Study on Sound Structure of Traditional Georgian Polyphony (1): Analysis of Its Temperament Structure

KAWAI Norie[1], MORIMOTO Masako[2], HONDA Manabu[2], ONODERA Eiko[3], NISHINA Emi[3,4] & OOHASHI Tsutomu[1]

1. Introduction

Traditional Georgian polyphony, one of the world's greatest musical treasures, is characterized by specific vocalization, singing style, and musical composition. Its harmonics are generated by its characteristic temperament structure. Although well known among researchers and practitioners of Georgian music, the structure of its temperament differs markedly from that of equal temperament and the musical notation of the five-line staff widely used in Western European music. Comprehensive, quantitative analyses of the temperament structure of Georgian polyphony and studies on the effect it has on listeners from the viewpoint of human science have yet to be carried out. One reason for this lacuna is that no music source suitable for pitch analysis has been available. To accurately analyze the pitch of musical tones, the sound source subjected to analysis should be in monotone, since it is otherwise difficult to analyze the precise pitch of each tone contained in the harmonics. However, since prevailing recordings of traditional Georgian polyphony routinely combine multiple parts of the polyphony, it has been extremely difficult to separate the precise pitch of individual tones.

Under such circumstances, we were very fortunate that among our collaborators Professor Anzor Erkomaishvili, President of the International Center of Georgian Folk Song, kindly provided us with a compact disk (CD) suitable for our research. On this CD entitled "*Teach Yourself Georgian Folk Songs - Megrelian Song*", published by the International Center of Georgian Folk Song [1], each voice of three vocalists, in groups of each singing one of three parts of a polyphony, has been separately recorded on an independent recording track. This digitally recorded sound source enabled us to accurately analyze the pitch of each tone of each vocal part separately, and then to compare these tracks with other temperaments, and even allowed us to shift the pitch of each tone individually.

Affording ourselves of this opportunity, we have undertaken this study to reveal the temperament structure specific to traditional Georgian polyphony and to measure its effect on

1 Foundation for Advancement of International Science, Tsukuba, Japan. 2 National Center of Neurology and Psychiatry, Tokyo, Japan. 3 The Graduate University for Advanced Studies, Chiba, Japan. 4. The Open University of Japan, Chiba, Japan.
Corresponding author: Prof. Tsutomu Oohashi, E-mail: oohashi@aqua.nifty.jp

listeners. In doing so, we first quantitatively compared the temperament of a certain traditional Georgian song with various other temperament structures, including equal temperament. Furthermore, we developed a special sound source in which the pitch of each tone of a Georgian song was electrically transformed into the pitch matching that of equal temperament. Using this sound source, we then examined, by means of a psychological evaluation, the subjective impression that traditional Georgian polyphony makes on listeners.

2. Analysis of the temperament structure of traditional Georgian polyphony

Methods

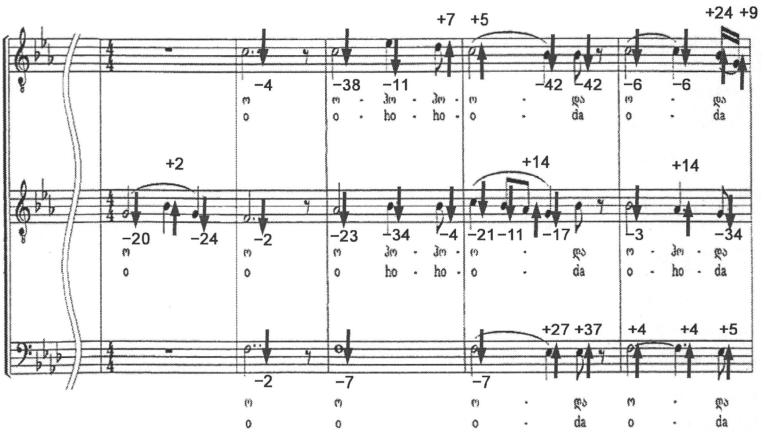
We chose a vocal composition entitled "O da" from the CD "*Teach Yourself Georgian Folk Songs - Megrelian Song*" as the subject of the present study. As mentioned above, this CD contain tracks on which three parts have been simultaneously recorded and three tracks on which each one of three vocal parts is separately recorded. Uploaded to a personal computer, the recording data of each part was analyzed with the software "Melodyne" (Celemony Software GmbH, München, Deutschland) to detect the pitch of individual monotones. This software enabled us to continuously detect the pitch of the fundamental tones. Since this software can also detect frequency changes due to vibrato and pitch fluctuation, the time width corresponding to a single musical note should be determined based on the song's musical score. Thus, the averaged frequency of a fundamental tone within the time width of each single musical note was taken as the pitch of that particular note.

Equal temperament, which can easily handle the pitch of a tone in a quantitative way, was employed as the assumed reference by which to express the pitch of individual tones. Each musical tone of the Georgian polyphonic composition was tentatively described using the nearest musical note in the assumed reference according to equal temperament. The pitch deviation of the original tone from the corresponding assumed reference tone was indicated by an upward or downward arrow and its values rendered in cents. The temperament of traditional Georgian polyphony was likewise compared to various other temperaments [2], such as just intonation (pure temperament), Pythagorean tuning and meantone temperament.

Results

Figure 1 shows a segment of the results of the "O da" pitch analysis. The direction of each arrow represents the direction of pitch deviation of the original "O da" tone from each corresponding reference tone according to equal temperament. Each value represents the degree of pitch difference, rendered in cents, between an original "O da" tone and each corresponding

reference tone. Most of the original tones exhibited pitch deviation of several or several tens of cents between each corresponding reference tone. Importantly, even when corresponding reference tones were the same in equal temperament, the degree and direction of pitch deviation of the original tones from the reference tone varied significantly depending on the vocal environment. All the data for our analysis will be presented elsewhere at a later date.



(Score cited from *"Teach Yourself Georgian Folk Songs - Megrelian Song"* [1])

Fig. 1 Results of pitch analysis of "O da" (segment)

Pitch analysis results for the whole epoch of "O da" were then summarized for each and every reference note and then shown in comparison with individual tones in equal temperament. Specifically, the mean and variation of pitch deviation of the original tones from the corresponding assumed reference tones were calculated for each and every assumed reference tone across the whole epoch of "O da". In Fig. 2, each box and error bar represents the averaged pitch shift with its standard deviation, respectively, from corresponding reference tone according to equal temperament. In order to reconcile the systematic difference in pitch between "O da" temperament and equal temperament, the whole scale was standardized relative to the tone of *ti*. This result shows that the pitch of individual "O da" tones deviated significantly from their corresponding reference tone according to equal temperament in varying degrees, suggesting that the temperament structure of traditional Georgian polyphony differs markedly from that of equal temperament.

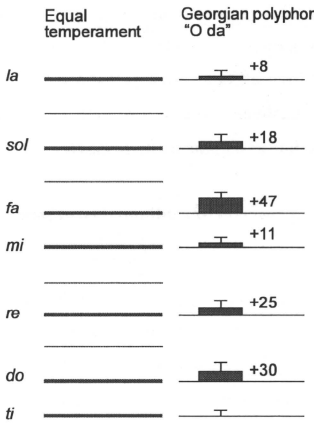


Fig. 2 Averaged pitch shift of "O da" from corresponding reference tone based on equal temperament

The temperament of the above song was then compared to various other temperaments to examine whether that used in "O da" was concordant. Figure 3 shows how each note of the various temperaments deviated relative to the mean pitch of each corresponding "O da" tone.

