PHASE AS AN ASSISTIVE FEATURE VECTOR FOR AUDIO CLASSIFICATION

Ioannis Paraskevas, Edward Chilton
(Centre for Vision Speech and Signal Processing)
School of Electronics and Physical Sciences, University of Surrey,
Guildford GU2 7XH, Surrey, United Kingdom

1 INTRODUCTION

Although a lot of research has taken place in the area of visual and speech recognition, the audio classification problem is not yet fully developed. The main reason is that the features that can be extracted from an audio utterance are not the same as the case for speech recognition. Up to now, the features extracted from the audio utterances have been magnitude and/or temporal related[1,2]. Magnitude related features are describing the evolution of the energy content of the signal, whereas the temporal related features convey information from the time domain of the signal. The phase content of a signal is ignored mainly due to difficulties in its computation. Discontinuities appearing in the phase spectrum of a signal are caused either by computational artifacts or by the signal itself. Phase discontinuities add to the phase spectrum, unwanted features that decrease the classification rate when used as part of a feature vector.

In the proposed method, the signal is transformed from the time to the frequency domain via the DTFT (Discrete-Time Fourier Transform). In this domain, phase is defined as the arctangent of the ratio of imaginary to real components of the fourier transform of the signal. Although the arctangent function is limited from – π to π, the phase spectrogram of a signal has to be a continuous function with unlimited range. After the phase spectrogram of a signal has been implemented, the resulting function has a lot of discontinuities because the angle values are wrapped around to zero. These 'phase jumps' are derived from the definition of phase (computational artifact) and not due to the nature of the signal itself. From now on, this kind of discontinuity will be referred as 'extrinsic' named after the source that causes it. The second kind of discontinuity is caused by the simultaneous zero crossing of the real and imaginary components of a signal in the frequency domain. In this case, the arctangent function produces an ambiguity (0/0). This case is equivalent to the existence of a pole/zero on top of the circumference of the unit circle at that frequency. The latter category of discontinuity will be referred as 'intrinsic', as it is caused by the internal structure of the signal and not due to computational artifacts.

In the proposed method, both sources of discontinuities are detected and removed. Then, statistical features are extracted from the magnitude and discontinuity-free phase spectrograms, of the signal. The statistical features derived from each spectrogram are used as target and test vectors to the classifier. The proposed method does not involve any a priori information[4]. The novelty of this technique is the method used to remove the phase discontinuities employing a different procedure for each case. The experimental results are showing that in certain classes, the discontinuity-free phase spectrogram outperforms the magnitude spectrogram. In the proposed method, the statistical analysis of phase is used as a feature extraction vector for classification, after its discontinuities are removed or compensated.

In section 2, theoretical discussion of the phase analysis via the Fourier transform and via the z-transform follows, while in section 3 a description of the statistical features that are extracted is given together with a description of the classifier. Finally in sections 4 and 5 results are reported and the advantages and disadvantages of each method are discussed.

2. METHOD

In this section, the implementation of the magnitude and discontinuity-free phase spectrograms are described. Initially, the signal is transformed to the frequency domain via the DTFT and the log-magnitude spectrogram is calculated by taking the logarithm of the absolute value of the complex representation of each acoustic utterance. Each signal is divided into equal-length frames (256 samples) and is windowed (Hanning window). In the case where the length of the last frame of the utterance is less than 256 samples the signal is zero-padded A frame length of 256 samples was adequate for these experiments as will be explained in section 3. The advantage of the spectrogram is that it provides a description of the energy evolution of the signal.

Up to now, phase has not been used as a feature extraction tool, due to the discontinuities that it exhibits. As mentioned before, the first kind is caused by the definition of phase (i.e. arctangent function) whereas the second type is due to discontinuities that exist in the structure of the signal. There are two methods employed to overcome both kinds of discontinuities. The first method uses the Fourier Transform, whereas the second one the z-transform. Each one of them has its advantages and disadvantages compared to the other, as will be analyzed in section 4.

2.1 Phase evaluation via the Fourier Transform

The signal is transformed into the frequency domain via the DTFT (Discrete Time Fourier Transform). In the frequency domain, phase is defined as:

$$f(\omega) = \tan^{-1}\left[\frac{\mathrm{Im}(S(\omega))}{\mathrm{Re}(S(\omega))}\right] \qquad \text{Eq. (1)}$$

where $0 < f(\omega) < 2\pi$, $S(\omega)$ is the complex Fourier spectrum of the data and Im(.), Re(.) are its imaginary and real components respectively.

