Quality Classification of Corn Tortillas using Computer Vision

Domingo Mery[*,a], Jorge J. Chanona–Pérez[b], Alvaro Soto[a], José Miguel Aguilera[a], Aldo Cipriano[a], Nayeli Veléz-Rivera[b], Israel Arzate-Vázquez[b], Gustavo F. Gutiérrez–López[b]

[a]*Pontificia Universidad Católica de Chile*
[b]*Instituto Politécnico Nacional de México*

Abstract

Computer vision is playing an increasingly important role in automated visual food inspection. However quality control in tortilla production is still performed by human operators which may lead to misclassification due to their subjectivity and fatigue. In order to reduce the need for human operators and therefore misclassification, we developed a computer vision framework to automatically classify the quality of corn tortillas according to five hedonic sub-classes given by a sensorial panel. The proposed framework analyzed 750 corn tortillas obtained from 15 different Mexican commercial stores which were either small, medium or large in size. More than 2300 geometric and color features were extracted from 1500 images capturing both sides of the 750 tortillas. After implementing a feature selection algorithm, in which the most relevant features were selected for the classification of the five sub-classes, only 64 features were required to design a classifier based on support vector machines. Cross validation yielded a performance of 95% in the classification of the five hedonic sub-classes. Additionally, using only 10 of the selected features and a simple statistical classifier, it was possible to determine the origin of the tortillas with a performance of 96%. We believe that the proposed framework opens up new possibilities in the field of automated visual inspection of tortillas.

Key words: Corn tortillas, computer vision, automated visual inspection, sensorial panel.

1. Introduction

Corn was the principal source of food for the pre-Columbian civilizations of the New World and today corn tortillas and derivative products are still the staple food of Mexico and Central America. Corn tortillas, corn chips and tortilla chips have also widely penetrated the Unites States market as well as various countries in Asia and Europe (Cortés-Gómez et al., 2005). In Mexico, where tortillas are consumed by 94% of the population, annual production and consumption is near 12 million tons of corn tortillas. Interestingly, 60% of those tortillas are processed in small stores called *tortillerias*) (Ayala-Rodríguez et al., 2009). The are three main levels of production and distribution of corn tortillas in Mexico: 1) small commercial scale (*tortillerias*), 2) medium commercial scale or supermarkets, and 3) large commercial scale or samples packed in plastic bags. Corn tortillas are produced from either traditionally milled nixtamal, which is a wet *masa*, or dehydrated *masa* flour. Both processes include the following steps: production of the nixtamal, elaboration of the dough (*masa*) followed by the shaping and baking of the tortillas (Arámbula-Villa et al., 2007). The use of modern machinery for commercial corn tortilla production has been utilized since the 1960s (Lind and Barham, 2004). A simplified scheme of typical machinery used in the small commercial scale *tortillerias* is shown in Fig. 1.

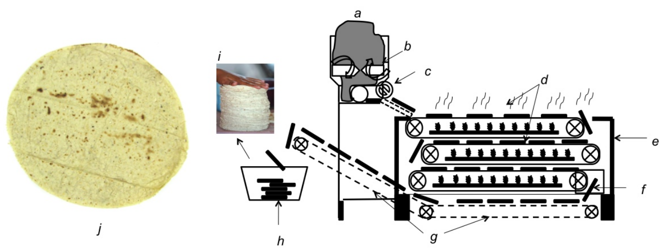

Figure 1: Schematic diagram of corn tortillas engine: a) input nixtalmalized dough, b) mixing, c) rolling and molding (shaping) , d) cooking, e) cooking chamber, f) output of corn tortillas from cooking chamber, g) cool belt, h) collector basket, i) product Staking, and j) corn tortilla image captured by computer vision system.

The nixtamalized *masa* has different thermal and rheological characteristics which requires a varying amount of heat during the cooking process. To produce adequately cooked tortillas, it is common for tortilla producers to manually regulate temperatures (heat flux) of *comal* or belt cooking to process each batch (Arámbula-Villa et al., 2007). However, this type of traditional production of corn tortillas has been modified by new processing technologies to make commercial-scale production a possibility. This practice has been successful in producing tortillas with sensory properties that differ from those produced traditionally. Despite the fact that corn tortilla production is now mechanized, tortilla machines are still manually controlled by human operators. Many aspects of the tortilla production still

*Corresponding Author:
Email address: dmery@ing.puc.cl (Domingo Mery)
URL: http://dmery.ing.puc.cl (Domingo Mery)

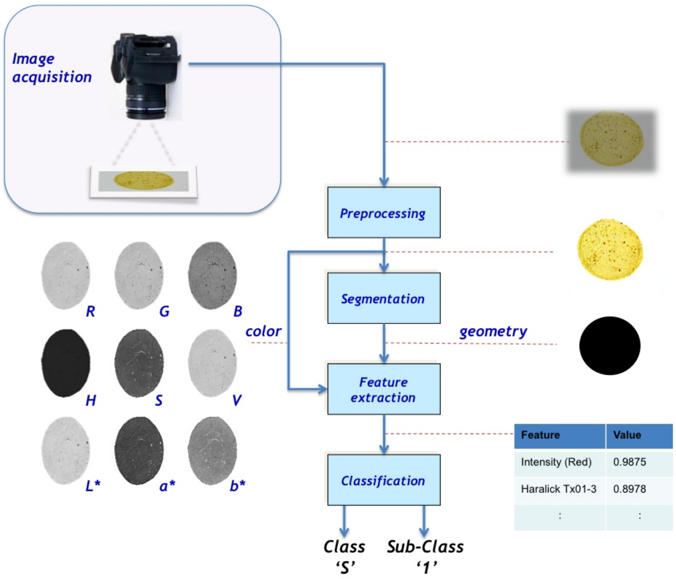

Figure 2: Computer vision schema used to determine tortilla quality.

depend on the operator's expertise such as well as dough input, mixing, shaping, baking and others process parameters.

