A Method for Dynamic Indexing of Large Image Databases

Yohanes Hendra, Raj P. Gopalan & M. Gopalan Nair
School of Computing, Curtin University of Technology
Western Australia
(yohanesh, raj, gopal)@cs.curtin.edu.au

ABSTRACT

The fascinating phenomenon of optical holography has been studied for nearly three decades. A new associated memory model based on the mathematical representation of optical holography was proposed in 1990. This approach was subsequently adapted to the problem of indexing large image archives. For this, trained associative memories based on the holographic model were used. The training process involved several iterations over the stimulus-response pattern associations during the encoding of the memory. This method, however, has severe limitations when it comes to dynamic indexing of images because of the need for retraining the associative memory whenever additions or deletions of images are made in the database.

In this paper we describe an encoding and decoding scheme for indexes of image archives based on a holographic model that does not involve training. As there is no training overhead, this method is suitable for dynamic indexing of image databases that are subject to frequent additions and deletions. We investigate the performance of this scheme using simulations and compare it with the method based on trained holographic memory model. We also report the results obtained with real image data. The untrained memory model with the retrieval method developed by us compares well with previous results for the effective operational range of the holographic memory. More importantly, the untrained memory gives robust performance when a large number of additions and deletions are made in the database.

1. INTRODUCTION

Optical holography is a fascinating phenomenon that has been studied for nearly three decades since Gabor [1] published his original paper. Sutherland [7] proposed a new associated memory model based on the mathematical representation of optical holography. He represented information as phase angle orientations on a Riemann plane, using the principle of information storage and expression in the form of stimulus-response patterns. The governing equations of his method formed a non-connectionist approach in which a large number of stimulus-response associations are enfolded onto a single memory element.

Khan [4] adapted Sutherland's approach to the problem of indexing large image archives. He mapped each image and its index to a stimulus-pattern association which were then encoded into an associative memory. Several such associations could be superimposed in the same memory using the analogy of optical holography. When a query image is presented, its index can be retrieved from the memory, by first mapping the image to a query stimulus pattern and then using it to retrieve the response pattern associated with a matching pattern that was previously encoded in the holographic memory. When encoding the holograph, a training process was applied to minimise the recall errors that would occur during retrieval. The training involved 10-30 iterations over the stimulus-response pattern associations such that the average recall error is below a given small value. Khan generalised the representation of stimulus-pattern associations to deal with colour images but most of his reported work is concerned with gray scale images [5].

Khan's work focussed on image archives that are encoded once and retrieved many times [5]. In a typical image database, however, both additions and deletions of images may be relatively frequent. For example, consider an image database that is part of a security system where images are used for confirming the identity of those authorised to enter a certain area. A security camera may be used to take a snap which would be compared with the images of authorised personnel, using a holographic index. As this security database is updated, new images may be added or existing ones deleted. A trained holographic index has severe limitations when it comes to dynamic indexing of images in such databases, because it is expensive to maintain the associative memory by retraining whenever additions or deletions of images are made in the database.

In this paper, we consider an untrained holographic model as the basis for indexing images. The untrained model has the advantage that it can be modified without the training overhead whenever additions and deletions of images are made in the database. However, it is necessary to deal with the higher recall errors in the untrained holograph in order to achieve reasonable retrieval performance. We describe an encoding and decoding scheme for such a model. The performance of this scheme is investigated using simulations and compared with that of Khan's method based on the trained holographic memory model [5]. We also report the performance of the untrained holograph using a sample of real image data. The results show that the untrained model with the retrieval method developed in this paper compares well with the trained model for the effective operational range of the holographic memory. More importantly, the untrained memory gives robust performance when a large number of additions and deletions are made in the database.

The rest of this paper is organised as follows: Section 2 describes the mathematical background of holographic associative memory. The basis of Khan's training algorithm and the method for indexing image archives are also briefly indicated. The holographic memory without training is described in Section 3. The performance of the untrained memory and its comparison with Khan's results are discussed in Section 4. Addition and deletion of patterns in the untrained holograph are discussed in Section 5, along with algorithms and simulation results. Section 6 is the conclusion of the paper which also contains some pointers for further research.

2. BACKGROUND

This section describes the computational model of a holographic memory and how it is used for information encoding and retrieval. This model is adapted from Khan [5] who derived it from the earlier work by Sutherland [7]. The representation of information using stimulus and response patterns, the mathematical expressions for encoding a large number of stimulus and response patterns into a single associative memory, and the expression for retrieving the response pattern associated with a given stimulus pattern are given below. The concept of a trained holographic memory and its application to indexing image archives are described briefly.

