Cepstrum Based Feature Extraction Method for Fungus Detection

Onur Yorulmaz [a] , Tom C. Pearson [b] , A. Enis Çetin [a]

[a] Electrical and Electronics Engineering Department, Bilkent University, 06800
[b] USDA-ARS-GMPRC, 1515 College Avenue , Manhattan, KS, USA, 66502

ABSTRACT

In this paper, a method for detection of popcorn kernels infected by a fungus is developed using image processing. The method is based on two dimensional (2D) mel and Mellin-cepstrum computation from popcorn kernel images. Cepstral features that were extracted from popcorn images are classified using Support Vector Machines (SVM). Experimental results show that high recognition rates of up to 93.93% can be achieved for both damaged and healthy popcorn kernels using 2D mel-cepstrum. The success rate for healthy popcorn kernels was found to be 97.41% and the recognition rate for damaged kernels was found to be 89.43%.

Keywords: Cepstrum Analysis, Image Processing, Fungus Detection in Popcorn Kernels, SVM

1. INTRODUCTION

All grain kernels are vulnerable to fungal infection if not dried quickly before loading into a storage bin. Popcorn poses a particularly difficult problem in that the drying rate must not be too fast and the level of kernel moisture not reduced too much or the kernels will not pop. Thus, popcorn processors have the difficult task of balancing the drying rate of incoming popcorn with the risk of inadequate drying and possible fungal infection. One type of prevalent fungus that infests popcorn at harvest time is the Penicillium fungi which causes a dark blue blemish on the germ of popcorn commonly referred to as "blue-eye" damage. While this type of infection is not a health risk, it does produce very strong off flavors, causing consumers to reject the popcorn. Given the difficulty of balancing drying rate while maximizing the number of kernels that will pop, a certain percentage of blue-eye damaged popcorn kernels are inevitable every year. The blemish caused by the fungal infection is small and currently no automated sorting machinery can detect it. Thus, new methods for detecting and removing this type of defect is needed by the popcorn industry. It is possible to detect popcorn kernels infested by the fungi using image processing. In this article an image processing method is developed for this purpose. The method is based on cepstral feature extraction [10] from popcorn images and SVM based classification of cepstral features.

Mel-cepstral analysis is a major tool for sound processing including important speech applications such as speech recognition and speaker identification [4], [8]. The two dimensional (2D) extension of the analysis method is also applied to images to detect shadows, remove echoes and establish automatic intensity control [4], [5], [6]. Recently the 2D cepstral analysis was applied to image feature extraction for face recognition [2] and man made object classification [3].

The 2D mel-ceptrum is defined as the 2D inverse Fourier transform of the logarithm of the magnitudes of 2D Fourier transform of an image. Cepstral analysis is useful when comparing two similar signals in which one of them is a scaled version of the other one. This is achieved through the logarithm operation.

As it is based on the magnitude of the Fourier transform, mel-cepstrum is shift invariant. If a given image is translated version of another, only the phase of the Fourier coefficients change. Therefore two images can be compared to each other using only the Fourier transform magnitude based mel-cepstrum.

The Mellin-ceptrum is similar to the mel-cepstrum method, where the difference is that a log-polar conversion is applied to the logarithm of the magnitudes of the Fourier coefficients to provide rotational invariance. In log-polar conversion, the magnitude values which are represented in Cartesian coordinate indices are converted to polar coordinate system

representation. As a result of this conversion, rotational differences become shift differences. Shift differences can be taken out by Inverse Discrete Fourier Transform (IDFT) that is followed by an absolute value operation so that the Mellin-cepstrum of an image becomes rotation invariant.

In [3], a grid technique is introduced to reduce the number of total features in the classification phase. A grid is a set of bins with different sizes covering the matrix of Fourier transform coefficients. By averaging the Fourier coefficients in bins, it is possible to reduce the number of feature parameters representing a given image. In [3] this combination is achieved through power averaging.

In this article, popcorn kernel images are represented in mel and Mellin-cepstral domain. The blue-eye damage is visible in visible light spectrum and thus it is detectable given adequate spatial and color resolution. Furthermore the location of the damage is usually in the middle of the germ of the kernel which is used by Pearson in his image processing based fungus detection method [1].

