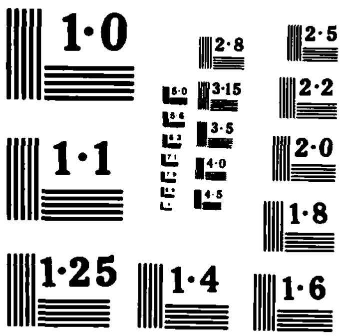
AD-A158 073 A MODEL OF AN EXPERT COMPUTER VISION AND RECOGNITION FACILITY WITH APPLICATIONS OF A PROPORTION TECHNIQUE (U) ARMY MILITARY PERSONNEL CENTER ALEXANDRIA VA 1/2

UNCLASSIFIED G E SHERMAN 15 MAY 85 F/G 9/2 NL



NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

A Model of an Expert Computer Vision and Recognition Facility With Applications of a Proportion Technique

George E. Sherman CPT
HQDA, MILPERCEN (DAPC-OPA-E)
200 Stovall Street
Alexandria, VA 22332

Final Report 15 May 1985

A thesis submitted to Kansas State University, Manhattan, Kansas, in partial fulfillment of the requirements for the degree of Master of Science, Department of Computer Science.

UNCLASS

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. A158 073	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Model of an Expert Computer Vision and Recognition Facility With Applications of a Proportion Technique		5. TYPE OF REPORT & PERIOD COVERED 15 May 1985
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) George E. Sherman, CPT, SigC		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Student, HQDA, MILPERCEN (DAPC-OPA-E), 200 Stovall Street Alexandria, VA 22332		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS HQDA, MILPERCEN, ATTN: DAPC-OPA-E, 200 Stovall Street Alexandria, VA 22332		12. REPORT DATE 15 May 1985
		13. NUMBER OF PAGES 115
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASS
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(Distribution A)

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This thesis is a model of human face recognition supported by computer vision and artificial intelligence techniques. Majority of research conducted at Kansas State University.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

"Expert Computer Vision"
"Computer Vision and Recognition"
"Human Face Recognition"
"Human Face Proportion"

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An expert vision system is proposed which infers, from the collection and evaluation of facial segments of a face, a set of possible identities. Retrievals from a data store .

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASS

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20 cont...

to accomplish this are facilitated by using the shape representations as keys to the records. The heuristic involves collecting the candidate identities into sets and then reducing the number of identities in those sets by set intersection and a weighting scheme. Sensitivity and tolerance factors in this expert are adjustable in terms of range values. The value obtained during evaluation of a facial segment is accepted or tolerated relative to the range of values expected for that segment. A retrieval effects a subrange whose width is determined by sensitivity to be the center and some number of adjacent values above and below that center value. Expectation ranges can be tailored to incorporate system and environmental variables.

A Model of an Expert Computer Vision and Recognition Facility With Applications of a Proportion Technique

by

GEORGE EDWARD SHERMAN

B.S., United States Military Academy, West Point, 1976

A MASTER'S THESIS

Submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

College of Arts and Sciences

Department of Computer Science

Kansas State University
Manhattan, Kansas 66506

1985

Approved by:

Major Professor

CONTENTS

	Page
PREFACE	i
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	iv
LIST OF FIGURES	v
Chapter	
1. INTRODUCTION	1
PHILOSOPHICAL SUPPORT	1
VISION CONCEPTS AND ALGORITHMS	1
EXPERT SYSTEM DEFINITION	8
RELEVANT CURRENT WORK	9
PROBLEM STATEMENT AND HYPOTHESES	15
PROPOSED SOLUTION	17
LIMITATIONS OF THE SOLUTION	19
2. LOW LEVEL MODEL	21
INTRODUCTION	21
DEFINITION OF TERMS	21
LOW LEVEL VISUAL DISCERNMENT	23
3. INTERMEDIATE PROCESSING PHASE MODEL	31
INTRODUCTION	31
TRANSFORMATION OF IMAGE QUERY	32
RELATING MODEL AND OBSERVATION	36
INTERMEDIATE PROCESSING ALGORITHM	39
4. HIGH LEVEL DISCERNMENT MODEL	42
PROBLEM DOMAIN CONCEPTUAL MODEL	42
TOP DOWN VISUAL DISCERNMENT MODEL	48

5. SIMULATED USE OF MODEL 55

INTRODUCTION 55

MODEL EXECUTION 55

6. RESULTS AND CONCLUSIONS 58

CHALLENGES TO METHOD 59

CONCLUSIONS 61

RECOMMENDATIONS FOR FURTHER STUDY 61

REFERENCES 65

SELECTED BIBLIOGRAPHY 68

APPENDIX A 72

APPENDIX B 81

APPENDIX C 86

APPENDIX D 90

PREFACE

The face has a number of unique characteristics. Artists have mastered the techniques of portraying such uniqueness. Perhaps, the most siginificant techniques in partitioning or decomposing the face are described in the treatise of Leonardo da Vinci. He discusses at length the concepts young artists should employ to master the mindset of facial primitives. Once young artists have such visual images fixed in their thinking, da Vinci contends, their painting efforts will become less troublesome.

Da Vinci lists the key characteristics for the face in his notebooks, that have been translated by McGurdy. This list provided insight into the development of a template for database partitioning and describing image segments. The following are quotations from da Vinci's work which provide the relevant thoughts.

> "The Order of Learning to Draw"
> First of all copy drawings by a good master made by his art from nature and not as exercises; then from a relief, keeping by you a drawing done from the same relief; then from a good model, and of this you ought to make a practice. (MS.2038,Bib. Nat.33r.)
>
> "Of the Way to Fix in Your Mind the Form of a Face"
> If you desire to acquire facility in keeping in your mind the expression of a face, first learn by heart the various different kinds of heads, eyes, noses, mouths, chins, throats, and also necks and shoulders. To take as an instance noses. They are of ten types: straight, bulbous, deep-set, prominent either above or below the centre, aquiline, regular, ape-like, round, and pointed. These divisions hold good as regards profile. Seen from in front noses are of twelve types: thick in the middle, thin in the middle, with the tip broad and narrow at the base, with nostrils broad or narrow, or high or low, and with the openings either distended or hidden by the tip. And similarly you will find variety in the other features; of which things you

ought to make studies from nature and so fix them in your mind. Or when you have to draw a face from memory, carry with you a small note-book in which you have noted down such features, and then when you have cast a glance at the face of the person you wish to draw, you can then look privately and see which nose or mouth has a resemblance to it, and make a tiny mark against it in order to recognise it again at home. Of monstrous faces I here say nothing, for they are kept in mind without difficulty. (MS. 2038, Bib.Nat. 26v.)

"Of the Parts of the Face"
If nature had only one fixed standard for the proportions of the various parts, then the faces of all men would resemble each other to such a degree that it would be impossible to distinguish one from another; but she has varied the five parts of the face in such a way that although she has made an almost universal standard as to their size she has not observed it in the various conditions to such a degree as to prevent one from being clearly distinguished from another. (C.A. 119 v.a.)

The most important terms used in the extracts above are listed below:

1) relief, 2) good model, 3) learn by heart the various different kinds, 4) when you have to draw a face from memory, 5) a small note-book in which you have noted down such features, 6) make a tiny mark against it in order to 'recognise it' again, 7) standard for the proportions of the various parts, 8) varied the five parts of the face.

In the context of the writing above these terms hint as to the nature of an expert vision system designed to recognize human facial features. They have influenced the proposed model in this thesis.

ACKNOWLEDGEMENTS

I wish to express my appreciation to my graduate committee for their guidance during my entire graduate program. The committee members -- Dr. Elizabeth Unger, Dr. Roger Hartley and Dr. Rich McBride -- have all been most helpful.

Special thanks go to Dr. Elizabeth Unger, committee chairperson, for many hours of constructive suggestions, encouragement, and meaningful input during the research and preparation phases of this thesis.

I thank Dr. Rich McBride for providing guidance on research materials concerning image data bases and concurrency and Dr. Roger Hartley for his suggestions and explanations that helped unravel the mysteries of expert systems and "this thing called AI."

A special thank you goes to my wife, Karen, for her encouragement during my total graduate program.

Finally, I thank my friends Dan and Randy for their editing assistance and printer support and the people often forgotten -- the univeristy librarians.

LIST OF TABLES

Table Page

2-1. Example ranges calculated from a small random sampling 28

5-1. Knowledge base of persons' characteristics 56

LIST OF FIGURES

Figure		Page
1-1.	C Language code for average gray scale computation	4
1-2.	Bayes decision rule expressed as inequalities	8
1-3.	The relationship between Top-down and Bottom-up modeling	11
1-4.	Fu web grammar representation of a scene	11
1-5.	University of Rochester query-oriented vision system	12
1-6.	A brief outline of SKETCH support components	14
1-7.	A representation of the entire proposed vision model	18
2-1.	Representation of proportion of facial features . . .	22
2-2.	Triad with internal angle for scalar calculation . . .	23
2-3.	A drawing of a general feature showing centroid calculation by integratrion	24
2-4.	A feature with perpendicular axes drawn	24
2-5.	Distance to perimeter of feature measured realtive to shifted axes	25
2-6.	Partitions are depicted as ranges of scalar values . .	27
2-7.	Example triad with primary angle marked	29
3-1.	Actions of region and feature experts at the intermediate processing level	32
3-2.	Regional and feature experts' relationship to knowledge store	33
3-3.	Region and feature experts' relationship to identity sets	34
3-4.	Blackboard sharing between experts	36
3-5.	Model to quadtree relationship	38
4-1.	Assignment of experts to conceptual model by partitioned level of control	43

conceptual model) and allow data structure independence. The a priori mechanism appears to be fulfilled by using expert-associated 'feature tables' that are seen as ranges for types of features [Rhodes].

Mathematical Tools:

The most useful tool for a combined represention of probabilities and a priori knowledge in the proposed model is Bayesian Statistics. Walpole and Meyers summarize the intent of this statistcal method as follows:

> The Bayesian approach to statistical methods of estimation combines sample information with other available prior information that may appear to be pertinent. The probabilities associated with this prior information are called subjective probabilities in that they measure a person's degree of belief in a proposition. The person uses his own experience and knowledge as the basis for arriving at a subjective probability. [Walpole]

The key word is belief, which applies to an expert discerning the probability that it has found the correct feature. A further implication of this belief is that the dimensions of the feature are within a reasonable range of expectation.

The majority of information available to support calculations in this thesis comes from terrain analysis sources. The majority of techniques emphasize filtering for LANDSAT applications. Filtering supports feature extraction. Concepts such as cluster training and the use of the Bayes rule support feature boundary definition during preprocessing. The probabilistic solutions applied to terrain feature and image interpretation can be applied to the 'terrain' of the human face.

The second source of information to support the calculations is quantitative cardiology. A great deal of research is being function called WHATISFACE. [Rhodes][Tucker][Hogg][Sowa]

The model offering the most specific information about structure and relating directly to the human face is SKETCH.

Eleven features are used to support the SKETCH system as modeled on the Penry Facial Identification Technique. These features include: face, nose, hairline, ears, chin, eyes, hair, mouth, jaw, eyebrows, and cheeklines [Rhodes].

Each feature can be modified in position, size and relative location to other features. All are individually addressable; with alterations to one feature made relative to another.

The template types are: face, nose, hairline, ears, chin, eyes, hair, mouth, jaw, eyebrows, cheeklines.

The modifier types are: fat, wide, thin, big, large, small, high, low, right, rightwards, left, leftwards, tall, short, up, down, upwards, downwards, slender, north, south, east, west.

Figure 1-6 A brief outline of SKETCH support components.

Relating SKETCH to this thesis, these entries can correspond to 'expert' partitions using templates. Within the experts and between experts the use of modifiers supports decision making. Modifiers can be described as numerical ranges for object proportions. The ranges are contained within the knowledge base.

Although SKETCH uses a five module approach for interactive support, the expert is kept completely external to the system, i.e., recognition and decisions are completely human driven. An important tool in SKETCH is the hash-dictionary look-up task. The components known as the 'feature table', 'display list' and "WEBFT" form the domain of a given expert (derived from the

Figure 1-3 summarizes the relationship between top-down and bottom-up with respect to the proposed recognition model.

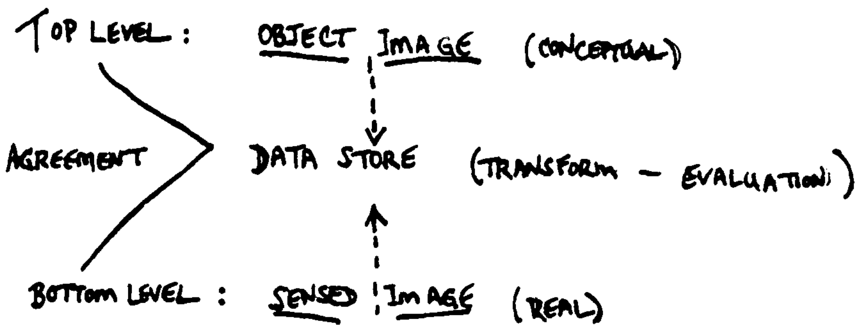


Figure 1-3 The relationship between Top-down and Bottom-up modeling.

The most important aspect of Fu's research is the concept of using web grammars to describe an object. The grammer is an abstract representation of an object's components, connected in order and relative spatial position, as they occur in the object. What results is a two-level hierarchy of representation with the top level being the object and the subordinate level being the compositional nature of the object. This grammar technique can prove to be useful in representing general knowledge about an object in the form of a grammar presentation model, see Figure 1-4.

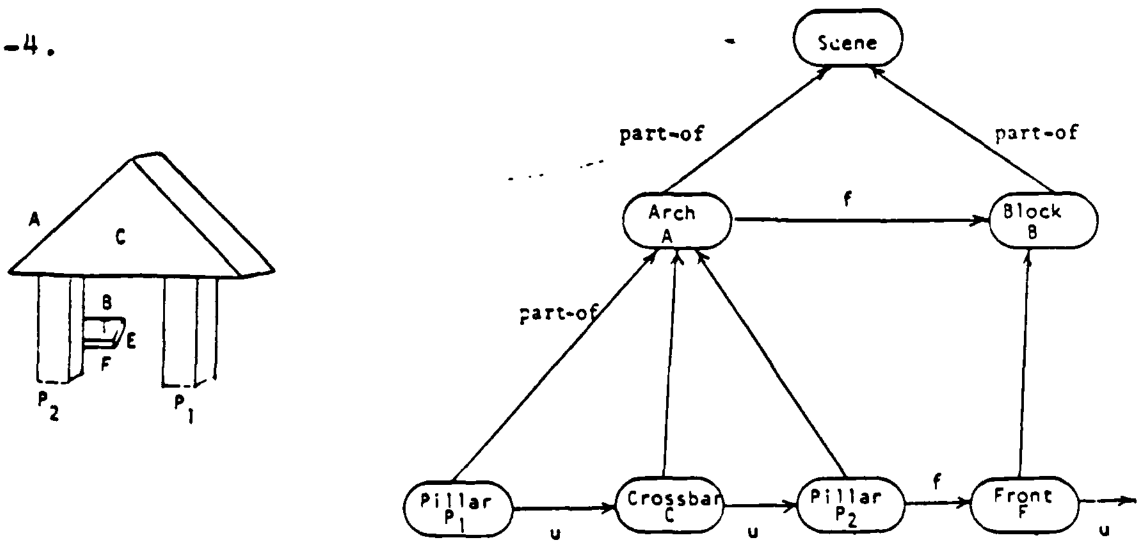


Figure 1-4 Fu web grammar representation of a scene. [Fu]

RELEVANT CURRENT WORK

A search of current literature yielded a number of possible techniques for representation of objects, recognition, image processing and calulation, image query forms, and image data base applications. The reviewed works ranged in focus from the conceptual modeling to the implementation levels of problem solving. Because of the complexity of the human face, however, only a few researchers have addressed the challenge of representing and recognizing the human face. The references that provided meaningful guidance germane to this thesis are described below in three sections models of vision, models of architectures for the support of vision systems and mathematical tools.

Models of Vision:

Marr supports the idea of considering vision at two levels. He asserts that computation must precede the design of analysis algorithms and that the two actions should not be mixed or confused. He further argues that the choice of representation affects the success of analysis [Cohen]. In some ways representation changes in phases, or as a metamorphosis, directed toward eventual understanding.

Marr's theory of vision is primarily oriented towards bottom-up processing from image to object. As mentioned in the discussion of segmentation, image interpretation is difficult without top-down models. However, new approaches tend to be more bottom-up in that they are based on physical properties of the world. The goal here is representation that allows constraints, provided in the real world, to be systematically exploited.

about its own inference process [Hayes-Roth] [Rich].

