QUALITY CONTROL SYSTEM FOR A BAKERY PRODUCT USING FUZZY LOGIC TO PREDICT CONSUMER PREFERENCES

A Thesis

Presented to

The Faculty of Graduate Studies

of

The University of Guelph

by

TERRENCE B. CHU

In partial fulfillment of requirements

for the degree of

Master of Science

May, 1999

0-612-43151-7

ABSTRACT

QUALITY CONTROL SYSTEM FOR A BAKERY PRODUCT USING FUZZY LOGIC TO PREDICT CONSUMER PREFERENCES

Terrence Chu
University of Guelph, 1999

Advisors:
Professor V. Davidson
Professor O. Basir

Food manufacturing is an important business worldwide. Producers are turning to automation to improve the rate and the quality of production, especially in the area of quality assurance inspection. The objectives for this research were to identify the behavior of the consumers in judging chocolate chip cookies, to develop a consumer model based on the judgment data, and to gauge the success of a consumer-behavior based automated inspection system.

Consumers rated chocolate chip cookies based on dough lightness, size, and percentage of chips visible. Furthermore, interactions between dough lightness and size, and percentage of chips and size were also significant influences on decision making. The automated inspection system developed for this research used fuzzy logic to successfully model consumer behavior as identified through integration theory. The system correctly classified the acceptability of eighty percent of the cookies tested.

Acknowledgements

I would like to thank the following people for their invaluable help throughout this thesis. I would not have succeeded without their support. I would like to extend my thanks to:

- My mother and father for their unconditional love and encouragement.
- Professor Davidson and Professor Basir whose guidance has kept the light at the end of the tunnel lit.
- Dale Scott, and Chris Semotok for helping me stay sane in the sunless computer lab.
- Joanne Rykes, Karen Conrad, and Allen Miller for their expertise and 'get it done' attitude.
- All of my family and friends for keeping me motivated.
- All the participants in the polls for their superior judgement of chocolate chip cookies.

Table Of Contents

List of Tables

List of Figures

1. Introduction

Food manufacturing is a major industry worldwide. For many food manufacturers the global community has become accessible due to technological advances in transportation and communication. Competition within the food industry has grown because of these advances. To remain competitive, manufacturers are enhancing production and processing techniques by introducing automated control and inspection as solutions to increase production, and improve quality (Sarkar, 1991).

In particular, manufacturers of baked goods have traditionally used manual inspection to perform quality assurance. Manual inspection introduces bottlenecks into the production process. Operators cannot work as fast as the machinery generating the products. More importantly, manual inspection relies solely on the inspectors' decision making. Unfortunately, the judgment of these inspectors may not reflect consumer opinions. For this reason, the process by which consumers judge food is of particular interest.

With an understanding of consumer behavior, a model that reproduces consumer decision-making can be developed. Quality assurance inspection can be automated by integrating this model with a machine vision system. An automated solution such as this improves production by allowing the manufacturer to assess product quality according to consumer opinion. Ultimately, the model facilitates objective and consistent consumer-based decision making during quality assurance.

Consumer behavior in product judgement is studied and quantified through psychophysics. One of the goals of psychophysics is to develop methodologies through which consumer judgement can be modeled. Specifically, through integration theory, integration and scaling have been identified as concepts that describe the mechanics of decision making (Lawless, 1990; Mcbride and Anderson, 1990).

Integration refers to the process by which people combine different food characteristics to render a judgement pertaining to the quality of the food. Characteristics such as size, shape, texture, color, taste, and odor are combined to produce an opinion about a food product. Food judgement is also a function of scaling. Scaling quantifies the influence each characteristic exerts on the final decision. In decision making, a significant characteristic is given more weight than one that is less significant.

Human decision making patterns are identified and modeled through integration theory. This model can be automated using fuzzy logic and then used to predict consumer behavior. In other words, integration theory can be translated into a fuzzy logic decision engine. Not all consumers will agree on the ideal combination of product characteristics. Fuzzy logic can account for this inconsistency and is a good candidate for consumer modeling because it can tolerate uncertainty and vagueness.

Machine vision is a general term that describes a system that uses both image processing and a decision engine to perform a duty. In a machine vision system, image processing identifies and quantifies characteristics such as the size and color of a product. Digitally captured images of this product are analyzed through techniques such as edge detection, perimeter following, and color analysis. These operations yield information including the shape, size, and color composition of the product. A consumer model can use these extracted features to assess this product's quality. This consumer model can be used in conjunction with image processing to produce an automated quality assurance inspection station.

The objectives of this thesis are to:

- Describe consumer behavior in judging chocolate chip cookies
- Develop an algorithmic model of the consumer behavior through fuzzy logic
- Implement and evaluate the success of an automated inspection system for chocolate chip cookies based on the consumer model

To accomplish these objectives, chocolate chip cookies with different physical characteristics (e.g. color, size, and shape) were presented to a group of subjects. These people represented a cross section of the consumer population. Each cookie's relative consumer appeal was expressed as a position along a line scale. The cookies' ratings were recorded as distances from the line scale's origin. Statistical analyses of the recorded distances provided insight into scaling and integration of the characteristics presented during judgement. Based on the statistical findings, a consumer decision model was developed and then tested against subsequent polls.

This thesis will present literature relevant to consumer modeling and automating quality assurance. The materials and methods employed in the study are described and explained. Then, experimental results are presented and analyzed. Finally, conclusions are drawn from the results, and recommendations for further study and improvements are presented.

2. Literature Review

Quality assurance is an important part of food manufacturing. Manufacturers such as commercial bakeries or fisheries require quality control personnel to identify and remove damaged or poor quality product. When judging the quality of a product, these inspectors identify defects and then classify the product's quality. This chapter will focus on the techniques and technologies that are required to automate defect identification and product classification.

In this literature review the characteristics of consumer behavior and the experimental techniques used to identify this behavior are discussed. Research into the effect of physical attributes on taste and apparent quality are presented. Then, basic elements in fuzzy logic are explained. The review continues with an examination of image-processing techniques that allow computers to quantify an object's characteristics. Research in combining image processing and artificial intelligence to automate food quality assurance are presented. Finally, deficiencies in the literature in tying consumer opinion to quality assurance are identified.

2.1. Consumer Behavior

When a food product is first developed, a consumer market is targeted and the product is developed to appeal to those consumers. Sensory evaluation helps the manufacturer to design and to control the quality of the product. But the sentient nature of perception and judgement makes understanding human behavior difficult. Through psychophysics and information integration theory, several stages have been developed to explain the decision making process (McBride and Anderson, 1990).

Valuation is the first stage in the decision making process. During this stage the physical stimuli of a food product are translated into psychological sensations. The stimuli are assembled and then scaled relative to one another according to importance. These stimuli include visual characteristics, fragrance, texture, taste, and sound.

The second stage of decision making is integration. The sensations identified during valuation are integrated into a single sensation. Other factors such as memory or context effects are also integrated into this sensation (Lawless, 1990). The sensation generated by integration is not observable, but it exists in integration theory as an intermediary stage between the observable stimuli and the response.

Finally, during the response stage an assessment of the quality of the food product is generated. The person identifies the acceptability or appeal of the food product. Through these three stages in integration theory, the decision making process can be systematically quantified and then modeled.

2.1.1. Types of integration

Integration theory provides us with the means to map stimuli to the physical responses. Through integration theory, it is possible to develop an understanding of consumer behavior. Statistical and graphical analysis of stimuli versus response are used to identify these relationships. Additive interaction between stimuli is a common relationship. An example of an additive relationship is the taste-odor integration between sweetness and orange aroma as presented by McBride and Anderson (1990). The total intensity of sensation was plotted versus nine combinations of taste and odor. The experiment produced the relationship shown in Figure 2.1. The additive response can be seen by the parallelism demonstrated by the three levels of sucrose and of orange aroma. The difference in response to each change in stimulus is consistent. This parallelism indicates an additive integration of the stimuli (McBride and Anderson, 1990).

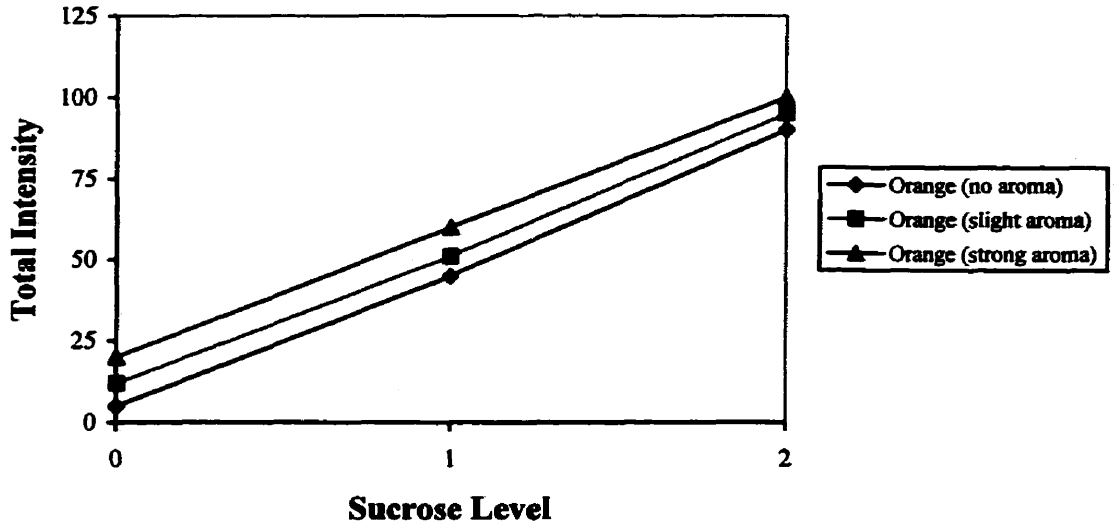


Figure 2.1: Total Intensity of Sensation versus Orange and Sucrose Levels

Other relationships that describe consumer behavior include subtractive, multiplicative, and nonlinear. A full description of each type of integration relationship is beyond the scope of this literature review but a more complete description can be found in McBride and Anderson (1990) and Lawless (1990).

2.1.2. Experimental Techniques in Identifying Consumer Behavior

Through integration theory, the multidimensional character of a food product can be examined. The influence of each characteristic on consumer appeal can be determined, and the ideal value or range for each characteristic can be identified. But, appropriate testing procedures need to be performed to acquire data that can be used for consumer judgment analysis and, ultimately, consumer modeling. In order to elicit a proper response to the food product, two procedures must be performed. Firstly, the food characteristics significant in decision making must be identified. Secondly, an appropriate test to record the consumer response to the product must be designed using those characteristics.

2.1.3. Identifying Significant Stimuli

The significant stimuli used for decision making can be identified through the repertory grid procedure (McEwan and Thomson, 1988). During this testing three examples are presented and the test subject chooses an outlier and then justifies the choice. These trials are repeated for all the characteristics that are perceivable in the food. Through this testing, the most influential characteristics, as well as the ranges of noticeable difference for each characteristic can be identified (McEwan and Thomson 1988).

For example, the dominant characteristic of an orange is the color. Different colors may generate different acceptability ratings and thus should be identified. For oranges the ranges may be green, green-orange, orange, and brown. By determining that the color is important and by identifying the different color ranges that are noticeable to the consumer, a test for the acceptability of an orange based on color can be created.

2.1.4. Testing for consumer appeal

To properly test a product, a complete set of characteristics and range of values within each characteristic need to be presented to the consumer. This set is generated using the characteristics and ranges identified through the repertory grid procedure.

Through studies in consumer behavior, and specifically in product testing, line scale tests used in conjunction with product rating methods have been shown to be the most successful in generating useful characteristic relationship information (McBride and Anderson, 1990).

The graphic line scale is essentially a line on a flat surface with markers along the line to anchor the test. Test subjects are then asked to place test samples of the product where they feel the sample fits best. For example, in a saltiness flavor test there would be a marker that denotes the saltiest example of the product, and one that denotes a version of the product with no salt. Testers may be asked to rate the saltiness of different samples of the product relative to these markers. In this test, different recipes can be tested for their apparent saltiness.

Another line test may be based on linguistic terms. Instead of using product samples as anchors, linguistic terms that represent concepts in human judgment may be used. This type of test is called an interval scale (Land and Shepherd, 1984). For example, an interval scale may use the concepts unacceptable and acceptable as anchors. In this case, the observers are asked to rate products in terms of their appeal. Presenting a complete set of characteristics for judgement can identify the qualities that influence a product's consumer appeal. The results of these tests can be analyzed through statistical means such as mean, standard deviation, and analysis of variance to identify consumer behavior (Land and Shepherd, 1984).

2.1.5. Research into the Effect of Appearance on Product Appeal

Christensen (1983) performed research on the effect of food color on perceived intensity, and quality of aroma and flavor. In the study, a group of testers were presented with pairs of normally colored and inappropriately colored food wherein each sample was rated for the quality of its aroma, flavor and texture. The foods tested included margarine, gelatin, bacon strips, an orange drink, and cheese. The study found that appropriately colored food was perceived as having more intense and superior aroma and flavor. Two groups of testers judged the pairs of foods. One group was blindfolded while the other was not. The sighted subjects judged the appropriately colored food as having a stronger aroma while the blindfolded group generated no response pattern. The research surmised that color influenced the anticipated oral and olfactory sensations because of the memory of previous eating experiences.

Dubose *et al.*, (1980) found that as the intensity of the color of the food increased, the perceived flavor increased accordingly. Furthermore, the experiment showed that specific colors induced flavors typically associated with those colors. The food stuffs studied included fruit flavored beverages and cake. For example, flavorless cakes were deemed to have increasingly acceptable lemon flavor when more yellow colorant was added to the cake. Ultimately, Dubose *et al.*, (1980) showed that memory effects are associated with visual perception and affect the judgment of consumers.

The research done by Dubose *et al.*, (1980) and Christensen (1983) showed that visual stimulus is a major factor that influences the consumer judgement of food. The physical attributes of a food product generate memories of past eating experiences. Ultimately, the overall quality of a food product can assessed based on visual features alone. Therefore, a consumer model that can successfully relate the visual features of a food product to the food's quality will provide a realistic assessment of consumer opinions.

2.2. Fuzzy Logic

Fuzzy logic is a diverse field of mathematics that has been applied in many areas such as control systems and decision support systems. Fuzzy inference systems are appropriate when the uncertainty of classification needs to be assessed, when the patterns for each class are ambiguous, and when it is difficult to define boundaries between classes (Cai and Kwan, 1998). In each circumstance, the domain of application has dictated the type of fuzzy implementation that is most appropriate. For example, fuzzy control systems are often developed with the help of an expert in the control environment. In this situation a rule-based fuzzy logic model of the expert is often used. On the other hand methods such as fuzzy-c means (Bezdek, 1981), or even fuzzy neural networks (Kuo *et al.*, 1993) are employed in circumstances where a fuzzy classifier needs to be built on ill-defined data. For example automated visual signature verification is an area where the data is ill defined. Each individual possesses unique hand writing characteristics. Common features or invariants between different individuals' signatures are hard to find. Furthermore, there is significant variation in signatures generated by the same individual. These challenges conspire to make personal signatures difficult to separate (Scott, 1998).

The focus in this section will be on fuzzy rule-based inference models. These models are appropriate for mimicking the rules governing consumer behavior as identified through integration theory. Fuzzy rule-based inference models generally consist of three processes, fuzzification, inference operations, and defuzzification (Cai and Kwan, 1998). During fuzzification crisp inputs are translated into the fuzzy domain through membership functions. Each crisp value is assigned a membership value to a fuzzy concept as defined by the membership functions. For example, the comfort level of a room according to an average person may be described by the membership functions shown in Figure 2.2

Ideally, people prefer the room to be warm. Temperatures below 16° C may seem cool to most people, while temperatures above 24° C may seem hot to many people. Rooms at 23° C are neither warm nor hot but somewhere in between. Fuzzification of a 23° C temperature would produce truth-values equal to 0.4 warm and 0.6 hot. After fuzzification the room is no longer described as a crisp temperature of 23°C, rather the room is described as being 0.4 warm and 0.6 hot.

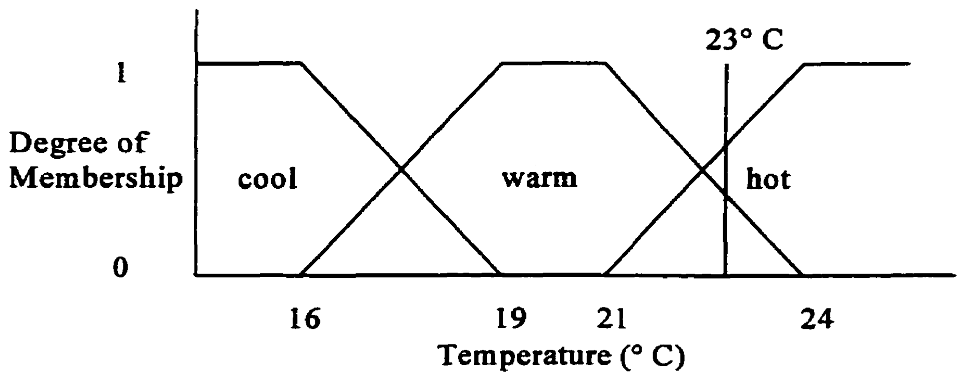


Figure 2.2: Fuzzy Membership Functions for the Temperature of a Room

These fuzzy values are then passed to "if...then" rules which combine and manipulate the variables to produce one or more fuzzy outputs. These rules represent the cognitive process that people perform when making decisions. Possible rules for a temperature controller using the ventilation fan-speed as the manipulated variable are shown in Figure 2.3.

• If temperature is **cool**	then fan-speed is **Slower**
• If temperature is **warm**	then fan-speed is **Unchanged**
• If temperature is **hot**	then fan-speed is **Faster**

Figure 2.3: Inference Rules for Fuzzy Temperature Controller

The antecedent in the rules is the temperature. The consequent, or the fuzzy output, is the change in the speed of the ventilation fan. The consequent is clipped to the same degree of membership as the antecedent. For a room temperature of 23° C these inference rules recommend fuzzy changes in fan speed as 0.4 unchanged and 0.6 faster.

Finally, defuzzification produces crisp or linguistic outputs from the fuzzy-consequences. In this case the crisp output is the change in revolutions per minute of the fan. Membership functions are used to map fuzzy consequences back to the crisp domain. The following example demonstrates a Sugeno-type defuzzification method that can be found in Terano *et al.* (1987). The membership functions for this controller are called singletons and are shown in Figure 2.4. The aggregation formula that generates the crisp change in fan speed can be seen in equation (2.1). The defuzzified fan speed is equal to +60 rpm.

$$\Delta_{\text{Fan Speed}} = \frac{\mu_{\text{Slower}} * \gamma_{\text{SlowerRPM}} + \mu_{\text{Unchanged}} * \gamma_{\text{UnchangedRPM}} + \mu_{\text{Faster}} * \gamma_{\text{FasterRPM}}}{\mu_{\text{Slower}} + \mu_{\text{Unchanged}} + \mu_{\text{Faster}}} \quad (2.1)$$

Where:

$\Delta_{\text{Fan Speed}}$ = the defuzzified change in fan speed in rpm
μ_{Slower} = the degree of membership in "Slower"
$\mu_{\text{Unchanged}}$ = the degree of membership in "Unchanged"
μ_{Faster} = the degree of membership in "Faster"
$\gamma_{\text{SlowerRPM}}$ = the defuzz. singleton rpm for Slower
$\gamma_{\text{UnchangedRPM}}$ = the defuzz. singleton rpm for "Unchanged"
$\gamma_{\text{FasterRPM}}$ = the defuzz. singleton rpm for "Faster"

This example has demonstrated a simple implementation of a min-max fuzzy logic temperature controller that manipulates fan speed using human decision making as a model. There are many ways to describe input and output membership functions. There are also many ways to infer consequences and to aggregate these consequences. These concepts in fuzzy logic theory, and their application, have already been studied in detail and can be found in Terano *et al.* (1987), Zimmerman (1996), De Silva (1995), Yen *et al.* (1995), and Pedrycz (1989).

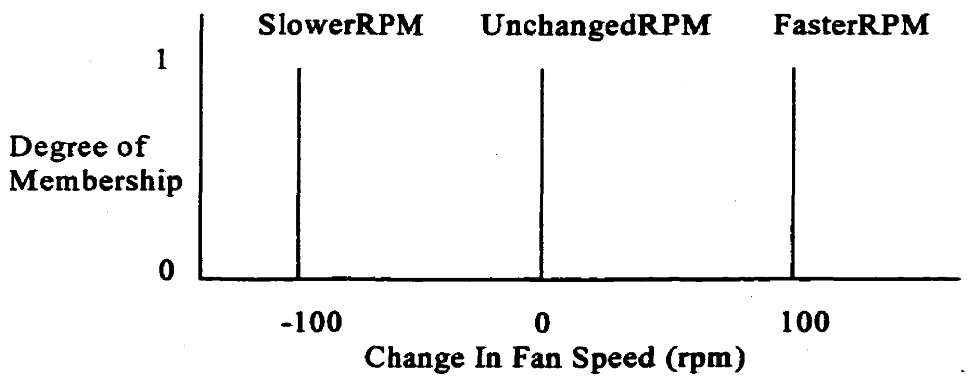


Figure 2.4: Defuzzification Membership Functions for Fuzzy Temperature Controller

2.3. Machine Vision Techniques

Machine vision is a general term that describes a system that processes images and then renders a decision about the extracted information. Essentially, machine vision combines both image processing and a decision engine to perform a duty. The role of image processing is to provide the necessary information for the decision engine to render a judgement.

Image processing refers to the quantification of information stored in a digital image. The field of image processing is very broad and uses many different techniques to extract and identify information in an image. These techniques include mathematical methods such as Fourier analysis to extract image information, as well as artificial intelligence techniques such as neural nets to categorize the extracted information. This section will focus on methods for enhancing and extracting object features such as color composition, shape, and size. Image processing techniques such as these will quantify an object's features in such a manner that they can be used by a decision engine.

