

立体映像とマルチチャンネルサウンドを用いた 複数人用マルチメディアコンテンツ環境に関する検討

中山 智量*¹ 北村 喜文*¹ 伊藤 雄一*¹ 岸野 文郎*¹

*¹ 大阪大学大学院情報科学研究科

本稿では、同じ場所と時間を共有している複数のユーザが、対面コミュニケーションを行う状況下で、全てのユーザにインタラクティブなマルチメディアコンテンツを提示する環境とその構築手法について述べる。具体的には、立体映像とマルチチャンネルサウンドを用いて、基本的に同一のコンテンツであるが、個人用にカスタマイズした情報も提示できる環境を構築するためのオーサリングツールの実装と、それを用いて作成されたマルチメディアコンテンツ提示環境を試作した結果について報告する。

A Study for Multimedia Content Environment using Stereoscopic Animations and Multichannel Sounds for Multiple Users

Tomokazu Nakayama*¹, Yoshifumi Kitamura*¹, Yuichi Itoh*¹, Fumio Kishino*¹

*¹ Graduate School of Information Science and Technology, Osaka University

We propose a method to construct an interactive multimedia content environment that allows multiple users to communicate in a face-to-face manner and share the same global time and space. We developed an authoring tool that can combine stereoscopic animations and multichannel sounds easily. In addition, we construct the multimedia content display environment where multiple users can see stereoscopic animation with individual sound.

1 はじめに

近年、同じ場所と時間を共有している複数のユーザが、同一のコンテンツを見ながら楽しむといった機会が増えてきている。例えば、複数人で競い合う TV ゲームで、ゲームの画面をみんなで見ながら楽しむ場面などが挙げられる。このようなコンテンツでは、ユーザの動きや操作などによってインタラクティブに情報が変化したり、同一のコンテンツ内でそれぞれのユーザに個人の情報が表示されることを望むユーザは多い。しかし、これまでそのような複数人用マルチメディアコンテンツを表示するシステムや、作成する手法については、検討されていなかった。

そこで本稿では、同じ場所と時間を共有している複数のユーザが、対面コミュニケーションを行う状況下で、全てのユーザに対して、基本的には同一のマルチメディアコンテンツをインタラクティブ

に表示することが可能な環境を構築した結果について述べる。具体的には、アニメーションとサウンドを読み込むことが可能なライブラリを用いて、複数人用のアニメーションの表示とマルチチャンネルサウンドの提示を実現する。また、同一のコンテンツに対して、個人のインタラクションに応じてカスタマイズした立体映像やサウンドを表示できる複数人用マルチメディアコンテンツ表示環境を構築するために、オーサリングツールを実装し、実際にそれを用いて環境を構築した結果について報告する。

2 複数人用コンテンツ提示システムとオーサリングツール

近年、対面コミュニケーションやコンテンツの共有の重要性について考えられるようになってきたことから、同じ時間や場所を共有している複数

のユーザが同一のコンテンツを楽しむことができるシステムが検討されている [1, 2]. そのようなシステムでは、コンテンツを共有するだけではなく、それぞれのユーザの動きや操作に応じてインタラクティブにその個人に特化した内容が表示されることを望むユーザも多い。しかし、同一のコンテンツを共有しながらも、そのような情報を表示するシステムは、今まで検討されていなかった。

一方、コンピュータの計算処理の高速化やネットワーク環境の整備により、映像や音声などのマルチメディアコンテンツに関する技術が普及してきている。オンラインゲームなどのコンテンツを例に挙げると、ユーザはネットワークを介して世界中の人々とコンテンツを共有することや、それらを用いて容易にコミュニケーションをすることが可能となっている [3]. また、車の設計会議や医療などの分野で用いられる立体映像 [4] や、エンタテインメントの分野で用いられるマルチチャンネルサウンド [5] などの様々なマルチメディアコンテンツが普及してきている。