After phase is calculated two problems arise. The first one is related to the discontinuities that the arctangent function presents (extrinsic discontinuities). Any phase value that is greater than 2π is wrapped around to zero[5]. Based on [5], any phase ‘jump’ over π between two consecutive phase angle values results to a compensation of ±2π. The basic concept of this technique, is the addition, where needed, of multiples of ±2π so as to smooth the transitions across branch cuts. Sometimes though, this method fails, especially when it applies to rapidly changing phase angles. Most of the researchers tend to ignore the discontinuities caused by the unwrapping algorithm, and they simply differentiate the unwrapped phase, as most of the features are encapsulated in the difference of phase rather than the phase itself.

The second category of discontinuities, that affects the classification rate, appears when both the real and imaginary parts of the signal are crossing zero simultaneously and causes phase ‘jumps’ of ±π. This kind of discontinuity, called ‘intrinsic’, is not detected by [5]. The simultaneous zero crossing of the real and imaginary part of the signal in the frequency domain is equivalent to the existence of a zero/pole exactly on the top of the unit circle [6]. Strictly, this kind of ambiguity is not a discontinuity but a rapid change of the phase angle. This ambiguity is caused by the signal itself and not due to computational artifacts as it occurs with the ‘extrinsic’ discontinuities. In terms of classification, rapid phase ‘jumps’ are affecting the classification rate. Phase ‘jumps’ caused by simultaneous crossing of zero of the real and imaginary part of the signal causes π or near π ‘jumps’. So, the algorithm designed by [5] does not always detect and smooth them, as its threshold is limited to π. So, the proposed algorithm can be summarized as: when there is a phase ‘jump’ greater than π or when there is a simultaneous zero-crossing (real and imaginary component), the ±2π compensation takes place [5] and the ±π compensation takes place respectively. So, initially, the 'extrinsic' discontinuities are detected and removed (±2π compensation) and then the 'intrinsic' ones (±π compensation). For the 'extrinsic' discontinuities the compensation rule can be seen in [5]. For the 'intrinsic' ones the compensation rule is: starting from the beginning of the phase spectrum, if the phase angle of the 'critical' point (i.e. simultaneous zero-crossing of the real and imaginary component) is greater compared to the phase angle of the point before it, then –π is added to the rest of the signal and vice versa.

Furthermore, there is one more theoretical aspect that has to be highlighted. This is the case when the 'intrinsic' discontinuity 'jump' is slightly higher than π and effectively it is detected and compensated by both 'extrinsic' and 'intrinsic' unwrapping methods. Analytically, when a 'critical' point (simultaneous zero crossing of the real and imaginary part) causes a phase 'jump' slightly higher than π then according to [5] a ±2π compensation takes place. This compensation though is incorrect as it should be limited to ±π only. If a frequency point is π (or slightly more than π) radians higher than the previous one (due to a simultaneous zero crossing) then a 2π radians will be subtracted. This subtraction of 2π is incorrect as it should be only π. Then in the second scan of the phase spectrum the proposed algorithm will detect that this particular phase point belongs to the 'critical' ones and comparing its value with the one before will detect that the 'critical' one is π radians lower and it will add to that and the rest of the signal π radians so as to compensate. The same case holds if a frequency point is π (or slightly more than π) radians lower than the previous one. This is the reason that the 'extrinsic' discontinuities [5] have to be removed first and the 'intrinsic' ones after.

In the next few lines, an example follows that demonstrates the case in which algorithm [5] fails to detect such phase 'jumps'. Consider a signal of multiple excitations, comprised of 256 samples of 26 equally spaced unit samples, repeating themselves every 10 samples whereas the rest samples are fixed to zero value [3]. Figure 1 shows the signal in the time domain, figure 2 shows the magnitude spectrum of the signal and figures 3 and 4 show how the real and imaginary part evolves across frequency

