

知的環境メディア

藤代一成 (正会員)

慶應義塾大学

Smart Ambient Media

Issei FUJISHIRO (Member)

Keio University

1. はじめに

画像コンテンツの高度化を目指した、コンピュータグラフィックスとビジョンの融合、すなわちビジュアルコンピューティング(Visual Computing, VC)²⁾の重要性が叫ばれるようになって久しい。本学会のVC委員会設立の狙いもまさにここにある。今日、その融合を旧来になく加速させているのは、コモディティGPUによる視覚・汎用計算能力の増強に加え、低価格かつ高性能な商用のセンサ類やVR/AR向けHMD装置等の入出力装置の普及にあると言っても過言ではない。

本稿では、それらの下支えによって実現されてきた新たな技術である知的環境メディアについて解説する。次章でその一般的な枠組みを示した後、3章で関連研究開発事例を6点紹介し、関連する諸概念にも言及する。さらに4章では、これらの事例を特徴づけ、知的環境メディアに共通する特徴とその可能性を明らかにする。最後に5章で本稿をまとめる。

本稿は、筆者が本学会会長就任時に寄稿した本誌巻頭言³⁾とそれに続く研究会招待講演⁴⁾が基になっている。

2. 知的環境メディアの枠組み

知的環境メディア(smart ambient media, SAM)とは、マルチモーダルセンシング—知的計算—適応的レンダリングを利用して、我々の日常生活の諸活動を巧みに支援する表示手法を総称したものである。この3フェーズは、時々刻々変化する実環境に適応するため、人間を中核に据えたループ：Human-in-the-Loopを構成する(図1)。知的計算には、流行の統計的機械学習も含まれる。情報通信分野では、センサネットワークを活用し、環境への働きかけを、ロボット等を含むアクチュエーションにまで拡充したコグニションサイクルによって予測・推奨情報を提供する、知的環境システム(intelligent ambient system)という汎化概念⁵⁾も知られている。



図1 知的環境メディア(SAM)の枠組み

Fig.1 Framework of Smart Ambient Media (SAM)

3. 知的環境メディアの事例

本章では、前章で定義した知的環境メディアの具体像を、筆者の研究室で近年研究開発した特徴的な事例を6点ピックアップして説明する。

3.1 錯視による裸眼立体視

図2は、L字型に直交配置された狭額ラスタディスプレイ2枚で実現された裸眼立体視像の例である。直角に開いたスケッチブックに描かれた物体が、あるアングルから鑑賞する際にポップアップする永井秀幸氏の作品群に触発され、スケッチブックを2枚のディスプレイモニターに取り替え、さらに汎用ウェブカメラを利用して観察者の顔追跡を実行し、描画立体の変形投影とスポットライトによる陰影付けをリアルタイムに調整することで視点拘束を解き、手軽な個人向け裸眼立体視環境を構築した⁶⁾。

この視覚トリックを成り立たせているのは、2枚のディスプレイを、予め観察者に各々「壁」と「床」に見立てさせる心理的前提知識に加え、レオナルド・ダ・ヴィンチの時代(15世紀)から利用されてきた、アナモルフォーシス(anamorphosis)と呼ばれる投影デザインによる「よい連続」

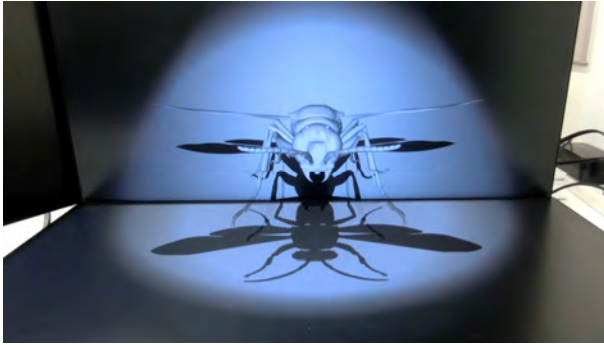


図2 直交配置ディスプレイを用いた裸眼立体視⁶⁾

Fig. 2 Naked-eye stereoscopic imaging using orthogonally arranged display monitors⁶⁾

(good continuation)のゲシュタルト効果、運動視差に基づく奥行き知覚効果の3点である。標準的なラスタディスプレイを組み合わせるだけの単純な仕掛けでも、裸眼立体視がさらに追究できることを本事例は示唆している。

