2005

Measuring Community Consensus in Facial Characterization Using Spatial Databases and Fuzzy Logic

James Lee Mastros
Virginia Commonwealth University

School of Engineering
Virginia Commonwealth University

This is to certify that the thesis prepared by James Lee Mastros entitled MEASURING COMMUNITY CONSENSUS IN FACIAL CHARACTERIZATION USING SPATIAL DATABASES AND FUZZY LOGIC has been approved by his committee as satisfactory completion of the thesis requirement for the degree of Master of Science of Science in Computer Science.

Lorraine M. Parker, PhD., Associate Professor of Computer Science

---

Jeff Elhai, PhD., Associate Professor of Biology

David Primeaux, PhD., Interim Chairman, Department of Computer Science

---

L. Thomas Overby, Ph.D., Assistant Dean of Graduate Affairs, School of Engineering

---

Robert J. Mattauch, Ph.D., Dean, School of Engineering

F. Douglas Boudinot, Ph.D., Dean, School of Graduate Studies

Date

MEASURING COMMUNITY CONSENSUS IN FACIAL CHARACTERIZATION

USING SPATIAL DATABASES AND FUZZY LOGIC

A Thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Computer Science at Virginia Commonwealth University.

by

JAMES LEE MASTROS

Bachelors of Science in Biology, Virginia Polytechnic Institute and State University 2001

Director: DR. LORRAINE M. PARKER
ASSOCIATE PROFESSOR OF COMPUTER SCIENCE

Virginia Commonwealth University
Richmond, Virginia
December 2005

Acknowledgement

I would like to thank Dr. Lorraine M. Parker for her guidance, without which this thesis would not have been completed.

I would also like to thank all the faculty and staff of the Department of Computer Science at Virginia Commonwealth University.

Table of Contents

List of Tables

Page

List of Figures

Abstract

MEASURING COMMUNITY CONSENSUS IN FACIAL CHARACTERIZATION

USING SPATIAL DATABASES AND FUZZY LOGIC

By James Lee Mastros, B.S Biology

A Thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Computer Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2005

Major Director: Dr. Lorraine M. Parker
Associate Professor of Computer Science

Spatial databases store geometric objects and capture spatial relationships that can be used to represent key features of the human face. One can search spatial databases for these objects, and seek the relationships between them, using fuzzy logic to provide a natural way to describe the human face for the purposes of facial characterization. This study focuses on community perception of short, average, or long nose length. Three algorithms were used to update community opinion of nose length. All three methods showed similar trends in nose length classification which could indicate that the effort to

extract spatial data from images to classify nose length is not as crucial as previously thought since community consensus will ultimately give similar results. However, additional testing with larger groups is needed to further validate any conclusion that spatial data can be eliminated.

CHAPTER 1 INTRODUCTION

In the past thirty years, storage of information has become increasingly vital to many areas of modern human life. Financial institutions and government agencies, to name a few examples, rely heavily on databases for their daily operations. It is from this need to store all types of information that there has grown specialized database management systems to administer specific types of data.

In many cases, data of one type such as financial data may have domain-specific requirements or operations that aid in the storage and retrieval of that data type. An example of financial data that requires additional retrieval and update capabilities are bank transactions. Bank transactions require the ability to rollback account changes in the event that an update to an account does not fully complete. In the event that a user tries to transfer money from a savings account into a checking account, there must be measures in place to prevent the money being lost if the transfer fails and does not update the checking account with the new balance.

Spatial data is one data type that has shown exceptional growth over the past thirty years, especially for use by geographic information services. Spatial data refers to information describing the location and relationships between objects in multidimensional space. An example of data containing spatial relationships can be seen from a simple question involving driving directions. Imagine a person driving a car and asking where the nearest gas station is within a five mile radius relative to their position [2]. They may also want to know several driving routes past an upcoming traffic jam that will still take them to their destination.

Those are just a few examples where an ordinary person could take advantage of spatial databases. But who else would use them? Shekhar and Chawala [2] describe three classes of spatial database users. There are business, scientific, and web users. Business users utilize spatial data with other collected information to make decisions about marketing strategies and company direction [2]. Scientific users are interested in analyzing spatial data about local and global environments [2]. Web users want to use spatial data in conjunction with easy-to-use tools to search maps and ask questions that are relevant to their everyday lives.

Although web users, using services such as MapQuest and Google Earth, have been able to harvest additional capabilities in Spatial Database Management Systems (SDMS), there is another class of users. Among these users are criminal justice or law enforcement users who are interested in the cataloging and mining of human facial characteristics for purposes of identification or analysis. An agency such as the Federal Bureau of Investigation might want to take a victim's description of a suspect's face and run that query through a spatial database to return potential matches of criminals with those facial features. Human facial features such as the distance between eyes or retina patterns have long been researched for identification purposes; within a field commonly referred to as biometrics.

However, the purpose of this paper is not to pick out a facial feature solely for identification purposes, but to select several facial objects and their relationship to each other for the purpose of facial characterization. Our Database Research Group is looking into translating fuzzy queries regarding facial characteristics into SQL statements that can search a spatial database system.

Fuzzy queries are questions that use subjective words such as long, short, or broad. Such terms are called subjective because it is unclear what constitutes a ‘short’ nose. The definition of a short nose in one part of the world may be completely different from that in another part of the world. Therefore, we have been examining how best to translate queries containing fuzzy statements such as “Give me all the suspects with a short nose and broad chin” into SQL statements. In order to accomplish this, community learning has been utilized to derive meanings for short and long before they are translated into queries. However, imagine for a moment the potential uses and benefits of this effort. A suspect’s facial data may be searched on a database containing millions of criminals and top results returned as images, helping a victim identify the suspect faster.

This paper utilizes MySQL, a spatial database management system (SDMS), along with geometric shapes that represent key features of the human face to evaluate several updating algorithms. These algorithms were used to record community opinion about whether images of human faces have short, average, or long noses. The current updating method, method A, used by the Database Research Group initializes the value used to classify nose length with a random weight. The weight is updated by a fixed amount based on user feedback with the final classification of nose length ultimately determined by how much feedback is obtained. However, this method failed to adequately represent community opinion. For example, if 300 users indicated that a nose was short, 30 users could specify it as long and override the classification of short. Using a newly developed algorithm, named the ‘steplock’ method or method B, it now is more difficult for 300 users to outweigh 30 users and thus ultimately represents community opinion better. An additional method, method C, uses spatial data but updates according to method A and is tested to determine if the effort to extract spatial data for purposes of nose length classification is indeed worthwhile.

1.1 Spatial Data and Spatial Indexing

Traditional relational databases have long been and still are the primary method of storing and retrieving data. However, specific types of data warrant specialized management systems to efficiently handle data of that type. But what in particular, does a spatial database package do that traditional relational packages do not?

In order to answer this, it is necessary to show how spatial data is different from other data. First, spatial data tends to be more complex when compared with data typically collected for business use [2]. As a result, the standard indexing methods are not suited to derive 2D and 3D relationships quickly or efficiently and often result in excess computation. For example, take a query “List all students located in zip code 23294”. A traditional relational database will have no retrieval problems and utilize a standard index to quickly find the results. However, if a professor asks “List all students twenty miles from the main campus” the DBMS will struggle. This is due to the distance relationship that must be derived before a search can be narrowed. In order to complete this query, the Database Management System (DBMS), must take each zip code and translate its position in longitude and latitude and compare against the position of the main campus [2]. Second, traditional databases lack the object methods or constructs needed to represent spatial objects [2]. Third, storage requirements for spatial data are also more complex; low-resolution satellite pictures of the United States can be 30MB or more [2]. Therefore, how does a spatial DBMS handle indexing of data that is larger and more complex than traditional data?

