

Monophonic Musical Instrument Sound Classification Using Impulse Response Modeling

Rutuja S Kothe¹, Dr D G Bhalke², Dr.Anupama A Deshpande³

JJTU, Rajasthan¹

AISSMSCOE, Pune²

JJTU, Rajasthan³

¹rskothe@gmail.com, ²bhalkedg2000@gmail.com, ³aadeshpande@gmail.com,

Article History: Received: 11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2021

Abstract: The field of music has promising commercial and social applications. Hence it has attracted the attention of researchers, engineers, sociologists and health care peoples. Therefore this particular research area has been selected.

In this manuscript the monophonic musical classification system using impulse response of the system is presented. In this research work 19 musical instruments monophonic sounds from 4 families are classified using WEKA classifier. The impulse response of all musical instruments and families are computed in Cepstral Domain. As Impulse response is used to model the body response of the musical instruments and helps to capture the information. It is different for different instruments. The features are extracted from impulse response and presented to WEKA Classifier.

The Musical instrument classification for individual instruments and family is verified using impulse response modeling. It is found that the impulse response is different for different instruments. It helps to easily distinguish between instrument to instrument and family to family. For individual instruments, the average classification accuracy has been obtained is 83.23% and 85.55% for family classification.

Keywords: Musical instrument classification; impulse response; Cepstral.

Introduction:

The musicology has not only attracted the attention of inherent artists but also, scientists, engineers, sociologist and specialist health care of human beings. since musical instruments are inseparable part of technology has always attracted attention of the technologist.

In view of this efforts have been made to present the database that has been used in the research work. Homo Sapiens particularly mankind has inherent capability to interpret and classify sound emanated by different creatures, instruments in the nature under varied condition and situation like entertainment involving singing, dancing that soothes mind of human beings. The aural system of human body collects the sounds emanated in the nature and through neurons of the nervous system of the ear (eustachian tube) transmits the impulses to the brain for further processing, analysis and understanding.

The receptors in the human brain analyses recognize and classify the sound impulses after processing the same for building the data base for further use. However, on account of some doubts certain questions still remain unclear uncomprehending and unanswered about how brain recognize the sounds. Also some questions like, Which type of sound information does brain receive from human auditory system? , What are the crucial features to be analyzed like redundant characteristics features of sound that are not useful and may cause confusion in recognition of genuine impulse or musical notes and their classification process?. The existing literature some algorithms are proposed which classify sounds using extracted features. However, none of these algorithms, match so closely to the recognition skill or ability of human brain through audio nervous system.

Literature Survey:

The Brown et al. [1] presented work on classification of musical sounds using computer. Again, Brown et al. [2], [3] used cepstral coefficients features extracted using constant-Q transform method with the help of k-means clustering classifier. The error rate observed is around fifteen percentages among oboe and saxophone. At the same time, Martin and Kim [4] designed a system which accurately classified fifteen musical instruments with the help of isolated notes. Author used training and testing data which was recorded from several musical instruments. An error rate of 28.40% was reported by the author. In addition, Marqueset al. [14] presented system using GMM and SVMs for monophonic instrument sound identification. An accuracy of 70% was reported for eight musical instruments (violin, lute, organ, bagpipes, trombone, piano, and harpsichord, clarinet). Consequently, Eronen and Klapuri [5], used features based on cepstrum, with twenty one other features which are Rise time, Decay time spectral spread, spectral centroid, Fundamental frequency, Amplitude

and frequency modulation rate for instrument sound classification. Classification accuracy ranging from 30% to 80% is being reported for thirty instruments playing monophonic musical sounds. Eronen et al. [6] performed experiment on wide range of features including MFCC and Linear Prediction coefficients. The author has critically analyzed 23 features and their relevance for sound classification. In all analysis MFCC feature found very effective in classification of instruments in most of the cases.

