

共有音場を活用したコミュニケーションの提案

大木 直人 阿部 圭一 岡田 謙一 松下 温
慶應義塾大学理工学部

近年、遠隔コミュニケーションの手段として注目されているサイバースペースを用いて、様々な形態のコミュニケーションを支援するためのシステム、Virtual Cocktail Party(VCP)を提案する。本システムは、臨場感のある音場を用い、あたかもユーザは仮想音場空間内を移動しているかのような感覚を得ることができる。この空間内で、ユーザはカクテルパーティ効果を利用して、自分の興味のある話題を見つけ参加することができる。さらにVCPは、グラフィカルな画像インタフェース、秘話機能等により快適なコミュニケーション環境を提供している。

A PROPOSAL OF COMMUNICATION TECHNIQUES APPLYING TO SHARED SOUND FIELD

Naoto Oki Keiichi Abe Ken-ichi Okada Yutaka Matsushita
Keio University
3-14-1, Hiyoshi, Kohoku-Ku, Yokohama 223, Japan

“*Virtual cocktail party*” is an application system for achievement of efficient and flexible telecommunication utilizing spatial audio interface. And, this system can support a number of physically separated users in a single shared audio (with supportive graphical) cyberspace. Users can circulate around this space as they wish, listening in or temporarily joining conversations and moving on as would minglers at a “real” cocktail party.

In this paper, we describe features of our system and techniques for implementation.

1 背景

テレコミュニケーションのインタフェース

近年、大都市に一局集中している都市機能の是正と、その解決策としての都市機能の郊外あるいは地方への展開が声高に議論されている。この実現のための手段として、郊外/地方と大都市、または郊外/地方相互の情報流通量の増大、品質の向上が求められている。そこで、大容量通信回線などの通信インフラストラクチャの発展により、様々な形態のテレコミュニケーションが研究され、また実用化されている。ここで、テレコミュニケーションにおいて重要であると思われるのが、ユーザインタフェースの問題である。どんなに優れたコミュニケーションツールも、インタフェース次第で使いものにならなくなってしまう可能性がある。

我々が理想とするテレコミュニケーションのためのインタフェースは、

1. 誰もが気軽に使えるような簡便で、直感的な操作環境
2. ユーザをコミュニケーションに没頭させる臨場感
3. 様々な形態のコミュニケーションが行える柔軟性、拡張性

を満たすものであると考える。これらの実現のために、本研究ではサイバースペースの概念を導入し、仮想音場空間を複数のユーザで共有することにより様々なコミュニケーションを可能にするシステムであるバーチャルカクテルパーティ(以下VCP)を提案する。

サイバースペース

まず、本研究について述べるにあたり、重要な概念であるサイバースペースについて触れておく。遠隔コミュニケーションの手段として、最近注目されている手法の一つがサイバースペースがある[1]。サイバースペースは、複数のユーザが相互にコミュニケーションを行うために提供される仮想空間を示す概念であり、声、顔、姿、しぐさで表現されるユーザと、テキスト、音声、画像などの話題情報、そしてそれらが存在する仮想コミュニケーション空間によって構成されている。サイバースペースの特徴を以下に実例を示しながら挙げる。

多対多コミュニケーション

従来のコミュニケーション手段では、1対1のコミュニケーション(通信)と、1対多のコミュニケーション(放送)が主となっていたが、サイバースペースにおいてはこれらを包括する多対多のコミュニケーションが可能となる。このとき、ユーザはある目的意識、例えば“これからAさんに会おう。”といったものを持ってサイバースペースに参加する場合もあるが、ただなんとなくぶらぶらとサイバースペース内を歩き廻って、偶然出会った人と会話を交わすようなコミュニケーションの形態も考えられる。サイバースペースを目指している多くのシステムでは、このような従来のコミュニケーションには見られなかった、目的意識を持たないコミュニケーションに関して注目しており、発見的コミュニケーションと呼ばれている。

日本において大手コンピュータネットワークがサービスを提供している habitat[1] では、ユーザはホストコンピュータに構築されている仮想的な街に、ネットワークを介してアクセスすることにより仮想世界の“住人”として参加できる。この仮想的な街をぶらぶらと歩き廻り、他の住人と仮想的に出会うことでコミュニケーションが成立する。このとき、複数のユーザが仮想的にサイバースペースのある場所に集まることによって、システムの許す限り何人もユーザが同時にコミュニケーションできる。