そして、これらマルチメディアコンテンツを提示するデバイスとして、様々なものが提案されている。立体映像を提示するデバイスとしては、現在までに CAVE[6]、CABIN[7] などの没入感を提示可能な巨大ディスプレイ、Workbench[8] などの表示面を水平にした方式、HMD などの利用者の頭部全面にディスプレイを提示する方式 [9, 10] などが提案されている。複数人で立体映像を見ることが出来るディスプレイとしては、IllusionHole[11] や VirtualShowcase[12] などが提案されている。また、サウンドを提示するデバイスとしては、特定の方向だけに音声を伝達する超指向性スピーカ [13] や、方向に依存しない無指向性スピーカなど、目的や用途に応じて様々なデバイスが提案されている。

このように、様々な複数人用のコンテンツを提示するための個々のデバイスについては検討されているものの、複数のユーザが同一時間や場所を共有しながら、マルチメディアコンテンツを、皆で楽しむ共有のものとしてだけでなく、それに加えて、個人用にカスタマイズしたものとして捉えることができる環境は今までになかった。また、これらのデバイスで表示されるコンテンツを作成するためのツールに関しては、通常の映像やサウンドを組み合わせたマルチメディアコンテンツをユーザが容易に作成することができるオーサリングツール [14, 15] は開発されているものの、立体映像とマルチチャンネルサウンドを組み合わせることのできるオーサリングツールは存在しない。

そこで本稿では、同じ場所と時間を共有してい

る複数のユーザが、対面コミュニケーションを行う状況下で、全てのユーザに対して、立体映像とマルチチャンネルサウンドを組み合わせて、基本的には同一のマルチメディアコンテンツを個人ごとにインタラクティブに表示することが可能な環境を構築する手法について述べる。

3 複数人用マルチメディアコンテンツ環境の表示手法

本章では、複数のユーザが同一の時間と場所を共有しながら、立体映像アニメーションとマルチチャンネルサウンドを体験できる複数人用マルチメディアコンテンツ環境を表示する手法について述べる。

3.1 立体映像アニメーション

構築する複数人用マルチメディアコンテンツ環境では、立体映像アニメーションとして、3D キャラクタを用いたアニメーションを表示する。ここでは、立体映像を用いたアニメーションの読み込みと表示を行うために、DirectGraphics ライブラリを用いる。このライブラリにより、市販の 3D モデリングソフトで作成したアニメーションモデルの表示と、複数のアニメーションの制御を実現する。

さらに、3次元トラックやカメラ等でユーザの視点位置を取得してから、立体映像を生成することにより、ユーザの動きに応じてインタラクティブに 3D キャラクタの向きが変化する立体映像アニメーションの表示を実現する。これを、DirectGraphics ライブラリとグラフィックスカードのマルチヘッド機能を用いることにより、それぞれ左眼用画像と右眼用画像の 2 つの画面を表示し、それぞれの眼用の 2 台のプロジェクタを用いて、偏光方式の立体視が可能な立体映像アニメーションの表示を実現する。

3.2 マルチチャンネルサウンド

構築する複数人用マルチメディアコンテンツ環境では、マルチチャンネルサウンドを用いて各ユーザに個別のサウンドを提示する。まず、DirectSound ライブラリを用いてマルチチャンネルサウンドに対応したバッファを用意する。例として、7.1 ch のマルチチャンネルサウンドに対応したバッファの構成を図 1 に示す。図 1 中、ch 1 から ch 8 までの正方形の領域は、それぞれの出力チャンネルから再生するサウンドを格納する領域 (16 bit) を表している。この例では、ch 1 から ch 8 のブロックが 1 サンプルである。DirectSound では、バッファを再生する場合、サンプル 1 を再生した後、サンプル 2 を再生し、その次にサンプル 3 を再生… のように、サンプルを順に読み込み出力するので、同じ