Commencing with earlier work, Agostini et al. made use of QDA features for classification of 27 monophonic musical instrument sounds. The results are verified using different classifiers such as Support vector machine, canonical discrimination, and KNN [8]. Authors reported an error rate of 7.2% for individual instruments. For instrument family classification an error of 3.1% is reported. Using pitch tracking algorithm. Again, Lee et al. [9], in 2002, used Hidden Markov Models (HMMs) to identify three different musical instruments which are violin, oboe and flute with 70% classification accuracy. Eronen et al. made use of MFCCs, and its derivative features and classifier used was HMM. The set of statistically independent features were extracted using ICA from the training data set. The recognition accuracy of 60% was reported using MFCCs and delta MFCCs for McGill university master sample database [12]. This dataset is a professionally recorded instrument samples and used in [5], [7],[9].

Meanwhile, Costantini et al. [13] used a Cepstral features, constant Q transform for classification of six instruments using neuro fuzzy model. The notes which are used are the fourth octave in the frequency range from 261.6 Hz to 522.2 Hz in their research work. An error rate of 15%-20% was reported for cepstrum-based features. Later, Essidet al. in 2004 [15], [16] performed musical instrument classification with less numbers of features that included Mel frequency Cepstral Coefficients (MFCCs). Further, Partridge and Jabri used (PCA) Principal Component Analysis [17] to increase the classification accuracy. Rodet et al. [18], [19] used 20 attributes for classification of real time solo performance. Using these features, authors reported 85% recognition accuracy for solo classification using k-Nearest Neighbor algorithm. Authors used 07 different instruments which are violin, cello, piano, flute, guitar, clarinet and bassoon, with a database of 108 recordings, played by different musicians. The importance of validation of cross database was analyzed in this research work.

Methodology:

In this section, musical instrument classification using impulse response modelling is described. The impulse response is computed in Cepstral Domain. Till now Cepstral domain features were widely used for speech processing applications, But not being used for sound processing applications. The Vocal tract and excitation response are separated in cepstral domain. This technique is used to model body of musical instruments. The impulse response is used as feature vector in this. The impulse response of the body of instrument is found using LPC coefficients. The procedure to compute LPC coefficients is shown in fig. 1.

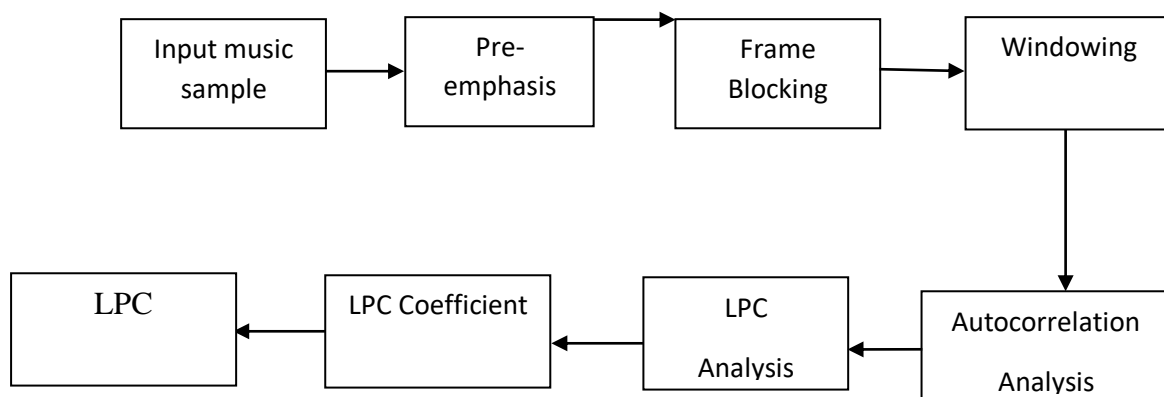


Fig1: Block diagram to compute LPC Coefficients

The concept behind the linear predictive analysis is that the music sample can be approximated as a linear combination of past music samples. Music data in the form of waves and these waves are divided into frames. Autocorrelation is used to measure the similarity between a signal and its delayed version. Cepstral coefficients represent essential information for speech, speaker, and musical instrument recognition. A differentiator is used to separate the delta coefficient and static coefficient.