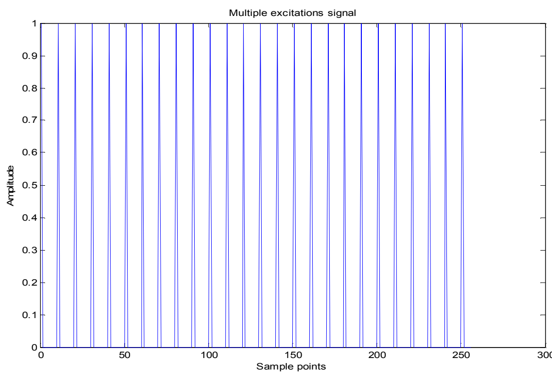


Fig. 1: Multiple excitation signal

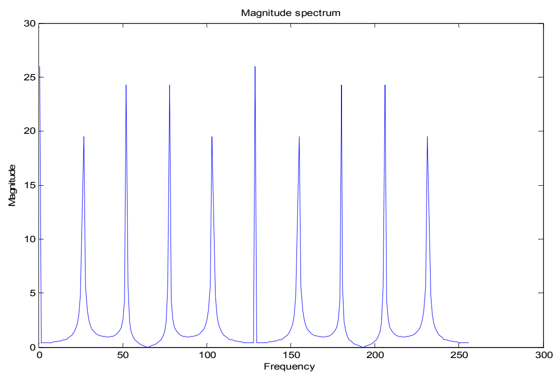


Fig. 2: Magnitude spectrogram of the multiple excitation signal

Finally, figure 5 shows that the wrapped and unwrapped phase is exactly the same. This example proves that there are cases in which the 'unwrapper' of [5] is not appropriate to detect and smooth rapidly changing phase angles that there value is slightly less than π. Also, comparing figures 3, 4 and 5, it is clear that these 'intrinsic' discontinuities occur when there is a simultaneous zero crossing of the real and imaginary part of the signal in the frequency domain. Finally, figure 6 shows that using the proposed technique the phase of the signal is unwrapped.

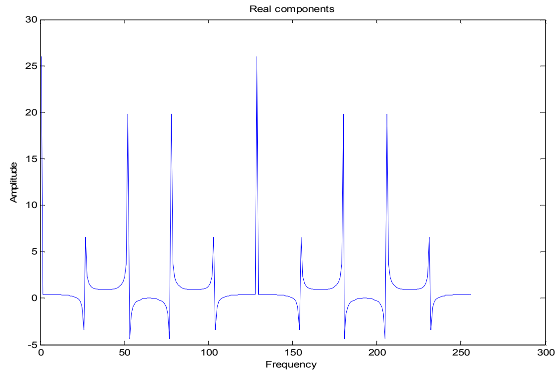


Fig. 3: Real part of the multiple pulses

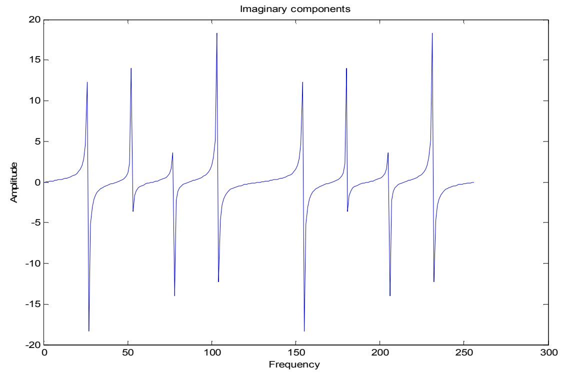


Fig. 4: Imaginary part of the multiple pulses

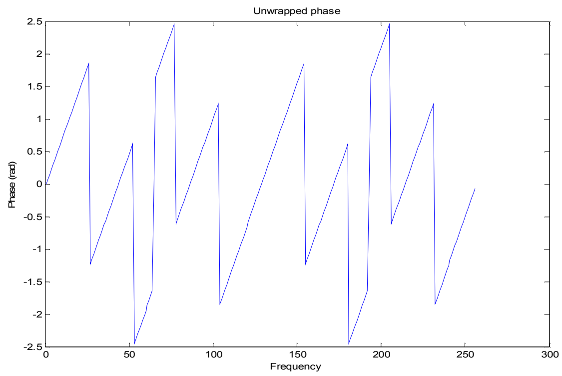


Fig. 5: Unwrapped phase of the multiple pulses (based on [5])