2.1. Representation of Information

The main idea of the holographic memory model is the principle of information storage (encoding) and retrieval (decoding) using stimulus-response pattern associations. Information to be stored in the holographic memory is represented by stimulus patterns. Each stimulus pattern is associated with a unique response pattern.

A stimulus pattern can be generally expressed as a set of elements, and is represented in the form:

$$S = (s_1, s_2, s_3, \dots, s_n).$$

Each element s_k represents the k^{th} piece of information or feature. These elements collectively define a pattern which is to be stored in the memory.

The holographic memory model approach adopted by Sutherland [7] and Khan [5] provides a mechanism to determine a *meta-knowledge* associated with each element, which allows encoding of a *degree of significance* for each element and forms an essential part of the basic representation. In this scheme, the k^{th} information element and its meta-knowledge are modelled as follows:

$$s_k = (\alpha_k, \beta_k).$$

Here, α_k corresponds to the measurement of the k^{th} element, and β_k represents the meta-knowledge about the k^{th} element. The above representation is then transformed into a complex number representation, denoted as a vector in two-dimensional space, by the following expression:

$$s_k = (\alpha_k, \beta_k) \Rightarrow \lambda_k e^{\hat{i}\theta_k}.$$

In this transformation, α_k is mapped into phase element θ_k in the range of $0 - 2\pi$ using a suitable function. The meta-knowledge, β_k, is mapped into a magnitude, λ_k, which is bounded within the unit interval [0, 1].

This model, which is a two-dimensional scheme, has been extended by Khan [5] into a multi-dimensional scheme given by the following expression:

$$s_k = (\alpha_k, \beta_k) \Rightarrow \lambda_k e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,k}\right)}. \tag{2.1}$$

Using this model, each α_k is mapped onto a set of phase elements $\theta_{j,k}$. Here, s_k becomes a vector within a unit hypersphere in a d-dimensional space. Each $\theta_{j,k}$ is the spherical projection or phase component of the vector along the dimension $\hat{i}_j$.

Based on this representation, each piece of information, s_k, may be composed of other sub-information. For example, in a colour image, the k^{th} pixel information may be built up by the values of three colours in RGB format corresponding to red, green and blue. The j^{th} colour of the k^{th} pixel is mapped to $\theta_{j,k}$.

A stimulus pattern with n elements of information is then represented as

$$[S^\mu] = \left[\lambda_1^\mu e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,1}^\mu\right)}, \lambda_2^\mu e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,2}^\mu\right)}, \dots, \lambda_n^\mu e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,n}^\mu\right)} \right],$$

where μ is the pattern index, and n is the number of elements in the stimulus pattern. The magnitude, λ_k^μ, corresponds to the degree of significance of the k^{th} element in the μ^{th} pattern.

The representation of a response pattern is similar to that of the stimulus pattern:

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

where R^μ is the response pattern associated with stimulus pattern S^μ, and m is the number of response pattern elements. The phase component $\phi_{j,k}^\mu$ represents the measurement of a response pattern element, and the magnitude $\gamma_{j,k}^\mu$ represents the expected *confidence* (system assigned significance) on $\phi_{j,k}^\mu$.

2.2. Encoding Process

Prior to the encoding process, the input information is mapped to a stimulus pattern as indicated in Section 2.1. The index number of input information is mapped to a response pattern using a suitable function.

Each stimulus and its response pattern is then associated producing a correlation matrix $[X]$, also called a stimulus-response pattern association, which is obtained as the inner product of the corresponding complex vectors:

$$[X^\mu] = [\bar{S}^\mu]^T [R^\mu]. \tag{2.2}$$

Here, $[\bar{S}^\mu]^T$ is the conjugate transpose of $[S^\mu]$. Since $[S^\mu]$ is a vector with n elements and $[R^\mu]$ is a vector with m elements, $[X^\mu]$ is an $n \times m$ matrix.

