

ハイパーメディア環境による次世代ホームシアター

森山 剛^{*1} 内田 孝幸^{*2} 栗林 英範^{*1} 久米 祐一郎^{*3} 田村 徹^{*3}

Next-Generation Home Theater with Hypermedia

Tsuyoshi Moriyama^{*1} Takayuki Uchida^{*2} Hidenori Kuribayashi^{*1}
Yuichiro Kume^{*3} Tohru Tamura^{*3}

Abstract: We propose an audiovisual theater system that has multimodal display, including a 3-way screen that covers full range of user's view (Ambient view), a three dimensional display before the screen (3-D view), an information panel at user's hand (Handy view), 5.1ch surround stereo audio around the user, and a tactile display. The proposed system realizes interactivity by sensing user's physiological activity, motion of the fingers, gaze direction, facial expressions, and vocal signs. It also has hypermedia interface that can access multimedia resources over the Internet. By overlapping 3-D display with the ambient view, it is capable of providing truthfully realistic reproduction of the audio and visual content in the room. It would enable as many applications including interior customization, education, entertainment, gaming, simulation, and teleconference. By allowing the capture studio to be outside the room that are accessible through the network or to be artificial environment, users can watch both online and offline content seamlessly.

1. はじめに

音楽や映画がデジタルデータとして流通し、これを視聴する環境の整備が必要とされている。従来のホームシアターは、サラウンド・ステレオと大型のディスプレイによって臨場感溢れる視聴覚環境を提供したが、近年ではさらにユーザに反応して表示を変える対話性、情報の種類に適したいくつかの表示形態を有するマルチ・モダリティ（音響、3-D表示を含む）に対する要求が高まっている。さらに従来はDVDディスクを読み出して再生するスタンドアロンの方式が一般的であるが、ブロードバンド・ネットワークを通じてビデオ配信が頻繁に行われる近年では、インターネット網を視聴覚コンテンツの保存場所と捉えるハイパーメディア環境を前提にしたシステムが実現可能となった。

3Dのシーンや物体を再生表示する装置として、従来いくつかの没入型の視聴方式が提案されてきた。単に大型の2D画面で視界を覆うものから、偏光眼鏡を介して両眼視差を与えるもの、DFD (Depth

Fused 3-D)原理によるもの⁸⁾、高速回転する板や水蒸気に投影するもの、プラズマ発光を用いるもの⁹⁾、ホログラムを用いるものがある。これらはディスプレイ技術単体で現実感を目指すものが多い。しかし、上述のマルチ・モダリティや対話性、ハイパーメディアといった機能の全てを有するシステムは、未だ見られない。

本論文では、本学ハイパーメディア研究センターで構築しつつある視聴覚システムについて述べる。本システムは、マルチ・モダリティと対話性、ハイパーメディアの全ての機能を有し、次世代のホームシアターとして生活のあらゆる場面でのメディア視聴を変え得る技術である。本システムは、2次元映像と3次元映像を重畳表示し、さらにハンディな情報ディスプレイとサラウンド・ステレオ、触覚提示により、臨場感と迫力、親近感を演出する。視線方向や表情、音声、手指動作を捉えて対話性を実現することで、ユーザの嗜好をより良く反映したり、ユーザに一体感を感じさせたりすることができる。さらに、ネットワークで接続された世界中のハイパ