Cepstral Domain Window Method

Here the procedure to compute cepstral features is described. The method used for the separation of the impulse response of the body of the instrument and the excitation response of the musical instrument is discussed. The algorithm to find the body response of the instrument is discussed here.

Algorithm to find body response :

In this algorithm, the entire note is used for processing.

Input musical sound signal

- Compute FFT of sound signal
- Plot the spectrum of signal
- Compute complex log spectrum
- Compute IFFT (Cepstrum)
- Plot cepstrum of signal
- Use window: cepstral Domain Windowing
- Take complex FFT of windowed cepstrum
- Find magnitude of FFT and take anti-log
- Compute complex IFFT which is time domain body response and excitation response of signal

The above algorithm is used to compute body response and impulse response of the different musical instruments. Mathematically the cepstrum can be represented by following equation.

$$Cepstrum = IFFT(\log(FT(m(n))))$$

The cepstrum analysis is used to separate the excitation source and instrument filter response. Excitation source and instrument filter response can be separated with the use of cepstrum analysis. Cepstrum values near origin gives information related to the instrument filter response. Cepstrum values away from origin gives excitation source information. The cepstrum for a note of an instrument is shown in Fig. 3.

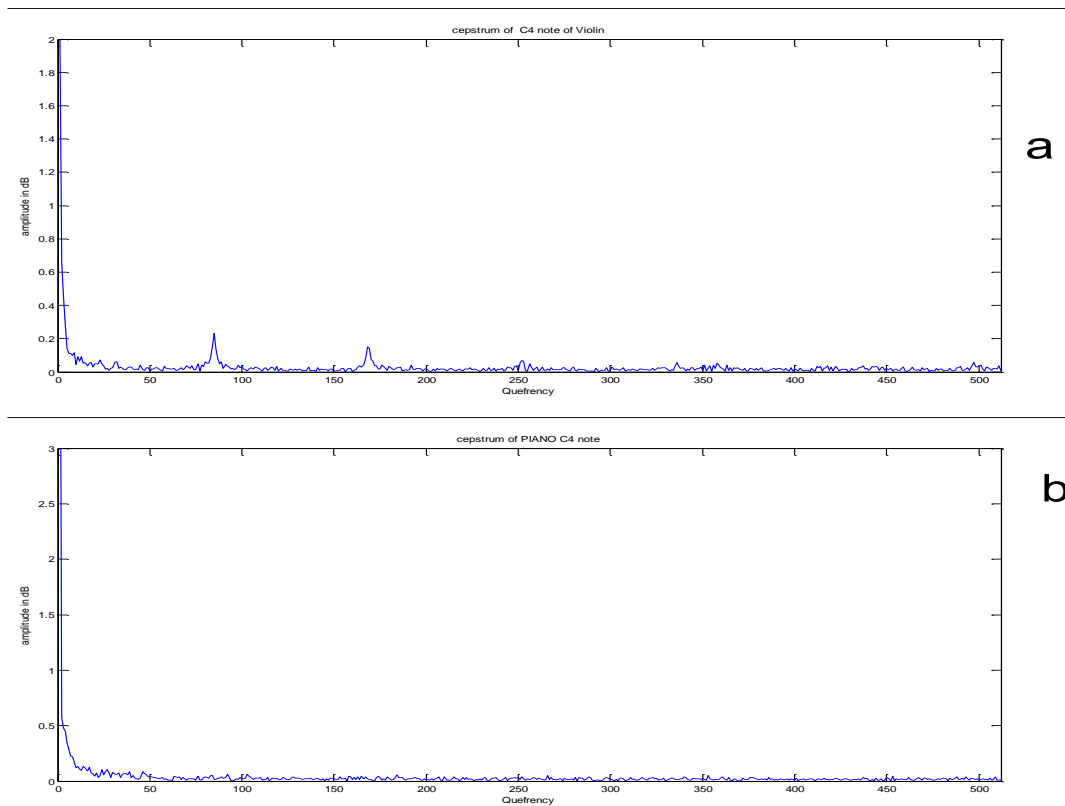
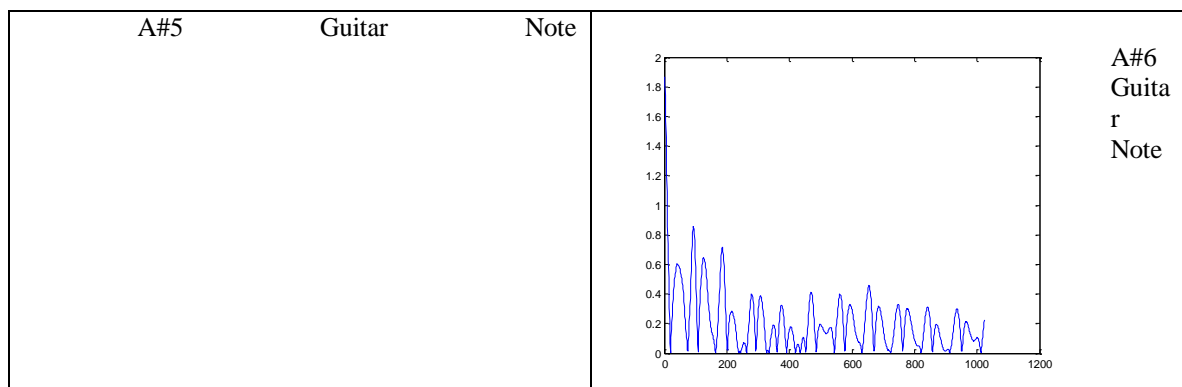


