellite transmissions. The "super highway" will offer another avenue of communication and may take advantage of the above technologies.

4. Compression schemes will continue to improve. The exact compression algorithms that will be used is still too difficult to predict. The evolving compression schemes will be incorporated offering increased performance. Lossy schemes will most likely be used to compress images unless there is some innovative lossless compression discovery.

5. The network that will evolve will most likely be a multi-access bus architecture for hospitals, allowing several users to be connected using a single bus. For the GPACS I would expect some sort of satellite network along

with the integrated services digital network utilizing fiber-optic cables.

6. Lastly, as GPACS continues to evolve so will the applications of a PACS within a hospital. Images will begin to play a bigger role within the hospital. Not only will the images be used to identify broken bones and tumors, but they will play more of a role in surgery and treatment. The increased dependency on images will increase the amount of attention and funds being invested in PACS which in turn will make a GPACS a reality.

There is still much to be done in the field of PACS and GPACS. The opportunities and improvements are endless. This thesis has given a background and introduction into the field of PACS. It also has shown the importance of a PACS while identifying areas where it needs to be improved along with recommendations for improvement.

APPENDIX A: Definition Tables

Table 1. Inter-Operability Definitions

Term	Definition	Abbreviation
Modalities	This a term used to describe image acquiring machines such as MRI, CT, Ultrasound etc.	
Standardization	This is establishing of a standard (something considered by an authority or general consent as an approved model)	
Reliability	The trustworthiness of something.	
Quality Assurance	The process of assuring that something or someone meets the specifications of quality previously established.	
Query	An inquiry about information contained in a location such as a database.	
International Standards Organization	An international organization that designates standards for industry to follow.	ISO
American National Standards Institute	The American organization that designates standards for the United States. This organization is the United States representative to the ISO	ANSI
American College of Radiology and the National Electrical Manufacturers Association	These two organizations created their own standard for image transfer to easily transfer images acquired by different modalities. This standard was the first of its kind to evolve.	ACR-NEMA
Hierarchy	A system of things ranked one above another.	
Random Access Memory	The memory in a computer that is used while the computer is operating. When the computer is turned off all information in the RAM is lost.	RAM
Versatile Message Transaction Protocol	A discipline followed when transmitting messages from one computer to another.	VMTP
File Transfer Protocol	A discipline followed when transmitting files from one computer to another.	FTP

Table 2. Network Definitions

Term	Definition	Abbreviation
Link	This is a connection established between two sites or nodes	-
Site	This term refers to a computer terminal or group of computers represented by on instance in a network. (Also referred to as a host or node)	-
Bottleneck	This is where a specific point in a sequence of events gets overwhelmed, slowing down the completion of other processes	-
Node	This term refers to a computer terminal or group of computers represented by on instance in a network. (Also referred to as a host or site)	-
Host	This term refers to a computer terminal or group of computers represented by on instance in a network. (Also referred to as a site or node)	-
Local Area Network	This is a network designed to cover a small geographical area, usually used in an office environment. LANs are generally very fast.	LAN
Wide Area Network	This is a network distributed over a large geographical areas connecting LANs	WAN
Bandwidth	This describes the range of frequencies available on a communication line.	BW
T1 Line	A link with the specified capacity of 1.44Mbits/sec	-
Broadband ISDN	A wider bandwidth used for specialized purposes on the ISDN.	B-ISDN
Packet	This term refers to a block of data that is transmitted containing part of the original data plus other transmission information specific to the data contained within.	-

Table 4. Database Definitions

Term	Definition	Abbreviation
Database	Computerized record keeping system	DB
Database Management System	The software layer between the users of a database and the actual physical database	DBMS
Data	Any information stored in a database, represented by an organized collection of bits	-
Relational Database	A database where the data is represented to the user as a collection of tables	RDB
Object-Oriented Database	A database where data is represented to the user as a set of objects	OODB
Query	An inquiry put to a database to retrieve the requested data contained in the database	-
Write Once Read Many	A term describing the capability to write to an optical disk only once. Then this disk can be read repeatedly.	WORM
Schema	A presentation of information in a diagram form and in our case a presentation of the way computer systems are organized.	-
Global Schema	This presentation holds a global diagram identifying all schemas available globally.	-
Autonomy	A self-governing state	
Kilobyte	Term referring to 1,000 bytes.	Kbyte, KB
Megabyte	Term referring to 1,000,000 bytes.	Mbyte, MB
Gigabyte	Term referring to 1,000,000,000 bytes.	Gbyte, GB
Terrabyte	Term referring to 1,000,000,000,000 bytes.	Tbyte, TB

Table 7. RDB Definitions

Term	Definition	Abbreviation
Standard Query Language	A language developed to query a relational database to allow a user to extract needed data.	SQL
RESTRICT	Extracts only the desired rows in a table, also commonly referred to as SELECT.	-
PROJECT	Extracts only the desired columns in a table.	-
PRODUCT	Builds a new table from two specified tables based on all combinations of rows.	-
UNION	Builds a new table consisting of all rows in either or both of two specified tables.	-
INTERSECT	Builds a new table consisting of all rows in both of two specified tables.	-
DIFFERENCE	Builds a new table consisting of all rows in the first but not the second of two specified relations.	-
JOIN	Builds a new table from two specified tables consisting of all possible combinations of rows, one from each of the two tables, such that two rows contributing to any given combination satisfy some specified condition. The resulting table is the Cartesian product of the two specified tables.	-
DIVIDE	Takes two tables, one binary and one unary, and builds a new table consisting of all values of one column of the binary relation that match (in the other column) all values of the unary relation.	-
Table	A systematic arrangement of data using rows and columns.	
Row	Data represented in horizontal fashion. Also referred to as a tuple when referencing a relational table.	
Column	Data represented in vertical fashion. Also referred to as a attribute when referencing a relational table.	
Cardinality	The number of rows in a table.	
Degree	The number of columns in a table.	
Primary Key	The element associated with data that identifies it uniquely.	
Atomic	The smallest semantic unit of data. This data, in reference to the relational model, has no internal structure and therefore cannot be broken down further. Ex : Scalars	

Table 12. OODB Definitions

Term	Definition	Abbreviation
Abstraction	An abstraction denotes the essential characteristics of an object that distinguish it from all other kinds of objects and thus provide crispy defined conceptual boundaries, relative to the perspective of the viewer.	-
Encapsulation	Encapsulation is the process of hiding all the details of an object that do not contribute to its essential characteristics.	-
Modularity	Modularity is the property of a system that has been decomposed into a set of cohesive and loosely coupled modules.	-
Hierarchy	Hierarchy is a ranking or ordering of abstractions.	-
Typing	Typing is the enforcement of the class of an object, such that objects of different types may not be interchanged, or at the most, they may be interchanged in only very restricted ways.	-
Concurrency	Concurrency is the property that distinguishes an active object from one that is not active.	-
Persistence	Persistence is the property of an object through which its existence transcends time (i.e. the object continues to exist after its creator ceases to exist) and/or space (i.e. the object's location moves from the address space in which it was created).	-