^{*1} 東京工芸大学工学部メディア画像学科助教 ^{*2} 東京工芸大学工学部メディア画像学科准教授

^{*3} 東京工芸大学工学部メディア画像学科教授 E-mail:moriyama@mega.t-kougei.ac.jp

キャプチャ・スタジオ anywhere & anytime



ビュー・スタジオ right now in this very room

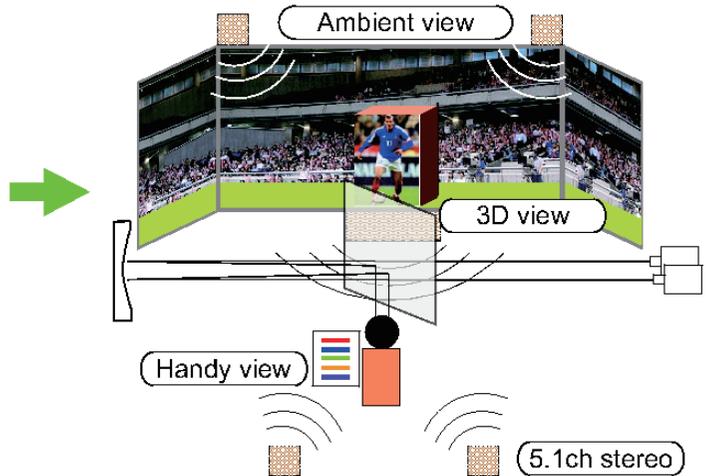


図 1 対話性を備えたマルチ・モーダルな次世代ホームシアター

メディア・コンテンツを部屋の中に再現することができる。

2. 提案する次世代ホームシアターシステム

提案する次世代ホームシアターシステムの概要を図1に示す。本システムは、視聴を行う「ビュー・スタジオ」と視聴コンテンツを取得する「キャプチャ・スタジオ」の2つから成る。ビュー・スタジオは、シーンの背景を表示するためにスクリーンを三面鏡のように配置した「Ambient view」(スクリーンごとにプロジェクタで投影)、2台のプロジェクタに両眼視差を与えてユーザに立体視させる「3-D view」、そして手元で詳細情報を表示する「Handy view」の3つのビューから成り、さらに5.1チャンネルのサラウンド・テレオによる音響ディスプレイ、触覚提示装置を備えている。また、ビデオカメラとマイクロフォン、触覚センサを備え、ユーザの視線方向や表情の変化、声、手指動作を捉え、そのパターンによってコンテンツの表示を制御することができる。キャプチャ・スタジオは、特定の物理的なスタジオを指すのではなく、本システム表示用に視聴コンテンツを記録する環境を総称する。視聴コンテンツを取得するプロセスと、それを表示するプロセスとを分けることによって、コンテンツへの

オンライン及びオフラインでのアクセスを可能にする。

3. 本システムの有する機能

本システムは、触覚/生体センサやカメラ、マイクロフォンを備えることで対話性の機能を実現し、前述の3つのビュー及びステレオ音響、触覚デバイスによりマルチ・モダリティ機能、さらにネットワークに接続することによりハイパーメディア機能を実現する。

3.1 対話性の機能

従来のホームシアターと最も異なる点は、本システムはユーザと対話的にその表示コンテンツを変化させる点である。

3.1.1 触覚による対話性

触覚提示において、皮膚上の複数の異なる点で触覚を刺激する場合、融合して知覚された位置が刺激強度の大きい位置に引きずられる現象をファントム・センセーション (以下 PhS) と呼ぶ¹⁾²⁾。PhS による位置知覚の様子を図2に示す。PhS を用いれば方向を含む空間情報の触覚提示が可能となり、映像情報と共に提示することでより現実感のあるメディア視聴 (例: レーシングゲームで、自分のコント

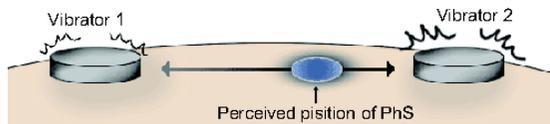


図 2 ファントムセンセーションの知覚位置

ロールする車体が映像中で衝撃を受ける場合、衝撃の方向を PhS で触覚提示する) が可能となる。また触覚センサを用いて手指による圧力と摩擦力を捉え、表示シーンの空間内を移動したり、表示シーンに対する視点を切り替えたりするインタフェースを実現する。

3.1.2 視線追跡による対話性

カメラで取得する画像列を処理して、顔及び顔部品の検出、頭部運動の追跡、さらに顔表情筋運動の解析を行い、顔表情と共に視線方向を推定する³⁾⁴⁾。視線方向追跡結果を用いて、Ambient view 上のどの点を注視しているかを求めれば、それをユーザが興味を持つ位置と解釈し、その点付近に表示されている映像オブジェクトに関する情報を Handy view に表示したり、その点付近に存在する音源を強調して擬似カクテルパーティ効果を演出したりできる。

3.1.3 生理反応による対話性

触覚センサに加えて、ユーザの発汗、脈波、呼吸の乱れを観測する。あらかじめ学習する規則に基づいてユーザの心理状態を解釈し、その結果に応じて、表示する視聴覚コンテンツを変化させることができる。例えばゲームプレイヤーの興奮を捉えて、ゲームキャラクターの表情をリアルタイムに変化させることが可能になる。