An expert system is often described as possessing expert rules and reasoning by manipulating symbols. A minimal functional capability of an expert system includes grasping fundamental domain principles. Expert systems derive their strength from an avoidance of 'blind search' and a reliance on weaker reasoning methods as 'functional reserves' when the expert rules fail. One luxury provided by the expert system is the ability to provide explanations for conclusions reached. The expert system described here far surpasses the capability of traditional numerical analysis programs. [Hayes-Roth]

An expert system's task domain can include monitoring, interpretation, prediction, and instruction. To accomplish any of these tasks, the expert system must accept problem terms and convert them into an internal representation appropriate for processing with its expert rules. By taking advantage of inference patterns, or expertise, an expert solves an assigned problem. The expert system's ability to interpret will become important in the development of concepts presented later.

All references to experts in this thesis imply the capabilities to make decisions and to interpret data based on prior knowledge of the problem domain. The expert vision system model proposed embodies many of the characteristics of a general expert system. The model proposed is primitive in some respects; the most important weakness is its inability to reason about its own inference processes and provide an explanation. The primary task of the proposed model is interpretation of visual images of the human face.

functions. These functions describe the probability of a pixel having some feature value given which class the pixel is in. Additionally, each function can be scaled by some a priori probability that a given class occurs in the image area of interest. The a priori probability mentioned above represents prior knowledge of the image area. Knowledge of the image area can be attained from either historical or experiential data. The Bayes decision rule is summarized in Figure 1-2 [Schowengerdt] [Walpole].

A pixel belongs to a class 1 if $p(x|1)p(1) > p(x|2)p(2)$
A pixel belongs to a class 2 if $p(x|2)p(2) > p(x|1)p(1)$

Figure 1-2 Bayes decision rule expressed as inequalities.

This rule works for all conditions, excluding the 'decision boundary' point, where the probablity functions intersect.

Bayes Theory will become useful in considering the peripheral features of the face; the features that will require some special attention include ears, hair, hairline, cheeks and jaw. Initial estimates for the dimensional ranges of these features will estabiish the probability functions for adjacent features. Statistical analysis will be useful in adjusting these probability functions. Bayesian Statistics will play a role in this thesis.

EXPERT SYSTEM DEFINITION

Defining an expert system is a difficult task. Generally, they differ from data processing systems and are defined as a system performing at an expert level using domain-specific problem-solving strategies, with the additional ability to reason an image, such as measures of spatial structure, may provide more useful information for classification. Thus it is wise to consider various pre-classification manipulations and transformations to extract the greatest amount of information from the original image. In some ways this can be considered as data normalization and feature extraction. Our overall goal is to extract features as homogeneous sub-regions within the total image area.

Under the circumstances described above, it seems appropriate to assign the task of classification to a computer. The potential for consistent and efficient analysis of a given surface by a computer promises a definite speed advantage over manual techniques.

Image classification is a decision-making process with data that can exhibit considerable statistical variance. This characteristic variance suggests the need to wisely employ mathematical tools from statistical theory. In reality, classifying a pixel into a particular class is a statistically intelligent guess. There is a probability of error here. It follows logically that a given decision made at pixel level should minimize some error criterion throughout the classified area, which is synonymous with a large number of individual pixel classifications. This goal can be termed 'maximum-likelihood', which is commonly known as Bayes optimal classification [Rich].

Bayes Theory concentrates on the problem of measuring some feature of a scene and deciding to which of two classes a pixel belongs. By calculating a relative frequency histogram of a feature we can approximate the continuous probability density

Detection of edges has been mentioned already in the discussion of thresholding; however, this becomes critical to our ability to enclose objects within well-defined boundaries. Centroid calculation and proportion analysis require these well-defined object boundaries.

Edge detection is a classical problem in image processing that is summarized as the detection of sudden changes in gray level from one pixel to another. Such changes usually indicate a boundary, or edge, between two distinctly different objects in the image. There are many approaches to this problem; one simple technique involves the thresholding principles described above.

As a specific example, gray level threshold can be applied to the gradient image of the face, or any other candidate surface, resulting in lines at edge boundaries. A compromise threshold must usually be accepted because a threshold that is too low results in many isolated pixels being identified and thick, poorly defined edge boundaries. A threshold that is too high results in thin, broken line segments. There are post-threshold processing techniques, called 'line thinning' and 'connecting', that help alleviate these problems. They appear to be partially successful but incur additional image processing costs. The ability to define boundaries provides the foundation for further computation and decision criteria.

Those aspects of remote sensor imagery that are used to define mapping classes are known as features. The simplest features, the pixel gray levels in each band of a multispectral image, are not necessarily the best features for accurate classification. Furthermore, more complex features derived from discernible regions. This technique of thresholding will be developed next.

Thresholding is a type of contrast manipulation that is not designed to enhance contrast. The objective is to use contrast rather than enhance it. Instead, it 'segments' an image into two classes defined by a single gray level threshold. The use of a binary threshold on certain types of images results in sharply defined spatial boundaries that may be used for masking portions of the image. Separate processing may then be applied to each of the two classes and the results recombined to alleviate the difficulties encountered with images. Thresholding may also be used as a simple classification algorithm, for example in a decision tree classifier. Thresholding applications are primarily intended to detect change in a pair of multitemporal image types [Schowengerdt].

Thresholding can be applied to objects registered using control points and geometric transformation. Changes in gray levels can be used to detect edges when set to detect differences in gray levels that exceed a decided magnitude. R. Schowengerdt argues that the selection of the 'best' threshold level is difficult and must usually be associated with a priori knowledge about the scene, or visual interpretation, to be meaningful. [Schowengerdt] In the case of the face, we are primarily interested in features that are decidedly identifiable by contrast. If it is possible to merge pixels into non-essential regions, this should be done to reduce the complexity of the image space.

```
int gray_average ()
{ int i,j ;
  int sum = 0 ;
  int avg ;
  for (i = 0; i < 10; i++)
    for (j = 0; j < 10; j++)
      sum + = PIXEL_ARRAY [i][j] ;
  avg = sum / (10 * 10) ;
  return (avg) ;
}
```

Figure 1-1 C Language code for average gray scale computation. [Lecky]

Other algorithms include convolutions and filters which are commonly used for smoothing and feature detection. Connectivity analysis, is reasonably simple to implement. Each pixel is examined and, if it is above a certain brightness, grouped with its bright neighbors. In this way contiguous features are identified. Any operation that has classically been applied to a data array can be applied to an array of pixel values, sometimes producing useful information.

Industrial applications usually require moderately fast, relatively fault-tolerant algorithms. These algorithms must be extremely stable, reliable, and repeatable to be of any use on a production floor. The algorithms used in industrial machines are highly specialized and development continues far past the prototype stage, usually over the course of several months at customer sites. Literally hundreds of algorithms and variations must be developed, implemented, and tested to arrive at the final version.

A digitized image, such as that of the human face, can be scanned to insure that the pixels offering information are kept while those below a given threshold are discarded. The goal here is to reduce the image search space to a collection of of nothing more than interpreting this massive data array. All software operations are standard. For example, to find the overall brightness level seen by the camera, we need only to compute the average value of all the elements in the array.

The real challenge lies in the manner in which the data sets containing large arrays are processed to extract useful information. The goal of representing the large arrays of data to avoid wasted storage directly relates to the extraction of information.

In vision applications, especially in research and small industrial companies, software generation is a day-to-day occurrence. Algorithms tend to be developed empirically. Thresholds, gains, and other inputs that may be dependent upon lighting or material variations must be automatically computed through the use of some algorithm. Vision machines must be given some notion of the difference between good and bad.

One of the most useful vision algorithms is the gray average which computes the average value a pixel takes on over a certain area. Another algorithm, equally widely known but more useful, is a variation of the gray average -- a simple weighting or moment, is added in for each gray level based on its distance from some base pixel value. This essentially amplifies the impact of very bright or very dark pixels. A third common algorithm is the template match, a "snapshot" technique in which a reference image is stored away and compared, pixel by pixel, to the current image. An example code segment used for the purpose of calculating the gray average is presented in Figure 1-1.

creates a "vision computer." The camera is equipped with standard lenses and filters and provided with a mounting bracket and a light source to create a desired image of an object. The camera converts the image that it sees into a video signal exactly like those found in an ordinary black and white television. The computer then uses special hardware to digitize that image, converting it to discrete number values representing the light intensity of each picture element, or pixel. A value of zero indicates that the camera sees no light coming from that particular region. If the camera barely senses light in a region, the pixel values for that area will be 1.

As more light is sensed, the pixel value increases until that pixel value of the camera becomes saturated with light, unable to measure any additional increase in brightness. The pixel value for this light level depends on the precision of the analog-to-digital conversion hardware in the computer and is typically 63, 127, or 255 for 6-, 7- or 8-bit precision, respectively. For most purposes, dividing the sensitivity range of the camera into 64 different brightness levels, or 'gray levels', is more than adequate, since a variation in light intensity of less than 5 percent of the range of the camera is seldom meaningful. Industrial performance variance is normally around 30 percent of the range of the camera, meaning that as few as ten camera responses or gray levels are necessary. Therefore, the most commonly used gray scale runs from 0 to 63.

Once the image has been converted to digital values, the pixels are automatically loaded into the computer's memory as a large two-dimensional array. Processing a video image consists

Chapter 1

INTRODUCTION

PHILOSOPHICAL SUPPORT

Understanding an image requires a priori knowledge of the task domain. Although features observed may be weak in detail, a person knows what to look for in the image. Image understanding is impossible without expectation.

In keeping with the principle above it is wise to split the task of vision into low and high level discernment. The low level discernment is synonymous with early processing. The high level detail discernment is done later in the processing, it is possible only after intermediate processing or segmentation. High level discernment embodies the handling of objects and relies heavily on domain-specific knowledge to construct descriptions of scenes.

Fu states that similarity measures, feature selection, and feature extraction are fundamental in recognizing human faces. He contends that automatic identification, classification, storage, and retrieval of human faces "could have considerable utility in many personnel, commercial, security, and law-enforcement applications" [Fu]. A review of the abilities and applications of current vision systems will prove useful in describing the starting point for solving the problem of human face recognition.

VISION CONCEPTS AND ALGORITHMS

A video camera, when coupled to a computer which has been equipped with hardware that enables it to read that camera,

4-2. Connectivity graph of facial features 44

4-3. Fu web grammer showing adjacency of human facial features 46

4-4. Sowa conceptual graph of the human face 47

4-5. Hierarchy of frames representation of the face 48

4-6. Pipeline diagram supported by queues 54

5-1. Person #10's alternate photograph 59

5-2. Person unknown to knowledge base 61

devoted to this area. The similarity between the face and the heart has primarily to do with the restricted domain that the two objects present. Work done by H. Sandler in measuring ventricular dimensions from video images has led to an improved version of the area-length method [Sandler]. Of more significant value is Sandler's reliance on statistical measures to match expectations of dimensions. The following statement provided meaningful guidance on a technique for representing 'expectation' in this thesis.

> Chamber dimensions are directly measured or derived from recorded images. Volumes calculated by these methods are corrected or adjusted by statistical equations for overestimation of actual ventricular volumes. When such corrections are not used, the resulting errors are so large as to make calculated volumes unreliable. [Sandler]

The ability to adjust visual recognition and control interpretations of images to capture 'actual' dimensions suggested a solution to face recognition and identification.

PROBLEM STATEMENT AND HYPOTHESES

The problem selected is to determine theoretically whether or not the human face can be represented and recognized through the use of proportional measures of the facial features and their proximity to one another. To provide a limited solution to this complex problem it will be necessary to develop some heuristics. The hypotheses are:

1. It is possible to use a technique that computes a shape-relative, or proportional, scalar in the spatial or picture domain for the purpose of feature representation.

The primary purpose for finding scalars is to extract characteristics that uniquely describe shapes independent of the global coordinate system. Assuming visual patterns exist that appear clustered, transforms can be used to produce scalars. The transforms ignore the relative spatial orientation of these clusters. The scalars can be coordinates or distances. Normalization techniques can be applied to these scalars. Proportional scalars can describe patterns or clusters. These proportional calculations can reduce clusters to scalars that are independent of an initial coordinate system.

2. It is possible to calculate a second scalar that expresses the relative position of features, within the domain of the face, with respect to each other. This scalar can be expressed as an angular measure that is bounded at the vertices by three feature-centroids. This three-centroid group is appropriately refered to as a triad.

3. Facial feature partitions can be expressed as ranges of scalar values. These ranges can delimit feature-based 'types.'

4. Instances of features can be expressed as subranges or subsets of range-delimited types. The subranges relate to associated identities of the person's face. Subranges can also be used to specify expectations quantitatively.

5. The notion of 'expectation' can be incorporated into an expert vision system. Expectation takes the form of a domain model, expressed as a hierarchy of frames, derived from a web grammar or connectivity graph of types.

The domain model allows the generation of hypotheses. The experts functionally guide the interfacing of hypotheses and the intermediate range representation of types. Guidance is defined as set operations and confirmation of hypotheses within the system ranges.

6. System ranges translate to data partitions which are actually collections of sets of like records. Data partitions, in turn, permit the generation of candidate sets of feature owners. These sets can be reduced by allowing interaction between experts thereby resolving the owner sets to the smallest (best possible) estimation of owner identity. This can be considered a leveled query.

PROPOSED SOLUTION

The method of human face recognition proposed is based on a model. The model provides structure to the method and is global relative to the task of recognition. The model is applied to pattern recognition activities through the controlled use of statistical classification. Once appropriate certainty levels are reached, a data store can be accessed for specific identification.

The human face represents structured data while the features of the face that participate in this structure can statistically vary in dimension. The structure of the face provides the global control for feature extraction. The variance of feature dimensions provides the foundation for statistical classification and identity matching. The coupling of structure model to statistical classification is the goal and intent of this thesis.

The model expert vision system proposed can be depicted in a level diagram as shown in Figure 1-7.

The query-oriented vision system of University of Rochester uses a layer technique. The layers are called the image data structure, the sketchmap and the model layer [Cohen]. Their system performs recognition as a series of query resolutions. Although their domain models are quite simple, this author believes that their techniques can be expanded to support more complex domains. The layers of the recognition mechanism are shown in Figure 1-5.

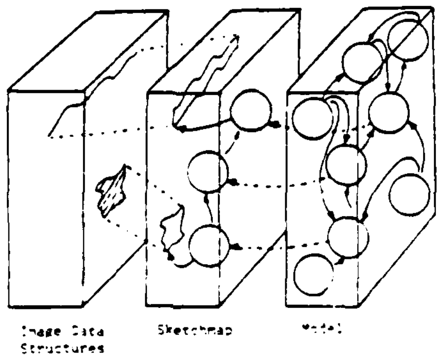


Figure 1-5 University of Rochester query-oriented vision system. [Cohen]

It is important to notice that the boundaries of transformations between representation forms are implied. It is still not clear in their writing as to how mapping between the levels occurs. These gaps between levels introduce the need for transform and mapping techniques which are proposed in this thesis.