図 1: マルチチャンネルサウンド対応バッファの構成

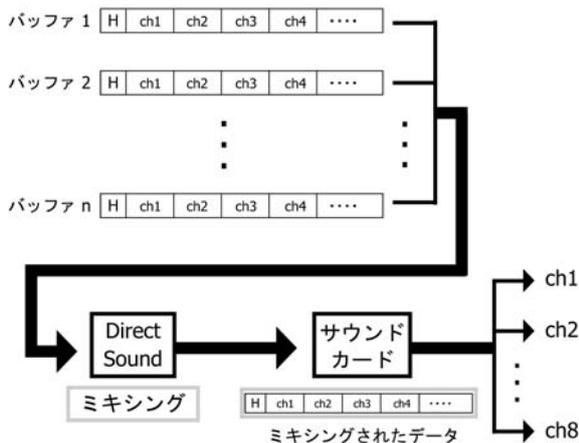


図 2: 複数のバッファのミキシング再生

サンプル内の出力チャンネルの各ブロックは、同じタイミングで再生される。例として、サンプリングレートが 44 kHz のサウンドでは、サウンドを 1 秒間再生するためにサンプルが 44,000 個分用意され、サンプル 1 から順に出力される。なお、図 1 の H のブロックには、ヘッダ情報として、バッファのフォーマット形式が書き込まれている。

また DirectSound ライブラリにより、複数のサウンドバッファを生成することができ、同時に複数のサウンドを出力する際には、複数のサウンドバッファを同時にミキシングして再生することができる。図 2 に示すように、再生される複数のバッファは、DirectSound によってミキシングされたデータとしてサウンドカードに出力され、サウンドカードのそれぞれの出力チャンネルからサウンドが出力される。

しかし、DirectSound では、バッファごとに再生や停止、ボリュームの変化を制御できるが、同じバッファ内でチャンネルごとに再生や停止、ボリュームの変化を制御することができない。つまり、同じバッファ内の異なるチャンネルごとにサウンドの再生や停止のタイミングや、サウンドのボリュームを変えたりできない。そこで、各ユーザーごとにサウンドの再生や停止のタイミング、サウンドのボリュームを変更可能とするため、図 3 の

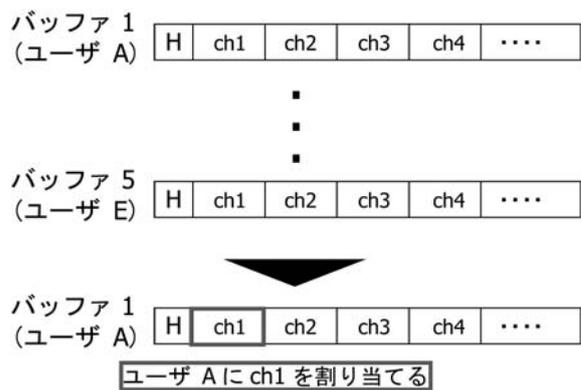


図 3: ユーザとバッファ、出力チャンネルとの対応づけ

ようにユーザーの数だけバッファを用意し、ユーザーごとに異なるバッファと出力チャンネルを割り当てる。図 3 は、ユーザー A～E までサウンドを出力すると仮定し、ユーザーそれぞれにバッファ 1～5 までを割り当て、さらにユーザー A には ch 1 を割り当てる場合を表す。

ユーザーに異なるバッファと出力チャンネルを割り当てた後、あるユーザーに割り当てられたバッファの出力チャンネルのブロックにのみ、そのユーザー用のサウンドデータを書き込み、それ以外のブロックには何も書き込まない。これは、DirectSound では、チャンネルブロックにデータが入っていない場合は、無音を出力するので、ユーザーに割り当てられた出力チャンネルからのみサウンドが出力するためである。これによって、出力したいユーザーに対応づけられたバッファを再生すれば、そのユーザーにのみサウンドを提示することができる。サウンドデータをバッファに書き込んでから、出力するまでのフローを図 4 に示す。図 4 は、ユーザー A とユーザー C にサウンドを出力すると仮定し、バッファ 1 とバッファ 3 のそれぞれのユーザーに割り当てられたチャンネルブロックにサウンドデータを書き込み、二つのバッファに割り当てられたサウンドがミキシングされて出力されるまでを表す。