自由度の高いインタフェース

サイバースペースでユーザに提供される仮想世界は、バーチャル・リアリティ技術などによって創り出されるものであるため、時間、物理的距離、果ては物理法則に至るまでの現実世界における様々な制約を受けない。このため、ユーザインタフェースの自由度は極めて高く、何らかのメタファにするなど、従来から研究されている様々なユーザインタフェース構築のための手法を用いることができ、また全く新しいインタラクションをも実現できる。

前出の habitat では、仮想空間は街、建物、道路、公園などの現実世界が、まるで家庭用ゲーム機のロールプレイングゲームに登場する街のように表示され、ユーザもまた、ゲームに登場する人物のように表示される。そしてユーザ同士のコミュニケーションは、漫画のようにユーザの頭上に現れるフキダシの中に表示されるセリフを通して行われる。このような、いわゆる“ゲーム世代”に親しみやすいインタフェース

を採用することによって直感的な操作が可能なことが、habitat が広く親しまれている理由の一つであると思われる。

コミュニケーション以外への応用

サイバースペースは、先に述べたようにインタフェース設計における自由度が高いために、人間同士のコミュニケーション手段以外のユーザインタフェースとしても応用することができる。例えば、M.Benedikt は、文献 [7] において、マルチメディアデータベースのビジュアルゼーションにサイバースペースを用いることを提案している。

2 Virtual Cocktail Party

Virtual Cocktail Party(以下 VCP) は、リアルタイムに変化する音像によって構成された仮想音場空間を、複数のユーザが共有する、音によるサイバースペースである。この VCP は、ユーザ同士の偶然の出会いによるコミュニケーション、および蓄積されている様々な情報の獲得を支援するためのユーザインタフェースを提供することが目的である。カクテルパーティの会場で見られるように、それぞれ数人のユーザを含むいくつかの小グループができて、それぞれの話について会話している様子を模しており、ユーザはこの会場の中をぶらぶらと歩き廻って自分が興味を持っている話題を見つけ、そのグループに参加する。

以下に、VCP の持つ特徴を列挙する。

臨場感のある共有音場空間

VCP は、主に音声によってブラウジングおよびコミュニケーションを行うが、その利点として、以下の点が挙げられる。

- 効率的なコミュニケーション

人間同士のテレコミュニケーション手段には、音声、テキスト、映像などが考えられるが、このなかでも音声によるコミュニケーションが、単位時間当たりの情報量が最も大きいと言われている。もっとも、これらを組み合わせたコミュニケーションメディアでは、音声単独のコミュニケーションよりも効率が上がることは考えられるが、そうした混合メディアの

なかでも、情報伝達に音声の果たす役割は非常に高い。

- 音声による情報の選択

人間の聴覚には、複数の音源が存在する環境でも、自分の興味のある音に対しては敏感に反応する能力があり、これはカクテルパーティ効果と呼ばれている。この効果は、複数の音源が空間的に離れている場合に、顕著に現れるとされている [4] ために、VCP では後述の MISF を用いて空間的な広がりを持った音場を提供している。

- 容易な空間制御

映像による臨場感、立体感の研究が進められているが、立体的な映像をリアルタイムに得るには、高速の描画専用コンピュータを必要とし、またディスプレイも、ヘッドマウントディスプレイなどの特殊なディスプレイ装置を必要とするのが一般的である。一方、音に関してはヘッドフォン、あるいはいくつかのスピーカで臨場感・立体感を提供する様々な方式が提案されている [5] [6]。これらの方式では、比較的簡便なハードウェアで立体的な音像定位が実現できる。また、我々の提案した ISF [2] では、高速な音像定位処理を行い、リアルタイムに音像の移動を行うことが可能である。

同期コミュニケーション

VCP での同期コミュニケーションとは、参加者が同時に VCP に参加し、互いにリアルタイムにコミュニケーションを行うことを指す。これに対して、後に述べる非同期コミュニケーションでは、必ずしも参加者がリアルタイムにコミュニケーションを行ってはいない。この同期コミュニケーションを支援するための次のような機能が VCP にはある。

