COMPUTER VISION ANALYSIS
OF CHOCOLATE CHIP COOKIE CHARACTERISTICS

by

PIERRE FAA

B.S., UNIVERSITY CENTER OF DSCHANG (CAMEROON), 1985

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Grain Science and Industry

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1988

Approved by:

Major Professor

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

AKNOWLEDGEMENTS

The author wishes to express his sincere gratitude to the following people:

Professor Joseph Ponte and Dr. Jon Faubion for their pati ence, guidance and support throughout my study program and this research.

Dr. Carol Klopfenstein for serving on my graduate committee.

Dr. Dallas Jonhson for his great help with the statistical analysis of results.

Dr. Laura Hansen for her valuable assistance in the sensory evaluation study.

Mr. Hugh Bright and Mr. Jerry McMaster for their assistance in the cookie formulation and baking procedure.

Ms Susan Saunders for her help in my getting familiar with the image analysis system.

The faculty, staff, and graduate students of the Grain Science Department for all their help and support during my years at Kansas State.

INTRODUCTION

It is estimated that 90% of all industrial inspection activities requiring vision will be done with computer vision systems within the next decade (Gevarter, 1982 cited by Zuech and Miller, 1987). Machine vision systems avoid the high cost of labor while providing objective, accurate and reliable measurements of product quality (Swientek, 1987). They also permit a 100% inspection of the product because of the high speed capability of most systems.

The purpose of this project was to study two parameters associated with chocolate chip cookie quality using image analysis. The parameters were: the number of chocolate chips on the cookie top surface and the total surface area occupied by the chips. The variability in those parameters was studied among cookies baked with chips of the same size, and between cookies with chocolate chips of different counts(sizes). Correlations among the weight of chips in the formula, the number of chips on the cookie top surface and the total area occupied by those chips were also determined. Because the efficiency and practical importance of image analysis would be enhanced if the analysis was associated with sensory attributes of cookies, a sensory panel was asked to evaluate cookies based on the amount of area covered by chips on the cookie top surface.

LITERATURE REVIEW

IMAGE ANALYSIS

Zuech and Miller (1987) define machine vision as "the process of producing useful symbolic descriptions of a visual environment from image data". Image analysis is a part of this process. The same authors define image analysis as "the process of generating a set of descriptors or features on which a decision about objects in an image is based". Simply stated, machine vision is the integration of television and computers to replace and/or improve some of the functions of human inspectors (Levine 1988). Machine vision systems have two primary elements. The first element is the imaging sensor. This can include cameras (vidicon or solid state), X-ray, ultra-violet, infrared, or ultrasonic sensors. The image from the sensor is stored by the second major element, the computer, as an array of pixels (picture elements) which contains information about the grey levels at each point in the image. The quantity of information that can be extracted from this array is limited by a number of factors: the number of pixels available in the array, the width of the grey scale used, the size and speed of the computer hardware available, and the imagination and knowledge of the system programmer (Levine 1988). The computer stores reference information, processes the image data, regulates control devices, and generates reports. In addition to the system components themselves, additional factors must be considered or optimized, one of which is specimen illumination. Some of the common constant illumination sources are tungsten, quartz halogen, quartz iodine, fluorescent and mercury (or xenon) arc lamps. In addition, various flash lamps, lasers and light emitting diode (LED) sources may be employed. Light intensity must be sufficient to obscure interferences from sources. Also important is the fact that the contrast of the object against its background must be greater than the local lighting variation around the feature of interest within the object. All of these factors are controlled, to a great extent, by the manner in which the specimen is illuminated. The range of resolution of machine vision systems is based on the density of its pixel array. This varies from 1 x 16 pixels to 1024 x 1024 pixels. The number of divisions in the gray scale affects the system sensitivity and varies as well, ranging from 2 to 256 (Levine 1988). The more divisions, the greater the potential sensitivity.

The amount of time required to obtain and process an image affects the way that the system can or must be employed. Matrix cameras generally require 1/60 second to scan an image and send it to the processor. However, at this speed only information relating to the presence or absence of an object can be processed by the central processing unit (CPU). For most applications of machine vision systems, more information than simple object presence or absence is required for analysis. Thus, slower throughputs are necessary. Linear array cameras can provide a greater resolution than matrix array cameras. However, this higher resolution results in a proportionately higher amount of data to process and longer processing times.

Gagliardi et al (1984) point out that a video inspection (image anlysis) system should be designed around the particular characteristics of the specimen to be inspected so as to allow the system to perform its inspection task with optimum speed and precision. If the system is to be used "in plant", the quality control functions required as part of a production line must be defined in terms that can be translated into system design criteria. The authors outline the general requirements for an inspection vision system suitable for use in production situations as follows.

1) It must minimize acceptance of unacceptable products and the rejection of acceptable products.

2) The system must operate at line speeds.

3) The system must tolerate the variations normally considered acceptable in the product.

4) The system must function properly in a factory environment.

5) The system must permit adjustment of the accept/reject parameters; in other words, it must be operator selectable.

6) It must be higly reliable with little or no downtime.

7) The system must have benefits that are consistent with system's cost.

According to Russ et al (1988) an ideal image analysis system for use in quality control should have a sufficiently flexible software, the ability to deal with difficult images and to apply automatic editing operations to the discriminated images. The system should also be able to store large amounts of data from many images and provide a statistical package of sufficient power to interpret data.

Applications

Although a new technology, machine vision has already found many applications in diverse situations. Some examples of machine vision applications in the food industry (Swientek, 1987) include:

1) The detection of bones in fish filets.

2) The screening of coffee beans and pistachios using ultra-violet (UV) light.

3) The classification of fish by species, size and weight.

4) The inspection of cuts of beef for automatic trimming of fat.

5) the measurement of sheeted dough thickness using structured light.

6) The automatic sorting of potatos according to length and diameter.

7) Package and container scanning for cap positioning, label placement, fill heights, seals, dents and defects.

8) The color analysis of snack products, bakery goods, french fries, fruits and vegetables.

9) The optical sorting of randomly diced carrots and potatoes.

10) The electro-optical grading and sorting of lemons, apples, potatoes, cucumbers, tomatoes and bell peppers.

Recently an increasing number of scientific studies have been conducted using image analysis. Unklesbay et al (1983) used computerized image analysis to determine the level of browness on bottom surface of pizzas. They characterized the technique as "objective measurement", mentioning that it could be useful in cases where rapid, non destructive testing for available lysine in baked food is needed. Heyne et al (1985) used computerized image analysis as an automated quality control technique for rapidly and precisely determining the protein quality of simulated pizza crusts without physically damaging them. Zayas et al (1986) used image analysis to discriminate variables of grain morphology in order to differentiate among wheat classes and varieties. These examples show the role that image analysis is coming to play in research and quality control.

COOKIES

In 1984, 2.006 billion pounds of cookies were produced in the United States resulting in average per capita consumption of 8.5 pounds of cookies (Bednarcyk, 1987). Clearly cookies are a major baked snack product. Usually, what first captures the attention of cookie consumers is their top surface appearance. In most cases, cookies are packaged so as to display this surface to the consumer. Advertisements emphasize features associated with the cookie top surface. Thus the top surface appears then to be an important part of a cookie's quality. Since the consumer evaluates this aspect, the top surface characteristics should, thus, be able to be anlyzed by image analysis. From the time of the Aztec empire under Montezuma II to the present, "xocoalt" which evolved later to "chocolate", has always held a special place among all foods. Chocolate has become the "infinitely and almost universally desirable" food (Morton and Morton, 1986). This partly explains the popularity of chocolate chip cookies. and it also suggests that the chocolate chips visible on the top surface of cookies are one of the most if not the most important top surface characteristics of chocolate chip cookies.

