

平成 12 年度
学士学位論文

サイバースペースにおける 音声通信処理方式の研究

A Study of Sound Communication System
for Cyber Space

1010412 辻 貴介

指導教員 島村 和典

2001 年 2 月 5 日

高知工科大学 情報システム工学科

要 旨

サイバースペースにおける 音声通信処理方式の研究

辻 貴介

従来の仮想空間における音声通信はサーバを介して行われていた。しかし、この方法ではサーバ負荷、ネットワーク経路に関する遅延が問題となる。本研究では、それを解決するためにマルチキャスト通信方式を用いた。

マルチキャスト通信を用いた場合、送信音声は1つで済むが、受信は再生音声分必要である。しかし、バンド幅、端末処理性能などにより、その数に制限がある。

音声の数を制限する方式として、従来は距離による選択を行っていたが、本研究では、距離とは別に、アバタの顔の方向と音声のボリュームを加えることで、より自然な音声対話を可能にしている。

キーワード 仮想空間, 音声通信, 多地点通信, ピア・ツー・ピア, マルチキャスト

Abstract

A Study of Sound Communication System for Cyber Space

Takasuke TSUJI

It goes through special server that conventional sound communications in Cyber Space. This system has some problems, which we load to special server and delay for the route of network. This study adopts multicast transport service to solve them.

In this multicast service based system case, the number of sound to send is only one, the number of sound to receive is necessary enough sounds to reproduce. But the number of sound to receive has a limit for network bandwidth and reproduction throughput.

The conventional systems limit the number of sound by distance between sender's avatar and receiver's avatar. The newly proposed system limits the number of sound by distance, face vector, and the sound level reproduced. This new limiting scheme might realize more natural sound communication than the conventional scheme.

key words CyberSpace, Sound Communication, Multipoint Communication, peer-to-peer, multicast

目次

第 1 章	はじめに	1
1.1	研究の背景	1
1.2	研究の目的	2
第 2 章	既存の技術	3
2.1	InterSpace	3
2.1.1	技術的内容	3
2.1.2	改善点と問題点	4
2.2	FreeWalk	7
2.2.1	技術的内容	7
2.2.2	問題点	7
第 3 章	採用する技術	9
3.1	特定サーバを設けない通信方式	9
3.2	multicast 通信方式	10
3.3	距離, 角度, ボリュームによる受信音声制御	11
第 4 章	システム内容	13
4.1	通信方式	13
4.2	制御方式	14
4.2.1	送信制御	14
4.2.2	受信制御	15
4.3	音の選択方式	15
4.3.1	扱う情報	15
4.3.2	音声 Volume の抽出	15

概念	15
実際の処理	15
4.3.3 処理方法	16
4.3.4 利点	16
4.4 再生音量制御	17
4.4.1 距離定数	17
4.4.2 角度定数	17
4.5 出力方法	19
4.6 音声処理システムの流れ	20
第5章 考察	23
5.1 トークグループによる音声対話方式	23
5.1.1 方式内容	23
5.1.2 利点	23
5.1.3 欠点	24
5.2 距離による音声選択方式	24
5.2.1 方式内容	24
5.2.2 利点	24
5.2.3 欠点	24
5.3 提案した音声選択方式	25
5.3.1 方式内容	25
5.3.2 利点	25
5.3.3 欠点	25
第6章 今後の課題	26
6.1 送信頻度	26
6.2 音声ボリューム判定間隔	26

6.3	発生予測	27
6.4	発生予測で扱う音声の数	27
6.5	高速に移動する音	28
第 7 章	まとめ	29
	謝辞	30
	参考文献	31
付録 A	仮想空間のシステム構成	32

目次

2.1	サーバからクライアントへミキシング処理の移行	5
2.2	優先順位リストの動作説明	6
2.3	送信先リストの動作説明	6
2.4	サーバを介さない音声通信	8
2.5	距離によるボリュームの設定	8
2.6	ミーティンググループの形成	8
3.1	サーバを設けない方式	10
3.2	アバタ間の距離関係	12
3.3	アバタ間の角度関係	12
3.4	アバタ間の距離と音声ボリュームの関係	12
3.5	自然な音声システムの例	12
4.1	3種類のマルチキャストアドレスの利用	14
4.2	音声 Volume の抽出概念	16
4.3	距離定数の設定	17
4.4	送信者から見た角度定数の設定	18
4.5	受信者から見た角度定数の設定	18
4.6	左右の音量比	19
4.7	LR 出力例	19
4.8	音声処理システムの流れ	20
A.1	仮想空間における各端末の機能別関連図	32

第 1 章

はじめに

1.1 研究の背景

ネットワークの高帯域化からマルチメディア通信の利用が容易になり、人々がネットワーク上で様々なコミュニケーションを行うようになった。

仮想空間によるコミュニケーションもその一つである。

従来、メディアの情報量は少なく、システムやコンテンツの開発、作成者が限られていたため、サービス毎に特定のサーバを設けて行うサービスが多かった。

しかし、年々PC利用者も増え、利用者の技術能力向上あるいはサービス利用の容易化している。また、各家庭のPC利用形態も、スタンドアロンからネットワーク常時接続となりつつあり、今後各サービスの利用者が増大するため、集中管理が困難になる。さらに、多くのマルチメディア情報はリアルタイムでの通信が必要となっている。

よって、サービス形態により、特定サーバの必要性が失われている。

音声通信においては、多人数同時会話が可能となり、ネットワークの高帯域化、端末処理性能の向上に伴い、その数は増加の傾向がある。

しかしながら、その数には制限があり、仮想空間システムのように不特定の相手との会話形態をとるサービスでは、最適な音声制御方式が必要となる。

1.2 研究の目的

従来の仮想空間における音声通信はサーバを介して行われていた。しかし、この方法ではサーバ負荷、ネットワーク経路に関する遅延が問題となる。それを解決するために、peer-to-peer 通信により、特定サーバを介さないで音声通信を行う。

この方法を採用すると、当然送受信する音声の数が増加する。送信音声を複数からひとつにする技術として、multicast 通信方式を用いることで解決する。

また、BandWidth, 端末処理性能により、仮想空間内で音声対話できる相手の数に制限がある。回線網の高速化、端末処理性能の向上に伴い、その数は徐々に増加している。しかし、再生可能な音声に制限があり、最適な音声選択方法を検討する必要がある。