Herrera-Corredor et al. (2007) concluded in a consumer study that overall acceptance of corn tortillas depended on chewiness as well as overall liking. Purchase intent was influenced mainly by appearance, plasticity, chewiness, taste, and overall liking. The study revealed critical sensory attributes and insight into how Mexican consumers weighted each characteristic when making purchasing decisions about corn tortillas. However, there is very little information available that defines the quality attributes of corn tortillas produced by mechanical machines. Mexican Regulations related to the manufacturing of corn tortillas (Norma-Mexicana, 2002; Norma-Oficial-Mexicana, 2009, 2002) are focused on aspects such as the quality of raw material and hygienic conditions during processing. These regulations focus on health and safety issues but do not give quality standard parameters such as: color, circularity, number of baking spots, different observable particles, size of corn grits, overall size, etc.

Taking the above into account as well as the necessity for more in-depth information regarding the control process of corn tortilla production, it became clear that computer vision techniques can be used to evaluate corn tortilla quality by quantifying their visual attributes. Computer vision systems have proven successful in the online measurement of several food products with applications ranging from routine inspection to complex, vision guided, robotic control ((Sun, 2008)).

The objective of this investigation is to utilize a computer vision framework, as shown in Fig. 2, to automatically determine the quality of corn tortillas. The steps involved in this framework are (Gonzalez and Woods, 2008):

- Image acquisition: A color digital image of the tortilla being tested is taken and stored in the computer.
- Pre-processing: The digital image is improved in order to enhance the details.
- Segmentation: The portion of the image containing the tortilla is found and isolated from the background of the scene.
- Feature extraction/selection: Significant features of the tortilla image are quantified.
- Classification: The extracted features are interpreted using a priori knowledge of the analyzed tortilla in order to determine its quality.

We used a supervised learning approach (Duda et al., 2001), where panelists use their experience to provide, in a training phase, a category label for each representative tortilla. Thus, the computer vision system is trained using the selected features and the supervised information received from panelists to classify a tortilla automatically.

2. Materials and methods

2.1. Sample selection of corn tortillas

Fifteen samples of corn tortilla production were selected from various places in Mexico City. Selection criterion as follows:
(1) Samples representing three levels of corn tortilla production available in the local Mexican market:
Small commercial scale (S): The small commercial scale samples were obtained from five different tortilla stores that used a partially mechanized process. All five stores produced tortillas using nixtamalized corn kernels obtained of the same nixtamal dough producer.
Medium commercial scale (M): The medium commercial scale samples were obtained from five different supermarkets that used a partially mechanized process similar to that used in the small commercial scale. However, the manufacturing process in the medium commercial scale supermarkets had better sanitary and production control than the small commercial scale tortilla stores. These samples were produced using nixtamalized corn flour.
Large commercial scale (L): The large commercial-scale samples packed in plastic bags were obtained from five different local supermarkets or small grocery stores. These samples were produced by a fully mechanized industrial process, made from industrially produced nixtamalized corn flour, and contained food additives used to enhance product texture and shelf life. Five different brands were used to obtain a wide range of visual attributes of corn tortillas.
(2) For each level of production, five samples (1,...,5) are selected which represent a wide range of desirable and undesirable visual appearance characteristics of corn tortillas. All samples used in this investigation exhibited complex physical characteristics such as black or burned areas, lack of uniform roundness, stripes across the surface caused by processing equipment, different visual textures, and different observable particle sizes of corn grits. These characteristics contributed to overall visual appearance in varying degrees depending on if the origin was small, medium or large commercial scale production.

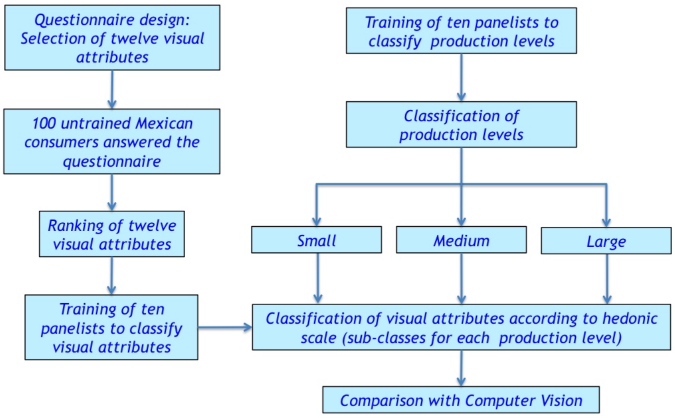

Figure 3: Diagram of human evaluation procedure (visual appearance).

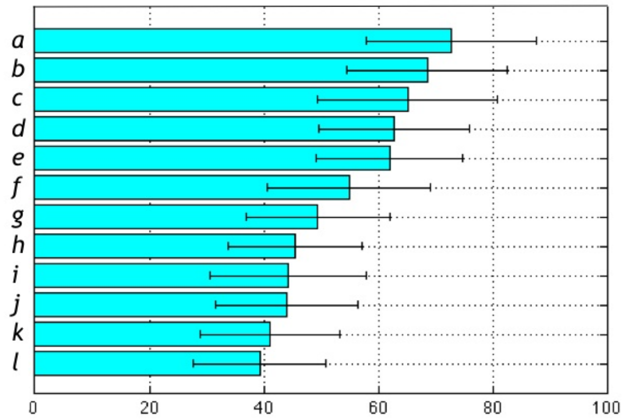

Figure 4: Score of the visual attributes: a) color, b) different observable particle size of corn grits, c) thickness, d) fractures or brakes, e) size, f) black/burned area due cooking, g) folds, h) circularity, i) *masa* residual, j) break in the superior film, k) stripes across the surface caused by processing equipment, and l) homogeneity of the borders on contours.

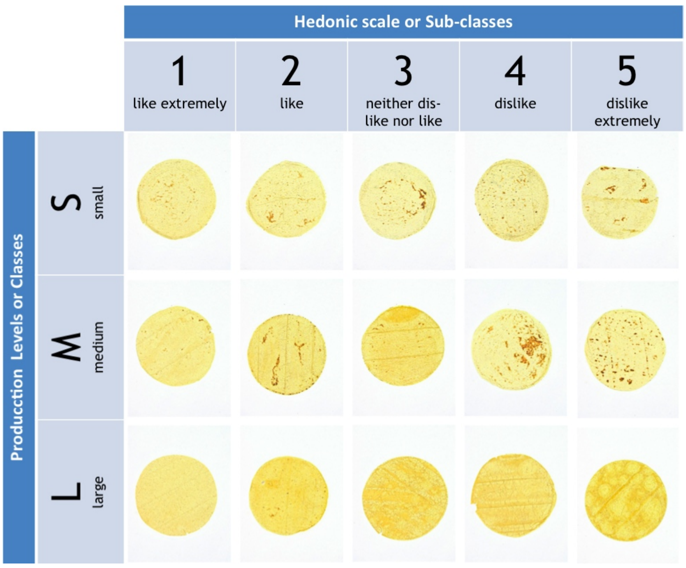

Figure 5: Gallery of corn tortilla images in their three production level (classes) and five sub-classes.

