

演奏情報の可視化による聴覚情報認知の拡大と 聴覚障害者を対象とする応用システム

16500138

平成 16 年度～平成 19 年度科学研究費補助金

(基盤研究 (C))研究成果報告書

平成 20 年 6 月

研究代表者 平賀 瑠美

筑波技術大学産業技術学部教授

はしがき

聴覚情報と視覚情報の間にある共通の認識を見出すことを、演奏情報の可視化を通して実現することを目的とし、4年間の研究を遂行した。この目的のために、人間の聴覚を特に視覚情報により支援するコンピュータ・ソフトウェアのための基本的な研究として以下の二点を行った。

A) 演奏情報の最適な可視化の追求。

B) 聴覚障害者が音あるいは画像から認知する事柄の解明。

この成果として、8編の論文を国際会議で、そのうち1本はセッションベストペーパーとして英文論文誌に、1本を口頭による国際会議で発表した。

聴覚障害者が音楽演奏を楽しむことを支援するシステムの構築を目指すことを念頭に置き、B)の音楽認知については、ドラムの即興演奏を用いて様々な実験を行った。聴覚障害者の音楽認知については世界的にもほとんど行われておらず、画像、マルチメディアを併用した音楽認知についての新たな興味深い知見を得ることができ、今後の研究の礎とすることができた。

研究組織

研究代表者：平賀瑠美(筑波技術大学産業技術学部教授)

研究分担者：加藤伸子(筑波技術大学産業技術学部准教授)

交付決定額(配分額)

(金額単位：円)

	直接経費	間接経費	合計
平成16年度	1,100,000	0	1,100,000
平成17年度	1,000,000	0	1,000,000
平成18年度	700,000	0	700,000
平成19年度	900,000	270,000	1,170,000
総計	3,700,000	270,000	3,970,000

研究発表

(1) 雑誌論文

1. Hiraga, R. and Kawashima, M.: Performance Visualization for Hearing Impaired Students, Proceedings of International Conference on Education and Information Systems: Technologies and Applications, 査読あり, Vol. 1, 2004, pp. 323-328.
2. Hiraga, R. and Matsuda, N.: Graphical Expression of the Mood of Music, The 2004 IEEE International Conference on Multimedia and Expo, 査読あり, 2004, 4-pages in CD-ROM Proceedings.
3. Hiraga, R. and Matsuda, N.: Visualization of Music Performance as an Aid to Listener's Comprehension, Proceedings of Advanced Visual Interfaces, 査読あり, 2004, pp. 103-106.
4. Hiraga, R., Yamasaki, T., and Kato, N.: Cognition of Emotion on a Drum Performance by hearing-impaired people, 11th International Conference on Human-Computer Interaction, 査読あり, 2005, 4-pages in CD-ROM Proceedings.
5. Hiraga, R., Yamasaki, T., and Kato, N.: Express Emotion by Hearing Impaired People through Playing of Drum Set, The 9th World Multi-Conference on Systemics, Cybernetics and Informatics, 査読あり, 2005, 4-pages in CD-ROM Proceedings.
6. Hiraga, R. and Kato, N.: Understanding Emotion through Multimedia--Comparison between Hearing-Impaired People and People with Hearing Abilities, Proceedings of the Eighth International ACM SIGACCESS Conference on Computers and Accessibility, 査読あり, 2006, pp. 141-148.
7. Hiraga, R., Kato, N., and Yamasaki, T.: Understanding emotion through Drawings--comparison between hearing-impaired people and people with normal hearing abilities, Proceedings of the 2006 IEEE International Conference on systems, Man and Cybernetics, 査読あり, 2006, pp. 103-108.
8. Hiraga, R., Yamasaki, T., and Kato, N.: Recognition of intended emotions in drum performances: differences and similarities between hearing-impaired people and people with normal hearing ability, Proceedings of the 9th International Conference on Music Perception and Cognition, 査読あり, 2006, pp. 219-224.

9. Hiraga, R. and Kawashima, M.: Performance Visualization for Hearing-Impaired Students, *Journal of Systemics, Cybernetics and Informatics*, 査読あり, 2006, 3:5, pp. 24-32.

(論文 1 がセッションベストペーパーとして論文誌に再掲載されたもの)

(2) 学会発表

1. Hiraga, R.: The catch and throw of music emotion by hearing-impaired people, International Conference on Music Communication Science, 2007 年 12 月 6 日, Sydney, Australia.

目次

1. はじめに	1
2. 研究成果	2
2.1. 音楽演奏の可視化	2
2.2. 演奏情報の認知.....	3
2.3. 論文概要	4
2.3.1. 音楽演奏の可視化.....	4
2.3.2. 演奏情報の認知.....	6
3. まとめ	8
付	9
A. 既発表論文.....	9
A.1 音楽演奏の可視化.....	9
A.1.1. Graphical Expression of the Mood of Music	9
A.1.2. Assisting Listeners' Appreciation with Performance Visualization.....	13
A.2 演奏情報の認知.....	17
A.2.1. Performance Visualization for Hearing Impaired Students.....	17
A.2.2. Cognition of Emotion on a Drum Performance by hearing-impaired people	23
A.2.3. Express Emotion through playing a Drum set by Hearing Impaired People	27
A.2.4. Understanding Emotion through Multimedia-comparison between hearing-impaired people and people with normal hearing abilities	31
A.2.5. Understanding emotion through drawings-comparison between hearing-impaired people and people with normal hearing abilities	39
A.2.6. The cognition of intended emotions for a drum performance: differences and similarities between hearing-impaired people and people with normal hearing ability.....	45
A.2.7. Performance Visualization for Hearing-Impaired Students (revision of A.2.1)	51
A.2.8. The catch and throw of music emotion by hearing-impaired people.....	60

1. はじめに

「演奏情報の可視化による聴覚情報認知の拡大と聴覚障害者を対象とする応用システム」というタイトルのもとに、聴覚情報と視覚情報の間にある共通の認識を見出すことを、演奏情報の可視化を通して実現することを目的とし、4年間の研究を遂行した。この目的のために、人間の聴覚を特に視覚情報により支援するコンピュータ・ソフトウェアのための基本的な研究として以下の二点を行った。

A) 演奏情報の最適な可視化の追求。

ピアノ演奏の可視化を行った。健聴者がどのような表示に演奏表情を理解し、聞いただけでは気付かなかったことに注意を喚起できるかという点、ならびに表情と同様非常に主観的であいまいな演奏のムードの可視化を行った。

B) 聴覚障害者が音あるいは画像から認知する事柄の解明。

感情を託したドラムの即興演奏を用いて様々な演奏認知の実験を行った。聴覚障害者の音楽認知については世界的にもほとんど行われておらず、画像、マルチメディアを併用した音楽認知についてまでの新たな興味深い知見を得ることができ、今後の研究の礎とすることができた。

この成果として、8編の論文を国際会議で、そのうち1本はセッションベストペーパーとして英文論文誌に、1本を口頭による国際会議で発表した。

4年間の研究成果は、聴覚障害者が音楽演奏を楽しむことを支援するシステムの構築を目指す際の基礎的な資料として今後も活用されるものである。

2. 研究成果

4年間の研究では、まず、演奏表情の可視化研究を行い、次に演奏認知の研究をいった。参考文献番号は付の論文に対応する。

2.1. 音楽演奏の可視化

人間が音楽を演奏する時は楽譜からの逸脱が生じ、それが音楽らしさや人間が好ましい、美しいと感じることのできる演奏を作り出す。そのような逸脱は、演奏を聴くことにより漠然と感じることはできるが、楽譜のどの箇所でのどのような楽譜からとの差異が生じているのかを正確に知ることは聴くことだけでは難しい。楽譜と演奏情報の差分を用いて、演奏表情の情報を数値化し、それを可視化することでより容易に表情の由来・原因を知ることができる。

本研究で行った演奏の可視化は、演奏情報がデジタル情報として記録されるいわゆる MIDI ピアノ（たとえば YAMAHA のピアノプレーヤー YP10, YP30 など）を演奏した結果の MIDI 情報を用いた。MIDI は音楽演奏の規格であり、規格にそった形で、演奏に関する基本的情報（音色あるいは楽器の種類、テンポ、拍子などと各音を構成する情報）を含むバイナリファイルとして保存される。本研究で用いた情報は、各音のオンセット（発音時刻）、オフセット（消音時刻）、鍵盤を押す速度で表される音の大きさである。

オンセット値を楽譜情報と比較するということは、メトロノームで正確に刻まれた時刻からどれくらい各音の演奏開始が外れるかということであり、これにより、楽譜上に明記されていなくても、メロディの構成から生ずる加速や減速といったテンポゆらぎを知ることができる。オンセット値とオフセット値を組み合わせ、楽譜情報と比較することで、楽譜上に記される各音の音価（4分音符・8分音符といった音符の楽譜上の長さ）が実際にはメロディやその中の各音符の役割により短めに弾かれたり、滑らかに演奏されているということが分かる。また、演奏に必要な間や、アーティキュレーション（呼吸間隔とでも呼べる繋がり・途切れのめりはり）が明らかになる。鍵盤速度による音の大きさは人間の耳に届く音圧とリニアないしは固定式で表される関係はないが、大きさの変化を知るのには利用できる値である。したがって本研究では、テンポゆらぎ、アーティキュレーション、音量の変化を可視化した。

演奏情報を可視化することにより、演奏情報の由来・原因を知ることがより容易になるため、ピアノ学習者が自分の演奏を視覚により理解したり、他者の演奏と比べたりすることができる可視化システムを念頭におき、プロトタイプを作成した[A.1.2]。このような局所的な表情の理解だけではなく、曲中のある単位（4小節、8小節、16小節など）の雰囲気がどのように変化していくかを見ていくこともできるように、ある程度の長さについてテクスチャとして可視化することも試みた[A.1.1]。

2.2. 演奏情報の認知

研究代表者や研究分担者が所属する筑波技術大学産業技術学部に入学者は、幼少より障害を持ち、音楽経験が制限されている者が多い。しかし、1997年より6年間行った「コンピュータ・ミュージック」の授業を通し、音楽に興味を持つ者が多いということを知った。コンピュータ・ミュージックの授業では、既存のアコースティック楽器ではなく、新しい電子楽器を学生に紹介し、学生自らがそれを使って演奏を発表した。電子楽器を用いた理由は、(1) コンピュータと接続することで、図形を表示できること、(2) 演奏方法が自由で容易に演奏状態に達することができるため、(3) 授業を行う者が音楽の専門家ではないため、である。

聴覚障害者が音楽に興味を持ち、自ら楽しんで楽器演奏をするということは明らかになったが、お互いに何かを感じあえる演奏となっているかという点については、打楽器を使ったサンバアンサンブル（バトゥカーダ）演奏を学生に課した時、演奏による一体感を感じたことから適切な環境があればお互いに音楽を楽しんで作っていけるはずであると確信した[A.2.1][A.2.7]。

このことにより、演奏を支援する環境の構築を目指し、必要なデータの収集を行った。聴覚障害者が取り組みやすいであろうということと振動による情報保障があり得るということから打楽器を用いたアンサンブルを想定し、アンサンブルで共有する情報は感情ならば音楽経験や音楽知識がなくても可能そうであることから、基本的な4つの感情“喜び”、“恐怖”、“怒り”、“悲しみ”を表現するところから着手した。

ここで必要なデータとは、おもに聴覚障害者が感情を区別して打楽器で演奏表

現ができるかどうか[A.2.3], 聴覚障害者は感情を表現した演奏に演奏者が意図したものを認識するかどうかとことについてである. これらについて, 聴覚障害者と健聴者のグループを被験者として比較実験を行った [A.2.2][A.2.4][A.2.5][A.2.6][A.2.8].

2.3. 論文概要

4年間に8編の論文を国際会議で, そのうち1本はセッションベストペーパーとして英文論文誌に, 1本を口頭による国際会議で発表した. 論文は大きく音楽演奏の可視化に関連するもの(2編)と演奏情報の認知に関するもの(7編と発表)に分けられる. 演奏情報の認知に関するものは, 聴覚障害を持つ学生への音楽の授業(2編), 聴覚障害者によるドラム演奏の分析と感情認識(2編), 感情をメディア上で表現した時の認識(3編), 聴覚障害者同士によるドラム演奏を通しての感情コミュニケーション(発表)に分けられる.

図1 研究の流れ・テーマと既発表論文(論文名の番号は付の論文番号に対応)に研究の流れを発表論文と併せ記す.

論文概要は以下の通りである.

2.3.1. 音楽演奏の可視化

A.1.1. Graphical Expression of the Mood of Music

楽曲中の各音についてのオンセット・オフセット・音量のデータを用いてテンポとアーティキュレーション情報を計算し, 値をRGBに変換, それらすべての音についての色を小さい方形のパネルに表す. 楽曲中のすべての音についての方形パネルを決められた順序で結合して大きい方形パネルを作った. 楽曲中の音の数によらず同一サイズの方形パネルを提示した.

音符の時系列を除くことで楽曲の雰囲気を表そうと試みたものである.

A.1.2. Visualization of Music Performance as an Aid to Listener's Comprehension

楽曲中の各音についてのオンセット・オフセット・音量のデータを用い

てテンポとアーティキュレーション情報を計算，楽曲中の各音の役割（音階上の重要性や拍子上の重要性）と合わせ，それらの計算値を各音について表す．演奏の傾向や音の役割が演奏に反映されているかどうかを一目で知ることができる可視化表現を試みたものである．

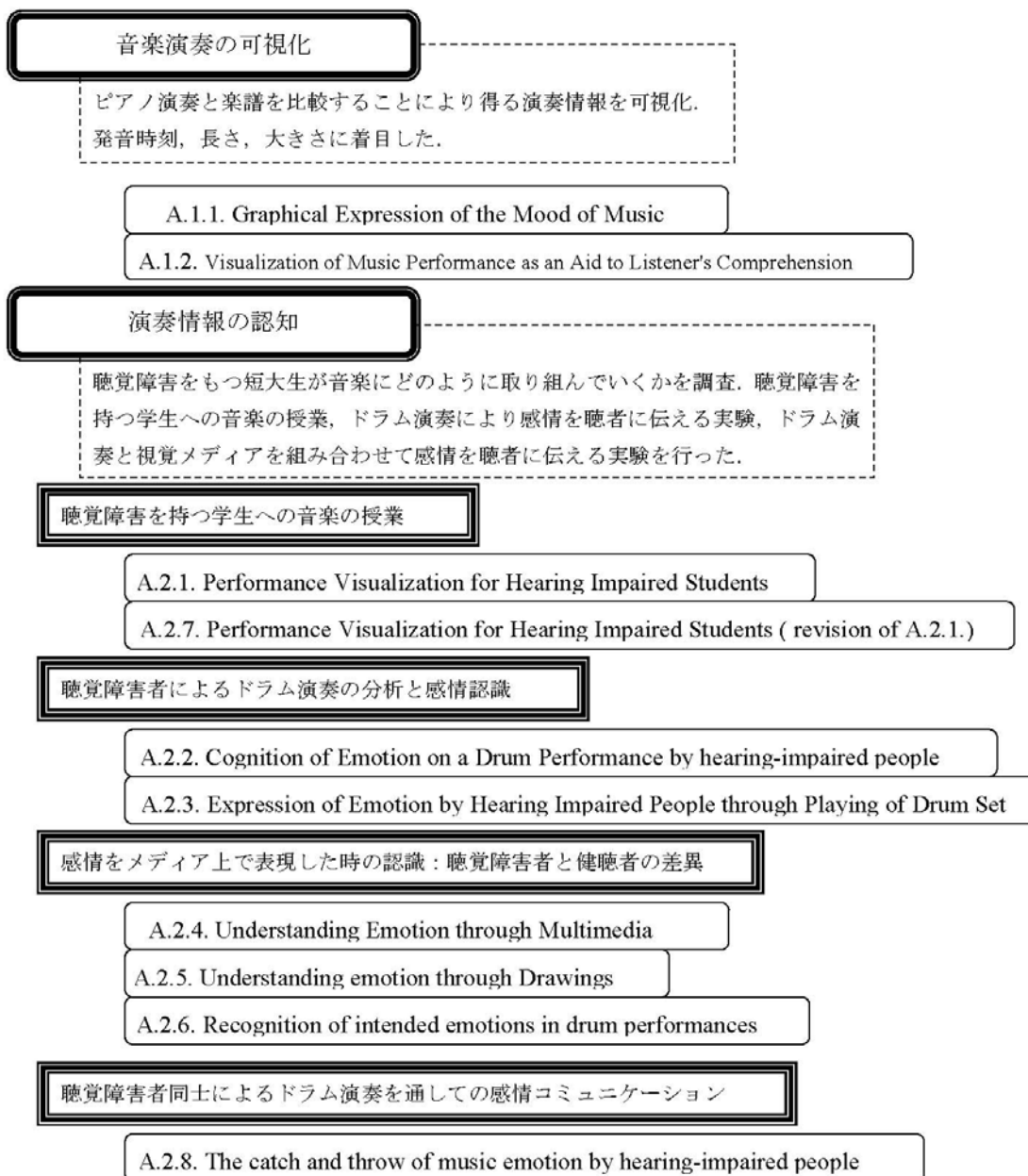


図 1 研究の流れ・テーマと既発表論文（論文名の番号は付の論文番号に対応）

2.3.2. 演奏情報の認知

➤ 聴覚障害を持つ学生への音楽の授業

A.2.1. Performance Visualization for Hearing Impaired Students

A.2.7 Performance Visualization for Hearing-Impaired Students (revision of A.2.1)

筑波技術短期大学（現筑波技術大学）において 1997 年から 6 年間行った「コンピュータ・ミュージック」の授業の総括と，学生がリズム認識を行うのにどのような提示が有効であるかの実験についての論文である．この実験では，音のみが最も有効，リズムの先読みができるものと音を合わせたものが次に有効であるという結果であった．

➤ 聴覚障害者によるドラム演奏の分析と感情認識

A.2.2. Cognition of Emotion on a Drum Performance by hearing-impaired people

A.2.3 で分析したドラム演奏を聴覚障害を持つ学生が聴いてそれぞれの演奏がどの感情を表すと認識するかに関する実験の報告である．喜び・恐怖・怒り・悲しみの各感情が演奏通り認識された割合はそれぞれ 56%，27%，57%，62%であった．分散分析により，恐怖以外の感情は，意図した感情を演奏に認識した割合が，他の 3 感情と認識した割合よりも有意に差があった．

A.2.3. Expression of Emotion by Hearing-Impaired People through Playing of Drum Set

聴覚障害を持つ学生が感情を込めて行った電子ドラム演奏データを重回帰分析と分散分析により行った．怒りの感情では，演奏時間，演奏中のビート数，平均音量が有効，悲しみでは平均音量とビート間隔が有効な変数という結果になった．平均音量と平均ビート間隔は感情の間で有意差が見られた．

➤ 感情をメディア上で表現した時の認識

A.2.4. Understanding Emotion through Multimedia-comparison between hearing-impaired people and people with normal hearing abilities

A.2.5. と A.2.6. で用いた画像と音響データを用い、音響のみ、音響と同じ意図の画像、音響と Microsoft の MediaPlayer が提供する動画エフェクト（アメーバと泉）という4種類の刺激を2つの被験者グループ（健聴学生、聴覚障害者学生）に提示した。分散分析により、恐怖の感情はすべての刺激タイプ、2つの被験者グループのいずれにおいても最も低い認識率であること、いずれの被験者グループにおいても感情の種類において認識に有意差があること、恐怖と他の3感情の間には認識に有意差があること、2つの被験者グループ間に有意差はないことがわかった。また、音響を画像と合わせた刺激が有意に他の刺激よりも高い認識を示したことがわかった。しかし、この場合でも2つのメディアは異なる感情を提示しているように感じたと記した被験者が複数いたことは特記すべきことであろう。聴覚障害者は動画の刺激を好む傾向が見られたが、必ずしも認識率向上にはつながらなかった。

A.2.5. Understanding emotion through drawings-comparison between hearing-impaired people and people with normal hearing abilities

3つの描画者グループ（健聴デザイン専攻学生、聴覚障害デザイン専攻学生、聴覚障害電子工学専攻学生）が描画した感情を意図した単純な抽象的線画をデータとし、3つの被験者グループ（健聴学生、聴覚障害デザイン専攻学生、聴覚障害電子工学専攻学生）が線画に認識する感情についての実験を行った。恐怖を意図した画像は3つの被験者グループにおいて最も低い認識率を示した。分散分析により恐怖を意図した画像は他の三つの感情を意図した画像と認識率において有意差があること、被験者が健聴者のグループの認識率は他の2つのグループと認識率において有意差があること、描画者グループにおいても健聴者による図形とそれ以外の図形では有意差があることが分かった。認識率の高い画像を3つの描画者グループから取り出したところ恐怖以外の3感情については、どのグループの描画も似ていた。

**A.2.6. The cognition of intended emotions for a drum performance:
differences and similarities between hearing-impaired people and people with
normal hearing ability**

3つの演奏者グループ（健聴プロ、健聴アマチュア、聴覚障害者）が感情を意図したドラム演奏したデータの感情の認識を2つの被験者グループ（聴覚障害、健聴）に対して行った。プロによる演奏の認識について聴覚障害者と健聴者に有意差が見られた。それ以外は、特記すべき有意差は聴覚障害者と健聴者の演奏から認識する感情についてなかった。恐怖の認識率が低いことはA.2.2の結果と一致していた。

A.2.8. The catch and throw of music emotion by hearing-impaired people

聴覚障害者2名が感情を意図したドラム演奏をしあい、感情によるコミュニケーションが可能かどうかを確かめる実験を行った。共有する感情から演奏を開始し、2名が交替で演奏、演奏中に受け取った感情から他の感情へ変化させた演奏をすると、その感情をもう1名が認識して演奏を開始する、ということを何回か繰り返し複数のセッションを行った。3割程度の認識間違いがあり、その多くは喜びへの変化とみなしていた。この実験の場合、障害の程度は影響がなかった。

3. まとめ

本研究は聴覚障害者の音楽認知・理解に対する基礎的理解を目指して行ったが、現状では世界的にみてもそのような研究は非常に少ない。さらに、この成果を基に実際に聴覚障害者のための演奏支援を行うという取り組みを目指すものは他にはない。したがって、この研究は大変独創的で、4年間にある程度の成果は得られたものと考えられる。

研究期間中においては、可視化研究は健聴者に対して行ったため、この研究成果を聴覚障害者に対して活用するまでにはいたらなかった。今後は、聴覚障害者が音楽を楽しむことが出来る環境の構築を引き続き目指す。このために、論文A.2.4.で示された動画像による音響情報の保障の可能性をより有意なものにすること、動画像提示のタイミング、楽譜のない音楽の演奏(即興)の理解、音楽的背景の少ない聴覚障害者が打楽器を叩くという状態からどのように音楽を作る状態に持っていくか、など、様々な解決すべき項目がある。

Graphical expression of the mood of music

(c) 2004 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.

Graphical Expression of the Mood of Music

Rumi Hiraga and Noriyuki Matsuda

Abstract—We propose a graphical method that a music performance is intended to create in the minds of the audience. Our graphical approach overcomes the problems associated with verbal labeling. Besides playing a melody and harmony, a music player tries to produce certain feelings in the audience by manipulating tempo, rhythm, articulation, and dynamic changes. Despite the linear nature of music, the produced mood does not necessarily preserve the temporal sequence, and is mentally representable in different forms. As a first approximation to such a representation, we have developed a plane method on which the above mentioned musical expression elements are projected. First, expression elements for all notes in a musical section were derived. They were then arranged according to the importance of notes in consideration of the musical structure.

Index Terms—Information visualization, Musical performance, Mood

I. INTRODUCTION

Musical mood has been left to the listeners' interpretation and has been described in subjective terms. Although it is of interest to elicit their verbal expressions or responses to verbal tags like the Semantic Differential (SD) method, many people encounter great difficulty at correctly translating the elicited mood into words or phrases unless they've been specially trained to do so. This poses a serious problem to the creation of a music database that contains mood as an attribute. The interface to retrieve performances by inputting a mood is different from that of contents searches based on melody.

The piano sonata Op. 27, No. 2, "Moonlight" by L. van Beethoven has been played by many famous pianists, and many listeners appreciate their differences in mood. Even performances by the same pianist sound different depending on the time and place of the performance. If the user wants to pick from the database one of several performances of a specific player that is representative of a certain mood, it is difficult to retrieve the desired data only with the player's name.

We propose a method to visualize musical mood as the first step in research on a new interface for musical performances. Once a mood is visualized for a music database, users are able to retrieve music by browsing the mood figures. The visualized mood is a clue with less ambiguity and subjectivity than verbal tags, because figures are drawn with expressive elements obtained from performance data.

We visualize a musical mood as a snapshot of a performance. Whereas a performance lasts a certain duration, a figure (a snapshot) should not necessarily be larger if a performance is longer, nor should it necessarily visualize performance data following the time order. Since mood is generated not only from a melody but from *all notes*

in the musical section, the figure includes information on all notes. Subjective evaluation of the proposed figure by subjects is a prerequisite for the next step of our research.

II. RELATED WORK

There are two types of music visualization: augmented score and performance visualization. Augmented score is either for composers, to put down their expressive intentions on a musical score [11], or for performers, to assist them in learning a musical piece [13]. In this paper, we restrict music visualization to mean performance visualization, to visualize performance data.

A work by Mazzola [9] and a proposal by Hiraga [5] are used as the complementary feedback of the performance. Complementary feedback with visual data helps players to understand their own performances. From the necessity for those who work on expressive performance, performance is visualized in order to help analyzing performance (Hiraga [3] [4] and Dixon [1]). The unsatisfactory usability of commercially available sequence software systems has led Watanabe [12] and Miyazaki [7] to propose new user interfaces for editing performance data. Foote proposed a checkerboard figure where two musical sections that resemble each other are turned black and white [2].

III. A PERFORMANCE MODEL

A. From a Score to a Performance

Expressive performances go far beyond simple mechanical translations of musical notes on a sheet (a still picture) into a performance (audio data). A musical sheet shows information of each note (the pitch and note value¹) and the relationship among notes (the time order is an example). We call these attributes *performance elements*. Players assemble them into the three essentials of music; melody, rhythm, and harmony. Given the essentials of music on a sheet that are common and have unique meanings to all players, each player instantiates performance elements differently with a *performance plan*. The performance plan is built from a complicated combination of the music essentials with undescriptive factors such as individual experience, knowledge, background or an era. The musical expression consists of tempo changes, articulations, or dynamics changes that we call *expression elements* and these are embodied in a performance (Figure 1).