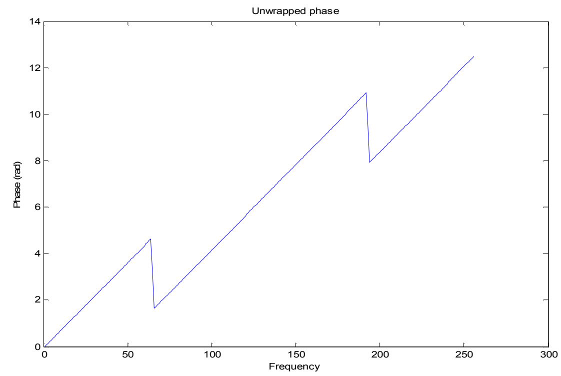


Fig. 6: Unwrapped phase of the multiple pulses (based on the proposed method)

Summarizing, the proposed method comprises two steps. Initially, the 'extrinsic' discontinuities are corrected based on algorithm [5]. Then, the phase spectrum is scanned again and having detected and stored the 'critical points' from before, the second level of compensation takes place, based on the 'critical' points only. The disadvantage of the proposed algorithm is that for non-artificial signals it is not always clear whether a correction should take place or not [5], i.e. if the discontinuities are caused due to rapidly changing angles or due intrinsic characteristics of the signal other than simultaneous approach to zero.

2.2 Phase evaluation via the z-transform

An alternative approach to phase analysis can be made in the generalized z-domain. The z-transform of each signal can be formed from a polynomial in powers of z with time data values as coefficients [7]. The roots of the polynomial are the zeros located in the z plane. This model is used to evaluate the signal phase response around the unit circle. The phase spectrogram in the z-domain is evaluated from the following algorithm,

$$f(\omega) = \tan^{-1}\left[\frac{\mathrm{Im}(S(\omega)) - \sin(w)}{\mathrm{Re}(S(\omega)) - \cos(w)}\right] \quad \text{Eq. (2)}$$

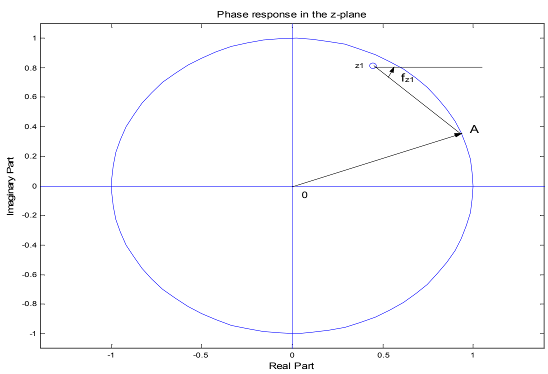
$$0 < f(\omega) < 2\pi$$

and

$$F_{TOTAL} = \sum_{k=1}^{l} f(\omega) \quad \text{Eq. (3)}$$

where,

$S(\omega)$ is the complex Fourier spectrum of the data, Im(.), Re(.) are its imaginary and real components respectively, *w* represents each frequency point on the unit circle, *k* the zeros used to evaluate the phase spectrogram and *l* the last zero to be taken under consideration. A schematic representation of the phase spectrogram calculation follows fig. 7,



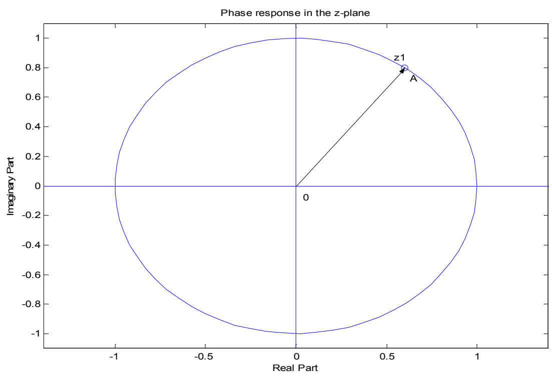


Fig. 7: Phase response of a zero in the z-plane

Fig. 8: Phase response of a zero located on the circumference of the unit circle in the z-plane

where A is the reference point, z1 is a zero point and fz1 is the argument of the zero with respect to A [7].
Analytically, the phase contributions from each zero z1 are evaluated according to (2), for each frequency A (3) around the unit circle. To evaluate the phase spectrogram, the unit circle has to be divided into a certain number of sampling intervals. The phase contribution is calculated based on eq. 2 and then the contributions from each zero are added up eq. 3 with respect to the first point. This value is stored and the same procedure is repeated for the next point on the unit circle. The distance between the first and the second point etc. depends on the sampling interval. This process terminates after the same procedure is repeated for the whole number of points in which the unit circle is divided into.