Image processing can be categorized into several levels of complexity. Basic operators that are used include contrast enhancement, edge detection, and transforming color-spaces. These operators reduce noise in the image and identify the information found in the image without consideration for the actual significance of the features. Essentially these operators perform operations on properties that are consistent across all images. Gunasekaran (1996) refers to this level of image processing as preprocessing. In effect, the image is being conditioned so that application specific information can be extracted from the image.

Contrast enhancement and edge detection are but two examples of low level image manipulation. Other operations include smoothing, converting to grayscale, and brightening. These operators can be used to accentuate information in the image in preparation for a more detailed analysis of the image.

Colors in digital images are represented by three components, namely, red, green, and blue. The standard RGB color representation in computers is difficult to manipulate for color recognition but can be transformed to more useable representations of color (Thomas and Connoly, 1986). The L*a*b* color space provides a baseline for describing colors and has been used for measuring pigmented materials such as those of the food industry (Volz, 1995). This color space provides a better description of color by separating lightness and hue. In the RGB color space, the lightness and hue are coupled to all three components.

The next stage in image processing is referred to as segmentation, representation and description (Gunasekaran, 1996). These operations include perimeter following, area classification, and color analysis. The information extracted by image processing operators identifies an object in the image through shape, size, and color composition. Ultimately, image processing extracts features that identify an object in a manner that can be manipulated by a decision engine.

2.4. Automated Inspection of Food Products

Food manufacturers understand the value of automated inspection over purely manual inspection. Research into automating basic processing tasks has produced systems that perform both image processing and classification with varying degrees of success. These works exemplify image processing and decision engine techniques that can be used for inspection research. Research has been performed to assess food quality using achromatic, or grayscale images. Unklesbay *et al.* (1983) assessed the nutritional value of pizza shells versus the browning of the shells. Pizza shells were baked, scanned by a monochrome camera, and then assessed for their nutritional value. The experiment successfully linked degree of browning to lysine content in the pizza shell via a linear mathematical model. The experiment showed that dough lightness could be used to identify nutritional value. Some inspection problems need more detail about the food to facilitate classification thus, as color digital camera technology becomes more commonplace, color features are being put to use in inspection research. McConnell *et al.* (1995) used color classification to classify baked and roasted food products. This research assessed the doneness of a food sample relative to a several reference examples of the product. The experiment demonstrated that many food products have multiple color attributes that need to be considered in inspection and judgment. The system employed a histogram analysis that was similar to a Bayesian maximum likelihood classification technique. Daley *et al.* (1995) studied poultry processing as a potential candidate for automated inspection. Color histogram analysis, pixel-level color categorization, and a mathematical filter were employed to identify defects. The goal of the experiment was to identify defects such as bruises, skin tears, and tumors. Ultimately, bruises and tumors were successfully identified by single pixel color analysis through neural networks, while skin tears were only identifiable through a mathematical filter.

Bell peppers have been sorted based on color (Shearer *et al.* 1990). Tao *et al.* (1995) sorted apples and potatoes through color analysis. Both studies used discriminant techniques to identify and classify the food. Other foods such as soybeans (Wigger *et al.* 1988), and peaches (Miller and Delwiche, 1989), have also been sorted based on color using basic mathematical techniques.

Luzuriaga *et al.* (1997) graded shrimp using color. The shrimp were differentiated based on amount of surface area demonstrating melanosis. A grade was assigned to the shrimp using a linear model that related the percentage of area exhibiting melanosis to the operator assigned shrimp grade.

Perrot *et al.* (1996) classified the color of cookies. Cookie color readings were captured and a fuzzy logic decision engine mapped the three dimensional numerical color reading to a linguistic color classification. The fuzzy engine modeled a human operator and was compared to a Bayesian classifier. The success rates for both types of classifiers were comparable, but the authors supported the fuzzy classifier because it more appropriately modeled the deterministic nature of human decision-making, i.e. decisions that are generated through rules. The fuzzy logic inference model is rule-based and properly represents the deterministic nature of the human decision making process.

Beyond color other physical features such as shape and size that can be used to identify an object. Howarth *et al.* (1992) quantified and classified carrot tip shapes using a Bayes discrimination procedure. The curvature profile was used to quantify the shape of the carrot tip. This profile was then passed to a discrimination procedure for classification. Heinmann *et al.* (1996) developed a machine vision system that graded potatoes based on size and shape. No statistical or intelligent techniques were used to grade the potatoes' size, or shape.

Ding and Gunasekaran (1994) used shape indices to describe food products including corn, crackers, and almonds. The system identified the quality of the shape of the product by comparing the sample to templates. Ultimately, the system identified misshapen or damaged products. The experiment successfully used both back-propagation neural networks and minimum indeterminate zone classifiers to classify the food products.

De Silva (1997) used physical features such as location of cut, length, smoothness of contour, and surface texture to classify the quality of a processed fish. These inputs were manipulated through fuzzy logic to generate accept or reject decisions. De Silva (1997) also presented an experiment where skeins of herring roe were graded based on visually quantified features including shape, size, and weight. The grading also accounted for the roe's ultrasound scanned firmness. The roe grading algorithm, based on fuzzy logic, facilitated size and grade classification based on the four identified features. In both experiments, the fuzzy logic algorithm modeled the expert online inspector for decision making.

The bodies of work discussed in this section show that the physical features foods can be identified and used to assess food quality. Furthermore, the papers show that the assessment of the foods can become complex, and that simple methods of relating visual attributes to food quality are needed. A fuzzy logic based consumer model can provide this simple means of assessing food quality.

2.5. Literature Deficiencies

The literature presented illustrates that there are several areas in automated inspection research that requires more study. Firstly, the classifiers in research are not informative about the product's quality according to the consumer. The relationship between product characteristics and consumer satisfaction should be investigated and applied to a classifier. Secondly, the applicability of fuzzy logic for modeling integration theory should be assessed. Finally, the applicability of fuzzy logic in automated inspection should be investigated further. Specifically, the ability of fuzzy logic to accept multiple mathematically unrelated inputs and generate multiple outputs should be investigated in an automated inspection application. Product quality and consumer behavior is seldom one dimensional therefore their multi-dimensional characters should be incorporated into the decision-making algorithms employed by inspection systems.

3. Methodology

3.1. Feature Extraction and Image Processing

The equipment used for capturing images was chosen for its functionality and flexibility. The list of equipment can be found in Table 3.1. The digital camera was chosen because it captured scenes more accurately than an average digital camera. It had 3 charge coupled device (ccd) sensors whereas most cameras only have one. Each ccd sensed red, green, and blue colors, respectively. The effective resolution at which the camera sensed the environment was triple that of most single ccd digital cameras. Many digital cameras have the red green and blue pixel sensors on the same ccd. These single ccd cameras do not reproduce the scene as accurately due to their lower effective sensing resolution.

The lens for the camera was chosen because it provided a flexible field of view and zoom. It was possible to adjust the amount of area sensed by the camera to suit different applications.

The grabber board was chosen because it was fully compatible with the camera and the video card in the computer. Furthermore, the grabber board was supplied with C language libraries that were used to interface with the camera.

For lighting, fluorescent lights were used because, relative to incandescent bulbs, the light produced was more consistent over the lifetime of the bulb. As well, the light produced was whiter. The selected lights provided a consistent environment and better color representation of objects for image capture.

Computer	Dell Dimension XPS M200s PC, Dell Computer Corporation, Round Rock, Texas, 78682.
	Microsoft Windows 95 service release version 2 and Microsoft Visual C++ 4.0, Microsoft, Redmond, WA, 98052-6399.
	Matrox Millenium video card
	Meteor RGB grabber board, Matrox Image Library Lite (MIL-Lite) 3.0, Forefront Graphics Corp, Downsview, ON, M3J 3K7.
Camera	JVC AA-P700 AC Adapter and JVC KY-F55BU 3ccd digital Camera, JVC Professional Products Co., Elmwood Park, NJ, 07407.
	Cosmicar/Pentax TV Zoom Lens, 8 mm - 48 mm, 1:1.0, Asahi Precision Co. Ltd., Wako-Shi, Saitama-Ken, 351-01.
	RGB cable
Calibration	Color Plates • White Minolta Chroma-Meter CR-300 Colorimeter calibration plate. • brown paint color sample squares.
Lighting	2 Noma work lights
	2 Sylvania 15 watt fluorescent bulbs, Danvers, MA, 01923.

Table 3.1 : Image Capture System

The physical setup of the system can be seen in Figure 3.1. The camera was supported on a stand that fixed the camera's position relative to the cookies. The lens was used to control the size of the field of view and light transmission. Images were captured using the program in appendix A. This program was developed in Microsoft C++ 4.0 and used the MIL-Lite 3.0 library to interface with the grabber board. The program displayed the field of view of the camera and allowed the user to capture the image displayed.

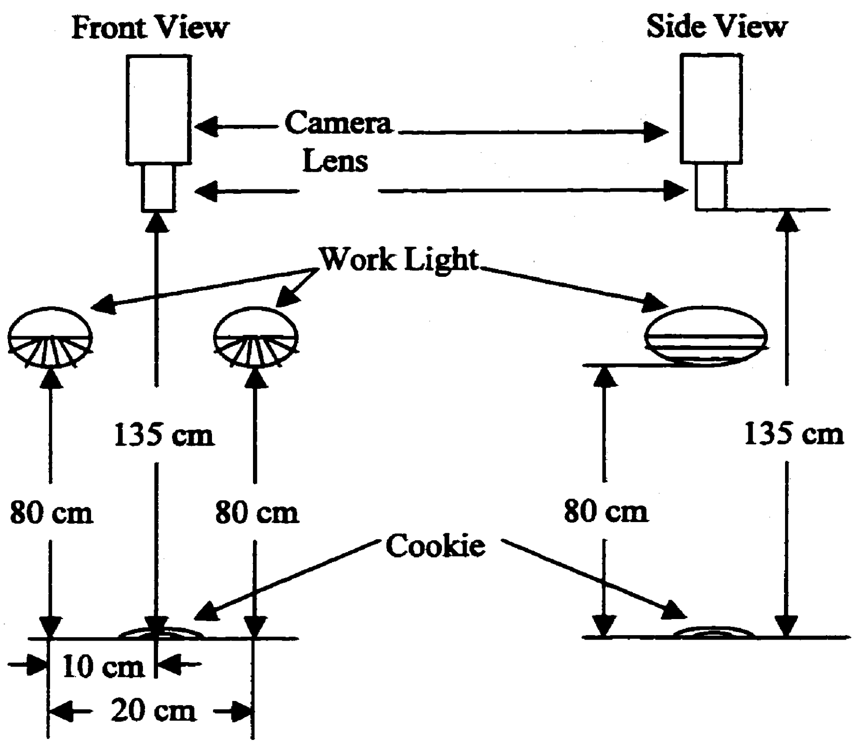


Figure 3.1 : Camera and Light Configuration

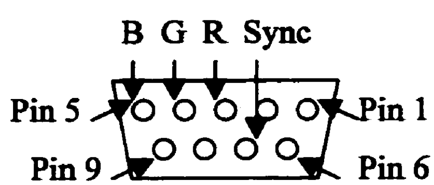


Figure 3.2 : Grabber Board, D-Sub Connector (9 pin female, viewed from front)

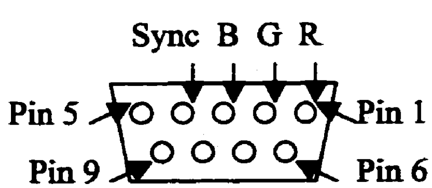


Figure 3.3 : Camera, D-Sub Connector (9 pin female, viewed from front)

An RGB Cable connected the camera to the grabber board in the computer. The configuration of the cable can be seen in Figure 3.2 and Figure 3.3.

The camera was aimed perpendicularly to the cookie surface to prevent from distorting the shape of the cookie. Direct lighting generated shadows around the edge of the cookie that interfered with the feature extraction. Therefore to minimize the shadows, the two lights were placed on both sides of the cookie and high as feasible.

The lights, camera and computer were warmed up for one hour before capturing images because the system's color sensitivity was not consistent until some time had passed. The source of this dynamic characteristic was not established.

The lights of the room were turned off and the work lights alone were used to illuminate the cookies. Size, focus, and sensitivity to light of the system were calibrated before each session of image capture. The field of view was adjusted so that 480 pixels represented 14.25 cm in the field of view, as shown in Figure 3.4. The size was calibrated by drawing two parallel lines exactly 14.25 cm apart on a sheet of paper. The field of view was adjusted so that the lines were at the edge of the visible field of view vertically. The focus was also adjusted so that the image sensed by the camera was sharp. Finally, the camera's sensitivity to lightness was calibrated by grabbing images of the four brown plates similar in hue to cookie dough. The plates' L* values were extracted and then compared to previously captured baseline images of the plates.

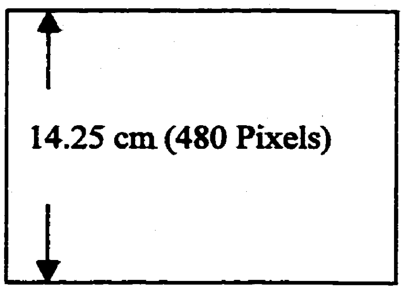


Figure 3.4 : Camera Field Of View Calibration

Images of the cookies were grabbed after the calibration. Each image contained only one cookie on a matte white background. Every cookie image was cropped and then passed to the feature extraction program for processing. The grabbed images were saved in "tiff" file format and were left uncompressed. Compressed images can generate color and pixel distortion and can be difficult to manipulate. A 24-bit color depth was used to provide a satisfactory color resolution for color analysis. This resolution provided approximately 16 million different colors.

Computer	SGI Indigo2 computer
	Irix 6.2 Operating System
	SGI Image Vision image processing toolkit
	C Visual Developer (CVD)
	cc (Unix ANSI C++ compiler)
Colorimeter	Minolta Chroma Meter CR-300 Colorimeter, Minolta Canada Inc., Mississauga, ON, L47-2H5

Table 3.2 : Image Processing System

The feature extractor was developed in C++ in a UNIX based environment and used the Image Vision library to facilitate the basic image processing functions. The program quantified cookie features from captured images including average diameter, major axis, minor axis, dough lightness excluding chocolate chips, and percentage of chocolate chips visible. The source code for the feature extractor can be found in appendix B.

The algorithms for identifying the average diameter as well as the major and minor axes required definitions of the perimeter and the center of gravity of the cookie. The areas in the image that were belonged to the cookie were used to calculate the center of gravity of the cookie. This calculation is shown in equations (3.1) and (3.2).

$$\text{Centroid}_x = \sum_{i=0}^{i=n} x_i / n \qquad (3.1)$$

$$\text{Centroid}_y = \sum_{i=0}^{i=n} y_i / n \qquad (3.2)$$

Where i = a pixel considered to be part of the cookie.
n = the number of pixels that are considered to be part of the cookie
x = x co-ordinate for pixel i
y = y co-ordinate for pixel i

A perimeter description of the cookie was generated and used to find the mean radius and the major and minor axes. The mean radius was found by taking the average linear distance between each perimeter pixel and the center of gravity. Another algorithm generated axes for the entire cookie. The longest axis was marked as the major axis, while the shortest axis was marked as the minor axis. The percentage difference between the major and minor axes was used to describe the shape of the cookie and was calculated using equation (3.3).

$$\% \text{ Difference Of Axis} = (A_{maj} - A_{min}) \Big/ A_{maj} \qquad (3.3)$$

Where A_{maj} = the length of the major axis
A_{min} = the length of the minor axis

Besides the linear dimensions of the cookie, the feature extractor also found the lightness of the dough. The lightness measure was an approximation of the L* value as measured by the Minolta colorimeter. The RGB values that described the image data were transformed first to CIE XYZ color space, and then to the L*a*b* space. The transformations from RGB to CIE XYZ are shown in equations (3.4), (3.5), and (3.6). The transformations from CIE XYZ to L*a*b* are shown in equations (3.7) through (3.12) (Ling *et al.*, 1996).

$$X = 0.4783*R + .2986*G + 0.1746*B \qquad (3.4)$$
$$Y = 0.30*R + 0.59*G + 0.11*B \qquad (3.5)$$
$$Z = 0.0197*R + .1601*G + 0.9077*B \qquad (3.6)$$

Where X = CIE X value; R = normalized Red value of pixel
Y = CIE Y value; G = normalized Green value of pixel
Z = CIE Z value; B = normalized Blue value of pixel

If $Y > 0.008856$ then
$$L^* = 116.0*Y^{1/3} - 16.0 \qquad (3.7)$$
If $Y <= 0.008856$ then
$$L^* = 903.3*Y \qquad (3.8)$$

$$a^* = 500*(F(X)-F(Y)) \qquad (3.9)$$
$$b^* = 200*(F(Y)-F(Z)) \qquad (3.10)$$

$$\text{If } \lambda > 0.008856 \text{ then } F(\lambda) = \lambda^{1/3} \qquad (3.11)$$
$$\text{If } \lambda <= 0.008856 \text{ then } F(\lambda) = 7.787*\lambda + 16/116 \qquad (3.12)$$

Where L* = L* value; λ = CIE X,Y, or Z value
a* = a* value
b* = b* value

The equations for RGB to CIE XYZ color transformations were based on Ling *et al.* (1996) but were modified to improve the color sensitivity of the system. The system was calibrated by adjusting the coefficients in equations (3.4) through (3.6) so that L* measurements sensed by the feature extractor agreed with the L* measurements sensed by a standard colorimeter.

Identifying the percentage of chocolate chips visible on the cookies was a two step process. First the average lightness of the non-chip area was calculated. To find the average dough lightness, the average L* value of all the pixels in the cookie with an L* value greater than 25 was calculated. Above this threshold every pixel was considered to be cookie dough. Then, using the average dough L* value, the dough-lightness-adjusted chip threshold was calculated using equation (3.13). All pixels below this threshold were considered to be chocolate chip. The calculation for finding the percentage of chocolate chips is shown in equation (3.14).

$$\text{Chip threshold} = 0.462(\lambda_{std} - 39.5)+16.0 \qquad (3.13)$$

Where λ_{std} = average L* value of cookie dough

$$\text{ChocolateChips} = (\text{TotalChipArea} / \text{TotalCookieArea})*100 \qquad (3.14)$$

Where TotalChipArea = total cookie area in cm^2 below chip-dough threshold
TotalCookieArea = total cookie area in cm^2

A sample set of cookies was baked in a conventional oven at 325°F several times and measured for dough lightness. With the chip/dough threshold adjustment, the Feature Extractor correctly identified higher percentages of chocolate chips as the cookies' dough became darker. The chip-threshold equation was found by changing the slope and intercept until the percentage chocolate chips sensed remained constant over all the bake sessions. A total of five bake sessions were used. The resulting range of dough lightnesses spanned the entire range of lightnesses used for polling.

3.2. Consumer Sensory Evaluation

The goal of the consumer evaluation was to identify and understand the behavior of the general population in judging chocolate chip cookies. Four characteristics were used to identify consumer behavior, including dough lightness, size, shape and amount of chocolate chips visible.

Class ranges within each characteristic were defined based on differentiability. Every attempt was made to separate classes as much as possible to provide the greatest observable difference between classes. At the same time, the range of values within each class was kept as narrow as possible so that cookies within each class appeared to be the same.

The variability of the cookies also guided the development of the classes. The cookies used for the experiment were produced at a commercial bakery. Only a small range of dough lightness was available because the baking process was fairly consistent. The size and shape of the cookie were also consistent because of the techniques used during production. The raw cookie dough was extruded through a die and then wire-cut. As a result, the raw cookie was a consistent flat disk. Inconsistencies in measuring and mixing the ingredients were minor and thus the variability of chips in the cookies was small. Compared to cookies that are hand made, the variability of the physical characteristics of the cookies used was low because the cookies were produced in a commercial bakery.

The dough lightness for these cookies was extended beyond normal production ranges manually to generate the desired observable differences between cookies. Dough lightness was manipulated by baking cookies to the required lightness. Acquiring different batches and brands of chocolate chip cookies also aided in extending the range of sizes. The fuzzy classes were defined by observable differences and are shown in Figure 3.5 through Figure 3.8.