Fig. 3: Cepstrum of a) C4 note of Violin b) C4 note of Piano

Impulse Response or Body Response Modeling

The figure 4 shows impulse responses or body responses of instruments.



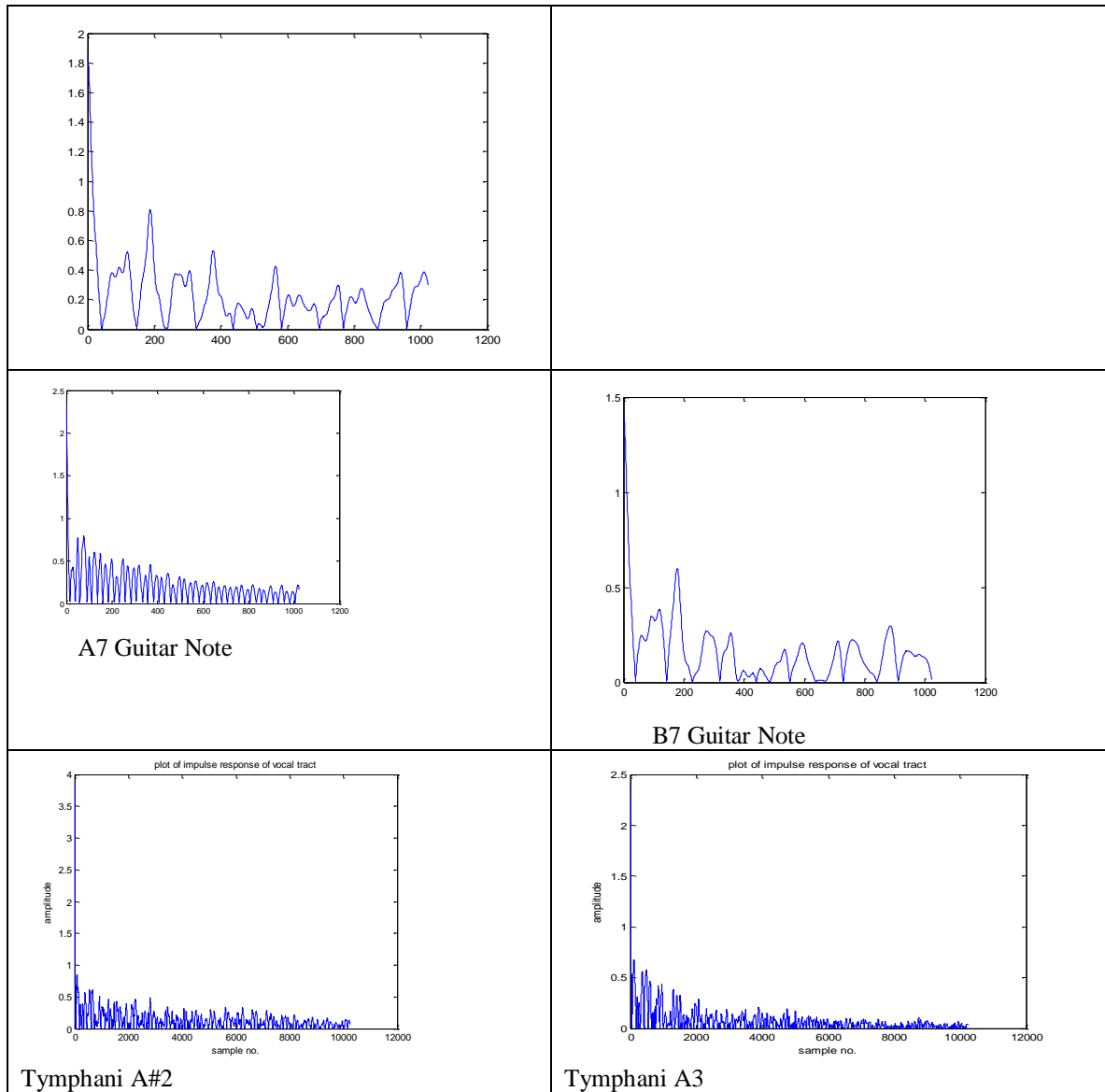


Fig. 4: The impulse responses or body responses for instruments.

Result:

The musical instruments are classified using features based on impulse response and WEKA classifier. The results in the form of confusion matrix are shown in table 1 and table 2 for individual instruments and family. Total 19 different musical instruments are used as shown in Table 1.

Table 1: Percentage accuracy in the form of Confusion matrix for individual instruments

| Instruments | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |
|------------------|----|-----|----|----|----|----|-----|----|-----|----|---|---|----|---|---|---|---|---|----|
| A=Saxophone | 90 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 10 |
| B=Bass | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C=Cello | 0 | 0 | 95 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 |
| D=Cornet | 0 | 0 | 5 | 80 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 5 | 0 | 0 |
| E=Eng Horn | 0 | 0 | 0 | 15 | 80 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 |
| F=F_Horn | 0 | 0 | 0 | 0 | 0 | 95 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 |
| G=Guitar | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 00 | 0 | 0 | 0 | 0 | 0 | 0 |
| H=Harpsichord | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 95 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 |
| I=Lute | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| J=Obeo classical | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 95 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| | | | | | | | | | | | | | | | | | | | |
|----------------|---|----|---|---|---|---|---|---|---|----|----|----|-----|----|----|-----|-----|----|----|
| K=Oboe d'Amore | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 85 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| L=Piano | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 85 | 0 | 0 | 0 | 0 | 15 | 0 | 0 |
| M=Drum | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 0 | 0 |
| N=Trombone | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 90 | 0 | 0 | 0 | 0 | 0 |
| O=Trumpet | 0 | 0 | 5 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 90 | 0 | 0 | 0 | 0 |
| P=Tuba | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 |
| Q=Tympani | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 |
| R=Viola | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 95 | 0 |
| S=Violin | 5 | 10 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 5 | 75 |
| Instruments | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S |

The percentage accuracy for instrument families are shown in table 2.

