HOLOGRAPHIC IMAGE ARCHIVE[1]

Javed I. Khan and David. Y. Y. Yun

Contact:

Laboratories of Intelligent and Parallel Systems
Department of Electrical Engineering
492 Holmes Hall, 2540 Dole Street
University of Hawaii at Manoa
HI-96822, USA

Phone: (808)-956-3868
Fax: (808)-941-1399
javed|dyun@hawaii.edu

SUMMARY

This paper presents an associative technique for content-based retrieval into image archive, based on a new computing paradigm called **Multidimensional Holographic Associative Computing** MHAC). Unlike any prior Artificial Associative Memory (AAM), MHAC has the unique ability to focus on any subset of pixels in the sample image and retrieve learned images based on the similarity of the visual objects. In addition, MHAC is adaptive, graciously accommodative of imprecision, efficient, parallelizable, scalable, and optically realizable. Together, these excellent properties of MHAC offer a promising novel approach to content-based search into massive image archives. The paper presents the needed transformational steps to incorporate this new mechanism into a complete image archival and retrieval system. This is the first associative search approach for content-based retrieval in image repository. The results show that this search system is capable of retrievals by using pattern objects as small as 15-10% of the query image frame at better than 90% accuracy. This demonstrates the potential of MHAC for handing content-based image applications far beyond the capability of current associative memories. The design, methodology and performance of this system have been illustrated in this paper through its application in managing a **Medical Image Archive** (MEDIA).

1 This paper has been accepted by the **Journal of Computerized Medical Imaging and Graphics**, and will appear in the special issue on Medical Image Databases in April 1996.

HOLOGRAPHIC IMAGE ARCHIVE

Javed I. Khan and David. Y. Y. Yun

Laboratories of Intelligent and Parallel Systems
Department of Electrical Engineering
University of Hawaii at Manoa, USA
javed|dyun@eng.hawaii.edu

ABSTRACT

This paper presents an associative technique for content-based retrieval into an image archive based on a new computing paradigm called Multidimensional Holographic Associative Computing (MHAC). Unlike any prior Artificial Associative Memory (AAM), MHAC has the unique ability to focus on any subset of pixels in the sample image and retrieve learned images based on the similarity of the visual objects. The paper presents the required steps to incorporate this new mechanism into a complete image archival and retrieval system and presents the performance of an implemented prototype content-based search system in querying into a medical image archive.

Key Words: Associative Memory, Content-based Retrieval, Attention.

1 INTRODUCTION

Image Information and Content-Based Retrieval: The process of content-based image retrieval (CBIR) can be defined as a technique where images are identified from a part of it. This part generally refers to an object or a group of objects in the image and the match is based on the similarity between these objects. For instance, a radiologist may want to find all the similar patient cases that have tumor of a certain shape to confirm a diagnosis at hand. In another situation, a radiologist can request for kidneys that are not of the normal shape like the one in a given example image. However, the problem is that these image objects and concepts used as "index" are perceptual and subjective.

The more an image is non graphical[2] (such as medical images), the concepts in it, objects (such as a tumor) and their relationships (such as anatomical, compositional, or spatial relations), tend to be increasingly subjective both in the senses of concreteness of their definability and precision of their measurability (in contrast, concepts in symbolic information can assume much firmer objectivity in both of the above senses).

Existing Retrieval Techniques: Because of the prohibitively high cost of searching into real images, in the recent years a number of techniques have been developed which can perform pseudo-content-based retrieval (such as QBIC [14], IIDS [3], PICQUERY [10], IDB [17]). These approaches rely on symbolic model of the 'content' of images to mediate the search. These techniques mostly use textual representations (natural language or keywords) in terms of a set of pre-defined attributes (color, shape, size, etc.) and their values [19, 8, 6]. Some other use machine detectable features (geometric moments, triangular cover, points of maximum curvature, etc.) to help automating the extraction process [9, 15, 14]. These symbolic descriptions are then stored into some form of data structures (ranging from plain text [8], to complex semantic net [19], or ingenious 2D-string [3]). Once such descriptive models are available, variants of conventional database techniques are used to search it.

However these psudo-content-based retrieval techniques face serious limitations because of their fundamental dependance on such intermediate model of "meaning" [12, 2, 5]. In the first level, it is difficult to design a sufficient language that can contain all the

2 Although, both graphical and natural images are sensed visually, possibly the process involved in the perceptualization is significantly different [12]. In a relative sense, graphical images (such as textual writings, engineering drawing) are more pre-abstracted, and convey refined information.

possible ranges of meaning. In the second level, even assuming the designability of such a language, the process of model extraction or interpretation itself seems to be indefinite and imprecise. Consequently, these psudo-content-based retrieval techniques can be used only for applications where images are near graphical and the archive is slow growing, static and small, allowing the luxury of extensive domain-specific modeling and tedious human involvement. Some innovative technique is much needed to deal with majority applications where none of the above applies.

In this paper we present an alternate approach for content based retrieval which can overcome the above limitations of the psudo-content based retrieval methods. It is based on direct visual similarity and does not require any adhoc intermediate modelling of "meaning", rather allows the individual inquirer (also referred as the user) to attach their own "meaning" dynamically during query. The search engine of this approach is based on a new associative computing model.

Associative Computing: Parallel and distributed models of Artificial Associative Memory (AAM) evolved from the phenomenal advancement of neural network research demonstrate two interesting characteristics which strongly suggest their potential applicability in content-based retrieval [7, 13]. These models are adaptive and thus can avoid the need for explicit symbolic model extraction. Also, computationally they are fast[3], inherently distributed and parallel, and thus can offer direct search ability into images. However, current AAMs still exactly can not be used for searching into image archive because they lack the ability to focus.

Focus, specially its post-learning dynamic specifiability, is necessary for visual query. For example, in a sample CT image depicting an abdominal cross-section, it is up to the inquirer to decide whether the object of interest is the "spinal column" or a "tumor" as his/her basis for similarity. Current AAMs cannot focus on either of them instead converge only to the statistically closest match based on all the pixels in the sample scene. For meaningful search, it is important that the memory be able to focus on demand on any image objects that the inquirer wants to emphasize as the basis for similarity, and if necessary should be able to retrieve different images from the same sample based on the object of focus.

Very recently [11] has demonstrated that a new associative computing mechanism called as *multidimensional holographic associative computing* (MHAC) allows post-learning dynamic specification of attention on the pixel fields in the sample. This model combines the critical capability of attention with the usual advantages of associative computing, and thus for the first time creates the opportunity to use associative computing for the content-based retrieval of images based on object similarity. Its computational mechanism is based on the principles of optical holography.

This paper presents how this new holographic computing mechanism and its ability to perform associative retrieval with changeable attention can be transformed into an image database and retrieval system. This paper is organized as follows. Section 2 first briefly presents the development of MHAC from optical holography. Section 3 then presents the architecture of the developed Content Associative Image Database Search System (CAIDSS), a prototype system that embodies MHAC. The functionality of CAIDSS has been explained through it's managing a medical image database called MEDIA. Sections 4 and 5 explain the important steps of the archival and retrieval process, including data preprocessing, index formulation, archive maintenance and the query formalism. Finally, section 6 provides the computational complexities of the procedures involved in CAIDSS.

3 Various classical pattern matching algorithms exists to find exact match in a pattern string. However these classical methods are impractical for large volume images. Because the multidimensional and indefinite nature of feature space makes it almost impossible to pre-order (sort) images that improves the search efficiency of these methods [18].

