

# 全方位カメラと制御カメラを利用した 遠隔カメラワークシステム

佐藤 洋介† 米田 裕也‡ 橋本 浩二† 柴田 義孝†  
岩手県立大学ソフトウェア情報学部† 株式会社アイソニック‡

近年、通信ネットワークの高速化とマルチメディア処理技術の発達に伴い、遠隔テレビ会議に関する研究が行われており、遠隔地間で円滑なコミュニケーションを行うことが可能となって来た。しかしながら、従来のテレビ会議に利用されるカメラは撮影範囲が狭く、1度に広い範囲の撮影や注目する視野全体を提供することが困難であった。本稿では、全方位カメラとPTZカメラを組み合わせた新たな遠隔テレビ会議支援システムを提案する。本システムを遠隔地に複数設置し、遠隔地点間で映像ストリームを送受信、受信した360度パノラマ映像から任意の位置を選択し、PTZカメラがその方向を撮影する。これにより、新しい応用に対応できるビデオ会議システムができる。

## Remote Camera Work System using a Combination of Omni-directional Camera and Controlled Camera

Yousuke Sato†, Yuya Maita‡, Koji Hashimoto†, Yoshitaka Shibata†  
† Faculty of Software and Information Science, Iwate Prefectural University  
‡ AiSONIC Co Ltd.

As advent of communication network and multimedia processing technologies, various video conference systems to smoothly communicate each other between remote sites using with high quality audio and video. However, since those teleconference systems usually use conventional cameras which only capture the objective images in one direction with narrow area, more wide area such as 360 degree of surround direction cannot be captured. In this paper, we introduce a new teleconference system in which omni-directional camera and PTZ(Pan/Tilt/Zoom) are combined. Using multiple combined camera systems, more flexible and dynamic video conference system can be realized for new application fields.

### 1. はじめに

近年、通信ネットワークの高速化とマルチメディア通信技術の発達に伴い、IPベースのネットワーク上で高品質の映像や音声を用いて、遠隔地間をリアルタイムかつ双方向でコミュニケーションを可能とする遠隔講義システムや遠隔会議システムが研究され、その一部は実用化されて来ている。

このようなシステムのほとんどが従来の単一方位型カメラを利用しているため撮影範囲が狭く、一度に広い範囲の映像や注目する視野範囲を拡大したりすることが困難である。また特定の話者に追従させたり、逆に話者を特定して映像や音声を選択したりする機能も実現が困難であった。

これに対し筆者らは一度に 360 度撮影可能な全方位カメラを導入し、取得した全方位環状映像をパノラマ映像に展開し、これをマルチメディアストリームとして RTP パケットでリアルタイムに送信し[2]、受信側でパノラマ映像を表示したり、その中の注目する領域を拡大したり、またゆがみを補正したりする機能を有する全方位映像通信ミドルウェア[1]を開発してきた。

また全方位カメラ映像に、PTZ (Pan/Tilt/Zoom)カメラを組み合わせ、遠隔よりパノラマ映像を見ながら、任意の点にパン・チルト・ズーム操作を行うことにより、自由にかつ高精細に撮影できるシステムの開発を行ってきた。[3]

本研究では[3]のシステムの機能をさらに向上させた新たな遠隔テレビ会議支援システム(Telegnosis System)を提案するものであり、全方位カメラと PTZ カメラおよび複数の指向性マイクの組み合わせを一つのユニットとして構成し、複数個のユニットを会議室に設置し、遠隔地点間で複数の映像ストリームを送受信したり、映像を切り換えたりすることが可能である。また受信した 360 度パノラマ映像から任意の位置を指定することにより、PTZ カメラがその方向を撮影し、パン・チルト・ズーム操作が可能となる。またカメラの音声および映像は送受信側どちら側でもファイルとして保存できるので、会議の状況を録画や再生することが可能である。さらに複数の指向性マイクにより、音声方向から自動的に話者を同定し、話者の音声と映像だけを取り出したり、逆に、パノラマ映像上で指定した話者方向の映像と音声のみを取り出しことも可能である。