Table 13(a). Compression Definitions

Term	Definition	Abbreviation
Entropy	A measure of disorder in a system. For the purposes of data compression, when data is compressed to an optimal form, this is considered the smallest value of entropy for that data. Generic formula is $S = k \ln P + c$, where S is the value of the measure for a system in a given state, P is the probability of occurrence of that state, k is a fixed constant and c is an arbitrary constant. Storer's definition of entropy for S over the radix r, $r > 1$, is $H_r(S) = \sum_{i=1}^t p_i \log_r(1/p_i),$ where p_i are the independent probabilities of each member of the set of S. The default for the radix is 2. The bottom line on entropy of data is that there is a minimum value of entropy for each data that we are trying to obtain, if we reach this value then the data has been compressed optimally, cannot be compressed in itself, any further.	-
Compression Ratio	This is the ratio used in data compression referring to the original size of the data divided by the compressed size of the data	-
Codewords	A string of characters representing a longer string of characters in textual compression.	-
Symbols	A single character representing an amount of data.	-
Lossless Compression	The act of compressing data in such a way that it can be recovered to its original state when uncompressed.	-
Lossy Compression	The act of compressing data that results in a degree of data being lost when the data is uncompressed.	-
Banding	The result of dividing an image up into one-dimensional rows or columns before compression (to make the compression an easier task). The "banding" occurs when each of these lossy compressed images are reunited, each band being a little different color from its neighbor.	-

Table 13(b). Compression Definitions

Term	Definition	Abbreviation
Noise	Irrelevant or meaningless data generated that is generated along with the desired data.	-
Orthogonal	Statistically independent	-
Quantization	Limiting the possible values to a specified range.	-

BIBLIOGRAPHY

Abramson, J. R., Krasner, B. H. "Data Compression in the PACS Environment", *Medical Imaging III: Image Capture and Display*, Roger H. Schneider, Samuel J. Dwyer III, R. Gilbert Jost, Editors, Proc. SPIE 1091, 42-49, 1989.

Acree, R.K., Ullah, N., Karia, A., Rahmeh, J.T., Abraham, J.A. "An Object-Oriented Approach for Implementating Algorithm-Based Fault Tolerance", *Phoenix Conference on Computers and Communications*, Twelfth Annual International, 210-216, Institute of Electrical and Electronics Engineers, Inc., Temple AZ, 1993.

Akrout, N., Diab, C., Prost, R., Goutte, R., Amiel, M. "Subband Directional Vector Quantization in Radiological Image Compression", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 203-212, SPIE-The International Society for Optical Engineering, Newport Beach CA, 1992.

Allen, L., & Frieder, O. "Exploiting Database Technology in the Medical Arena.", *IEEE Engineering in Medicine and Biology.*, MAR V 11 N 2, 42-49, 1992.

Anderson, M.P., Brown, D.G., Loew, M.H. "Evaluation of Medical Image Compression by Gabor Elementary Functions", *Medical Imaging V: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1444, 407-412, 1991.

Anderson, D., Boyd, D., Grosky, W., Ashany, R., Golshani, F., Pizzarello, T. "Multimedia Systems", *Phoenix Conference on Computers and Communications*, Tenth Annual International March 27-30, 719-724, IEEE Computer Society Press, Scottsdale AZ, 1991.

Anson, L.F. "Fractal Image Compression", *Byte*, Oct, 195-202, 1993.

Apiki, S. "Compressing with Fractals", *Byte's Essential Guide to Windows*, Spring, 95-96, 1993.

Atwood, T. "ODMG-93: The Object DBMS Standard", *Object Magazine*, Sep-Oct, 37-44, Sigs Publications Inc., New York NY, 1993.

Aubry, F., Bizais, Y., Chameroy, V., Gibaud, B., Forte, A.M., Di Paola, R., Scaabin, J.M. "Object-oriented Model for Medical Image Database", *Medical Imaging V: PACS System Design and Evaluation*, R. G. Jost, Editor, Proc. SPIE 1446, 168-176, 1991.

Aubry, F., Chameroy, V., Giron, A., Di Paola, R., Gibaud, B., Bizais, Y., Vital, D., Lienard, R., Todd-Pokropek, A.,

Kranz, R., Deconinck, F., Ratib, O. "MIMOSA : Conceptual Modelling of Medical Image Management in an Open System Architecture", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan , Volume 14, 1199-1201, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Badal, D.Z. "Neural Network Recognition of Human Face Images Stored in the Database", *Phoenix Conference on Computers and Communications*, Twelfth Annual International, 552-557, Institute of Electrical and Electronics Engineers, Inc., Temple AZ, 1993.

Barnsley, M.F. "Fractal Modeling of Real World Images", *The Science of Fractal Images*, Editors Heinz-Otto Peitgen and Dietmar Saupe, Chapter 5, 219-242, Springer-Verlag New York Inc., New York NY, 1988.

Barnsley, M.F. Fractal Image Compression, A. K. Peters, Wellesly MA, 1993.

Barnsley, M.F. Fractals Everywhere, 2d ed, Academic Press, San Diego CA, 1993.

Bell, D., Grimson, J. Distributed Database Systems, Addison-Wesley, Reading MA, 1992.

Beuscart, R., Grave, C., Wartki, M., Serbouti, S., Beuscart-Zephir, M.C. "Computer Supported Cooperative Work For Medicine Imaging", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan , Volume 14, 1213-1214, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Bilow, S.C. "Reflections on the State of Object-Oriented Literature", *Journal of Object-Oriented Programming*, June, 72-73, Sigs Publications Inc., New York NY, 1993.

Bilow, S.C. "Reflections on the State of Object-Oriented Literature.", *Journal of Object-Oriented Program*, Jun v6 n3, 72-80, 1993.

Bizais, Y., Aubry, F., Gibaud, B., Forte, A.M., Di Paola, R., Scaabin, J.M. "A Qualitative Approach to Medical Image Data Bases", *Medical Imaging V: PACS System Design and Evaluation*, R. G. Jost, Editor, Proc. SPIE 1446, 156-167, 1991.

Black, U. TCP/IP and Related Protocols, McGraw-Hill Inc., New York NY, 1992.

Blinn, J. "What's the Deal with the DCT?", *IEEE Computer Graphics & Applications*, July, 78-83, 1993.

Booch, G. Object-Oriented Design, Benjamin/Cummings, Redwood City CA, 1991.

Booker, E. "Rebirth of Fractals: Fractal Equations Find Raison D'etre in Data Compression", *Computerworld*, July v27 n27, 28, 1993.

Busse, T. "IS: A Prescription for Spiraling Health Costs; Billing Systems are Just the Start.", *InfoWorld*, Feb 15 v15 n7, 1-2, 1993.

Cardenas, A. F. "Pictorial/Visual Access to Multimedia/Heterogenous Databases", *IEEE*, EH0336, 88-91, 1991.

Chameroy, V., Di Paola, R. "Towards Multidimensional Medical Image Coding", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 261-272, SPIE-The International Society for Optical Engineering, Newport Beach CA, 1992.

