

全方位映像通信のためのミドルウェアの研究

米田祐也 橋本浩二 柴田義孝

岩手県立大学大学院ソフトウェア情報学研究科

本稿において全方位映像を用いた双方向通信やマルチキャスト通信のためのミドルウェアシステムについて述べる。全方位映像は通常の単方位カメラと比較して 360 度の広い視野を持つため監視や TV 会議システムを効率的に実現できる。しかし、全方位映像は見やすくするための展開処理が必要であり容易に利用できないため、本ミドルウェアは全方位映像利用や効率的な映像転送をサポートすることより全方位映像の利用環境を提供する。

Research of Middleware for Omnidirectional Video Communication

Yuya Maita, Koji Hashimoto, Yoshitaka Shibata

Graduate school of Software and Information Science, Iwate Prefectural University

In this paper, a middleware for Omnidirectional video communication system is introduced. The omnidirectional image has a wider view than a single directional camera and able to realize surveillance system or TV conference in small room. In this paper, we describe the architecture and functions of the middleware for omnidirectional image control and effective video transmission system.

1. はじめに

近年、高性能な PC が一般的に普及しインターネットを介しての映像 / 音声の配信が行われるようになってきたが、これらの映像は通常の単方位カメラ映像利用しており広い視野で監視したり狭い空間内で遠隔 TV 会議を実現するには困難であった。本研究ではこれらの問題点を解決するために全方位カメラを導入し 360 度の視界を有する会議システムや監視システムのためのミドルウェアを開発している。本ミドルウェアにおいては利用できる計算機やネットワーク資源を考慮し動的に全方位映像[2]や拡大画像の品質を保証することが可能となる。そして、プロトタイプシステムとして以下 TV 会議システムを構築し評価について述べる。

岩手県立大学大学院ソフトウェア情報学研究科

Graduate school of Software and Information Science,

Iwate Prefectural University

2. システムの概要

本システム図 1 に示すようにサーバ・クライアント型で構築されており、全方位映像（環状画像）の取得、画像処理、送受信機能を果たす。そして、全方位映像は周囲 360 度の視界を有するといった特徴があり、これによりユーザの要求に応じて自由に視点を設定することできる仕組みを提供する。視点の制御

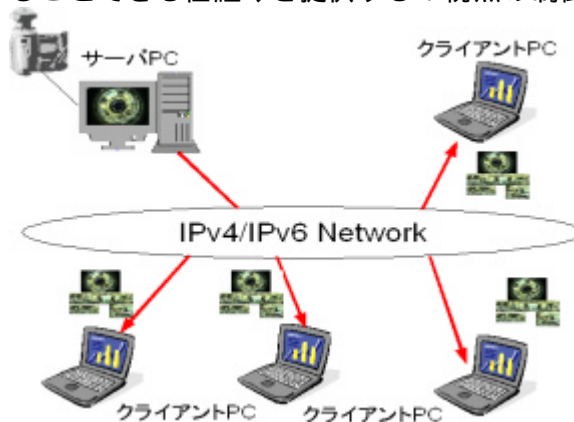


図 1 クライアント展開処理型

は上下左右と拡大縮小が可能であたかも通常の単方位カメラを可動させて使うかのように視点の操作が可能である。

サーバ側では全方位映像は PAL レンズを装着した DV(Digital Video) [3] カメラから IEEE1394 インターフェースで PC に取り込まれる。取り込まれた全方位映像はユーザからの要求に応じて展開処理を行わないもしくは展開処理を施し MidField システム [1] を用いて DV, MPEG4, Motion-JPEG の圧縮符号化方式で各クライアント側へ配信される。ユーザの要求により視点の制御をサーバ側で行うことによりユーザにはサーバで設定した視点の映像を提供することができる。

クライアントではサーバ側から配信された映像を受信し、クライアント側で全方位画像を展開してパノラマ映像を表示、またはサーバ側で展開されたパノラマ映像を受信しそのままの形で表示する。クライアント側で展開処理を行うことによりクライアントでは任意の方向の映像を自由に拡大縮小などの操作が可能となる。

2.1 . 全方位映像

全方位カメラはいくつかのメーカーより発売されているが、本研究での全方位映像の撮影は DV カメラに PAL レンズ(図 3)を装着(図 2)し、IEEE1394 で PC に接続することで全方位映像を取り込む。取り込まれる全方位映像を環状画像と言い 720 × 480 ピクセル 30 fps で周囲 360 度と垂直方向に 60 度(水平を基



