CLASSIFICATION OF CLOSED AND OPEN-SHELL (TURKISH) PISTACHIO NUTS USING DOUBLE TREE UN-DECIMATED WAVELET TRANSFORM

Nuri F. Ince1, Fikri Goksu1, Ahmed H. Tewfik1, Ibrahim Onaran2, A. Enis Cetin2, Tom C. Pearson3

ABSTRACT

Due to low consumer acceptance and possibility of having immature kernels, closed-shell pistachio nuts have to be separated from open-shell pistachio nuts. The feasibility of a system using non-contact impact acoustics as a means for classifying closed-shell nuts from open-shell nuts has already been shown with better discrimination performance than a mechanical system. The accuracy of an impact acoustics based system is determined by the signal processing and feature extraction procedures. In this paper a new time-frequency plane feature extraction and classification algorithm was developed to discriminate between open and closed-shell Turkish pistachio nuts. The proposed approach relies on the analysis of impact acoustics signal of a pistachio nut which is emitted by dropping it from a certain height onto a metal plate. Features are extracted by decomposing the acoustic signals into time and frequency components using Dual-Tree un-decimated wavelet transform. The most discriminative features from the dual tree nodes are selected by a greedy search algorithm that combines a structural pruning and classifier feedback. This approach requires no prior knowledge of the relevant time or frequency content of the acoustic signals. The proposed algorithm used a small number of features and achieved a classification accuracy of 91.5% on the validation data set. A previously implemented algorithm, which uses maximum signal amplitude, absolute integration and gradient features, achieved an 82% classification accuracy on the same dataset. The results obtained show that the time-frequency features extracted from impact acoustics can be successfully used in open and closed shell Turkish pistachio classification.

Keywords: Impact Acoustic, Turkish Pistachio, Wavelet Transform, Classification

INTRODUCTION

Closed shell pistachio nuts could be rejected by consumers because they are difficult to open and may contain immature kernels. Therefore their separation from open shell nuts is crucial. Closed shell pistachio nuts are currently separated from open shell nuts by mechanical devices called “pinpickers.” They can inadvertently damage the kernel of open shell nuts by inserting a needle into the kernel meat. The hole created by the needle can give the appearance of an insect tunnel, leading to rejection by the consumer. In addition, according to (Pearson et al., 2001), approximately 5 to 10% of all U.S. open shell pistachio nuts are incorrectly classified by mechanical devices as having a closed shell, costing the industry $3.75 to $7.5 million per year in lost revenue. Therefore high accuracy classification systems are needed in the industry.

Recently, a new non-contact system based on impact acoustic emission has been proposed for food kernel inspection which overcomes several limitations of the approaches summarized above. This system was designed to separate pistachio nuts with closed shells from those with open shells (Pearson, 2001). In this system the pistachio nuts are impacted onto a steel plate and the resulting acoustic signals are recorded. During the off-line training (learning) phase a total of 359 features are extracted. They are based on the observations from open and closed shell pistachio nut sound signals and use the absolute value of the signal magnitude, the absolute value of the gradient, or both. Among those 359 features the best 3 are selected using both linear and non-linear discriminant analyses. The selection of only the 3 best features is due to the real-time processing constraint. This constraint requires that the nut under examination be classified before the next nut comes into the system. Using these 3 selected features on the validation set the classification accuracy was approximately 97% with a throughput of 40 nuts per second, proving better performance than traditional mechanical devices.

There is also a number of different non-contact classification devices designed for pistachio nuts. Their systems are based on processing images rather than sound signals. For a description of those image-based devices and a short comparison of their performances with that of the device based on sound signals, we refer the reader to the references in (Pearson et al., 2001), and (Ghazanfari et al., 1997). In this paper we are interested in the system based on sound signals. The pistachio study using impact acoustics emphasized the importance of signal processing stages. In particular it has been shown that having a priori information about relevant time and frequency content of the signals has important effects on classification accuracy. However the adjustment of these parameters is demanding. Furthermore the differences in size between nuts from region to region result in sound signals with different characteristics. This makes it necessary to develop an adaptive classification system which can adjust its parameters for the given signal (nut type).

