

# CAFIIR: An image based CBR/IR application

Arcot Desai Narasimhalu  
Institute of Systems Science  
National University of Singapore  
Email desai@iss.nus.sg

## Abstract

In this paper we describe a multimedia application called Computer Aided Facial Image Inferencing and Retrieval (CAFIIR) system. This system uses both Case Based Reasoning and Information Retrieval Techniques. In CAFIIR we use fuzzy measures to represent characteristic features of a human face. This paper describes a method designed to implement inferencing using fuzzy measures. It also describes how CAFIIR handles CHF values that may span across two or more Fuzzy sets. The paper illustrates how Dempster Shafer theory can be useful for extracting weights for individual characteristic features when weights are given for combined features. Some parts of the application have been implemented while others are still under implementation.

## Introduction

The application reported in this paper is a Computer Aided Facial Image Inferencing and Retrieval (CAFIIR) system. The system has photographs of human faces and text records related to these faces. The input to CAFIIR system is either a graphics sketch of a human face or a digitized image of a human face and some text description. The CAFIIR application classifies human faces into subclasses as is normally done in a Case Based Reasoning (CBR) [Bain, 1986] application. It uses composite indexes and similarity measures, just as in any Information Retrieval (IR) [Salton, 1991] application, on human faces to shortlist a relevant set of faces from a class of faces that correspond to a target face. The rest of the paper is organized as follows. Section 2 provides a description of the CAFIIR application. Section 3 discusses the issues faced in the design of the CAFIIR application. Section 4 describes the CAFIIR system architecture. Section 5 summarizes the findings reported in this paper.

## The CAFIIR application

In this section we will describe the characteristic features of a human face, provide a brief description of the application, and discuss why we consider this ap-

plication to be a multimedia application that straddles across CBR and IR technologies.

## Characteristic Features of Human Faces

Human faces have several characteristic features (CHFs). Some of these CHFs are visual in nature. Features such as eyes, nose, outline of a face are visual characteristics (VCs). There is a second set of features of a human face that forensic experts and plastic surgeons are interested. These are called Anthropometric landmarks (ALs) [Jurgens *et al.*, 1990]. Anthropometric landmarks are points on the surface of a human face that are measured by plastic surgeons for the purposes of study and manipulation. There is a third set of features in a human face. These are of interest to orthodontists and forensic experts. These are called Cephalometric landmarks (CLs) [McNamara, 1984] and are points on the hard tissues of a human face. There is a fourth set of characteristics on a human face which do not occur in all human faces. These are features such as scars, moles etc. which can very distinctly identify a human face. We call these features Special features (SFs) for lack of a better name. Different combinations of the four features described above are relevant for different applications. There are also seldom considered global features such as the combination of nose, lips and chin that are called Global characteristics (GCs). For the purposes of this paper we do not consider skin texture as a feature.

## Description of the CAFIIR application

Once a human face is computerized, there are several possible applications that can make use of the computerized face data. Examples are Forensic, Crime investigation, Plastic surgery and Orthodontistry. The CAFIIR system described in this paper is a crime investigation application using human faces. In a crime investigation application, there is a repository or database of criminals. Each entry in this database consists of a set of photographs and other related information in text format. Some of the related information are structured fields such as Name, Address, Age and Sex of a criminal. There is another set of information

that is basically in free form text which describes the nature of crime, scene of crime, weapons used and outcome of the crime (attempted crime, abandoned crime, successful crime, injuries, death, etc.). The faces and their records in this collection are called the reference collection.

Whenever a crime investigation begins, the victim is interviewed and either a graphical sketch of the face of the suspect or a photofit image of the face is generated. In CAFIIR, we refer to this sketch or photofit image as the target face. The target face comes along with text information which is a description of the crime. This description is usually elicited from the victim and in general describes details such as the scene of the crime, the sequence of events and a description of the suspect(s). These descriptions are usually submitted in text form. The application is required to use the graphical sketch / photofit image and the text descriptions to produce a ranked list of suspects.