Cookies are made from soft wheat flour and characterized by a formula high in sugar and shortening and relatively low in water. Cookies vary in formula and in type of manufacture. Hoseney (1986) classifies cookies by the way the dough is placed on the baking band as follows:

Rotary-mold cookies in which the dough is forced into molds on a rotating roll. A popular example of this is the Oreo cookie.

Cutting-machine cookies in which the dough is made into a continuous sheet and the product cut from it. Typical examples of cutting machine cookies are animal cookies and gingerbread cookies.

Wire-cut cookies in which a relatively soft dough is extruded through an orifice and cut to size, usually by a reciprocating wire. This type is the most popular in the

U.S. (Bright 1987) and includes many combinations of ingredients such as chocolate chips, chocolate chip with peanut butter, nuts and cinnamon, raisin and raisin paste.

Sugar wafers; this type is considered a cookie only because it doesn't fit elsewhere. The formula contains no sugar, essentially no fat, and a high amount of water.

Another classification scheme (Bright 1987) includes, in addition to rotary mold and wire cut cookies, deposited cookies (the most popular of this type being Danish Butter Cookie) and extruded cookies.

MATERIALS AND METHODS

Cookie Baking

The formulation used for baking chocolate chip cookies was as shown in Table I (adapted from Bright, 1987). A single batch yielded 18 cookies and required 200g of flour. The flow chart in Figure 1 describes the method of baking (adapted from Bright, 1987). Creaming and mixing were done in an Hobart N-50 mixer. Cookies were deposited on trays with a 26 cm^3 scoop and baked at 350^0F (177^0C) for 11 minutes.

Chocolate Chips

Chocolate chips were obtained from Ambrosia Chocolate Company (Milwaukee WIS. 53203). Three sizes were used:

Table I

Formulation for the baking of chocolate chip cookies

INGREDIENTS	% (flour weight basis)
Flour (pastry)	100
Shortening	50
Brown sugar	54
Granulated sugar	54
Whole egg (dry)	8
Water	30
Sodium bicarbonate	1.0
Salt	1.5
Chips	66

Figure 1.   Flow chart of baking procedure.

Sugar, Shortening
Salt, Eggs

↓

Cream, 60 sec. (speed 1)

↓

Add Water

↓

Cream, 60 sec. (speed 1)

↓

Cream 60 sec. (speed 2)

↓

Scrape Bowl

↓

Cream 3 min. (speed 3)

↓

Add flour and
Sodium Bicarbonate

↓

Mix, 2 min. (speed 1)

↓

Add Chips

↓

Mix, 15 sec. (speed 1)

1) The large size or 1000 count chips had a count per pound of 1000 ± 100 chips.

2) The medium size or 2000 count chip had 2000 ± 1000 chips per pound.

3) The small size or 4000 count chip had 4000 ± 200 chips per pound.

All counts per pound were verified by weighing three different samples of 100 chips for each size. For the large size chips, the three measures were 45.9g, 45.7g and 45g. This gave an average weight of 45.5g, that is about 996 chips per pound. The three measures of the medium size chips were 22.8g, 22.9g, and 23g. This gave an average weight of 22.9g, that is about 1978 chips per pound. For the small size chips, 11.2g, 11.4g and 11.3g were obtained, thus giving an average weight of 11.3g, that is about 4009 chips per pound. All these counts per pound fall in the range specified by the manufacturing company. Those data are summarized in Table II.

Image Analysis

The image analysis system used in all the studies reported below consisted of:

TV camera (Panasonic model wv-150, Matsushita Communication Industrial Co. Ltd. Japan)

TV Monitor (RCA, model TC1910, RCA Closed Circuit

Table II

Average number of chocolate chips per pound for each chip count.

COUNT	AVERAGE NUMBER OF CHIPS
1000	996
2000	1978
4000	4009

Video Equipment, Lancaster PA)

Apple II E based image analysis computer (Dapple System, Suuyude CA)

Lighting system composed of two 50 watt tungsten lamps positioned oppositely at each end of a horizontal rod

A 8510 PRINTER (Itoh Electronics, Inc. Japan)

Figure 2 illustrates this system.

The method used for the analysis of all cookies consisted of the acquisition of separate images of individual cookies by the computer, their modification (if necessary) in order to isolate the feature(s) of interest (that is chips), measurement of these feature(s) of interest, and calculation of required statistical data (total, mean value, and variance). The precision of this image analysis system was assessed by analyzing nine "Chips Ahoy" cookies from Nabisco for their area, the number of visible chips on the top surface, and the total surface area of those chips. Each cookie was analyzed three times for the above parameters. The low standard deviations obtained suggested a good reproducibility of the system(see Appendixes A, B, and C).

Cookie Baking: Chip Count Studies

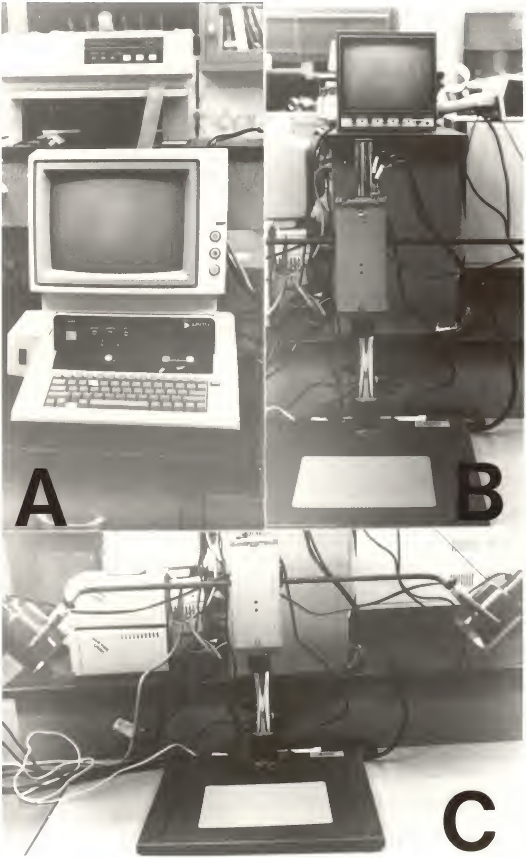
Eight batches of cookies from each chip count (1000, 2000, 4000) were baked and three cookies randomly picked

Figure 2. Picture of image analysis system.

A = Computer and Printer

B = Video camera and Tv Monitor

C = Lighting System


A
B
C

from each batch (using random numbers). All were then image-analyzed to quantitate the number of chips visible on the cookie top surface and the total area occupied by those chips.

Results were interpreted using the SAS 5.16 (1985) program for the analysis of variance.

Cookie Baking: Chip Weight Studies

Batches of cookies were baked with a variation in the weight of chocolate chips in the formula from 118g to 146g (118, 120, 122,...144, 146g). In each batch, three cookies were randomly picked. Each cookie was image-analyzed to determine the number of visible chips on the cookie top surface, and the total area occupied by those chips. The statistical program SAS 5.16 (1985) was used to test the null hypothesis that rho, the population correlation coefficient, was equal to zero for each of the correlations that were to be evaluated. The method described by Snedecor and William (1980) was used to obtain a confidence interval for rho from its estimate r.

Preference Testing

Sensory Panel Composition

Thirty four subjects (aged 18 to 42) served as panelists. All were untrained, volunteers, and either students, faculty or staff at Kansas State University.

Sample Preparation

Cookies were baked with two different treatments using the formula and procedure described previously. In the first treatment cookie batches were baked with equivalent weights of chips (132g) in the formula, but with the chip size being varied from 4000 count, to 1000 count. In the second treatment, cookie batches were baked with chips of the same size (2000 count), but with the weight of chips in the formula being varied (118g, 132g, and 146g).

