

いという要請がある。このような映像提示に適した表示装置として、我々は、多視点对応裸眼立体視表示装置を注目している。現在商品化されているものとしては、Actuality社の開発した全方位型表示装置（Perspecta）と三洋電機が開発した多視点方式表示装置がある。前者は、全周からの鑑賞が可能な全方位型表示装置（360-degree-viewable volumetric 3D display）であり、透明な球体中にある円形の回転式スクリーンにボリュームデータのスライス像（198スライス、768x768）が投影されることにより立体映像が表示される。これらの表示装置はCG映像の表示に用いられており、カメラで撮影された実映像表示には用いられていない。遠隔協調空間では、相手の表情を伝えるために多視点カメラで撮影された映像を表示する必要がある。多視点映像向け裸眼立体視表示技術に関する研究の詳細は参考文献[5]をご覧ください。上記の研究概要およびその他の取り組みは当研究室のホームページに記載されている（<http://www.viz.media.kyoto-u.ac.jp/>）。

参考文献およびURL

[1] <http://www.viz.media.kyoto-u.ac.jp/>

[2] 小山田耕二、酒井晃二，“ボリュームコミュニケーション技術を使った協調可視化環境の構築”、可視化情報、Vol.24, No.95, 2004（10月）、pp.228-233

[3] N. Sakamoto, K. Koyamada, K. Sakai and K. Kikugawa, “Voxelization of Hexahedral Cell with the Two-Pass Rasterization Technique,” The 4th IASTED International Conference on VISUALIZATION, IMAGING, AND IMAGE PROCESSING (VIIP 2004) , pp. 178 – 181, 2004.

[4] Y. Watashiba, J. Nonaka, N. Sakamoto, Y. Ebara, K. Koyamada and M. Kanazawa, “A Streaming-based Technique for Volume Rendering of Large Datasets,” Proceedings of the IASTED CGIM2003, pp. 187-192, 2003

[5] H. Hazama, N. Sakamoto, H. Horii, Y. Ebara and M. Koyamada, “Multi-Viewpoint Videos Merging System Using Auto-Stereoscopic Display in Tele-Immersion,” The 4th IASTED International Conference on VISUALIZATION, IMAGING, AND IMAGE PROCESSING (VIIP 2004) , pp. 719 – 724, 2004.

情報メディア工学講座 複合メディア分野（中村(裕)研究室）

「映像コンテンツの自動取得 ～人間の行動の観測による映像の自動撮影・編集～」

コンピュータや通信技術の進歩により、映像(動画)やそれを基にしたマルチメディアコンテンツが教育コミュニケーションの手段として重要な位置を占めるようになってきました。企業、教育機関、個人やその他のコミュニティでも、映像を製作することに対する関心が高まっています。しかし、映像の撮影は、世界で起こっている出来事の一部(時間、空間的な一部分)を切り出し、編集する行為であり、真面目に取り組むとかなり難しい問題だとも言えます。単純に撮り流したホームビデオが、他人にとって見るに耐えない代物となることから、それはよくわかります。

この問題に対する我々のアプローチは、映像の自動撮影、自動編集を行うシステムを構築すること、得られる映像データの新しい利用方法をサポートすることです。そのために、図のような自動化スタジオを構築しています。現在は、プレゼンテーション、トレーニング、ミーティングのような教育に関連する場面を想定し、そこでの実応用を目的としています。

この研究のポイントの一つは、自動カメラマンの機能をシステムに持たせたことです。システムが撮影対象、つまり人間の顔や手先などを追跡しながら撮影しますが、人間のカメラマンとは違い、種々のセンサからの計測結果、画像認識、音声認識の結果を使って、複数の自動カメラを制御することによって、映像データを取得します。

また、ただ撮影しただけでは、多くのカメラから得られる大量の生データが残るだけです。そこで、複数のカメラで撮影されたデータの中から最も重要そうな部分、つまり視聴者(ユーザ)に見せるべき部分を検出し、自動的に映像の編集をする機能についても研究を行っています。注目すべき部分は、時間的・空間的に常に変化するため、人間の行動(体の動きや発話等)を観測し、動作・発話・物体の動きなどを統合的に認識するマルチモーダルな認識技術が重要なポイントとなっています。

さらに、得られたデータは、様々な視点から人間の行動の様々な部分をとらえたものとなっています。つまり、様々な人の様々な要求に対して適切な情報を提供するという意味からも、面白いデータとなっており、質問応答などの研究への応用を試みています。

このような技術が進めば、人間の先生の講義、実演、アドバイスなどを自動撮影し、その言動や知識を仮想化(様々な形で利用コンピュータの中に取り込む)することが可能になります。本や現在のマルチメディア教材との違いは、先生を仮想化することによって、質問に答えたり、利用者の状況に合わせてアドバイスする機能を持たせることが可能になる点です。これによって、自動化システムが生徒の様子をじょうずに見守りながら必要かつ的確なアドバイスをするバーチャル実験室やトレーニングルームの実現も夢ではないと考えています。

<参考文献>

- 尾関、中村、大田、机上作業シーンの自動撮影のためのカメラワーク、信学論D-II-J86, 2003
- 尾形、中村、大田、制約充足と最適化による映像編集モデル、信学論D-II-J87, 2004
- 尾関、中村、大田、注目喚起行動に基づいた机上作業映像の編集、信学論D-II-J88, 2005

