

多視点映像を用いた遠隔オフィス間 テレワーク実験環境の構築

山崎 貴宏

事業環境の変化やICTの進展により、オフィス環境におけるコミュニケーションスタイルやワークスタイルが多様化している。場所にとらわれないワークスタイルとしての遠隔オフィス間のテレワークや、保守現場・製造現場と事務所などの遠隔地とのリアルタイムコミュニケーションを実現する、テレワーク環境に対する期待も高まっている。我々は、地理的に離れたオフィス間で働く人同士が、同じオフィスにいるかのような感覚を得て仕事ができるようなテレワーク環境の実現を目指し、これを「超臨場感テレワークシステム」と呼んでいる。図1は、超臨場感テレワークシステムの実現イメージを示している。このシステムでは、オフィス内に配置した複数のカメラやマイク、センサーを用いることで、遠隔地へオフィス空間の状態を伝え、距離を意識しないコミュニケーションの実現を目指している。



距離を意識しないコミュニケーションの例としては、遠隔オフィス間での立ち話の実現が挙げられる。同じオフィスにいるオフィスワーカー同士で立ち話を開始する場合は、相手の状況（忙しそうか、話しかけてよさそうか）を察して、会話を開始するという行為が自然

に行えるが、遠隔地の場合では、相手の状況を察することができないため、タイミングよく立ち話を開始することは難しい。超臨場感テレワークシステムでは、遠隔地の俯瞰映像による様子見や、忙しさの度合いといったオフィス状況の推定結果を利用することにより、話しかけてもよいかを判断し、自然に会話を開始できる。距離を意識しないコミュニケーションのもう一つの例としては、臨場感のある映像による対話がある。まるでそこにいるかのような臨場感の高い映像表現を用いることで、同じオフィスで働いている感覚をより高めることが可能となる。

上で挙げた例のうち、臨場感テレワークシステムにおける対話のシーンに着目すると、図1では右上で、大画面共同作業端末を用いた遠隔コラボレーションを行っている様子で表現している。ここでの対話を“窓越し感覚の対話”と表現している。“窓越し感覚”とは、遠隔地にいる相手の映像が、まるでそこにいるかのように表示され、ディスプレイがあたかもガラス窓であるかのように感じるような映像表現である。“窓越し感覚の対話”により、相手の存在感が感じられるだけでなく、視線や身振りなども実際に会って話をしているときに近い状態で伝わるため、共同で作業を行う上で円滑なコミュニケーションが期待できる。

相手がそこにいるかのような映像表現としては、立体映像表示が考えられる。立体映像表示のうち、左右の2視点を用いたステレオ立体映像表示は、3D映画などで一般的になりつつある。ステレオ立体映像表示では、左右2視点による両眼視差により、立体感を感じることができるが、表示する視点はあくまで2視点のみであり、“窓越し感覚”を実現できない。

“窓越し感覚”を実現する立体映像表現には、両眼視差による立体感に加え、見る位置に応じた視点の変化で生じる運動視差による立体感の表現が挙げられる。このような表現を行うためには、2視点だけではなく、さらに多くの視点が必要となる。

そこで、遠隔地との“窓越し感覚の対話”が可能な映像表現を実現するため、多視点映像を撮影し、遠隔地へ

伝送するシステムの開発を行った。本稿では、開発したシステムと、多視点映像を用いた“窓越し感覚の対話”を実現するための取り組みについて紹介する。

立体映像表示のための 多視点映像伝送技術の課題

“窓越し感覚の対話”を実現するような、立体映像表示を行うためには、同期のとれた複数視点の映像が必要となる。一般的に、複数視点の映像取得には、撮影同期機能を持つ複数の単眼カメラを組み合わせ、カメラアレイを設置する方法がとられる。

取得した多視点映像を遠隔地へ伝送する際には、H.264 MVC¹⁾(以降では、MVCと表記する)で圧縮符号化し、伝送する方法が広く用いられている。MVCは多視点映像の圧縮符号化を行うために規格化されており、符号化した結果の多視点ストリームには、視点間の同期を維持する構造を持つ。