(a)   (b)

Figure 1. (a) Healthy popcorn and (b) a blue-eye damaged popcorn kernel.

Pearson has developed a method [1] for detection of blue-eye damaged popcorn images. In his approach approximate location of the defect is used to find a curve which may contain a "valley like" shape in the presence of fungus. A row of the image signal is taken from an approximated y coordinate and a simple curve comparison is applied to make a decision. The germ color is white and the damage is observed as a dark region. Thus in [1] only the red channel is used considering that the damage is more visible on this channel. In our new approach using cepstral features, the red channel is used for the same reason. Pearson's approach was able to detect 74% of the blue eye damage in which the detection rates of healthy kernels was 91%. In this work, we are able to improve the detection rates which are presented in the experimental results in Section 4.

The rest of the paper is organized as follows; in Section 2, the cepstrum based feature extraction method is described. In Section 3, the Support Vector Machine (SVM) classification method which is used as a classification engine is briefly explained. The experimental procedure and the experimental results are presented in Section 4.

2. CEPSTRUM BASED FEATURE EXTRACTION

The 2D extension of the cepstral Analysis is defined as follows:

$$\tilde{x}(p,q) = F_2^{-1}(\log(|X(u,v)|^2)), \qquad (1)$$

where p and q are the 2D cepstrum domain coordinates, F_2^{-1} denotes the inverse 2D Fourier transform, and $X(u, v)$ denotes the Fourier transform coefficient of a given signal x at frequency locations given by u and v. In practice the Fast Fourier Transform (FFT) is used to calculate the Fourier transform of a signal while the Inverse Fast Fourier Transform

(IFFT) is used to calculate the inverse Fourier transform.

As an extension of 2D cepstral analysis method, the mel-cepstrum method applies grids to the Fourier transform of the signal and sums the energies of Fourier transform components within grid cells before computing the logarithm [2]. The goal of this is to reduce the size of the data and to emphasize some frequency bands. In the mel-cepstrum method, there is also a weighting process that emphasizes the important frequency values and reduces the contribution of the noise to the final decision. Five different grids are designed and studied for best grid selection. The weights are also selected to increase the higher frequency components' contribution. The DC value was multiplied by a small weight value where other high frequency grid cell energies are multiplied by relatively larger numbers. A sample grid and corresponding weight matrix are shown in Figure 2.

(a)   (b)
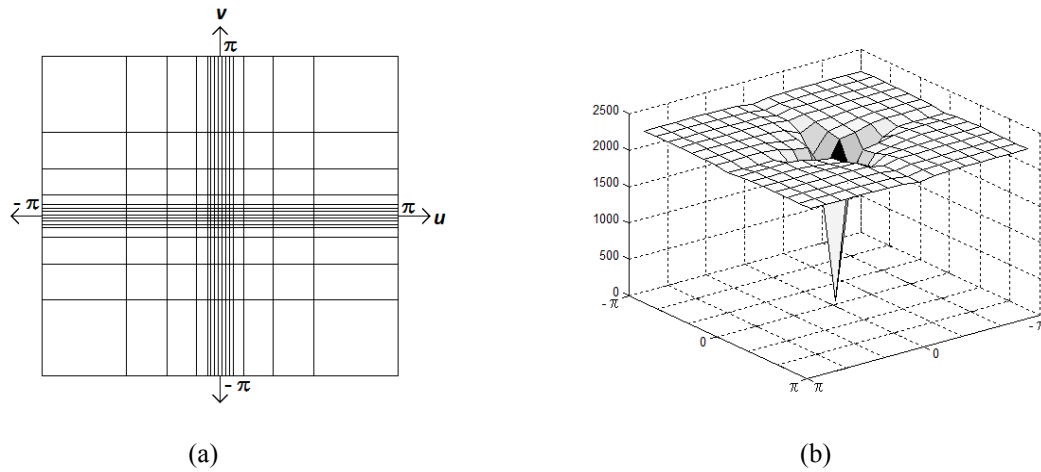
Figure 2. (a) A sample grid and (b) its corresponding sample weight.

