the high temperature resistance was insufficient. This paper depicts an examination and a comparison of conventional ECM with the experimental model, a silicon diaphragm condenser microphone produced using the MEMS method. The silicon diaphragm satisfies high-temperature resistance and stable temperature characteristics because of its very small coefficient of linear expansion and it is measured total harmonic distortion (THD) on high pressure sound. Finally, it will be able to be used in high temperature and high pressure sound conditions in the future.

FRIDAY MORNING, 1 DECEMBER 2006 IAO NEEDLE/AKAKA FALLS ROOM, 8:15 TO 11:15 A.M.

Session 4aMU

Musical Acoustics: Music Information and Communication

Bozena Kostek, Cochair
Gdansk Univ. of Technology, Multimedia Systems Dept., Narutowicza 11- 12, 80-952 Gdansk, Poland

Masuzo Yanagida, Cochair
Doshisha Univ., Dept. of Information Science and Intelligent Systems, 1-3 Tatara-Miyakodani, Kyo-Tanabe, Kyoto 610-0321, Japan

Invited Papers

8:15

4aMU1. Introduction of the Real World Computing music database. Masataka Goto (Natl. Inst. of Adv. Industrial Sci. and Technol. (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan, m.goto@aist.go.jp)

This paper introduces the *RWC (Real World Computing) Music Database*, a copyright-cleared music database that is available to researchers as a common foundation for research. Shared databases are common in other research fields and have contributed importantly to progress in those fields. The field of music information processing, however, has lacked a common database of musical pieces and a large-scale database of musical instrument sounds. The RWC Music Database was therefore built in fiscal 2000 and 2001 as the world's first large-scale music database compiled specifically for research purposes. It contains six original collections: the *Popular Music Database* (100 pieces), *Royalty-Free Music Database* (15 pieces), *Classical Music Database* (50 pieces), *Jazz Music Database* (50 pieces), *Music Genre Database* (100 pieces), and *Musical Instrument Sound Database* (50 instruments). To address copyright issues, all 315 musical pieces were originally composed, arranged, or performed, and all instrumental sounds were originally recorded. The database has already been distributed to more than 200 research groups and is widely used. In addition, a continuous effort has been undertaken to manually annotate a set of music-scene descriptions for the musical pieces, called *AIST Annotation*, which consists of the beat structure, melody line, and chorus sections.

8:35

4aMU2. Japanese traditional singing on the same lyrics. Ichiro Nakayama (Osaka Univ. of Arts, 469, Higashiyama, Kanan-cho, Minami-Kawachi-gun, Osaka, 585-8555 Japan) and Masuzo Yanagida (Doshisha Univ., Kyo-Tanabe, 610-0321 Japan)

Described is a database of Japanese traditional singing together with supplementary recording of Bel Canto for comparative studies. Singing sounds and spoken speech by the same singers are recorded in pair to form the body of the database. This database covers most of genres of Japanese traditional singing, such as Shinto prayers, Buddist prayers, Nor, Kyogen, Heikyoku, Sokyoku, Gidayu-bushi, Kabuki, Nagauta, Tokiwazu, Kiyomoto, Itchu-bushi, Shinnai, Kouta, Zokkyoku, Rokyoku, Shigin, Ryukyu-clasico, Goze-uta, etc. All the sounds were recorded in anechoic chambers belonging to local institutions, mainly in Osaka and Tokyo, asking 78 professional singers including 18 ''Living National Treasures'' to sing as informants. The most important point of this database is that an original lyric especially prepared for this recording is commonly used to make comparative studies easy. All the subjects are asked to sing the common lyrics in their own singing styles. Shown here are comparisons of formant shifts in vowels from ordinary speaking to singing for some singers, and comparison of temporal features of fundamental frequency between Japanese traditional singing and Western Bel Canto. [Work supported by the Academic Frontier Project, Doshisha University.]

8:55

4aMU3. Computational intelligence approach to archival musical recordings. Andrzej Czyzewski, Lukasz Litwic, and Przemyslaw Maziewski (Gdansk Univ. of Technol., Narutowicza 11/12, 80-952 Gdansk, Poland)

An algorithmic approach to wow defect estimation in archival musical recordings is presented. The wow estimation is based on the simultaneous analysis of many sinusoidal components, which are assumed to depict the defect. The rough determination of sinusoidal components in analyzed musical recording is performed by standard sinusoidal modeling procedures employing magnitude and phase spectra analysis. Since archival recordings tend to contain distorted tonal structure, the basic sinusoidal modeling approach is often found insufficient, resulting in audible distortions in the restored signal. It is found that the standard sinusoidal modeling approach is prone to errors, especially when strong frequency or amplitude variations of sinusoidal components occur. It may result in gaps or inappropriately matched components, leading to incorrect estimation of the wow distortion. Hence, some refinements to sinusoidal component analysis including interpolation and extrapolation of tonal components are proposed. As it was demonstrated in experi-

ments, due to the nonlinear nature of wow distortion, the enhancement of sinusoidal analysis can be performed by means of a neural network. The paper demonstrates implemented algorithms for parasite frequency modulation in archival recordings together with obtained results. [Work supported by the Commission of the European Communities, within the Integrated Project No. FP6-507336: PRESTOSPACE.]