秘話機能

ある参加者グループで、グループ外の参加者には知られたくない話題を話す場合に、秘話機能が用いられる。この秘話機能には 2 つのレベルがあり、

1. 秘話を行っているグループの存在が、グループ外の参加者から分かるが、会話の

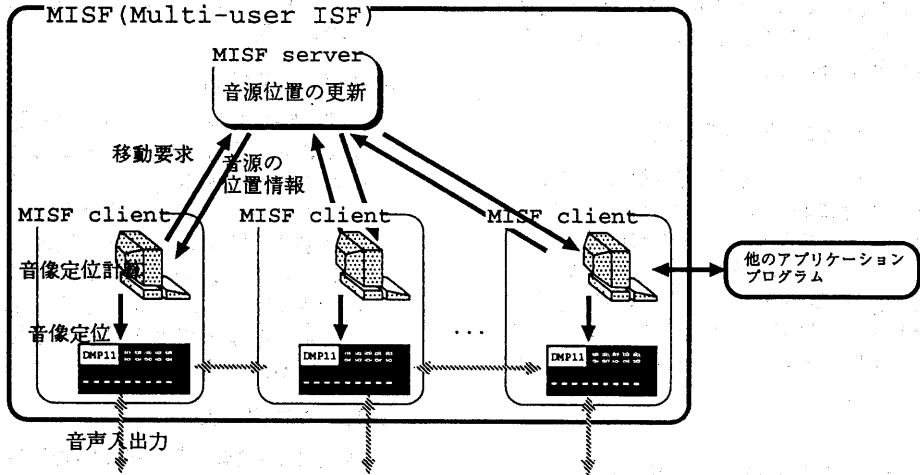


図 1: Architecture of MISF

内容は分からない

2. 秘話を行っているグループの存在自体が分からない

が選択できる。

非同期コミュニケーション

VCPでは、前述の同期コミュニケーション以外にも、非同期コミュニケーションをもサポートしている。ここで、非同期コミュニケーションとは、実際にはVCPの端末にユーザが向かっていなくても、予め自分のメッセージを登録しておくことによって、仮想パーティ空間に自分の声がある位置に定位して聴こえてくる。

例えば、新しいデータベースシステムを扱っている営業マンA氏が、VCPでこの売り込みを図っているとします。A氏は、このデータベースシステムの長所などを説明した音声データを作っておき、予めVCPのシステムに登録しておく。すると、A氏が実際には端末に向かっていなくても、仮想カクテルパーティ会場内で、A氏の“分身”が、このデータベースシステムの説明を喋ってくれるのである。

また、非同期コミュニケーションにおける音源は、必ずしも参加者の“分身”である必要はなく、システムが創り出した“仮想人物”でも構わない。例えば、カクテルパーティ会場内を

自動的に歩き回って音声や画像で様々な案内を行う、仮想案内嬢が組み込まれている。

3 実装

本システムVCPは、リアルタイムにユーザの動きに応じた音像定位を創り出すシステムであるISF(Interface with Interactive Sound Field)[2]を拡張し、複数ユーザに対応させたMISF(Multi-user ISF)を用いて共有音場空間を創り出している。

ISF

ISFは、仮想音場空間におけるユーザの位置に応じて、音場空間をリアルタイムに変化させ、あたかもユーザが音場空間内を移動しているかのような感覚を持たせることができる音インタフェースである。ISFは、他のアプリケーションプログラムからのユーザ、音源の仮想空間における位置情報に従って、ユーザの位置からの各音源の方角、距離を求め、その結果に基づいて、ハードウェアで音像を定位させる。音像定位の方法としては、相関係数変化法[9][10]、および、エフェクタを用いる方法[8][3]を選択して用いることができる。これらの手法は、HRTF(Head Related Transfer function:音源から聴取者の耳までの空間の音響的性質を表す伝達関数)を用