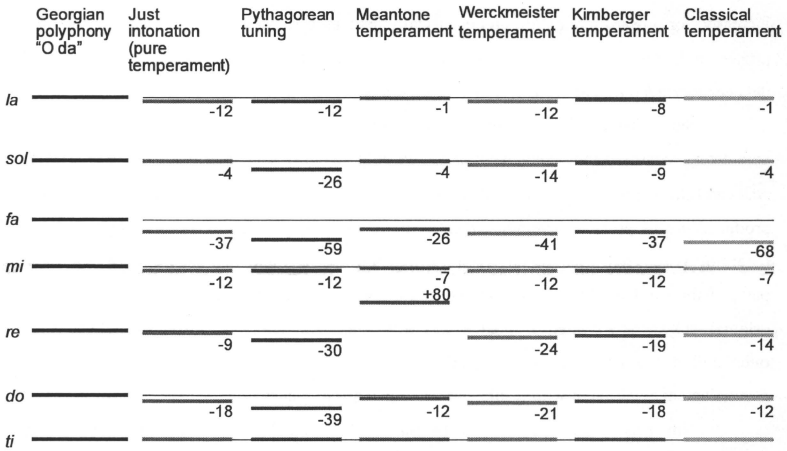


Fig. 3 Deviation with other temperaments in comparison to "O da" Temperament

Notably, the temperament of "O da" does not fully correspond to any other temperament, suggesting that the temperament structure of traditional Georgian polyphony exhibits unique characteristics, distinguishing it from all other existing temperaments. Interestingly, although the temperament of "O da" was relatively close to just intonation (i.e., pure temperament), it exhibited considerable pitch deviation at semitonic differences (e.g., *do* and *fa*).

3. Psychological effect of the temperament of traditional Georgian polyphony on listeners

Methods

The analysis above suggests that Georgian polyphony can be characterized as having a unique temperament structure. Based on a new approach, we proceeded to examine how its specific temperament affected listeners. Using state-of-the-art technology for electrical acoustic processing, we developed a sound source in which only the pitch of each note of original music shifts with no change in any other factor relative to sound quality. We then compared the pitch-shifted sound source with the original music in terms of a listener's psychological response.

The pitch of each "O da" tone was independently shifted to the nearest tone of equal temperament. Pitch shift was separately performed for each of the three parts of "O da" using "Melodyne" software on a personal computer. First, the direction and degree of pitch shift of each "O da" tone were determined by the above-mentioned analysis so as to correspond to equal temperament. The number of waves of the corresponding sound signal was electrically increased for upward pitch shift, and electrically decreased for downward pitch shift. Then the processed data were elongated or shortened in time so that the temporal length of the generated data corresponded to that of the original music tones.

Since this processing might significantly change the tone color, or timbre, various parameters regarding tone color were extracted beforehand. Using such parameters, tone color was carefully adjusted at the final stage of the pitch shift process. The three pitch-shifted parts produced in this process were subsequently mixed and downloaded to a CD. This digitalized recording is referred to as "O da" Pitch-Shifted Into Equal Temperament. Similarly, the three parts of the original "O da" were mixed and downloaded to a CD, which then served as the original "O da". We then evaluated the differences in effect on listeners between the pitch-shifted "O da" and the original "O da". Below we report our findings from a psychological experiment with results obtained on a questionnaire.

Nineteen subjects participated (8 male and 11 female; mean age, 44.9 y.o.) in this experiment. The original "O da" and the pitch-shifted "O da" were presented serially. The

subjects filled out a questionnaire to rate the sound quality in terms of 21 elements, each identified by a pair of contrasting Japanese terms (e.g., "natural" vs. "artificial"). Each element of each contrasting pair of terms was graded on a scale of 1 to 5. After a short interval, the two sound sources were presented in reverse order. The order of presentation was alternated among subjects.

The two sound sources were presented to subjects with identical sound presentation systems consisting of a CD player (HSCD-20, Action Research Inc., Tokyo, Japan), mixing console (9098i, AMEK, Potters Bar, UK), power amplifier (P-1000, Accuphase Laboratory Inc., Yokohama, Japan) and speakers (OOHASHI MONITOR, a special speaker developed by Prof. Tsutomu Oohashi).

Results

Compared to the "O da" Pitch-Shifted Into Equal Temperament, the original "O da" with the specific temperament of traditional Georgian polyphony was judged as more pleasant with statistical significance in terms of the following three elements of sound quality: "clear and limpid", "feel at ease" and "likable" (Fig. 4). Among other evaluation elements, there was also a tendency for the original "O da" to be more positively perceived by listeners than the "O da" Pitch-Shifted Into Equal Temperament.

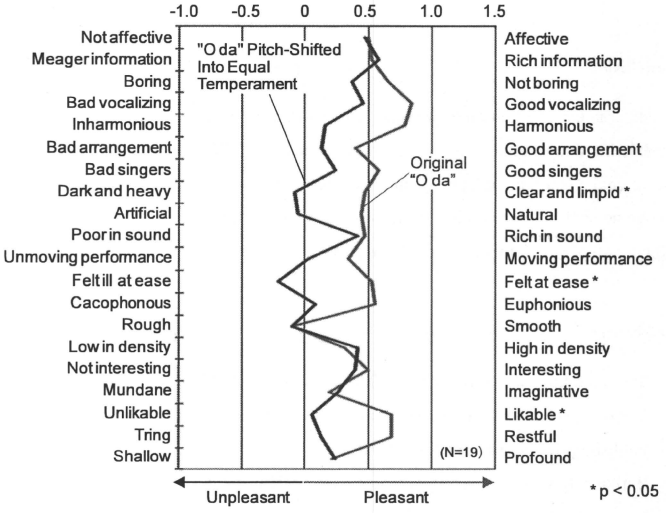


Fig. 4 Subjective impressions of Original "O da" and "O da" Pitch-Shifted Into Equal Temperament

4. Discussion

The present study shows that each tone of the temperament of traditional Georgian polyphony differs in various ways from that of equal temperament. Its temperament also differs from any existent temperaments including just intonation (i.e., pure temperament). Additionally, although the temperament of traditional Georgian polyphony somewhat corresponds to just intonation, it shows significant differences in pitch from just intonation at portions with a semitonic difference.

It is noteworthy that the pitch of each tone of traditional Georgian polyphony does not have a fixed relationship with equal temperament but varies considerably depending on the situation. This finding suggests the existence of a surprising mechanism: the structure of Georgian polyphony temperament, namely, its unique harmonic structure and temporal sequence of tones, exhibits the flexibility to appropriately adjusted within a certain range the pitch of each tone, depending on its mutual relationship with other voices simultaneously singing other parts. Psychological evaluation revealed that, with statistical significance, the temperament of Georgian polyphony having such characteristics is more "clear and limpid", "feel at ease" and "likable" than equal temperament. This suggests that the specific temperament structure of traditional Georgian polyphony contributes to strongly inducing a positive emotional response in listeners.