また、これらのマルチチャンネルサウンドは、市販のマルチチャンネル対応のサウンドカードと、ヘッドホン、サラウンドスピーカ、指向性スピーカなどのサウンド提示デバイスを用いて、ユーザーごとに出力する。

4 複数人用マルチメディアコンテンツ環境の構築手法

立体映像アニメーションとマルチチャンネルサウンドを組み合わせた複数人用マルチメディアコ

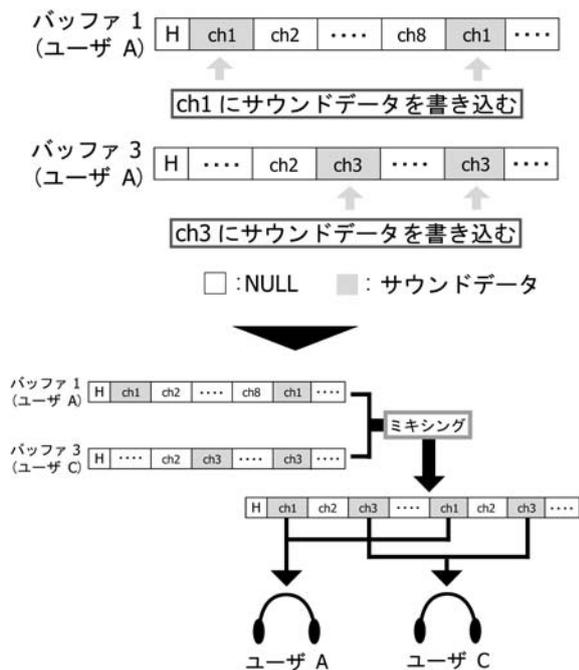


図 4: サウンドの書き込みと再生

コンテンツ環境を構築するためには、アニメーションとマルチチャンネルサウンドの編集を同一時間軸上で扱うことができるツールが必要である。そこで、アニメーションの開始や停止、サウンドの開始や停止等を同一時間軸上のイベントとして扱うイベントドリブン方式を採用し、これらのアニメーションとマルチチャンネルサウンドを組み合わせることができるオーサリングツールを開発する。イベントとしては、アニメーションの開始や停止、サウンドの開始や停止、サウンドのボリューム変化などが挙げられる。コンテンツを表示する側では、イベントの種類に応じて処理を行う。

複数人用マルチメディアコンテンツ環境では、ユーザの動きなどのインタラクションに応じて、立体映像アニメーションとマルチチャンネルサウンドが変化するコンテンツを表示する。オーサリングツールで作成するコンテンツとしては、次のようなものが挙げられる。

- (ア) ユーザの立ち位置に応じて立体映像アニメーションが変化するコンテンツ
- (イ) 立体映像アニメーションの向きがユーザに対して追従するコンテンツ
- (ウ) ユーザの立ち位置に応じて異なるサウンドを出力するコンテンツ
- (エ) ユーザごとに異なるサウンドを出力するコンテンツ