Sample Presentation

Plastic plates with three compartments were used to present cookies to panelists. Plates contained either one cookie from each chocolate chip size or one cookie from each chip weight. Within each treatment, each cookie was coded with a three digit number, and randomly placed in one of the compartments on the plate to avoid selection bias. Each plate was wrapped with plastic wrap to preserve freshness.

Panelists were asked to evaluate (see questionaire in Figure 3) the cookies according to how much space the chips occupied on the top surface of the cookie and according to how well they liked the amount of space covered by chips. Each panelist evaluated two plates of cookies; one of each treatment.

Figure 3. Sensory evaluation questionaire.

CHOCOLATE CHIP COOKIES EVALUATION

AGE (OPTIONAL):
SEX:
NATIONALITY:

Evaluate the cookies left to right.

1. Evaluate the top surface of each cookie according to the amount of space occupied by chips. Check the box that best describes how much of the surface is covered by chips:

	COOKIE ______	COOKIE ______	COOKIE ______
Very much covered	____	____	____
Much covered	____	____	____
Moderately covered	____	____	____
Slightly covered	____	____	____
Not covered	____	____	____

2. Please rate each cookie checking one statement on the following scale, to indicate how well you like the amount of space occupied by chips on the top surface of the cookies:

	COOKIE ______	COOKIE ______	COOKIE ______
Like very much	____	____	____
Like moderately	____	____	____
Like slightly	____	____	____
Neither like/dislike	____	____	____
Dislike slightly	____	____	____
Dislike moderately	____	____	____
Dislike very much	____	____	____

3. Comments:

For the statistical analysis of the panel response results, numerical values were assigned to the scales used in the questionnaire. In the first question, "very much covered" was assigned the highest value of 1. For the second question, the scale ranged from 7 for "like very much" to 1 for "dislike very much". Analysis of variance with mean separation using the LSD method (SAS 5.16, 1985) was conducted on each question.

RESULTS AND DISCUSSION

VARIABILITY OF COOKIE TOP SURFACE CHARACTERISTICS AS A FUNCTION OF CHOCOLATE CHIP COUNT

The purpose of these experiments was to study the variability in the number of chips (on the cookie top surface) and the total area occupied by those chips, among cookies baked with chips of the same size, and between cookies baked with chips of different sizes.

Tables III - V present the data obtained. Figures are of independant measurements of three cookies. Figures 4 - 9 illustrate the results obtained within each chip count for both chip number and area. These graphs show that the technique found that variability existed in both the average number of chips and the average total visible area of those chips. Figure 10 and 11 show the regroupment of

Table III

Mean number of chips and mean total area occupied by chips for cookies baked with large size chocolate chips.

BATCH	MEAN NUMBER OF CHIPS	MEAN TOTAL AREA (cm^2)
I_1	5.3	0.553
I_2	3.7	0.571
I_3	3.3	0.420
I_4	2.7	0.450
I_5	5.3	0.910
I_6	7.7	1.506
I_7	6	0.922
I_8	3.7	0.893

I= 1000 chip count (large size)

Table IV

Mean number of chips and mean total area occupied by chips for cookies baked with medium size chocolate chips.

BATCH	MEAN NUMBER OF CHIPS	MEAN TOTAL AREA (cm^2)
II_1	10.3	1.196
II_2	8	1.344
II_3	7.3	1.246
II_4	10.3	1.429
II_5	10	1.757
II_6	7	0.922
II_7	8	1.387
II_8	6.3	0.734

II= 2000 chip count (medium size)

Table V

Mean number of chips and mean total area occupied by chips for cookies baked with small size chocolate chips.

BATCH	MEAN NUMBER OF CHIPS	MEAN TOTAL AREA (cm^2)
IV_1	12	1.628
IV_2	9.7	1.357
IV_3	11	1.553
IV_4	15	2.036
IV_5	14.7	1.851
IV_6	14	1.562
IV_7	13.3	1.448
IV_8	15	1.881

IV= 4000 chip count (small size)

Figure 4. Mean number of chips, as a function of cookie batch, detected on the cookie top surface for 1000 count chips.

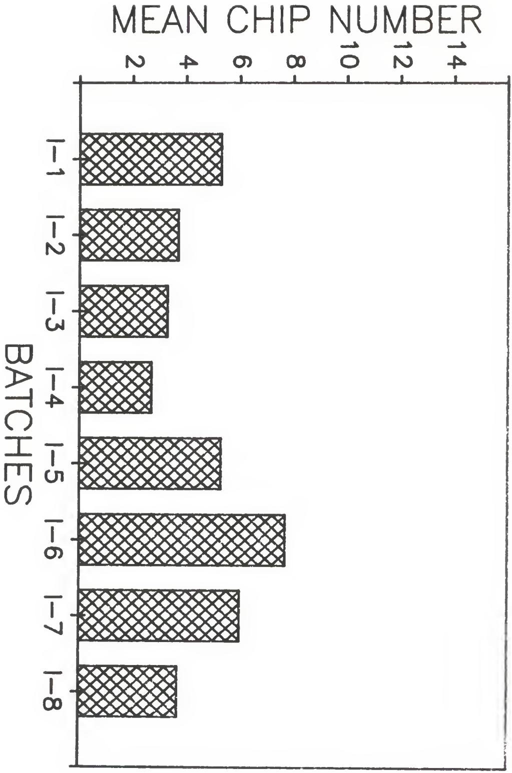

MEAN CHIP NUMBER
2
4
6
8
10
12
14
I-1
I-2
I-3
I-4
I-5
I-6
I-7
I-8
BATCHES

Figure 5. Mean total area of chips (cm^2), as a function of cookie batch, detected on the cookie top surface for 1000 count chips.

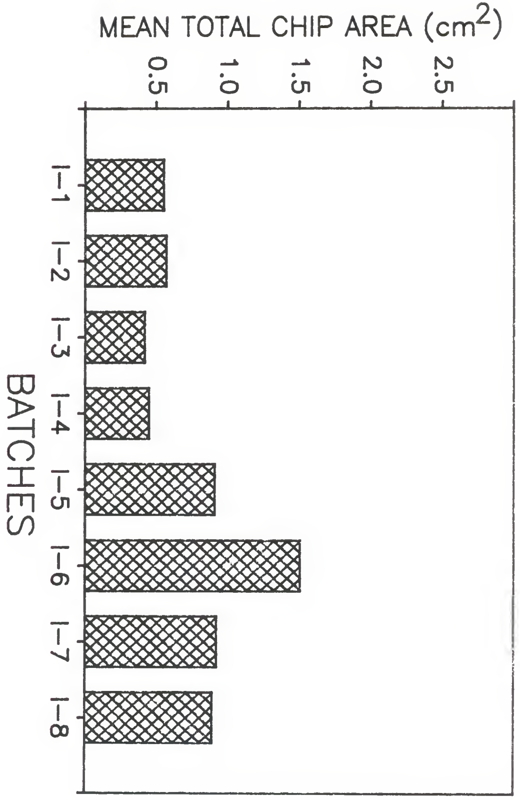

MEAN TOTAL CHIP AREA (cm2)
0.5
1.0
1.5
2.0
2.5
I-1
I-2
I-3
I-4
I-5
I-6
I-7
I-8
BATCHES

Figure 6. Mean number of chips, as a function of cookie batch, detected on the cookie top surface for 2000 count chips.