第 2 章

既存の技術

本章では、音声通信処理を行う仮想空間コミュニケーションシステムのうち、音声通信を利用したもので、さらに音声選択制御を行っているものとして InterSpace と FreeWalk について述べる。

2.1 InterSpace

InterSpace とは、多数のユーザが、ネットワークに接続した PC により仮想空間を共有し、その仮想空間内で映像・音声・テキストチャット等マルチメディアを用いたコミュニケーションを行えるシステムである。

ここでは、NTTサイバースペース研究所が 2000 年 2 月 21 日に発表した、「共同仮想空間における音声コミュニケーション方式の検討」[1] について述べる。

2.1.1 技術的内容

インターネット等のデジタルネットワークを利用した、クライアント・サーバ型の多地点間音声コミュニケーションシステムを開発した。

内容は、従来サーバで行われていたミキシングを、クライアント側で行うことにより、サーバ負荷を減らすシステムの開発である。

具体的内容を次ページに記す。

- 従来の技術

1. クライアントは入力音声を A/D 変換し、圧縮 (GSM) してサーバに送信する.
2. サーバは受け取った音声パケットを伸長してミキシング処理を行った後に再度圧縮し、結果をクライアントに送信する.
3. クライアントは受信した音声パケットを伸長、D/A 変換し、音声出力する.

- 改善後の技術

1. クライアントからサーバへ、従来方式 1 と同じ音声パケットを優先順位リスト、送信先の情報と共に送信する.
2. サーバはミキシングせず、要求に従いクライアントに音声を送信する.
3. クライアントはサーバから受け取った音声パケットを伸長、ミキシング、D/A 変換し、音声出力する.

2.1.2 改善点と問題点

- 改善点

1. ミキシング処理をサーバからクライアントへ移行したため、インタースペース方式において以下のような点が改善された。サーバからクライアントへミキシング処理の移行を図 2.1 に示す.

- サーバでの負荷が軽減した

- サーバでの遅延が短縮された
 - ミキシング人数の制約がゆるめられた
 - 音像定位処理をクライアントで実施可能になった
2. 「優先順位リスト」, 「送信先リスト」を利用することで, 話し手と聞き手の意志を反映した.

● 問題点

1. 全ユーザの音声データ, 「優先順位リスト」, 「送信先リスト」をサーバが処理するため, サーバに負荷が集中してしまう
2. 遅延 (0.6 ~ 0.7sec ; 通信時間を除く) が大きい
3. 特定サーバによる管理集中型であるため, 全ユーザはサーバに依存してしまう
4. 受信音声の数の増加

優先順位リスト, 送信先リストの動作説明を, それぞれ図 2.2 及び図 2.3 に示す.

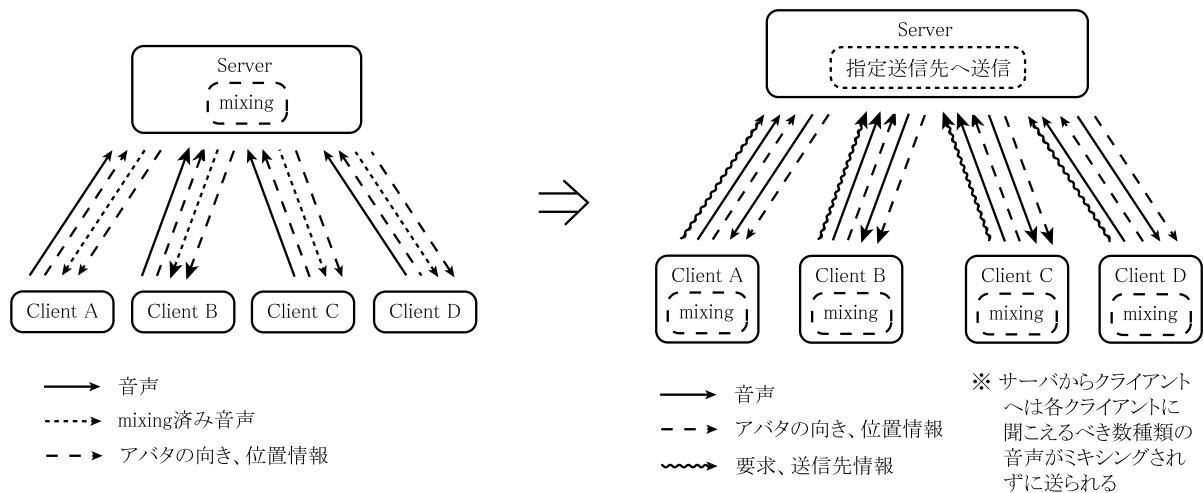


図 2.1 サーバからクライアントへミキシング処理の移行

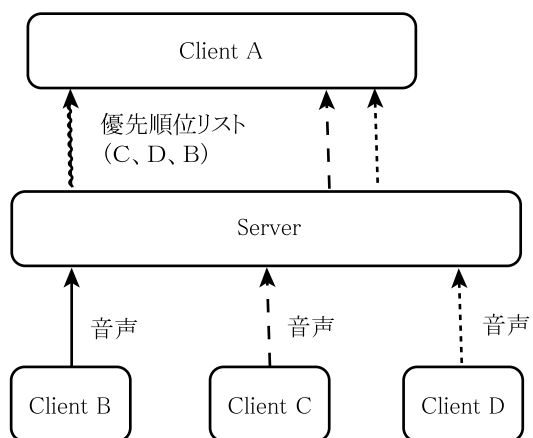


図 2.2 優先順位リストの動作説明

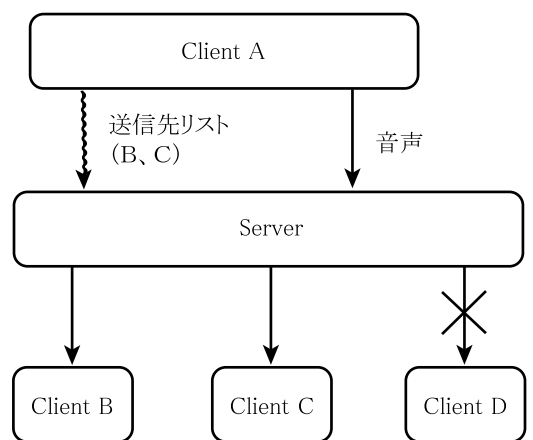


図 2.3 送信先リストの動作説明

2.2 FreeWalk

2.2.1 技術的内容

FreeWalk は京都大学で開発されたシステムである。

FreeWalk では、アバタ（ユーザの仮想空間内での化身）の位置情報についてはサーバを介して通信が行われているが、ビジョン（ユーザの顔情報）や音声などのマルチメディア情報に関する通信は、特定サーバを介さないでユーザ同士がユニキャストにより通信を行っている。サーバを介さない音声通信を図 2.4 に示す。