Architectures for the Support of Vision Systems:

The candidates for the architecture of an expert vision system include 1) a model known as SKETCH, 2) a biomedical neuron cell recognizer known as Lewis Tucker's Expert Vision System, 3) a model-based University of Brighton program known as WALKER and 4) a simple system with the human providing the entire expert

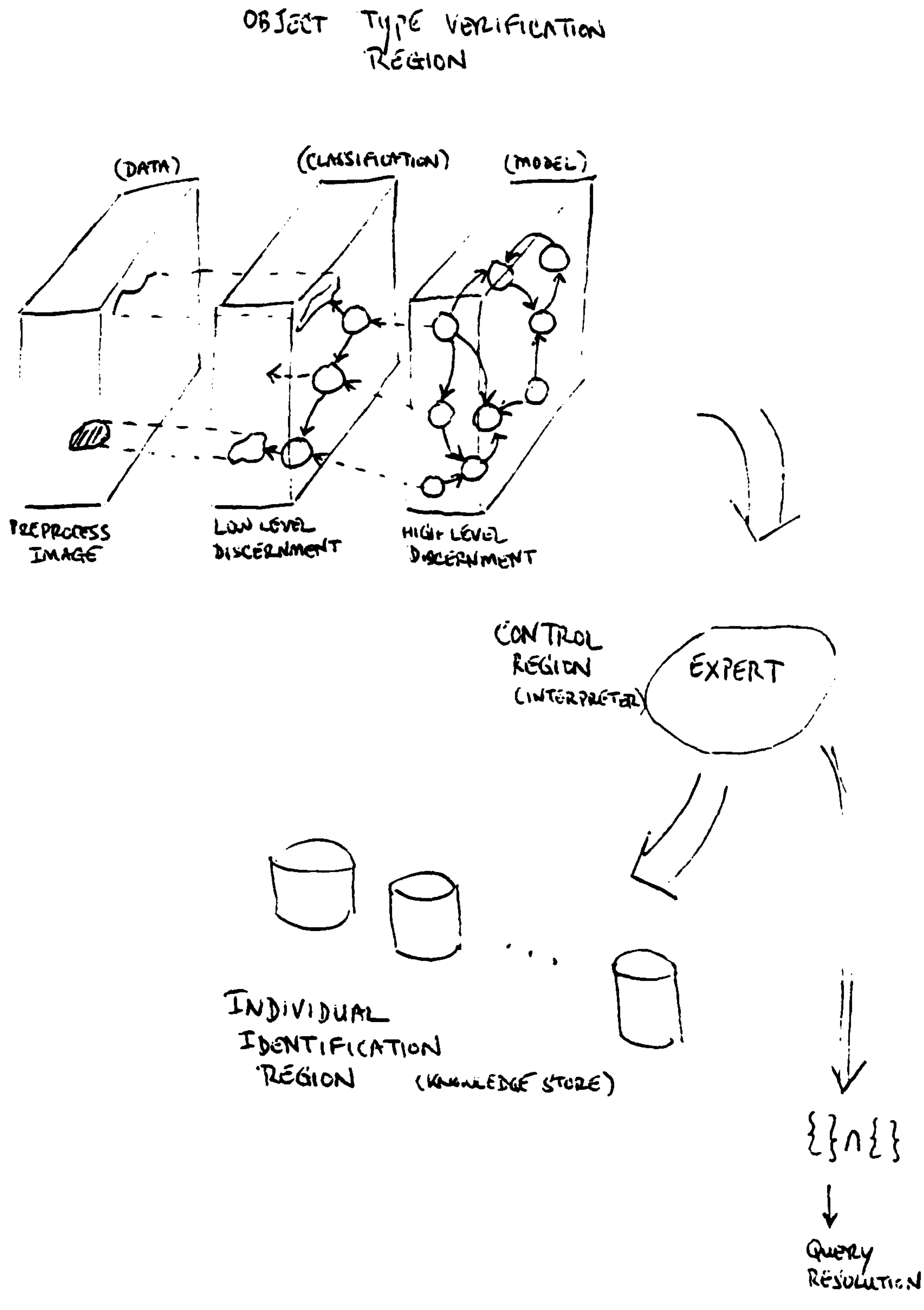


Figure 1-7 A representation of the entire proposed vision model. [Cohen]

LIMITATIONS OF THE SOLUTION

The model itself is theoretically supportable; however, the realization of the mechanisms will depend a great deal on the ability to clearly view and digitize facial images. Experimentation with real data will provide the support for fine tuning the model. Whether the model will be successful in identifying one individual remains to be proven, although it is theoretically possible. At worst the model will provide a useful heuristic for reducing matching algorithm workloads by limiting the search sets of candidate images.

The clarity of images, whether from photograph or video sources, is best supported by frame-grabbing technology. Frame-grabbing is defined as a video 'snapshot' technique that allows very fast recordings of video images. The speed of digitization allows computation to proceed nearly immediately. Frame-grabbing can occur at speeds of 1/30 of a second which allows multiple samplings of the surrounding environment. Although this equipment is not available in the KSU Department of Computer Science, preprocessed images from two research sources provided a limited starting point. Analysis of these images was directed toward determining whether individual features are discernable.

This thesis consists of five additional chapters. Chapter 2 provides the formal definitions supporting the proposed concepts and an indepth discussion of low level discernment. Chapter 3 discusses image query components, actions accomplished at the transitions between levels, intermediate vision model levels, and the intermediate processing algorithm. Chapter 4 provides a discussion of problem domain modeling, high level discernment objectives, and a summary top-down procedural vision model. Chapter 5 provides a limited simulation of the vision model which involves the use of preliminary investigations of the statistical characteristics of human faces, and facial features, in support of statistical classification. The simulation also makes use of structural model and knowledge store integration. Chapter 6 provides a discussion of the results of the model design, conclusions and suggestions for future work. A brief discussion of challenges to the model is also included in the last chapter.

Chapter 2

LOW LEVEL MODEL

INTRODUCTION

The discussion of the system model begins with low level discernment. Low level discernment is primarily computational in nature involving the calculation of dimensional characteristics of facial features viewed by the vision system. The primary goal of low level computation is to provide proportional and angular scalars for use in data base transactions during image query processing. This chapter provides formal definitions in support of concepts discussed in this thesis. Additionally, the lower level of the expert vision model is developed in behavioral detail with supporting explanations for the choices of mathematical tools.

DEFINITION OF TERMS

With respect to Sampling Theory:

DEFINITION 1. Range

Range is defined as a simple computation of variability of a random sample. Formally, this appears as follows:

> The range of a random sample x1, x2, ..., xn, arranged in increasing order of magnitude, is defined by the statistic
>
> xn - x1 [Myers]

DEFINITION 2. Variance

Variance is defined as a measure of variability that considers the position of an observation relative to the sample mean. Formally, this appears as follows:

> If x1, x2, ..., xn represent a random sample of size n, then the sample variance is defined by the statistic
>
> S**2 = For i:1..n { SUM ((xi - xmean)**2) } / (n -1) [Myers]

DEFINITION 3. Centroid

Centroid is defined as the center of mass of an object having constant density. In formula terms it is expressed as follows:

Given $m = \iint_R SD \, dA$, where m is mass, R is the region, SD is the surface density function, with respect to dA (changes in area)

and Given $Mx = \iint_R y*SD \, dy \, dx$, where Mx is the moment of x

and Given $My = \iint_R x*SD \, dy \, dx$, where My is the moment of y

Then the centroid is (xbar, ybar) where

xbar = My/m and ybar = Mx/m . [Goodman]

DEFINITION 4. Proportion

Given a shape S in a plane P, let p = f(S) be a point in P computed from shape S according to procedure f, where f determines the centroid. Also, let a and b be line segments, each having one endpoint at p and the other on shape S. Hold the line segments a and b perpendicular at p. The proportional ratio is R = a/b. Then R is a shape-relative ratio of S. See Figure 2-1.

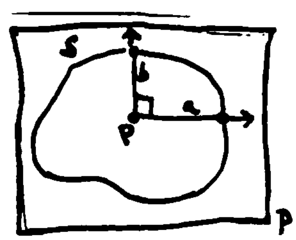


Figure 2-1 Representation of proportion of facial features. [Morrill]

DEFINITION 5. Triad and Angular Measure

Triad is defined generally as a group of three things. Things in the context of this thesis are actually centroids. Therefore a triad is a grouping of three selected centroids, forming a trinity. In presentation form it appears in Figure 2-2.

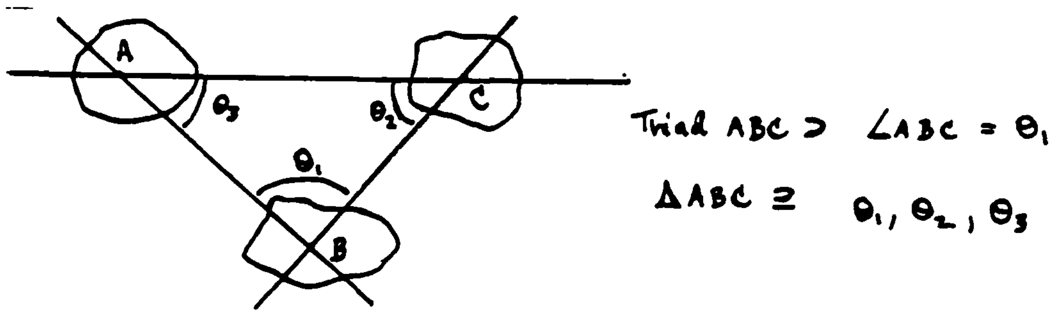


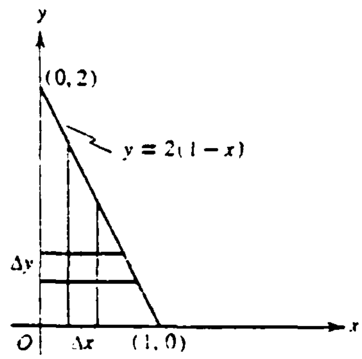
Figure 2-2 Triad with internal angle for scalar calculation.

LOW LEVEL VISUAL DISCERNMENT

In this expert vision system a suitable registration tool is required for identifying regions and objects. A registration tool is a calculation technique that provides dimensional measures with registration being the association of the measures with the region or object being examined. Registration is critical to the identification of unique objects. A computational form is required that offers a reasonable degree of uniqueness. Uniqueness, in this sense, is understood as the ability to distinguish between types of elements. Additionally, a major goal of uniqueness is to establish set membership and to reduce sets when functionally possible. By registering elements, and by providing a means for restricted grouping in the form of sets, we will be able to establish identity.

One technique for unique registration of an element is the calculation of the element's centroid. An element has a characteristic shape and content. By reducing the shape and content to a single centroid, we describe a unique center of mass for that element. The centroid is formed by taking n multiple integrals over n-space, and is expressed as a coordinate based on the coordinate space within which the calculation is performed. This centroid is object-unique, in the context of the object.

Equation expressions for the centroid were presented previously in Definition 3; however, a diagram of an example object with an indication of the centroid calculation provides further clarity. See Figure 2-3.



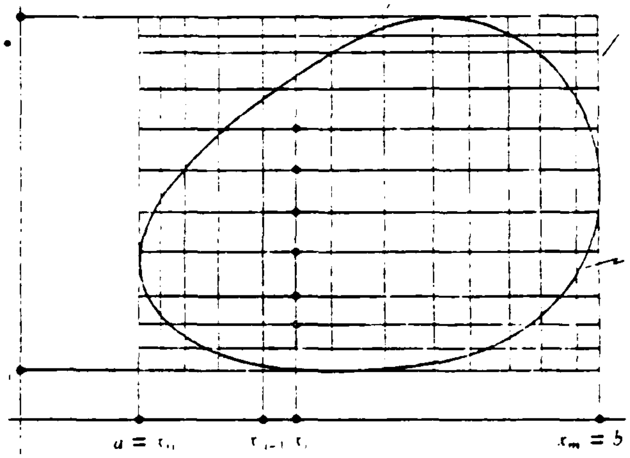


Figure 2-3 A drawing of a general feature showing centroid calculation by integration. [Goodman]

The centroid for a given element is not always unique in the element's environment. The element might have the same coordinate value for center of mass (centroid) as some other element. This can happen if centroids are computed from a set of coordinate planes that are aligned for each element individually, see Figure 2-4. Our problem now is to include sufficient information with the centroid to preserve uniqueness of the element beyond its own context.

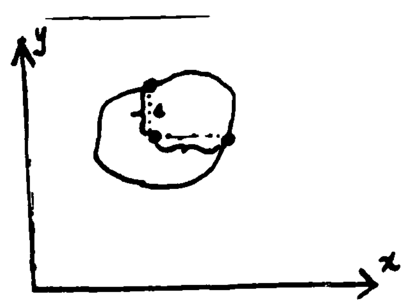


Figure 2-4 A feature with perpendicular axes drawn.

The technique proposed in this thesis involves the calculation of the centroid followed by the examination of the image for distances from that centroid to perimeter sections of that element, see Figure 2-5. Once these distance values have been computed their ratio can be used to record the proportionality of the element. It is important to emphasize that the proportion of the element can be used to uniquely describe it. If the proportions of the object are similar to other elements, then these elements can be grouped into sets and appropriately named for their 'type.' The grouping of elements into sets, based on similarity, implies a range of proportional values. This concept of ranges will become important when the object experts verify the type of the object they are viewing.

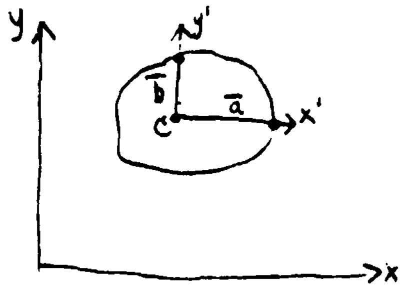


Figure 2-5 Distance to perimeter of feature measured relative to shifted axes.

The human face possesses a minimum of eleven critical features, as listed in Figure 1-6, that can guide the task of partitioning [Rhodes]. The differences in the proportionality of these features can be used to our advantage. There are two attributes of the elements that insure a high degree of mutual exclusion between partitions: adjacency and proportionality. Proportionality has been explained above, and can be thought of as moving the coordinate planes of scene space from outside the element to within it. The origin of the new coordinate space is the element's centroid. The distances to the perimeter lie along these coordinate axes. In a sense, the element carries its unique proportions in the form of a ratio. [DaVinci] This ratio is used as a 'key' which can support retrieval from a knowledge or data store. The retrieval action might result in a response being a subset, with the number of elements greater than or equal to one. in the event the ratio is within a proper range. The query might be better handled as a many-membered subset at lowlevel, with the highlevel experts making the decision as to which candidate identities to retain. The data stores can be partitioned then into eleven groups thereby alligning the knowledge store and experts. Proportionality ratios provide the retrieval link needed to assemble candidate identities.

The ability to use adjacency at the low level visual discernment requires proportional measures that cross partition boundaries. Although high level discernment processes are provided knowledge of adjacency, from the conceptual model, this knowledge is not sufficient to discern beyond the 'type' of object being viewed. The approach proposed here is a technique using the relative positions of centroids in face-image-space. If coordinate planes are used within the domain of the system, or scene the elements within the system can be individually addressed by their centroids. The complexity of the face can be reduced by calculating angles between these element centroids in the face-image-space. Each human face exhibits a combined proportional geometry. If this geometry can be reduced to simple ratios of proportion, then retrieval and association can be accomplished easily. The question at this point, is whether precision is critical at this low level or should combinatorial techniques at the high level handle the intersection of identity member sets? A tradeoff in complexity of computation occurs at this point.

With the above knowledge, it is possible to calculate data base keys that map to data base partitions. These partitions can contain pointers to human identitites. Therefore, a query causes navigation to a partition and allows the return of a set of pointers to candidate identities. This series of query actions occurs assuming that a key provided by the low level discernment model is within the range of tolerance for a given target partition, see Figure 2-6. The next major issue is the control of the query to insure that access to a data base partition occurs only when the query key has a value within the expected range of the object-partition.

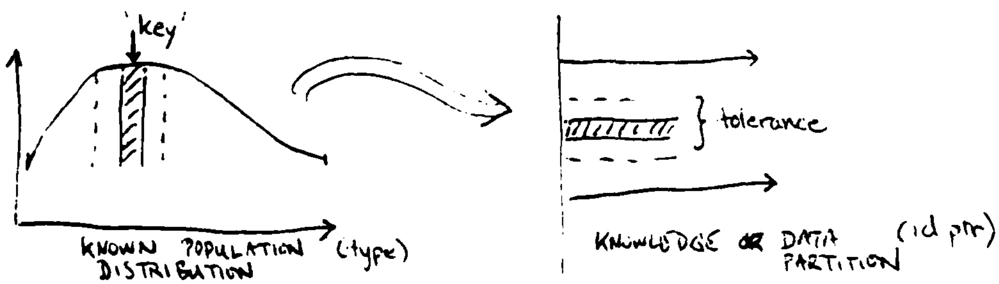


Figure 2-6 Partitions are depicted as ranges of scalar values.

The facial element groups are described as ranges of measurement ratios. Within these groups some examples of proportion ratios appear as shown in Table 2-1.

3) Hypothesize and test the candidate identity. There might be more than one candidate from a limited retrieval of owners.

4) Hypothesize at 'parent level', based on interrelationships of features (centroid axes and distance measures) who the best possible candidates are. Remember, the candidate sets are reduced as the hierarchy is traversed in the upward direction. Resolution of the image query will occur in this way.

5) Share composite hypothesis with subordinate experts for shared goal. This might aid those experts that have not resolved features to a desired clarity.

6) Consider the effects of unique identifiers such as scars or mal-formed characteristics that might reduce the search space of subordinate experts.

The algorithm above also provides a concise review of the intermediate processing portion of the vision model. The intermediate level goals are three-fold: 1) Provide computational characteristics of the image features being viewed, 2) confirm the hypothesis or expectation of an expert that a feature type at a given location matches the model prediction and 3) after confirming a feature an expert should provide a set of candidate identities for the owner.