The extrinsic discontinuities can be overcome if the unwrapping algorithm is substituted by the geometric evaluation of the phase. In this case, although each phase component still uses the arctangent function, its value is constrained to $\pm\pi$ radians and no ambiguity arises i.e. phase is not wrapped around zero and consequently 'extrinsic' discontinuities do not arise.

In the proposed method the signal is transformed to the z-domain and its zeros (roots of the signal) are plotted around the unit circle. The roots lying on the circumference of the unit circle are causing the ‘intrinsic’ discontinuities [6]. So, by removing the zeros on, or close to, the unit circle and reconstructing the phase with those remaining zeros, the discontinuities should do not exist. From figure 8, it is clear why zeros lying on the circumference of the unit circle are causing discontinuities [7]. In that case, based on Equation (2), the imaginary and the real part of the signal become equal to the sin and cos respectively and consequently the 0/0 ambiguity arises.

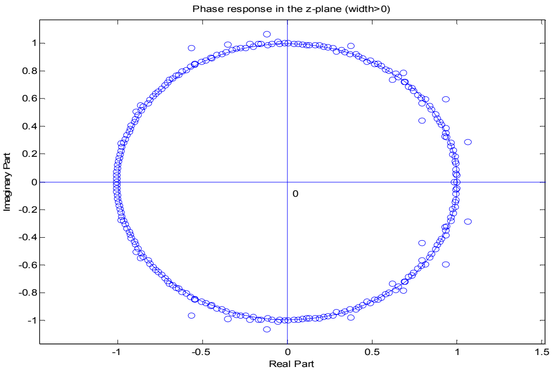


Fig. 9: Zeros location in the z-plane of a single 256 samples long frame from class (vii)

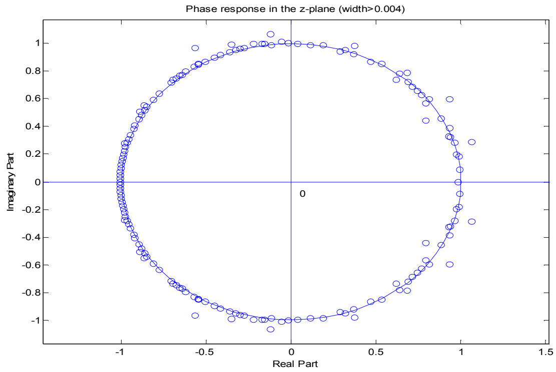


Fig. 10: Zeros location in the z-plane of a single 256 samples long frame from class (vii), 'ring' width greater than 0.004

Figures 9 and 10 demonstrate how the zeros of a 256 samples frame from an utterance of class (vii) are distributed around the unit circle, for different 'ring' widths. The circles are denoting where the zeros are located. Some of them are located exactly on top, or very close to the circumference of the unit circle. So, by drawing a 'ring' with a certain width, the ones located very close or on top of the unit circle are removed. Of course, there is a trade-off between the zeros that have to be removed and the amount of information loss caused due to their removal. So, it is important to detect the width of the 'ring' so as to obtain the optimum classification rate.

3 CLASSIFIER AND STATISTICAL FEATURE EXTRACTION

3.1 Classifier

The Mahalanobis classifier [10] was employed in order to classify the acoustic patterns. The Mahalanobis distance is defined as:

$$d(x_t, x_r) = (x_r - x_t)C_r^{-1}(x_r - x_t)^T \quad \text{Eq. (4)}$$

where x_t and x_r are the reference and the test vectors respectively and C_r is the covariance of the reference data.

A codebook is derived from the mean values of each class, so each class is represented by each codeword. Then, the distance between each test pattern and the codeword is calculated using distance metric. Therefore, the test pattern is classified based on the minimum distance between itself and the codeword. The Mahalanobis distance classifier is employed to estimate the classification rate in all three spectrograms (magnitude, phase via the F.T. and phase via the z-transform).

3.2 Statistical feature extraction

After all three spectrograms are formed, the statistical feature extraction process follows. The information that each spectrogram conveys has to be compressed before it will be presented to the classifier. An effective way to compress information, is the calculation of statistical features from each spectrogram. The eight statistical values calculated were [8,9] :

i) variance, ii) skewness, iii) kurtosis, iv) entropy, v) range, vi) inter-quartile range,vii) median and viii) mean absolute deviation.

