

M-085

全方位画像に対応したストリーミング再生プレイヤーの開発 Developing Streaming Player for Hyper Omni Vision

西尾 佳祐[†] 岩井 儀雄[†] 長原 一[†] 谷内田 正彦[†]
Keisuke Nishio Yoshio Iwai Hajime Nagahara Masahiko Yachida

概要

動画のストリーミング配信技術が向上するにつれて、クライアント側で高解像度の映像を自由な視点で見られる動画への需要が高まってきている。本論文では360度の情報を一度に獲得できる全方位視覚センサ [1] を用いて高品質で高視野な動画を生成し、視点情報に基づいて必要な動画のみを配信するシステムについて述べる。本研究ではネットワークを通じて視点情報をサーバとやり取りし、高解像度でユーザが見たい領域をインタラクティブにストリーミング再生するプレイヤーを開発する。

1. はじめに

近年、ブロードバンドの普及やパーソナルコンピュータの性能向上に伴い、ストリーミング動画配信が可能となってきた。インターネット環境も、ADSL(Asymmetric Digital Subscriber Line) から光ファイバ(Fiber To The Home)へと移行しつつあり、より大容量のデータ転送が可能となると予想される。これにより、高精細かつ広視野な動画配信や、インターネットを介したサーバとクライアントのインタラクティブな動画配信のニーズが高まってきている。

我々は、高解像度の全方位動画を多チャンネル配信するための受信クライアントの構築を目的とする。しかし、本研究で想定する高解像度全方位動画は、既存のプロトコルで配信するにはデータ量が大きすぎるので、提案システムでは全方位動画を複数の低解像度動画に分割し、クライアントとサーバが視点情報をやり取りすることで、ユーザの見たい方向の動画のみを複数の伝送路により多チャンネルで配信し、表示することにより実現する。本研究では、このようなシステムを実現するために、既存の通信プロトコルを拡張し、クライアントがユーザの見たい部分をサーバに指示し、受信した複数の動画を合成し、ユーザに提示できるようなストリーミングプレイヤーの開発を目指す。

2. 想定システム

2.1 全方位高精細動画配信システム

本研究で目標とするシステムの概略図を図1に示す。複合センサカメラ [2, 3] によって撮像された高解像度低フレームレートの動画と低解像度高フレームレートの動画の二つの動画から、高解像度高フレームレートの動画 [4, 5] を生成する。生成された高解像度高フレームレートの動画はサーバに分割保存され、クライアントに配信される。ここでクライアントに配信する際、ユーザが見たい領域のみを多チャンネルで配信する。クライアント側では、多チャンネルで配信された複数のの部

分画像を合成して一つの高解像度高フレームレートの全方位画像を作成し、それを幾何変換によって透視投影画像に変換してユーザが見たい領域をモニタに出力する。ユーザが別の視点を見たい場合は、変更した視点情報をサーバ側に伝え、サーバはそれに対応した部分画像を再び多チャンネルで受信することによってユーザが見たい部分を見られるようなインタラクティブな動画配信システムを構築する。

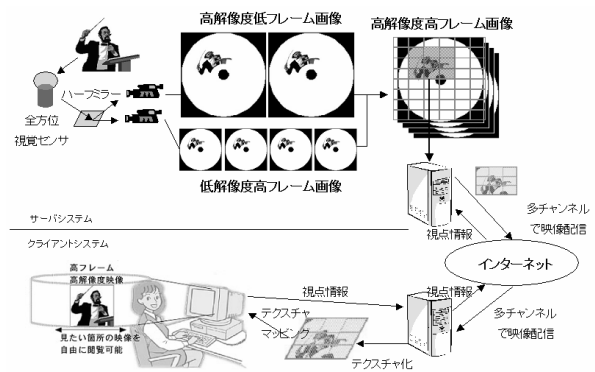


図 1: システム概略図

2.2 インタラクティブ配信

本節では、全方位動画のインタラクティブ配信の有用性について述べる。インターネットの通信帯域は有限であるため、図2に示すように全方位動画を全て配信してしまうと、ユーザの見えていない部分が必ず存在し、通信帯域の無駄になる。そこで視点情報のやり取りによって、クライアントは見ている方向の高解像度動画のみを受信する。このサーバ・クライアント間のインタラクティブ配信によって大幅に通信帯域を節約できる。