しかしながら、MVCは視点間予測により効率的な圧縮を実現する符号化方式のため、符号化処理に必要な演算量が大きく、それは、処理する視点数に比例して大きくなる。さらに、伝送される多視点ストリームはすべての視点を含む1ストリームという形態であるため、送信側、受信側それぞれ1つの計算機ですべての視点の符号化・復号処理を行うこととなり、分散して処理を行うことが難しい。そのため、多くの視点の伝送・表示をリアルタイムに処理を実現するのが難しい、という問題があった。

遠隔コミュニケーションにおいて、“窓越し感覚の対話”を実現するためには、立体映像表示に必要な視点の多視点映像を同期のとれた形で撮影し、リアルタイムに伝送、処理可能なシステム構成が必要となる。

システム概要

我々は、5視点映像の取得・符号化、送信を行う多眼映像センサーと、これを複数用いた多視点映像のリアルタイム伝送を行うシステムを開発した。

多眼映像センサーは符号化方式に少ない演算量で符号化を行うことができるDistributed Video Coding²⁾(以降ではDVCと表記する)方式を採用することにより、同期のとれた5視点映像の符号化・伝送が可能となっている。

開発した多視点映像伝送システムでは、複数の多眼映像センサーからそれぞれ、多視点ストリームが送信

されるため、伝送する視点数を増やす場合でも負荷を分散することができる。

開発したシステムでは、8台の多眼映像センサーを使い、40視点の映像取得・伝送と、伝送された40視点の映像のうち、5視点分のリアルタイム映像出力を実現している。

システムの出力する多視点映像を、立体映像表示が可能な表示装置へ接続することで、“窓越し感覚の対話”を実現する環境が構築可能になっている。

システム構成

図2に開発した多視点映像伝送システムの構成を示す。送信側では、多眼映像センサーを用いて、多視点映像を取得し、遠隔地へ伝送する。受信側では、中継サーバーで受信した多視点ストリームのうち、多眼映像センサー単位の5視点ストリームを選択し、出力端末においてリアルタイムに復号・映像出力を行う。

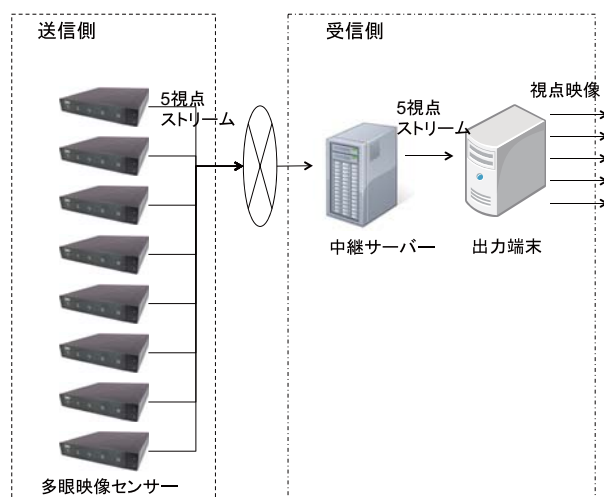


図2 システムの構成

構築したシステムでは多眼映像センサー単位の5視点のリアルタイム出力を実現しているが、出力端末を増やすことで、処理負荷を分散しつつ、さらに多くの視点を表示する拡張も可能となっている。

以下では、構築したシステムの特徴的な構成要素である、多眼映像センサーについての説明を行う。

多眼映像センサー

多眼映像センサーは、筐体に固定した5つのカメラモジュールから映像を取得し、符号化、ネットワーク送信

を行うネットワークカメラである。図3に示す通り、5つのカメラモジュールは、直線状に配置しており、その間隔は67mmと人間の目の間隔に近いものとなっている。そのため、自然な立体視のための視点映像の取得が可能になっている。