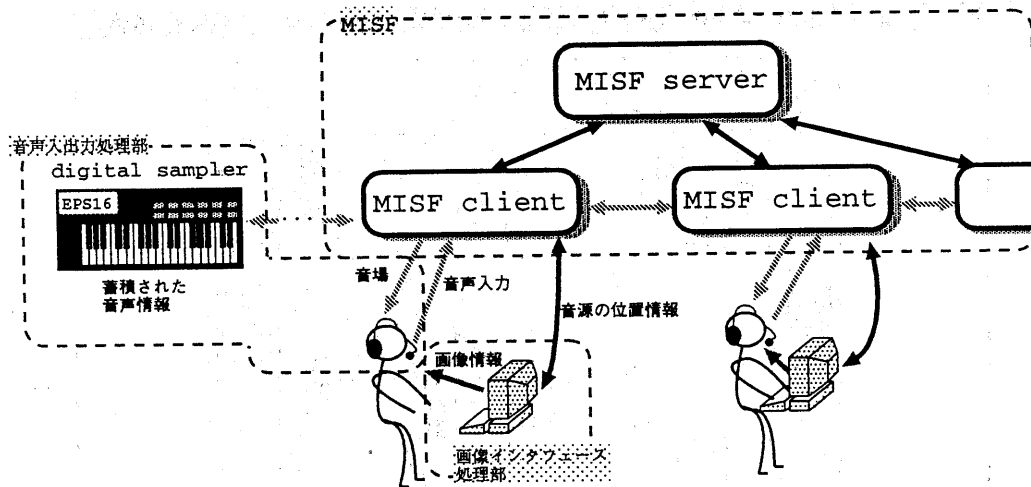


図 2: Overview of Virtual Cocktail Party

いるなどの他の方法に比べて、音像定位のための処理時間が短いという利点がある。このため、ISF では、複数の音源の音像定位をほぼリアルタイムに変化させることができる。

MISF

MISF は、ISF を複数のユーザが 1 つの音場空間を共有する事ができるように ISF を拡張した音インタフェースである。音像定位の手法に関しては ISF と同様の方法を用いているが、ISF と大きく異なる点は、複数ユーザへの共有音場空間の提供を効果的に行うために、サーバ・クライアントモデルを採用していることである(図 1 参照)。

MISF サーバは、仮想空間内における全ての音源の位置を管理しており、MISF クライアントの音源位置の移動要求(仮想空間内におけるユーザの移動)に応じて、その位置を更新し、新たな音源配置を MISF クライアントに提供する。MISF クライアントは、MISF サーバにユーザからの移動要求を伝達し、MISF サーバから提供される各音源の位置情報に従って、ユーザの位置からの各音源の方向、距離などを計算するソフトウェアと、ユーザの音声を入力し、また、立体的な音をユーザに提供するためのハードウェアによって構成されている。

VCP

前述の MISF を用いて Virtual Cocktail Party は、図 2 のような構成を持つ。

音声入出力処理部

VCP で扱う音声データには 2 種類ある。ユーザが同期に発声する音声と、あらかじめ蓄積され、非同期に発声される音声である。このうち、前者はユーザの声をマイクロフォンで拾音し、MISF の音像定位部に送られ、立体的な音声に加工されて各ユーザのヘッドフォンに伝えられる。また、後者については、あらかじめ音声データはデジタルサンプラに蓄積されている。このデジタルサンプラは、エンソニックジャパン社の EPS-16plus を用いており、32 種類の音声データをデジタル録音し、同時に 8 つの独立した音声ラインに出力することが可能である。この蓄積された音声データが、MISF の音像定位部に送られ、各ユーザのヘッドフォンに伝えられる。