図 2 全方位映像撮影カメラ



図 3 PAL レンズ

準に-20 度から+40 度)の視界を有するドーナツ状の映像である。このドーナツ状の環状画像を PC 内のミドルウェアによる展開処理を施しユーザにとって容易に理解可能なパノラマ映像を配信する。

PAL レンズを用いての全方位映像はレンズの特性上光学ズームなどを行うことができない。しかしながら全方位映像のような広い視野を持つ映像では一部を切り出して拡大して確認するといった使い方が必要なことから特定部分のズームはソフトウェアによるデジタルズームで実現している。しかし、デジタルズームであるため大きく拡大するに従って画素が荒くなってしまふなどの問題があるので拡大率の上限を設けてある。

2.2 . 環状画像

PAL レンズを用いて撮影する全方位映像はドーナツ状の環状画像で撮影され、この画像の外円と内円の円周が展開映像の最大横幅でリング幅が最大縦幅となる。

2.3 . 提供する全方位映像

PAL レンズを装着したビデオカメラから取得した映像はドーナツ状の環状画像(図 4 環状画像)でありそのままの形状では視覚的に理解しにくいので次に示すようなパノラマ画像や

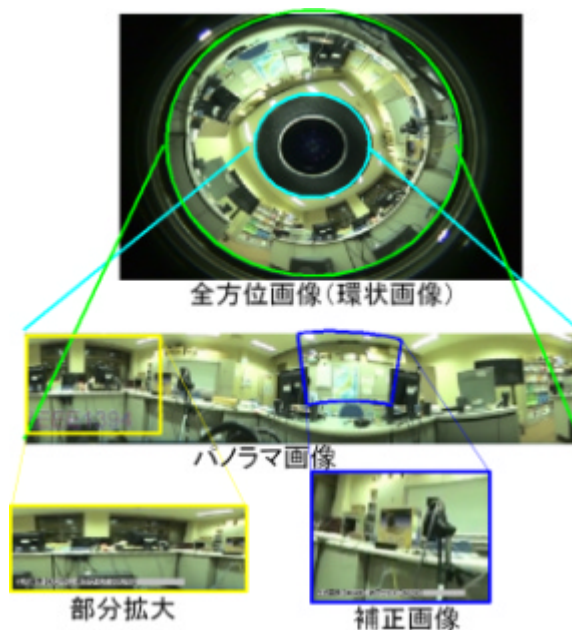


図 4 全方位映像

部分拡大画像や補正画像に展開する変換処理を行う。

パノラマ画像は PAL レンズの視野をすべて利用した 360 度のパノラマ画像である。

部分拡大画像はパノラマ画像を部分的に切り出し拡大処理を行った映像を提供する。ユーザは必要な部分の拡大映像を得ることができる。

補正画像は全方位映像の歪みを取り除き通常の単方位カメラで撮影したような映像の提供を行う。PAL レンズで撮影した全方位映像はその特性上ゆがみが生じておりカメラから得られる環状画像の内側と外側では撮影されている面積が異なっている。そのため単純に展開処理を行った場合、環状画像の外側の部分は詰まったようになり、内側の部分は横に引き伸ばしたような映像へと展開されるため、このゆがみを補正しあたかも通常のカメラで撮影したかのような映像にする。

これらの展開処理はカメラから取得した環状画像の各フレームに対して行われるのでユーザに対してはこの展開処理を施した映像が提供される。

2.4. 全方位映像展開プロセス

全方位映像の展開プロセスは図 5 に示すような形で行われる。

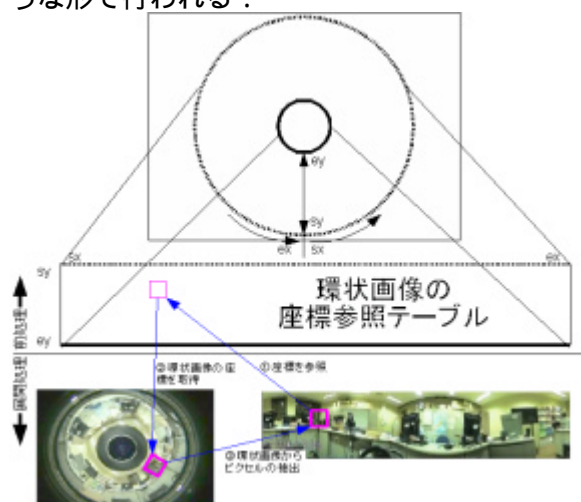


図 5 全方位画像の展開処理

1. 座標変換テーブルの作成。全方位映像

の展開処理は極座標変換により実現されているためすべてのフレームに対して極座標変換を行うと計算量が非常に多くなり効率が悪い。そこで座標変換テーブルを環状画像のサイズとレンズの視野角などの環状画像パラメータを元に生成し極座標変換の計算を削減し展開処理時間を短縮する。