2 MULTIDIMENSIONAL HOLOGRAPHIC ASSOCIATIVE MEMORY

2.1 Concept: Bimodal Associative Memory

Let a stimulus pattern is denoted by a symbolic vector $S^{\mu} = \{s_1^{\mu}, s_2^{\mu}, \ldots . s_n^{\mu}\}$. Each of the individual elements in this vector represents a piece of *information* (subscript refers to the element index and superscript refers to pattern index). The values of these elements correspond to a measurement obtained by some physical sensor.

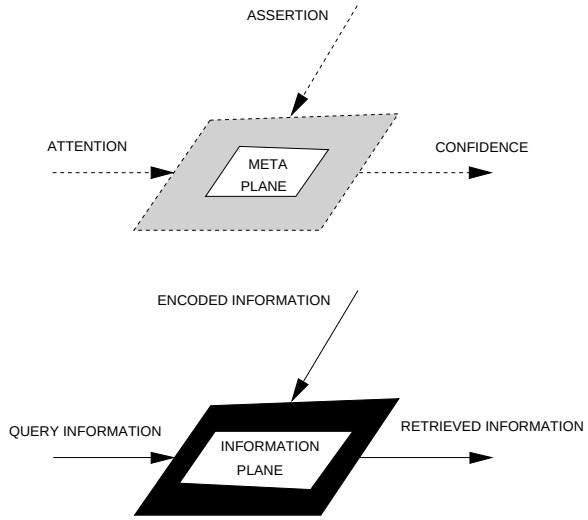


Fig-1 Information Flow Model of Bimodal Memory

A general memory has three information channels (as shown in the bottom plane of Fig-1). The first is the encoder, where information is received during learning. The second is the decoder, where query template information is received from inquirer. The third is the output channel, where for each query, the memory generates the response.

A conventional memory processes only the above measurement components of the involved information elements. In contrast, the novel formalism we would like to propose adopts an additional meta-knowledge plane (as shown in the upper plane in Fig-1) to accommodate dynamic attention.

The linguistic interpretation to the quantities of this meta-plane varies depending on the channel. For the encoding channel this meta-knowledge corresponds to a form of *assertion* from the encoder. For the query pattern it corresponds to a form of *attention* on the part of inquirer. For the memory response it corresponds to the *confidence* on the retrieved information as assessed by the memory itself.

Thus, in this new formalism each information element is modeled as a bi-modal pair $s_k^{\mu} = \{\alpha_k^{\mu}, \beta_k^{\mu}\}$. Where α represents the *measurement* of the information elements and β represents the *meta-knowledge* associated with it.

This memory is expected to maintain the following specific relationships between the *measurement* and the *meta-knowledge* components of information elements:

Inflow Expectation: *During retrieval, more importance should be given to a piece of information that has higher attention value than to a piece with lower attention in the query, and vice verse.*

Outflow Expectation: *If a query element demonstrates high degree of measurement resemblance to that of a priory encoded stimulus pattern, then memory should retrieve the associated response measurement with higher degree of confidence, and vice verse.*

The above expectations constitute the behavioral definition of the memory formalism which incorporates possibility of imperfection in all associated measurements. An associative computational method based on optical holography [4] can realize this special memory.

2.2 Optical Holography

Let us consider a simplified version of the hologram synthesis process. As shown in Fig-2, imagine two wavefronts incident on a medium in plane P. Let, one of these is a plane wave of light (at an angle to the optical axis) indicated by $S = K.e^{-j\alpha x}$, and the other is the wavefront $R(x, y)$, that contains the complex variations to be encoded on the hologram. If the two wavefronts are coherent (such condition is generally met by using a single LASER source, mirrors and beam splitters), then the intensity incident on plane P is given by (bar indicates complex conjugation):

$$\begin{aligned} I(x,y) &= |\, R(x,y) + Ke^{-j\alpha x} \,|^2 \\ &= |\, R(x,y) \,|^2 + K^2 + K\overline{R}(x,y)e^{-j\alpha x} + KR(x,y)e^{+j\alpha x} \\ &= I_A + I_B + I_C + I_D \qquad \ldots.(1) \end{aligned}$$

Consequently, the medium used in plane P will have a transmittance t(x,y) that is proportional to I(x,y).

Interestingly, the encoded wavefront R(x,y) can now be retrieved from the hologram t(x,y), if we reilluminate it by the stimulus wavefront that was used in the hologram recording process. For such illumination, the light leaving the hologram in Fig-2 is given by:

$$O(x,y) = t(x,y).Ke^{-j\alpha x} = K.I(x,y)e^{-j\alpha x}$$
$$= K(K^2 + |R(x,y)|^2)e^{-j\alpha x}$$
$$+K^2\overline{R}(x,y)e^{-j2\alpha x} + K^2R(x,y)$$
$$= (A+B) + C + D \qquad \ldots.(2)$$

Due to the diffraction on the holographic plate three beams leave the holographic plate at three directions. A viewer placed in the direction of D sees only the wavefront R(x,y). This is exactly the one that was supposed to emerge from the real object. This wavefront creates two slightly different images in the two eyes of the viewer, just in the way the actual R(x,y) directly reflected from the object would have created. As a result when the viewer looks towards the holographic film he/she sees the hologram of the actual object in its original place with all its 3-D feel. The same holograph can be used to obtain the wavefront S from R(x,y) in almost same manner. These ideas were originally suggested by Gabor in Late 40's [4]. The actual 3D hologram had to wait 10 years till the perfection of LASER.

Associations between Objects: Hologram can also store association between a pair of objects. If instead of the plain wave, now the light reflected from a second object placed in the scene becomes the stimulus wave $S(x,y)$, then the interference pattern created on P stores the association between these pair of objects. If the plane P is reilluminated by the light reflected from one of the object then the light leaving the plane becomes the one reflected from the other. Hologram has two additional characteristics which make it even more interesting.

Multiple Associations: The first among these is the ability to store multiple associations on the same holographic film. An interference pattern incident from a second pair of objects can now be super imposed on the earlier interference pattern on plane P. If any of these four objects is used to reilluminate the holograph, then the recollected wavefront looks like the one associated with this object. A large number of associations between object pairs thus can be stored on the same photographic film.

Robustness: The other interesting property of hologram is its ability to reconstruct the image from partial information. During recording, each point on plane P receives a reflected beam from all the points of both the objects. At the same time the illuminating beam reflected from each point of both the objects reaches all the points on P. During retrieval, each point of P illuminated by each reflected beam from S(x,y) independently reconstructs the original wavefront R(x,y).

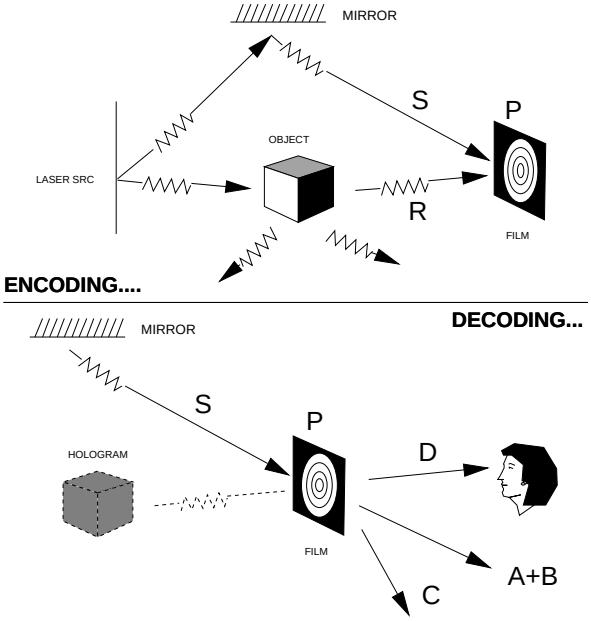


Fig-2 Recording and Decoding in Holography

The effect of such independence is twofold. First, if only a part of the hologram film is used, even then the actual R(x,y) is reconstructed (except for a slight loss of the viewing angle). The hologram also has robustness in a different sense. Instead of dividing the film, if only a part of the stimulus beam S(x,y) is used to illuminate the film (and the rest is blocked), even then the original wavefront R(x,y) appears (except for some loss in brightness).