Several stimulus-response pattern associations can be superimposed onto the same matrix space as follows:

$$[X] = \sum_{\mu=1}^{p} [X^\mu] = \sum_{\mu=1}^{p} [\bar{S}^\mu]^T [R^\mu], \tag{2.3}$$

where p is the number of stimulus and response pattern associations (or the number of patterns for short). Here $[X]$ is called the *holograph matrix.*

2.3. Decoding Process

The purpose of the decoding process is to retrieve the response pattern associated with a given query stimulus pattern. For the decoding process, the query stimulus pattern $[S^q]$ is defined in a similar way to the encoded stimulus patterns:

$$[S^q] = \left[\lambda_1^q e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,1}^q\right)}, \lambda_2^q e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,2}^q\right)}, \dots, \lambda_n^q e^{\left(\sum_j^{d-1} \hat{i}_j \theta_{j,n}^q\right)} \right].$$

The degree of significance of each query stimulus element is determined by setting the corresponding magnitude. The value of one indicates that the query element fully contributes to the decoding process. Lower values indicate lower contributions of the query element. If the magnitudes are set to zero, the elements contribute nothing to the decoding operations.

The retrieval of the corresponding response pattern is performed by computing the inner product of the query stimulus pattern and the holograph matrix as

$$[R^q] = \frac{1}{c}[S^q][X], \tag{2.4}$$

where c is the normalisation factor defined by:

$$c = \sum_{k=1}^{n} \lambda_k^q.$$

If all elements in the query stimulus pattern $[S^q]$ match an existing stimulus pattern $[S^t]$, the retrieved response pattern $[R^q]$ will be close to the associated response pattern $[R^t]$. Here, closeness is reckoned in terms of the phase values and the magnitudes (confidence level) of each element [5].

If the query stimulus pattern only partially matches an existing pattern, then the retrieved confidence level is expected to be lower than one. It would be even much lower if the query stimulus pattern does not match any existing pattern.

The above characteristics correspond to the capability of optical holography to reconstruct the reference beam from the partial wave fronts of an object, where a noisy version of the reference beam would be returned [5].

2.4. Load Factor

Theoretically, the number of patterns (p) that can be stored in the holographic memory for reliable performance, never exceeds the number of stimulus pattern elements (n). This characteristic is expressed as *load factor* (L), which is the ratio of p to n, for $p \le n$ and given as (p/n). Here, L is in the range 0–1.0. It was originally defined for basic associative memory by Haykin [2].

2.5. Principal and Crosstalk Components

To understand how the retrieval of a response pattern works, the right term of Equation 2.3 is grouped as *principal* and *crosstalk* components [5]. This is done by substituting [X] from Equation 2.2 into Equation 2.3 as follows:

$$\begin{aligned}\left[R^{q}\right] &= \frac{1}{c}\left[S^{q}\right]\left[\overline{S}^{t}\right]^{T}\left[R^{t}\right]+\frac{1}{c}\sum_{\substack{\mu=1\\ \mu\neq t}}^{p}\left[S^{q}\right]\left[\overline{S}^{\mu}\right]^{T}\left[R^{\mu}\right]\\ &= \left[R^{q}_{\text{principal}}\right]+\left[R^{q}_{\text{crosstalk}}\right].\end{aligned} \tag{2.5}$$

Here, [S^t] is an existing stimulus pattern, which closely matches the query stimulus pattern [S^q], and [R^t] is the original response pattern, which is associated with the stimulus pattern [S^t] in the encoding process. When the query stimulus pattern fully matches the existing stimulus pattern, the terms $\frac{1}{c}\left[S^{q}\right]\left[\overline{S}^{t}\right]^{T}$ becomes unity, thus the principal component is the same as the original response pattern.

The crosstalk, given by the second term, is considered as the noise or interference that comes from other stimulus-response pattern associations, because the query stimulus pattern [S^q] does not match the other encoded stimulus patterns, [S^μ], where $1 \le \mu \le p$, $\mu \ne t$.

2.6. Trained Holographic Memory

In the holographic memory with training, the latest state of the holographic memory is obtained after iteratively 'learning' the stimulus-response associations. The learning process is performed by applying the current stimulus pattern to the current state of the holographic memory using the decoding expression (Equation 2.4). The difference between the assigned response pattern and the retrieved response pattern is then calculated and associated to the stimulus pattern. This is the main part where the 'error' correction is made. This association is then superimposed on the holograph. The process is repeated for all patterns iteratively until a satisfactory level of recall error is reached.

The following is the mathematical expression of the training algorithm used by Khan [5] for encoding the holographic memory:

$$[X]=[X]+\alpha\left[\overline{S}\right]^{T}\left([R]-\frac{1}{c}[S][X]\right),$$

where α is a *learning constant* in the range 0 – 1. The training algorithm used by Khan [5] is adapted from Hebbian's Differential Learning [6]. The performance characteristics of the trained holographic model have been studied in detail by Khan [5]. He also investigated the application of trained holographs for searching image archives. Due to space limitations we do not summarise his results here. However, in Section 4, we compare the performance characteristics of untrained holographs with those obtained by Khan for trained holographs.