In each grid cell the Fourier transform magnitudes are summed as follows;

$$g(m,n) = \frac{1}{\text{size}(B(m,n))} \sum_{u,v \in B(m,n)} |X(u,v)|, \tag{2}$$

where $g(m, n)$ is the value of the grid in coordinate locations m and n, and $B(m, n)$ denotes the (m, n)-th grid bin. The size($B(m, n)$) gives the total number of frequency components that are in the bin $B(m, n)$. Thus, the procedure of grid usage on the Fourier transform of an image can be illustrated as in Figure 3.
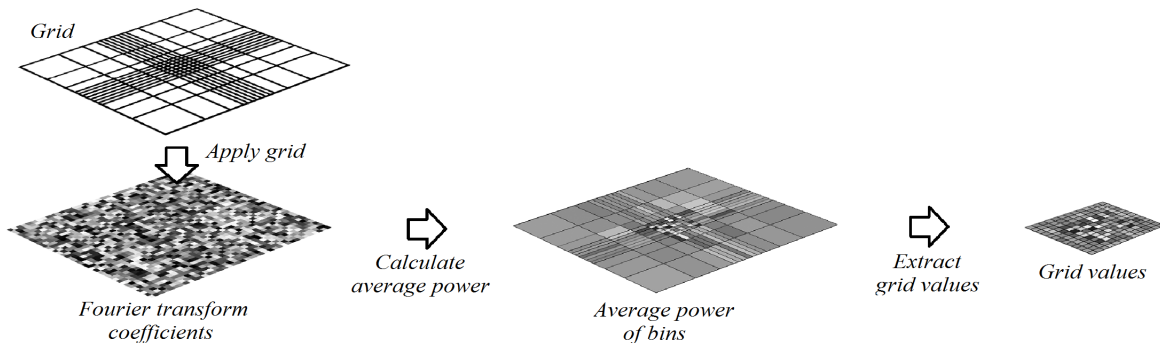

Figure 3. The process of applying a grid to the 2D Fourier transform coefficients of a given image.

After finding the grid values, the cepstral feature parameters are calculated as follows:

$$\tilde{x}(p,q) = F_2^{-1}(\log(|g(m,n)|)) \tag{3}$$

Thus, the mel-cepstrum procedure is summarized in six steps:

1. 2D Fourier transform of the given NxN image matrices are taken. Here N is selected to be a power of 2 such that $N=2^k$ where k is an integer. The images are padded with zeros before Fourier transform so that the width and height values are increased to N.
2. The absolute value of the Fourier coefficients are taken to apply grids. The grid bins are placed so that each frequency component resides inside a certain bin according to the grid implementation.
3. The grid features are taken by averaging. In this part of the procedure, the average of the elements of each bin are calculated. Then the averages are taken to be single features to be used in the process. While taking the features, their locations in the grid are also untouched, so that an MxM grid features matrix is calculated.
4. Each component of the grid features matrix is multiplied by a coefficient taken from designed weights matrix.
5. Logarithm of the grid features are taken to make use of the cepstral analysis method.
6. The inverse 2D Fourier transform is taken to find the final mel-cepstrum features.

After taking the absolute values of 2D Fourier coefficients, the procedure becomes translation invariant, however, it is still sensitive to rotational transformations. The rotational invariance is achieved in another method called Mellin-cepstrum that is an extension of the mel-cepstrum. In Mellin-cepstrum, after the application of grid process, a log-polar conversion is applied to achieve rotational invariance. The log polar conversion is defined as;

$$p(r, \theta) = g(e^r \cos(\theta), e^r \sin(\theta)), \tag{4}$$

where g(m, n) is the calculated average power of the Fourier transform inside the bin at coordinates m and n: the grid value, p is the frequency coefficient representation in polar coordinates and r and θ are the parameters of polar coordinate system. Here the parameters of the grid coefficient matrix may be non-integer values. This problem can be solved by finding the approximation of the values in those points through interpolation. The points in polar space that correspond to the outside of the Fourier matrix can be taken as zero.