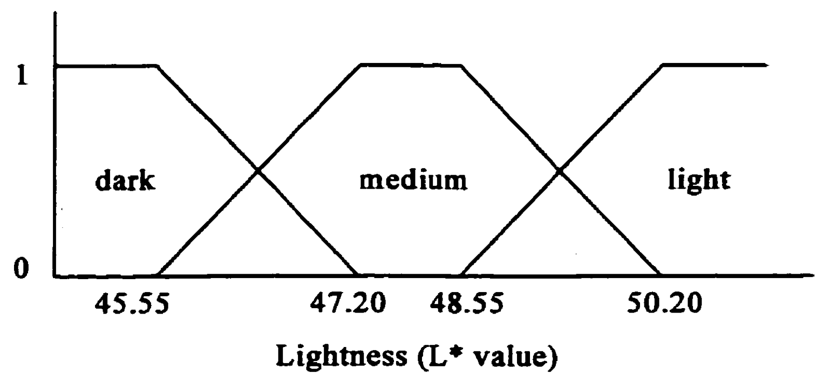


Figure 3.5 : Dough Lightness Membership Function

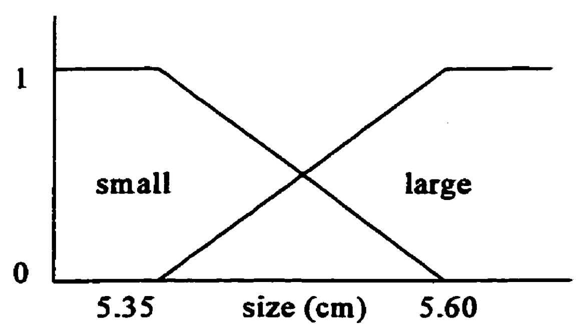


Figure 3.6 : Cookie Size Membership Function

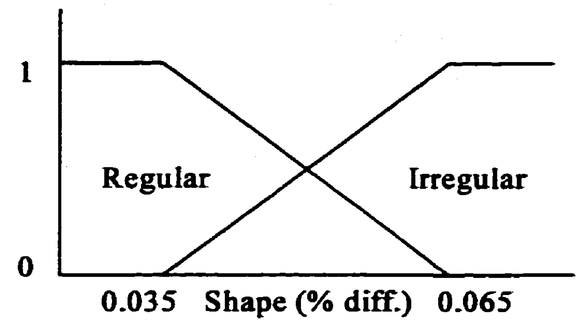


Figure 3.7 : Cookie Shape Membership Function

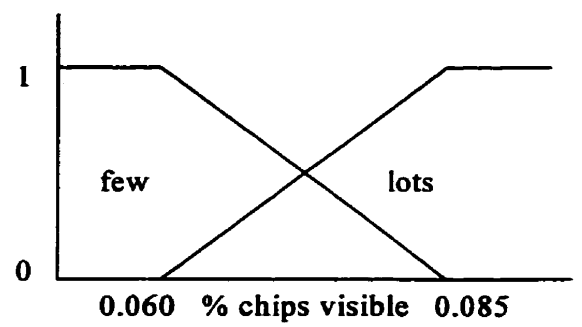


Figure 3.8 : Percentage Chocolate Chips Visible Membership Function

The fuzzy classes chosen for the four characteristics resulted in a universe of twenty-four cookies. Cookies with complete membership in one class for each of the four characteristics were called archetypes. Cookies with intermediate values between classes were considered partial members of multiple archetypes.

Twenty-four archetypes were used to identify consumer behavior. The archetypes were separated into two blocks because there were too many cookies to present to the consumer judges all at once. The blocking confounded the interaction between sizes, shapes, and chip area. Each block was presented to the consumer sensory evaluation panel to rate using the line scale shown in Figure 3.9. The linguistic descriptions of quality provided a natural and intuitive means for the test subjects to rate the cookies. These panels consisted of untrained judges all of whom were selected to represent a cross section of the Canadian population. Untrained judges were selected because training may introduce biases into the decision making process of consumers. Without training, these judges could generate an accurate representation of the behavior of an average consumer. After the consumer judges evaluated the cookies, the ratings were recorded as measurements of the cookies' distances from the origin. The rating for each cookie was recorded as an average distance between the left and right edge of the cookie relative to the origin.

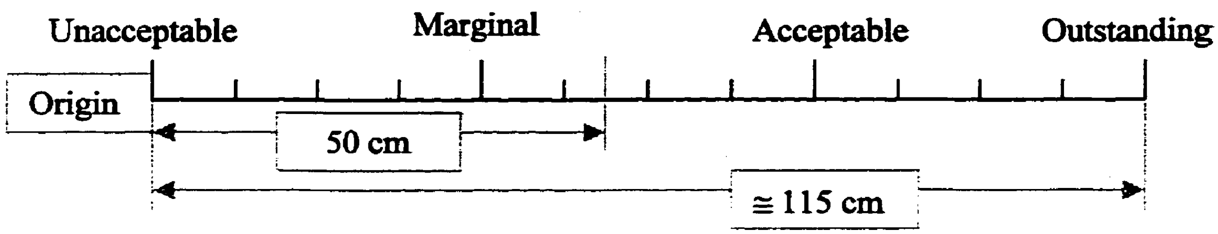


Figure 3.9 : Polling Acceptability Scale Setup

Three polls were performed to acquire data for consumer modeling and model validation. Each poll consisted of 30 consumer judges. Two different sets of cookies were presented to the consumer judges. Both sets of cookies represented entire range of permutations of characteristics for archetypes. The first set of cookies was used during the first two repetitions of polling. Many of the same judges who participated in the first repetition also participated in the second repetition. The consistency of consumer judgement was recorded by using the same people and cookies in two separate polls.

The second set of cookies was used for the third poll repetition. The majority of the judges who participated in the third poll did not participate in either of the first two polls. The consistency of the consumer judgement between different test subjects was recorded by the third repetition.

Upon completion of the three poll repetitions, an analysis of the mean ratings for regular and irregular cookies showed that shape was not a significant influence on the consumer ratings. For consumer modeling, the ratings for both regular and irregular cookies were combined. As well, an error was found in the feature extractor after completing the first three polls. This error affected the accuracy with which the cookies' dough lightness and chips were sensed but size and shape were not affected. As a result, 40% of the cookies that were originally considered archetypes were actually outside of the defined ranges for their classes. These cookies were discarded and a new set of archetypes was established from the remaining archetype cookies. The discarded cookies were used for validating the consumer models and were called Validation Set 1. The universe of archetypes was reduced to twelve permutations by eliminating the shape characteristic. The resulting twelve archetypes consisted of three classes of dough lightness, two classes of size, and two classes of percentage of chocolate chips visible. A list of the ratings for the twelve archetypes can be found in Appendix E.

A general linear model (GLM) procedure (SAS, 1998) was performed on the cookie ratings for all three polls to identify the effect of the characteristics, and two-factor combinations thereof, on the ratings. The analysis also identified the significance of the poll repetitions, the individual panelists, and the effect of the blocking. The GLM program can be found in Appendix C and was written for SAS. The results of the GLM procedure were used to guide the development of the consumer models.

3.3. Consumer Evaluation for Validation

A new set of cookies was established for validation and was called Validation Set 2. Three blocks of ten cookies were selected and can be found in Appendix D. Efforts were made to create each set of cookies to represent a diverse set of characteristic combinations. The cookies used were partial members of two or more archetypes. During the first validation poll, the each set of cookies was presented to twenty people for judgement. A second validation poll was performed where the sets were presented to thirty people each.

3.4. Fuzzy Logic Decision Engines

Two different fuzzy logic engines were developed for comparison. The first engine was a traditional fuzzy logic engine based on a Sugeno type fuzzy techniques while the second was based on a weighted-average technique.

The membership functions for defuzzification were the same diagrams depicting the characteristic classes, Figure 3.5 through Figure 3.8. Both the traditional, and the weighted-average consumer models used the same membership functions to fuzzify the cookie characteristics, however, the engines differed in how they manipulated and combined the membership grades.

3.4.1. Traditional Fuzzy Engine

The membership grades for dough lightness, size, and percentage chips visible were used by twelve rules. Each rule represented one permutation of characteristics in the universe of archetypes. The truth-value for a rule was defined as the minimum membership of all the rule's antecedents, lightness, size, and percentage chips and was calculated by equation (3.15).

$$t_i = \mathrm{Min}\,(\mathrm{lightness}_i, \mathrm{size}_i, \mathrm{chips}_i) \qquad (3.15)$$

Where t_i = the truth value for rule i
Lightness = the degree of membership in lightness characteristic for rule i
Size = the degree of membership in size characteristic for rule i
Chips = the degree of membership in chips characteristic for rule i
i = the i^{th} rule in set of rules

The truth-values of the consequents were then clipped to the minimum truth-value of the rule's antecedents. The consequents for each rule were the mean rating and the variance of the ratings from the associated archetype. The consequent mean values were then aggregated to generate the prediction of the crisp mean as seen in equation (3.16).

$$\text{Crisp-mean} = \left(\sum_{i=0}^{i=n} t_i * M_i\right) \div \left(\sum_{i=0}^{i=n} t_i\right) \qquad (3.16)$$

Where Crisp-mean = the predicted mean of the consumer ratings
n = the number of consequents in the fuzzy engine
i = the i^{th} consequent in set of consequents
t_i = maximum truth value for consequent i
M_i = mean rating of archetype i

The variance with the maximum truth-value was used as the predicted crisp variance, as shown in equation (3.17). The predicted variance provided an assessment of the variability of consumer ratings about the predicted mean.

Crisp-variance = δ^2_i of rule with greatest t_i for all i in n (3.17)

Where Crisp-variance = the predicted variance of consumer ratings
δ^2_i = the variance for rule i
t_i = the truth value for rule i
n = the number of rules in the fuzzy engine
i = the i^{th} rule in set of rules

3.4.2. Weighted Average Fuzzy Relation Engine

A cookie that was judged by the system was called a sample. To find the predicted consumer opinion, the archetypes to which the sample had partial belonging were found. These archetypes made up the decision space for the sample. This decision space was called a context.

All of the archetypes in the context were then compared based on the mean and variances of their polled consumer ratings. If any two archetypes had mean ratings that were farther than 8 cm apart, or if their variances differed by more than 150 cm^2, then they were considered to be rated differently by the consumer. If the two archetypes were considered to be different, then the physical characteristics that differentiated the archetypes were given comparative weights. Each archetype in the context began with weights equal to 1.0 for all characteristics. Then the weights of the characteristics were changed based on the ratings of the archetypes.

For example, a sample X had partial belonging to archetypes A (light, large, lots of chips: mean 80, variance 500), B (light, large, few chips: mean 70, variance 400), C (medium, large, few chips: mean 40, variance 250) and D (medium, large, lots of chips: mean 80, variance 450). These archetypes make up the context for sample X. Sample X's fuzzy description was light ($Truth_{light}$ = 0.7), and a member medium ($Truth_{medium}$ =

0.3), large ($Truth_{large}$ = 1.0) and small ($Truth_{small}$ = 0.0), and finally, lots of chips ($Truth_{lotsChips}$ = 0.8), and few chips ($Truth_{fewChips}$ = 0.2).

To determine the characteristic weights for archetype A, A was compared to all other archetypes in the context. The comparisons and resulting weights can be seen in Table 3.3.

Archetypes for Comparison	Results of Comparison	Differing Characteristics
A (light, large, lots of chips: mean 80, var. 500) *Vs* B (light, large, few chips: mean 70, var. 400)	Δ mean > 8 cm $\therefore$ rated differently	chips (weight +1.0)
A (light, large, lots of chips: mean 80, var. 500) *Vs* C (medium, large, few chips: mean 40, var. 250)	Δ mean > 8 cm & Δ var. > 150 cm^2 $\therefore$ rated differently	lightness (weight +1.0) & chips (weight +1.0)
A (light, large, lots of chips: mean 80, var. 500) *Vs* D (medium, large, lots of chips: mean 80, var. 450)	Δ mean < 8 cm & Δ var. < 150 cm^2 $\therefore$ rated the same	N/A
Resulting attribute weights for Archetype A: Lightness Weight = 2.0 Size Weight = 1.0 Chips Weight = 3.0		

Table 3.3 : Technique for Setting Archetype Characteristic Weights for Weighted Average Fuzzy Engine

This process was repeated for archetypes B, C and D. The characteristic weights for B were 3.0 for lightness, 1.0 for size, and 3.0 for chips. The weights for C were 3.0 for lightness, 1.0 for size, and 3.0 for chips. Finally, the characteristic weights for archetype D were 2.0 for lightness, 1.0 for size, and 3.0 for chips.

Next, the weighted-average of the characteristics was used to determine the degree of belonging of the sample to the archetypes in the context. The equation for calculating the degree of belonging can be seen in equation (3.18).

$$\text{Degree of Belonging to Archetype} = (w_L m_L + w_s m_s + w_c m_c) \div (w_L + w_s + w_c) \qquad (3.18)$$

Where
w_L = weight for lightness characteristic
w_s = weight for size characteristic
w_c = weight for %chip characteristic
m_L = degree of membership in lightness characteristic
m_s = degree of membership in size characteristic
m_c = degree of membership in %chip characteristic

The degree of belonging to an archetype was used as a weight to produce the predicted mean. The equation for calculating the predicted mean can be seen in equation (3.19). The calculation for variance was similar to that of the traditional fuzzy engine. The variance of the archetype to which the sample had the highest belonging was used as the predicted variance. The calculation for the variance can be seen in equation (3.17).

$$\text{Predicted Mean} = \left(\sum_{i=a}^{i=b} w_i * \mu_i\right) \div \left(\sum_{i=a}^{i=b} w_i\right) \qquad (3.19)$$

Where
a = the first archetype in the current context
b = the last archetype in the current context
w_i = the weight of archetype number i
μ_i = the mean of the consumer rating distribution for archetype i
i = the i^{th} rule in set of rules

To continue the example, the degree of belonging for X to each archetype was calculated using equation (3.18). The degrees of belonging are 0.8 archetype A, 0.53 archetype B, 0.36 archetype C and 0.67 archetype D. Using equation (3.19), the final predicted crisp mean was calculated as 71.6. The final predicted variance was 500 because sample X belonged to archetype A the most.

3.4.3. Accept/Reject Decision Making

Both engines used a normal distribution that was fitted to the predicted mean and variance to represent the predicted distribution of consumer ratings for a cookie. The predicted distributions were used to make decisions about the acceptability of a product sample.

These decisions were made by choosing a target rating, and by choosing a target percentage of the population above that rating. If at least the target percentage of the population judged the cookie equal to or greater than the target rating then the cookie was accepted. For this experiment, the target rating was 50 cm and the target percentage of population was 70%. 50 cm was approximately the mid-point between acceptable and marginal. Cookies where 70% or more of the population rated the cookie to be 50 cm or better were accepted. These criteria accept the risk that at most 30% of the population may deem this cookie marginal or worse.

3.4.4. Tuning the Consumer Models Using the Outlook

To account for the variability of consumer decision making and to accommodate different target consumers, both engines also allowed the predictions of consumer ratings to be shifted up or down. The coefficient used to shift the predictions was called the Outlook. Positive Outlooks shifted the prediction up, while negative ones shifted the predictions down the rating scale. The calculation for shifting the means can be seen in equation (3.20). The confidence interval of the mean for each archetype was calculated using the individual archetype means from each poll and can be found in Appendix G.

$$\text{Archetype-mean}_i = \mu_i + O * C_i \text{ for all } i \text{ in } n \qquad (3.20)$$

Where Archetype-mean$_i$ = the resulting shifted mean for archetype i
μ_i = the mean for archetype i before shifting
O = the outlook coefficient
C_i = Confidence interval for archetype i
i = the archetype number
n = the entire universe of archetypes

The mean for each archetype was shifted before performing the rules or weight assignments. The Outlook should vary between – 1.0 to + 1.0. At a value of 1.0 in either direction, the archetype means are shifted to the limit of the confidence interval. In effect, the shifting accounted for the degree of uncertainty in the consumer evaluation data.

3.4.5. Validation of Consumer Models

The cookies in Validation Set 1 (Appendix E) and Validation Set 2 (Appendix D) were used to perform validation testing on the consumer models. First the ratings for the Set 1 were compared to the predicted ratings. Then the ratings generated by Set 2 were compared to the models' predicted ratings. The performances of the consumer models were demonstrated by these tests. The cookies in Set 1 and Set 2 had characteristics that were between the class ranges for archetypes. Therefore, the performances of the consumer models in predicting consumer behavior for cookies between archetypes could be displayed by comparing the actual versus the predicted ratings on these cookies.

Furthermore, the efficacy of adjusting the outlook in fine tuning the acceptance of cookies was also displayed during validation.

4. Results and Discussion

4.1. Feature Extraction and Image Processing

The feature extractor was evaluated for accuracy in quantifying dough lightness, size, shape, and chocolate chips. Its ability to extract the lightness of the dough was tested by comparing the lightness values of four brown plates as sensed by the system, versus a colorimeter. The system's ability to recognize average diameter, major and minor axes, and percentage of chips visible were tested by viewing the locations of the extracted features relative to the original image. In addition, the extracted percentage of chips visible was also tested over a range of dough lightness values to ensure consistent identification.

To establish the performance of the system in sensing dough lightness, the feature extractor was tested against a colorimeter. The L^* value for four brown plates of similar lightness and hue to cookie dough were measured by both the feature extractor algorithm and colorimeter, and then compared.

The performance of the system in extracting the lightness of different shades of brown can be seen in Figure 4.1. The graph shows that the system did not reproduce exactly the true L^* values of the brown colored plates as sensed by the colorimeter. However, the system was sensitive enough to adequately distinguish between different lightness values in the light brown to dark brown color regions similar to cookie dough.

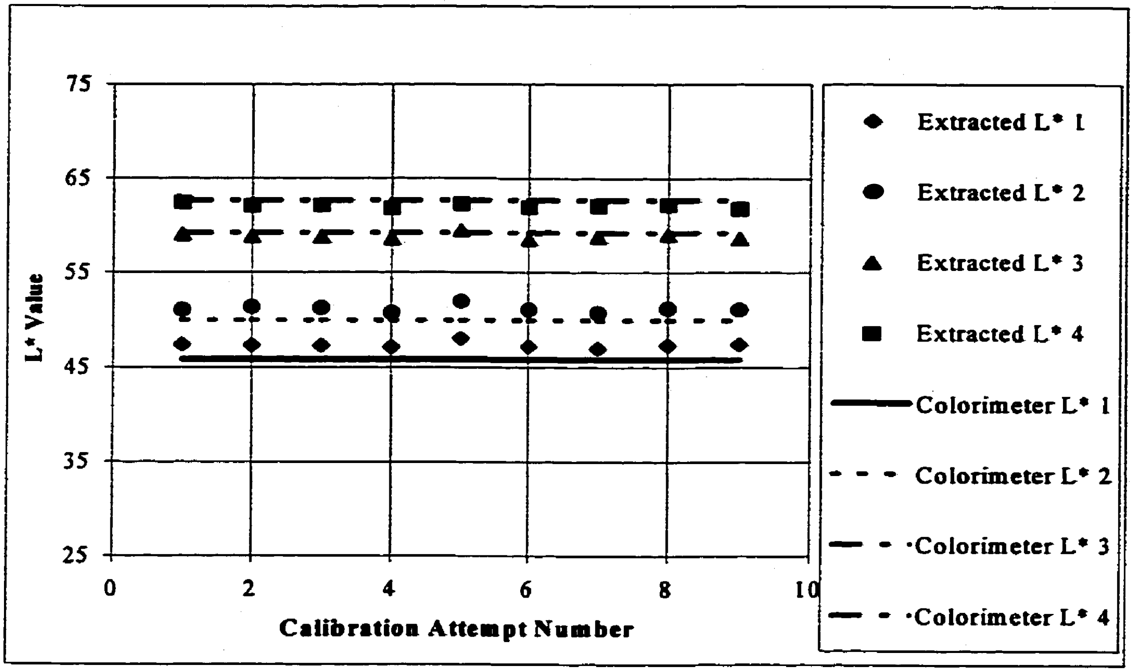


Figure 4.1 : Graph of L* value Measurement for Feature Extractor versus Colorimeter

The accuracy of the feature extractor in identifying the average diameter and major and minor axes was verified visually and manually by comparing the sensed dimensions to those measured by hand. The visual inspection included viewing the extracted perimeter of the system, and then comparing the found perimeter to the actual cookie image.

The system's ability to extract the physical dimensions of the cookies can be seen in Figure 4.2, through to Figure 4.4. The original image of the cookie is shown in Figure 4.2 while the extracted perimeter in Figure 4.3. The perimeter image shows that the feature extractor faithfully reproduced the edge of the cookie. In Figure 4.4, any pixel considered outside the cookie is colored dark gray. The processed image shows that the feature extractor found the area and the center of gravity, labeled c.o.g, of the cookie accurately.

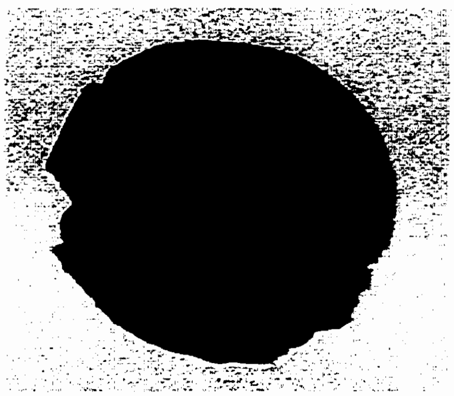


Figure 4.2 : Original Image of Cookie Before Processing

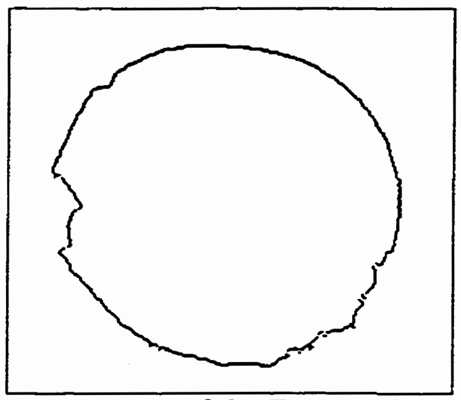


Figure 4.3 : Image of the Extracted Perimeter

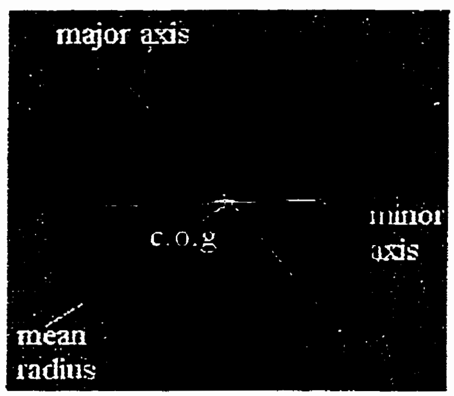


Figure 4.4 : Image of Cookie with Mean Radius, Major and Minor Axis Extracted

The average radius of the cookie is shown as the dark circle overlaid on the image. The major and minor axes are also labeled. The image shows that these features were properly extracted. However, the accuracy of the feature extractor in measuring the length of extracted features in centimeters depended on the calibration of the lens. The system generated results that were consistent with the actual dimensions based on manual measurements with a ruler. The accuracy of the pixel-to-distance and the square pixel-to-area translations were maintained by ensuring that the field of view displayed exactly 14.25 mm per 480 pixels.