また、音場制御については、送信者のアバタと受信者のアバタとの間の距離により音声のボリュームを設定し、ミーティンググループの形成が可能である。距離によるボリュームの設定を図 2.5 に、ミーティンググループの形成を図 2.6 に示す。

2.2.2 問題点

- 送受信情報が多い

ユーザ同士のユニキャストによる音声通信では、送受信共に会話相手分の情報が必要であり、多人数音声対話では、その情報量は莫大なものとなる。

- 受信音声制御が最良でない

InterSpace 同様、距離のみでは最適な受信音声制御を成すとは言えない。

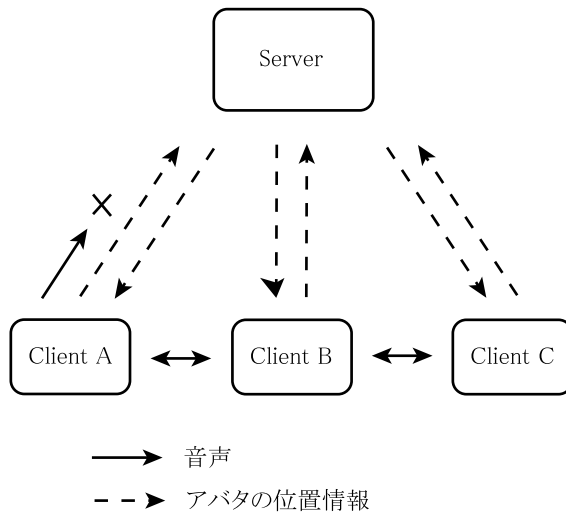


図 2.4 サーバを介さない音声通信

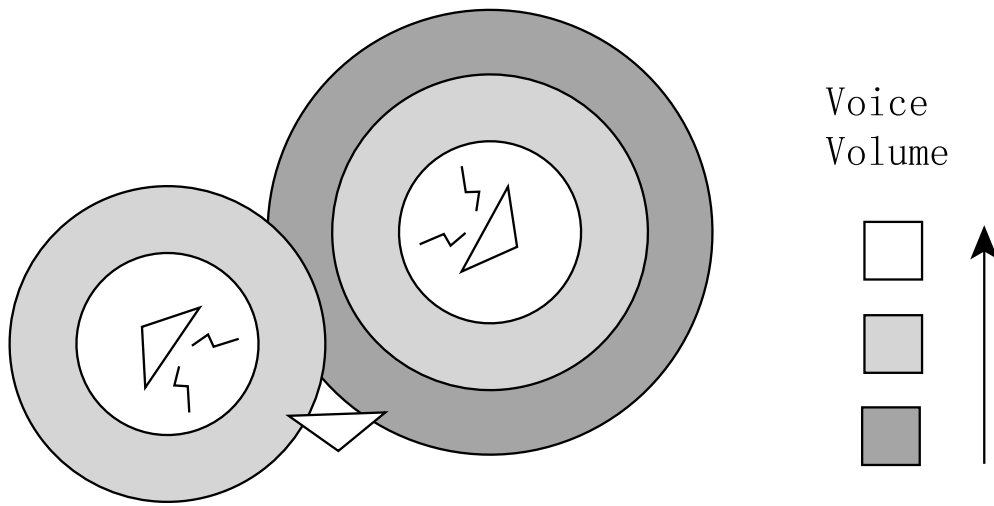


図 2.5 距離によるボリュームの設定

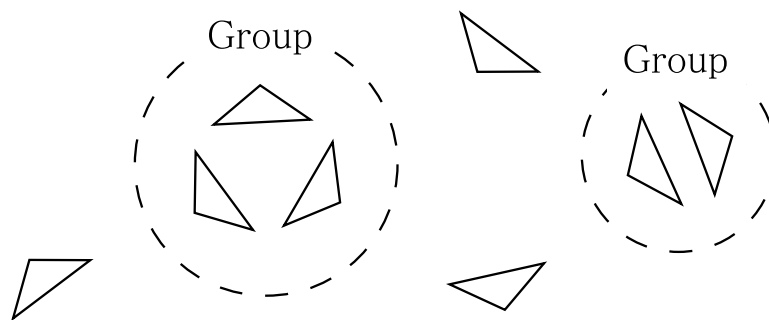


図 2.6 ミーティンググループの形成

第3章

採用する技術

本研究では次にあげる3つの技術(特定サーバを設けない通信方式, multicast 通信方式, 距離, 角度, 音声ボリュームによる受信音声制御)を重点におく.

3.1 特定サーバを設けない通信方式

以下に特定サーバを設けた場合の問題点をあげる.

- サーバ処理負荷の増大

全ユーザの音声情報を特定サーバで処理するため, 音声情報あるいはユーザ数が増加するの比例してサーバ処理量も増加する.

これにより処理遅延も生じる.

- サーバ付近のネットワーク負荷の増大

全ユーザの音声情報が集中するため, サーバ付近のネットワークで情報量が増大する.

- サーバに依存

集中管理型であるため, 全ユーザはサーバに依存してしまう.

- 通信経路による遅延

サーバを介することで, 通信経路が余分に必要になる場合がある. それに伴い遅延が生じる.