In most relational database systems, a B-tree is utilized for indexing. However, due to the nature of spatial data, B-trees are insufficient and in some cases can result in the loss of spatial neighbor information. For example, consider a two dimensional matrix of size 4 by 4 with integer values inside each location. At any position, there will be a certain subset of values that are the chosen position’s neighbors. If ordering is used as a part of B-tree conditions for indexing, the neighbors might change and result in the loss of previous neighbors at those locations [2]. Therefore, B-trees were modified into R-trees to more effectively handle multidimensional objects. An R-tree is similar to a B-tree, but specially designed to be height-balanced for multiple dimensions [2]. It encloses each geometric object with the smallest single rectangle that encloses the object [11], also known as the minimum bounding rectangle (MBR). In other words, spatial objects are divided into bounding cells by a fixed size grid [2], and then the cells it intersects with are recorded as a rectangular object. Rectangular objects are simpler representations and thus faster in searching and used as the spatial key. R-trees can also capture spatial neighbor information in the branches of the R-tree. See figure 1.1 for an example showing that matrix position with 2 has neighbors 4,6, and 8.

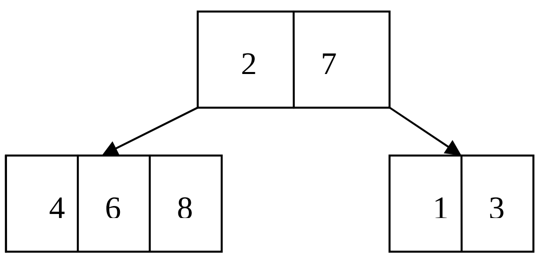

Figure 1.1 - Spatial relationships can be captured in the branches. Number 2 is neighbors with numbers 4,6, and 8.

Generally, spatial indexing is accomplished in two ways. The first is through data structures designed to hold spatial objects [2]. The second is to transform spatial objects into one-dimension so they can be used with the default one-dimensional index used with B-trees [2], often implemented with the Z-order or Hilbert curve algorithm to reduce processing time of spatial queries [2]. For more information concerning these algorithms, please refer to cited reference [2]. In addition, spatial indexes often utilize ‘buckets’ to organize objects. A bucket or container will have those objects that fall into specified regions or categories and are used for faster retrieval of data.

Spatial data is usually represented or described through geometric shapes. Dimensionless spatial object types are *points* while one-dimensional types include *lines*, *linestrings*, and *curves*. *Points*, for example, can represent buildings or other locations while *lines* can be used for state boundaries, railroads, or power lines. Two-dimensional shapes include *polygons* and *circles*. City outlines or districts are good examples of objects that can be represented by *polygons* [11].

The built-in functions involving these shapes can be broken up into two groups: topological and nontopological [2]. Topological refers to characteristics that involve relationships between objects that are not affected by transformations. For example, if you draw two squares on a balloon and then stretch the balloon, the squares still connect. Nontopological examples are perimeter, area, and length. If you stretch the balloon, these properties will be affected [2]. Examples of topological functions include *endpoint*, *touches*, and *contains*. Endpoint will return the point at the end of an arc and touches returns what shapes are touched by a shape. Contains returns those shapes enclosed within another shape. Nontopological functions include those that compute Euclidean distance or area.

In addition, most spatial packages allow collections of shapes. The advantage of providing collections is that additional geometric relationships involving set theory can be used such as intersection and difference. Examples of such a queries would be, “Give me all the states which interstate 95 transverses” or “What counties border Henrico county in Virginia?” The *intersection* function returns *true* for each state object that intersects with the *linestring* object representing I-95.

1.2 Spatial Database Management Systems

Currently, there are several spatial database management systems (SDMS) available. Commercial packages include Oracle, Sybase, and SQL Server. Oracle’s flagship product, Enterprise Database Server 10g, has very robust spatial extensions. Autometric, a division of Boeing, which has developed Spatial Query Server, has as an add-in for Sybase databases. As far as open-source options, MySQL is the most popular open source database package and includes support for two dimensional geometric shapes.

Each of the above packages conforms to Open Geographic Information System (OGIS) standards. OGIS is an international organization that consists of more than 250 companies, agencies, and universities that are actively engaged in solving common issues and problems found when working with spatial data [10].

The two predominant formats supported by OGIS for representing spatial data in the above SDMS are Well-Known Text (WKT) and Well-Known Binary (WKB). The WKT format is designed to use ASCII values to represent geometric shapes. For example, to represent a *Point* in WKT one would use *Point*(x, y), a syntax similar to object constructs in object oriented languages such as Java. A collection of shapes would be represented as *GEOMETRYCOLLECTION* (*POINT*(4 35), *LINESTRING*(13 13, 40 40)) [10]. The WKB format utilizes binary streams of data to represent geometric shapes. It is formatted as follows: “one-byte unsigned integers, four-byte unsigned integers, and eight-byte double-precision numbers (IEEE 754 format)” [10]. A WKB example that would translate from hexadecimal into binary and represents a *Point* object is 0100010000011000000000D02C000000000000E05F [10].

At the time of writing, the latest version of MySQL is 5.0 and it does not support more than two dimensional (2D) spatial objects. Oracle supports both 2D and 3D objects. Sybase, which can also handle 2D and 3D data, takes a unique approach and categorizes its shapes into three categories [24]. The first category contains shapes with a fixed number of parts such as point or rectangle. The second category consists of a collection of shapes having multiple parts such as mixed set of points, lines, and polygons. And the last category, consisting of lines and polygons, are those that are specified by the user when an object is instantiated and have a fixed number of parts [23].

MySQL 5.0 uses two methods for optimizing spatial searches used in constructing spatial indexes [20]. The first is to search for all points within a desired region and the second is to search for all objects that are contained within a region. MySQL uses R-Trees with quadratic splitting to index spatial columns [4]. Oracle also uses fast R-trees and quadtree indexing [11]. Sybase utilizes the above methods, but emphasizes clustering techniques for improved efficiency and also reduces the number of page reads required to do a search [23].

1.3 Nasal Length Measurement

In order to classify a facial attribute such as nose length into a 'short' category, additional information about possible lengths of the nose is needed. Unfortunately, the number of studies researching facial attribute measurements is limited [30]. Such studies have been done either on populations having a small age range or having few individuals as subjects. Having additional information concerning facial feature measurements could help in the diagnosis of dysmorphic syndromes such as Down syndrome. Down syndrome can be hard to diagnose in some cases because initial impressions of the patients face and symtoms may be misleading. The depressed nasal bridge may appear to be spaced widely when they are actually around the normal bridge length [30]. Yet another application for facial measurements is for plastic surgery analysis. Plastic surgeons may be able to better judge if their operation is a success by comparing lengths and relationships to other facial features [5].

In this collective study, measurements for the head, face, orbits, nose, lips, mouth, and ears were collected from 2326 Caucasian children and adults [27]. There were also 235 children and adults from China and 132 African American adults from the United States. In this study, male nose lengths ranged from 2.2cm to 5.0cm and females ranged from 2.1cm to 4.4cm.

The most recent study, done in 2002, measured nose length, nasal protrusion, and philtrum length for 2500 healthy individuals. All individuals were of central European decent with ages ranging from birth to 97 years old [30]. Figure 1.2 illustrates the technique of how measurements were taken from each of the participants which will help clarify how this study is interpreting 'nose length'.