2.7. Indexing Image Archives

As mentioned earlier, the main inputs in creating a holographic memory are a set of stimulus patterns and their associated response patterns. These inputs are obtained from external measurements. For the image domain, an image corresponds to a stimulus pattern. The intensity values of pixels become the measurements of the stimulus pattern. The response pattern can be obtained from the corresponding index number indicating the location of the image in a database.

The encoding process begins with the mapping of the image information, i.e., pixel intensities, into a stimulus pattern. Subsequently, a response pattern is mapped from the image's index number. Next, the stimulus pattern is associated with the response pattern, which produces a stimulus-response pattern association. The stimulus-response pattern association is then superimposed onto the holograph. To retrieve from the image archive it is necessary to find the index number of the required image or pattern. Given a query image, its index number is identified by applying the decoding process to the holographic index.

3. HOLOGRAPHIC MEMORY WITHOUT TRAINING

The untrained holograph is produced from a straight implementation of the mathematical expressions explained in Section 2, without any attempt to minimise the recall error in the decoding process. This is in contrast to the training process employed by Khan for encoding holographs [5]. Without the training process, each association is encoded into the holographic matrix by adding it to the existing associations. No modification is made to the matrix to compensate for recall errors that would be caused by crosstalk.

Given a query stimulus pattern, the retrieval process is aimed at finding the response pattern that is associated with it in the holograph. As explained in Section 2, the retrieved response pattern is described as a vector sum of the principal and crosstalk vectors. Ideally, the crosstalk vector should be close to zero, in terms of its vector magnitude, for a correct retrieval. However in practice, this is not so. The magnitude and orientation of the crosstalk vector are dependent on a number of factors such as the number of patterns encoded in the holograph and the distribution of stimulus and response elements of the patterns.

In a trained holograph, the recall error caused by the crosstalk component is compensated by the learning algorithm during encoding so that the crosstalk effects are minimised. Therefore, the retrieved response pattern for any given query pattern is closer to the original response pattern associated with the matching pattern. However, it is possible to keep the effects of crosstalk minimal only up to a certain level of loading of the holograph.

In an untrained holograph, there are no attempts to compensate for the recall error during the encoding process, and consequently the crosstalk effects still remain. This causes the retrieved response vectors to move away from the original response vectors. The difference between the phase value of a retrieved response vector and of the original response vector is referred to as recall error. Correct retrieval is possible while the recall error is low enough to be able to determine the original response vector from a given retrieved response vector. In Section 4, we discuss the factors affecting the performance of the untrained holograph based on simulation experiments.

3.1. Encoding Algorithm for Untrained Holograph

The algorithm for encoding an untrained holograph is shown in Figure 3.1. The complex matrices of the stimulus pattern, the response pattern, the stimulus-pattern association, and the holograph are [S], [R], [X] and [H] respectively.

```
1  FOR i = 1 to p patterns DO
2  BEGIN
3    GET ith pattern information;
4    MAP ith pattern to Stimulus Pattern
       [S] = SPMap(pattern-i);
5    MAP ith patt. idx. number to Response Patt.;
       [R] = RPMap(pattern-index);
6    COMPUTE ith Stimulus-Response Patt. Assoc.
       [X] = conjugate(transpose([S])) × [R];
7    ADD the ith stimulus-response
     patt. assoc.onto the Holograph
       [H] = [H] + [X];
8  END
9  SAVE the Holograph [H]
```

Figure 3.1: Algorithm for encoding untrained holgraph.

In line 3, the pattern information is obtained. For an image, it consists of the intensity values of pixels. Next, the pattern information is mapped to a stimulus pattern using a suitable function (line 4). In the case of an image, the intensity values are mapped to elements of the stimulus pattern. Subsequently, the pattern index number is mapped to a unique response pattern using an appropriate function (line 5). In line 6, after the stimulus and the response patterns are obtained, their association is computed using Equation 2.2, and then the association is superimposed onto the holographic matrix (line 7). Lines 3 – 7 are repeated for p number of patterns. After that, the holographic memory is saved (line 9).