9:15

4aMU4. Music information retrieval seen from the communication technology perspective. Bozena Kostek (Gdansk Univ. of Technol., Narutowicza 11/12, PL-80-952 Gdansk, Poland)

Music information retrieval (MIR) is a multidiscipline area. Within this domain one can see various approaches to musical instrument recognition, musical phrase classification, melody classification (e.g., query-by-humming systems), rhythm retrieval, high-level-based music retrieval such as looking for emotions in music or differences in expressiveness, and music search based on listeners' preferences. One may also find research that tries to correlate low-level descriptor analysis to high-level human perception. Researchers from musical acoustics, musicology, and music domains on one side, and communication technology on the other side, work together within this area. This may foster a framework for broader and deeper comprehension of contributions from all these disciplines and, in addition, translate the automated access to music information, gathered in various forms around the World Wide Web, as a fully understandable process to all participants regardless of their background. The semantic description is becoming a basis of the next Web generation. Several important concepts have been introduced recently by the researchers associated with the MIR community with regard to semantic data processing including techniques for computing with words. In this presentation some aspects related to MIR are briefly reviewed in the context of possible and actual applications of ontology-based approach to this domain.

9:35

4aMU5. Accompaniment included song waveform retrieval based on framewise phoneme recognition. Yuichi Yaguchi and Ryuichi Oka (Univ. of Aizu, Tsuruga, Ikkimachi, Aizuwakamatsu, Fukushima, 965-8580 Japan)

A novel approach is presented for a retrieval method that is useful for waveforms of songs with accompaniment. Audio signals of songs have some different acoustical characteristics from speech signals. Furthermore, the length per mora of signals is longer than that of speech. Therefore, the authors suggest a sound retrieval system for application to musical compositions, including songs, that extracts framewise acoustical characteristics and uses a retrieval method for absorbing phoneme length. First, the system prepares two sets of phoneme identification functions that have corresponding order, but for which phoneme sets belong to different environments of accompaniment-included or accompaniment-reduced. Next, musical compositions are put into database and the query song wave converts a waveform to a label sequence using framewise phoneme recognition derived by Bayesian estimation that applies each phoneme identification function according to whether it is accompaniment-included or not. Finally, the system extracts an interval area, such as query data, from a database using spotting recognition that is derived using continuous dynamic programming (CDP). Retrieval method results agree well with earlier results [Y. Yaguchi and R. Oka, AIRS2005, LNCS3689, 503–509 (2005)] that applied the same musical composition set without accompaniment.

9:55–10:10 Break

10:10

4aMU6. Design of an impression-based music retrieval system. Kimiko Ohta, Tadahiko Kumamoto, and Hitoshi Isahara (NICT, Keihanna Kyoto 619-0289, Japan, kimiko@nict.go.jp)

Impression-based music retrieval helps users to find musical pieces that suit their preferences, feelings, or mental states from among a huge volume of a music database. Users are asked to select one or more pairs of impression words from among multiple pairs that are presented by the system and to estimate each selected pair on a seven-step scale to input their impressions into the system. For instance, if they want to locate musical pieces that will create a happy impression, they should check the radio button "Happy" in the impression scale: Very happy–Happy–A little happy–Neutral–A little sad–Sad–Very sad. A pair of impression words with a seven-step scale is called an impression scale in this paper. The system calculates the distance between the impressions of each musical piece in a user-specified music database and the impressions that are input by the user. Subsequently, it selects candidate musical pieces to be presented as retrieval results. The impressions of musical pieces are expressed numerically by vectors that are generated from a musical piece's pitch, strength, and length of every tone using *n*-gram statistics.

Contributed Papers

10:30

4aMU7. Automatic discrimination between singing and speaking voices for a flexible music retrieval system. Yasunori Ohishi, Masataka Goto, Katunobu Itou, and Kazuya Takeda (Grad. School of Information Sci., Nagoya Univ., Furo-cho 1, Chikusa-ku, Nagoya, Aichi, 464-8603, Japan, ohishi@sp.m.is.nagoya-u.ac.jp)

This paper describes a music retrieval system that enables a user to retrieve a song by two different methods: by singing its melody or by saying its title. To allow the user to use those methods seamlessly without changing a voice input mode, a method of automatically discriminating between singing and speaking voices is indispensable. We therefore first investigated measures that characterize differences between singing and speaking voices. From subjective experiments, we found that human listeners discriminated between these two voices with 70% accuracy for 200-ms signals. These results showed that even short-term characteristics such as the spectral envelope represented as MFCC can be used as a discrimination cue, while the temporal structure is the most important cue when longer signals are given. According to these results, we then developed the automatic method of discriminating between singing and speaking voices by combining two measures: MFCC and an $F0$ (voice pitch) contour. Experimental results with our method showed that 68.1% accuracy was obtained for 200-ms signals and 87.3% accuracy was obtained for 2-s signals. Based on this method, we finally built a music retrieval system that can accept both singing voices for the melody and speaking voices for the title.