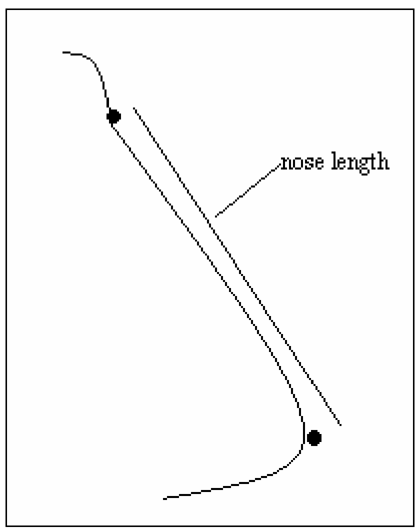

Figure 1.2 – Reveals how nose length is interpreted with the study.

These results show that nose length measured in subjects, between the ages of 15 and 80, ranged from 5.0cm to approximately 5.9cm. It also supports the common conception that older adults have larger noses. Measurements were also taken from males and females and on average, males had larger nose lengths than females. This is not that significant on the surface, but may hold a key if statistical analysis of the data is needed to group nasal lengths into short, average, or long categories. A comparison between the two studies, both the Farkas and the Zankl study, reveal that nasal length falls between 2 and 5.9cm from birth to 97 years old.

CHAPTER 2 EXPERIMENT OVERVIEW

2.1 Project Overview

We are looking into translating fuzzy queries regarding facial characteristics into SQL statements that can be used to search a spatial database system. Fuzzy queries are questions with subjective words such as long, short, or broad. They are subjective because it is unclear what constitutes a 'short' nose. The definition of short in one part of the world may be completely different from that in another part of the world.

Fuzzy words used to describe the face are of particular importance to the research group because they may better represent human thought and communication [25]. Therefore, the database group is researching into translating statements containing these fuzzy words such as "Give me all the suspects with a short nose and broad chin" into SQL queries. Currently, users can only query images based on one attribute such as eyes, chin, or nose. However recent work by the group, successfully demonstrated compound attribute querying [26]. Attributes can also have up to two modifiers. An example of a two modifier statement is a 'very short nose'.

In order to accomplish this, community learning was utilized to derive meanings for fuzzy words before they were translated into queries [25]. Example of fuzzy words include very, slightly, medium, short, average, and long. Users evaluated a set of images and provided feedback on whether a particular image met an attribute condition such as 'short'. As a result of their feedback, the community of users dictated the underlying value for a 'short' fuzzy descriptor.

Additional research for this paper explored what points, features, and objects of the human face should be captured and represented geometrically. Using MySQL, a total of 22 points were ultimately selected to represent the nose, eyes, lips, forehead, and chin (see Figure 2.1). The selection of the points came from an initial set of queries composed by the group on a limited number of facial attributes listed above. Point positions were selected to best capture metrics such as length, width, and area. There may be additional points added as the need to query other attributes grows.

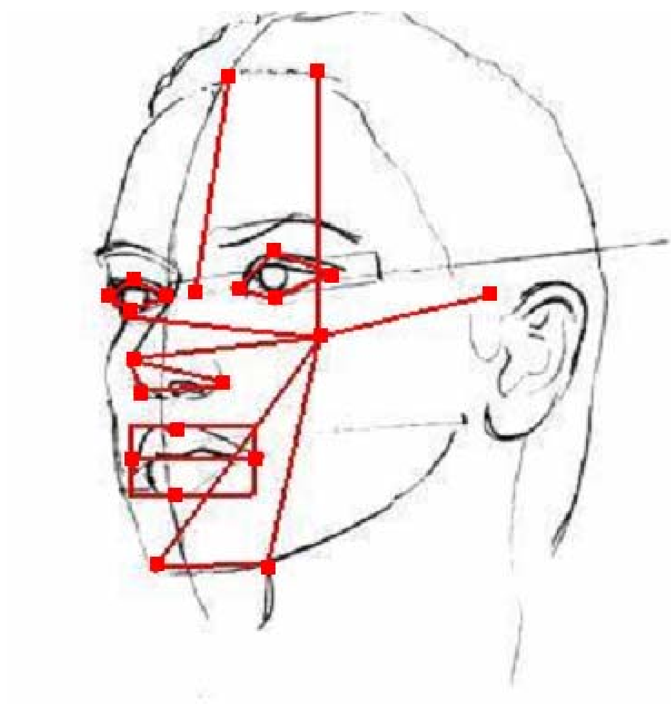

Figure 2.1 – Twenty-two points were initially selected to represent a limited set of facial attributes.

MySQL was chosen because it is open source, easy to use, and supports spatial objects. See Table 2.1 for the RELVARS of spatial objects made up by these points. Although only five shapes are represented, additional points and facial features may be selected and stored as the need to query these features becomes greater.

The next phase of research involved integrating the ability to represent facial features spatially with existing work of the database group into translating statements containing fuzzy words into SQL queries. It was in this integration process that several problems immediately appeared. The previous system did not need human subjects because random weights were used to represent descriptors of an attribute such as color or nose length. However, the new method obtains an actual value for nose length by measuring the attribute directly from a human subject. It is possible to determine nose length or eye color proportions of red, green, and blue from an image, but a hands-on measurement is preferred.

Unfortunately, to do so required additional human testing approval and special equipment. For this study, nose length was easier to work with because it is less ambiguous than eye color. Nose length is simply represented in mm or cm, rather than a composition for red, green, or blue. Nose length itself is not devoid of ambiguity, however, because of the variation of human facial features and the complexity in determining the exact locations of the points on the human face.

Schema: Face	
Table: Eyes	
Field	Type
leftEye	GeometryCollection(Line, Line, Polygon)
rightEye	GeometryCollection(Line, Line, Polygon)
midPoint	Point(x,y)
SSN*	Varchar (9)
Table: Nose	
Tip	Polygon(Line, Line, Line)
Alvar	Linestring(Point(x,y), Point(x,y))
Peak**	Linestring(Point(x,y), Point(x,y))
SSN*	Varchar (9)
Table: Mouth	
vermillionHorizontal	Linestring(Point(x,y), Point(x,y))
vermillionVertical	Linestring(Point(x,y), Point(x,y))
Lips	Polygon(Line, Line, Line, Line)
SSN*	Varchar (9)
Table: Chin	
Mandiblepeak	Point(x,y)
Peak	Polygon(Line, Line, Line)
SSN*	Varchar (9)
Table: Forehead	
Frontal	Linestring(Point(x,y), Point(x,y))
Vertical	Linestring(Point(x,y), Point(x,y))
Horizontal	Linestring(Point(x,y), Point(x,y))
Widowspeak	Point(x,y)
SSN*	Varchar (9)
* primary key, **spatial index – a spatially derived feature such as nose length is chosen as an index.	

Table 2.1 – Database RELVAR constructed with MySQL to represent the face spatially

There is also the problem of how the underlying value representing nasal length will be updated. Since the current updating system is lacking the ability to adequately represent change in community opinion, a new system needed to be designed to accommodate both retrieval and updating of spatial data as well as to better represent community consensus concerning nose length. Additional information on how this is accomplished is given in Section 2.3.

Thus the scope of this paper is to determine how the ability to represent the human facial attributes spatially will interact and fit into existing work done by our Database Research Group. In particular, how will nose length be used to select and update images as users vote their opinion on short, average, or long noses?