1 Department of Electrical and Computer Engineering, University of Minnesota, USA.
2 Department of Electrical and Electronics Engineering, Bilkent University, Turkey.
3 United States Department of Agriculture – Agricultural Research Service, USA
In this paper we tackle these problems by developing an adaptive time-frequency plane feature extraction and classification algorithm which utilizes the same impact acoustic system. Based on the observations from previously mentioned algorithms above, our proposed algorithm does not depend on a priori knowledge of time-frequency content of the signals under examination. Hence, it is more universally adaptable to other types of applications and uses features extracted from both time and frequency content of signals concurrently. In this respect the proposed algorithm extracts adaptive time-frequency features from impact acoustics by using a dual tree undecimated wavelet transform. It locates most discriminant features among classes by a greedy algorithm using a Linear Discriminant Analysis (LDA) classifier. In particular, we tested our proposed system to discriminate between open and closed shell pistachio nuts collected from the Gaziantep region of Turkey.

**MATERIALS AND SYSTEMS**

The apparatus used to record sound signals and sort the pistachio nuts is based on the system in [Pearson et al., 2001] and shown in Figure 1, below. From a higher place a nut is fed down on a declining slide where at the end it drops onto a steel impact plate. The resulting acoustic signal from the impact is recorded with a highly directional microphone. The thickness of the impact plate and the microphone are selected in a way such that the unwanted noise during the sound signal recordings is minimized. Output of the microphone is digitized with 100 kHz sampling frequency by a sound card attached to a PC and stored for further analysis. The sound signals available this way are analyzed in an off-line manner for feature extraction. After features are selected and the decision rules are set, the system is run for validation. In real time implementation a decision is given and either the nut is diverted by an air valve to one stream or no action is taken and the nut goes to another stream. Obviously, a timing constraint is present and this is also one of the reasons it is necessary to give a good decision as fast as possible. Although we do not validate the performance of our data in real-time these time constraints have to be taken into consideration for practice.

![Figure 1. Impact sound classification system.](image)

For each type (closed and open) of pistachio 200 recordings were obtained using this setup. Each recording was 256 samples long. Sample sound signals are given in Figure 2.

![Figure 2. Sample impact sound signals of an open-shell (left) and a closed-shell pistachio (right).](image)

The variations between open and closed-shell pistachio sound signals can be observed in these time domain signals. But as can be seen in the results section, a method which is based only on time domain signal variations yields lower classification accuracy. Consequently, it is a better strategy to use information both from time and frequency domains. To give a flavor of the time and frequency content of the signals, a Short Time Fourier Transform (STFT) is computed for each sound signal. 256-point FFT is used to compute the STFT. The absolute values of these STFT’s are averaged over 200 signals for each class. Logarithm of these averages one for each class is depicted in Figure 3. Not only do we have the time variations but also the frequency variations for both classes of signals as well.
Feature Extraction

In the last several years there has been a growing interest to explore the time-frequency plane for classification by using adaptive strategies. The local discriminant bases (LDB), an algorithm of (Saito et al.,) has been proposed to achieve this task. This approach first represents a given signal for each class in a redundant manner in a single pyramidal tree structure by wavelet packets (WP) or cosine packets (CP). The pyramidal tree structure is pruned from bottom to top such that the discrimination power between expansion coefficients in the nodes of the tree is maximized. Once the tree is pruned a complete representation of the signal is obtained. This is followed by sorting the expansion coefficients according to their discrimination power and inputting them to a classifier for the final decision. This powerful method has been used in several applications and successful results have been obtained. However the LDB algorithm has many drawbacks. First the WP/CP’s which are used to represent the signal in the nodes of the tree structure does not satisfy the shift invariance property. Briefly, a shift in the signal is reflected by unpredictable changes in the expansion coefficients. This behavior is not appropriate for pattern recognition applications. Furthermore, since a single tree is used, this algorithm can only adapt in a single axis. Several studies have shown that the adaptation in both axes is crucial (Ince et al., 2006). Finally the pruning and feature sorting stages do not account for the interactions/relations between different time-frequency cells. The obtained complete representation may not be the best subset for the classifier.

Here, as a first step, we use un-decimated wavelet transform (UDWT) to achieve a shift invariant signal representation. Unser (1995) has shown that the classification results obtained with UDWT is superior to WT. Furthermore we use a dual tree to adapt both in time and frequency axes. Finally we use a greedy approach to select features from the redundant representation. We explain these techniques in detail in the following sections.

Un-Decimated Wavelet Transform

The discrete Wavelet Transform (DWT) and its variants have been extensively used in 1D and 2D signal analysis due to their good localization properties both in time and frequency domains (Vetterli et al., 2001). Here we consider the 2-band DWT without downsampling. This means number of samples in a subband at any level is same as that of in the original signal. The subbands computed by UDWT and the original signal constitute the frequency branch of the double tree undecimated DWT.