4. OBSERVATIONS

In this section, the procedure followed will be described as well as practical issues that arise due to computational artifacts. Initially (time-domain), each audio utterance is normalized and then divided into frames of equal length (256 samples). In all three cases (magnitude spectrogram, phase spectrogram via the z-transform and phase spectrogram via the Fourier transform), the last frame of the utterance is zero-padded when the number of samples is less than 256.

4.1 Database

The database consists of ten different classes of gunshot sounds. Namely, i) firing a revolver with echo, ii) firing a .22caliber handgun, iii) firing an M-1 rifle, iv) firing a World War II German rifle, v) firing a cannon, vi) firing a 30-30 rifle , vii) firing of a .38 calibre semi-automatic pistol , viii) firing a lever action Winchester rifle, ix) firing a 37mm anti-tank gun and x) firing a pistol. The length of each sample was not the same, due to the non-artificial nature of the data. Each class consisted, on average, of ten audio samples. Seven of them were used as training data, whereas the rest was used as test data.

4.2 Log-magnitude spectrogram

The signal is transformed into the frequency domain, via the DTFT (Discreet-Time Fourier Transform). Analytically, the signal is divided into frames (256 samples each), windowed (Hanning window) and transformed into the frequency domain. Finally, the log-magnitude value is calculated for each frequency point. The number of frequency points in which the spectrum is divided into, equals the signal's number of samples.

4.3 Phase spectrograms

4.3.1 Phase spectrogram via the Fourier Transform

As mentioned in section 2.1, after the signal has been transformed into the frequency domain, the phase spectrogram is implemented based on Equation (1). A conventional phase unwrapping algorithm was applied [5] to the 'wrapped' phase so as the 'extrinsic' phase discontinuities to be removed. Then, the 'intrinsic' discontinuities were removed by applying the $\pm\pi$ compensation whenever there was a simultaneous zero crossing of the real and imaginary component of the signal. The zero-crossing of the real and imaginary component of the signal is equivalent to a phase 'jump' from either the first to third or second to fourth quadrant and vice versa. These 'jumps' are causing π (or near π) phase angle changes. Figures 11 and 12 are showing the phase spectrograms implemented via [5] and via the proposed method, respectively.

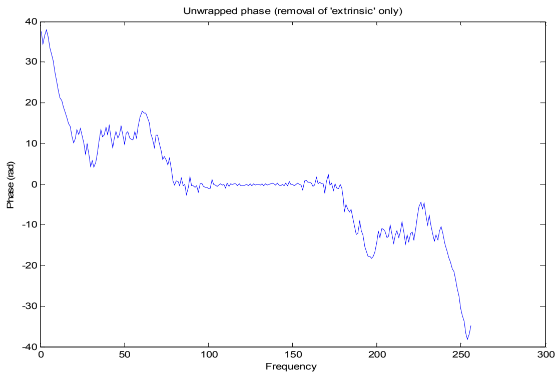


Fig. 11: 'Unwrapped' (removal of 'extrinsic'only, based on [5]) phase of a single 256 samples long frame from class (v)

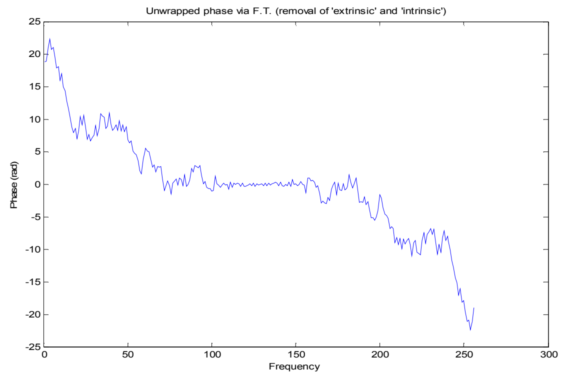


Fig. 12: 'Unwrapped' (based on the proposed method) phase of a single 256 samples long frame from class (v)

The advantage of this technique is that, after the appropriate compensation, all the samples are used for the construction of the phase spectrogram i.e. there is no information loss. Its disadvantage is the lack of precision that the limited accuracy produces [11]. Analytically, it is not possible to specify the exact area that a zero-crossing occurs, as the sampling rate cannot be infinite. Preliminary experimental results are showing that the classification rate reaches its optimum value when the number of frequency points (i.e. frequency rate) is similar to the frame size (i.e. 256 sample points). Finally, based on experimental results, it is shown that, although 'intrinsic' discontinuities are caused due to rapidly changing angles (i.e. because of the internal structure of the signal) and not due to computational artifacts, their existence reduces the classification score.