The log-polar procedure is also given as follows:

1. 2D Fourier transform of the given NxN image matrices are taken. Here N is selected to be a power of 2 such that $N=2^k$ where k is an integer. The images are padded with zeros before Fourier transform so that the width and height values are increased to N.
2. Logarithm of the absolute value of the Fourier coefficients are taken to make use of the cepstral analysis method.
3. Grids are applied to resulting features. The grid bins are placed so that each frequency component resides inside a certain bin according to the grid implementation.
4. The grid features are taken by averaging. In this part of the procedure, the average of the elements of each bin are calculated. Then the averages are taken to be single features to be used in the process. While taking the features, their locations in the grid are also untouched, so that an MxM grid features matrix is calculated.
5. Each component of the grid features matrix is multiplied by a coefficient taken from designed weights matrix.
6. Log-polar transformation is applied to make the classification method rotation invariant.
7. The inverse 2D Fourier transform is taken.
8. To remove imaginary parts in coefficients, absolute values of the resulting coefficients are taken.

3. CLASSIFICATION

To classify mel-cepstrum features, a Support Vector Machine (SVM) is used [9]. The SVM is a supervised machine learning method that is developed by Vladimir Vapnik [7]. The method projects the feature vectors to a higher dimensional space to achieve a detectable distinction between given classes. In this work, the RBF kernel of the SVM was used. The features that are calculated at the end of the mel and Mellin-cepstrum methods, are applied to the SVM as vectors of features.

Since SVM is a supervised machine learning method, it has to be trained using manually classified prior data. During the training and test phases, various grid sizes are also studied. Experimental results are presented in the next section.

4. EXPERIMENTAL RESULTS

We tested the algorithm using two different data sets. The first data set includes 238 healthy and 238 damaged kernels of the same kind and the same harvest year. The images are taken at once with a Canon Powershot G11 digital camera and individual kernel images are extracted through a program written in Matlab. Pictures are taken with different exposure levels and the best exposure level emphasizing the difference of the two classes is selected. The example kernel images are shown in Figure 4.
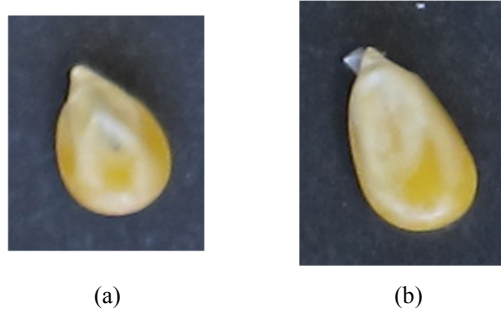

(a)   (b)

Figure 4. Popcorn kernel image samples (a) with and (b) without blue-eye damage. Notice that the image sizes may vary.

200 of the images from each of the classes were used for training and 38 are used for testing. Image sizes ranged from 113x225 to 155x164. To make use of the FFT algorithm, image sizes were expanded to 256 by adding zeros to the matrix. The leave one out principle is used in the tests. Both mel and Mellin-cepstral features are tested on the red color channel of the images. To achieve the best result five different grids were applied. The sizes of the grid value matrices and the number of resulting grid features are shown in Table 1. The grid size corresponds to the number of bins in horizontal and vertical directions in a grid. Since inverse Fourier transform was applied to the results of the grids which are real numbers, the grid feature values are symmetric with respect to the origin. The number of resulting features are also shown in Table 1.

Table 1. Grids, their grid sizes, Number of resulting features and the success rate of mel and Mellin-cepstral features on the first data set. The best results are achieved with mel-cepstrum using the NonUniform_Grid4.

Grid Name	Grid Size	Number of Resulting Features	Overall Success of mel-cepstrum	Overall Success of Mellin-cepstrum
NonUniform_Grid1	49 x 49	1225	83.82%	77.94%
NonUniform_Grid2	39 x 39	780	86.76%	79.41%
NonUniform_Grid3	35 x 35	630	85.29%	67.65%
NonUniform_Grid4	29 x 29	435	92.65%	88.24%
NonUniform_Grid5	11 x 11	66	67.21%	62.55%

It is observed that the best results are achieved using mel-cepstral features with "NonUniform_Grid4" which uses 435 cepstral coefficients. As the number of resulting features decreases from 1225 to 425, success rates increase because the noise like high frequency values are eliminated by the non-uniform grids. However, the success rates starts to decrease after 435 because of the fact that the number of features becomes insufficient to define classes as shown in Table 1.