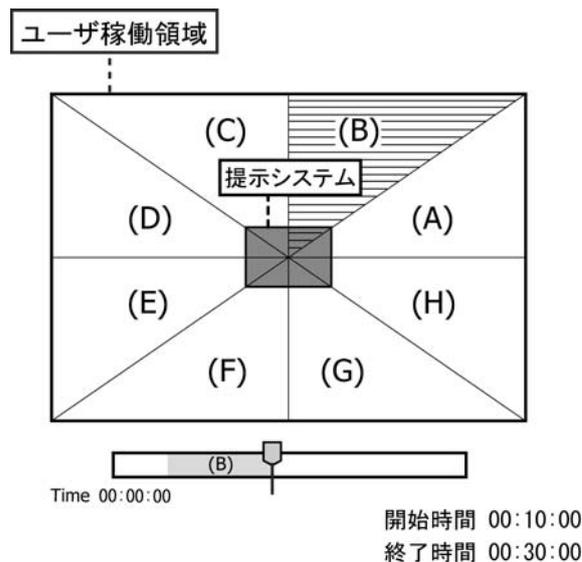


図 5: ユーザの立ち位置に応じてアニメーションが変化するコンテンツを編集する機能の例

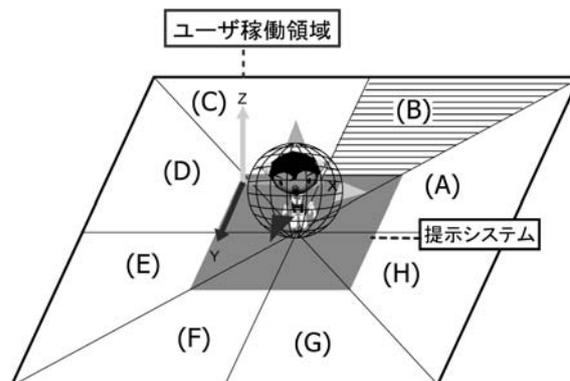


図 6: 立体映像の位置と回転を編集する機能の例

- (オ) ユーザごとに異なるサウンドを出力し、また立ち位置に応じて異なるサウンドを出力するコンテンツ

ここでは、これらのコンテンツを作成できるオーサリングツールに必要な機能について検討する。

(ア) の場合、ユーザがある領域内に入るとアニメーションが変化するといったインタラクションとリアクションを実現するため、図 5 に示すように、アニメーションが変化する領域を (A) ~ (H) に分割し、その領域ごとにアニメーションの種類とその開始時間、終了時間を決定する機能と、図 6 に示すように、表示するキャラクターの移動や回転方向を決定する機能が必要であると考えられる。

(イ) の場合、追従の対象となるユーザを選択し、アニメーションの開始時間と終了時間を決定する

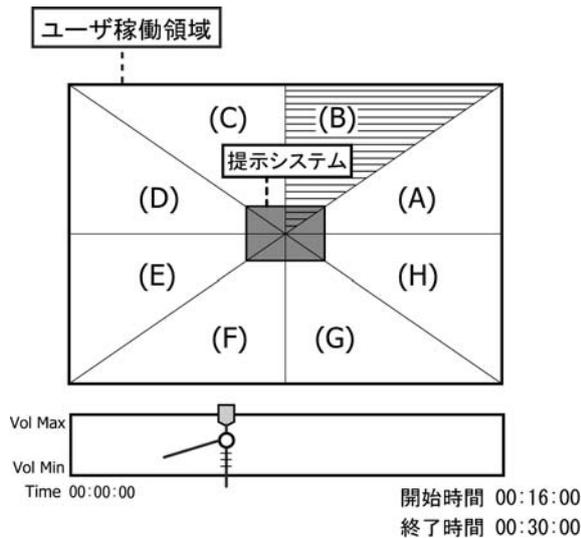


図 7: 場所によって異なるサウンドを出力するコンテンツを編集する機能の例

機能が必要であると考えられる。

(ウ) の場合、ユーザがある領域内に入るとサウンドが変化することを実現するため、図 7 に示すように、サウンドが変化する領域を指定し、その領域ごとに出力するサウンドの種類や開始時間、終了時間、サウンドのボリューム等を決定する機能が必要であると考えられる。

(エ) の場合、ユーザごとにサウンド編集領域を割り当て、ユーザごとに出力するサウンドの種類や開始時間、終了時間、サウンドのボリュームを決定する機能が必要であると考えられる。

(オ) の場合、(ウ) と (エ) のそれぞれについて検討した機能を組み合わせ、サウンドを提示する範囲をマウス等で指定し、各ユーザごとにその場所で出力するサウンドの種類や開始時間、終了時間、サウンドのボリュームを決定する機能が必要であると考えられる。