図3 多眼映像センサー

多眼映像センサーの特徴として、符号化方式に低演算量で処理可能なDVC方式を採用していることが挙げられる。DVCは映像の圧縮符号化を低演算量で行うことが可能である方式のため、5視点のHD映像の撮影と圧縮符号化、ネットワーク送信の処理をFPGA 1チップで実現している。この特徴により、5台のカメラモジュールの起動タイミングを揃えることが可能となり、それぞれのシャッタータイミングの同期を実現している。さらに、符号化と多重化も同じFPGA内で行うため同期を保ったまま処理を行うことが可能になっている。多重化方式には、MVCで規定された多視点ストリームの構造を独自拡張した形式で実現している。この構造を用いることで、視点間の同期を維持したまま、DVC方式で符号化したデータを伝送することが可能となっている。

“窓越し感覚の対話”の実現への取り組み

テレワークにおける、“窓越し感覚の対話”を実現するために、開発した多視点映像伝送システムに独立行政法人 情報通信研究機構の開発した裸眼立体視ディスプレイ³⁾を接続する取り組みを行っている。

裸眼立体視ディスプレイでは、複数のプロジェクタを用いて、多視点映像を特殊なスクリーンに重ねて投射し、スクリーンの特性により、見る位置に応じた視点の映像が表示される。図4に、裸眼立体視ディスプレイの視点とプロジェクタの対応のイメージを示す。

裸眼立体視ディスプレイにおける、見る位置に応じた、視点映像の表示の違いを、図4を用いて説明する。観察

者がAの位置でスクリーンを見ている場合、スクリーンの特性により、右目にはプロジェクタ1の出力する視点映像が映り、左目にはプロジェクタ2で出力する視点映像が映る。左右の目それぞれに視点の異なる映像が映るため、両眼視差による立体感を感じることができる。次に、位置Aから位置Bに移動した場合、右目にはプロジェクタ1の視点映像が見えている状態から、移動に伴い、プロジェクタ2, 3, 4の視点映像と滑らかに切り替わって見える。これにより、観察者は運動視差による立体感を得ることができる。

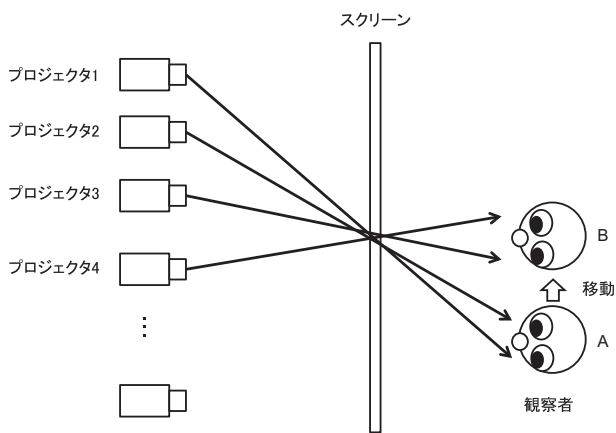


図4 裸眼立体視ディスプレイの構成イメージ
文献⁴⁾掲載の図を基に作成

この表示方式により、従来の2眼式の3Dディスプレイで必要であった、特殊なメガネを用いることなく、両眼視差を感じることができる上に、運動視差による立体感も同時に体感できる。

開発した多視点伝送システムと、裸眼立体視ディスプレイの接続には、多眼映像センサーのカメラの並びに対応するように、出力端末から出力される視点映像をそれぞれのプロジェクタへ入力する。この接続により、多眼映像センサーで撮影した空間を再現するような立体映像が表示される。

図5では、開発した多視点映像伝送システムを用いて、多眼映像センサーで撮影した多視点映像を裸眼立体視ディスプレイへ表示した際の様子を示している。この図では、異なる3地点でのスクリーンに映る視点映像の見え方の違いを表しており、人物と、人物の持つ物体の位置関係の違いから、見る位置によって視点が切り替わっていることがわかる。