5. Conclusion

The temperament of traditional Georgian polyphony is shown to significantly differ from any other existing temperaments including equal temperament. The specific temperament structure of traditional Georgian polyphony thus seems to play an important role in generating its attractive acoustic characteristics. In pursuing this study, we make note of the attractiveness of Georgian polyphony through a scientific approach.

6. References

[1] The International Centre for Georgian Folk Song, CDs *"Teach Yourself Georgian Folk Songs - Megrelian Song"* (Collection of sheet of music and Four Compact Discs), Publishing House Sakartvelos Matsne, 2005.

[2] Barbour, J. M., "Tuning and Temperament - A Historical Survey", Michigan State University Press, 1972.

Study on Sound Structure of Traditional Georgian Polyphony (2): Quantitative Analysis of Its Fluctuation Structure

MORIMOTO Masako[1], HONDA Manabu[1], NISHINA Emi[2], KAWAI Norie[3] & OOHASHI Tsutomu[3]

1. Introduction

The Western European concept of music that provides the foundation for conventional music and audio theory regards an articulated musical tone having a stationary sound signal structure as the basic musical component. By way of contrast, in applying the sound communication model in the communication science domain, we postulate a biological concept of music in which music is defined as follows: music is an artificial sound system activating a neuronal auditory system and reward-generating system endowed with an informational structure that, at the macro-temporal level, provides sustaining patterns encoded by genes and cultures while at the micro-temporal level, continuously changes and is thus non-stationary [1-3]. This definition of music as being essentially non-stationary in character and thus offering a continuously changing informational structure at the micro temporal level is therefore quite at odds with the Western European concept of music, which regards stationary "musical tones" as the constituent musical component.

We previously examined the hyper-symbolic sound structure of traditional Georgian polyphony and compared it with singing voices from other cultures by using the maximum entropy spectral analysis method (MESAM) to visualize and observe the spectral fluctuation at the micro-temporal level. We discovered that the sound structure of traditional Georgian polyphony contains rich temporal fluctuation of power spectra at the micro-temporal level. Our previous method of analysis, however, was simply based on a qualitative inspection of the spectral array without providing any quantitative information regarding the complexity or degree of spectral fluctuation. Therefore, in the present study, we have developed an index that makes it possible to quantitatively evaluate the complexity of the fluctuation of the sound spectrum.

---

1 National Center of Neurology and Psychiatry, Tokyo, Japan. 2 The Open University of Japan, Chiba, Japan. 3 Foundation for Advancement of International Science, Tsukuba, Japan. Corresponding author: Prof. Tsutomu Oohashi, E-mail: oohashi@aqua.nifty.jp

2. Methods

We analyzed two traditional Georgian songs: *Chakrulo* (solo) and *Khasanbegura* (trio). These were recorded with a 4939 microphone (Brüel & Kjær, Nærum, Denmark) and Y. Yamasaki's high-speed sampling, one-bit coding signal processor . This recording system has a sampling frequency of 3.072MHz and a good response over 100kHz. For the comparison, we also analyzed an operatic solo recorded in DVD-audio [4].

Power spectral analysis at a micro-temporal level was carried out by MESAM, applying the mathematical formulae of the maximum entropy method and the power spectral estimation from the autoregressive model [2,3]. First, we digitally sampled the recorded singing voice data using DAQ Card-6062E and the device's software (National Instruments Co., Austin, TX, USA) with a sampling frequency of 250 kHz. The power spectrum of the sound data was calculated for every 20-msec epoch with an overlap of 10 msec by the maximum entropy method, which is known to be suitable for precise spectral estimation from a short period of data. The frequency resolution was 500 Hz. The estimated power spectra were displayed in a three-dimensional array. The spectral estimation and three-dimensional display were made by MATLAB (The MathWorks, Inc., Natick, MA, USA).

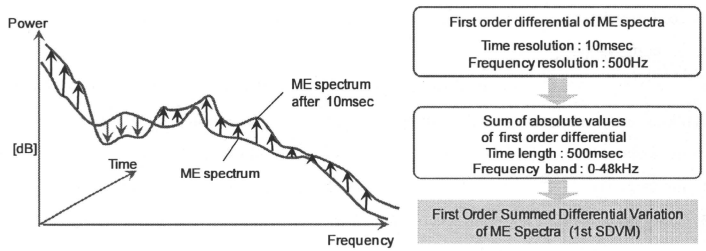


Fig. 1 The first Order Summed Differential Variation of ME Spectra (1st SDVM)

We then developed two indices by which we could quantitatively evaluate the complexity of the fluctuation of the sound spectrum. The first index, The First Order Summed Differential Variation of ME Spectra (1st SDVM) Index (Fig. 1), quantifies the degree of temporal fluctuation of the ME spectra. The ME spectra, drawn in 10 msec intervals with a frequency resolution of 500 Hz, were differentiated in a temporal direction. That is to say, the first order differential between a certain power spectrum and the next power spectrum at each frequency was calculated in dB. Absolute values of the calculated first differentials, namely, the

size of the difference in power regardless of increase or decrease, were summed across all frequencies, from 0 to 48 kHz. The summed values were then integrated across all the analysis epochs of 500 msec and represented as the 1st SDVM. This index is thought to represent not only minor changes at the macro temporal level corresponding to musical notes, but also more complex changes at the micro temporal level that cannot be described by musical notes, such as reverberation and attack.

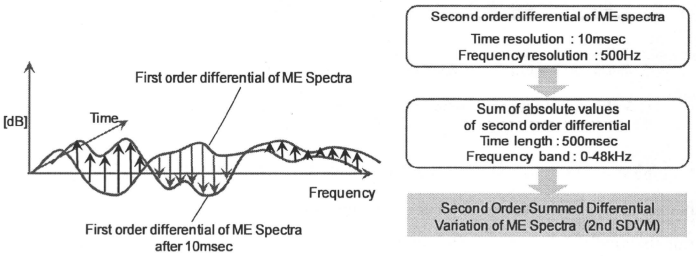
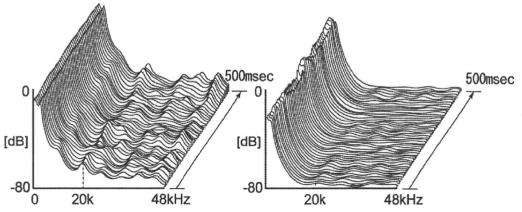


Fig. 2 The Second Order Summed Differential Variation of ME Spectra (2nd SDVM)

Although the 1st SDVM reflects the size of changes in power spectra, it does not reflect the complexity of such changes. In other words, the 1st SDVM shows a greater value for a monotonic increase, monotonic decrease or instantaneous major change corresponding to a musical note. To evaluate the complexity of changes in power spectra at the micro-temporal level, namely, the complexity in ups and downs of the power spectral array, we developed a second index, The Second Order Summed Differential Variation of ME Spectra (2nd SDVM) (Fig. 2). Similar to the 1st SDVM, these second order differentials were calculated and their absolute values were summed across the whole frequency range of 0 to 48 kHz and at the time epochs of 500 msec.

3. Results

Figure 3 shows ME spectral arrays of solo singing of traditional Georgian polyphony and that of solo singing of traditional Western opera. The traditional Georgian singing voice contains a rich, inaudible high-frequency component even beyond 40 kHz. In addition, even while a single keynote continues, the inaudible high-frequency component remains non-stationary and changes in complex ways. By contrast, the operatic *bel canto* singing voice does not contain any inaudible high-frequency components and is characterized by a periodical change of power spectra in an audible range, originating from a vibrato.