MISF

前述の通り、MISF では音声入出力部に入力された音声信号を、仮想空間の各ユーザの位置に従って立体的に配置する。

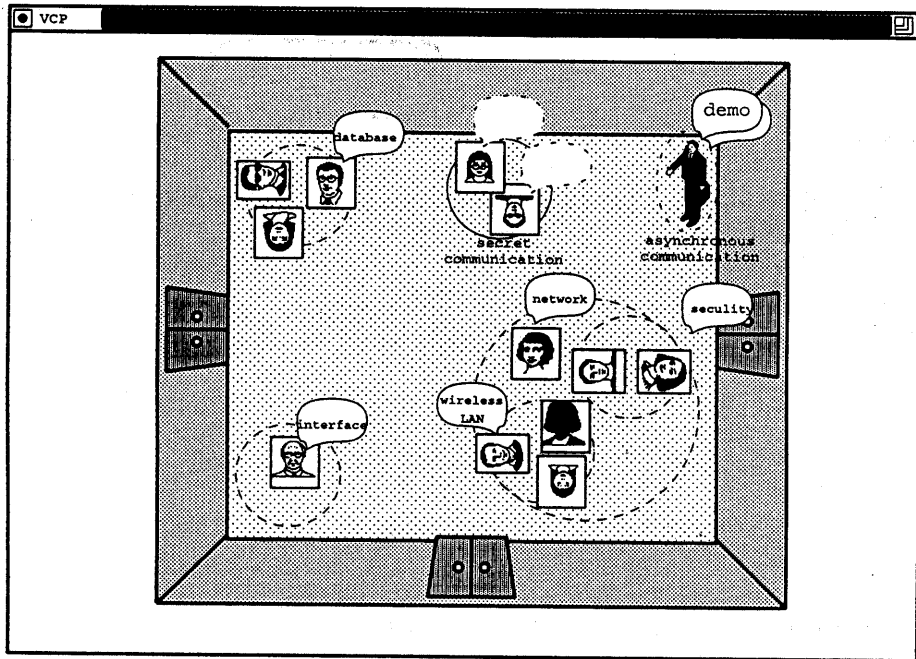


図 3: Image of graphical interface of VCP

画像インタフェース処理部

VCP では、音声インタフェースだけではなく、画像インタフェースもまた提供されている。画像インタフェースは、カクテルパーティ会場内をユーザが移動する場合に、その移動を入力するための手段として用いている。現バージョンで用意されているインタフェースは、カクテルパーティ会場を上から見た様子を模しており、参加者の位置は、それぞれのアイコンで示されている(図3参照)。ユーザは自分自身のアイコンをマウスでドラッグすることによって、カクテルパーティ会場内を移動する。このとき、他人のアイコンを移動することはできない。

また、画像インタフェースは、参加者相互のコミュニケーションに用いることもできる。例えば、参加者 A と参加者 B が会話をしているときに、A が静止画あるいは動画で話題の説明をしたい場合には、A と B の端末にその画像データが表示される。この画像データには相互に書き込みをする事もできるし、B がその情報を自分のデータとして保存しておくことも可能である。

4 結論

本論文で提案した Virtual Cocktail Party は、音による様々なテレコミュニケーションを実現するサイバースペースを指向したアプリケーションシステムである。これを用いることによって、様々な形態のコミュニケーションに対応できる柔軟なインタフェースが実現できる。

現在、試作システムが稼働中であり、今後この試作システムを用いて、ユーザインタフェースに関する定量的な評価を行う予定である。それとともに、音の立体感の向上、画像インタフェースの改良を行う予定である。

謝辞

本研究を行うにあたり、御助力を頂いているエンソニックジャパンインコーポレイテッドの方々に感謝いたします。

参考文献

- [1] Chip Morningstar, F. Randall Farmer, "The Lessons of Lucasfilm's Habitat", *Cyberspace:First Steps*, MIT press, pp.273-301, 1992
- [2] 阿部, 大木, 亀倉, 岡田, 松下, "インタラクティブな臨場感を持った音場インタフェース", 第45回情報処理学会全国大会, 1992
- [3] 大木, 亀倉, 阿部, 岡田, 松下, "人工現実感を用いた音情報検索", 情報処理学会情報メディア研究会資料, 11-1, pp.1-8
- [4] Cherry,E.C., "Some experiments on the recognition of speech with one or two ears", *J.Acoust.Soc.Am.*, 22, pp.61-62, 1953
- [5] 小泉, "音場制御", *テレビジョン学会誌*, Vol.45, No.4, pp.474-479, 1991
- [6] 森本, "音の距離定位について", *日本音響学会誌*, Vol.34, No.6, pp.356-361,1978
- [7] M.Benedikt, "Cyberspace : Some Proposals", *Cyberspace:First Steps*, MIT press, 1992
- [8] M.Cohen, "Extending the Notion of a Window System to Audio", *COMPUTER*, August, 1990, pp66-72
- [9] 黒住, 大串, "相関係数変化法による新しい音像の拡がり感制御方式", *電子通信学会論文誌*, Vol.J67-A, No.3, pp204-211, 1984
- [10] 黒住, 二階堂, 大串, "相関係数変化法による音像の距離感制御方式", *電子通信学会論文誌*, Vol.J67-A, No.9, pp872-879, 1984