Mean radius and axes confidence factors were used as indicators to show whether or not a full search around the perimeter was achieved. Sporadically, there was noise in the image to such a degree that the program could not identify the entire perimeter, and thus some searches were terminated prematurely. Noise can be described as an area of low contrast between the cookie edge and the background. Shadows or other environmental phenomena generated these areas of low contrast. Failures to accurately extract cookie features were resolved by re-capturing the cookie or by retouching the image by hand. The images were retouched by replacing the low contrast background areas and the shadows cast by the cookies with white space.

The success rate for complete searches for both the mean radius and axes was about 80 % without retouching the images. However, when taking into account retouched images, the success rate for complete mean radius and axis extraction approached 95 %. The success rate including retouching is acceptable given the random edge profiles of the cookies.

Visual feedback was also used to verify that the chocolate chip areas were properly extracted. Consistent chip extraction was tested over the entire spectrum of dough lightness-values. The test was conducted by baking a cookie several times and extracting the chocolate chips after each bake. The performance of the extractor in finding the chip area can be seen in Figure 4.5. Extracted chip areas are identified in the image as the gray areas covering the chocolate chips. Relative to the original cookie image in Figure 4.2, the processed imaged shows that the chips were properly extracted while very little dough was mistaken for being chocolate chip.

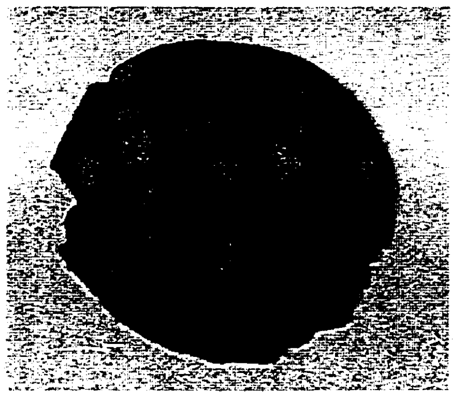


Figure 4.5 : Image of Extracted Chocolate Chips

The performance of the feature extractor in extracting chips from the cookie after each bake can be seen in Figure 4.6. The graph shows that despite the cookie's changing dough-lightness, the percentage of chips identified was fairly consistent. The graph also shows the performance of the chip threshold algorithm. Without the chip threshold correction performed by the program, the amount of chips sensed by the system would have been erroneous and would have increased as the cookie darkened.

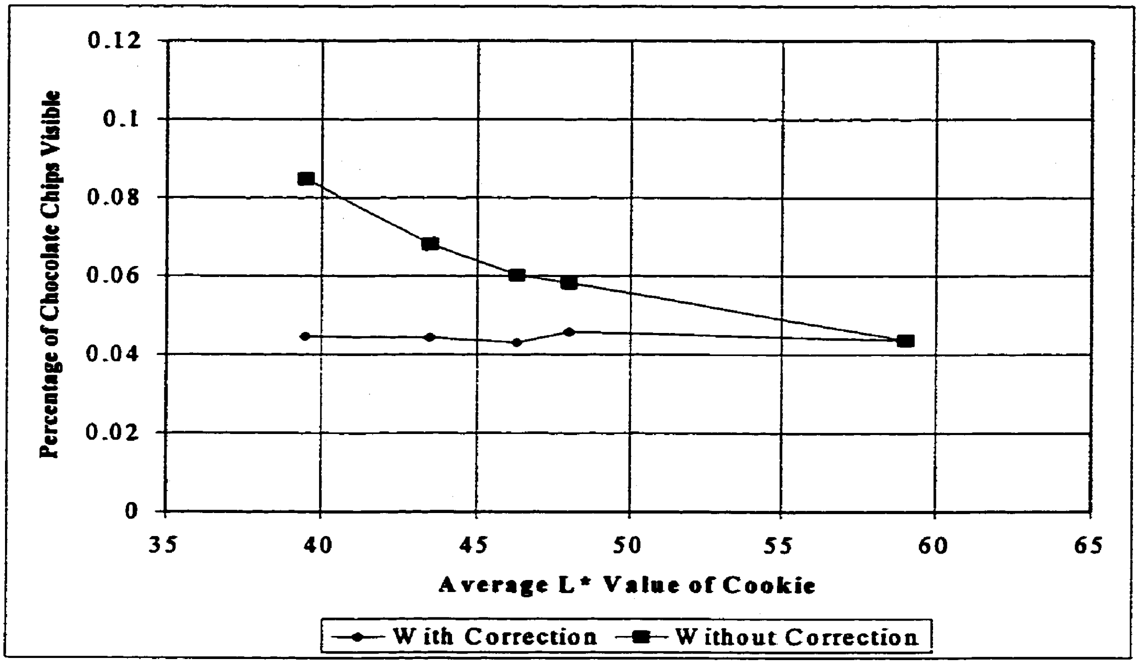


Figure 4.6 : Percentage of Chocolate Chips Sensed vs Average Cookie L* Value

The graphs and images show that the feature extractor identified cookie features accurately. The accuracy of the system depended on the calibration of the system. Repeatability is important when calibrating this system given that the feature extractor is very sensitive to improper calibration. Poor calibration of the size of the field of view generates erroneous cookie size quantification by the feature extractor. Errors in adjusting the color sensitivity of the camera change the extracted lightness of the cookie dough, and can affect the extracted percentage of chips. Careful attention was paid to each calibration performed, but none the less, there existed a small degree of variability in the measurements.

4.2. Consumer Sensory Evaluation

To evaluate consumer behavior, a line scale was used to display the judgement of the consumers. The scale provided linguistic descriptions to which consumers could relate their judgement. For analysis the scale also facilitated numerical measurements of the consumer ratings. The relationship between the visual stimuli of a cookie and the consumer responses could be identified using these measurements.

Analysis of the consumer behavior for shape showed that the ratings for regularly shaped cookies were similar to those for irregularly shaped cookies. Therefore, shape was not a very important feature in consumer judgement of chocolate chip cookies. The ratings for regular and irregular shaped cookies were combined resulting in twelve archetypes. The general trend of consumer opinions for the archetypes was displayed by combining and graphing the ratings from all three polls. These graphs can be seen in Figure 4.7 and Figure 4.8.

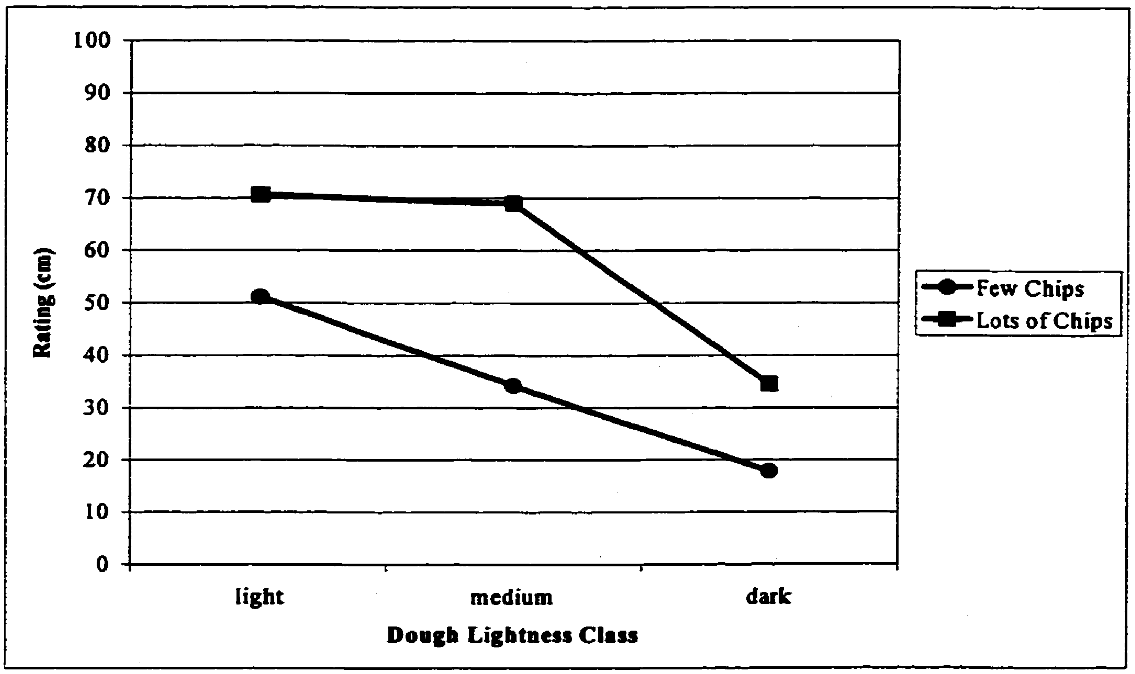


Figure 4.7 : Mean Archetype Ratings Versus Dough Lightness for Small Size Archetypes

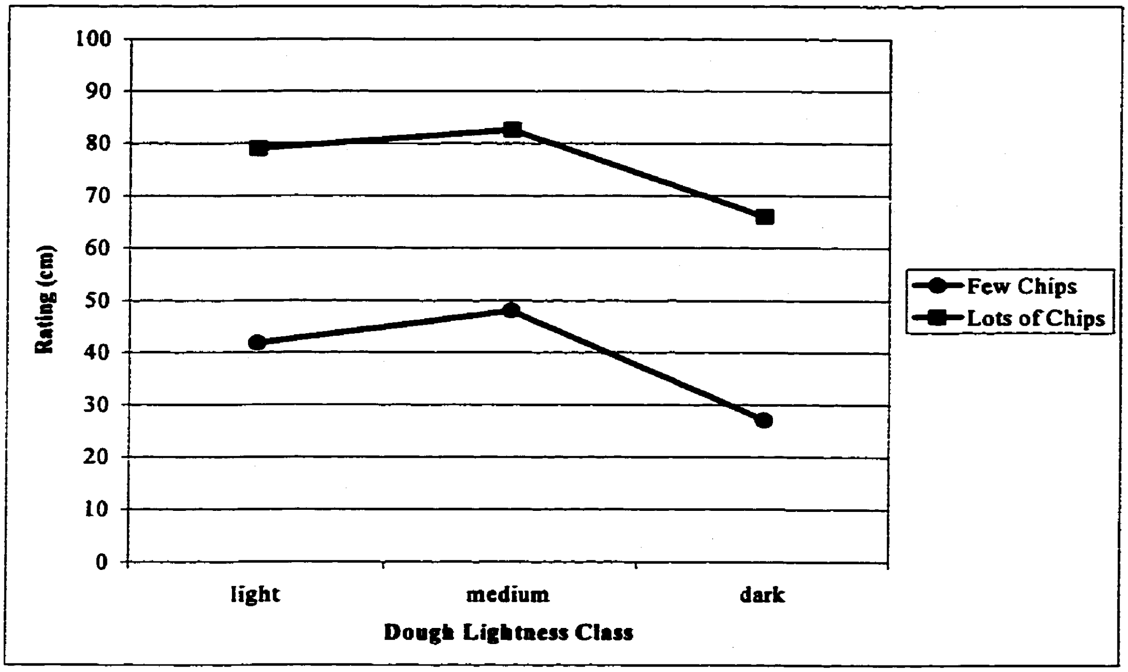


Figure 4.8 : Mean Archetype Ratings Versus Dough Lightness for Large Size Archetypes

The graphs showed that dark cookies were less desirable than light or medium cookies. Furthermore, the graphs showed that a small size cookie was rated lower than a large one. While size had a weaker influence on consumer judgement than lightness, it was still important to the decision-making process. Finally, people had a preference for lots of chips over cookies with few chips. A large difference between the mean ratings for the classes of chocolate chips showed that chips were significant factors in decision-making.

A more detailed analysis was needed to establish the significance of interactions between characteristics, classes, and the ratings of the cookies. Information about the effect of panelists, blocking and poll repetitions on the consumer ratings was also needed. The results of a general-linear-model (GLM) analysis of variance on the data can be seen in Table 4.1. The procedure was used to identify the significant influences on the consume ratings. The analysis showed that dough lightness, size, and chips were directly related to the cookie rating. The analysis also showed that poll-repetitions, individual panelists, and certain interactions were also related to the consumer ratings.

Total Degrees of Freedom : 1188				
Mean Square Error : 488				
Source		DF	Mean Square	F Value
Main effects	Poll Repetition	2	2433	4.98*
	Test Individual	29	2205	4.52*
	Cookie Set	1	880	1.80
	Dough Lightness	2	100714	206.31*
	Cookie Size	1	31457	64.44*
	% Chips Visible	1	209331	428.80*
Inter-actions	Lightness & Size	2	8252	16.90*
	Lightness & Chips	2	1452	2.97
	Size & Chips	1	13957	28.59*

* significant at 95% confidence interval

Table 4.1 : GLM Analysis of Sensory Panel Data

The GLM analysis also indicated that the model had an R^2 of less than 0.5. The low R^2 term shows that statistically, there is a weak correlation between image characteristics and consumer ratings.

These results were used to guide further data analysis. Based on the graphs in Figure 4.7, and Figure 4.8, integration rules were used to describe the trends exhibited by the ratings for each characteristic. For example, the ratings for medium and dark cookies exhibited an additive integration function for dough lightness in this range. The consumer ratings increased from the dark cookie to the medium lightness cookie. However, the light cookies were rated inconsistently. The graphs of consumer ratings indicated that for the range of dough lightness used for this experiment, the consumers did not deem light cookies to be better than medium lightness cookies.

The integration function for size was additive. Larger diameter cookies were rated consistently higher than smaller cookies. The light, small, and few chips archetype was an exception to the rule. Consumer rating variability may have contributed to these unique consumer ratings.

Finally, the relationship between consumer ratings and percentage of chips visible was also additive. The ratings for cookies with lots of chips were higher than cookies with few chips for all combinations of dough lightness and cookie size.

The graphs of archetype means displayed the trends of the consumer opinions, but more information about the ratings was needed to understand consumer judgment. The ratings for any given cookie were distributed along the line scale. Statistical analyses on the ratings for each archetype can be found in Appendix F and G. A summary of the analyses, including the mean, variance, and skewness can be seen in Table 4.2.

When assessing the consumer judgement for a cookie, it was important to consider the variance of the ratings because many of the distributions spanned the entire line scale from unacceptable to outstanding. The variance of the distributions ranged from 204 cm^2 to 865 cm^2. A small variance indicated that there was good agreement between consumers when rating that particular cookie.

Archetype Rating Statistical Summary			
light, small, few chips		*light, small, lots of chips*	
Mean	51.13	Mean	70.56
Sample Variance	503.92	Sample Variance	496.26
Skewness	0.32	Skewness	-0.50
medium, small, few chips		*medium, small, lots of chips*	
Mean	34.27	Mean	68.90
Sample Variance	354.45	Sample Variance	475.89
Skewness	0.26	Skewness	-1.18
dark, small, few chips		*dark, small, lots of chips*	
Mean	17.86	Mean	34.55
Sample Variance	204.27	Sample Variance	582.51
Skewness	0.94	Skewness	0.79
light, large, few chips		*light, large, lots of chips*	
Mean	41.83	Mean	79.00
Sample Variance	673.66	Sample Variance	559.11
Skewness	0.42	Skewness	-1.03
medium, large, few chips		*medium, large, lots of chips*	
Mean	48.05	Mean	82.57
Sample Variance	424.84	Sample Variance	366.53
Skewness	-0.156	Skewness	-1.21
dark, large, few chips		*dark, large, lots of chips*	
Mean	27.07	Mean	66.00
Sample Variance	370.16	Sample Variance	864.79
Skewness	1.01	Skewness	-0.28

Table 4.2 : Summary of Statistical Analysis of Archetype Rating Distributions

The skewness of the distributions ranged from approximately – 1.2 to + 1.0. These values indicated that many of the distributions were asymmetric about their means.

The statistical analysis of the consumer evaluation ratings continued with an investigation of the confidence intervals of the archetypes. The confidence intervals assessed the reliability of the mean for an archetype based on the variation over the three polls, each including 30 consumers.

The reliability of each overall archetype mean was measured as a 95 percent confidence level on the mean rating from each poll for that archetype. The confidence intervals can be seen in Table 4.3, and are quite high for some of the archetypes. The light, large archetypes, and the medium, large, few chips archetype had large confidence intervals because the variance of the means for these cookies was high given the amount of data available. A confidence interval was not calculated for the medium, small, lots of chips archetype because data from only a single poll represented the consumer opinion for that archetype. When using these archetypes to evaluate consumer behavior, the variability of their ratings needed to be taken into account. The large confidence intervals for these archetypes indicated that the potential range of data for these archetypes might be more diverse than for other archetypes.

However, in general the confidence intervals for most of the archetypes were reasonably small, approximately 20 cm or less. These values therefore indicated that the majority of the consumer data was an acceptable representation of consumer judgment. Furthermore, these values also show that predictions of consumer behavior using this set of data should be reasonably accurate.

Cookie type	95% Conf. Interval	lower bound	mean	upper bound
light, small, few chips	5.93	45.2	51.1	57.1
light, small, lots of chips	5.82	64.7	70.6	76.4
light, large, few chips	20.12	21.7	41.8	62.0
light, large, lots of chips	58.13	20.9	79.0	137.1
medium, small, few chips	10.32	23.9	34.3	44.6
medium, small, lots of chips	NA	NA	68.9	NA
medium, large, few chips	53.26	-5.2	48.1	101.3
medium, large, lots of chips	4.24	78.3	82.6	86.8
dark, small, few chips	21.60	-3.7	17.9	39.5
dark, small, lots of chips	22.47	12.1	34.5	57.0
dark, large, few chips	4.22	22.9	27.1	31.3
dark, large, lots of chips	8.22	57.8	66.0	74.2

Note: all measurements are in centimeters

Table 4.3 : Comparison of Confidence Intervals for Archetype Means

The graphs of the archetype-means showed the trend of consumer judgment for dough lightness, size, and chips in detail. The GLM analysis identified dough lightness, size, chips, poll repetitions, and panelists as significant influences on the ratings of the cookies. A statistical analysis of the archetypes' consumer data showed that the ratings were distributions along the line scale. Finally, the confidence intervals showed that the consumer judgement data provided an acceptable representation of consumer opinion.

4.3. Fuzzy Logic Decision Engines

Two consumer models were developed based on the data acquired during sensory evaluation. Both consumer models used fuzzy logic to imitate consumer responses to dough lightness, size, and chocolate chips. Because the sensory evaluation and statistical analysis showed that consumer responses were distributions along the line scale, normal distributions were used to represent the predicted consumer behavior. The normal distributions were constructed from predicted the means and variances generated by the fuzzy engines.

The models also allowed the predicted ratings to be fine tuned. An analysis of the confidence intervals for the ratings showed that there was some uncertainty in the consumer data acquired. An adjustment called an "Outlook" was used to tune the model to account for this uncertainty when predicting consumer behavior. The predicted means were shifted up or down the rating scale using this tuning variable.

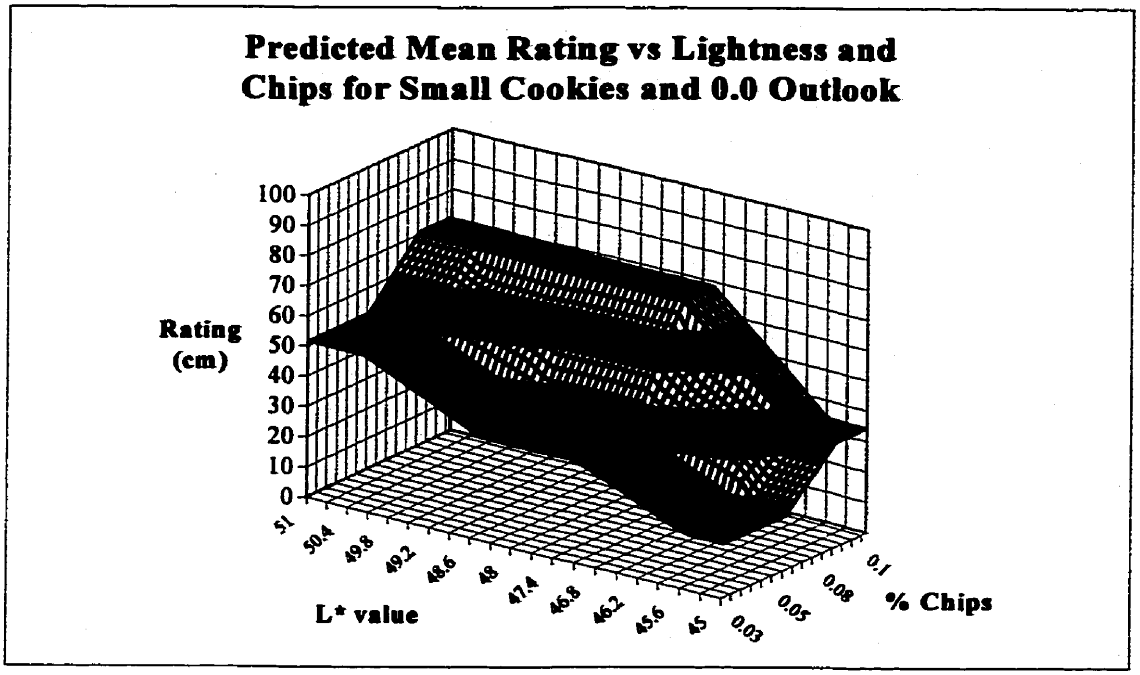
The first consumer model developed was a traditional min-max, rule-based, fuzzy logic decision engine. The second consumer model was a weighted-average, rule-based fuzzy logic decision engine. The performances of the models are discussed in the following sections.