立体視による没入感(illusion of immersion)は、複雑な対象の継続的な視覚分析を支援し、新たな知識獲得につながる確率を高めることが知られている。現在、本システムを利用した没入的視覚分析論(immersive analytics)への発展⁷⁾を検討している。

3.2 協同可視化のハイブリッド表示方式

図3は、タイル状ディスプレイとシースルーHMDを組み合わせた対話型可視化プロトタイプシステム⁸⁾を用いて、大規模3次元流動データの可視化を実行している様子である。前節のような個人用途と異なり、大規模化する数値シミュレーションや計測データを、プロジェクトに関わる多人数が同時に視覚的探索できる協同可視化(collaborative visualization)には、タイル状ディスプレイ環境の利用が現状技術による現実解の一つである。しかし効果的な可視化を実現するには、可視化制御パラメタのチューニングが必須である。そのために、高解像度のタイル状ディスプレイ全域に対して試行錯誤を繰り返すのは余りに非効率的である。そこで、代表ユーザがシースルーHMD(Metal)を用いて、典型的な部分ビューを選んで予備的な視覚解析を実行し、より適切な可視化制御パラメタ値を推定した後に、その視覚効果を全体に波及させるハイブリッド方式を考案したのである。

20年前のHMDの流行が一時的なものに終わってしまった主要因の一つは、効果的なコンテンツ生成を支援する当時の計算資源に限界が存在したからである。技術が進んだ現在でも、解析・表示する自然界の対象の複雑度に際限はないため、まさに同じ状況がタイル状ディスプレイに対しても生じている。この研究事例には、「同じ轍は踏まない」精神に則り、解像度や精細度の点で格段に向上したHMDを併用することで、窮状を効果的に回避しようとする強かな戦略が窺える。



図3 タイル状ディスプレイとシースルーHMDを組み合わせた協同可視化⁸⁾

Fig. 3 Collaborative visualization using a tiled display and see-through HMDs⁸⁾

3.3 自由度の高いパペトリ

現在でも、3D キャラクタアニメーション制作において、モーションキャプチャへの依存度は依然として高い。しかし、演者と自由にデザインされたキャラクタの骨格は必ずしも一致していないため、キャプチャされたモーションをキャラクタの骨格に合わせる変換処理：モーションリターゲットング(motion retargeting)が必要になる。近年、Microsoft Kinectをはじめとしたリアルタイムモーションキャプチャデバイスが登場し、インタラクティブなCGコンテンツの需要が高まっている。このキャラクタ(アバタ)の操作にもモーションリターゲットングが用いられているが、特にインタラクティブに様々なアバタを操作することを強調したモーションリターゲットングは、パペトリ(motion puppetry)と呼ばれている。しかし、操作できるアバタはヒトもしくはヒトにほぼ類似する骨格をもつものが殆どであった。近年提案された統計的機械学習をベースとした幾つかの手法では、ヒト以外の骨格のアバタへのリターゲットングを扱ってはいるものの、インタラクティブ性に欠けていたり、学習用のデータ以外に必要なメタデータが膨大であったりといった課題を抱えていた。

本研究では、既存手法に対してユーザとアバタの骨格の類似性を考慮することで、学習用のアニメーションデータと簡単な入力だけで、インタラクティブかつユーザの直観に沿うようなパペトリを可能にしている。さらに、アバタモデルの部位を手続き的に分類することで、ユーザの指定によるアバタの任意の部位へのパペトリを実現した⁹⁾。図4に示す実行例では、演者の足踏み動作が羊の歩行に転化されると同時に、演者の挙手によって羊の耳を交互に動かすような動作の対応付けが自然に実現されていることがわかる。