以上のような問題点より、特定サーバを設ける方式は好ましくない。

よって、クライアント・サーバ方式よりもピア・トゥー・ピア方式による音声通信方式の方が有効である。サーバを設けない方式を図 3.1 に示す。

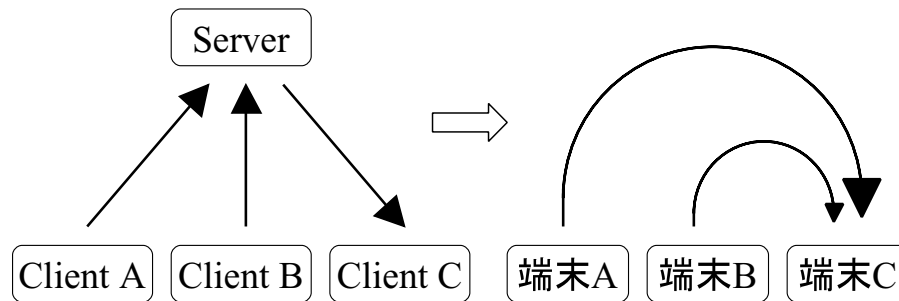


図 3.1 サーバを設けない方式

3.2 multicast 通信方式

ピア・トゥー・ピア方式で音声通信を行った場合、ユーザ端末の処理負荷が増え、新たに以下のような課題点が生じる。

1. 複数音声を送受信しなければならない
2. 音声情報が増えるにつれ、処理負荷も増大する

multicast 通信方式を用いることにより、上の課題点 1 のうち、送信する音声の数を 1 つにすることができる。

ユーザ A の音声を、あるマルチキャストアドレス (MAA とおく) に送信する。

ユーザ A 以外のユーザが、ユーザ A の音声を取得したい場合に、MMA を指定して音声情報を受信する。

3.3 距離、角度、ボリュームによる受信音声制御

受信音声が増えるにつれ、ユーザの端末処理負荷が増大することから、受信する音声情報を制限する必要がある。

既存の技術では、距離による音声制御が各種行われてきた。

しかし、その方法では、一定空間（音声を受信する空間）内に存在するユーザが音を発していない場合には、音声がかんたえず、会話が成立しない場合が生じる。

また、不必要な音声までを受信し、処理しなければならない。

音声制御に、アバタ間の距離、アバタ同士の角度及び音声ボリュームを用いることで、それらは解決する。

アバタ間の距離関係を図 3.2、アバタ間の角度関係を図 3.3、アバタ間の距離と音声ボリュームの関係を図 3.4 に示す。

図 3.2 の A は B よりも C に近く、音がかんたえやすい。

図 3.3 の矢印は音の方向を示す。A は C の方向を向いているが、B は C の方向を向いていない。この場合、A の音の方がかんたえやすい。

図 3.4 の矢印は音の方向を示し、長さは音の強さを表す。B は A よりも遠いが、音が大きいのでかんたえやすい。

具体的な内容は第 4 章で述べる。

この方式により、遠くにいるユーザでも、大きな声で話しているとその音がかんたえ、より自然なシステムを提供している。

自然な音声システムの例を図 3.5 に示す。

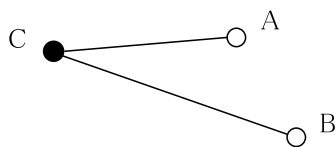


図 3.2 アバタ間の距離関係

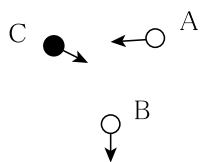


図 3.3 アバタ間の角度関係

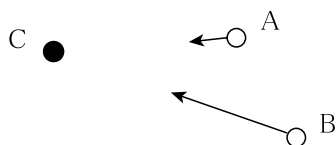
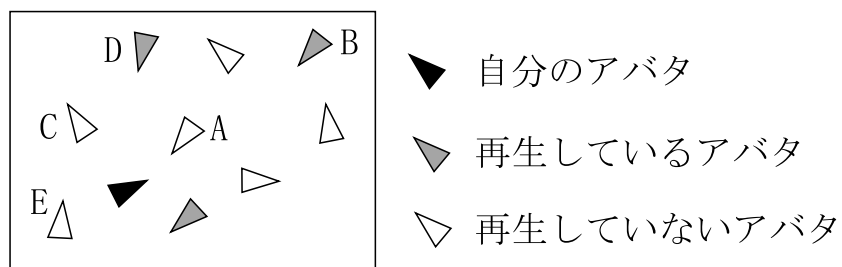


図 3.4 アバタ間の距離と音声ボリュームの関係



- A は近くにいますが音を出していない（無音）
- B は遠くにいますが大きな音を出している
- C は近くにいますが反対を向いている
- D は遠くにいますが自分に向かっている
- E は近くにいますが後ろにいる

図 3.5 自然な音声システムの例

第 4 章

システム内容

本章では、扱う音としてアバタの音声を取り上げたが、音声以外の音も同様に扱うことができる。

ただし、通常のアバタより高速移動する音については除外している。

そのような音への対処法に関しては、第 5 章の今後の課題でふれる。

4.1 通信方式

仮想空間における、音声処理に関する情報で送受信で扱うものは、音声データ、音声ボリューム、アバタの位置情報、アバタの顔の向き情報及び時間を表す情報である。

これらを以下の 3 種類のパケットに分け、それぞれ異なったマルチキャストアドレスに送信する。

ただし、以下の 3 は空間管理情報のパケットである。

1. 音声パケット

音声情報及び時間を表す情報

2. 音声の Volume を表すパケット

ボリュームを表す情報及び時間を表す情報

3. 空間管理情報パケット

アバタの位置情報、アバタの顔の向き情報及び時間を表す情報

上記3は同一空間に存在するユーザあるいは利用者全員から受信する。

2については同一空間、すなわち音声通信可能領域に存在するユーザあるいはそれ以下のユーザから受信する。

そして、1については再生する音声のみを受信する。

3種類のマルチキャストアドレスの利用を図4.1に示す。

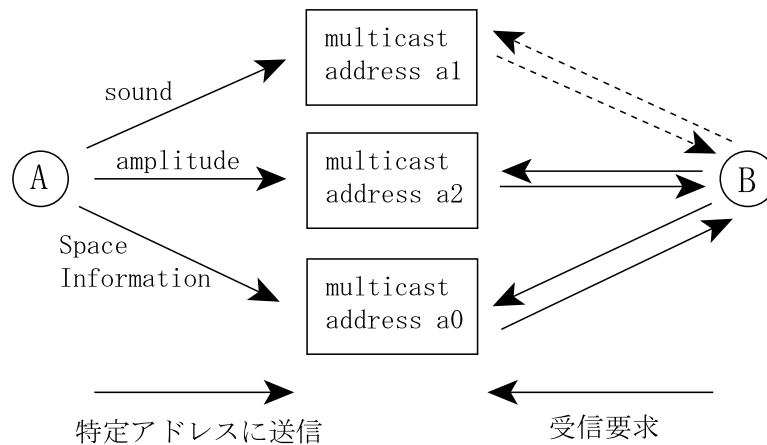


図 4.1 3種類のマルチキャストアドレスの利用

4.2 制御方式

4.2.1 送信制御

一定以下の Volume の音については送信しない。

ユーザが声を発していない場合でも、周辺雑音等が入り、多少の音声ボリュームが生じる場合がある。このような音声を送信しても音量が極めて低ければ受信要求がこないため、必要がない。

ただし、音声ボリュームに関するデータは送信する。

4.2.2 受信制御

空間の位置情報, 顔の向き情報, Volume をもとに一定人数分以下のデータを取得する.
上記情報は, 蓄積した一定時間前のデータを扱う.