4.3.1. Traditional Fuzzy Engine

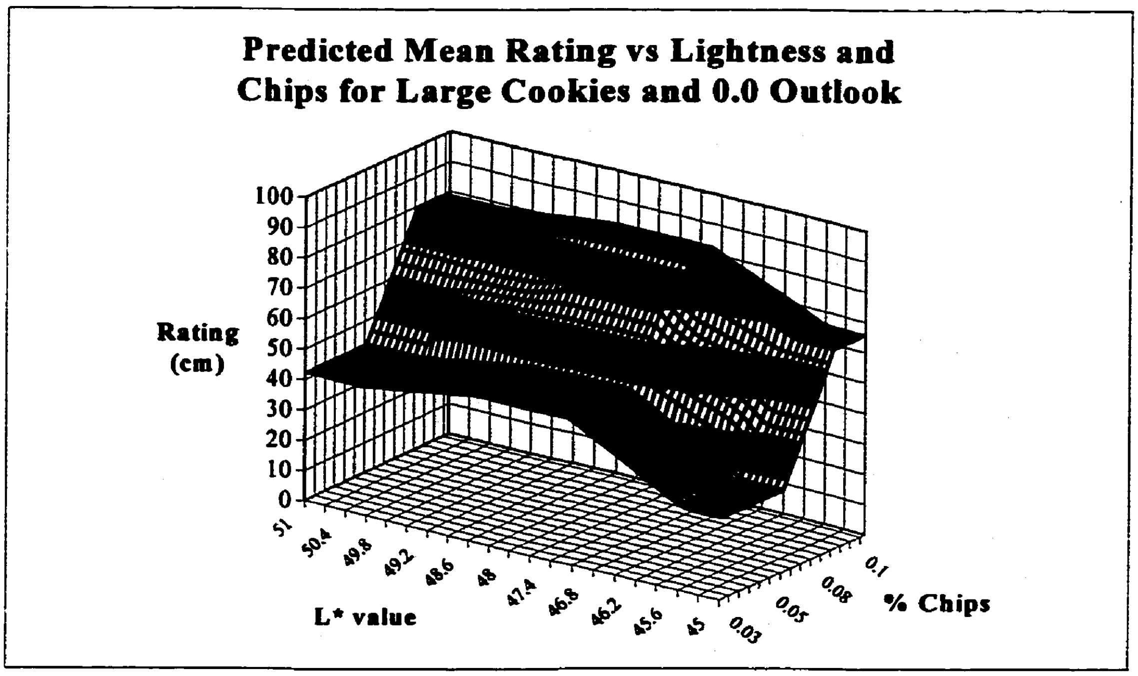
The traditional fuzzy consumer model reproduced trends that were similar to those identified in the sensory evaluation data in Figure 4.7 and Figure 4.8. In other words, light and medium cookies were given higher ratings than dark cookies. Cookies with a lot of chips were rated higher than those without. Finally, large cookies were given higher ratings than small cookies.

The predicted consumer ratings produced by the traditional fuzzy system can be seen in Figure 4.9. The means for each archetype summarized in Table 4.2 were reproduced by the archetype areas, seen as the flat regions in the decision space. Regions of transition between the predicted archetype means were sloped surfaces. The predicted mean ratings in the transitions were gradually interpolated between archetypes.

The variance in the ratings for cookies was not interpolated between archetypes, rather the variance of the archetype with the greatest degree of truth was used as the predicted variance. Figure 4.10 shows that the decision models reproduced the variances in the archetype ratings as summarized in Table 4.2. The predictions of consumer behavior variance were important when judging cookies on the borderline of rejection. For example, if the predicted mean response for a cookie was acceptable and the predicted variance was small, then most likely the cookie would have been accepted. However, if the cookie's ratings had the same predicted mean but a large variance, then the likelihood of rejecting the cookie would have been higher.



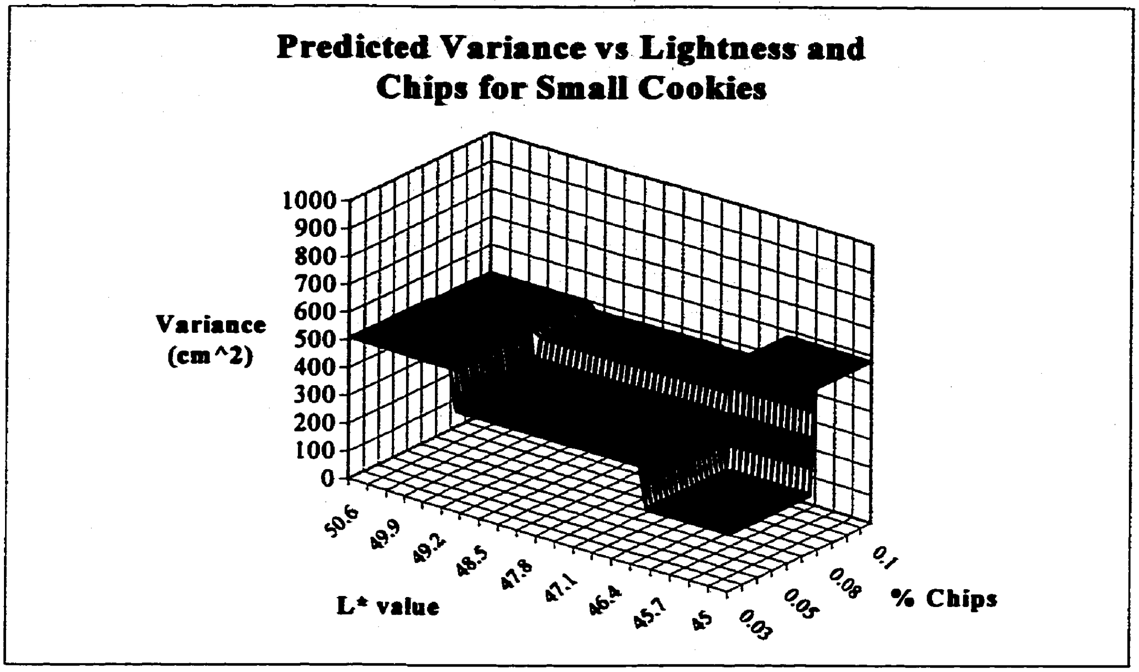
a)



b)

Figure 4.9 : Predicted Means for the Traditional Fuzzy Engine

a)



b)

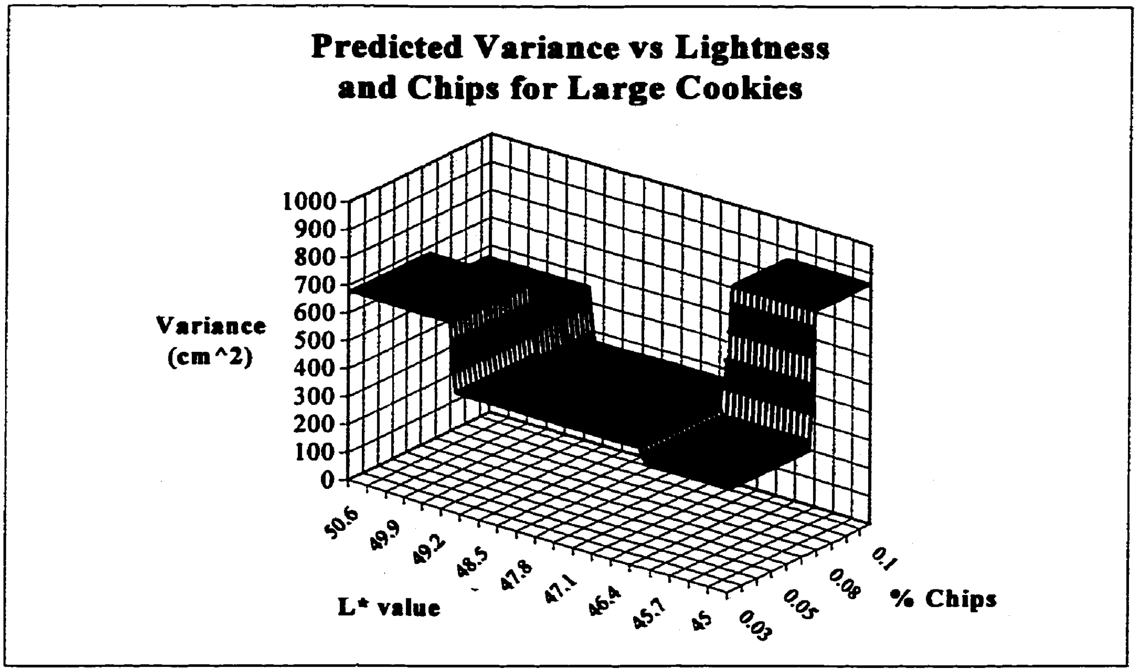


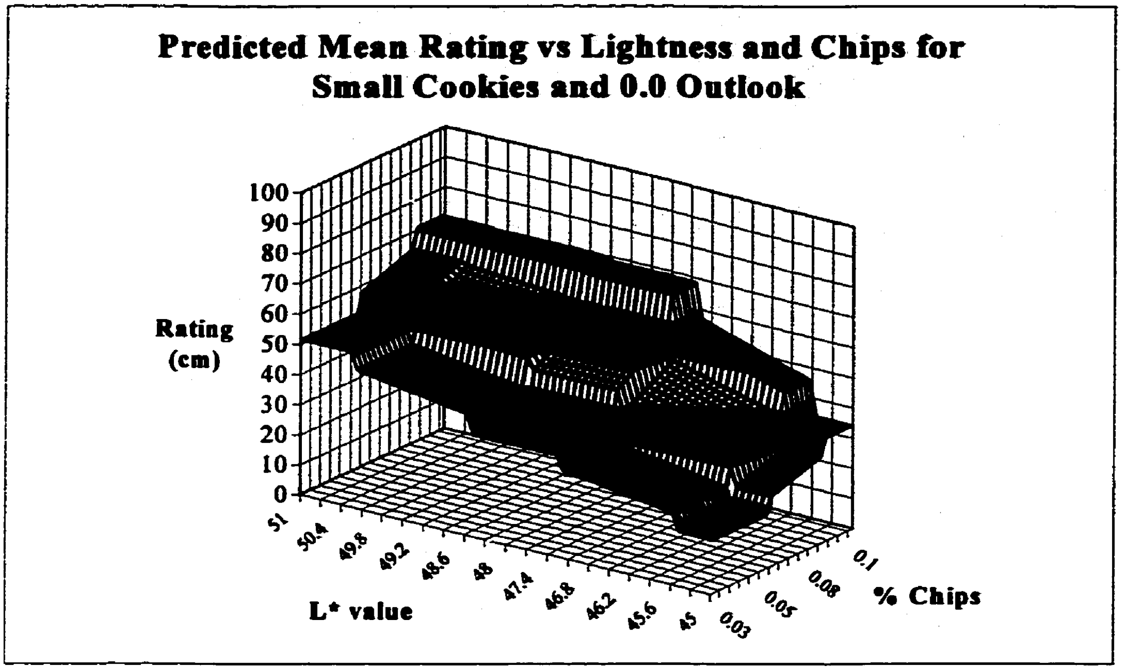
Figure 4.10 : Predicted Variances for the Traditional Fuzzy Engine

4.3.2. Weighted-Average Fuzzy Engine

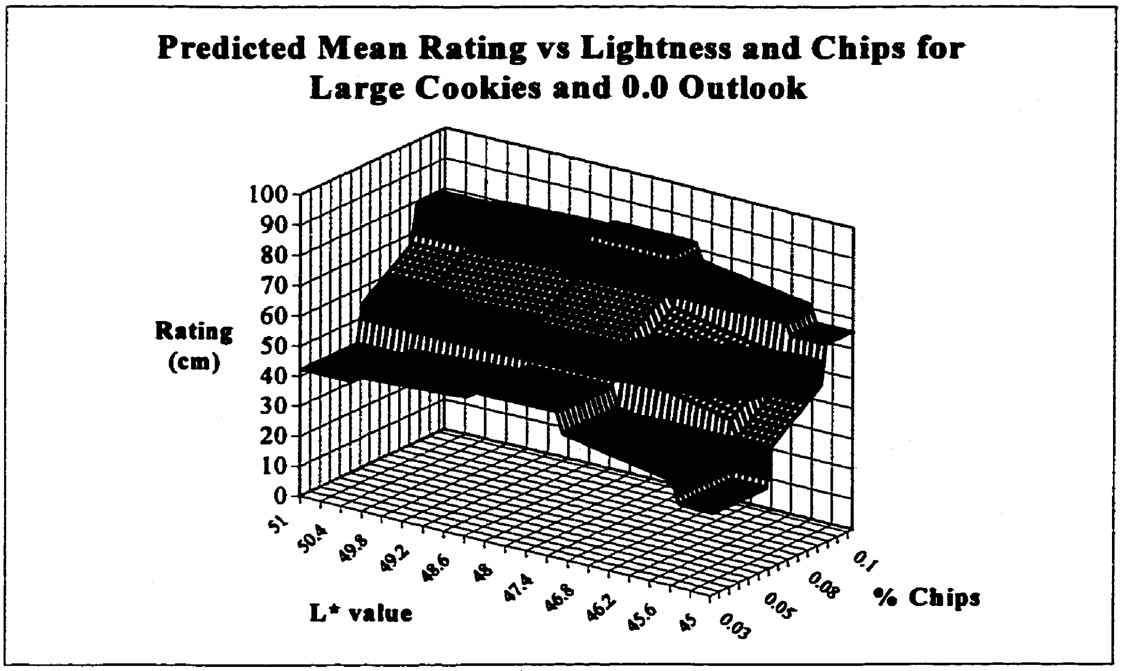
The performance of the weighted-average consumer model was similar to the traditional fuzzy engine. Light and medium cookies were rated higher than dark cookies, and large cookies were rated higher than small ones. As well, the model predicted that cookies with lots of chips would be rated higher than cookies with few of chips. Figure 4.11 shows that the model predicted trends in the mean consumer rating that were similar to those observed from the sensory evaluation data in Figure 4.7 and Figure 4.8. Furthermore, the predicted means for archetypes matched the actual ratings in Table 4.2.

The difference between the traditional engine and the weighted-average engine was in the treatment of the transitions between archetypes. The transitions for the weighted-average fuzzy engine were a series of steps with different slopes. These stepped transitions are in contrast to the traditional fuzzy engine, where the transitions were gradual. Each transition plane in the weighted average consumer model represented a different context where a cookie belonged to two or more archetypes.

The weighted-average fuzzy engine accentuated the differences between each context. The transition contexts brought the predicted means closer to the average mean for all the archetypes in that context. Essentially, the engine provided a compromise between binary decision making and pure fuzzy interpolation between archetypes. The engine classified cookies as either being a pure archetype or a transition cookie.



a)



b)

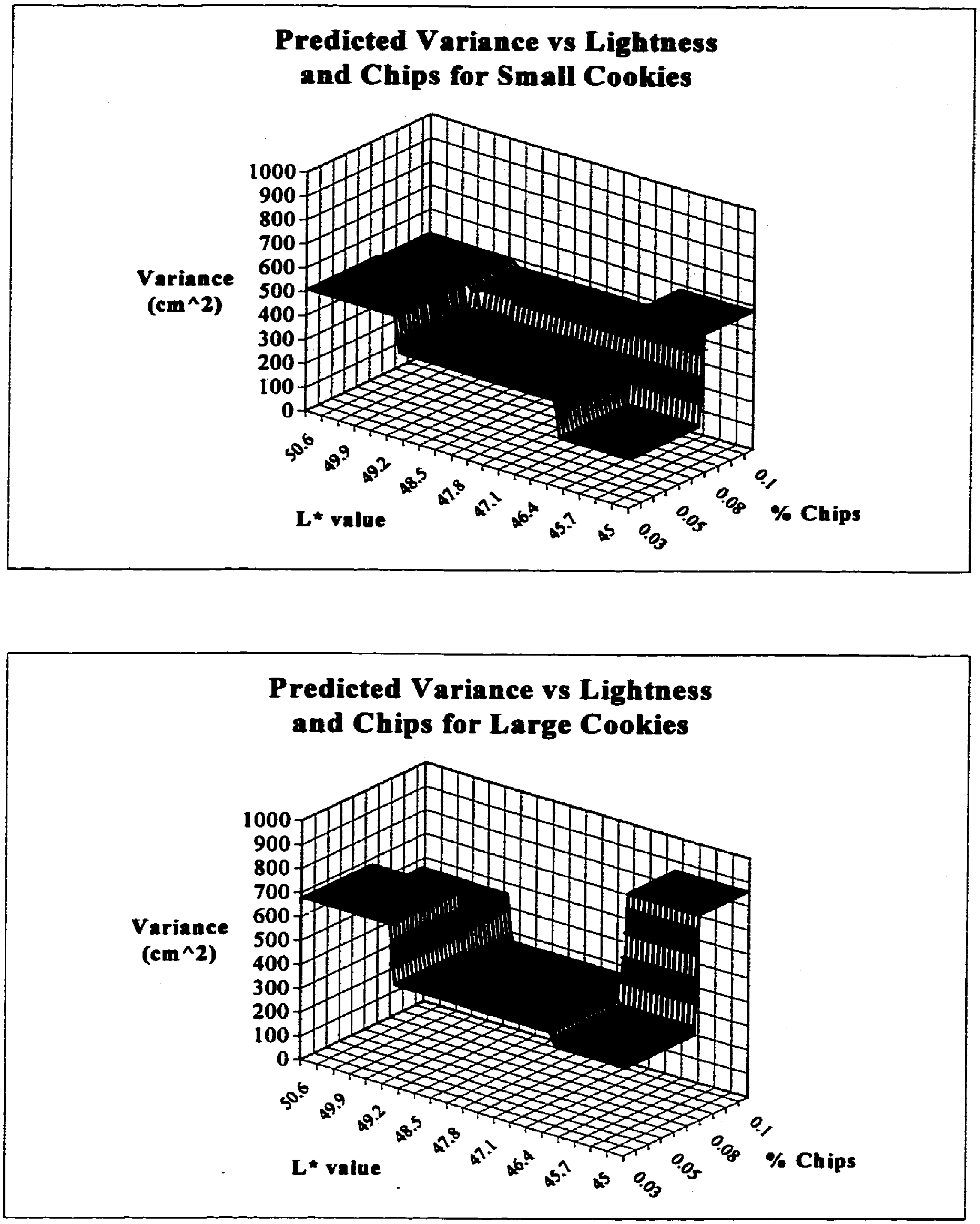
Figure 4.11 : Predicted Means for the Weighted-Average Fuzzy Engine

The predictions of the variance for the weighted-average fuzzy engine were generated in the same manner as the traditional engine and can be seen in Figure 4.12. The variance was not interpoated, rather, the variance of the archetype with the greatest degree of truth was used as the predicted variance.

4.3.3. Distributions of Predicted Consumer Responses

For both the traditional, and the weighted-average fuzzy engines, a normal distribution was used to represent the predicted consumer opinions. The normal distribution was constructed from the predicted mean and variance. Figure 4.13 compares the predicted distribution of consumer ratings for a light, large cookie with lots of chips to actual ratings recorded during sensory evaluation. The skewness of the actual ratings was approximately – 1.0. Although the skewness was high, the maximum point of the predicted distribution was acceptably close to that of the actual distribution. This archetype distribution, even with a high degree of skewness, had its maximum within 15 cm of the predicted maximum.

The error generated by approximating a skewed distribution with a symmetric normal distribution was acceptable because the consumers' rating for any given cookie can change during a single day. For example, the consumer judgement of any cookie may be affected by the hunger of a test subject. Christensen (1983) showed that hunger affects peoples' food choices. A cookie judged before a meal may have been considered acceptable, while after a large meal, the cookie may have then been judged as only marginally acceptable.



a)

b)

Figure 4.12 : Predicted Variances for the Weighted-Average Fuzzy Engine

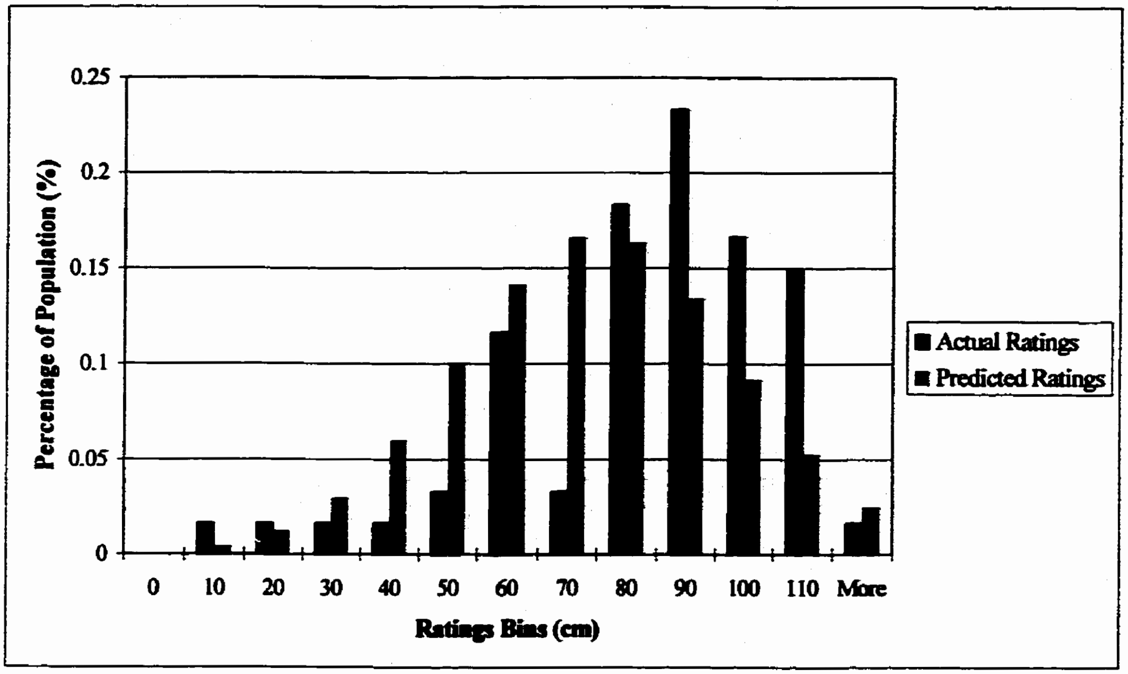


Figure 4.13 : Actual and Predicted Ratings for Light, Large, and Lots of Chips Archetype

To fine-tune the predictions, both consumer models used the Outlook to shift the predicted mean. Positive Outlooks shifted the predictions up the rating scale, while negative Outlooks shifted the predicted distributions downward.