Time Segmentation

In order to decrease the dimensionality and preserve the energy-based information in each subband, like in the frequency decomposition tree, every subband is segmented into orthogonal time segments at each level with a pyramidal tree structure successively. In each time segment the sum of the squares of the samples, energy, is computed as one feature to be used in off-line training. The time segmentation explained above forms the second branch of the double tree. From then on we keep the index information of the dual tree structure to be used in the later stage for dimension reduction via pruning.

To summarize this section the reader is referred to the double tree structure in Figure 4. This double tree uses 1-level in both planes. The vertical middle boxes are the frequency subbands. Box 1 represents unfiltered original signal, box 2 represents low pass filtered signal and box 3 high pass filtered signal. Each of these subbands is segmented in time into 3 segments, as shown. Segment 1 covers the whole subband, segment 2 covers the first half of it and segment 3 the second part of it. The parameters for deciding the number of features are the number of levels for frequency and time segmentation. Let $T$ be the number of levels in time and $F$ the number levels in frequency, there will be $2^{T+F-1}$ subbands (including the original signal) and $2^{T+F-1}$ time segments for each subband. This will make the total number of features $(2^{F+1})T(2^{T+1})F$. 

![Figure 3. Averaged STFT images of open-shell (left) and a closed-shell pistachios (right).](image)
Feature Selection

As explained in the previous section, the dual tree has total \( NF = (2^{F+1}-1)(2^{T+1}-1) \) number of features for each sound signal where \( F \) is the frequency level and \( T \) the time level. The constructed set of features forms a redundant dictionary. As emphasized above, the selection of features from a redundant dictionary is critical. The strategy for feature selection is based on evaluation of the actual classification performance. We quantify the efficiency of each feature set by evaluating its classification accuracy by a cost measure and we use this cost to reformulate our dictionary.

Three different types of methods are considered for feature selection. The general structure of the algorithm for all three methods is given in Figure 5. The left most box is the dictionary of feature vectors. LDA on the right is used both for classification and extracting the relationship among combinations of features. This output is fed to a cost function to measure the discrimination power for that combination of features between classes. This measure will be used to select the best feature combination among other feature combinations. In this study, Fisher discrimination (FD) criteria,

\[
FD = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \tag{1}
\]

and Miss-Classification (MC) rate evaluated by the LDA are used as cost functions to quantify the discrimination between two classes.

Type I

Type I is a greedy search method. All of the feature vectors from the dictionary (from each class) enter to the LDA one by one and a corresponding cost is measured for each using the cost function. After this search is done over all \( NF \) feature vectors, the best feature is selected by comparing cost values of each feature vector. In the next step the second best feature vector which will do the best in combination with the first selected one is searched over the remaining feature vector set one by one. This procedure is run until the desired number of features is reached. Type I uses all the boxes and connections in Figure 5 except the feedback from the cost function to the dictionary. Since no dimension reduction is implemented this approach has high computational complexity.

Type II

Type II is a modified version of type I with an additional pruning module for dimension reduction. As in type I, feature vectors are fed to the LDA one by one and corresponding costs are measured. By comparing the cost values over all feature vectors the best feature is selected.

After the first feature is selected now we use the feedback path from the cost function to the dictionary as in Figure 5. The index of the selected feature corresponds to a node on the double tree which has a frequency tree index and a time tree index in that subband. In the frequency tree the nodes (subbands) which are not orthogonal to the selected frequency index are removed. Similarly in the time tree, the nodes which are not orthogonal to the selected time index are removed. This way only “good” potential feature vectors are kept in the dictionary; hence, the dictionary is pruned based on the last selected feature. Now the next
feature which will do best in combination with the first selected one is searched on the pruned dictionary. This procedure is run until the desired number of features is reached. Therefore, the only difference between type II and type I is that pruning is done on the dictionary based on the selected features.

Type III

This type is the simplest one. It does not use the LDA or a feedback path as in Figure 3. Instead, using cost function FD, a cost value is computed for each node on the double tree individually. Then a pruning algorithm as in (Saito et al.,) is run on the double tree from bottom to top to find the nodes with maximum discrimination power measured by the FD cost function. The sound signals from each class are analyzed in an off-line manner for feature extraction by one of the three methods explained in the preceding sections. After features are selected and the decision rules are set the system is run for validation.

