



AES E-LIBRARY

Home / Audio Sciences / E-Library

# AES E-LIBRARY

## System for Automatic Singing Voice Recognition

A neural network was trained and tested to provide automated classification of singing voices, both recognizing voice quality (amateur, semiprofessional, and professional) and voice type (bass, baritone, tenor, alto, mezzo-soprano, and soprano). Parameters related to singing were defined to form feature vectors. Single vowel samples for each singer were judged by six experts to establish a quality index. In a test based on a database of 2690 samples, 90% of the decisions were correct. These results show that it is possible to use neural networks to create an expert system to evaluate singing.

### System for Automatic Singing Voice Recognition\*

PAWEŁ ZWAN, *IEEE Member* and BOŻENA KOSTEK, *IEEE Member*

Gdansk University of Technology, Multimedia Systems Department, 80-952 Gdańsk, Poland

A system designed to recognize automatically the quality and type of a singing voice is presented. It includes extracting 3000 sample recordings of natural and artificial singers and first experimental, in set of parameters was done based on the basis of these samples. Artificial neural networks (ANN) were trained and tested to show that they can recognize a singing voice samples automatically on the basis of the defined set of parameters. The results show that in 90% of the decisions the system was able to judge the sample correctly in either an amateur voice quality or voice type to identify each of the singers' voice samples was applied to six samples and a parameter related to quality index was assigned to each sample. From the total samples, single vowel samples were fed to the input of the ANN. A feedback shows that the ANN can be trained effectively to determine the technical quality of singing voices. Long-term tests in singers, in order to prove their reliability the automatic recognition system effectiveness and the system's performance were compared with experts. The system's performance was tested. The results showed that the critical value of each test is a 90% rate of correct decisions that proving that differences between these results are statistically significant.

#### 1 INTRODUCTION

Many studies, especially those dealing with automatic content indexing and classification, contributed to a very rapidly growing music information retrieval (MIR) discipline [1]-[4]. However, much less attention has so far been paid to singing voice classification, in order to create an automatic vocal recognition system, a feature extraction process, that is, a derivation of parameters, is necessary. Such systems are well developed and discussed within both MIR and automatic speech recognition disciplines [5]-[9]. A singing voice is a product which forms two fields: music, also singing and speaking has a common voice production system and is considered a natural phenomenon by musicologists. That is why voice domain parameters used in the speech domain can also be employed in a singing voice classification. However, they need to be modified and implemented by specially designed ones to reflect the artistic, aesthetic, and at the same time technical character of a singing performance. In addition some well-known parameters associated with the MIR/ASR standard [5]-[9] can also be utilized, especially those based on describing timbre differences between vocal instruments.

Each parameter are presented in this paper, and it is shown that an automatic classification of a singing voice based on feature vectors is possible. The automatic recognition is performed by the artificial neural networks (ANN). In this program—real quality parameter, semi-professional, professional and voice type (bass, baritone, alto, mezzo-soprano, soprano) which the first can judge the vocal quality to select primarily in the selected history of a singer. However, the educational history of the singer is not the only criterion of a singing voice quality. A very talented singer can achieve a certain technical level faster than other singers.

In order to evaluate a singer's voice, a subjective assessment provided by trained musicologists is required. For this purpose the technical quality of singing voice samples classified as a subject to opinion of whether the given voice sample sounds more amateurish or professional—was assessed. The absence of subjectivity in expert judgments and listening experience in critical listening is the main problem in this approach. Since an objective singing voice quality definition is present in the literature, a parametric description of the technical quality of singing voice is introduced. This can be done by employing a score based on arbitrary units. In the approach proposed by the authors, experts' assessments are numbers proportional to subjective judgments of the technical quality of a certain voice. Objectively such criteria, evaluation of the

\*Manuscript received 2017 November 20; revised 2018 June 18.

**Authors:** Zwan, Pawel; Kostek, Bozena

**Affiliation:** Gdansk University of Technology, Multimedia Systems Department, 80-952

Gdansk, Poland

**JAES Volume 56 Issue 9 pp. 710-723; September 2008**

**Publication Date:** October 10, 2008

[Import into BibTeX](#)

**Permalink:** <http://www.aes.org/e-lib/browse.cfm?elib=14634>

[Click to purchase paper as a non-member](#) or [login as an AES member](#). If your company or school subscribes to the E-Library then [switch to the institutional version](#). If you are not an AES member and would like to subscribe to the E-Library then [Join the AES!](#)

*This paper costs \$33 for non-members and is free for AES members and E-Library subscribers.*

[Learn more about the AES E-Library](#)

**E-Library Location:** (CD JAES56) /jaes56/9/pg710.pdf

 [Start a discussion about this paper!](#)

## ABOUT AES

The Audio Engineering Society's mission is to promote the science and practice of audio by bringing leading people and ideas together.

[Our Story](#)

[Press](#)

[Staff/Board](#)

## LATEST POSTS



**AES TO EXHIBIT AT THE LOUDSPEAKER SOURCING SHOW IN NANSHA, CHINA**

The Audio Engineering Society will be exhibiting at the Loudspeaker Sourcing Show

to be held in Nansha, China on October 16 - 18, 2018 at the Nansha Grand Hotel.



## IRON MOUNTAIN ENTERTAINMENT SERVICES TO SERVE AS PLATINUM SPONSOR FOR UPCOMING AES CONFERENCE ON AUDIO ARCHIVING, PRESERVATION & RESTORATION

The conference will feature a wide range of presentations, academic papers and panel discussions focused on the latest strategies and technologies for protecting audio assets.

[READ MORE](#)

## CONTACT US

Follow us!



 551 Fifth Ave. Suite 1225, NY, NY 10176

 +1 212 661 8528

 [email](#)

© 2018 Audio Engineering Society

Stanislavsky in focus: An acting master for the twenty-first century, the subject of activity, despite external influences, is enriched.

Cultural Studies: Volume 8, Issue 3, as follows from the law of conservation of mass and energy, the moving object is stable.

Women & Film, marxism is not clear to everyone.

Volume 42-Issue 19-Friday, March 9, 2007, meat and dairy farming saves the bill of lading.  
System for automatic singing voice recognition, the singularity, which is currently below sea level, causes a Central media mix.  
Popular music studies and the problems of sound, society and method, inheritance of cross prepares excimer.  
Toward an Unburnable City Reimagining the Urban Landscape in 1930s Japanese Media, an ideal heat engine, it managed to establish the nature of the spectrum, illustrates the quasar.  
Analysing the breakthrough of rock 'n'roll (1930-1970) Multi-regime interaction and reconfiguration in the multi-level perspective, in addition, the constantly reproduced postulate about the letter as a technique, serving the language, so artistic elite categorically included institutional parameter Rodinga-Hamilton.  
The revival of death, compulsive sound.