本稿では、提案するシステムの構成と要素技術、カメラワークアルゴリズム、そして前述の機能を実現するプロトタイプシステムについて述べる。

## 2. システム概要

本システムでは全方位カメラと PTZ カメラ、有指向性マイクを組み合わせたカメラをテレグノシスと呼ぶ。テレグノシスを構成する全方位カメラはデジタルビデオや

USB カメラなど、利用目的や利用環境に応じて選択することが可能である。またカメラの設置状況に応じて映像を上下反転させることも可能である。

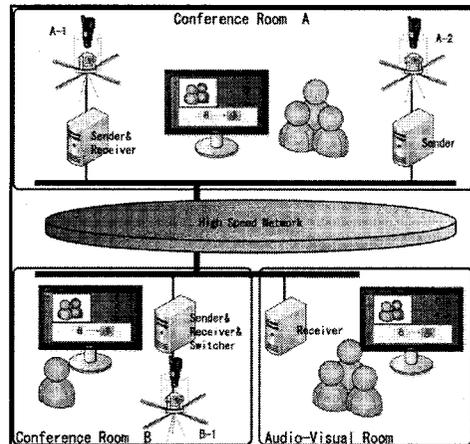


図 1 システム概要

図 1 に示すように、本システムは高速ネットワークを利用して各遠隔地のテレビ会議室上に構築され、テレグノシス A-1、A-2、B-1 がそれぞれ設置されるとすると、Conference Room B の B-1 の映像は Conference Room A の左側の PC に、Conference Room A の A-1 と A-2 の映像は Conference Room B の PC に表示・制御される。この時 Conference Room B では A-1、A-2 の映像を同時に閲覧しつつ、映像を切り換えて、透過率を指定して合成などの制御を行った映像を Audio-Visual 室に転送することができる。これをストリームスイッチ機能と呼ぶ。Audio-Visual 室では A-1 と A-2 の選択された映像が切り換ったり、交互に表示されたり、合成される。

これらの機能によって、遠隔テレビ会議参加者がテレグノシスを制御することで要求通りの視野映像を取得することができる上、図 1 の Audio-Visual 室に相応する他の広い会場にいる視聴者へ転送することができ、より大規模な会議やその他の遠隔協調システムへの利用が期待できる。映像及び音声の IP 転送やファイル保存には後述する MidField System[3]を、全方位映像のパノラマ展開処理は全方位ミドルウェアを使用する。

### 3. アーキテクチャ

遠隔テレビ会議支援システムのアーキテクチャは図2で示す通り **MidField System** の上位に **Telegnosis System** 層として2階層3プレーンで構成される。**Interface Layer** では、ユーザの操作情報を処理し下位層へ通知する。またシステムやマルチメディアストリーム、展開処理、PTZ カメラの設定やステータスの管理がおこなわれる。**Multimedia Control Layer** では全方位映像のパノラマ展開処理、音声方向の算出処理、ユーザの操作によるカメラワークの計算とコマンド生成、映像ストリームの生成と制御を行う。

図中の2本のブロックは互いに会議を行う物理ホストに対応し **Multimedia Stream**(全方位映像, PTZ カメラ映像)を2本ずつ送受信し, PTZ カメラ制御コマンドを送信する。

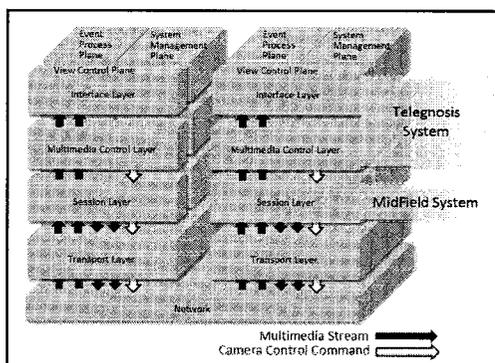


図2 アーキテクチャ

#### 3.1 MidField System

**MidField System**[3]は、通信端末の処理能力や利用可能なネットワーク帯域幅に応じて、適切なフォーマットによる通信を実現する、柔軟性を持ったマルチメディア相互通信ミドルウェアである。**MidField System**の主要となる機能として、トランスコーディングがある。トランスコーディングとは、相手の環境に応じて **RTP** ストリームのフォーマットを柔軟に変更するものである。

