

共同研究

電子透かし技術を応用した音響バリアフリー システムの開発と形成的評価

—「パーソナル音響キャプションデコーダ」の実用化に向けて—

(共同研究機関：東北大学電気通信研究所)

(平成 19 年度～ 20 年度)

研究成果報告書

平成 21 年 3 月

独立行政法人
国立特別支援教育総合研究所

まえがき

聴覚障害がある場合には、サイレンなどの避難勧告音や、駅など公共の場において流れる公共空間音放送（PA）など、音による重要な情報を得ることが困難となる。

本研究は、平成16年度から平成18年度まで国立特別支援教育総合研究所と東北大電気通信研究所が共同研究*として実施した「電子透かし技術を応用した障害者のための情報補償システムの開発－音響の情報バリアフリー化に向けて－」の研究成果とともに、“パーソナル音響キャプションデコーダ”の実用化に向けて、新たなデータの埋め込み手法の導入、建屋内での放送設備を用いた空気伝搬耐性の評価を進めたものである。その意味では基礎的研究であるが、これらの成果が、近い将来、支援機器として普及されることを目指しており、その意味で教育現場に直接に貢献する“実際的研究”と考えている。

データハイディング技術を用いて、音情報にデータを埋め込み、それを携帯端末によって検知することで、例えば、当該音信号を聴取可能な場所にいる聴覚障害者が、健聴者と同様の報知内容を視覚的に得ることを可能とする「音のユニバーサルデザイン」の実現を目的としており、障害者基本法に規定された「情報の利用におけるバリアフリー化」に対応する研究課題でもある。

本研究期間内においては、実用化に向けて着実な進展があったものと考えるが、普及までの道のりを考えれば、ここがスタート地点かもしれない。各位より忌憚のないご意見を頂けると幸いである。

独立行政法人国立特別支援教育総合研究所
総括研究員 棟 方 哲 弥

国立大学法人東北大学電気通信研究所
教授 鈴木 陽一

*「共同研究」とは、障害種別等に係る専門的研究に関し、研究所の研究員等又は研究員等のチームと外部機関等とが、共通の課題について共同して行う研究であって、研究所の実際的・総合的研究と外部機関等における基礎的・理論的研究を融合し、障害のある子どもの教育の充実に資するものをいう（独立行政法人国立特別支援教育総合研究所共同研究実施規則より）。

研究体制

本共同研究は、以下の研究体制で行われた（なお、甲と乙の記述は共同研究契約書による）。

（甲）独立行政法人国立特殊教育総合研究所

総括研究員 棟方哲弥

（乙）東北大学電気通信研究所

教 授 鈴木陽一

大学院情報科学研究科 半田浩規

現 情報通信研究機構ユニバーサルメディア研究センター / ATR 認知情報科学研究所

西村竜一（前 東北大学・助教授）

国立特別支援教育総合研究所の研究協力者、研究協力機関は以下のとおりである。

研究協力者

筑波大学附属聴覚特別支援学校・教諭 板橋安人

筑波大学附属聴覚特別支援学校・教諭 武林靖浩

筑波大学附属聴覚特別支援学校・教諭 竹村 茂

筑波大学附属聴覚特別支援学校・教諭 横山知弘

研究協力機関

筑波大学附属聴覚特別支援学校

東京都立葛飾ろう学校

上記の他に、国立特別支援教育総合研究所研究支援員山口俊光氏には、システムの実装ならびに研究資料作成を担当いただいた。ここに記して謝意を表する。

目次

まえがき	i
研究体制	ii
第Ⅰ章 研究の内容と実施経過	
1. 研究の趣旨及び目的	1
2. 研究実施状況と成果	2
3. 今後の課題	3
4. 研究成果の公表	3
第Ⅱ章 パーソナル音響キャプションデコーダの実用化に向けて	
1. エコー法とエコー拡散法によるデータハイディング	5
2. 実験室における空気伝搬耐性評価	6
3. 建屋内の放送設備による実用化のための空気伝搬耐性評価	10
4. 聴覚特別支援学校における校内放送のメッセージ	24
5. エコー拡散法によるプログラムのPDAへの実装	25
6. まとめと今後の課題	26
文 献	28
第Ⅲ章 論文編	31
第Ⅳ章 資料編	49

第Ⅰ章 研究の内容と実施経過

第Ⅰ章 研究の内容と実施経過

1. 研究の趣旨及び目的

聴覚障害がある場合には、サイレンなどの避難勧告音や、駅など公共の場において流れる公共空間音放送(PA)など、音による重要な情報を得ることが困難となる。

本研究は、平成16年度から平成18年度まで東北大学電気通信研究所と共同研究として実施した「電子透かし技術を応用した障害者のための情報補償システムの開発－音響の情報バリアフリー化に向けて－」の研究成果である“パーソナル音響キャプションデコーダ”の実用化に向けて、新たなデータの埋め込み手法の導入などを進めたものである。

東北大学電気通信研究所が持つデータハイディング技術を用いて、音情報にデータを埋め込み、それを携帯端末によって検知することで、当該音信号を聴取可能な場所にいる聴覚障害者が、健聴者と同様の報知内容を視覚的に得ることを可能とするシステムの開発を目的としている。その意味では、障害者基本法に規定された「情報の利用におけるバリアフリー」に対応する研究課題の1つである。

前回の研究で開発した“パーソナル音響キャプションデコーダ”は、PDAタイプの携帯コンピュータ(165mm×95mm, 600g, 重さは補助バッテリー含む。)を本体として動作するもので、スピーカから発せられる音に埋め込まれた電子透かしを解読して、文字メッセージを表示することを可能とした。評価実験では、屋内で最小40dB程度という、図書館内の騒音レベルとされる程度の小さな音圧のデータから電子透かしを認識しており、音源からの距離は77mを記録した。屋外では50-55dB、音源から32mの距離で認識が可能であった。これは、音電子透かしの空気伝搬の耐性に関する当初の予想を大きく上回った成果となつたが、その一方で、伝達する情報量をより多くすること、建屋内の放送設備による実証実験を行うことなど、実用化に向けた研究の推進が必要であった。それらが本研究期間の目的となっている。

一方、このような機器のプロトタイプの試作からβ版(製品の試用版)の開発を行うためには、比較的大きな費用が必要となる。本研究所の1課題としての予算規模での実用化が難しいことから、所内ヒアリングにおいて、実用化に向けた基礎的な研究を進めながら、積極的に外部の競争的資金に申請する方向性が決定した。このため国際学会を含め、研究成果の発表により外部の評価を積極的に受ける方針とした。

将来は、携帯電話のアプリケーション化や、データ埋め込みに関する音響のバリフリーアクション化を進めることで、公共の場でのアナウンス放送などの音のユニバーサルデザイン化につながる研究と考えている。

2. 研究実施状況と成果

“パーソナル音響キャッシュデコーダ”の実用化に向けて伝送可能な情報量をより多くするための基礎研究（東北大学）と実機への実装と評価（国立特別支援教育総合研究所）、建屋内の放送設備による実証実験（国立特別支援教育総合研究所）、搭載するメッセージを決定するための分析（国立特別支援教育総合研究所と筑波大学附属聴覚特別支援学校）を実施した。

まず、東北大学電気通信研究所では、前年度に実装したエコー法を改善したエコー拡散法を用いる実験システムを開発し、その評価実験を実施した。その結果、従来4bps程度であった伝送速度を16bpsに向上することが可能であることを確認した。次いで、このプログラムを用いて、国立特別支援教育総合研究所では、新たにSONYのPDAタイプのコンピュータであるVGN-UX92SをリナックスOSとして実装し動作を確認した。また、エコー法による埋め込み情報の検出率が高いことも今回改めて明らかになった。

一方、国立特別支援教育総合研究所では、建屋内の放送設備を用いた実証実験を行った。ここでは、空气中を伝搬した音を、廊下、及びフロア内の室内に設置した11台のOLYMPUS製LINEAR PCM RECORDER LS-10によって録音し、そのデータをMATLABを用いてケプストラム変換を行って埋め込まれたデータのビット情報を確認した。その結果、複数のスピーカから同時に音が再生される条件（実用的な場面設定）において、電子透かしを附加したデータが読み取り可能である事実を確認した。これと平行して、筑波大学附属聴覚特別支援学校では、校内の文字放送システムの運用を通じて、学校内で必要となるメッセージの種類と分量などを検討した。

本研究期間においては、パーソナル音響キャッシュデコーダの実用化に向けた研究を実施することで、その実用化に向けて着実に成果を上げたものと考える。

外部への発表として電子情報通信学会・情報システムソサイエティ・第1回マルチメディア情報ハイディング研究会（2007年5月）での口頭発表と実機のデモンストレーション発表、電子情報通信学会・教育工学研究会（ET）（2009年1月）での口頭発表、IEEE / Computer Society / International Conference on Intelligent Information Hiding and Multimedia Signal Processing（2007年11月）で口頭発表を行つ

た。さらに、これらの研究成果をまとめ、国際学術雑誌である International Journal of Innovative Computing, Information and Control に投稿し採録が決定した。外部の競争的資金の獲得に向けて必要となる一定の評価が得られたと考えている。また、科学研究費補助金に申請を行った。

3. 今後の課題

本研究の成果は、試作から β 版の開発を経て、世の中に普及し、教育現場へ大きく貢献するものと考えている。その意味で、上述してきた外部資金の獲得とそれによる開発研究は不可欠である。将来的には、携帯電話用のアプリケーション化や、データ埋め込みに関する音響のバリフリーに関する規格化を進めることで、公共の場でのアナウンス放送などの音のユニバーサルデザイン化に発展させることが重要である。

4. 研究成果の公表

(1) 研究成果報告書等の刊行物

・報告書名称：「電子透かし技術を応用した音響バリアフリーシステムの開発と形成的評価－「パーソナル音響キャプションデコーダ」の実用化に向けて－」（国立特別支援教育総合研究所共同研究報告書、平成21年3月）（本報告書）

・Webサイト上の掲載アドレス：www.nise.go.jp

・掲載年月日（予定）：平成21年6月（予定）

(2) 学会発表等

① 国内学会発表

音データハイディングによるユニバーサル情報提示システム実現の試み（電子情報通信学会・情報システムソサイエティ・第1回マルチメディア情報ハイディング研究会、2007年5月），報告者：半田浩規（東北大学）・西村竜一（情報通信研究機構）・棟方哲弥（特総研）・鈴木陽一（東北大学）

音電子透かしを用いた聴覚障害者のための情報提示（実物デモンストレーション）（電子情報通信学会・情報システムソサイエティ・第1回マルチメディア情報ハイディング研究会、2007年5月），報告者：国立特別支援教育総合研究所・東北大学電気通信研究

所

聴覚特別支援学校における文字校内放送システムの運用と展望（電子情報通信学会・教育工学研究会（ET）, 信学技報 ET2008-84 (2009-1), 2009年1月），報告者：横山知弘（筑波大附属聴覚特別支援学校）・棟方哲弥（特総研）・板橋安人・竹村 茂・武林靖浩（筑波大附属聴覚特別支援学校）

②国際学会発表（査読付き）

（パーソナル音響キャプションデコーダの開発に関する論文）

A Portable Acoustic Caption Decoder Using IH Techniques for Enhancing Lives of the People Who Are Deaf or Hard-of-Hearing – System Configuration and Robustness for Airborne Sound –, Proceedings of Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Vol. 2, 271–274, 2007. (Tetsuya Munekata, Toshimitsu Yamaguchi, Hiroki Handa, Ryouichi Nishimura, and Yo-iti Suzuki)

③学会誌掲載（査読付き）

（パーソナル音響キャプションデコーダの開発に関する論文）

Tetsuya Munekata, Toshimitsu Yamaguchi, Hiroki Handa, Ryouichi Nishimura, and Yo-iti Suzuki, A Portable Acoustic Caption Decoder Using IH Techniques for Enhancing Lives of the People Who Are Deaf or Hard-of-Hearing – System Configuration and Robustness for Airborne Sound –, International Journal of Innovative Computing, Information and Control, ISSN 1349-4198 (採録決定) , 2009.

④学術図書掲載（東北大大学における“エコー拡散法”に関する論文）

Ryouichi Nishimura, Yo-iti Suzuki and Byeong-Seob Ko: Advanced Audio Watermarking based on Echo Hiding: Time-Spread Echo Hiding, Digital Audio Watermarking Techniques and Technologies: Applications and Benchmarks, Idea Group Reference, pp. 123–151, [ISBN-13] 978-1599045139, 2008.

第Ⅱ章

パーソナル音響キャプションデコーダの実用化 に向けて

1. エコー法とエコー拡散法によるデータハイディング

—人間に知覚できないほど反射音を人工的に付加して情報を伝達する手法—

電子透かし技術とは、人間の知覚（視覚、聴覚）特性を利用し、静止画、動画、オーディオ等のデジタルコンテンツに対して、コンテンツとは別の情報を、人間に知覚できないように埋め込む技術をいう。「密かに情報を埋め込む技術」ということでは、以前より、「紙幣の透かし」やステガノグラフィ（深層暗号ともいう）という暗号技術が使われていたが、この「情報を埋め込む技術」自体が注目を浴びるようになってきたのは、インターネットが広く普及するようになった極く最近のことである。（「電子透かし」技術の動向と調査対象の概観 標準技術集 特許庁総務部企画調査課技術動向班 http://www.jpo.go.jp/shiryou/shiryou_list.htm より）

例えば、駅のホームでのアナウンスが行われる際に、聴覚障害者への配慮が十分でない場合には、伝達すべき情報を伝えることができない。この場合に、このアナウンス信号に電子透かしを埋め込んでおき、聴覚障害者が電子透かしを検出できる端末を用いるとすれば、その端末のディスプレイ上で、視覚的にアナウンス情報を伝えることが可能になる。

このために解決すべき課題は、アナウンスの音声が空气中を伝搬して、端末側で収音する際の雑音の影響、つまり空気伝搬耐性（アナログ耐性）の特性を明らかにする必要のあること、また、その処理をリアルタイムで行わなければならないことである。

さて、音響に関する電子透かし技術の1つに、エコー法がある。エコー法とは、人間の聴覚特性における継時マスキング*を利用して、その知覚されない時間内に人工的にエコーを付加し、情報を埋め込む手法であり、原音に対して、ビット情報を反射音とし

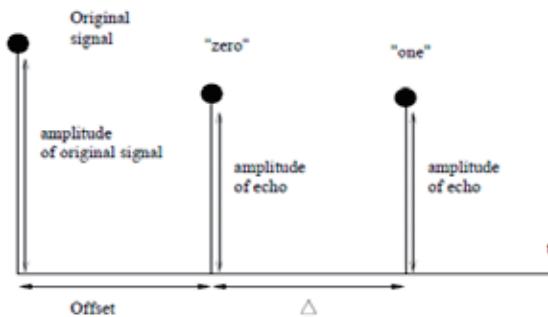


図 2-1 エコー法のエコーカーネル（原音のエコーを特定のオフセットで畳み込む。）

*マスキング現象は、JISに「ある音の最小可聴値が他の音の存在によって上昇すること」と定義される。すなわち、ある音に別の音が加わることによって、人が知覚しづくなる、あるいは知覚できなくなる現象であり、継時マスキングとは、ある音の直前・直後に別の音を発生させた時に起こるマスキング現象である。

て、埋め込めこむ。これが埋め込まれた音データは、これに対して演算（ケプストラム変換）を行うことで、音データから埋め込んだ情報を取り出すことができるという手法である。

さらに、エコー拡散法とは、高ら(2002)によって提案された手法であり、データハイディングを行った音信号に対して、無歪のまま検出処理を行うと、エコー法よりもよい検出性能を有するとされる。この手法は、エコー法のエコー信号となるインパルスをPN系列を用いて図2-2(右)のように時間軸上に拡散させる、エコー法を発展させた手法である。また、また、検出は、ケプストラム変換および、PN系列との相互関係により行う。

図2-2のように、エコー拡散信号の振幅は、エコー法における振幅の $1/L$ の大きさでエコー法と同等の検出性能を見込むことができ、同等の検出性能を有するように振幅を決定した場合、拡散長が長いほうが埋め込む総エネルギーを小さくでき、聴感上の音色に対する影響を少なくすることができる利点がある。

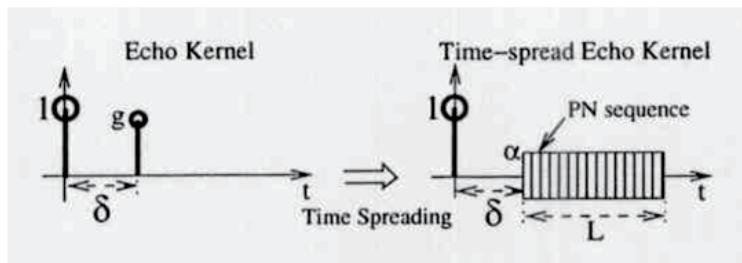


図2-2 エコー法のエコーカーネルとエコー拡散法のエコーカーネル

出典：「PN系列を用いたエコー拡散透かし手法の提案」、「日本音響学会2002年春期研究発表会講演論文集-I-」(2002年3月18日)，高秉燮，西村竜一，鈴木陽一著，(社)日本音響学会。536頁 図1 PN系列を用いたエコー拡散カーネルより。

2. 実験室における空気伝搬耐性評価

—エコー法とエコー拡散法の比較実験—

ところで、エコー法やエコー拡散法によって埋め込まれたデータは、他のデータハイディング手法と同様に、空気を介することによって検出精度に影響を与えると考えられる。そこで、システムの性能向上のため、エコー拡散手法がエコー法よりも空気を介し

た場合、検出精度のよい手法であるのかを、エコー法を拡散長が1のエコー拡散法とみなし、最適な拡散長について検討することによって行った。

原信号に対して、エコー法、および拡散長を変化させたエコー拡散法を用いてデータハイディングを行い、作成したユニバーサルな音信号を音源として、ラウドスピーカ及びマイクロホンを用いて空気伝搬させ、検出性能の評価を行った。

検出性能を決定する評価指標 d' は、大きいほど検出性能が高いことを示している。

雑音を付加しない場合、および雑音を付加した場合の実験結果を図2-3に示す。

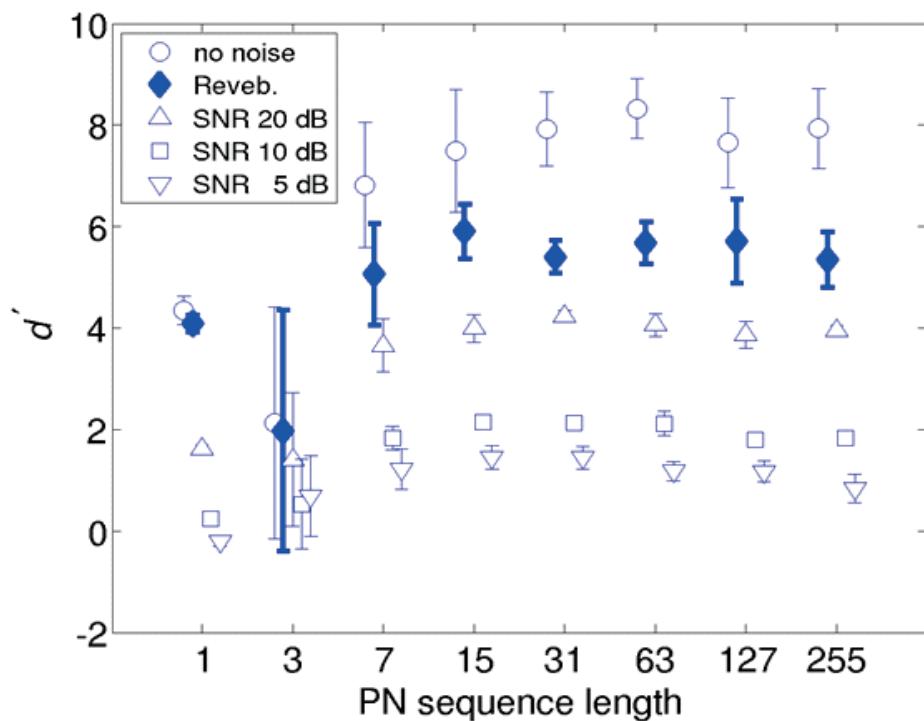


図2-3 エコー拡散法の拡散長と検出性能

図2-3からは、PN sequence lengthと書かれた拡散長でいえば、63という値が検出性能が高いことが理解される。Reverbと書いたデータは、反響音の起こる通常の部屋での条件であり、その他は無響室のデータである。

図2-3によればエコー法（拡散長を1のエコー拡散と見なす。）は、エコー拡散法に劣る手法のように見える。すなわち、よりこれにより、エコーを埋め込む際に要する総エネルギーを等しくした場合、エコーカーネルが複数のエコー拡散法は、エコーカーネルが1本のエコー法より空気を介しても検出精度のよい手法であることが示唆されたことになる。

しかしながら、上記の結果は、PN系列との相互相関をとった場合であった。実際にエコー法は、埋め込み情報を検出する際に、PN系列との相関をとる必要はなく、ケプストラム変換のみにより埋め込んだ情報を検出することが可能である。

上記とは別に、その条件で検証したデータを表2に示す。

表2 PN系列との相関をとらない場合のエコー法とエコー拡散法(PN=63)の比較

	無響室				反射音有り
	no_noise	SNR20	SNR10	SNR5	no_noise
エコー法	8.13	4.16	2.46	1.16	4.55
(標準偏差)	0.14	0.06	0.07	0.03	0.19
エコー拡散(PN=63)	6.52	3.31	1.96	0.93	5.68
(標準偏差)	3.19	1.63	0.94	0.45	0.37

その結果は、以下のようなことであった。すなわち、エコー法は、エコー拡散法で最も検出率の良いPN(拡散長)が63と同等、あるいは、それ以上の検出率があり、加えて、検出のばらつきが少ない手法であるという結果であった。(これは後の屋内における検出実験において、実感と異なる実験結果であったことから、再確認したことで明らかになった事実である。)

今後の実用化に向けて、より適切な検出手法を選定する必要があること、さらには、検出手法以前に、空気伝搬した音データが、実際に、どのように変成するのか、といった基礎的なデータの必要性を示しているものとも考えられた。

一方、これとは別に、さらに、原信号とエコー信号の時間間隔を切り替える間隔を短くすることにより、どの程度ビットレートを増やしても正確に伝達が可能なのかの検討を行った。その結果、エコー拡散法を用いて音質に影響を与えない範囲内でデータハイディングを行うと、無響室においては16bpsまでビットレートを増加させても正確に検出された。

これにより、伝送速度において、本研究が目指す音響キャッシュデコーダの性能を

向上させることが可能であることが明らかになった。

引用文献：音データハイディングによるユニバーサル情報提示システム実現の試み（電子情報通信学会・情報システムソサイエティ・第1回マルチメディア情報ハイディング研究会，2007年5月），報告者：半田浩規（東北大学）・西村竜一（情報通信研究機構）・棟方哲弥（特総研）・鈴木陽一（東北大学）

（なお、この節の記述は、主に上記の半田浩規・西村竜一・棟方哲弥・鈴木陽一(2007)によっているが、後半のエコー法の優位性については、その後の実験によって得られた修正データを合わせて記述して考察した。）

3. 建屋内の放送設備による実用化のための空気伝搬耐性評価

目的

前回の共同研究における実験（国立特殊教育総合研究所，2007）では、单一のスピーカにより発せられた音が空気伝搬しながらも、電子透かし情報をキャッシュンデコーダで検出することを報告したが、とりわけ屋内では、建屋内の放送設備を用いた実験が不可欠と考えられた。すなわち、実際の放送設備では、複数の天井埋め込み型等のスピーカから、同時に原音が発声することから、データの変成等、電子透かしの読み取りに影響を与えるのではないかと考えた。そこで、実際の放送設備を使って、空气中を伝搬した音を、廊下、及びフロア内の室内で録音し、そのデータに埋め込まれたデータのビット情報を確認することで実用化の可能性を探ることとした。

方法

評価実験の場所として、国立特別支援教育総合研究所管理棟1Fを選定した。これは直線距離で70mを越える廊下があることに加えて、大きさの異なる大小の部屋があること、数カ所の曲がり角があるなど、実用化に向けて、様々な条件を確認することが出来ると判断したからである。（前回の共同研究と同様である。）

原音(white noise)に電子透かしと埋め込んだ。具体的には、エコー法により、120 bitsのデータを埋め込んだ30秒間の音声データ(サンプリング周波数は44.1kHz、ビットレートは1,411kbpsのwav形式のファイル)を繰り返しモードで再生しデータを採取した。埋め込みと分析には、MATLAB Version 7.1.0.246(R14) Service Pack 3を用いた。

実際の埋め込みについて、説明図を作成した。表2-1は、実際に原音に埋め込んだ120bitsのビット列を示している。はじめは、”0”ビットが6個続いて、次に”1”ビットが6個続き、”0”ビットが3個というように長方形で示した。図2-4は、原音をケプストラム変換したグラフである。左側には、ピークが現れず、平坦であることがわかる。表2-1にしたがって、透かしの埋め込みを上記のとおり行った音（空気伝搬していない状態）をケプストラム変換すると図2-4の右のようにピーク部分が、”0”と”1”に、表2-1のとおりに現れることが理解される。

表2-1 原音(White Noise)30秒間に埋め込まれたデータビット(4bps)

1~15bit	16~30bit	31~45bit	46~60bit	61~75bit	76~90bit	91~105bit	105~120bit
y(1)=0	y(16)=0	y(31)=0	y(46)=0	y(61)=0	y(76)=1	y(91)=1	y(106)=1
y(2)=0	y(17)=0	y(32)=0	y(47)=0	y(62)=0	y(77)=0	y(92)=1	y(107)=1
y(3)=0	y(18)=1	y(33)=0	y(48)=0	y(63)=0	y(78)=0	y(93)=0	y(108)=1
y(4)=0	y(19)=1	y(34)=1	y(49)=0	y(64)=0	y(79)=0	y(94)=0	y(109)=0
y(5)=0	y(20)=1	y(35)=1	y(50)=1	y(65)=0	y(80)=0	y(95)=0	y(110)=0
y(6)=0	y(21)=0	y(36)=1	y(51)=1	y(66)=1	y(81)=0	y(96)=0	y(111)=0
y(7)=1	y(22)=0	y(37)=0	y(52)=1	y(67)=1	y(82)=1	y(97)=0	y(112)=0
y(8)=1	y(23)=0	y(38)=0	y(53)=0	y(68)=1	y(83)=1	y(98)=1	y(113)=0
y(9)=1	y(24)=0	y(39)=0	y(54)=0	y(69)=0	y(84)=1	y(99)=1	y(114)=0
y(10)=1	y(25)=0	y(40)=0	y(55)=0	y(70)=0	y(85)=0	y(100)=1	y(115)=0
y(11)=1	y(26)=1	y(41)=0	y(56)=0	y(71)=0	y(86)=0	y(101)=0	y(116)=0
y(12)=1	y(27)=1	y(42)=1	y(57)=0	y(72)=0	y(87)=0	y(102)=0	y(117)=0
y(13)=0	y(28)=1	y(43)=1	y(58)=1	y(73)=0	y(88)=0	y(103)=0	y(118)=0
y(14)=0	y(29)=0	y(44)=1	y(59)=1	y(74)=1	y(89)=0	y(104)=0	y(119)=0
y(15)=0	y(30)=0	y(45)=0	y(60)=1	y(75)=1	y(90)=1	y(105)=0	y(120)=0

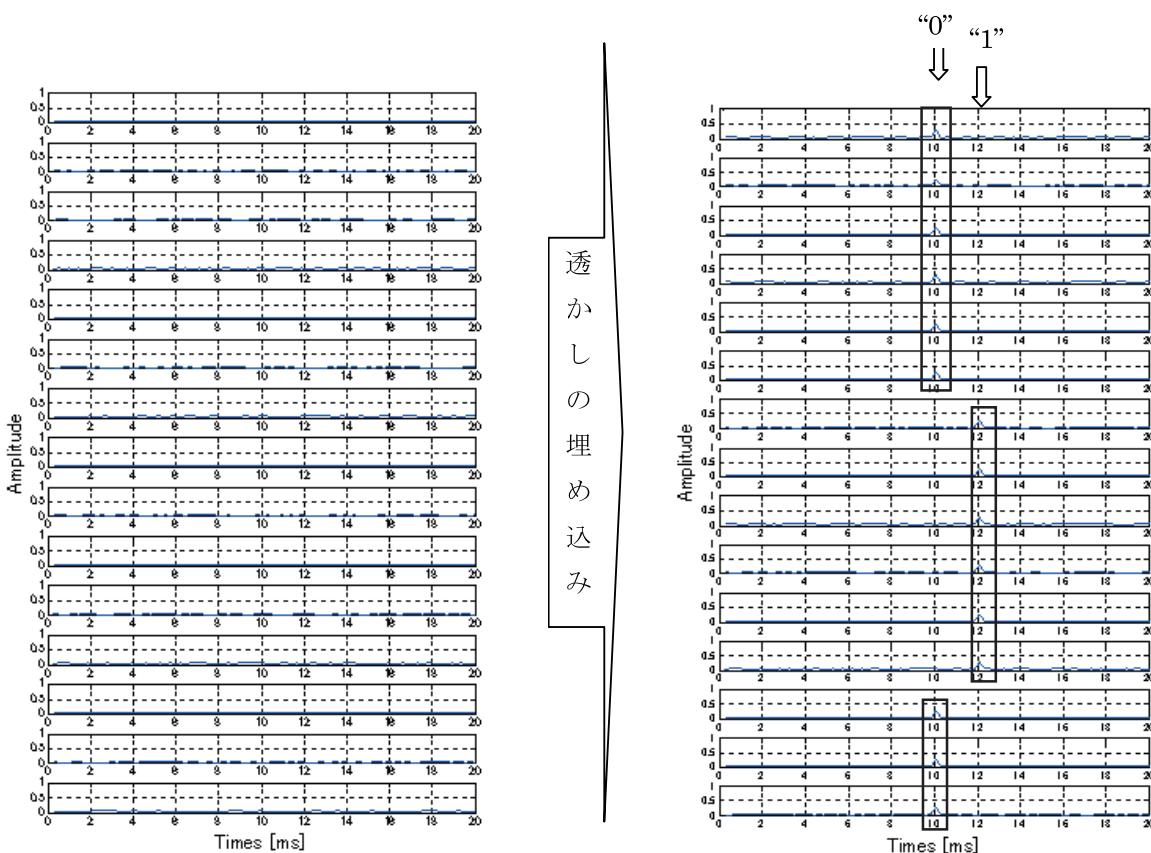


図 2-4 透かし埋め込み前のデータ（左）と表2-1 の bit を埋め込んだデータ（右）

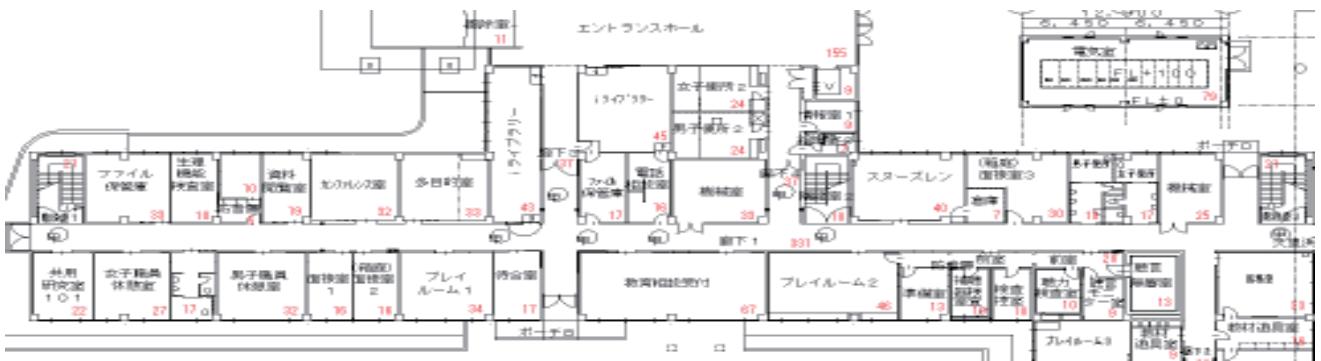


図 2-5 実験を行った当時の建屋内のフロアプラン



図 2-6 実験を行った建屋内の様子
 (図2-5の廊下部分の左端から右を正面に向いた様子)



図 2-7 IC レコーダーとその設置の様子

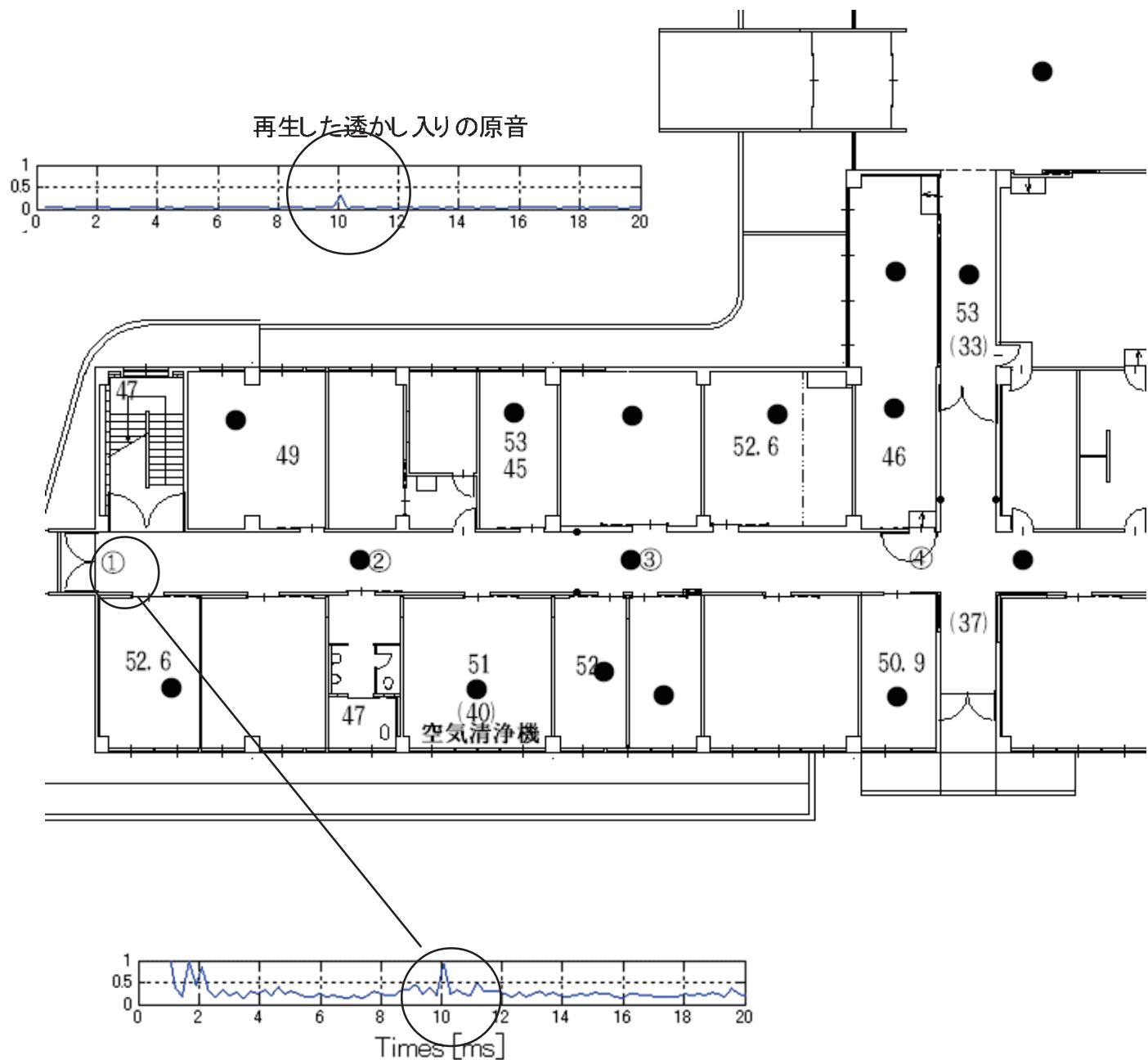
録音は、IC レコーダー OLYMPUS LINEAR PCM RECORDER LS-10 を三脚で 90cm の高さで測定した。台数は、11 台であり、それぞれの条件において、図中に①から⑪までで示される。

また、音圧レベルの測定は、小野測器製精密騒音計 LA-500 を用いた。測定範囲は 27 ~ 130dB であり、周波数範囲は 20 ~ 12500Hz である。

業務用放送設備は PANASONIC WL-7550 シリーズ認定番号 94-018、システムコントローラ WK-850、Mixer WV-M60A であった。天井のスピーカは、Victor 製の天井埋込型スピーカー SB-X166 であり、その仕様は、16cm ダイナミック型、定格入力 6W、出力音圧レベル 92dB/W(1m)、周波数特性 50Hz ~ 14kHz、偏差 20dB であった。図中では、●で示した場所の天井に埋め込まれていた。

実験を行うにあたり、連続して原音を繰り返し再生する必要から、放送設備への入力は、CD プレーヤなどの再生装置の代わりに、EPSON 製 Endeavor Pro 2500、Windows2000 Service pack 4.0、Pentium 4 3.2G M-AUDIO 社の DELTA Audiophile 2496 とした。また、再生ソフトウェアは、Sony Sonic Gate を用いた。

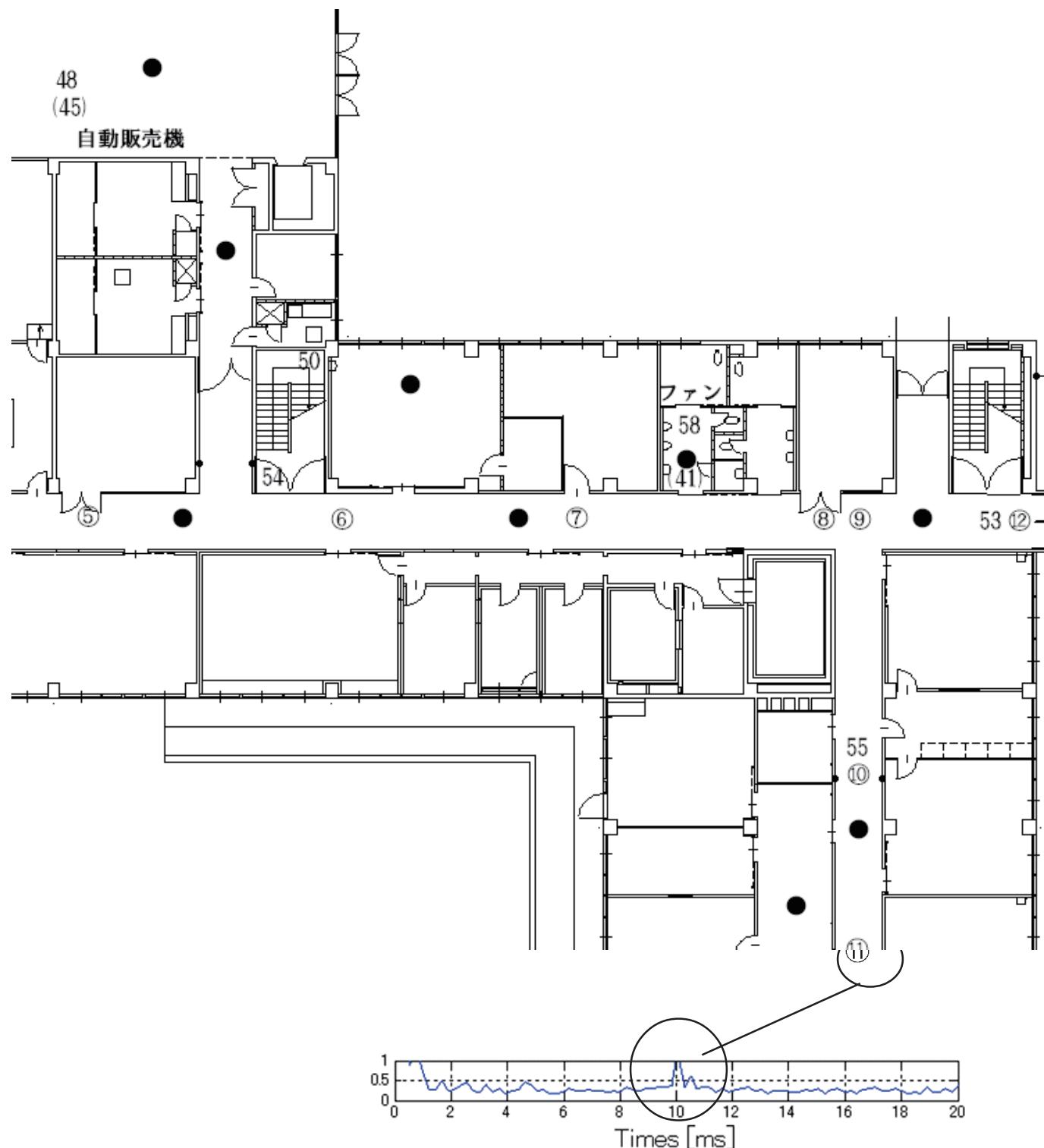
結果を次ページの図 2-8 以降に示す。なお、実際に音響キャプションデコーダーで読み取ることのできた音圧レベルとは別に、①から⑪までの IC レコーダーに記録された、それぞれの地点における録音データの一部を切り取って、それぞれにケプストラム変換を行った。その一部として、図 2-8 に中で、廊下が十字に交差して、複数の部屋に接する④について、図 2-9 以降に、グラフを示す。



①における空気伝搬後の録音データのケプストラム変換データ

図 2-8 放送設備による空気伝搬の測定

●は天井埋め込み型スピーカであり、①～⑪はICレコーダ LS-10をセットした位置を、数値は検出を確認した時の音圧を、()内は放送を入れる前の音圧状態を示す。記述の無い部分は23-30dBの範囲であった。



⑪における空気伝搬後の録音データのケプストラム変換データ

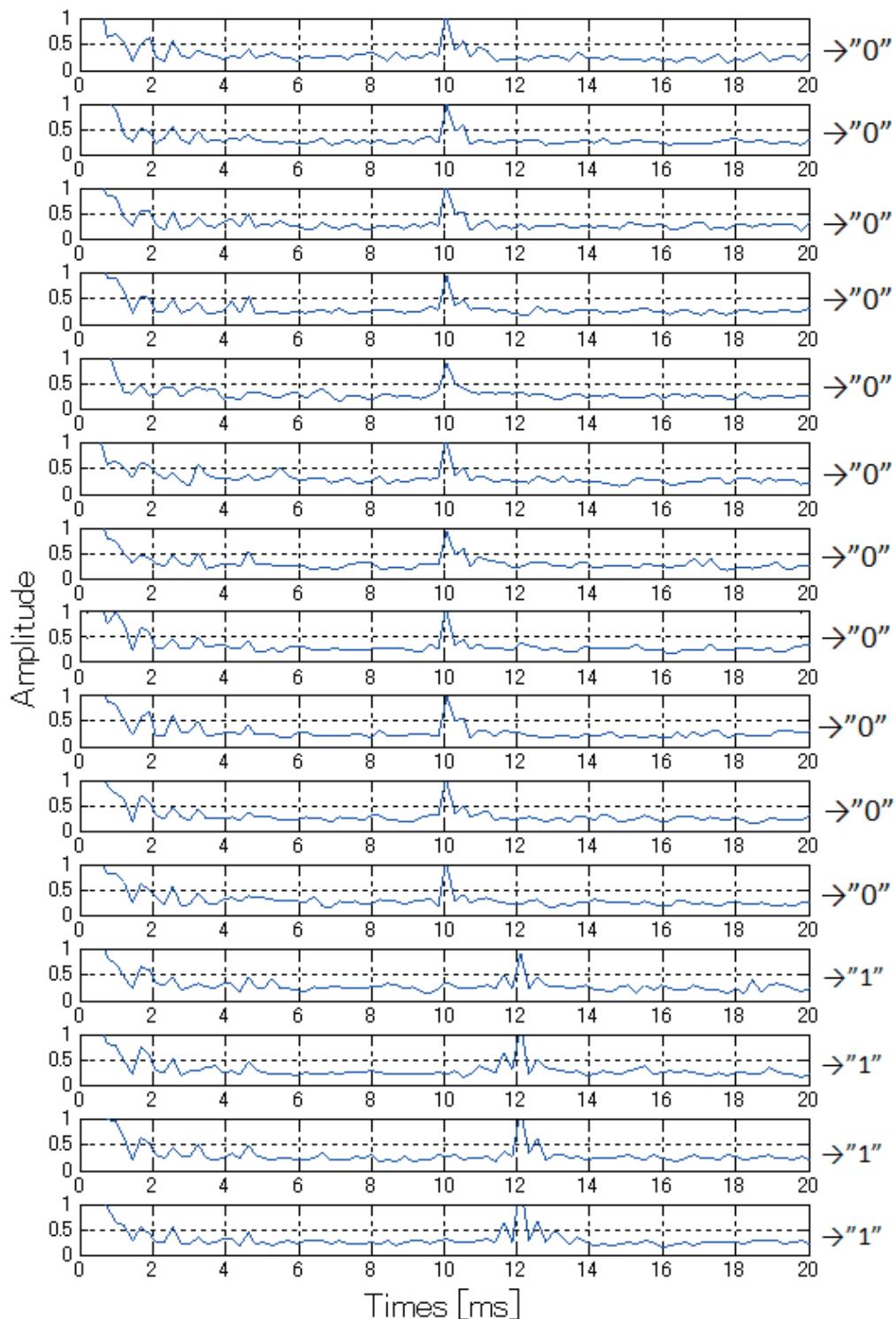


図 2-9 ④の IC レコーダに記録されたデータのケプストラム分析結果（その1）

説明：前のループの”0”と連続して”0”が6個、その後に”1”が6個、続いて”0”が5個、”1”が3個と続く原音(表2-1)が検出されている。

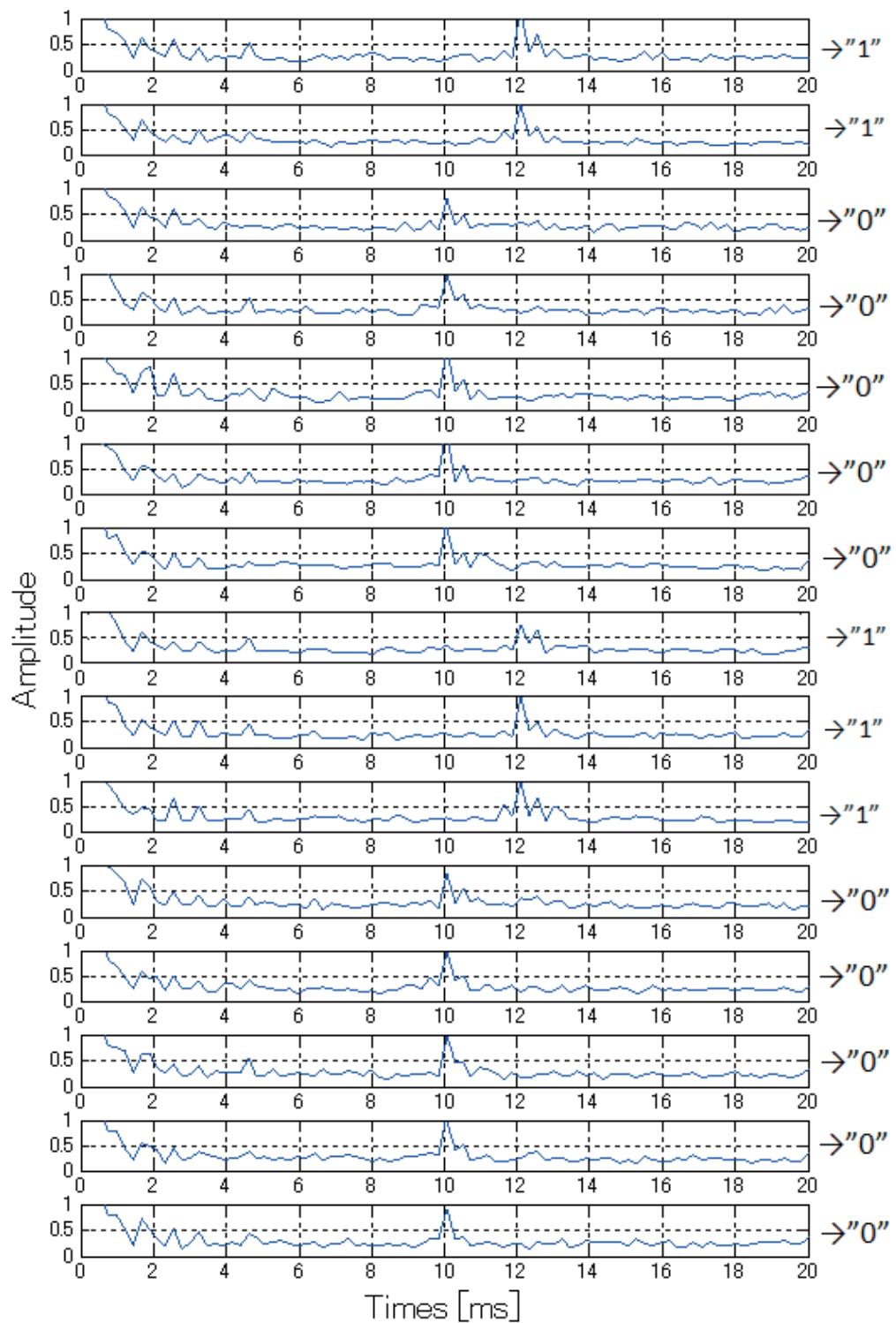


図 2-10 ④の IC レコーダに記録されたデータのケプストラム分析結果（その2）

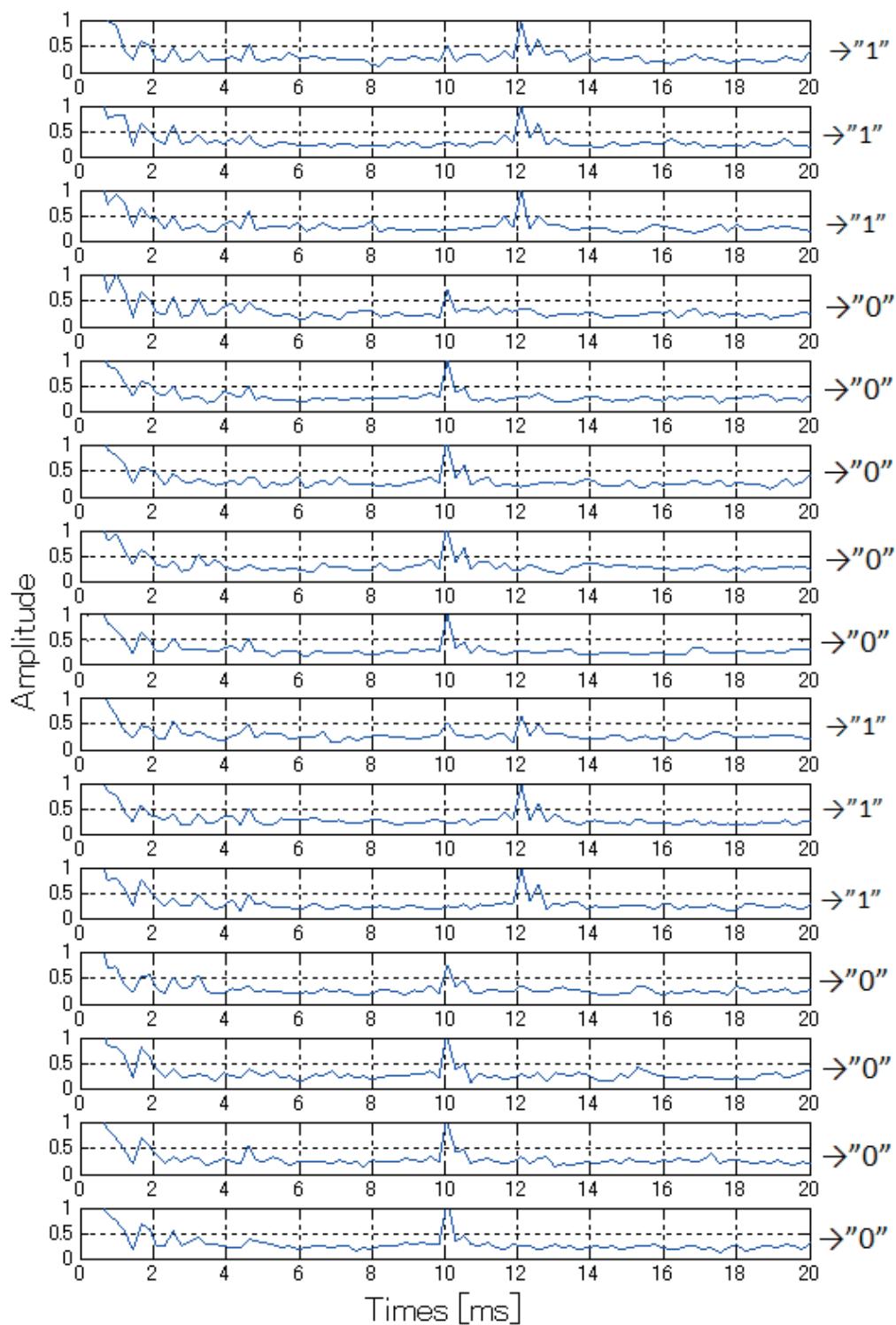


図 2-11 ④のICレコーダに記録されたデータのケプストラム分析結果（その3）

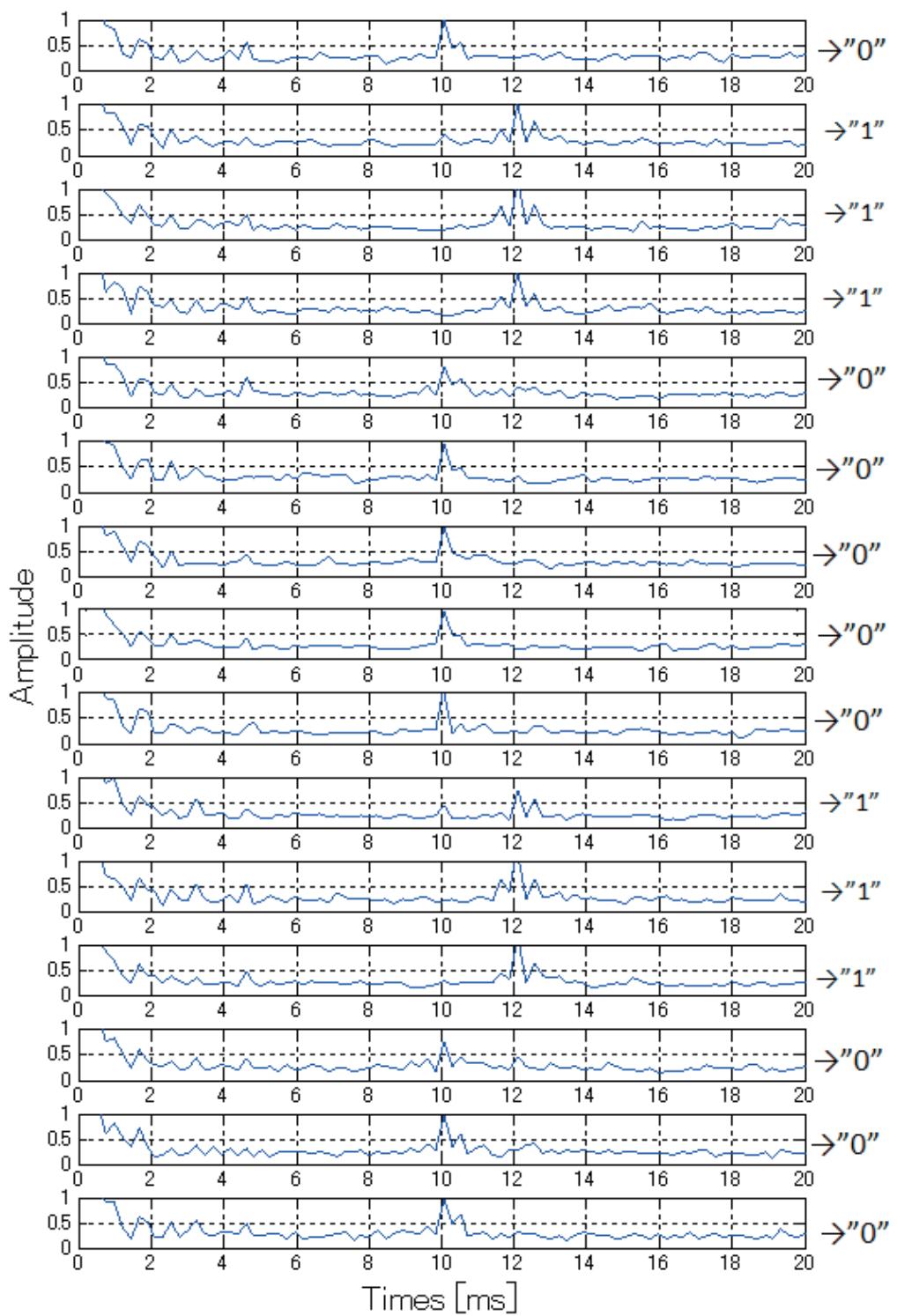


図 2-12 ④のICレコーダに記録されたデータのケプストラム分析結果（その4）

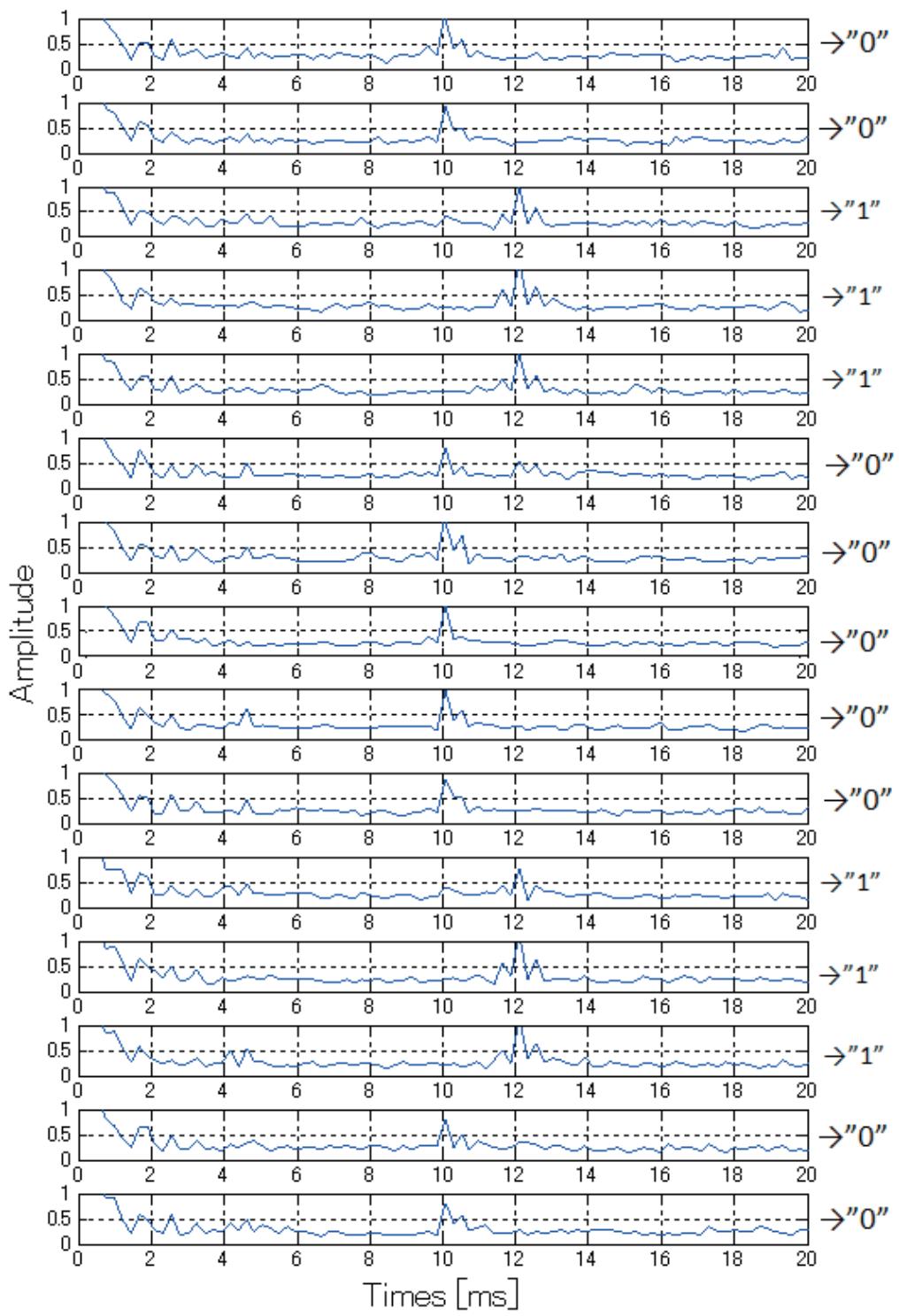


図 2-13 ④のICレコーダに記録されたデータのケプストラム分析結果（その5）

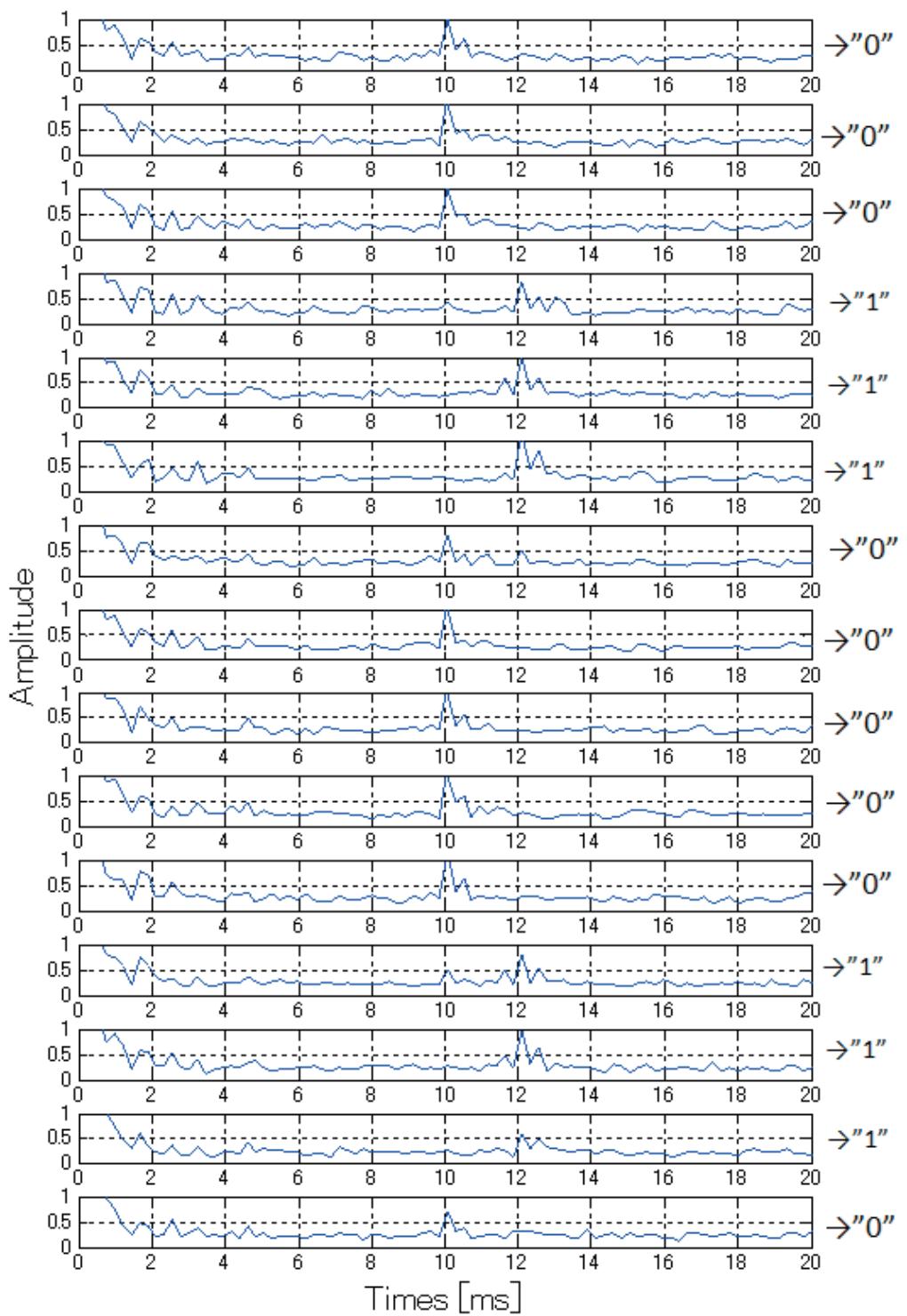


図 2-14 ④のICレコーダに記録されたデータのケプストラム分析結果（その6）

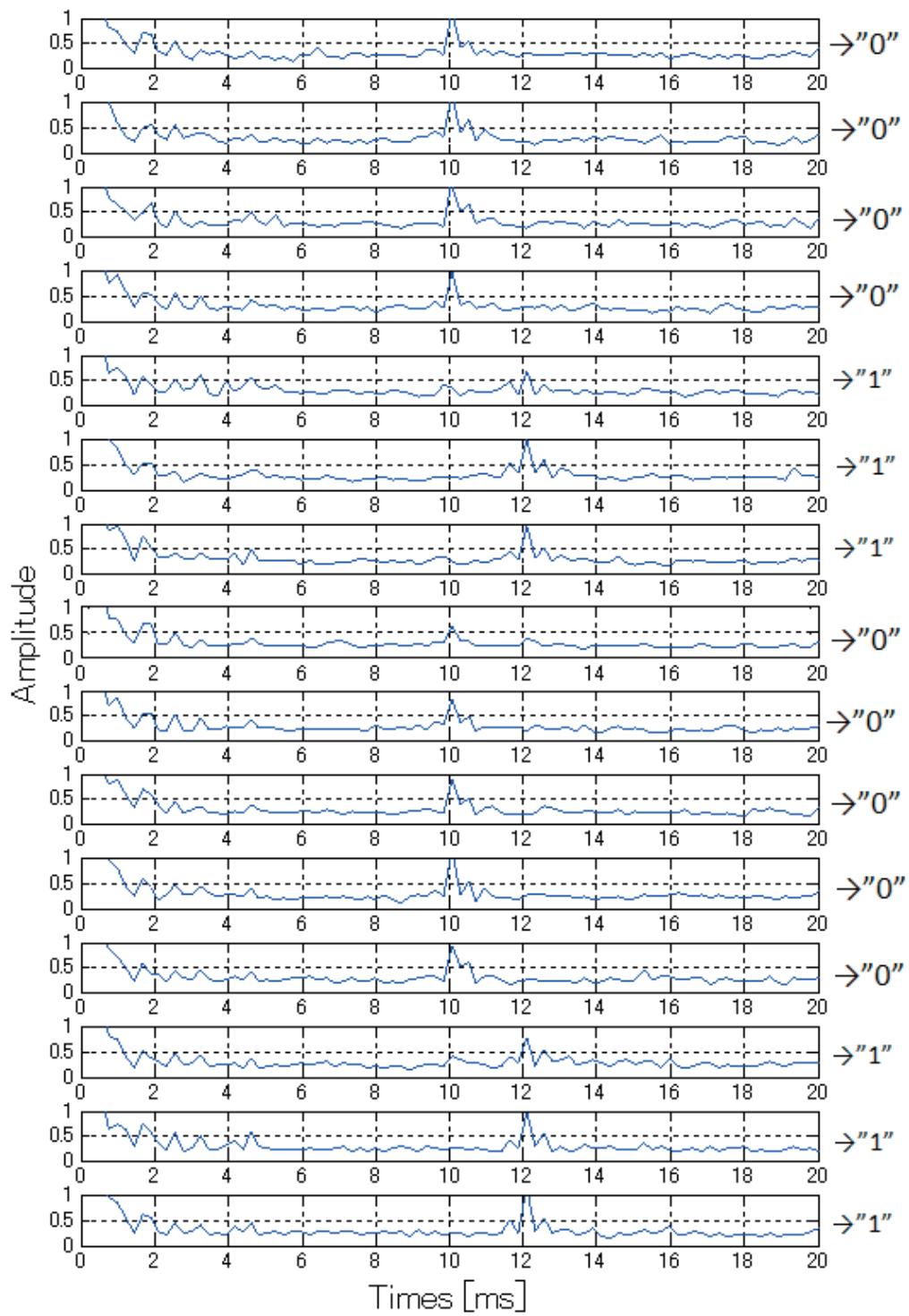


図 2-15 ④の IC レコーダに記録されたデータのケプストラム分析結果（その7）

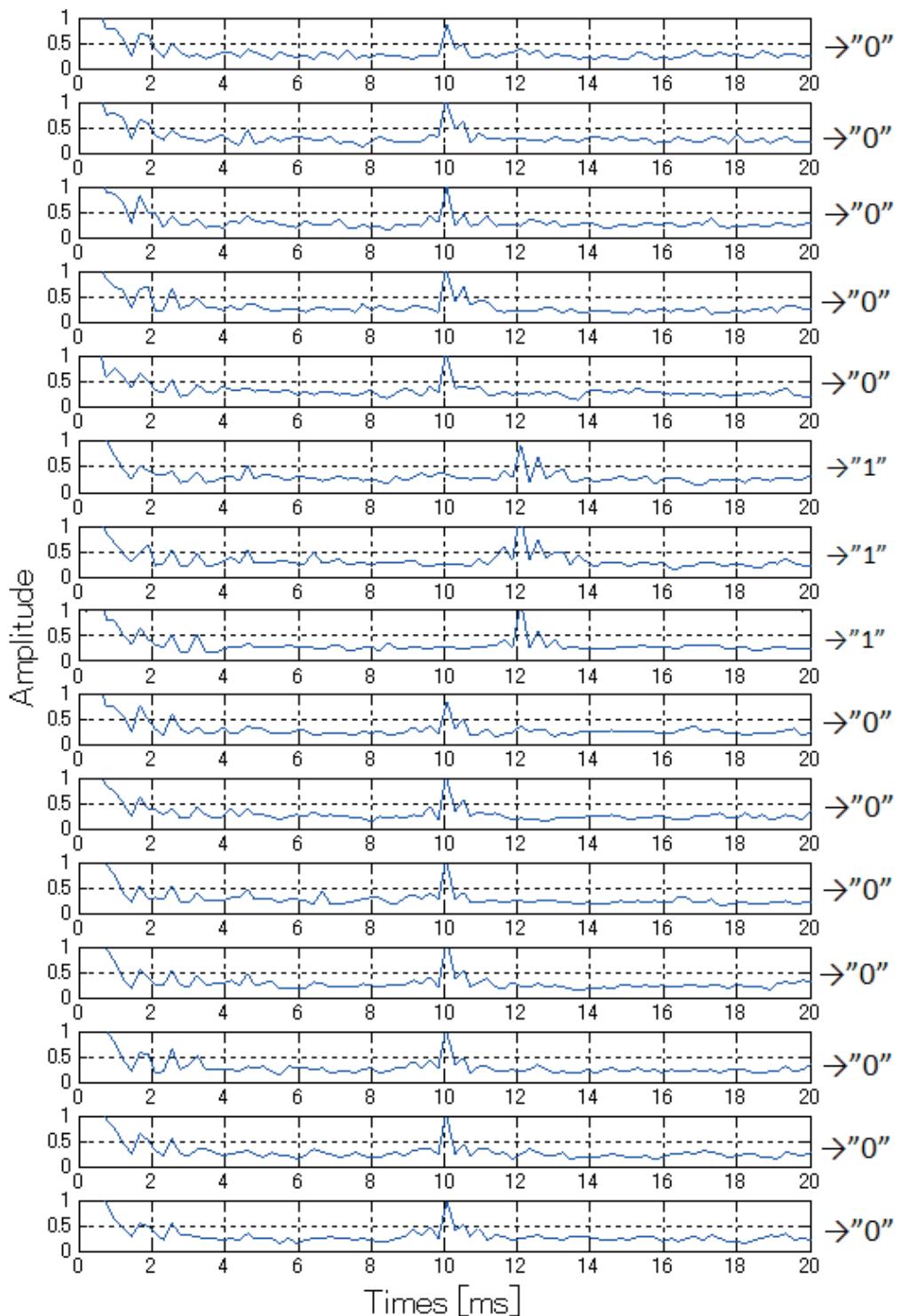


図 2-16 ④のICレコーダに記録されたデータのケプストラム分析結果（その8）

説明：上から6番目から”1”が3個,”0”が5個で1セット終了で、残りは”0”が続いている。

結論

パーソナル音響キャプションデコーダの実用化に向けて、複数の天井埋め込み型等のスピーカから同時に原音が発生する実際の放送設備を用いた実験を行った。

その結果、放送設備による原音が聴取可能な場所においては、音電子透かしのデータを解読することが可能であることを確認した。これは、複数のスピーカからの音が同時に聞こえる場所においても同様であった。

さらに、空气中を伝搬した音を、廊下、及びフロア内の室内で録音し、ケプストラム変換した結果、そのデータに埋め込まれたデータのビット情報を読み取れる状態で、各位置に届いていることを確認した。

これらのことから、建屋内の実際の放送設備による空気伝搬においても、エコー法による電子透かしは有効であることが示唆された。

4. 聴覚特別支援学校における校内放送のメッセージ

本研究の研究協力校である筑波大学附属聴覚特別支援学校の校内文字表示システムにおいて2004年4月から2008年12月まで(このうち2004年4月から2006年12月は、前回の研究で分析を行っている。)の間に実際にメッセージが配信されている。

前回は800件のデータについて①配信データ番号、②配信表タイトル、③配信内容、④配信期間、⑤配信時間帯、⑥配信場所、⑦フラッシュ(視覚的警告)が記録について分析を済ませている(国立特殊教育総合研究所、2007)。

今回のメッセージを加えて、キャプションデコーダに搭載するメッセージについて、検討した。

第1に、メッセージの数についてである。メッセージをそのまま利用する場合であれば、配信されたメッセージ数の合計は953件であり、10bitのデータを正確に送ることができれば、最大で1024種類のメッセージを選択することが可能となる。

第2に、メッセージの構成要素についてである。前回までのメッセージは、最大文字数は67文字であった一方で、今回の153件のメッセージの場合には37文字であった。発話内容として前回は見られていない”！”の使用が5%見られているほかは、前回の枠組みに一致していた。すなわち、メッセージを要素から構成する場合には、発話に関

する分類については 5 種類 (3 bit) で 98% 程度をカバー可能であること。時間帯に関する分類であれば 6 種類 (3 bit) で、全体のカバーが可能となること。これに日数の 30 種類 (5 bit), 曜日の 7 種類 (3 bit), 場所については 7 種類 (3 bit) で 8 割程度をカバーするとなれば、18 bit が必要となること。さらに人の指名についての情報量が加わる。これについては、在籍生徒数を 256 名、8 bit となれば、合計で 26 bit となること。

Web 上などから更新されるデータを読み込むようなシステムが考えられるが、この場合には、より少ない 12 bit の情報があれば、特定の 16 種類の Web サイトのアドレスを記憶し、それぞれの URL 毎に 16 ページの中から、1 ページにある 16 項目のメッセージテーブルの 1 つを利用するシステムが構築可能である。

これらについて、16 bps (1 秒間に 16 bit) の情報を音に載せることが可能であるとすると、実際には、通信のエラー処理を含めて必要な bit 数は増加すると考えられるが、単純計算で、放送時間が 1 秒間以内に上記の情報を送信することが可能となる。

なお、153 件の追加データのうち、全く同じメッセージについては重複を避けて整理したものを資料編に掲載する。

5. エコー拡散法によるプログラムの PDA への実装

東北大学電気通信研究所では、前年度に実装したエコー法を改善した拡散エコー法を用いる実験システムを開発し、その評価実験を起こった。その結果、従来 4 bps 程度であった伝送速度を 16 bps に向上することが可能であることを確認した。これを受けて、国立特別支援教育総合研究所では、新たに SONY の PDA タイプのコンピュータである VGN-UX92S に実装し動作を確認した。

基本的には、エコー法のデコードプログラムを拡散エコー法に置き換える作業であったが、プログラムのモジュールの役割の変更などに対応するため、デコードプログラムの呼び出し方法を変更するなどして対応した。

前回の PDA タイプの携帯コンピュータである VGN-UX90S の後継機種にあたる VGN-UX92S (165mm × 95mm, 600g, 重さは補助バッテリー含む。) を採用した。なお、VGN-UX92S は、前回同様にハードディスクドライブ等の可動部品を持たない PC であり、バッテリによる連続駆動時間が長く、衝撃に対しても比較的ロバストであるという特徴をもっている。

音電子透かし入りのアナウンスを効率よく得るため、PC本体のマイク端子には小型マイクロフォンである ECM-DS30P (SONY 製) を使用した。これには、マイクロフォンとしてバッテリを必要とせずメンテナンス性に優れているという特徴がある。オペレーティングシステムとしては、東北大学とリソースを共有するために、Linux を採用している。

本システムは、電子透かしの入った音声をデコードする部分と、デコード結果を画面上に表示する部分の 2 つからなる。

一方で、今回の実装は、エコー拡散法によるシステムの評価実験用であり、表示部分はシステムによってデコードされたビット数列を表示する機能を有する。今回の実験で明らかになったエコー法の有用性を含めて、さらに検証を進めることで、データ伝送速度が高く検出率の良いシステムを実装させることが必要である。

6. まとめと今後の課題

聴覚障害がある場合には、サイレンなどの避難勧告音や、駅など公共の場において流れる公共空間音放送(PA) など、音による重要な情報を得ることに困難を生ずる。

本研究は、平成 16 年度から平成 18 年度まで東北大学電気通信研究所と共同研究として実施した「電子透かし技術を応用した障害者のための情報補償システムの開発－音響の情報バリアフリー化に向けて－」の研究成果である“パーソナル音響キャプションデコーダ”の実用化に向けてシステムの改善と建屋内の放送設備を用いた実証実験を行った。

エコー法を改善したエコー拡散法によって、従来 4bps 程度であった伝送速度を 16bps に向上することが可能であることを確認した。国立特別支援教育総合研究所では、建屋内の放送設備を用いた実証実験を行った。ここでは、空气中を伝搬した音を、廊下、及びフロア内の室内に設置した 11 台の OLYMPUS 製 LINEAR PCM RECORDER LS-10 によって録音し、そのデータを MATLAB を用いてケプストラム変換を行って埋め込まれたデータのビット情報を確認した。その結果、複数のスピーカから同時に音が再生される条件(実用的な場面設定)において、電子透かしを付加したデータが読み取り可能である事実を確認した。

筑波大学附属聴覚特別支援学校では、校内の文字放送システムの運用を継続し、学校

内で必要となるメッセージの種類と分量などを検討した。

今後は、実際に多数の児童生徒が歩き回る学校生活の中での実証研究が必要であろう。既に、屋外で草むらを歩き回りながらデータを読み取ることができる試用を行っているが、実用化に向けて、さまざまな場面における実験が必要と考えられる。また、声や音楽などのさまざまな音データによる検出力の検討や、事前にデータが埋め込まれたメッセージ以外の利用（この場合には、リアルタイムに電子透かしを埋め込むシステムが必要となる）の実現に向けた研究が必要である。

音環境のユニバーサルデザインを進めるためには、このような電子透かしを埋め込むための標準規格の提案なども必要と考えている。

文 献

- 1) 国立特殊教育総合研究所：電子透かし技術を応用した障害児者のための情報補償システムの開発－音響の情報バリアフリー化に向けて－(平成16年度～平成18年度)研究報告書， 共同研究「電子透かし技術を応用した障害児者のための情報補償システムの開発－音響の情報バリアフリー化に向けて－」(平成16年度～平成18年度)研究報告書， 2007.
- 2) Tetsuya Munekata, Toshimitsu Yamaguchi, Hiroki Handa, Ryouichi Nishimura, and Yo-iti Suzuki, A Portable Acoustic Caption Decoder Using IH Techniques for Enhancing Lives of the People Who Are Deaf or Hard-of-Hearing – System Configuration and Robustness for Airborne Sound –, Proceedings of Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Vol.2, 271-274, 2007.
- 3) Tetsuya Munekata, Toshimitsu Yamaguchi, Hiroki Handa, Ryouichi Nishimura, and Yo-iti Suzuki, A Portable Acoustic Caption Decoder Using IH Techniques for Enhancing Lives of the People Who Are Deaf or Hard-of-Hearing – System Configuration and Robustness for Airborne Sound –, International Journal of Innovative Computing, Information and Control, ISSN 1349-4198 (採録決定)， 2009.
- 4) 電波を利用した聴覚障害者支援システムに関する調査研究会：電波を利用した聴覚障害者支援システムに関する調査研究会報告書， 2006.
- 5) 半田浩規：エコーに基づく音電子透かしの空気伝搬耐性に関する検討， 東北大卒業論文， 東北大学工学部情報工学科， 2007.
- 6) 横山知弘， 武林靖浩：校内 LAN を活用した文字放送システムの導入について， 筑波大学附属聾学校紀要第 27 卷， 121-128, 2005.
- 7) 横山知弘， 武林靖浩， 三好博文， 梶川雄司：校内 L A N を活用した文字放送システム (L-CAST) の運用について， 筑波大学附属聾学校紀要第 28 卷， (web: <http://www.deaf-s.tsukuba.ac.jp/ceremony/kiyo2005.pdf>)， 2006.
- 8) W3C, Core Techniques for Web Content Accessibility Guidelines 1.0. [Http://www.w3.org/TR/WCAG10-CORE-TECHS](http://www.w3.org/TR/WCAG10-CORE-TECHS).

- 9) Federal Communications Commission. Closed Captioning. <http://www.fcc.gov/cgb/consumerfacts/closedcaption.html>.
- 10) Research Committee on Support System for Persons with Hearing Impairment using Radio Wave technology [in Japanese]. The Final Report, 2006.
- 11) Yoiti Suzuki, Ryouichi Nishimura and Hao Tao. Audio watermark enhanced by LDPC coding for air transmission. Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing: pp.23-26, 2006.
- 12) Akira Nishimura. Data hiding in speech sounds using subband amplitude modulation robust against reverberations and background noise. Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing: pp.7-10, 2006.
- 13) W. Bender, D. Gruhl, N. Morimoto, A. Lu. Techniques for data hiding. IBM SYSTEMS JOURNAL, vol. 35, NOS 3&4: 313-336, 1996.
- 14) J. S. Pan, Y. C. Hsin, H. C. Huang, and K. C. Huang, Robust image watermarking based on multiple description vector quantisation, IET Electronics Letters, vol. 40, no. 22, pp. 1409-1410, Oct. 2004.
- 15) J. S. Pan, M. T. Sung, H. C. Huang, and B. Y. Liao, Robust VQ-based digital watermarking for the memoryless binary symmetric channel, IEICE Trans. on Fundamentals of Electronics, Communication and Computer Sciences, vol. E-87A, no. 7, pp. 1839-1841, Jul. 2004.
- 16) J. S. Pan, H. C. Huang, L. C. Jain, and W. C. Fang (editors), Intelligent Multimedia Data Hiding: New Directions, Springer, Germany, Apr. 2007.

第Ⅲ章 論文編

A Portable Acoustic Caption Decoder Using IH Techniques for Enhancing Lives of the People Who Are Deaf or Hard-of-Hearing - System Configuration and Robustness for Airborne Sound -

Tetsuya Munekata*, Toshimitsu Yamaguchi*, Hiroki Handa**, Ryouichi Nishimura***, and Yôiti Suzuki**

*The National Institute of Special Needs Education (NISE), Japan**
*Tohoku University***

*National Institute of Information and Communications Technology (NICT) / Advanced Telecommunications Research Institute International (ATR), Japan****

{munekatt,yamaguchi}@nise.go.jp, {handa, yoh}@ais.riec.tohoku.ac.jp, ryou@nict.go.jp

Abstract

A Portable Acoustic Caption Decoder enables a person who is deaf or hard-of-hearing to have access to audio announcements in public places if the auditory presentation accompanies embedded transcripts. Despite of discouraging results of studies indicating the low robustness of a signal that is transmitted through the air, our decoder successfully detected signals emitted from a loudspeaker located 77 meters away. Embedded bits were conveyed by white-noise in 16-bits digital ($fs=44.1\text{ kHz}$; $1,411\text{ kbps}$) at 4 bps using echo-hiding technique. The sound pressure level of the 77-meter point was as low as 46 dB . Autocepstrum diagrams of the point showed that each of the embedded bits keep good “spikes”. Although, further experiments are needed, the results will encourage researchers to use “airborne stego-sound” to explore more practical applications.

1. Introduction

One of the main characteristics of Information Hiding (IH) technology is, needless to say, the capability to accompany transcript data without perceptual degradation of original sources for convenient levels. This is important when we establish quality universal design (UD) features for multimedia.

Compared to other fields, such as Web content and TV programs, the awareness level of accessibility in the field of audio public announcements still lags considerably behind. For example, W3C insists “auditory presentation must be accompanied by text

transcription” and recommends the SMIL specification [1], while the U.S. Federal Communication Commission mandates Closed Captioning for TV as well as DTV [2]. The specification is also fixed.

It is necessary to establish standards and concrete specifications for implementation in the field of public audio announcement to enhance the lives of people who are deaf or hard-of-hearing. In fact, there are several proposed techniques for “audio for all”. These include radio waves and the Internet as carriers [3]. However, the authors believe that audio information must be transmitted by identical media of the original source, which in this case is sound.

The chief uniqueness of this approach is based on the idea that sound conveys information exclusively to the people who are within the range of that sound. This is different from other approaches such as using radio waves and telephones via the Internet. This allows people in the same place at the same time to appreciate the same information conveyed by the specific information of the sound. Figure 1 illustrates this idea and some trade-off with regard to current technologies.

In the field of information hiding research, a few recent studies have already focused on enhancing the lives of people who are deaf or hard-of-hearing using technology [4][5]. Since these studies concern low robustness, they proposed reasonable strategies to improve the robustness of embedded transcript information.

Along these lines, to promote the practical use of the promised technology, we developed a prototype portable acoustic caption decoder for enhancing the lives of students of schools for the deaf and hard-of-

hearing by ensuring that each of them receives announcements much more easily in the school.

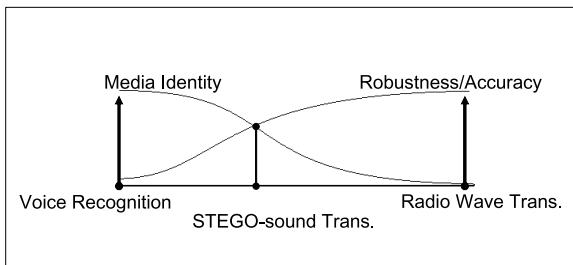


Figure 1. Possible technologies for UD in the field and current trade-off

2. Students' needs and requirements for the new system

Because the primary aim was to develop a device that enables each student to have good access to audio information in schools, we conducted a survey as follows.

The School for the Deaf, University of Tsukuba, one of the special schools for children affiliated to the university, started to use character-based broadcast systems (CBBS) in 2004. The CBBS system displays information through its character display terminal (CDT), which is fixed on the wall. Each CDT displays 11 two-bytes-characters at a time with a character size of 35.7mm in height. A flash light is attached to arouse attention. The system has pros and cons. We therefore conducted (1) a questionnaire survey regarding students' satisfaction and the usability of CBBS to determine students' needs for a new system and (2) analysis of 800 messages that have been delivered by CBBS in the school so far.

2.1 Result of the questionnaire survey

112 students belonging to the school's high school department participated and responded to a questionnaire of 10 items regarding the quality of CBBS and readability of CDTs. The students were divided into two groups. Group 1 consisted of students in a classroom where a CDT is set up. Group 2 consisted of students belonging to a class without any CDTs. (N.B. a CDT is attached to a wall of each floor of the school building.) Although the results were not statistically significant, the main results were as follows:

1. Group 2 seems to be less satisfied with the number of the characters presented by the CDT than Group 1 ($\chi^2(1)=3.41$, $p<.10$).

2. Group 2 seems to be less satisfied with the scrolling speed of the characters presented by the CDT than Group 1 ($\chi^2(1)=3.64$, $p<.10$).

Overall, 71.5% of the students were not fully satisfied with the functions of the CDT. The results also show that they need information everywhere in the school (e.g. they need CDT in the bathroom, gym, and playground).

The following comment was very important. "I am afraid that, in an emergency, we need CBBS terminals wherever students are. Without this system, since we do not have sufficient information, we would hardly know what is happening."

2.2 Results of message analysis

A total of 800 messages delivered by CBBS at the school were examined. We categorized and classified them by dates, duration, and places to which they were delivered, and then determined the function of the contents. As a result, it was concluded that, assuming that we use message ID codes and the number of students is 256, 26 bits of information can represent all the messages broadcasted.

3. System configuration

3.1 Hardware

In this study, our portable acoustic caption decoder consisted of a mobile computer with a 4.5 inch WSVGA display screen and 512MB of RAM with 16GB flash memory (SONY VGN-UX90S) and a small and compact electret condenser microphone (SONY ECM-DS30P) so that its size was as small as 165mm×95mm and its weight was 600g including an auxiliary battery. The microphone's response bandwidth was 100-10000 Hz and its sensitivity 40dB.

3.2 Software configuration

3.2.1 Decoding thread

The decoding thread detected and identified whether the incoming data represent "Zero" or "One" based on echo-hiding technique [6] and then sent them to a queue. It was assumed that the incoming sound was as follows: 44.1kHz; 1,411kbps; containing 120 bits of embedded data in 30 seconds (i.e. 4bps).

3.2.2 Displaying thread

Data from the queue were continuously received and manipulated by the display thread so that messages appeared on the display monitor without any loss of data sequence.

The Display monitor was divided into two domains, one indicating the receiving status and the other indicating messages in floating characters. Colors and sizes can be changed according to users' preferences.

4. Experiments

Two experiments were carried out to evaluate practical performance levels of the prototype decoder. Since our primary goal is to provide a portable caption decoder in a school setting, we chose two venues for the experiments: (1) a floor with a corridor and various rooms; (2) a courtyard similar to a school playground.

In these experiments, we used a 30-second white-noise containing 120 bits of information. Matlab (ver.7.1.0.246) was used to embed a bits sequence based on echo-hiding technique. The sound source was provided by a Sony Walkman NW-E40S connected to a loudspeaker (Sony SRS-T77). Sound pressure levels were measured by a sound level meter (Ono-sokki LA 500).

4.1 Experiment 1

Experiment 1 was carried out in the first floor of an office building. The floor plan is shown in Figure 2 (on the last page of this paper for technical reasons). The corridor was 2.3meters wide and 2.25mmeters high. All the doors were left open. The floor consisted of more than 30 rooms with various doors so that the results of the experiment would be practical.

We checked 23 points including the corridor as well as some of the rooms.

In Figure 2, numbers indicate the sound levels at the points where detection was successful. The sound level of the source was 77dB. Although we never missed any bits at these indicated points, we checked whether we could detect three continuous bytes (24bits) successfully. We confirmed that, at the end of the corridor, located 77 meters from the sound source, the decoder continuously received correct bit trains at 46dB. Also, the decoder detected the bit train correctly in a room with a sound pressure level of less than 40dB.

As shown in Figure 3, autocepstrum diagrams indicate that each of the embedded bits still kept good "spike" even when sound data traveled a long way. For

the spikes of bit "one", data from the 77-meter-point (see Figure 3(b)) were more stable and "mature" than that from the 5-meter-point (see Figure 3(a)).

Even when the decoder detected correct bit stream, we found some low spikes in those diagrams.

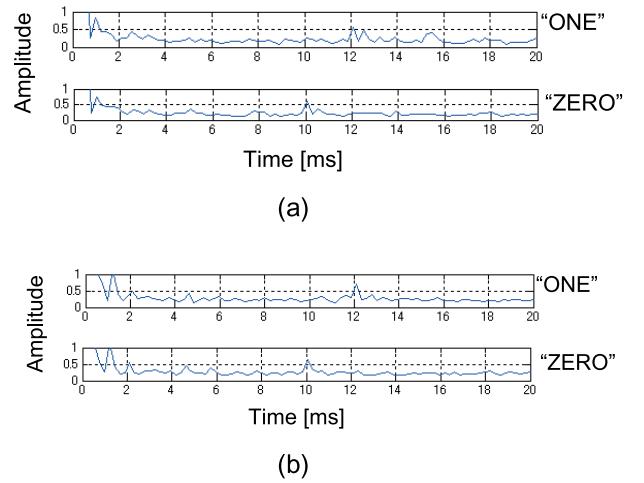


Figure 3. Sample of autocepstrum peaks at (a) 5-meter-point and (b) 77-meter-point

4.2. Experiment 2

Experiment 2 was carried out on a square field of lawn as shown in Figure 4. Experiment conditions were exactly the same as in Experiment 1 except for the sound pressure level of the source.

The decoder continuously detected the bit stream correctly within a 32-meter line from the sound source where the sound pressure level was over 50-55dB. The weather was fine and the wind came from the southwest at 6m/s (from the bottom in the figure).



Figure 4. Experiment venue 2

5. Discussion

Recent studies have reported that the robustness of airborne sound is not sufficient if the technology is based on a simple echo-hiding method [4][5]. Using airborne sound itself, there are also many uncertainties because when a signal is transmitted through the air and re-sampled with a microphone, it is subjected to possibly unknown nonlinear modifications [6] as well as various additive noises.

Although the results obtained here suggest an optimistic prognosis for using this technology for practical applications, further investigations will be needed to clarify what really happened and did not happen during the transmission process.

Furthermore, since the ultimate goal of the authors is to establish universal design (UD) environments for public audio announcements, it is particularly important to establish accessibility standards for airborne audio messages so that the technology can be used for everyday applications.

6. Conclusion

In this study, a portable acoustic caption decoder was developed based on echo-hiding technique and evaluated using airborne sound. The decoder successfully detected embedded data at a 77-meter point from a loudspeaker. The minimum sound pressure level was less than 40 dB. The results demonstrated not only the robustness of embedded data in airborne sounds but also the capability of HI technologies accompanying a mobile computer to play a role of useful assistive technology.

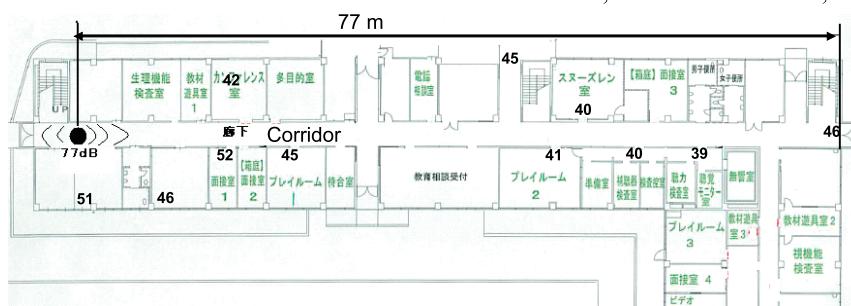
Further investigations are needed to clarify the parameters that affect performance levels as well as establish accessibility standards for airborne audio messages.

Acknowledgments

The authors wish to thank Mr. Shun-ichiro Abe, a doctoral student of Tohoku University, for his help in developing real-time decoding software program, we would not achieve the goal of this report. We also wish to express our gratitude to the many individuals at the affiliated school for the deaf at the University of Tsukuba. This research was partly supported by (1) a grant for collaborative research projects from the National Institute of Special Needs Education along with Res. Inst. of Electrical Communication, Tohoku University and (2) the Strategic Information and Communication R & D Promotion Programme (SCOPE), Ministry of Internal Affairs and Communications.

References

- [1] W3C. Core Techniques for Web Content Accessibility Guidelines 1.0.
<http://www.w3.org/TR/WCAG10-CORE-TECHS>.
- [2] Federal Communications Commission. Closed Captioning.
<http://www.fcc.gov/cgb/consumerfacts/closedcaption.html>.
- [3] Research Committee on Support System for Persons with Hearing Impairment using Radio Wave technology [in Japanese]. The Final Report, 2006.
- [4] Yōiti Suzuki, Ryōichi Nishimura and Hao Tao. Audio watermark enhanced by LDPC coding for air transmission. *Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing*: 23-26, 2006.
- [5] Akira Nishimura. Data hiding in speech sounds using subband amplitude modulation robust against reverberations and background noise. *Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing*: 7-10, 2006.
- [6] W. Bender, D. Gruhl, N. Morimoto, A. Lu. Techniques for data hiding. *IBM SYSTEMS JOURNAL*, VOL 35, NOS 3&4: 313-336, 1996.



聴覚特別支援学校における文字校内放送システムの運用と展望

横山 知弘[†] 棟方 哲弥[‡] 板橋 安人[†] 竹村 茂[†] 武林 靖浩[†]

† 筑波大学附属聴覚特別支援学校（筑波大学附属聾学校）〒272-8560 千葉県市川市国府台 2-1-1

‡ 独立行政法人国立特別支援教育総合研究所 〒239-8585 神奈川県横須賀市野比 5-1-1

E-mail: †{yokoyama, itabashi, takemura, takebayasi}@deaf-s.tsukuba.ac.jp, ‡ munekatt@nise.go.jp

あらまし 聴覚特別支援学校における文字校内放送システムの導入が進み始めた。聴覚障害教育に必要な情報保障には、授業や指導における情報保障にとどまらず、校内放送などの生活全般における支援が重要である。今後の特別支援教育の進展に合わせて、様々な教育の場への普及が期待される。先進校である筑波大学附属聴覚特別支援学校における校内 LAN を用いた文字放送システムは、その導入から 5 年目を迎える。本稿では、その利用者である生徒を対象としたアンケート調査結果から、今後の学校における情報保障について展望する。

キーワード 聴覚障害、校内放送、情報保障（補償）、アンケート調査、文字表示装置

Implementation and Perspectives on the use of character based broadcasting systems at the Schools for the Deaf or Hard of Hearing

Tomohiro YOKOYAMA[†] Tetsuya MUNEKATA[‡] Yasuto ITABASHI[†] Shigeru TAKEMURA[†]
and Yasuhiro TAKEBAYASHI[†]

† School for the Deaf, Tsukuba University 2-2-1 Konodai, Ichikawa City, Chiba, 272-8560 Japan

‡ the National Institute of Special Needs Education 5-1-1 Nobi, Yokosuka Kanagawa, 239-8585 Japan

E-mail: †{yokoyama, itabashi, takemura, takebayashi}@deaf-s.tsukuba.ac.jp, ‡ munekatt@nise.go.jp

Abstract Implementations of character based school broadcasting systems have been emerged. It is important for the school for the Deaf or Hard of Hearing to assure information accessibility not only limited in the teaching activities but also in daily activities such as utilizing announcements in the school buildings. It has been for 5-years since the School for the Deaf, University of Tsukuba, one of the pioneers, first launched its character based school broadcasting system. In this paper we present the results of a research survey of the students and discuss regarding future perspectives of this field.

Keyword Deaf or Hard of Hearing, School broadcasting systems, Accessibility, Survey Study, Character Display Terminals

1. はじめに

情報保障（補償）は、障害者基本法における情報機器のバリアフリーや参議院で付帯決議などを待つまでもなく、障害のある人の生活に欠かせない重要な課題である。ここでは、音響のバリアフリーに焦点をあて、学校内を含めた公共の場などにおける音声アナウンスを文字情報に変換するシステムについて検討する。

電波を利用した聴覚障害者支援システムに関する調査研究会(2006)によれば、共用品推進機構の「聴覚障害者に対するニーズ調査」(2003)は、聴覚障害者のニーズを 11 種類にグルーピングしている。これに基づくと本研究のテーマは「施設内外の放送・案内・呼び出しを理解、対応したい（グルーピング項目の 5）」「（家庭や職場、学校の）機器の報知音、通知音を知る（同 6-1）」「事故災害時の情報がいつでもどこでも欲

しい（同 7）」「ラジオの音声を理解する（同 8-2）」「わかりやすさ、見やすい表示を充実させてほしい（同 9）」に対応する重要な課題と考えられる。

例えば、学校においては、校内放送がこの 1 例であろう。前田（2005）は、当時、筑波大学附属聾学校と東京都立葛飾ろう学校が、既に、文字表示システムを用いて校内文字放送を開始したことを紹介している。また、国立特殊教育総合研究所（当時）(2006)が行った特別支援教育に必要な教育設備整備の在り方に関する調査研究では「フラッシュライト付き電光掲示板」を聾学校の 15.2%が保有し、83.3%が必要な設備に上げていた。さらに、今後必要な設備として「校内放送文字表示システム」を回答した学校のあることが示されていることなどから、さらに普及が求められる分野であり、先の 2 校の実践が先駆的であることが伺われる。

火災発生など、緊急時の警報を第一の目的として整備されたシステムを一般の校内放送の利用へと発展させた事例（横山、武林, 2005），一般的な内容の情報提供の例として校内 LAN に Web の掲示板を設置して情報を校内で共有する試み（加藤, 2005）や、携帯電話で閲覧可能なホームページを利用する試み（同じく、加藤, 2005）が始まっている。

東京都立葛飾ろう学校では、プラズマディスプレイを用いた「見える校内放送」（例えば、伊藤, 2006 など）であり、筑波大学附属聾学校では、表示装置として LED や蛍光表示管（例えば、横山、武林, 2005; 横山、武林, ほか 2006）を用いている。2 校ともに、校内 LAN に接続し、廊下や教室等、校内に配置された表示装置にメッセージが配信されるシステムを有している。

筑波大学附属聾学校は 2004 年 4 月よりシステムを稼働させており、当該システムは固定式の文字表示装置が廊下と教室に設置されており、メッセージの内容によって、その配信場所と配信時間帯を事前に設定可能な機能を有している。また、継続して生徒へのアンケート調査が行われており、複数年にわたる配信メッセージの記録がある。

次節以降では、同校における導入の経緯について記述し、生徒を対象に行ったアンケート調査の結果について報告する。

2. 筑波大学附属聴覚特別支援学校（筑波大学附属聾学校）における文字放送表示システム

2.1. 開発の背景と経緯

筑波大学附属聾学校の正門を入ってすぐの所に中庭がある。高等部棟も中庭に面して建っている。

もし火災が起きた時には、この中庭に消防車が数多くやってくる。したがって、本校では、消防車との接触事故を防ぐために、この中庭を避難場所としない取り決めがなされてきた。

高等部の通常の避難場所グラウンドへは、中学部棟の横を通って歩道橋で避難する。しかし、もしも火災が起きた場所が中学部棟であったら、この避難場所と方法では、生徒を安全に避難させるどころか、かえって危ない目に遭わせることになってしまう。そこで、中学部棟で火災が起きた場合には小学部のグラウンドに避難することになっているが、その時も、中庭を突っ切らないように注意をして避難するようにしている。

つまり、高等部では火災が起きた場所によって避難する場所を使い分けているのである。だから、もしも火災が起きた場合には、火災が起きたことだけではなく、どこが燃えているのかを、正確に、素早く伝えることが、必要不可欠になっている。

どの場所で火災が起きたのか。その情報を伝えるための手段は、防災委員長でもある教頭による「音声」放送を聴こえる教員が聴き取って、生徒に伝える。と言う方法が従来の方法であった。

ところで、高等部には現在、聴こえない教員が 3 人いる。当然だが、聴こえない教員は教頭の「音声」による放送を聴き取ることはできない。生徒を避難誘導しようというその時に、極めて重要な情報である「火災が起きた場所」についての情報を得ることができないのである。そこで、高等部では次のような 2 つの約束をしていた。

約束 1：火災が起きたことを知ったら、教頭の音声放送をしっかりと聴き取って、自分が授業している教室の両隣の教員と放送の内容を確認しあおう。

約束 2：授業中でない場合には、建物の外に出る前に聴こえない教員の部屋に火災が起きた場所を知らせに行こう。

このような約束をして聴こえない教員に情報が伝わるように訓練を積み重ねてきた。

しかし、放送を聴き取ってから別の部屋にいる聴こえない教員に放送の内容を伝えるという方法では聴こえない教員に情報が伝わるのがどうしても遅れがちになる。

また、高等部ということもあって、休み時間は教室内は生徒だけで居ることも多いのである。

そのため、火災が起きた場所を目で見える形で、素早く伝えることができる方法が是非とも必要だと言うことは、高等部の教員共通の思いであった。

一方、学校の防犯対策が見直され始めたが、その中の課題の 1 つに、学校に不審者が入ってきた時に、どのようにして全校に知らせるか、と言うことがある。

公立の小学校など、聴こえる生徒の学校では、不審者が学校に入ってきたときには暗号文を音声で放送したり、その時だけに用いると約束した音楽を放送したりするそうである。どちらの方法も不審者本人に全校への通知を気づかれて、不審者が逆上することを防ぐために「できるだけ不自然にならないように」工夫された方法だと思う。しかし、音は正面からだけではなく、後ろからも横からも全ての方向から入ってくる。このようなものを使って不審者侵入を全校に伝えつつ、不審者本人に気づかれないようにすることが、果たしてできるのであろうか。

そんな時、もし情報を音ではない方法、目で見える形で伝えることができる方法があれば、不審者を刺激せずに情報を伝えることができるのではないだろうか。目で見える情報。具体的に言うと文字による情報である。文字を放送する装置があれば、火災が起きた場所を知らせることも、不審者が学校に入ったことを

知らせることもできるはずである。

そのように考えて、改修工事の際には是非とも文字を LAN 回線を利用して放送する装置を教室に設置したいと考え、情報の収集を始めたところコストがかさみがちな LED 以外の表示方式の技術を持つ企業として、千葉県茂原に研究施設を持つ双葉電子工業という企業が浮かび上がった。双葉電子工業は LED に換わる表示方式として蛍光表示管（VFD）を提唱しており、その蛍光表示管は音響機器、自動車の各種メーターに採用されるなど、耐久性・信頼性の面でも本校の防災を担う装置に相応しいと考え、双葉電子工業の担当者とさっそく連絡を取り、実機をみせてもらうことになった。2002 年度の 3 月のことである。

当時はまだ小型の装置では RS232C による接続のものしかなかったが、LAN 回線で文字メッセージを送受信・表示できる試作機 NMB もみせてもらうことができた。NMB の反応速度はすこぶる速く、防災装置としての潜在能力を十分に持っているように思われた。

つづく 2003 年度の 6 月に双葉電子工業から NMB を借りて、実際の避難訓練で使ってみた。聴こえない教員の協力のもと、文字を放送する装置の効果の大きさを確認した。また、双葉電子工業には、聾学校の放送装置として用いる場合に改良してもらいたいこと伝えた。単なる抽象論ではなく、実際の装置を元に具体的な改良ポイントを校内で検討できた価値は大きく、そのことが導入以後のスムーズな運用を生んでいると言っても過言ではない。

避難訓練での試用後に校内で具体的な改良の要望について意見を交わした後、双葉電子工業にいくつかの要望を伝えたが、それらの要望のうち最も強くお願いしたことは、放送の非常放送との速やかな連動、入力から表示までの時間が速やかであること、そして重要な情報が配信されたことを知らせる手段として、フラッシュランプが点灯するような機能をつけて欲しいと言うことであった。

これらの要望をもとに双葉電子工業では聾学校における放送の未来像を探るという研究目的で本校仕様の試作機を製作してくれることになった。試作機が完成したのは 8 月であった。

完成した本校仕様の試作機を聴こえない教員と共にみて、そこからさらに改良して欲しいポイントを伝えた。その時伝えたことは「情報の重要性に応じて、フラッシュランプの光り方を使い分けることができるようすること」である。非常時なのかそれとも重要ではあるが日常連絡なのかそのあたりの表現をフラッシュランプを点灯させる回数で表現できるようにお願いしたのである。

双葉電子工業はその要望に快く応じてくれて、12 月

には 8 月に求めたとおりの表示端末が試作機として完成されていた。試作機の完成披露と同じ日に併行開発中のソフトウェアのインターフェースや機能についての最終的な打ち合わせを双葉電子工業のプログラム開発者と行った。その時新たな要望として、防災情報表示機としての情報伝達速度を損なうことなく、「月間予定表のファイルから放送内容を入力できるようにしてほしいこと」を伝えた。

このような度重なる改良要求や、完成間近になってからの仕様追加のお願いを行ったにもかかわらず双葉電子工業の技術者たちはその要望に快く、そして着実に応じてくれた。

できあがったシステムの完成度は大変高く、事前に要求した内容全てを満たすものであった。

当初は 1 階から 4 階までの廊下に 2 台ずつと職員室、事務室に計 10 台の装置が設置された。

設置直後の 2004 年度 5 月 24 日に行われた避難訓練において、一切人の手を介さず非常ベルとの連動が確立していることを確認した。その際、全ての装置の非常ベル鳴動から装置作動開始までの時間を計測したところ、平均で 2.49 秒であった。さらにこの年の 12 月に双葉電子工業の協力により装置が増やされ、21 台になった。全ての高等部普通科教室に文字で情報を伝える装置が設置されたのである。

増設後の 2005 年にも避難訓練における測定を行ったが作動開始時間の平均は 2.59 秒と誤差の範囲内であり、増設にともなう遅延は出ておらず、今後全校規模で文字表示装置を増設した場合にも装置数が増えることによる遅延は発生しないものと考えている。

筑波大学附属聾学校の放送、文字放送表示システムは、災害時に備える防災装置として設置された。そして防災装置として災害時に備えつつ、防災装置としての点検も兼ねて毎日の放送を行っている。これまでに放送した内容の多くは、月間予定表にも掲載されている各種日程と、諸連絡さらに各種注意事項の再確認で、その内容は校内開発のインターフェース「配信予定表」を用いることで順調に数を伸ばし 2006 年 6 月 5 日の時点で 854 種類に達している。

また、2006 年 6 月 5 日 16:20 に訓練以外での初の緊急放送をおこなった。学校の近辺で強盗事件が発生、犯人が刃物を持って逃走中という情報を全校に散った生徒に対して配信することを管理職より依頼を受け、直ちに配信。生徒をいち早く掌握することに貢献した。

3. 文字放送表示システムに関するアンケート調査

3.1. システムについて

文字放送全体のシステムについては横山、武林

(2005), 横山, 武林ら (2006) が詳しい。ここでは文字表示装置について以下に述べる。

この分析では2種類の文字表示装置について検討する。1つは、文字の大きさが96mm角×12文字表示の可能なLED(赤, 緑, オレンジ)方式(以下, Aタイプと呼ぶ。)と、もう1つは、文字の高さが35.7mm×11文字表示の可能な「青緑発光」の蛍光表示管(以下, Bタイプと呼ぶ。)方式である。

それぞれが校内LANに接続されて、教員などが入力する配信内容、配信スケジュールにより、指定された場所、時間帯に正確に配信・表示される。

3.2. アンケート調査の項目

分析データの基になったアンケート調査の各項目の概要は以下の通りである。

表1 アンケート調査の各項目の概要

項目1	学年と所属
項目2	あなたのホームルーム教室には校内文字放送端末がありますか
項目3	あなたは校内文字放送端末をどんな時に見ますか
項目4	あなたの校内文字放送端末の見方は、(7. 全ての文を最後まで読む。4. 自分には関係ないとわかった時点で読むのをやめる)
項目5	あなたは校内文字放送端末からどんな情報を入手していますか
項目6	校内文字放送端末が設置されてからどんなことが便利になったと思いますか
項目7	あなたが校内文字放送端末から入手したい情報にはどんなものがありますか
項目8	校内文字放送端末は高等部の建物の中では、十分な数が配置されていると思いますか
項目9	校内文字放送端末を増設するしたらどこが良いと思いますか
項目10	校内文字放送端末の読みやすさについて(この装置の文字の大きさ、この装置に一度に表示される文字数(1行あたりの文字数)、この装置で表示される1つ1つのメッセージの文字数、この装置の文字が流れる速度は)
項目11	校内文字放送端末のスクロールについて(「表示している」ことがわかりやすい、情報量を増やすために適切な方法である、読み取るのに苦労はしない、情報量を減らしても良いので固定表示に換えた方が良い)
項目12	自由記述

3.3. 回答者について

回答者112名のうちわけは、表2の通りであった。

表2 回答者の学年(専攻科は高等部修了者)

	普通科			専攻科	
	1年	2年	3年	1年	2年
回答者数	27	25	26	18	16
割合(%)	24.1	22.3	23.2	16.1	14.3

ここで、普通科と回答している生徒の教室には文

字表示装置が設置されており、専攻科と回答している場合には、それがない場合である。なお、専攻科はさらに、造形学科とビジネス学科の2つの学科から構成される。本来、問題とする属性以外は条件統制が必要であるが、ここでは、回答者の聞こえの状態がグループ間でそれぞれ均等に分散すると仮定して、上記の②にある文字表示装置に読み取りにくさを感じているか否かについて分析を進めた。

なお、造形学科とビジネス学科とした回答数は、それぞれ18人、16人であった。

3.4. 文字表示装置の表示方法の適切さについて

図2は、文字表示装置の文字の大きさやスクロールの速さ、文字数などの問い合わせに対する回答として「適切」以外と回答した項目を図1に示す。縦軸が、「適切」以外と回答のあった項目数であり、横軸は個々の回答者を示す番号で、項目数の少ないものから昇順にソートしたグラフである。

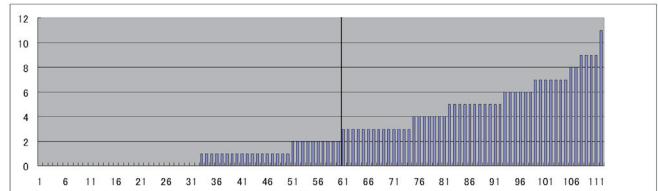


図1 文字表示装置の表示方法の適切さについて
(横軸:回答者連番、縦軸:「適切」以外とした項目の数)

図1に示すように、文字表示装置の属性について全てに「適切」と回答している者は、32名(28.5%)であった。中央値付近は、項目数が3であった。

3.5. 文字表示装置の表示端末数の適切さについて

表3 教室(ホームルーム)における文字表示装置設置の有無
×表示装置(端末)数への満足

	有り	無し
十分な数が配置されていると思う	72	17
十分な数が配置されているとは思わない。増設が必要である	3	15

観測度数が5を下回るため、直接確率計算を行った結果、人数の偏りは有意であった(両側検定: $p < .05$)。表3によれば、教室(ホームルーム)における文字表示装置設置の無い場合には、端末を増設する必要があると回答して、そうでない場合には、十分に配置されていると回答する傾向があるといえる。

3.6. 文字表示装置の表示文字数の適切さについて

表4 教室(ホームルーム)における文字表示装置設置の有無
×Aタイプに一度に表示される文字数(一行あたりの文字数)

	有り	無し
適切	70	26
要改善	8	8

χ^2 検定の結果、人数の偏りは有意傾向であった ($\chi^2(1)=3.41$, $p<.10$)。表 4 によれば、教室(ホームルーム)における文字表示装置設置の無い場合には、A タイプの端末に表示される文字数が適切でないと回答して、そうでない場合には、適切であると回答する傾向があるといえる。

3.7. 校内文字放送端末の文字のスクロールについて

表5 教室(ホームルーム)における文字表示装置設置の有無
×校内文字放送端末の文字のスクロールについて

	有り	無し
適切	62	22
要改善	11	10

χ^2 検定の結果、人数の偏りは有意傾向であった ($\chi^2(1)=3.64$, $p<.10$)。表 5 によれば、教室(ホームルーム)における文字表示装置設置の無い場合には、校内文字放送システムの文字スクロールの方式に改善の余地があると回答し、そうでない者は、適切であると回答する傾向があるといえる。

アンケート調査では「校内文字放送端末を増設するとしたらどこが良いと思いますか」を聞いている。

これに関して、回答は 53 件あった。そのうち教室に設置されていない群(専攻科)の 23 件中 18 件が「専攻科教室」と予想通りであった。他に「トイレ」7 件、「体育館」17 件、「グランド」や「プール」という記述があった。さらに、自由記述部分には「学校は広いので災害が起こった時に、人がトイレや実習室など、表示装置のない場所にいた場合に、耳が聞こえないのでは何があったかわからない」との記述が見られた。

音声による緊急放送が校庭を含めて、校舎の隅々まで伝搬する特質があることを考えると、文字表示器の視覚情報による警報の限界について検討する必要があることを示唆している。

3.8. B タイプに一度に表示される文字の大きさについて

表6 表示方法に「ほぼ満足」/「なんらかの改善を期待」
×B タイプに一度に表示される文字の大きさ

	有り	無し
適切	49	6
要改善	27	18

χ^2 検定の結果、人数の偏りは有意傾向であった ($\chi^2(1)=11.48$, $p<.10$)。表によれば、表示方法に「ほぼ満足している」場合には、文字表示システム(B タイプ)の文字の大きさが適切であると回答し、そうでない場合には、改善の余地があると回答する傾向があるといえる。

なお、上記の他に、“あなたの校内文字放送端末の見方は、(ア. 全ての文を最後まで読む. ブ. 自分には関係ないとわかった時点での読むのをやめる)”において、改善を期待する群において、途中で読むことを止める傾向がやや ($p=.177$) みられた。

自分に関係のない情報を読まないということは、ある意味でリテラシーが育っているとの解釈が可能である一方で、校内放送は、自分のことばかりでなく、その放送を知った人が、対象となる人に伝えるべきものもある。読み取ることが辛いために、読むことへの意欲が弱くなった可能性は否定できない。また、音声のアナウンスが、”自然に耳に入る”ことに比べて、文字表示装置を見続ける必要があることから、不要と理解された時に、意識的に、早めに、読む作業を中断するという理由も考えられよう。

3.9. アンケート結果から示唆されること

総合すると以下のようなことが示唆される。

第一に、個別化された情報提示の必要性についてである。図 1において明らかのように、文字表示方法について全てに満足していた者(「適切」と応えた回答)は全体の 1/3 以下であったこと、さらに、表示方法にほぼ満足していると思われた群における自由記述に「文字の色を変えて表現できるように」、「文字を大きくして欲しい」、「設置位置が高く、わかりづらい」などの改善案が記述されており、全員に 1 つの最適なサイズや表示方法を提供することの難しさが示されている。固定され、”平均化された” 最適なサイズや表示方法が否定されているわけではないが、その一方で、個別化された情報提示方法の実現が望まれていることが示唆される。

なお、 χ^2 検定の結果は、偶然に有意になる項目があることに留意しなければならないが、教室に端末が設置されている場合とそうでない場合に、表示の適切さの評価の傾向に差異が見られた。すなわち、シス

ムへの慣れの程度によって、読みやすさの評価が変化することが予想される。

4. 文字放送表示システムの配信内容について

筑波大学附属聾学校において配信されたデータには、①配信データ番号、②配信表タイトル、③配信内容、④配信期間、⑤配信時間帯、⑥配信場所、⑦フラッシュ（視覚的警告）が記録されている。分析の結果は、以下のような内容であった。

4.1. 配信の時間帯

配信の時間帯は、図2にあるように①始業前、②1時間目休み時間、③2時間目休み時間、④3時間目休み時間、⑤昼休み、⑥5時間目休み時間、⑦清掃時間、⑧15:35-17:45、⑨17:45-18:00、⑩18:00以降となっている。ここでは、始業前ののみ、午前中、昼休み、午後と放課後、全時間帯を指定の5つの時間帯に分けて集計した。複数設定が可能なため、重複がある。

表7 時間帯による配信メッセージ数(2004年4月から2006年12月)

n=800	始業前ののみ	午前中	昼休み	午後と放課後	全時間帯を指定
配信メッセージ数	20	785	5	466	94
割合(%)	2.5	98.1	0.6	58.3	11.8

4.2. 配信継続期間

表8 指定された配信期間

n=799	1日	2日	3日	4日	~1週間	~1ヶ月	~1年	1年以上
配信メッセージ数	408	44	29	26	191	95	4	2
割合(%)	51.1	5.5	3.6	3.3	23.9	11.9	0.5	0.25

表8の“1年程度以上”的内容は、”まもなく下校時間です。18:00までには下校しましょう”（ほぼ1年）と”昼休みの体育館利用は13時15分までです”（536日間）であった。一方、1日のものは、”本日5日お昼休み体育館使用種目はバスケットボールです。”などの情報であった。

4.3. 配信内容

配信内容について分類を検討した。入江（2004）などによる発話に関する分類（発話内行為の分類）に対応させた。結果を表10に示す。ここで最も多い「断定」とは、”です”となるメッセージであり、意味的に言えば、他の分類になり得る余地があると考えられる。

表9 指定された配信期間

n=799	1日	2日	3日	4日	~1週間	~1ヶ月	~1年	1年以上
配信メッセージ数	408	44	29	26	191	95	4	2
割合(%)	51.1	5.5	3.6	3.3	23.9	11.9	0.5	0.25

5. 今後の展望

システムの運用を進める中で、文字によるアナウンスが、生徒の行動の“定着化”に寄与すること、また、そのためにメッセージが持つべき内容があること、手話の得意な子どもと、そうでない子どもでは表示される文字を読む態度に差異があるとの印象を得ている。

本アンケートで明らかになった個別化の必要性や、固定表示装置のある場所以外での情報保障のための個人携帯端末の開発、教育メディアとしての活用の方途を探ることなどと合わせて今後の課題と考える。

6. 問合先

横山知弘

筑波大学附属聴覚特別支援学校

(筑波大学附属聴学校)

〒272-8560 千葉県市川市国府台 2-1-1

TEL : (047) 371-4135 FAX : (047) 372-3672

E-mail : yokoyama@deaf-s.tsukuba.ac.jp

文 献

- [1] 電波を利用した聴覚障害者支援システムに関する調査研究会：電波を利用した聴覚障害者支援システムに関する調査研究会報告書、2006。
- [2] 前田直弘：井戸端会議のメーリングリスト—the Teachers of School for the Deaf 2-mailing list—、聴覚障害、Vol.60, No.650, 28-36, 2005.
- [3] 加藤友仁：生活に生かせる視覚情報の活用能力を育成する、聴覚障害、Vol.60, No.649, 17-27, 2005.
- [4] 横山知弘・武林靖浩・木村和弘・廣瀬由美・平山望武、校内 LAN を活用した文字放送システムの運用について、筑波大学附属聴学校紀要、2006 年
- [5] 横山知弘：校内 LAN を活用した文字放送システムの導入について、聴覚障害、Vol.60, No.649, 28-34, 2005.
- [6] 横山知弘、武林靖浩：校内 LAN を活用した文字放送システムの導入について、筑波大学附属聴学校紀要第 27 卷、121-128, 2005.
- [7] 横山知弘、武林靖浩、三好博文、梶川雄司：校内 LAN を活用した文字放送システム（L-CAST）の運用について、筑波大学附属聴学校紀要第 28 卷
- [8] 入江幸男：2004年度1学期 「問答の意味論的 分析」
http://www.let.osaka-u.ac.jp/~irie/kougi/tokusyu/2004ss/2004ss02introduction.htm, 2004.
- [9] 伊藤 守：ろう学校における IT の活用について、平成 17 年度 E スクエア・エボリューション成果発表会分科会 G 資料、2006.
- [10] 国立特殊教育総合研究所：文部科学省委託研究「特別支援教育に必要な教育設備整備の在り方に関する調査研究」の盲・聾・養護学校における特殊教育設備整備のあり方に関する調査研究報告書、2006.

**A PORTABLE ACOUSTIC CAPTION DECODER USING IH
TECHNIQUES FOR ENHANCING LIVES OF THE PEOPLE WHO
ARE DEAF OR HARD-OF-HEARING - SYSTEM CONFIGURATION
AND ROBUSTNESS FOR AIRBORNE SOUND -**

TETSUYA MUNEKATA¹, TOSHIMITSU YAMAGUCHI¹
HIROKI HANNA², RYOUICHI NISHIMURA³ AND YÔITI SUZUKI²

¹The National Institute of Special Needs Education (NISE)

5-1-1 Nobi, Yokosuka, Japan

{ munekatt; yamaguchi }@nise.go.jp

²Tohoku University

2-1-1 Katahira, Aobaku, Sendai, Japan

{ handa; yoh }@ais.rie.c.tohoku.ac.jp

³National Institute of Information and Communications Technology (NICT)

Advanced Telecommunications Research Institute International (ATR)

2-2-2 Hikari-dai, Seika-cho, Souraku-gun, Kyoto, Japan

ryou@nict.go.jp

Received February 2008; revised July 2008

ABSTRACT. *A Portable Acoustic Caption Decoder enables a person who is deaf or hard-of-hearing to have access to audio announcements in public places if the auditory presentation accompanies embedded transcripts. Despite of discouraging results of studies indicating the low robustness of a signal that is transmitted through the air, our decoder successfully detected signals emitted from a loudspeaker located 77m away. Embedded bits were conveyed by white-noise in 16-bits digital ($fs=44.1\text{kHz}$; $1,411\text{kbps}$) at 4bps using echo-hiding technique. The sound pressure level of the 77m point was as low as 46dB . Autocepstrum diagrams of the point showed that each of the embedded bits keep good “spikes”. We also confirmed that embedded bits sequence can be detected in a whole floor of a building using its real PA system. Although, further experiments are needed, the results will encourage researchers to use “airborne stego-sound” to explore more practical applications.*

Keywords: Intelligent information hiding, Air-borne sound, Universal design, Disability

1. Introduction. One of the main characteristics of information hiding (IH) technology is, needless to say, the capability to accompany transcript data without superfluous perceptual degradation of original sources. This is important when we establish quality universal design (UD) features for multimedia.

Compared to other fields, such as Web content and TV programs, the awareness level of accessibility in the field of audio public announcements still lags considerably behind. For example, W3C insists “auditory presentation must be accompanied by text transcription” and recommends the SMIL specification [1], while the U.S. Federal Communication Commission mandates Closed Captioning for TV as well as DTV [2]. The specification for this closed captioning is also fixed.

It is necessary to establish standards and concrete specifications for implementation in the field of public audio announcement to enhance the lives of people who are deaf or hard-of-hearing. In fact, there are several proposed techniques for “audio for all”. These include radio waves and the Internet as carriers [3]. However, the authors believe that to

realize true universal-design communications, audio information must be transmitted by identical media of the original source, which in this case is sound.

The chief uniqueness of this approach is based on the idea that sound conveys information exclusively to the people who are within the range of that sound. This is different from other approaches such as using radio waves and telephones via the Internet. This allows people in the same place at the same time to appreciate the same information conveyed by the specific information of the sound. Figure 1 illustrates this idea and some trade-off with regard to current technologies. Regarding Public Announcement (PA) system, the voice recognition technology has not yet been good enough for its accuracy in this field while radio wave transmission technology has very high robustness and accuracy but using different medium. We believe that data hiding technology is promising to establish some universal design features in this field.

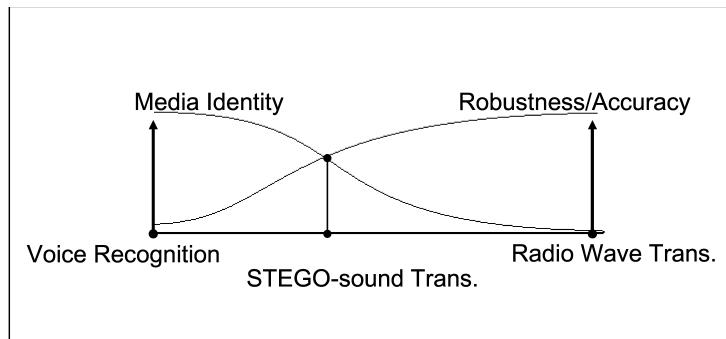


FIGURE 1. Possible technologies for UD in the field and current trade-off

In the field of information hiding research, a few recent studies have already focused on enhancing the lives of people who are deaf or hard-of-hearing using technology [4,5]. Since these studies concern low robustness, they proposed reasonable strategies to improve the robustness of embedded transcript information.

Along these lines, to promote the practical use of the promised technology, we developed a prototype portable acoustic caption decoder for enhancing the lives of students of schools for the deaf and hard-of-hearing by ensuring that each of them receives announcements much more easily than at present in the school.

2. Students' Needs and Requirements for the new System. Because the primary aim was to develop a device that enables each student to have good access to audio information in schools, we conducted a survey as follows.

The School for the Deaf, University of Tsukuba, one of the special schools for children affiliated to the university, started to use character-based broadcast systems (CBBS) in 2004. The CBBS system displays information through its character display terminal (CDT), which is fixed on the wall or ceiling. Each CDT displays 11 two-bytes-characters at a time with a character size of 35.7mm in height. A flash light is attached to arouse attention. The system has pros and cons. We therefore conducted (1) a questionnaire survey regarding students' satisfaction and the usability of CBBS to determine students' needs for a new system and (2) analysis of 800 messages that have been delivered by CBBS in the school so far.

2.1. Result of the questionnaire survey. In this survey, 112 students belonging to the school's high school department participated and responded to a questionnaire of 10 items regarding the quality of CBBS and readability of CDTs. The students were divided

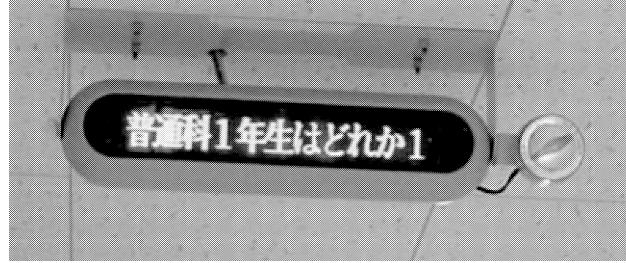


FIGURE 2. CDT is fixed in the ceiling at the school (The sentence on the display reads “The first class pupils in general department should....”)

into two groups. Group 1 consisted of students in a classroom where a CDT is set up. Group 2 consisted of students belonging to a class without any CDTs. (N.B. a CDT is attached to a wall of each floor of the school building.) Although the results were not statistically significant, the main results were as follows.

- 1) Group 2 seems to be less satisfied with the number of the characters presented by the CDT than Group 1 ($\chi^2(1) = 3.41, p < .10$).
- 2) Group 2 seems to be less satisfied with the scrolling speed of the characters presented by the CDT than Group 1 ($\chi^2(1) = 3.64, p < .10$).

Overall, 71.5% of the students were not fully satisfied with the functions of the CDT. The results also show that they need information everywhere in the school (e.g. they need CDT in the bathroom, gym, and playground).

The following comment was very important. “I am afraid that, in an emergency, we need CBBS terminals wherever students are. Without this system, since we do not have sufficient information, we would hardly know what is happening.”

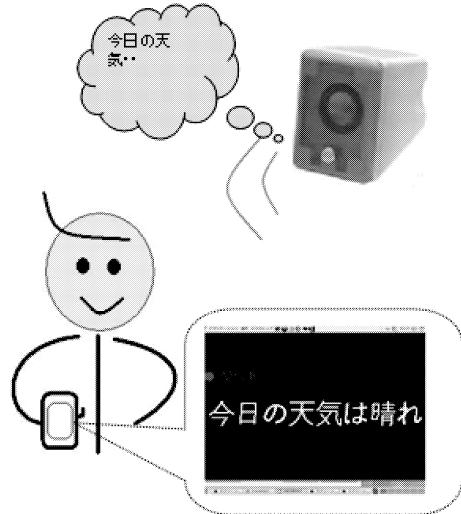


FIGURE 3. Concept design for the system (The sentence reads “It will be fine today...”)

Following keywords were derived: Portable, Personal, Individualized, Anytime, Anywhere, etc. Figure 3 shows derived concept of Acoustic Caption Decoder.

2.2. Results of message analysis. A total of 800 messages delivered by CBBS at the school were examined. We categorized and classified them by dates, duration, and places to which they were delivered, and then determined the function of the contents. As a result, it was concluded that, assuming that we use message ID codes and the number of students is 256, 26bits of information can represent all the messages to be broadcasted.

3. System Configuration.

3.1. Hardware. In this study, our portable acoustic caption decoder consisted of a mobile computer with a 4.5inch WSVGA display screen and 512MB of RAM with 16GB flash memory (SONY VGN-UX90S) and a small and compact electret condenser microphone (SONY ECM-DS30P) so that its size was as small as 165mm × 95mm and its weight was 600g including an auxiliary battery. The microphone's response bandwidth was 100 to 10000Hz and its sensitivity 40dB.

3.2. Software configuration.

3.2.1. Decoding thread. The decoding thread detected and identified whether the incoming data represent “Zero” or “One” based on echo-hiding technique [6] and then sent them to a queue. It was assumed that the incoming sound was as follows: sampling frequency of 44.1kHz, quantization of 16bits; containing 120bits of embedded data in 30 seconds (i.e. 4bps).

3.2.2. Displaying thread. Data from the queue were continuously received and manipulated by the display thread so that messages appeared on the display monitor without any loss of data sequence.

The Display monitor was divided into two domains, one indicating the receiving status and the other indicating messages in floating characters. Colors and sizes can be changed according to users' preferences.



FIGURE 4. Appearance of the portable acoustic caption decoder (The sentence on the displays reads “The deadline of the registration is today...”)

4. Experiments. Three experiments were carried out to evaluate practical performance levels of the prototype decoder. Since our primary goal is to provide a portable caption decoder in a school setting, in the first two experiments the sound was radiated from a portable loudspeaker and two venues were chosen as follows: (1) a floor with a corridor and various rooms; (2) a courtyard similar to a school playground. In the third experiments, to examine the usability of the proposed system in more practical conditions, the real PA system of the school.

In these experiments, we used a 30-second white-noise containing 120bits of information. Matlab (ver.7.1.0.246) was used to embed a bits sequence based on echo-hiding technique. The sound source was provided by a Sony Walkman NW-E40S connected to a loudspeaker (Sony SRS-T77) in Experiment 1 and Experiment 2 while the PA system of the school (Panasonic's WL-7550) was used in Experiment 3. Sound pressure levels were measured by a sound level meter (Ono-sokki LA 500).

4.1. Experiment 1. Experiment 1 was carried out in the first floor of an office building. The floor plan is shown in Figure 2. The corridor was 2.3m wide and 2.25m high. All the doors were left open. The floor consisted of more than 30 rooms with various doors so that the results of the experiment would be practical.

We checked 23 points including the corridor as well as some of the rooms.

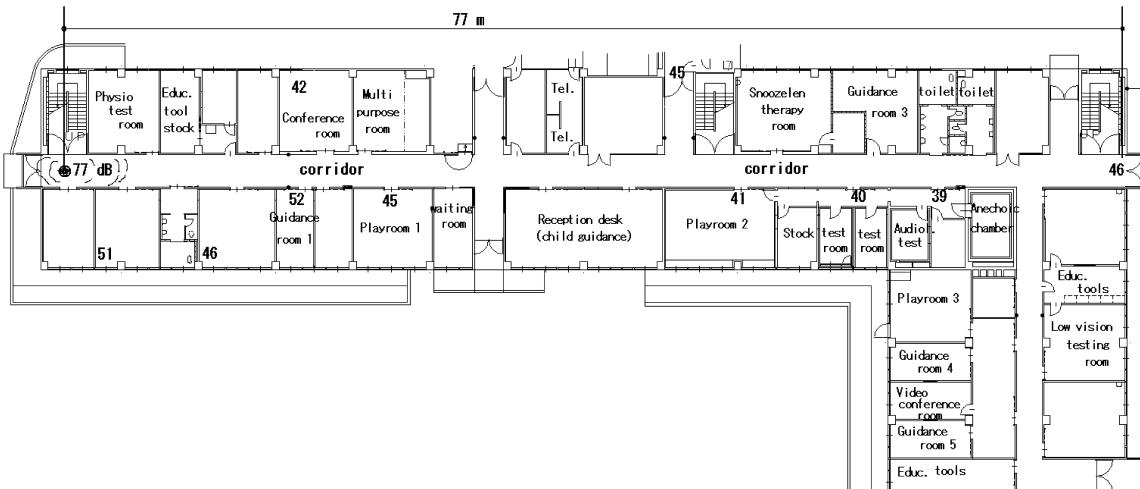


FIGURE 5. Floor plan of experiment venue 1
 (The numbers indicate sound pressure level of the points where detection was successful)

In Figure 5, numbers indicate the sound levels at the points where detection was successful. The sound pressure level at the distance of 30cm from the sound source was 77dB. Although we never missed any bits at these indicated points, we checked whether we could detect three continuous bytes (24bits) successfully. We confirmed that, at the end of the corridor, located 77m from the sound source, the decoder continuously received correct bit trains at a sound pressure level of 46dB. Also, the decoder detected the bit train correctly in a room with a sound pressure level of less than 40dB.

As shown in Figure 6, autocepstrum diagrams indicate that each of the embedded bits still kept good “spike” even when sound data traveled a long way. Peaks at 12ms and 10ms denote “one” and “zero”, respectively in this study. For the spikes of bit “one”, data from the 77-m-point (see Figure 6(b)) were more stable and “mature” than that from the 5-m-point (see Figure 6(a)).

Even when the decoder detected correct bit stream, we found some low spikes in those diagrams.

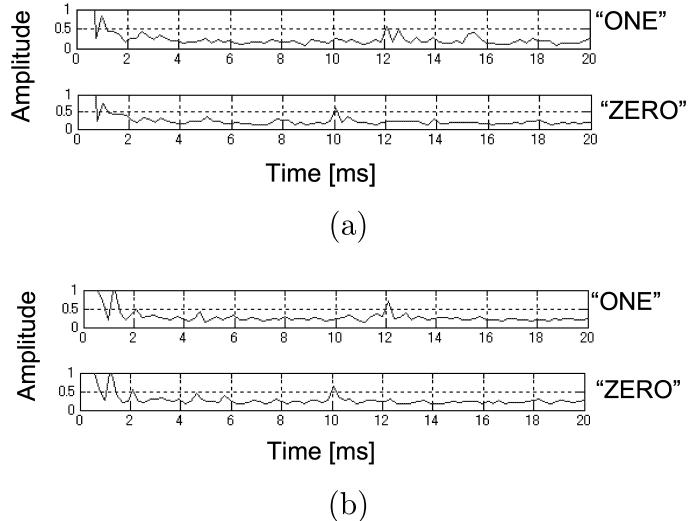


FIGURE 6. Sample of autocepstrum peaks at (a) 5-m-point and (b) 77-m-point

4.2. Experiment 2. Experiment 2 was carried out on a square field of lawn as shown in Figure 4. Experiment conditions were exactly the same as in Experiment 1 except for the sound pressure level of the source.

The decoder continuously detected the bit stream correctly within a 32-m line from the sound source where the sound pressure level was over 50 to 55dB. Outskirts of this region where the sound pressure level was lower than 50dB, errors were detected. The weather was fine and the wind came from the south-west at 6m/s (from the bottom in the figure).

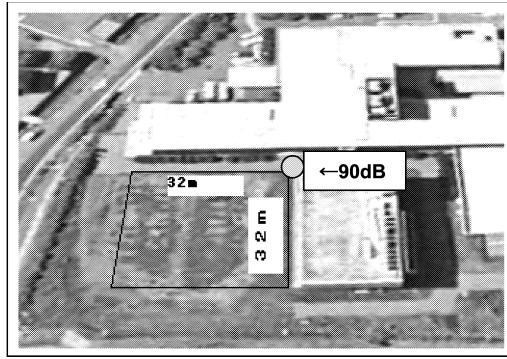


FIGURE 7. Experiment venue 2

4.3. Experiment 3. In this experiment, we used the same sound data as in the previous experiments and use the real PA system, Panasonic's WL-7550 instead of use potable speaker system. The PA system equipped with its mixer WU-M60A. Originally, a 16cm dynamic ceiling speaker, Victor's SB-X166, is built in the ceiling every 10m in the corridor and each of the rooms has same dynamic loudspeaker except in the rest room indicated below. The sound source was provided by an analog audio output of a sound card of

PORTABLE ACOUSTIC CAPTION DECODER

a computer to WL-7550PA system located the central operating room of the building. Sound pressure levels were measured by the sound level meter.

It was confirmed that embedded bits sequence were detected successfully in a whole floor of this building shown below except one room, a toilet without a loudspeaker which has 5cm thick steel door (as indicated asterisk 1) and in which the sound pressure level was as low as 40dB maximum in this experiment. On the other hand, embedded data was successfully detected in a small compartment of a toilet (as indicated asterisk 2 in the figure). The sound pressure level was 54dB when the decoder detected the bits stream. This toilet has a ventilation fan so that its background noise level was 44dB in the compartment.

In Figure 8, two kinds of numbers are seen in a room layout. The number with parenthesis is the background noise level while the other indicates the sound pressure level when the PA system is generating the sound source sequence. Rooms without numbers were closed during this experiment. The criteria of the correctness used in this case were the same as in Experiment 1.

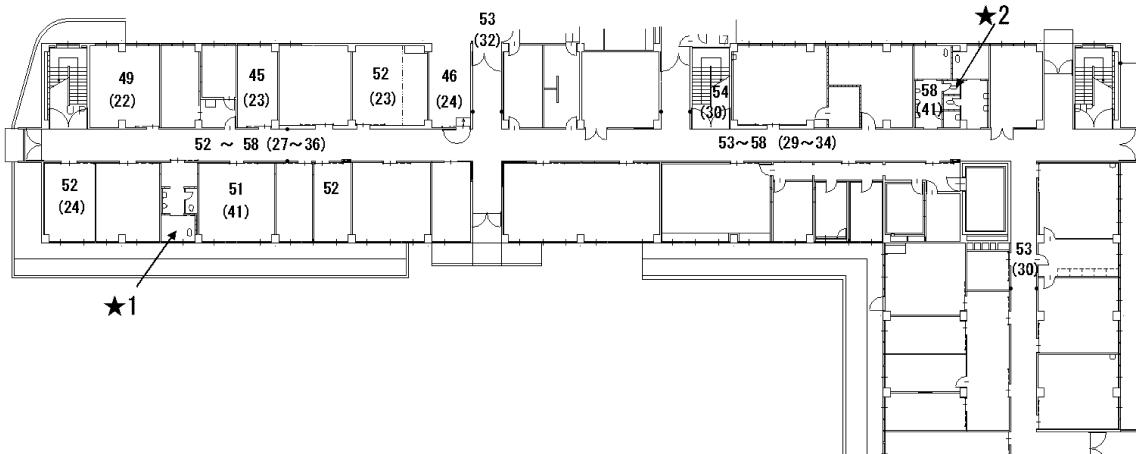


FIGURE 8. Experiment venue 3 (the same floor as Experiment 1 was conducted)

5. Discussion. Recent studies have reported that the robustness of airborne sound is not sufficient if the technology is based on a simple echo-hiding method [4,5]. Using airborne sound itself, there are also many uncertainties because when a signal is transmitted through the air and re-sampled with a microphone, it is subjected to possibly unknown nonlinear modifications [6] as well as various additive noises.

Also, in the echo method, audio watermarks are embedded by convolving a host audio signal with an echo kernel. Time-lag between the first echo and the direct sound takes discrete values enabling digital watermarking. However, in this method, if the coefficient is set to a large value to make embedded watermarks robust, the sound quality of the host signal becomes quite unnatural by the comb-filtering by the echo kernel, which is referred to as the coloring effect. If air propagation can be a very severe condition for audio watermarking, large coefficient value should be used to ensure robust detection but this causes significant sound quality degradation.

In the other hand, an innovative scheme for robust image watermarking based on vector quantization for transmitting over noisy channels was proposed [7-9]. This kind of robust watermarking technique can be applied to acoustic signals with various additive noises we are facing.

Although the results obtained here suggest an optimistic prognosis for using this technology for practical applications, further investigations will be needed to clarify what really happened and did not happen during the transmission process.

Furthermore, since the ultimate goal of the authors is to establish universal design (UD) environments for public audio announcements, it is particularly important to establish accessibility standards for airborne audio messages so that the technology can be used for everyday applications.

6. Conclusions. In this study, a portable acoustic caption decoder was developed based on echo-hiding technique and evaluated using airborne sound. The decoder successfully detected embedded data at a point 77m apart from a small portable loudspeaker with a sound pressure level of less than 40dB. It was also confirmed that embedded bits sequence was detected in a whole floor of a building using the real PA system equipped for the school. The results demonstrated not only the robustness of embedded data in airborne sounds but also the capability of HI technologies accompanying a mobile computer to play a role of useful assistive technology.

Further investigations are needed to clarify the parameters that affect performance levels as well as establish accessibility standards for airborne audio messages.

Acknowledgment. The authors wish to thank Mr. Shun-ichiro Abe, a graduate school student of Tohoku University, for his help in developing real-time decoding software program, we would not achieve the goal of this report. We also wish to express our gratitude to the many individuals at the affiliated school for the deaf at the University of Tsukuba. This research was partly supported by (1) a grant for collaborative research projects from the National Institute of Special Needs Education along with Research Institute of Electrical Communication, Tohoku University and (2) the Strategic Information and Communication R & D Promotion Programme (SCOPE), Ministry of Internal Affairs and Communications.

REFERENCES

- [1] W3C, Core techniques for Web content accessibility guidelines 1.0,
[Http://www.w3.org/TR/WCAG10-CORE-TECHS](http://www.w3.org/TR/WCAG10-CORE-TECHS)
- [2] Federal communications commission, closed captioning,
<http://www.fcc.gov/cgb/consumerfacts/closedcaption.html>
- [3] Research committee on support system for persons with hearing impairment using radio wave technology, *The Final Report*, 2006 (in Japanese).
- [4] Y. Suzuki, R. Nishimura and H. Tao, Audio watermark enhanced by LDPC coding for air transmission, *Proc. of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp.23-26, 2006.
- [5] A. Nishimura, Data hiding in speech sounds using subband amplitude modulation robust against reverberations and background noise, *Proc. of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp.7-10, 2006.
- [6] W. Bender, D. Gruhl, N. Morimoto and A. Lu, Techniques for data hiding, *IBM SYSTEMS JOURNAL*, vol.35, Nos.3-4, pp.313-336, 1996.
- [7] J. S. Pan, Y. C. Hsin, H. C. Huang and K. C. Huang, Robust image watermarking based on multiple description vector quantisation, *IET Electronics Letters*, vol.40, no.22, pp.1409-1410, 2004.
- [8] J. S. Pan, M. T. Sung, H. C. Huang and B. Y. Liao, Robust VQ-based digital watermarking for the memoryless binary symmetric channel, *IEICE Trans. on Fundamentals of Electronics, Communication and Computer Sciences*, vol.E-87A, no.7, pp.1839-1841, 2004.
- [9] J. S. Pan, H. C. Huang, L. C. Jain and W. C. Fang (eds.), *Intelligent Multimedia Data Hiding*, New Directions, Springer, Germany, 2007.

第IV章 資料編

配信内容

本日 5・6hは新入生歓迎会があります。

高2・3年の皆さん→12:40に上履き持参で寄宿舎食堂へ集合。

高1の皆さん→12:50に上履き持参で寄宿舎ラウンジに集合。

試験配信 1

試験配信 2

試験配信 3

英語検定申し込み締め切り日→5月10日(木)

明日、10日(木)は英語検定申し込み手続きの締め切り日です。

本日10日(木)は英語検定申し込み手続きの締め切り日です。

本日、木曜日は黒板消しクリーナーを整備する日です。

週番長さん→本日はゴミを捨てる日です。

副週番長さん→本日はゴミを捨てる日です。

漢字検定申し込み締め切り日→4月28日(土)

明日、28日(土)は漢字検定申し込み手続きの締め切り日です。

本日28日(土)は漢字検定申し込み手続きの締め切り日です。

校風美化委員に連絡→本日はモップ交換日です。

昼休みの体育館利用は13時15分までです。

生徒に連絡→試験週間です。

職員室・理科準備室・主事室・印刷室へは入室を禁止します。

校風美化より→29日放課後、2階西（体育館）側にWAXを塗布します

作業中、通行のご不便をおかけしますが、ご理解とご協力をお願いします

校風美化より→6月5日放課後、2階東側にWAXを塗布します

作業中、通行のご不便をおかけしますが、ご理解とご協力をお願いします

校風美化より→12日放課後、1階一生徒玄関にWAXを塗布します

作業中、通行のご不便をおかけしますが、ご理解とご協力をお願いします

校風美化より→19日放課後、3階西（体育館）側にWAXを塗布します

作業中、通行のご不便をおかけしますが、ご理解とご協力をお願いします

校風美化より→26日放課後、3階東側にWAXを塗布します

作業中、通行のご不便をおかけしますが、ご理解とご協力をお願いします

生徒に連絡→試験週間です。

職員室・理科準備室・主事室・印刷室へは入室を禁止します。

英語検定申し込み締め切り日→9月18日(火)

連休明けの18日(火)は英語検定申し込み手続きの締め切り日です。

本日18日(火)は英語検定申し込み手続きの締め切り日です。

普通科生徒会より→後期生徒会費を集金中です。

生徒に連絡→試験週間です。

職員室・理科準備室・主事室・印刷室へは入室を禁止します。

高等部普通科の生徒の皆さんへ→

本日はディベートに関するアンケートの〆切日です→

各学級委員長はアンケートを集めて→

昼休みまでに江幡まで提出をお願いします（^u^*）

校風美化より→明日・明後の2日間、1階のワックス作業を行います。

15日(木)のWAX作業は、担当班B班。作業区域は1階西(体育館側)

本日のWAX作業は、担当班B班。作業区域は1階西(体育館側)

16日(金)のWAX作業は、担当班C班。作業区域は1階東です。

本日のWAX作業は、担当班C班。作業区域は1階東です。

作業中、通行のご不便をおかけしますが、ご理解とご協力をお願いします

生徒に連絡→試験週間です。

職員室・理科準備室・主事室・印刷室へは入室を禁止します。

英語検定申し込み締め切り日→12月17日(月)

来週の月曜日、17日は英語検定申し込み手続きの締め切り日です。

本日17日(月)は英語検定申し込み手続きの締め切り日です。

生徒に連絡→入試問題を作成中です。

職員室・理科準備室・主事室・印刷室への入室は教員の許可を得て下さい

生徒に連絡→試験週間です。

職員室・理科準備室・主事室・印刷室へは入室を禁止します。

校風美化より→明日の放課後、西階段1~4階にWAXを塗布します。

通行のご不便をおかけしますが、ご理解とご協力をお願いします。

校風美化より→本日放課後、西階段1~4階にWAXを塗布します。

通行のご不便をおかけしますが、ご理解とご協力をお願いします。

漢字検定申し込み締め切り日→4月25日(金)

明日、25日(金)は漢字検定申し込み手続きの締め切り日です。

本日25日(金)は漢字検定申し込み手続きの締め切り日です。

英語検定申し込み締め切り日→5月13日(火)

連休明けの13日(火)は英語検定申し込み手続きの締め切り日です。

本日13日(火)は英語検定申し込み手続きの締め切り日です。

生徒に連絡→試験週間です。

職員室・理科準備室・主事室・印刷室へは入室を禁止します。

生徒に連絡→試験週間です。

職員室・理科準備室・主事室・印刷室へは入室を禁止します。

生徒に連絡→試験週間です。

職員室・理科準備室・主事室・印刷室へは入室を禁止します。

生徒に連絡→試験週間です。

職員室・理科準備室・主事室・印刷室へは入室を禁止します。

校風美化より→本日放課後、2階廊下にWAXを塗布します。

作業中、通行のご不便をおかけしますが、ご理解とご協力をお願いします

校風美化より→本日放課後、生徒用玄関・1階トイレ前にWAXを塗布します。

作業中、通行のご不便をおかけしますが、ご理解とご協力をお願いします

校風美化より→本日放課後、職員用玄関にWAXを塗布します。

作業中、通行のご不便をおかけしますが、ご理解とご協力をお願いします

今週は、今日(水曜日)に黒板消しクリーナーを整備してください。

英語検定申し込み締め切り日→9月16日(火)

来週の火曜日、16日は英語検定申し込み手続きの締め切り日です。

ただ今 英語検定申し込み受付中！

取れる資格は、早めに取りましょう！

本日は英語検定申し込み手続きの締め切り日です。

ただ今 漢字検定申し込み受付中！

次の金曜日、10月3日は漢字検定申し込み手続きの締め切り日です。

本日は漢字検定申し込み手続きの締め切り日です。

体育祭の写真を電子掲示板にて公開中。1、3階、2、4階が同一内容

高1が団結して制作した映画「僕らの絆」をご覧ください。

他に教室でおもちゃ展示をやっています。ぜひ来てね。高1☆3階へGO!

ただ今 漢字検定申し込み受付中！

取れる資格は、早めに取りましょう！

本日は漢字検定申し込み手続きの締め切り日です。

明日は漢字検定申し込み手続きの締め切り日です。

ただ今 英語検定申し込み受付中！

取れる資格は、早めに取りましょう！

本日は英語検定申し込み手続きの締め切り日です。

明日は英語検定申し込み手続きの締め切り日です。

共同研究契約書

(契約項目表)

1. 甲	独立行政法人国立特別支援教育総合研究所 理事長 小田 豊		
2. 乙	国立大学法人東北大学電気通信研究所 所長 矢野雅文		
3. 研究題目	電子透かし技術を応用した音響バリアフリーシステムの開発と形成的評価 —「パーソナル音響キャプションデコーダ」の実用化に向けて—		
4. 研究目的	障害者基本法に書かれた情報の利用におけるバリアフリー化を目指す一環として聴覚障害者を主な対象とした情報補償システムの開発を目指す。		
5. 研究内容	乙が開発した音響分野における電子透かし等の技術を用いて、音声情報に対して情報バリアフリーに必要な量の文字情報等を付加させるための手法、実用化に向けたシステムの研究を共同で行う。		
6. 研究分担	区分	氏名	所属部局・職名
	※甲における責任者 甲	棟方哲弥	企画部・ 総括研究員
	※乙における責任者 乙	鈴木陽一	人間情報システム研究部門・ 教 授
7. 研究スケジュール	甲乙協議の上、推進する。		
8. 研究実施場所	独立行政法人国立特別支援教育総合研究所、国立大学法人東北大学電気通信研究所		
9. 研究期間	平成19年4月1日 から 平成21年3月31日 まで		
10. ノウハウの秘匿期間	ノウハウが生じた時点より1年間		

甲と乙は、上記契約項目表記載の共同研究（以下「本共同研究」という。）を実施するにつき、次の各条のとおり共同研究契約（以下「本契約」という。）を締結し、この契約の締結を証するため、本契約書2通を作成し、甲、乙それぞれ1通を保管するものとする。

平成19年 4月 1日

(甲) 神奈川県横須賀市野比5丁目1番

独立行政法人国立特別支援教育総合研究所

理事長 小田 豊



(乙) 宮城県仙台市青葉区片平2丁目1-1

国立大学法人東北大学電気通信研究所

所長 矢野 雅文



(定義)

第1条 本契約書において、次に掲げる用語は次の定義によるものとする。

- 一 「研究成果」とは、本契約に基づき得られたもので、実績報告書中で成果として確定された本共同研究の目的に関する発明、考案、意匠、著作物、ノウハウ等の技術的成果及び研究の過程で生じた有体物をいう。
 - 二 「知的財産権」とは、次に掲げるものをいう。
 - イ 特許法(昭和34年法律第121号)に規定する特許権、実用新案法(昭和34年法律第123号)に規定する実用新案権、意匠法(昭和34年法律第125号)に規定する意匠権、商標法(昭和34年法律第127号)に規定する商標権並びに外国における上記各権利に相当する権利
 - ロ 特許法に規定する特許を受ける権利、実用新案法に規定する実用新案登録を受ける権利、意匠法に規定する意匠登録を受ける権利、商標法に規定する商標登録を受ける権利、並びに外国における上記各権利に相当する権利
 - ハ 著作権法(昭和45年法律第48号)第2条第1項第10号の2のプログラムの著作物及び同号の3のデータベースの著作物(以下「プログラム等」という。)の著作権並びに外国における上記各権利に相当する権利
 - ニ 技術情報のうち秘匿することが可能なものであって、かつ、財産的価値があるものの中から、甲乙協議の上、特に指定するもの(以下「ノウハウ」という。)
- 2 本契約書において「発明等」とは、次に掲げるものをいう。
- 一 特許権の対象となるものについては発明
 - 二 実用新案権の対象となるものについては考案
 - 三 意匠権、商標権、及びプログラム等の著作権の対象となるものについては創作
 - 四 ノウハウを対象とするものについては案出
- 3 本契約書において、知的財産権の「実施」とは、特許法第2条第3項各号に規定する行為、実用新案法第2条第3項に規定する行為、意匠法第2条第3項に規定する行為、商標法第2条第3項に規定する行為、著作権法に規定する複製、上映、公衆送信、頒布及び翻案並びにノウハウの使用をいう。
- 4 本契約書において「専用実施権」とは、次に掲げるものをいう。
- 一 特許法、実用新案法及び意匠法に規定する専用実施権、商標法に規定する専用使用権
 - 二 プログラム等の著作権に係る著作物について独占的に実施をする権利
 - 三 第1項第2号ニに規定するノウハウについて独占的に実施をする権利
- 5 本契約書において「通常実施権」とは、次に掲げるものをいう。
- 一 特許法、実用新案法及び意匠法に規定する通常実施権、商標法に規定する通常使用権
 - 二 特許法に規定する特許を受ける権利、実用新案法に規定する実用新案登録を受ける権利、意匠法に規定する意匠登録を受ける権利、商標法に規定する商標登録を受ける権利の対象となるものについて非独占的に実施をする権利
 - 三 プログラム等の著作権に係る著作物について非独占的に実施をする権利
 - 四 第1項第2号ニに規定するノウハウについて非独占的に実施をする権利
- 6 本契約書において「研究担当者」とは、表記契約項目表6.に掲げる者及び本契約第2条第3項に該当する者をいう。また、「研究協力者」とは、研究担当者以外の者であって本共同研究に協力する者をいう。

(様式 2) (規則第5条関係)

(共同研究に従事する者)

- 第2条 甲及び乙は、表記契約項目表6.に掲げる者を本共同研究の研究担当者として参加させるものとする。
- 2 甲及び乙は、必要があると認めるときは、相手方の同意を得て、共同研究に従事する者を客員研究員等として相手方のもとに派遣するものとする。
- 3 甲及び乙は、甲又は乙に属する者を新たに本共同研究の研究担当者として参加させようとするときはあらかじめ相手方に通知し同意を得るものとする。

(研究協力者の参加及び協力)

- 第3条 甲又は乙のいずれかが、共同研究遂行上、研究協力者の参加又は協力を得ることが必要と認めた場合、相手方の同意を得た上で、当該研究担当者以外の者を研究協力者として本共同研究に参加させることができ
- る。
- 2 研究担当者以外の者が研究協力者となるに当たっては、当該研究担当者以外の者を研究協力者に加えるよ
- う相手方に同意を求めた甲又は乙（以下「当該当事者」という。）は、研究協力者となる者に本契約内容を遵守させなければならない。
- 3 研究協力者が本共同研究の結果、発明等を行った場合は、第9条の規定を準用するものとする。

(実績報告書の作成)

- 第4条 甲及び乙は、双方協力して、本共同研究の実施期間中に得られた研究成果について、実績報告書を甲乙協議して定める時期までにとりまとめるものとする。

<実績報告書の内容例>

- (1) 研究題目
(2) 研究成果の概要
(3) 研究成果の今後の活用方法
(4) 研究経費の支出実績

(研究経費の分担)

- 第5条 共同研究の実施に係る費用は、共同研究の分担に応じ、甲及び乙がそれぞれ分担するものとする。ただし、共同研究を遂行するに当たり当事者の一方にとって著しく負担となる費用については、両者で協議して定める。

(施設及び設備の使用等)

- 第6条 甲及び乙は、本共同研究の用に供するため、それぞれの所有に係る施設及び設備のうち甲乙協議の上で指定するものについて、無償又は有償で使用させることができるとともに、共同研究に必要な研究機器の持ち込みを相互に認めるものとする。

(研究の中止又は期間の延長)

- 第7条 天災その他やむを得ない理由のため、共同研究の継続が困難となったときは、甲及び乙は協議の上、本共同研究を中止し、又は研究期間を延長することができる。この場合において、甲又は乙はその責を負わないものとする。

(様式 2) (規則第5条関係)

(知的財産権の帰属等)

第8条 共同研究の結果生じた発明等に係る知的財産権は、甲乙それぞれの貢献度を踏まえて、双方が所有するものとする。

- 2 共同研究の結果得られた研究成果の有体物等の所有権は、甲乙協議の上、その帰属等を決定するものとする。

(知的財産権の出願等)

第9条 甲及び乙は、本共同研究の結果、発明等が生じた場合には、迅速に、相互に通報するとともに、帰属の決定、出願等事務が円滑に行われるよう努めなければならない。

- 2 甲及び乙は、甲及び乙に属する研究担当者が共同研究の結果共同して発明等を行った場合において、当該発明等に係る出願等を行おうとするときは、当該知的財産権に係る甲及び乙の持分等を定めた共同出願等契約を締結の上、共同して出願等を行うものとする。ただし、甲又は乙が相手方から特許等を受ける権利を承継した場合は、甲又は乙が単独で出願等をするものとする。
- 3 甲又は乙はそれぞれ、甲又は乙に属する研究担当者が、本共同研究の結果それぞれ独自に発明等を行った場合においては、単独所有とし、単独で出願等の手続きを行うものとする。ただし、出願等に先立ち、それぞれ相手方の同意を得なければならない。
- 4 前二項の規定は、甲が発明者から発明等の権利を承継しない場合には甲に適用しない。

(ノウハウの指定)

第10条 甲及び乙は、本共同研究の結果生じた発明等に係る共有となった知的財産権（以下「共有に係る知的財産権」という。）のうち、ノウハウに該当するものについて、協議の上、速やかに指定するものとする。

- 2 ノウハウの指定に当たっては、秘匿すべき期間を明示するものとする。
- 3 前項の秘匿すべき期間は、甲乙協議の上、決定するものとし、原則として、表記契約項目表10.に掲げる期間とする。ただし、指定後において必要があるときは、甲乙協議の上、秘匿すべき期間を延長し、又は短縮することができる。

(外国出願)

第11条 前条の規定は、外国における発明等に関する知的財産権（著作権及びノウハウを除く。）の設定登録出願、権利保全（以下「外国出願」という。）についても適用する。

- 2 甲及び乙は、外国出願を行うにあたっては、双方協議の上行うものとする。

(実施契約)

第12条 甲と乙は、本共同研究の実施により発明等が生じた場合には、速やかに相互に通知し、当該発明等の実施に関する実施契約（以下「実施契約」という。）を締結するものとする。

(実施の許諾等)

第13条 甲及び乙は、共有に係る知的財産権の自らの持分を譲渡し、それを目的として質権を設定し、又は専用実施権を設定し、若しくは通常実施権を許諾しようとする場合には、それぞれ事前に相手方の同意を得なければならない。

- 2 甲又は乙は、前項の規定における通常実施権の許諾については、正当な理由がない限り、相手方に同意す

(様式 2) (規則第5条関係)

るものとする。

(独占的実施)

第14条 実施契約において、甲は、本共同研究の結果生じた発明等に係る甲が承継した知的財産権（本条第3項に規定するものを除く。以下「甲に帰属する知的財産権」という。）を、乙又は乙の指定する者に限り、実施契約による一定期間、独占的に実施（甲が第三者への実施許諾を行わないことをいい、以下「独占的実施」という。）させることができるものとする。

- 2 実施契約において、甲は、甲乙協議の上、共有に係る知的財産権を、乙又は乙の指定する者に限り、実施契約による一定期間、独占的に実施させることができるものとする。
- 3 甲は、乙又は乙の指定する者から前二項に規定する独占的に実施できる期間（以下「独占的実施期間」という。）を更新したい旨の申し出があった場合には、更新の是非及び更新する期間について、甲乙協議の上、定めるものとする。

(第三者に対する実施の許諾)

第15条 甲は、前条の場合において、乙又は乙の指定する者が、独占的実施期間中、契約で定めた年次以降において正当な理由なく実施しないとき、又は当該知的財産権を独占的に実施させることができが公共の利益を著しく損なうと認められるときは、乙又は乙の指定する者以外の者に対し、当該知的財産権の実施を許諾することができるものとする。

(実施料)

第16条 甲に帰属する知的財産権を乙又は乙の指定する者が実施しようとするときは、原則として、実施権者は実施契約で定める実施料を甲に支払わなければならない。

- 2 甲及び乙は、共有に係る知的財産権を実施しようとするときは、原則として、当該知的財産権に係る甲及び乙の持分に応じ実施契約で定める実施料を相手方に支払わなければならない。
- 3 共有に係る知的財産権を第三者に実施させた場合の実施料は、当該知的財産権に係る甲及び乙の持分に応じて、それぞれに配分するものとする。

(出願等費用)

第17条 甲及び乙は、共同研究の結果生じた発明等に係る知的財産権を共有する場合には、当該知的財産権の出願等及び権利保全に必要な費用（弁理士費用、出願料、維持費等）を、それぞれ持分に応じて負担するものとする。

(情報交換)

第18条 甲及び乙は、本共同研究の実施に必要な情報、資料を相互に無償で提供又は開示するものとする。ただし、甲及び乙以外の者との契約により秘密保持義務を負っているものについては、この限りではない。

- 2 提供された資料は、本共同研究完了後又は本共同研究中止後相手方に返還するものとする。

(知的財産権の放棄)

第19条 甲又は乙は、共有に係る知的財産権を放棄しようとする場合には、放棄する前に、その旨を相手方に報告しなければならない。

(様式 2) (規則第5条関係)

(物件に係る権利の帰属)

第20条 共同研究を行うために取得した物件に係る権利は、その費用を負担した者に帰属する。ただし、特別な事情のあるときは、この限りではない。

(秘密の保持)

第21条 甲及び乙は、本共同研究において知り得た一切の情報を秘密として扱い、相手方の事前の同意なしに、それらを第三者に開示してはならない。ただし、それらの情報が次のいずれかに該当するものである場合は、この限りではない。

- 一 開示を受け又は知得した際、既に自己が保有していたことを証明できる情報
- 二 開示を受け又は知得した際、既に公知となっている情報
- 三 開示を受け又は知得した際、自己の責めによらずに公知となった情報
- 四 正当な権限を有する第三者から適法に取得したことを証明できる内容
- 五 相手方から開示された情報によることなく独自に開発・取得していたことを証明できる情報
- 六 書面により事前に相手方の同意を得たもの

(研究成果の通知及び公表)

第22条 甲及び乙は、本共同研究の成果について相手方に通知しなければならない。

2 甲及び乙は、共同研究の成果の全部又は一部を発表しようとするときは、あらかじめ他の当事者の同意を得なければならない。

(契約の解除)

第23条 甲及び乙は、次の各号のいずれかの事実が生じ、かつ相当な期間を定めて催告後同期間内に是正されないとときは本契約を解除することができるものとする。

- 一 相手方が本契約の履行に関し、不正又は不当の行為をしたとき
- 二 相手方が本契約に違反したとき

(損害賠償)

第24条 甲又は乙は、前条各項に掲げる事由によって、又は甲、乙、研究担当者もしくは研究協力者が故意もしくは重大な過失によって、相手方に損害を与えたときには、その損害を賠償しなければならない。

(契約の有効期間)

第25条 本契約の有効期間は、表記契約項目表9.に掲げる期間とする。

2 本契約の失効後も、第4条、第8条から第22条、第24条及び第28条の規定は、当該条項に定める期間又は対象事項が全て消滅するまで有効に存続する。

(協議)

第26条 この契約に定めのない事項について、これを定める必要があるときは、甲乙協議の上、定めるものとする。

(様式 2) (規則第5条関係)

(書面による同意)

第27条 甲及び乙は、本契約において相手方の同意を必要とする場合には、原則として書面をもって行うものとする。

(裁判管轄)

第28条 本契約に関する訴えは、甲を所在地とする横浜地方裁判所の管轄に属する。

共同研究（共同研究機関：東北大学電気通信研究所）
電子透かし技術を応用した
音響バリアフリーシステムの開発と形成的評価
—「パーソナル音響キャプションデコーダ」の実用化に向けて—
(平成19年度～平成20年度)
研究成果報告書

平成21年3月発行

研究代表者 棟方哲弥（国立特別支援教育総合研究所）・鈴木陽一（東北大学電気通信研究所）
発行 独立行政法人国立特別支援教育総合研究所
〒236-8585 神奈川県横須賀市野比五丁目1番1号
TEL 046-839-6837 FAX 046-839-6909