### **Why is CAFIIR a CBR/IR application**

In CAFIIR, the reference collection is organized into classes. The classification is decided based on how people would describe suspects and crimes. For example, description based on ethnic background or skin color are found to be the easiest descriptors used by victims to describe suspects. Other descriptors often used are the facial outline, hairstyle and special features. While these classifications are based on features of a human face, there is another set of orthogonal descriptors used for further subclassification. These features include types of crime, types of weapons used and places of crime. Whenever a crime is reported, a quick assessment is made on which of the descriptors can be used for the purposes of investigation. The search is narrowed down to a subclass of criminals. Since the reasoning proceeds on the basis of previously known cases, CAFIIR qualifies to be labeled as a CBR application. CAFIIR uses certain concepts such as similarity measures, relevance ranking and search techniques that are used in Information Retrieval (IR) applications and hence can be considered to be an IR application.

### **Why is CAFIIR a multimedia application?**

CAFIIR uses text data type for storing descriptions. It uses image data type for storing photographs of criminals and also for storing photofit images. It uses graphical sketches to visualize a victim's description of an assailant. It is also highly interactive and allows users (such as police officers) to control the interaction. Further, it uses cognitive models of the the suspects as described by the victims for the purposes of representation. These cognitive models are used to derive the task model. It also uses user models in terms of what features they are likely to notice and what levels of confidence can be placed on their descriptions. All these factors combined makes CAFIIR a multimedia application. CAFIIR has the potential to integrate other

data such as finger prints and genetic codes in the future. When 3D laser scanners are easily accepted for common use, photographs may be supplemented by 3D scans of criminals' faces.

### **Issues in the design of CAFIIR**

The issues considered during the design of CAFIIR include the following:

- List of CHF's to be considered
- Relative importance of the CHF's
- Domain values for the CHF's
- Classification of CHF's
- Composite classification codes
- Prototypicality
- Types of inference mechanisms
- Support for views

We will consider each of these issues in some detail in this section.

### **The list of CHF's to be considered**

At the highest level, CAFIIR has two sets of CHF's to consider. The first set is the set of features from a human face. The second set is the set of features from the details of a crime.

The visual landmarks of a human face considered for CAFIIR application includes:

- Hairstyle
- Eyes
- Eyebrows
- Nose
- Chin
- Lips
- Special Features
- Facial outline
- Some combinations of the above items

While all these features can be used for inferencing, it is important to see how permanent or reliable these CHF's are. For example, it is found that many victims can easily describe an assailant's hairstyle. However, it is common knowledge that hair style is easily alterable either through processing or through accessories such as a wig. Another example is the eyelids. Eyelids carry a lot of details that can be very useful for identification purposes. They have single fold or double folds. Each of them can be either complete (i.e. from one end to another) or incomplete. While these features may be useful for a forensic application, it is impossible to expect a victim to have had the time to study these features of the assailant in any reasonable detail. This is where the expertise of the police officers comes in useful. Given their experience in investigating several

hundreds of cases, they are able to suggest an useful list of CHF's in a human face. For the CAFIIR application, special features, facial outline, eyebrows, eyes, nose and the lip-chin combination were shortlisted in that order. The order implies that special features are likely to more accurately divide the search space into relevant and non-relevant reference collections than outline of a face. It also implies that there are bound to be cases where additional evidence obtained by considering a nose is important for the purposes of effective identification.

The second set of CHF's have to be extracted from the text descriptions. Some of the information from the text descriptions can be extracted and coded. Type of crime and type of weapon are two examples of CHF's that can be found in text descriptions and lend themselves to be extracted and coded. Where the number of values that a CHF can take is very large, such values are not translated into codes but are left as text for later use by a free text search engine. Even where some CHF's are extracted and coded, police officers retain the right to leave such information embedded within the rest of the text description.