The Outlook shifted the means of the archetypes by using their archetypes' confidence intervals. The uncertainty in the sensory evaluation data was incorporated into the consumer models through the Outlook. Large confidence intervals represented more uncertainty in the data, and thus generated large shifts in the mean. On the other hand, archetypes with small confidence intervals were good representations of consumer opinion and thus shifted only slightly with the Outlook. The graph in Figure 4.14 shows how the Outlook can alter the leniency of the system. If a negative Outlook was used, the system shifted the prediction down the rating scale. A shift up the scale was observed for positive Outlooks.

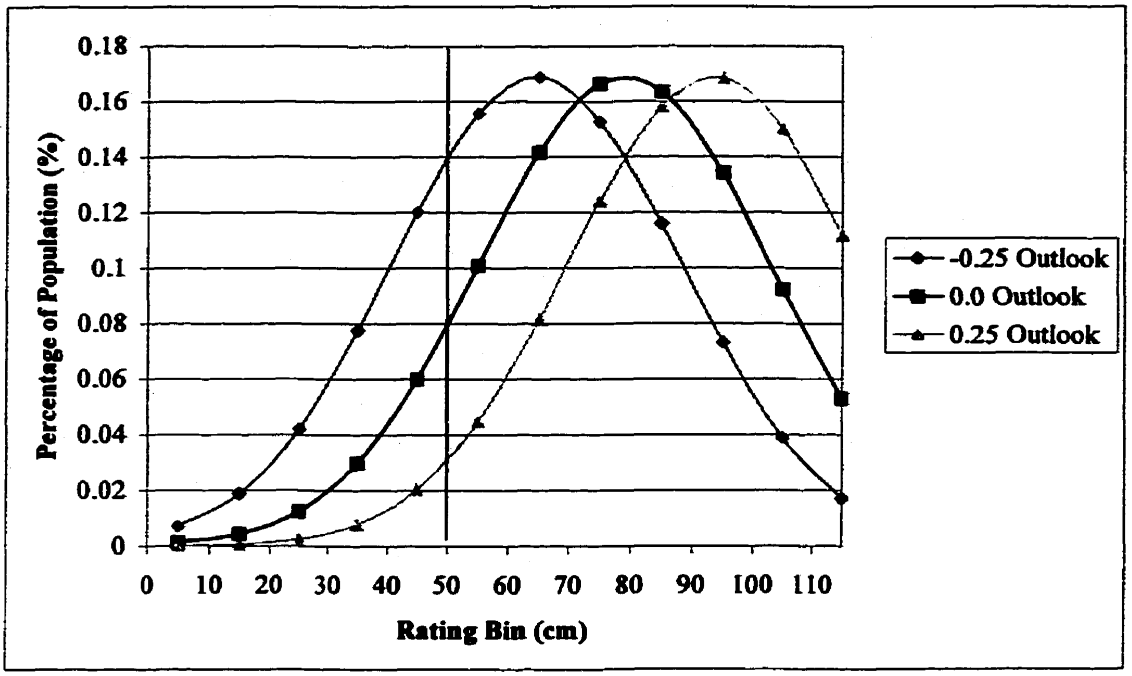


Figure 4.14 : The Effect of Different Outlooks on the Predicted Consumer Ratings for a Light, Large, Lots of Chips Cookie (Traditional Fuzzy Engine)

The intended uses of the consumer models were for assessing the risk of accepting or rejecting a cookie. To assess the risk of accepting a cookie, a manufacturer would decide on a target rating to perform the risk assessment. In this example a target rating of 50 cm was chosen. This rating represented approximately the halfway point between acceptable, and marginally acceptable on the line scale.

The results of the system's prediction for a Light, Large, Lots of Chips cookie can be seen in numerically in Table 4.4 and graphically in Figure 4.14. The table shows the effect of adjusting the Outlook. The risk of accepting this cookie using a 0.0 outlook was that 12 percent of the population would consider the cookie marginal or worse. An outlook of –0.25, generated a prediction that 27 percent of the population would consider the cookie marginal. By tuning the engines using the Outlooks, the systems could be tailored to suit different target consumers. Negative Outlooks would cater to customers that were very picky about their food. Conversely, positive Outlooks would cater to customers that were less discriminating about the food they consumed.

Outlook Value	**Target Rating (cm)**	**Percentage of Population**	
		Above 50 cm	Below 50 cm
-0.25	50	0.73	0.27
0.00	50	0.88	0.12
0.25	50	0.96	0.04

Table 4.4 : Predicted Rating Distribution for a Light, Large Cookie, with Lots of Chips using the Traditional Fuzzy Engine

For example, if the cookies were being prepared for consumption as snacks on long airline flights, the cookie manufacturer may use a more positive Outlook on the judgement of the cookies. In this situation, the passengers on the plane may be less critical of food quality therefore the manufacturer may include a wider range of cookies than usual.

On the other hand, if the cookies were being presented for advertising or promotion, the manufacturer may make the judgement of the system stricter by using a negative outlook. In this situation ensuring high quality is necessary because the cookies would be under close scrutiny. A negative Outlook is used to bias the predicted ratings to represent a more critical consumer.

Care must be taken when adjusting the Outlook. If large Outlooks approaching ± 1.0 are used, the predictions generated by the system may no longer be a good representation of the consumer opinions. Although at the limit of the Outlooks the shifts are still within the 95 percent confidence interval of the mean, the quality of the data should be taken into account. If there is little data representation for an archetype, that archetype will have a large confidence interval. Therefore, a large shift in the mean would result from any adjustments to the Outlook. These large shifts may produce poor predictions in consumer opinion.

4.3.4. Consumer Modeling and Fuzzy Logic

Fuzzy rule-based models are very similar to human thinking and behavior. Terano *et al.* (1987) refer to fuzzy logic as "human simulation". In fact, the structure of fuzzy logic as implemented in these models is very similar to the structure of decision-making as defined by integration theory. First, crisp inputs are analogous to physical stimuli. Secondly, membership functions mimic the translation of the stimuli into psychological sensations. Thirdly, inference rules are similar to the integration stage of decision making. Finally, defuzzification and crisp outputs are comparable to the response stage.

4.3.5. Validation of Consumer Models (Validation Set 1)

The consumer models were assessed for their accuracy in predicting consumer ratings by comparing the predicted consumer ratings to the actual consumer ratings. The first cookie set used for this assessment was Validation Set 1. These cookies were presented to the consumers at the same time as the archetypes but were not used for modeling. Many cookies were partial members of two or more archetypes. The performance of the models in predicting cookies that were partial members of two or more archetypes can be seen in the following analysis.

The predicted and actual ratings were compared by graphing the predicted percentage versus the actual percentage of the population that was above a 50cm rating. These graphs can be seen in Figure 4.15 and Figure 4.16. A cookie was acceptable if 70 percent or more of the population rated the cookie 50 cm or better. The graphs show that the traditional and the weighted-average fuzzy engines performed well on the validation cookies. The accept/reject error rate on the validation cookies for both engines was approximately 27 percent. These graphs show that consumer behavior can be effectively reproduced by predicting both the mean and variance.

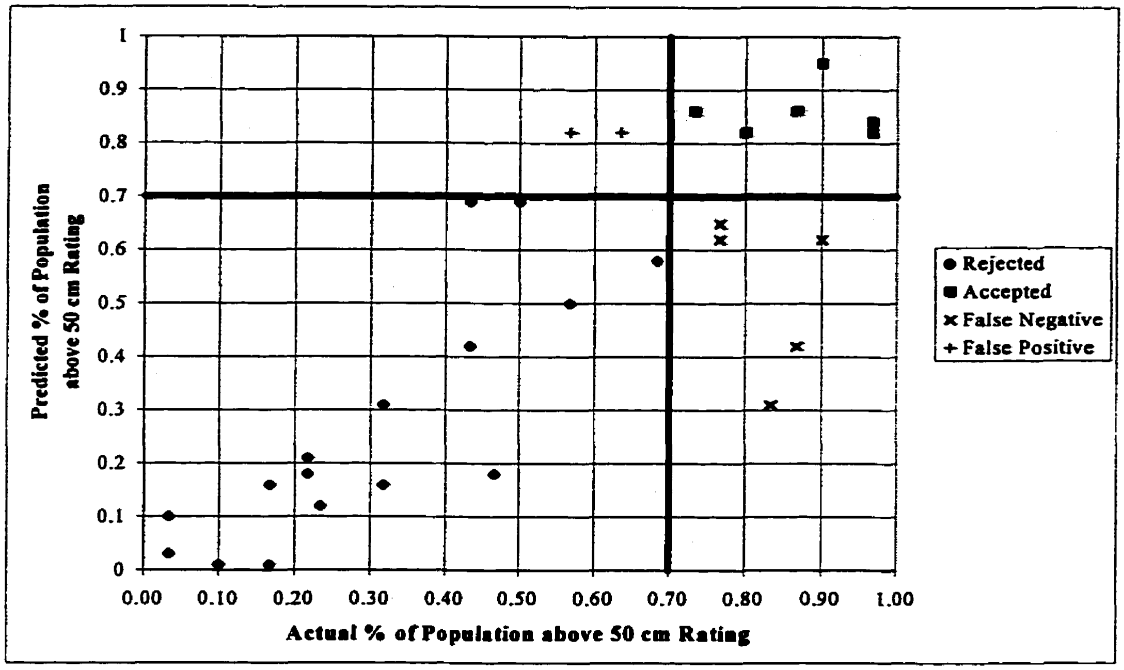


Figure 4.15 : Predicted Versus Actual Accepted and Rejected Validation Cookies for the Traditional Fuzzy Engine Using 0.0 Outlook (Validation Set 1)

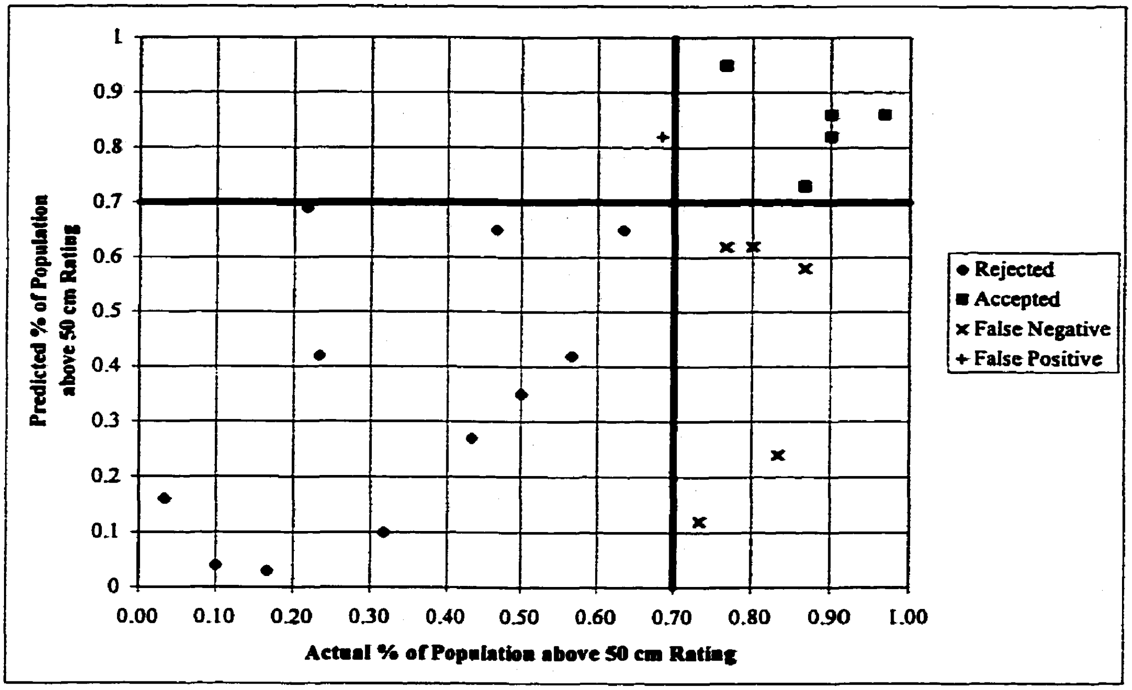


Figure 4.16 : Predicted Versus Actual Accepted and Rejected Validation Cookies for the Weighted-Average Fuzzy Engine Using 0.0 Outlook (Validation Set 1)

The graphs also show that the accept/reject condition was very strict. Very few cookies were considered acceptable. The models only accepted about 20 percent of the cookies. By changing the Outlook, the predictions can be tuned to provide a higher rate of acceptance without changing the accept/reject conditions. The Outlook can be adjusted to account for the uncertainty in the data upon which the model was built. A positive outlook generated an optimistic prediction of the consumer behavior.

Figure 4.17 and Figure 4.18 demonstrate the effects of adjusting the Outlook. As compared to the results shown in Figure 4.16 and Figure 4.17, the number of false negatives was reduced significantly by changing the Outlook from 0.0 to 0.4. With an Outlook of 0.4, the consumer models properly classified approximately 80 % of the cookies while the number of false negative classifications were reduced by 40 % for the traditional engine and 60 % for the weighted-average engine.

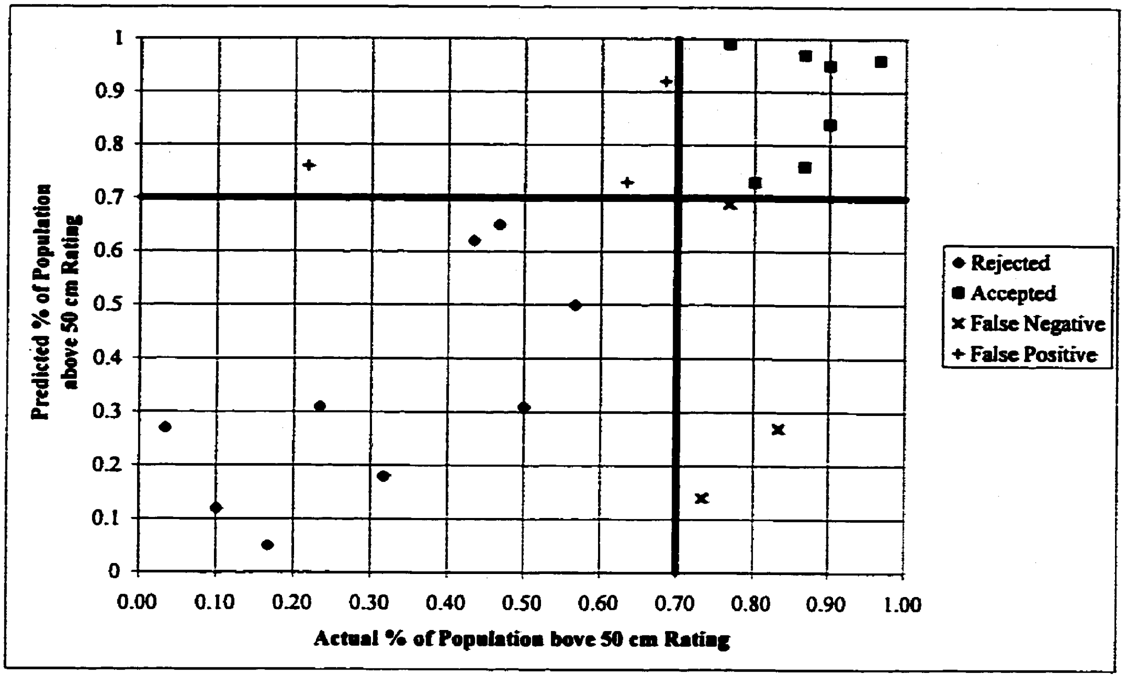


Figure 4.17 : Predicted Versus Actual Accepted and Rejected Validation Cookies for Traditional Fuzzy Engine Using 0.4 Outlook (Validation Set 1)

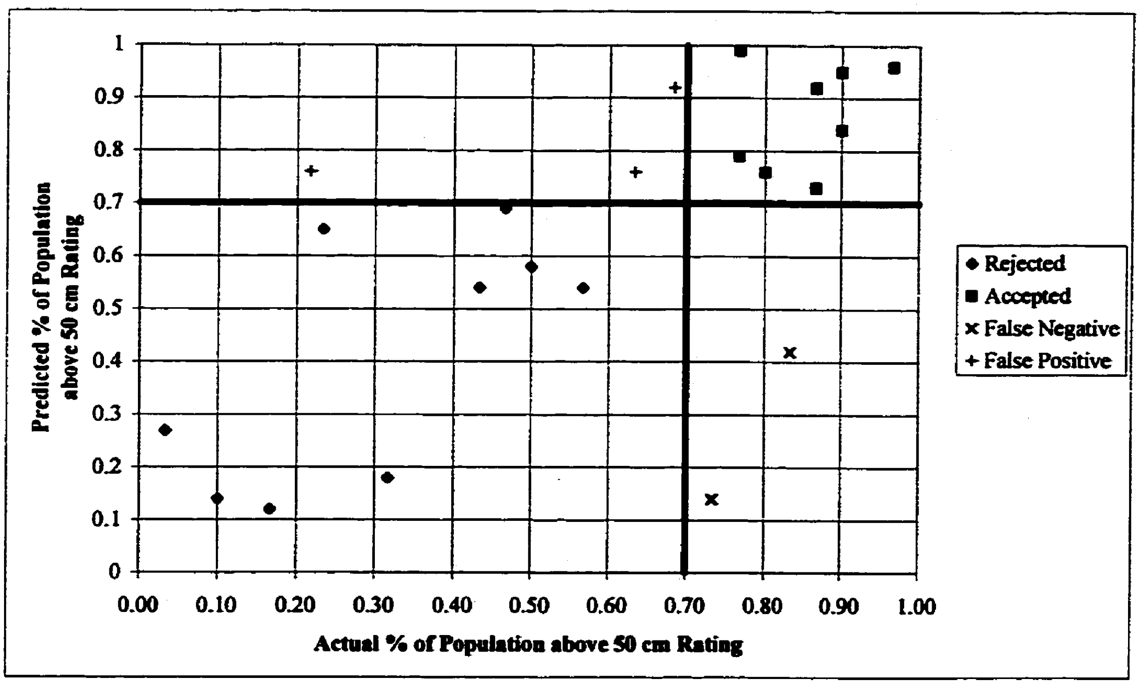


Figure 4.18 : Predicted Versus Actual Accepted and Rejected Validation Cookies for Weighted-Average Fuzzy Engine Using 0.4 Outlook (Validation Set 1)

To further examine the performance of the consumer models, differences between the predicted and the actual percentage of population above a 50cm rating were also calculated and can be seen in Figure 4.19. If the models were perfect in predicting consumer behavior then the difference would equal zero. If the difference was negative, then the fuzzy engines under-estimated the consumer rating. If the difference was positive then the models over-estimated the actual rating.

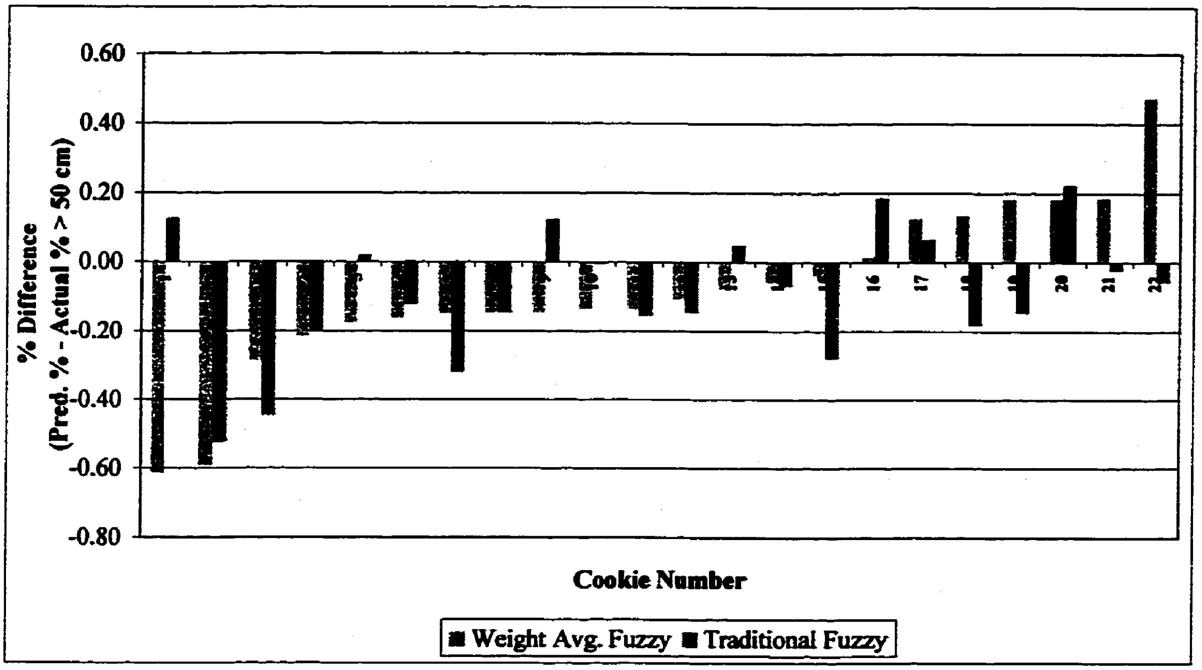


Figure 4.19 : Performance of the Consumer Models' Predictions of Percentage Population Above 50 cm Rating with 0.0 Outlook (Validation Set 1)

The graph shows that neither the traditional nor the weighted-average fuzzy systems were perfect in predicting consumer ratings. The average of the absolute error from zero was calculated for both consumer models. The traditional fuzzy engine produced on average ±16 % error while the weighted-average fuzzy engine produced ± 20 % error. There were no discernable relationships between cookie characteristics and the errors generated by the system. The graph shows that there were no significant differences in the performances of either consumer model.

