IDENTIFICATION OF MUSICAL INSTRUMENT SOUNDS USING MULTILAYER PERCEPTRON

[1]SUSHEN R. GULHANE, [2]SURESH D. SHIRBAHADURKAR, [3]BHAGWAT D. JADHAV

[1]D. Y. Patil Institute of Technology, D. Y. Patil College of Engineering, Pune, India
[2]Zeal College of Engineering, Pune, India, [3]Rajarshi Shahu College of Engineering, Pune, India
E-mail: sushenrgulhane1@rediffmail.com, [2]s_shir00@yahoo.co.in, [3]bdjadhav1979@gmail.com

Abstract - In this paper, Timbral feature extraction of Musical Instrument sound is done. Seven timbral features are extracted using Music Information Retrieval (MIR) toolbox 1.6. & analyzed and found the result for Musical Instrument Identification. After extracting these timbral features, they are analyzed and using classifier Multilayer Perceptron, musical instruments sounds are identified.

Keywords - Spectral Skewness, Spectral Entropy, Spectral Centroid, Spectral Spread, Spectral Kurtosis, Spectral Regularity, Flatness

I. INTRODUCTION

Feature extraction is the process of extracting the features of the audio signal. In the last few decades, a number of features are developed for the analysis of audio sound signals. There are a lot of temporal, spectral and Cepstral features available for analysis of the sound signal. The temporal features are called as time domain features), Spectral Features are called as Frequency-domain features and Cepstral features are called Frequency domain features. One of the first application domains of audio analysis was speech recognition [14]. The analysis of music and general purpose environmental sounds gained importance in today day's novel application areas [28]. Different research fields evolved, such as audio segmentation, musical instrument identification and classification, music information retrieval (MIR), and environmental sound recognition (ESR). Each of these areas developed its specific description techniques (features). The no. of researchers are using time-domain and spectral-domain features for the classification the musical instruments.[6][18][21][26][27].A recent effort to standardize multimedia description tools led to the MPEG-7 standard. MPEG-7 is an ISO/IEC standard for multimedia content description [13]. The standard defines low-level descriptions techniques (including audio) as well as high-level tools for multimedia processing [15]. The numbers of features are extracted from the musical instrument sound signals. These features are compared and find the appropriate features required for the classification of musical instrument.[28] The huge number of existing features makes the selection of the most appropriate feature set for a task difficult [15]. The Statistical analysis of the sound signals will be helpful for the identification of independent features. In this paper, Timbral i.e. spectral feature extraction of Musical Instrument sound is done. The no. of timbral features are extracted using Music Information Retrieval (MIR) toolbox 1.6. After extracting the features, they are analyzed and found seven appropriate timbral features required for Musical Instrument Identification. For database generation of Musical Instrument Sounds, 8 Musical Instruments are considered. Out of which, 2 Indian Woodwind Musical Instruments, 2 Western Woodwind Musical Instruments, 2 Indian String Musical Instruments, and 2 Western String Musical Instruments and each have 10 samples. So in all eighty samples database is considered. A quantitative analysis of timbral features of musical instrument sound signals has been performed. Different description techniques have been investigated by statistical data analysis in order to identify similarities and redundancies [15]. The objective of feature extraction is to find the appropriate features i.e. prominent features required for musical instrument identification [28].

II. THEORY

Basically, there are three types of features which can be considered for the identification of musical instrument sounds. They are Time-Domain Features (Temporal Features), Spectral Features (Timbral Features) and Cepstral Features [28]. These are discussed as follows:

2.1 Time- Domain Features (Temporal Features)

The time- domain features are the features related to the shape of the envelope of a node. They are measured as attack time(A), decay time(D), sustain time(S) and release time(R) as shown below; The terms related to the shape of an envelope will be discussed as; the instant of the time at which note starts is called onset time. When the note starts, the amplitude of the envelope will continue to rise for some time and then it will start decreasing. One has to find the time difference between the start of a note (onset time) and the time when the envelope goes to its maximum value; called the end of attack time. The difference will be then called attack time. The

amplitude of the envelope starts decreasing from the end –of- attack as the time proceed in the forward direction. One has to identify the time instant when the Envelope amplitude drops to 25% of its maximum value at the end of attack instant. The time domain envelops can be described by the four-time values and roll-off rate. They are Attack time, Decay time, Sustain time, Release time [28].