3.2. Mapping and Retrieval of Response Patterns

The choice of functions to map the stimulus and response patterns has significant bearing on the recall error for a given level of loading of the holograph. A detailed discussion of the mapping functions can be found in Hendra [3]. Each image is given a unique index number sequentially, which serves as the identification of the corresponding image in the database. Using an appropriate function, this index number is mapped to a unique response pattern that consists of m elements. As discussed in Section 2.1, each element of the response pattern can also be regarded as a vector, where the phase values of the vector in polar form correspond to a predefined interval in the $0 - 2\pi$ range. If q is the number of intervals, the following formula is employed to determine the number of elements m used in the response pattern:

$$m = \log_q p,$$

where p is the maximum number of patterns to be encoded. For example, if $q = 8$ and $p = 512$, m would be 3. In this case, there would be 3 elements, each with 8 intervals, and each of these three elements could have 1 out of 8 possible phase values. They represent a series of unique combinations of elements. The index number for a given pattern is mapped to one of these unique element combinations.

On retrieval, the response pattern elements may deviate from their assigned positions due the effect of crosstalk. An error threshold level is used to perform correct retrievals. If the number of intervals for a response pattern element is q, the phase difference between two intervals is $2\pi/q$. The *error threshold* level, ϕ_T is the half way point of this distance, given by $d=\pi/q$. Let ϕ_i^t be the phase value of the original response vector and ϕ_i^q be the phase value of the retrieved response vector of the i^{th} response pattern element, where $1 \le i \le m$ and m is the number of elements in the response pattern. In simulations to determine the performance characteristics, the recall error of the i^{th} element, where $1 \le i \le m$, denoted by ϕ_i^e, is calculated by:

$$\phi_i^e = \min\left\{\left|\phi_i^t - \phi_i^q\right|, 2\pi - \left|\phi_i^t - \phi_i^q\right|\right\}.$$

To determine if a retrieval is correct, first the maximum recall error is calculated by:

$$\phi_{\max}^e = \max\left\{\phi_i^e \mid 1 \le i \le m\right\}.$$

If $\phi_{\max}^e \le \phi_T$ then it is a correct retrieval, otherwise it is an incorrect retrieval.

4. PERFORMANCE

The performance of the untrained holograph is studied using a number of measures, and also compared with the trained holograph of Khan [4], [5].

4.1. Performance Measures

Three main measures are used to quantify the performance of the untrained holograph. The first measure is *signal-to-noise ratio* (SNR), which is calculated as the mean-squared error of the retrieved and original phase values:

$$SNR = 20\log\left(\frac{2\pi}{\text{mse}}\right) \qquad \text{mse} = \sqrt{\frac{1}{m}\sum_i^m\left(\phi_i^\mu - \phi_i^T\right)^2}$$

In the above expression, m is the number of response pattern elements, ϕ_i^μ is the phase value of the μ^{th} pattern being measured, and ϕ_i^T is the retrieved phase value. Here, the SNR unit is *db* (decibel). It was used by Khan [5] as the main measure of holographic performance. A high SNR level means a low recall error and vice versa.

The second measure is the *percentage of correct retrieval* (PCR), which is used to show the retrieval performance in terms of how many correct retrievals are achieved per 100 retrievals. It is calculated as :

$$PCR = \frac{\text{number of correct retrievals}}{p} \times 100,$$

where p is the number of encoded patterns.

In the encoding process, the magnitudes of the original response vectors are set to one. Based on the magnitudes of the retrieved response vectors, the third measure called *mean normalised confidence* (MNC) is calculated using the following expression:

$$MNC = \frac{\sum_i^m \gamma_i^{T(\mu)}}{m},$$

where m denotes the number of elements in the response pattern, and μ denotes the pattern number. Here, $\gamma_i^{T(\mu)}$ denote the magnitudes of the i^{th} element of the retrieved response patterns.

The MNC reflects the confidence level of the existence of the query pattern [5]. A value close to one will be returned when a query pattern *fully matches* an encoded pattern, and less than 0.5 for an unmatched query pattern.

4.2. Retrieval Characteristics of Untrained Holograph

For the following results, we used randomly generated patterns, with various sizes in terms of number of pixels (n = 512, 1024, 4096 and 10000). In the encoding we used 2 intervals for mapping each response pattern element ($q = 2$).

Figure 4.1 shows the percentage of correct retrieval achieved (the scale on the left y-axis). Almost 100% correct retrieval has been achieved up to a load factor of 0.2, which is comparable to the operational range of trained holographs [5]. The MNC levels are close to one (the scale on the right *y*-axis). These results also confirm the observation of Khan [5], that the holograph capacity is dependent on the load factor, irrespective of the number of elements in the stimulus patterns.