Traditional Georgian singing voice (solo) *Bel canto* singing voice (solo)

Fig. 3 ME spectral array of traditional Georgian singing voice and *bel canto* singing voice (solo)

Figure 4 shows 1st SDVM and the 2nd SDVM of Georgian and *bel canto* singing. Both indexes show greater values in Georgian singing than in *bel canto*. Such evidence supports the notion that traditional Georgian singing offers a more complex power spectral change.

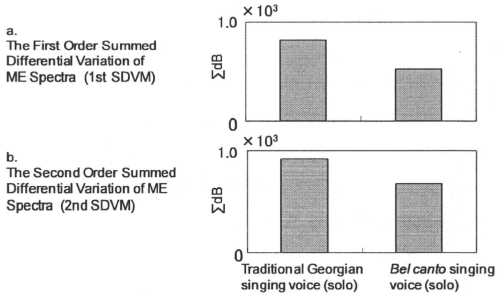


Fig. 4 Summed Differential Variation of ME Spectra
of traditional Georgian singing voice and *bel canto* singing voice (solo)

Figure 5 compares ME spectral arrays of solo and trio singing of traditional Georgian polyphony. We selected *Khasanbegura* for the trio. This piece contains more inaudible high-frequency components than does the solo. Moreover, spectral changes appear more prominent in the trio performance. Figure 6, with greater 1st SDVM and 2nd SDVM in the trio than in the solo corroborates such an interpretation. This finding suggests that the combination of three voices of traditional Georgian polyphony produces a more complex fluctuation in sound structure than does a solo performance.

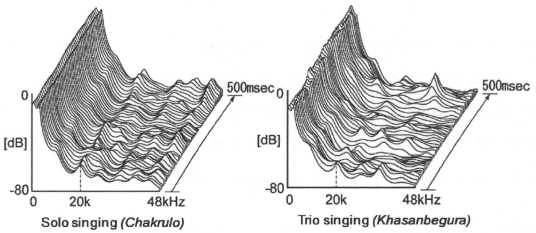


Fig. 5 ME spectral array of solo and trio singing in traditional Georgian polyphony

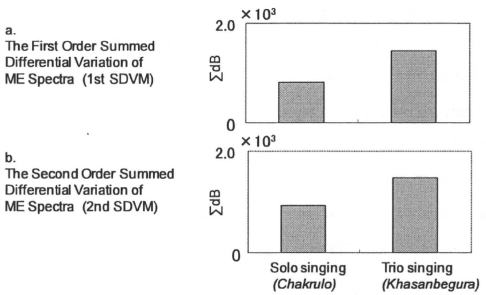


Fig. 6 Summed Differential Variation of ME Spectra of solo and trio singing in traditional Georgian polyphony

4. Discussion

In this study, we have developed indexes for the quantitative evaluation of fluctuation of sound structure utilizing our previously developed ME spectral array method. The 1st SDVM Index includes both rough changes at the macro temporal level corresponding to musical notes and more subtle changes at the micro temporal level, such as reverberation and attack. At the same time, the 2nd SDVM more directly reflects the complexity of fluctuation of the ME spectral array at the micro-temporal level.

Using these two indexes, we compared traditional Georgian singing and *bel canto* singing both in solo performances. We found that both the 1st and 2nd SDVM showed greater values for traditional Georgian singing than for *bel canto*. In addition, such complexity was enhanced in trio performances in which three different voices sang together. The epochs analyzed in the present study correspond to a single musical note and thus did not contain any

changes in musical interval. Since *bel canto* singing contains vibrato, the rough change of the ME spectral array in the audible range below 20 kHz, upon visual inspection, seems more pronounced than in the Georgian singing. Nevertheless, both the 1st and the 2nd SDVM show greater values in traditional Georgian singing than in *bel canto*. This suggests that traditional Georgian singing contains an extremely rich fluctuation structure at the micro-temporal level, which surpasses the simple, periodic change of power spectrum in *bel canto*.

These characteristics of the sound structure of traditional Georgian singing correspond well to the biological concept of music proposed by Oohashi, which postulates that the essential characteristic of music is a non-stationary, continuously changing informational structure at the micro-temporal level. The Western European concept of music regards music as characterized by articulated musical tones with a stationary sound signal structure as a basic component; music and musical score can be mutually transposed. By way of contrast, as Georgians say, a Georgian folk singer never sings any phrase exactly the same way twice [5]. It is interesting that non-Western styles of singing, including the Georgian, emphasize the importance of improvisation by the singer rather than the musical score of the composer. The improvisation component may thus contribute to the unconsciously perceived, extremely complex fluctuations of the sound structure at the micro-temporal level .

Recent auditory physiology has revealed that the stationary character of sounds, such as stationary pitch and strength, is processed at an earlier stage in the auditory nervous system, whereas the non-stationary, complex character of sounds, such as fluctuation in amplitude and frequency, is processed throughout the whole auditory system including the higher brain system, which includes the thalamus and cerebral cortex [6]. Notably, the expressional strategy of traditional Georgian polyphony seems quite musically appropriate because it provides the listener with extremely rich non-stationary fluctuation that strongly activates the whole auditory nervous system.

5. Conclusion

We have developed two Summed DifferentialVariation of ME Spectra indices, which makes it possible to quantitatively evaluate the complexity of the fluctuation of the sound spectrum. Using these indices, we quantitatively compared the complexity of fluctuation of the power spectra at the micro-temporal level of traditional Georgian polyphony with that of the *bel canto* vocal performance. The power spectra of traditional Georgian polyphony are shown to offer more complex fluctuation than does a *bel canto* performance. This finding endorses the singular

importance of complex fluctuation as an expressional strategy of traditional Georgian polyphony.

6. References

[1] Oohashi, T., "*Oto to Bunmei* (Sound and Civilization)", Iwanami Shoten, 2003.

[2] Nishina, E. et al., "Hyper-symbolic sound structure of *Kartuli* Polyphonia", Proceedings of the Second International Symposium of Traditional Polyphony, 2004.

[3] Morimoto, M. et al., "Transcultural study on frequency and fluctuation structure of singing voices", Proceedings of the 18th International Congress on Acoustics (ICA2004), 55-58, 2004.

[4] Cura, J., DVD-audio *"Dramatic Hero-Verismo Opera Arias"*, Warner Music Japan, WPAS-10010, 2001.

[5] The International Centre for Georgian Folk Song, Preface to CDs *"Teach Yourself Georgian Folk Songs - Megrelian song"* (Collection of sheet music and Four Compact Discs), Publishing House Sakartvelos Matsne, 2005.