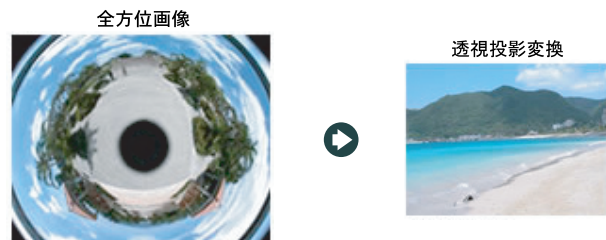


図 2: 再生結果

2.3 多チャンネル化

前節でインタラクティブ配信について述べたように、サーバからはユーザの見たい方向の高解像度動画が

[†]大阪大学大学院基礎工学研究科

部分配信される。本研究では、 4096×4096 画素の全方位高解像度動画の配信を目指している。しかし、本研究で想定している高解像度の部分動画を一つの伝送路で配信できる通信プロトコルが存在しない。現在インターネット配信などで多く利用されている MPEG-4 でも、メインプロファイルのレベル 4 で最大 1920×1088 画素 (HDTV) と規定されている [7, 8]。

そこで、図 3 に示すように、全方位動画を複数の低解像度動画に格子状に分割し、配信サーバの動画データベースに蓄積する。サーバはクライアントに要求されたユーザの見たい方向 ($\alpha_1, \alpha_2, \alpha_3, \alpha_4$) に対応する動画のみを複数の伝送路を用いて多チャンネルでクライアントに転送する。本研究では、サーバから配信される全方位動画を多チャンネルで受信するために、通信プロトコル (RTSP) を改良し、多チャンネル受信クライアントを構築する。

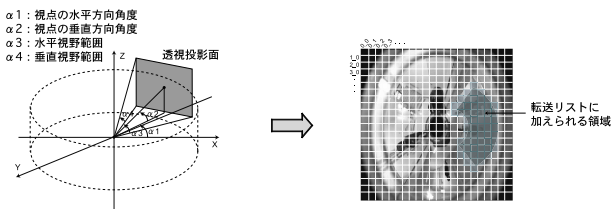


図 3: 視点情報と多チャンネル配信

3. システム設計

この節では想定システムで実際に動画を配信するために、具体的なプロトコルの改良とクライアント側のプレイヤーの開発について述べる。

3.1 プロトコルの改良

一般的な動画配信システムでは、単一のクライアントは単一のチャンネルで受信することになり、多チャンネルで受信することはできない。本節では、多チャンネルインタラクティブ配信を実現するための拡張 RTSP について述べる。本研究の多チャンネル動画配信に対応する RTSP も基本的に既存の流れに基づいているが、全方位動画を多チャンネルで配信するために独自拡張を行う。RTSP (Real Time Streaming Protocol) は、Real Networks 社、Netscape Communications 社、Columbia 大学が共同で策定したセッションコントロールプロトコルで、IETF により RFC2326 として勧告されている [9]。本研究において拡張した RTSP の流れを図 4 に示す。拡張 RTSP では、新たに DESCRIBE (全体的な記述情報) と SET_PARAMETER を追加し、さらに視点情報に対応する複数のファイルに対して、DESCRIBE, SETUP, PLAY を繰り返すことで複数の伝送路を確保し、多チャンネル受信を実現する。

また、クライアントは再生の途中で視点を変更すると、再び SET_PARAMETER を介して、対応する低解像度動画のファイルリストを受信する。新たに必要になったファイルに対して、再びサーバに DESCRIBE, SETUP, PLAY を請求する。このような流れで、新たな視野範囲に対応する動画ファイルを受信することができる。