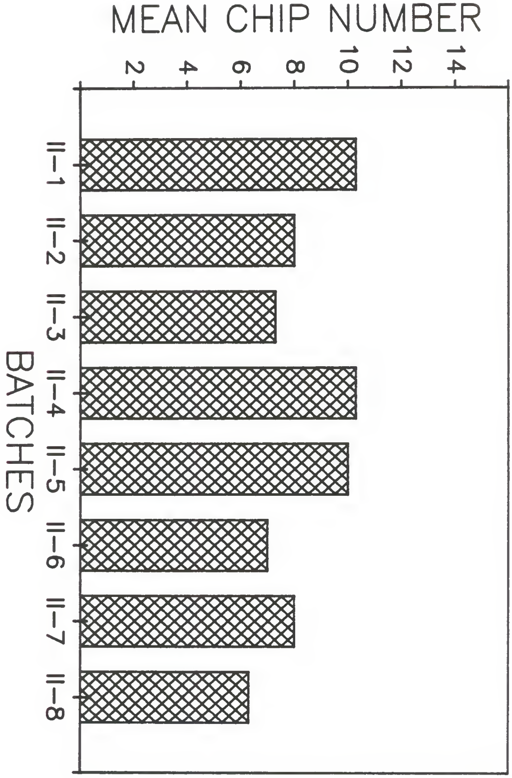

MEAN CHIP NUMBER
2
4
6
8
10
12
14
II-1
II-2
II-3
II-4
II-5
II-6
II-7
II-8
BATCHES

Figure 7. Mean total area of chips (cm^2), as a function of cookie batch, detected on the cookie top surface for 2000 count chips.

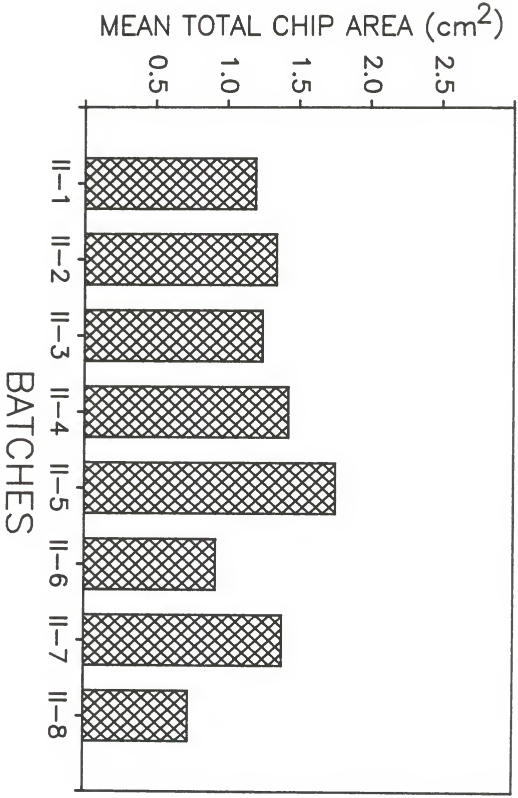

MEAN TOTAL CHIP AREA (cm2)
0.5
1.0
1.5
2.0
2.5
II-1
II-2
II-3
II-4
II-5
II-6
II-7
II-8
BATCHES

Figure 8. Mean number of chips, as a function of cookie batch, detected on the cookie top surface for 4000 count chips.

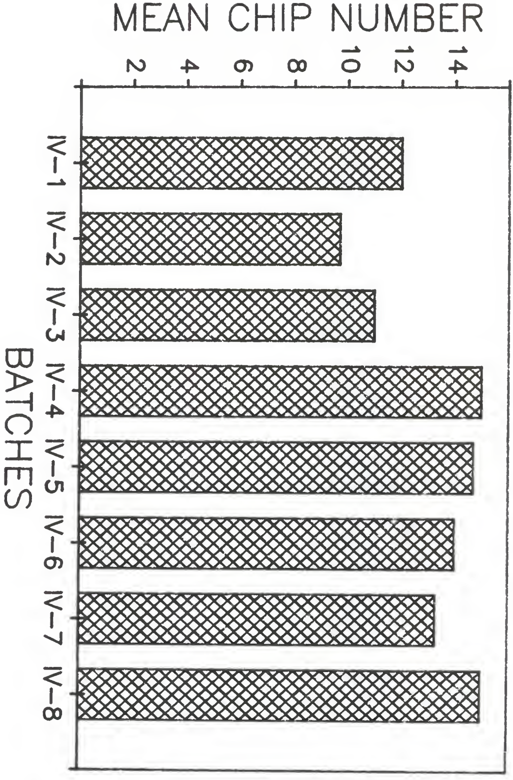

MEAN CHIP NUMBER
2
4
6
8
10
12
14
IV-1
IV-2
IV-3
IV-4
IV-5
IV-6
IV-7
IV-8
BATCHES

Figure 9. Mean total area of chips (cm^2), as a function of cookie batch, detected on the cookie top surface for 4000 count chips.

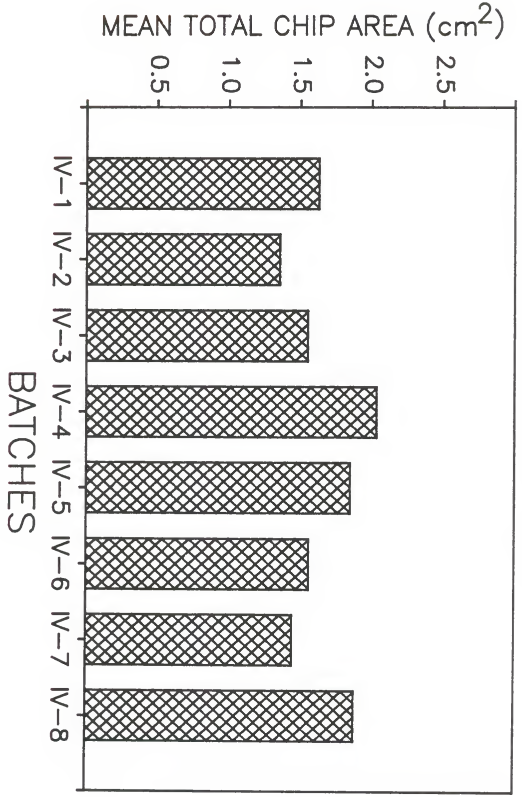
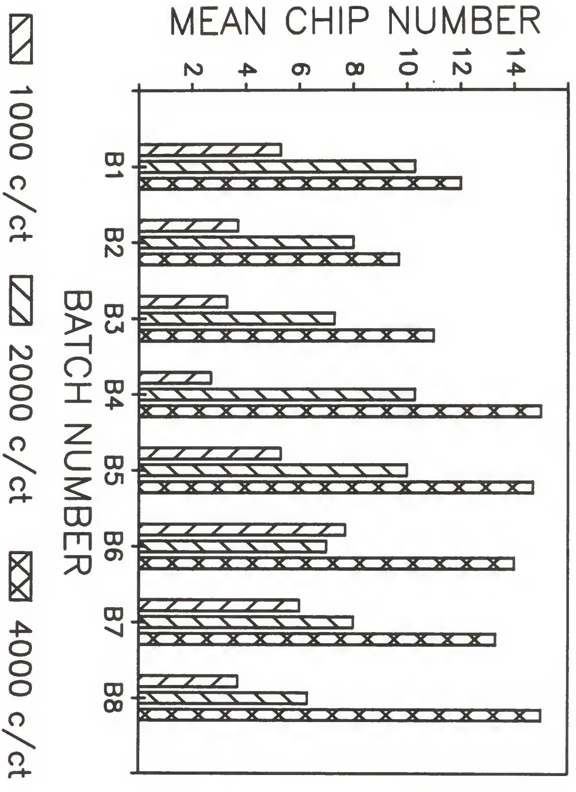
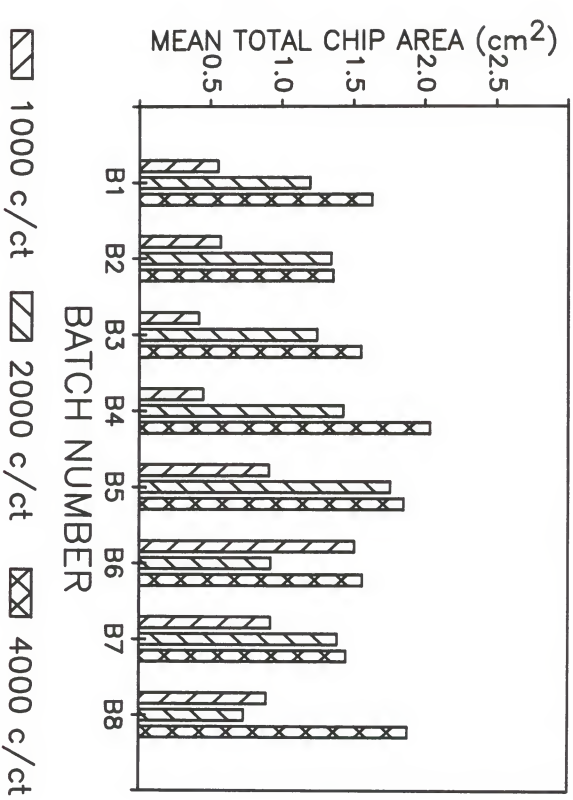