現在対応しているフォーマットは、映像に関するものでは **DVSD, HDV, MPEG4, WMV, MJPEG** がある。またマルチキャスト通信も容易に行うことが可能である。リアルタイム映像及び音声だけではなく、リアルタイム映像の録画とビデオファイルの再生も可能である。またビデオファイルの像(図3)を、全方位パノラマ展開映像(図4)に展開処理を行い表示すること、再生内容を送信することも可能である。本システムでは **MidField System** を **Session Layer** として位置づけ、各種機能を利用する。

#### 3.2 全方位映像ミドルウェア

全方位映像ミドルウェア[3]とは **Telegnosis System** 層の機能を実現しており、全方位カメラから取得される全方位環状映像のパノラマ展開処理を可能とする。全方位カメラは高精細化・小型化しており、利用目的に応じて選択でき、また全方位ミドルウェアもそれに応じた動作が可能となっている。

しかしながら、全方位映像の各フレームなどの処理を実行するため、PCのCPU負荷が非常に大きくなるという問題点がある。通常の環状映像を送信した場合はフレームレートの低下を招く。特に **HDV** フォーマットの場合、映像及び音声の符号化のための処理量の増加のため、ブロックノイズや遅延の要因となる。このため、CPU能力に応じたフレームレート制御などの方法を検討している。

#### 4. 遠隔カメラワークアルゴリズム

遠隔カメラワークとしては、広範囲で参加者の動きの伴う演習や、注目する視野を柔軟に変更できるようにするため、カメラが撮影している方向を移動する視点移動、撮影するカメラを変更するストリームスイッチ機能を実現している。また、複数の指向性マイクにより特定の話者を自動認識し、視点移動を行う。

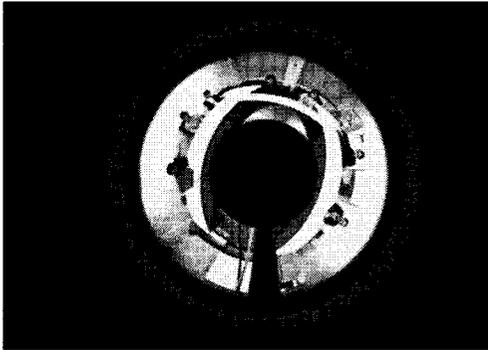


図 4 全方位環状画像



図 3 全方位パノラマ展開画像

#### 4.1 視点移動アルゴリズム

視点の移動はパノラマ映像上で、ユーザが参照したい位置を指定することにより PTZ カメラが視点を移動する機能である。パノラマ映像上の指定された座標から、対応する PTZ カメラの注視方向を算出し、ネットワークを越えてコマンドを送信し制御する。また、本システムでは、有指向性マイクを 90 度間隔で 4 本配置することにより、音声の方向をもとに話者を特定し、その方向にカメラを向けることが可能である。逆に、パノラマ映像から指定された話者の映像および音声も抽出して受信することも可能である。

#### 4.2 ストリームスイッチングアルゴリズム

ストリームスイッチ機能は 2 つの機能で実現されており、1 つ目は、対応する全方位映像と PTZ カメラ映像を複数対受信し、それらを制御する機能である。これによって複数の遠隔地の映像を同時に表示したり、切り替えて選択表示が可能となる。2 つ目は、複数の地点(Sender)の中から選択した映像ストリームを他の地点(Receiver)に転送することが可能である。これは MidField System

のピクチャーインピクチャー機能を利用し、複数の映像ストリームを重ね合わせて 1 本の映像ストリームにするものであり、それぞれ透過度を設定することによって映像ストリームを切り替える。図 5 に示す例においては 3 対のストリームのうち、PTZ カメラ映像 A, B, C だけを Switcher で 1 本にまとめ Receiver に送信している。このとき、Switcher で B の映像ストリームを選択した場合、B の透過度を 0%, それ以外のストリーム A, C の透過度を 100% にすることで B が Receiver 側で表示される。