The Phenomena of Holography: The properties associated with the hologram are extremely intriguing to our senses. Here we will try to list them. In essence a hologram can (i) record and accurately regenerate 3D complex waveforms, (ii) learn associations between two waveforms, (iii) enfold enormous number of multiple associations, (iv) regenerate waveforms from a part of the hologram film, and finally (v) regenerate waveforms even

when excited by partial stimulus wavefront. The implications of these fascinating characteristics for us can be stated in the following way.

It the first property which makes hologram a 3D waveform recording medium (for which it is most famous). The addition of second property makes it a memory. Further addition of the third property makes it an associative memory. Similarly, the fourth property makes it a distributed associative memory. Finally, the last property makes it a distributed associative memory with the ability to focus. Next, the essential computation of holography which can yield all five of these properties digitally is identified.

Essence of the Computation: As evident by now, the key to the reconstruction process lies in the 3rd and 4th terms of (1). A real hologram operates with two constraints; the first one is (i) that the optical setup performs addition of two wavefronts and the other one is (ii) that the recording medium can store only positive valued functions. But a computer can easily evaluate products and store complex-valued spatial functions. Therefore, the 3rd term is directly captured and the 4th term is evaluated if needed since it is conjugate of the 3rd. The above simplification not only reduces the amount of data that has to be stored in the medium but also eliminates the interference from unwanted terms. The most important fact is that I_D retains all the critical properties of a holography.

2.3 Realization: Computational Representation

In MHAC scheme, a generalized form of the above computational model is adopted to associate one response label pattern (RLP) with every image to be stored. In this scheme, RLP is substituted for R, and the image is substituted for S. Later during recollection, the image or any of it's parts is used to reconstruct the full RLP.

In holography an element of the pattern is represented as a 2D complex value. We adopt a generalized formalism and represent the elements in the form of *multidimensional complex numbers* (MCN) as a point on the surface of a hypersphere. In this formalism the MCN phase represents the *measurement* and magnitude represents the *meta-knowledge* component of each piece of information. Thus:

$$s_k = (\alpha_k, \beta_k) \Rightarrow \lambda_k e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,k}\right)} \ldots. \quad (3)$$

Here, each α_i^k is mapped onto an angular span defined by the phase elements θ_i^k in the range of $\pi \geq \theta \geq \pi$ through a suitable transformation, where each $s(\lambda_k, \theta_1^k, \theta_2^k, \ldots \theta_{d-1}^k)$ is a d-dimensional vector. Each of the θ_j^k is the spherical projection (or phase component) of the vector along the dimension $\hat{i}_j$. λ_k becomes the magnitude of this vector. Thus, a stimulus with *n* elements and a response with *m* elements are represented respectively as:

$$[S^\mu] = \left[\lambda_1^\mu e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,n}^\mu\right)}, \lambda_2^\mu e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,n}^\mu\right)}, \ldots. \lambda_n^\mu e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,n}^\mu\right)} \right]$$

$$[R^\mu] = \left[\gamma_1^\mu e^{\left(\sum_j^{d-1} \hat{i}_j \phi_{j,1}^\mu\right)}, \gamma_2^\mu e^{\left(\sum_j^{d-1} \hat{i}_j \phi_{j,2}^\mu\right)}, \ldots. \gamma_m^\mu e^{\left(\sum_j^{d-1} \hat{i}_j \phi_{j,m}^\mu\right)} \right]$$

Here, in the response pattern the phasor ϕ represents the measurement of the expected associated response pattern elements from the memory system, and γ represents the expected system confidence on ϕ.

2.4 Encoding

The association between each individual stimulus and its corresponding response is analogous to I_D and is computed as following:

$$[X^\mu] = [\overline{S}^\mu]^T \cdot [R^\mu] \quad \ldots.(4)$$

The associations derived from a set of stimuli and a set of corresponding responses are superimposed on a super matrix X which is referred as Holograph.

$$[X] = \sum_\mu^P [X^\mu] = \sum_\mu^P [\overline{S}^\mu]^T [R^\mu] \quad \ldots.(5)$$

2.5 Retrieval

During recall, an excitory stimulus pattern $[S^e]$ is obtained from the query pattern:

$$[S^e] = \left[\lambda_1^e e^{\left(\sum_j^{d-1} i_j \theta_{j,1}^e\right)}, \lambda_2^e e^{\left(\sum_j^{d-1} i_j \theta_{j,2}^e\right)}, \ldots. \lambda_n^e e^{\left(\sum_j^{d-1} i_j \theta_{j,n}^e\right)} \right]$$

The decoding operation is similarly performed as following:

$$[R^e] = \frac{1}{c}[S^e] \cdot [X] \quad \text{....(6)}$$

$$where, \quad c = \sum_{k}^{n} \lambda_k$$

The basic associative memory characteristics of this model explaining how equations (4), (5), and (6) together can correctly retrieve original stored response despite superimposition of the associations in (5) is given in [11].

As expected, this new computational model has advantages similar to the conventional AAMs. In fact, as shown in equation (6), a search into thousands of stored images through this technique requires a single complex matrix multiplication. The most important capability of post learning dynamic focus of MHAC is given by the external adjustability of each individual magnitude of the query stimulus elements.

3 DESIGN OF IMAGE RETRIEVAL SYSTEM

3.1 System Overview

This section now demonstrates the detailed steps that have been developed to transform the technique of MHAC and its ability to perform associative retrieval with changeable attention into a complete search mechanism for content based querying into image database. A prototype system called *Content Associative Image Database Search System* (CAIDSS) has also been implemented. CAIDSS involves three principal components; (i) storage and access subsystem (SAS), (ii) associative encoding subsystem (AES), and (iii) associative query subsystem (AQS). Fig-3 presents the schematic.

The *storage and access subsystem* (SAS) is concerned with space efficient compact storage of the images and access speed. It is functionally independent from the rest of the search system, and will not be discussed further.

The *associative encoding subsystem* (AES) is concerned with the generation and maintenance of the holographic archive abstract of the images. It involves the following major steps, (i) image pre-processing and stimulus pattern generation, (ii) Generation and assignment of *response label patterns* (RLP), (iii) Holograph training, and (iv) Holograph characterization.

The *associative query subsystem* (AQS), allows a direct content based search into the image archive. The principal responsibilities of this subsystem are; (i) query interface, (ii) search optimization, and (iii) associative recall. It also communicates with the SAS if physical access to the image is demanded. The functionalities of AES and AQS are illustrated in the next two sections.

4 ENCODING PROCESS

4.1 Stimulus Generation

The principal design issues at this stage are: (i) domain knowledge based *semantic content refinement*, and (ii) system architectural constraint based *data-mapping*.

Semantic refinement: The objective of this step is to remove the undesirable and to enrich the desirable components of information. For medical images, the regions can generally be segmented into regions of (a) bone tissue, (b) soft tissue, (c) fat tissue, (f) air. Typically, queries are based on the first two or three of these segments. At the preprocessing stage of MEDIA, the input images are trimmed off from the air segment identified with a Gaussian classifier. MEDIA sets assertion valueλ=0.05 for pixels corresponding to the air segments (mainly dark backgrounds with no reflectance of x-ray), which is typically large for CT-SCAN or MRI images but bears no cognitive importance.