MEAN TOTAL CHIP AREA (cm2)
0.5
1.0
1.5
2.0
2.5
IV–1
IV–2
IV–3
IV–4
IV–5
IV–6
IV–7
IV–8
BATCHES

Figure 10. Mean number of chips visible on the cookie top surface as a function of batch. Here batches had different chip counts.



MEAN CHIP NUMBER
2
4
6
8
10
12
14
B1
B2
B3
B4
B5
B6
B7
B8
BATCH NUMBER
1000 c/ct
2000 c/ct
4000 c/ct

Figure 11. Mean total area (cm^2) of chips visible on the cookie top surface as a function of batch. Here batches had different chip counts.


MEAN TOTAL CHIP AREA (cm2)
0.5
1.0
1.5
2.0
2.5
B1
B2
B3
B4
B5
B6
B7
B8
BATCH NUMBER
1000 c/ct
2000 c/ct
4000 c/ct

batches with different chip counts in order to affect the between batch comparison of mean number of chips and the mean total chip area. These figures tend to suggest that the smaller the chip size, the higher the number of visible chips on the cookie top surface, and the total surface area of those chips.

The statistical analysis of these results is summarized in Tables VI - IX.

Image Analysis of Chip Number

The p-value of 0.6616 (Table VI) for batches within a chip count shows that there were no significant differences in the number of chips in cookies from different batches of cookies having the same chocolate chip size. Thus, even though Figures 4, 6, and 8 showed apparent differences, this variability was not statistically significant. Since the chips were, preumably, randomly distributed in the dough, this result demonstrates the precision of the image quantification technique. However, the difference was significant for the number of chips between cookies of different chip counts as suggested by the very low p-value of 0.0001 and the LSD test. The latter showed that cookies with the smallest chip size had the highest number of visible chocolate chips, followed by cookies with medium chip size, whereas the 1000 chip count

Table VI

Analysis of variance for variability in chip number,on the cookie top surface, within and between chip counts.

Source of variation	p-value
Batches with a single chip count	0.6616
Batches of different chip counts	0.0001

DF batch: 7 DF error: 48 DF chip count: 2

MS batch: 10.4107 MSE: 14.6111 MS chip count: 422.6806

F value 0.71 F value: 28.93

Table VII

LSD test between batches of differing chip counts for the mean number of chips visible on the cookie top surface.

Chip count	Mean*
IV	13.083 a
II	8.417 b
I	4.708 c

* means with the same letter are not significantly different.

Alpha = 0.05 DF = 14 MSE = 8.7441 LSD = 1.8308

Table VIII

Analysis of variance for variability in the total area occupied by chips on the cookie top surface within and between chip counts.

Source of variations	p-value
Batches with a single chip count	0.7062
Batches of different chip counts	0.0004

DF batch: 7	DF error: 48	DF chip count: 2
MS batch: 0.1938	MSE: 0.2947	MS chip count: 4.7232
F value: 0.66		F value: 16.03

Table IX

LSD Test between batches of differing chip counts for the mean total area occupied by chips visible on the cookie top surface.

Chip count	Mean*
IV	1.664 a
II	1.252 b
I	0.778 c

* means with the same letter are not significantly different.

Alpha = 0.05 DF = 14 MSE = 0.3295 LSD = 0.3554 cookies had the lowest number of chips visible. This confirms the presumptions from Figure 10. It also demonstrates the ability of the image analysis procedure employed here to discriminate small differences in surface characteristics.

Image Analysis of the Total Area of Visible Chips

The total area of visible chips followed the same trend as did the number of chips. Batches prepared the same way showed a high p-value, suggesting that there were no significant differences in the total area occupied by chips in cookies from batches of the same chip count. Again, this is in spite of the fact that figures 5, 7 and 9 did show some variability. The low p-value of 0.0004 and the LSD test results suggests that chip size significantly affected the total area occupied by visible chips. Thus, cookies with the smallest chip size possessed the highest mean total area of chips. The lowest mean total area of visible chips was observed in cookies with the largest chip size, confirming observations from Figure 11.

Results from these studies show that there was no significant variability in either the number of chips or the total area of visble chips between batches of cookies containing equal amount of chips of the same count. Significant differences could be measured in the number of chips and the total area of chips visible between batches containing equal amounts of different sized chips.

CORRELATION BETWEEN MEASURED TOP SURFACE CHARACTERISTICS AND THE FORMULA WEIGHT OF CHOCOLATE CHIPS

This study was undertaken to determine:

1) The correlation between the amount of chocolate chips in the cookie formula and the number of visible chips on the top surface of the cookie.

2) The correlation between the amount of chips in the formula and the total surface area of visible chips on the cookie top surface.

3) The correlation between the number of chips on the top surface of the cookie and the total area of those chips.

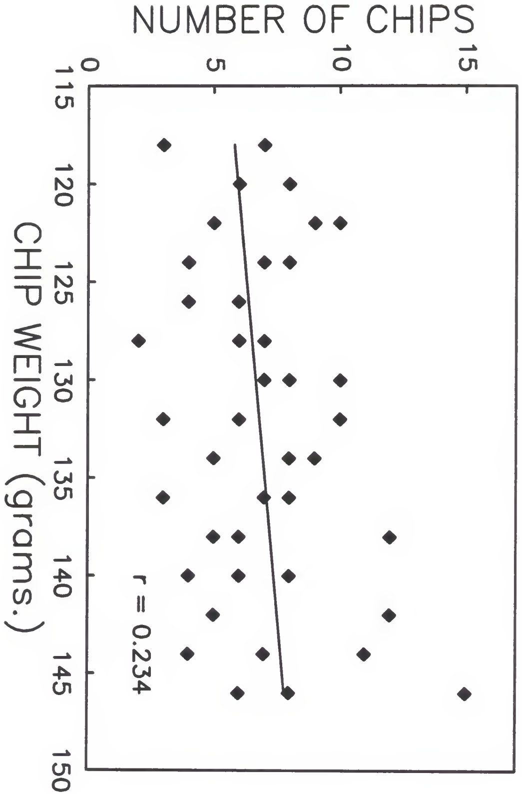
Table X presents the mean number and the mean total surface area of visible chocolate chips as a function of the weight of chips in the cookie formula. All Means are the result of independant measurements of three cookies. Figures 12 and 13 present the data in graphical form. Both tend to show a random distribution of visible chip number and total area of visible chips as function of the amount of chips in the formula. Figure 14 gives a merest suggestion of ordered distribution. Specifically, the total area of chips appears to increase with the number of visible chips. To make sure that results from Figures 13 and 14 were not affected by

Table X

Mean number and mean total area of visible chips on the cookie top surface as a function of the weight of chocolate chips in the formula.