B. Time Span Tree

The result of music analysis, in other words, the musical structure behind the surface information on a musical sheet, is said to affect building a performance plan. Several music analysis models have been proposed by musi-

R. Hiraga is with Bunkyo Univ., 1100 Namegaya, Chigasaki, Japan, email: rhiraga@shonan.bunkyo.ac.jp

N. Matsuda is with Univ. of Tsukuba, 1-1-1 Tennoh-dai, Tsukuba, Japan, email: mazda@sk.tsukuba.ac.jp

¹ The note value is the quantized duration of a note specified on a sheet.

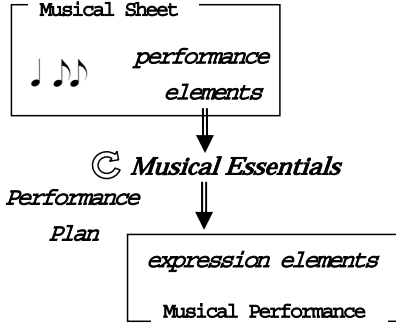


Fig. 1. Performance elements, music essentials, performance plan, and expression elements

cologists [8][10]. The Time Span Tree (TST) [8] evaluates the importance of notes in a musical section and uses the importance to build a tree structure.

A TST is derived by comparing the importance (impressiveness) of neighboring notes. For comparison of two notes \mathcal{N}_i and \mathcal{N}_{i+1} , let's assume that we decide \mathcal{N}_i is more important than \mathcal{N}_{i+1} taking the key, harmony, and rhythm into consideration. We can make a weighted binary tree with two leaves (\mathcal{N}_i and \mathcal{N}_{i+1}), a node on which the weaker leaf depends, and a root that represents the stronger note. By keeping a comparison of all leaves in a musical section, we will have a TST for the section. The TST thus indicates the snapshot of a music that listeners appreciate.

Let's consider which note impresses us the most in the first two measures of Etude Op. 10, No. 3 by F. Chopin (Figure 2). Take a look at the consecutive two notes G#4² and F#4. Both are sixteenth notes in the second half of the first beat in the second measure of the sample score. G#4 is more important than F#4 from the point of beat and key (E major). In this case, G#4 is weighted and we call it the *head* of the TST consisting of G#4 and F#4. Next we compare the head G#4 and another G#4 just before it. Although at the irregular beat position, the G#4 in the previous position becomes the head of the new TST consisting of three notes. In this way, we will see the most impressive note is the last G#4 of the musical section. That note becomes the head of the TST.

Since obtaining a TST has not been automated, we need to analyze each musical section manually. We should also mention that a TST does not give the complete order to all leaves (notes in a section).

IV. VISUALIZING THE PERFORMANCE MOOD

As described in Section I, the performance mood does not depend on the length of a musical section or the order of appearance of notes, and it consists of all notes. We visualize the mood using similar sized squares whose textures represent the mood.

By looking at the sample score, we see that the left hand part has the different rhythm from the right hand part



Fig. 2. A sample score: The first two measures of Etude Op. 10, No. 3 by F. Chopin and a part of its TST.

(Figure 2). A way to involve all notes is the serialization. Serialization also gives the notes the order of importance in the musical section. TST is used to generate information for the serialization.

Each performed note has expression elements. Taking the degree or level of each expression element into account, a color is assigned to each note. A small colored square for the note represents a fragment of the musical mood of the section. By using the serialization information, small squares are arranged along the zigzag line into a bigger square for representing the mood of the section. The bigger squares are the same size for all musical sections however long or many notes they include. Namely, we condense all expression elements into a figure. If each colored square is displayed on a score in place of a note, the longer musical section lengthens the figure and a score with sparse notes always gives the impression of weakness or sparseness regardless of whether it has more various impressions for the performance.

A. Process of visualizing Mood

The visualization process consists of the following steps.

1. *Obtain the value of expression elements.* First, calculate the deviation of the performance elements of each note by comparing performed information with the score information. The expression elements are then calculated from the deviation of performance elements; for example, the local tempo of a note \mathcal{N}_i is derived by subtracting the expected onset time on a musical sheet from the actual onset time, and then dividing it by the note value of the previous note, \mathcal{N}_{i-1} , for regulation. The details of the calculation are described in [6].
2. *Get a colored small square for each note using the values of the expression elements.* First, we assign a color from a two-dimensional colormap (Figure 3) to each pair of expression elements. For example, if the tempo value is bigger and the articulation is smaller for a note, a reddish color is selected for the note. In

² We follow the style that calls the note C at the center as C4.

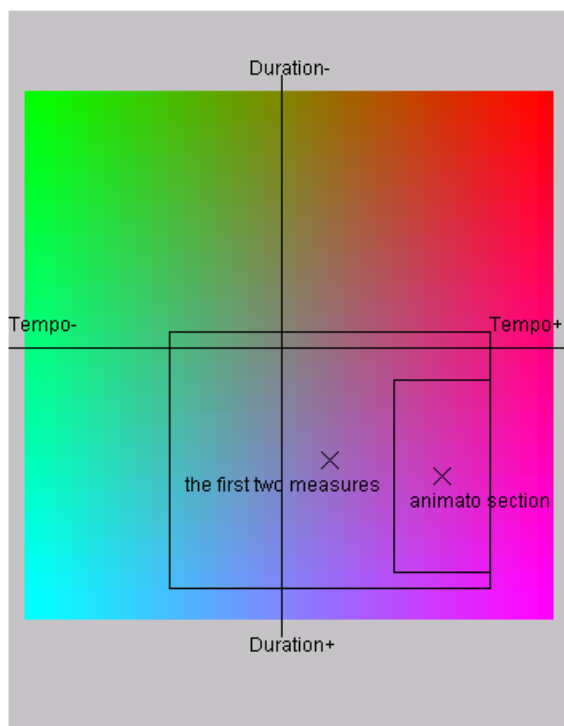


Fig. 3. Colormap. The average and standard deviation of two musical sections are indicated.

this way, each note is represented as a small colored square. The size of the square is decided by the number of notes in a musical section in order to make the size of the figure the same, independent of the length and the number of notes in the section.

3. *Serialize the notes in a multi-voice sections.* First, obtain a TST using a musical sheet. Then, starting from the head of a section, compare two nodes in the level of the next importance manually to give the complete order, since leaves in the TST (notes in the musical section) are in the partial order.
4. *Arrange expression elements into a bigger square.* Like the coefficients of DCT (Discrete Cosine Transform) are arranged on a square by doing a zigzag scan (Figure 4), the serialized expression elements represented as a small square are arranged into a bigger square. In this square, the most impressive note comes at the left-top corner followed by the note of the next highest importance.

In this way, we will be able to perceive the musical mood at a glance with the multi-colored square texture.

B. Examples

Here we show example figures for a performance of Etude Op. 10, No. 3 by F. Chopin. A professional pianist played the Yamaha Piano Player, which has a MIDI³ recorder on it. The expression elements are calculated using the recorded performance data in MIDI format.

³ Musical Instrument Digital Interface, a digital data format of performances.

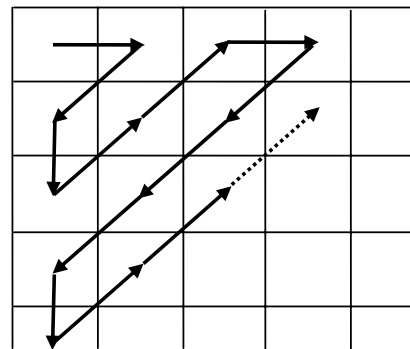


Fig. 4. The colored squares are arranged along the zigzag line.

We have certain options when preparing a colormap. In the example figures, we used the colormap with the following calculation (Figure 3).

```

for (int k = 0; k <= 32; k++) {
  for (int j = 0; j <= 32; j++) {
    g.setColor(new Color(tab[j], tab[32 - j], tab[k]));
    g.fillRect(10 + j * 10, 50 + 10 * k, 10, 10);
  }
}
    
```

A color is assigned to each note using the value of two of the expression elements. In the example figures, these elements are tempo change and articulation.

As described in Section IV-A, we have to serialize notes using the partial order TST. The strategy for serialization in this example is to compare nodes for the melody first, for the bass part, then the tenor, and finally the alto part.

In order to make a square using notes in a musical section, the number of notes in a section should be adjusted to a square number. We select the biggest square number that is not bigger than the number of notes. Namely, if the number of notes in a section is $\#(\mathcal{N} \in Sec)$, we need a square number \mathcal{S} that satisfies $\mathcal{S} \leq \#(\mathcal{N} \in Sec)$. This means that some notes of the less importance in the section will not be shown. For example, if a section involves 52 notes, a $7*7$ square is used for the visualization. If there are 26 notes, a $5*5$ square is shown.

The first example shows the first two measures of Etude No. 10, Op. 3. It is shown in a $6*6$ square (Figure 5). The small square at the left-top shows the expression elements of the most important note (the last G#4 in the second measure in Figure 2). The small colored squares for each note are arranged along the zigzag line. The second example shows a different section in the same piece where *animato* is indicated on the musical sheet (Figure 6).

The first example consists of many colors while the *animato* section is colored red. If we wished to verbally express the impression that we obtain from these two figures, we could say that the first two-measure section is played in *rubato* (by observing the color variation), while the *animato* section is played hastily. The averages and standard deviations of each performance are shown on Figure 3, specified by “X”s and rectangles respectively.

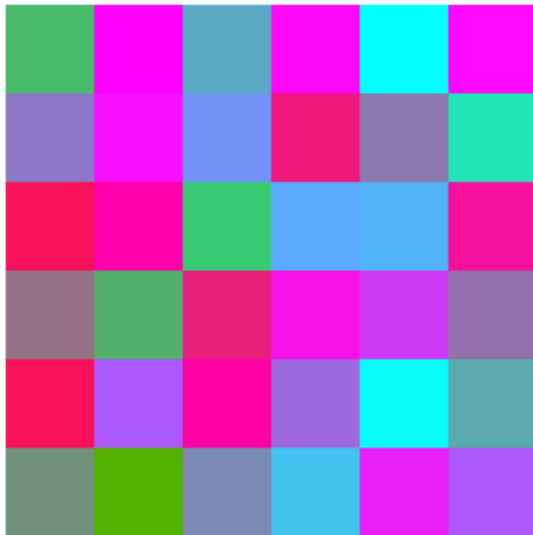


Fig. 5. Example visualization of the performance mood (1): the first two measures of Etude 10-3.

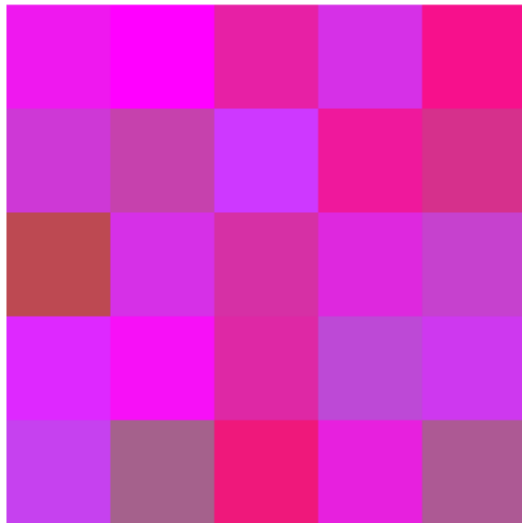


Fig. 6. Example visualization of the performance mood (2): two beats of the animato section in Etude 10-3.

V. DISCUSSION

Our new method for visualizing the mood of a performance visualizes multi-voice performances at a glance by using the information about the importance of notes. Moreover by releasing a musical performance from its temporal sequence, musical sections can be shown as squares of the same size regardless of how long or how many parts they have. Although we verbally explained two figures (Figures 5 and 6) in Section IV-B, our intention is to give a non-verbal representation of musical impression.

We do not insist on the visualization figure shown in the previous section being the best for showing the mood of music. We still need to evaluate the figure before we can attempt to resolve the following issues.

- Use of colors: The impression of color is different depending on the person. Therefore it may not be suitable to give a performance impression with colors. The way of making the colormap should also be well considered.
- Performance elements to visualize: When a third expression element, say dynamics change, is shown in the same figure, there are several possibilities to extend the visualization of the current figure. One is to provide a colormap on each surface of a cube, say x-axis for tempo, y-axis for duration, and z-axis for dynamics. Each surface of the colormap cube shows a combination of two performance elements. A performance could be shown on a cube, the three visible faces, showing the atmosphere for the tempo change and articulation like in Figure 5, the atmosphere for the tempo change and dynamics change, and atmosphere for the dynamics change and articulation.
- Categorization of the mood of music. In order to understand a performance's mood at a glance, we need to make many more performance figures and categorize them for the future user interface.

VI. ACKNOWLEDGMENTS

We show our gratitude to Dr. Hirata for his valuable advice in TST. This research is supported by The Ministry of Education, Culture, Sports, Science and Technology through a Grant-in-Aid for Exploratory Research, No. 13878065.

REFERENCES

- [1] S. Dixon, W. Goebel, and G. Widmer. The performance worm: Real time visualization of expression based on langner's tempo-loudness animation. In *Proc. of ICMC*, pages 361–364. ICMA, 2002.
- [2] J. Foote. Visualizing music and audio using self-similarity. In *Proc. of ACM MultiMedia*, pages 77–80. ACM, 1999.
- [3] R. Hiraga. Case study: A look of performance expression. In *Proceedings of IEEE Visualization*. IEEE, 2002.
- [4] R. Hiraga, S. Igarashi, and Y. Matsuura. Visualized music expression in an object-oriented environment. In *Proc. of ICMC*, pages 483–486. ICMA, 1996.
- [5] R. Hiraga and M. Kawashima. Performance visualization for hearing impaired students—a report of the preliminary experiment. In *Proceedings of EISTA*. IIS, 2004.
- [6] R. Hiraga and N. Matsuda. Visualization of music performance as an aid to listeners. In *Proceedings of AVI*, 2004.
- [7] R. Hiraga, R. Miyazaki, and I. Fujishiro. Performance visualization—a new challenge to music through visualization. In *Proc. of ACM MultiMedia*, pages 239–242. ACM, 2002.
- [8] F. Lerdahl and R. Jackendoff. *A Generative Theory of Tonal Music*. The MIT Press, 1996.
- [9] S. Müller and G. Mazzola. The extraction of expressive shaping in performance. *Computer Music Journal*, 27(1):47–58, 2003.
- [10] E. Narmour. *The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model*. The University of Chicago Press, 1992.
- [11] D. Oppenheim. Compositional tools for adding expression to music. In *Proc. of ICMC*, pages 223–226. ICMA, 1992.
- [12] A. Watanabe and I. Fujishiro. tutti: A 3d interactive interface for browsing and editing sound data. In *Proc. of The 9th Workshop on Interactive System and Software*. Japan Society for Software Science and Technology, 2001.
- [13] F. Watanabe, R. Hiraga, and I. Fujishiro. Brass: Visualizing scores for assisting music learning. In *Proc. of ICMC*, pages 107–114. ICMA, 2003.

Visualization of music performance as an aid to listener's comprehension

©ACM, 2004. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in AVI '04 Proceedings of the working conference on Advanced visual interfaces, 2004, 103-106

<http://doi.acm.org/10.1145/989863.989878>

Visualization of Music Performance as an Aid to Listener's Comprehension

Rumi Hiraga
Bunkyo University
1100 Namegaya, Chigasaki
253-85500, Japan

rhiraga@shonan.bunkyo.ac.jp

Noriyuki Matsuda
University of Tsukuba
1-1-1 Tennoh-dai, Tsukuba
305-8573, Japan

matsuda@sk.tsukuba.ac.jp

ABSTRACT

We present a new method for visualizing musical expressions with a special focus on the three major elements of tempo change, dynamics change, and articulation. We have represented tempo change as a horizontal interval delimited by vertical lines, while dynamics change and articulation within the interval are represented by the height and width of a bar, respectively. Then we grouped local expression into several groups by k-means clustering based on the values of the elements. The resulting groups represented the emotional expression in a performance that is controlled by the rhythmic and melodic structure, which controls the gray scale of the graphical components. We ran a pilot experiment to test the effectiveness of our method using two matching tasks and a questionnaire. In the first task, we used the same section of music, played by two different interpretations, while in the second task, two different sections of a performance were used. The results of the test seem to support the present approach, although there is still room for further improvement that will reflect the subtleties in performance.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces—*Graphical user interface*; J.5 [Arts and Humanities]: Performing arts

1. INTRODUCTION

We started a project to use music visualization to enhance listening comprehension. In this paper, we propose a new visualization method that shows a “snapshot” of musical expression in order to understand the performance more concretely. In this method, we propose the elements to visualize, the way to estimate them, and the relationship between each element and the figured component.

An important thing to keep it in mind is that expressive performance goes far beyond simple mechanical translations

of notes on a sheet (a still picture) to the actual performance (audio data). It is easy to conceptualize this point if you compare sentence readings by both synthesized and natural human voices. While the former is a flat translations of (verbal) signs, the latter adds depth to them by varying tempo, accent, pauses, and so forth.

It is the depth of the music performance that we attempt to visualize through the primary elements of tempo, articulation, and dynamics. The distributions of these elements throughout the entire performance provide the basis for the secondary element that uses the dependency between these elements to create a picture. Introducing the secondary element as a visual element is a new idea. Our present goal is to develop a cross-modal comprehension model of music performance in both audio and visual forms. An intermediate level piano student can easily recognize the differences between his/her teacher's style and his/her own by comparing the corresponding graphic presentations. In the first practical application of our method, we conducted experiments in which participants were asked to match the audio stimuli to the graphic representations.

2. RELATED WORKS

There are figures that visualize musical performance in MIDI data format¹ in sequence software systems². One example is the piano-roll figure. Since it shows only a limited number of parameter values, it is not easy to visualize musical expression from the figure.

From the perspective of information visualization, there are two types of music visualization, augmented score, and performance visualization. Augmented scores are intended to assist composers in documenting expressive intentions on a musical score [10] or to assist performers in learning a piece of music [12]. Performance visualization technology was originally developed out of necessity to assist researchers who work on music performances (Hiraga [3] [4] or Dixon [2]), and was highly analytical for that reason. A work by Mazzola [8] is used as the complementary performance feedback. Because of the difficulties in using product sequence software systems, Watanabe [11] and Miyazaki [5] proposed a new user interface to edit performance data.

¹Musical Instrument Digital Interface, a digital format for performance data.

²The purpose of sequence software systems is to create musical performance.

3. VISUALIZATION TO ASSIST IN UNDERSTANDING PERFORMANCE

Performance visualization should clarify the unspecified expression information on a musical score because it is added by the musician at the time of performance. Two issues relevant to our visualization are choosing and arranging expression elements.

3.1 Choosing Visualization Parameters

If we describe a performance in physical terms, such as frequency, the absolute moment when a note is performed, or decibel, we are not able to understand the performance in the cognitive terms described by melody, rhythm, and phrasing. A MIDI performance, using terms such as onset time³, offset time⁴, and velocity⁵, does not express our emotions in an expressive performance.

We chose quantifiable local expression elements, consisting of *tempo change*, *articulation*, and *dynamics change* that can be appreciated qualitatively during a performance, because they have an affinity with music cognition. These are the basic elements that influence the human emotion of the listener, and are called expressive cues by Bresin [1]. We call them primary elements.

Since listeners appreciate the grouping structure in performances as described by Lerdahl [7] or Narmour [9], it is desirable for performance visualization to reflect the grouping structure. Using the dependency between the primary expression elements, we manipulated them as a set to represent the degree of expression. We grouped the primary expression elements by k-means clustering into several degrees of movement, based on the values of the elements. Each group reflects the degree of expressive movement controlled by the rhythmic and melodic structure. The grouping is the secondary element.

We extracted notes from a melody in the MIDI formatted performance data so that no two notes would start at the same time. We assigned an integer number to each of the notes, according to the time of their appearance. We wrote the i_{th} note in a performance as $\mathcal{N}_{p,i}$, with the attributes of starting time ($S_{p,i}$), ending time ($E_{p,i}$), and velocity ($V_{p,i}$). Since $\mathcal{N}_{p,i}$ is an instantiation of a note on a score, there is a corresponding note on the score. For each note played, $\mathcal{N}_{p,i}$, $\mathcal{N}_{sco,i}$ represents the note on a score. The note value (the duration of a note) is shown on the sheet as a quarter note or eighth note. We wrote the note value of $\mathcal{N}_{sco,i}$ as $Val_{sco,i}$ and the starting time as $S_{sco,i}$.

We calculated the local tempo change, local articulation, and local dynamics change. We wrote them as $Tempo_{p,i}$, $Artc_{p,i}$, and $Dy_{p,i}$ respectively for the i_{th} note. Because $S_{p,i}$, $E_{p,i}$, and $V_{p,i}$ are represented as relative values that are independent from tempo⁶ and to remove the note value

³Onset time is the moment a note is played.

⁴Offset time is the moment a note finishes playing.

⁵Because the velocity to play a keyboard affects the dynamics, the term velocity is used for MIDI dynamics. The velocity value is from 0 to 127. These numbers do not represent specific decibels.

⁶By only changing the tempo indication in MIDI data, we

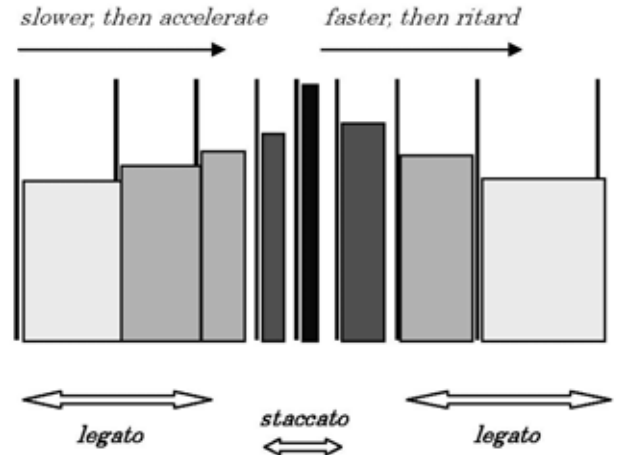


Figure 1: An explanatory figure of performance visualization

effect, denominators were used for regulation. The following is a description of the three primary elements for the i_{th} note and the secondary element.

Tempo : $Tempo_{p,i} = (S_{p,i} - S_{sco,i})/Val_{sco,i-1}$

If a performance follows a score precisely, then $Tempo_{p,i} = 0$, or in other words, the local tempo deviation is zero. If it is played faster than expected, $Tempo_{p,i} < 0$, otherwise $Tempo_{p,i} > 0$. The performance accelerates when we obtain $Tempo_{p,*} < 0$ for contiguous notes, the performance accelerates. $Tempo_{p,*} > 0$ for contiguous notes means *ritardando*, and otherwise, possibly in *rubato*.

Articulation : $Artc_{p,i} = (E_{p,i} - S_{p,i+1})/Val_{sco,i}$

If the i_{th} note is sustained after the $i + 1_{th}$ note starts, $Artc_{p,i} > 0$, the i_{th} note is played in *legato*. If $Artc_{p,i} < -0.5$, then the note is played in *staccato*.

Dynamics change : $Dy_{p,i} = (V_{p,i} - V_{p,i-1})/V_{p,i-1}$

If $\mathcal{N}_{p,i}$ is played softer than $\mathcal{N}_{p,i-1}$ (local *diminuendo*), then $Dy_{p,i} < 0$, otherwise it is played louder (local *crescendo*).

Degrees of change : We clustered the sets of the three expression elements $\mathcal{E}_{p,i} = \{Tempo_{p,i}, Artc_{p,i}, Dy_{p,i}\}$ for $\mathcal{N}_{p,i}$ into four groups according to the result of the simple k-means clustering algorithm [6]. Within the three-dimensional space for tempo, articulation, and dynamics change, clustering calculates the distance from the no expression point. $\mathcal{E}_{p,*}$ is grouped according to the distance from the origin. If $\mathcal{E}_{p,i}$ is in the farther group, then $\mathcal{N}_{p,i}$ is played with more expression differences than $\mathcal{N}_{p,i-1}$.

can replay the performance data by stretching or shrinking along time.

A.1.2. Visualization of Music Performance as an Aid to Listener's Comprehension

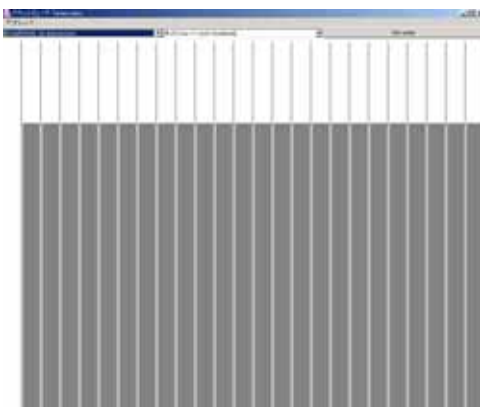


Figure 2: Visualization of a performance that follows a score precisely.



Figure 3: Visualization of a live performance (Beethoven's piano sonata, Op. 13 "Pathetique").

3.2 Arranging Parameters for Visualization

The three primary elements for expression are mapped on a two-dimensional graph, where the x-axis shows the time and the y-axis shows the relative dynamics change of $\mathcal{N}_{p,i}$ (Figure 1). A vertical bar indicates the start of a note. The space between the two vertical bars is assigned to the note. Tempo is shown as the interval between two vertical bars. We do not use absolute time or tick⁷. In a figure, each $\mathcal{N}_{p,i}$ is given a unit interval that varies according to $Tempo_{p,i}$. If the interval between two vertical bars is narrow, then the local tempo accelerates, otherwise it is in ritard. A rectangle between two bars shows the articulation and dynamics change of $\mathcal{N}_{p,i}$ in the width and height. If $\mathcal{N}_{p,i}$ is played in *legato*, the width is wider than the interval. If $\mathcal{N}_{p,i}$ is played louder than $\mathcal{N}_{p,i-1}$, the height is taller than the previous rectangle.

The secondary element decides the gray scale of a bar and rectangle. Each clustered group is assigned a gray scale. The most expressive group is given the darkest value.

⁷Tick is a unit to count relative time in MIDI data format. Usually a quarter note is assigned 480 ticks.

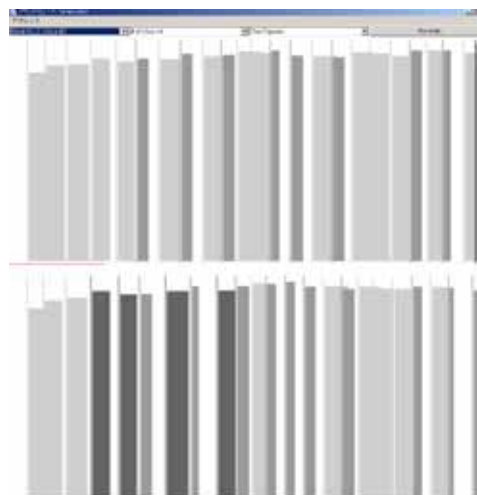


Figure 4: Differences in expression of two performances of the same piece (from Chopin's Etude Op. 10, No. 3).

3.3 Examples

When a performance follows a score precisely, then all the vertical intervals and rectangles are the same (Figure 2). The figure appears the same for any musical pieces regardless of the different rhythms and melodies on their scores.

When we listen to the Beethoven piano sonata, Op. 13 "Pathetique", we are impressed by several notes. With the figure, we are able to visualize the notes that have the biggest impact on us. In addition to the change of tempo, articulation, and dynamics changes, we can see the repeating patterns in the figure that are emphasized using the different gray scales (Figure 3). The darker rectangle means that a note has a stronger expressive movement. The repeating pattern shows us the player's phrasing.

4. EXPERIMENT

We ran a pilot experiment to test the effectiveness of our method using two matching tasks and a questionnaire. All subjects had studied the piano for more than five years. The tasks consisted of matching auditory records to their graphical displays. In the first task, the same section of music (Chopin's Etude Op. 10, No. 3) was played with two different interpretations. In the second task, two different sections of Beethoven's "Pathetique" were played by a single pianist. In both tasks, the subjects looked at the two figures while listening to their corresponding performances, then matched the figures and performances. They were shown a sample performance and its figure with an explanation of the expression elements and visual components before the experiment.

In the first task, the two performances resembled each other (Figure 4). We can see the resemblance especially in the articulation and dynamics changes in the figures. In this task, the subjects had difficulty in matching the figures and performances. One reason was the lack of clues in the figure that would help them locate the point to watch while listening to a performance.

A.1.2. Visualization of Music Performance as an Aid to Listener's Comprehension



Figure 5: Matching performance to its figure (from two sections Beethoven's "Pathétique").

In the second task, the upper figure shows the first four measures of the sonata, while the lower shows the second four measures (Figure 5). Since the two sections have different musical meanings (the upper in a minor chord while the lower was in a major chord) and the expression differences were well reflected in the figure, the subjects were able to distinguish the figures for each section.

5. DISCUSSION

Considering the early development stage of our research, we were pleased with our results for the expressive elements. We are encouraged by the support for our present approach and realize the need for more elements that will enhance the listeners' understanding and appreciation of music. Also, the questionnaire responses indicated a need to extend the method to incorporate pitch and other elements.

Performance visualization shows great potential for several applications.

- Learning assistance for music performance.
The difficulty in clarifying the subtle expression differences in a performance should be resolved with the visual clues for making a connection between the figure and the performance. However, more information on pitch and timing are required.
- Animated interior reflecting music performance.
Music and visual effects have been linked to arts like opera and ballet. Many animations also have musical accompaniment. Together, music and visual effects can greatly enhance the listeners' enjoyment and emotional responses. Our figure approach, if used as a kind of animated wallpaper, will use information from the actual performance to amplify the listeners' emotions.
- Visual music data mining by the mood of the performance.
Current research has enabled music retrieval by contents. This means that we can find a piece of music

by inputting a melody, and data mining for a piece of music that closely resembles it. Currently, if we try to retrieve a piece of music by its mood, we need to prepare tagged data. Visualizing performance expression frees us from retrieving by tags containing subjective terms such as "warmly performed" or "solemn performance". With the mood shown on a visualized performance figure, we will be able to access music reflective of any atmosphere we desire.

Acknowledgements

This research is supported by The Ministry of Education, Culture, Sports, Science and Technology through a Grant-in-Aid for Exploratory Research, No. 13878065.

6. REFERENCES

- [1] R. Bresin and A. Friberg. Emotional coloring of computer-controlled music performances. *Computer Music Journal*, 24(4):44–63, 2000.
- [2] S. Dixon, W. Goebel, and G. Widmer. The performance worm: Real time visualization of expression based on langner's tempo-loudness animation. In *Proc. of ICMC*, pages 361–364. ICMA, 2002.
- [3] R. Hiraga. Case study: A look of performance expression. In *Proceedings of IEEE Visualization*. IEEE, 2002.
- [4] R. Hiraga, S. Igarashi, and Y. Matsuura. Visualized music expression in an object-oriented environment. In *Proc. of ICMC*, pages 483–486. ICMA, 1996.
- [5] R. Hiraga, R. Miyazaki, and I. Fujishiro. Performance visualization—a new challenge to music through visualization. In *Proc. of ACM MultiMedia*, pages 239–242. ACM, 2002.
- [6] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu. The analysis of a simple k-means clustering algorithm. In *Symposium on Computational Geometry*, pages 100–109, 2000.
- [7] F. Lerdahl and R. Jackendoff. *A Generative Theory of Tonal Music*. The MIT Press, 1996.
- [8] S. Müller and G. Mazzola. The extraction of expressive shaping in performance. *Computer Music Journal*, 27(1):47–58, 2003.
- [9] E. Narmour. *The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model*. The University of Chicago Press, 1992.
- [10] D. Oppenheim. Compositional tools for adding expression to music. In *Proc. of ICMC*, pages 223–226. ICMA, 1992.
- [11] A. Watanabe and I. Fujishiro. tutti: A 3d interactive interface for browsing and editing sound data. In *Proc. of The 9th Workshop on Interactive System and Software*. Japan Society for Software Science and Technology, 2001.
- [12] F. Watanabe, R. Hiraga, and I. Fujishiro. Brass: Visualizing scores for assisting music learning. In *Proc. of ICMC*, pages 107–114. International Computer Music Association, 2003.

Performance Visualization for Hearing Impaired Students

Rumi Hiraga
Bunkyo University
1100 Namegaya, Chigasaki
253-85500, Japan

rhiraga@shonan.bunkyo.ac.jp

Mitsuo Kawashima
Tsukuba College of Technology
4-3-15, Amakubo, Tsukuba
305-0005, Japan

kawasima@a.tsukuba-tech.ac.jp

ABSTRACT

We have been teaching computer music to hearing impaired students of Tsukuba College of Technology for six years. Although students have hearing difficulties, almost all of them show an interest in music. Thus, this has been a challenging class to turn their weakness into enjoyment. We thought that performance visualization is a good method for them to keep their interest in music and try cooperative performances with others. In this paper, we describe our computer music class and the result of our preliminary experiment on the effectiveness of visual assistance. Though it was not a complete experiment with a sufficient number of subjects, the result showed that the show-ahead and selected-note-only types of performance visualization were necessary according to the purpose of the visual aid.

Keywords

Hearing Impaired, Computer Music, Music Performance, Visual Feedback

1. INTRODUCTION

We have been teaching computer music to hearing impaired students of Tsukuba College of Technology (TCT) for six years. Students with the hearing impairments of more than 100 decibels are qualified to enter the college and get a quasi-bachelor degree in three years. They learn architecture, design, computer software, or computer hardware according to their major to obtain useful skills. This style resembles that of Gallaudet University and the National Technical Institute for the Deaf at the Rochester Institute of Technology (NTID).

There are many professional musicians with visual impairments, moreover, there are several activities to assist those people with computer software such as WEDELMusic [12]. Though it is not surprising that there are very few professional musicians with hearing impairments, the number of them is not zero. Some of them are talented deaf musicians, like Evelyn Glennie, a famous percussion soloist, who even has absolute pitch.

The computer music class is open to students of all specialties but mainly those of the computer hardware course have taken the class. This is not a required subject. Not necessarily all the professors at the college agree on the importance of the class. On the other hand, we came to know that not a small number of students have an interest in music, independent of the degrees of their handicap and personal experience with music. Thus given the computer assistance for them to understand and enjoy music, their quality of life (QOL) is considered to be improved. We thought performance visualization would be a good method for such assistance. Since the research of performance visualization is not a mature area and currently there is no suitable user

interface to assist students, we need a good performance visualization system for them. In order to design and build such a system, we conducted a preliminary experiment on cooperative musical performance using visual assistance.

2. COMPUTER MUSIC CLASS

We set the purpose of the computer music class to allow students to understand and enjoy music in order to broaden their interest [5]. In other words, the class was more music oriented (and amusement oriented), not computer oriented. Considering that the class should meet the requirements of the college, especially for the computer hardware course, the purpose above is not necessarily appropriate. The reason for setting such a purpose is to get rid of the difficulty of keeping students' interest, especially in an area that they have not experienced much in their lives. If we start teaching them from computer perspective such as the format of MIDI¹ or the structure of synthesizers, they will have conversations in sign language, or even worse no students may register for the class.

Making students continue to move their bodies with music is the most effective way to keep the class active. Thus, the computer has been used as a tool for assisting them in enjoying music in the class, not as a tool with which to develop new computer music software or hardware systems.

Materials

Because it was not possible for teachers who did not receive special education in music to teach conventional acoustic musical instruments to students, we benefited from the newly developed MIDI instruments. Furthermore, we were able to connect several machines and instruments with MIDI. The following are the hardware and software systems we used in the class.

- Miburi R2 (Yamaha): A MIDI instrument² with sensors. Sensors are attached to a shirt which a performer puts on. When a performer moves his/her elbows, shoulders, wrists, and legs, sound that corresponds to the position and its movement is generated from the sound generator of Miburi. The sound generator provides several drum sets, tonal sound colors, and SFX sound (murmuring sound of a stream, the sound of gun fire, a bird song, etc.). The good points in using Miburi for students were as follows:

¹Musical Instrument Digital Interface, a digital format for performance data.

²A MIDI instrument generates MIDI data when a player plays it. It has a MIDI terminal to connect with another MIDI instrument or a PC. It needs a sound generator either inside or outside the instrument.

A.2.1. Performance Visualization for Hearing Impaired Students

- With simple body movement, you can generate music.
 - It is a new instrument in which playing methods are not difficult and not established.
 - Miburi performers can communicate by looking at each other's movement.
 - Since MIDI data is generated by playing the Miburi, their movement is reflected synchronously to visualization if systems are connected. Through the visualization, students understand their movement and its result as music.
- XGWorks (Yamaha): A sequence software system to make performance data in MIDI.
 - VISISounder (NEC): An animation software system whose action is controlled by MIDI data. It prepares several scenes and animation characters beforehand. For example, a frog at a specific position in the display jumps when a sound "C" comes, while another frog jumps with "G." Using this software, students were directly able to feel their performance with Miburi through visualization. They liked it very much.
 - MIDI drum set and MIDI keyboard (Yamaha): MIDI instruments.
 - Music table (Yamaha): A MIDI instrument, originally designed for music therapy for elderly people. Pads are arranged on the top of the table on which people pat. There is a guide lamp for indicating the beat.

Though we tried an actuator that is used inside a speaker system for the haptic feedback purpose, it was not suitable to use because it heats up as sound was generated.

An unfortunate thing in using these products is that some of them had a short life. In the past six years when we taught the class, Miburi and VISISounder, which were the most suitable materials for the class, disappeared from the market. Although there are several other MIDI instruments and animation systems with MIDI data at the research level, products are more reliable and end user oriented.

Students' presentation

The class is held in a school term. There were ten or eleven weeks in a term. Every year we asked students to make a musical presentation at the final class. The following is an excerpt from the list of students presentation.

- A dramatic play using Miburi. Accompanied by SFX sounds, a student played out her daily life in sign language. For example, the barking of the dog was heard accompanying the sign language for a dog made by wrist movement.
- A music performance using Miburi. With a tonal sound, a student played the "Do-Re-Mi song." Her performance controlled characters of VISISounder.
- A Japanese Taiko (drum) performance using Miburi. Though it is a completely virtual performance, the change of drum sets was musically very effective. Usually a Taiko player uses one to three Taikos in an actual play, a player with Miburi can use many more types of Taiko as if all of them are around him/her.
- Samba performance using a Music table and a drum set. Seven students played three different rhythm patterns that cooperatively made Samba percussion performances (Baticuada). One student stood up and

played as a conductor by performing a basic rhythm pattern. Playing Baticuada gave students the sense of unity in music.

- Some students used sequence software in order to perform accompaniment music for Karaoke. They sang using the sign language accompanied by the music.

After their presentations, many students indicated on a questionnaire that they would like to play in an ensemble or they enjoyed playing with other students.

3. RELATED WORK

Although there are several studies of aiding visually handicapped people in their musical activities, there are very few for hearing impaired people. We conducted the experiment described in Section 4 from the viewpoint of performance visualization. Thus, in this section, we describe research on performance visualization.

Sobieczky visualized the consonance of a chord on a diagram based on a roughness curve [10]. Hiraga proposed using simple figures to help users analyze and understand musical performances [3][4][6]. Smith proposed a mapping from MIDI parameters to 3D graphics [9]. Foote's checkerboard type figure [1] shows the resemblance among performed notes based on the data of a musical performance. 3D performance visualization interface is proposed for users to browse and generate music using a rich set of functions [7][11].

These performance visualization works have different purposes such as for performance analysis and sequencing. So far, there has been no work for cooperative musical performance.

4. EXPERIMENT

Outline

In order to determine a more suitable visualization interface for performance feedback to support cooperative musical performances by hearing impaired people, we investigated the characteristics of animated performance visualization proposed by commercial systems and a prototype system by a student. The investigation was done by a usability test of each performance visualization.

The purpose of the test was to see the playing timing of each subject with a guided animation that is controlled by MIDI data of a model performance. Namely, subjects played a MIDI instrument looking at the animation and their performances were recorded, then we compared the performances with a model performance. The time differences were calculated between the onset time³ and of subjects' performances and the model performance.

Subjects

Three students (call them SA, SB, and SC) and a technical staff member (call her SS) were the subjects of the experiment. Students were in a sense exceptional among all students regarding their musical experience because two of them were members of a pop music club and had performance experiences, and the other had been learning play the

³Onset time is the moment a note is played by a keyboard or a drum. It is the time of a MIDI message of "Note On" is generated. The message includes the note number (pitch) and the velocity (volume) of the note on. We can see from the note number which drum pad is patted or which key on a keyboard is played.

A.2.1. Performance Visualization for Hearing Impaired Students



Figure 1: Rhythm A and B

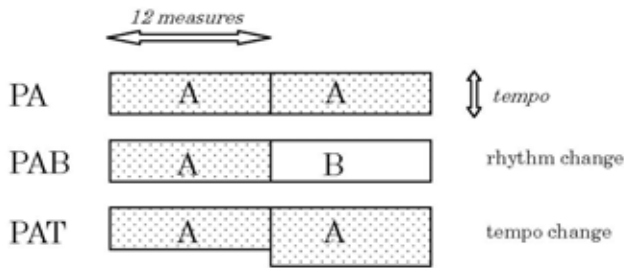


Figure 2: Three types of model performance: PA, PAB, and PAT

piano for six years. They were assigned different instruments and tried to play cooperatively with a model performance using feedback.

Model performances

We used two rhythm patterns, A and B (Figure 1), then prepared three types of model performance, PA, PAB, and PAT, by combining them (Figure 2). PA repeats rhythm A for twenty-four measures with tempo $MM=108^4$. PAB repeats rhythm A for twelve measures then changes rhythm to B with the constant tempo $MM=108$. PAT repeats rhythm A for twenty-four measures with tempo $MM=108$ for the first half, then with tempo $MM=140$ for the second half.

Feedback types

The experiment used four types of feedback: three types of visual feedback and one type of sound feedback. These feedback types were exclusively given to subjects. They were as follows.

1. VISISounder. We used a scene that clearly showed the difference among performed notes by the movement of characters (either a monkey or frogs) (Figure 3). A monkey in the center corresponds to the performance of a model performance and frogs to those by subjects. Characters pop up when an instrument is played. Since a frog character was assigned to individual subjects, we could distinguish subjects through the animation.
2. XGWorks. Although XGWorks has several visualization forms for performance, we used a “drum window” (Figure 4). In the drum window, each line corresponds to a type of drum, such as a Conga. When a rhythm changes or a tempo changes, a drum used by a model performance changes accordingly. A cursor indicated the place of a model performance on the display.

A big difference in the visualized performance on XGWorks from the other two types of visual feedback is that subjects are able to predict the rhythm (show-ahead feedback). In PAB, the rhythm change from the

⁴ $MM=108$ means that there are a hundred and eight beats in a minute, namely a beat takes 0.556 ($60/108=0.556$) second. The larger the number, the faster the tempo.

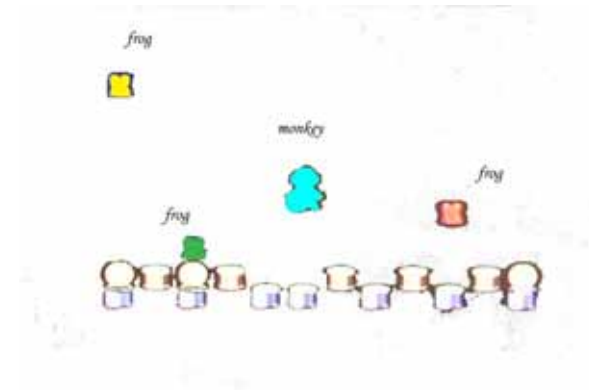


Figure 3: A snapshot of VISISounder, a monkey for a model performance (center) and three frogs for performances by subjects.

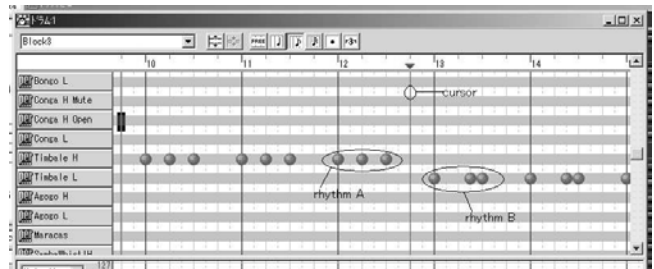


Figure 4: XGWorks: rhythm changes from A to B at the thirteenth measure

thirteenth measure was shown on the display, therefore, subjects could see the change of the rhythm before the cursor came to the position. Although the tempo change was also indicated by using a different type of drum, the degree of tempo change could not be shown.

Other differences are that a model performance is shown as continuous cursor movement, and performances by subjects are not shown on the window.

3. Virtual Drum. Virtual Drum is a program using direct API calls and Mabry Visual Basic MIDI IO controls, originally freeware [2]. A student partially modified the program in order to make it a game program for scoring a performer’s playing timing with respect to a model performance.

In Virtual Drum, a model performance appears in the

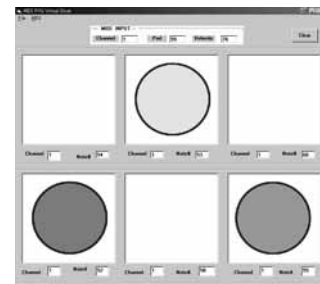


Figure 5: Virtual Drum: a model performance (circle above) and performances by subjects (two circles below)

A.2.1. Performance Visualization for Hearing Impaired Students

	VISI	XGW	VD	Sound
PA	AregVISI	AregXGW	AregVD	AregSound
PAB	ABVISI	ABXGW	ABVD	ABSound
PAT	ATVISI	ATXGW	ATVD	ATSound

Feedback types are abbreviated as follows: VISI for VISISounder, XGW for XGWorks, VD for Virtual Drum, and Sound for sound only.

Table 1: Twelve experimental sessions

upper boxes and performances by subjects in the lower boxes (Figure 5).

4. Sound only. The model performance is not visualized but only performed.

Sessions

Combining three types of model performance and four types of feedback, the experiment consisted of twelve sessions as shown in Table 1. Subjects were informed about the twelve sessions and practiced PA, PAB, and PAT only by clapping by themselves without a model performance before the experiment.

5. RESULT

We obtained the time difference between a subject performance and the model performance. The average and standard time difference for a session were calculated using the performed beats in twenty-four measures by all subjects as shown in Table 2.

The average of the time difference between a subject performance and the model performance for each beat was shown as a line graph for the rhythm patterns of PA (Figure 6), PAB (Figure 7), and PAT (Figure 8). Each line shows a session whose name is specified in Table 1. In the graphs, X-axis showed the beat. Since three notes were performed in every measure of the two rhythm patterns, beat number four was not the fourth beat of the first measure but the first beat of the second measure. Therefore, the beat number thirty seven (the first beat of the thirteenth measure) was the changing point of the rhythm in PAB and the tempo in PAT. The Y-axis showed the time difference counted by “ticks.” In the experiment, a beat consisted of 480 ticks. Therefore, tempo MM=108 meant a beat was played every 556 ms^5 and a tick roughly corresponded to 1 ms^6 .

The results from Table 2 and the figures are as follows.

1. *VISISounder*. The average and the standard deviation of the feedback of VISISounder are rather large.
2. *XGWorks*. Both the average and standard deviation of the feedback of XGWorks compare fairly well to those of other types of feedback.
3. *Virtual Drum*. Though with a small average, Virtual Drum has the largest standard deviation. This means the subjects’ performances waver.
4. *Sound feedback*. The smallest standard deviation value was obtained from the sound feedback for two of the three model performances. This is also found in the small movement of a line for the session ATSound in the three graphs (Figures 6, 7, and 8). On the other hand, the sound feedback average is rather large.

⁵ $60/108 = 0.556$

⁶ $60/108/480 = 0.00118$

The average and standard deviation of four measures before and after the rhythm change and the tempo change, namely the ninth to twelfth measures and thirteenth to sixteenth for PAB and PAT are shown in Table 3. Data of the ninth to twelfth measures show the steadiness of subjects performances after performing several repeats of a rhythm pattern with a regular tempo.

For the rhythm change (PAB), ABVISI made a big difference before and after the change, while for the tempo change (PAT), ATXGW made a big difference.

6. DISCUSSION

In discussing the time, we have to notice the basic numbers, such as, we are able to perceive multiple vocalizations when the time lag is over 20 ms or due to MIDI hardware and display redrawing. In the experiment, we do not need to take those numbers into consideration, because the precision of the time is the next step. Here we would like to see the tendency between the subjects’ performances and feedback types.

We are able to see that the sound method gives better feedback from the point of view of standard deviation than other types of feedback from the result shown in Table 2. It can be interpreted that once subjects form the performance model of the rhythm and tempo within themselves, it is more comfortable and easier for subjects to keep playing it. Of course, this result comes from the fact that the subjects are less impaired. The next good result is using the feedback of XGWorks. In spite of this, subjects did not appreciate the show-ahead of tempo and rhythm with the moving cursor of XGWorks. On the other hand, we are also able to see in Table 3 that the show-ahead visualization by showing the change in rhythm and tempo is useful as judged from the result of the smallest standard deviation obtained using XGWorks for PAB and PAT. Though with the worse result, they well appreciate the animation of VISISounder. These observations show that it is important to show something fun in the visual aid for cooperative performance.

From the experimental results, we came to the conclusion that the important thing in designing performance visualization for cooperative performance is the show-ahead of the tempo. Animation that shows only the important notes for cooperation concerning musical structure will reduce the physical burden.

Visualization with the purpose of game animation is not suitable for accompaniment. Performance visualization should be designed according to its purposes. The new user interface will be the combination of continuous information for the tempo and discrete information of the musical structure. The following is future work.

- Since the experiment was with a small number of subjects and not a variety of subjects, we need to ask more people with different musical experiences and levels of hearing impairment.
- We have to make it clear how long the subjects are affected by the change of rhythm or tempo.
- On the questionnaire after the experiment, subjects made comments on four types of feedback. They say looking at the display for the movement makes them fatigued. Therefore, we should take the physical burden caused by the feedback into consideration. Also, we should notice that animation should not always be given attention.

A.2.1. Performance Visualization for Hearing Impaired Students

PA (rhythm A, tempo regular)				
	AregVISI	AregXGW	AregVD	AregSound
average	165.3645833	5.40625	21.18981481	40.69444444
std. dev.	77.43551312	49.27962463	177.1019602	55.83921558
PAB (rhythm A and B, tempo regular)				
	ABVISI	ABXGW	ABVD	ABSOUND
average	54.39236111	33.69097222	-14.65277778	63.33333333
std. dev.	92.12372867	61.83359502	122.43531	40.74482599
PAT (rhythm A, tempo changes)				
	ATVISI	ATXGW	ATVD	ATSOUND
average	70.12820513	56.2275641	22.12820513	79.31196581
std. dev.	85.34492414	64.03764601	127.8672001	29.7305038

Table 2: The average and standard deviation of the twelve sessions

PAB (rhythm A and B, tempo regular) measure 9–12					measure 13–16			
ABVISI	ABXGW	ABVD	ABSOUND		ABVISI	ABXGW	ABVD	ABSOUND
-16.79166667	16.22916667	-11.5	48.11111111	average	41.1875	69.875	23.94444444	70.63888889
51.2869416	34.39979822	196.8497094	11.66305861	std. dev.	151.920962	79.06826021	66.34510834	82.6630569
PAT (rhythm A, tempo changes) measure 9–12					measure 13–16			
ATVISI	ATXGW	ATVD	ATSOUND		ATVISI	ATXGW	ATVD	ATSOUND
97.60416667	26.64583333	-1.055555556	54.38888889	average	122.75	124.1875	64.36111111	105.3056
31.02280289	30.22650446	114.6288551	16.55650462	std. dev.	116.1130425	69.42402036	117.1951199	37.81412764

Table 3: The average and standard deviation of four measures before and after the rhythm change (above) and the tempo change (below).

- Besides, in order to create less physical burden because of the reason above, there are other good reasons to visualize a part of the performance. They are (1) not all notes in a musical piece are given the same role and importance, and (2) a report by a music researcher indicated that a phrase can be analyzed to a tree structure according to the degree of prominence of each note [8]. The prominence of notes gives performers important information on performance. Therefore, a possible new performance visualization could show animation only at important notes, such as the first beat of every or every other measure.
- Though we could see that the showing-ahead type of performance visualization is effective as far as the tempo is regular, the sudden change in the cursor movement of XGWorks according to the tempo change is difficult to follow for subjects. A reason for the difficulty is that the movement is different from that of a human conductor who controls tempo smoothly. It is necessary to suggest the change of tempo in a smoother manner by referring to the movement of a human conductor.

7. ACKNOWLEDGMENTS

We appreciate Y. Ichikawa for her great support in preparing the musical instruments, data, and many other things. This work is supported by The Ministry of Education, Culture, Sports, Science and Technology through a Grant-in-Aid for Scientific Research (#14580243).

8. REFERENCES

- J. Foote. Visualizing music and audio using self-similarity. In *ACM Multimedia99*, pages 77–80. ACM, 1999.
- Gould Academy. http://intranet.gouldacademy.org/music/faculty/virtual/virtual_instruments.htm
- R. Hiraga. Case study: A look of performance. In *Proceedings of IEEE Visualization*, pages 501–504. IEEE, 2002.
- R. Hiraga, S. Igarashi, and Y. Matsuura. Visualized music expression in an object-oriented environment. In *Proc. of ICMC*, pages 483–486. ICMA, 1996.
- R. Hiraga and M. Kawashima. Computer music for hearing impaired students. *Technical Report of SIGMUS, IPSJ*, 42:75–80, October 2001.
- R. Hiraga and N. Matsuda. Visualization of music performance as an aid to listener’s comprehension. In *Proceedings of AVI*, 2004.
- R. Hiraga, R. Miyazaki, and I. Fujishiro. Performance visualization—a new challenge to music through visualization. In *Proceedings of ACM MultiMedia*, pages 239–242. ACM, 2002.
- F. Lerdahl and R. Jackendoff. *Generative Theory of Tonal Music*. The MIT Press, 1983.
- S. M. Smith and G. N. Williams. A visualization of music. In *Proceedings of IEEE Visualization*. IEEE, 1997.
- F. Sobieczky. Visualization of roughness in musical consonance. In *Proceedings of IEEE Visualization*. IEEE, 1996.
- A. Watanabe and I. Fujishiro. tutti: A 3d interactive interface for browsing and editing sound data. In *The 9th Workshop on Interactive System and Software*. Japan Society for Software Science and Technology, 2001.
- WEDELMusic. <http://www.wedelmusic.org/>

A.2.1. Performance Visualization for Hearing Impaired Students

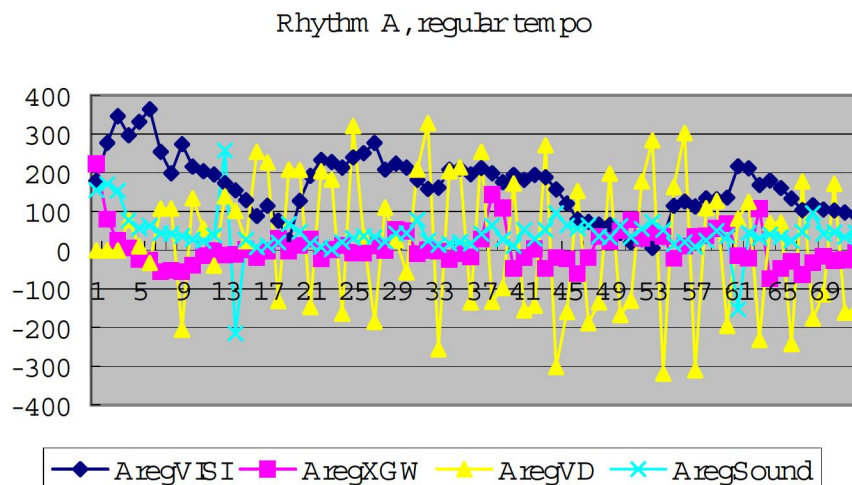


Figure 6: PA (rhythm A, regular tempo)

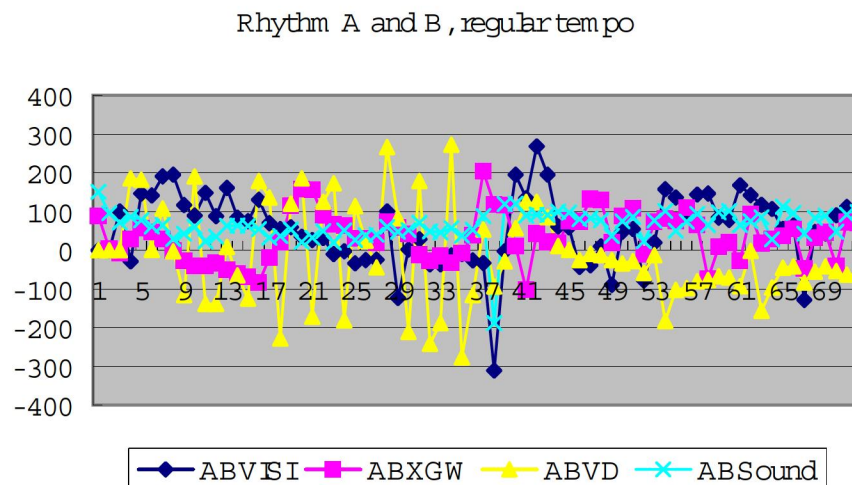


Figure 7: PAB (rhythm A and B, regular tempo)

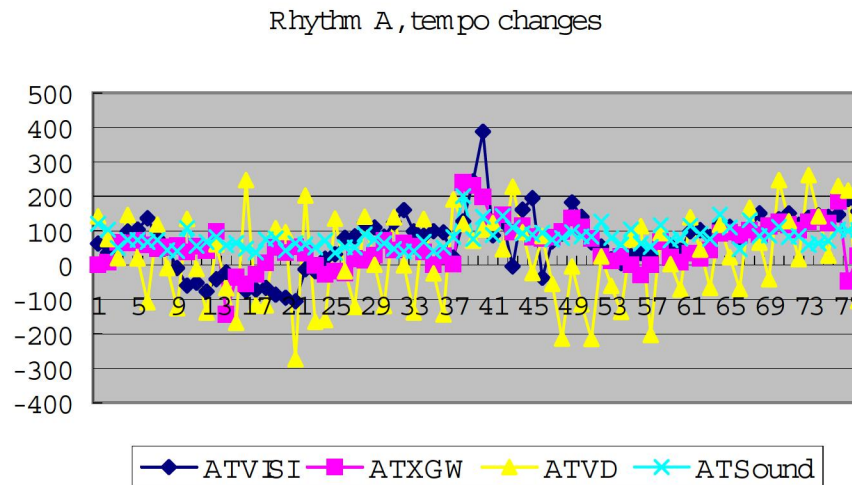


Figure 8: PAT (rhythm A, tempo changes)

Cognition of Emotion on a Drum Performance by hearing-impaired people

Rumi Hiraga

Bunkyo University
1100 Namegaya, Chigasaki
253-8550, Japan
rhiraga@shonan.bunkyo.ac.jp

Teruo Yamasaki

Osaka Shoin Women's University
958 Sekiya, Kashiba
639-0255, Japan
Yamasaki@gamelans.org

Nobuko Kato

Tsukuba College of Technology
4-3-15, Amakubo, Tsukuba
305-0005, Japan
nobuko@a.tsukuba-tech.ac.jp

Abstract

With the purpose of building a performance assistance system for hearing-impaired people and normal hearing ability to play together music, we need to know how hearing-impaired people express and understand music. This poster describes an experiment about how hearing-impaired people understand an "emotion" in a drum performance with an intended emotion. With this experiment, the possibility of the communication through musical performance among hearing-impaired people was shown.

1 Introduction

With six years of experience with teaching computer music to hearing-impaired students at Tsukuba College of Technology ([TCT](#)), we believe that the hearing-impaired have an interest in music and anxiously hope to enjoy music. As a deaf person majoring in music, Whittaker also describes a similarity in interests and enjoyment of music for both the hearing-impaired and people with normal hearing (Whittaker86). Thus we set our goal to propose an assistance system for the hearing-impaired that play instruments in an ensemble style. Performance visualization is a good candidate for use in such an environment, because it complements the listening feedback with visual cues. In spite of that, our previous experiment, using visual cues to follow the tempo, showed that a simple media transformation from the performance data to visual figures was not effective in giving excellent cues for a performance to the hearing-impaired (Hiraga04). Thus, we needed to better understand what information from a performance would be usable as visual feedback and decide how to efficiently visualize it.

We conducted experiments on how hearing-impaired people express an intended emotion on a drum performance and how they understand the emotion in a drum performance. This paper describes about the cognition of emotions in drum performances played by hearing-impaired people. The result shows the possibilities of the communication through musical performance among hearing-impaired people and the visual cues that induce some emotions work better in cooperative performances. The experiment of how hearing-impaired people express emotions on a drum performances and the analysis of the recorded performances are described in another paper by Hiraga (Hiraga05). Subjects played the drum set with one of the emotions, joy, fear, anger, and sadness. The results were almost the same as the results of the analyses of the experiments in which musically untrained adult players with normal hearing abilities were subjects. Using the performances played by hearing-impaired people, Yamasaki conducted an experiment of how people with normal hearing ability understand intended emotions in drum performances played by hearing-impaired people (Yamasaki05). The results suggest that hearing impaired people can communicate basic emotions through musical performances.

2 Experiment

The purpose of the experiment was to understand how the hearing-impaired people understand an intended emotion through a musical performance. The experiment is based on Yamasaki's experiment with kindergarten children (Yamasaki04).

A.2.2. Cognition of Emotion on a Drum Performance by Hearing-Impaired People

2.1 Subjects

We asked eleven students, 10 male and 1 female, ages 19 to 21, of Tsukuba College of Technology (TCT) to be our subjects. TCT is a three-year college of a National University Corporation for the hearing and visually-impaired. Hearing loss of over 90 decibels (dBs) qualifies for application to the hearing-impaired division of TCT. Among the five majors available in the hearing-impaired division, all our subjects major in electronics engineering. Their levels of hearing loss are as follows; one of them has 80-90 dBs of hearing loss. Although it is hard to listen to loud speech with this level of hearing loss, the subject can hear voices and sounds that are near their ears. Two of them have 90-100 dBs of hearing loss. It is generally said that people with over 90 dBs of hearing loss cannot hear loud voices near their ears. The other eight subjects have 100-110 dBs of hearing loss. Their ages of hearing loss are as follows; five innate, one under the age of two, three under the age of three, one under the age of four, and the last one at the age of seven. Of those that wear a hearing aid, five put it on all the time, two almost all the time, two occasionally, and the others never put it on.

2.2 Apparatus

We conducted the experiment in a wooden-floor gymnasium. A MIDI drum set, a Yamaha DD-55, was connected to a sequence software system, Yamaha XGWorks, on a Personal Computer, an IBM ThinkPad A31p (Intel Pentium 4, 1.70 GHz, with 1 GB RAM). We used XGWorks to record the performance data into a standard MIDI file (SMF) when a student played the DD-55. In an SMF, we can find the information for the timing of each beat, the strength of each beat in terms of velocity, and the kind of drum pads that were used. Only two of the six drum pads that were the player's side were used. The timbre of the left pad was a snare, and that of the right was a floor tam. A portable speaker set, a Roland MA-8, was connected to the DD-55.

2.3 Procedure

We asked the subjects to play the drum set using two of the drum pads for one of the four emotions; joy, sad, anger, or fear. We specified one of the four emotions for use in a random order. The subjects played a performance with an emotion one at a time. Before the experiment, we gave them the following guidance.

1. We explained to the students the ability of music to convey a player's emotion, and that the player and listeners can share this emotion through the music. We told them that among the three elements of music (melody, harmony, and rhythm), we were focusing on only rhythm for this experiment.
2. The students practiced handclapping the two and three beats sequences for a few minutes to get the hang of it.
3. We indicated the four feelings, joy, fear, anger, and sadness to the subjects. We asked them to imagine a scene that invoked a given emotion in their mind, one at a time. Next, we told them to move their body with the emotion. Altogether it took about 15 minutes.
4. We divided the students into two groups, players and listeners. Each group consists of six and five subjects.
5. Students in the player group practiced playing the drum set in order. Each of them played for just about 10 seconds. Then, the students in the player group that were on standby with their backs to the playing position in order to avoid imitating other players.
6. We gave each student in the listener group a sheet of paper with check-boxes on it that represented the emotion they felt in a performance. They sat on a chair about 10 meters from a player. Some needed to sit on the floor just in front of the player because of their individual hearing problems. All listeners turned their backs to the player, in order to avoid seeing the facial expressions of the player during the performance.

After this guidance, we started the following steps of the experiment:

1. The emotion was randomly indicated to each player, then after a few seconds, the player started playing the DD-55.
2. After each performance, the students in the listening group made a mark on a check sheet they were given that represented the emotion they felt from the performance.
3. Each player played the four emotions separately. Each player played an emotion then waited for all the other players to play. After that, they played the next emotion.

A.2.2. Cognition of Emotion on a Drum Performance by Hearing-Impaired People

After all the students in the player group performed the four emotions, the students changed roles. In this way, we gathered sixty answers for a set of performances with an intended emotion (five listeners for six performances and six listeners for five performances).

3 Result

For sixty performances of each intended emotion, the recognized emotions by hearing-impaired listeners are shown in Table 1. The analyses of the correct rate, the chi-square test, and Ryan's procedure are as follows.

- For an intended emotion, the correct rate of listening cognition were 56%, 27%, 57%, and 62% for joy, fear, anger, and sadness respectively.
- Chi-square values, $df=3$ and $p<0.00001$, of the listeners' cognition showed the significance except fear. The Chi-square values are 38.27, 41.73, and 50.00 for the intended emotion of joy, anger, and sadness, respectively.
- When the alpha level is .05, the nominal level for the step 2, 3, and 4 are .025, .0125, and .0083, respectively. The Chi-square values of two emotions for an intended emotion of joy, anger, and sadness are shown in Table 2 (a), (b), and (c), respectively. Numbers in the parenthesis shows the steps. Ryan's procedure showed that the intended emotions were recognized at the significantly higher rate than other three emotions when joy, anger, and sadness were intended. The table shows that there are no significances between the cognition of fear and anger at listening to joy-intended performances, between fear and joy at listening to anger-intended performances, and between anger and joy at listening to sadness-intended performances.

4 Discussion

Fear was the only intended emotion whose chi-square test does not show the significance even at $p=.05$. According to Table 1, fear is recognized as anger and sadness. In the result of analyzing drum performances with intended emotions (Hiraga05), we can see the resemblance between fear and sadness regarding performing factors of mean velocity. They are played much more softly compared to joy and anger and Fisher's LSD test revealed the mean velocity of anger was significantly higher than that of fear ($anger>fear$). Similarly, the significance was shown for $joy>fear$, $anger>sadness$, and $joy>sadness$. On the other hand, we cannot explain about recognizing the intended fear as anger from the performance analysis. One possible explanation is fear is understood with more variety than other emotions depending on individual subject.

By investigating the case of fear and acquire the explanation of the result, we will be able to assume that it is very likely that hearing-impaired people can communicate the emotion through the drum performance and use visual cues for emotions for a performance assistance system. In order to get the insight about the system to be used by both hearing-impaired people and people with normal hearing ability for cooperative performance, we are going to conduct another experiment of the cognition of emotion with the larger number of hearing-impaired listeners and compare the result with that of the experiment with normal hearing ability as subjects.

4. Acknowledgement

We appreciate Y. Ichikawa for her great support in preparing the musical instruments, data, and many other things. The Japan Society for the Promotion of Science supported this research through a Grant-in-Aid for Scientific research Exploratory Research, No. 16500138.

5. References

Hiraga04

Hiraga, R. and Kawashima, M. (2004). Performance Visualization for Hearing Impaired Students, *Proc. of the 3rd International Conference on Education and Information Systems: Technologies and Applications (EISTA2004)*. 323-328.

Hiraga05

Hiraga, R., Yamasaki, T., and Kato, N. (2005). Expression of Emotion by Hearing-Impaired People through Playing of Drum Set, *Proc. of the 9th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI2005)*.

A.2.2. Cognition of Emotion on a Drum Performance by Hearing-Impaired People

Tsukuba College of Technology

<http://www.tsukuba-tech.ac.jp/college.htm>

Whittaker86

Whittaker, P. (1986). Musical Potential in the Profoundly Deaf, *Music and the Deaf*.

Yamasaki04

Yamasaki, T. (2004). Emotional Communication through Music Performance played by Young Children, *Proc. of International Conference of Music Perception and Cognition (ICMPC04)*.

Yamasaki05

Yamasaki, T., Hiraga, R., and Kato, N. (2005). Emotional communication through music performance played by hearing-impaired people, *The Neurosciences and Music-II*.

Table 1: Cognition of Emotion

Intended Emotion	Recognized Emotion			
	Joy	Fear	Anger	Sadness
Joy	34	11	14	1
Fear	8	16	18	18
Anger	17	9	34	0
Sadness	4	16	3	37

Table 2(a): Chi-square values of two recognized emotions when **Joy** was intended in the performance

	Sadness	Fear	Anger	Joy
Sadness		.00389 (2)	.0008 (3)	2.4327e-08 (4)
Fear			.5485 (2)	.0006 (3)
Anger				.0039 (2)

Table 2 (b): Chi-square values of two recognized emotions when **Anger** was intended in the performance

	Sadness	Fear	Joy	Anger
Sadness		.0027 (2)	3.7380e-05 (3)	5.5112e-09 (4)
Fear			.1167 (2)	.0002 (3)
Joy				.0173 (2)

Table 2 (c): Chi-square values of two recognized emotions when **Sadness** was intended in the performance

	Anger	Joy	Fear	Sadness
Anger		.7055 (2)	.0029 (3)	7.6213e-08 (4)
Joy			.0073 (2)	.2.5535e-07 (3)
Fear				.0039 (2)

この部分は以下の論文で構成されていますが、著作権者(著者、出版社、学会等)の許諾を得ていないため、筑波技術大学では電子化・公開しておりません。pp.27-30

Expression of Emotion by Hearing-Impaired People through Playing of Drum Set

The 9th World Multi-Conference on Systemics, Cybernetics and Informatics,
4-pages in CD-ROM Proceedings, 2005

Understanding emotion through multimedia: comparison between
hearing-impaired people and people with hearing abilities

©ACM, 2006. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in Assets '06 Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility, 2006, 141–148
<http://doi.acm.org/10.1145/1168987.1169012>

Understanding Emotion through Multimedia

Comparison between Hearing-Impaired People and People with Hearing Abilities

Rumi Hiraga
Faculty of Information and Communication
Bunkyo University
1100 Namegaya, Chigasaki, 253-8550, Japan
rhiraga@shonan.bunkyo.ac.jp

Nobuko Kato
Faculty of Industrial Technology
Tsukuba University of Technology
4-3-15 Amakubo, Tsukuba, 305-8520, Japan
nobuko@a.tsukuba-tech.ac.jp

ABSTRACT

We conducted an experiment to determine the abilities of hearing-impaired and normal-hearing people to recognize intended emotions conveyed in four types of stimuli: a drum performance, a drum performance accompanied by a drawing expressing the same intended emotion, and a drum performance accompanied by one of two types of motion pictures. The recognition rate was the highest for a drum performance accompanied by a drawing even though participants in both groups found it difficult to identify the intended emotion because they felt the two stimuli sometimes conveyed different emotions. Visual stimuli were especially effective for performances whose intended emotions were not clear by themselves. The difference in ability to recognize intended emotions between the hearing-impaired and normal-hearing participants was insignificant. The results of this and a series of experiments will enable us to better understand the similarities and differences between how people with different hearing abilities encode and decode emotions in and from sound and visual media. We should then be able to develop a system that will enable hearing-impaired and normal-hearing people to play music together.

Categories and Subject Descriptors

J.5 [Arts and Humanities]: Performing arts; J.4 [Social and Behavioral Sciences]: Psychology

General Terms

Human Factors

Keywords

Hearing-impairment, Emotion, Recognition, Drum performance

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ASSETS'06, October 22–25, 2006, Portland, Oregon, USA.
Copyright 2006 ACM 1-59593-290-9/06/0010 ...\$5.00.

1. INTRODUCTION

Our six years of teaching hearing-impaired students at the Tsukuba College of Technology, now Tsukuba University of Technology (NUTUT [4]), how to use computers to play music has shown us that many hearing-impaired people are interested in and enjoy playing music, especially with others. As a deaf person who had majored in music, Whittaker described a similarity in musical interests among hearing-impaired people and people with hearing abilities [14].

Thus, we set as our goal the development of a system that will enable hearing-impaired people to play music in an ensemble comprising both hearing-impaired people and people with hearing abilities. Besides widely used with music in performing arts [2], visualized music cues give hearing-impaired people more information about the music. We plan to design the system to assist users with music performance visualization.

We will initially use drums as the primary instrument in our system because they require simpler body movements and less knowledge of music. Moreover, drums are generally easier for people to play, and drum performances are usually easier to recognize than other types of musical performances, such as performances by piano. Playing the drums also has a healing effect [7]. However, the results of a previous experiment showed that following even a simple rhythm and tempo on the basis of visual cues can be somewhat burdensome for hearing-impaired people, particularly having to pay close attention to visual cues to keep up with the rhythm and tempo [9]. We concluded that communicating an intended emotion through drum playing might be the best approach for a performance assistance system because a musical performance that focuses on an emotion favors freedom over accuracy.

Our system will assist users improvising with drum instruments with intended emotions to get the feel of unity. Before we can design our system and construct a prototype, we needed to improve our understanding of how hearing-impaired people interpret drum performances and what types of visual stimuli would be the most useful to them. We thus conducted an experiment to evaluate how well hearing-impaired and normal-hearing people recognize an intended emotion. We used four types of stimuli: a drum performance, a drum performance accompanied by a drawing expressing the same intended emotion, and a drum performance accompanied by one of two types of motion pic-

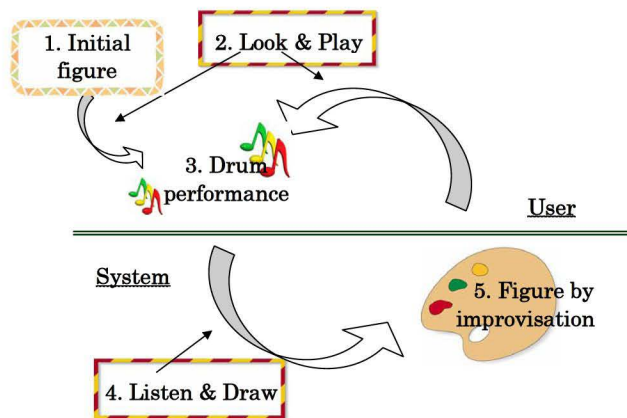


Figure 1: System concept.

tures. The sounds and drawings used in this experiment were generated in previous experiments by both hearing-impaired and normal-hearing people with an intended emotion in mind, so the results of this experiment should show how well the two types of people can identify each other's intended emotions and use those emotions to communicate. Comparison of the recognition rates between the hearing-impaired and normal-hearing participants showed that the highest recognition rate was for a drum performance accompanied by a drawing expressing the same intended emotion although the participants sometimes had difficulty in identifying the intended emotion because they felt the two stimuli sometimes conveyed different emotions.

2. BACKGROUND

2.1 System concept

Our target music performance assistance system, called the “performance enhancement machine” (PEM), will generate visual cues that enable users, even those without any musical training, to enjoy playing the drums and to feel a sense of unity by playing with others. The basic concept is illustrated in Figure 1. All the players simply look at an initial drawing chosen by the leader to determine which emotion is to be emphasized in the performance at first and play their instruments as a group with that emotion in mind. The system analyzes the sounds in their generated performance, identifies the dominant emotion, and generates a representative drawing of it. The players look at the generated drawing to determine which emotion to emphasize and play their instruments again as a group with that emotion in mind. There is thus a cycle of group performance and system drawing, leading to the players harmonizing their performances and playing in better unison. That is, in a sense, PEM is a system that realizes user-machine interaction through emotion.

The generated drawings are simply cues to the users suggesting which emotion to emphasize. They do not specify a performance rule or act as a substitute musical score. The users can play their instruments freely and improvise. Because the cues help clarify the intended emotion, the users can get the feeling of playing in unity.

2.2 Related works

This experiment is one in a series of experiments we are conducting on recognition related to sound and visual information with hearing-impaired people and normal-hearing people. So far, we have conducted experiments on encoding emotions in drum performances [10][11], recognizing the intended emotions in a drum performance [12], and encoding and decoding intended emotions in drawings [8].

In the experiment on recognizing the intended emotion in a drum performance, we found that hearing-impaired listeners did not differentiate between the drum playing of hearing-impaired people, of normal-hearing people with no training in playing drums (amateurs), and professionals. Listeners with normal hearing, on the other hand, could differentiate performances by professionals from those by other performers. There were no significant differences between hearing-impaired and normal-hearing people in recognizing the intended emotions in performances by hearing-impaired people and amateurs.

In our experiment on decoding intended emotions in drawings, hearing-impaired people had better recognition rates than normal-hearing people, although the difference was not significant.

Bresin and his colleague developed a system that renders a musical performance with an intended emotion [5] and Friberg developed a system that analyzes a musical performance and decodes the emotions it expresses [6]. Though Juslin surveyed the recognition of emotion through music [13], little research has been done on understanding intending emotions through drum performances specifically [15] or with hearing-impaired people.

3. EXPERIMENT

3.1 Methods

We used the four basic emotions commonly used in experiments on music perception: joy, fear, anger, and sadness. To determine how well hearing-impaired people and normal-hearing people recognize an intended emotion, we used four types of stimuli generated in a previous experiment: a drum performance, a drum performance accompanied by a drawing expressing the same intended emotion, and a drum performance accompanied by one of two types of motion pictures (Windows Media Player's amoeba effect or fountain effect). The drum performances and drawings were encoded with one of the four emotions by both hearing-impaired and normal-hearing people in our previous experiments.

Subjects either listened to or listened to and watched the presented stimuli and then decided which emotion they felt in the stimuli. As in our other experiments, we focused on comparing the ability of hearing-impaired and normal-hearing subjects to recognize the emotion in the stimuli.

3.2 Material

3.2.1 Drum performances (sound)

The drum performances were recorded in previous experiments. We asked three groups of people to play a drum set so as to convey a particular emotion. The three groups were hearing-impaired people [10][11], people with normal hearing abilities who had no training in drum performance (we call them amateurs), and people with normal hearing abilities who are professional drummers. The number of players

A.2.4. Understanding Emotion through Multimedia

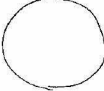











Hearing-impaired college students		College students with normal hearing abilities	
Electronics major	Design major	Design major	
Joy			
			
Fear			
			
Anger			
			
Sadness			
			

Figure 2: Drawings used as stimuli.

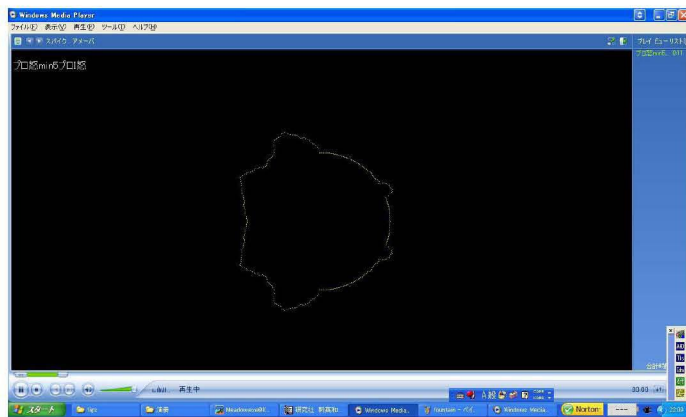


Figure 3: Example amoeba effect.

in each group was 11, 5, and 2, respectively. Since each player did one performance for each of the four emotions, there were 44, 20, and 8 performances per group, respectively. The length of a performance varied from about 10 seconds to over 60 seconds. Hearing-impaired college students played a MIDI drum set, Yamaha DD-55, and their performances were recorded as standard MIDI files (SMFs) with a sequence software system, Yamaha XGWorks. Other performances were played with a tam and recorded into a DAT recorder, Sony TCD-D10 PRO II, through a sound-level meter, RION NL-20.

We calculated the recognition rates for each performance. If the listener perceived the same emotion as that intended by the performer, the trial was scored as correct. For each of the three groups of performers, we identified the performances with the best and worst recognition rates for each emotion. These 24 performances (three groups * four emotions * two qualities) were used as the sound stimuli in the present experiment.

3.2.2 Drum performances paired with drawings (drawing)

We paired each of the 24 drum performances used as sound stimuli with a drawing that conveyed the same emotion. The drawings were also from a previous experiment [8]. We asked three groups of people to draw simple pictures conveying an emotion. The three groups were hearing-impaired college students whose major was electronics, hearing-impaired college students whose major was design, and college students with normal hearing whose major was design. The number of people in the groups was 14, 11, and 7, respectively. From these samples, we chose the one with the highest recognition rate for each emotion for each group in the previous experiment. We excluded drawings that represented concrete objects, such as the sun and tear drops, even though they were with the highest recognition rate. The 12 drawings (three groups * four emotions) are shown in Figure 2.

Except for conveying the same emotion, the pairings were random. Since there were 24 performances and only 12 drawings, each drawing was used twice. The drawings were presented using Windows PowerPoint during the first half of a performance and gradually withdrawn during the second half.

3.2.3 Drum performances paired with motion pictures (amoeba and fountain)

We also paired the drum performances with motion pictures: the Windows Media Player amoeba and fountain effects. Although these effects were controlled by and synchronized with the sound data, the resulting animations did not convey any particular emotion themselves. We chose amoeba (Figure 3) because its representations looks a little like some of the drawings we used. We also wanted to use pictures that are quite different in shape and movement from amoeba. We chose fountain (Figure 4) because it uses fewer colors than other pictures that are different from amoeba.

The order of performances was random. The order was the same through the four stimulus categories. The four stimulus categories were presented to subjects in the order of sound, drawing, amoeba, and fountain.

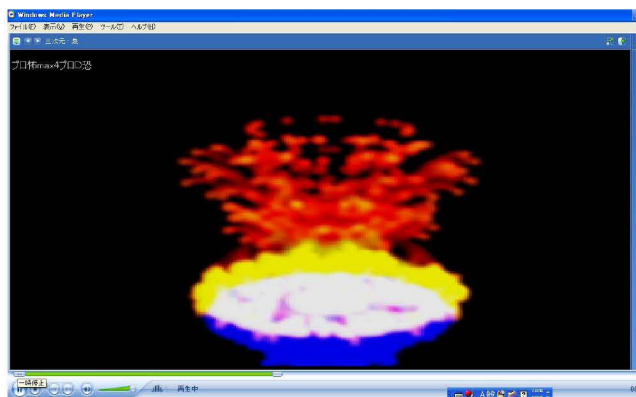


Figure 4: Example fountain effect.

3.3 Subjects

The subjects were hearing-impaired college students and college students with normal hearing. The 11 hearing-impaired subjects comprised 3 men and 8 women¹ (ages 18–22), and the 15 normal-hearing subjects comprised 13 men and 2 women (ages 20–24).

The hearing-impaired subjects were all students in the hearing-impaired division at NUTUT. All had a hearing loss of more than 100 decibels. We surveyed their musical experience in terms of karaoke, music-related games (such as “Dance, Dance, Revolution [1]” and “Drum Master [3]”), dance, and music-related club activity. Of the 11 hearing-impaired subjects, 10 had experience in karaoke, 10 in games, and 7 in dance. Two of them belonged to a dance club and two to a Japanese drum (taiko) club.

3.4 Procedure

The hearing-impaired subjects were tested in a wood-floor gymnasium. They sat on the floor where there was a hearing compensation device (Figure 5). The normal-hearing subjects were tested in a classroom.

We gave the subjects check sheets and instructed them to mark which of the four emotions they recognized from each stimulus. They were presented the 24 stimuli in each category one after the other, about 12 minutes per category, with a 5-minute break between categories. During each break, they prepared a self-judgment report. After viewing the stimuli in all the categories, they summarized how they felt about the experiment.

4. RESULTS

4.1 Recognition rates and ANOVA

We calculated the recognition rates for the stimulus categories and used them for two-way analysis of variance (ANOVA) on arcsine-transformed data. In the following results, significant difference is considered less than a 5 percent probability. We formed three ANOVA analyses where factors were as follows.

1. Intended emotion (four levels: joy, fear, anger, and sadness) and stimulus category (four levels: sound, drawing, amoeba, and fountain).

¹Three subjects (1 man and 2 women) did not participate in the part of the experiment using the fountain stimulus.



Figure 5: Hearing-impaired subjects being tested on wood floor.

2. Intended emotion (four levels) and subject group (two levels: hearing-impaired and normal hearing).
3. Stimulus category (four levels) and subject group (two levels).

Figure 6 shows the recognition rates for emotions by subject type. Figure 7 shows it for the subject types by stimulus category, and Figure 8 shows it for the subject types by intended emotion. Table 1, 2, and 3 show χ^2 values of each ANOVA above.

From Figure 6, Table 1, and Ryan’s procedure, we obtained the following results.

- Fear was the most poorly recognized emotion by both subject groups for all stimuli.
 - There was a significant difference in the recognition rates between emotions by both subject groups.
 - Although the ordering of the recognition rates by emotion differed between subject groups, Ryan’s procedure showed that there was a significant difference between recognizing fear and recognizing the other three emotions in both subject groups.
 - The drawing stimuli produced the highest recognition rates for both subject groups regardless of the intended emotions.
 - The recognition rates for the hearing-impaired subjects, in descending order, were for the drawing, fountain, sound, and amoeba stimuli. For the normal-hearing subjects, they were for the drawing, fountain, amoeba, and sound stimuli.
- The subjects with normal hearing showed a significant difference in recognition rates between the amoeba and sound stimuli, while the difference for the hearing-impaired subjects was insignificant.
- The recognition rates for the drawing stimuli were significantly higher than for the other three categories for both subject groups.

From Figure 7, Table 2, and Ryan’s procedure, we obtained the following results.

A.2.4. Understanding Emotion through Multimedia

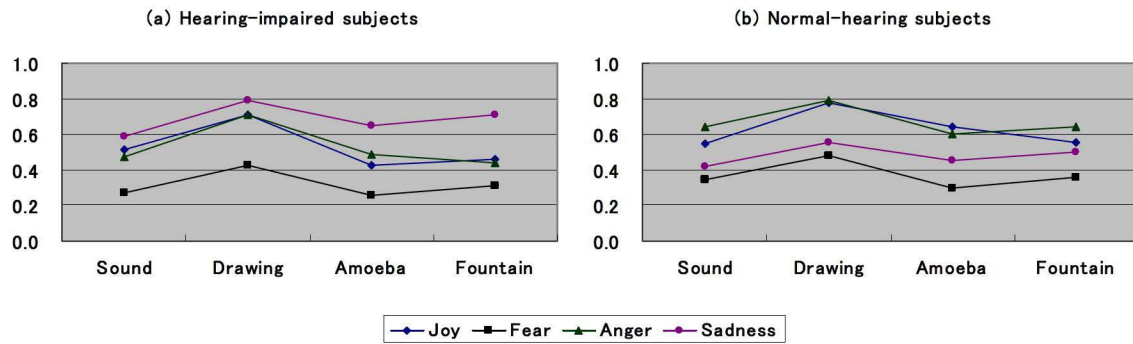


Figure 6: Recognition of emotion by category of stimulus.

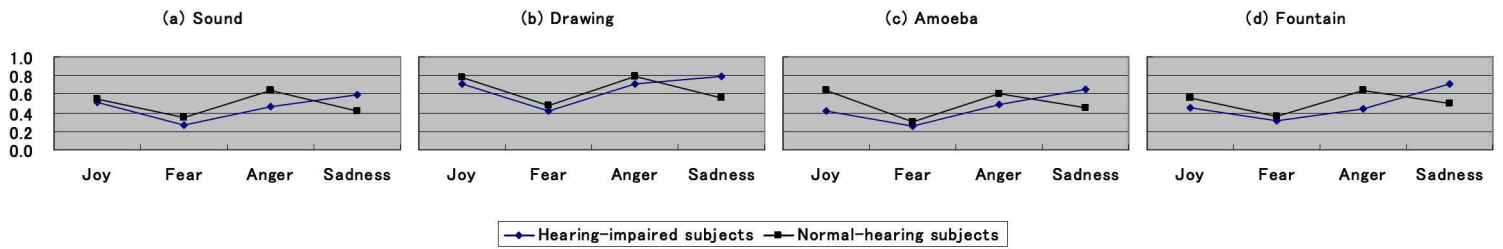


Figure 7: Recognition by subjects for each emotion.

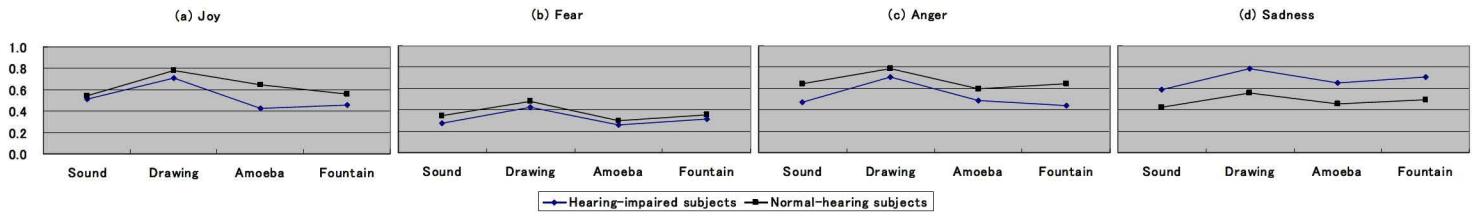


Figure 8: Recognition by subjects for each category of stimulus.

Table 1: χ^2 values of ANOVA (1). Main effects are intended emotion and stimulus category.

	Subjects	
	Hearing impaired	Normal hearing
Main effect (A): Intended emotion	70.82*	87.31*
Main effect (B): Stimulus category	30.27*	27.55*
Interaction of two effects above	5.13	6.05

*: the significance at $p \leq .05$.

Table 2: χ^2 values of ANOVA (2). Main effects are intended emotion and subject group.

	Stimulus categories			
	Sound	Drawing	Amoeba	Fountain
Main effect (A): Intended emotion	24.68*	39.14*	34.34*	21.03*
Main effect (B): Subject group	0.50	0.04	1.34	0.59
Interaction of two effects above	9.69*	12.07*	14.59*	12.12*

*: the significance at $p \leq .05$.

A.2.4. Understanding Emotion through Multimedia

Table 3: χ^2 values of ANOVA (3). Main effects are stimulus category and subject group.

	Intentional emotion			
	Joy	Fear	Anger	Sadness
Main effect (A): Stimulus category	23.75*	10.73*	20.31*	9.70*
Main effect (B): Subject group	6.51*	1.82	12.61*	25.12*
Interaction of two effects above	2.96	0.10	1.31	0.50

*: the significance at $p \leq .05$.

- Ryan’s procedure showed that the recognition rates for fear were significantly lower than for the other three emotions for all four types of stimuli, while there were no significant differences between any two of the other three emotions for any types of stimuli.
- For the sound and drawing stimuli, the recognition rates in descending order were for anger, joy, sadness, and fear. For the other two categories (amoeba and fountain), the rates in descending order were sadness, anger, joy, and fear.
- There was no significant difference between subject groups in recognizing intended emotions for all four types of stimuli.

From Figure 8, Table 3, and Ryan’s procedure, we obtained the following results.

- Subjects with normal hearing had higher recognition rates for emotions other than sadness than the hearing-impaired subjects.
- There was no significant difference between subject groups in recognizing fear.
- Ryan’s procedure showed that recognition with the drawing stimuli differed significantly from the other three types of stimuli for all four emotions. The exception was that there was no significant difference between the drawing and fountain stimuli for recognition of sadness.
- Though the recognition rates differed among emotions by types of stimulus, Ryan’s procedure showed that the difference between the amoeba and fountain stimuli was not significant.

4.2 Self-judgment

The post-experiment self-judgment investigated how difficult the subjects found the experiment.

4.2.1 Difficulty

Subjects checked one of the five degrees of difficulty (from 5 for “very easy” to 1 for “very difficult”).

- None of the hearing-impaired subjects checked 5 (very easy), two checked 4, three checked 3, five checked 2, and one checked 1 (very difficult.)
- The corresponding numbers for normal-hearing subjects were 0, 2, 1, 10, and 2.
- Less than one-third (3 out of 11) of hearing-impaired subjects felt the experiment was difficult, while about 80% with normal hearing felt it was difficult.

4.2.2 Preferences

The subjects also indicated the types of stimuli in which they felt it was the easiest and the most difficult to recognize the intended emotion (Figure 9). The hearing-impaired subjects strongly preferred motion picture stimuli, while the normal-hearing subjects about equally preferred drawing and motion picture stimuli. The preference for motion picture stimuli among hearing-impaired subjects was indicated with our previous experiment in following tempo and rhythm, even though the stimuli did not yield a good result [9].

4.2.3 Stimulus categories

Seven hearing-impaired subjects and eight normal-hearing subjects indicated that they sometimes recognized different emotions between the performance stimulus and in the drawing stimulus in a drawing category pair, even though the intended emotions were the same. Six hearing-impaired subjects and seven normal-hearing subjects indicated that the emotions in the amoeba stimuli were easier to recognize than those in the drawing stimuli. Six hearing-impaired subjects and seven normal-hearing subjects indicated that the emotions in the fountain stimuli were easier to recognize than those in the amoeba stimuli.

The subjects were asked to specify the easiest and the most difficult emotion to recognize for each stimulus category. More subjects with normal hearing found fear the easiest to recognize and sadness the most difficult to recognize for all the categories than did the hearing-impaired subjects.

5. DISCUSSION

5.1 Recognition of intended emotion

5.1.1 Lowest recognition rate for fear

Since fear had the lowest recognition rate in our previous experiments on recognition of intended emotions with performances and drawings, it is not surprising that fear had the lowest recognition rate for all stimulus categories. The reason for this is not clear. A possible explanation is that fear is not easy to encode into any type of media.

5.1.2 Significant difference in recognition of emotions

Because our previous experiments showed the same result, it was not surprising that there was a significant difference in recognition rates among emotions for all stimulus categories and subject groups. This is a serious problem for our planned system. We need to find a way to improve the recognition rate so as to eliminate significant differences in recognition among emotions.

A.2.4. Understanding Emotion through Multimedia

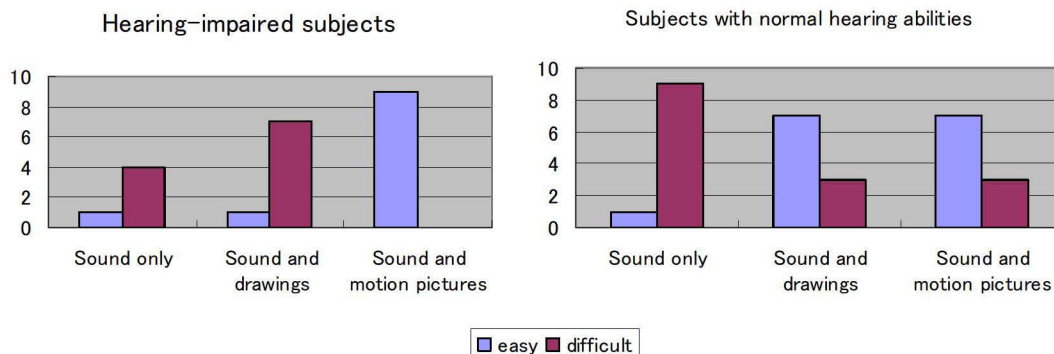


Figure 9: Preferences for types of stimuli.

5.1.3 Higher recognition rate for sadness by hearing-impaired subjects

Only sadness was better recognized by the hearing-impaired subjects, while the three other emotions were better recognized by the normal-hearing subjects. The self-judgment reports show that more hearing-impaired subjects felt it was easy to recognize sadness, though fewer felt it was easy with the fountain stimulus. For all stimulus categories, more subjects with normal hearing found sadness difficult than did hearing-impaired subjects. The self-judgment report results correspond to the recognition rate results, at least in this case.

5.2 Drawing stimulus category

5.2.1 Highest recognition rate

It is noteworthy that 7 out of 11 hearing-impaired subjects and 8 out of 15 normal-hearing subjects mentioned that the emotion they perceived from a performance sometimes differed from the one they perceived from the paired drawing, although the performance and drawing had the same intended emotion.

In spite of that, the recognition rate with the drawing stimuli was the highest of all types for all intended emotions because we used the drawings with the best recognition rate from a previous experiment. Although this may be the reason, we cannot explain why some subjects found a conflict between the sound and visual stimuli or why some of them reported that they used the sound stimulus more than the visual one in deciding which emotion to mark for the drawing category.

5.2.2 Drawing stimulus only

We conducted a supplemental experiment one month later to try and clarify how drawings are recognized. We randomly arranged the same drawings in Figure 2 and asked the same subjects² to identify the intended emotion for each drawing. The result with the recognition rate was as follows.

- There was no significant difference in the average recognition rates between both the two groups.
- There was no significant difference in the intended emotion and hearing ability factors in the ANOVA analysis.

²Eight of the 11 hearing-impaired subjects and all 15 of the normal-hearing ones participated.

- Since we used performance data with the best and worst recognition rates in a previous experiment [12] and drawing data with the best recognition rate in another previous experiment [8], we analyzed the recognition rates by ANOVA where factors were stimulus category (drawing and drawing-only) and best-worst performance data (obtained by splitting the performance data set into the best-recognized and worst-recognized groups). We obtained the following results, which were common to both subject groups.

- There was a significant difference between the best and worst performances.
- The subordinate test showed that the simple main effect of the best-worst factor for the drawing stimuli was significant.
- The subordinate test showed that the simple main effect of the stimulus category factor for the worst performance data was significant.

These results indicate that the recognition rates for the worst performance set increase when subjects listen to them along with visual information. It means that visual stimuli are effective in recognizing emotions for performances whose intended emotions were not clear by themselves.

5.3 Subjects' self-judgment

The self-judgment reports described in 4.2 were not consistent with the recognition rates for the stimulus categories. As described in 4.2.3, the emotions in the amoeba stimuli were easier to recognize than those in the drawing stimuli, and the emotions in the fountain stimuli were easier to recognize than those in the amoeba stimuli. This inconsistency may be due to the way we presented the inquiry. More consistent results might have been obtained if we had asked the participants to simply order the types of stimuli by how easy it was to recognize the emotions in them.

The self-judgment reports were also inconsistent regarding the ease and difficulty of recognizing the four emotions. They showed that fear was not necessarily the most difficult emotion to recognize, in fact, fear was the easiest emotion for the normal-hearing subjects to recognize except in the sound category.

Contrary to our prediction that the fountain stimuli would convey impressions of joy and anger because of its colors

A.2.4. Understanding Emotion through Multimedia

(red, yellow, blue, and white) and dynamic movements, their presentation did not affect the recognition of emotion. Only one of the 23 subjects reported that she found it difficult to differentiate between joy and anger in the fountain category.

5.4 Future work

Before we can actually build our performance assistance system, we have to more specifically determine how the system will analyze musical performances and use the results to draw pictures expressing the intended emotion (“Listen & Draw” in Figure 1). For that purpose, we need to understand the following things in particular.

- The physical characteristics of performances and drawings that identify the intended emotion. Then we can confirm that the encoding rules between hearing-impaired people and normal-hearing people are similar.

We will also be able to use the physical characteristics of both types of stimuli to dissolve the significant difference in recognizing emotions.

- The method to map physical characteristics of performances to those of drawings. Then we can make the system artificially generate a drawing expressing the emotion identified in the performance.
- The timing to generate a drawing based on the analysis of a performance. We do not want the system to distract players from their performances by presenting a drawing too early or make them uneasy by presenting it too late. We may have to introduce the basic concepts of music, such as a tempo and a measure, to the system without requiring players understand them.

The further research on the recognition of sound by hearing-impaired people may improve the usability of the system. It includes the following things.

- We will investigate the recognition of intended emotions in performances in relation to the degree of hearing impairment. In the experiment reported here, we simply divided the participants into two groups: hearing impaired and normal hearing. However, there could be gradations in recognition ability related to the degree of impairment and the amount of musical experience.
- We will then investigate how learning to play the drums can change the encoding and decoding processes for hearing-impaired people.

6. CONCLUSIONS

We conducted an experiment to determine the abilities of hearing-impaired and normal-hearing people to recognize intended emotions in four types of stimuli: a drum performance, a drum performance accompanied by a drawing expressing the same intended emotion, and a drum performance accompanied by one of two types of motion pictures. The recognition rate was the highest for a drum performance accompanied by a drawing even though subjects in both groups found it difficult to identify the intended emotion because they felt the two stimuli sometimes conveyed different emotions. Visual stimuli were especially effective for performances whose intended emotions were not clear by

themselves. The difference in ability to recognize intended emotions between the hearing-impaired and normal-hearing subjects was insignificant.

After we more specifically determine how the system will analyze musical performances and use the results to draw pictures expressing the intended emotion, we will construct and test a prototype of our performance assistance system.

7. ACKNOWLEDGMENTS

The Japan Society for the Promotion of Science supported this research through a Grant-in-Aid for Scientific Research, Exploratory Research No. 16500138.

8. REFERENCES

- [1] Dance Dance Revolution freak. <http://www.ddrfreak.com/>.
- [2] Digital Image Processing with Sound. <http://dips.dacreation.com>.
- [3] Drum Master. <http://www.namco.com/games/taiko/>.
- [4] National University Corporation Tsukuba University of Technology. <http://www.tsukuba-tech.ac.jp>.
- [5] R. Bresin and A. Friberg. Emotional coloring of computer-controlled music performances. *Computer Music Journal*, 24(4):44–63, 2000.
- [6] A. Friberg. pDM: an expressive sequencer with real-time control of the KTH music performance rules. *Computer Music Journal*, 30(1):37–48, 2006.
- [7] R. L. Friedman. *The Healing Power of the Drum*. White Cliffs Media, 2000.
- [8] R. Hiraga, N. Kato, and T. Yamasaki. Understanding emotion through drawings: comparison between people with normal hearing abilities and hearing-impaired people. In *Proceedings of IEEE SMC 2006*, 2006 (to appear).
- [9] R. Hiraga and M. Kawashima. Performance visualization for hearing impaired students—a report of the preliminary experiment. In *Proceedings of EISTA 2004*, 2004.
- [10] R. Hiraga, T. Yamasaki, and N. Kato. Cognition of emotion on a drum performance by hearing-impaired people. In *Proceeding CD of HCII 2005*, 2005.
- [11] R. Hiraga, T. Yamasaki, and N. Kato. Expression of emotion by hearing-impaired people through playing of drum set. In *Proceedings of WMSCI 2005*, 2005.
- [12] R. Hiraga, T. Yamasaki, and N. Kato. The recognition of intended emotions for a drum performance: differences and similarities between hearing-impaired people and people with normal hearing abilities. In *Proceedings of ICMPC 2006*, 2006 (to appear).
- [13] P. N. Juslin and J. A. Sloboda. ed. *Music and Emotion: Theory and Research*. Oxford University Press, 2001.
- [14] P. Whittaker. *Musical potential in the profoundly deaf*. Music and the Deaf, West Yorkshire, BK, 1986.
- [15] T. Yamasaki. Emotional communication through performance played by young children. In *Proceedings of ICMPC 2004*, 2004.

Understanding emotion through drawings comparison between
hearing-impaired people and people with normal hearing abilities

(c) 2006 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.

Understanding Emotion through Drawings

the comparison between people with normal hearing abilities and hearing impairment

Rumi Hiraga, Nobuko Kato, and Teruo Yamasaki

Abstract—With the purpose of building a performance assistance system with visual cues for hearing-impaired people and people with normal hearing abilities to play music together, we need to know how hearing-impaired people express and understand music. We have conducted a series of experiments about how hearing-impaired people understand an emotion in a drum performance with an intended emotion. With the experiments, the possibility of the communication based on emotion through musical performance. Then we need to know the usability of visual interface. In this paper, we describe an experiment on how hearing-impaired people understand an emotion in small simple drawings with an intended emotion. The results suggest the similarity and the difference on the cognition of emotion between hearing-impaired people and people with normal hearing abilities.

I. INTRODUCTION

With six years of experience with teaching computer music to hearing-impaired students at Tsukuba College of Technology (now National University Corporation, Tsukuba University of Technology, hereafter we call it NUTUT [1]), we believe that the hearing-impaired people have an interest in music and anxiously hope to enjoy music. As a deaf person majoring in music, Whittaker also describes a similarity in interests and enjoyment of music for both the hearing-impaired and people with normal hearing abilities [2]. Thus we set our goal to propose an assistance system for the hearing-impaired people to play instruments in an ensemble style; ensemble both by hearing-impaired people and people with normal hearing abilities. Performance visualization is a good candidate for use in such an environment, because it complements the listening feedback with visual cues. In spite of that, our previous experiment, using visual cues to follow the tempo, showed that a simple media transformation from the performance data to visual figures was not effective in giving excellent cues for a performance to the hearing-impaired [3]. To find out an appropriate information to visualize is an issue to build the system.

Since the drum performance is more familiar to hearing-impaired people, at least students of NUTUT because they know Japanese drum, and understanding it does not relate to pitch, we plan to use drums with our performance assistant system. With the system, both hearing-impaired people and people with normal hearing abilities can enjoy music and feel the unity by playing together with little music techniques. We

found emotional communication through music performance meets the purposes of the system, thus, we needed to better understand how hearing-impaired people express drum performances, how they understand them, and whether there are any differences in playing and understanding performances between hearing-impaired people and people with normal hearing abilities. Although some researchers have worked on emotion carried on music by subjects of normal hearing abilities [4], there have been no research on the issue.

So far, we have conducted experiments on drum performances. They are experiments on how hearing-impaired people express an intended emotion on a drum performance [5] and how they understand the emotion in a drum performance [6]. We restrict emotions to joy, fear, anger, and sadness. Drum performances that are the objects of an encoded emotion both by hearing-impaired people and people with normal hearing abilities were analyzed physically and we found that performances with an intended emotion are similar between the two types of players. As for the cognition of an emotion through performances, the correct rate was the lowest for the fear-intended performance set. Fear was the only intended emotion whose χ^2 test does not show the significance even at $p=.05$, while other three performance sets show very high values. We also compared the understanding the emotion in a drum performance by hearing-impaired people and people with normal hearing abilities [7]. The results suggested that hearing impaired people can communicate basic emotions through musical performances.

In order to improve performers' satisfaction by using the system, visual cues are another point of the system. In this paper, we describe an experiment of usability of visual interface that may assist understanding emotion in performance. The experiment shows how hearing-impaired people and people with normal hearing abilities recognize an emotion in a small simple drawing with an intended emotion. The results suggest the similarity and the difference on the cognition of emotion between hearing-impaired people and people with normal hearing abilities.

II. EXPERIMENT

A. Outline

We conducted an experiment of how hearing-impaired people and people with normal hearing abilities understand an intended emotion through a drawing. The experiment consists of the two steps; subjects drew four simple monochrome line drawings, each of the drawings is with one of the emotion of joy, fear, anger, and sadness, at the first step.

R. Hiraga is with Faculty of Information and Communication, Bunkyo University rhiraga@shonan.bunkyo.ac.jp

N. Kato is with Faculty of Electrical Engineering, Tsukuba University of Technology, nobuko@a.tsukuba-tech.ac.jp

T. Yamasaki is with the Department of Humanity, Shoin Women's College, yamasaki@gamelans.org

A.2.5. Understanding Emotion through Drawings

Then subjects look at collected drawings and specifies an emotion they feel in each drawing at the second step. In this paper, we focus on the second step of the experiment, namely, the cognition of an emotion in drawings.

We prepared three sets of drawings by different categories of people; (1) college students with normal hearing abilities whose major is design, (2) hearing-impaired college students whose major is electronics, and (3) hearing-impaired college students whose major is design. The number of people of each category is 7, 14, and 11. Since each person drew four drawings with emotions, there are 28, 56, and 44 drawings in each drawing set.

B. Subjects

Subject to look at three drawing sets consists of three groups; (1) college students with normal hearing abilities whose major is not design, (2) hearing-impaired college students whose major is , and (3) hearing-impaired college students whose major is design. The number of each subject group is shown in Table I.

Some of the hearing-impaired subjects drew drawings prior to the experiment, while subjects with normal hearing abilities who looked at drawings and people with normal hearing abilities who drew drawings are completely different group.

Hearing-impaired people are all students of hearing-impaired division of NUTUT. Hearing loss of over 90 decibels (dBs) qualifies for application to the hearing-impaired division of NUTUT.

C. Procedure

We prepared sheets for three inquiries that correspond to three drawing sets. Each sheet includes drawings and check-boxes for four emotions. Figure 1 is a sheet of inquiry. Subjects made a mark on a check-box that represented the emotion they felt from the drawing.

III. RESULT

A. Correct rate

We use the correct rate of cognition of an intended emotion. The correct rate was analyzed according to the subject groups who looked at drawings and the difference of drawing groups.

Figure 2 and Figure 3 show the correct rate by each subject group and that of each drawing group respectively. Figure 2 (a) shows how people with normal hearing abilities recognize an emotion from drawings with an intended emotion. Each line shows a drawing group. Figure 3 (b) shows how drawings by hearing-impaired people of electronics major are recognized by subjects. Each line shows a subject group.

Fear-intended drawings by all the three drawing groups show the lowest correct rate in the cognition of all the three subject groups.



















	<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲
	<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲
	<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲
	<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲
	<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲
	<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲		<input type="checkbox"/> 喜 <input type="checkbox"/> 怖 <input type="checkbox"/> 怒 <input type="checkbox"/> 悲

Fig. 1. A sample sheet of inquiry

B. Analyses of variance

We use the correct rate of cognition of an intended emotion for two types of two-way analyses of variance (ANOVA) where factors are as follows. ANOVA were used on arcsine transformed data.

- 1) Emotional intention (4 levels: joy, fear, anger, and sadness) and subject groups (3 levels: people with normal hearing abilities, hearing-impaired people of electronics major, and hearing-impaired people of design major).
- 2) Emotional intention (4 levels) and drawing groups (3 levels: people with normal hearing abilities of design major, hearing-impaired people of electronics major, and hearing-impaired people of design major).

Table II and Table III show χ^2 values of the above two types of ANOVA respectively.

Ryan's procedure shows the following results.

- In both ANOVA, the main effects of emotion shows the significant difference of the cognition of fear from all other three emotions.
- In the first ANOVA, the main effect of subject groups is significant for drawings by normal hearing abilities (de-

A.2.5. Understanding Emotion through Drawings

TABLE I
NUMBER OF SUBJECTS TO LOOK AT DRAWINGS.

Subjects	Drawing by		
	normal hearing abilities design major	hearing-impaired electronics major	hearing-impaired design major
normal hearing abilities	34	34	11
hearing-impaired electronics major	19	19	10
hearing-impaired design major	10	10	9

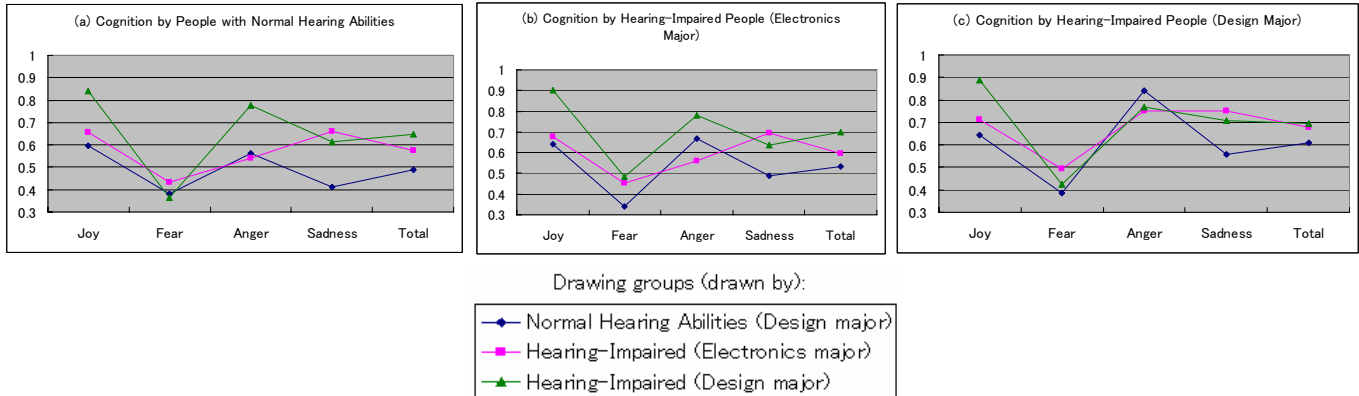


Fig. 2. Correct Rate of each subject group: (a) Cognition by people with normal hearing abilities, (b) Cognition by hearing-impaired people (electronics major), (c) Cognition by hearing-impaired people (design major)

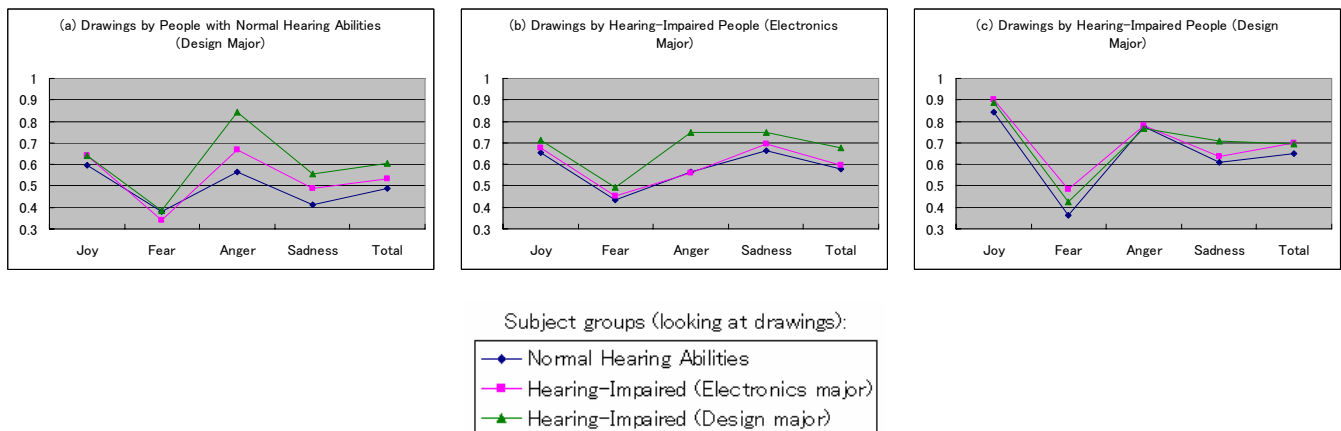


Fig. 3. Correct Rate of each drawing set: (a) Drawings by people with normal hearing abilities (design major), (b) Drawings by hearing-impaired people (electronics major), (c) Drawings by hearing-impaired people (design major)

TABLE II

χ^2 VALUES OF ANOVA (1). MAIN EFFECTS ARE EMOTIONAL INTENTION AND SUBJECT GROUPS. * SHOWS THE SIGNIFICANCE AT $p \leq .05$.

	Drawing by		
	normal hearing abilities design major	hearing-impaired electronics major	hearing-impaired design major
Main effect (A): Emotional intention	93.35*	106.42*	183.45*
Main effect (B): Subject groups	15.49*	23.11*	3.82
Interaction of the above two effects	14.32*	9.63	3.58

A.2.5. Understanding Emotion through Drawings

TABLE III

χ^2 VALUES OF ANOVA (2). MAIN EFFECTS ARE EMOTIONAL INTENTION AND DRAWING GROUPS. * SHOWS THE SIGNIFICANCE AT $p \leq .05$.

	Subjects		
	normal hearing abilities	hearing-impaired electronics major	hearing-impaired design major
Main effect (A): Emotional intention	139.40*	110.27*	96.82*
Main effect (B): Drawing groups	49.47*	42.26*	7.98*
Interaction of the above two effects	47.26*	32.10*	24.50*

sign major) and by hearing-impaired people (electronics major).

In both drawing groups, the intended emotions were recognized in the order of hearing-impaired people (design major) > hearing-impaired people (electronics major) > people with normal hearing abilities (design major) from the higher correct rate.

The procedure also shows there is the significant difference between subject groups of people with normal hearing abilities and other two groups of hearing-impaired people.

- In the second ANOVA, the main effect of drawing groups is significant for all subject groups. In all subject groups, the intended emotions were recognized in the order of hearing-impaired people (design major), hearing-impaired people (electronics major), then people with normal hearing abilities from the higher correct rate.

The procedure also shows there is the significant difference between drawing groups of people with normal hearing abilities (design major) and other two groups of hearing-impaired people.

C. Drawings of the high/low correct rate

We selected drawings by three groups that are the highest and the lowest correct rate for each emotion. It appeared that those drawings selected by three subject groups are almost the same, except the lowest correct rate drawings for fear. Figure 4 shows the selected drawings drawn by people with normal hearing abilities (design major). Selected drawings drawn by other two groups (hearing-impaired people of electronics major and design major) show the same tendency, namely drawings with the highest and the lowest correct rates for each emotion by three subject groups are almost the same.

In the selection of drawings, especially that of the highest correct rate, we have to mention that there are many more drawings of concrete objects. "Tear drops" for fear for example, in drawing sets by hearing-impaired people of electronics major and design major compared to that by people with normal hearing abilities. On the abstract level of drawings and their cognition is discussed in Section V.

IV. DISCUSSION

A. Visual information

From both ANOVA, we can distinguish the hearing-impaired people and people with normal hearing abilities in the sense of encoding an intended emotion into a drawing

and the cognition of emotion-intend drawings, even though it is not the sufficient experiment.

We may be able to assume that hearing-impaired people use different way of understanding an image or make much use of information in an image, in spite of that humans commonly have visual intelligence [8].

B. Comparison with the previous experiment with drum performances

We compare the experiment with the previous experiment of cognition of emotion-intended drum performances. In the previous experiment, we asked two groups of subjects, hearing-impaired people and people with normal hearing abilities [5][6][7] using three types of performance sets played by hearing-impaired people, people with normal hearing abilities who have no musical training (amateur), and professionals musicians.

We used the correct rate of cognition of an intended emotion for two types of two-way analyses of variance (ANOVA) where factors were (1) emotional intention (4 levels) and subject groups (2 levels: people with normal hearing abilities and hearing-impaired people) and (2) emotional intention (4 levels) and player groups (3 levels: hearing-impaired people, amateurs, and professionals).

1) *Correct rate*: In the previous experiment, the lowest correct rate was fear recognized both by hearing-impaired people and people with normal hearing abilities for all player groups. The highest correct rate varied in the previous experiment.

It indicates that fear is difficult to encode into media objects and there is less commonality on fear among people than other emotions.

2) *ANOVA*: The following is the comparison of ANOVA of the experiment with drawings and performances.

- Same as the experiment with drawings, the main effect of emotional intention was significant in the two ANOVA in the experiment with drum performances.
- A predictable result was that the main effect of hearing abilities showed the significant difference for all three player groups.
- An interesting result was that the main effect of player groups showed the significant difference for subjects with normal hearing abilities and they recognize performances by professionals, hearing-impaired people, and amateurs in the higher order. Performances by professionals were recognized significantly from other

A.2.5. Understanding Emotion through Drawings











Correct Rate	Intended Emotion			
	Joy	Fear	Anger	Sadness
The highest				
The lowest		(a)  (b)  (c) 		

Fig. 4. The Drawings of the highest and the lowest correct rate for each emotion (drawings by people with normal hearing abilities of design major). The drawings of the lowest correct rate for fear is different by each subject group, (a) people with normal hearing abilities, (b) hearing-impaired people (electronics major), and (c) hearing-impaired people (design major)

two performance sets by subjects of normal hearing abilities.

- A notable result was that subjects with normal hearing abilities differentiated the performances by professionals, while subjects of hearing-impaired do not.

C. Concrete drawing

Some of the correct rates are very high. The correct rate of the joy-intended drawings by hearing-impaired people (design major) is a typical example. It is 0.84 recognized by subjects with normal hearing abilities, 0.9 by hearing-impaired subjects of electronics major, and 0.89 by hearing-impaired subjects of design major.

A reason of this high correct rate is that the drawing sets include concrete drawings. The set includes “heart” and “the sun” for joy-intend drawings and “tear drop” for sadness-intended drawings, for example. Those concrete drawings are included mainly in drawings by hearing-impaired people, both electronics major and design major. As described in Section II-B, the number of people who drew drawings for the experiments were 14 for electronics major and 11 for design major (both hearing-impaired people). It means that there are 14 joy-intended drawings in the drawing set by electronics major.

If we exclude concrete drawings from the drawing sets, the number of drawings gets fewer. Table IV shows the number of drawings before and after extracting those drawings from the two drawing sets.

We compared the correct rate of before and after extracting drawings. For both drawing sets, the correct rate reduced. We analyzed the correct with ANOVA where main effects are subjects to look at drawings and the drawing sets before and after the extraction for drawings by electronics major and design major. The result is shown in Table V. It shows that those drawing sets before and after the extraction make the significant difference in the correct rate.

We are not sure whether to use or restrict using concrete drawings in our system. If we pursue synesthesia, then

we should exclude those drawings. If we use drawings for increasing the correct rate, those drawings can be involved.

D. Future work

In order to design and build our system for hearing-impaired people and people with normal hearing abilities to play drum performances together, we have to conduct several other analyses and experiments to understand the cognition of sound and visual objects.

- Analyze physically the drawing information. With the analysis, we can categorize emotion-intended drawings. We also need to understand the reason why some emotion-intended drawings of the highest and lowest resemble (joy and sadness) while some do not as shown in Figure 4.
- Jointly use the sound and the visual objects in an experiment to see whether the correct rate is improved by the appropriate combination of the two types of objects.
- In order to support cooperative performance, the animated image for timing guide should be investigated.

The planned system is similar to the framework of Lee’s design process analysis in which a designer is given image stimuli to generate his/her own creative work [9]. In our system, the first stimuli are images with which users render drum performances. The images is a kind of music score that guides and amplifies users’ emotion. The effective use of image and sound objects is a key to our system.

V. ACKNOWLEDGEMENT

The Japan Society for the Promotion of Science supported this research through a Grant-in-Aid for Scientific Research, Exploratory Research No. 16500138.

REFERENCES

- [1] <http://www.tsukuba-tech.ac.jp/>
- [2] P. Whittaker, *Musical Potential in the Profoundly Deaf*, Music and the Deaf, 1986.

A.2.5. Understanding Emotion through Drawings

TABLE IV

THE NUMBER OF DRAWINGS BEFORE AND AFTER EXTRACTING CONCRETE DRAWINGS

Drawn by	Emotion							
	Joy		Fear		Anger		Sadness	
	before	after	before	after	before	after	before	after
Hearing-impaired electronics major	14	9	14	10	14	10	14	5
Hearing-impaired design major	11	6	11	8	11	6	11	5

TABLE V

χ^2 VALUES OF ANOVA. MAIN EFFECTS ARE SUBJECT GROUPS AND DRAWING SETS BEFORE AND AFTER THE EXTRACTION. * SHOWS THE SIGNIFICANCE AT $p \leq .05$.

	Drawing by	
	hearing impaired electronics major	hearing impaired design major
Main effect (A): Subject groups	17.77*	4.27
Main effect (B): Before and after the extraction	7.28*	21.32*
Interaction of the above two effects	0.23	0.03

- [3] R. Hiraga and M. Kawashima, *Performance Visualization for Hearing Impaired Students*, Proc. of the 3rd International Conference on Education and Information Systems: Technologies and Applications (EISTA 2004), pp. 323–328, 2004.
- [4] P. N. Juslin and J. A. Sloboda ed., *Music and Emotion: Theory and Research*, Oxford University Press, 2001.
- [5] R. Hiraga, T. Yamasaki, and N. Kato, *Expressin of Emotion by Hearing-Impaired People through Playing of Drum Set*, Proc. of the 9th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2005), 2005.
- [6] R. Hiraga, T. Yamasaki, and N. Kato, *Cognition of Emotion on a Drum Performance by Hearing-Impaired People*, Proc. of the 11h International Conference on Human-Computer Interaction (HCII 2005), 2005.
- [7] R. Hiraga, T. Yamasaki, and N. Kato, *Communication through drum performances: Exploring the cognition of an intended emotion for a drum performance*, in preparation.
- [8] D. D. Hoffman, *Visual Intelligence: How We Create What We See*, W W Norton & Co Inc., 1998.
- [9] S. H. Lee, *A Study of Design Approach by an Evaluation based on the Kansei Information, "Images"*, doctoral thesis, University of Tsukuba, 1998.

Recognition of intended emotions in drum performances: differences and similarities between hearing-impaired people and people with normal hearing ability

Rumi Hiraga

Faculty of Information and Communication,
Bunkyo University
Chigasaki, Japan
rhiraga@shonan.bunkyo.ac.jp

Teruo Yamasaki

Faculty of Human Science,
Osaka Shoin Women's University
Osaka, Japan
yamasaki@gamelans.org

Nobuko Kato

Faculty of Industrial Technology,
Tsukuba University of Technology
Tsukuba, Japan
nobuko@a.tsukuba-tech.ac.jp

ABSTRACT

We plan to propose a performance assistance system for hearing-impaired people and people with normal hearing abilities to play music together. Using the system, people will play percussion instruments and communicate their emotions with the visual cues. To design the system, we need to understand the similarities and differences of encoding and decoding the intended emotions of drum performances between people with different hearing abilities. We describe an experiment comparing recognition of drum performances intended to express a particular emotion between hearing-impaired subjects and subjects with normal hearing abilities. The most remarkable result was that subjects with normal hearing abilities distinguished performances by professional drummers from performances done by others, while hearing-impaired subjects did not.

Keywords

Drum performance, Emotion, Hearing-impaired people

INTRODUCTION

Having taught computer music classes for six years to hearing-impaired students of Tsukuba College of Technology (now Tsukuba University of Technology [NUTUT]), we believe that the hearing-impaired people are interested in playing music. As a deaf person undertaking a music major, Whittaker argued that hearing-impaired people and people with normal hearing abilities had similar interests in music [Whittaker]. There is even a deaf professional musician

who is a percussion soloist [Evelyn Glennie].

We set as a goal proposing an assistance system to enable hearing-impaired people to play music in an ensemble with people with normal hearing abilities. Visualization of music is widely used in interactive music performance and gives hearing-impaired people more information about the music. Thus our system will assist users with performance visualization.

Since the healing effects of drum performances are known [Friedman], drum performances seem to be easier for hearing-impaired people to recognize than other musical performances, and the simplicity in playing drum instruments, we plan to use drums in our system, at least in the beginning. On the other hand, as shown in one of our previous experiments, following simple rhythms and tempos with visual cues can be somewhat burdensome for hearing-impaired people compared to cues only with sound [Hiraga2004]. For our system to be enjoyable, it must be accessible to users with little music technique and knowledge. Thus we infer that emotional communication with drum performances following their emotions might work with the performance assisting system.

To design the system, we need to understand how hearing-impaired people understand drum performances and what kinds of visual cues are useful for the system. In this paper, we describe an experiment on recognition of intended emotions expressed in drum performances by hearing-impaired subjects and subjects with normal hearing abilities. The experiment is one in a series on the recognition of sound and visual information by hearing-impaired people and people with normal hearing abilities. So far, we have conducted experiments on encoding emotions to drum performances [Hiraga2005(a), Hiraga2005(b)], encoding emotions to drawings, and decoding emotions with drawings [Hiraga2006(a)]. The most remarkable result of the current experiment was that subjects with normal hearing abilities distinguished performances by professional drummers from those by others, while hearing-impaired subjects did not.

Proceedings of the 9th International Conference on Music Perception & Cognition (ICMPC9). ©2006 The Society for Music Perception & Cognition (SMPC) and European Society for the Cognitive Sciences of Music (ESCOM). Copyright of the content of an individual paper is held by the primary (first-named) author of that paper. All rights reserved. No paper from this proceedings may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information retrieval systems, without permission in writing from the paper's primary author. No other part of this proceedings may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information retrieval system, with permission in writing from SMPC and ESCOM.

EXPERIMENT

Outline

For emotional communication through drum performances, we chose four emotions; joy, fear, anger, and sadness. We conducted an experiment to determine how hearing-impaired people and people with normal hearing abilities understand an intended emotion through a drum performance. The experiment followed the standard paradigm for most studies of emotional expression. It consisted of two steps; first subjects played drums intending to express a particular emotion, then in the second step, subjects listened to collected performances and specified which emotion they felt in response to each performance. In this paper, we focus on the second step.

We prepared sets of performances by three different types of players; (1) hearing-impaired college students, (2) people with normal hearing abilities who have no training in playing drums (we call them amateurs), and (3) professional drummers with normal hearing abilities. There were 11, 5, and 2 of these types of players respectively. Since each player did a performance for each emotion, there were 44, 20, and 8 performances in each performance set. The length of a performance varies from about 10 seconds to just over 60 seconds. Hearing-impaired college students played a MIDI drum set, Yamaha DD-55, and their performances were recorded as standard MIDI files (SMFs) with a sequence software system, Yamaha XGWorks. Other performances were played with a tam and recorded into a DAT recorder, Sony TCD-D10 PRO II, through a sound-level meter, RION NL-20.

Subjects

Two groups of subjects listened to the three performance sets consisted of two groups; (1) hearing-impaired college students and (2) college students with normal hearing abilities. There were 10 hearing-impaired subjects (9 males and 1 female, age 20-22), 15 hearing-impaired subjects (12 males and 3 female, age 20-22) for amateur and professional performances and 33 subjects with normal hearing abilities (20 males and 13 females, age 21-26). The number of subjects in each group for a performance set is listed in Table 1.

Some of the hearing-impaired subjects had played drum performances with an intended emotion about a year before the listening experiment. No subjects with normal hearing abilities had played drum performances in the first step of the drum experiment.

Hearing level and music experiment of hearing-impaired subjects

Hearing-impaired subjects were all students of the hearing-impaired division at NUTUT, which requires a hearing loss greater than 100 decibels to qualify for enrollment.

Table 1. Number of subjects for a performance set. Numbers in the parentheses show the number of players.

Subject:	performance set		
	hearing-impaired (11)	amateurs (5)	professionals (2)
hearing-impaired	10	15	15
normal hearing abilities	33	33	33

We surveyed their musical experience in terms of Karaoke, music related game such as “Dance, Dance, Revolution [DDR]” and “Drum Master [Taiko]”, dance, and music related club activity. The number of subjects who have experienced Karaoke, game, and dance is 13, 13, and 9, respectively out of 15 subjects. Three of them belonged to a dance club, two to a Japanese drum (Taiko) club, and one to a song club that uses sign language.

Procedure

We gave subjects a check sheet to mark one of the four emotions they felt to be expressed by a performance. Subjects were instructed on the use of the sheet and then listened to 72 consecutive performances for about 25 minutes. The experiment with hearing-impaired subjects was executed in a wooden-floored gymnasium. The subjects sat on the floor where there was a device that compensated for hearing loss. The experiment with normal hearing subjects took place in a classroom.

RESULTS

Correct recognition rate

Figure 1 (a) shows the rate at which hearing-impaired subjects correctly recognized the intended emotions expressed by performances by three types of players. Figure 1 (b) shows the correct recognition rate for subjects with normal hearing abilities. The correct recognition rate is shown from another perspective in Figure 2 (a), Figure 2 (b), and Figure 2 (c). They show the correct recognition rates for performances by hearing-impaired people, amateurs, and professionals respectively.

We can observe the following things from Figure 1 and 2:

- (1) Comparing Figure 1 (a) and (b), it is noteworthy that subjects with normal hearing abilities show a higher correct recognition rate of intended emotion in performances by professionals than hearing-impaired subjects do. The distinction is also clear in Figure 2 (c) where the correct recognition rates for performances by professionals by the two groups of subjects are shown.

A.2.6. Recognition of intended emotions in drum performances

ICMPC9 – International Conference on Music Perception and Cognition - Proceedings

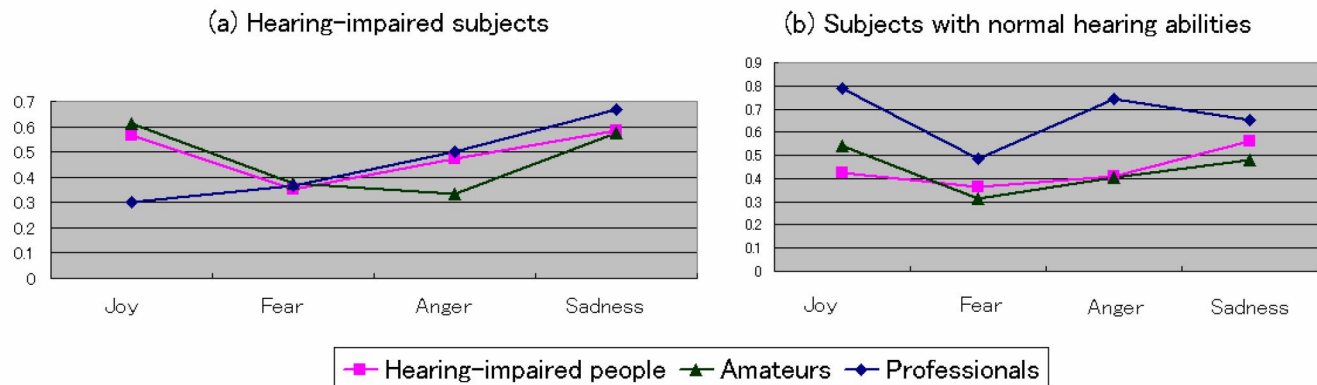


Figure 1. The correct rate of cognition of performances by three types of players.

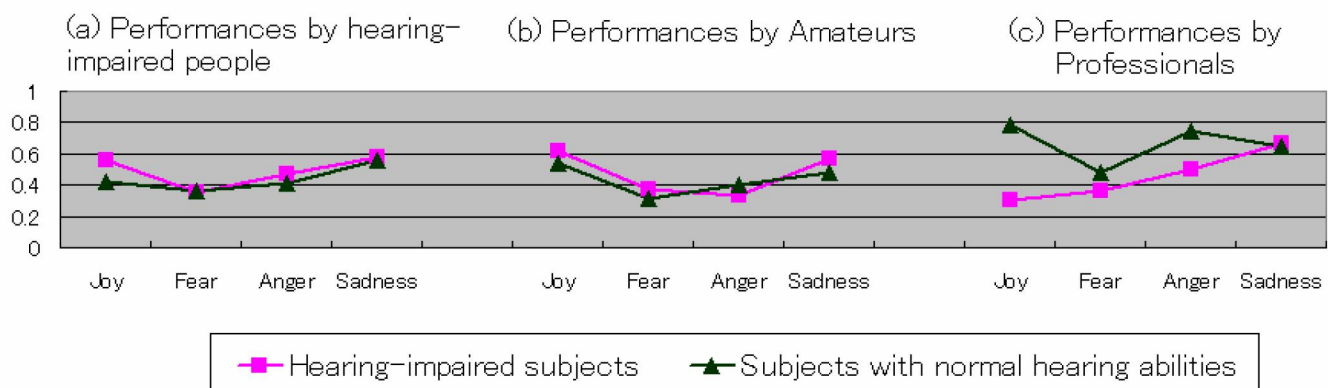


Figure 2. Comparison of the correct rate of cognition by hearing-impaired subjects and subjects with normal hearing abilities

- (2) Performances by hearing-impaired people are understood better by hearing-impaired subjects than by subjects with normal hearing abilities (Figure 2 (a)).
- (3) The correct recognition rate for fear is lowest for most of the combinations of performer and subject sets.

Analysis of variance

We used the correct recognition rate for an intended emotion for three separate two-way analyses of variance (ANOVA) where the factors were as follows. ANOVA was used on arcsine-transformed data. Significant difference is considered less than a 5 percent probability.

- (1) Emotional intention (four levels: joy, fear, anger, and sadness) and subject groups (two levels: hearing-impaired people and people with normal hearing abilities.)

- (2) Emotional intention (four levels) and performer groups (three levels: hearing-impaired people, amateurs, and professionals).
- (3) Subject groups (two levels) and performer groups (three levels).

Table 3, 4, and 5 show the χ^2 values for the three separate ANOVAs.

Table 4 shows a remarkable difference of recognition according to performance sets between subject groups. Subjects with normal hearing abilities were able to distinguish performance sets. The performance sets were recognized in order by professionals, hearing-impaired people, and amateurs based on the higher correct recognition rate. Ryan's procedure showed the following results:

- (1) The main effect of performance sets was significant differences in correct recognition rates between performances by professionals and amateurs and perform-

A.2.6. Recognition of intended emotions in drum performances

ances by professionals and hearing-impaired people, while there was no significant difference between performances by amateurs and hearing-impaired people.

- (2) In all levels of performance sets, the simple main effect of emotional intention showed significant difference.
- (3) In all levels of emotional intention, the simple main effect of performance sets showed significant difference.

Table 2. The number shows the subjects who felt easy or difficult to recognize each emotion (hearing-impaired subjects/subjects with normal hearing abilities).

	Joy	Fear	Anger	Sadness
The easiest	7/10	1/2	1/16	3/5
The most difficult	0/11	8/14	3/2	2/6

Self judgment

We did a post-experiment inquiry in which we asked how difficult the subjects found the experiment. Subjects checked one of the five levels of difficulty (5 for “very easy” and 1 for “very difficult”), the easiest emotion to recognize through the experiment, and the most difficult emotion. They checked an item after the experiment. Though not all of the hearing-impaired subjects answered, one subject checked 5 (very easy), one checked 4, three checked 3, five checked 2, and one checked 1. The number of subjects with normal hearing abilities checked each level was 1, 0, 0, 24, 8 from very easy to very difficult.

Table 2 shows the number of subjects that found it easy or difficult to recognize an intended emotion in performances. Joy was the easiest emotion to recognize for seven subjects and fear was the most difficult for eight for hearing-impaired subjects.

Subjects with normal hearing-abilities described the difficulties in distinguishing “joy and anger” (15 subjects), “sadness and fear” (24 subjects), and “anger and fear” (12 subjects). Ryan’s procedure for ANOVA No. 2 (main effects are emotional intention and performer group) showed that there were no significant difference between the cognition of “joy and anger” and “anger and fear” in the multi-level comparison on the main effect of emotional intention and the multilevel comparison on the all performer groups. It also showed that there was significant difference between the recognition of “sadness and fear.”

DISCUSSION

Distinguishing performances by professionals

It is noteworthy that subjects with normal hearing abilities distinguished performances by professionals from those by other performers while hearing-impaired subjects did not.

Although we need to analyze the sound data of drum performances physically and compare them among performer groups, the ways of expressing intended emotions were similar across the three types of performers; for example, playing louder and faster when expressing anger. This might suggest that the way of encoding an emotion into a drum performance is similar whether or not a person has a hearing impairment. On the other hand, many of the people with normal hearing abilities felt that performances by professionals are somehow different from performances by other types of performers, just by listening to them. If the decoding process used a similar kind of information as the encoding, then hearing-impaired subjects use the same kind of information, in both the encoding and the decoding process, while subjects with normal hearing abilities utilize another kind of information in decoding.

To investigate this phenomenon, we can conduct an experiment with artificially generated music performances with characteristics used to express each emotion, then gradually change the value of performance characteristics, and observe the difference in correct recognition rates between hearing-impaired subjects and subjects with normal hearing abilities.

Low correct recognition rate for fear

The lowest correct recognition rate among four emotions was for fear, except for performances by professionals listened to by hearing-impaired subjects. Another experiment of recognition of intended emotion in drawings [Hiraga2006(a)] showed the lowest correct rate was for fear for any combination of drawing sets (drawn by people with normal hearing abilities who undertaking design major, hearing-impaired people undertaking electronics major, and hearing-impaired people undertaking design major) and subject groups (subjects with normal hearing abilities, hearing-impaired subjects undertaking electronics major, and hearing-impaired subjects undertaking design major). This implies that fear is difficult to encoding into, at least two media; music and drawing.

Figure 3 shows the distribution of correct recognition rates. It shows that only the correct rate of fear is distorted to lower values.

Future work

After conducting an experiment on recognition of drawings by subjects with different hearing abilities [Hiraga2006(a)], we conducted an experiment of giving subjects stimuli of combined information; performance with drawings and performance with motion pictures [Hiraga2006(b)]. Still, there is a lot to do to design our system.

A.2.6. Recognition of intended emotions in drum performances

ICMPC9 – International Conference on Music Perception and Cognition - Proceedings

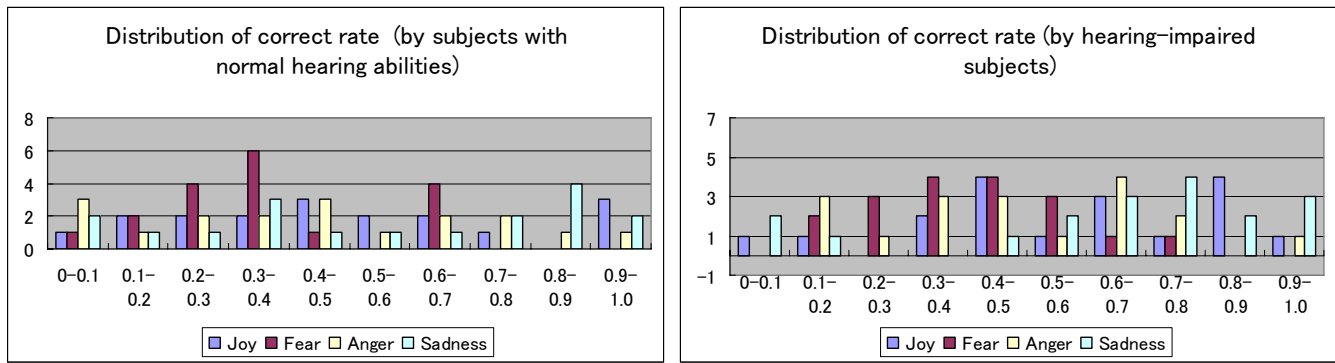


Figure 3. Distribution of correct recognition rate

- (1) We must physically analyze performances and clarify performance characteristics according to the intended emotion. Then we can confirm that the encoding rule between hearing-impaired people and people with normal hearing abilities is similar and that is the rule by nature.

The analysis for drawings that we obtained in another experiment is also required.

- (2) Artificially generate music performances and drawings with intended emotions. Then we will conduct an experiment with the two kinds of information as stimuli for subjects to recognize emotion.
- (3) Find out the appropriate mapping and timing of combining visual and sound information.
- (4) Investigate recognition according to level of impairment.
- (5) Investigate how the encoding and decoding processes for hearing-impaired people may change after the training of drum performances.

Though there is a thorough survey of emotion and its communicability in music [Juslin], there has been little research on emotional communication through music dedicated to drum performances [Yamasaki2004] or emotional communication with hearing-impaired people.

Our research found an interesting phenomenon with regard to recognition of drum performances by professionals – hearing-impaired people do not distinguish the performance from performances by amateurs and hearing-impaired players while people with normal hearing abilities do.

ACKNOWLEDGMENTS

The Japan Society for the Promotion of Science supported this research through a Grant-in-Aid for Scientific Research, Exploratory Research No. 16500138.

REFERENCES

[NUTUT] National University Corporation Tsukuba University of College. <http://www.tsukuba-tech.ac.jp>

[Whittaker] P. Whittaker. Musical potential in the profoundly deaf. *Music and the Deaf*, 1986.

[Evelyn Glennie] Evelyn Glennie, <http://www.evelyn.co.uk/homepage.htm>

[Friedman] R. L. Friedman. *The Healing Power of the Drum*. White Cliffs Media, 2000.

[Hiraga2004] R. Hiraga and M. Kawashima. Performance Visualization for Hearing-Impaired Students. *Proc. of EISTA 2004*, 2004.

[Hiraga2005(a)] R. Hiraga, T. Yamasaki, and N. Kato. Expression of Emotion by Hearing-Impaired people through Playing of Drum Set. *Proc. of WMSCI 2005*, 2005.

[Hiraga2005(b)] R. Hiraga, T. Yamasaki, and N. Kato. Cognition of Emotion on a Drum Performance by Hearing-Impaired People. *Proc. of HCII 2005*, 2005.

[Hiraga2006(a)] R. Hiraga and N. Kato. Understanding emotion through drawings – comparison between people with normal hearing abilities and hearing-impaired people. In preparation.

[DDR] Dance Dance Revolution Freak, <http://www.ddrfreak.com/>

[Taiko] Drum Master, <http://www.namco.com/games/taiko/>

[Hiraga2006(b)] R. Hiraga and N. Kato. Understanding emotion through multi-media – comparison between people with normal hearing abilities and hearing-impaired people. In preparation.

[Juslin] P. N. Juslin and J. A. Sloboda ed. Communicating Emotion in Music Performance: A Review and Theoretical Framework. in *Music and Emotion: Theory and Research*. Oxford University Press. 2001.

A.2.6. Recognition of intended emotions in drum performances

ICMPC9 – International Conference on Music Perception and Cognition - Proceedings

[Yamasaki2004] T. Yamasaki. Emotional Communication through Performance played by Young Children. *Proc. Of ICMPC 2004*, 2005.

Table 3. χ^2 values of ANOVA (1). Main effects are emotional intention and subject groups. * shows the significance at $p \leq .05$.

	Performances by		
	Hearing-impaired people	Amateurs	Professionals
Main effect (A): Emotional intention	33.194*	12.750*	10.406*
Main effect (B): Subject groups	3.937*	0.477	14.896*
Interaction of the above two effects	4.284	1.418	12.009*

Table 4. χ^2 values of ANOVA (2). Main effects are emotional intention and performer groups. * shows the significance at $p \leq .05$.

	Subjects	
	Hearing-impaired subjects	Subjects with normal hearing abilities
Main effect (A): Emotional intention	26.803*	30.112*
Main effect (B): Performer groups	0.546	75.955*
Interaction of the above two effects	11.077	25.745*

Table 5. χ^2 values of ANOVA (3). Main effects are subject groups and performer groups. * shows the significance at $p \leq .05$.

	Emotional intention			
	Joy	Fear	Anger	Sadness
Main effect (A): Subject groups	4.449*	0.232	2.195	0.859
Main effect (B): Performer groups	2.166	2.377	1.639	5.739
Interaction of the above two effects	38.878*	2.507	6.530*	0.593

Performance Visualization for Hearing-Impaired Students

Source	Journal of Systemics, Cybernetics and Informatics, Vol.3, No.5, pp.24-32, 2005
--------	---

Performance Visualization for Hearing-Impaired Students

Rumi HIRAGA

Faculty of Information and Communication, Bunkyo University
1100 Namegaya, Chigasaki 305-0032, Japan

and

Mitsuo KAWASHIMA

Faculty of Industrial Technology, Tsukuba University of Technology
4-3-15 Amakubo, Tsukuba 305-0005, Japan

ABSTRACT

We have been teaching computer music to hearing impaired students of Tsukuba College of Technology for six years. Although students have hearing difficulties, almost all of them show an interest in music. Thus, this has been a challenging class to turn their weakness into enjoyment. We thought that performance visualization is a good method for them to keep their interest in music and try cooperative performances with others. In this paper, we describe our computer music class and the result of our preliminary experiment on the effectiveness of visual assistance. Though it was not a complete experiment with a sufficient number of subjects, the result showed that the show-ahead and selected-note-only types of performance visualization were necessary according to the purpose of the visual aid.

Keywords: Hearing Impaired, Computer Music, Music Performance, and Visual Feedback.

1. INTRODUCTION

We have been teaching computer music to hearing impaired students of Tsukuba College of Technology (now National University Corporation, Tsukuba University of Technology) for six years. Students with the hearing impairments of more than 100 decibels are qualified to enter the college and get a quasi-bachelor degree in three years. They learn architecture, design, computer software, or computer hardware according to their major to obtain useful skills. This style resembles that of Gallaudet University and the National Technical Institute for the Deaf at the Rochester Institute of Technology (NTID).

There are many professional musicians with visual impairments, moreover, there are several activities to assist those people with computer software such as WEDELMusic [1]. Though it is not surprising that there are very few professional musicians with hearing impairments, the number of them is not zero. Some of them are talented deaf musicians, like Evelyn Glennie, a famous percussion soloist, who even has absolute pitch.

The computer music class is open to students of all specialties but mainly those of the computer hardware course have taken the class. This is not a required subject. Not necessarily all the professors at the college agree on the importance of the class. On the other hand, we came to know that not a small number of students have an interest in music, independent of the degrees of their handicap and personal experience with music. Thus

given the computer assistance for them to understand and enjoy music, their quality of life (QOL) is considered to be improved. We thought performance visualization would be a good method for such assistance. Since the research of performance visualization is not a mature area and currently there is no suitable user interface to assist students, we need a good performance visualization system for them. In order to design and build such a system, we conducted a preliminary experiment on cooperative musical performance using visual assistance.

2. COMPUTER MUSIC CLASS

We set the purpose of the computer music class to allow students to understand and enjoy music in order to broaden their interest [2]. In other words, the class was more music oriented (and amusement oriented), not computer oriented. Considering that the class should meet the requirements of the college, especially for the computer hardware course, the purpose above is not necessarily appropriate. The reason for setting such a purpose is to get rid of the difficulty of keeping students' interest, especially in an area that they have not experienced much in their lives. If we start teaching them from computer perspective such as the structure of synthesizers or the format of Musical Instrument Digital Interface (MIDI), a digital format for performance data, they will have conversations in sign language, or even worse, no students may register for the class.

Making students continue to move their bodies with music is the most effective way to keep the class active. Thus, the computer has been used as a tool for assisting them in enjoying music in the class, not as a tool with which to develop new computer music software or hardware systems.

Materials

Because it was not possible for teachers who did not receive special education in music to teach conventional acoustic musical instruments to students, we benefited from the newly developed MIDI instruments. Furthermore, we were able to connect several machines and instruments with MIDI. A MIDI instrument generates MIDI data when a player plays it. It has a MIDI terminal to connect with another MIDI instrument or a PC. It needs a sound generator either inside or outside the instrument.

A.2.7. Performance Visualization for Hearing-Impaired Students (revision of A.2.1.)



Figure 1. Taiko performance with Miburi



Figure 2. Batucada performance

The following are the hardware and software systems we used in the class.

- Miburi R2 (Yamaha): A MIDI instrument with sensors. Sensors are attached to a shirt which a performer puts on. When a performer moves his/her elbows, shoulders, wrists, and legs, sound that corresponds to the position and its movement is generated from the sound generator of Miburi. The sound generator provides several drum sets, tonal sound colors, and SFX sound (murmuring sound of a stream, the sound of gun fire, a bird song, etc.).

The good points in using Miburi for students were as follows:

- With simple body movement, you can generate music.
- It is a new instrument in which playing methods are not difficult and not established.
- Miburi performers can communicate by looking at each other's movement.
- Since MIDI data is generated by playing the Miburi, their movement is reflected synchronously to visualization if systems are connected. Through the visualization, students understand their movement and its result as music.
- XGWorks (Yamaha): A sequence software system to make performance data in MIDI.
- VISISounder (NEC): An animation software system whose action is controlled by MIDI data. It prepares several scenes and animation characters beforehand. For example, a frog at a specific position in the display jumps when a sound "C" comes, while another frog jumps with "G." Using this software, students were directly able to feel their performance with Miburi through visualization. They liked it very much.
- MIDI drum set and MIDI keyboard (Yamaha): MIDI instruments.
- Music table (Yamaha): A MIDI instrument, originally designed for music therapy for elderly people. Pads are arranged on the top of the table on which people pat. There is a guide lamp for indicating the beat.

Though we tried an actuator that is used inside a speaker system for the haptic feedback purpose, it was not suitable to use because it heats up as sound was generated.

An unfortunate thing in using these products is that some of them had a short life. In the past six years when we taught the class, Miburi and VISISounder, which were the most suitable materials for the class, disappeared from the market. Although there are several other MIDI instruments and animation systems with MIDI data at the research level, products are more reliable and end user oriented.

Students' presentation

The class is held in a school term. There were ten or eleven weeks in a term. Every year we asked students to make a musical presentation at the final class. The following is an excerpt from the list of students presentation.

- A dramatic play using Miburi. Accompanied by SFX sounds, a student played out her daily life in sign language. For example, the barking of the dog was heard accompanying the sign language for a dog made by wrist movement.
- A music performance using Miburi. With a tonal sound, a student played the "Do-Re-Mi song." Her performance controlled characters of VISISounder.
- A Japanese Taiko (drum) performance using Miburi. Figure 1 shows the performance. Though it is a completely virtual performance, the change of drum sets was musically very effective. Usually a Taiko player uses one to three Taikos in an actual play, a player with Miburi can use many more types of Taiko as if all of them are around him/her.
- Samba performance using a Music table and a drum set. Seven students played three different rhythm patterns that cooperatively made Samba percussion performances (Batucada). Figure 2 shows the performance. One student stood up and played as a conductor by performing a basic rhythm pattern. Playing Batucada gave students the sense of unity in music.
- Some students used sequence software in order to perform accompaniment music for Karaoke. They sang using the sign language accompanied by the music.

After their presentations, many students indicated on a questionnaire that they would like to play in an ensemble or they enjoyed playing with other students.

A.2.7. Performance Visualization for Hearing-Impaired Students (revision of A.2.1.)

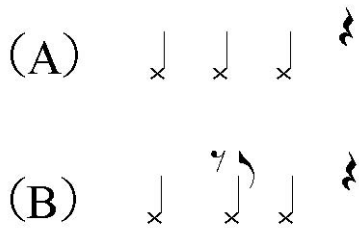


Figure 3. Rhythm A and B

3. RELATED WORKS

Although there are several studies of aiding visually handicapped people in their musical activities, there are very few for hearing impaired people. We conducted the experiment described in Section 4 from the viewpoint of performance visualization. Thus, in this section, we describe research on performance visualization.

Sobieczky visualized the consonance of a chord on a diagram based on a roughness curve [3]. Hiraga proposed using simple figures to help users analyze and understand musical performances [4][5][6]. Smith proposed a mapping from MIDI parameters to 3D graphics [7]. Foote's checkerboard type figure [8] shows the resemblance among performed notes based on the data of a musical performance. 3D performance visualization interface is proposed for users to browse and generate music using a rich set of functions [9][10].

These performance visualization works have different purposes such as for performance analysis and sequencing. So far, there has been no work for cooperative musical performance.

4. EXPERIMENT

Outline

In order to determine a more suitable visualization interface for performance feedback to support cooperative musical performances by hearing-impaired people, we investigated the characteristics of animated performance visualization proposed by commercial systems and a prototype system by a student. The investigation was done by a usability test of each performance visualization.

The purpose of the test was to see the playing timing of each subject with a guided animation that is controlled by MIDI data of a model performance. Namely, subjects played a MIDI instrument looking at the animation and their performances were recorded, then we compared the performances with a model performance. The time differences were calculated between the onset time of subjects' performances and the model performance. Onset time is the moment a note is played by a keyboard or a drum. It is the time of a MIDI message of "Note On" is generated. The message includes the note number (pitch) and the velocity (volume) of the note on. We can see from the note number which drum pad is patted or which key on a keyboard is played.

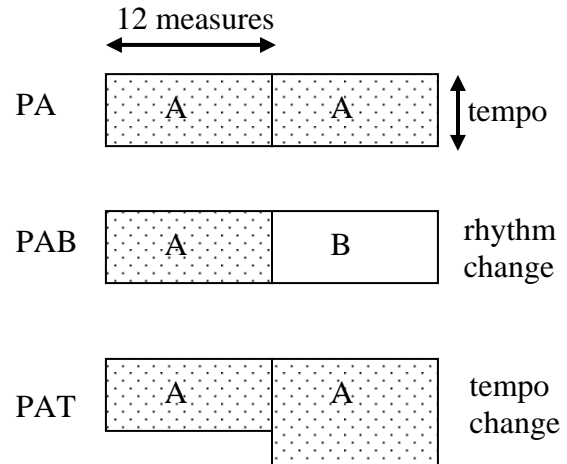


Figure 4. Three types of model performance: PA, PAB and PAT

Subjects

Three students (call them SA, SB, and SC) and a technical staff member (call her SS) were the subjects of the experiment. Students were in a sense exceptional among all students regarding their musical experience because two of them were members of a pop music club and had performance experiences, and the other had been learning play the piano for six years. They were assigned different instruments and tried to play cooperatively with a model performance using feedback.

Model performances

We used two rhythm patterns, A and B (Figure 3), then prepared three types of model performance, PA, PAB, and PAT, by combining them (Figure 4). PA repeats rhythm A for twenty-four measures with tempo $MM=108$. $MM=108$ means that there are a hundred and eight beats in a minute, namely a beat takes 0.556 ($60/108=0.556$) second. The larger the MM number, the faster the tempo. PAB repeats rhythm A for twelve measures then changes rhythm to B for another twelve measures with the constant tempo $MM=108$. PAT repeats rhythm A for twenty-four measures with tempo $MM=108$ for the first half, then with tempo $MM=140$ for the second half.

Feedback types

The experiment used four types of feedback: three types of visual feedback and one type of sound feedback. These feedback types were exclusively given to subjects. They were as follows.

1. VISISounder.

We used a scene that clearly showed the difference among performed notes by the movement of characters (either a monkey or frogs) (Figure 5). A monkey in the center corresponds to the performance of a model performance and frogs to those by subjects. Characters pop up when an instrument is played. Since a frog character was assigned to individual subject, we could distinguish subjects through the animation.

A.2.7. Performance Visualization for Hearing-Impaired Students (revision of A.2.1.)

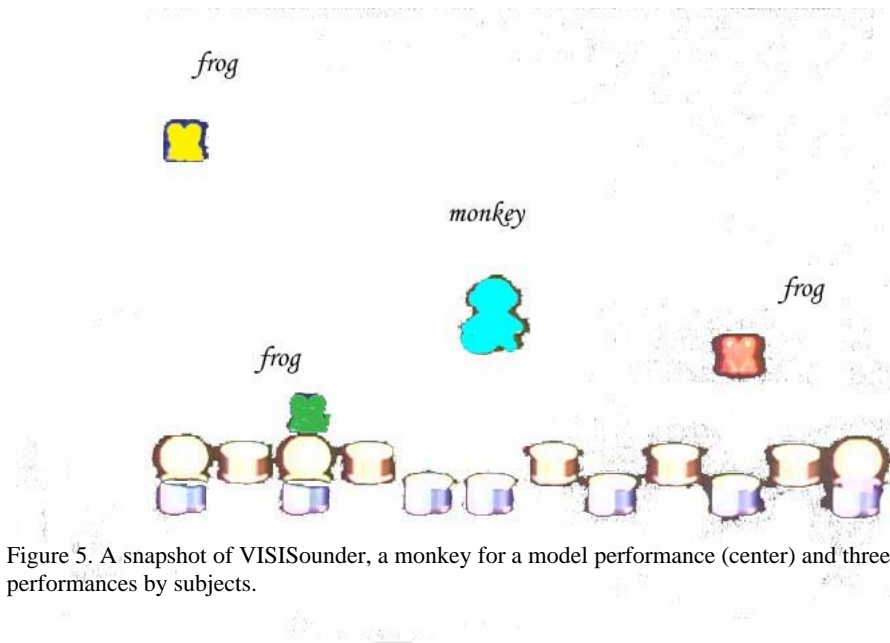


Figure 5. A snapshot of VISISounder, a monkey for a model performance (center) and three frogs for performances by subjects.

2. XGWorks.

Although XGWorks has several visualization forms for performance, we used a “drum window” (Figure 6). In the drum window, each line corresponds to a type of drum, such as a Conga. When a rhythm changes or a tempo changes, a drum used by a model performance changes accordingly. A cursor indicated the place of a model performance on the display.

A big difference in the visualized performance on XGWorks from the other two types of visual feedback is that subjects are able to predict the rhythm (show-ahead feedback). In PAB, the rhythm change from the thirteenth measure was shown on the display, therefore, subjects could see the change of the rhythm before the cursor came to the position. Although the tempo change was also indicated by using a different type of drum, the degree of tempo change could not be shown.

Other differences are that a model performance is shown as continuous cursor movement, and performances by subjects are not shown on the window.

3. Virtual Drum.

Virtual Drum is a program using direct API calls and Mabry Visual Basic MIDI IO controls, originally freeware [11]. A student partially modified the program in order to make it a game program for scoring a performer's playing timing with respect to a model performance.

In Virtual Drum, a model performance appears in the upper boxes and performances by subjects in the lower boxes (Figure 7).

4. Sound only.

The model performance is not visualized but only performed.

Sessions

Combining three types of model performance and four types of feedback, the experiment consisted of twelve sessions as shown in Table 1. Subjects were informed about the twelve sessions and practiced PA, PAB, and PAT only by clapping by themselves without a model performance before the experiment.

5. RESULT

We obtained the time difference between a subject performance and the model performance. The average and standard deviation of time difference for a session were calculated using the performed beats in twenty-four measures by all subjects as shown in Table 2.

The average of the time difference between a subject performance and the model performance for each beat was shown as a line graph for the rhythm patterns of PA (Figure 8), PAB (Figure 9), and PAT (Figure 10). Each line shows a session whose name is specified in Table 1. In the graphs, X-axis showed the beat. Since three notes were performed in every measure of the two rhythm patterns, beat number four was not the fourth beat of the first measure but the first beat of the second measure. Therefore, the beat number thirty seven (the first beat of the thirteenth measure) was the changing point of the rhythm in PAB and the tempo in PAT.

The Y-axis showed the time difference counted by “ticks.” In the experiment, a beat consisted of 480 ticks. Therefore, tempo MM=108 meant a beat was played every 556 ms ($60/108=0.556$) and a tick roughly corresponded to 1 ms ($60/108/480=0.00118$).

A.2.7. Performance Visualization for Hearing-Impaired Students (revision of A.2.1.)

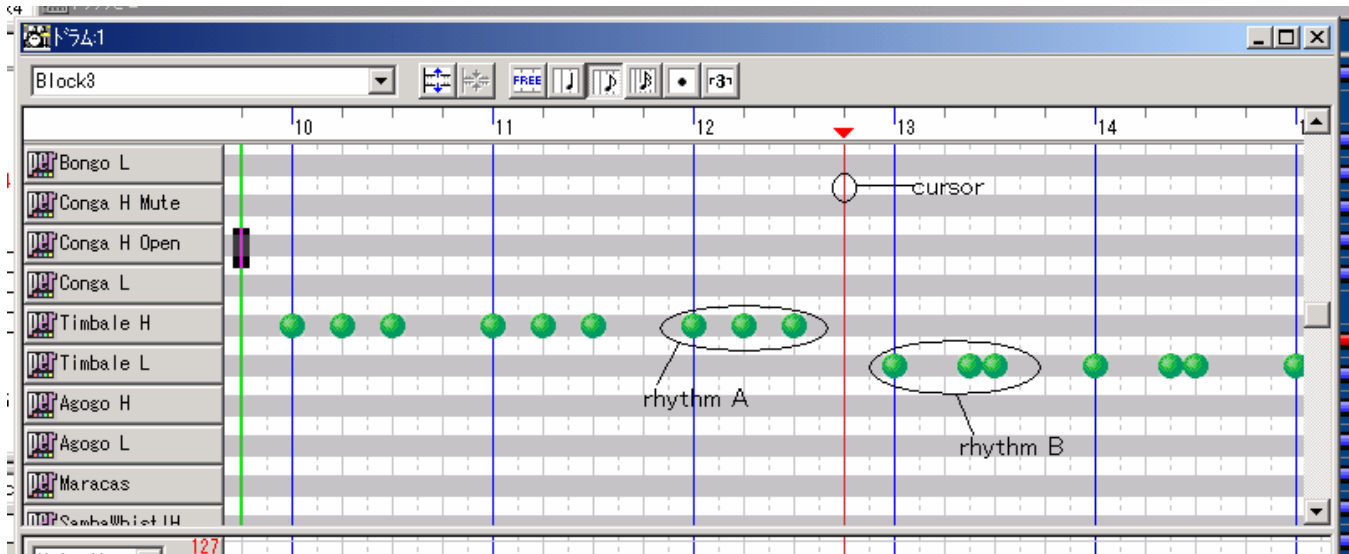


Figure 6. XGWorks. Rhythm changes from A to B at the thirteenth measure

The results from Table 2 and the figures are as follows.

1. VISISounder.

The average and the standard deviation of the feedback of VISISounder are rather large.

2. XGWorks.

Both the average and standard deviation of the feedback of XGWorks compare fairly well to those of other types of feedback.

3. Virtual Drum.

Though with a small average, Virtual Drum has the largest standard deviation. This means the subjects' performances waver.

4. Sound feedback.

The smallest standard deviation value was obtained from the sound feedback for two of the three model performances. This is also found in the small movement of a line for the session A*Sound (* is either null, "B", or "T") in the three graphs (Figures 8, 9, and 10). On the other hand, the sound feedback average is rather large.

The average and standard deviation of four measures before and after the rhythm change and the tempo change, namely the ninth to twelfth measures and thirteenth to sixteenth for PAB and PAT are shown in Table 3. Data of the ninth to twelfth measures show the steadiness of subjects performances after performing several repeats of a rhythm pattern with a regular tempo.

For the rhythm change (PAB), ABVISI made a big difference before and after the change, while for the tempo change (PAT), ATXGW made a big difference.

6. DISCUSSION

In discussing the time, we have to notice the basic numbers, such as, we are able to perceive multiple vocalizations when the time lag is over 20 ms or due to MIDI hardware and display redrawing. In the experiment, we do not need to take those numbers into consideration, because the precision of the time is the next step. Here we would like to see the tendency between the subjects' performances and feedback types.

We are able to see that the sound method gives better feedback from the point of view of standard deviation than other types of feedback from the result shown in Table 2. It can be interpreted that once subjects form the performance model of the rhythm and tempo within themselves, it is more comfortable and easier for subjects to keep playing it. Of course, this result comes from the fact that the subjects are less impaired. The next good result is using the feedback of XGWorks. In spite of this, subjects did not appreciate the show-ahead of tempo and rhythm with the moving cursor of XGWorks. On the other hand, we are also able to see in Table 3 that the show-ahead visualization by showing the change in rhythm and tempo is useful as judged from the result of the smallest standard deviation obtained using XGWorks for PAB and PAT. Though with the worse result, they well appreciate the animation of VISISounder. These observations show that it is important to show something fun in the visual aid for cooperative performance.

From the experimental results, we came to the conclusion that the important thing in designing performance visualization for cooperative performance is the show-ahead of the tempo. Animation that shows only the important notes for cooperation concerning musical structure will reduce the physical burden.

A.2.7. Performance Visualization for Hearing-Impaired Students (revision of A.2.1.)

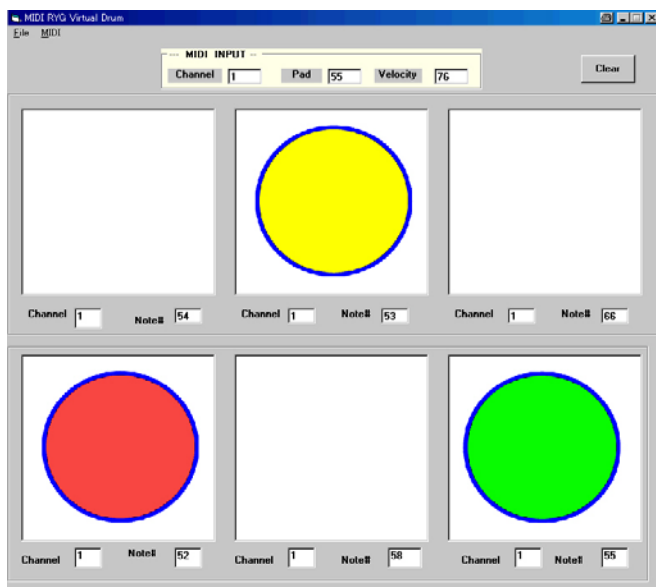


Figure 7. Virtual Drum: a model performance (circle above) and performances by subjects (two circles below)

Visualization with the purpose of game animation is not suitable for accompaniment. Performance visualization should be designed according to its purposes. The new user interface will be the combination of continuous information for the tempo and discrete information of the musical structure. The following is future work.

- Since the experiment was with a small number of subjects and not a variety of subjects, we need to ask more people with different musical experiences and levels of hearing impairment.
- We have to make it clear how long the subjects are affected by the change of rhythm or tempo.
- On the questionnaire after the experiment, subjects made comments on four types of feedback. They say looking at the display for the movement makes them fatigued. Therefore, we should take the physical burden caused by the feedback into consideration. Also, we should notice that animation should not always be given attention.
- Besides, in order to create less physical burden because of the reason above, there are other good reasons to visualize a part of the performance. They are (1) not all notes in a musical piece are given the same role and importance, and (2) a report by a music researcher indicated that a phrase can be analyzed to a tree structure according to the degree of prominence of each note [12]. The prominence of notes gives performers important information on performance. Therefore, a possible new performance visualization could show animation only at important notes, such as the first beat of every or every other measure.

- Though we could see that the showing-ahead type of performance visualization is effective as far as the tempo is regular, the sudden change in the cursor movement of XGWorks according to the tempo change is difficult to follow for subjects. A reason for the difficulty is that the movement is different from that of a human conductor who controls tempo smoothly. It is necessary to suggest the change of tempo in a smoother manner by referring to the movement of a human conductor.

7. ACKNOWLEDGEMENTS

We appreciate Y. Ichikawa for her great support in preparing the musical instruments, data, and many other things. This work was supported by The Ministry of Education, Culture, Sports, Science and Technology through a Grant-in-Aid for Scientific Research, No. 14580243.

8. REFERENCES

- [1] WEDELMusic, <http://www.wedelmusic.org/>
- [2] R. Hiraga and M. Kawashima, "Computer music for hearing-impaired students", **Technical Report of SIGMUS**, IPSJ, no. 42, 2001, pp. 75-80.
- [3] F. Sobieczky, "Visualization of roughness in musical consonance", **Proceedings of IEEE Visualization**, IEEE, 1996.
- [4] R. Hiraga, "Case study: A look of performance", **Proceedings of IEEE Visualization**, IEEE, 2002, pp. 501-504.
- [5] R. Hiraga, S. Igarashi, and Y. Matsuura, "Visualized music expression in an object-oriented environment", **Proceedings of International Computer Music Conference**, ICMA, 1996, pp. 483-486.
- [6] R. Hiraga and N. Matsuda, "Visualization of music performance as an aid to listener's comprehension", **Proceedings of Advanced Visual Information**, 2004.
- [7] S. M. Smith and G. N. Williams, "A visualization of music", **Proceedings of IEEE Visualization**, IEEE, 1997.
- [8] J. Foote, "Visualizing music and audio using self-similarity", **Proceedings of ACM Multimedia99**, ACM, 1999, pp. 77-80.
- [9] R. Hiraga, R. Miyazaki, and I. Fujishiro, "Performance visualization -- a new challenge to music through visualization", **Proceedings of ACM Multimedia02**, ACM, 2002, pp. 239-242.
- [10] A. Watanabe and I. Fujishiro, "tutti: a 3D interactive interface for browsing and editing sound data", **The 9th Workshop on Interactive System and Software**, JSSST, 2001.
- [11] Gould Academy, http://intranet.gouldacademy.org/music/faculty/virtual/virtual_instruments.htm
- [12] F. Lerdahl and R. Jackendoff, **Generative Theory of Tonal Music**, The MIT Press, 1983.

A.2.7. Performance Visualization for Hearing-Impaired Students (revision of A.2.1.)

	VISI	XGW	VD	Sound
PA	AregVISI	AregXGW	AregVD	AregSound
PAB	ABVISI	ABXGW	ABVD	ABSOUND
PAT	ATVISI	ATXGW	ATVD	ATSound

Table 1. Twelve experimental sessions. Feedback types are abbreviated as follows: VISI for VISISounder, VD for Virtual Drum, and Sound for sound only.

PA (rhythm A, tempo regular)

	AregVISI	AregXGW	AregVD	AregSound
Average	165.36	5.41	21.19	40.69
Std. dev.	77.44	49.28	177.19	55.84

PAB (rhythm A and B, tempo regular)

	ABVISI	ABXGW	ABVD	ABSOUND
Average	54.39	33.69	-14.65	63.33
Std. dev.	92.12	61.83	122.44	40.74

PAT (rhythm A, tempo changes)

	ATVISI	ATXGW	ATVD	ATSound
Average	70.13	56.23	22.13	79.31
Std. dev.	85.34	64.04	127.87	29.73

Table 2. The average and standard deviation of the twelve sessions.

PAB (rhythm A and B, tempo regular)

measure 9-12				measure 13-16				
ABVISI	ABXGW	ABVD	ABSOUND		ABVISI	ABXGW	ABVD	ABSOUND
-16.79	16.23	-11.50	48.11	Average	41.19	69.88	23.94	70.64
51.29	34.40	196.85	11.66	Std. Dev.	151.92	79.07	66.35	82.66

PAT (rhythm A, tempo changes)

measure 9-12				measure 13-16				
ABVISI	ABXGW	ABVD	ABSOUND		ABVISI	ABXGW	ABVD	ABSOUND
97.60	26.65	-1.06	54.39	Average	122.75	124.19	64.36	105.31
31.02	30.23	114.63	16.56	Std. Dev.	116.11	69.42	117.20	37.81

Table 3. The average and standard deviation of four measures before and after the rhythm change (above) and the tempo change (below).

A.2.7. Performance Visualization for Hearing-Impaired Students (revision of A.2.1.)

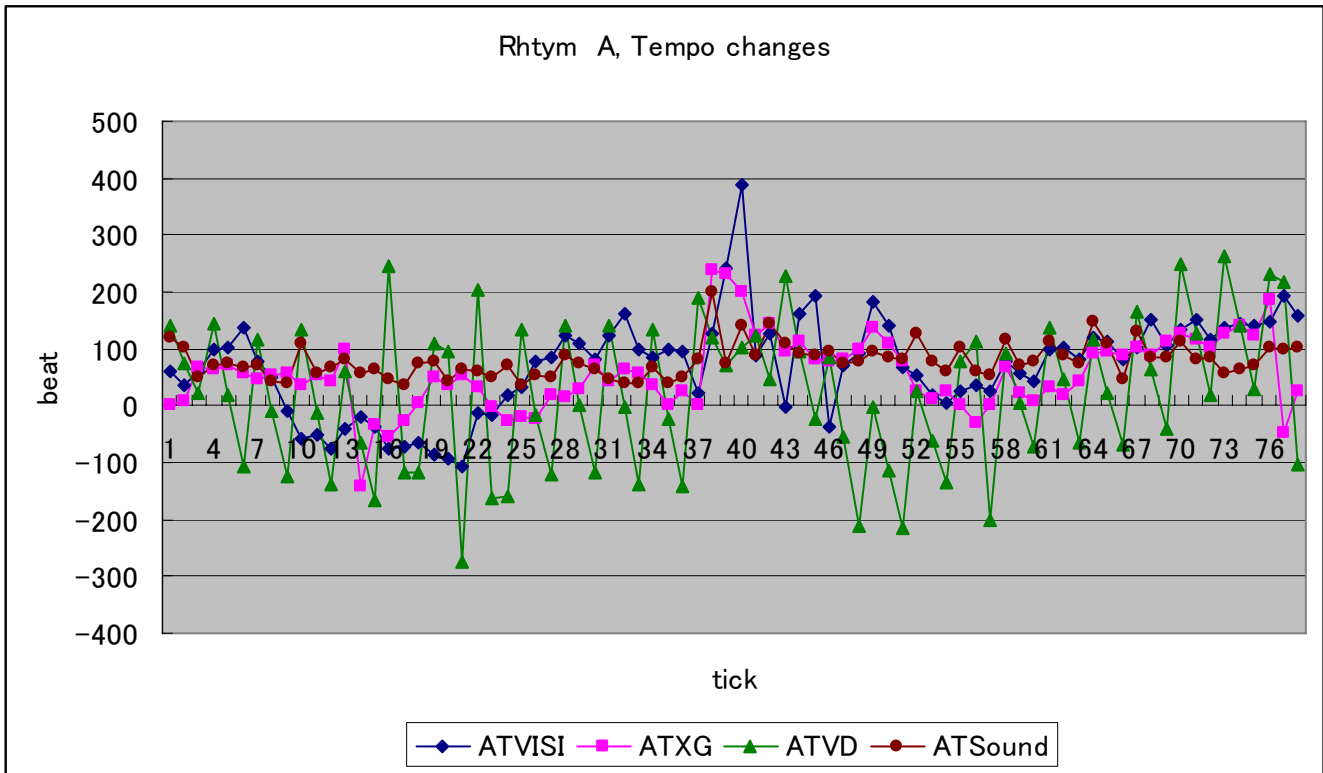


Figure 8. PA (rhythm A, tempo regular)

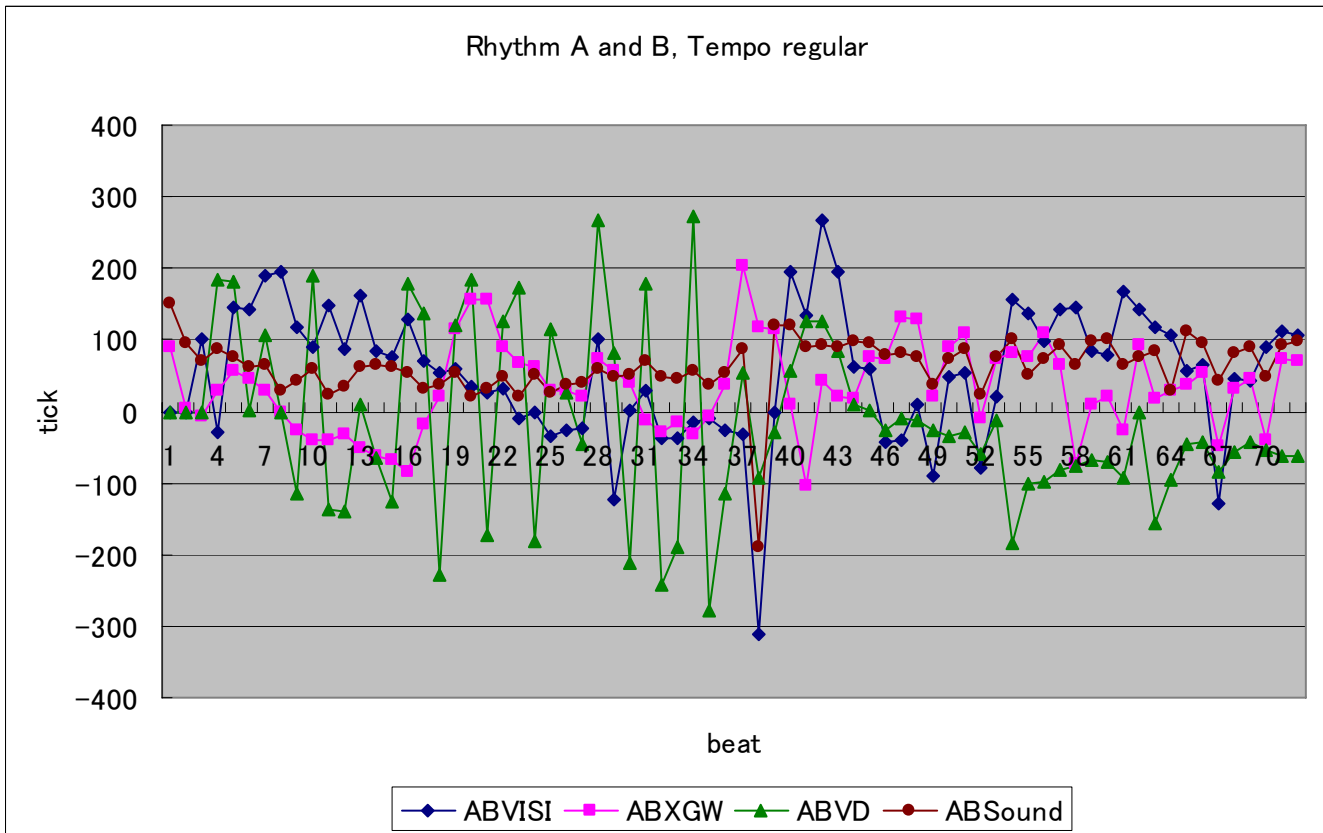


Figure 9. PAB (rhtym A and B, tempo regular)

A.2.7. Performance Visualization for Hearing-Impaired Students (revision of A.2.1.)

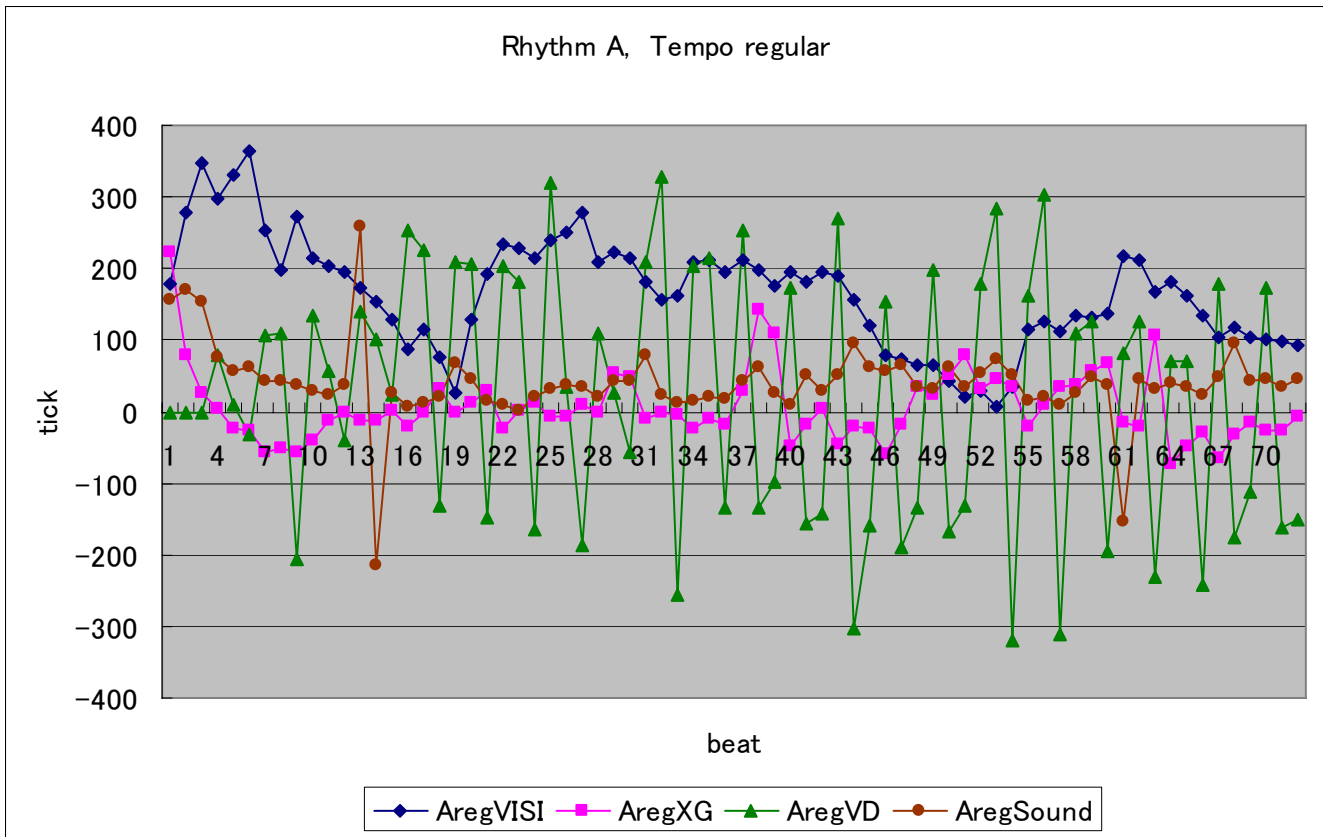


Figure 10. PAT (rhythm A, tempo changes)

A.2.8. The catch and throw of music emotion by hearing-impaired people

The catch and throw of music emotion by hearing impaired people

From the teaching experience of computer music to hearing-impaired college students (HI), we believe that they have interest in music. Thus we set our purpose to build a system that supports them playing music together by sharing emotions -- joy, sadness, fear, and anger -- with a drum set using visual cues.

We have conducted a series of introductory experiments to find out the similarities and differences between people with and without hearing-disabilities in playing a drum set with an emotion and in recognizing an emotion in a performance.

For our purpose, we need to understand the way how HI can recognize the other's performance and what kind of visual cues are useful for the purpose. Thus issues relate to music recognition, music acoustics, disabilities education, and multimedia issues in computer science.

The current issues are to find the possibility of dynamic emotion exchange in music performance and the relevance of the level of disabilities. We conducted an experiment: two HI played their own drum set by turns, starting from the emotion they felt in another player's performance. During a performance, the player was free to change their emotion.

1. The catch and throw of emotions was mostly done well.
2. When a player changed an emotion during his performance, the new emotion tended to mislead the other player into recognizing it as "joy."
3. One of the HI could feel a beat of over 70dB, the other could listen to sound around 30dB with a hearing aid. The hearing level did not affect much on music communication.

Musical communication that represents an emotion through the drum performance seems to be formed. As a next step toward our system, some experiments that quantitatively investigate the relationship between the hearing abilities, HI's recognition of sound, their musical experiences, and the use of visual cues.