4.3 音の選択方式

4.3.1 扱う情報

1. アバタの位置情報
2. アバタの顔の向き
3. 音の大きさ (Volume)

4.3.2 音声 Volume の抽出

概念

音声データは, 一定間隔で sampling されている.

その音声データを連続して複数集めると, ある時間区間の音声波形を作ることができる. この音声波形の最大振幅をだし, それをボリュームとして生の音声データと別に multicast で送信する.

音声 Volume の抽出概念を図 4.2 に示す.

実際の処理

sampling 毎に音声レベルを抽出する.

それを例えば, 1つの音声パケットに入れる音声の時間間隔 (20ms) を一区切りとし, 音声レベルの大小比較を繰り返すことにより, その間隔での最大音声レベルを割り出す.

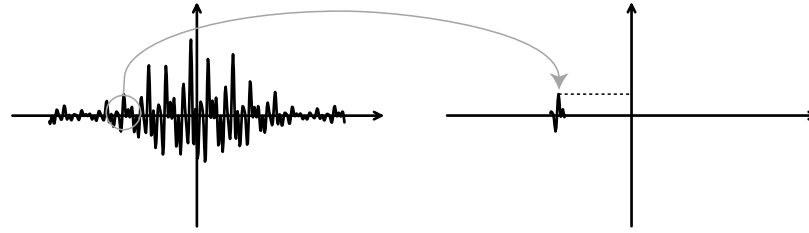


図 4.2 音声 Volume の抽出概念

4.3.3 処理方法

前で述べたボリュームデータとアバタ間の距離、角度から、聞こえる音の比を割り出し、入力された音の大きさと合わせて、実際に聞こえる音の大きさを導く。
その情報を比べることで、複数の音声から、出力音声を選択する。
受信者は、これにより選択された音声のみを受信する。

4.3.4 利点

ある一定範囲内に人がいなくても、それ以上の広い範囲で発生している音を受信することができる。
また、負荷がかかるだけの不必要なデータを受信、処理しなくてすむ。

4.4 再生音量制御

4.4.1 距離定数

アバタ間の距離により Volume を設定する.

距離定数 に対し, 距離に応じてその逆数をかけることにより, 距離による再生音量制御を行う.

これにより, 距離が近いほど音は大きく, 距離が遠いほど音は小さく聞こえるようになる.

距離定数の設定を図 4.3 に示す.

Distance :

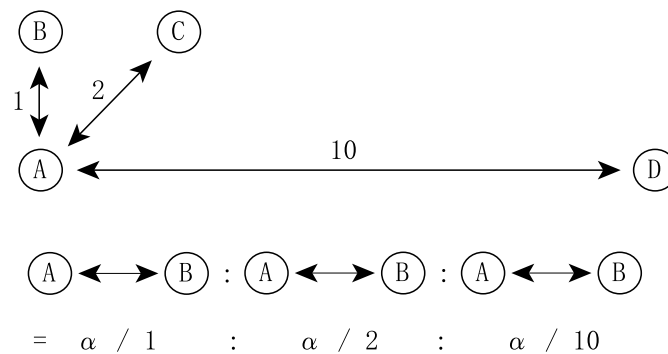


図 4.3 距離定数の設定

4.4.2 角度定数

アバタの位置関係により Volume を設定する.

アバタの位置関係の説明変数の一つとして角度定数を用いる. ここで, 角度定数とは, 着目する送信者を原点とし, 送信者の顔の方向を正とした場合, 正の方向を π ラジアン方向, 負の方向を 0 方向とし, その間の数量化を行った値とする.

また, 着目する受信者を原点とした場合も, 同様に定める. 送信者から見た角度と受信者から見た角度, それぞれに対し音の大きさを変える.