3.2 マルチ・モダリティの機能

本システムでは、Ambient view で環境（遠景）、3-D view で環境中の特定オブジェクトや領域（中景）、Handy view でユーザの興味の対象に関する詳細情報（近景）、5.1 チャンネル・サラウンド・ステレオ音響、そして触覚の5つのモダリティを有する。

Ambient view では、三面鏡のように配置した視界全体を覆うスクリーンに、マルチ・プロジェクタでつなぎ目を意識させない滑らかな像を表示する。

3-D view ではコンテンツの内容やユーザの注視

点に応じた映像オブジェクト（例：山林シーンにおける鳥や木々）を表示する。両眼視差を与えたプロジェクタ2台の映像を、凹面鏡を経由してユーザの左右の眼に入射し、特殊な眼鏡を装着させることなく立体視させる⁵⁾。ハーフミラーをユーザと Ambient view の間に配置することで、Ambient view と 3-D view に表示されるコンテンツが重畳され、一体となって見える。

Handy view では有機エレクトロルミネッセンス (OLED) を用いて、映像オブジェクトの詳細な情報を表示する。有機 EL 素子は有機薄膜に直流電圧を印加することで自発光する特徴をもつため、薄型、軽量であり視野角依存性がない。有機膜が約 $0.2\mu\text{m}$ と薄いため、可視領域にほとんど吸収がなく、有機膜の上部、下部ともに透明な電極を設けることで、非発光時に透明な発光素子、透明有機 EL 素子 (TOLED: Transparent Organic Light Emitting Device) を作製することができる⁶⁾。さらに、赤、緑、青の透明有機 EL 素子を積層した素子によりカラー表示も行うこともできる⁷⁾。

3.3 ハイパーメディア機能

ハイパーメディアとは、ハイパーテキストの概念を音や映像といったメディア全般に拡張した概念で、ハイパーリンクを介して、ネットワークに接続されたあらゆるメディアにアクセスできる技術である。本システムでは、視聴覚コンテンツを再生するコンピュータをネットワークに接続することによって、ブロードバンド・ネットワーク上に、アクセス可能な状態で存在する視聴覚コンテンツを表示することができる。

4. 提案システムの開発

提案するシステムは、現在 (2007年9月現在) 建設中であり、これまで3章で述べた要素技術それぞれの開発してきた。これらを組み込んだビュー・スタジオは、2008年春に完成及び公開の予定である。現在建設中のビュー・スタジオは、Ambient view 用に、縦 $1,540\text{mm}$ 、横 $2,050\text{mm}$ の正面スクリーン、横幅をその半分にした側面スクリーン (左右)、3-D view 用に 700mm 四方の凹面鏡をそれぞれ配置し、Ambient view 用プロジェクタ3台、3-D view 用プロ



図 3 2次元 PhS の提示装置

ジェクタ 2 台をそれぞれ 1 台の PC に接続したものである。うち 1 台の PC は音響用ミキサを通して、5.1 チャンネルのスピーカに接続している。以下、3 章で述べた要素技術の実験結果について述べる。

4.1 PhS による方向提示装置の実装

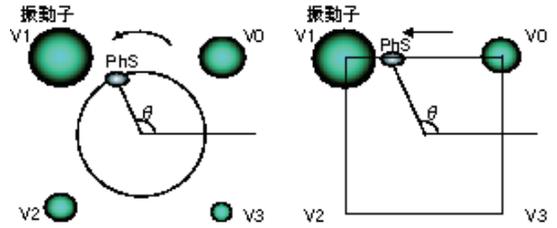
図 3 に示すように、市販のゲームの背面に 4 個の振動子を取り付け（両手の人差指と小指で触れる）、両手の掌内に 2 次元 PhS を生起させて、PhS の偏移によって方向を提示する装置を実装した。PhS の知覚位置を制御する際、図 4(a)に示すように、4 個の振動子の振動強度和を一定に保つように振動強度を変えることで方向提示できたが、知覚される感覚が広がってしまった。そこで図 4(b)に示すように、同時に振動させる振動子の数を減らして、隣接する振動子間で直線的に PhS を生起させて方向提示することで、PhS の感覚を局在化し方向定位能を向上させることに成功した。また振動強度和を変化させて方向定位能を測定したところ、振動強度を小さくすると方向定位能がやや低下しても、おおよその方向は知覚させられることがわかった。

4.2 透明フレキシブル両面発光有機 EL デバイスの作製

図 5 に作製した透明フレキシブル両面発光有機 EL デバイスを示す。薄膜素子の利点を活かし、プラスチックフィルムに作製することで、折り曲げ可能な素子を作製した⁶⁾。