### Relative importance of the CHF's

A short list of CHF's was presented in the previous section. Expert crime investigators have spent several years and have accumulated sufficient experience in being able to quantify the relative importance of these CHF's. They can, for example, assign weights (normalized to range between 0 and 1) to each of the CHF's. The weight of a CHF is the probability that its use will help in identifying the suspect. A CHF with a higher weight is defined to be more discriminating in dividing the reference collection into relevant and irrelevant groups. All the expert crime investigators may not assign the same weights to each of the CHF's.

These weights are determined also by considering how invariant a CHF can be. For example, while the lips in a human face may be considered to be a CHF, it is not a very reliable measure since some faces may have their lips being bitten into, while others might have their lips in a normal position, while some others can have parted lips revealing teeth. The weight assigned to lips will hence have to be rather low.

### Domain values for CHF's

Each selected CHF will have a domain of values that it can assume. This domain of values is not always the set of natural numbers or integers. In several cases, the domain will contain abstract values. For example, the values for the CHF 'eyes' can be from the set (*big, small, average, medium, narrow, bulging, deepset*). Notice that some of the values such as *average* and *medium* may be declared as synonymous. Also, notice that the values are such that they can be further divided into two separate subsets. In this case, these two subsets can be (*big, small, average, medium*)

and(*narrow, bulging, deepset*).

Whether narrow should fall in the first set or the second will be decided by the crime investigation experts. In this case, the first set refers to the size of the eye and the second set refers to additional description. Thus an eye can be either *small, bulging* or *big, narrow*. For the sake of completeness one can also include the value *null* in each of the subsets. This is included to allow for only one value from either of the two subsets to be used for assigning a value to the CHF 'eye' of a face. The domain values of a CHF in fact defines subclasses of that CHF.

When the values assumable by a CHF is abstract, we represent them using Fuzzy measures [Zadeh, 1987]. Some of the fuzzy measures can be partially ordered. For example the first subclass can be partially ordered (in descending order) into (*big, medium, average, small*). Such subclasses will obey the axioms of fuzzy measures [Zadeh, 1987]. However, consider the subclass (*narrow, deepset, bulging*). There is no natural partial order in this case. And, introducing any partial order can only be artificial and lead to endless debates. We describe the first subset as using regular fuzzy measures. We describe the second subset as using irregular fuzzy measures.

### Classification of CHF's

When a criminal's face is submitted to the CAFIIR system, it is first classified before it is registered into the reference collection. The classification of a face is achieved by first classifying the selected CHF's. Individual CHF classifications are then used to derive the overall classification. In order to classify a CHF, it has to be extracted with clear boundaries. This requires good segmentation algorithms. When CHF's have to be extracted from a large number of photographs, availability of either automatic or semiautomatic segmentation algorithms will be useful. However, the image of a human face is a gray scale image. Moreover, different visual features are highlighted or lowlighted depending on the position of the light source. While every attempt is normally made to keep the position of the light source constant, the angle at which a face is presented to the camera cannot be as easily controlled. Given the differences in lighting and camera angles thereby resulting in unclear feature boundaries, it was decided to use operator assisted segmentation. In CAFIIR, a set of visual landmarks are first fixed by an operator before the features are segmented and extracted. These landmarks correspond to and are determined by some of the CLs and ALs close to the VCs.

Once a feature is extracted, it needs to be classified. It is easy to classify those features that assume numeric values. For others, there can be either automatic classification using neural network technology, automatic classification using some mensuration or expert assisted classification. For unsupervised classification using neural networks, one can select a neural net-

work such as Kohonen net [T.Kohonen, 1990], supply the input nodes with the segmented CHF. With sufficient training, the system will converge to spatially organized classes. This approach is designed to be used in CAFIIR to classify GCs or combinations of CHF. This approach has obvious limitations such as long training time for large samples and the possibility that the classes generated may not have an exact correspondence with the classes desired. Even when the training is successful, there is no clear understanding on why certain values of CHF were grouped into a class.