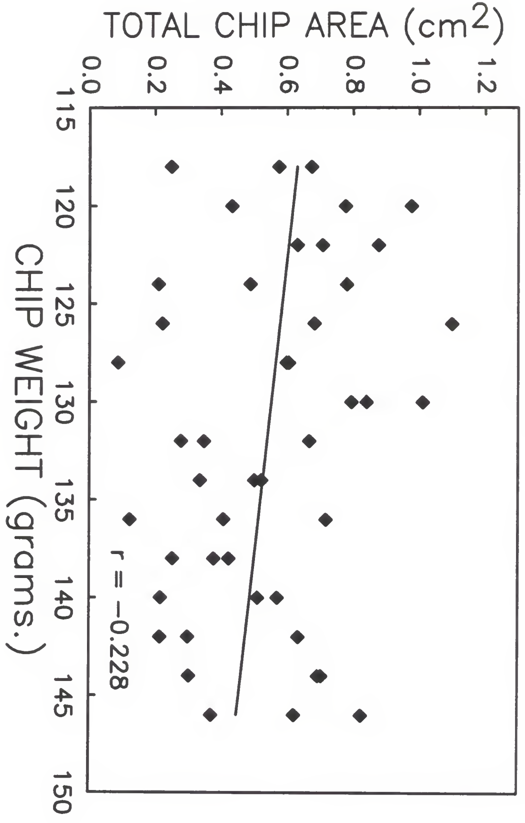
WEIGHT (g)	MEAN CHIP #	MEAN AREA (cm2)
118	4.3	0.499
120	7.3	0.728
122	8	0.738
124	6.3	0.493
126	5.3	0.668
128	5	0.430
130	8.3	0.882
132	6.3	0.432
134	7.3	0.453
136	6	0.416
138	7.7	0.352
140	6	0.433
142	7.3	0.384
144	7.3	0.567
146	9.7	0.607

Figure 12. Number of chocolate chips visible on the cookie top surface as a function of the weight of chocolate chips in the formula.



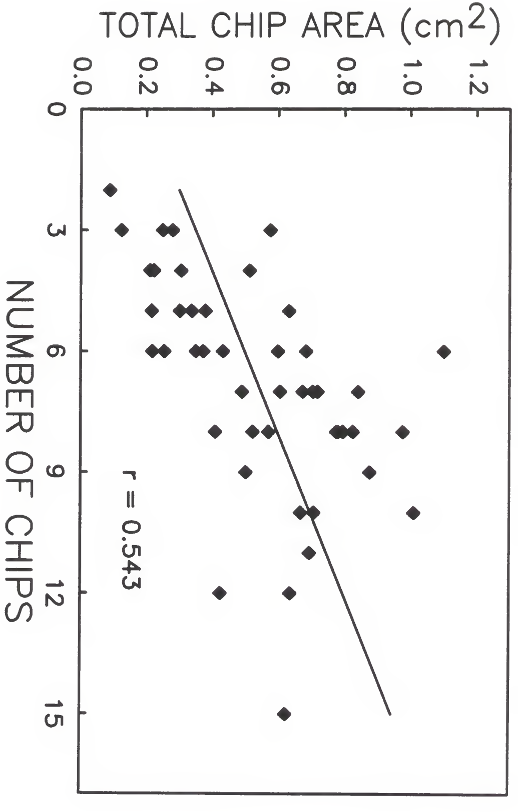
NUMBER OF CHIPS
0
5
10
15
115
120
125
130
135
140
145
150
CHIP WEIGHT (grams.)
r = 0.234

Figure 13. Total visible chip area (cm^2) on the cookie top surface as a function of the weight of chocolate chips in the formula.



TOTAL CHIP AREA (cm2)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
115
120
125
130
135
140
145
150
CHIP WEIGHT (grams.)
r = −0.228

Figure 14. Total visible chip area (cm^2) on the cookie top surface as a function of the number of chocolate chips on the cookie top surface.


TOTAL CHIP AREA (cm2)
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0
3
6
9
12
15
NUMBER OF CHIPS
r = 0.543

the cookie area, the total area occupied by visible chips on the cookie top surface was expressed as the percentage of cookie area. Figures 15 and 16 which were then obtained showed essentially the same trend, suggesting that cookie area did not affect the total surface area occupied by visible chips on the top surface of cookies. Table XI presents the results of the t test for rho for the correlations studied. There was not enough evidence ($p = 0.1224$) to reject the null hypothesis that rho = 0 for the correlation between the weight of chips in the formula and the number of visible chips on the cookie top surface. Thus, it may be concluded that the amount of chips in the cookie formula did not significantly affect the number of visible chips on the cookie top surface. This was somewhat surprising. Because of the high p-value, the same conclusion can be drawn for the correlation between the weight of chips in the formula and the total area of visible chips on the cookie top surface. It may be that the number of visible chips on the cookie top surface depends, to a certain extent, (among factors not controlled) on the completely random mixing of chocolate chips with the cookie dough, and the way cookie dough is deposited on the baking tray. The sampling across batches may explain the high variability which resulted in those non significant

Figure 15. Total visible chip area (cm^2), as a percentage of cookie area, on the cookie top surface as a function of the weight of chocolate chips in the formula.

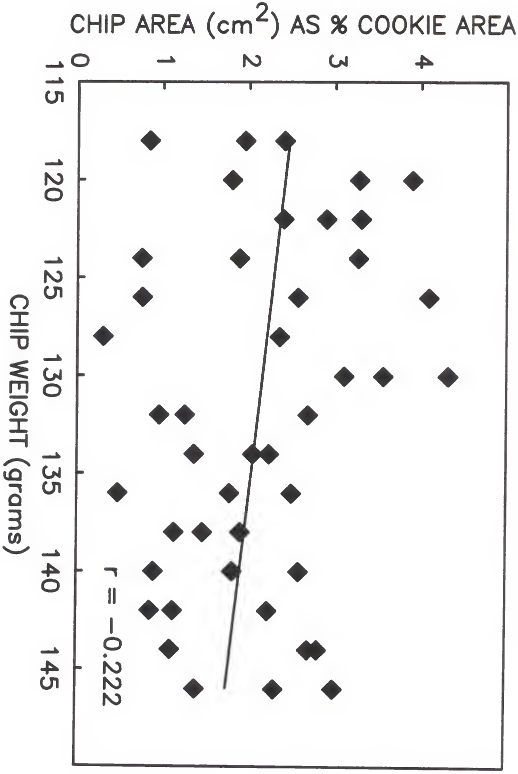

CHIP AREA (cm2) AS % COOKIE AREA
0
1
2
3
4
115
120
125
130
135
140
145
CHIP WEIGHT (grams)
r = −0.222

Figure 16. Total visible chip area (cm^2), as a percentage of cookie area, on the cookie top surface as a function of the number of chocolate chips on the cookie top surface.