2.2. Questionnaire and Classification testing

Corn tortillas exhibit several complex sensory characteristics. Consumers associate varying degrees of importance to sensory attributes in order to use sound judgment when purchasing tortillas. According to report a by Herrera-Corredor et al. (2007) the Mexican market considers the overall visual appearance of the tortilla as one of the most important sensory attributes when making purchasing decisions. However, literature has not yet been able to define the specific characteristics that qualify a tortilla's appearance as acceptable.

According to our framework proposed in Fig. 3, in order to gain an understanding of consumer's preferences, regarding visual appearance acceptability and provide criterions of classification based in the quality of corn tortillas, a questionnaire was designed to establish which attributes were relevant to the visual appearance of the samples (see visual attributes in Fig. 4). The questions were based on visual attributes mentioned in previous works (Herrera-Corredor et al., 2007; Bejosano et al., 2005) as well as relevant characteristics observed in several samples. It should be noted that Mexican Food Regulation currently has no visual requisites for corn tortilla production. Untrained Mexican consumers (N = 100), who regularly ate corn tortillas, answered the questionnaire. Most of the participating consumers reside in Mexico City and the State of Mexico. Each of the 100 participants chose and classified twelve attributes according to their visual preference in decreasing order. Color was the most important attribute for consumers (around of 70%) while the homogeneity of the borders or contours was the attribute of least importance to consumers (around 40%). The results are presented in Fig. 4 as preference percentages for each attribute.

Ten trained panelists participated in the study which is more than the minimum of seven panelists recommended by Larmond (1977). They were randomly selected men and women from the Escuela Nacional de Ciencias Biológicas in Mexico City. The visual appearance of corn tortillas used in this study varied greatly depending on which of the three levels of production it came from (Fig. 5). A brief description of these differences was provided to the ten panelists and they were asked to classify the corn tortillas into one of three production levels or classes: small (S), medium (M) and large (L). The samples were shown randomly to panelists and each one evaluated ten sets of tortillas, fifteen per set, by studying both sides (superior and inferior). Each set of fifteen corn tortillas shown to panelists contained samples corresponding to the three production levels (S,M,L) as well as tortillas from at least five different sample points. Due to the easy classification of the production levels (100% of samples were classified correctly by the panelists), this experiment was carried out to supervise the training of the computer vision framework (Fig. 3).

Panelists were briefed about the questionnaire, particularly the visual attributes of corn tortillas and their meanings, and sample handling during evaluation. For each production level, the panelists were asked to evaluate acceptability of each one of the visual attributes selected based on the results obtained in the questionnaire about corn tortillas which follows the order in

Fig. 4. A set of five corn tortillas was presented randomly to panelists and they were asked to classify each set. They determined five sub-classes according to the 5-point hedonic scale (1:*like-extremely*, 3:*neither dislike nor like*, 5:*dislike extremely*) proposed by Peryam and Pilgrim (1957), as shown in Fig. 5. Fresh corn tortillas were used to ensure product freshness. Fifty sets of corn tortillas were evaluated on both sides and each panelist tested ten sets from each small, medium and large production level or class. A total of 750 corn tortillas were evaluated on both sides. Questionnaire and visual appearance tests were conducted at the Escuela Nacional de Ciencias Biolgicas (Mexico). They were conducted in individual cabins illuminated with cool, natural, fluorescent lights. The visual appearance test was written in Spanish.

2.3. Computer Vision

Two classification problems were distinguished: *i)* determination of hedonic scale (sub-class), and *ii)* determination of production level (class). The framework assigned a tortilla image to one of the five sub-classes (1:*like-extremely*, 2:*like*, 3:*neither dislike nor like*, 4:*dislike* or 5:*dislike-extremely*) and additionally, to one of the three classes (S, M or L) as shown in Fig. 5. In a training process, representative tortilla images were collected and segmented, features from each image were extracted and analyzed in order to select only those features relevant to the classification tasks. In this *supervised* training (Duda et al., 2001), the sub-class and class determined by the classifiers should coincide with the sub-class and class given by the sensorial panel and the production level respectively. Afterwards in a testing process, only the selected features are extracted to automatically classify the samples.

The computer vision framework used to automatically classify corn tortillas consists of five steps (Fig. 2): image acquisition, pre-processing, segmentation, feature extraction/selection and classification. Below, each step is explained in detail. Finally, the validation technique of the framework is explained.

2.3.1. Image Acquisition and Pre-processing

A computer vision system similar to the one described in (Pedreschi et al., 2004) was employed to capture the tortilla images (3456 × 2592 pixels in RGB color and JPEG format). Samples were illuminated using four fluorescent lamps TL-D deluxe, natural daylight, 18 W/965 (Philips, Mexico) with a color temperature of 6500 K (D65, standard light source commonly used in food research). Lamps (60 cm long) were arranged in the form of a square, 35 cm above the sample and at a 45 angle in relation to the sample. A total of 1500 images were obtained, two images (front and back) from each corn tortilla. Each sample was placed in front of the camera in the same position and orientation at a distance of 21 cm. To capture the images, a color Digital Camera (CDC) model Power Shot SX110 IS (Canon, NY, USA) with 9 megapixels was positioned vertically over the sample. Images were taken on a matte gray, bright background using the following camera settings: manual mode with lens aperture at f = 2.8 and speed 1/15 s, no zoom nor flash. The camera was connected to the USB port of a PC (Dell Precision 380 Pentium 4, Intel R, 3 GHz, 2 GB RAM, 80 GB hard disk) with Remote Capture Software DC installed (version 35.1, Canon, China). In order to calibrate the digital color system, a similar framework to the one investigated in (León et al., 2006) was used, the color values of 35 color charts were measured for each chart, the L*a*b* color values were measured using a colorimeter. Additionally, a RGB digital image was taken of each chart, and the R, G and B color values of the corresponding regions were measured using the ImageJ v 1.34s program which computes the mean values for each color channel. Finally, the images were resized to 0.1 times the size of the original images using cubic interpolation (Gonzalez and Woods, 2008).