今回は、オーサリングツールの第一段階として、立体映像アニメーションとマルチチャンネルサウンドを同期させ、(エ) のようなコンテンツを編集することが可能なオーサリングツールを試作した。

5 実装例

5.1 複数人用マルチメディアコンテンツ環境の表示手法

システムの実装例として、立体映像を表示するデバイスとして円偏光を用いた多人数共有型立体表示装置 IllusionHole[16]、マルチチャンネルサウンドを提示するデバイスとして Bluetooth タイプの無線ヘッドホンをユーザの数だけ用いて、複数人

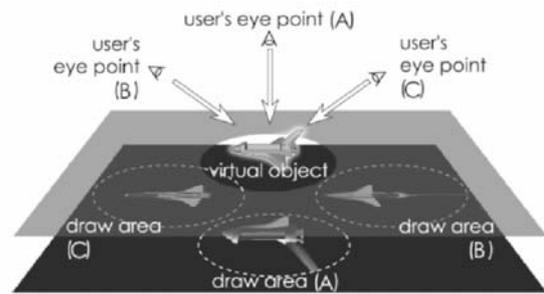


図 8: IllusionHole の原理

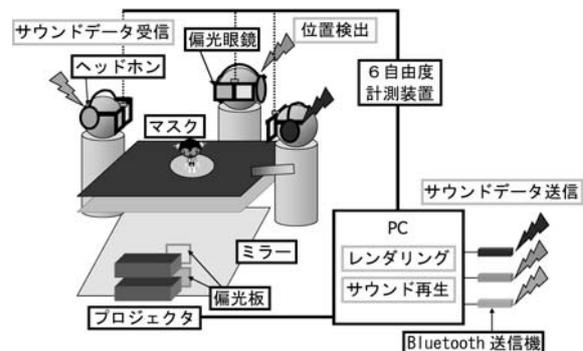


図 9: システム構成

用マルチメディアコンテンツ表示環境を構築した。

IllusionHole は、図 8 のように一般のディスプレイ装置と、中央に穴の開いたディスプレイマスクから構成される。ディスプレイマスクをディスプレイ面から適当な距離だけ離れた位置に設置することで、各ユーザが異なるディスプレイ上の画像表示領域を観察することを可能にする。この際、ユーザの視点位置を検出し、各ユーザの画像表示領域には両眼視差を用いた立体視用の画像を提示する。各ユーザはマスクホールを通して観察するため、自分の画像表示領域は観察できるが、他人の画像表示領域はマスクに隠されて観察することができない。これにより、単一のディスプレイで複数のユーザの視点移動に適切に対応した歪みのない立体映像を表示することができる。実装例で用いた IllusionHole は、立体視の実現手法として円偏光フィルタを用いたものを使用する。

システム構成を図 9 に示す。各ユーザは、視点位置を取得するための 3 次元トラッカと、立体映像を見るための偏光メガネを装着する。また、PC のサウンドカードの出力チャンネル部分に Bluetooth 送信機を接続し、各ユーザはそれぞれの出力チャンネルに対応した無線ヘッドホンを装着する。以

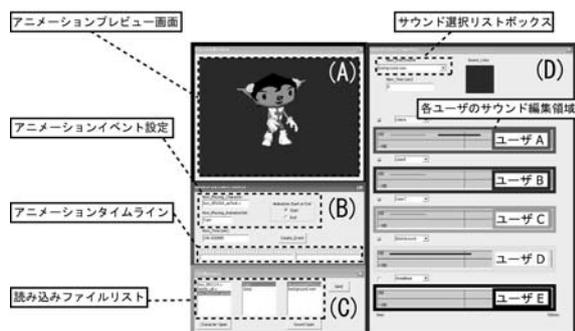


図 10: オーサリングツールの GUI

上で述べたシステム構成を用いることで、各ユーザに観察方位に応じた立体映像とユーザごとに異なるサウンドを表示し、同一のコンテンツに対して個人用にカスタマイズした情報も表示することが可能となった。