In the next chapter the sources and design issues concerning the controlling domain model and processing of high level hypotheses will be discussed. High level discernment will be discussed in terms of candidate set operations above the regional expert level. The relationship between intermediate processing and high level discernment will be discussed with an emphasis on expert process control.

data store. The measure of the quadtree tile reduces to this numerical identifier. An example of quadtree processing of a human face is shown at Appendix D [Omolayole].

INTERMEDIATE PROCESSING ALGORITHM

The tasks accomplished at the intermediate phase of processing in the visual model can be summarized as an algorithm. The previous discussion introduced the concepts of proportionality, ratio, and dimension. The low level mechanisms can work to provide the ratios and dimensions while the high level mechanisms can concentrate on 'relating' these dimensions to the domain model. The relating referred to here is actually the hypothesis scheme already introduced in this thesis (see Chapter 2, Table 2-1). Experts must be able to verify, or disprove, hypotheses and decide consequent actions. Relating dimensional information to specific identities inherently supports the task of hypotheses verification of feature-type, and allows the recognition system to discern to a much higher capability. We do not merely want to ascertain that we have found a nose; rather we want to continue the analysis to the point of identifying the owner. The following algorithm, in draft, describes the interactions of high and low level vision mechanisms.

Intermediate Algorithm:

1) Form hypothesis about features based on expected location and approximate shape, referring to the conceptual graph.

2) Explore the boundary of the feature, calculating centroid and centroid axes. The use of quad-trees at this point is critical.

Quadtrees can provide this interface. The relationship of the model to the quadtree hierarchy is depicted in Figure 3-5.

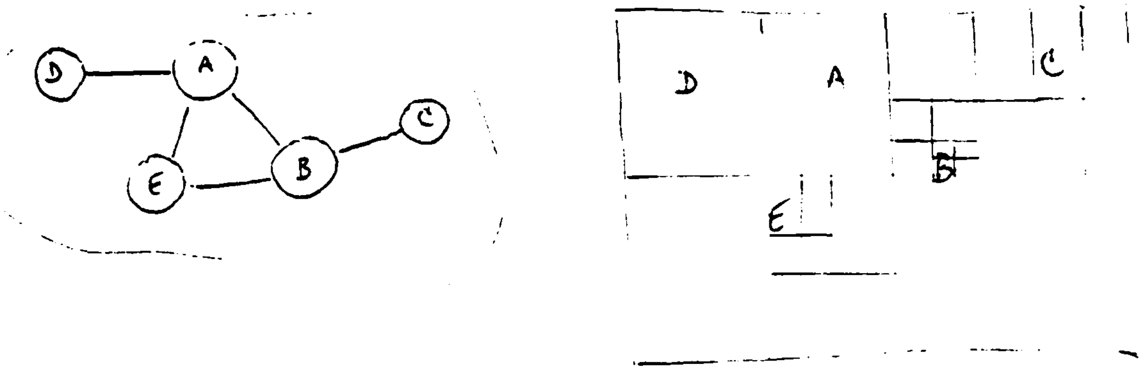


Figure 3-5 Model to quadtree relationship.

The threshold and edge detection techniques, described previously as preprocessing techniques in Chapter 1, play an important role in the activities of the tiles formed by quadtree techniques. The pixel information within a tile or window, at a given level of the quadtree, can be processed using edge detection techniques for boundary and level controls [Shu-Xiang][Omolayole][Grosky].

Quadtrees facilitate large chunks of ambiguous regions. Limiting the regions to quads allows an economy of storage space and computational complexity as compared to individual pixel management. If a quad-group is used for representation, then it follows that quad averaging is possible without losing resolution. A quadtree can be evaluated to a single average value. If the tree is summed or consolidated to a given level, for example an eye or ear, the region then has a unique, or possibly unique, numerical identifier which can be hashed to a resulting in an association with candidate owners (see Figures 2-2 and 2-6).

The actions of an expert can be described as a series of the following three steps: hypothesis generation, result testing, and verification or rejection of the original hypotheses. This series of actions would probably be sufficient to discern that the image is a face or a facial feature, however, it is not complete to the point of being able to recognize that the face belongs to a particular person. A really strong hypothesis will depend on the low level computational analysis of the region in question. The conceptual model can steer the hypothesis steps, while proportional and triad-angle analyses can support the collection of candidate identities. Once candidates have been gathered, the region experts can compare and decide on the validity of these candidates thereby forming a best-hypothesis as to facial region identity.

The feature expert must be controlled during the viewing of an image. The control mechanism is the conceptual model. The model forms an expectation of feature type, location, and characteristics. The issue of viewing must be addressed in order to decide to what degree of resolution the image is to be investigated. The resolution of the viewing mechanism can start effectively from the model level using quadtree techniques. The model level becomes the root level of the quadtree. Starting at the model level avoids the magnitude of individual pixel storage. At a machine primitive level the steering of the 'eye' of an expert is performed by the model. The model must be interfaced to the image of interest for any further actions to occur.

The image has been abstracted to collections of proportion ranges between features, within features, and allows a best estimate when directed by the conceptual model to options for image clarification. Directing by the conceptual model is actually a control issue. A feature expert can anticipate, or expect, its feature based on proportionality criteria. A region expert expects the relative positioning of its features based on geometric criteria. The assembly of features and regions requires control by experts and eventual supervision by some superior expert to resolve an image query.

RELATING MODEL AND OBSERVATION

The experts of this proposed system model must be given sufficient 'expertise' to allow classification and recognition of facial features and eventually the entire face. This expertise must occur as a result of careful integration of low and high level discernment mechanisms. Integration involves development of a causal relationship between the expert and low level analysis techniques. Expertise, as a causal relationship, is analogous to the hypothesis and confirmation cycle. This cycle will be defined and expanded in the following discussion.

A feature hypothesis is centered on feature 'proportionality' within itself (see Table 2-1 and Figure 2-6). The dimensions of a feature provide characteristic values that can be used to associate a feature type with candidate owners.

A region hypothesis is based on features and, more importantly, on the proximity of features. As a feature is described in terms of its own proportionality it is then thought of in terms of proportionality within the scene of the face, image identities based on the information it was given. Figure 3-4 shows the blackboard sharing concept described above.

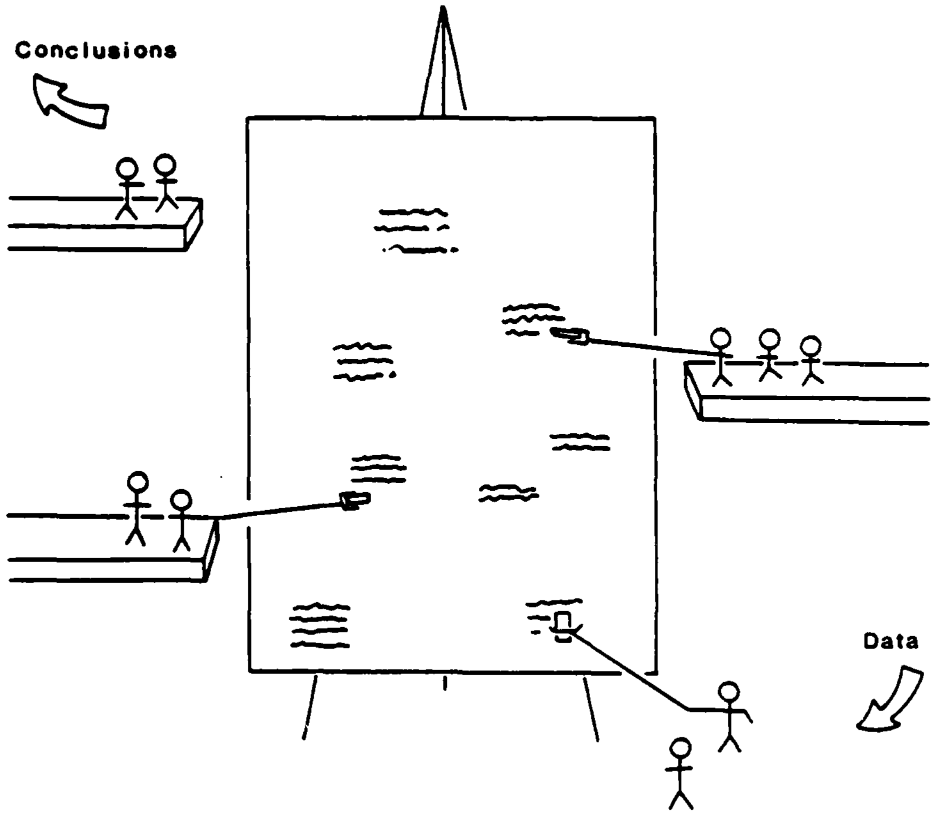


Figure 3-4 Blackboard sharing between experts. [Winston]

The upward transition and synthesis of hypotheses can be analogous to recognition. Recognition in this case is the combining of percepts, obtained from a sensory icon, and eventual reconstruction within the control environment of relationships specified in the conceptual graph [Sowa]. Recognition in this proposed vision model is described as obtaining image data from a vision device and reconstructing the image by relating precepts within the constraints of the conceptual model of the problem domain.

members common to two sets and -1 applied to set members not common to to both sets. This weight scheme provides a wider distribution. At a level arbitrarily set to some value (for example -5) all members having weights below this value are dropped. This can be described as a join, with selection implemented as a restricted range of weight. The retention level value can be adjusted to whatever value experimentation shows is practical and useful.

Beginning at the regional expert level, there is a need for a work area that supports set operations for the purpose of identity resolution. The concept of blackboards or scratch areas, as described by Winston, must be developed to support the intercommunication of experts in the hierarchy. Sharing work (scratch) areas allows sharing information about region-types, but more importantly the distribution of possible identities. The blackboard described here is more than a shared variable. The blackboard is actually partitioned into conceptual regions, "allowing the formation of interest groups of procedures that can pay special attention to the messages of their associated region." [Winston] The blackboard, as a control metaphor, preserves the levels of the hierarchy of experts and relates clearly to the original conceptual model of the human face. The candidate identities become meaningful as the collection and intersection of the candidate sets propagates upwards in the hierarchy of experts. This propagation is facilitated, in a controlled manner. by the blackboard. With each level of the hierarchy, the hypothesis of identity becomes more sound. The top level object expert will eventually offer a set of candidate possible by the feature experts. Once the regional characteristic scalar is accepted query processing continues. The collection of identity pointers resolves a partial query by producing the region owner set. The region owner set can be improved one step further, however, since there are a total of four candidate identity sets at this level of processing. These sets are actually collections of pointers from the three feature experts subordinate to the region expert and the additional set provided by the region expert after accessing the data base. The region expert resolves the query by intersecting the four sets to obtain a set of most frequently reoccuring members. This action is analogous to a join and selection sequence during query processing in a common data base. At this point, the region expert has fulfilled its role in the intermediate stage of recognition.

A summary of feature and region experts relationships to the identity sets is provided in Figure 3-3.

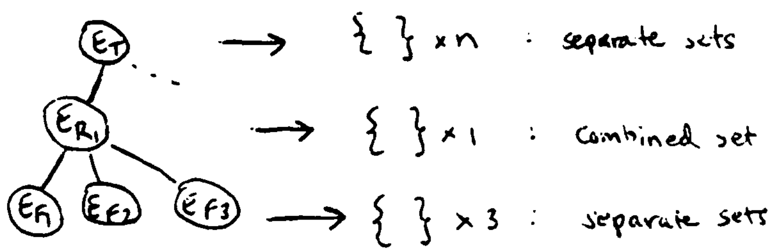


Figure 3-3 Region and feature experts' relationship to identity sets.

Set manipulations play a critical role in the region expert after retrieval of identity sets. The proposed model restricts set manipulations to simple set intersection. The simplest application of weights is a strict occurrence count of set members. Weights are simple to implement with +1 applied to set knowledge store.

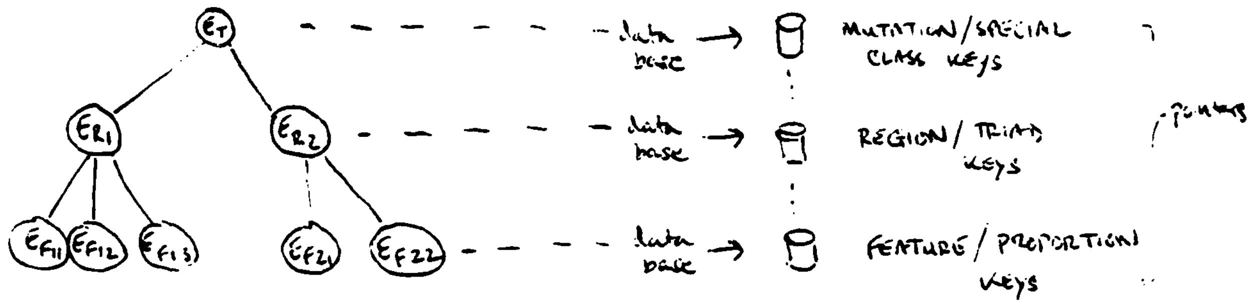


Figure 3-2 Regional and feature experts' relationship to knowledge store.

For each expert there is a matching data base partition. The expert for a feature is a discernment function with conceptual and computational components. An expert is manifested at low level discernment as a matching of expectation of feature type to the calculated value of proportion. If the hypothesis of feature type reasonably matches the resulting scalar, i.e., the range of feature values contains the scalar, then the expert verifies the hypothesis and allows a data base retrieval to occur. This interface and verification phase is made possible through the use of statistical methods.

The expert completes the feature confirmation phase and continues with the collection of candidate identity pointers thereby resolving a partial query. The collection of pointers produces the feature owner set. At this point the feature expert has fulfilled its role in the intermediate stage of recognition.

The regional expert participates in the query transformation one level above the feature experts in the hierarchy of process-controlled image analysis. It performs a region confirmation phase based upon the triad and angular scalar derivation made

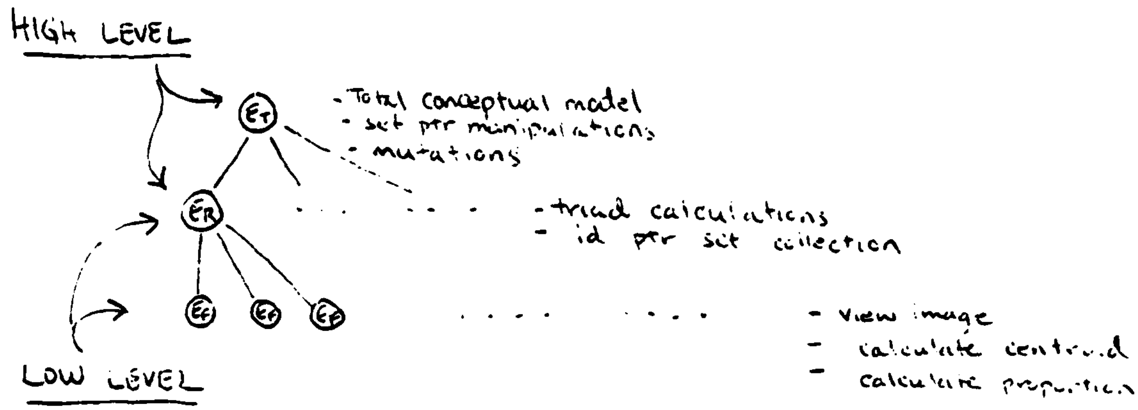


Figure 3-1 Actions of region and feature experts at the intermediate processing level.

TRANSFORMATION OF IMAGE QUERY

Recalling the definition of expert provided in Chapter 1, an expert is qualified simply as a process entity with the following attributes: 1) has accessible knowledge, 2) makes decisions upon invocation, 3) is object relative, 4) is goal oriented, 5) controls discerning functions pertinent to the identity sets, 6) only passes completed work on, and 7) has expectations.

Experts exist for each of the features in the human face and are supported by statistical methods and conceptual models. Using a proper modeling approach to support percept matching to conceptual structures is the first step towards recognition. The conceptual uniqueness of features and the ability to match expectations to what is observed allows controlled retrieval and placement functions to occur. The knowledge store becomes functionally linked to the experts and the reconstruction, or recognition of the face, is made possible. Figure 3-2 shows the relationship between the feature and regional experts and the

Chapter 3

INTERMEDIATE PROCESSING PHASE MODEL

INTRODUCTION

Query formation based on the scalar values discussed in Chapter 2 might be very useful, since the image features are transformed from collections of pixels to a range of values that further translate to a data base partition. This transformation is analogous to a data base design effort involving entities and attributes. An entity such as the left eye occurs on most human faces; however, it can have different attribute proportions relative to the individual human face. This attribute occurs as a subrange. The attribute of spatial position of features within a given region of adjacency, in the domain of the human face, also represents a subrange. This subrange occurs within the range of values comprising a partition identified with a specific triad. This concept can be summarized as a transformation from pixel composition to a data base which briefly codifies our knowledge of the human face and the owners of faces known.