RESULTS

We tested the proposed approach on the pistachio acoustic signals. We used a 2 times 2 fold cross validation method to estimate the classification performance. Basically half of the data set is used for training and the rest for testing. Then the test and train sets are swapped. This experiment is repeated 2 times. We use a frequency level $F=3$ and time level $T=3$ for the dual tree. After calculating the energy features in the nodes of the dual tree, they are converted to log scale. Table 1 shows the classification accuracies obtained with the proposed methods. The classification error obtained with the base line algorithm which is based on the time domain characteristics of the signals is given as BA. The comparison of all 3 types with their minimum errors and the number of features used to reach them are given. The Type-I and Type-II approaches which use the feedback from the classifier have outperformed the Type-III and BA approaches. As indicated before, the Type-III only uses the same tree pruning method of original LDB algorithm and does not account for the interactions between features. The classification accuracies obtained strongly indicate that the evaluation of the classification performance of the combined features in the training stage is important and should be preferred to evaluating the individual discrimination power of the features. Furthermore we note that Type-II approach used always a smaller number of features than the other approaches to achieve the minimal error.

Table 1. Open-Closed shell pistachio nut minimum classification errors (minimum) for the proposed types. NoF stands for the number of features used to reach minimum error.

<table>
<thead>
<tr>
<th>Type</th>
<th>Error (%)</th>
<th>NoF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>8.25</td>
<td>17</td>
</tr>
<tr>
<td>Type II</td>
<td>8.25</td>
<td>8</td>
</tr>
<tr>
<td>Type III</td>
<td>11.5</td>
<td>31</td>
</tr>
<tr>
<td>BA</td>
<td>18</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2. Effect of cost function for Type II.

<table>
<thead>
<tr>
<th>Cost Function</th>
<th>Error (%)</th>
<th>NoF</th>
</tr>
</thead>
<tbody>
<tr>
<td>FD</td>
<td>8.25</td>
<td>8</td>
</tr>
<tr>
<td>MC</td>
<td>11.25</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 1 shows the classification accuracies obtained with the proposed methods. The classification error obtained with the base line algorithm which is based on the time domain characteristics of the signals is given as BA. The comparison of all 3 types with their minimum errors and the number of features used to reach them are given. The Type-I and Type-II approaches which use the feedback from the classifier have outperformed the Type-III and BA approaches. As indicated before, the Type-III only uses the same tree pruning method of original LDB algorithm and does not account for the interactions between features. The classification accuracies obtained strongly indicate that the evaluation of the classification performance of the combined features in the training stage is important and should be preferred to evaluating the individual discrimination power of the features. Furthermore we note that Type-II approach used always a smaller number of features than the other approaches to achieve the minimal error.

To see the effects of the cost functions on the classification accuracy, the results for Type II with FD and MC as cost functions are given in Table 2. The selected feature set by evaluating FD has outperformed that of the MC and uses less number of features. This, we believe is due to the discrete nature of the MC compared the continuous nature of the FD when evaluating discrimination. The classification error curves versus the number of features for both results in Table 1 are given in Figure 6. Since each selected feature index is used to prune the dual tree structure, the orthogonalization of the representation space helps the classifier to use small number of features to achieve minimal error rate. Although the obtained classification accuracies of type-I and type-II are comparable, the number of used features carries significant importance for real-time applications. It reduces computational complexity.

Figure 6. Classification curves for all three types (Left), effect of cost function for Type II (Right).
In order to make an additional connection to the time frequency images of each class given in Figure 3, we prepared a time frequency map of the first 8 features of Type II. It reached the minimal error with this set. The intensity of each segment on the map is related with that feature’s individual discriminative power measured by FD cost function. The resulting map is depicted in Figure 7.

![Time Frequency Map](image)

Figure 7. The discriminant time-frequency map of the first 8 features of Type II. The darker regions have more discrimination power.

**CONCLUSION**

In this paper we described a new adaptive time-frequency plane feature extraction and classification algorithm to classify open and closed shell pistachio nuts. The system is based on the analysis of sound signals from impact acoustics. We applied the algorithm to particular pistachio nuts from Turkey. The results we obtained show that the algorithm is superior to a previous algorithm applied in this area. Based on the observations from the previously mentioned algorithm above, the proposed algorithm does not depend on a priori knowledge of time-frequency content of the signals under examination. Furthermore its adaptation capability to both time and frequency content of signals concurrently makes the algorithm a universal method for food kernel inspection which can resist the variability between nuts from region to region with respect to size and weight. One of the critical aspects one should consider is the environmental noise that can disturb the impact signal. Currently the authors are exploring other impact signals such as vibration which may be more robust to interference from neighboring sorters compared to acoustic signal.

**REFERENCES**