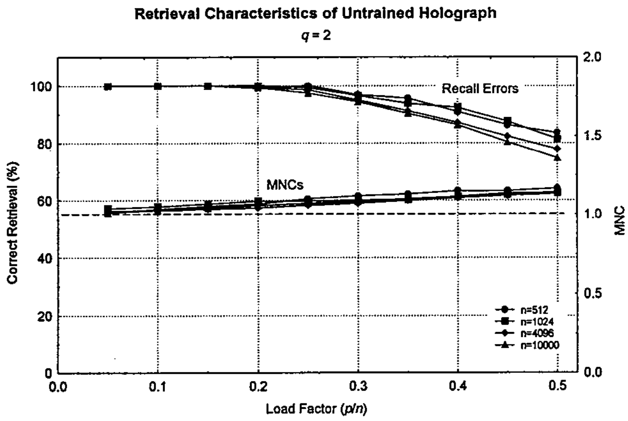

Figure 4.1: Retrieval characteristics of untrained holograph.

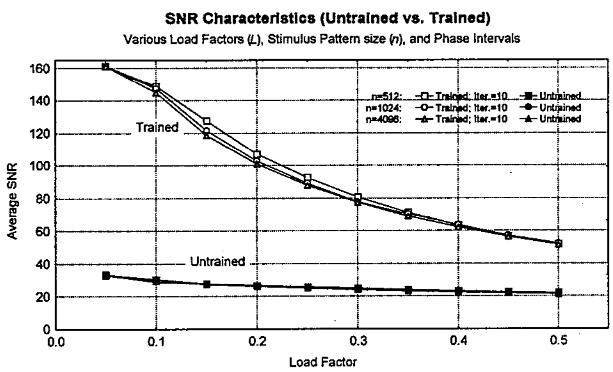

Figure 4.2: Recall errors and MNCs before deletions.

However, Khan [5] used only the SNR measure and not the percentage of correct retrieval to study the performance of trained holographs. For the sake of comparison, we also present the SNR levels for the above retrieval along with the results of retrievals obtained from a trained holograph in Figure 4.2 (for this purpose, the same set of patterns were re-encoded using the training process). As expected, the SNR level of the trained holograph is higher than that of the untrained one. Despite the lower levels of SNR for the untrained holograph, the correct retrieval levels, as shown in the previous graph, are close to 100%, up to a load factor of 0.2. This is because, we use a threshold to determine correct retrievals from the untrained holograph as explained in Section 3.2. With q=2, the threshold level is $\pi/2$ which has been found to provide the best performance for untrained holographs. The details of experiments to establish the optimal number of intervals are given in [3].

4.3. Experimental results on real images

The method described in the previous sections has been tested using real images. The result of one of the tests is described below. Figure 4.3 shows examples of the images used which are in greyscale format, displaying human faces with various expressions. The image size is 120x160, and the number of images used was 456, giving a load factor of 0.024.

Figure 4.4 presents the recall error, the MNC and the asymmetry level for each frame. The frames are arranged in the descending order of recall error levels. The images with recall errors below the threshold level of 90 degrees are retrieved correctly. Those with recall errors above the threshold cannot be retrieved correctly.

Figure 4.3: Examples of real images encoded in holograph.

Khan [5] had identified asymmetry of stimulus pattern as a factor affecting the correct retrieval of response patterns from the holograph. Asymmetry measures the lack of symmetry in the distribution of mapped intensities in the stimulus pattern. It can range from 0 to 1, where 0 corresponds to the lowest level. When the asymmetry level of a stimulus pattern is higher, the recall error is likely to be higher. In the simulations of Section 4.2, the patterns were generated with random intensity values for pixels of each pattern keeping the asymmetry level of stimulus patterns very low. For the real images used in the present example, the asymmetry levels range from 0 to 0.4. It may be seen from Figure 4.4, that images corresponding to asymmetry values of 0 to 0.2 are retrieved correctly as the recall errors of these patterns fall below the threshold. The asymmetry level can be altered by choosing appropriate mapping functions. But a detailed discussion of this aspect is beyond the scope of this paper.