2.2 Spectral Features (Timbral Features)

Timber is nothing but the colour of sound or tone. Features related to it called as spectral features or timbral features. There are a lot of spectral features available to extract. Some of them are explored as follows:

2.2.1 Centroid: It represented the location of gravity i.e. magnitude spectrum.

2.2.2 Spread: It describes the average deviation of the rate map around its Centroid which is commonly associated with the bandwidth of the signal.

2.2.3 Skewness: It measures the symmetry of the spectrum around its arithmetic mean. The feature will be zero for silent segment and high for voiced speech where substantial energy is present around the fundamental frequency.

2.2.4 Kurtosis: It is a statistical measure that defines how heavily the tails of a distribution differ from the tails of a normal distribution.

2.2.5 Flatness: It is ratio of geometric mean to arithmetic mean which can be used to distinguish between harmonic (close to zero) and noisy signal (close to one).

2.2.6 Regularity: It is factor distinct from harmonic relations in auditory grouping [29].

2.2.7 Entropy: is a measure of its spectral power distribution. The concept is based on the Shannon entropy, or information entropy, in information theory.

2.3 Cepstral Features

There are 39 Cepstral coefficients out of which 12 are Cepstral coefficients, 12 are Delta Cepstral coefficients, 12 are Double Delta Cepstral Coefficients, 1 is energy coefficient, 1 is delta energy coefficient and 1 is double delta coefficient. 12 Cepstral coefficients + 1 Energy coefficient (13 coefficients) represents the information about the vocal tract filter. So only 13 coefficients are enough.

Block Diagram: Following figure shows the block diagram of feature extraction.

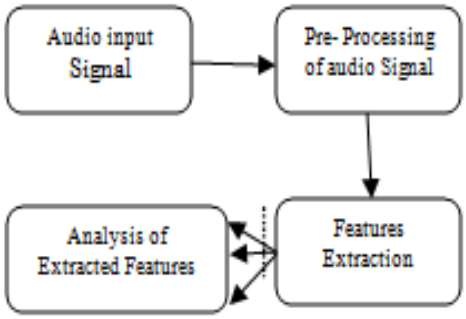


Fig.1 Block diagram of feature extraction

There are no. of audio signal file formats. In any one formats, one record the audio signal & will be fed to the pre-processing stage. The pre-processing stage will pre-process the audio input signal & fed to the feature extraction block. By using different feature extraction techniques, features are extracted and the results are compared.

III. EXPERIMENT RESULTS

For database generation of Musical Instrument Sounds, 8 Musical Instruments are considered. Out of which, 2 Indian Woodwind Musical Instruments, 2 Western Woodwind Musical Instruments, 2 Indian String Musical Instruments and 2 Western String Musical Instruments and each have 10 samples. So in all 80 samples database is created. The numbers of experiments are performed for feature extraction of audio signals. For experimentation, audio input is taken. In proposed system, audio features are extracted from the sound samples by using;

A. Timbral feature extractors

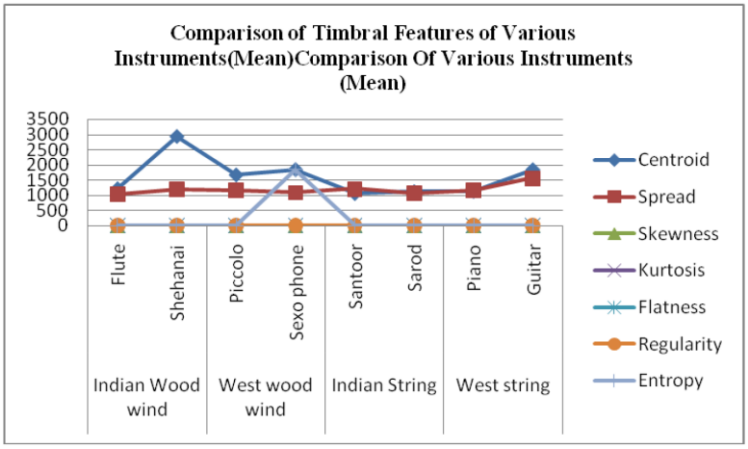
The experiments are done to extract the timbral features. The results are shown below:

Type of Instruments	*Instrument*	*Centroid*	Spread	Skewness	Kurtosis	Flatness	Regularity	Entropy
Indian Wood wind	Flute	1213.547	1042.74	1.66	5.46	0.186667	0.48	0.693333
	Shehanai	2942.213	1202.12	0.086667	2.046667	0.12	0.733333	0.786667
West wood wind	Piccolo	1669.86	1175.427	1.36	4.073333	0.2	1.566667	0.653333
	Sexo phone	1848.313	1103.913	0.84	3.32	0.106667	0.666667	1848.313

Indian String	Santoor	1057.167	1218.74	2.02	6.14	0.206667	0.64	0.8
	Sarod	1118.047	1078.193	1.82	6.146667	0.226667	1.026667	0.833333
West string	Piano	1129.533	1180.467	1.613333	4.893333	0.26	0.706667	0.866667
	Guitar	1842.167	1571.467	0.586667	2.033333	0.4	0.693333	0.9

Table 1: Comparison of Timbral Features of Various Instruments(Mean)Comparison Of Various Instruments (Mean)

Table 1 shows the comparison of different `timbral features of Various Musical Instruments. The detailed feature extraction of eight instruments is shown below:



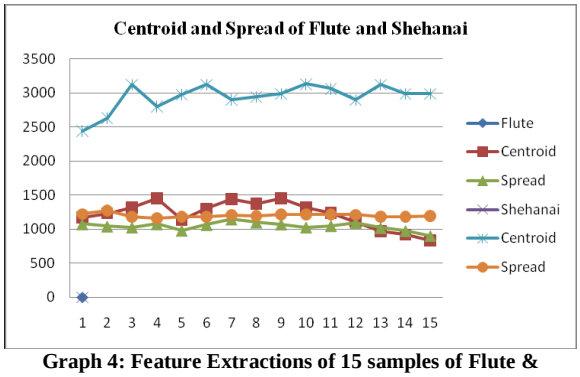
Graph 3: Comparison of Timbral Features of Various Instruments (Mean)

1. Indian Woodwind Musical Instruments

Here Flute and Shehanai instruments under Indian Woodwind Musical Instruments are considered.

Flute	Centroid	Spread	Shehanai	Centroid	Spread
Flute	1164.7	1074.5	Shehanai	2440.6	1222.6
	1229	1041.7		2635.3	1272.1
	1315.6	1026.9		3124.6	1185.2
	1449.4	1076.5		2800.7	1161.4
	1130.4	979.4		2978	1183.4
	1297.7	1064		3121.5	1181.9
	1437.1	1145.1		2900.8	1211.1
	1367.1	1099.5		2944.8	1195.2
	1443.9	1071.4		2989.3	1216.7
	1318	1027.6		3131.2	1212.3
	1233.6	1048.3		3063.1	1214.5
	1104.6	1093.2		2900	1211.3
	966.6	1017.8		3124.6	1185.2
	916.4	974.1		2988.1	1188
	829.1	901.1		2990.6	1190.9

Table 2: Spread & Centroid Features extraction for 15 samples



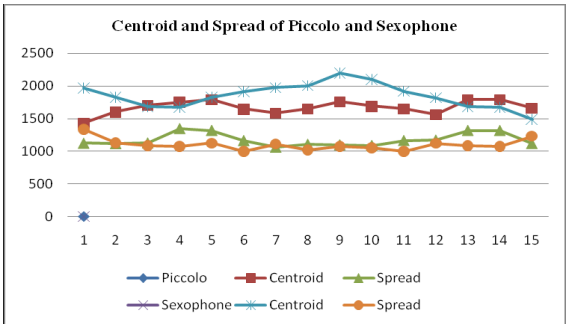
Graph 4: Feature Extractions of 15 samples of Flute & Shehanai

Graph 4 is drawn from experimental results shown in Table 2. This graph shows the differentiation of Flute & Shehanai using Spread and Centroid features.