本研究が提案する柔軟なパペトリは、今後エンタテインメントゲームやVRの分野等で利用されることが期待されている。

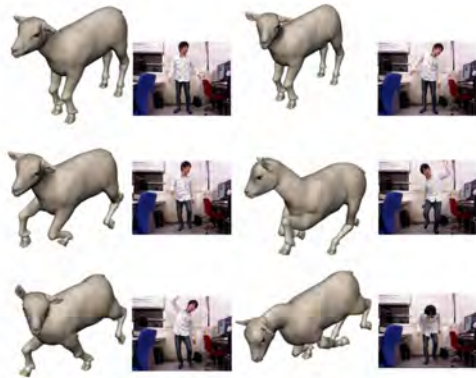


図4 骨格の類似性を考慮したパベトリ⁹⁾

Fig. 4 Puppetry taking skeletal similarity into account⁹⁾

3.4 流体の直感的形状制御

近年、流体シミュレーションはさまざまなCG映像作品に用いられている。砕波のように、流体を写實的に表現するだけでなく、映像の内容や演出に合わせて非写實的に挙動する流体がとり扱われることも多い。例えば、2016年に公開されたウォルト・ディズニー映画の『モアナと伝説の海』に登場する主人公モアナが幼少期に戯れる波のような流体を実現するためには、流体らしさを維持する一方で、擬人的な挙動も制御しなければならない。写実性とともにもこのような可制御性も加味した流体形状制御には数多くのパラメータが存在するが、制作者に流体シミュレーションの知識がなければ、安定にかつ正しく操作することは容易なタスクではない。

本研究では、流体シミュレーションおよび流体制御に用いるパラメータ調整を特定のハンドジェスチャに対応させた直接操作により、直感的な流体形状制御を試みている。実際、開発システムは、Leap Motion によって両手のジェスチャを取得し、左右のハンドジェスチャの組合せから特定のコマンドを認識し、その指示内容に応じて流体形状を制御しながら流体の粒子(SPH: Smoothed Particle Hydrodynamics)シミュレーションを実行する処理を反復する¹⁰⁾。

図5は、最新の3D流体形状制御システム¹¹⁾の操作例を示している。画面に投影されているような両手のジェスチャで流体の一部を選択し、筒状に引き上げている様子がわかる。

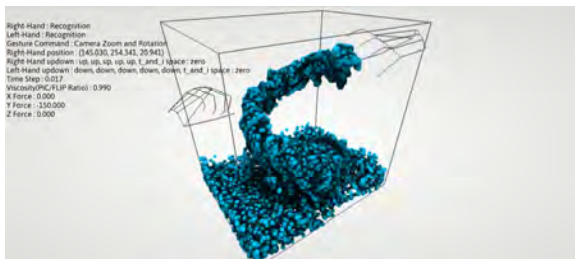


図5 3次元流体の形状制御¹¹⁾

Fig.5 Shape control of 3D fluid¹¹⁾

3.5 注視する画像特徴に応じたオーディオ再生

現在、デジタルサイネージやARなど、さまざまな日常生活レベルのデバイスインタフェースとして視線追跡を利用する手法が数多く実用化されている。非接触かつその存在を利用者に気付かせない「隠的」視線追跡により、臨場感を損ねず、利用者が作品に集中しやすい環境を作り出している。このような視覚インタフェースは、オーディオコンテンツの制御にも利用できる可能性があると考えられる。

本研究では、図6に示すように、汎用的なウェブカメラ1台から得られたユーザの注視位置に従って、再生するオーディオコンテンツを選択可能にする手法を提案している^{12),13)}。例えば美術館で絵画鑑賞する際のオーディオガイドを想定してみよう。通常は、絵画ごとに決まった解説ナレーションやBGMが流れるが、同一の絵画でも制作のエピソードやモチーフは部分ごとに異なるはずである。そこで、絵画を鑑賞する人たちの視線を追跡し、その先の局所の特徴に合わせて解説の内容やBGMをきめ細かく切り替える仕組みを実現した。

このように粒度の細かいオーディオ制御は、次世代の標準になると考えられる。

3.6 ビジュアルとの共演

AMP (Ambient Music co-Player) システムは、即興演奏を行う電子ギターのスロ奏者の演奏情報をリアルタイムに読み取り、それに呼応した映像生成を行うシステムである^{14),15)}。AMPシステムは、生演奏をアフェクティブな基準に従って評価し、その結果に応じた映像を生成し、奏者にインスピレーションを与える。インスピレーションを得た奏者は即興的に自身の演奏を変化させ、AMPシステムはまたそれに呼応して映像を変化させていくといったアフェクティブループ(affective loop)を形成することで、呼称のとおり奏者の「共演者」としての役割の実現を目指している。