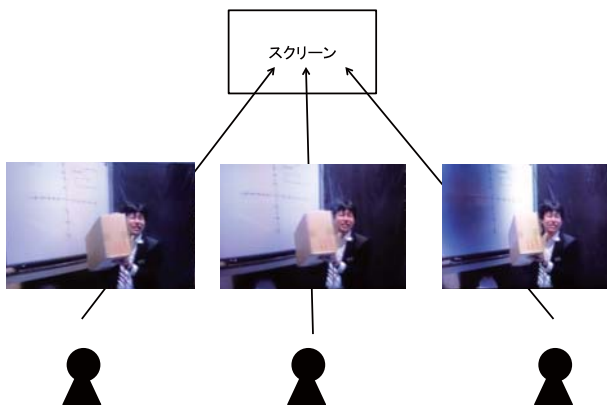


図5 裸眼立体視ディスプレイでの表示

今後の展開

現在、開発した多視点映像伝送システムと接続し、立体映像の表示を行っている。開発した多視点映像伝送システムでは40視点の映像伝送を実現しており、出力端末をさらに増やすことで、リアルタイム処理を維持したまま、出力する視点数を増やすことが可能な構成となっている。さらに、裸眼立体視ディスプレイで表示可能な視点数は40視点よりも多い。そこで、5視点を越えた、さらに多くの視点数を用いることで、広い範囲の空間を再現するような、立体映像表示を行うことを検討している。この場合、多眼映像センサー単位の5視点では生じなかった問題、例えば、多眼映像センサー間の同期やカメラの位置関係の調整などの課題があると考えており、これらの課題を解決するための技術開発を進める。同時に、構築したシステムを東京-大阪間のテレワーク環境として運用し、実証実験を行っていく。

多視点映像の伝送と裸眼立体ディスプレイの接続による応用例としては、オフィス間のテレワークといった使い方のほかにも、遠隔会議として用いられているテレプレゼンスへの応用はもちろん、銀行などの窓口業務を遠隔地で行う、バーチャル窓口のような使い方も可能となる。また、見る位置を変えることで自然な視点の移り変わりが表現できることを利用して、ロボットの遠隔操作にも利用可能だと考える。例えば、遠隔操作を行う際に、作業領域が死角となっているような場面では、頭の位置を動かすことにより、作業領域を確認する、という自然な操作を実現できると考える。このほかにも、デジタルサイネージや、ショールームといった商品説明への利用や、ライブ中継のパブリックビューイングなどのエンターテインメントへの応用も考えられる。

おわりに

地理的に離れたオフィス間のコミュニケーションといった、遠隔地コミュニケーションにおいて、臨場感の高い映像を用いた対話を行うために必要となる、多視点映像の伝送を実現するシステムを構築した。

本稿では、多眼映像センサーを用いて取得した40視点映像の伝送するシステムについて説明し、構築したシステムと、裸眼立体視ディスプレイを接続することによる“窓越し感覚の会話”の実現のための取り組みについて紹介した。

今後は、裸眼立体視ディスプレイで表示する視点数を増やすといった改良を行うとともに、構築した環境を用いて、テレワークを行い、多視点映像伝送・表示による効果を検証する実証実験を進めていく。

謝辞

本稿で開発したシステムの一部は、独立行政法人情報通信研究機構の委託研究「革新的な三次元映像による超臨場感コミュニケーション技術の研究開発」の成果である。◆◆

参考文献

- 1) Multiview video coding (MVC): ISO/IEC 14496-10:2009 Annex H.
- 2) K. Sakomizu, et al.: A real-time system of distributed video coding, in Picture Coding Symposium (PCS), 2010. pp. 538-541, 2010.
- 3) <http://www.nict.go.jp/press/2008/press-20080930-3.pdf> (2014年8月20日)
- 4) 岩澤昭一郎：世界初！200インチ自然裸眼立体視ディスプレイ, NICT NEWS, 2011年11月号, No.410, pp1-4. による立体視の実現

● 筆者紹介

山崎貴宏：Takahiro Yamasaki. 研究開発センタ メディア処理技術研究開発部