2.2 Experiment Overview

The purpose of this research study was to evaluate different methods of assigning initial weights to attributes within a fuzzy database. The Database Research Group originally decided on a plan of initializing weights and then having the community provide feedback in order to adjust the weight to reflect their opinions. This plan developed since the DBRG was trying to adapt methods from machine learning to the problem at hand. Within the context of machine learning, it is usual to use randomly assigned weights as the initialization process. Since, however, this is not truly a machine learning environment, there are other alternatives. One such alternative is to use a voting mechanism. Another is to use some factual information to predict the value such as data deduced from the image. As an initialization process, direct voting uses an expert or community subset to predict the consensus of the community. Direct voting can be used in the process of community learning, from the point of initialization onwards, but is not explored further in this study. The quality of this process will depend on how well the expert reflects the community as a whole. The question is whether this is any better than just choosing a random weight, or whether a better initial value can be deduced from the image.

The first method assigns random weights to represent nose length. The reasoning behind this approach is that the use of a random weight will place it randomly within the short, average, or long classification range. The community subsequently provides feedback on nose length thereby increasing or decreasing this value, ultimately placing the value within the classification range where the community believes it belongs. The advantage is that it also easy to implement and does not require obtaining any measurements from human subjects or from facial images. The disadvantages of this approach are not with the assigning of random weights but how those weights are updated according to community opinion. Thus, if not properly implemented, the opinion of the few will outweigh the great. In Method A, for example, 30 users providing feedback that a nose is long can outweigh the 300 users' opinions of short.

The next approach assigns a value for nose length derived from an image or measured directly from a human face and provides two advantages. The physical length of the nose is known, thereby allowing additional study into the differences into how the community views a statistically long nose to a perceptually long nose. There is also the ability to analyze how nose lengths from one image initially compare with other human noses. When utilizing random weights, the weight is meaningless in the beginning and cannot to compare to another random weight until after the data altered by a community's opinion. This requires that a training phase take place, since only feedback from an expert or a group of users will move the weights to values that reflect the communities' opinion. The physical nose length measurement will allow you to make such comparisons from the start. The disadvantage of this approach is that is more difficult to obtained length values. The difficulty is due to the time and effort needed to measure large numbers of individuals manually. If this project is to be used with millions of faces, it is not feasible to manually measure each individual's nose length. And if automated measurement is employed, accuracy and precision to obtain exact points of measurement within images will be challenging. Method B falls into this category. Method B allows for handling of large numbers of images without training (or an expert) being needed. Also, method B does not distinguish between the actual length of a nose and the perceived length of the nose by the community.

The last method, direct voting, involves asking the expert or each user of the community subset whether a nose contained within an image is short, average, or long. If the user agrees, a vote of yes is recorded for that category. If the user disagrees, a vote of no is recorded. All votes of yes and no are summed and the higher number dictates its classification. For example, if 10 users voted a nose to be short and 5 users disagreed, then the nose would be classified short. A derivative of this approach may assign different weights for a vote of short, average, or long. Overall, the advantage of this approach is that it is relatively easy to implement. A disadvantage is that it assumes that the expert, or community subset, will correctly represent the community as a whole. In order to overcome this, a methodology must be in place to allow continuous user feedback.

Another objective of the experiment was to determine if the effort to extract spatial attribute data, such as nose length, from images offers advantages when compared to simply assigning random weights to nose length. The decision to assign a random weight was made at the start of work toward this project, but there may be other methods that do not assign values at all. If the community reaches the same consensus with real data that it does with random weights then there may not be an absolute need to store spatial data. However, it should not be eliminated if community consensus is reached more quickly with its use. A second objective is to improve the method in which values representing nose length are updated as the users vote on the image.

The previous method of updating the attribute weights, under specific conditions, did not adequately handle changes in community opinion of the attributes. For example, if 300 users provided feedback that an image had a short nose, a smaller group of 50 users could change it to a long classification. The method did not scale weights proportionally to user feedback. Further descriptions of these methods are provided in Section 2.3. The new method is designed to better represent community opinion as well as to give an indication of confidence of the image classification acquired by votes.

The testing environment consisted of 15 computers loaded with an application written in Visual Basic.Net. Four test sessions were conducted on different days with approximately 40 users providing feedback. Ideally, a test session with 60 users would be preferred because of the range of the three classification boundaries are separated by approximately 10. Therefore to go from one extrema to the other, 30 each direction, a minimum of feedback from 60 users is needed.

An earlier application used to collect user feedback was originally written for earlier work done by the Database Research group; however it was redesigned and retooled with additional features for the specific need of this experiment. The database backend was also adapted to MySQL instead of Microsoft's SQL Server. The main purpose of the program is to present images of human faces to the users based on predefined criteria. The user then provides feedback whether the faces meet criteria such as faces with short, average, or long noses. The user can choose if the nose length from images 'meet the criteria', 'shorter', 'longer', or 'no preference'.

The application also contained logging features that tracked and recorded how users voted on images and which method of updating was being evaluated. This helped for analysis of user voting and difference between updating methods. It also provided data for further studies by other members of the database group. There were several user design changes as well. Although not the main scope of the study, the user design changes were aimed at providing users with a more enjoyable experience when evaluating and giving feedback about the images. Examples of helpful features include scrolling windows and progress indicators.

Instruction sheets containing screenshots and detailed descriptions were provided to users to alleviate confusion and provide step-by-step assistance when executing the image evaluation application. Users were subsequently asked to answer a survey indicating their opinion of the user interface for work done by another member of the research group.

The 15 images of human faces displayed to users and stored in the database throughout testing were selected based on how well an image displayed both the forehead and chin from 40 images obtained freely on the internet. The reason for this selection criterion is explained further on in this chapter. Fifteen images were chosen to keep the test between 5 and 10 minutes and to help ensure that images received adequate votes. These images are free of any legal or human subject restrictions and have been used by the Database Research Group in the past. Images were included if the complete forehead and chin were visible, due to the technique used to extract nasal length from the images and normalize them. Using a common proportion found in the human face, the distance between the forehead and the tip of the chin is proportional to nasal length [28] (See Figure 2.2). For example, if the image had a distance from the forehead to chin of 1.29 and 0.32 was the distance between the point on the nose at eye level and the tip of the nose, and then the proportion would be 1.29/.32 which equals 4.03. This helped normalize the values to alleviate discrepancies between subject distances to the camera. Ultimately, it would be better if procedures for testing of human subject were approved and a set of images obtained from individuals measured for exact lengths. The image points were selected as carefully and consistently as possible via a computer mouse, but existing inconsistencies may affect how images were originally classified as short, average, or long.

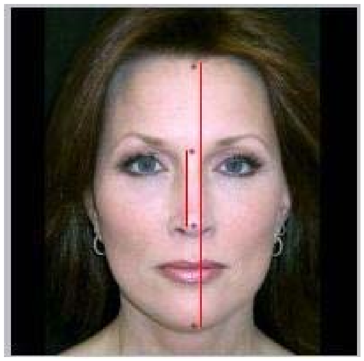

Figure 2.2 – the proportion of the top of the forehead to the bottom of the chin divided by the length of the nose from the nasal tip to the midpoint of the eyes is proportional to our interpretation of nasal length as shown in Figure 1.3

2.3 Methodology

The experiment utilized three different updating methods. The method's main function recorded each user vote and updated the weight or value of the attribute being evaluated. This value ultimately indicated whether a subset on an image was categorized as having a short, average, or long nose. A method uses simple average was not tested due to time restraints but could be an avenue of future research.

2.3.1 Updating Method A

Method A is the original method used by the Database Research group to adjust image weights. It is a simple implementation originally designed to show proof of concept. It divides the short, average, and long categories into three continuous ranges. The short range begins at 0.1 and ends at 0.33, the average from 0.34 to 0.66, and the long from 0.66 to 1.00. See Table 2.2 below.