The following sections provide detailed descriptions of image query processing, the interface of the conceptual model components to image processing computations, and interactions between feature and region experts. At the intermediate level, the region and feature experts provide the most active processing state of the entire vision model. For the purpose of clarity an overview of the intermediate processing level is presented in Figure 3-1.

representing the image in this fashion is that these scalars may be considered representative of the image features and yet remain quite small in terms of information storage space. This independent and possibly unique representation of shapes allows database storage and key finding issues to be handled, in support of image processing and recognition. The next chapter will discuss query processing, the relationship of keys from different experts to the data store, and the intermediate phase of image processing.

The centroid and shape-boundary have thus been consolidated into one expression P, which is a scalar value. The implication of reducing the original feature to a representative scalar is that the result is actually a candidate key. This will become important in a later discussion of data retrieval.

The next task is to consider regions of the face, and the relationship of features within that region, to each other. The role of another scalar, previously described as an angular measure, is introduced as the measure relating feature identities. The technique for calculating this scalar is described in the following steps.

The centroids xa and ya can be used to discern patterns within a more complex object by mapping the centroids of at least three patterns into a triad as illustrated in Figure 2-7.

```
A----------B          A, B and C are centroids
 \        /                  of features.
   \  *  /
      C
```

Figure 2-7 Example triad with primary angle marked.

3. Calculate the angular disposition of the centroids:

```
Given C is chosen as the  relative origin,
 then  (slope CA) - (slope CB)  gives  slope differential
Since  slope is op / adj  or cos 0
 then  inverse cos = 0 for the angular relationship of
             A and B around basis C.
```

The results of these calculations are three successive scalars representing a shape and its relationship to other shapes to form an image. These three scalars are: 1) picture plane points or centroids, 2) shape-related proportional ratios and 3) angular measures between the centroids. The benefit of

Table 2-1

	Low Value	High Value	Median	Variance
Left Eye -	0.083	0.786	0.53	0.04
Right Eye-	0.375	0.800	0.54	0.02
Mouth -	0.083	0.500	0.21	0.01

Table 2-1 Example ranges calculated from a small random sampling.

The low level discernment system computes the ratio of what is expected to be an ear or eye, etc.; the value computed is tested for membership in that partition's range and finally a retrieval is made within a defined sub-range. The result is two-part: the verification that the ratio is within a range (therefore it should be a member of a partition that matches the hypothesis) and secondly the key maps to a subrange or memberset of identities of possible owners of that feature.

The actions of low level discernment have been described previously and can be summarized as a stepwise algorithm that results in a collection of candidate keys. The previous centroid, triad, and angle definitions are applied stepwise in this algorithm.

Suppose shape S is represented by a discrete set of points
{(xi,yi), i=1,...,n}

1. The centroid is calculated as follows:

xa = 1/n * SUM xi

ya = 1/n * SUM yi

2. The proportion of the object is calculated as follows:

Dx = delta x = xlimit - xa
Dy = delta y = ylimit - ya

P = Dy / Dx

Chapter 4

HIGH LEVEL DISCERNMENT MODEL

INTRODUCTION

High level discernment is described as the use of conceptual models to provide expectations and knowledge concerning the problem domain. The conceptual model provides a concise representation of the scene in the problem domain with specific reference to objects and features and relationships between them. The use of the term 'high level' expresses the notion of knowledge and deliberateness in accomplishing interpretation leading to recognition. In this visual recognition model the top level expert is limited to expertise concerning the human face. Expertise is derived solely from the conceptual model of the problem domain, probabilistic decision criterion, and prior knowledge. The purpose of expertise is to support decision making during image interpretation. This chapter describes the development of the conceptual model from conceptual tools and establishes processing behavior of the top level of the vision model. The levels of processing and use of expertise are presented in a proposed theoretical model.

PROBLEM DOMAIN CONCEPTUAL MODEL

The use of percepts to build concept graphs has a direct correspondence to partitioning the face and building facial type-clas͜es.

The actual partitioning of the face into 'sub-regions' is possible because of well defined features and object symmetry.

Representation of the features, or regions, can then take the form of segments within a conceptual graph. These segments are the inner-workings of the 'expert for a given region'. Consequently, the regions have corresponding experts, i.e., ear expert, nose expert, chin expert, etc. It should be noted that a synthesis of experts, by partitioning or dividing the knowledge domain, allows these experts to conduct independent processing at their level. A graphical presentation of partitioning is shown in Figure 4-1.

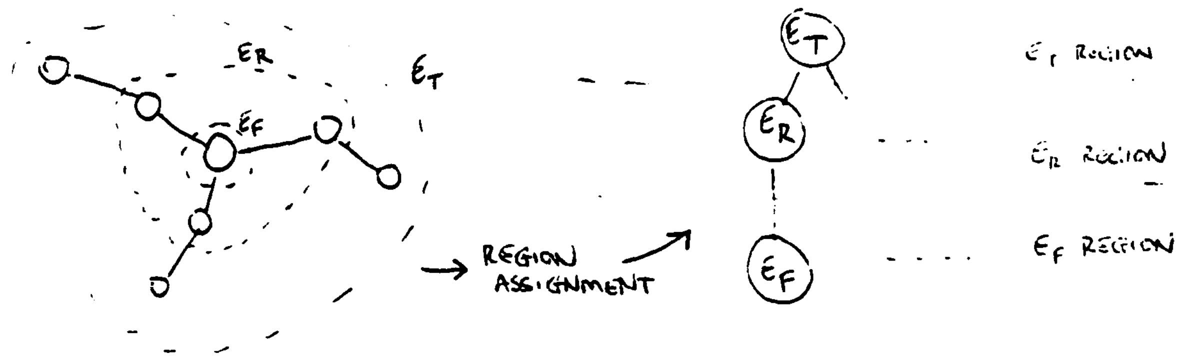


Figure 4-1 Assignment of experts to conceptual model by partitioned level of control.

As the higher, or parent level, is reached processing control is centered in the parent. If control is then localized at this level concurrency will be divided into more manageable processing areas. Adjacency is important to a structured scan of the object in question. If regions yield positive identification then the next issue is the role of 'influence' between the regions. Can the region aid the recognition activitiy of an adjacent region, or should it be discouraged, inorder to prevent errors or disruption of a potentially correct identification?

As you move up the hierarchy the need for concurrency controls becomes imperative. Sub-regions will eventually be well-described and contribute to a major-region identification. Rules for adjacency synthesis, or combination, will become necessary. Graphs for such a mechanism of control can begin at the early stage known as the 'conceptual model.' The interfacing of concepts (conceptual sub-graphs) can take the form of a hierarchy. Looking at a brief diagram of the processing strategy, the form of pipe-line intersections to 'scratch pads' captures this idea in a primitive form. [Sowa]

The connectivity graph (CG) for the face is representable in simple terms of object position. The graph is normalized by using adjacency rules for component objects. See Figure 4-2.

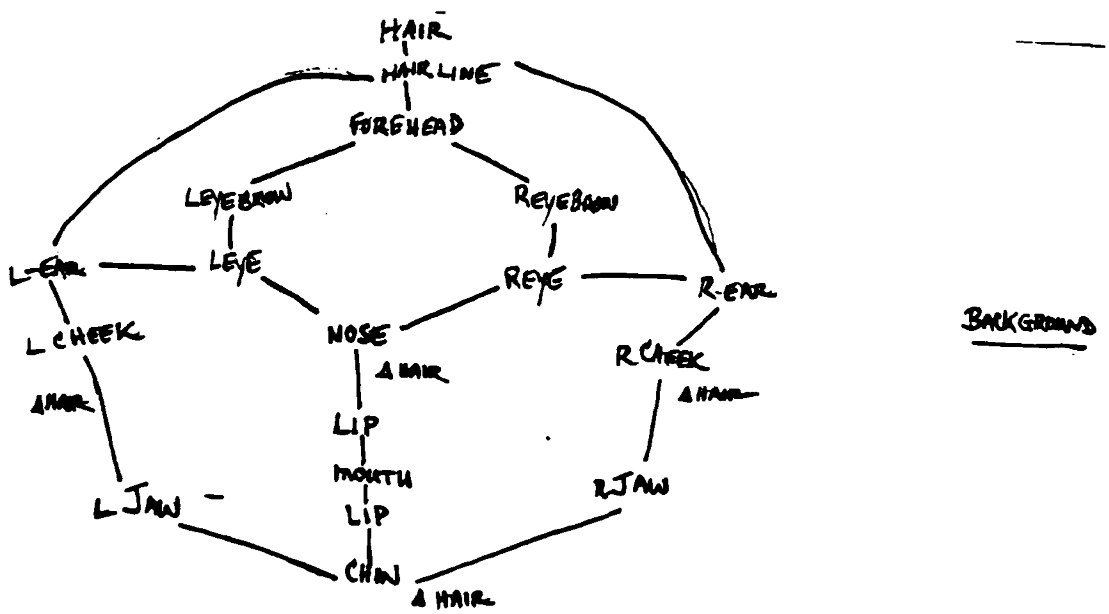


Figure 4-2 Connectivity graph of facial features.

Although the CG is abstract and flexible there must be specific values or realizations of the elements within the graph to support recognition. The knowledge environment will play a very important role in this realization. Knowledge of specific values will be incorporated into the CG in a manner described in the remainder of this chapter.

This discussion points towards a hierarchical image representation that is used by a knowledge-driven system. The task of recognition at a high level can be conceptualized in terms of such a hierarchy. Recognition is accomplished when the model (image hierarchy) is reasonably well matched to the image domain.

The image domain for this thesis is the human face. The scene model is an image hierarchy composed of background and the critical features of the human face. The objectives of the vision system are to initially identify facial features and then associate these features with a candidate owner-identity using expert processes working together, concurrently.

By combining the hierarchical tree structure and the relational graph structure a picture can be described and summarized. As classified by Fu, this hierarchical graph is actually a 'web'. More specifically, a web is a derivation diagram for a context-free web grammar [Fu]. Describing a picture is similar to forming the derivation diagram of a web grammar. Fu emphasizes that the set of underlying grammar rules for the construction of the derivation diagram are called the syntax of the described picture.

The similarity of Fu's, Minsky's, and Sowa's modeling tools provides strong support for a hierarchical graph representation of the human face. The goal of such a representation is to thoroughly depict the concept of a human face. Upon examining the human face we see that it possesses region attributes and relational properties such as adjacency, composition, and proportion. Fu might call these properties the natural grammar of facial net, see Figure 4-3. Sowa's conceptual graphs would capture the relational properties as position information of features in the image space, see Figure 4-4. With full consideration of the strengths of expression offered by Fu and Sowa, a human face model can be sufficiently represented.

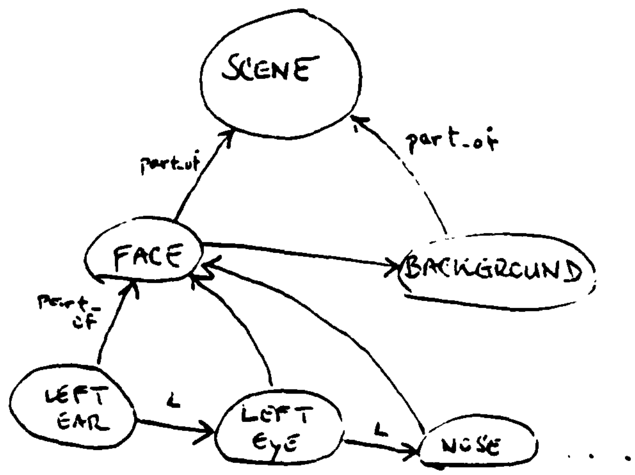


Figure 4-3 Fu web grammar showing adjacency of human facial features.

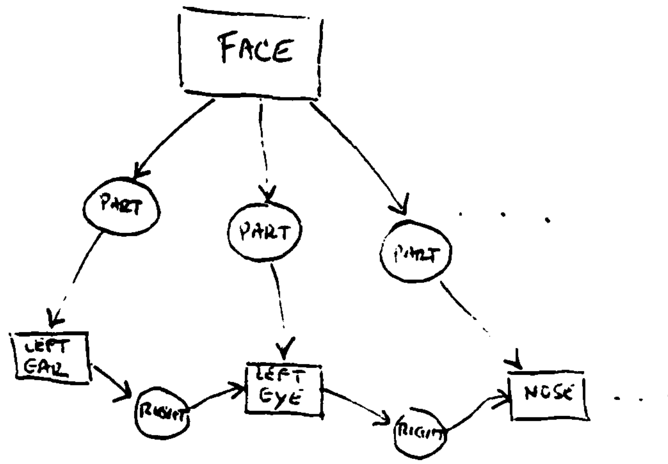


Figure 4-4 Sowa conceptual graph of the human face.

Frames, brought to prominence by Minsky, are often presented as components to semantic nets or conceptual graphs. A frame hierarchy then, is used to model information in terms of the conceptual hierarchy. Within the hierarchy of the human face model it will be important to consider both regions and objects. Following the recommendations of L. Tucker, the proposed face model will be composed of both region and object frames. The frame hierarchy will capture information pertaining to topological relationships between regions and objects. The remaining criteria for completing the hierarchy graph are: segmentation and adjacency. These criteria complete the a priori knowledge base of the vision system. A detailed frame hierarchy can accurately describe a picture or pictured-concept, see Figure 4-5.

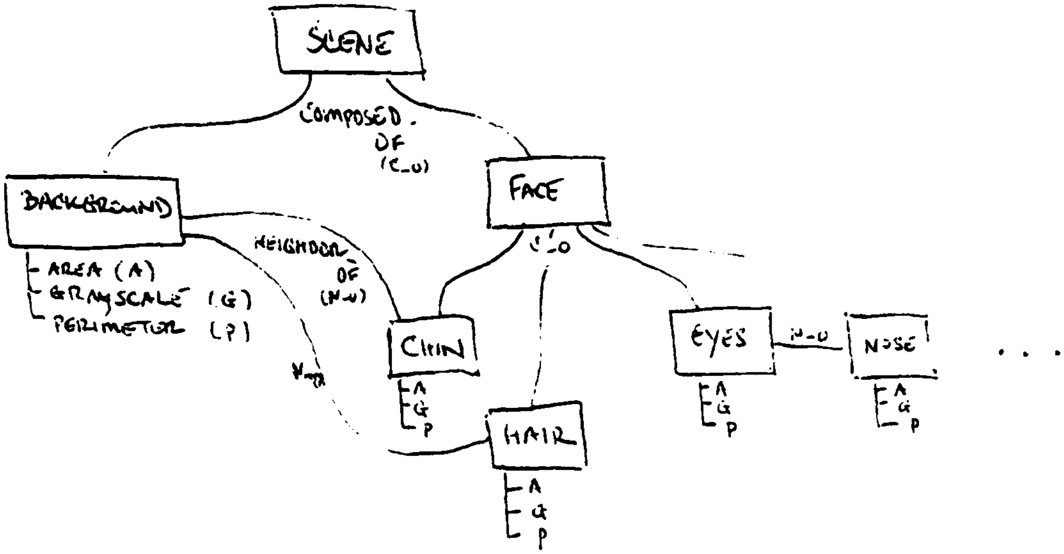


Figure 4-5 Hierarchy of frames representation of the face.

This graph will support the broadcasting of information between regions and permits a relaxation-labeling process to influence the final scene interpretation.

The low level vision model must be integrated with this high level model for experts to interact up the hierarchy during a reconstruction of the abstract image. These two levels must grow towards each other in a synthesis caused by partition association. If the association between concept and computation is to occur, the partitions must be placed and fitted at an abstract intermediate level.

In the concurrent vision model proposed, the scratch areas support the experts as they traverse the CG forming the necessary abstractions. The intermediate formations between the experts are subabstractions of the whole concept.

TOP DOWN VISUAL DISCERNMENT MODEL

This high level model is patterned closely after the technique of Lewis W. Tucker. Tucker's system recognizes biological cells in simple image domains [Tucker]. The following adaptation of his system is provided to illustrate the mechanisms necessary to support an expert vision system capable of recognizing a human face. This model serves as our suggestion for the processing behavior of such an expert vision system.

Purpose

Successive approximation is a viable approach to problem solving. The approximation can be accomplished based on conceptual model-driven image segmentation. At the implementation level this technique is manifested as quadtree segmentation. These segments are used by goal-oriented expert processes to form successive scene and image interpretations. The interpretations are formed from hypotheses that are tested and verified. It becomes obvious that expert processes should accomplish these tasks concurrently. Concurrence and system parallelism, inherent to the independent experts, requires special control strategies. A conceptual graph will provide a control structure for the integrated expert vision system. Successive approximation can be summarized as a cycle in which error terms are calculated and a selection of succeeding options occurs. Finally, the cycle is terminated when the error term is reasonably small.