The second approach that is being pursued by the CAFIIR application is to translate visual characteristics into classes using some rules based on a model of human face. The actual model used in the CAFIIR system is quite complex. We illustrate the second approach using a simple model for the CHF 'eye' and the related set of rules. For example, if  $h$  is the height of an eye and  $w$  its width, one can define a ratio  $S$  such that  $S = w/h$ . Once  $S$  is known, the class to which that eye belongs to can be determined easily. For example, when  $S$  is

- $\geq 3$  it belongs to the subclass *long*
- $< 3$  and  $\geq 1.5$  it belongs to the subclass *average*
- $< 1.5$  it belongs to the subclass *small*

The third approach is to ask an expert to look at the extracted feature and classify it under one of the previously defined classes. One of the major problems with this approach is that human beings often use descriptions such as 'I think that this eye is almost of average size'. Or sometimes they will say that they believe that an eye might qualify for inclusion in multiple classes. CAFIIR uses the concept of Fuzzy sets to represent such ambiguous descriptions. Use of Fuzzy sets allows multiple set inclusions which are necessary for an application such as CAFIIR. For example, an user can describe an extracted CHF such as an eye to belong to the class *long* with a possibility of 0.8 and to belong to another class such as *average* with a possibility of 0.5.

While Fuzzy set representation is useful in classifying a CHF, it may be impractical to expect an expert to be available for classifying every CHF in every face stored in the reference collection. Hence, in CAFIIR it is proposed to get some input from users for a normative set of CHF. This reference set or normative set will give a prototype image for every domain value (or subclass) of each of the CHF. For example, one can obtain a prototype for the CHF 'eye' and the domain value *small*. From this prototype one can calculate the values of  $w$  and  $h$  for a small eye. The expert can also specify a membership function for the CHF 'eye' which will give membership inclusion values if the ratio of  $w$  to  $h$  deviates from the norm. We will discuss membership inclusion values in a later section. However, there are some CHF, especially global ones, for which it is difficult to obtain such accurate measurements. For such CHF it is planned to use self-organizing neural nets

such as the LEP [Wu, 1990] for generating the classifications.

### Composite classification codes and similarity measures

We discussed the issue of classifying individual CHF in the previous section. Once this step is completed we need to design a composite index to represent each of the faces. A simple way will be to generate a code for each CHF. For example, in the case of 'Eye', the code 'A' can represent the value *Big*, the code 'B' can represent the value *Average*, etc.

Once individual features have been coded, a composite code is developed by concatenating the codes corresponding to the different CHF. For example, a composite code to represent 'eye' and 'nose' can be A:B, which is interpreted as the face having big eyes and average sized nose. Where the set inclusion is fuzzy, this will be modified to 0.8A,0.5B:0.7B which is interpreted as the eyes belong to the class *Big* with a possibility of 0.8 and to class *Average* with a possibility of 0.5 and that the nose belongs to the class *Average* with a possibility of 0.7.

Once such a composite code is developed, it is next very important to decide when two composite indexes are said to be similar or to be able to compute some distance measure. We will illustrate this by an example.

Let  $f_1, f_2, f_3$  be the set of CHF representing the text and image features. Let T be the target face and  $R_1, R_2, R_3$  be the reference faces in a collection. We first compute the similarity between T and the three reference faces. let  $d_{i,j}$  represent the similarity of CHF  $i$  in T to the same CHF in the  $j$ th face in the reference collection.  $d_{i,j}$ s are normalized and will take values between 0 and 1. Values closer to 1 represent a strong similarity and values closer to 0 represent a weak similarity. Then D is a similarity matrix with the columns representing reference faces and the rows representing the CHF. In this example, D is a 3 x 3 matrix.

The table below presents one such set of values:

CHF	Faces		
	R1	R2	R3
1	0.8	0.7	0.5
2	0.6	0.5	0.6
3	0.3	0.4	0.2

We call this the similarity table  $T_s$ . Once these are listed we can follow the inferencing / matching process described in a later section.

### Prototypicality

As in CBR, we designate one or more faces as prototypes or representatives of a class. Prototypicality has been examined in great detail by [Tversky, 1984]. We use the notions of prototypicality as defined in Tversky. If  $P(a, \Lambda)$  denotes the degree of prototypicality of face  $a$  with its class  $\Lambda$  where the cardinality of  $\Lambda$  is  $n$ ,  $P(a, \Lambda) = p_n(\lambda \sum f(A \cap B) - \sum (f(A - B) + f(B - A)))$  where the summations are over all  $b$  in  $\Lambda$ .