[6] Moore, B. C. J., "Hearing", Academic Press, 1995.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Development of a Brain PET System, PET-Hat: A Wearable PET System for Brain Research

Seiichi Yamamoto, *Member, IEEE*, Manabu Honda, Tutomu Oohashi, Keiji Shimizu, and Michio Senda

***Abstract*—Brain functional studies using PET have advantages over fMRI in some areas such as auditory research in part because PET systems produce no acoustic noise during acquisition. However commercially available PET systems are designed for whole body studies and are not optimized for brain functional studies. We developed a low cost, small, wearable brain PET system named PET-Hat dedicated for brain imaging. It employs double counter-balanced systems for mechanical supports of the detector ring while allowing the subject some freedom of motion. The motion enables subject to be measured in the sitting position and move relatively freely with the PET during acquisition. The detector consists of a Gd_2SiO_5 (GSO) block, a tapered light guide and a flat panel photomultiplier tube (FP-PMT). Two types of GSO are used for depth-of-interaction (DOI) separation allowing the use of a small ring diameter without resolution degradation. The tapered light guide allows the use of larger GSO blocks with fewer FP-PMTs. Sixteen detector blocks are arranged in a 280 mm diameter ring. Transaxial and axial field-of-view (FOV) are 20 cm and 4.8 cm, respectively. Energy resolution of the block detectors was $\sim$ 15% full width at half maximum (FWHM) and timing resolution was $\sim$ 4.6 ns FWHM. Transaxial resolution and axial resolution at the center of the FOV were $\sim$ 4.0 mm FWHM and $\sim$ 3.5 mm FWHM, respectively. Sensitivity was 0.7% at the center of the axial FOV. Scatter fraction was $\sim$ 0.6. Hoffman brain phantom images were successfully obtained. We conclude that the PET-Hat is a promising, low cost, small, wearable brain PET system for brain functional studies.**

***Index Terms*—Brain, GSO, PET, PSPMT, wearable.**

I. Introduction

In the early stage of human activation studies, positron emission tomography (PET) was used and many interesting brain functional insights were obtained [1]–[4]. After the introduction of the functional magnetic resonance imaging (fMRI) [5]–[6], most of these brain functional studies were shifted from PET to fMRI because the latter does not require positron radionuclides and thus does not require injections and has no radiation exposure. Furthermore, the activation sensitivity to the stimulation is generally higher than PET.

One drawback of the fMRI studies is the acoustic noise from the gradient sequence which makes it difficult to use for auditory experiments in human studies. Acoustic noise in MRI is usually more than 100 dB sound presser level (SPL) requiring headphones or bone conduction speakers for the auditory stimulation for fMRI studies, making it quite different from natural auditory conditions

Another drawback of the fMRI studies is that subject needs to lie in narrow and deep tunnel of the MRI making most of subjects uneasy, especially for subjects of claustrophobic show difficulty in measuring in MRI [7] and [8].

Brain functional studies using PET have advantages over fMRI in some areas such as auditory research of the brain because recent PET systems basically produce no acoustic noise with acquisition. However commercially available PET systems are designed for whole body studies and are not optimized for brain functional studies. Most of the commercial available PET systems are for imaging human body thus the diameter of the detector ring is large enough to image the human whole body increasing the cost and reducing the sensitivity of the PET system [9]–[12]. In addition, these commercially available PET systems measure subject lying on the bed in the tunnel of the PET. In the case of PET/CT system, the length of the tunnel became longer [9], [11] and the similar drawback to MRI system may be serious for claustrophobic subjects.

In the brain functional studies using delicate auditory stimulation, fMRI may not be a candidate for the imaging modality because of the serious acoustic noises and narrow spaces in the MRI measurements. PET will have an advantage for these applications. Commercially available whole-body PET systems are better, but like the MRI, subjects are measured while lying on the bed in the relatively long tunnel. In addition, the acoustic noise level in the tunnels of PET systems is much smaller than in MRI but relatively high from such as the cooling fans of the electronics in the gantry of the system.

For the measurements of sensitive stimulation such as the detection of hypersonic effect [13], subject must be measured in a silent and relaxed condition where the only target stimulus activates the subject. For the relax condition, it will be better to be measured in the sitting position. And if the detector ring can move with the subject's head, the subject may feel more relaxed during PET measurement while minimizing the head movement.

Some PET systems dedicated for brain measurements have been developed [14]–[18]. However in most of the PET system, subject must be lying on the bed during measurement while one

Manuscript received May 03, 2010; revised August 26, 2010; accepted January 03, 2011. This work was supported by the Japan Science and Technology Association through Core Research for Evolutional Science and Technology (CREST).

S. Yamamoto is with the Kobe City College of Technology, Nishi-ku, Kobe 651-2194, Japan (e-mail: s-yama@kobe-kosen.ac.jp).

M. Honda is with the National Center of Neurology and Psychiatry, Tokyo 187-8502, Japan (e-mail: honda@ncnp.go.jp).

T. Oohashi is with the Foundation for Advancement of International Science, Ibaraki 305-0062, Japan (e-mail: oohashi@fais.or.jp).

K. Shimizu and M. Senda are with the Institute of Biological Research and Innovation, Kobe, Hyogo 650-0047, Japan (e-mail: Shimizu@ibri.org; senda@ibri.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNS.2011.2105502

0018-9499/$26.00 © 2011 IEEE

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

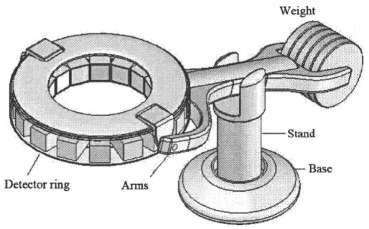


Fig. 1. Conceptual drawing of the PET-Hat.

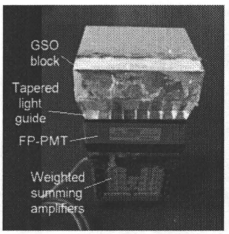


Fig. 2. Photograph of GSO block detector (from the top, GSO block, tapered light guide and FP-PMT) used for the PET-Hat. In photograph, printed boards of weighted summing amplifiers are shown under FP-PMT.

of the PET systems can measure in the sitting or standing position [18] but subject cannot move freely during measurement with the PET system. If the subject can move relatively freely while measurement, new neurological data that are impossible to measure such as measurements of blood flow changes during body or heads movement may become possible. Trying to satisfy these demands on the PET brain studies, we have developed a low cost, small and wearable PET system named PET-Hat.

II. Materials and Methods

A. Conceptual Design of the PET-Hat

Fig. 1 shows the conceptual drawing of the PET-Hat. The PET-Hat employs double counter-balanced systems in the mechanical supports. The detector ring of the PET system is supported by arms around which the detector ring can freely rotate during acquisition because the detector ring is balanced in the arms. The arms are supported by a stand and the detector weight is counterbalanced, which allows flee up and down motion. In addition, the stand can freely rotate around the base of the stand. These three motions enable subject to move relatively freely with the PET detector ring during acquisition by softly connecting the detector ring with the subject head.

B. GSO DOI Detector Block of the PET-Hat

The block detector for the PET-Hat consists of a Gd_2SiO_5 (GSO) block, a tapered light guide and a flat panel photomultiplier tube (FP-PMT). GSO was selected for the scintillators because the decay times can be controlled by the Ce concentration. Two types of GSO are stacked in the depth direction to form the depth-of-interaction (DOI) detector [19] and [20]. The DOI detection makes it possible to minimize the ring diameter of the PET system because it can reduce the spatial resolution degradation at off-center of the field of view (FOV). The tapered light guide is used to increase the size of the GSO blocks and reduce the number of FP-PMTs used for the PET-Hat.