これにより、自分が相手に向かっている場合、あるいは相手が自分の方を向いている場合にそれぞれの音声が届きやすくなる。

送信者から見た角度定数の設定を図 4.4 に、受信者から見た角度定数の設定を図 4.5 に示す。

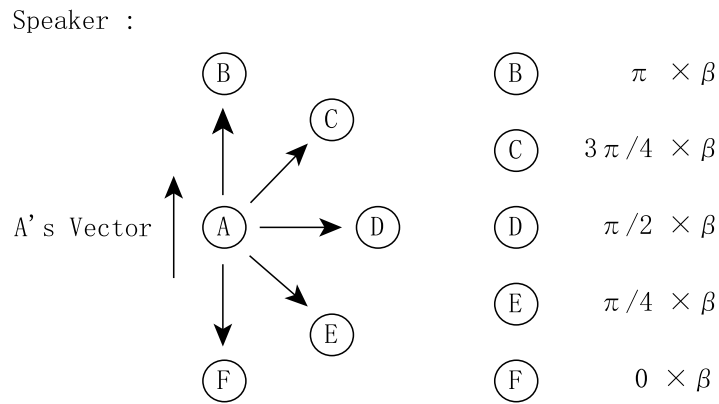


図 4.4 送信者から見た角度定数の設定

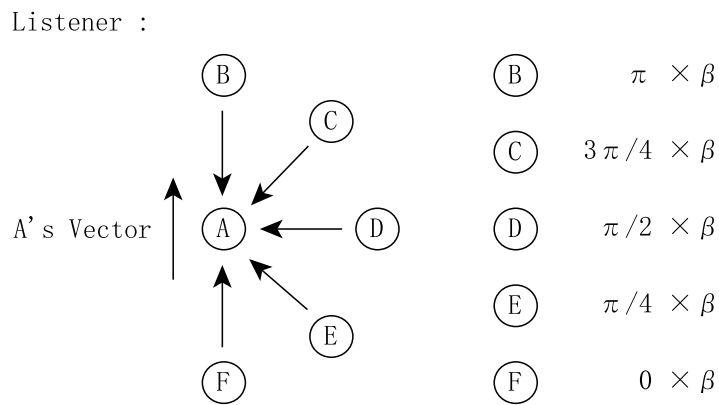


図 4.5 受信者から見た角度定数の設定

4.5 出力方法

再生音量制御の部分で受信した音声をアバタの顔の方向及び位置関係から音量の調節をした。

次に出力音源ここでは LR にあわせて音量の調節をする。

LR 出力により、音源の位置により左右の違いがわかるようにする。

そして受信した音声をすべて（その時間に出力されるものすべて）ミキシングし出力する。

左右の音量比を図 4.6 に、LR 出力例を図 4.7 に示す。

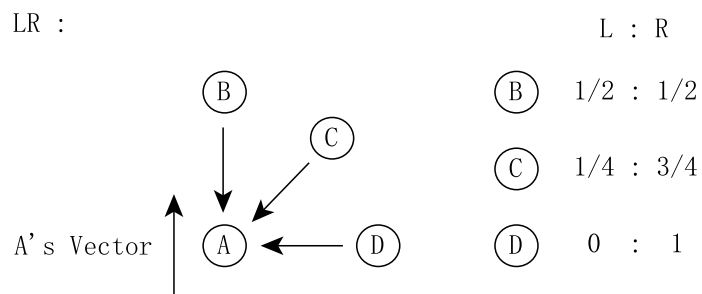


図 4.6 左右の音量比

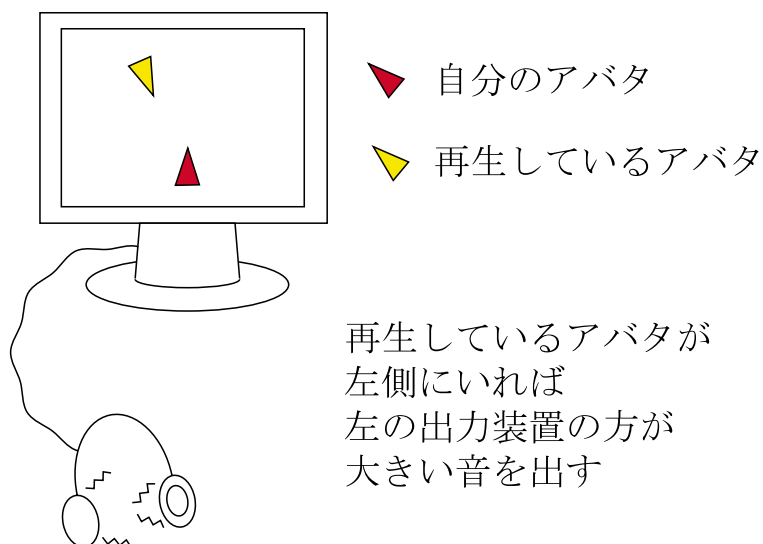


図 4.7 LR 出力例

4.6 音声処理システムの流れ

音声, 及び音声処理に必要な位置, 顔の向き情報の入力から, 音声出力までの全体の流れを説明する.

音声処理システムの流れを図 4.8 に示す.