Note A is the set of values for the chosen set of CHF's corresponding to face *a* and B is the set of values for the corresponding set of CHF's for a face *b*, where *a* and *b* are different faces from the class A. The face(s) that have maximum value of  $P(a, \Lambda)$  is (are) considered to be the prototype(s) for the class. Here  $p_n$  is a constant which is introduced to represent the effect of the size of the class on prototypicality. It normally assumes a value of  $1/n$ .  $\lambda$  is a constant that is introduced to represent a bigger weight for similarity among features of two faces being compared as opposed to the dissimilarity among features.  $\lambda$  is normally set to be greater than one. This results from the observation made by Tversky that common features are weighted more heavily for the purposes of prototypicality than in judgements of similarity. For the purposes of fine tuning it is necessary to fine tune the constant  $\lambda$  for each of the classes. As a result, there can be a different value of  $\lambda$  for each class.

### Inference in CAFIIR

In this section we will define the inference techniques designed for CAFIIR. We will first describe inference techniques for composite classification codes. After this we will explain how this scheme can be extended for inference in the presence of fuzzy values for the CHF's. CAFIIR application does not require only exact matches or unification. The final selection is done by a user (victim or criminal expert). Hence the objective of matching / inference is to derive a ranked shortlist of both exact and approximate matches.

**Inference for normal composite classification codes.** Let us revisit the table mentioned in the section on composite classification codes. In normal IR techniques, one would try and compute the similarities between the target object (or the query) and the members of the reference collection using either Boolean or other search techniques [Salton and McGill, 1983].

In CAFIIR we introduce an inference technique where the matching is based on the relative ranking of the reference faces with respect to the target face across all CHF's. For this purpose we revisit the table  $T_r$  and replace its values as described below. We convert the absolute values of similarities for each CHF into relative ranking of similarities. A face with the highest similarity value for a CHF gets the rank 1 and the one with the next highest value gets rank 2 etc. This ranking scheme allows for ties. For example, for CHF2, faces 1 and 3 tie for the first rank. The modified table (shown below) contains integer values. We call this the similarity rank table  $T_{sr}$ .

CHF	Faces		
	R1	R2	R3
1	1	2	3
2	1	2	1
3	2	1	3

This table is interpreted as follows. For CHF 1, face 1 is most similar to the target face, face 2 is less similar

and face 3 is the least similar. Similar interpretation applies for other CHF's as well.

The inferencing to determine which of the faces in the reference collection is closest to the target face across all the CHF's is computed as follows.

Let  $R_{i,j}$  be the similarity rank of the *j*th face for the *i*th CHF. We have previously mentioned that each of the CHF's are given a weighting which is based on perceived reliability and the capability to discriminate the relevant faces in the reference collection from the irrelevant faces. Let the weighting factor for the CHF *i* be  $W_i$ . Then the similarity measure  $SM_j$  for a reference face *j* with respect to a given target face across all the *i* CHF's is given by:

$$SM_j = \sum W_i R_{i,j}, \text{ for all } i.$$

For the example shown above, let us assume the following: Let CHF 1 be a special feature and its weight  $W_1 = 0.8$ . Let CHF 2 be the facial outline and its weight  $W_2 = 0.6$ . Let CHF 3 be the eyes and its weight  $W_3 = 0.3$ .

$$\text{Then } SM_1 = W_1 R_{1,1} + W_2 R_{2,1} + W_3 R_{3,1}$$

Substituting the values from the Similarity rank table  $T_{sr}$ , we get,

$$SM_1 = 0.8 * 1 + 0.6 * 1 + 0.3 * 2 = 2.0$$

The values for  $SM_2$  and  $SM_3$  can also be calculated similarly and are

$$SM_2 = 3.1 \text{ and } SM_3 = 3.9$$

Once the  $SM_j$ 's have been calculated, they will be ranked in the descending order. The face *l* that is selected as the one that is most similar to the target face is

$$l \text{ such that } S_l = \min(S_j) \text{ for all } j.$$

In this example, R1 turns out to be the face most similar to the target face.