Data mapping: The MCN representation specifically requires; (i) dynamic range of the input data distribution to be mapped into the 0-2π range, and (ii) sufficient symmetrical distribution. In the raw format, an image is represented by a two dimensional array of pixels with discrete intensity values in the range of 0-255. Typically image demonstrates normal distribution of intensity. The spiral mapping transform specified in equation (7) has been used to obtain symmetric distribution in the range of -π to +π.

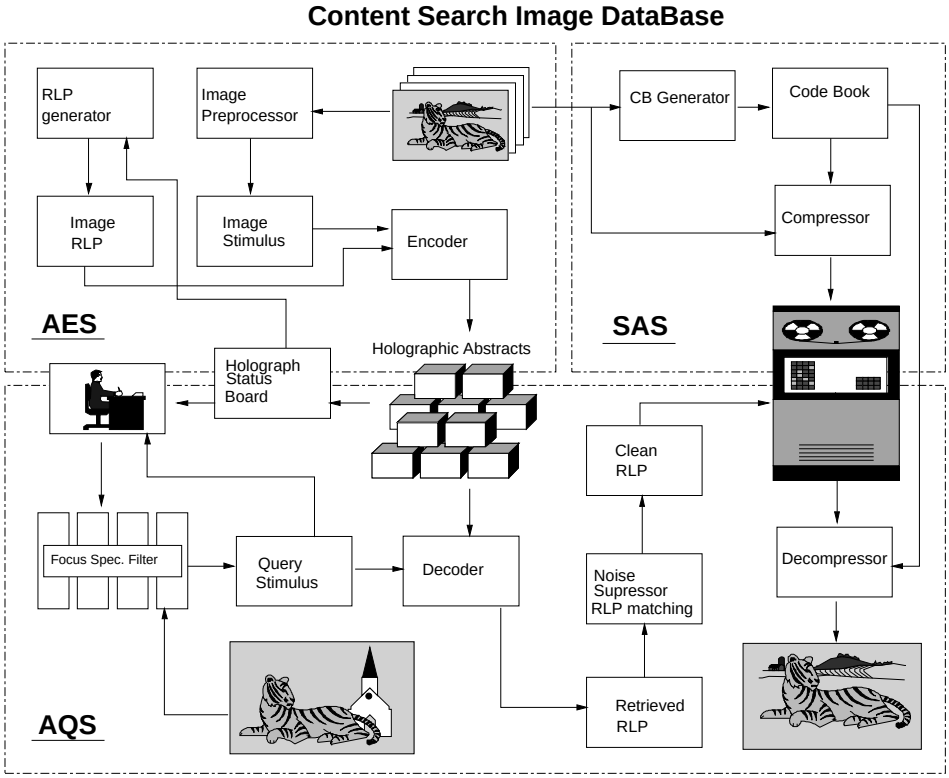


Fig-3 CAIDSS Architecture

$$\theta_i = m(I_i) = I_i.k_{spread}.\frac{2*\pi}{I_{\max} - I_{\min}} - \pi \qquad \ldots.(7)$$

k_{spread} is the spreading coefficient. To maintain the uniqueness of mapping in 2^n level intensities it should be an odd number (2 is the only prime factor). Other functions can also be used besides (7) as long as they are not dependant on image frame characteristics. In object oriented content-based retrieval, one of the basic possibility is that the query image frame may differ considerably from the encoded one outside the region of focus. Therefore, functions such as sigmoidal mapping [16], or histogram equalization [13], which are often used in neural networks for data mapping, should not be used here.

4.2 Response Label Associations

Response Label Patterns (RLP) are internal index used by the Associative Query System (AQS) to identify images. The principal design issues are (i) the length of RLPs. (ii) and the assignment of RLPs.

RLP length: The length of the RLP (m) is decided by number of image frames (P) and the quantization interval in the dynamic range of θ. Given a quantization interval q, the length is given by equation (8). The quantization interval q depends on the expected loading of the holograph.

$$m = \log_{\left(\frac{2\Pi}{q}\right)}(P) \qquad \ldots.(8)$$

RLP assignment: A simple sequential assignment (such as assigning intervals 1,2,3,.. to images 1,2,3) of RLPs tends to asymmetrically overcrowd the dynamic range of θ for a holograph that contains fewer patterns than the designed limit computed by (8). To avoid such overcrowding, a Reverse Modulus Weight code (RMWC) has been devised to assign RLPs to the images. Fig-4 shows three stages of a valid assignment sequence. Such RWMC always ensures maximum separability among the

assigned RLPs, and thus reduces error. The technique of obtaining such symmetric assignment is to maximize the distance between each consecutive code pair. This is done by altering the maximally weighted digit. If *u* is a sequence number then it is mapped to a polar phase *v* by the transform (9):

$$u = a_2.d^2 + a_1.d^1 + a_0.d^0$$

$$\Rightarrow (a_2.d^0 + a_1.d^1 + a_0.d^2) * \frac{2\Pi}{d} = v \qquad \ldots.(9)$$

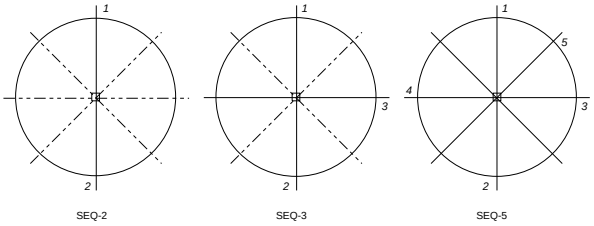


Fig-4 RMWC Code

4.3 Training Algorithm

The training is performed in three stages and corresponds to equations (4) and (5) and is given in Fig-5. A differential training approach is adopted [16], where before encoding, in the first stage, the stimulus *S* is decoded for performance improvement.

The prototype holographic archive called MEDIA encodes 64 CT scan and MRI images each of size 256x256 pixels. Fig-11 shows some samples. Fig-6 shows the convergence characteristics of average recall SNR (left y-axis) and *mean normalized confidence* (MNC) (right y-axis) as the training (x-axis) progressed. As can be seen, the encoding process attained the peak performance within only 20 iterations.

Define complex patterns ***RR, IMG, R, R***$_{diff}$*;*
Define holograph ***H, A***$_{diff}$*;*
a. Initialize the Holograph Randomly;
b. Set t=0; iter=1 {
c. While (t < ALL-IMAGES) {
i. Read the t^{th} *RLP,* ***RR=ReadPattern(t)****;*
ii. Read the t^{th} *image,* ***IMG=ReadImage(t)****;*
iii. Convert the Image ***IMG*** *to stimulus pattern* ***S****;*
iv. Decode for ***R*** *by applying* ***S*** *to* ***H****,* ***R=S.H****;*
v. Compute the difference, ***R***$_{diff}$***= RR - R****;*
vi. Generate the difference association
A$_{diff}$***= conjugate(transpose(S)).R***$_{diff}$*;*
vii. Super impose the association on the holograph
H*= alpha*****H****+****A***$_{diff}$*;*
viii. Set t=t+1; Repeat from step c; }
d. Compute average recall error
E *for all images;*
e. Set t=0; iter=iter+1;
f. If (E > EPSILON) Repeat from step b; }
g. Save the holograph;