2. 環状画像の補正。DV カメラで撮影した全方位映像は縦方向につぶれたような楕円形の環状画像で撮影されるため縦方向に引き伸ばし 720 × 520 ピクセルにすることで真円の環状画像となる。

3. パノラマ画像・部分拡大画像・補正画像への展開。座標変換テーブルを参照し環状画像から必要なピクセルを抽出する。

4. 補間処理。抽出したピクセルの近隣 4 ピクセルを参照し線形補間処理を適用する。これにより極座標変換による画素間の色のばらつきを低減しなるべくスムーズな展開画像にする。

5. 次回以降のフレームは 2 から繰り返す。

3. アーキテクチャ

本システムは図 6 に示すようにアプリケーション層、全方位展開層、同期処理層、セッション層、MidField System 層の 5 層から構成される。

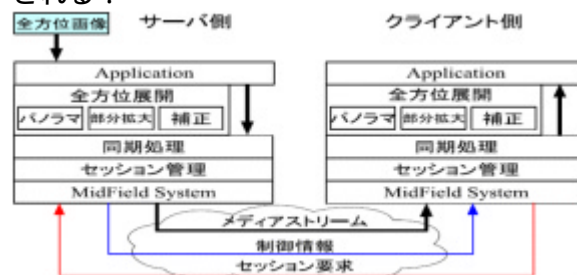


図 6 アーキテクチャ

・アプリケーション層は TV 会議や遠隔監視などのアプリケーションであり環状画像の取り込み及び表示機能を実現している。

・全方位展開層では環状画像からパノラマ、部分拡大、補正画像に展開処理を行う。

・同期処理層ではパノラマ、部分拡大、補

正画像の展開タイミングの同期をとる。

- ・セッション層では各メディアストリームを管理する。

- ・MidField System 層では後述するように、映像転送のためのミドルウェアでありネットワークを通じてメディアパケットの転送を行う。

4. MidFieldシステム

MidField System は柔軟なマルチメディア通信を実現するためのミドルウェアで、多様なユーザ環境において、ユーザの QoS 要求を考慮しトランスコーディング機能を有する通信サービスの提供を可能とする分散マルチメディアシステムである。

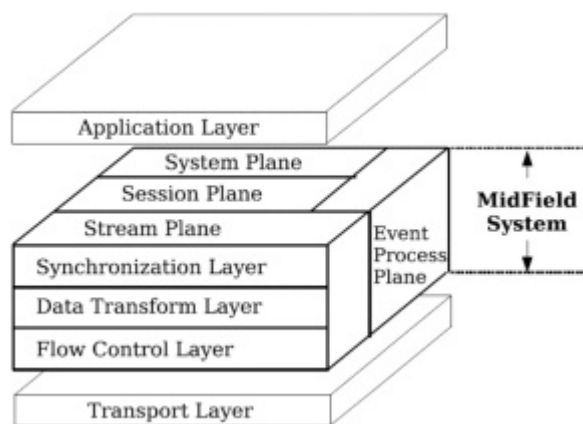


図 7 MidField システムアーキテクチャ

MidField System はトランスポート層の上位層に 3 階層 4 プレーンで構成(図 7)されておりアプリケーションに対し柔軟なマルチメディア通信を提供する。Stream Plane ではマルチメディアストリーム転送(メディアの同期, データ変換, フロー制御), Session Plane では相互通信セッション管理, System Plane では資源管理, Event Process Plane ではシステム内部イベント処理を行っている。

本システムではマルチメディアストリーム転送機能を使用して遠隔地点間で全方位映像の送受信を行う。図 8 に示すように全方位映像モジュールを MidField システムの Data Transform Layer に導入することで MidField システムで全方位映像を扱うことが可能となる。