2. Western Woodwind Instruments

Piccolo	Centroid	Spread	Sexophone	Centroid	Spread
Piccolo	1429.9	1122.3	Sexophone	1965.8	1338.1
	1600.6	1116.1		1824.8	1127.5
	1706	1126.5		1677.7	1087
	1751.7	1344.5		1671	1079.5
	1790.2	1317.2		1824.5	1126.1
	1645.1	1163.6		1912.2	1000.1
	1585.1	1061.4		1973.3	1114.4
	1645.5	1106.8		2005.8	1025.1
	1761.8	1098.2		2198.4	1077.7
	1689.3	1088.2		2100.2	1055.8
	1646.3	1163.9		1917.9	1002.5
	1559.7	1170.8		1818.3	1125.3
	1788.6	1316.7		1677.7	1087
	1789.4	1316.6		1673.5	1078.4
	1658.7	1118.6		1483.6	1234.2

Table 3: Spread & Centroid Features extraction for 15 samples



Graph 5:Feature Extractions of 15 samples of Piccolo & Sexophone

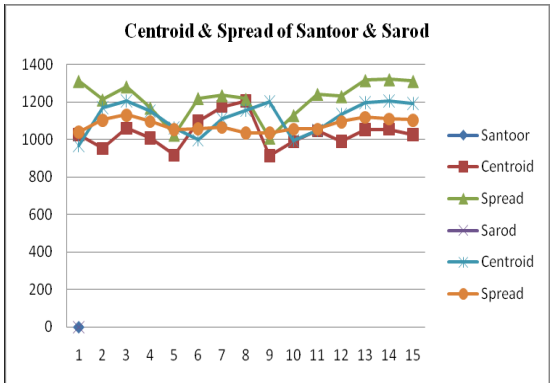
Here, Piccolo and Sexophone under Western Woodwind Instruments are considered for comparison purpose.

3. Indian String Musical Instruments

Here, Santoor and Sarod under Indian String Musical Instruments are considered. Graph 6 is drawn from Table 4.

Santoor	Centroid	Spread	Sarod	Centroid	Spread
	1025.5	1311.1		966.8	1040.3
	951	1211.5		1169.1	1105.1
	1059.3	1281.5		1204	1132
	1006.5	1167.4		1148.4	1097.4
	916	1023.1		1061.2	1055
	1096.2	1219.2		996.1	1059.3
	1176.7	1234.4		1110.9	1064.9
Santoor	1206.5	1215	Sarod	1154	1037
	911.7	1007		1200.9	1035.3
	987.4	1127.9		993.4	1056.9
	1046.6	1241.5		1047.9	1057.3
	990.3	1229.6		1131.4	1095.7
	1052.6	1315.3		1193.6	1120.3
	1054.3	1320		1201.9	1111.5
	1025.5	1311.1		1191.1	1104.9

Table 4: Spread & Centroid Features extraction for 15 samples



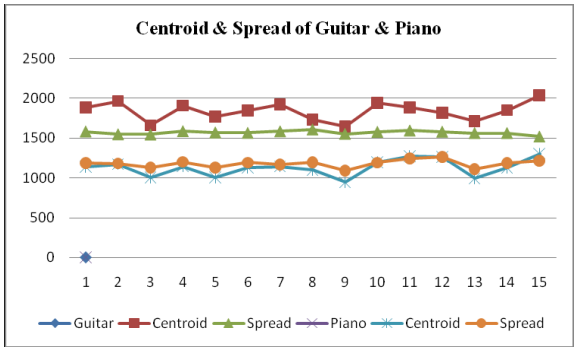
Graph 6: Feature Extractions of 15 samples of Santoor & Sarod

4. Western String Musical Instruments

Here, Guitar and Piano under Westeren String Musical Instruments are considered for comparison purpose.

Guitar	Centroid	Spread	Piano	Centroid	Spread
	1888.6	1585.2		1136.9	1187.2
	1967.8	1549.7		1170.7	1182.3
	1663.8	1546.6		1005	1132.1
	1911.6	1589		1143	1202.1
	1773.6	1572.4		1005.2	1129
	1852.1	1567.7		1131.3	1194.7
	1924.4	1585.8		1142.5	1171.1
Guitar	1736.4	1609.9	Piano	1099	1199
	1648.1	1555.4		944.4	1089.8
	1946.8	1580.2		1197.6	1194.8
	1888.7	1596.5		1274	1243.3
	1822.8	1582		1264.7	1267.2
	1713.4	1564.9		991.7	1109.7
	1852.8	1565.4		1131.8	1188.5
	2041.6	1521.3		1305.2	1216.2

Here Graph 7 is drawn from Table 5.