Fig-5 Training Procedure of CAIDSS

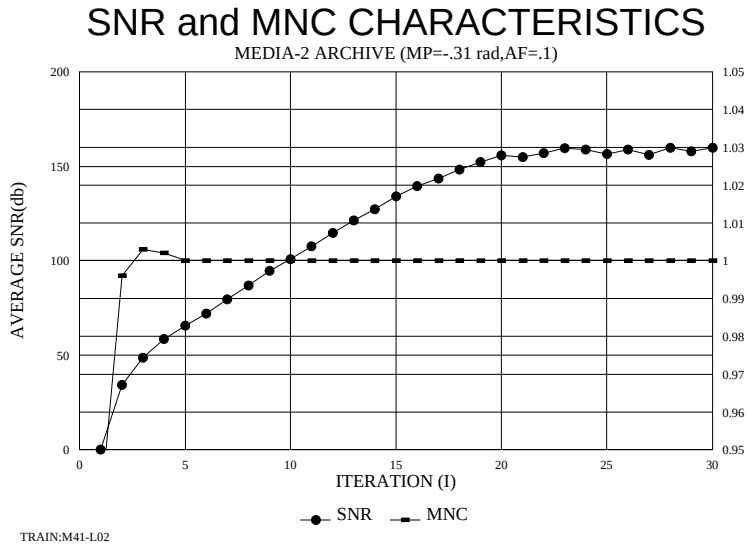


Fig-6 Training Characteristics of MEDIA

4.4 Holograph Characterization

Two tests (called ZOOM test and SCAN test) have been developed to monitor the status of holographs. ZOOM test probes the distinguishability among the encoded images by applying focus windows of decreasing dimensions at the center of the template. SCAN test probes the distinguishability at various spatial locations with a fixed size focus window.

The six pulse-trains in Fig-7 show the accuracy of each pattern recall for six widow sizes (70x70,80x80,90x90,100x100, and 110x100) during ZOOM test (wrong detection triggers a jump in the pulse). As shown, it retrieves each of these images correctly when the window is 100x100 pixels wide (pulse train C3). When the window drops to 70x70 there have been five wrong detections (image 9, 38, 41, 43, 45) out of 64 images. Fig-8 summarizes the average accuracy. The line-plot (right y-axis) shows the attention strength, the bar-plot (left y-axis) shows the average accuracy, and the label on the bar shows the actual number of accurate detections for each of these zoom windows.

The result of SCAN test is similarly shown in Fig-9 and Fig-10. SCAN test probes the distinguishability at five spatial locations with 75x75 focus window (at the image frame center and the centers of its four quadrants). As can be seen, the distinguishability among the images varies (between 3-10 errors, for less than 10% focus) with the region of spatial attention, which is quite natural.

When a holograph is saturated, the tracking tests provides necessary indication for switching to alternate holographs. In conventional database maintenance, occasionally optimization processes are triggered when estimated upper bound of access time exceeds certain threshold. In contrast, for associative archives the main concern is the accuracy of retrieval (access time is almost independent), thus, lower bound of accuracy acts as the trigger. Also during matching, test results help in verifying the correctness of the results or in preempting a query when it is launched into a frame region with low distinguishability.