4.3.2 Phase spectrogram via the z-transform

The signal, in the time domain, is divided into frames of 256 samples. So, the signal forms a polynomial of power 255 with 256 coefficients. The roots of this polynomial are evaluated and plotted in the z-plane [7]. As mentioned in section 2.2, the phase spectrogram of the signal is constructed from the roots of the polynomial. The roots lying on (or very close to) the circumference of the unit circle are introducing ambiguities [6] ('intrinsic' discontinuities'). So, the phase spectrogram is constructed by removing the zeros that exist on top or very close to the unit circle.

The disadvantage of this method arises from the lack of accuracy introduced when the roots of the polynomial are calculated [11, sect. 7]. The higher the order of the polynomial, the higher the round-off error and consequently the zeros are not located in the exact place on the z-plane. To keep this kind of inaccuracy in low levels, the frame length is limited to 256 samples.

In both cases, 4.3.1 and 4.3.2, after the phase spectrograms are implemented, their difference is calculated. According to experimental results, the difference of the phase spectrogram provides higher classification scores.

Summarizing, the proposed method consists of the following steps. After each audio event is normalized, the magnitude spectrogram and phase spectrograms (via the Fourier transform and via the z-transform) are implemented. Then, eight statistical features are extracted from each of the three curves derived for each acoustic utterance. Finally, the statistical features derived, are introduced to a Mahalanobis distance classifier.

5 EXPERIMENTAL RESULTS

The three methods (log-magnitude spectrogram and phase spectrograms) will be applied to the database and the overall scores, of each stream independently, will be presented. Moreover, how all three streams appropriately combined, increase the overall classification rate will be shown. Also, analytic results for these classes that the phase spectrograms outperform the magnitude spectrogram will be provided.

Based on the method described in Section 2.2 (phase spectrogram via the z-transform), the classification rates for various 'ring' widths are provided in table 1. The range of values for the width of the 'ring' is between 0 up to 0.001. The maximum 'ring' width was limited to 0.001, as further increase would result in unreliable results due the limited number of zeros taken under consideration. Table 1 demonstrates the overall classification scores for all ten classes. As it is clear, the classification score reduces as the 'ring' width increases. The decrease of the number of zeros taken under consideration, affects the classification score. It is important to note that the classification score reaches its highest value when the 'ring' width is 0.0003. In this value, the zeros located on the circumference of the unit circle as well as most (if not all) of the zeros located very close to the circumference of the unit circle are removed ('intrinsic' discontinuities).

	classif. scores (%)
ring width	disc. free phase via the z-trans
>0	78.6
>0.00003	80.0
>0.00006	78.6
>0.0001	74.3
>0.0003	70.0
>0.0006	68.6
>0.001	64.3

Table 1: Overall classification scores in % for various 'ring' widths

The highest overall classification score was obtained when the log-magnitude spectrogram was employed, followed by the phase spectrogram via the z-transform. Analytically, the overall results are, log-magnitude spectrogram: 91.4%, phase spectrogram via the z-transform: 80.0% and phase spectrogram via the F.T.: 75.7%. Although the log-magnitude spectrogram provided overall the highest classification score, there were certain classes in which one or both the phase spectrograms outperformed it (table 2). Analytically, at least one of the phase spectrograms performed better or similar to the log-magnitude spectrogram in the second, fourth, fifth and tenth classes.

	classif. score in (%)		
	magnitude spectr.	phase spectr. (via F.T.)	phase spectr. (via z-tr.)
class 2	83.3	91.7	91.7
class 4	63.6	81.8	72.7
class 5	81.8	90.9	81.8
class 10	91.7	83.3	91.7

Table 2: Classification scores in % for classes 2, 4, 5 and 10 using the Fourier magnitude spectrogram and the and the phase spectrograms independently

The combination of the three classification streams in certain ways could reinforce the reliability of the system and/or increase the recognition rate. In this part, it will be described possible ways to combine the three classification streams (Fourier magnitude spectrogram, phase spectrogram via the F.T. and phase spectrogram via the z-transform). The first way is to combine the Fourier magnitude with either the phase via the F.T. or the phase via the z-transform classification streams (i.e. first and second or first and third branches (figure 13)). In that way, when both classification streams coincide to a certain decision, the utterance is classified to that class. So, although the overall score is reduced, the reliability of the system is increased considerably. Experimental results show that the percentage of both streams to misclassify the same utterance is very low.