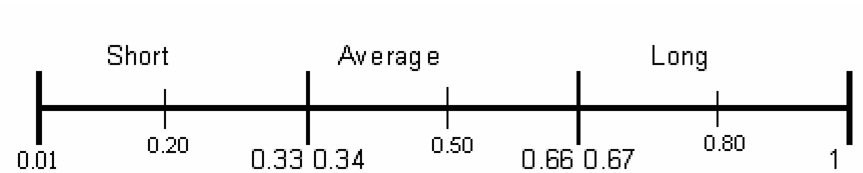

Figure 2.3 – Range used in method A to classify nose length as short, average, or long.

0.0 is not possible due to the semantic meaning of having a nose length of 0. Although, it is possible for a human to not have a nose, for the purposes of this study, it is not. There is also a threshold value located in the middle of each range that is utilized by this method. A better description of the threshold value would be a ‘convergence point’. It serves to help separate votes over time by moving image points that ‘meet criteria’ closer to the indicated threshold value. The threshold values chosen for short were 0.2, 0.5 for average, and 0.8 for long. This merely serves to increase separation between groupings of data points. The midpoints within each range would be a good choice for threshold values and were not selected as an oversight.

Initialization of this method involves assigning random weights between 0.1 and 1.0 to each image. When a user votes the image in the shorter or longer direction, a step size of 0.03 is used. Different step sizes can be used, however 0.03 gives each range about 8 to 10 steps. If a user votes that the image 'meets the criteria', then the image is moved 0.03 steps toward the threshold value of the range. For example, if the current image value is 0.31 and it is voted as 'meeting the criteria' of short, then the value would be 0.28. If the value is 0.18 and voted as meeting the criteria, then it will move to 0.21. Additional measures are taken to ensure that values never exceed the upper bounds of 1.0 and lower bounds of 0.1. For example, if a vote is at 1.0 and is voted 'larger' then it simply remains at 1.0.

Although a simple method to implement, this method suffers from a fundamental flaw. It does not adequately represent change in community opinion. For example, if 300 users vote that a nose is short and the value is moved to the far left extrema 0.1, it only takes 30 users voting that it is long to reclassify it as a long nose. The method relies heavily on the principle that attribute values will move back and forth and the community will eventually reach a consensus, but with the inability to hold the greater opinion of 300, it doesn't do so effectively.

2.3.2 Updating Method B

Method B, also named the 'steplock' method, was redesigned to assign a weight proportional to how users voted and made it harder for 30 users to override the greater opinion of 300 users. The method is named ‘steplock’ because of the way values become sticky or locked into a classification range the more users vote on that classification. Furthermore, the name also stems from the fixed number of steps inside each range and the slopes between the transitions from one classification to another. For example, there are two slopes in Figure 2.4; one slope from short to average and another from average to long. The reason they are not visible is that in the first step of the algorithm, there are no steps on the slope and the slope is equal to 0; it initially appears similar to the traditional method. Steps are added between these transition points when users consistently vote an image in a category in which it is located. This concept is explained in further detail below.

However, before the algorithm is explained, explanation of how the boundary and image values are selected is needed. The boundaries between short, average, and long categories are chosen similarly to the traditional method.

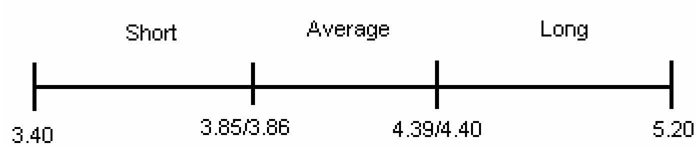

Figure 2.4 – Range used in method A to classify nose length as short, average, or long (units are in cm).

The lower and upper boundaries for use with real nose lengths were obtained from research into the upper and lower limits of human nose length. In one study, researchers from the United Kingdom tracked nose length, nasal protrusion, and philtrum length from birth to 97 years [30]. Their finding showed that nose length, as shown below, grows over the course of human aging and older adults tend to have larger noses than young adults. In order to determine which lower and upper values were used, a subset of the full 0 to 97 years was selected. Since most of the images in the evaluation program were young adult to senior citizens, only nose lengths from 15 years to 80 years were considered. Values ranged from 5.0 cm to 5.9 cm over these age groups (see Figure 2.5). If a database contained images predominately of children, then a larger range should be used. As you can see in Figure 2.5 shown below, nasal growth slows and plateaus.

Ultimately, values from this study could not be used to classify the upper and lower boundaries due to the inability to have human subjects and thus real nose length values. Therefore an alternative method was used using images taken from a set used by the Database Research Group. As described above, the proportion of forehead to chin divided by the two points on the nose normalized the images and provided a value for each image. The values for nose length were sorted and upper and lower extremas identified. Using these values, an approximately equal distribution for short, average and long values was determined. Ideally, these boundaries would be representative of real nose length, however, due to experimental constraints this will be left for future work.

In addition to boundary values, the steplock method uses a ‘perceived value’ as the underlying value that is used in updating and selecting images. A ‘perceived value’ initially equal to the real length extracted from an image, but changes with community opinion. The reason for making a distinction between real and a perceived value is that when evaluating a particular feature of the human face, other features may play a part in how the user views that attribute. For example, faces with nose values in the upper range of 5.9 cm might not be perceived as having long noses by the community due to an individual having a wide face or a broad chin. This new ability to separate real values from community perceived values opens another avenue of future research into how facial attributes relate to each other.

Step 1

An image's relative nose length is extracted from an image. The perceived nose length is set equal to the relative length and used to initially categorize images. Classification boundaries are as shown in Figure 2.4.

Step 2

Users are presented with a screen showing the image and allowed to rate the image from 4 potential choices. These are 'meets criteria', 'shorter', 'longer', and 'no preference'. The first choice 'meets criteria' means that the image fits the specified criteria and is interpreted as a vote of confidence in the current category. One step is added on the slopes between category transition points. Two additional fields are also updated. The field 'current votes' keeps tally on the number of votes or user feedback given to the current category the image is located in. The second field is the total votes an image has received. The current vote divided by the total votes gives the confidence level that an image is correctly classified. For example, if an image has 10 votes in its current field of short and

40 total votes then there is a 25% confidence level that the image is correctly classified as having a short nose. Details on how these fields are updated are given below.

Step 3

As users vote that an image meets criteria, the steps on the slope between short and average increases. This can be seen in Figure 2.5 where the slope between the transition point of 3.85 to 3.86 has increased as the result of adding steps.

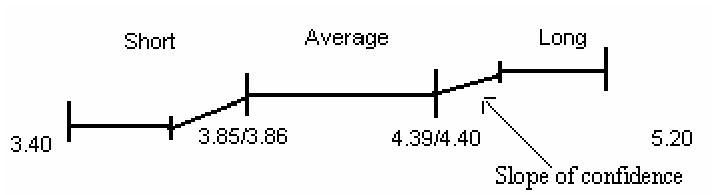

Figure 2.5

If a vote is not for the current category and the number of steps on the slope is greater than 0, the number of steps is subtracted by 1. If there are no steps on the slope, the value simply increments by 0.05. If the image is allowed to move, meaning there are no steps on the slope, the image value is checked to detect whether the image has changed categories. If so, the current vote field is set to 1 and the total votes category is incremented by 1.

Step 4

If a nose is in the short category and receives a vote of ‘shorter’, it is interpreted as a vote of confidence and is treated as it meets criteria. The reasoning is that if the criteria is to rate images with a short nose and the user votes ‘shorter’, the nose is still considered short.