The second dataset includes various popcorn kernels from previous harvest years. This provide further robustness to the algorithm towards the changes relative to the year of harvest and the seed variety. This dataset contains 398 healthy and 510 damaged kernels.

The images are taken in two different shooting modes: reflectance and transmittance. Reflectance images capture the light reflecting from the kernels, which are similar to the images in the first dataset. On the other hand, transmittance images are obtained in such a way that the light that passing through the kernels is captured. In other words, in

transmittance images, the light source is behind the kernels and in reflectance images, the source is in front of the kernels. Example images from the second dataset are shown in Figure 5. All of these images were aquired with a document scanner at a resolution of 4780x2950 pixels.

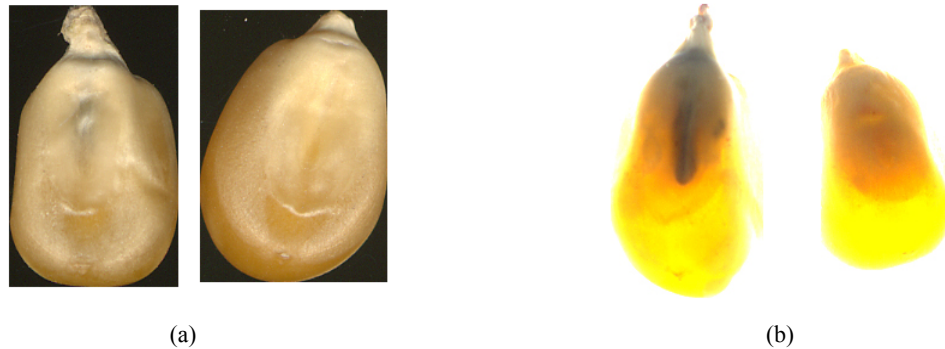| (a) | (b) |

Figure 5. (a) A pair of damaged and healthy kernels in reflectance capturing mode and (b) a pair of damaged and healthy kernels in transmittance capturing mode.

Having kernels from different varieties and years reduced the recognition rates below 90% in both reflectance and transmittance images. To increase the recognition rates three intensity based features are added which are; the mean of the pixel intensity values in red channel, the difference between mean and the minimum of the intensity values and the number of pixels with intensity values less than a given threshold i. The value i is determined by a Matlab program that finds the best i to maximize the detection rate. The second parameter related to the minimum intensity value is ignored in the tests with reflectance images since it reduced success rates because of the dark background. To determine the above mentioned intensity based parameters, popcorn kernel images were cropped in proportion with its size to focus on the approximate location of the suspected fungal infection regions. This procedure is explained in Figure 6.

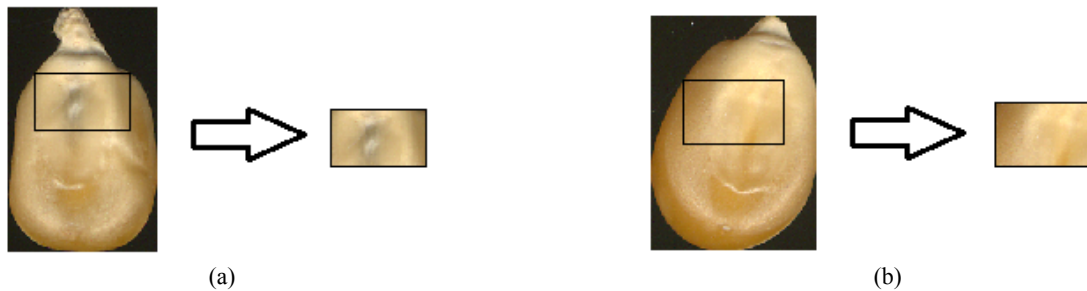| (a) | (b) |

Figure 6. Cropping of suspected fungus regions from (a) a damaged kernel and (b) a healthy kernel. Cropping is done on each image according to the left, right, top and bottom margins, which are calculated as a percentage of the height and width values of the popcorn image. Thus the margin values vary proportional to the height and width values of each image.

NonUniform_Grid4 is selected since it was the one producing the highest recognition rates in the tests with the first dataset. For both transmittance and reflectance images, 84% of the images are used for training and 16% are for testing. Resulting success rates are shown in Table 2.