For the specific purpose of image recognition, successive approximation can be used as the vision expert studies the image-scene at the center of interest. Resolution of the image begins with an initial assumption about the object label. Object label and interpretation are used inter-changeably. This is the first hypothesis made by an expert. As the model-to-image mapping develops, any dissimilarity causes a new group of hypotheses to be formed and verified. A conceptual model predicts image characteristics while the sight mechanism returns what is actually seen. This cycle of hypothesis testing and verification is the foundation of the 'causal' mechanism in the vision expert.

Semantic labels can be applied to image regions within a scene. These labels are obtained from a set of labels presented in the form of a conceptual graph. The relationship between the graph and scene is formed by partitioning the object-center of the scene into recognizable regions. In this way meaningful regions, or scene components, provide a structured approach to the task of recognition.

If the search space of the object domain is reasonably well known, then segmentation into small regions is possible. The steering mechanism for such segmentation is a priori knowledge. The model we have given the expert is this knowledge. This knowledge provides the root of 'expectation' in the expert vision system.

There are three fundamental reasons for choosing an expert vision system which embodies the general capabilities previously described:

1) "The total computational effort required to solve a particular scene analysis problem should be decreased by the integration of high- with low-level processing." [Tucker]

2) "Model knowledge permits the selective application of computationally-expensive operators only when needed. The uniform application of image-filtering operators is avoided because planning is employed for a more judicious allocation of power." [Tucker]

3) "An expert vision system could be designed to take on a more active role -- capable of executing efficient search algorithms for specific objects -- by probing the environment for detail, rather than the traditional passive role of evaluation and interpretation." [Tucker]

Region and Object Experts

The design of expert processes must begin with a review of their functional requirements. The order of the following tasks implies the level at which hypotheses will be formed and a possible chronology for problem solution.

1) Estimate the probability that a given scene contains a FACE.

2) Begin hypothesis formation by computing the most likely label (region or object identifier) for a given quadtree tile.

3) Generate the region adjacency list for a given region.

4) Search for a facial feature at an expected location.

5) Refine borders between regions, or as appropriate, clarify boundaries of objects.

6) Merge poorly defined regions with BACKGROUND or NEUTRAL regions, until well defined borders result.

7) Evaluate how well a particular scene interpretation matches the prediction by the model.

8) Collect candidate identities of owners.

9) Find the connected components of a given region.

10) Compare and retain most probable candidate identities of owners.

11) Split quadtree tiles having a high edge feature component.

The strategy of partitioning the work steps described above will provide for better concurrent processing and control. As suggested previously, the experts can be formed to align with the elements of the human face model. This approach allows each expert to be designed with respect to properties of one particular frame. The complexity of image objects and regions is then forced to higher levels in the hierarchy. Sharing information concerning owner identities should be conducted at the higher levels as well. The total number of experts at the bottom of the tree (opposite of root) is currently three, expandable to eleven (see Figure 1-6). It should be noted that knowledge and the ability to hypothesize are now localized, thereby concentrating the expertise of the vision system at nodes within the frame hierarchy.

Model Matching and Hypothesis Testing

Given a particular segmentation, experts must evaluate how well the existing regions fit the predictions supplied by the model in order to improve upon the current interpretation. Using model-matching techniques and relaxation labeling, each expert is able to assign a confidence value for the given label based upon its intrinsic characteristics (size, shading, shape, centroid, and proportion), its structural composition and its relationship to its immediate neighborhood. Procedural information, given as a series of pattern invoked rules, trigger actions for relabeling the existing regions or the search for a better segmentation.

The actions just described reflect a deliberate integration of low and high level vision mechanisms. These mechanisms interact in a controlled manner relative to the structure of the conceptual model. Since the system is goal oriented, the expectations at the high level, must be supported by the sub-goals of the low level. These sub-goals are provided in the form of computable characteristics that support the uniqueness and correctness of the expert's hypotheses. The division of steering actions and direct-sight functions provides an important independence of function in this vision system model. The reasons for insuring this independence will become clear in later discussions about job control and system management.

Job Control and the System Manager

Concurrent processing is an objective of this vision system; it is a problem to decide the degree to which expert processes are allowed to execute independently. Establishing a system manager that enforces processing policies is a potential solution [Hwang]. In terms of the blackboard concept discussed in chapter 3 the process scheduler (inherent to the blackboard) might satisfy the role of system manager.

The manager should ensure that experts execute only after stating their processing requests. This 'check-in' scheme allows a full-span control of the system-program state. Such control might be realized by using "request-centered control" defined by Winston as a situation when a system's procedures know their own purpose, before responding to system requests. Strategies for segmentation and information sharing (or job mix) can occur only as the manager permits by the requests that are submitted. Each level of the expert hierarchy can make requests of subordinate experts. In many ways the manager synchronizes the dynamic character of the face model-to-image mapping relative to the frame hierarchy by providing goal-directed requests. In its simplest role the manager prioritizes process categories by recognizing the degree of expertise at each level. Although a complete solution to management is not presented here the goal of the manager is still simply to coordinate the experts, thus preventing conflicts.

The manager serves another purpose as well, that being the pivot between the low and high level activities of the vision model. As activities concentrate at the object frame level, picture related computation becomes intense, meaning that the system requests are detailed and restricted to the appropriate interest groups in the blackboard. Alternately, as the need for hypothesis verification and region coordination occurs intensity in inter-communications increases at higher levels of the hierarchy. The actions of the intermediate and high level interest groups become more critical in providing resultant conclusions. The processing of experts, or interest groups, on these different levels seems to be somewhat independent, but remains controlled. The manager is aware of the global process (and goals) and can make decisions about the proper course of computation and search in the image space. This entire plan of action can be summarized as a goal-directed process-packaging technique supported by a blackboard system. An initial solution to the sequencing control strategy might be a policy of assigning a queue to each expert, thereby establishing 'pipes' for concurrent-like processing using interleaving. Although sequencing control might be useful in sharing centroid and angle calculation resources, this partial solution is not an attempt to describe the total behavior of the scheduler in the blackboard system. A diagram of this concurrent-like control is presented in Figure 4-6.

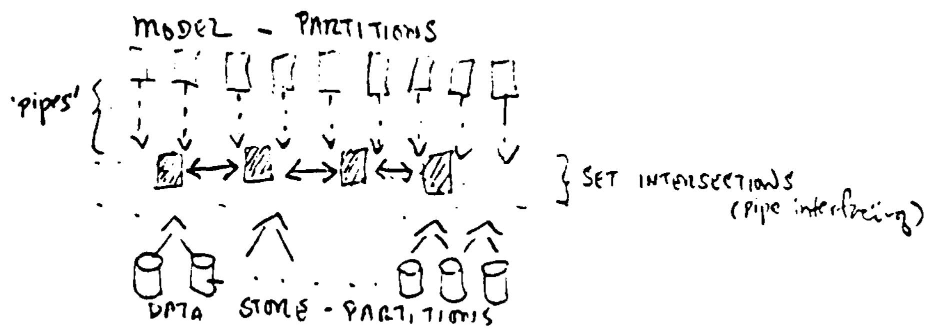


Figure 4-6 Pipeline diagram supported by queues.

concentrate on those processes finishing partition analysis with the intention of quickly providing a total or near total reconstruction. A rapid 'near total' reconstruction could provide a best-guess under time restricted operating conditions.

Intersecting positive identification sets provides an estimate of identity faster than trying to serially identify from a long list of images contained in the database. Knowledge can be represented as a set of rules directing such an optimal intersection. The real question is one of trying to incorporate an 'adaptive' characteristic into the vision system during run-time.

Mutations and major alterations should be considered to help determine whether the image is disguised or has naturally mutated. Such analysis also highlights uniqueness in an image and provides an early image partition access or whole image 'handle'. This situation might be considered an orientation tool for structure oriented communication between identification processes in the proximity of a mutation.

dictionary (DD) might incorporate conceptual models into the data store access mechanism more effectively. The dynamic characteristic of the DD is seen as a traversal or mapping to a group 'type' key. The keys can be retained in a knowledge or rule-based section. This strategy supports an a priori knowledge of the viewer image. The goal is to perform an image query in a concurrent data dictionary environment in support of more than one possible sight mechanism.

A mapping from transform information (perceived image) to the entry keys can be depicted in a number of ways. The most critical issue is the assumed support of concurrent or parallel architecture-strategies at the design level. The model itself, at this level, is actually a cause-and-effect chain with some branching and adjacent linking.

The concept of visual search or saccade can be dynamically adapted to the core of this proposed vision system model. The rapid movement of the eye from one point to another is called saccade. During the search process the eye fixes on a point and then jumps rapidly to another point. The use of model driven quadtrees, as discussed in Chapter 3, is analogous to saccade as it occurs in human vision. The concurrent expert processes will exist at flow states that differ, as each attempts to recognize the assigned partition of the facial image. The perceptual level of the model provides quantified information pertaining to partitioned attributes of the facial image. There are a number of expert processes at the perceptual level. These expert processes can ask each other if they have finally made an effective identification, or a supervisor expert process can system is badly diminished.

Tools to Consider:

Facial Schemata for each individual showing unique combination of key attributes.
Providing collection (select and join) scratch areas for best fit or look-ahead during evaluation of retrievals.

Partitioning database into attribute groups with extensive characteristic information contained in the partitions.

Early screening of mutation or racial considerations.

Abort ability due to diverse or unusual centroid results.

Tagging of entity-attribute associations to individuals, found to be redundant in the scratch areas, will facilitate building the 'focus' set of probable identification. By intersecting data sets and then discarding mutually exclusive attribute associations the 'recall' process speed can be increased. Increasing the speed of the recall process means identifying best possible matches and using the reduced set in a structured (or deliberate) manner to check for adjacent feature identification in a concurrent environment.

The final reduced set will provide a better starting point for search and detailed match between observed object and the stored abstract of the object and a faster mathematical tool can be considered.

Transforms differ in clarity and accuracy. They can be ranked in terms of their benefit to perception and recognition. An analogy that has been adopted in this thesis research is one of center vision and peripheral vision. Objects might be detectable as they move into view, without initial identification. Experimenting with the peripheral to center vision transition of an object as it actually moves into 'clear view' shows that an object is not initially discernable to the human. Some of the transforms considered offer the peripheral 'warning' or alert of an approaching object. Other transforms are fast and detailed offering center vision capabilites. An investigation of transform usage policies might prove useful.

Techniques to Consider:

Group keys will need to be expressed in a brief or reduced form for computational requirements. Establishing a data

Object representation depends on image clarity which is made possible by well defined object boundaries. It has been assumed in this thesis that the image features are bounded, i.e. an edge has been detected that encloses the domain of the feature. If the object is not totally enclosed, or well defined, the image can be averaged or checked for ratios of location based on neighboring features. An alternate technique for extracting the geometry of a feature should be considered in providing the proposed vision model additional strength and flexibility.

CONCLUSIONS

The modeling approach to percept matching and conceptualizing used in this thesis is based on relative geometries known in simple terms as proportionality. This proportionality concept is the foundation for uniqueness which allows retrieval and placement. The knowledge store is functionally linked to the experts through the use of retrieval and placement. Reconstruction of a face's identity from stored segments is a deliberate process involving evaluation and decision-making thereby making this model a expert visual recognition mechanism.

RECOMMENDATIONS FOR FURTHER STUDY

The decision mechanism for the experts should include an effective method for cost, merit, or weight labeling. Some decision supportive heuristic for weight assessment should be incorporated into the intersection and comparison of candidate sets. If sets are not reasonably reduced at the expert level, the decision capability (or discerning power) of the vision the purpose of developing thesis concepts. If images are canted or turned, but still preserve a frontal view, then simple adjustments in orientation and calculation could be made with some modifications to the vision model. If a facial image is presented in a diagonal view the proposed model will fail. The conceptual model of the face remains useful in situations of varied facial orientations. The primary limitation of this vision model is its sole dependence on proportion and triad angles established for a frontal view. There is potential for extending the proportion technique by using rotational transforms that adjust image geometry.

If the uniqueness of an object can be captured in an independent expression then the object has been represented without the need for massive reconstruction. The concept of object representation centers around centroid and proportion techniques. Ideally, because of the relative uniqueness of values created in these two ways, objects and features can be individually identified. This technique relies on the inherent differences in measures of features being examined. Noise sensitivity and the issues of variables, such as human hair, can render some calculations useless. The goal of object representation is to produce a unique numerical value, or key, for features that allows hashing into the data store or knowledge base. This hashing action will only occur when a key has been obtained. Until the key is obtained the expert system remains incapable of building identity sets at a given level which results in partial or total failure in processing an image query.

thereby supporting detection and recognition.

An expert vision system is proposed which infers, from the collection and evaluation of facial segments of a face, a set of possible identities. Retrievals from a data store to accomplish this are facilitated by using the shape representations as keys to the records. The heuristic involves collecting the candidate identities into sets and then reducing the number of identities in those sets by set intersection and a weighting scheme. Sensitivity and tolerance factors in this expert are adjustable in terms of range values. The value obtained during evaluation of a facial segment is accepted or tolerated relative to the range of values expected for that segment. A retrieval effects a subrange whose width is determined by sensitivity to be the center and some number of adjacent values above and below that center value. Expectation ranges can be tailored to incorporate system and environmental variables.

CHALLENGES TO THE METHOD

There are some problems experienced in trying to represent a facial feature in the most 'unique' manner possible. Representation must be accomplished without relying on exhaustive quantities of information about the image. We want to avoid exhaustive calculations that perform various transforms and slowly provide a measure of the object. Some of these issues are discussed in this section.

This vision model attempts to solve a small portion of the complex problem of automated human face recognition. The human face is viewed from a direct frontal view orientation for

Chapter 6

RESULTS AND CONCLUSIONS

This thesis considers the human face as a planar shape consisting of planar shapes, and presents a new approach to face recognition, from a frontal view, based on the observation that matching of a new face with a database of known faces is a much easier task when the intrafacial shapes are represented in terms of centroids and angles. Representation is defined as finding scalar values that correspond to facial features. These scalar values are potentially useful in the design, storage and retrieval of information in data bases. Additionally, representation offers the advantage of applying shape matching algorithms which use the scalars.

It is shown that classification of planar figures can be achieved using shape-relative ratios computed from given figures. Representation of groups of planar figures is also presented as a technique for shape matching. The groups suggested are defined as triads, i.e., composed of three centroids. A scalar value for the angle between the centroids is computed, which represents the proportional position of facial features. Example calculations of the centroid, proportion ratios and relational angle are provided.

In image segmentation, shape is often the basis for the detection and recognition of patterns such as lines, edges, corners and more general images such as facial features. Quantifying shape as the proportion of objects, is an important aspect in facial image processing systems. The scalar values of facial features define ranges of values for the image segments, Initial formation of the triad and angle calculations from the Mouth (angle between Left and Right Eye) yield:

44 degrees, angle measured manually from xerox copy

Initial Query of Triad (angle) yields:

44 -> Partition Range 45 to 53 -> Value not in Range
Value matches NULL but is close to Persons #1,7,8,
Identity Set = {}

Regional expert set manipulations to reduce candidate set yield:

Identity Set = {4,7,9}
Identity Set = {}
Identity Set = {1,2,6}
Identity Set = {}

-> 1 Occurences of #9
1 Occurences of #7 50% NULL Identity Set
1 Occurences of #6
1 Occurences of #4
1 Occurences of #2
1 Occurences of #1

-> Best Estimation: #1,2,4,6,7,9
2nd Best Choices: NULL (new person) (* Learn This Face *)

-> Reduced Region Candidate Set = {1(1),2(1),4(1),6(1),7(1),9(3)} with counters
or
{} reduced by set intersection

The third query example uses a photograph of a person not known by the system, i.e., a person not in the knowledge base (see Figure 5-2).

Figure 5-2 Person unknown to knowledge base.