5.2 複数人用マルチメディアコンテンツ環境の構築

試作したオーサリングツールの外観を図 10 に示す。図 10 (A) で示す部分は、アニメーションのプレビューを表示する画面を表し、図 10 (B) で示す部分は、アニメーションの開始・停止を制御する画面を表している。図 10 (C) で示す部分は、読み込むアニメーションファイル、サウンドファイルを管理する画面を表し、図 10 (D) で示す部分は、サウンドを制御する画面を表している。ここでは、ユーザごとに個別のサウンドを出力可能なコンテンツを作成できるように、ユーザごとにサウンド編集領域を割り当てる。

次に、アニメーションのコンテンツを作成する手順を以下に示す。まず、図 10 (C) で読み込むアニメーションを選択し、図 10 (B) のスライダーを操作して、アニメーションの開始・終了の時間を決定し、イベントを作成する。作成したコンテンツは、スライダーを移動することにより、図 10 (A) のプレビュー画面で確認できる。

サウンドのコンテンツを作成する手順を以下に示す。まず、図 10 (D) に示すリストボックスからサウンドを選択する。そして、編集領域内でサウンドを再生する時間から、停止する時間までラインを引く。ここで、編集領域内で作成したサウンドの種類と描画するラインの色が、コンテンツ作成者に対して直感的にわかるように、サウンドの種類ごとにラインの色を割り当てる。また、編集領域内の高低はボリュームの大小と対応しており、編集領域内で作成したラインの傾斜が、ボリュームの変化を表している。ラインの長さは、サウン



図 11: 試作したコンテンツを体験している様子

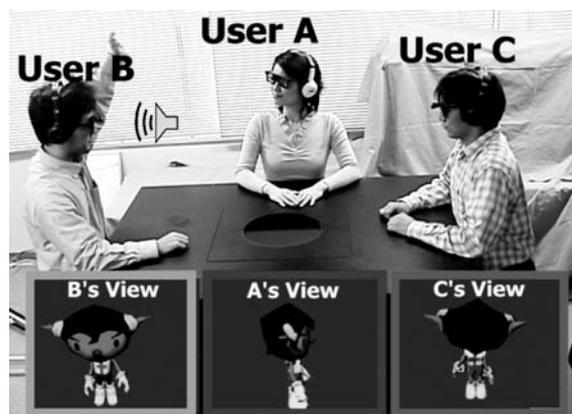


図 12: 試作したコンテンツの一場面

ドの再生時間を表しているため、コンテンツ作成者は、サウンドコンテンツを直感的に作成することができると考えられる。

以上の操作を繰り返すことで、アニメーションとサウンドを組み合わせた複数人用マルチメディアコンテンツを作成することができる。これをコンテンツデータとして保存し、立体映像とマルチチャンネルサウンドが表示できる環境においてコンテンツデータを読み込むことで、作成したコンテンツを表示できる。

5.3 マルチメディアコンテンツの表示

本章で構築した複数人用マルチメディアコンテンツ表示環境でコンテンツを表示した。3人がコンテンツを見ている様子を図 11 に示す。コンテンツとしては、IllusionHole 上に表示されたキャラクターが IllusionHole 自身の機能や特徴を説明するものであり、説明の際には、各ユーザごとに異なるサウンドが出力される。表示したコンテンツの一場面を図 12 に示す。ユーザ A, B, C のそれぞれの位置からは、キャラクターが図中の下の各領域の

ように見えている。この場面では、キャラクターはユーザ B と対面し、そのユーザに対して「手を挙げてください」という内容の音声を出力している。このとき、ユーザ B にのみ音声が出力され、他のユーザにはその音声が出力されないが、キャラクターがユーザ B に正対していることを容易に認識することができる。