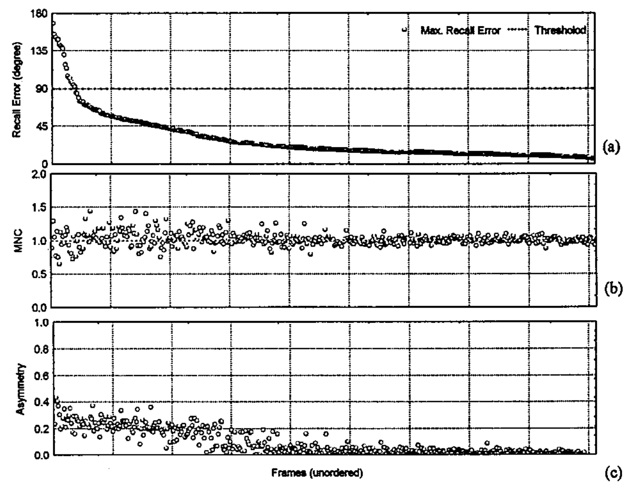

Figure 4.4: Recall errors, MNC and asymmetry of images.

5. ADDITIONS AND DELETIONS OF PATTERNS

Additions of new images and deletions of existing images may be performed frequently in a dynamic database. It is also necessary to make corresponding changes to the holographic index. If this is not possible, the holographic capacity in terms of its available load factor cannot be used effectively. Therefore the holographic indexing scheme should have the capability of dynamic indexing by adding and deleting patterns.

5.1. Addition and Deletion Algorithm

The trained holographic memory model developed by Khan [5] is not suitable for efficient dynamic indexing. This is because 10-30 iterations over all patterns are needed to initially encode the holograph, and further training would be need for additions and deletions to maintain the holograph. The stimulus-response pattern associations stored in the holograph are dependent on one another, because of the training process to minimise the recall error. As there is no training overhead for untrained holographs, they are less expensive to encode. In addition, the associations encoded in an untrained holograph can be individually removed without any residual effects on the remaining data. So the untrained holographic memory is more suitable for dynamic indexing.

Equations 2.2 and 2.3 indicate that we can add stimulus-response associations one at a time, which confirms the ability to dynamically add patterns to the holographic memory. Figure 5.1 shows the algorithm for adding a new pattern to a holograph. In line 1, a response pattern not used in the holograph need to be found first, since several patterns may have been previously deleted from the holograph.

```
1  Set [R] to a response pattern
   not used in Holograph [H];
2  Map new pattern I to stimulus pattern [S];
3  Compute association
   [X] = conjugate(transpose([S])) * [R];
4  Add the association [X] to [H], [H]=[H]+[X];
```

Figure 5.1: Algorithm to add a new pattern to holograph.

We can also delete a stimulus-response pattern association from the holographic index. The expression for deletion is given below:

$$[X']=[X]-[X^{\mu}], \qquad (5.1)$$

where $[X']$ is the state of the holograph after the μ^{th} association, $[X^{\mu}]$, has been subtracted from the old holograph $[X]$. Figure 5.2 shows the algorithm for deleting a pattern from a holograph.

```
Map I to stimulus pattern [S]
Set [R] to the corresponding
  response pattern of [S]
Compute association
  [X] = conjugate(transpose{[S]}) * [R]
Subtract (delete) the association [X]
  from [H], [H]=[H]-[X]
```

Figure 5.2: Algorithm to delete a pattern from holograph.

5.2. Simulation Results

Several experiments were performed to study the dynamic capabilities of the holographic indexes. However, only one sequence of experiments is described here. A set of patterns is first encoded into the untrained holograph, and then several patterns are deleted and several new patterns added. The recall errors and MNCs of the patterns during the retrieval are observed.