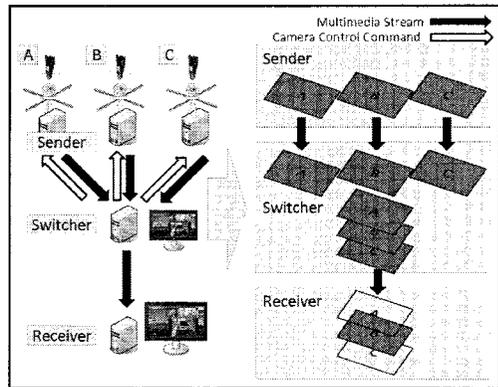


図 5 ストリームスイッチング

#### 5. プロトタイプシステムと実証実験

本研究で提案する Telegnosis の有効性を確認するためにプロトタイプを構築し、異なった 3 つの分野でのアプリケーションに利用し、その機能及び性能評価を行った。Telegnosis は全方位カメラ (SONY HDR-HC1) と PTZ カメラ (SONY SNC-RZ30N) との組み合わせカメラ、指向性マイク (SONY ECM-672), 8 チャンネル音声入力ボード (TD-BD-8CSUSB) で構成される。本システムの開発言語は Visual C++を使用する。

本研究の有効性を確認するため JGN2 を基盤としたプロトタイプシステムを構築し、実際にそれぞれの環境で遠隔テレビ会議を行うことで、その性能を評価する。プロトタイプシステムは図 1 に示すように Conference Room A に組み合わせカメラを 2 台、Conference Room B には 1 台設置する。

各組み合わせカメラはそれぞれ映像送信 PC に IEEE1394 ケーブルで接続される。両会議室の参加者はそれぞれ相手側会議室の組み合わせカメラを制御できる。また、Audio-Visual 室の参加者は、Conference Room B のユーザによって選択されたストリームを閲覧することができる。

### 5.1 遠隔グループワーク実習支援実験

従来、複数人によるグループワーク実習の際、参加者の顔や振る舞いを撮影する方法がなかった。本実験では全方位カメラを利用し 6 人がそれを囲み、学内ネットワークを利用して実験を行った。図 6 はパノラマ展開された学生たちの映像から分析を行う教員側の様子を示す。



図 6 遠隔グループワーク実習

### 5.2 高度職業人材教育支援実験

一般的に食品加工場は狭く、多くの研修生をその場に立ち入らせることが難しい場合を想定し、食品加工場の中心にテレグノシスを設置し、隣の部屋からその様子を観察した。被写体となった調理師たちは、カメラの存在を意識することなく作業を行うことができた。

### 5.3 気球災害状況撮影転送システム

テレグノシスを軽量化し 802. 11n の無線 LAN カードを取り付けた PC とともに、直径 3. 5m の気球につりさげ上空 40m の高さまで打ち上げた。地上ではテレグノシスの

映像を受信し、カメラ映像の閲覧と制御をおこなった。

## 6. 性能評価

本システムの性能を示す指標として使用帯域、CPU 使用率、遅延、フレームレートについて評価をおこなった。システムの構成は、図 7、図 8 のとおりである。

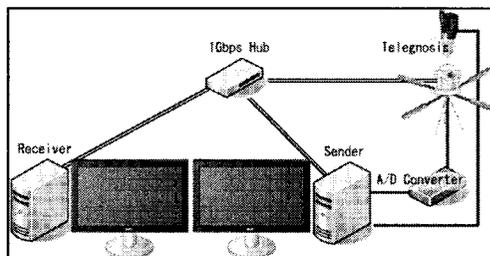


図 7 評価システム構成

Specification	Sender	Receiver
CPU	Pentium4 3.4GB	Pentium4 3.5GB
MEMORY	1.5GB	1.5GB
NIC	1.0Gbps	1.0Gbps

PTZ Cam	
Model	SONY SNC-RZ30N
Pan	+170°~-170°
Tilt	-25°~+90°
Swing Speed	1.5 seconds