4.3 高精細目領域モデルを用いた視線追跡

目領域の構造及び運動による見えの多様性をモデル化したモデルを用い、視線方向を追跡する手法を実装した³⁾。図 6 に頭部運動の追跡及び視線の追跡の結果例を示す。本手法では、眉毛と目の両端点のみを手入力で与え、それらの点を基に上まぶた小領域、下まぶた小領域、目の内部領域の 3 つの領域



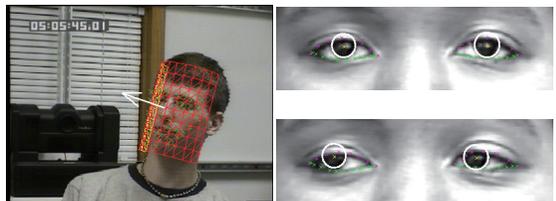
(a) 4 振動子を同時制御し (b) 2 振動子を同時制御して円形状に提示して正方形形状に提示

図 4 同時に振動させる振動子の数を変えた場合の PhS による方向提示



(a) 非発光時 (b) 発光時

図 5 透明フレキシブル両面発光有機 EL 素子 (フィルム厚み 125 μm, 有機膜厚み約 0.2 μm)



(a) 頭部運動追跡 (b) 瞳追跡(上は第 1 フレーム, 下は第 12 フレーム)

図 6 視線追跡の結果例

を決める。次にあらかじめ学習サンプルを用いて生成しておいた部分空間にそれぞれの領域画像を投影し、部分空間内で最も近い学習サンプルを各々求める。学習サンプルには目の構造(例:二重の厚さ)の情報が埋め込まれており、これを用いて目領域モデルを個人化する。個人化した目領域モデルの運動パラメータをフレームごとに探索することで、目の運動を追跡する。実験の結果、本目領域モデルを用いた目の運動追跡は、FACS (Facial Action Coding System) に挙げられている目の全ての運動について正しく追跡が行え、また、モデルの自動個人化を行

った場合でも、手動個人化を行った場合と同等の精度で目の運動追跡を行えることを示した。

5. 本システムの応用例

本システムを用いることで実現可能な、これまでは実現できなかった新しいメディア視聴の具体例を示す。

5.1 スポーツ観戦への応用

スポーツ観戦の例として、サッカーの試合を本システムで視聴する場合を考える。フィールド向こう側の観客席が Ambient view いっぱいに映し出され、その手前のハーフミラー上にボールを競り合う選手近傍の領域が 3-D 表示され、Handy view には選手の情報の他、フィールドを真上から俯瞰した長方形の中に、チーム別に色分けられた選手の点が動く様子が表示される。さらにスタジアム中に響く観客の歓声が、5.1 チャンネル・サラウンド・ステレオからユーザを囲むように再生される。

複数台のカメラを切り替える仕組みと、指のなぞる動作を触覚センサで位置の移動量に変換する仕組みとを組み合わせることによって、スタジアムを囲むように配置したカメラを指で高速に切り替え、選手がボールを競り合う様子の良く見える視点を選択したりすることができる。

5.2 自然観賞とインテリアへの応用

森林の映像を Ambient view に、手前の 3-D view に森の木々や鳥、苔むす岩場やせせらぎを表示し、実際に森林の中で集音した音をサラウンド・ステレオで再生する。森林以外にも、美しい海の広がるビーチや、名所旧跡等観光地、水中や宇宙空間、さらには人工的に作り出された仮想空間、ミクロの世界を再現することにより、室内の景観を入れ替えることができる。

自宅のリビング、画廊やオフィスの廊下に設置することで室内環境を好みのものに変え、特に寝たきりで動けない人が使用すれば、気分転換や癒しの効果が期待できる。

5.3 ゲーム及びシミュレータへの応用

本システムをドライビング・シミュレータとして

用いる場合を考える。ユーザは Ambient view に表示される前方の道路状況を見ながら車両を操舵する。そこへ看板や歩行者など路上に現れるオブジェクトが 3-D view に表示され、路面に応じた走行音や対向、後続車両の音がサラウンド・ステレオで再生される。長時間の運転で眠気が生ずると瞬き回数が減少することを検出し、ドライバを覚醒するための仕組みを実験することもできる。