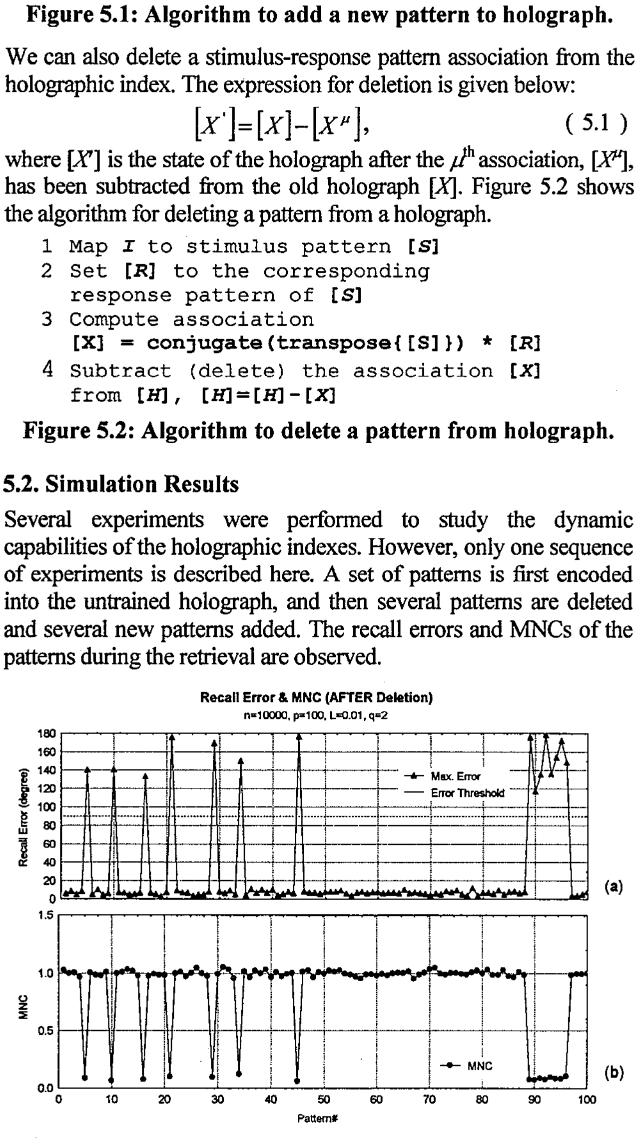

Figure 5.3: Recall errors and MNCs after deletions.

A set of 100 random patterns of 10000 elements were encoded into an untrained holograph, i.e., with a load factor of 0.01. On retrieving all patterns, the recall errors were between 5 and 10 degrees and MNCs around 1. Next, pattern numbers 5, 10, 16, 21, 29, 34, 45, and 89–96 were deleted from the holograph. All patterns were then retrieved again, including the deleted ones. Figure 5. (a) and (b) show the recall errors and MNCs after the deletions. The low MNC levels of deleted patterns indicate that the patterns do not exist any more in the holograph.

After the deletion, five new patterns were added, reusing the response patterns previously assigned to the first five deleted patterns (pattern number 5, 10, 16, 21 and 29). The new patterns could be retrieved correctly, indicated by low recall error and MNC values close to one. The recall errors and MNCs of other patterns were unaffected.

6. CONCLUSION AND FURTHER WORK

We described a refinement to the holographic associative memory originally proposed by Sutherland [7] and subsequently adapted to indexing image archives by Khan [5]. This research was motivated by the need to reduce the overhead in encoding and maintaining the index when it is used in a database that is subject to a large number of additions and deletions. We have developed an encoding and retrieval scheme for a holographic index without training. The performance of this model is comparable to that of the trained holograph of Khan. We also discussed the behaviour of the model when additions and deletions are made for maintaining a dynamic index for images.

The work reported in this paper has been extended to the study of video indexing which will be reported in another paper. Although multidimensional representation had been proposed by Khan, it has not been studied in detail. We are currently investigating untrained holographs using multidimensional complex number representation which can deal colour images, instead of only the gray scale images that have been studied so far.

7. REFERENCES

[1] Gabor, D., "Associative holographic memories," *IBM Journal of Resarch and Development*, Vol. 3, 1969, pp.156-159.

[2] Haykin, S., "*Neural Netoworks: A Comprehensive Foundation*," Macmillan College, 1994.

[3] Hendra, Y., "*Content-Based Retrieval of Information from Image and Video Databases Using A Holographic Memory Model*," Masters Thesis, School of Computing, Curtin University of Technology, Perth, Western Australia, 1999.

[4] Khan, J. I., "Characteristics of multidimensional holographic associative memory in retrieval with dynamically localizable attention," *IEEE Transactions on Neural Networks*, Vol. 9, No.3, 1998, pp. 389-406.

[5] Khan, J. I. (1995). *Attention Modulated Associative Computing and Content-Associative Search in Image Archive*, PhD. Dissertation, Department of Electrical Engineering, University of Hawaii, Hawaii.

[6] Klopf, A.H., "Drive-Reinforcement Learning: A Real Time Learning Mechanism for Unsupervised Learning", *Proc. Of 1st IEEE Conf. On Neural Networks*, Vol. II, N.J., 1987, pp.441-445.

[7] Sutherland, J. G., "A holographic model of memory, learning and expression," *International Journal of Neural Systems*, Vol. 1, No. 3, 1990, pp.259-267.