Chen, P.P. "The Entity-Relationship Model: Towards a Unified View of Data", *ACM Transactions on Database Systems*, MAR v1 n1, 1976.

Chen, P.P. ed Entity-Relationship Approach to Information Modeling and Analysis, North-Holland The Netherlands, 1983.

Chen, J., Flynn, M.J. "The Effect of Block Size on Image Quality for Compressed Chest Radiographs", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 252-260, SPIE-The International Society for Optical Engineering, Newport Beach CA, 1992.

Cho, P. S., Huang, H.K., Tillisch, J. "Centralized vs. Distributed PACS for Intensive Care Units", *Medical Imaging III: PACS System Design and Evaluation*, Roger H. Schneider, Samuel J. Dwyer III, R. Gilbert Jost, Editors, Proc. SPIE 1093, 387-391, 1989.

Chuang, K.S., Huang, H.K. "Comparison of Chi-Square and Join-Count Methods for Evaluating Digital Image Data", *UCLA Medical Imaging Division and Clinical PACS Project*, FEB, 12-17, Los Angeles CA, 1993.

Clementini, E., Di Felice, P. "Multiparadigm Query Strategies in Object-Oriented Geographic Databases", *Phoenix Conference on Computers and Communications*, Tenth Annual International March 27-30, 295-301, IEEE Computer Society Press, Scottsdale AZ, 1991.

Cockroft, G., Hourvitz, L. "NeXTstep: Putting JPEG to Multiple Uses", *Communications of the ACM*, APR v34 n4 , 45, 1991.

Codd, E.F. "A Relational Model of Data for Large Shared Data Banks", *Communications of the ACM*, JUN v13 n6, 377-387, 1970.

Codd, E.F. "A Database Sublanguage Founded on Relational Calculus", *Proc. ACM SIGFIDET Workshop on Data Description, Access and Control*, Nov, San Diego CA, 1971.

Codd, E.F. "Extending the Database Relational Model to Capture More Meaning", *ACM Transactions on Database Systems*, DEC v4 n4, 397-434, 1979.

Codd, E.F. The Relational Model for Database Management Version 2, Addison-Wesley, Reading MA, 1990.

Computer Letter "Squeeze Play", *Computer Letter*, Jul, 1-6, Technologic Partners, 1993.

Cormack, G.V. "Data Compression on a Database System", *Communications of the ACM*, v28 n12, 1336-1342, 1985.

Crowther, W.R., Heart, F. E., McKenzie, A.A., McQuillan, J.M., Walden, D.C. "Issues in Packet Switching Network Design", *Proceedings of the AFIPS National Computer Conference*, 161-176, 1975.

Date, C. J. Introduction to Database Systems, Vol 2, Fifth Ed, , Addison-Wesley, Reading MA, 1983.

Date, C. J. Introduction to Database Systems, Vol 1, Fifth Ed, , Addison-Wesley, Reading MA, 1990.

Date, C.J. Relational Database Writings 1985-1989, Addison-Wesley, Reading MA, 1990.

Davros, W.J., Gaskill, J.W., & Mun S.K. "Quality Assurance for Image Display Devices in a Hospital-wide IMACS Network", *Medical Imaging III: Image Capture and Display*, Roger H. Schneider, Samuel J. Dwyer III, R. Gilbert Jost, Editors, *Proc. SPIE* 1091, 301-304, 1989.

Del Bimbo, A. Lucarella, D. "A Fuzzy Object Retrieval System for Image Understanding", *Phoenix Conference on Computers and Communications*, Tenth Annual International March 27-30, 835-841, IEEE Computer Society Press, Scottsdale AZ, 1991.

Demongeot, J., Cinquin, P. "PACS and Multi-Modality Image Registration", *Proceedings of the Annual International*

Conference of the IEEE Engineering in Medicine and Biology Society, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan, Volume 14, 1207-1208, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Demurjian, S.A., Hu, M.Y., Ting, T.C., Kleinman, D. "Towards an Authorization Mechanism for User-Role Based Security in an Object-Oriented Design Model", Phoenix Conference on Computers and Communications, Twelfth Annual International, 195-202, Institute of Electrical and Electronics Engineers, Inc., Temple AZ, 1993.

DeWitt, D., Gray, J. "Parallel Database Systems: The Future of High Performance Database Systems", Communications of the ACM, June v35 n6, 85-99, 1992.

Dickstein, D. "For End Users Only: How To Choose Your Database Management System", Data Based Advisor, Vol. 11, No. 5, May, 49-57, Data Based Solutions, Inc., San Diego CA, 1993.

Dietrich, R.B., Boechat, M.I., Huang H.K. "Experience with Phosphor Imaging Plates: Clinical Experience in Pediatric Radiology", Medical Imaging III: Image Capture and Display, Roger H. Schneider, Samuel J. Dwyer III, R. Gilbert Jost, Editors, Proc. SPIE 1091, 242-244, 1989.

Doll, D.R. "Telecommunications Turbulence and the Computer Network Evolution", Computer, Vol. 7 No. 2, Feb, 13-22, 1974.

Donnelly, J.J. "Selection of Image Acquisition Methods", Medical Imaging V: Image Capture, Formatting, and Display, Yongmin Kim, Editor, Proc. SPIE 1444, 351-356, 1991.

Durkin, M. A., Oosterwijk, H. J., Przybylowicz, J. E. "Image Quality Issues With Viewing and Printing in a Network Environment", Medical Imaging VI: Image Capture, Formatting, and Display, Yongmin Kim, Editor, Proc. SPIE 1653, 111-116, SPIE-The International Society for Optical Engineering, Newport Beach CA, 1992.

Elmasri, R., Navathe, S. Fundamentals of Database Systems, Benjamin/Cummings, Redwood City CA, 1989.

English, L.P. "Object Databases at Work: What are Object Database Management Systems - and What are They Good For?", DBMS, Oct v5 n11, 44-51, 1992.

Equitz, W. "Fast Algorithms for Vector Quantization Picture Coding", ICASSP, 18.1.1, 1987.

Erdekian, V.V., Trombetta, S.P. "Display Systems for Medical Imaging", *Medical Imaging V: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1444, 151-158, 1991.

Eskow, D. "How to Get Commitment from Distributed Databases", *Corporate Computing*, Feb v2 n2, 181-182, Ziff-Davis Publishing Company, 1993.

Esser, P.D., Halpern, E.J. "Quality Assurance of PACS systems with Laser Film Digitizers", *Medical Imaging V: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1444, 100-103, 1991.

Feingold, E., Seshadri, S.B., Arenson, R.L. "Folder Management on a Multimodality PACS Display Station", *Medical Imaging V: PACS System Design and Evaluation*, R. G. Jost, Editor, Proc. SPIE 1446, 211-216, 1991.

Finkelstien, R. "Breaking the Mold.", *Database Programming & Design*, Feb v6 n2, 48-54, Miller Freeman Publications, 1993.