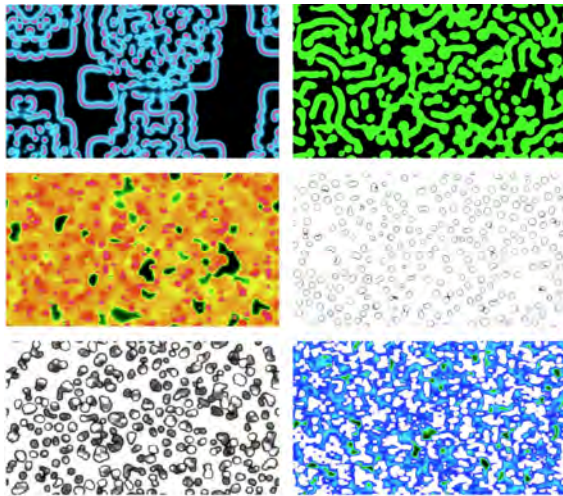


図6 視線追跡によるオーディオインタフェース^{12),13)}

Fig.6 Eye tracking-based audio interface^{12),13)}



(a) システムと共演する奏者



(b) 生成されたテクスチャ画像群

図7 Ambient Music co-Player (AMP)の実行情例^{13), 14)}

Fig.7 Execution example of Ambient Music co-Player (AMP)^{13), 14)}

図7に実行例を示す。図7(a)は共演の様子である。図7(b)は、実際の演奏によりAMPシステムが生成した反応-拡散系テクスチャ画像の例である。演奏の音楽パラメタから分類された感情カテゴリにシフトし、適切な拡散-反応系パラメタが割り振られて自動的にテクスチャ生成が行われている。重要なことは、演奏の同一興奮度が暫く継続すると、それを変化させる方向にAMPシステム自身が誘導するようなテクスチャを生成する点である。演奏のイニシアティブを互いにとりあうような双方向ビジュアルコミュニケーションが成立している。

4. 知的環境メディアの特徴

表1は、前章で紹介した知的環境メディアの研究開発事例6点を、その実現目的と、マルチモーダル入力系、知的計算の計3つの観点から特徴づけている。上から順に2件ずつ、表示高度化、アニメーション制御、視聴覚融合という目的に分類できる。また、個々に目的によって必要となる入力系は相異なるが、そこから得た情報から利用者の意図や感情を同定し、適応的な表示につなげていくループ処理には共通点を見出すことができる。

表1 3章の知的環境メディア事例の特徴分析

Table 1 Feature analysis of SAM examples in Chapter 3

節	実現目的	入力系	知的計算特徴
1	簡易裸眼立体視	ウェブカメラ	錯視
2	ハイブリッド表示	RGB-Dカメラ	座標系同期通信
3	柔軟なパベトリ	KINECT2	動作分類の機械学習
4	直観的流動形状制御	Leap Motion	ジェスチャ認識
5	視覚-聴覚融合	RGBカメラ	注視対象特定
6	双方向視覚情報通信	Waveデータ分析	アフェクティブループ

5. おわりに

画像・映像情報の効率的かつ効果的な伝送と蓄積に関する技術の追求と標準化は、ビジュアル情報処理の中核的ゴールの一つである。しかし今日の画像・映像情報のライフサイクルを鑑みれば、視覚データの生成から消費まで処理のパイプラインを拡大し、各フェーズ間の相互関係に目を向け、それを上手に利用していく必要がある。特に、消費の問題を本格的に取り扱うには、画像・映像情報の消費者である人間の知覚・認知の仕組みをよく知らなければならない。知覚心理学や脳科学に関する最新の知見を積極的に援用すれば、生成や伝送・蓄積の前フェーズに対しても大きな効果が与えられる。本稿で論じた知的環境メディアは、この拡大パイプラインをループ化し、リアルタイム処理化する試みに他ならない。

謝辞

本稿に示した事例は部分的に、科研費新学術領域計画研究25120014、基盤研究(A)26240015, 17H00737, 17H00738、挑戦的萌芽研究15K12034の支援による成果である。