Likewise, if an nose is in the long category and receives a vote of ‘longer’ then it is treated as meets criteria. This may change if additional modifiers such as ‘slightly’ or ‘very’ are implemented. Noses in the average category receiving a vote of shorter or longer, simply increments or decrements the underlying value by 0.05 if no steps are on the slope.

Step 5

If a user chooses ‘no preference’, values are not altered, and the total number of votes is not incremented.

2.3.3 Updating Method C

Method C differs from Method A in three ways. Relative nose lengths extracted from images and used in the steplock method are substituted for the random weights used by method A. Classification boundaries are also setup with similar values to the steplock method. The threshold values also reflect the difference in boundary values and the value for short is 3.6, average is 4.1, and long is 4.7. The step size of 0.05 remains the same to keep the number of steps within each range between 8 and 10.

CHAPTER 3 RESULTS

A total of four feedback sessions were conducted to gather data to evaluate the updating algorithms. Each session lasted approximately 3 hours with method A receiving feedback from a maximum of 46 users and a minimum of 35. The steplock method received feedback from a maximum of 39 and a minimum of 35. The control method received feedback from a maximum of 53 users and a minimum of 35. The difference in user feedback per method is due to users not fully completing each testing session. In some cases due to time restraints on the evaluator, the complete listings of images were not completed. Users were encouraged to finish, however it was not a requirement.

The numerical difference in user feedback is also due to inconsistencies in user's interpretation of the application interface. For example, several users did not see the complete list of images to be evaluated because of a design flaw with the scroll bar. The scroll bar appeared correctly on the screen but due to the color of the pane, several users did not scroll down to see the remaining images. As a result, when the user submitted their feedback the images that were missed were submitted with the default selection of 'no preference'. In these cases, users simply re-submitted feedback on missed images because a vote of 'no-preference' did not after user opinion about nose length.

Method A had six images out of the fifteen that did not change from their original random classifications. Initially, there were three images in the short category, 8 in average, and 4 in the long category. After testing, there were 2 images in short, 9 in average, and 4 images in long. The steplock method had 9 images that did not change classifications. This behavior is as expected because the steplock method makes it more difficult to change classification, especially if a large group of users consistently voted an image into a category. The initial classification distribution of the steplock method had 6 in the short category, 8 in average, and 1 in the long category. After testing, 4 were short, 6 average, and 5 were long. The control method had 7 images that did not change classifications. Initially, 6 were classified as short, 8 were average, and 1 long. This method was initialized similar to the steplock method. After testing, 8 were short and 7 were average; there no images with noses classified as long.

Analysis of the results between methods showed that when comparing method A to the steplock method that they agreed on 40% of their original image classifications (see Figure 3.1a,b,c). However, they agreed on 80% of their classifications after testing. Method C initially agreed with 40% of method A's classifications and 33% after testing. The initial and final classification between each method might also be due to where an image's nose length initially falls into each classification range. The farther away a nose length value is from a classification boundary, a greater number of votes will be needed reach consensus. This alone is insufficient to fully evaluate the methods; therefore image by image analysis is needed to determine trends.

Image 5 was initialized with 0.23 and ended with 0.56 with the nose length classification changing from short to average. The steplock method initialized image 5 with a value 4.0 and as average. It ended at 4.1 and was still average after 36 votes. Method C initialized image 5 as average and ultimately as short. Values started at 4.0 but fluctuated back and forth from 4.0 to 3.85. For this image to be classified as average, 3.86 is needed: a difference of 0.01 and therefore 1 vote. Overall, in each method, the image 5 moved toward the average category (see Table 3.1 a,b,c).

Method A

Image #	Initial	Final	Initial Classification	Final Classification
5	0.23	0.56	short	average
6	0.26	0.8	short	long
10	0.35	0.77	average	long
11	0.4	0.58	average	average
12	0.52	0.76	average	long
17	0.39	0.57	average	average
18	0.5	0.53	average	average
20	0.8	0.56	long	average
21	0.73	0.49	long	average
22	0.68	0.8	long	long
27	0.47	0.29	average	short
29	0.66	0.6	average	average
30	0.27	0.27	short	short
38	0.44	0.5	average	average
39	0.83	0.56	long	average

	Short	Average	Long
initial	3	8	4
final	2	9	4

Table 3.1a – Method A initial and final random weights after obtaining community opinion whether an image contained a short, average, or long nose.

Method B

Image #	Initial	Final	Initial Classification	Final Classification
5	4	4.1	average	average
6	4.2	4.4	average	long
10	3.7	4.4	short	long
11	3.7	4	short	average
12	3.7	4.45	short	long
17	4	4	average	average
18	4.1	4.1	average	average
20	5.2	4.65	long	long
21	4.2	3.8	average	short
22	4.1	4.4	average	long
27	3.6	3.6	short	short
29	4	4.05	average	average
30	3.5	3.5	short	short
38	4.3	4.25	average	average
39	3.7	3.8	short	short

	Short	Average	Long
initial	6	8	1
final	4	6	5

Table 3.1b Method B initial and final extracted nose length after obtaining community opinion whether an image contained a short, average, or long nose.

Similar analysis was performed on all images. Images 6 and 10 both moved toward the 'longer' direction in all three methods. Although, the control method showed them as average, both images were within 1 to 2 votes from long. Image 11 was average in all methods. Image 12 moved toward 'longer' in all three categories with the traditional method moving from average to long, and the steplock method moving from short to long. The control method, as with image 10, moved from short to average, but was not far from long. Image 17 and 18 showed similar movement and the differences in classification were minor.

Method C

Image #	Initial	Final	Initial Classification	Final Classification
5	4	3.85	average	short
6	4.2	4.35	average	average
10	3.7	4.25	short	average
11	3.7	4.05	short	average
12	3.7	3.9	short	average
17	4	3.85	average	short
18	4.1	3.85	average	short
20	5.2	3.75	long	short
21	4.2	3.65	average	short
22	4.1	4.05	average	average
27	3.6	3.4	short	short
29	4	3.9	average	average
30	3.5	3.4	short	short
38	4.3	3.95	average	average
39	3.7	3.5	short	short

	Short	Average	Long
initial	6	8	1
final	8	7	0

Table 3.1c – Method C initial and final extracted nose length after obtaining community opinion whether an image contained a short, average, or long nose.

Image 20 moved in a downward trend toward ‘short’ across all three methods, however the classifications did not all agree; additional votes may negate any of these differences. Image 21 also showed a ‘shorter’ trend with two out of the three methods agreeing on a short classification. Two methods classify image 22 as long while control strongly classifies it as average. Image 27 was a solid short across all three methods with two of the methods starting out with values in the short category. Images 29 and 38 were average in all three methods. Image 30 was short with all three methods and Image 38 was average in all three. Image 39 was short in two methods and average in the traditional method, but all three showed a downward trend.

ID	PLength	NumSteps	NumVotes_Current	NumVotes_Total	Confidence Level
5	4.1	22	27	34	79.4%
6	4.4	26	29	35	82.9%
10	4.4	1	3	34	8.8%
11	4	0	4	34	11.8%
12	4.45	9	16	34	47.1%
17	4	27	31	35	88.6%
18	4.1	23	29	35	82.9%
20	4.65	0	12	35	34.3%
21	3.8	0	4	29	13.8%
22	4.4	10	17	31	54.8%
27	3.6	20	26	32	81.3%
29	4.05	6	17	29	58.6%
30	3.5	10	21	32	65.6%
38	4.25	8	15	23	65.2%
39	3.8	9	20	33	60.6%

Table 3.2 – Table used to record user feedback with Method B. Plength refers to this studies interpretation of nose length. NumSteps is the number of steps on the slope between classification ranges. NumVotes_Current is the number of votes an image is voted in its currently located category. NumVotes_Total is the total number of votes received on an image. Confidence Level is the NumVotes_Current divided by NumVotes_Totoal.