2.3.2. Segmentation

The tortilla region of each image was segmented using the algorithm developed for food images proposed in (Mery and Pedreschi, 2004). The method has three steps:

1. Computation of a high contrast grey value image from an optimal linear combination of the RGB color components.
2. Estimation of a global threshold using a statistical approach.
3. Morphological operation in order to fill the possible holes presented in the segmented binary image.

2.3.3. Feature extraction and selection

Features extracted from each segmented tortilla region were divided into two families: *geometric* and *color* features (see Table 1).

Geometric features provide information on the size and shape of the tortilla region. Size features, such as area, perimeter, height and width, are given in pixels. Shape features are usually coefficients without units. In our approach, 54 geometric features of the following four groups were extracted from each region:

1. Standard: Simple shape and size information like area, perimeter, orientation, Euler number and solidity among others (MathWorks, 2003).
2. Elliptical: Shape and size information extracted from a fitted ellipse to the boundary of the region (Fitzgibbon et al., 1999).
3. Fourier Descriptors: Shape information invariant to scale, orientation and position based on Fourier descriptors (Zahn and Roskies, 1971; Chellappa and Bagdazian, 1984; Persoon and Fu, 1977).
4. Invariant moments: Shape information invariant to scale, orientation and position based on Hu (1962) moments, Flusser and Suk (1993) and Gupta and Srinath (1987).

Color features provide information about the color intensity of a tortilla region. In our approach, 227 features per color channel were extracted, *i.e.*, 227 × 10 = 2270 features for ten color channels (as shown in Fig. 2): gray; red, green and blue

Table 1: Extracted Features

Family	Group	Name and references
Geometric	Standard	Center of gravity *i*, Center of gravity *j*, Height, Width, Area, Perimeter, Euler Number, Equivalent, Diameter, Major Axis Length, Minor Axis, Length, Orientation, Solidity, Extent, Eccentricity, Convex Area and, Filled Area (MathWorks, 2003). Danielsson factor (Danielsson, 1978). Roundness (Hartmann, 1996).
	Elliptical	Major axis, Minor axis, Eccentricity, Orientation, Centre i and Centre j (Fitzgibbon et al., 1999).
	Fourier Descriptors	Descriptors (0,..., 15) (Zahn and Roskies, 1971).
	Invariant Moments	Hu (1,...,7) (Hu, 1962). Flusser (1,...,4) (Flusser and Suk, 1993). Gupta (1,...,3) (Gupta and Srinath, 1987).
Color (g=Gray, R,G,B, H,S,V, L*,a*,b*)	Standard	Mean Intensity, Standard deviation Intensity, Standard deviation Intensity with Neighbor, Mean Laplacian and Mean Gradient (Nixon and Aguado, 2008)
	Statistical textures	$Tx(k, p)$ (mean/range) for k=1. Angular Second Moment, 2. Contrast, Correlation, 4. Sum of squares, 5. Inverse Difference Moment, 6. Sum Average, 7. Sum Variance, Entropy, 8. Sum Variance, 9. Entropy, 10. Difference Variance, 11. Difference Entropy, 12., 13. Information Measures of Correlation, and 14. Maximal Correlation Coefficient, and p=1,...5 pixels (Haralick, 1979; Sonka et al., 1998).
	Filter Banks	DFT (1,2;1,2) and DCT (1,2;1,2) (Gonzalez and Woods, 2008). Gabor (1,...,8;1,...8), max(Gabor), min(Gabor), Gabor-J (Kumar and Pang, 2002).
	Invariant Moments	Int-Hu (1,...,7) Hu moments with intensity information (Hu, 1962).

(from RGB color space); hue, saturation and value (from HSV color space); L*, a* and b* (from L*a*b* color space). The following four groups of color features were used.

1. Standard: Simple intensity information related to the mean, standard deviation of the intensity in the region, mean first derivative in the boundary, and second derivative in the region (Nixon and Aguado, 2008).
2. Statistical textures: Texture information extracted from the distribution of the intensity values based on the approach proposed by Haralick (1979). They are computed utilizing *co-occurrence matrices* that represent second order texture information (the joint probability distribution of intensity pairs of neighboring pixels in the image), where mean and range of the following variables were measured: Angular Second Moment, Contrast, Correlation, Sum of Squares, Inverse Difference Moment, Sum Average, Sum Entropy, Sum Variance, Entropy, Difference Variance, Difference Entropy, Information Measures of Correlation, and Maximal Correlation Coefficient.
3. Filter banks: Texture information extracted from image transformations like Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) (Gonzalez and Woods, 2008), and Gabor features based on 2D Gabor functions, *i.e.*, Gaussian-shaped bandpass filters, with dyadic treatment of the radial spatial frequency range and multiple orientations, which represent an appropriate choice for tasks requiring simultaneous measurement in both space and frequency domains (usually 8 scale and 8 orientations) (Kumar and Pang, 2002).
4. Local binary patterns: Texture information extracted from occurrence histogram of *local binary patterns* (LBP) computed from the relationship between each pixel intensity value with its eight neighbors. The features are the frequencies of each one of the histogram's 59 bins. LBP is very robust in terms of gray-scale and rotation variations (Ojala et al., 2002).

In total, 54 geometric and 2270 color features, *i.e.*, n=2324 features, were extracted from each tortilla image. Afterwards, the features were selected in order to decide on the relevant features for the classification tasks, namely, the determination of the five sub-classes and the three classes. The n extracted features were arranged in an n-vector: $\mathbf{w} = [w_1...w_n]^\mathsf{T}$ that corresponds to a point in a n-dimensional feature space. The features were normalized as

$$f_{ij} = \frac{w_{ij} - \mu_j}{\sigma_j} \tag{1}$$

for $i = 1, ..., N$ and $j = 1, ..., n$, where w_{ij} denotes the j-th feature of the i-th feature vector, N is the number of samples, and μ_j and σ_j are the mean and standard deviation of the j-th feature. The normalized features have zero mean and a standard deviation equal to one. Those features that provide information about the position in the image of the tortilla (*e.g.* centers of gravity), constant features (*e.g.*, Euler number because tortillas have no holes) and high correlated features (*e.g.*, color features extracted from hue channel and gray value) were eliminated.