Fisher, P., Grover, B., Brauer, G. & Ritchie, G. "Digital Image Display Station Performance Requirements Based on Physician Experience with a Prototype System", *Medical Imaging III: Image Capture and Display*, Roger H. Schneider, Samuel J. Dwyer III, R. Gilbert Jost, Editors, Proc. SPIE 1091, 325-331, 1989.

Fotouhi, F., Lee, T, Grosky, W.I. "The Generalized Index Model for Object-Oriented Database Systems", *Phoenix Conference on Computers and Communications*, Tenth Annual International March 27-30, 302-308, IEEE Computer Society Press, Scottsdale AZ, 1991.

Frank, L. Database Theory and Practice, Addison-Wesley, Reading MA, 1988.

French, J.C., Jones, A.K., Pfaltz, J.L. "A Summary of the NSF Scientific Database Workshop", *IEEE*, Aug, 175-181, 1991.

French, B. "The OOP Primer for Xbase Programmers", *Data Based Advisor*, Vol. 10 No. 12 Dec, 95-98, Data Based Solutions Inc., San Diego CA, 1992.

Frense, F., Le Gall, G., Barillot, C., Gibaud, B., Manens, J.P., Toumoulin, C., Lemoine, D., Chenal, C., Scarabin, J.M. "A Stereostatic Multibeam Radiation Therapy System in a PACS Environment", *Medical Imaging V: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1444, 26-36, 1991.

Fuochi, A. "How Would You Like Your Database?", *Computing Canada*, 28 SEP v18 n20, 25, Pleshman Publications Ltd., Canada, 1992.

Garcia, A., Davis, T. "A Compiled List of Current RDBMS", *Data Based Advisor*, Vol. 11, No. 5, May, 58-83, Data Based Solutions, Inc., San Diego CA, 1993.

Garcia-Molina, & H., Lindsay, B. "Research Directions for Distributed Databases.", *Database Engineering*, Volume 9, 196-197, 1990.

Gelish, A. "You Can't Just Plug It In: Digital Imaging Networks/Picture Archiving and Communication Systems (DIN/PACS) Installation", *Medical Imaging V: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1444, 363-372, 1991.

Gibaud, B., Bizais, Y., Morcet, N., Gandon, Y., Buhé, T., Forte, A.M., Aubry, F., Chabriaux, J., Chameroy, V., Di Paola, R., Ratib, O., Pokropek, A.T., Kranz, R., Guitel, M., Vital, D., Ramond, J.P. "Toward Modelling Image Management in PACS: Lessons From a Preliminary Review of End-Users Requirements", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan, Volume 14, 1202-1203, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Goeringer, F. "Medical Diagnostic Imaging Support Systems for Military Medicine", *Medical Imaging V: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1444, 340-350, 1991.

Goeringer, F. "An Electronic Beachhead in Medical Imaging: An Interview with Fred Goeringer", *Siemens Review*, Interview, May, 5-11, 1992.

Goodenough, D. J., Atkins, F.B., Dyer, S.M. "Automated Techniques For Quality Assurance of Radiological Image Modalities", *Medical Imaging V: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1444, 87-99, 1991.

Graham, I. "Interoperation: The Tools of the Trade", *Object Magazine*, Sep-Oct, 75-76, Sigs Publications Inc., New York, NY, 1993.

Gray, R.M. "Vector Quantization", *ASSP Magazine*, Apr, 4-29, 1984.

Gray, J.E. "Use of the SMPTE Test Pattern in Picture Archiving and Communication Systems(PACS)", *Medical Imaging*

V: Image Capture, Formatting, and Display, Yongmin Kim, Editor, Proc. SPIE 1444, 118-124, 1991.

Gray, P.M.D., Kulkarni, K.G., Paton, N.W. Object-Oriented Databases: A Semantic Data Model Approach, Prentice-Hall International (UK) Ltd, Englewood Cliffs NJ, 1992.

Griffiths, J.M. ISDN Explained: Worldwide Network and Applications Technology, (Second Edition), John Wiley & Sons Ltd, West Sussex England, 1992.

Guo, M. "An Object-Oriented SQL (OSQL) Based On Association Pattern Query Formulation", Phoenix Conference on Computers and Communications, Twelfth Annual International, 231-237, Institute of Electrical and Electronics Engineers, Inc., Temple AZ, 1993.

Gupta, R., Horowitz, E. Object-Oriented Databases with Applications to CASE, Networks, and VLSI CAD, Prentice-Hall, Englewood Cliffs NJ, 1991.

Halpern, E.J., Esser, P.D., "An Improved Phantom for Quality Control of Laser Scanner Digitizers in PACS", Medical Imaging V: Image Capture, Formatting, and Display, Yongmin Kim, Editor, Proc. SPIE 1444, 104-115, 1991.

Hauser, S., Mittu, R., Rivera, G., Thoma, G. "Performance Study of a LAN Based Image Server", Phoenix Conference on Computers and Communications, Tenth Annual International March 27-30, 630-635, IEEE Computer Society Press, Scottsdale AZ, 1991.

Held, G., Data Compression: Techniques and Applications, Hardware and Software Considerations, John Wiley and Sons, New York NY, 1983.

Heldman, R. K. Global Telecommunications : Layered Networks' Layered Services, McGraw-Hill Inc., New York NY, 1992.

Hillman, B.J., Fajardo, L.L. "Scientific Assessment of the Clinical Utility of Phosphor Plate Computed Radiology", Medical Imaging III: Image Capture and Display, Roger H. Schneider, Samuel J. Dwyer III, R. Gilbert Jost, Editors, Proc. SPIE 1091, 245-251, 1989.

Ho, B.K.T., Chan, K.K., Huang, H.K. "Integrated System Design for High-speed Image Decompression and Display", Medical Imaging III: Image Capture and Display, Roger H. Schneider, Samuel J. Dwyer III, R. Gilbert Jost, Editors, Proc. SPIE 1091, 332-336, 1989.

Ho, B.T.K., Ma, M., Chen, D.T. "Mathematical Model to Quantify JPEC Block Artifacts", UCLA Medical Imaging

Division and Clinical PACS Project, FEB, 22-27, Los Angeles CA, 1993.

Horspool, R.N. "Improving LZW", DCC '91 Data Compression Conference, , 332-341, IEEE Computer Society Press, Los Alamitos CA, 1991.

Howard, P.G., Vitter, J.S. "New Methods for Lossless Image Compression Using Arithmetic Coding", Data Compression Conference, Apr, 257-266, IEEE Computer Society Press, Los Alamitos, California, 1991.

Hsu, C., Rattner, L. "Metadatabase Solutions for Enterprise Information Integration Problems", Database, Vol 24, No. 1, Winter , 23-35, SIGBIT, 1993.

Huang, L.J., Valentino, D.J., Taira, R.K. "Software Development and Management for a Picture Archiving and Communication System (PACS)", UCLA Medical Imaging Division and Clinical PACS Project, FEB, 97-104, Los Angeles CA, 1993.

Huffman, D.A. "A Method for the Construction of Minimum-Redundancy Codes", Proceedings of IRE, v40, 1098-1101, 1952.