Another measure of the accuracy of the consumer models is the average error. On average, the traditional fuzzy engine under-predicted the consumer ratings by – 9.0% while the weighted-average fuzzy engine under-predicted by –8.0%. The under estimation of the systems can be attributed to the large variance used to describe the ratings of the archetypes. The results from each of the three modeling polls generated different ratings. Each archetype had a broad overall range of ratings because these results were combined. These large variances generated predicted rating distributions that were wider than the actual rating distributions for any single poll. A percentage of the predicted distributions proportionate to the large variances fell below the 50 cm mark because these predicted distributions were wider than the validation polls.

The criteria used for rejecting a cookie were very strict to demonstrate the systems' ability to classify the cookies. However, in a manufacturing environment the producer would likely use more lenient accept/reject criteria. From a manufacturing standpoint, rejecting cookies that are definitely acceptable to the consumer is a loss of profit. However, manufacturers are likely to overlook selling a small percentage of cookies that are marginally acceptable. Therefore, a manufacturer could use more lenient rejection criteria to usefully classify the cookies. Further fine-tuning of the consumer models can be accomplished with the Outlook. If necessary, a positive Outlook will allow the consumer models to make fewer false negative errors while rejecting the majority of the specimens that would be rated poor by the consumer.

4.3.6. Validation of Consumer Models (Validation Set 2)

In addition to the cookies in Validation Set 1, the predicted and actual consumer ratings for the cookies in Validation Set 2 were also compared. The results can be seen in Figure 4.20 and Figure 4.21. The rate of successful classification for both fuzzy engines was similar for both Validation Set 1 and Validation Set 2. For the cookies in Set 2, the consumer models both correctly classified the cookies 80 percent of the time.

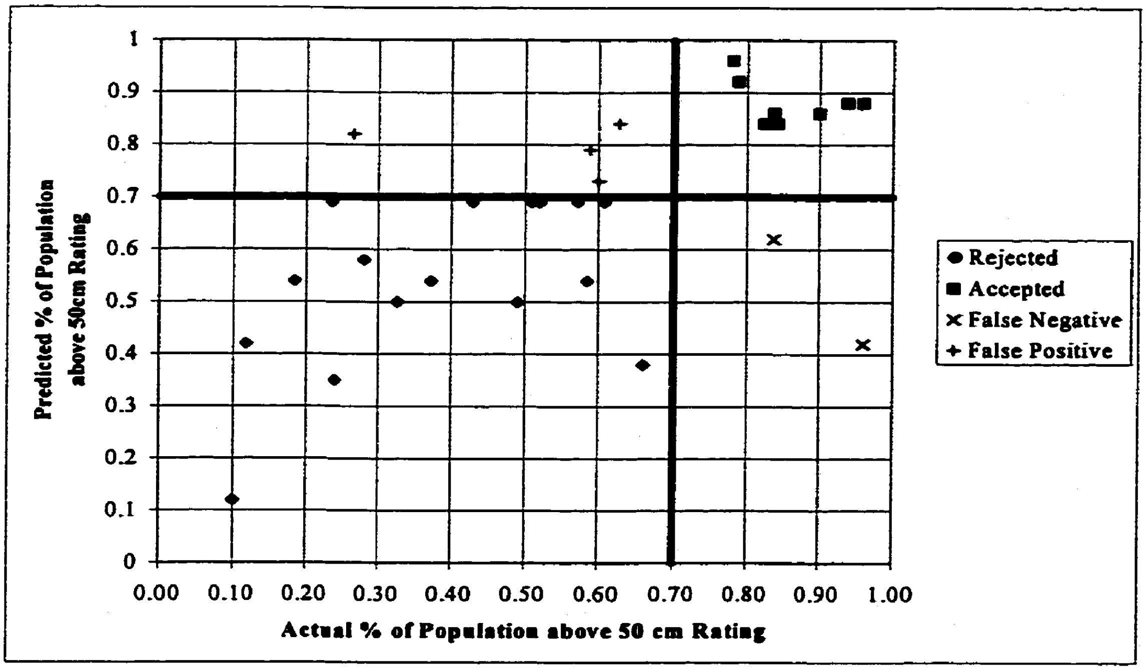


Figure 4.20 : Predicted Versus Actual Accepted and Rejected Cookies for Traditional Fuzzy Engine Using 0.0 Outlook (Validation Set 2)

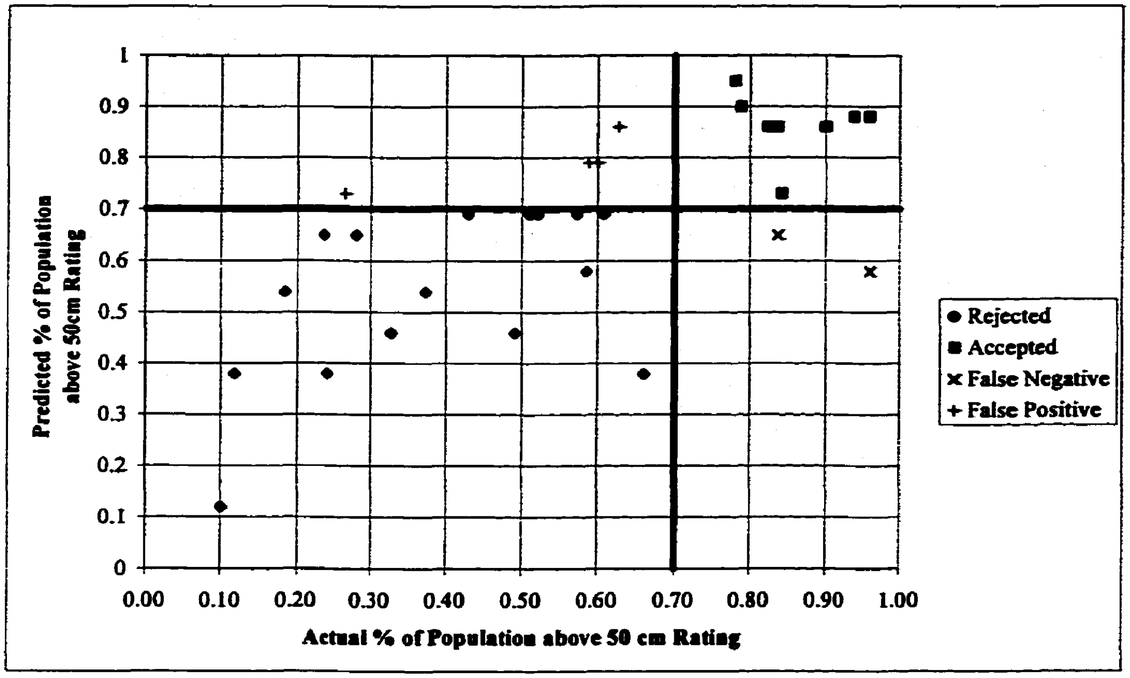


Figure 4.21 : Predicted Versus Actual Accepted and Rejected Cookies for Weighted-Average Fuzzy Engine Using 0.0 Outlook (Validation Set 2)

The graph in Figure 4.22 shows the consumer models' accuracy in predicting the consumer rating for cookies used for Validation Set 2. The average absolute error for both models was approximately ± 18 percent. The average overall error for the models was approximately +10 percent. In other words, on average, the consumer models overestimated the consumer ratings for Validation cookies by 10 percent. The models tended to overestimate the ratings for cookies that were judged to be poor by the panelist judges. In contrast, average error for both fuzzy engines on Validation Set 1 was approximately –9%. The different panelists and cookies used to gather data for Validation Set 2 may have contributed to this difference in error.

The trend in errors did not affect the classification of the models for this experiment because the target cutoff rating was 50 cm. However if a lower target rating was chosen, for example 25 cm, the models' overestimation of poorly rated cookies may adversely affect the performance of the consumer models.

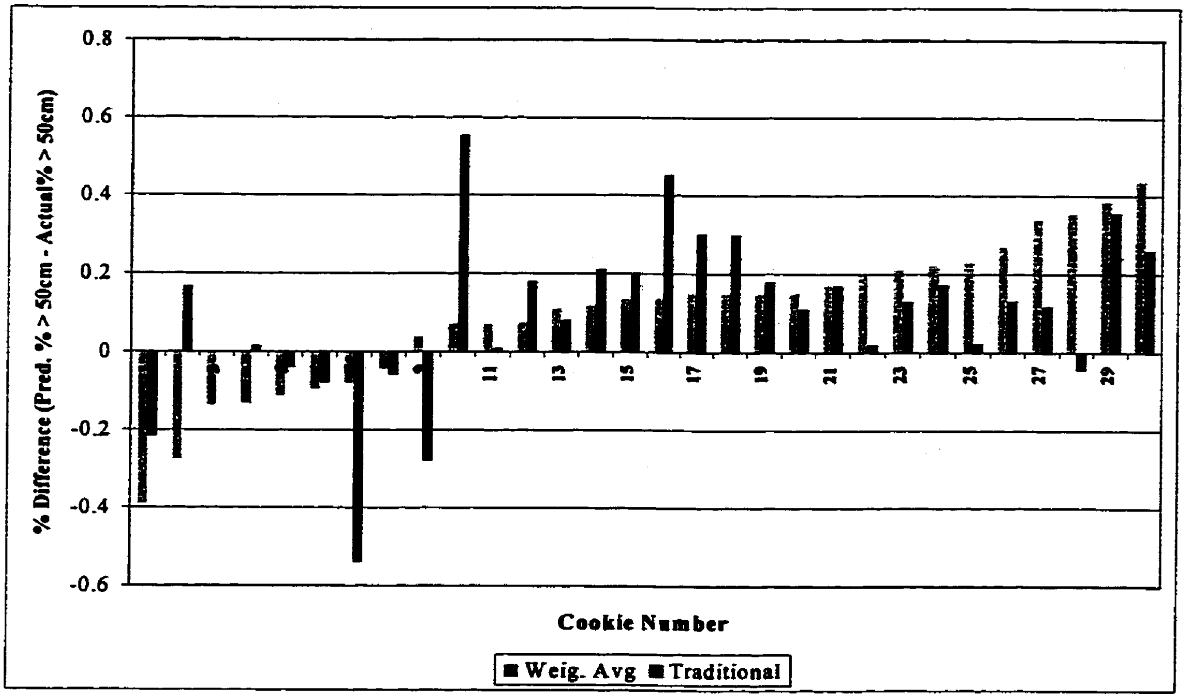


Figure 4.22 : Performance of the Consumer Models' Predictions of Percentage Population Above 50 cm Rating with 0.0 Outlook (Validation Set 2)

The results have shown that it was possible to predict consumer behavior for chocolate chip cookies based on the ratings for 12 archetype cookies. Both models were able to predict the range of consumer opinions for cookies. Finally, the systems were also shown to be adaptable to different target consumers and different manufacturer's needs. The system was able to maintain an acceptable error rate while reducing false negative decisions by manipulating the Outlook.

5. Conclusions and Recommendations

Conclusions

Consumer Behavior

The sensory evaluation model successfully identified the behavior of consumers in judging chocolate chip cookies. The study has shown that the significant that consumers used to judge chocolate chip cookies were dough lightness, size, and percent chocolate chips visible. Furthermore, interactions between dough lightness and size, and chocolate chips and size were notable influences on decision making. Specifically, light and medium lightness cookies were preferred over dark cookies. Large cookies were preferred over small cookies, and finally, cookies with lots of chips were preferred over those with few chips.

The sensory evaluation also showed that consumer decision making for any particular cookie is a trend that is best described as a distribution of opinions. In addition, the study showed that there were variations in the consumer ratings between each poll. The study also showed that there was no discernable pattern to the variance of the ratings for the archetypes. Ultimately, when characterizing consumer judgement for cookies, the uncertainty of consumer behavior needed to be taken into account.

Fuzzy Logic Consumer Models

Fuzzy logic was successfully used to model consumer behavior and its uncertainty by emulating the decision-making stages as identified by integration theory research. The models have shown that crisp inputs into the fuzzy engine are similar to physical stimuli as sensed by people. Furthermore, membership functions are analogous to the translation of the stimuli into psychological sensations. The fuzzy inference rules represent the integration stage of decision making. Finally, defuzzification and the resulting crisp outputs are congruous with the response stage of consumer behavior.

The research has also shown that, through fuzzy logic, it was possible to classify the acceptability chocolate chip cookies based on the consumer ratings for 12 archetype cookies. Both the traditional fuzzy engine and the weighted-average fuzzy engine accounted for the variability of consumer judgement by representing the predicted ratings as distributions. Fitting normal distributions to predicted mean ratings and variances for cookies generated these rating distributions.

The traditional fuzzy engine gradually interpolated the mean rating for cookies between archetypes. In contrast, the weighted-average fuzzy engine used a series of steps to represent the mean ratings in transitions between archetypes. The two fuzzy models predicted the variance of the ratings by using the variance of the archetype to which a sample cookie was most similar. The predicted rating distributions were used to classify the acceptability of cookies given a target rating and a target percentage of the population above that rating.

The accuracy of the classification performed by the fuzzy consumer models was good when compared to a statistical model of the consumer data. The statistical model generated an R^2 of 0.48, which indicated that the mathematical model could only establish a weak correlation between the cookie characteristics and the actual consumer ratings. Furthermore, the nature of consumer judgement could was more appropriately represented through fuzzy logic. The consumer evaluation indicated that the consumer response was additive and deterministic. The behavior was deterministic because the responses to each individual archetype were best described by a rule specific to that archetype. The mathematical model could not represent the deterministic behavior of consumer judgement as appropriately as the fuzzy models.

Automated Inspection System

The results have shown that these fuzzy consumer models could be applied to an inspection system to successfully identify and classify the quality of cookies according to the consumer. The image-processing component of the system was able to quantify cookie features important to decision making. The fuzzy consumer models were then able to use these features to classify a cookie with acceptable accuracy. The inspection systems correctly classified the acceptability of 80% of the cookies while only 7 percent of decisions were false-negative misclassifications. The low percentage of false-negative errors indicated that the systems could assist quality assurance in a manufacturing setting.

Manufacturers who may use this inspection system will also need to adjust the predicted ratings to account for changing consumer behavior. In certain situations, consumers may be critical about the products they buy, while at other times the consumer may be less particular. It was possible to tune the performances of the inspection systems using the Outlook variable. The tuning shifted the predicted ratings by taking into account the variability of consumer ratings. Positive Outlooks shifted the predicted ratings up for an optimistic prediction of consumer opinion, while negative Outlooks generated pessimistic predictions. The Outlook variable provided manufacturers with a means to adjust the performance of the system to suit their needs.

Recommendations

Currently the automated inspection system is specific to cookies. The system extracts the physical features of chocolate chip cookies and then uses a consumer decision model for these cookies to make a judgement. However, manufacturers, especially those in the food industry, seldom produce only a single product. For the system to be viable for manufacturing, it should be able to recognize multiple products, and render judgement based on consumer models specific to each product.

The real-time considerations of manufacturing also need to be addressed before this system can be implemented in a factory setting. The system currently takes seven seconds to process an image. The time to extract features from a cookie needs to be reduced significantly before this automated quality assurance system can be implemented in a factory.

Furthermore, the system as it stands is strictly a decision support system. In other words, the system can only aid quality assurance to decide whether to accept or reject a cookie, given a risk factor. The next step in developing the quality assurance system is to incorporate solution recommendation. The system should be extended to keep track of trends in dough-lightness, size, and chips, and recommend possible remedies for deficiencies in the cookies.

Finally, the system should also be extended to close the control loop. In other words, using the remedy recommendations, the system should be able to automatically adjust the cookie mixture, oven temperature, or bake time. Therefore, the system would not only monitor the quality of cookies, but it would also control the entire baking process to maximize cookie quality according to the consumer, and minimize waste.

6. References

[1] Ballard, D.H. and Brown, C.M. 1982. *Computer Vision*. Prentice-Hall Inc, Englewood Cliffs, New Jersey, USA.

[2] Batchelor, B.G. and Whelan, P.F. 1997. *Intelligent Vision Systems for Industry*. Springer-Verlag, London, England.

[3] Bezdek, J.C. 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York, New York, USA.

[4] Booth, D.A. 1990. Designing Products for Individual Customers. *Psychological Basis of Sensory Evaluation*. Elsevier Science Publishers Ltd, Essex, England. 163-193.

[5] Cai, L.Y. and Kwan, H.K. 1998. Fuzzy Classification Using Fuzzy Inference Networks. *IEE Transactions on System, Man, and Cybernetics, Part B: Cybernetics*, 28(3): 334-347.

[6] Chou, P.B., Rao, A.R., Sturzenbecker, M.C., Wu, F.Y. and Brecher, V.H. 1997. Automatic Defect Classification for Semiconductor Manufacturing. *Machine Vision and Applications*, 9(4): 201-214.

[7] Chamberlin, G.J. and Chamberlin, D.G. 1980. *Color, its Measurement, Computation, and Application*. Heyden and Son Ltd, Philadelphia, Pennsylvania, USA.

[8] Christensen, C.M. 1983. Effects of Color on Aroma, Flavor, and Texture Judgements of Foods. *Journal of Food Science*, 48(3): 787-790.

[9] De Silva, C.W. 1997. Some Recent Development of Intelligent Quality Assessment In Automated Fish Processing. *Proceedings of the 7th International Fuzzy Systems Association World Congress*, 4: 370-375.

[10] De Silva, C.W. 1995. *Intelligent Control: Fuzzy Logic Applications*. CRC Press Inc, Boca Raton, Florida, USA.

[11] Di Nola, A., Sessa, S., Pedrycz, W. and Sanchez, E. 1989. *Fuzzy Relation Equations and their Applications to Knowledge Engineering*. Kluwer Academic Publishers, Boston, Massachusetts, USA.

[12] Ding, K. and Gunasekaran. S. 1994. Shape Feature Extraction and Classification of Food Material Using Computer Vision. *Transactions of the ASAE*, 37(5): 1537-1545.

[13] Dowell, F.E. 1998. Automated Color Classification of Single Wheat Kernels Using Visible and Near-Infrared Reflectance. *Cereal Chemistry*, 75(1): 142-144.

[14] Dubose, C.N., Cardello, A.V. and Maller, O. 1980. Effects of Colorants and Flavorants on Identification, Perceived Flavor Intensity, and Hedonic Quality of Fruit-Flavord Beverages and Cake. *Journal of Food Science*, 45: 1393-1415.

[15] Francis, B. 1990. A New Vision of Quality Contrtol. *Datamation*, 36(7): 70-72.

[16] Francis, F.J. and Clysdesdale, F.M. 1975. *Food Colorimetry, Theory and Applications*. Avi Publishing Co., Westport, Conneticut, USA.

[17] Gunasekaran, S. 1996. Computer Vision Technology for Food Quality Assurance. *Trends in Food Science and Technology*, 7(8): 245-256.

[18] Hall, L. O'Higgins, and Kandel, A. 1986. *Designing Fuzzy Expert Systems*. Koln, TUV, Rheinland.

[19] Heinemann, P.H., Pathare, N.P. and Morrow. C.T. 1996. An Automated Inspection Station for Machine Vision Grading of Potatoes. *Machine Vision and Applications*, 9(1): 14-19.

[20] Hoang, K. and Nachimuthu, A. 1996. Image Processing Techniques for Leather Hide Ranking in the Footwear Industry. *Machine Vision and Applications*, 9(3): 119-129.

[21] Howarth, M.S., Brandon, J.R., Searcy, S.W. and Kehtarnavaz, N. 1992. Estimation of Tip Shape for Carrot Classification by Machine Vision. *Journal of Agricultural Engineering Research*, 53(2): 123-139.

[22] Hunt, R.W.G. 1987. *Measuring Color*. Halsted Press, Toronto, Ontario, Canada.

[23] Johnson, R.A. 1994. *Miller and Freund's Probability and Statistics for Engineers*. Prentice-Hall Inc, Englewood Cliffs, New Jersey, USA.

[24] Klir, G.J. and Yuan, B. 1995. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall Inc., Upper Saddle River, New Jersey, USA.

[25] Kuo, Y.H., Kao, C.I. and Chen, J.J. 1993. A Fuzzy Neural Network Model and its Hardware Implmentation. *IEEE Transactions on Fuzzy Systems*, 1(3): 171-183.

[26] Land, D.G. and Shepherd, R. 1991. Scaling and Ranking Methods. *Sensory Analysis of Foods*, Elsevier Science Publishers Ltd, Essex, England. 141-177.

[27] Lawless, H. 1990. Applications of Experimental Psychology in Sensory Evaluation. *Psychological Basis of Sensory Evaluation*, Elsevier Science Publishers Ltd, Essex, England. 69-91.

[28] Liao, K. Pulsen, M.R., Reid, J.F., NI, B.C. and Bonifacio-Maghirang, E.P. 1993. Corn Kernel Breakage Classification by Machine Vision Using a Neural Network Classifier. *Transactions of the ASAE*, 36(6): 1949-1953.

[29] Ling, P.P., Ruzhitsky, V.N. and Lee, T. 1996. Measuring the Color of Food. *Chemtech*, 26(3): 46-53.

[30] Luzuriaga, D.A., Balaban, M.O. and Yeralan, S. 1997. Analysis of Visual Quality Attributes of White Shrimp by Machine Vision. *Journal of Food Sciences*, 62(1): 113-118,130.

[31] Mamdani, E.H. and Gaines, B.R. 1981. *Fuzzy Reasoning and its Applications*. Academic Press, Toronto, Ontario, Canada.

[32] Mancini, L. 1998, Keeping and Eye on the Line. *Food Engineering*, 70(30): 89-96.