10:45

4aMU8. Various acoustical aspects of an Asian (South) Indian classical music concert. M. G. Prasad (Dept. of Mech. Eng., Stevens Inst. of Technol., Hoboken, NJ 07030), V. K. Raman (Flautist, Germantown, MD 20874), and Rama Jagadishan (Edison, NJ 08820)

An Asian (South) Indian classical music concert is an integrated acoustical experience for both the audience and the player(s). A typical concert team, either vocal or instrumental, consists of a main vocalist (or an instrumentalist) accompanied by a violinist, up to three percussion instrument players, and a reference drone. The concert is comprised of many songs. Each song has two main parts, namely Alapana and Kriti. The Alapana is an elaboration of a raga (tune) and the Kriti refers to the lyrics of the song. The violinist actively follows and supports the main musician during the concert. The percussion player(s) are provided an opportunity to present a solo of their rhythmic skills. The players and the audience communicate emotionally and intellectually with each other. Elements such as aesthetics, rhythm, skill, and emotional aspects of the players are evaluated and appreciated by the audience. This talk will present various aspects of a concert that brings about an integrated and holistic experience for both the audience and the player(s). Some samples from live vocal and instrumental music concerts will be presented.

11:00

4aMU9. Musical scales, signals, quantum mechanics. Alpar Sevgen (Dept. of Phys., Bogazici Univ., Bebek 34342, Istanbul, Turkey)

Scales, being finite length signals, allow themselves to be treated algebraically: key signatures are related to the ''ring'' property of the scale labels; cyclically permuted scales and their mirror images have the same number of sharps and flats; and complementary scales (like major and pentatonic) have their sharp and flat numbers exchanged. A search for minimum principles to select among all possible scales those employed in music yields two possibilities: (a) minimize total number of accidentals and (b) minimize frequency fluctuations in a scale. Either of these minimum principles helps filter those scales employed in music from the universe of all scales, setting up very different criteria than the harmonic ratios used by musicians. The notes of the scales employed in music seem to prefer to stay as far apart from each other as possible. Operators that step through the multiplet members of scales with N semitones form a complete set of operators together with those that step through their eigenvectors. The mathematics reveals the discrete Fourier transformations (DFT) and is identical to finite state quantum mechanics of N-level Stern-Gerlach filters worked out by J. Schwinger.

FRIDAY MORNING, 1 DECEMBER 2006 MAUI ROOM, 7:30 A.M. TO 12:15 P.M.

Session 4aNS

Noise and Architectural Acoustics: Soundscapes and Cultural Perception I

Brigitte Schulte-Fortkamp, Cochair
Technical Univ. Berlin, Inst. of Technical Acoustics, Secr TA 7, Einsteinufer 25, 10587 Berlin, Germany

Bennett M. Brooks, Cochair
Brooks Acoustics Corp., 27 Hartford Turnpike, Vernon, CT 06066

Invited Papers

7:30

4aNS1. Soundscape in the old town of Naples: Signs of cultural identity. Giovanni Brambilla (CNR Istituto di Acustica ''O.M. Corbino'' Via del Fosso del Cavaliere 100, 00133 Roma, Italy), Luigi Maffei, Leda De Gregorio, and Massimiliano Masullo (Second Univ. of Naples, 81031 Aversa (Ce), Italy)

Like all cities in Magna Grecia, the ancient Neapolis was built along three main parallel, tight, and straight streets called *decumani*. Since then and during the following centuries, commercial and handicraft activities, as well as social life, have been developed along these streets. The narrow ground rooms forced shopkeepers to occupy the main street to show their merchandise using vocal appeals to magnify their product, and handicrafts to work directly on the street (hammering, sawing, etc.). Music artists had their performance on the streets too. The soundscape in the area was a strong symbol of the Neapolitan cultural identity. Nowadays *decumani* have kept the main features of the past but some of these are overrun by road traffic. To investigate in which way the traffic noise has modified the soundscape perception and cultural identity, sound walks were registered during day and night time. A number of residents were interviewed and laboratory listening tests were carried out. Despite the congested urban environment and high sound levels, preliminary results have shown that part of the residential population is still able to identify the soundscape more related to Neapolitan historical identity.

7:50

4aNS2. Soundscape design in public spaces: Concept, method, and practice. Hisao Nakamura-Funaba and Shin-ichiro Iwamiya (Kyushu Univ., 4-9-1.Shiobaru, Minami-ku, Fukuoka 815-8540, Japan)

Soundscape design of public spaces necessitates consideration of whether a space has important meaning for the user. It is essential to imagine an ideal sound environment of that space. We designed an actual soundscape from the viewpoint of the environment, information, and decoration. In many cases, producing some silence in an environment becomes the first step of soundscape design. There is neither a special technology nor a technique when designing. It merely requires use of a general technology and techniques concerning sound according to location. A key point is knowledge of know how to coordinate these technologies and techniques. For instance, silence was made first at the renewal project of Tokyo Tower observatory through cancellation of its commercial and call broadcasting functions and installation of sound-absorbing panels to the ceiling. Next, suitable and small sounds were added at various points. Guests can take time to enjoy viewing as a result.