In feature selection, a subset of m features ($m < n$) that leads to the smallest classification error is selected. The selected m features were arranged in a new m-vector $\mathbf{z} = [z_1...z_m]^\mathsf{T}$.

The features can be selected using several algorithms previously investigated, such as Sequential Forward Selection (SFS) (Jain et al., 2000), Forward Orthogonal Search (Wei and Billings, 2007), Selection through Identification of Critical Variables of Principal Components (Mao, 2005), Ranking by Class Separability Criteria (MathWorks, 2007), and Combination with Principal Components (Duda et al., 2001). In our experiments the best performance was achieved using the well-known Sequential Forward Selection (SFS) algorithm. This method selects the best single feature and then adds one feature at a time, in combination with the selected features, to maximize classification performance. The iteration is halted once no considerable improvement in the performance is achieved by adding a new feature. By evaluating selection performance we ensure: *i)* a small intraclass variation and *ii)* a large interclass variation in the space of the selected features. For the first and second conditions the intraclass-covariance $\mathbf{C}_b$ and interclass-covariance $\mathbf{C}_w$ of the selected features $\mathbf{Z}$ are used respectively. Selection performance can be evaluated using the spur criterion for the selected features $\mathbf{S}$:

$$J(\mathbf{Z}) = \mathrm{spur}\left(\mathbf{C}_w^{-1}\mathbf{C}_b\right). \tag{2}$$

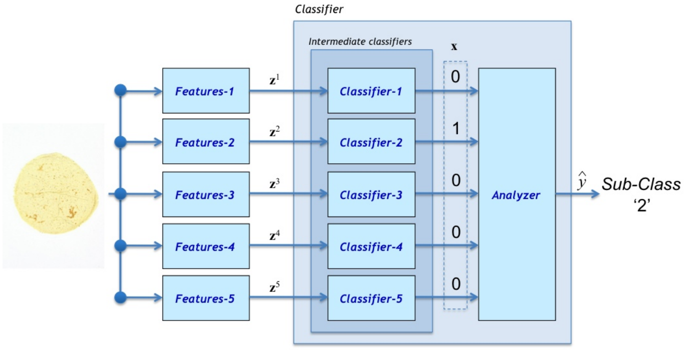

Figure 6: Classification strategy for the hedonic scale classification. In this example, only Classifier-2 has an output '1', *i.e.*, the determined sub-class is '2'.

The larger the objective function J, the higher the selection performance.

2.3.4. *Classification*

Two classification problems were distinguished (Fig. 5): *i)* determination of hedonic scale (sub-class), and *ii)* determination of production level (class). The classifiers assigned a feature vector $\mathbf{z}$ to one of the determined sub-classes and one of the determined classes. In order to find the best performance for each classification problem, the framework was tested on a bank of classifiers, such as support vector machines (SVM) (Shawe-Taylor and Cristianini, 2004), linear and quadratic discriminant analysis (Webb, 2005), k-nearest neighbor (Duda et al., 2001), neural networks (Bishop, 2006), boosting (Viola and Jones, 2004), and minimal and Mahalanobis distance (Duda et al., 2001).

***i)* Hedonic scale classification (sub-class):** The first classification problem was very difficult to solve using a simple classifier, however, high performance was achieved with SVM implemented in the Bioinformatics Toolbox™ of Matlab (MathWorks, 2007).

Typically, SVM transforms a two-classes feature space, where the classes overlap, into a new enlarged feature space where the classification boundary is linear. Thus, a simple linear classification can be designed in the transformed feature space in order to separate both classes (Shawe-Taylor and Cristianini, 2004). The original feature space is transformed using a function $h(\mathbf{z})$, however, for the classification only, the kernel function $K(\mathbf{z},\mathbf{z}') = \langle h(\mathbf{z}), h(\mathbf{z}')\rangle$ that computes inner products in the transformed space is required. In this case, the best classification was obtained using a *Gaussian Radial Basis* (RBF) function kernel defined by (Hastie et al., 2003):

$$K(\mathbf{z},\mathbf{z}') = e^{-\|\mathbf{z}-\mathbf{z}\|^2} \tag{3}$$

where the linear boundary, *i.e.*, the separating hyperplane in the transformed space, is computed using Least–Squares approach (MathWorks, 2007).

Since SVM is able to classify only two classes, the five-sub-classes problem was split into five *one-against-all* classification problems. In this case, intermediate "Classifier-i" was trained to determine if a tortilla belonged to sub-class i, for $i = 1, ..., 5$. Intermediate "Classifier-i" yields $x_i = 1$ if the classifier determined that the tortilla belongs to sub-class i, and $x_i = 0$ otherwise. Afterwards, the five outputs given by the intermediate classifiers were analyzed to estimate the final sub-class of the tortilla, as shown in Fig. 6. The final sub-class $\hat{y}$ was calculated as the average of the corresponding outputs of the intermediate classifiers:

$$\hat{y} = \frac{x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5}{x_1 + x_2 + x_3 + x_4 + x_5}. \tag{4}$$

For example, if the five intermediate classifiers are $\mathbf{x} = (0,0,1,1,1)$, the final sub-class is $\hat{y} = (3+4+5)/3 = 4$. If no intermediate classifier detects a sub-class, *i.e.*, $\mathbf{x} = (0,0,0,0,0)$, then the final sub-class is $\hat{y} = 3$, because sub-class '3' is the choice where the classification error $|\hat{y} - y|$ is expected to be minimal. In this approach, a final sub-class with decimals, *e.g.*, $\hat{y} = 4.5$ for $\mathbf{x} = (0,0,0,1,1)$, is allowed, this means that the final sub-class is between two sub-classes of the hedonic scale. Ideally, $\hat{y} = y$, however in this pattern recognition problem, a small classification error $e = |\hat{y} - y| > 0$ was obtained.