The initial calculation of centroids yields:

Left Eye:	a point to the upper side of the pupil
Right Eye:	a point at left upper edge of the pupil
Mouth:	a point centered along the length of the mouth and almost on the horizontal

The initial calculation of proportions yields:

Left Eye:	0.2 / 0.5 = 0.400	(cm)
Right Eye:	0.15/ 0.5 = 0.300	
Mouth:	0.05/ 0.6 = 0.083	

Initial Query of Left Eye yields:

0.400 -> Partition Range 0.083 to 0.786 -> Value is in Range
Value matches NULL but is close to Persons #4,7,9,
Identity Set = {4,7,9}

Initial Query of Right Eye yields:

0.300 -> Partition Range 0.375 to 0.800 -> Value not in Range
Value matches NULL but is close to Person #4
Identity Set = {}

Initial Query of Mouth yields:

0.083 -> Partition Range 0.083 to 0.500 -> Value is in Range
Value matches Persons #1,6 and is close to Person #2
Identity Set = {1,2,6}

Initial Query of Left Eye yields:

0.500 -> Partition Range 0.083 to 0.786 -> Value is in Range
Value matches Persons #4,7 and is close to Persons #6,8
Identity Set = {4,6,7,8}

Initial Query of Right Eye yields:

0.667 -> Partition Range 0.375 to 0.800 -> Value is in Range
Value matches NULL and is close to Persons #1,3,6,10
Identity Set = {1,3,6,10}

Initial Query of Mouth yields:

0.300 -> Partition Range 0.083 to 0.500 -> Value is in Range
Value matches NULL and is close to Persons #3,4,5,7,10
Identity Set = {3,4,5,7,10}

Initial formation of the triad and angle calculations from the Mouth (angle between Left and Right Eye) yield:

50 degrees, angle measured manually from xerox copy
Initial Query of Triad (angle) yields:

50 -> Partition Range 45 to 53 -> Value is in Range
Value matches Person #9 and is close to Persons #3,4,6,10
Identity Set = {3,4,6,9,10}

Regional expert set manipulations to reduce candidate set yield:

Identity Set = {4,6,7,8}
Identity Set = {1,3,6,10}
Identity Set = {3,4,5,7,10}
Identity Set = {3,4,6,9,10}

-> 3 Occurences of #10
3 Occurences of #6
3 Occurences of #4
3 Occurences of #3
2 Occurences of #7
1 Occurences of #9
1 Occurences of #8
1 Occurences of #5
1 Occurences of #1

-> Best Estimation: #10,6,4,3
2nd Best Choice: #7

-> Reduce' Regi n Candidate Set = {3(3),4(3),6(3),10(3)} with counters
or {} reduced by set intersection

->   Best Estimation: #10
      2nd Best Choices: #9 and #3

-> Reduced Region Candidate Set = {3(3),9(3),10(4)} with counters or {10} reduced by set intersection

The second query example uses Person #10's photograph that was not previously used in constructing the knowledge base (see Figure 5-1).

Figure 5-1 Person #10's alternate photograph.

The initial calculation of centroids yields:

Left Eye: a point centered on the pupil

Right Eye: a point near center of the pupil

Mouth: a point centered along the length of the mouth and slightly above the horizontal

The initial calculation of proportions yields:

Left Eye: 0.05/0.10 = 0.500 (cm)

Right Eye: 0.05/0.075 = 0.667

Mouth: 0.075/0.25 = 0.300

The initial calculation of proportions yields:

Left Eye: 0.2 / 0.7 = 0.286 (cm)

Right Eye: 0.4 / 0.7 = 0.571

Mouth: 0.4 / 1.9 = 0.211

Initial Query of Left Eye yields:

0.286 -> Partition Range 0.083 to 0.786 -> Value is in Range
Value matches Person #10 and is close to Person #9,
Identity Set = {9,10}

Initial Query of Right Eye yields:

0.571 -> Partition Range 0.375 to 0.800 -> Value is in Range
Value matches Person #10 and is close to Persons #1,3,7,9,
Identity Set = {1,3,7,9,10}

Initial Query of Mouth yields:

0.211 -> Partition Range 0.083 to 0.500 -> Value is in Range
Value matches Person #10 and is close to Persons #3,5,7,
Identity Set = {3,5,7,10}

Initial formation of the triad and angle calculations from the Mouth (angle between Left and Right Eye) yield:

49 degrees, angle measured manually from xerox copy

Initial Query of Triad (angle) yields:

49 -> Partition Range 45 to 53 -> Value is in Range
Value matches Persons #3,10 and is close to Persons #5,9,
Identity Set = {3,5,9,10}

Regional expert set manipulations to reduce candidate set yield:

Identity Set = {9,10}
Identity Set = {1,3,7,9,10}
Identity Set = {3,5,7,10}
Identity Set = {3,5,9,10}

-> 4 Occurences of #10
3 Occurences of #9
3 Occurences of #3
2 Occurences of #7
2 Occurences of #5
1 Occurences of #1

presented at Appendix C. The calculation of centroids was simulated by measuring the approximated centers of features. The angles of the triads formed above were measured manually by protractor. The results of preliminary measurements and calculations is shown in Table 5-1 which also represents the knowledge base.

Table 5-1 (cm)	Proportional Ratios Left Eye	Right Eye	Mouth	Triad Angles
Person#1	.4/.7=.571	.3/.6 =.500	.1/1.2=.083	46 deg
Person#2	.5/.8=.625	.4/.9 =.444	.2/1.8=.111	53
Person#3	.55/.7=.786	.65/1.1=.591	.4/1.9=.211	49
Person#4	.3/.6=.500	.3/.8=.375	.3/1.1=.273	51
Person#5	.5/.8=.625	.8/1.0=.800	.3/1.3=.231	48
Person#6	.5/.9=.556	.7/.9 =.778	.1/1.2=.083	51
Person#7	.3/.6=.500	.4/.8 =.500	.4/1.4=.286	46
Person#8	.4/.7=.571	.4/1.0=.400	.25/1.4=.179	45
Person#9	.3/.9=.333	.4/.8 =.500	.2/1.3=.154	50
Person#10	.2/.7=.286	.4/.7 =.571	.4/1.9=.211	49

Table 5-1 Knowledge base of persons' characteristics.

MODEL EXECUTION

Preprocessing was accomplished by using a xerox copier set to 'very light' in an effort to simulate thresholding. The images obtained are included at Appendix A with the markings of centroids, proportional axes, and triads.

The first query example uses Person #10's photograph, as a test, that was previously used in constructing the knowledge base.

The initial calculation of centroids yields:

Left Eye: a point to the upper left side of the pupil

Right Eye: a point on the right edge of the pupil

Mouth: a point centered along the length of the mouth and slightly above the horizontal

Chapter 5

SIMULATED USE OF MODEL

INTRODUCTION

It is possible to provide a clear impression of capabilities of the model by presenting a sample execution. The model will be demonstrated through the use of an example photograph and the mathematical tools developed in the previous chapters. The demonstration of the vision model will proceed stepwise from the preprocessing phase to the resolution of query. The algorithms discussed in previous chapters serve as outlines for the processing steps in this simulation. It should be understood that this demonstration is a primitive simulation of the model.

The statistical information base used in these calculations is presented at Appendix B. The statistics were processed by a Pascal language program written to support this simulation.

The query photographs were chosen using the following criteria: 1) a known photograph used to construct the knowledge base, 2) an additional photograph (not used to construct the knowledge base) of a person known to the system, and 3) a photograph of a person not known to the system.

The knowledge base was constructed from a random sampling with a population total of 10. The photographs used were 8" X 10" black and white. The photographs were processed to threshold levels leaving features as the predominant image fragments. This processing of photographs was accomplished through the use of a xerox copier. An example of the xerox threshold simulation is

REFERENCES

REFERENCES

Cohen, Paul R. and Feigenbaum, Edward A., "The Handbook of Artificial Intelligence", Volume 3, HeurisTech Press, Stanford, California, 1982.

Fu, K.S., "Picture Syntax" in "Pictorial Information Systems" from Lecture Notes in Computer Science, Vol. 80, Springer-Verlag, 1979.

Goodman, A.W. "Modern Calculus with Analytic Geometry", MacMillan Company, New York, 1970.

Grosky, W.I., and Jain, R. "Optimal Quadtrees for Image Segments", presented in "Pattern Analysis and Machine Intelligence", IEEE Transactions, Vol. PAMI-5, No. 1, 1983, pp. 77-83.

Hayes-Roth, F., Waterman, D.A., and Lenat, D.B. "Building Expert Systems", Addison-Wesley Publishing Company, Inc., Massachusetts, 1983.

Hogg, D., "Model-based vision: a program to see a walking person", from "Image and Vision Computing", Vol. 1, No. 1, February 1983, pp. 5-20.

Hwang, K., and Briggs, F.A. "Computer Architecture and Parallel Processing", McGraw-Hill Book Company, New York, 1984.

Lecky, J.E., "Forth for Computer Vision in Industrial Applications", from "Robotics Age", April 1985, pp. 25-28.

McCurdy, Edward, "Leonardo da Vinci's Notebooks", Empire State Book Company, New York, 1923.

Mitiche, Amar and Aggarwal, J. K., "Contour Registration by Shape-Specific Points for Shape Matching", from Computer Vision, Graphics, and Image Processing, (Vol.22.No.3, June 1983) 396-408.

Morrill. W.K., "Analytic Geometry", International Textbook Company, Pennsylvania, 1964.

Omolayole, J.O. and Klinger, A. "A Hierarchical Data Structure for Storing Pictures" in "Pictorial Information Systems" from Lecture Notes in Computer Science, Vol. 80, Springer-Verlag, 1979.

Rich, E., "Artificial Intelligence", McGraw-Hill Book Company, New York, 1983.

Rhodes, M.L. and Klinger, A. "Conversational Text Input for Modifying Graphics Facial Images" from "Fuzzy Reasoning and its Applications", Academic Press, 1981.

Rosenfeld, Azriel and Kak, Avinash C., "Digital Picture Processing", Academic Press, Inc., New York, 1982.

Rosenfeld, Azriel, "Recent Trends In Image Analysis", "Cybernetic Systems: Recognition, Learning, Self-Organisation", Research Studies Press LTD., England, 1984. pp. 87-93.

Sayre, Kenneth M., "Recognition: A Study in the Philosophy of Artificial Intelligence", Univ. Notre Dame Press, 1965.

Schowengerdt, R.A., "Techniques for Image Processing and Classification in Remote Sensing", Academic Press, Inc., New York, 1983.

Sequeira, E.G., "Structured vs. Unstructured Scan Path in Static Visual Search Performance", Master's Thesis, KSU, 1979.

Sandler, H., "Left Ventricular Volume, Mass and Related Measures; Usefulness in Determination of Ventricular Function" from "Quantitation in Cardiology", Leiden University Press, 1972.

Shu-Xiang, L. and Loew, M.H. "Boundary Chain Codes from Quadcodetrees", presented in Proceedings of the Workshop on Computer Vision: Representation and Control, IEEE, Maryland, 1984, pp. 214-218.

Sowa, John F., "Conceptual Structures", Addison-Wesley, Reading, Massachusetts, 1984.

Tucker, L.W., "Control Strategy for an Expert Vision System using Quadtree Refinement", presented in Proceedings of the Workshop on Computer Vision: Representation and Control, IEEE, Maryland, 1984, pp. 214-218.

Walpole, R.E., and Myers, R.H. "Probability and Statistics for Engineers and Scientists", MacMillan Company, New York, 1972.

Winston, P.H. "Artificial Intelligence", Addison-Wesley Publishing Company, Massachusetts, 1984.

SELECTED BIBLIOGRAPHY

SELECTED BIBLIOGRAPHY

Aho, A.V., Hopcroft, J.E. and Ullman, J.D., "The Design and Analysis of Computer Algorithms", Addison-Wesley Publishing Co., Massachusetts, 1974.

Artwick, Bruce A., "Applied Concepts in Microcomputer Graphics", Prentice-Hall, Inc., New Jersey, 1984.

Blaser, A., "Data Base Techniques for Pictorial Applications" from Lecture Notes in Computer Science, Vol. 81, Springer-Verlag, 1979.

Borghesi, P., Cappellini, V. and Del Bimbo, A., "A Digital Processing Technique for Recognition of Moving Objects", "Cybernetic Systems: Recognition, Learning, Self-Organisation", Research Studies Press LTD., England, 1984, pp. 111-120.

Brigell, M., "Simultaneous and Sequential Illusions of Extent: Tests of a Neural Model of Perceptual Information Integration", Doctoral Dissertation, KSU, 1977.

Caianiello, E.R. and Musso, G. (editors), "Cybernetic Systems: Recognition, Learning, Self-Organisation", assembled papers from the International Workshop on Cybernetic Systems at Salerno in December 1981, Research Studies Press LTD., Letchworth, Hertfordshire, England 1984.

Campbell, Hugh G., "Linear Algebra with Applications", Appleton-Century-Crofts, New York, 1971.

Cantoni, V. and Musso, G., "Shape Recognition Using the Hough Transform", "Cybernetic Systems: Recognition, Learning, Self-Organisation", Research Studies Press LTD., England, 1984, pp. 121-128.

Cohen, Donald N., "Knowledge Based Theorem Proving and Learning", UMI Research Press, Computer Science: Artificial Intelligence, Ann Arbor, Michigan, 1981.

Cohen, Paul R. and Feigenbaum, Edward A., "The Handbook of Artificial Intelligence", Volume 3, HeurisTech Press, Stanford, California, 1982.

Conrad, Michael, "Robot Vision Research", Wayne State Univ., Detroit, Michigan, 1984 (Professor of Computer Science).

Di Zenzo, Silvano, "A Note on the Estimation of A-Priori Probabilities in Pattern Recognition", "Cybernetic Systems: Recognition, Learning, Self-Organisation", Research Studies Press LTD., England, 1984, pp. 95-102.

Dorf, Richard C., "Modern Control Systems", Addison-Wesley Publishing Company, Massachusetts, 1974.

Fu, K.S., "Picture Syntax" in "Pictorial Information Systems" from Lecture Notes in Computer Science, Vol. 80, Springer-Verlag, 1979.

Garibotto, G. and Tosini, R., "3-D Pattern Recognition Techniques", "Cybernetic Systems: Recognition, Learning, Self-Organisation", Research Studies Press LTD., England, 1984, pp. 103-109.

Gass, Saul I., "Linear Programming", McGraw-Hill Book Company, New York, 1969.

George, F.H., "The Brain as a Computer", from the International Series of Monographs on Pure and Applied Biology, Dept. of Psych., Univ. of Bristol, Pergamon Press, Oxford, 1961.

Gilman, Leonard and Rose, Allen J., "APL An Interactive Approach", John Wiley & Sons, Inc., New York, 1984.

Goodman, A.W. "Modern Calculus with Analytic Geometry", MacMillan Company, New York, 1970.

Grosky, W.I., and Jain, R. "Optimal Quadtrees for Image Segments", presented in "Pattern Analysis and Machine Intelligence", IEEE Transactions, Vol. PAMI-5, No. 1, 1983, pp. 77-83.

Hayes-Roth, F., Waterman, D.A., and Lenat, D.B. "Building Expert Systems", Addison-Wesley Publishing Company, Inc., Massachusetts, 1983.

Hogg, D., "Model-based vision: a program to see a walking person", from "Image and Vision Computing", Vol. 1, No. 1, February 1983, pp. 5-20.

Hwang, K., and Briggs, F.A. "Computer Architecture and Parallel Processing", McGraw-Hill Book Company, New York, 1984.

Ichbiah, Jean, Barnes, John and Firth, Robert, "ADA Programming Course", Honeywell, Minneapolis, Minnesota, 1980.

Janson, H.W., "History of Art", Prentice-Hall, Englewood Cliffs, New Jersey, 1973.

Johnson, R.M., "The Processing of Visual Form Information as a Function of Repeated Brief Exposures to Visual Stimuli", Doctoral Dissertation, KSU, 1975.

Krutch, John, "Experiments in Artificial Intelligence for Small Computers", Howard W. Sams & Co., Inc., Indiana, 1981.

Lakshmivarahan, S., "Learning Algorithms Theory and Applications", Springer-Verlag, Heidelberg, 1981.

Law, Averill M. and Kelton, W. David, "Simulation Modeling and Analysis", McGraw-Hill, 1982.

Lecky, J.E., "Forth for Computer Vision in Industrial Applications", from "Robotics Age", April 1985, pp. 25-28.

Lewis, William E., "Problem Solving Principles for ADA Programmers", IBM Corp., Hayden Book Company, Inc., New Jersey, 1982.

Marshall, Garry, "Programming with Graphics", Prentice-Hall, Inc., New Jersey, 1983.

McCurdy, Edward, "Leonardo da Vinci's Notebooks", Empire State Book Company, New York, 1923.

Mitiche, Amar and Aggarwal, J. K., "Contour Registration by Shape-Specific Points for Shape Matching", from Computer Vision, Graphics, and Image Processing, (Vol.22,No.3, June 1983) 396-408.

Morrill, W.K., "Analytic Geometry", International Textbook Company, Pennsylvania, 1964.

Nfer. Nelson, "Catalog of Clinical Tests and Procedures", subdivision 'Tests of Vision and Visual Perception', Windsor, England, 1984-85.