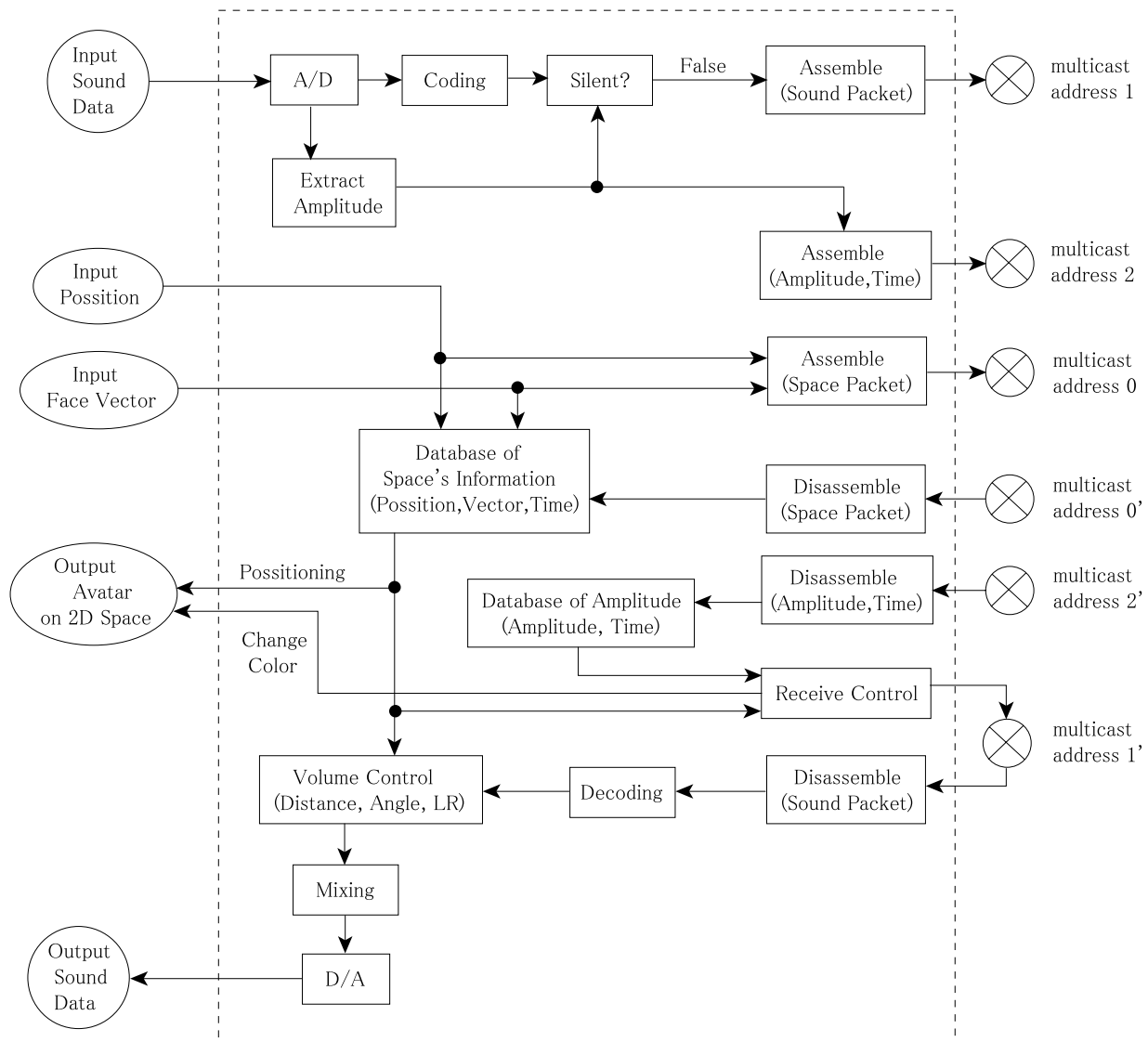


図 4.8 音声処理システムの流れ

以下は、前ページ図 4.8 の説明である。

入力された音声は、まずアナログからデジタルに変換（A/D 変換）される。

次に、符号化され、別の行程で得た音のボリュームデータをもとに、ある一定よりも小さい音であれば送信しない（送信制御）。

ある一定以上の音であれば、それを時間の情報と共にマルチキャストアドレス 1 へ送る。

A/D 変換後の音声データからは音のボリュームも抽出する。

抽出された音声ボリュームは送信制御に使われる。

また、音声データとは別に、時間情報と共にマルチキャストアドレス 2 へ送る。

アバタの動きはジョイスティック等で入力される。

入力されたデータから、アバタの位置情報、顔の向きがわかる。

位置情報と顔の向きは、空間情報として、時間の情報と共にマルチキャストアドレス 0 へ送られる。

また、受信制御、音量制御、アバタの表示のために空間情報のデータベースに蓄えられる。

同一空間あるいは利用者全員のアバタの空間情報をマルチキャストアドレス 0 ' から受け取り、空間情報のデータベースに蓄える。

また、同一空間すなわち音声通信可能領域に存在するアバタあるいはそれ以下のアバタの音声ボリュームの情報をマルチキャストアドレス 2 ' から受け取り、それを音声ボリュームのデータベースに蓄える。

空間管理及び音声ボリュームのデータベースより一定時間前で一定時間分の、それぞれのアバタの位置情報、顔の向き情報、音声ボリュームの情報を抽出する。

そして、それらの情報より、再生するアバタの音声を選ぶ。

選んだ音声データのみを、マルチキャストアドレス 1 ' より受信する。

その音声データを、複合化し、それぞれのアバタの位置情報及び顔の向き情報から音量を調節する。

そして、同一時刻の音声データをミキシングし、デジタルからアナログへ変換して、音声出力をする。

また、音声の出力とは別に、受信制御した情報をもとに、出力するそれぞれのアバタに変化、例えば色の変化などを加える。

<注意>

マルチキャストアドレス0から2は、あるユーザが送信するアドレスである。

マルチキャストアドレス0'から2'は、あるユーザが受信するアドレスで、こちらは単一ではなく、複数のアドレスを指す。ここでは、まとめてマルチキャストアドレス0'のように表している。

第5章

考察

再生に扱える音声の数に限りがある場合、音声の数をある一定以下にしなければならない。

既存の技術では、トークグループによる音声対話方式、及び距離による音声の選択方式が主として用いられている。

本研究では、音声選択に距離、顔の方向、及び音声ボリュームを用いることで、より自然な音声対話を可能にした。

5.1 トークグループによる音声対話方式

5.1.1 方式内容

トークグループによる音声対話方式では、音声対話をする場合に、あるトークグループに参加するかたちを採っている。

誰かと会話をしたい場合、その人が参加しているトークグループに参加要請を出す。許可されれば、その時点から、そのトークグループに参加したことになり、そのグループ内のアバタと自由に対話できる。

トークグループから抜きたい場合は、参加終了を提示しグループから抜ける。

5.1.2 利点

- 好きな人同士で会話ができる

- 聞かれたくない人に話の内容を聞かれない

5.1.3 欠点

- トークグループに参加するのに手間がかかる
- 非現実的である
- 自由なコミュニケーションシステムとしては不適切である
- 他のサービス、例えばミーティングルームなどでも同じようなことが可能である

5.2 距離による音声選択方式

5.2.1 方式内容

アバタ間の距離による音声選択方式では、自分から近いアバタの音声を優先的に受信し、再生する。

あるアバタと会話をしたければ、そのアバタに近づくことで、そのアバタの音声が聞こえるようになる。

5.2.2 利点

- 近くの音が聞こえるため、自然である
- 遠くの聞こえない音を受信しなくて済む

5.2.3 欠点

- 向かい合っていないなくても聞こえる
- 不必要な小さい音まで採用する

5.3 提案した音声選択方式

5.3.1 方式内容

再生音の選択に、アバタ間の距離、アバタの顔の方向、及び音のボリュームを用いる。

5.3.2 利点

- 向かい合っているアバタの音声を優先的に受信する
- 遠くであっても、大声で話している人の音を聞くことができる

5.3.3 欠点

- 処理負荷が増える
- 送受信する情報が増える

上の欠点について、処理負荷は増えても微量であり問題はない。

また、受送信する情報は増えるが、音声データの情報をそのまま受け取るよりもはるかに小さいし、受信してからの処理が軽減される。

上記3つの方式について、仮想空間内で最も自然に音声対話が可能であるものは、今回提案した方式であるといえる。

第 6 章

今後の課題

6.1 送信頻度

音声に関する情報で送信する情報は第 4 章第 1 節で記した。

音声情報, 音声 Volume 情報及び空間管理情報を別のマルチキャストアドレスに送信する理由は, 送受信の頻度の差が考えられるからである。

それぞれの送受信頻度を変えることでどのような効果をもたらすかを検証する必要がある。

6.2 音声ボリューム判定間隔

音声のボリューム, アバタ間の距離, 角度より, 一定間隔の音量を割り出し, その大小により音声の選択を行う。

しかし, ボリュームの間隔により会話が終了してしまったアバタの音声を選択してしまう可能性がある。これを防ぐために, 最適な音声ボリューム判定間隔を割り出す必要がある。