Additionally, features that were relevant for one intermediate classifier were not necessarily relevant for other intermediate classifiers. For this reason, *ad-hoc* features for each intermediate classifier were selected. The extraction of the selected features $\mathbf{z}^i$ for intermediate "Classifier-i" were extracted in "Features-i" block (Fig. 6). The selection of the features was performed using the SFS method explained in Section 2.3.3 for the five *one-against-all* classification problems.

***ii)* Level production classification (class):** The second classification problem was easy to solve, because there was enough visual information to determine efficiently which of the three production levels a tortilla belonged to, *i.e.*, classes S, M or L. Many of the mentioned classifiers achieved a satisfactory performance using less than ten features selected by the SFS approach. For example, the Mahalanobis classifier obtained a performance greater than 96%.

For the Mahalanobis classifier, a mean value $\bar{\mathbf{z}}_c$ of the training samples was calculated for each class: c=S, M, L. A test feature vector $\mathbf{z}$ was assigned to class 'c' if the Mahalanobis distance between $\mathbf{z}$ and $\bar{\mathbf{z}}_c$ defined as:

$$d_c(\mathbf{z}, \bar{\mathbf{z}}_c) = (\mathbf{z} - \bar{\mathbf{z}}_c)^{\mathsf{T}}\mathbf{C}_c^{-1}(\mathbf{z} - \bar{\mathbf{z}}_c) \tag{5}$$

is minimal, where $\mathbf{C}_c$ is the covariance matrix of class 'c'. The Mahalanobis classifier takes into account errors associated with prediction measurements, such as noise, by using the feature covariance matrix to scale features according to their variances (Duda et al., 2001).

2.3.5. *Validation*

The performance of the classifier was defined as the ratio of the tortillas that were correctly classified to the total number of tortillas. The performance was validated using cross-validation, a technique widely implemented in machine learning problems (Mitchell, 1997). In cross-validation, the samples are divided into F folds randomly. $F - 1$ folds are used as training data and the remaining fold is used as testing data to evaluate the performance of the classifiers. We repeated this experiment F times rotating train and test data. The F individual performances from the folds are averaged to estimate the final performance of the classifiers.

3. Experimental Results

As explained in Section 2, 250 corn tortillas were examined for each production level: small (S), medium (M) and large (L). A sensorial panel categorized each tortilla in five hedonic sub-classes (see Fig. 5). Two images for each tortilla, *i.e.*, 250 × 3 × 2 = 1500 images, were captured. Each image was segmented according to the approach explained in Section 2.3.1, and 2324 (geometric and color) features were extracted as detailed in Table 1[1].

After the feature extraction, 75% of the samples of each class were randomly chosen to perform the feature selection. Results obtained in the two mentioned classification problems: *i)* determination of hedonic scale (sub-class), and *ii)* determination of production level (class), are discussed below.

***i)* Hedonic scale classification (sub-classes):** This classification problem, with five-sub-classes, was split into five *one-against-all* classification problems. Five data sets were built, one for each class, where the samples were labeled with '1' for the selected class, and '0' for the rest. In the data sets, the samples labeled with '1' were replicated three times in order to obtain balanced classes. Forty features for each data set using the SFS were selected. Five SVM intermediate classifiers (Fig. 6) were designed for the first m = 5, 10, 15, ..., 40 selected features. The final sub-class was determined by the computing equation (4).

The performance was evaluated using cross-validation with $F = 10$ folds as explained in Section 2.3.5. The results for different values of m are shown in Table 2, where the following statistical variables were computed:

- h: coincidence ratio between classifier $\hat{y}$ and sensorial panel y (ideally $h = 1$),
- $\bar{e}$: the mean of the classification error $e = |\hat{y} - y|$ (ideally $\bar{e} = 0$),
- σ_e: the standard deviation of the classification error (ideally $\sigma_e = 0$), and
- η: average performance of the intermediate classifiers (ideally $\eta = 1$).

For $m = 15$ extracted features our classifier had very high performance ($\eta = 0.95$). In average for this case, the sub-class given by our classifier $\hat{y}$ will be the same given by the sensorial panel $y \pm 0.28$. Table 3 shows the selected 15 features for each class, where there are 64 features without repetitions. It is recommended that the reader refer to Table 1 and Section 2.3.3 to see a description of the features.

Table 2: Performance of the classification of hedonic sub-classes using cross-validation with 10 folds.

m	h	$\bar{e}$	σ_e	η
5	0.227	1.162	1.390	0.806
10	0.531	0.704	1,071	0.877
15	0.825	0.284	0.672	0.950
20	0.929	0.108	0.399	0.981
25	0.969	0.045	0.223	0.991
30	0.985	0.021	0.147	0.996
35	0.991	0.012	0.106	0.998
40	0.995	0.006	0.052	0.999

The rest of the classifiers mentioned in Section 2.3.4 were tested in this experiment, however the obtained performance was inferior in all cases, *e.g.*, as the value h for $m = 15$ was between 0.220 (for linear discriminant analysis) and 0.799 (for a neural network) in comparison to 0.825 (for SVM) as shown in Table 2.

***ii)* Level production classification (classes):** For the second classification, where the production level S, M or L is to be determined, two sets of features selected using SFS were analyzed. Set-1 consisted of the best ten features selected from the 2324 extracted features. On the other hand, in order to reduce the total number of extracted features, Set-2 contained the best ten features chosen from the 64 features selected for the hedonic scale classification problem, *i.e.*, only features from Table 3 were selected. The performance of the classification using the Mahalanobis classifier and the selected features for Set-1 and Set-2 were validated using cross-validation with $F = 10$ folds as explained in Section 2.3.5. The results are shown in Table 4. Set-1 considered only nine features because the classification performance was not improved by considering the tenth selected feature. It is evident that Set-1 achieved a higher performance (almost perfect yielding $\eta = 0.995$), because the features were selected from all the extracted features, whereas for Set-2 only the selected features of the first classification problem were available. In spite of that, Set-2 achieved a very high performance ($\eta = 0.96$). As mentioned before, it is recommended that the reader refer to Table 1 and Section 2.3.3 to see a description of the features.