Table 2: Percentage accuracy in the form of Confusion Matrix using Impulse response

| | String | Wood wind | Brass | Percussion |
|------------|--------|-----------|-------|------------|
| String | 97.5 | 1.25 | 1.25 | 0 |
| Woodwind | 1.25 | 92.5 | 6.25 | 0 |
| Brass | 1 | 2 | 96 | 1 |
| Percussion | 0 | 0 | 0 | 100 |

Conclusion:

The Musical instrument classification for individual instruments and family is verified using impulse response modeling. It is found that the impulse response of instrument varies from instrument to instrument. It helps to easily distinguish between instrument to instrument and family to family. Impulse response is used to model the body response of the musical instruments. For individual instruments, the average classification accuracy has been obtained is 83.23% and 85.55% for family classification.

References:

1. J. Brown, "Cluster-based probability model for musical instrument identification," Journal of the Acoustical Society of America, vol. 101, pp. 3167, 1997.
2. J. Brown, "Musical instrument identification using autocorrelation coefficients", in Proceedings of the International Symposium on Musical Acoustics, 1998.
3. J. C. Brown, "Computer identification of musical instruments using pattern recognition with cepstral coefficients as features", The Journal of the Acoustical Society of America, vol. 105, no. 3, p. 1933, 1999.
4. Martin, K.D. and Kim, Y.E., "Musical Instrument Identification: A Pattern- Recognition Approach", 136th Meeting of the Acoustical Society of America, October, Norfolk, VA., 1998.
5. Eronen and A. Klapuri, "Musical instrument recognition using cepstral coefficients and temporal features", IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 2, 2000, pp. II753.
6. Eronen, "Comparison of features for musical instrument recognition", IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics, pp. 19-22, 2001.
7. G. Agostini, M. Longari, and E. Pollastri, "Musical instrument timbres classification with spectral features", IEEE Workshop on Multimedia Signal Processing, pp. 97, 2001.
8. J. C. Burges, "A tutorial on support vector machines for pattern recognition", Data Mining and Knowledge Discovery, vol. 2, no. 2, pp. 121-168, 1998.
9. J. Lee and J. Chun, "Musical instruments recognition using hidden Markov model", Asilomar Conference on Signals, Systems and Computers, vol. 1, p.196, 2002.
10. L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition", Proceedings of the IEEE, vol. 77, no. 2, pp. 257-286, 1989.
11. Eronen, "Musical instrument recognition using ICA-based transform of features and discriminatively trained HMMs", 7th International Symposium on Signal Processing and its Applications, vol. 2, p. 133, 2003.
12. McGill University Master Samples. [Online]. Available: <http://www.music.mcgill.ca/resources/mums/html/mums.html>
13. G. Costantini, A. Rizzi, and D. Casali, "Recognition of musical instruments by generalized min-max classifiers", in IEEE Workshop on Neural Networks for Signal Processing, pp. 555-564, 2003.
14. J. Marques and P. J. Moreno, "A study of musical instrument classification using Gaussian mixture models and support vector machines", Compaq Cambridge Research Laboratory, Tech. Rep., 1999.

15. S. Essid, G. Richard, and B. David, "Musical instrument recognition on solo performance", in Proc. Eur. Signal Processing Conf., pp. 1288–129, 2004.
16. S. Essid, "Efficient musical instrument recognition on solo performance music using basic features", in Proc. AES 25th Int. Conf., 2004.
17. M. Partridge and M. Jabri, "Robust principal component analysis", in IEEE Signal Processing Society Workshop, vol. 1, pp. 289–298, 2000.
18. A. Livshin and X. Rodet, "Instrument recognition beyond separate notes - indexing continuous recordings", in Proc. Int. Computer Music Conf., 2004.
19. A. Livshin and X. Rodet, "Musical instrument identification in continuous recordings", in Proc. of the 7th Int. Conference on Digital Effects, pp. 222 – 227, 2004.