**Inference on Composite Classification codes with Fuzzy values.** The inference discussed so far is when abstract or fuzzy descriptions such as *big* or *small* map into point values. This approach is useful when a similarity measure for a CHF is computed based on a single value. As discussed in the section on classification of CHF's, it is often not possible for users to give a single value for a CHF. Instead they may give multiple values with different possibility values for each of these values. Moreover, the mapping between some physical metrics such as *S* and the abstract classes such as *big* or *small* may not be non overlapping. For example, the range of values for *S* may turn out to be as follows:

- $\geq 3$  it belongs to the subclass *long*
- $< 3.5$  and  $\geq 1.5$  it belongs to the subclass *average*
- $< 1.7$  it belongs to the subclass *small*

Notice that the boundary values between *long* and *average* and between *average* and *small* overlap.

In such cases, it is necessary to determine the fuzzy membership function. Let us consider a simple example for a *S* that falls in the first category *long*. A Crime

investigation expert can help in determining the fuzzy membership function as shown in the following table.

Value	Membership function
3.6	1.0
3.5	0.97
3.4	0.95
3.3	0.92
3.2	0.91
3.1	0.9
3.0	0.89
2.9	0.0

A table such as the one shown above is called a Fuzzy Membership function table or  $T_{fm}$ . Fuzzy membership functions ought to be calculated for every shortlisted CHF in the reference collection.

As soon as a CHF is extracted, a relevant  $T_{fm}$  is consulted which will translate the measured values into a fuzzy class and a possibility value. Assume that the value of  $S$  for an 'eye' in a face was found to be 3.2. This implies that this eye belongs to two classes, viz., 'long' and 'average'. The possibility value for 'long' can be looked up the  $T_{fm}$  table presented above and found to be 0.91. A similar look up into the  $T_{fm}$  table for *average* can provide a possibility value for this eye's membership into the fuzzy class *average*.

We will next discuss how to handle inference in the context of multiple fuzzy set inclusion. For every CHF, create a table which lists the faces from which the CHF was obtained as the rows and the different fuzzy sets defined for that CHF as the columns. We call such a table, a table containing fuzzy membership values for a CHF or simply as  $T_{fmcHF}$ . We present one such table below for the CHF 'eye'. It is therefore called  $T_{fmeye}$ . For the purposes of the following discussions we shall simply refer to it as T.

Values of CHF Eye

Face No	Big	average	Small	SD	Rank
T	0.9	0.2	0.6	0	-
1	0.9	0.3	0.5	0.8	4
2	0.9	0.4	0.6	0.4	2
3	0.9	0.3	0.6	0.2	1
4	0.9	0.5	0.6	1.0	6
5	0.9	0.2	0.5	0.6	3
6	0.9	0.4	0.5	1.0	7
7	0.9	0.1	0.6	0.2	1
8	0.8	0.3	0.5	1.7	8
9	0.8	0.2	0.6	0.9	5
10	0.7	0.5	0.4	3.6	9

This table contains 11 rows,  $R_T$ , the row corresponding to the target face, and rows  $R_1$  through  $R_{10}$  which are the rows corresponding to (say) ten faces from the reference collection. The values for all the eleven rows are derived from the fuzzy membership function tables for the CHFs. This is done by first extracting the CHFs for both the target and the reference faces. Next, the different parameters are measured. The fuzzy set membership values are derived using the relevant rules and

$T_{fms}$ .