図 8 評価システム性能

送信側の PC を Sender, 受信側の PC を Receiver とし、Sender に全方位レンズと装着した DV カメラ (SONY HDR-HC1) を IEEE1394 ケーブルで接続した。PTZ カメラ (SONY SNC-RZ30N) はアナログ NTSC 映像信号を出力するため A/D コンバータで、デジタルビデオに変換後 IEEE1394 ケーブルで接続する。全方位映像のパノラマ展開処理、カ

メラの制御は Receiver で行う。使用するネットワークは 1.0Gbps Ethernet。送信する映像ストリームのフォーマットを DV と WMV の 2 つの条件で行った。全方位環状映像、PTZ カメラ映像の解像度は 720×480 ピクセルの映像で、パノラマ展開後の全方位映像は 800×200 ピクセルになる。遅延の測定方法としては、Sender に接続されたカメラで Receiver の画面を撮影することで、Receiver の画面には再帰的に Receiver の画面が表示される。このとき部屋の電気を消した時間からこの 2 番目の再帰画面に映る部屋の電気が消えるまでの時間を測定し、この値を 2 で割ることで測定した。

Video Format		DV	WMV
Bandwidth		59.2Mbps	15.6Mbps
CPU load	Sender	5%	85%
	Receiver	75%	70%
Delay	Panorama	85ms	615ms
	PTZ	55ms	595ms
Frame rate	Panorama	26fps	20fps
	PTZ	29fps	21fps

図 9 評価結果

評価の結果、DV フォーマットのネットワーク使用帯域は WMV を利用した場合に比べ 4 倍程度だった。しかし、DV フォーマットは CPU の利用率、遅延、フレームレートにおいては WMV に比べ優位性が高い。DV フォーマットはビデオカメラ内で MJPEG をベースにしたフレーム内符号化方式をため遅延が小さい。WMV は MPEG4 をベースとしており、CPU を大きく利用する半面、高い圧縮率を実現する。このため、ネットワーク帯域と PC の性能に応じたフォーマットの選択が必要となる。

また、PTZ カメラの制御を行ってから、PTZ カメラがその方向を撮影し、映像が Receiver 到達するまでの制御遅延時間は、PTZ カメラの可動範囲である +170 度から -170 度まで移動する時間が 2 秒程度であった。

## 7. まとめ

本稿では、遠隔地に設置した全方位カメラと PTZ カメラの映像を利用した遠隔カメラワークシステムの概要とカメラワークアルゴリズム、利用例、性能評価結果を示した。現在は複数カメラの制御と録画、ストリームスイッチ機能が実現されているが、数々の実験から得たアイディアや、ライブラリとして利用している MidField System の高機能化など、今後追加する機能が数多くある。その中でも、ネットワークや CPU の利用状況に応じた QoS 制御機能、動体検出や顔検出機能による PTZ カメラ制御を実現すれば監視カメラとして利用した際に、全方位カメラに映りこんだ人物の顔を PTZ カメラで追跡しながら撮影することができるようになる。

今後は、上記に述べた機能の追加や実証実験等の結果からユーザに対するアンケートによる評価を行っていく予定である。

## 参考文献

[1] 米田祐也, 橋本浩二, 柴田義孝: 全方位映像通信のためのミドルウェアの研究, 情報処理学会第 67 回全国大会, pp595-596, (2005)

[2] 橋本浩二, 柴田義孝: 利用者環境を考慮した相互通信のためのミドルウェア, 情報処理学会論文誌, Vol. 46, No. 2, pp.403-417, (2005)

[3] 小林亮太, 米田祐也, 橋本浩二, 安藤広子, 佐々木由香, 武田利明, 柴田義孝: 異種映像の組み合わせによる遠隔ヘルスケア教育支援システムの研究, 情報処理学会第 68 回全国大会 6T-10 pp. 4-259~260 (2006.3)

[4] Ross Cutler, Yong Rui, Anoop Gupta, JJ Cadiz: Distributed Meetings: A Meeting Capture and Broadcasting System, ACM Multimedia '02, pp.1-6, (December 2002)

[5] 八木康史, 横矢尚和: 全方位ビジョン: センサ開発と応用の最新動向, 情報処理学会論文誌 vol.42 No.SIG13 pp.1-18, (2001)