Hughes, J.G. Object-Oriented Databases, Prentice-Hall International (UK) Ltd, Englewood Cliffs NJ, 1991.

Hurlow, M. "Lab Notes of Image Compression and Virtual Memory", MACWORLD, March, 93-94, 1992.

Hurson, A.R., Pakzad, S., Cheng, J. "Object-Oriented Database Management Systems: Evolution and Performance Issues", IEEE Computer, Feb , 48-60, 1993.

Itoh, Y., Ishigaki, T., Sakuma, S., Hirose, M., Fukatsu, H., Itoh, S., Horikawa, Y., Shimamoto, K., Tadokoro, T., Ikeda, M., & Itoh, Kengo "Influence of CRT Workstation on Observer's Performance", Computer Methods in Biomedicine, v37 , 253-258, 1992 .

Jackson, S., Szasz, I. "The Use of Data Compression in Telecommunication of Nuclear Medicine Images", Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan , Volume 14, 1229, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Jacquet, P., & Mühlethaler, P. " Machnet: A Simple Access Protocol For High Speed or Long Haul Communications ", SIGCOMM '90, 90-100, 1990.

Kahn, R. "Resource-Sharing Computer Communications Networks", *Proceedings of the IEEE*, Vol. 60 No. 11 Nov , 1972-1407, 1972.

Kajiwar, K. "JPEG Compression For PACS.", *Computer Methods and Programs in Biomedicine*, V. 37 n 4 May , 343-351, 1992.

Kao, T., Shieh, S.H., Wu, L.C. "Dynamic Radionuclide Images Compression Based on Principal Components Analysis", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan , Volume 14, 1227-1228, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Karadimce, A.P., Urban, S.D. "CDOL: A declarative Platform for Developing OODB Applications", *Phoenix Conference on Computers and Communications*, Twelfth Annual International, 224-230, Institute of Electrical and Electronics Engineers, Inc., Temple AZ, 1993.

Kim, W., Lochovsky, F.H. Object-Oriented Concepts, Databases, and Applications, Addison-Wesley, Reading MA, 1989.

Kimme-Smith, C. "Glandular Tissue Contrast in CCD Digitized Mammograms", *UCLA Medical Imaging Division and Clinical PACS Project*, FEB, 51-56, Los Angeles CA, 1993.

Kishore, S., Khalsa, S., Seshadri, S. B., Arenson, R. L. "A Digital Film Library Implementation", *Medical Imaging V: PACS System Design and Evaluation*, R. G. Jost, Editor, Proc. SPIE 1446, 188-194, 1991.

Kishore, S., Khalsa, S., Feingold, E., Seshadri, S. B., Arenson, R. L. "PACS Software Testing", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 117-125, SPIE-The International Society for Optical Engineering, Newport Beach CA, 1992.

Kitanosono, T., Kurashita, Y., Honda, M., Hishida, T., Konishi, H., Mizuno, M., and Anzai, M. "The Use of Multimedia in Patient Care", *Computer Methods in Biomedicine*, v37, 259-263, 1992.

Klien, S.A., Carney, T. "'Perfect' Displays and 'Perfect' Image Compression in Space and Time", *Human Vision, Visual Processing, and Digital Display II*, SPIE Vol. 1453, 190-205, 1991.

Korzeniowski, P. "Gateways Link Legacy, Distributed Databases; Connectivity Level Varies Among DBMSs; Early

Users Balk at Price, Inconsistency.", *Software Magazine*, Nov v12 n15, 85-90, Sentry Publishing Company Inc., 1992.

Kreysig, E. Advanced Engineering Mathematics, (Fifth Edition), John Wiley and Sons, New York, New York, 1983.

Lawton, G. "Protecting Integrity of Distributed Data; Two-Phase Commit as a DBMS Service Not Widely Adopted; Coding a Fallback", *Software Magazine*, Jan v13 n1, 71-76, Sentry Publishing Company Inc., 1993 .

Lee, H., Rowberg, A. H., Frank, M. S., Choi, H. S., Kim, K. "Subjective Evaluation of Compressed Image Quality", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 241-251, SPIE-The International Society for Optical Engineering, Newport Beach CA, 1992.

Leehan, J.A., Ismael Magña M., Duchêne, J., Kanz, R., Lerallut, J.F. "A Hardware Image Compression Subsystem for a NuBus-based Workstation in a Picture Archiving and Communication System Environment", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan , Volume 14, 1217-1218, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Lelewer, D.A., Hirschberg, D.S. "Data Compression", Department of Information and Computer Science, Technical Report No. 87-10, , University of California, Irvine CA, 1987.

Ligier, Y., Ratib, O., Girard, C., Perrier, R., Logean, M. "Portable and Extensible Software for the Manipulation of Medical Images", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan , Volume 14, 1209-1210, Institute of Electrical and Electronics Engineers, Inc., Paris, France, 1992.

Lin, E., Lu, C.C. "A Study on Objective Fidelity For Progressive Image Transmission Systems", *Human Vision, Visual Processing, and Digital Display II*, SPIE Vol. 1453, 206-209, 1991.

Linde, Y., Buzo, A., Gray, R.M. "An Algorithm for Vector Quantizer Design", *IEEE Trans. Comm.*, Com-28(1):84-95, 1980.

Liu, B.J., Chuang, K.S. "A Sigma-Median Filter", *UCLA Medical Imaging Division and Clinical PACS Project*, FEB, 18-21, Los Angeles CA, 1993.

Lo, S. B., Krasner, B. H., Mun, S. K., Horii, S. C. "Hybrid Coding for Split Gray Values in Radiological Image Compression", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 194-202, SPIE-The International Society for Optical Engineering, Newport Beach CA, 1992.

Lodder, H., van Poppel, B.M., Bakker, A.R. "Prefetching: PACS Image Management Optimization Using HIS/RIS Information", *Medical Imaging V: PACS System Design and Evaluation*, R. G. Jost, Editor, Proc. SPIE 1446, 227-233, 1991.

Loomis, M.E.S. "Object and Relational Technologies: Can They Cooperate?", *Object Magazine*, Jan-Feb, 35-40, Sigs Publications Inc., New York NY, 1993.

Lui Sheng, O., et al "Distributed Database Design and Modeling for PACS", *Medical Imaging IV: PACS System Design and Evaluation*, SPIE Vol. 1234, 1990.

Lui Sheng, O., Chen Garcia, O. "The Design of Medical Image Databases: A Distributed Approach", *IEEE 1990 Phoenix Conference on Computers and Communications*, Ninth Annual International, 1990.

Lunt, T., Fernandez, E. "Database Security", *IEEE*, Aug, 227-234, 1991.

Lynch, T.J. Data Compression: Techniques and Applications, Lifetime Publications, Belmont CA, 1985.

Lynne, K.J. "Distributing the Server Function in a Multi-Ring PAC System", *Medical Imaging V: PACS System Design and Evaluation*, R. G. Jost, Editor, Proc. SPIE 1446, 177-187, , 1991.

MacLeod, J.N. "Performance of Encoding/Decoding the Transfer Syntax", *Phoenix Conference on Computers and Communications*, Tenth Annual International March 27-30, 644-650, IEEE Computer Society Press, Scottsdale AZ, 1991.