6 おわりに

立体映像を用いたアニメーションとマルチチャンネルサウンドを組み合わせたオーサリングツールを開発し、複数人用マルチメディアコンテンツ表示環境を構築した。今後は、異なるライブラリを用いて時分割立体視対応のアニメーション表示や、ユーザのいる場所やその他のインタラクションによって異なるサウンドが聞こえるなどのマルチメディアコンテンツ表示環境の構築を予定している。また、本環境の特徴を生かしたいくつかのコンテンツを試作し、複数のユーザがそこでとる行動や心理状況の変化などについても観測し、検討したい。

謝辞 本研究の一部は、総務省 戦略的情報通信研究開発推進制度 (SCOPE)、及び文部科学省 21 世紀 COE プログラム (研究拠点形成費補助金) の助成を受けた。

参考文献

- [1] Paul Dietz, Darren Leigh: “DiamondTouch: a multi-user touch technology,” Proceedings of the 14th annual ACM Symposium on User Interface Software and Technology (UIST’01), pp. 219–226 (2001).
- [2] Chia Shen, Neal Lesh, Baback Moghaddam, Paul Beardsley, Ryan Scott Bardsley: “Personal digital historian: user interface design,” CHI ’01 extended abstracts on Human factors in computing systems, pp. 29–30 (2001).
- [3] ラグナロクオンライン
<http://www.ragnarokonline.jp/>
- [4] 原島博: 3次元画像と人間科学, オーム社 (2000).
- [5] Dolby Laboratories
<http://www.dolby.com/>
- [6] Carolina Crus-Neira, Daniel J. Sandin, and Thomas A. DeFanti: “Surround-screen projection-based virtual reality: the design and implementation of the CAVE,” Computer Graphics, Annual Conference Series (Proceedings of SIGGRAPH 1993), pp. 135–142 (1993).
- [7] 廣瀬通孝, 小木哲朗, 石綿昌平, 山田俊郎: “多面型全天周ディスプレイ (CABIN) の開発とその特性評価,” 電子情報通信学会論文誌, Vol. J81-D-II, No. 5, pp. 888–896 (1998).
- [8] M. Agrawala, A. C. Beers, B. Frohlich, and P. Hanrahan: “The two-user responsive workbench: support for collaboration through individual views of a shared space,” Computer Graphics, Annual Conference Series (Proceedings of SIGGRAPH 1997), pp. 327–332 (1997).
- [9] 澤田一哉: “多様化する高臨場感没入型視覚ディスプレイ,” 映像情報メディア学会誌, Vol. 53, No. 7, pp. 927–931 (1999).
- [10] 小木哲朗: “没入型ディスプレイの特性と応用の展開,” ヒューマンインタフェース学会論文誌, Vol. 1, No. 4, pp. 43–49 (1999).
- [11] Yoshifumi Kitamura, Takashige Konishi, Sumihiko Yamamoto, and Fumio Kishino: “Interactive stereoscopic display for three or more users,” Computer Graphics, Annual Conference Series (Proceedings of SIGGRAPH 2001), pp. 231–239 (2001).
- [12] Oliver Bimber, Bernd Frohlich, Dieter Schmalstieg, L. Miguel Encarnacao: “The virtual showcase,” IEEE Computer Graphics and Applications, Vol. 21, No. 6, pp. 48–55 (2001).
- [13] 三菱電機エンジニアリング株式会社
<http://www.mee.co.jp/pro/sales/kokodake/kokodake.html>
- [14] Macromedia Director MX 2004
<http://www.macromedia.com/jp/software/director/>
- [15] Adobe Premiere 1.5 Pro
<http://www.adobe.co.jp/products/premiere/main.html>
- [16] 北村喜文, 中山智量, 中島孝司, 山本 澄彦, 岸野文郎: “偏光フィルタを用いた多人数共有型立体表示装置の試作と検討,” 電子情報通信学会技術研究報告, Vol. 105, No. 608, pp. 113–118 (2006).