The fifth column in this table is actually the similarity distance (SD) measure between the target face (row T) and each of the ten reference faces. The values under that column are calculated using the formula given below.

$$SD_T = 0;$$

$$SD_i = \sum W_j * (T_{T,j} \sim T_{i,j}) \text{ for } j = 2, 3 \text{ and } 4.$$

Here  $W_j$  is the weight assigned to the fuzzy class described in the  $j$ th column. For our purposes, we determined  $W_j$  to be  $T_{T,j} * 10$ . This scheme, in a sense, adequately reflects the relative membership strengths of the fuzzy classes for a CHF. The row which has the smallest SD value has an eye which is closest to the eye in the target face and is ranked number 1. The face with the next higher CHF value is ranked second and so on until all the faces are ranked. The ten faces in the table above are ranked for the CHF Eye. Once ranks for all the CHFs are computed, the algorithm used previously for ranking faces using composite classification codes can be applied to determine the overall ranking of the reference faces over all CHFs.

### Support for Views

In an earlier section we mentioned that different expert users can have their own views on which of the CHFs are relevant for a given application and what weights they should be assigned. While a system can have a core set of CHFs that is a union of CHFs required by all the users of the system, views are implemented by controlling the relative weights. For example, if a crime investigation expert relies more heavily on the facial outline than another expert, then the weights assigned to that CHF by the first expert will be greater than that assigned by the second expert. So, where multiple experts are expected to use the system, we build a CHF value vector corresponding to each expert. When an expert starts using the system, his CHF value vector will be loaded in for the purposes of inference. All the weights in a CHF value vector are initially set to zero. When an expert assigns relative weights to different CHFs, these weights are captured in his/her CHF value vector. Since this vector is used during runtime only, an expert can modify the relative weights any time he or she desires.

In CAFIIR, we use Dempster-Shafer theory [Shafer, 1976] for handling incomplete information. The following discussion illustrates our approach.

Let there be evidence that the combinations of CHFs Special features (SF), Facial outline (FO) and Eyes (EY) can reliably identify a victim with a probability 0.95. We call this evidence  $E_1$ . This probability is in fact the weight assigned to this CHF by an expert investigator.

Let there also be a second evidence  $E_2$  which states that the probability that a criminal can be identified using SF and the CHF 'Nose' (NO) is 0.8. For the purposes of this discussion let us assume that these are the

only four CHF's used in the application. It has been established that human faces are recognized due to a combination of CHF's and not by any one CHF alone. In other words, this means that the weights assigned to any one CHF cannot be considered to be independent. If these are the only evidences supplied by an expert crime investigator, and if it is required to have the individual probability (which translates to weight for the CHF), for the CHF SF, then it can be calculated using Dempster Shafer theory as shown in the following table.

	SF,FO,EY	ALL
	0.95	0.05
SF,NO	SF	SF,NO
0.8	0.76	0.04
ALL	SF,FO,EY	ALL
0.2	0.19	0.01

In the above table ALL refers to all the four CHF's being discussed. The above calculation gives us a mechanism to extract a separate weight for the CHF SF. However, it is to be noted that there is still a combined value that for the CHF's SF, FO and EY.

In CAFIIR, we try to use this approach for extracting individual weights for some of the CHF's when individual weights cannot be identified by an expert. The ability to decluster the weight of a CHF from the weights of combined CHF's is considered to be a necessary process to be able to define views which require weights for individual CHF's.

### System architecture

The system architecture is presented in figure 1. Besides the features discussed in the previous sections, the system is designed to have the capability to either age a face or photofit a face using predefined templates for individual CHF's. The aging and photofit modules determine the source of the target face and do not affect the rest of the design of CAFIIR. The reference face collection is organized into two parts. The first part contains natural faces. The second part contains aged and photofit faces. The search can be directed to either one or both of these collections. Details of system design can be found in [Wu *et al.*, 1993].

### Summary

We have described the CAFIIR system in this paper. Classification of CHF's, neural networks for the classification of compound CHF's and the photofit modules have already been implemented. The inference mechanism as described in this paper will be integrated in the near future. We find this application to span across Expert Systems, Multimedia databases, Information Retrieval, Case Based reasoning and multimedia technologies. As in both CBR and IR, CAFIIR applies filtering, feature extraction, classification and storage for the registration phase and filtering, feature extraction, classification, similarity distance calculation and ranking for

the query phase. In most IR systems there is little emphasis on classification and in most CBR systems there is little emphasis on indexing. In CAFIIR, the indexing is based on classification information and hence it spans across both CBR and IR techniques. The system at present works with less than hundred faces. We expect a first version of the complete system to be developed by June 1993. After that phase, we expect to concentrate on issues related to scaling up. We are also examining alternatives to the Dempster-Shafer theory that will allow us to better extract individual weights from collective weights.