Magedson, B. "Building OT From the Bottom UP", *Object Magazine*, Jan-Feb, 45-47, Sigs Publications Inc., New York NY, 1993.

Mandelbrot, B.B. The Fractal Geometry of Nature, W.H. Freeman, New York, New York, 1977.

Manduca, A. "A Wavelet-Based Tool for Medical Image Compression", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 495-503, SPIE-The International Society for Optical Engineering, Newport Beach CA, 1992.

Manduca, A. "Interactive Wavelet-Based Image Compression with Arbitrary Region Preservation", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan, Volume 14, 1224-1226, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Marglin, S. I. "Clinical Aspects of Quality Assurance in PACS", *Medical Imaging V: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1444, 83-86, 1991.

Martin, J., Odell, J.J., "Object-Oriented Analysis and Design", *Object-Oriented Databases*, Ch. 13, 197-215, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1992.

Matheus, R.; Van Goor, J. M. Noothoven "European Community: Standardization in Medical Informatics and Imaging.", *Computer Methods and Programs in Biomedicine*, V. 37 n 4 May, 333-341, 1992.

Maydell, U.M., MacGregor, M.H. "Performance of the ACR/NEMA Protocol", *Medical Imaging III: Image Capture and Display*, Roger H. Schneider, Samuel J. Dwyer III, R. Gilbert Jost, Editors, Proc. SPIE 1091, 68-76, 1989.

McDonald, C., Barnett, G. "Hospital Information Systems.", Shortliffe, E., Perreault, L., Wiederhold, G., and Fagan, L. (1990) *Medical Informatics: Computer Applications in Health Care*. Addison-Wesley Publishing Company, Inc., Chapter 7, 219-243, 1990.

McDonald, C., Barnett, G. "Medical-Record Systems.", Shortliffe, E., Perreault, L., Wiederhold, G., and Fagan, L. (1990) *Medical Informatics: Computer Applications in Health Care*. Addison-Wesley Publishing Company, Inc., Chapter 6, 181-218, 1990.

McGoveran, D. "Two-Phased Commit or Replication?", *Database Programming & Design*, May, 34-44, Miller Freeman Inc., San Francisco CA, 1993.

McNeill, K. M., Capp, M. P., Dallas, W. J., Martinez, R. "Standardization of Network Addressing in Picture Archiving and Communications Systems Utilizing the ISO OSI protocols.", *Computer Methods and Programs in Biomedicine*, V. 37 n 4 May, 311-318, 1992.

McNitt-Gray, M. F., Pietka, E., Huang, H. K. "Preprocessing Functions for Computed Radiography Images in a PACS Environment", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 94-102,

SPIE-The International Society for Optical Engineering,
Newport Beach, California, 1992.

McNitt-Gray, M. F., Sayre, J.W., Huang, H. K., Razavi, M. "A Pattern Classification to Segmentation of Chest Radiographs", UCLA Medical Imaging Division and Clinical PACS Project, FEB, 1-11, Los Angeles CA, 1993.

Micallef, J. "Encapsulation, Reusability, and Extensibility in Object-Oriented Programming Languages", Journal of Object-Oriented Programming Languages, APR/MAY v1 n1, 1988.

Minato, K., Komori, M., Nakano, Y., Yonekura, Y., Sasayama, S., Takahashi, T., Konishi, J., Abe, M., Sato, K., Hosoba, M. "Off-Line Image Exchange Between Two PACS Modules Using the "ISAC" Magneto-Optical Disk", Medical Imaging V: PACS System Design and Evaluation, R. G. Jost, Editor, Proc. SPIE 1446, 195-198, 1991.

Moll, T., Genin, G., Diab, C., Akrouit, N., Prost, R., Bremond, A., Jacquemet, P. "Compressed Mammograms by Block Subband Coding: Medical Assessment in the Detection of Microcalcifications", Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan, Volume 14, 1221-1223, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Navio, J., Martí, Fernandez, G., Pulido, N., Dueñas, A., Muñoz, A., Gonzalez, M.A., Salvador, C.H. "An ODA Implementation for Managing Patient Folders in Hospital Transplantation Units", Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan, Volume 14, 1230-1231, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Newling, H. "JOOP Guide to Object-Oriented Consulting : A Comprehensive Guide to Over 150 consultants Throughout the World that Offer OOP-Related Services", Journal of Object-Oriented Programming, vol. 6, No. 5, Sep, 54-60, Sigs Publications Inc., New York, New York, 1993.

Ni, L. M., McKinley, P. K. "A Survey of Wormhole Routing Techniques in Direct Networks", IEEE Computer, FEB, 62-76, 1993.

Owen "", 1982.

Panza, R. "Welcome to RDAville", LAN Magazine, Dec v7 n12, 81-89, Miller Freeman Inc., 1992 .

Pedersen, E. "Distributed Data Management is at Hand", *MIDRANGE Systems*, Dec 8, v5 n23, 37, Professional Press Inc., 1992 .

Perl, Y., Maram, V., Kadakuntla, N. "The Cascading of the LZW Compression Algorithm with Arithmetic Coding", *Data Compression Conference*, Apr, 277-286, IEEE Computer Society Press, Los Alamitos CA, 1991.

Peterson, H.A., Peng, H., Morgan, J.H., Pennebaker, W.B. "Quantization of Color Image Components in the DCT Domain", *Human Vision, Visual Processing, and Digital Display II*, SPIE Vol. 1453, 210-223, 1991.

Peterson, L., Turner, C. " Image Transfer: An End-to-End Design.", *SIGCOMM' 92*, 258-268, , MD,USA, 1992.

Prior, F. W., Nabijee, K. H. "Information Management for Data Retrieval in a PACS", *Medical Imaging III: PACS System Design and Evaluation*, Roger H. Schneider, Samuel J. Dwyer III, R. Gilbert Jost, Editors, , Proc. SPIE 1093, 488-497, 1989.

Pronios, N.B., Yovanof, G.S. "Effects of Transmission Errors on Medical Images", *Medical Imaging V: PACS System Design and Evaluation*, R. G. Jost, Editor, Proc. SPIE 1446, 108-128, 1991.

Quillin, E. "", *Medical Diagnostic Imaging System Manager at Wright-Patterson AFB, OH*, Personal Interview, Nov, 1992.

Raghuram, A., Morgan, T.W., Rajaraman, B., Ronen, Y. "A Model for the Performance of Locking Algorithms in Distributed Database Systems", *Phoenix Conference on Computers and Communications*, Tenth Annual International March 27-30, 137-144, IEEE Computer Society Press, Scottsdale, Arizona, 1991.

Ratib, O., Ligier, Y., Girard, C., Perrier, R., Logean, M., Vurlod, J.F. "A Picture Archiving and Communication System Based on an Open Distributed Architecture", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan , Volume 14, 1197-1198, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Resnikoff, H. L. "A Guide to Wavelets and Their Application", *Curves and Surfaces in Computer Vision and Graphics III*, Joeseeph D. Warren, Editor, Proc. SPIE 1830, 13-22, 1992.