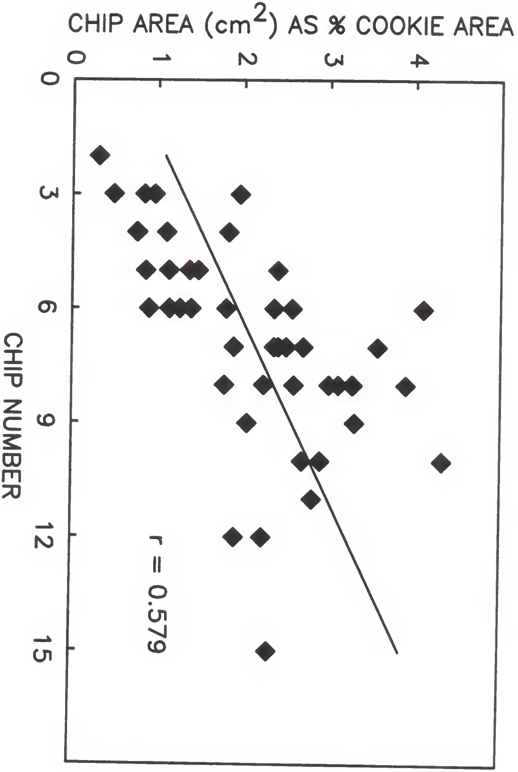

CHIP AREA (cm2) AS % COOKIE AREA
0
1
2
3
4
0
3
6
9
12
15
CHIP NUMBER
r = 0.579

Table XI

T-tests for rho (the population correlation coefficient).

Correlations	p-value
Weight * Number of Chips	0.1224
Weight * Area	0.1324
Number of Chips * Area	0.0001

correlations. Also, the variation in the amount of chips in the formula may not have been large enough to overcome the effect of the other parameters (most of which are difficult to control) on the number of visible chips on the cookie and the total area of those chips. The p-value of 0.0001 for the correlation between the number of visible chips on the cookie top surface and the total area of those chips, leads to the conclusion that there is a significant correlation between the number of visible chips on the cookie top surface and the total area of those chips. A 95% confidence interval for rho was determined from its estimate r of 0.543 and was found to be 0.327 < rho < 0.735. A positive correlation, thus, exists between the number of visible chips on the cookie and their total area, and we may be 95% certain that the correlation coefficient is found in the interval from 0.327 to 0.735. This interval tends to suggest a relatively low correlation coefficient. A major factor that may have contributed in weakening the correlation coefficient is the chip coverage on the top cookie surface. There is a great variability in the way chips are exposed on the surface of cookies. Thus, there may be have been few chocolate chips on the cookie top surface but each had exposed a large proportion of its surface. It may also happen that there are higher numbers of visible chips on the cookie top surface but each chip presents only a small percent of its surface for exposure. The conical shape of chips probably increases this variability of surface exposure.

It was found that there was no significant correlation between the weight of chocolate chips in the formula and the number of chips on the cookie top surface. Also, no significant correlation was found between the amount of chips in the cookie formula and the total area of visible chips on the cookie top surface. A positive correlation was found between the number of chocolate chips on the cookie top surface and the total area occupied by those chips.

SENSORY EVALUATION OF COOKIES

The objective of this study was to test the ability of humans to differentiate among chocolate chip cookies based on the amount of space covered by chips on the cookie top surface, so that comparison can be made with results obtained from image analysis. It was also intended to evaluate the extent that this amount of space influenced the preference of panelists, in order to see how consumer preference can be related to quantitative data from image analysis.

Results obtained and the statistical analysis of those results are summarized in Tables XII-XXI.

Cookies containing chips of different sizes

No significance differences ($p = 0.1360$, Table XVI) were noted by panelists in the way the different sized chips covered the top cookie surface. One explanation for this lack of discrimination may be that, cookies made with smaller chips might have a greater number of chips per cookie and, therefore, more chips on the top surface than the cookies made with larger chips. Thus the area of chip coverage could be similar. Results from image analysis of cookies from different chip sizes showed significant differences in the total area of visible chips on the cookie top surface. Possibly, the ability of machine vision

Table XII

Panel responses; amount of surface covered by visible chocolate chips of different sizes.

RESPONSE	CHIP SIZE		
	SMALL	MEDIUM	LARGE
Very much covered	6	3	2
Much covered	12	8	4
Moderately covered	7	12	16
Slightly covered	3	8	9
Not covered	3	0	0

Table XIII

Hedonic response; amount of surface covered by visible chocolate chips of different sizes.

RESPONSE	CHIP SIZE		
	SMALL	MEDIUM	LARGE
Like very much	11	7	1
Like moderately	7	8	8
Like slightly	3	10	9
Neither like/dislike	6	3	8
Dislike slightly	0	2	1
Dislike moderately	2	1	2
Dislike very much	2	0	2

Table XIV

Panel responses; amount of surface covered by visible chocolate chips by weight in the formula (all chips were 2000 count).

RESPONSE	CHIP WEIGHT		
	LOW	NORMAL	HIGH
Very much covered	0	5	8
Much covered	3	6	11
Moderately covered	6	14	10
Slightly covered	18	9	5
Not covered	7	0	0

Table XV

Hedonic response; amount of surface covered by visible chocolate chips by weight in the formula (all chips were 2000 count).

RESPONSE	CHIP WEIGHT		
	LOW	NORMAL	HIGH
Like very much	0	6	13
Like moderately	7	9	11
Like slightly	5	12	4
Neither like/dislike	5	4	5
Dislike slightly	6	0	0
Dislike moderately	4	2	1
Dislike very much	7	1	0

Table XVI

Analysis of variance; amount of surface covered by visible chocolate chips of different sizes.

Source of variation	p-value
Chip size	0.1360

DF chip size = 2 DF error = 90

MS chip size = 2.0753 MSE = 1.0172

F value = 2.04

Table XVII

Analysis of variance; hedonic response on the amount of surface covered by visible chocolate chips of different sizes.

Source of variation	p-value
Chip size	0.0777

DF chip size = 2 DF error = 90

MS chip size = 6.5269 MSE = 2.4824

F value = 2.63

Table XVIII

Analysis of variance; amount of surface covered by visible chocolate chips by weight in the formula.

Source of varition	p-value
Chip weight	0.0001

DF chip amount = 2 DF error = 99

MS chip amount = 20.2059 MS error = 0.9251

F value = 21.84

Table XIX

LSD test; amount of surface covered by visible chocolate chips by weight in the formula.

Chip weight	Mean*
High	3.647a
Normal	3.206a
Low	2.147b

* means with the same letter are not significantly different

alpha level = 0.05   DF = 99   MSE = 0.9251   LSD = 0.46288

Table XX

Analysis of variance; hedonic response on the amount of surface covered by visible chocolate chips by weight in the formula.

Source of variation	p-value
Chip weight	0.0001

DF chip amount = 2 DF error = 99

MS chip amount = 48.8922 MSE = 2.3464

F value = 20.84

Table XXI

LSD test; hedonic response on the amount of surface covered by visible chocolate chips by weight in the formula.

Chip amount	Mean*
High	5.853a
Normal	5.206a
Low	3.529b

* means with the same letter are not significantly different

alpha level = 0.05 DF = 99 MSE = 2.3464 LSD = 0.73717 system to discriminate is greater than human vision. Alternatively, the perception of areas covered was altered by some other unrelated visual cue. No significant differences (p = 0.00777, Table XVII) in panelists preference were found among treatments with regard to how well panelists liked the amount the chips covered the top surface of cookies. Because panelists could not discern differences between treatments, it would be unlikely for them to have preferences between treatments. However, it is difficult to determine whether or not the panelists based their preference only on the amount of space covered by chocolate chips, or if other factors such as the texture or the color of the cookie, affected their judgement. Furthermore, some panelists commented that, not only do they like a lot of chocolate chips on their cookie, but they would also like the chips to be evenly distributed throughout the cookie surface.

<u>Cookies containing different amounts of chips of the same size</u>

Significant differences (p = 0.0001, table XVIII) were found between batches of cookies made with different amount of chips. Panelists reported that high and normal amounts of chips covered the cookie top surface more than low level of chips (Table XIX). Apparently, consumers can more easily discriminate differences in area covered by chips when the amount of chocolate chips in the formula is altered but the size of chips is held constant. Mean separation (Table XIX) shows that panelists observed no significant differences in surface coverage of cookies baked with the high amount of chips and those with the normal amount of chips. This might be due to experimental factors difficult to control such as, random mixing of chips with the cookie dough, deposition of cookies on the baking tray, or, more likely, the limits of human discrimination.