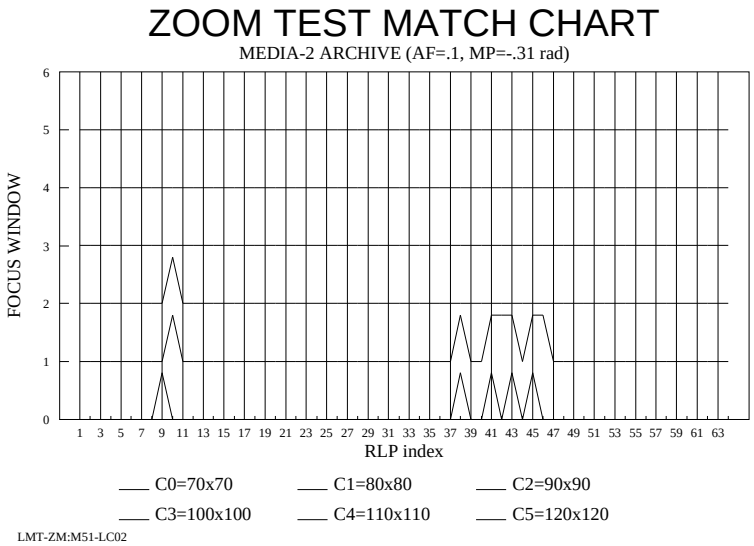


Fig-7 Pattern Distinguishability in ZOOM test

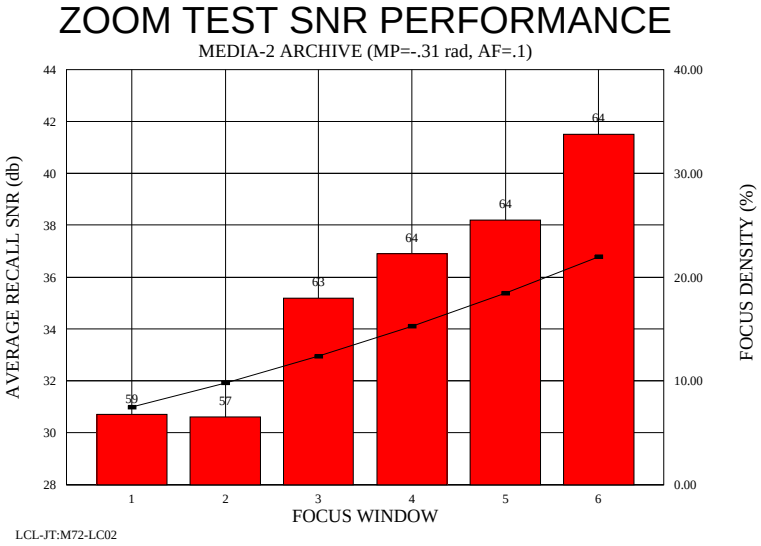


Fig-8 ZOOM test in MEDIA

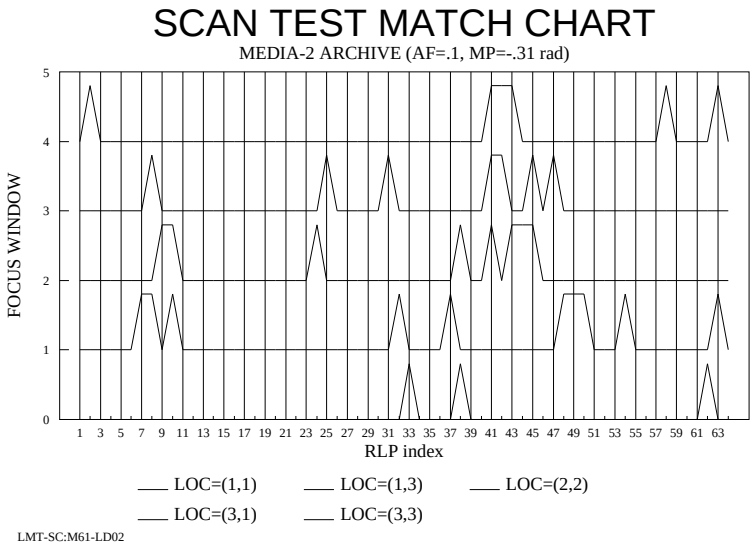


Fig-9 Pattern Distinguishability in SCAN test

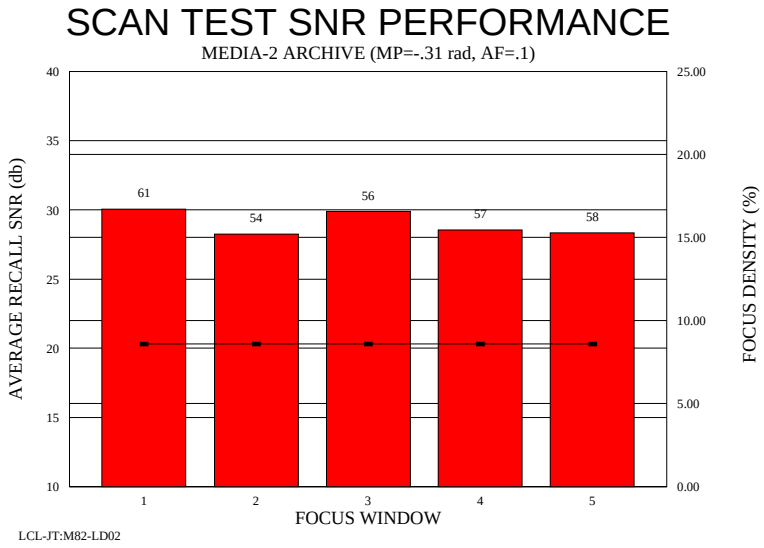


Fig-10 SCAN test in MEDIA

5 DECODING PROCESS

Users' query is conveyed to the AQS through a combination of pictorial examples, focus shields and logical connectives. In the first level, the user specifies a set of elementary objects through defining focus shields on various query image frames. The user can then perform query involving objects or defining composite objects by logically combining these elementary objects. The following sub-sections consecutively discuss (i) the procedure for defining the elementary objects, (ii) specification of complex objects, and finally, (iii) the search process.

5.1 Elementary Objects

A sample scene generally contains many regions of interest. CAIDSS helps the inquirer to decide how he/she wants to define these objects and, which objects to use as the basis of similarity.

Prior to launching a query, the user is expected to specify the elementary basis of his/her search through a set of image and mask pair. Mask represents the attention shield that defines the object inside this example image. In the current implementation, mask is specified by defining a rectilinear window filtering on the spatial and color dimensions. The mask for grey scale images requires 6 parameters (xmax, xmin, ymax, ymin, imax, imin) and color images require 10 parameters to define range boundaries. Table-1 shows the specifications of some objects from Fig-12(a). The attention strength or the size of the individual objects is given by column f and is given in equation (10):

$$f = \frac{\sum_{i}^{n} \lambda_i}{n} \qquad \ldots.(10)$$

5.2 Composite Objects

Once the elementary objects are specified, more compound objects can be defined using logical connectives. The following is an example.

$$(Q1.OBJ1) \cup (\neg .(Q2.OBJ2)) \rightarrow (Q3.OBJ3)$$

In this formalism *OBJ1, OBJ2* are component objects and *OBJ3* is the composite objects. *Q1, Q2, Q3,* ... are fuzzy quantifiers, with values assigned between [0,1]. In reality they can be assigned linguistic multi-valued quantifiers such as *HIGH, LOW*, etc. with a membership function set to translate them into analogue values. The logical union, intersection and complementation operations are defined according to fuzzy set functions given in equation (11).

$$f_{A \vee B}(x) = \max(f_A(x), f_B(x))$$

$$f_{A \wedge B}(x) = \min(f_A(x), f_B(x))$$

$$f_{\bar{A}}(x) = 1 - f_A(x) \qquad \ldots.(11)$$

5.3 Compound Search

Just like the objects, the search can also be logically compounded. Search compounding generally involves multiple decoding. Individual decoding results are logically assimilated in AQS. Compound search allows users to perform searches when the query image differs from the expected image in some procedurally definable abstract sense (such as the sought object may be approximately half the size than the sample).

$$(HIGH.OBJ8.INREGION(region)) \rightarrow (OBJ9)$$

For example, the above query results in a search for OBJ8 in a specific region on the image and refers to specific spatial translation invariancy in 2D. The procedural specification of the term "*INREGION()*" can be given (or selected from library) by the inquirer.

It should be noted that repeated multiple search is not as dreadful as it is in conventional database. In a conventional database, compound search may mean searching enormous number of images repeatedly. Whereas in CAIDSS each search means decoding into one holograph.

5.4 The Search Process

The process of obtaining query stimulus patterns is identical to the encoding process, except that now the magnitudes of the composite elements are obtained from the mask. The raw RLPs are then retrieved by direct associative recall using equation (6). If the specified object is not present in the database, the recalled raw RLP shows distinctively low magnitude of the MCNs. Otherwise, it resembles the actual RLPs of the matching image. A linear search is performed in NSU among the RLPs to find the closest match. The search produces a list of matching RLPs in order of their numerical proximity using the distance measure given below:

$$D(R^Q, R^\mu, \Lambda^Q) = \left[\sum_{i}^{N} \lambda_i^Q dist(r_i^Q, r_i^\mu) \right] \qquad \ldots(12)$$

MASK#	Object	xmax	xmin	ymax	ymin	f
1	Basi-occipital	192	-064	167	-117	.033
2	Foramen Magnum	170	-076	118	-035	.041
3	Foramen Ovale and Spinosul	230	-169	229	-154	.024
5	Jugular Foramen and Carotid Canal	074	-021	146	-093	.048

Table-1 Masks for Objects of Focus

MASK#	Object	Match#	Image	xsft	ysft	SNR (db)	MNC
1	Basi-occipital	1.1 1.2	A35 A33	-60 -10	-90 100	36.21 33.42	0.956 0.895
2	Foramen Magnum	2.1	A33	-10	100	32.27	1.123
3	Foramen Ovale and Spinosul	3.1	A35	-60	-90	33.33	1.092
5	Jugular Foramen and Carotid Canal	5.1 5.2	A34 A33	-130 -10	-20 100	31.92 37.46	1.095 1.183

Table-2 Results of Retrieval

5.5 Query in MEDIA Archive

Fig-12(a) shows a typical sample image which can be used as a pictorial example for query. Let us consider the four objects shown by the focus windows in it (it would not be any different if another inquirer wants to define his/her objects in a different way). Table-1 shows the mask specifications of these index objects. The last column shows the focus strengths (f) of these masks. As evident the typical features or objects with perceptual indices quite often fall below 10-4% of the total image.

Given, the sample image and the mask definitions, the AQS searches at various spatial locations of the MEDIA holograph. During decoding, The matches numbered 1.1 , 1.2, 2.1, 3.1, 5.1, and 5.2 of stored images A33, A35, and A34 (shown in Fig-12(b)) are respectively retrieved by the system from the MEDIA as match. As evident, although none of these stored pictures have statistical similarity with the query image, but each match closely on the basis of respective cognitive objects.

The search process of CAIDSS uses a 10x10 grid resolution for spatial search into MEDIA. The associative retrieval generates response at these locations. Search at each grid point generates a response with a specific MNC value. As explained before, high MNC indicates possible match. The 5th and 6th columns of Table-2 show the grid locations (through the shift values) for which a retrieval MNC greater than 0.5 have been measured. The 7th and 8th columns respectively show the corresponding MNC and the accuracy of retrieved RLPs.

6 SEARCH COMPLEXITIES

Time Complexity: *For CAIDSS, the RCA search complexity for any p-image database of n pixels each, and RLP length d, is* $O(d \cdot n) \approx O(n \cdot \log p)$.