Rosenfeld, A., Kak, A.C. Digital Picture Processing, (Second Edition) Volume 1, 116-208, Academic Press, Inc., San Diego CA, 1982.

Roth, M.A., Van Horn, S.J. "The Advantages of Data Compression", *Database Programming & Design*, July, 54-60, Miller Freeman Inc., San Francisco CA, 1993.

Routhier, J. "Entity-Relationship Versus Object-Oriented Modeling and the Underlying DBMS", *Graduate Thesis - AFIT*, Dec, 1992.

Rowberg, A. H. "Need for Quality Assurance Related to PACS", *Medical Imaging V: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1444, 80-82, 1991.

Saito, T., Kudo H. "A Reversible Data Compression Scheme for CT Image Archiving", *Medical Imaging III: Image Capture and Display*, Roger H. Schneider, Samuel J. Dwyer III, R. Gilbert Jost, Editors, Proc. SPIE 1091, 60-67, 1989.

Salazar, A., Forero, M.G., Goudin, S., Langevin, F., Fauchet, M. "A Low Cost Image Transfer System for Small Medical Centers", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan, Volume 14, 1204-1205, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Saranummi, N., Akisada, M., Irie, G., Kiuru, A., Mattheus, R., Okabe, T., Olsson, S., Viitanen, J. "User Requirements and standards for PACS. 2nd Japan-Nordic PACS Symposium.", *Computer Methods and Programs in Biomedicine*, V. 37 n 4 May, 237-245, 1992.

Sayre, J., Aberle, D. R., Boechat, M. I., Hall, T. R., Huang, H. K., Ho, B. K., Kashfian, P., Rahbar, G. "Effect of Data Compression on Diagnostic Accuracy in Digital Hand and Chest Radiography", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 232-240, SPIE-The International Society for Optical Engineering, Newport Beach CA, 1992.

Sayre, J.W., Lee, J.S., Stewart, B.K., Cox, G.G., Alberle, D.R., Dwyer, S.J. "Techniques for Multiple-Signal Multiple-Reader Evaluations", *UCLA Medical Imaging Division and Clinical PACS Project*, FEB, 83-96, Los Angeles CA, 1993.

Sedgewick, R. Algorithms, Addison-Wesley Company Inc., Reading MA, 1983.

Seiter, C. "The Big Squeeze", *MACWORLD*, January, 128-135, 1991.

Seiter, C. "Image Compression Matures", MACWORLD, March, 146-151, 1992.

Shah, A.A., Fotouhi, F., Grosky, W. "Renovation of Complex Objects in the Temporal Object System", Phoenix Conference on Computers and Communications, Twelfth Annual International, 203-209, Institute of Electrical and Electronics Engineers, Inc., Temple, Arizona, 1993.

Sheu, P.C.-Y., Silver, D. "Extending Object-Oriented Databases With Problem Solving and Visualization", Computers & Graphics, Vol. 17, No. 4, 447-455, Pergamon Press Ltd., Great Britan, 1993.

Shindyalov, I., Bourne, P. "The Shape of Databases to Come", DEC Professional, NOV v11 n12 , 38, 1992 .

Silberschatz, A., Stonebraker, M., Ullman, J.D., Eds "Database Systems: Achievements and Opportunities", Report of NSF Workshop in Palo Alto, published as a Technical Report TR-90-22, FEB, Computer Science Department, University of Texas, Austin TX, 1990.

Silberschatz, A., Peterson, J. L., Galvin, P. B. Operating System Concepts, (Third Edition), Addison-Wesley Company Inc., Reading MA, 1991.

Simon, B. "How Lossy Compression Shrinks Image Files", PC Magazine, July, 371-382, 1993.

Sirota, W. "Compression Engines", PC World, Nov v11 n11, 66-72, 1993.

Smith, D.V., Smith, S., Cawthon, M. "MDIS Workstation Issues: Clinical Perspective", Medical Imaging V: Image Capture, Formatting, and Display, Yongmin Kim, Editor, Proc. SPIE 1444, 357-362, 1991.

Sriram, P., Marcellin, M. "Image Coding For Telecommunications", Phoenix Conference on Computers and Communications, Twelfth Annual International, 376-377, Institute of Electrical and Electronics Engineers, Inc., Temple AZ, 1993.

Stallings, W. ISDN an Introduction, Macmillan Publishing Company, New York, New York, 1989.

Stern, T. "An Algorithm for the Construction of a Critical Clustered Exponential Code for Use in Image Data Compression", Phoenix Conference on Computers and Communications, Tenth Annual International March 27-30, 747-753, IEEE Computer Society Press, Scottsdale AZ, 1991.

Stewart, B.K., Mankovich, N.J., Sayre, J.W., Dwyer, S.J., Gold, R.H., Kangarloo, H. "Technical Development of Shared Radiology Educational Resources via Teleconferencing", *UCLA Medical Imaging Division and Clinical PACS Project*, FEB, 57-68, , Los Angeles CA, 1993.

Stonebraker, M. Readings in Database Systems, Morgan Kaufman Publishers, Inc., San Mateo CA, 1989.

Stonebraker, M. "Architecture of Future Data Base Systems", *IEEE EH0336*, AUG , 202-207, 1991.

Storer, J.A. Data Compression: Methods and Theory, Computer Science Press, Rockville MD, 1988.

Storer, J.A., Reif, J.H. Proceedings of Data Compression Conference, Mar, Snowbird UT, 1991.

Storer, J.A., Cohn, M. Proceedings of Data Compression Conference, Mar, Snowbird UT, 1992.

Stytz, M. R., Frieder, O. "Three-Dimensional Medical Imaging Modalities: An Overview", *Critical Reviews in Biomedical Engineering*, Volume 18 Issue 1, 27-54, 1990.

Stytz, M.R., Frieder, O. "Computer Systems for Three-Dimensional Diagnostic Imaging: An Examination of the State of the Art", *Critical Reviews in Biomedical Engineering*, Volume 19, Issue 1, 1-45, 1991.

Stytz, M.R., Frieder, G., Frieder, O. "Three-Dimensional Medical Imaging: Algorithms and Computer Systems", *ACM Computing Surveys*, v 23 n 4 DEC, 421-499, 1991.

Suarez, P.F. "Face Recognition with the Karhunen-Loève Transform", *Thesis, AFIT/GE/ENG/91D-54*, 1991.

Taft, D.K. "Military Hospitals Expect Imaging to Save Them Millions", *Government Computer News*, 2 MAR 1992 v11 n5, 36-37, 1992.

Taira, R.K., Chan, K.K., Stewart, B.K., Weinberg, W.S. "Reliability Issues in PACS", *Medical Imaging V: PACS System Design and Evaluation*, R. G. Jost, Editor, Proc. SPIE 1446, 451-458, 1991.

Taira, R.K., Breant, C.M. "The Knowledge-Based Object-Oriented PICQUERY Language", *UCLA Medical Imaging Division and Clinical PACS Project*, FEB, 69-82, , Los Angeles CA, 1993.