また, 次に説明する発生予測に必要な分だけの間隔をとると, 上に記した問題も解消される。

6.3 発生予測

ボリュームによる音声選択を行った場合、発生はじめの音をひろいきれない可能性がある。なせなら、扱うボリュームが一定時間前のものであり、それ以前の音声も反映するためである。発生はじめでも、十分大きい音でない場合は、発生終了の音より小さい可能性があり、その場合にその音は反映されない。

それを解消するために、音声ボリュームの変動から、発生音を決定する方法がある。音声ボリュームが小さくなれば発生終了の確率が高く、逆に大きくなれば発生はじめの確率が高くなる。このことより、簡単な線形予測で判定することが可能である。

6.4 発生予測で扱う音声の数

仮想空間にログインしているユーザの数、同一空間すなわち音声通信可能領域に存在するアバタの数、音声受信を行って再生する音の数及び発生予測で扱う音声ボリューム情報の数は、それぞれ異なる。

この中で、発生予測で扱う音声情報について、どの程度の数を取ることが最良であるかを検証する必要がある。

現在、同時再生可能な音声の数は 16 とされており、これは音声ボードの性能のためである。この数はその性能の向上から今後増えると予想される。

また、端末の処理性能も向上すると考えられる。

それらの数の変動から、発生予測で扱う音声情報の数も変動する可能性がある。

その数についても検証する必要がある。

6.5 高速に移動する音

高速に移動する音については、距離の変化量が大きくなるため、本研究で提案した方式では音をうまく拾いきれない可能性がある。

その対処法としては、新たに音源の動くスピードに関するパラメータを用意し、それを音声ボリュームと同様に multicast で受送信する。

受信したデータより、移動地点を予測することで違和感のない音の出力を行う。

これは、位置情報の送受信頻度にも関わる。

音の移動速度が遅ければ、位置情報を送受信する頻度は少なくて済むが、移動速度が速ければ、位置情報を頻繁に送受信しなければならない。

第7章

まとめ

従来の仮想空間における音声通信では Client/Server 通信あるいは peer-to-peer 通信により行われていた。

本研究では, multicast 通信を採用することで, 特定のサーバを使わないでより高速に, また, peer-to-peer 通信とは違い, 送信する情報をひとつに絞ることが可能になった。

また, 本研究では, BandWidth や端末処理性能などにより, 再生できる音声に限りがあることに着目した。

それを対処する再生音声選択方式に扱う要素として, 距離, 角度, 及び音のボリュームの3つを利用している。

それにより, 我々の日常生活の感覚により自然な仮想空間を実現する手法を提案している。

謝辞

本研究を行うに際し、多大なる御指導、御鞭撻を頂いた本情報システム工学科の島村 和典教授に深く感謝いたします。

また、本情報システム工学科の清水 明宏助教授、福本 昌弘講師並びに通信・放送機構、高知通信トラフィックリサーチセンターの加藤 寛治研究員に深く感謝いたします。

また、通信・放送機構、高知通信トラフィックリサーチセンターの神田 敏克研究員、高松希匠研究員、本研究室院生の中平 拓司氏に深く感謝いたします。

参考文献

- [1] 小長井 俊介, 森内 万知夫, 石橋 聡; 共有仮想空間における音声コミュニケーション方式の検討, ISSN 0913-5685, 信学技報 Vol.99 No.647, MVE99-73, pp.61-66.(2000.2.21)
- [2] Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura, and ToruIshida, *Kyoto University*, "FreeWark: A3D Vitual Space for Casual Meetings," IEEE MultiMedia, MediaSpaces, Volume6 Number2, April-June 1999, pp.20-28.

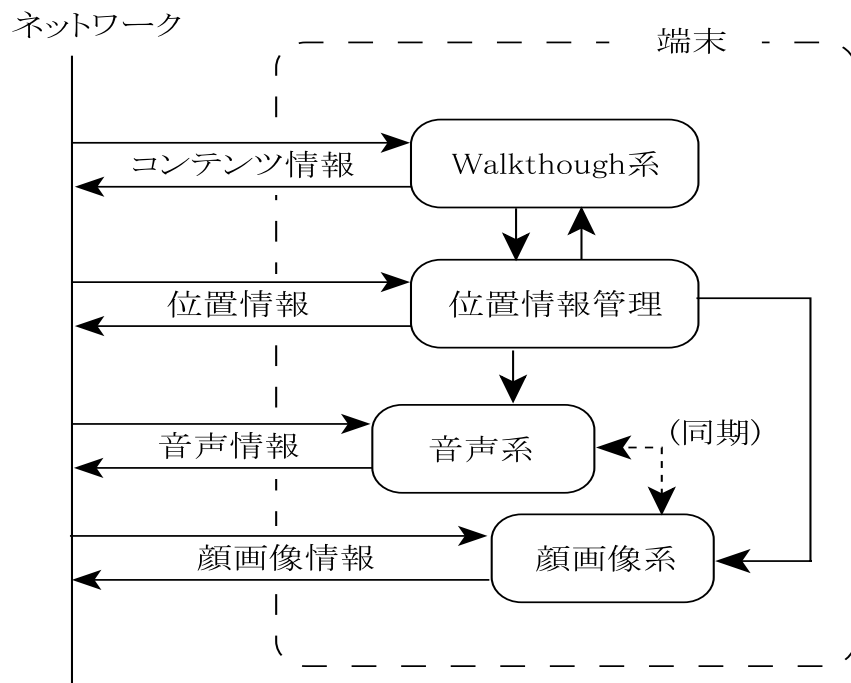
付録 A

仮想空間のシステム構成

仮想空間システムは大きく分けて、音声系、顔画像系、Walkthrough系、そして位置及びベクトル情報管理の4つの機能に分けられる。

上記4機能に分けることにより、他者への影響を小さくすることができる。

中でも位置、ベクトル、ここではアバタの顔の向き情報の管理は、他の3機能と関連が深い。仮想空間における各端末の機能別関連図を図 A.1 に示す。



※ 位置情報管理の出入矢印はすべて位置(& ベクトル)情報

図 A.1 仮想空間における各端末の機能別関連図

位置情報はクライアントのユーザインタフェース、例えばマウスなどにより入力されたものが、Walkthrough 系で処理され、音声、顔画像に送られる。

音声系では、位置、ベクトル情報を受信する音声の決定、音場制御の際に利用する。加えて、時間を表す情報も必要である。

また、音声と顔画像に同期をもたせる技術も考えられるが、本研究では扱っていない。