Derivation: Let us consider that the length of RLP is m. The search process involves (a) computation of pattern, (b) holographic decoding, and (c) RLP matching. The complexities of the corresponding stages are as following. The cost of pattern computation = $O(n)$. The cost of holographic decoding = $O(nm)$. This is an inner product matrix operation involving complex matrix multiplication. RLP matching = $O(mp)$. It is a linear search with relatively very small pattern length. Thus, the cost is negligible. Generally, $m \approx O(\log p)$. Thus the complexity of the overall search process is.

$$O(n) + O(n \cdot \log p) + O(p \cdot \log p) = O(n \cdot \log p) \quad \ldots.(13)$$

Space Complexity: *For CAIDSS, the RCA search into a p-image database of n pixels each, and RLP length d, requires* $M = m \cdot d \cdot (n + p)$ *space for holographic encoding.*

Derivation: Let us consider that the length of RLP is *m* and that each complex element requires *d* bytes for representation. The space, required by the holograph is *m.n.d*. Some additional space is also required by the RLPs. Which is *p.m.d*. Thus the total space requirement is:

$$M = m \cdot d \cdot (n + p) \quad \ldots.(14)$$

In practice $p \approx n$, and 4-12 bytes are sufficient for images with 256x256x256 full colors.

7 CONCLUSIONS

7.1 Scalability

Extensive simulations and analysis already conducted have confirmed the scalability of MHAC beyond the MEDIA prototype. As an example, it has been demonstrated that, given reasonable symmetry 1000-2000 images with 4K pixels each can be associatively enfolded on a MHAC memory of 12K bytes and can be recalled with less than 4% error. It should also be observed that much of the involved computations are potentially realizable on optical architecture, which makes this approach even more suited to large scale image databases.

7.2 Bound of Representation

A profound, nevertheless interesting issue pertaining to this new approach is that what pictorial concepts can be searched automatically? One of the critical requirements for our approach is that user's index concept must be expressible in terms of finite set of pixels before any search. Indeed, It is quite impossible to come up with a general pixel-format representation of concepts such as "tree", "hill", "ocean", etc. We believe that this is probably a fundamental limitation of all fully automatic content based search mechanisms. In conventional automatic approaches (pseudo-content-based techniques), this limitation is equally vivid. The initial model extraction requires a filter program to obtain the model. In the absence of such pixel-format representations, there also cannot be such a filter program.

7.3 MHAC and Conventional Approaches

The key difference that distinguishes CAIDSS from the conventional model based approaches lies in the way that they attach "meaning" to the content.

Consider the sample CT-scene of Fig-12(a) that depicts some objects of interest in a horizontal anatomical cross-section of human skull [1]. Here, hardly any region of interest has any fixed definable concrete boundary. (there is nothing sacred in the numbers shown in Table-1). A different interpreter may begin with a different set of objects (such as "the region around *soft palate*"), or some other finer objects (for example, the *foramen magnum* can be spatially decomposed into the *medulla oblangata*, the *vertebral arteries* and *cerebellar tonsils*). The relationships among these objects are also similarly subjective (such as the relationship *cerebellar tonsils* "lies within" *foramen magnum*). Despite such subjectivities in the definitions of these index concepts, an intermediate model based approach has to guess these interpretations (and choices) from the very beginning during model building, which is quite unrealistic[4]. In contrast, CAIDSS avoids any adhoc attachment of such "meaning" during encoding.

It should be distinguished that holographic representation (strictly speaking, another intermediate representation) does not store any meaning but emulates (instead of really searching all the images, which indeed is very expensive) a dynamically weighted search during retrieval.

---

4 The implication of this distinction in handling "meaning" can be best understood, in the concern expressed by one of the pioneers in this field; S. K. Chang in an evaluation report titled "Where Do We Go From Here?" wrote [2]: "Spatial entities (object) and relationships (image features) in image do not carry any semantic meaning by themselves.. Associating semantic meanings by naming will cause some problems with image information. First, the same image could be interpreted in different ways. Second, the same image could be used in different ways during different time periods. ... directly associating semantic meanings to image entities and relationships will severely limit the usage of image information".

However, CAIDSS does not downplay the importance of meaning. A meaning structure is convenient, no matter how subjective it is, to the inquirer, as the search objective becomes complex and sophisticated. CAIDSS approach removes the need for intermediate model, but it empowers user to search the image archive directly with his own perception of meaning into the image.

7.4 The Challenges of Content Based Search

Collective experience over two decades evinces that most likely there is no panacea to solve the complex problem of image information management using image contents. On the whole, content based search is not only a database issue, but it is also closely related to the understanding our own mechanism of representation and perception. Conventional model based approaches tend to be more useful for applications where the image concepts are well defined and extensive domain-specific modeling is possible. Images of graphical natures, such as texts, maps, circuit drawing, etc. tend to fall in such category. In contrast, the approach demonstrated in this work has its advantage when the objects are difficult to describe or model, the content does not show any unambiguously distinguishable structure, the volume of images is enormous, and examples with visual similarity at the object level are available. Medical diagnostics imagery, stellar images, fingerprint, satellite or planetary landscape images tend to fall in this category. It has been long anticipated that associative memories can provide a whole new effective means for content-based image management. The work presented in this paper shows that such an approach is finally realizable.

7.5 Acknowledgement

Finally the authors would like to acknowledge the support of Dr. Jay Cook, Dr. Farhad Ahmed, Dr. Ahmed Jamal, and Mr. Jim Francios. A part of this research has been funded by ACTS and Supercomputing in Remote, Co-operative Medical Triage Support and Radiation Treatment Planning project of ARPA under research grant DABT 63-93-C-0056.

8 REFERENCES

1. Cahill, D. R. & Orland M. J., *Atlas of Human Cross-Sectional Anatomy*, Lea & Febiger, Philadelphia, 1984.

2. Chang, S.K., Arding Hsu, "Image Information Systems: Where Do We Go From Here?", *IEEE Trans. on Knowledge and Data Engineering*, v.4, n.5, October 1992, pp431.

3. Chang, S. K., C. W. Yan, D. Dimitroff, T. Arndt, "An Intelligent Image Database System", *IEEE Trans. on Software Engineering*, v.14, n.5, pp681, May 1988.

4. Gabor, D., "Associative Holographic Memories", *IBM Journal of Research and Development*, 1969, I3, p156-159.

5. Gupta, A., T. Weymouth, R. Jain. semantic Queries in Image Databases, E. Knuth, L. M. Wegner (eds.), *IFIP Trans. on Visual Database Systems*, II, A7, 1992, pp201-215

6. Halin, G., N. Mouaddib, "An Object Oriented Approach to Design a Content-Based Image Retrieval Model", *Image Storage and Retrieval Systems*, SPIE, v.1662, pp100, 1992.

7. Hinton, G.E., J. A. Anderson, *Parallel Models of Associative Memory*, Lawrance Erlbaum, NJ, 1985.

8. Hibler, J. D., C. H. C. Leung, K. L. Mannock, M. K. Mwara, "A System for Content-Based Storage and Retrieval in an Image Database", *SPIE, V1662 Image Storage and Retrieval Systems*, pp80, 1992.

9. Hu, M. K., "Visual Pattern Recognition by Moment Invariants", *IRE Transactions on Information Theory*, IT-8, 1962, pp28-32.

10. Joseph, T., A. F. Cardenas, "PICQUERY: A High Level Query Language for Pictorial Database Management", *IEEE Transactions on Software Engineering*, v.14, n.5, pp630, May 1988.

11. Khan, J. I., "Attention Modulated Associative Computing and Content Associative Search in Images", *Ph.D. Dissertation*, Department of Electrical Enginnering, University of Hawaii, July, 1995.