Takizawa, M., Ando, Y., Imai, M., Imamura, K., Kuranishi, M., Mori, Y., Maeda, T., Inamura, K., Akisada, M. "User and Manufacturer's Requirements for IMAC Standardization in

Japan", *Computer Methods and Programs in Biomedicine*, v37 , 247-252, 1992 .

Tanenbaum, A.S. Computer Networks, (Second Edition), Prentice-Hall, Englewood Cliffs NJ, 1988.

Thomson, D. "Interfacing Objects With The Relational DBMS", *Database Programming & Design*, August, 32-41, Miller Freeman Inc., San Francisco CA, 1993.

Trefler, M., Russell, E. "Ongoing Evaluation of PACS in a Clinical Setting", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 466-469, SPIE-The International Society for Optical Engineering, Newport Beach, California, 1992.

Tsichritzis, D., Lochovsky, F.H. Data Models, Prentice-Hall, Englewood Cliffs NJ, 1982.

Tsur, Shalom "Data Dredging", *IEEE*, Aug , 242-247, 1991 .

Udupa, J. K., Hung, H. M., Odhner, D., Goncalves, R. J. "3DVIEWNIX: A Data-, Machine-, and Application-Independent Software System For Multidimensional Data Visualization and Analysis", *Medical Imaging VI: Image Capture, Formatting, and Display*, Yongmin Kim, Editor, Proc. SPIE 1653, 185-191, SPIE-The International Society for Optical Engineering, Newport Beach CA, 1992.

Ullman, J. Principles of Database and Knowledge-Based Systems, Vol 1 & 2, , Computer Science Press, Rockville MD, 1989.

Umeda, T., Inamura, K. Inamoto, K., Kondoh, H., Kozuka, T. "Multi-Media PACS Integrated with HIS/RIS Employing Magneto-Optical Disks", *Medical Imaging V: PACS System Design and Evaluation*, R. G. Jost, Editor, Proc. SPIE 1446, 199-210, 1991.

Urbano, L., Gibaud, B., Coatrieux, J.L., Duvauferrier, R., Lucas, C. "3-Dimensional Medical Image Compression : A First Approach to the Application of the ADCT-ISO", *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Edited by Jean Pierre Morucci, Robert Plonsey, Jean Louis Coatrieux and Swamy Laxminarayan , Volume 14, 1219-1220, Institute of Electrical and Electronics Engineers, Inc., Paris France, 1992.

Valentino, D.J., Manager, Clinical PACS Project, Department of Radiological Sciences at University of California at Los Angeles CA, Personal Interview, Aug, 1993.

van Gennip, E. M. S. J., Bakker, A. R. and Greberman, M.
"Do the Benefits Outweigh the Costs of PACS? The Result of
an International Workshop on Technology Assessment of PACS.
", *Computer Methods and Programs in Biomedicine*, V. 37 n 4
May , 265-271, 1992.

Van Name, M.L., Catchings, B. "Two-phase Commit: Don't
Distribute Data Without It", *PC Week*, Nov 30 v9 n48, 80,
Ziff-Davis Publishing Company, 1992 .

Van Name, M.L., Catchings, B. "Distributed Databases: Don't
Believe All You Hear", *PC Week*, FEB v10 n5 , 63, Ziff-Davis
Publishing Company, 1993.

Varma, S. "Objects and Databases: Where Are We Know?",
Database Programming & Design, May, 60-64, Miller Freeman
Inc., San Francisco CA, 1993.

Vasan, R. "Relational Databases and Objects: A Hybrid
Solution", *Object Magazine*, Jan-Feb, 41-44, Sigs
Publications Inc., New York NY, 1993.

Vetter, R. J., Du, D. H. C. "Distributed Computing with
High-Speed Optical Networks", *IEEE Computer*, FEB, 8-18,
1993.

Wallace, G.K. "Digital Image and Video Standards",
Communications of the ACM, APR v34 n4 , 31-44, 1991.

Waltz, M. "Keeping Multiple Images Organized.", *MacWeek*,
March v6 n11, 59-60, Coastal Associates Publishing L. P.,
1992.

Wang, J., Huang, H.K. "Film Digitization Aliasing Artifacts
Caused by Grid Line Patterns", *UCLA Medical Imaging Division
and Clinical PACS Project*, FEB, 40-50, Los Angeles CA, 1993.

Yang, Y. "Change Management in Object-Oriented Databases",
Phoenix Conference on Computers and Communications, Twelfth
Annual International, 238-244, Institute of Electrical and
Electronics Engineers, Inc., Temple AZ, 1993.

Yonekawa, H., Chuang, K.S., Taira, R.K., Huang, H.K.
"Scatter Removal Using Modified Air-Gap Method: Spectral
Composition Technique", *UCLA Medical Imaging Division and
Clinical PACS Project*, FEB, 28-39, Los Angeles CA, 1993.

Zaniolo, C., Ait-Kaci, H., Beech, D., Cammarata, S.,
Kerschberg, L., Maier, D. "Object-Oriented Database Systems
and Knowledge Systems", *Expert Database Systems, Proceedings
from the First International Workshop*, L. Kerschberg, Ed.,
Benjamin/Cummings, Redwood City CA, 1986.

Zhao, W., Srinivasan, B., Avram, C. "Prototyping Distributed Applications Based on RPC", Phoenix Conference on Computers and Communications, Twelfth Annual International, 517-523, Institute of Electrical and Electronics Engineers, Inc., Temple AZ, 1993.

Zip, J., Lempel, A. "A Universal Algorithm for Sequential Data Compression", *IEEE Transactions on Information Theory*, Vol. IT-23, 337-343, 1977.

Zip, J., Lempel, A. "Compression of Individual Sequences via Variable Rate Coding", *IEEE Transactions on Information Theory*, Vol. IT-24, 530-536, 1978.

REPORT DOCUMENTATION PAGEForm Approved
OMB No. 0704-0188

Report Number 29, 298 for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE April 1994	3. REPORT TYPE AND DATES COVERED Final
4. TITLE AND SUBTITLE Global Picture Archiving and Communications Systems (GPACS): An Overview			5. FUNDING NUMBERS None.
6. AUTHOR(S) 1Lt Russell C. Ellis			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AFIT/ENG			8. PERFORMING ORGANIZATION REPORT NUMBER GCS/ENG/94J-02
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSORING / MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release: Distribution Unlimited			12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 words) The world is currently experiencing a computer revolution. With the expansion and development of new computer technology, the age of video is upon us. The medical world is also currently undergoing a dramatic change. With the development and perfection of non-invasive techniques of probing the body, large amounts of computer data is collected and much of it is used to generate images. The organization and management of these images is the topic of this thesis. The discussion of storage schemes, databases, communications and data compression techniques gives the reader a background and some insight to the current state of the art in Picture Archiving and Communication Systems (PACS) and how this schema will some day be expanded to a global scale called Global PACS (GPACS).			
14. SUBJECT TERMS Database, encryption, computer networks			15. NUMBER OF PAGES 174
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited