

第Ⅱ章

パーソナル音響キャプションデコーダの開発

第Ⅱ章 パーソナル音響キャプションデコーダの開発

1. はじめに

障害者基本法では、情報機器等における障害者の利便の確保を国と地方自治体の責務としており、参議院内閣委員会の同法の付帯決議には、情報機器等のみならず、コンテンツや通信サービスについてのバリアフリー化の実現に向けて万全を期することが明記された。

本研究が対象とする音響のバリアフリーの実現のためには、さまざまな情報補償の方法が考えられよう。例えば、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部会（2006）は、音声情報の代わりに、電波を使って、PDAに情報配信を行うシステムの“構想例”を述べている。実現に至っていないシステムについて意見を述べることは控えるが、当該構想が音とは別の電波という媒体を用いること、基本的に、伝える相手が特定できると仮定しているところは、本研究と本質的に異なると考えている。本研究の着眼点は、大変にユニークである。それは、音による情報伝達の特性、すなわち、音が聞こえる範囲にのみ伝わる特性に着目した点である。一般に、情報は、人を特定して伝達されるものであると考えられる。しかしながら、これは、所在に関係なく常にそのPDAを持つ者に何かを伝える、ということではない。

その一例として、ある浜辺の防災用スピーカから流れる津波注意報を考えると、この警報の伝えられるべき相手は、その時間帯に、その場所にいる不特定多数の人となる。音声の警報は、空間を広がることで、視覚情報とは異なり、特定の位置に常に注意を払っている必要なく、その音声の聞こえる範囲の人に向けて発せられる。

このことは警報に限らない。ある本を読んだことがある。そこには、聴覚障害の人が大人になって初めて気がついたことが書いてあったと記憶している。その人が子どもの頃、周りの健聴の子どもたちは、夕方の5時になると、誰に言われるわけではなく、家に帰って行くことが不思議で仕方がなかった。それが夕方のサイレンであり、あるいはサイレンの代わりに「七つの子」の曲であったことを、大人になって、はじめて知ったというものであった。

この音の情報伝達の媒体としての特性を、そのままに活かして、さらに聴覚障害のある場合などの情報補償を行うことが、本研究の主眼である。このための手法として構想したものが音の電子透かし技術である。しかしながら、研究の開始時点において、空気伝搬させた音の電子透かしの解析は未知の領域であった。

2. 研究の経過

研究の開始当時において、東北大学電気通信研究所（以下、通研）における電子透かしの解読のためのリアルタイム処理は、高性能のCPU（演算装置）を搭載したデスクトップコンピュータで実行されており、主には音楽CDなどの著作権データを扱う技術として研究されており、空気伝搬させた音の電子透かしは行われていなかった。

研究を開始した平成16年から平成17年度においては「LDPC符号による音電子透かしの耐性強化に関する研究」として、当時修士課程に在籍していたカクトウ氏により、透かし情報の解読精度の向上方法が明らかにされた。この間、国立特殊教育総合研究所（以下、特殊研）では、実証機に搭載可能なプログラムの完成を待ちながら、通研とのデータ互換を行うためのシステムの準備と合わせて、その応用方法について研究協力校である筑波大学附属聾学校と検討を進めた。第1章に述べた通り、同校が独自に運用を始めていた「校内文字情報表示端末システム」（廊下と教室に電光掲示板が固定設置されたタイプ）の利用状況等から、提示すべきメッセージ等について検討を行ってきた。

このような状況の中、通研において、空気伝搬させた電子透かしデータのリアルタイムの解読処理を行う研究室実演が公開されたのは平成18年10月であった。同研究所において、左右のスピーカから出力される別々の電子透かしデータを、一般的なノートパソコンによって解読することに成功した。そこでは多くの人が出入りする実用に近い状況で電子透かし情報の解読に成功した一方で、その段階で、常に動作が安定していたわけではなかった。このソースプログラムとCライブラリの提供を受け、特殊研においてPDAへの実装に取り組んだ。

次節では、プログラムが実装された「パーソナル音響キャプションデコーダ」の仕様、機能について述べ、認識率などの機能評価について報告する。

3. パーソナル音響キャプションデコーダ

3.1 システム概要

本共同研究の成果として開発された「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の解読して、聾学校の校内文字放送システムで用いられるメッセージを表示することが可能となっている。

図3と図4はシステムの外観を示す。



図3 システムの外観



図4 システムの外観 (寸法)

VGN-UX90S は、ハードディスクドライブ等の可動部品を持たない PC である。よって、バッテリー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) が装備されている。これには、マイクロフォンとしてバッテリー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 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、デコード結果をメッセージとしてが面に表示する部分の 2 つからなる。

デコード部分、表示部分ともに C 言語で記述されている。これは、特に音声すかしのデコードにおいて高速な演算性能が求められるためである。それに合わせ、表示部も C 言語で開発することとした。

これら、2 つの部分はそれぞれ独立したスレッドとして動作し、どのようなタイミングでアナウンスが流れてももれなく情報を取得できるように工夫されている。

それぞれ部分についての詳細は以降の節で述べる。

3.2 デコードスレッド

デコードスレッドは、入力された音声から電子透かしを抽出する機能を持っている。デコードスレッドでは、エコー法のアルゴリズムに基づき電子透かし (8 ビット) を抽出する。今回は 8 ビット中の 4 ビットを使ってメッセージを表現することとした。

残りの 4 ビットをメッセージコードとして表示スレッドに伝達する。ちなみに、この 4 ビットは、将来の拡張用—メッセージ種類の追加、メッセージの優先度の表現、部分的な点滅や色の変更などの装飾—に使用することができると考えている。特殊研は通研よりソースコードの提供を受け、今回の成果物作成のため一部を改変して使用している。

3.3 表示スレッド

デコードスレッドから送られた電子透かし情報は、一度キューに収められる。キューからデコードされた 4 ビットの電子透かしを順に取り出して、表示処理を行う。

この仕組みにより、一度に大量の情報が伝送されてきた場合でも、メッセージを確実に表示できるように工夫してある。4 ビットの電子透かしに対応するメッセージをデータファイルか

ら探し出して画面に表示する。

画面表示は、大きく分けて2つの領域からなる。図4の画面で「受信中」と赤字で表示されている部分は、音電子透かしが付与されたアナウンスを、受信していることを利用者に知らせるためのインジケータである。

受信中は赤く表示され、非受信中は何も表示されない。もう一方の領域が、メッセージ表示部である。この部分に放送の内容が表示される。メッセージは、右から左に向かってスクロールするようになっている。

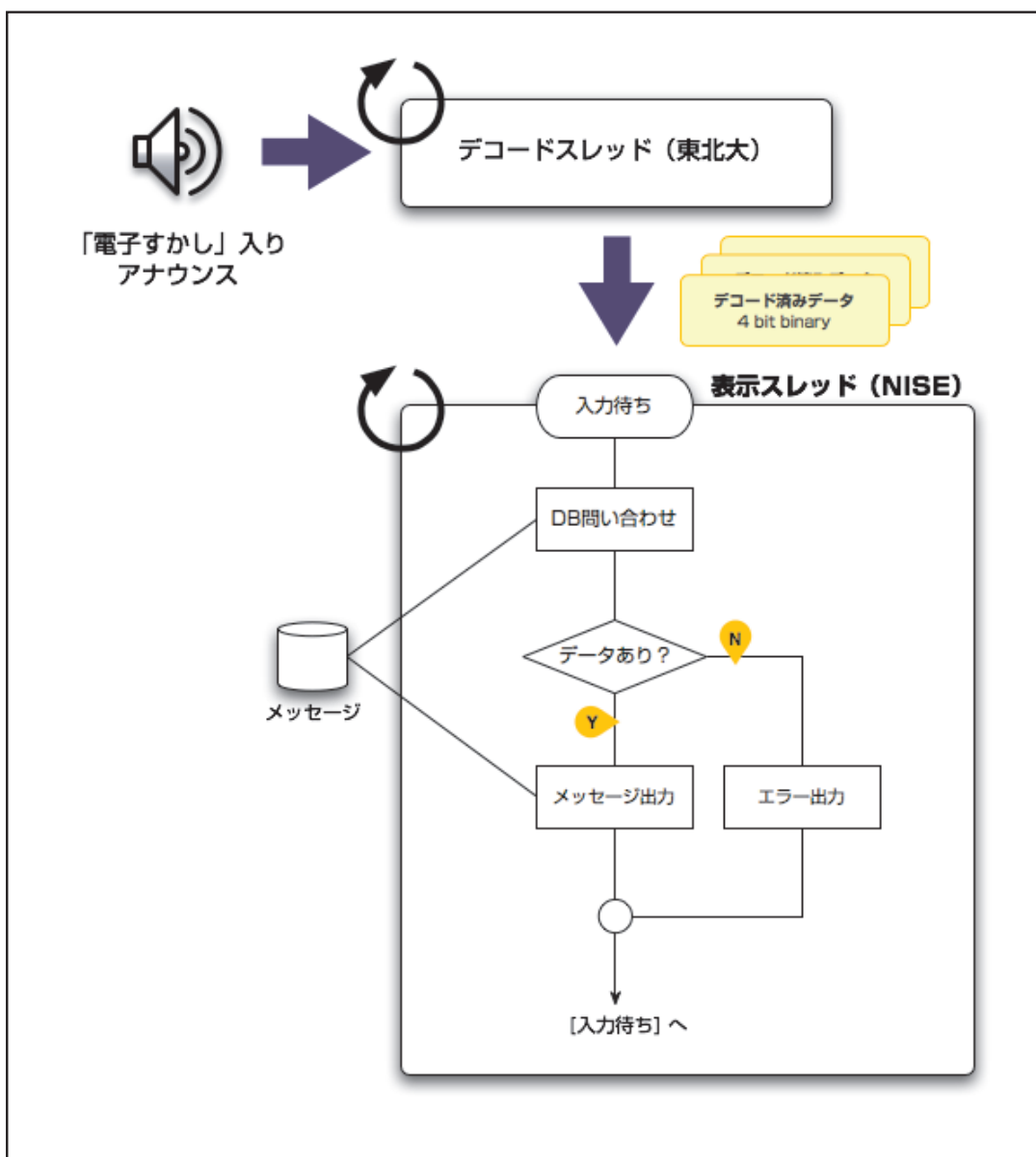


図5 動作説明図(フローチャート)

3.4 本システムの動作について

前のページの図5と、図6は、本システムの動作を説明したものである。

スピーカから流れる「電子透かし」は、現在のところ、エコー法という方式を用いて音声信号に組み込んでいる。

図5のスレッドと書かれている部分は、デコード処理と表示処理を平行して（マルチスレッド）処理を行うことでビット認識のタイミングを逃さないように工夫した。デコードスレッドと書かれている部分のリアルタイム処理部分は東北大学大学院阿部俊一郎氏によるものであり、表示スレッドのプログラムは国立特殊教育総合研究所の山口俊光によるものである。



図6 動作説明図（操作）

現在は4bps（1秒間に4ビットの情報）で電子透かし情報が組み込まれており、本デコードシステムは、このデータを読み取って、対応するメッセージを画面上に表示する仕様になっている。

第1章で明らかなように、表示する文字の大きさや、色、表示する文字の個数など、それぞれの利用者に個別化して提供することが望まれる。本システムは、個人が利用する端末であり、それらの個別化が可能な構成となっている。

エコー利用型の電子透かしは、人間が、直接音と反射音（エコー）の時間差が短い場合に、両方が融合して聞こえることを利用しており、図7に示すように、直接音と反射音の時間差を変えることで情報を埋め込む手法である。詳細は、第IV章に説明を行う。

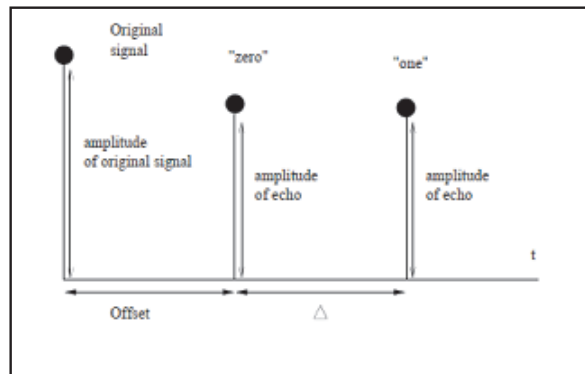


図7 エコー法によるビット表現
(カクトウ氏の修士論文より引用)

4. 評価実験

4.1. 目的

本システムが音電子透かしの受信ならびに埋め込まれたコードの実用に耐える解読機能を有することを、学校等の実際の建物や、グラウンドなどを想定した状況で行う評価実験により明らかにする。

4.2 方法

エコー法にて電子透かしを埋め込んだホワイトノイズをデータとして使用した。サンプリング周波数は44.1kHz、ビットレートは1,411kbpsのwav形式のファイルであり30秒間でファイルサイズが5.05MBである。これを繰り返しモードで再生して用いた。

データは、4bpsでデータが埋め込まれており、8bitを1単位として読み込み、その中の4bit部分にあるデータが8回正しく認識できた場合を“正しく認識”したと判断した。



図8 音源として使用したスピーカと音楽プレーヤ



図9 音圧を測定した精密騒音計

図8は音源である。実験における装置の可搬性を考慮してスピーカは単3電池4本で駆動する携帯可能な製品（Sony製SRS-T77）とした。また、同様な意図により、ホワイトノイズの再生はSony Walkman NW-E40Sを用いた。

音圧レベルの測定は、図9にある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。

4.3. 結果と考察

4.3.1 屋内における評価

国立特殊教育総合研究所の実験室と無響室においてシステムの動作を確認した後、屋内における評価実験の場所として、国立特殊教育総合研究所研究管理棟 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。平面図は、図10のとおりであった。

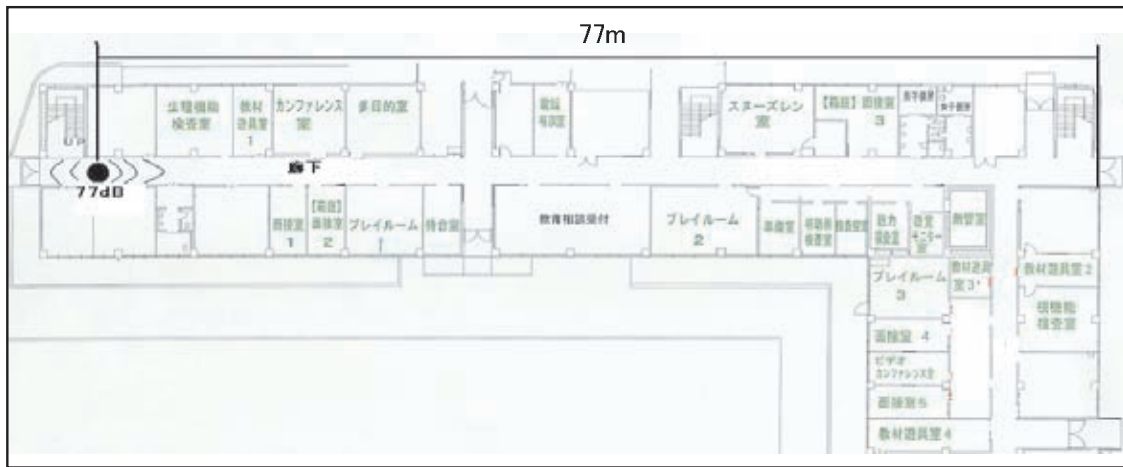


図10 屋内での評価実験場所（国立特殊教育総合研究所研究管理等1F）

図10の左端に見える黒丸を音源の場所として、図に示した位置において、音圧と認識率を確認した。結果は、図11のとおりである。赤い楕円は、電子透かしを正しく認識した測定場所であり、数字は、そのときの音圧レベルであり、赤い×印の入った四角は認識不能領域である。なお、青色に塗ってある部屋はドアを閉めて壁と同様としたものである。

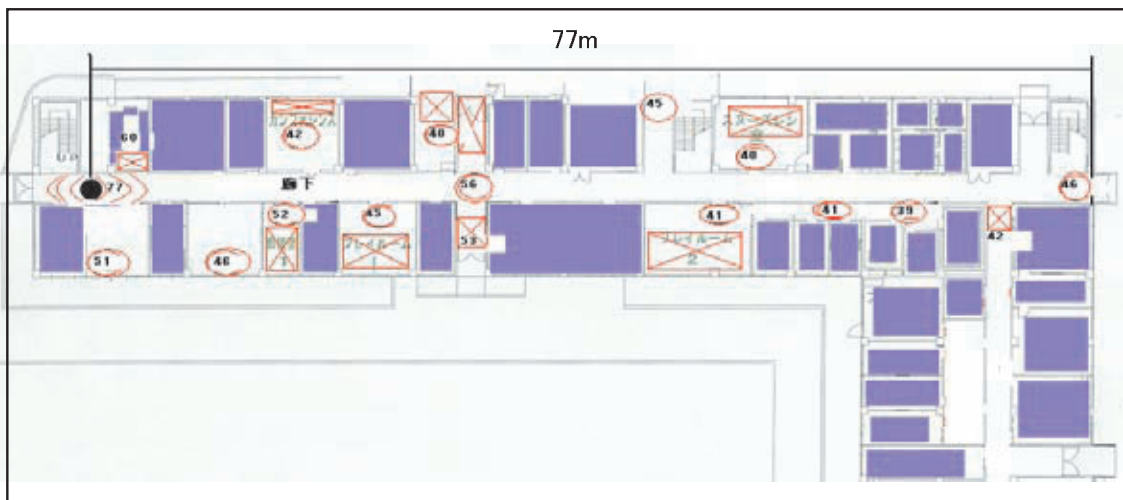


図11 屋内での評価実験結果

右端にある“46”と書いてある場所は音源から77mの距離が最大値であったが、実際には、物理的に、その先に進むことができない限界値であった。

結果から、40dB程度までは電子透かしを正しく認識することが観測された。一方、図にあるように60dBを越える音圧において、認識が行われない場所があった。その拡大図が図12である。図に書かれた60と52は、それぞれ音圧であり60dBと52dBの音圧があったことを考えると十分に認識可能な音圧レベルを思われた。エコー法は人工的に作成したエコーを透かしとして埋め込む手法であるため、実際の壁などにより生成される現実のエコーと干渉が起きていると考えられた。

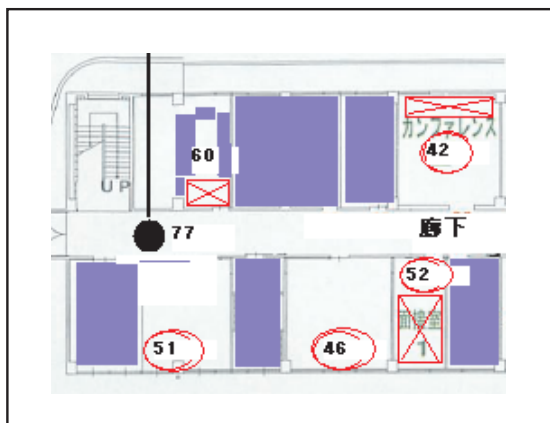


図12 屋内での評価実験結果（一部拡大）

より生成される現実のエコーと干渉が起きていると考えられた。

このことから、屋内において認識を行うためには、認識を行わせようとする部屋の大きさによる現実のエコーと、電子透かしとして人工的に付加するエコーの特徴に配慮する必要があると考えられた。

4.3.2 屋外における実機評価

屋外での評価は研究所の南庭で行われた。図13は、南庭の様子である。南庭は草などが生え、また、煉瓦の歩道が作られていることなどから、ここでは、実地の様子を説明する目的で、Google Earthによる研究所の写真を利用した。

図13に四角に囲った部分は1辺が32mの部分である。音源が90dBのときに、この領域は40～50dB程度の音圧が届く範囲であった。その一方で、音源を止めた際の音圧レベルは、35～60dBであり、実験者の耳にはホワイトノイズの存在が確認されていたが、音圧レベルからするとホワイトノイズのレベルと殆ど同程度の雑音があったことになる。



図13 屋外での評価実験場所（国立特殊教育総合研究所南庭）



図 14 屋外での評価実験結果

実験結果を、図 14 で説明する。電子透かしが正しく認識された部分は、図 14 に示した 50dB 以上のエリアにおいてであった。40～50dB のエリアでは認識のエラーが生じた。また、図 14 の緑色の部分は草地であり、すこし赤茶色に見える部分は煉瓦の床である。この煉瓦の部分では、40～50dB のエリアにおいても、電子透かしが正しく認識された。

5. 結論

屋内の実験では最小で 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、46dB の音圧では、音源からの距離が直線ではあったが 77m を記録し、廊下の長さの限界が無ければ、最小で 40dB の音圧で認識していることから、さらに大きな距離での認識が可能であったのではないかと考えられた。

その一方で、屋内での利用にあたっては、人工的に付加したエコーと部屋の大きさを考慮すべき事が明らかになった。

屋外では 50-55dB、距離として音源から 32m の位置で認識が可能であった。このことは、学校であれば、グラウンドの屋外放送への適用や、サイレンなどの警報への適用の可能性を示した結果と考えられる。

今回の実験では、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て、実用化に迫る成果を上げたものと考えている。