Table 2. Healthy, damaged and overall recognition rates of mel and Mellin-cepstrum based classification method on transmittance and reflectance mode popcorn kernel images in the second dataset.

	Reflectance mode images			Transmittance mode images		
	Healthy	Damaged	Overall	Healthy	Damaged	Overall
Mel-cepstrum on NonUniform_Grid4	86.33%	78.85%	83.07%	97.41%	89.43%	93.93%
Mellin-cepstrum on NonUniform_Grid4	82.28%	78.28%	80.53%	91.14%	89.51%	90.43%

The experiments showed that the transmittance images are more successful to detect the damage. Also it is seen that Mellin-cepstrum did not provide a great improvement since kernels in the images do not have an significant rotational motion. The best results are achieved with mel-cepstrum with transmittance images which has shown an overall weighted recognition rate of 93.93%.

5. CONCLUSION

It is experimentally observed that it is possible to classify regular and fungus corrupted popcorn kernels using image processing. Popcorn kernel images are obtained using regular cameras. Cepstral parameters are extracted from popcorn kernel images and they were classified using support vector machines. Since the cepstrum is a shift and amplitude invariant, it is ideally suited for machine vision based sorting systems. Recognition results of 93.9% and 89.4% for healthy and infected kernels, respectively, represent a significant improvement over the previous method [1] using simple image processing steps. While results from both the new method presented here and in [1], were based on popcorn kernels from one harvest year and seed variety, this can be utilized in a real processing setting as kernels are processed all within one year and usually one variety at a time. More work is needed for adapting the method to slight variations in kernels from year to year and for different seed varieties. Different growing conditions can also effect seed morphology and color so further study is needed in this area.

Classification methods based on transmittance images show promise of higher classification results and appear to be more adaptable to more variation in seed size, morphology, and color. However, transmittance images may be more difficult to acquire in a high speed sorting application.

REFERENCES

[1]   Pearson, T. C., "Hardware-based image processing for high-speed inspection of grains," Computers and Electronics in Agriculture 69(1), 12-18 (2009).
[2]   Cakir, S. and Cetin, A. E., "Mel- and Mellin-cepstral Feature Extraction Algorithms for Face Recognition", The Computer Journal, (2010).
[3]   Eryildirim, A., Cetin, A.E., "Man-made object classification in SAR images using 2-D cepstrum," IEEE Radar Conference, 1-4 (2009).
[4]   Toreyin, B. U., Cetin, A. E.,"Shadow detection using 2D cepstrum," Proc. SPIE 7338 , (2009).
[5]   Chen, L. F., Liao, H. Y. M., Ko, M. T., Lin, J. C., Yu, G. J., "A new LDA- based face recognition system which can solve the small sample size problem," Pattern Recognition 33(10), 1713-1726 (2000).
[6]   Qin, J. and He, Z.-S. , "A SVM face recognition method based on Gabor featured key points," Machine Learning and Cybernetics 8, 5144-5149 (2005).
[7]   Boser, B. E., Guyon, I. M., Vapnik, V. N. , "A training algorithm for optimal margin classifiers," Proceedings of the fifth annual workshop on Computational learning theory, 144-152 (1992).
[8]   Yeshurun, Y. and Schwartz, E. "Cepstral filtering on a columnar image architecture: a fast algorithm for binocular stereo segmentation,"IEEE Transactions on Pattern Analysis and Machine Intelligence, 759-767 (1989).
[9]   Chang, C. C., Lin, C. J., "LIBSVM: a library for support vector machines," http://www.csie.ntu.edu.tw/~cjlin/libsvm/, (2001).
[10]   Bogert, B. P., Healy , M. J. R, Tukey, J. W., "The Quefrency Alanysis of Time Series for Echoes: Cepstrum, Pseudo Autocovariance, Cross-Cepstrum and Saphe Cracking," Proceedings of the Symposium on Time Series Analysis 15, 209-243 (1963).
[11]   Cetin, A. E., Ansari, R., "Convolution-based framework for signal recovery and applications," J. Opt. Soc. Am. A 5, 1193-1200 (1988).