Graph 7: Feature Extractions of 15 samples of Guitar & Piano

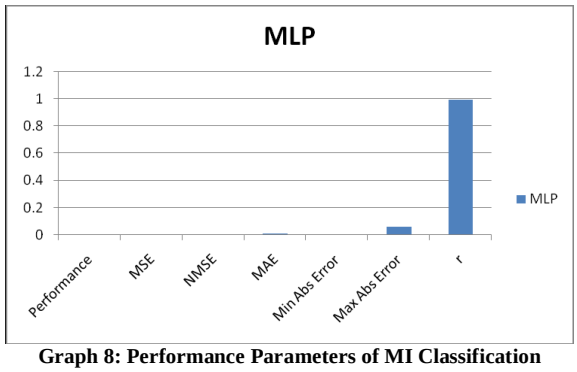
From Table 2 & Graph 3, it is concluded that only Spread and Centroid features are useful for identification of Musical Instruments sound.

B. Classification of Musical Instrument Using MLP

MLP is a Multilayer Perceptron classifier using for classification of various musical instrument sounds. For classification, following performance parameters are used. They are MSE, NMSE, MAE, Min Abs error, Max Abs error,r and Percentage Accuracy. The available database is divided into three section as 60% database is used for training, 25% database is used for testing and training and 15% database is used for cross validation. The results are shown below;

Classifiers	MLP
Performance	
MSE	0.000484829
NMSE	0.00498667
MAE	0.012844864
Min Abs Error	0.000198727
Max Abs Error	0.063507315
r	0.997796717
Percent Correct	100

Table 6: Classification of MI using MLP



Graph 8: Performance Parameters of MI Classification With multilayer perceptron for 8 instruments, accuracy is reached to 100%.

IV. CONCLUSIONS

The no. of experiments are performed to extract Timbral features which are required to identify the musical instruments. Table 1 show that only Centroid

and Spread features distinguishing the instruments. From the experimental results, it has been concluded that Timbral Features' Spread and Centroid can be enough to identify the musical instruments. Using Multilayer Perceptron, 100% accuracy is getting for 8 instruments.

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references)

[2] D. Giannoulis and A. Klapuri, "Musical Instrument Recognition in Polyphonic Audio Using Missing Feature Approach," IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, Vols. VOL. 21, NO. 9,, SEPTEMBER SEPTEMBER 2013.

[3] S. Essid and G. Richard, "Musical Instrument Recognition by Pairwise Classification Strategies," IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, Vols. VOL. 14, NO. 4, JULY 2006.

[4] T. H. Park, "Towards Automatic Musical Instrument Timbre Recognition," Princeton University, New Jersey,USA, November 2004.

[5] O. Lartillot and P. Toiviainen, "A Matlab Toolbox for Musical Feature Extraction from Audio," in International Conference on Digital Audio Effects, Bordeaux, France, 2007.

[6] S. H. Deshmukh and S. G. Bhirud, "North Indian Classical Music's Singer Identification by Timbre Recognition using MIR Toolbox, "International Journal of Computer Applications, vol. 91 – No.4, p. 0975 – 8887, 2014.

[7] Wieczorkowska, Z. W. Ra´s and R. Lewis, "Multi-way Hierarchic Classification of Musical Instrument Sounds"..

[8] S. Essid, G. Richard and. B. David, "Musical Instrument Recognition by Pairwise Classification Strategies," IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, Vols. VOL. 14, NO. 4, JULY 2006.

[9] S. Essid, G. Richard and B. David, "Efficient musical instrument recognition on solo performance music using basic features," in AES 25th International Conference, London, U.K., 2004.

[10] Eronen, "COMPARISON OF FEATURES FOR MUSICAL INSTRUMENT RECOGNITION," Tampere, Finland.

[11] D. M. Chandwadkar and M. . S. Sutaone, "Selecting Proper Features and Classifiers for Accurate Identification of Musical Instruments," International Journal of Machine Learning and Computing, vol. 3, no. 2, April 2013.