サーバ側ではカメラから取り込まれたメディ

アストリームは MidField システムの Stream Plane の Data Transform Layer で全方位映像の展開処理及びエンコーダによりトランスコーディングが行われる。

クライアント側では受信したストリームがデコーダにより復元され展開処理が必要な場合は全方位映像展開モジュールにより展開処理が行われる。また, 展開処理が必要ない場合はデコードされた後そのまま次のレイヤーへの処理が移る。

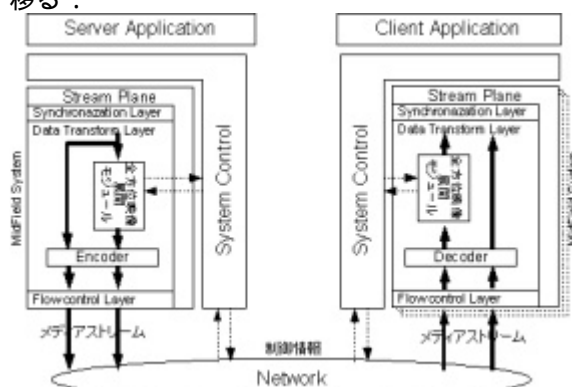


図 8 MidField システムへのモジュールの組み込み

5. ストリーム制御手法

ストリームの制御は大きく分けて 2 つの方法を用いている。1 つはクライアント展開処理型, 2 つめはサーバ展開処理型の 2 つを用いてストリームの制御を行う。

1 つめのクライアント展開処理型(図 9)はサーバ側で PC に接続された全方位映像カメラから環状画像を DV フォーマットで取得しそのままの状態 MidField システムを用いてクライアントへマルチキャストを行う。クライアントはサーバからのマルチキャストストリームを受信し環状画像を受け取り必要に応じて展開処理を行う。この方法を用いた場合は各クライアントで自由に展開映像を生成で

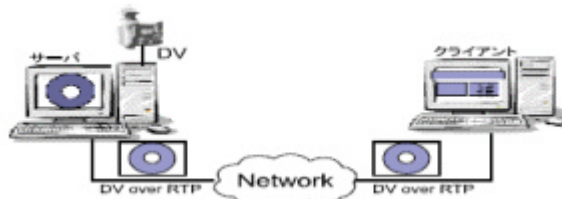


図 9 クライアント展開処理型

き圧縮処理などがないため画質もよいと言った利点があるが展開処理は多くのマシンパワーを必要とするため非力な環境では十分な品質が得られないことや DV ストリームで配信されるためネットワーク帯域を約 29Mbps 必要とすると言った欠点も考えられる。

2つめのサーバ展開処理型(図 10)はサーバ側で PC に接続された全方位映像カメラから環状画像を取得し各種展開処理を行い展開映像を生成し MidField システムのトランスコーディング機能により送信時に MPEG4 にエンコードされクライアントへマルチキャストされる。クライアントはサーバからのマルチキャストストリームを受信し展開映像を取得することができる。この方法を用いた場合はクライアントで展開処理を行う必要はなく比較的的非力な環境でも十分な品質を得ることができストリームは MPEG4 にエンコードされ配信されるため約 1.4Mbps のネットワーク帯域があれば受信することが可能である。しかし、サーバで一括して展開処理を行っているため視点制御機能に一部制限が出ることやサーバには非常に高い負荷がかかるといった欠点もある。

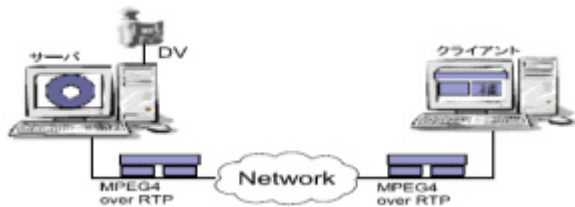


図 10 サーバ展開処理型

この 2 つの手法を組み合わせることでユーザーの環境に応じて適切なストリーム制御を行い全方位映像を提供する。

6. QoS制御

現在はこれらの制御手法に加えて CPU 利用率やネットワーク帯域を考慮した QoS 制御を導入する。QoS 制御を行うことでユーザーの環境に応じて適切な品質の映像の提供を可能にする。

CPU 利用率が高くなりすぎると全方位映像の展開処理にフレームレートの低下や通信パ

ケットの送受信に影響を与える。CPU 利用率を考慮することにより PC の負荷状態を知ることができるため CPU 利用率に応じて展開映像サイズの変更やフレームレートを変更する。