ロール・プレイング・ゲームの場合では、ユーザがプレイ中にはらはらすると、生体センサが脈波や呼気の乱れを検出して、キャラクタ CG の顔が額に汗をかかせたり、ユーザが疲れてくるとキャラクタを眠そうにさせたりする仕組みを実現することができる。

5.4 教育教材への応用

本システムを、動植物の仕組みや動きをつぶさに見せる環境として用いる場合を考える。素粒子や天体といったコンテンツは、3-D view に表示することによって立体的な把握が容易で、かつ興味を引き立てると考えられる。また対象についての詳細な情報を Handy view 上に常に参照することができる。

また遠隔地で行われている講義を受講する場合、ハイパーメディアを生かして、講義の内容に関してインターネットから取得可能な情報(イラストやCG、音を含む)と合わせてビュー・スタジオに提示させることができる。講師にとっても、受講生の様子を取得するカメラやマイクロフォンの情報と、それらから推定される講義への関心度や疑問のサインを合わせて提示させることにより、より効果的な講義を遠隔地から実現できるようになる。

6. むすび

ユーザの視界を完全に覆うように配置される大型スクリーン (Ambient view)、その手前に配置される立体ディスプレイ (3-D view)、そしてユーザの手元に配置される情報ディスプレイ (Handy view)、これに 5.1 チャンネル・サラウンド・ステレオによる音響ディスプレイと触覚提示を加えた没入型視聴環境を提案した。本システムは、ネットワーク越しにマルチメディア・コンテンツを視聴できるハイパーメディア機能、ユーザのアクションをコンテン

ツの提示に反映させる対話性功能, さらに 3 種の view と立体音響, 触覚という複数のチャンネルでの提示を行うマルチ・モダリティ機能の全てを実現する。これまで室内で映画の再生を行っていたホームシアターに代わって, 世界中の情報やマルチメディア・コンテンツにつながる窓として, 外界のメディア世界への玄関「ホーム・メディア・ポータル」として, これからのメディア視聴を変え得る技術である。また本システムは, パターン認識や感性, マルチ・モダリティを対象とする多くの学際的な研究課題を実現するプラットフォームとしても有用である。本システムは 2008 年春に公開の予定である。

謝辞

本研究の一部は, 文部科学省ハイテク・リサーチ・センター整備事業(平成 17 年度~平成 21 年度), ならびに平成 19 年度双葉電子科学財団の助成により遂行した。

参考文献

- 1) 久米祐一郎, 中島勇一, “2 次元ファントムセンセーションによる方向提示と電子ゲームへの応用”, 映像情報メディア学会技術報告, Vol.29, pp.21-24, 2005.
- 2) 久米祐一郎, 礪部正利, “多次元ファントムセンセーションの基礎特性”, 日本バーチャルリアリティ学会第 5 回大会論文集, pp.241-242, 2000.
- 3) Tsuyoshi Moriyama, Takeo Kanade, Jing Xiao, Jeffrey F. Cohn, “Passive Eye Monitoring”, chapter 2, pp.17-40, Springer, March 2008.
- 4) 森山剛, 金出武雄, Jeffrey F. Cohn, 小沢慎治, “Active appearance model 探索による目領域構造の自動抽出”, 電子情報通信学会 画像の認識・理解シンポジウム(MIRU2006), pp.830-835, Vol.7, 2006.
- 5) 畑田豊彦, 生理光学 15, No.73, O plus E, pp.98-109, 1985.
- 6) T. Uchida, T. Mimura, S. Kaneta, M. Ichihara, M. Ohtsuka, and T. Otomo, “Transparent organic light emitting devices fabricated by the Cs incorporated RF magnetron sputtering deposition”, Japanese Journal of Applied Physics, Vol.44, No.8, pp.5939-5942, 2005.
- 7) 市原正浩, 内田孝幸, 金田真吾, 田村徹, 大塚正男, 大友俊夫, “RGB 各透明有機 EL を積層させたフルカラー表示画素の初期的検討”, 日本写真学会誌, Vol.69, No.1, pp.55-58, 2006.
- 8) 伊達宗和, “DFD 方式”, 電気学会誌, Vol. 127, No. 9, pp.594-596, 2007.
- 9) Hidei Kimura, Taro Uchiyama, Hiroyuki Yoshikawa, “Laser produced 3D display in the air”, ACM SIGGRAPH 2006 Emerging technologies, pp.20, 2006.