The steplock algorithm provides the confidence level an image belongs inside in current category based on user votes. The results in Table 3.2 show images 5, 6, 17, 18, and 27 have confidence levels around 80%; some reaching as high as 88.6%. Image 18, for example, has 29 votes in its current category of average and 35 votes overall, therefore 31/35 = 82.9 confidence. These images were classified similarly across traditional and steplock methods, but several images in the control were a few votes short. The control method shows the same trend as traditional and steplock, therefore the discrepancy in classification should be similar if additional votes were given. The 11.8% confidence value of image 11 has the same classification of average across all three methods.

Another pattern of interest is whether or not the selected set of images used for this study of older individuals were repeatedly reported to have larger noses than images of younger persons. Likewise, were younger individuals voted as having short noses? Such opinion would follow the recent research into measuring human nasal length from birth to 97 years old in a group of individuals of European decent [30]. This study reveals that nasal growth continues through human aging, and that older individuals have larger noses than younger. To test if the user community follows this trend, the following were identified subjectively as older individuals (see Images 38, 29, and 6 in Figure 3.1).

Image 38 Image 29 Image 6

Figure 3.1 Images 38, 29, and 6

Referring to this figure, image 6 was classified as long in both method A and the steplock methods. Method C was .15 away from long and therefore less than a vote away. Images 29 and 38 were voted average with all three methods. Although image 6 was voted 'long', images 29 and 38 were solidly in the average category. Therefore, with the user feedback obtained, support for the pattern observed in the Zankl study of older individuals having longer noses when compared with images selected for this study is inconclusive. Perhaps with additional data and a larger sample size, a stronger conclusion can be made. In addition, this does not take into account other facial features that contribute or take away from the belief in a longer nose.

Image 18

Image 30

Figure 3.2

Images 18 and 30 were identified subjectively as younger individuals. Image 18 was voted as average with traditional and steplock methods, but short with the control method. The short value, however, was .01 from average and less than 1 vote away. Image 30 was voted as short across all three methods. This also does not prove or disprove that younger individuals are identified as having shorter noses.

CHAPTER 4 CONCLUSIONS

The process of evaluating community opinion of nasal length with method A, method B, and method C yielded interesting results. All three methods showed similar trends in behavior for each image and, in the majority of cases, the final classification was consistent across all three methods. In the instances where the final classifications differed, it was usually only a short distance within the classification range away from being consistent.

This is significant because it implies that the effort needed to extract real spatial data is not as crucial to classifying nose length as previously thought. Community consensus with random initial weights appears to mirror user feedback with spatial values. It is important to note that due to human subject restrictions, actual spatial data was not used. However, if the proportions exploited to derive nasal length from images do indeed reflect actual nose length, then the results should still hold.

But before concluding that spatial data may not needed when classifying nose length, a larger group of users must evaluate the images. Overall, each image received feedback from 30 to 50 users with a target of 60 users. It is likely that having a maximum of 50 users' feedback will not provide accurate image classification. Thus it may be that the results obtained are merely a 'snapshot' in larger community opinion that supports that spatial data is not essential to classifying nose length. The original target of 60 users was selected so a value representing nose length had enough user feedback to reach the endpoint of the short classification range to the endpoint of the ‘large’ classification. This approach was purely subjective and additional testing sessions with larger groups is needed to better understand how the underlying updating algorithms reflect community consensus. The greater number of votes recorded will provide higher confidence in the image’s classification. The concept of a ‘community’ in these studies refers to a set of individuals with something in common, such as a shared geographical location. Clearly, the size of the community can vary depending on its defining attribute.

Many reasons remain for using spatial data even if it is not required for community classification of nose length. An advantage to using spatial data is that no training session is needed to prepare nose lengths before users give feedback and additional analysis of relationships between facial attributes is possible through the use of real and perceived values. However, the savings in any efforts to formulate a precise and accurate method of point extraction on human faces is enormous. With this information, VCU’s Database Research Group may continue forward without having to pursue such an endeavor. Spatial data also allows the instant classification of an image’s nose length by revealing where an extracted nose length falls into previously established user opinion about nose length.

A disadvantage to using spatial data is that several facial features such as the ears and nose continue to grow throughout life. This has implications with data integrity because the spatial objects representing the human nose stored in the database can become invalid over time. This can also happen with Method A if images are not updated regularly to reflect current community opinion. If a new image is introduced into the image data set, it may be possible to determine a method to normalize the value of the new nose length based on other nose length values with similar characteristics or age to determine how it would be initially classified.

There is also concern in both cases that the boundaries initially selected are correct. In this study, boundaries were selected to be approximately three equal ranges. However, it may be that the classifications such as short and long are smaller ranges; it may be 20% are short, 20% long, and 60% fall into the average category. From the studies measuring nasal features from birth to 97 years old, boundaries may also need to be separate for female and males. Since males, on average, have larger nasal lengths than females, a better classification may be possible. Boundary values could also be established through statistics of previous or new measurement studies. Machine learning could be used to help analyze relationships and separation of the image data.

All three updating methods showed similar trend behavior, but which out of the three methods should be used? Overall, the steplock method seemed to be a better method because it is less prone to the shifting flaw of Method A. As mentioned earlier, if 300 users provided feedback that a nose was 'small' and 30 users indicated that the same nose as 'long' in method A, the nose classification changed to long and thus does not adequately reflect the greater community opinion. The steplock method proportionally provides greater weight to the 300 users and makes it harder for opinion to change if it is consistently voted.

CHAPTER 5 FUTURE WORK

The ability to search and analyze specific facial features has far reaching implications. Our Database Research Group is initially applying this effort to criminal image searching, but it may equally apply to other fields. Physicians, for example, could use this to aid in the diagnosis of dysmorphic syndromes or defects that are ordinarily difficult to diagnosis from appearance. Plastic surgeons could use this for pre and post-surgery comparisons to measure the overall success of the procedure. However, these potential applications will require additional research to become a reality.

Effort can be put into researching the analysis of other facial features besides nasal length such as mouth width or forehead height. One of the advantages of using spatial objects is that they can be dynamically exploited to provide measurements such as perimeter and area. This is due to the entire geometric object being stored in the database. These values and the relationships between other facial features could provide valuable information for a variety of fields. One of the fields heavily studied that examines facial attribute relationships is human attraction and beauty studies. Relationships could be derived that a broad chin affects human perception of nasal length or that a human nose above 5.2cm is considered unattractive through subjective testing.

Boundary values used in the study should be refined through statistical analysis or machine learning. Improved boundaries points could better reflect community consensus and strengthen the ability of the overall system. The steplock method also has values that can be adjusted in future studies in order to expedite achievement of consensus. Example values include the number of steps added to the slope between classifications and the number of steps removed when users vote against the image's current category. Likewise, the size of the step per vote can be adjusted; initially it was set to 0.03 for traditional and 0.05 for steplock and control. These values were initially selected to provide an equal amount of steps per classification range. A larger or smaller set size could yield different results.

Updating methods besides the steplock method and the traditional method may be tested and compared to better understand any potential advantages or disadvantages between the methods. An algorithm involving weighted averages, currently being research by VCU's Database Research Group, is one such method.