The sizes of the GSO scintillators are $4.9\ \text{mm} \times 5.9\ \text{mm} \times 7\ \text{mm}$ for inner layer (GSO with 1.5 mol% Ce: decay time of 40 ns) and $4.9\ \text{mm} \times 5.9\ \text{mm} \times 8\ \text{mm}$ for outer layer (GSO with 0.4 mol% Ce: decay time of 80 ns), respectively. The inner layer means the layer closer to subject and the outer closer to the PSPMTs Light output difference between these two types of GSO were within 5%.

Depth length of these GSO scintillators was reduced to minimize the weight of the block detector for increasing the safety and decreasing the inertia of the PET-Hat. These GSO scintillators are combined into 11×8 matrix to form a block with size 55 mm (transaxial) $\times$ 48 mm (axial) . The GSO block is optically coupled to a FP-PMT through the tapered light guide. For the FP-PMTs, Hamamatsu H8500, 2-inches 8×8 multi-anode type [21] were used. The tapered light guide has $48\ \text{mm} \times 48\ \text{mm}$ area in the bottom (near to the FP-PMT) surface and $55\ \text{mm} \times 48\ \text{mm}$ in the top (near to the GSO block) surface and 8×8 tapered cells are combined with multi-layer optical film (ESR: 3M) between them. Fig. 2 shows the assembled GSO block detector with GSO block, tapered light guide and FP-PMT.

C. Configuration of PET-Hat

Sixteen GSO DOI block detector was arranged in a ring with diameter 280 mm. The signals from each GSO block detector is weighted summed and is fed to 100 MHz analog to digital (A-D) converters of the data acquisition system [22] and signals are integrated with two different integration time (120 ns and 320 ns) [23], calculating the position using the Anger principle by field programmable gate array (FPGA). Also coincidences are measured digitally among eight groups (2 block detectors for 1 group) and stored in list mode to the personal computer (PC). The data acquisition system is basically the same as that used for small animal PET systems [22]. Time window was set to 16 ns and lower energy window to 350 keV. The gain of the FP-PMTs was manually tuned to be similar level before acquiring the position map of the block detectors for setting the position boundaries and energy windows. Data for normalization was measured using a 24 cm diameter ring source phantom containing F-18 solution.

Fig. 3 shows the developed PET-Hat system. It consists of a detector ring with double counter balanced arm, reclining chair and a notebook PC. End-shields made of two layers of 2 mm thick tungsten contained rubber were pasted at the lower edge of the detector ring. The end-shield covered scintillator blocks and the length was 2 cm.

The data acquisition system is encased under the detector ring. The control of the PET-Hat as well as data processing including image reconstruction is controlled by the notebook PC by wireless communication with a desk-top PC under the detector ring.

Fig. 4 shows photograph of the PET-Hat with a subject. Subject can sit on the reclining chair and the PET-Hat can be set

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

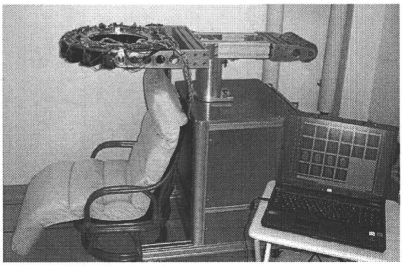


Fig. 3. Developed PET-Hat system.

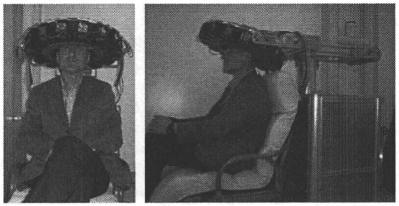


Fig. 4. PET-Hat system with subject; front view (left) and side view (right).

from the top. By softly connecting the subject's head to detector ring using such as a straw hat that is attached at the detector ring of the PET-Hat, the subject can move relatively freely with the detector ring of the PET-Hat during acquisition.

Parts of a movie showing the movement of the PET-Hat with subject are shown in Fig. 5. The horizontal movement of the head and detector ring of the PET-Hat is shown in Fig. 5(a), the vertical movement in Fig. 5(b). The subject and the detector ring were connected with a straw hat that was connected with the detector ring of the PET-Hat.

D. Performance Evaluation of PET-Hat

1) Spatial Resolution: Spatial resolution measurements were made using a 1 mm diameter spherical shape Na-22 point source (radioactivity: 300 kBq) positioned at the center, 0 cm, 4 cm, 6 cm and 8 cm from the center of the FOV. Random coincidences were subtracted using the delayed data. At each position, more than 100 k counts were accumulated. List mode data were sorted into sinograms, after single slice rebinning with maximum ring difference of 4 and 2D filtered back-projection with ramp equivalent real space filter was used for reconstruction. Images were made with and without DOI correction.

2) Axial Resolution: Axial resolution was measured using the same Na-22 point source (1 mm diameter spherical shape point source, with radioactivity of 300 kBq). Images of the point source were reconstructed using 2D filtered back-projection with ramp equivalent real space filter and coronal images were re-sliced and evaluated.

3) Sensitivity: Sensitivity was measured by moving a Na-22 point source (1 mm diameter spherical shape point source, with radioactivity of 300 kBq) in the axial direction and true coincidence rates were measured as a function of axial position.

(a) (b)

Fig. 5. Parts of a movie showing movement of the PET-Hat with subject; horizontal movement (a) and vertical movement (b).

4) Scatter Fraction: Scatter fraction was measured using a National Electrical Manufacturers Association (NEMA) 20 diameter, 70 cm long phantom using F-18 solution (radioactivity: $\sim$ 10 MBq) contained in the tube. The phantom was positioned at the center of the FOV. Scatter fraction was evaluated inside the FOV (20 cm). Scatter fraction was evaluated based on the NEMA NU 2001 standard [24].

5) Count Rate Performance: Count rate performance was measured using a NEMA standard 20 cm diameter, 70 cm height phantom contained $\sim$ 74 MBq F-18 solution. Following the decay of F-18, prompt, delayed and prompt minus delayed count rate were measured.

Noise equivalent count rate was also evaluated using the following formula with $k = 2$ because we used delayed coincidence for random correction.

$$\mathrm{NECR} = \frac{(\mathrm{T} \times \mathrm{T})}{(\mathrm{T} + \mathrm{kR} + \mathrm{S})}$$

where

T	true count rate
R	random rate
k	Delayed coincidence fraction
S	scatter rate

6) Hoffman Brain Phantom Imaging: The Hoffman brain phantom [25] contained 20 MBq of F-18 solution was positioned at the center of the FOV of the PET-Hat and measurements were made for 2 hours and total counts of $\sim$ 50 Mc were acquired. Data were reconstructed by 2D filtered back-projection using the normalization data. Analytical correction and single value subtraction were used for attenuation correction and scatter correction, respectively.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

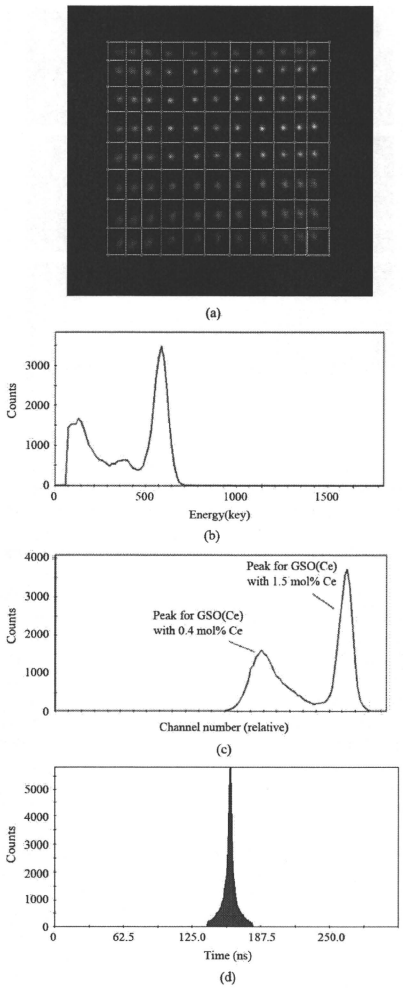


Fig. 6. Position map of the GSO block detector with position boundary (a) energy spectrum (b) pulse shape spectrum (c) and timing spectrum (d) of the GSO DOI block detector.