4. Conclusion

The need for more information about the control process of making corn tortillas, evaluation of quality and visual appearance by means of quantitative methods can be satisfied using computer vision. This non-destructive technique objectively measures color and geometric patterns in non-uniformly colored surfaces, and also determines other physical features such as image texture, morphological elements, and defects in order to automatically classify tortilla quality. The promising results

[1] All image segmentation, feature extraction, feature selection, classification and validation approaches are implemented in *Balu Matlab Toolbox* – Group of Machine Intelligence, Department of Computer Science, Catholic University of Chile (download in `http://dmery.ing.puc.cl`).

Table 3: Selected features for each hedonic sub-class ([•] means the color channel).

m	1	2	3	4	5
1	Int-Hu(5) [S]	$Tx(14,1)$(mean) [a*]	$Tx(14,1)$(mean) [a*]	$Tx(14,1)$(mean) [a*]	$Tx(14,1)$(mean) [a*]
2	Int-Hu(6) [S]	DFT(1,1) [H]	Gabor(5,6) [B]	$Tx(14,3)$(range) [B]	DFT(2,2) [H]
3	Hu(4)	DCT(2,1) [V]	DFT(1,1) [R]	$Tx(12,3)$(mean) [H]	Std Intensity [S]
4	Int-Hu(7) [S]	DCT(2,2) [a*]	$Tx(2,1)$(mean) [g]	$Tx(14,2)$(mean) [B]	$Tx(14,4)$(range) [V]
5	Int-Hu(2) [S]	$Tx(14,1)$(mean) [H]	DCT(2,2) [a*]	$Tx(12,1)$(mean) [S]	$Tx(14,2)$(range) [L*]
6	Hu(3)	DCT(2,2) [S]	DCT(2,1) [V]	MajorAxisLength	DFT(2,2) [a*]
7	MajorAxisLength	DCT(1,2) [V]	DCT(2,2) [R]	$Tx(12,4)$(range) [a*]	Fourier-descriptor(5)
8	Gupta(2)	DCT(1,2) [R]	DFT(2,2) [R]	$Tx(13,3)$(mean) [S]	$Tx(14,3)$(range) [g]
9	Fourier-descriptor(0)	$Tx(5,1)$(mean) [G]	DCT(2,1) [a*]	$Tx(14,3)$(range) [S]	MajorAxisLength
10	Solidity	Gabor(2,3) [g]	Int-Hu(2) [a*]	$Tx(13,1)$(mean) [B]	Gabor(1,7) [H]
11	Fourier-descriptor(2)	$Tx(3,1)$(mean) [L*]	$Tx(6,1)$(mean) [R]	$Tx(12,3)$(mean) [a*]	Fourier-descriptor(14)
12	Int-Hu(3) [S]	DCT(1,1) [b*]	Std Intensity [R]	Laplacian [a*]	DCT(1,1) [b*]
13	Fourier-descriptor(5)	DCT(2,2) [V]	$Tx(12,4)$(range) [S]	Gabor(8,7) [g]	DCT(2,2) [S]
14	$Tx(13,1)$(mean) [G]	DCT(2,2) [H]	$Tx(14,5)$(range) [B]	DFT(2,2) [a*]	DCT(2,1) [B]
15	$Tx(12,2)$(mean) [H]	DCT(2,2) [g]	$Tx(14,4)$(range) [B]	DFT(1,2) [a*]	$Tx(12,2)$(range) [a*]

Table 4: Performance using cross-validation with 10 folds and selected features of production level classification ([•] means the color channel).

Features	Set-1	Set-2
1	Area	Int-Hu(1) [S]
2	Mean Intensity [R]	Mean Intensity [R]
3	Gabor(2,6) [b*]	DFT(1,1) [H]
4	Mean Intensity [g]	$Tx(12,1)$ (mean) [H]
5	Std Intensity [R]	$Tx(12,2)$ (mean) [H]
6	Gabor(1,7) [H]	DCT(1,1) [g]
7	$Tx(12,1)$(mean) [a*]	$Tx(6,1)$(mean) [b*]
8	Gabor(2,7) [H]	Gabor(1,6) [H]
9	DFT1 [g]	Std Intensity [R]
10	–	DFT(1,2) [R]
η	0.9947	0.9620

outlined in this work demonstrate that, using a very large number of features combined with efficient feature selection and classification approaches, a very high classification rate in tortilla quality control can be achieved. The key idea of the proposed method was to select, from a large universe of features, only those features that were relevant for the separation of the classes. The method was tested in three different tortilla production levels with tortillas of five different hedonic sub-classes yielding a performance of 95% in accuracy using 64 features and support vector machines. Additionally, using only 10 of the selected features and a simple statistical classifier based on Mahalanobis distance, it was possible to determine with 96% accuracy the production level of the tortillas.

Additionally, the contribution to computer vision technology is the use of the state-of-art algorithms in the automated quality classification of corn tortillas. In our case, we use a general framework that extracts more than 2300 features and only 64 of them are selected. We believe that the proposed approach opens new possibilities not only in the field of automated visual inspection of tortillas but also in other similar food products like breads and cookies among others.

Acknowledgments

This work was supported by grant No. R0308LAC003 from LACCIR Virtual Institute: http://www.laccir.org, grant No. 20091151 from the Instituto Politécnico Nacional (IPN, Mexico) and grant No. 59730 from CONACYT–Mexico.

References

Arámbula-Villa, G., Gutiérrez-Árias, E., Moreno-Martínez, E., 2007. Thermal properties of maize masa and tortillas with different components from maize grains, and additives. Journal of Food Engineering 80 (1), 55 – 60.

Ayala-Rodríguez, A., Gutiérrez-Dorado, R., Milán-Carrillo, J., López-Valenzuela, S. M.-R. J., Valdez-Ortiz, A., Paredes-López, O., Reyes-Moreno, C., 2009. Nixtamalised flour and tortillas from transgenic maize (zea mays l.) expressing amarantin: Technological and nutritional properties. Food Chemistry 114 (1), 50–56.

Bejosano, F., Joseph, S., Lopez, R., Kelekci, N., Waniska, R., 2005. Rheological and sensory evaluation of wheat flour tortillas during storage. Cereal Chemistry 82 (3), 256–263.

Bishop, C. M., 2006. Pattern Recognition and Machine Learning. Springer.