Overall, the issue that affected the overall conclusion of this study the most is the number of votes received for each image. A second study could be conducted with a larger number of participants to validate and reinforce conclusions that spatial data is not explicitly needed. This situation could be solved through additional testing sessions or through rewriting the evaluation application with a web programming language such as ASP.NET.

This would allow for a larger number of individuals to take the test at any time, but doesn't allow participants to clarify session instructions throughout the test .

Literature Cited

Literature Cited

[1] Crouse, B., and Pandrinath, A., (2004). Spatial Databases: A Directed Research Project. Virginia Commonwealth University.

[2] Shekhar, S., and Chawla, S., (2003). Spatial Databases: A Tour, Prentice Hall.

[3] Scognamillo, R., Rhodes, G., Morrone, C., (2003). A feature-based model of symmetry detection. The Royal Society. Volume 270, 1727-1733.

[4] Patnik, V., Rajan, S., Sanju, B. (2003). Anatomy of A Beautiful Face & Smile. J. Anat. Soc. India 51(1) 3-5.

[5] Web. The Rhinoplasty Center, Anatomy of the Nose. http://www.therhinoplastycenter.com/index.html. (2005.09.14)

[6] Web. Oracle 10g Documentation. Spatial Databases. http://www.oracle.com/technology/products/spatial/pdf/qt.pdf (2005.09.14)

[7] Zhang, et al., (2004). Comparison Between Geometry-Based and Gabor-Wavelets-Based Facial Expression Recognition Using Multi-Layer Perception. Third IEEE International Conference on Automatic Face and Gesture Recognition, April 14-16 1998, Nara Japan, IEEE Computer Society, pp. 454-459.

[8] Web. Learn to draw, Introduction to drawing people. http://www.learn-to-draw.com/drawing-people/08-eyeline.htm (4/5/2005)

[9] Web. 3D Faces and Face proportions. http://www.sierravista.wuhsd.k12.ca.us/basicart/faces.htm

[10] OpenGIS Consortium and Feature Specification. www.opengis.org (2004.11.08).

[11] Web. Oracle Spatial Users Guide and Reference. http://www.oracle.com/technology/products/spatial/index.html

[12] Web. Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Fourth_dimension

[13] Web. Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Tesseract

[14] Lachhwani, V. and Kim, Eok. GIS as a Tool for Facial Recognition www.uwm.edu/Dept/GIS/flyers/vikascompetition04.doc

[15] Web. Dictionary. www.dictionary.com

[16] Web. 3 Faces and Face Proportion http://www.sierravista.wuhsd.k12.ca.us/basicart/faces.htm

[17] Web. Drawing the Portrait. http://www.portrait-artist.org/face/

[18] Web. Human Lips. http://www.lipaugmentation.com/lipauggraphics/myologyfibers_Orbicularisoris.gif

[19] Date, C.J. (2004). Spatial Databases. An Introduction to Database Systems. 8th ed. Page 875-876.

[20] Web. MySQL 5.0 manual (2005) http://www.browardphp.com/mysql_manual_en/manual_Introduction.html

[21] Lee, Kuo, Hsu et al. (2004). 3-D Face recognition system based on feature analysis and support vector machine, IEEE.

[22] Balasuriya, L. and Kodikara, N., (2001). Frontal View Human Face Detection and Recognition.

[23] Web. Spatial Query Server 4.3 Manual.

[24] Sybase. Spatial Query Server. http://sismissionsystems.boeing.com/support/downloads/sqs34212SunSyb12_0.tar

[25] Joy, K. and Dattari, S. Implementing a Fuzzy Relational Database Using Community Defined Membership Values, Proceedings of the 2005 ACM SE Conference, Vol. 1, pages 264-26

[26] Crabtree, P., Fuzzy Relational Database and Querying System With Compound Query Capabilities (2005). A Directed Research Project. Virginia Commonwealth University.

[27] Deutsch CK, D'Agostino RB Jr. Statistical Appendix. In: Anthropometry of the Face and Head in Medicine, second edition, Farkas LG, ed., Raven Press, New York, September 1994.

[28] Feingold, Murry and Bossert, William. Normal Values for selected physical parameters: An Aid to Syndrome Delineation. Birth Defects: Original Article Series. Vol. X No. 13, 1974.

[29] Cunningham, M. R. (1986). Measuring the physical in physical attractiveness: quasi experiments on the sociobiology of female facial beauty. /Journal of Personality and Social Psychology, 50/(5), 925-935.

[30] Zankl, Andreas. Eberle, Lukas. Molinari, Luciano, and Albert Schinzel. Growth Charts for Nose Length, Nasal Protrusion, and Philtrum Length From Birth to 97 Years. American Journal of Medical Genetics 111:388-391 (2002).

[31] Mastros, James L. Spatial Database Feasibility for Facial Characterization using Fuzzy Logic Queries. A Directed Research Project. Virginia Commonwealth University.

APPENDIX A

high	low	modifier	threshold
3.85	3.4	short	3.6
4.39	3.86	average	4.1
5.2	4.4	long	4.7

Method B Range Values

high	low	modifier	threshold
3.85	3.4	short	3.6
4.39	3.86	average	4.1
5.2	4.4	long	4.7

Method C Range Values

high	low	modifier	threshold
0.33	0	short	0.2
0.66	0.34	average	0.5
1	0.67	long	0.8

Method A Range Values

APPENDIX B

ID	PLength
5	3.85
6	4.35
10	4.25
11	4.05
12	3.9
17	3.85
18	3.85
20	3.75
21	3.65
22	4.05
27	3.4
29	3.9
30	3.4
38	3.95
39	3.5

Method C - Final nasal length values

ID	PLength	NumStepsS_to_A	NumStepsA_to_L	NumVotes_Current	NumVotes_Total
5	4.1	22	0	27	34
6	4.4	26	0	29	35
10	4.4	1	0	3	34
11	4	0	0	4	34
12	4.45	9	0	16	34
17	4	27	0	31	35
18	4.1	23	0	29	35
20	4.65	0	0	12	35
21	3.8	0	0	4	29
22	4.4	10	0	17	31
27	3.6	20	0	26	32
29	4.05	6	0	17	29
30	3.5	10	0	21	32
38	4.25	8	0	15	23
39	3.8	9	0	20	33

Method B - Final nasal length values

ID	Weight
5	0.56
6	0.8
10	0.77
11	0.58
12	0.76
17	0.57
18	0.53
20	0.56
21	0.49
22	0.8
27	0.29
29	0.6
30	0.27
38	0.5
39	0.56

Method A - Final nasal length values

APPENDIX C

ImageID	method	Count of method
10	control	43
	steplock	39
	traditional	39
	Total	121
11	control	45
	steplock	37
	traditional	39
	Total	121
12	control	41
	steplock	38
	traditional	38
	Total	117
17	control	46
	steplock	36
	traditional	39
	Total	121
18	control	43
	steplock	36
	traditional	39
	Total	118
20	control	36
	steplock	35
	traditional	35
	Total	106
21	control	38
	steplock	39
	traditional	39
	Total	116
22	control	37
	steplock	38
	traditional	35
	Total	110
27	control	35
	steplock	35
	traditional	40
	Total	110
29	control	44
	steplock	36
	traditional	42
	Total	122
30	control	35
	steplock	35
	traditional	46
	Total	116
38	control	40
	steplock	36
	traditional	39
	Total	115
39	control	35
	steplock	35
	traditional	38
	Total	108

- Count of user feedback on nose length per image -