参考文献

- 1) 藤代 一成: “コンピュータグラフィックス”, 電子情報通信学会100年史, 第2部, D-5-2, pp. 374-377 (2017).
- 2) 藤代 一成, 國井 利泰: “ビジュアルコンピューティングとは?”, カラーエイジ, No. 8, pp. 14-17 (1992).
- 3) 藤代 一成: “縦の絆・横の鎖” (随想), 画像電子学会誌, Vol. 45, No. 3, p. 259 (2016).
- 4) 藤代 一成: “ビジュアル情報処理のパイプラインと研究スパイラル再考” 画像電子学会第278回研究会招待講演, 映像情報メディア学会技術報告, AIT2016-139, Vol. 40, No. 7, pp. 9-10 (2016).
- 5) 大槻 知明, 渡辺 尚, 中澤 仁, 川島 英之, 滑川 徹, 山本 高至: “知的環境とセンサネットワーク —アンビエントセンサネットワーク—”, 電子情報通信学会誌, Vol. 96, No. 7, pp. 495-500 (2013).
- 6) 井阪 建, 藤代 一成: “L字型表示面を用いた錯視による裸眼立体

映像生成” (動画付き研究速報), 映像情報メディア学会誌, Vol. 70, No. 6, pp. J143-J146 (2016).

- 7) I. Fujishiro, N. Sawada, M. Nakayama, H.-Y. Wu, K. Watanabe, S. Takahashi, M. Uemura: “TimeTubes: Visual Exploration of Observed Blazar Datasets”, to appear in Journal of Physics: Conference Series: Proceedings of International Meeting on High-Dimensional Data-Driven Science (HD³-2017) (2018).
- 8) K. Nagao, Y. Ye, C. Wang, I. Fujishiro, K.-L. Ma: “Enabling Interactive Scientific Data Visualization and Analysis with See-Through HMDs and a Large Tiled Display”, Proc. of VR2016 Workshop on Immersive Analytics, IEEE Xplore (2016).
- 9) N. Kaneko, R. Takahashi, I. Fujishiro: “Motion Puppetry Taking Skeletal Similarity into Account”, Proc. of 28th International Conference on Computer Animation and Social Agents (CASA2015) (2015).
- 10) Y. Hayakawa, Issei Fujishiro: “2D Fluid Shape Design by Direct Manipulation”, Proc. of the 5th IEEEJ International Workshop on Image Electronics and Visual Computing (IEVC2017), 3B-2 (2017).
- 11) 早川 雄登, 藤代 一成: “直接操作による 3D 流体シミュレーションの制御”, 情報処理学会研究報告, Vol. 2017-CG-168, No. 26 (2017).
- 12) F. Shimizu, I. Fujishiro: “Selection of Localized Audio Track Based on Eye-Tracking Technologies with Application to Musical Art Gallery”, Proc. of the 5th IEEEJ International Workshop on Image Electronics and Visual Computing (IEVC2017), 5C-1 (2017).
- 13) 清水 文也, 藤代 一成: “視線の隠的追跡手法とそのエンタテイメント応用”, 情報処理学会研究報告, Vol. 2017-CG-168, No. 25 (2017).
- 14) 小林 杏理, 藤代 一成: “即興演奏に呼応するアフェクティブな映像生成システム”, 映像情報メディア学会技術報告, AIT 2016-42, Vol. 40, No. 11, pp. 261-264 (2016).
- 15) A. Kobayashi, Issei Fujishiro: “An Affective Video Generation System Supporting Impromptu Musical Performance”, Proc. of Cyberworlds 2016, pp. 17-24, IEEE Xplore (2016).



藤代 一成 (正会員)

1985年 筑波大学大学院博士課程工学研究科電子・情報工学専攻修士号取得退学後, 東京大学理学部情報科学科助手着任. 1988年 理学博士 (東京大学). 筑波大学助手・講師, お茶の水女子大学助教授・教授, 東北大学教授を経て, 2009年より慶應義塾大学理工学部情報工学科教授, 現在に至る. 2015年同塾特選塾員, 2017年より同大学院理工学研究科情報工学専攻主任教授を兼務. ビジュアルコンピューティング, 特に応用可視化と可視化ライフサイクル管理, 知的環境メディアに関する研究に従事. 本学会では, VC 委員長, 会長等を歴任. 現在, 日本学術会議連携会員, 日本工学会・情報科学国際交流財団理事, 芸術科学会・画像情報教育振興協会評議員.