III. Results

A. Performance of the GSO DOI Block Detector

Fig. 6(a) shows the position map of the GSO block detector. Gamma photons form Cs-137 (661-keV) were irradiated from ~ 5 cm from the top of the GSO DOI block detector. The position map showed clear separation of all the GSO scintillators.

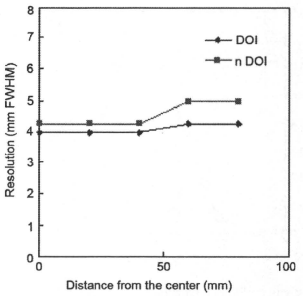


Fig. 7. Transaxial resolution as a function of distance from center.

The separation is good enough to divide each GSO scintillator using square boundaries. Fig. 6(b) shows an energy spectrum for one of the GSO scintillators in the block detector for Cs-137 gamma photons. The spectrum showed a single peak although the GSO scintillators consist of dual layer GSO with different decay times. Energy resolution was 15% full width at half maximum (FWHM).

Fig. 6(c) shows a pulse shape spectrum of one of the GSO scitnillators of the block detector for Cs-137 gamma photons. The pulse shape spectrum showed good separation of these two types of GSO with different decay times. The right peak in Fig. 6(c) is the GSO with 1.5 mol % Ce and left is the with 0.4 mol% Ce. The peak to valley (P/V) ratio among these peaks was 14. With this P/V ratio, the percent error in separation of two layers is almost zero.Fig. 6(d) shows the timing spectrum measured using a positron source between GSO block detectors. Timing resolution was 4.6 ns FWHM. The timing spectrum showed wider distribution at the bottom area so time window was set relatively wider (16 ns).

B. Performance of PET-Hat System

1) Spatial Resolution: Fig. 7 shows transaxial resolution as a function of distance from the center. Transaxial resolutions at the center of the FOV were 4.0 mm FWHM with DOI correction and 4.3 mm FWHM without DOI correction at the center of the FOV and 4.2 mm FWHM with DOI correction and 5.0 mm FWHM without DOI correction at 8 cm from the center of the FOV.

2) Axial Resolution: The axial resolution at the center of the FOV was 3.5 mm FWHM.

3) Sensitivity: Sensitivity profile as a function of the axial position is shown in Fig. 8. Sensitivity for point source was 0.72% at the center of the axial FOV. The count rate outside the axial FOV (48 mm) is from the scatter coincidence between detector blocks when the source is outside FOV.

4) Scatter Fraction: Scatter fraction as a function of slice number is shown in Fig. 9. Average scatter fraction was 0.6.

5) Count Rate Performance: Count rate characteristic is shown in Fig. 10. The maximum prompt minus delayed count rate was ~ 12 kcps and NECR was 0.82 kcps within the measured activity range.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

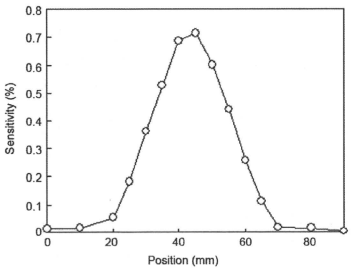


Fig. 8. Sensitivity profile as a function of the axial position.

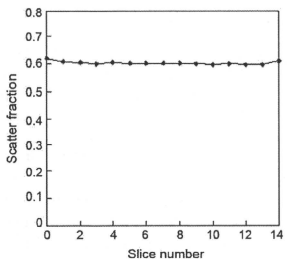


Fig. 9. Scatter fraction as a function of the slice number.

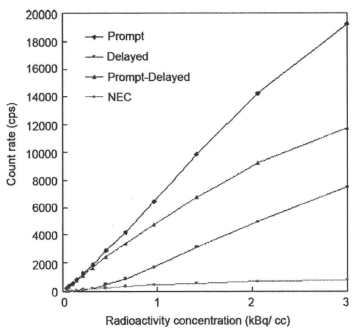


Fig. 10. Count rate characteristic measured using 20 cm diameter, 70 cm height cylindrical phantom.

6) Images of the Hoffman Brain Phantom: Images of the Hoffman brain phantom at the central 9 slices are shown in Fig. 11. In the images we can observe the small structures of the gray matter regions of the phantom.

IV. Discussion

We successfully developed a wearable brain PET system. The PET-Hat could move relatively freely with subject movement.

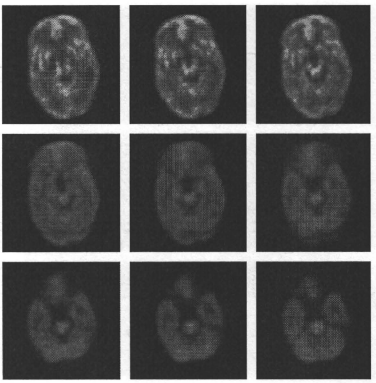


Fig. 11. Hoffman brain phantom images contained F-18.

However the rotation or sideways neck motion produced the sift of position in the straw hat (connecting part of PET-Hat and head) that may produce image degradation from subject movements. For more free movement, rotation of the detector ring with the subject can add one more free movement of the subject with some sacrificing of the increase of the weight of the detector ring. The increase of the weight increases the inertia thus increases force for starting and stopping the movement of the detector ring.

The sensitivity of the system can be increased by several ways; increase the depth of the scintillators, use of the dense, high atomic number scintillators such as LGSO, LYSO, or LSO and increase the axial FOV. These attempts to increase the sensitivity also increase the weight of the detector ring that will require more safety mechanism for the mechanical support such as counter balanced system.

In this PET-Hat system, the effect of DOI detection was not very obvious because the depth of the GSO scintillators were relatively short, 7 mm and 8 mm. If we select GSO scitnillators with longer depth, the difference of the spatial resolution with and without DOI detection would be more attractive. However in this case, the weight of the detector ring would more heavy.

The scatter fraction of the system was relatively high, higher than the whole body PET systems [9]–[12]. The reasons are the end-shield of the detector ring is set only the lower side of the detector ring and its thickness and length are small. For the brain studies, only the scatter from the lower side of the ring will be important because there is no activity on the upper side of the detector ring. Thus the scatter contribution of the human studies will be smaller than that used NEMA phantom. The use of an additional gamma shield from the body such as gamma absorb apron may be useful for the human studies.