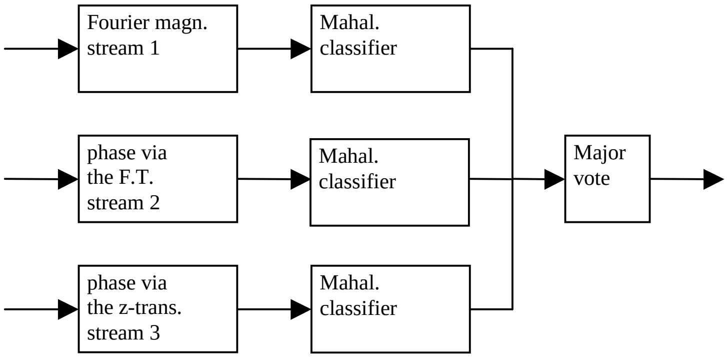


Fig. 13: Decision rule involving all three classification streams

The other combination is to use all three streams simultaneously and if two or more of them coincide (major vote), then classify the utterance to a certain class (figure 13). In that way, although the overall score is reduced compared to the Fourier magnitude stream, the classification rate for the four classes (table 2), is increased. For the four classes, when the three streams were used independently, the classification score of each one of them is, for the log-magnitude spectrogram: 80.4%, for the phase spectrogram via the z-transform: 84.8% and for the phase spectrogram via the Fourier Transform: 87.0%. Based on the decision rule (figure 13), the classification score for the four classes is increased to 95.7% which is higher compared to any of the three independent streams.

Summarizing, in this paper a new feature extraction method was presented that uses phase in the feature extraction process. The phase points that cause ambiguities or discontinuities are detected and removed. Two different techniques are presented i.e. phase spectrogram via the z-transform and phase spectrogram via the F.T. Both techniques were applied to a database of different kinds of gunshots and the experimental results demonstrated that there were certain classes in which the classification score derived via the phase spectrograms outperformed the score derived via the log-magnitude spectrogram, which is the widest used and most reliable technique for audio classification. This shows that phase is an important part of the information that a signal conveys and could be employed as an additional feature vector to assist the classification process. Finally, the appropriate combination of two or more of the classification streams provides a more reliable system and/or increases the classification score.

References

1. Y. Wang, and Z. Liu, “Multimedia Content Analysis using both Audio and Visual Cues,” IEEE Signal Processing Magazine, November 2000.
2. T. Zhang, and C.C.J. Kuo, “Audio content analysis for online audiovisual data segmentation and classification”, IEEE Transactions on Acoustics, Speech, and Audio Processing, vol. 9, no. 4, May 2001.
3. Sanjeev Sarpal, “Efficient System Identification based on Root Cepstral Deconvolution”, Ph.D. thesis, University of Surrey, January 2003, ch. 7.

4. C.L. Nikias, and A.P. Petropulu, Higher-order spectra analysis: a nonlinear signal processing framework, Prentice Hall signal processing series, 1993, ch. 6.
5. J.M. Tribolet, “A new phase unwrapping algorithm”, IEEE Transactions on Acoustics, Speech and Signal Processing, volume ASSP-25, pages 170-177, April 1977.
6. Hamid Al-Nashi, “Phase Unwrapping of Digital Signals,” IEEE Transactions on Acoustics, Speech and Audio Processing, vol. 37, no. 11, November 1989.
7. J.G. Proakis and D.G. Manolakis, Digital Signal Processing Principles, Algorithms, and Applications, Macmillan Publishing Company, 1992, ch.4 and 5.
8. E. Mansfield, Basic Statistics with Applications, W.W.Norton and Co., 1986.
9. A. Papoulis, Probability and Statistics, Prentice-Hall, Inc., 1990, ch. 12.
10. P.C. Mahalanobis, “On the generalized distance in statistics,” Proceedings of the National Institute of Science of India, 12 (1936) 49-55.
11. G.E. Forsythe, M.A.Malcom, and C.B.Moler, Computer Methods for Mathematical Computations, Prentice-Hall, 1977.

ACKNOWLEDGEMENTS

The author would like to thank Professor Josef Kittler and Dr. Philip Jackson for reviewing an earlier version of that paper and making useful suggestions.