[12] P. Herrera, G. Peeters and . S. Dubnov, "Automatic Classification of Musical Instrument Sounds," Journal of New Music Research, vol. 32, 2003.

[13] Glowacz, W. Glowacz and A. Glowacz, "Sound Recognition of Musical Instruments with Application of FFT and K NN Classifier with Cosine Distance".

[14] ISO/IEC 15938 (2002). Information Technology – Multimedia Content Description Interface. First Edition.

[15] Rabiner, L. & Juang, B. (1993). Fundamentals of speech recognition. New York: Prentice-Hall.

[16] Dalibor Mitrovic,Matthias Zeppelzauer, and Horst Eidenberger," Analysis of the Data Quality of Audio Descriptions of Environmental Sounds", Journal of Universal Knowledge Management, vol. 1, no. 1 (2006).

[17] P. Shinde, . V. Javeri and . O. Kulkarni, "Musical Instrument Classification using Fractional Fourier Transform and KNN Classifier," International Journal of Science, Engineering and Technology Research (IJSETR), vol. 3, no. 5, May 2014.

[18] G. Agostini, M. Longari and E. Pollastri, "Musical instrument timbres classification with spectral features," in Proc. International Workshop Multimedia Signal Processing, Cannes, France, Oct. 2001.

[19] K. Jensen and J. Amspang, "Binary Decision Tree Classification of Musical Sound," in ICMC Proceedings, 1999.

[20] G. Mazarakis, . P. Tzevelekos and . G. Kouroupetr, "Musical Instrument Recognition and Classification Using Time Encoded Signal Processing and Fast Artificial Neural Networks".

[21] H. Fletcher, "Loudness, Pitch and Timber of Musical Tones and their Relations to the Intensity, the Frequency and the Overtone Structure," JASA, vol. 6/2,59-69.

[22] S. K. Banchhor and A. Khan, "Musical Instrument Recognition using Spectrogram and Autocorrelation," International Journal of Soft Computing and Engineering (IJSCE),ISSN: 2231-2307, vol. 2, no. 1, March 2012.

[23] Sarfaraz Masood ; Shubham Gupta ; Shadab Khan,"Novel approach for musical instrument identification using neural network", 2015 Annual IEEE India Conference (INDICON), ISSN: 2325-9418, March 2016.

[24] Mohammad Adeli ; Jean Rouat ; Sean Wood ; Stéphane Molotchnikoff ; Eric Plourde,"A Flexible Bio-Inspired Hierarchical Model for Analyzing Musical Timbre",IEEE/ACM Transactions on Audio, Speech, and Language Processing, ISSN: 2329-9304,Vol:24,Issue:5,May 2016.

[25] Yoonchang Han ; Jaehun Kim ; Kyogu Lee ,"Deep Convolutional Neural Networks for Predominant Instrument Recognition in Polyphonic Music", IEEE/ACM Transactions on Audio, Speech, and Language Processing, ISSN : 2329-9304,Volume: 25, Issue:1,Jan.2017.

[26] Shubham Bahre ; Shrinivas P. Mahajan ; Rohan T. Pillai,"Novel audio feature set for monophonie musical instrument classification", 2017 International Conference on Recent Innovations in Signal processing and Embedded Systems (RISE), ISBN:978-1-5090-4760-4, June 2018.

[27] Monica S. Nagawade, Varsha R. Ratnaparkhe, "Musical instrument identification using MFCC", 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT),ISBN: 978-1-5090-3704-9, January 2018.

[28] Sushen R. Gulhane, Sanjay S. Badhe, Dr. Suresh D. Shirbahadurkar," Cepstral__MFCC__Feature_and_Spectral__Timbral__Features_Analysis_for_Musical_Instrument_Sounds",2018 2nd IEEE Global Confernece on Wireless Computing and Networking(GCWCN-2018),ISBN-978-1-5386-5201-5, 23-24 Nov 2018.

[29] Roberts B1, Bailey PJ., "Spectral regularity as a factor distinct from harmonic relations in auditory grouping", J Exp Psychol Hum Percept Perform. 1996 Jun;22(3):604-14.

★★★