[33] McBride, R.L. and Anderson, N.H. 1990. Integration Psychophysics. *Psychological Basis of Sensory Evaluation*, Elsevier Science Publishers Ltd, Essex, England. 93-115.

[34] MCConnell Jr., R.K. and Blau, H.H. 1995. Color Classification of Non-Uniform Baked and Roasted Foods. *Food Processing Automation IV*, ASAE, St. Joseph, Michigan, USA.

[35] Miller, M.K. and Delwiche, M.J. 1989. A Color Vision System for Peach Trading. *Transactions of the ASAE*, 30(4): 1484-1490.

[36] Ni, B., Paulsen, M.R., Liao, K. and Reid, J.F. 1997. Design of and Automated Corn Kernel Inspection System for Machine Vision. *Transactions of the ASAE*, 40(2): 491-497.

[37] Pedrycz, W. 1989. *Fuzzy Control and Fuzzy Systems*. John Wiley and Sons Canada Ltd, Rexdale, Ontario, Canada.

[38] Perrot, N., Trystram, G., Le Guennec, D. and Guely, F. 1996. Sensor Fusion for Real Time Quality Evaluation of Biscuit During Baking. Comparison between Bayesian and Fuzzy Approaches, *Journal of Food Engineering*, 29(3): 301-315.

[39] Ross, T.J. 1995. *Fuzzy Logic With Engineering Applications*. McGraw Hill Inc., Toronto, Canada.

[40] Sarkar, N.R. 1991. Machine Vision for Quality Control in the Food Industry. *Instrumental Methods for Quality Assurance In Foods*. Marcel Dekker Inc., New York, New York, USA.

[41] SAS Institute Inc. 1990. The SAS system release 6.07. Cary, NC, USA.

[42] Scott, D.C. 1998. *A Data Fusion Approach to Verifying Hand-Written Signatures on Bank Cheques*, University of Guelph.

[43] Shearer, S.A. and Payne, F.A. 1990. Color and Defect Sorting of Bell Peppers Using Machine Vision. *Transactions of the ASAE*, 33(6): 2045-2050.

[44] Simonton, W. and Pease, J. 1993. Orientation Independent Machine Vision Classification of Plant Parts. *Journal of Agricultural Engineering Research*, 53(1): 231-243.

[45] Tao,Y., Heinemann, P.H., Varghese, Z., Morrow, C.T. and Sommer III, H.J. 1995. Machine Vision for Color Inspection of Potatoes and Apples. *Transactions of the ASAE*, 38(5): 1555 1561.

[46] Terano, T., Asai, K. and Sugeno, M. 1987. *Fuzzy Systems Theory and Its Applications*. Academic Press Inc, Toronto, Ontario, Canada.

[47] Thomas, W.V. and Connoly, C. 1986. Applications of Color Processing in Optical Inspection. *Applications of Digital Image Processing*, 65: 116-122.

[48] Unklesbay, K., Unklesbay, N. Keller, J. and Grandcolas, J. 1983. Computerized Image Analysis of Surface Browning of Pizza Shells. *Journal of Food Science*, 48: 1119-1123.

[49] Volz, H.G. 1995. *Industrial Color Testing, Fundamentals and Techniques*. VCH Verlagsgesellschaft mbH., New York, U.S.A.

[50] Von Altrock, C. 1995. *Fuzzy Logic and Neurofuzzy Applications Explained*. Prentice-Hall, Upper Saddle River, New Jersey, USA.

[51] Waksman, A., and Rosenfeld, A. 1997. Assessing the Condition of a Plant, *Machine Vision and Applications*, 10(1): 35-41.

[52] Wang, D., and Keller, J.M. 1998. Use of Fuzzy-Logic-Inspired Features to Improve Bacterial Recognition Through Classifier Fusion, *IEEE Transactions on System, Man, and Cybernetics, Part B: Cybernetics*. 28(4): 583-591.

[53] Wigger, W.D., Paulsn, M. R., Litchfield, J.B., and Sinclair, J.B. 1988. *Classification of fungal-damaged soybeans using color-image processing*, ASAE Paper No. 88-3053. ASAE, St. Joseph. Michigan, USA.

[54] Yang. Q, 1994. An Approach to Apple Surface Feature Detection by Machine Vision, *Computers and Electronics in Agriculture*, 11(2): 249-264.

[55] Yen, J., Langari, R., Zadeh, L.A. 1995. *Industrial Applications of Fuzzy Logic and Intelligent Systems*. Institute of Electrical and Electronics Engineers Inc, New York, New York, USA.

[56] Zimmermann, H.J. 1996. *Fuzzy Set Theory and its Applications*. Kluwer Academic Publishers, Boston, Massachusetts, USA.

7. Appendix A : Image Capture Program Source Code

The source code for the image capture program can be found on the attached disk under the directory "Appendix A : Image Capture Program Source Code".

8. Appendix B : Feature Extraction Program and Fuzzy Engine Source Code

The source code for the feature extraction program and fuzzy engines can be found on the attached disk under the directory "Appendix B : Feature Extraction Program and Fuzzy Engine Source Code".

9. Appendix C : GLM Program and Data

SAS Program for GLM analysis of cookie characteristic significance:

```
DATA dat;

infile 'archrate.txt' ;
INPUT REPNUM PANELIST BLOCK L D C CODE $ R1 R2 Y;
OPTIONS LINESIZE=80 PAGESIZE=66;
TITLE 'Archetype Analysis: (archrate.txt)';
PROC GLM;
 CLASS REPNUM PANELIST BLOCK L D C;
 MODEL Y = REPNUM PANELIST BLOCK L D C L*D L*C D*C;
```

GLM analysis of archetype data:

Archetype Analysis:(archrate.txt)
09:10 Wednesday, June 17, 1998
General Linar Models Procedure
Class Level Information

Class	Levels	Values
REPNUM	3	1 2 3
PANELIST	30	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
BLOCK	2	1 2
L	3	1 2 3
D	2	1 2
C	2	1 2

Archetype Analysis:(archrate.txt)
09:10 Wednesday, June 17, 1998
General Linear Models Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	41	545266.55937	13299.18437	27.24	0.0001
Error	1188	579952.68023	488.17566		
Corrected Total	1229	1125219.23960			

R-Square	C.V.	Root MSE	Y Mean
0.484587	44.45562	22.094697	49.700569

Source	DF	Type I SS	Mean Square	F Value	Pr > F
REPNUM	2	4865.14550	2432.57275	4.98	0.0070
PANELIST	29	63940.84326	2204.85666	4.52	0.0001
BLOCK	1	880.27576	880.27576	1.80	0.1796
L	2	201427.90637	100713.95319	206.31	0.0001
D	1	31457.40220	31457.40220	64.44	0.0001
C	1	209331.28104	209331.28104	428.80	0.0001
L*D	2	16503.78955	8251.89478	16.90	0.0001
L*C	2	2903.34999	1451.67500	2.97	0.0515
D*C	1	13956.56570	13956.56570	28.59	0.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
REPNUM	2	10227.56811	5113.78405	10.48	0.0001
PANELIST	29	63940.84326	2204.85666	4.52	0.0001
BLOCK	1	207.52951	207.52951	0.43	0.5145
L	2	176572.39162	88286.19581	180.85	0.0001
D	1	23318.64408	23318.64408	47.77	0.0001
C	1	167809.97074	167809.97074	343.75	0.0001
L*D	2	17364.02622	8682.01311	17.78	0.0001
L*C	2	3136.59671	1568.29835	3.21	0.0406
D*C	1	13956.56570	13956.56570	28.59	0.0001

10. Appendix D : Selected Cookies for Polls 1, 2, 3, and Validation Polls

Cookie Set #1 for Consumer Poll 1 and 2

Cookie Code	File	L* Value	Diameter (cm)	Shape	% chip
1111	85c.tif	52.13	5.14	0.04	0.06
1112	84c.tif	50.44	5.08	0.05	0.12
1121	108c.tif	50.47	5.22	0.11	0.06
1122	119c.tif	49.87	5.32	0.11	0.12
1211	158c.tif	51.84	5.54	0.06	0.07
1212	55c.tif	49.96	5.52	0.07	0.12
1221	17c.tif	50.47	5.6	0.11	0.05
1222	51c.tif	49.49	5.7	0.10	0.12
2111	89c.tif	47.01	5.26	0.04	0.06
2112	116c.tif	48.25	5.22	0.08	0.14
2121	98c.tif	47.88	5.3	0.11	0.04
2122	60c.tif	48.09	5	0.13	0.14
2211	176c.tif	46.73	6.06	0.06	0.05
2212	140c.tif	44.97	5.64	0.06	0.15
2221	175c.tif	46.73	5.96	0.13	0.06
2222	56c.tif	47.37	5.7	0.14	0.13

Cookie Code	File	L* Value	Diameter (cm)	Shape	% chip
3111	37c.tif	41.61	5.18	0.06	0.08
3112	1c.tif	39.6	5.3	0.06	0.12
3121	83c.tif	43.64	5.22	0.11	0.08
3122	63c.tif	40.46	5.24	0.14	0.15
3211	172c.tif	38.96	5.88	0.06	0.09
3212	128c.tif	41.35	5.66	0.06	0.13
3221	174c.tif	39.6	6.04	0.13	0.10
3222	130c.tif	38.58	5.7	0.12	0.14

Cookie Set # 2 for Consumer Poll #3

Cookie Code	File	L* Value	Diameter (cm)	Shape	% chip
1111	233c.tif	51.03	5.38	0.061	0.076
1112	245c.tif	50.03	5.29	0.065	0.116
1121	223c.tif	50.09	5.08	0.137	0.076
1122	242c.tif	49.39	5.29	0.118	0.119
1211	199c.tif	52.94	5.96	0.062	0.055
1212	288c.tif	49.26	6.09	0.052	0.083
1221	186c.tif	53.26	6	0.151	0.034
1222	260c.tif	51.1	5.93	0.106	0.097
2111	239c.tif	47.18	5.16	0.067	0.075
2112	216c.tif	47.67	5.36	0.062	0.144
2121	238c.tif	46.97	5.36	0.107	0.073
2122	248c.tif	47.29	5.34	0.112	0.134
2211	189c.tif	46.53	6.08	0.060	0.047
2212	265c.tif	47.31	6.19	0.07	0.119
2221	181c.tif	45.89	6.06	0.119	0.051
2222	266c.tif	46.22	5.94	0.12	0.108

Cookie Code	File	L* Value	Diameter (cm)	Shape	% chip
3111	210c.tif	42.29	5.16	0.049	0.087
3112	222c.tif	42.49	5.26	0.038	0.124
3121	250c.tif	42.11	5.1	0.171	0.103
3122	206c.tif	43.56	4.81	0.189	0.147
3211	185c.tif	40.83	5.89	0.052	0.085
3212	286c.tif	42.62	5.91	0.05	0.143
3221	179c.tif	43.04	5.63	0.185	0.063
3222	280c.tif	42.7	5.97	0.142	0.154

Validation Set 2

Set #	Cookie #	L* value	Diameter (cm)	% Chips
1	504	44.48	5.72	0.097
1	505	44.76	5.73	0.095
1	507	46.19	5.71	0.135
1	509	49.72	5.13	0.058
1	508	50.10	5.24	0.073
1	506	50.46	5.42	0.116
1	291	51.27	6.26	0.062
1	503	52.87	5.35	0.059
1	240	56.98	5.41	0.087
1	223	57.47	5.23	0.059
2	518	44.04	5.75	0.047
2	517	45.29	5.86	0.102
2	519	45.79	5.72	0.119
2	516	46.83	5.41	0.064
2	515	48.72	5.61	0.131
2	514	51.18	5.27	0.064
2	267	51.72	6.10	0.098
2	323	55.60	5.62	0.060
2	242	56.17	5.44	0.105
2	227	56.36	5.63	0.081
3	527	45.97	5.72	0.072
3	525	46.54	5.64	0.118
3	526	50.22	5.47	0.072
3	146	50.26	5.67	0.071
3	524	51.29	5.46	0.096
3	528	51.40	5.26	0.058
3	523	51.96	5.40	0.048
3	292	52.18	6.12	0.088
3	364	52.88	5.08	0.085
3	529	53.68	5.60	0.074

11. Appendix E : Selected Archetypes for Modeling and Validation Cookies

Archetypes Used For Consumer Modeling

Cookie Image #	L* value	Diameter (cm)	% Chips
	Dark	**Small**	**Few**
Pics/Poll1/37c.tif	42.70	5.17	0.045
Pics/Poll2/210c.tif	41.81	5.16	0.057
	Dark	**Small**	**lots**
Pics/Poll2/222c.tif	43.00	5.26	0.090
Pics/Poll1/63c.tif	43.65	5.23	0.113
Pics/Poll2/206c.tif	44.02	4.81	0.114
	Dark	**Large**	**Few**
Pics/Poll2/179c.tif	42.43	5.63	0.039
Pics/Poll1/172c.tif	39.03	5.88	0.041
Pics/Poll1/174c.tif	39.85	6.04	0.048
Pics/Poll2/185c.tif	40.36	5.89	0.049
	Dark	**Large**	**Lots**
Pics/Poll2/286c.tif	43.65	5.91	0.102
Pics/Poll1/128c.tif	44.23	5.66	0.104
Pics/Poll2/280c.tif	43.84	5.97	0.115
	Medium	**Small**	**Few**
Pics/Poll1/98c.tif	47.23	5.30	0.030
	Medium	**Small**	**Lots**
Pics/Poll2/248c.tif	48.51	5.34	0.117
	Medium	**Large**	**Few**
Pics/Poll1/175c.tif	47.37	5.97	0.042
	Medium	**Large**	**Lots**
Pics/Poll2/265c.tif	48.29	6.19	0.092
Pics/Poll1/140c.tif	47.59	5.64	0.131
	Light	**Small**	**Few**
Pics/Poll1/85c.tif	52.02	5.14	0.054
Pics/Poll1/108c.tif	50.85	5.21	0.056
	Light	**Small**	**Lots**
Pics/Poll2/245c.tif	51.15	5.29	0.104
Pics/Poll1/119c.tif	50.58	5.33	0.104
Pics/Poll2/242c.tif	50.50	5.29	0.106
Pics/Poll1/84c.tif	51.82	5.08	0.110
	Light	**Large**	**Few**
Pics/Poll2/186c.tif	52.84	6.00	0.030
Pics/Poll2/199c.tif	52.95	5.96	0.047
	Light	**Large**	**Lots**
Pics/Poll1/51c.tif	51.10	5.70	0.108

Validation Set 1

Cookie Image #	L* value	Diameter (cm)	% Chips
Pics/Poll1/83c.tif	45.09	5.14	0.060
Pics/Poll2/250c.tif	42.09	5.10	0.068
Pics/Poll1/1c.tif	40.85	5.30	0.073
Pics/Poll1/89c.tif	46.39	5.26	0.049
Pics/Poll1/130c.tif	41.39	5.69	0.103
Pics/Poll2/238c.tif	47.07	5.36	0.060
Pics/Poll2/239c.tif	47.22	5.16	0.061
Pics/Poll2/189c.tif	45.93	6.07	0.034
Pics/Poll1/176c.tif	46.62	6.06	0.038
Pics/Poll2/266c.tif	46.70	5.94	0.068
Pics/Poll1/60c.tif	49.52	5.30	0.119
Pics/Poll2/216c.tif	49.24	5.36	0.128
Pics/Poll2/288c.tif	49.66	6.09	0.062
Pics/Poll1/56c.tif	49.27	5.70	0.115
Pics/Poll2/223c.tif	50.45	5.08	0.068
Pics/Poll1/17c.tif	51.16	5.55	0.042
Pics/Poll2/233c.tif	51.55	5.38	0.068
Pics/Poll1/158c.tif	51.71	5.54	0.069
Pics/Poll1/55c.tif	51.53	5.47	0.108
Pics/Poll2/260c.tif	51.83	5.93	0.081
Pics/Poll2/181c.tif	45.23	6.06	0.038
Pics/Poll1/116c.tif	50.22	5.22	0.127

12. Appendix F : Archetype Rating Distribution Statistical Analysis

Archetype Rating Statistical Summary

light, small, few chips	
Mean	51.127083
Standard Error	2.0492364
Median	50
Mode	64.75
Standard Deviation	22.44826
Sample Variance	503.92437
Kurtosis	-0.166228
Skewness	0.3170567
Range	105.5
Minimum	7.75
Maximum	113.25
Sum	6135.25
Count	120

light, small, lots of chips	
Mean	70.558056
Standard Error	1.6604332
Median	74.75
Mode	79
Standard Deviation	22.277049
Sample Variance	496.2669
Kurtosis	0.2071795
Skewness	-0.502408
Range	110.75
Minimum	9
Maximum	119.75
Sum	12700.45
Count	180

medium, small, few chips	
Mean	34.270833
Standard Error	2.4305487
Median	35.75
Mode	2.5
Standard Deviation	18.826949
Sample Variance	354.45401
Kurtosis	-0.253747
Skewness	0.2601642
Range	75.75
Minimum	2.5
Maximum	78.25
Sum	2056.25
Count	60

medium, small, lots of chips	
Mean	68.9
Standard Error	3.9828223
Median	72.125
Mode	89
Standard Deviation	21.814816
Sample Variance	475.88621
Kurtosis	1.3137355
Skewness	-1.184082
Range	87.75
Minimum	6
Maximum	93.75
Sum	2067
Count	30

dark, small, few chips	
Mean	17.855556
Standard Error	1.5065296
Median	13.75
Mode	2.5
Standard Deviation	14.292194
Sample Variance	204.26682
Kurtosis	0.3133167
Skewness	0.9414982
Range	57.25
Minimum	2.25
Maximum	59.5
Sum	1607
Count	90

dark, small, lots of chips	
Mean	34.547917
Standard Error	2.2032378
Median	32.375
Mode	2.75
Standard Deviation	24.135261
Sample Variance	582.51081
Kurtosis	0.6421596
Skewness	0.7937109
Range	119.25
Minimum	2.25
Maximum	121.5
Sum	4145.75
Count	120

light, large, few chips	
Mean	41.833333
Standard Error	3.3507613
Median	36.5
Mode	33
Standard Deviation	25.954885
Sample Variance	673.65607
Kurtosis	-0.874213
Skewness	0.4235661
Range	94.25
Minimum	2.75
Maximum	97
Sum	2510
Count	60

light, large, lots of chips	
Mean	79
Standard Error	3.0526281
Median	87
Mode	88.5
Standard Deviation	23.645555
Sample Variance	559.11229
Kurtosis	1.1263467
Skewness	-1.0297
Range	118.25
Minimum	4.75
Maximum	123
Sum	4740
Count	60

medium, large, few chips	
Mean	48.05
Standard Error	2.6609676
Median	50.875
Mode	23
Standard Deviation	20.611766
Sample Variance	424.84492
Kurtosis	-0.268441
Skewness	-0.155615
Range	95.75
Minimum	3
Maximum	98.75
Sum	2883
Count	60

medium, large, lots of chips	
Mean	82.572222
Standard Error	2.0180453
Median	85.25
Mode	90
Standard Deviation	19.144859
Sample Variance	366.52562
Kurtosis	1.7680676
Skewness	-1.21114
Range	89.75
Minimum	23.75
Maximum	113.5
Sum	7431.5
Count	90

dark, large, few chips	
Mean	27.072222
Standard Error	1.434034
Median	22.875
Mode	3
Standard Deviation	19.239586
Sample Variance	370.16165
Kurtosis	0.6539948
Skewness	1.0118803
Range	85.5
Minimum	2.5
Maximum	88
Sum	4873
Count	180

dark, large, lots of chips	
Mean	65.989583
Standard Error	2.6845102
Median	68.75
Mode	68.75
Standard Deviation	29.407336
Sample Variance	864.79138
Kurtosis	-0.686253
Skewness	-0.275007
Range	128.5
Minimum	2.75
Maximum	131.25
Sum	7918.75
Count	120

13. Appendix G : Archetype Mean Confidence Interval Data

Statistical Information for Archetype Means

light, small, few chips

	means
	52.48
	53.59
	52.85
	45.58
mean of means	51.13
var of means	13.87
Confidence Level(95.0%)	5.92648

light, small, lots of chips

	means
	68.90833
	74.55333
	66.01667
	79.90833
	66.20833
	67.75333
mean of means	70.55806
var of means	30.70446
Confidence Level(95.0%)	5.815082

medium, small, few chips

	means
	35.08333
	33.45833
mean of mean	34.27083
var of means	1.320313
Confidence Level(95.0%)	10.32375

medium, small, lots of chips

	means
	68.9
mean of means	68.9
var of means	0
Confidence Level(95.0%)	NA

dark, small, few chips

	means
	13.51667
	12.18333
	27.86667
mean of means	17.85556
var of means	75.6112
Confidence Level(95.0%)	21.60076

dark, small, lots of chips

	means
	29.53
	25.68
	27.38
	55.60
mean of means	34.55
var of means	199.444
Confidence Level(95.0%)	22.47201

light, large, few chips

	mean
	40.25
	43.41667
mean of means	41.83333
var of means	5.013889
Confidence Level(95.0%)	20.11807

medium, large, few chips

	means
	52.24167
	43.85833
mean of mean	48.05
var of means	35.14014
Confidence Level(95.0%)	53.25995

dark, large, few chips

	means
	27.83333
	27.71667
	24.65833
	20.95833
	28.25
	33.01667
mean of means	27.07222
var of means	16.18494
Confidence Level(95.0%)	4.221926

light, large, lots of chips

	means
	83.575
	74.425
mean of means	79
var of means	41.86125
Confidence Level(95.0%)	58.13064

medium, large, lots of chips

	means
	80.60833
	83.7
	83.40833
mean of means	82.57222
var of means	2.913912
Confidence Level(95.0%)	4.240472

dark, large, lots of chips

	means
	67.075
	65.125
	72.125
	59.63333
mean of means	65.99
var of means	26.65696
Confidence Level(95.0%)	8.215553