Omolayole, J.O. and Klinger, A. "A Hierarchical Data Structure for Storing Pictures" in "Pictorial Information Systems" from Lecture Notes in Computer Science, Vol. 80, Springer-Verlag, 1979.

Pearl, J., "Heuristics", Addison-Wesley Publishing Company, Massachusetts, 1984.

Pratt, Terrence W., "Programming Languages: Design and Implementation", Prentice-Hall, Inc., New Jersey, 1984.

Pray, W.C., "The Effect of a Prior Two-Dimensional Experience on the Drawing Spatial Ability of Third Grade Pupils", Master's Thesis, KSU, 1977.

Pringle, L.R., "Mental Transformations of Possible and Impossible 4-Cornered Tori", Master's Thesis, KSU, 1977.

Rich, E., "Artificial Intelligence", McGraw-Hill Book Company, New York, 1983.

Rhodes, M.L. and Klinger, A. "Conversational Text Input for Modifying Graphics Facial Images" from "Fuzzy Reasoning and its Applications", Academic Press, 1981.

Rosenfeld, Azriel and Kak, Avinash C., "Digital Picture Processing", Academic Press, Inc., New York, 1982.

Rosenfeld, Azriel, "Recent Trends In Image Analysis", "Cybernetic Systems: Recognition, Learning, Self-Organisation", Research Studies Press LTD., England, 1984, pp. 87-93.

Sayre, Kenneth M., "Recognition: A Study in the Philosophy of Artificial Intelligence", Univ. Notre Dame Press, 1965.

Schowengerdt, R.A., "Techniques for Image Processing and Classification in Remote Sensing", Academic Press, Inc., New York, 1983.

Sequeira, E.G., "Structured vs. Unstructured Scan Path in Static Visual Search Performance", Master's Thesis, KSU, 1979.

Silliman, B., "Visual Conceptualizing in Note Review", Master's Thesis, KSU, 1981.

Sandler, H., "Left Ventricular Volume, Mass and Related Measures; Usefulness in Determination of Ventricular Function" from "Quantitation in Cardiology", Leiden University Press, 1972.

Shu-Xiang, L. and Loew, M.H. "Boundary Chain Codes from Quadcodetrees", presented in Proceedings of the Workshop on Computer Vision: Representation and Control, IEEE, Maryland, 1984, pp. 214-218.

Sowa, John F., "Conceptual Structures", Addison-Wesley, Reading, Massachusetts, 1984.

Tucker, L.W., "Control Strategy for an Expert Vision System using Quadtree Refinement", presented in Proceedings of the Workshop on Computer Vision: Representation and Control, IEEE, Maryland, 1984, pp. 214-218.

Walpole, R.E., and Myers, R.H. "Probability and Statistics for Engineers and Scientists", MacMillan Company, New York, 1972.

Wilson, J.C. and Wilkinson, T.A., "WATCOM APL Users' Guide", Canada, 1984.

Winston, P.H. "Artificial Intelligence", Addison-Wesley Publishing Company, Massachusetts, 1984.

APPENDIX A

Person #1

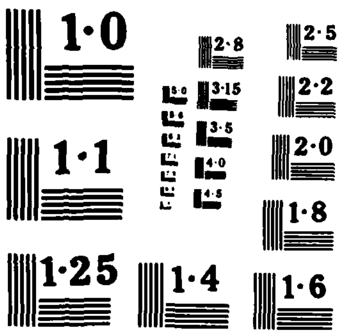
Person #2

Person #3

Person #4

AD-A158 073 A MODEL OF AN EXPERT COMPUTER VISION AND RECOGNITION FACILITY WITH APPLICATIONS OF A PROPORTION TECHNIQUE (U) ARMY MILITARY PERSONNEL CENTER ALEXANDRIA VA G E SHERMAN 15 MAY 85 2/2

UNCLASSIFIED F/G 9/2 NL



NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

Person #5

Person #6

Person #7

Person #8

Person #9

Person #10

APPENDIX B

```
Statistics   "Lefteye Proportional Ratio"

Total Population (enter 0 if unknown) = 0
Enter each value frequency pair.
Omitted frequencies default to 1.
Enter 0 0 after the last pair.

Pair 1: 0.625

Pair 2: 0.083

Pair 3: 0.786

Pair 4: 0.500

Pair 5: 0.625

Pair 6: 0.556

Pair 7: 0.500

Pair 8: 0.571

Pair 9: 0.333

Pair 10: 0.286

Pair 11: 0 0


Results tabulated as follows:
Total Population: unknown
Number of Samples: 10
Sum of Samples:        4.86
mean:        0.49
Sum of Squares:        2.73
Mean deviation:        0.15
median:        0.53
variance:        0.04
variance with Shep. Corr.:        0.03
Standard Deviation:        0.19
Std. Dev. with Shep. Corr.:        0.18
Unbiased estim. of variance:        0.04
Std. Dev. using that variance:        0.20
Probable error:        0.13
Standard error of mean:        0.06
Coeff. of variation:     39.38%
Range:  7.03000000000000e-01
Max Value:  7.86000000000000e-01
Min Value:  8.30000000000000e-02
```

```
Statistics   "Righteye Proportional Ratio"

Total Population (enter 0 if unknown) = 0
Enter each value frequency pair.
Omitted frequencies default to 1.
Enter 0 0 after the last pair.

Pair 1: 0.444

Pair 2: 0.571

Pair 3: 0.591

Pair 4: 0.375

Pair 5: 0.800

Pair 6: 0.778

Pair 7: 0.500

Pair 8: 0.400

Pair 9: 0.500

Pair 10: 0.571

Pair 11: 0 0


Results tabulated as follows:
Total Population: unknown
Number of Samples: 10
Sum of Samples:       5.53
mean:       0.55
Sum of Squares:       3.24
Mean deviation:       0.11
median:       0.54
variance:       0.02
variance with Shep. Corr.:       0.02
Standard Deviation:       0.14
Std. Dev. with Shep. Corr.:       0.14
Unbiased estim. of variance:       0.02
Std. Dev. using that variance:       0.14
Probable error:       0.09
Standard error of mean:       0.05
Coeff. of variation:     24.68%
Range:  4.25000000000000e-01
Max Value:  8.00000000000000e-01
Min Value:  3.75000000000000e-01
```

Statistics "Mouth Proportional Ratio"

Total Population (enter 0 if unknown) = 0
Enter each value frequency pair.
Omitted frequencies default to 1.
Enter 0 0 after the last pair.

Pair 1: 0.111

Pair 2: 0.500

Pair 3: 0.211

Pair 4: 0.273

Pair 5: 0.231

Pair 6: 0.083

Pair 7: 0.286

Pair 8: 0.179

Pair 9: 0.154

Pair 10: 0.211

Pair 11: 0 0

Results tabulated as follows:
Total Population: unknown
Number of Samples: 10
Sum of Samples: 2.24
mean: 0.22
Sum of Squares: 0.62
Mean deviation: 0.08
median: 0.21
variance: 0.01
variance with Shep. Corr.: 0.01
Standard Deviation: 0.11
Std. Dev. with Shep. Corr.: 0.11
Unbiased estim. of variance: 0.01
Std. Dev. using that variance: 0.12
Probable error: 0.07
Standard error of mean: 0.04
Coeff. of variation: 49.41%
Range: 4.17000000000000e-01
Max Value: 5.00000000000000e-01
Min Value: 8.30000000000000e-02

```
Statistics   "Triad Angle Lefteye-Mouth-Righteye"

Total Population (enter 0 if unknown) = 0
Enter each value frequency pair.
Omitted frequencies default to 1.
Enter 0 0 after the last pair.

Pair 1: 53

Pair 2: 46

Pair 3: 49

Pair 4: 51

Pair 5: 48

Pair 6: 51

Pair 7: 46

Pair 8: 45

Pair 9: 50

Pair 10: 49

Pair 11: 0 0


Results tabulated as follows:
Total Population: unknown
Number of Samples: 10
Sum of Samples:     488.00
mean:     48.80
Sum of Squares:   23874.00
Mean deviation:       2.04
median:      49.00
variance:       5.96
variance with Shep. Corr.:      5.88
Standard Deviation:        2.44
Std. Dev. with Shep. Corr.:      2.42
Unbiased estim. of variance:      6.62
Std. Dev. using that variance:      2.57
Probable error:       1.65
Standard error of mean:       0.81
Coeff. of variation:      5.00%
Range:  8.00000000000000e+00
Max Value:  5.30000000000000e+01
Min Value:  4.50000000000000e+01
```

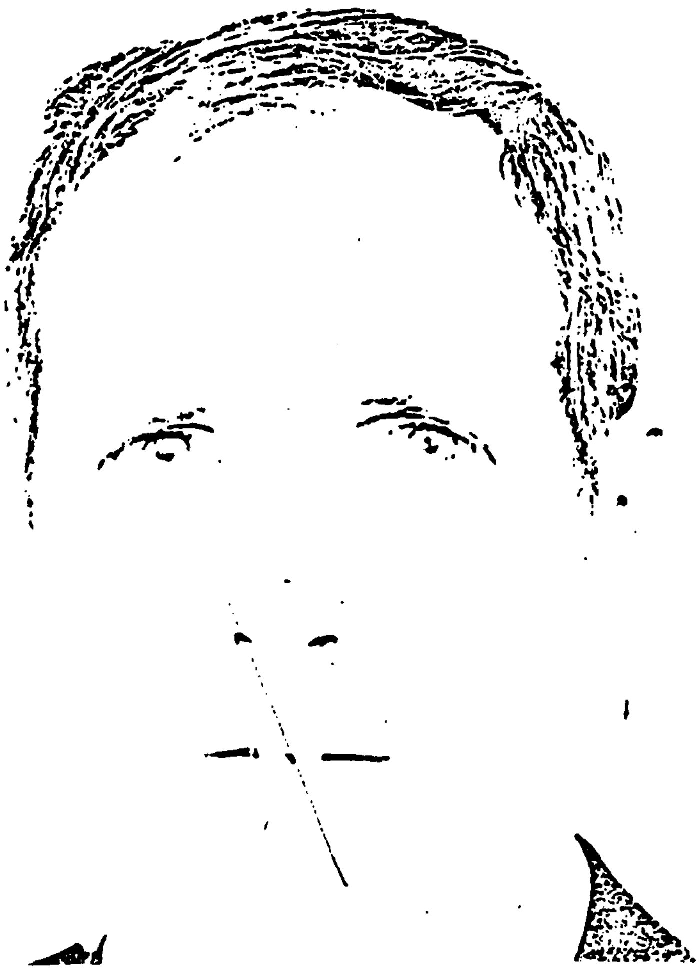
APPENDIX C

APPENDIX C

Phase 1 Threshold set to first level of lightness.

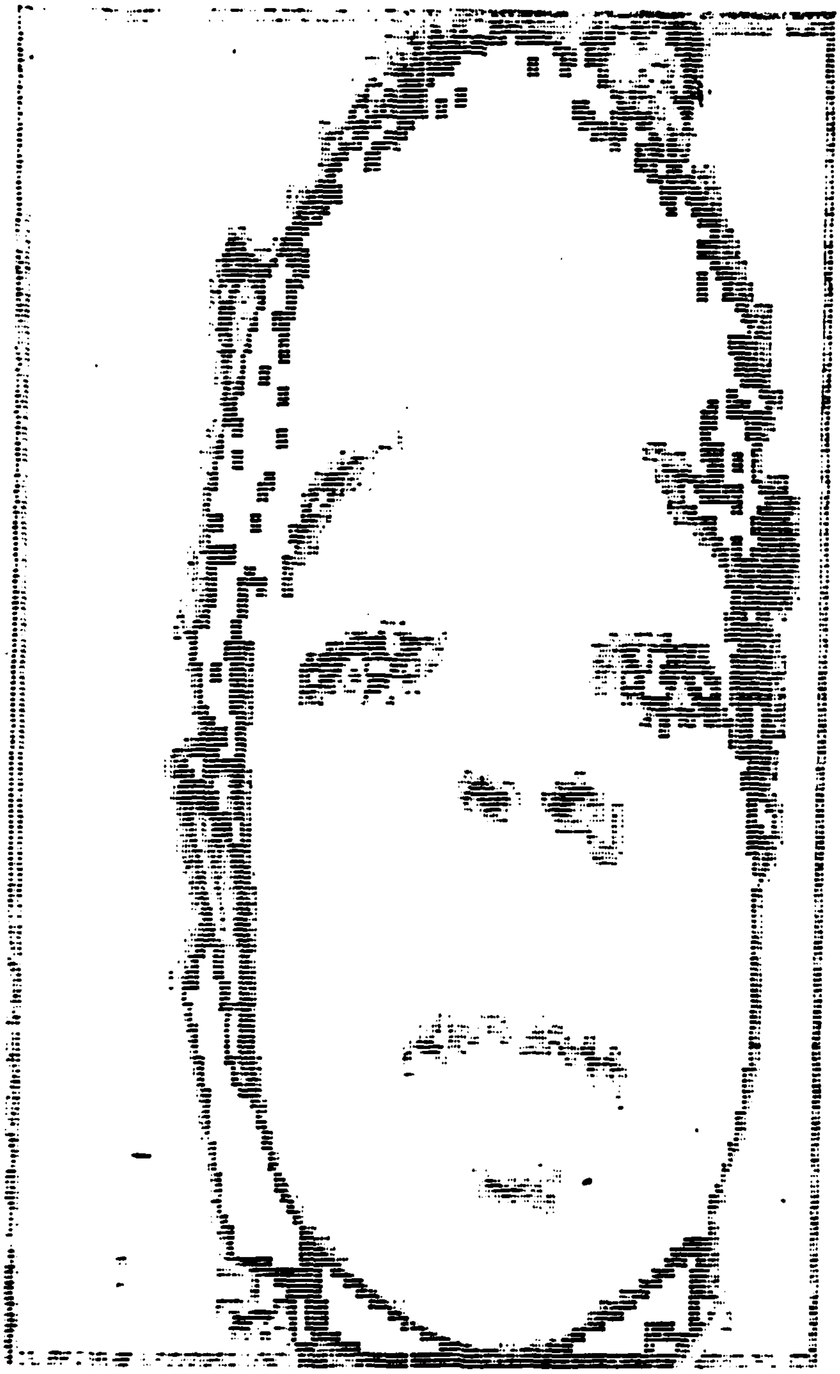
Phase 2 Threshold set to second level of lightness.

Phase 3 Threshold set to third level of lightness.
(This was assumed to be the most realistic level.)



APPENDIX D

Quadtree Analysis



A Model of an Expert Computer Vision and Recognition Facility With Applications of a Proportion Technique

by

George E. Sherman

B.S., United States Military Academy, West Point, 1976

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

College of Arts and Sciences

Department of Computer Science

Kansas State University
Manhattan, Kansas 66506

1985

Abstract

A MODEL OF AN EXPERT COMPUTER VISION AND RECOGNITION FACILITY WITH APPLICATIONS OF A PROPORTION TECHNIQUE

This thesis considers the human face as a planar shape consisting of planar shapes, and presents a new approach to face recognition, from a frontal view, based on the observation that matching of a new face with a database of known faces is a much easier task when the intrafacial shapes are represented in terms of centroids and angles. Representation is defined as finding scalar values that correspond to facial features. These scalar values are potentially useful in the design, storage and retrieval of information in data bases. Additionally, representation offers the advantage of applying shape matching algorithms which use the scalars.

It is shown that classification of planar figures can be achieved using shape-relative ratios computed from given figures. Representation of groups of planar figures is also presented as a technique for shape matching. The groups suggested are defined as triads, i.e., composed of three centroids. A scalar value for the angle between the centroids is computed, which represents the proportional position of facial features. Example calculations of the centroid, proportion ratios and relational angle are provided.

In image segmentation, shape is often the basis for the detection and recognition of patterns such as lines, edges, corners and more general images such as facial features. Quantifying shape as the proportion of objects, is an important

aspect in facial image processing systems. The scalar values of facial features define ranges of values for the image segments, thereby supporting detection and recognition.

An expert vision system is proposed which infers, from the collection and evaluation of facial segments of a face, a set of possible identities. Retrievals from a data store to accomplish this are facilitated by using the shape representations as keys to the records. The heuristic involves collecting the candidate identities into sets and then reducing the number of identities in those sets by set intersection and a weighting scheme. Sensitivity and tolerance factors in this expert are adjustable in terms of range values. The value obtained during evaluation of a facial segment is accepted or tolerated relative to the range of values expected for that segment. A retrieval effects a subrange whose width is determined by sensitivity to be the center and some number of adjacent values above and below that center value. Expectation ranges can be tailored to incorporate system and environmental variables.