ネットワーク利用帯域は配信する映像のサイズや DV や MPEG4 や Motion-JPEG などの転送フォーマットによって変化する。接続するネットワークの帯域に応じて適切な転送フォーマットや映像サイズで提供することによりユーザーの環境に適切に対応する全方位映像環境を提供することが可能となる。

7. プロトタイプシステムと評価

プロタイプシステム(図 11)としてビデオ会議システムを構築し実験を行った。



図 11 プロトタイプ構成図

使用言語は C++(Microsoft Visual C++6.0)を使用しメディア処理と管理には DirectX9.0b (DirectShow)を用いて構築する。

実行環境 PC のスペックは CPU が Pentium4 2.8GHz でメモリを 2GByte 搭載しネットワークインターフェースカード(NIC)は 100Mbps のマシンで DV カメラ (SONY DCR-PC300K)を IEEE1394 ケーブルで接続している。

このビデオ会議システムは双方向で全方位映像と音声の送受信を行う。全方位映像の配信の形式として 1 つはサーバから環状画像を DV フォーマットでクライアントへマルチキャスト配信をする形式とサーバ側でパノラマ画像、2 つめは補正画像への展開を行った後に展開した映像を MPEG4 でクライアントへ配信する形式の 2 種類を用いてデータの測定を行っている。

DV フォーマットを用いて全方位映像の転送を行った場合は解像度が 720 × 480 の映像で

約 29Mbps の帯域を必要とし、MPEG4 で全方位映像の転送を行う場合は解像度が 360 × 240 の映像で約 1.4Mbps の帯域を必要とする。なお、音声については DV フォーマット利用時は映像と音声为一体となった形で転送され、MPEG4 フォーマット利用時は PCM フォーマットで送信され約 64kbps の帯域を必要とする。

評価として遠隔ビデオ会議システムをアプリケーションとして構築した。これは全方位映像と音声を双方向で共有し、それぞれに相手側のパノラマ画像、部分拡大画像、補正画像を表示する。そして、画質やフレームレートや遅延時間を会議システムとして性能評価を行った結果、フレームレートについてはパノラマ画像(800x200 ピクセル)は 8fps、部分拡大画像(360x240 ピクセル)は 9fps、補正画像(360x240 ピクセル)は 9fps で展開処理が可能であった。

遅延時間に関しては音声と映像共に約 0.8 秒程度であった。この遅延は DV フレームのパケット化と DV フレームの再構築と各画像への展開処理と展開処理の同期処理に伴いこの程度の遅延が発生していると考えられる。

画質はカメラから被写体の距離が 1m 以内であれば人の顔が識別可能だった。この点については DV の解像度が 720x480 ピクセルであり環状画像を展開処理した場合の有効解像度が約 1/3 になるためと考えられる。しかしながら、会議システムとして用いた場合、複数人を同時撮影可能であり視点制御もソフトウェアで行うためスムーズな視点切り替えが可能である。

8. まとめ

本論文では全方位映像配信利用のためのモデルウェアの提供とプロトタイプシステムによる会議システムの構築と評価について述べた。

本システムでは全方位映像の利用環境の提供を行い MidField システムと組み合わせてユーザ

の環境に応じて適切にトランスコーディングを行い効率的に全方位映像を配信する。

プロトタイプシステムでは会議システムとして評価を行いフレームレートや画質の評価を行った。

9. 今後の課題

今後は QoS 制御の導入や IPv6 への対応、画像処理の最適化を行いフレームレートの改良を行う。画質についても周波数フィルタなどを組み合わせることにより画質を向上させる予定である。また、1000 万画素を超える CCD も製品化されており、解像度の向上も期待される。

全方位映像は会議システム、監視システム、ロボットの視覚センサーのような多くの分野に応用が考えられる。今後はこのような応用アプリケーションに利用していく。

参 考 文 献

- [1] Koji Hashimoto and Yoshitaka Shibata "Design of a Middleware System for Flexible Intercommunication Environment", International Conference on Advanced Information Networking and Applications (AINA2003), pp.59-64, 2003
- [2] 八木康史, 横矢直和 全方位ビジョン: センサ開発と応用の最新動向, 情報処理学会誌 vol.42 No.SIG13pp.1-18, 2001
- [3] 久保田幸雄 デジタルビデオ読本 オーム社, 1995