### Acknowledgements

The detailed design and implementation of the CAFIIR system is being carried out by the FACE group in ISS. This FACE group is made up of Wu Jian Kang, Ang Yew Hock, Lam Chiam Prong and S. Krishnamoorthy. We are grateful to the Singapore Police Force for sharing their expertise and requirements with us for the purposes of carrying out this research.

### References

- [Bain, 1986] W. Bain. *Case-Based Reasoning: A Computer Model of Subjective Assessment*. PhD thesis, Department of Computer Science, Yale University, New Haven, Connecticut, 1986.
- [Jurgens *et al.*, 1990] H.W. Jurgens, H.A. Aure, and U. Pieper. International data on anthropometry. Technical report, International Labor Office, 1990.
- [McNamara, 1984] J. A. McNamara. A method of cephalometric evaluation. *American Journal of Orthodontics*, 86(6):272-283, December 1984.
- [Salton and McGill, 1983] G. Salton and M.S. McGill. *Introduction to Modern Information Retrieval*. McGraw-Hill, 1983.
- [Salton, 1991] G. Salton. Developments in automatic information retrieval. *Science*, 253, 1991.
- [Shafer, 1976] G. Shafer. *A Mathematical theory of evidence*. Princeton University Press, 1976.
- [T.Kohonen, 1990] T.Kohonen. The self-organizing map. *Proc. IEEE*, 78(9):1464-1480, September 1990.
- [Tversky, 1984] A. Tversky. Features of similarity. *Psychological Review*, 84(4):327-354, 1984.
- [Wu *et al.*, 1993] J.K. Wu, Y.H. Ang, C.P. Lam, S.K. Moorthy, and A.D. Narasimhalu. Facial image retrieval, identification and inference system. Submitted to ACM MM93, 1993.
- [Wu, 1990] J.K. Wu. Lep - learning based on experiences and perspectives. In *Proc. of International Conference on Neural Networks*, Paris, 1990.
- [Zadeh, 1987] L.A. Zadeh. Fuzzy measures. In R.R. Yager, S. Ovchinnikov, R.M. Tong, and H.T. Nguyen, editors, *Fuzzy sets and Applications*. John Wiley and Sons, 1987.

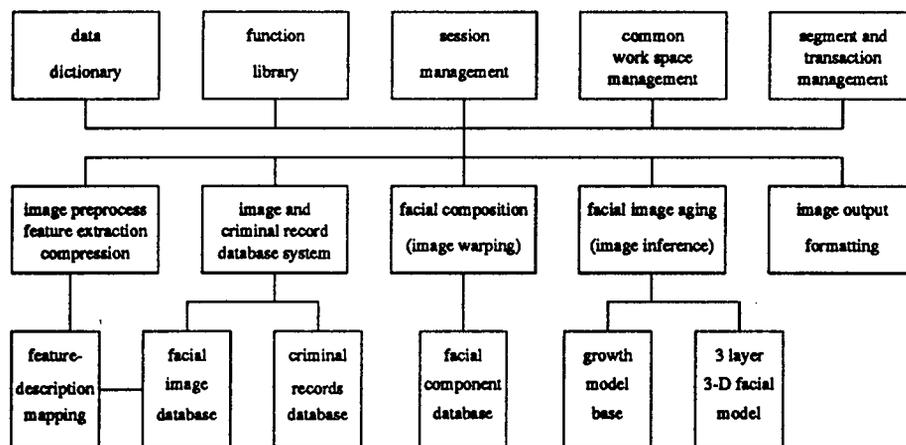


Figure 1: The System Architecture of CAFIIR