Chellappa, R., Bagdazian, R., 1984. Fourier coding of image boundaries. IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-6 (1), 102–105.

Cortés-Gómez, A., Martín-Martínez, E. S., Martínez-Bustos, F., Vázquez-Carrillo, G. M., 2005. Tortillas of blue maize (zea mays l.) prepared by a fractionated process of nixtamalization: analysis using response surface methodology. Journal of Food Engineering 66 (3), 273 – 281.

Danielsson, P.-E., 1978. A new shape factor. Computer Graphics and Image Processing 7, 292–299.

Duda, R., Hart, P., Stork, D., 2001. Pattern Classification, 2nd Edition. John Wiley & Sons, Inc., New York.

Fitzgibbon, A., Pilu, M., Fisher, R., 1999. Direct least square fitting ellipses. IEEE Trans. Pattern Analysis and Machine Intelligence 21 (5), 476–480.

Flusser, J., Suk, T., 1993. Pattern recognition by affine moment invariants. Pattern Recognition 26 (1), 167–174.

Gonzalez, R., Woods, R., 2008. Digital Image Processing, 3rd Edition. Pearson, Prentice Hall.

Gupta, L., Srinath, M. D., 1987. Contour sequence moments for the classification of closed planar shapes. Pattern Recognition 20 (3), 267–272.

Haralick, R., 1979. Statistical and structural approaches to texture. Proc. IEEE 67 (5), 786–804.

Hartmann, I., 1996. Mustererkennung. Skriptreihe Regelungstechnik und Bildverarbeitung, Technische Universität Berlin.

Hastie, T., Tibshirani, R., Friedman, J., August 2003. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, corrected Edition. Springer.

Herrera-Corredor, J., Saidu, J., Khachatryan, A., Prinyawiwatkul, W., Carballo-Carballo, A., Zepeda-Bautista, R., 2007. Identifying drivers for consumer acceptance and purchase intent of corn tortilla. Journal of Food Science 72 (9), S727–S730.

Hu, M.-K., 1962. Visual pattern recognition by moment invariants. IRE Trans. Info. Theory IT (8), 179–187.

Jain, A., Duin, R., Mao, J., 2000. Statistical pattern recognition: A review. IEEE Trans. Pattern Analysis and Machine Intelligence 22 (1), 4–37.

Kumar, A., Pang, G., 2002. Defect detection in textured materials using gabor filters. IEEE Trans. on Industry Applications 38 (2), 425–440.

Larmond, E., 1977. Laboratory methods for sensory evaluation of food. Publication 1284, Canadian Department of Agriculture: Ottawa, Canada.

León, K., Mery, D., Pedreschi, F., León, J., 2006. Color measurement in L*a*b* units from RGB digital images. Food Research International 39 (10), 1084 – 1091, physical Properties VI.

Lind, D., Barham, E., 2004. The social life of the tortilla: food, cultural politics, and contested commodification. Agric and Human Values 21, 47–60.

Mao, K., April 2005. Identifying critical variables of principal components for unsupervised feature selection. IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics 35 (2), 339–344.

MathWorks, January 2003. Image Processing Toolbox for Use with MATLAB: User's Guide. The MathWorks Inc.

MathWorks, 2007. Matlab Toolbox of Bioinformatics: User's Guide. Mathworks Inc.

Mery, D., Pedreschi, F., 2004. Segmentation of colour food images using a robust algorithm. Journal of Food Engineering(accepted April 2004).

Mitchell, T., 1997. Machine Learning. McGraw-Hill, Boston.

Nixon, M., Aguado, A., 2008. Feature Extraction and Image Processing, 2nd Edition. Academic Press.

Norma-Mexicana, 2002. NMX-FF-034/1-SCFI-2002: Non Industrialized Food Products for Human Consumption - Cereals - Part I: White Corn for Alkaline Process of Corn Tortillas and Nixtamalized Corn Products - Specifications and Test Methods. Secretaría de Economía, Estados Unidos Mexicanos.

Norma-Oficial-Mexicana, 2002. NOM-187-SSA1/SCFI-2002: Productos y servicios. Masa, tortillas, tostadas y harinas preparadas para su elaboración y establecimientos donde se procesan. Especificaciones sanitarias. Información comercial. Métodos de prueba. Secretaría de Economía, Estados Unidos Mexicanos.

Norma-Oficial-Mexicana, 2009. NOM-019-ENER-2009: Eficiencia térmica y eléctrica de máquinas tortilladoras mecanizadas. Límites, método de prueba y marcado. Secretaría de Economía, Estados Unidos Mexicanos.

Ojala, T., Pietikainen, M., Maenpaa, T., Jul 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. Pattern Analysis and Machine Intelligence, IEEE Transactions on 24 (7), 971–987.

Pedreschi, F., Mery, D., Mendoza, F., Aguilera, J., 2004. Classification of potato chips using pattern recognition. Journal of Food Science 69 (6), E264–E270.

Persoon, E., Fu, K., 1977. Shape discrimination using Fourier descriptors. IEEE Trans. Systems, Man, and Cybernetics SMC-7 (3), 170–179.

Peryam, D., Pilgrim, F., 1957. Hedonic scale method for measuring food preferences. Food Technology 11 (9), 9–14.

Shawe-Taylor, J., Cristianini, N., 2004. Kernel Methods for Pattern Analysis. Cambridge University Press.

Sonka, M., Hlavac, V., Boyle, R., 1998. Image Processing, Analysis, and Machine Vision, 2nd Edition. PWS Publishing, Pacific Grove, CA.

Sun, D.-W. (Ed.), 2008. Computer Vision Technology for Food Quality Evaluation. Academic Press.

Viola, P., Jones, M., 2004. Robust real-time object detection. International Journal of Computer Vision 57 (2), 137–154.

Webb, A., 2005. Statistical Pattern Recognition. Wiley, England.

Wei, H.-L., Billings, S., Jan. 2007. Feature subset selection and ranking for data dimensionality reduction. Pattern Analysis and Machine Intelligence, IEEE Transactions on 29 (1), 162–166.

Zahn, C., Roskies, R., 1971. Fourier descriptors for plane closed curves. IEEE Trans. Computers C-21 (3), 269–281.