AN INTERACTIVE COMPUTER SYSTEM FOR RETRIEVING FACES

J.W. SHEPHERD

1. INTRODUCTION

The evidence of eyewitnesses plays an important role in the work of the police and security services. The use of identification parades and mug-shot albums is an integral part of the investigative aspect of police work, while the evidence of eyewitnesses often forms a major part of the prosecution's case. It is therefore not surprising that for many years police forces have tried to improve the efficiency of methods of obtaining identification evidence from eyewitnesses.

The use of illustrative aids has a long history (see Davies 1981 for a review) and these are still popular with the police forces in most countries. Those in use in the United Kingdom, the United States and some parts of Continental Europe usually either involve some variant of a jig-saw principle in which separate parts of the faces (hair, eyes, mouth, etc.) are selected from a number of alternatives and are assembled to make a composite of the face, or involve an artist producing an impression of the target from the description provided by the witness.

These methods share the common difficulty that they require the witness to recall all the major features of the face of the target or at least be able to recognise each of the major facial elements from among a large number of alternatives. However, people often find great difficulty in recalling the features of faces even of familiar people, and, as Klatsky and Forrest (1984) report, subjects were unable to report details about the faces of people they had just been shown. For example, they were often unable to report whether the person's mouth had been open or closed. Failure to select the appropriate feature in composite construction may lead to grossly misleading end products, even if other features are a good fit. An additional problem of many such systems is that the range of alternatives available in comparison with the population of features they are supposed to sample is so small that a good match for any feature is difficult to make. The advent of computer techniques (see Laughery, Rhodes and Batten 1981 for a review) has enabled more flexibility to be introduced into the construction process but even these do not overcome the limitations of recall in the witness. For even when subjects are asked to reproduce from memory a Photofit composite they have just seen they are seldom able to make a perfect match (Ellis, Shepherd and Davies 1975).

The alternative to asking a witness to provide a composite construction or verbal description of a suspect is to ask him to search for a photograph of the person from among a collection of mugshots. The advantage of a photograph is that instead of having to recall a face, a
witness has only to recognize that suspect, a task at which people in general are highly efficient. Most police forces retain collections of photographs of previously convicted offenders, and these are frequently assembled into albums through which witnesses can search in an attempt to identify a suspect. Apart from the requirement that the suspect must have a previous conviction for this method to be used, the main problem with it is that the witness may be required to look through a very large number of photographs, possibly a thousand or more, with the attendant risks of interference, change in criterion or fatigue (Davies, Shepherd and Ellis, 1979; Laughery, Alexander & Lane, 1971). Where the police make up a special album, for particular kinds of offence or categories of person, there are considerable time demands on the police personnel.

Ideally a witness should be presented with a subset of faces selected by searching the complete set of faces available on the basis of information provided by the witness and by other means of investigation, so that only the most likely items in the collection are included in the set to be searched. Such a preliminary search is a task for which a computer is clearly ideally suited.

The application of computers to this search problem has taken two main forms. One approach has been to obtain some form of graphic input, from an artist’s impression or Identikit or Photofit construction, and run this through a pattern matching algorithm. This approach does not surmount the problem of detail recall facing the witness. The alternative is to use a coded verbal description of the suspect to search through a database of coded descriptions of the faces on file (e.g. Goldstein, Harmon & Lesk, 1971), and to produce a list of the faces which fit the initial description most closely (see Laughery, Rhodes and Batten, 1981 for a review). This is the approach which has been the basis of the method adopted at Aberdeen, where FRAME, Face Retrieval And Matching Equipment, has been developed in collaboration with the Home Office.

2. THE FRAME SYSTEM

In its present form FRAME comprises a videodisc on which photographic images of 1000 faces are stored, a database of 1000 records of these faces coded on 50 attributes or parameters, and a program which compares a set of input parameters for a face with each of the records in the database, ranks records in their order of similarity to the input, and displays the corresponding facial images via a video disc player on a television monitor in this order.

Setting up the system involved three stages: assembling a collection of photographs to provide the basis of the prototype system, coding the photographs to form the database on which the search program operates, and evaluating the system.

2.1. Assembling the collection

Two considerations entered into decisions about the nature of the collection. Since this was to be a prototype for a police system the size of the collection should be such that it would give some indication of its applicability to the mug-shot system which might be held by a moderately sized police force. It was decided that 1000 faces would meet this requirement, for although many police forces would have photographic collections larger than this, the size of the collection to be searched could be reduced by preliminary screening, for example using the nature of the offence. The second guiding consideration was that the sample of faces should be representative in age distribution, and the presence of