The image quality of the Hoffman brain phantom shown in Fig. 11 was not very attractive. One reason is the low sensitivity of the system with the small axial FOV (44 mm) and the short scintillators depth (15 mm). The other reason is the low NECR of the system because of the high scatter fraction and random

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

rate. Using the filtered back-projection for the image reconstruction is another reason of the limited quality of the phantom images. Optimization of the system parameters such as lower energy level or time window may improve the image quality in somehow. Also applying an iterative image reconstruction will improve the image quality of the PET system.

V. Conclusion

We have successfully developed the PET-Hat for brain research. The PET-Hat could be measured in the sitting position of the subject and could move with the subject. We conclude that the PET-Hat is promising, low cost, small size, wearable brain PET system for brain functional studies.

Acknowledgment

The authors would like thank to Mr. Ohta of Eiko-Sangyo Co. for designing and constructing the mechanical part of the PET-Hat system.

References

[1] P. T. Fox, M. A. Mintun, M. E. Raichle, F. M. Miezin, J. M. Allman, and D. C. Van Essen, "Mapping human visual cortex with positron emission tomography," *Nature*, vol. 323, pp. 806–809, Oct. 1986.

[2] S. E. Petersen, P. T. Fox, A. Z. Snyder, and M. E. Raichle, "Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli," *Science*, vol. 249, no. 4972, pp. 1041–1044, 1990.

[3] M. Corbetta, F. M. Miezin, S. Dobmeyer, G. L. Shulman, and S. E. Petersen, "Attentional modulation of neural processing of shape, color and velocity in humans," *Science*, vol. 248, no. 4962, pp. 1556–1559, 1990.

[4] P. T. Fox, M. E. Mintun, M. E. Raichle, and P. P. Herscovitch, "A noninvasive approach to quantitative functional brain mapping with H2 (15)O and positron emission tomography," *J. Cereb. Blood Flow Metab.*, vol. 4, no. 3, pp. 329–333, 1984.

[5] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," *Proc. Natl. Academy Sciences USA*, vol. 87, no. 24, pp. 9868–9872, 1990.

[6] S. Ogawa, T. M. Lee, and G. Barrere, "The sensitivity of magnetic resonance image signals of a rat brain to changes in the cerebral venous blood oxygenation," *Magn. Reson. Med.*, vol. 29, no. 2, pp. 205–210, 1993.

[7] I. Eshed, C. E. Althoff, B. Hamm, and K. G. Hermann, "Claustrophobia and premature termination of magnetic resonance imaging examinations," *J. Magn. Reson. Imag.*, vol. 26, no. 2, pp. 401–404, 2007.

[8] S. Thorpe, P. M. Salkovskis, and A. A. Dittner, *Magn. Reson. Imag. Claustrophobia in MRI: The Role of Cognitions*, vol. 26, no. 8, pp. 1081–1088, 2008.

[9] M. Teräs, T. Tolvanen, J. J. Johansson, J. J. Williams, and J. Knuuti, "Performance of the new generation of whole-body PET/CT scanners: Discovery STE and discovery VCT," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 34, no. 10, pp. 1683–1692, 2007.

[10] K. Matsumoto *et al.*, "Performance characteristics of a new 3-dimensional continuous-emission and spiral-transmission high-sensitivity and high-resolution PET camera evaluated with the NEMA NU 2-2001 standard," *J. Nucl. Med.*, vol. 47, no. 1, pp. 83–90, 2006.

[11] S. Surti, A. Kuhn, M. E. Werner, A. E. Perkins, J. Kolthammer, and J. S. Karp, "Performance of philips gemini TF PET/CT scanner with special consideration for its time-of-flight imaging capabilities," *J. Nucl. Med.*, vol. 48, no. 3, pp. 471–480, 2007.

[12] G. Brix *et al.*, "Performance evaluation of a whole-body PET scanner using the NEMA protocol. National electrical manufacturers association," *J. Nucl. Med.*, vol. 38, no. 10, pp. 1614–1623, 1997.

[13] T. Oohasi *et al.*, "In audible high-frequency sound affect brain activity: Hypersonic effect," *J. Neuriphysiol.*, vol. 83, no. 6, pp. 2548–2558, 2000.

[14] J. S. Karp, S. Surti, M. E. Daube-Witherspoon, R. Freifelder, C. A. Cardi, L. E. Adam, K. Bilger, and G. Muehllehner, "Performance of a brain PET camera based on anger-logic gadolinium oxyorthosilicate detectors," *J. Nucl. Med.*, vol. 44, no. 8, pp. 1340–1349, 2003.

[15] H. W. de Jong, F. H. van Velden, R. W. Kloet, F. L. Buijs, R. Boellaard, and A. A. Lammertsma, "Performance evaluation of the ECAT HRRT: An LSO-LYSO double layer high resolution, high sensitivity scanner," *Phys. Med. Biol.*, vol. 52, no. 5, pp. 1505–1526, 2007.

[16] E. Yoshida *et al.*, "Design and inital evaluation of a 4-layer DOI-PET system: The jPET-D4," *Igaku Butsuri.*, vol. 26, no. 3, pp. 131–140, 2006.

[17] D. L. Bailey, T. Jones, T. J. Spinks, M. C. Gilardi, and D. W. Townsend, "Noise equivalent count measurements in a neuro-PET scanner with retractable septa," *IEEE Trans. Med. Imag.*, vol. 10, no. 3, pp. 256–260, 1991.

[18] M. Watanabe *et al.*, "A new high-resolution PET scanner dedicated to brain research," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 3, pp. 634–639, Jun. 2002.

[19] S. Yamamoto and H. Ishibashi, "A GSO depth of interaction detector for PET," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 3, pp. 1078–1082, Jun. 1998.

[20] S. Yamamoto, "A dual layer DOI GSO block detector for a small animal PET," *Nucl. Instrum. Methods Phys. Res. A*, vol. A598, pp. 480–484, 2008.

[21] R. Pani, M. N. Cinti, R. Pellegrini, C. Trotta, G. Trotta, L. Montani, S. Ridolfi, F. Garibaldi, R. Scafè, N. Belcari, and A. D. Guerra, "Evaluation of flat panel PMT for gamma ray imaging," *Nucl. Instrum. Methods Phys. Res. A*, vol. A504, pp. 262–268, 2003.

[22] S. Yamamoto, H. Mashino, H. Kudo, K. Matsumoto, and M. Senda, "A dual layer GSO PET system for small animal: K-PET II," *IFMBE Proc.*, vol. 14, pp. 1712–1715, 2007.

[23] S. Yamamoto, "Optimization of the integration time of pulse shape analysis for dual layer detector with different amount of Ce," *Nucl. Instrum. Methods Phys. Res. A*, vol. A587, pp. 319–323, 2008.

[24] M. E. Daube-Witherspoon *et al.*, "PET performance measurements using the NEMA NU 2-2001 standard," *Nucl. Med.*, vol. 43, no. 10, pp. 1398–1409, 2002.

[25] E. J. Hoffman, P. D. Cutler, W. M. Digby, and C. J. Mazziotta, "3-D phantom to simulate cerebral blood flow and metabolic images for PET," *IEEE Trans. Nucl. Sci.*, vol. 37, no. 2, pp. 616–620, Apr., 1990.