Significant differences were found ($p = 0.0001$ Table XX) in the way panelists liked the amount of space covered by chocolate chips on cookies when the amount of chips in the formula was varied and chip size held constant. Cookies baked with high and normal amounts of chips were prefered to cookies from a formula with a low level of chips. This indicates that consumers prefer cookies with a greater amount of chips on the top surface. This also shows how consumer prefernce can be tied with the quantitative image data from image analysis. Specifications meeting consumers preference can be set, and their uniformity in the production process controlled by image analysis.

Consumer evaluation of chocolate chip cookies found no significant differences in the way chocolate chips of different counts (1000, 2000, and 4000) covered the cookie top surface, and in how well the chip coverage was liked. Significant differences were found in the way chocolate chips of different amounts (size held constant) in the cookie formula covered the cookie top surface and in the way the chip coverage was liked. The study suggests that the more the chips occur on the top cookie surface the better the cookie is liked. Possibly consumers prefer many small chips on the surface instead of a few large chips. The study also suggests that humans set their limits more broadly than necessary in evaluating the amount of area covered by chips.

CONCLUSIONS

This study showed that image analysis can be used as a means of monitoring the uniformity in the number of chips and the total area covered by chips on the cookie top surface. But the image analysis system used in this study was not able to make a distinction between chips and dark spots on the cookie. Image analysis thus needed the help of human vision.

Image analysis of chocolate chip cookies showed that the variability in the number of chips and the total area of chips was significant only between batches of different chip sizes. However, a human evaluation did not find any significant difference in the way chocolate chips of different chip sizes covered the cookie top surface. It thus seems that, quantitatively, human vision is less acute than machine vision, or that humans set their limits more broadly than necessary. The fact that machine vision is more acute than human vision indicates that the former should be valuable for quality control in production processes. The image analysis conducted in this study also indicated a correlation between the number of visible chips on the cookie top surface and the total area covered by those chips. Bakers may need to find an economical way of controlling the number of visible chips on the cookie top surface. They may for example deposit chips on each single cookie prior to baking instead of mixing them into the whole dough. This study suggested that consumers prefer cookies whose top surface is more extensively covered by chips, and in a uniform manner. Possibly consumers prefer many evenly distributed small chips on the cookie top surface instead of a few large chips. It may thus be useful to design a way of monitoring the even distribution of chips on the top surface of cookies. Cookies may be divided in sections or strips, and, the number of chips and the total area covered by these chips in each strip obtained through image analysis. Plots of strips against the number of chips and the total area covered by the chips, and statistical interpretations would then help assess the uniformity in chips distribution on the cookie top surface. Also, In order to associate human vision and preference with machine vision more efficiently, it might be better to start with a consumer evaluation of chocolate chip cookies first, then use the same cookies for image analysis.

LITERATURE CITED

BEDNARCYK, E. N. 1987. Nutrient Contributions of Cookies and Crackers. Cereal Foods World 32(5): 367-369.

BRIGHT, H. 1987. Snacks Biscuits and Cookie Seminar. King's College, Cambridge.

GAGLIARDI, R. G., SULLIVAN, D., and SMITH, F. N. 1984. Computer-Aided Video Inspection. Food Technology 38(4): 53-59.

GEVARTER, W., B. An Oveview of Computer Vision. National Bureau of Standards, Report No. NBSIR 82-2582, September 1982.

HEYNE, L., UNKLESBAY, N., UNKLESBAY, K., and KELLER, J. 1985. Computerized Image Analysis and Protein Quality of Simulated Pizza Crusts. Can. Inst. Food. Sci. Technol. J. 18(2): 168-173.

HOSENEY, C. R. 1986. Principles of Cereal Science and Technology. 251-256.

LEVINE, L. 1988. Machine Vision: A New Technology. Cereal Foods World 33(2): 242.

MORTON, M., MORTON, F. 1986. Chocolate. An Illustrated History. p. 2. Crown Publishers, Inc. New York.

RUSS, C. J., STEWART, D. W., and RUSS, J. C. 1988. The Measurement of Macroscopic Images. Food Technology 42(2): 94-102.

SNEDECOR, G. W., and COCHRAN, W. G. 1980. Confidence limits and tests of hypotheses about rho in: Statistical Methods, Seventh Edition. The Iowa State University Press: Ames, IA.

SWIENTEK J. R. 1987. Machine Vision Systems. Food Processing 48(4): 68-80.

UNKLESBAY, K., UNKLESBAY, N., KELLER, J., and GRANDCOLAS, J. 1983. Computerized Image Analysis of Surface Browning of Pizza Shells. J. Food Sci. 48(4): 1119-1123.

ZAYAS, I., LAI, F. S., and POMERANZ, Y. 1986. Discrimination Between Wheat Classes and Varieties by image analysis. Cereal Chem. 63(1): 52-56.

Zuech, N., and Miller, R. K. 1987. Machine Vision. Fairmont Press, Inc.

APPENDIX A

Mean cookie area and standard deviation.

Cookie	Area(cm^2)	Standard deviation
1	25.540	0.008
2	25.540	0.109
3	25.401	0.111
4	25.781	0.343
5	26.483	0.266
6	25.518	0.460
7	24.395	0.107
8	25.971	0.337
9	23.553	0.041
10	25.158	0.072

APPENDIX B

Mean number of visible chips (on the cookie top surface) and standard deviation.

Cookie	Chip number	Standard deviation
1	9.7	0.6
2	9.0	0.0
3	14.0	0.0
4	9.0	0.0
5	14.7	0.6
6	13.7	0.6
7	17.0	1.0
8	9.0	0.0
9	7.0	0.0
10	12.7	0.6

APPENDIX C

Mean total area (cm^2) occupied by visible chips (on the cookie top surface) and standard deviation.

Cookie	Chip area	Standard deviation
1	1.916	0.016
2	2.056	0.037
3	2.343	0.157
4	1.627	0.048
5	2.272	0.066
6	1.909	0.035
7	2.456	0.130
8	1.887	0.136
9	1.317	0.010
10	2.862	0.055

COMPUTER VISION ANALYSIS
OF CHOCOLATE CHIP COOKIE CHARACTERISTICS

by

PIERRE FAA

B.S., UNIVERSITY CENTER OF DSCHANG (CAMEROON), 1985

AN ABSTRACT OF A THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Grain Science and Industry

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1988

Two parameters of chocolate chip cookies were studied through image analysis: the number of chocolate chips on the cookie top surface, and the total surface area occupied by these chips.

For this study, the variability in those parameters were studied among batches of cookies of the same chip size, and between batches of cookies from different chip counts. No significant differences were found in the number of chips and the total area occupied by chips between batches of the same chip count. However, the variability in the number of chips and the total area occupied by chips was significant between batches different chip sizes.

Correlations among the weight of chips in the formula, the number of chips on the cookie top surface, and the total area occupied by those chips were also studied. There wsa no significant correlation between the weight of chocolate chips in the formula and the number of chips on the cookie top surface. No significant correlation was also found between the amount of chips in the cookie formula and the total area of visible chips on the cookie top surface. A positive significant correlation was found between the number of chocolate chips on the cookie top surface and the total area occupied by those chips.

A sensory panel was asked to evaluate cookies varying in the amount of surface area covered by chips. From that evaluation, it was found that there were no significant differences in the way chocolate chips of differing chip count (1000, 2000, and 4000) covered the cookie top surface, and in how well the chip coverage was prefered. Significant differences were found in the way chocolate chips of different amounts (size held constant) in the formula covered the cookie top surface and in the way the chip coverage was liked.