12. Khan J. I.,& D. Yun, "Searching into Amorphous Information Archive", *International Conference on Neural Information Processing*, ICONIP'94, Seoul, October, 1994, pp739-749.

13. Kulkarni, A. D., *Artificial Neural Networks for Image Understanding*, Van Nostrand Reinhold, New York, 1994, pp15.

14. Niblack, W., R. Barber, W. Equitz, M. Flickner, et al., "The QBIC Project: Querying Images By Content Using Color, Texture, and Shape", *SPIE, v.1908*, pp173, 1993.

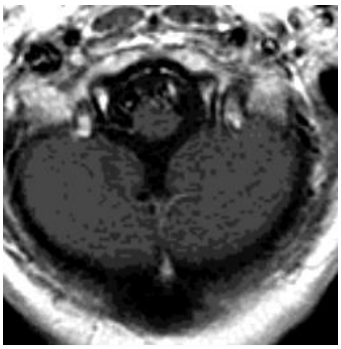
15. Rabitti, F., P. Savino, "Automatic Image Idexation to Support Content-Based Retrieval", *Information Processing & Management* v.28 n.5, pp547, 1992.

16. Sutherland, J., "Holographic Models of Memory, learning and Expression", *International J. Of Neural Systems*, 1(3), pp356-267, 1990.

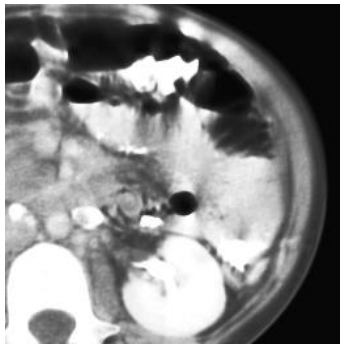
17. Turtur, A., F. Prampolini, M. Fantini, R. Guarda, M. A. Imperato, "IDB: An image database system", *IBM Journal of Research & Development*, v.35, n,1/2, pp88, January/March 1991.

18. Tucci, M., G. Costagliola, S. K. Chang, "A Remark on NP-Completeness of Picture Matching", *Information Processing Letters*, v.39, n.5, pp241, September 1991.

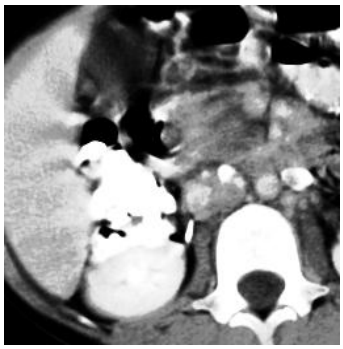
19. Yamane, J., M. Sakauchi, "A Construction of a New Image Database System which Realizes Fully Automated Image Keyword Extraction", *IEICE Trans. Infor. & Sys.*, v.E76-D, n.10, pp1216, October 1993.



SKL−DN:A33:256x256

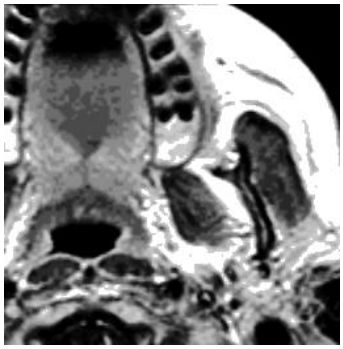


ABD−RT:A26: 256x256



ABD−LT:A25:256x256

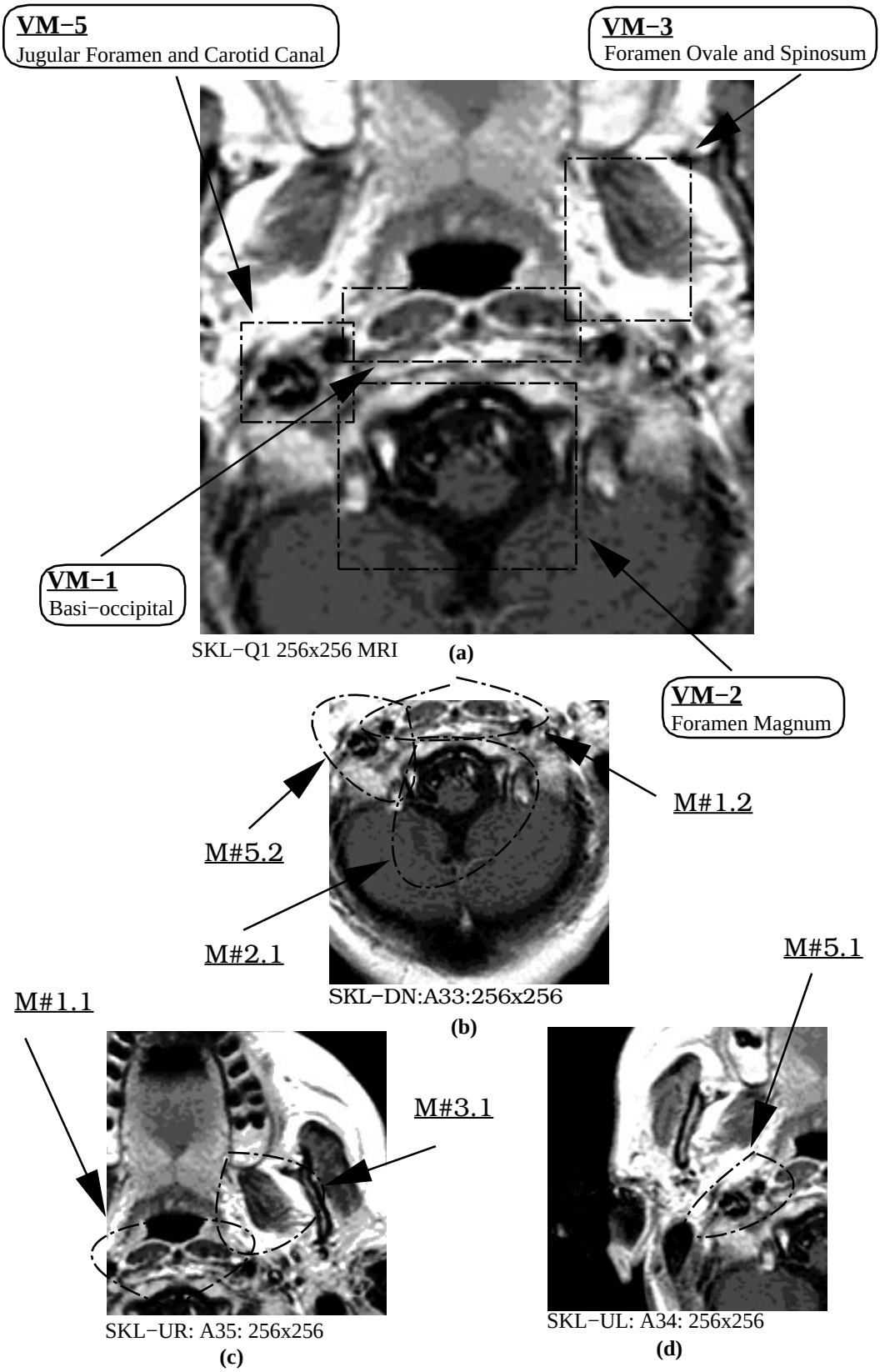
SKL−UL:A34:256x256



SKL−UR:A35:256x256

LNG−LT:A29: 256x256

Fig− 11 Sample Images in MEDIA



SKL−Q1 256x256 MRI **(a)**

SKL−DN:A33:256x256 **(b)**

SKL−UR: A35: 256x256 **(c)**

SKL−UL: A34: 256x256 **(d)**

Fig− 12 Sample Queries in MEDIA (Skull)