- DESCRIBE (全体的な記述情報)

一般的なストリーミングサーバは、クライアントの DESCRIBE 請求により、クライアントが受信したい単一のメディアファイル記述情報をクライアントに回答する。提案する拡張プロトコルでは、まずクライアントは受信したい全方位動画のディレクトリを指定して DESCRIBE 請求を行うことにより、全体的な記述情報をサーバから得る。記述情報の中には、透視投影変換に必要な全方位センサのパラメータ、全方位動画を構成する分割された全ての低解像度動画のファイルリスト、各低解像度動画の位置とサイズ情報が記述されている。

- SET_PARAMETER

提案する拡張プロトコルでは、クライアントは全方位動画の見たい部分の視点情報を SET_PARAMETER を介してサーバに送る。視点情報は、“viewPoint: Angle1_Angle2_Angle3_Angle4” の形でサーバに送信する。Angle1, Angle2, Angle3, Angle4 の 4 つのパラメータはそれぞれ、視点の水平方向角度 α_1 、視点の垂直方向角度 α_2 、視野の水平範囲角度 α_3 、視野の垂直範囲角度 α_4 を意味する。サーバはその視点情報をもとに対応するファイルを選択する。サーバからはその視点情報に対応するファイルリストが返される。



図 4: 提案システムの RTSP の流れ

3.2 実装

本研究では MPlayer+liveMedia[10] をベースプログラムとした。MPlayer+liveMedia の通信モジュールと映像生成モジュールを中心に、多チャンネル受信機能を実装し、配信された複数の動画の合成、同時再生とキーボード操作による視点変更を可能とするプレイヤーを開発する。多チャンネル配信において、接続開始から MPlayer の通信モジュールと映像生成モジュールの処理の流れを図 5 に示す。

3.2.1 通信モジュール

MPlayer+liveMedia では、ストリーミングサーバとのやり取りを open_stream, demux_open という 2 種類の

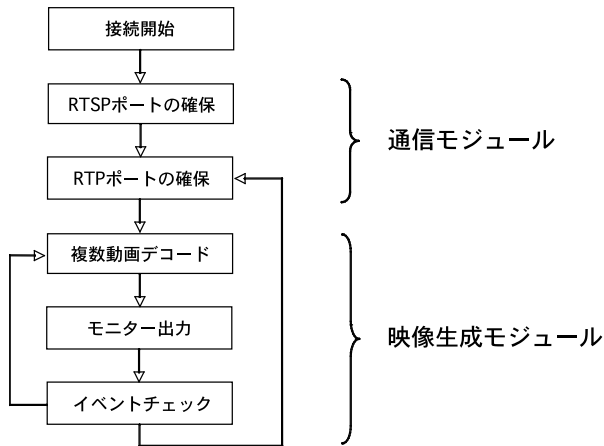


図5: プレイヤの処理の流れ

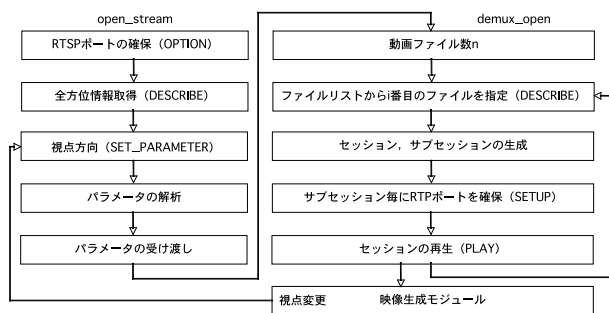


図6: 多チャンネル受信の処理フロー

通信モジュールが担っている。図6に通信モジュールでの処理の流れを示す。モジュールには、各 RTSP Method に対応した関数が用意されており、その関数を拡張する。本研究で提案する多チャンネル動画配信に必要なとされる通信モジュールの役割は、まず RTSP でのやりとりを行なうためのポートを確保し、サーバから受け取った全方位パラメータや視点情報に対応するファイルリストを解析する。次に多チャンネルで受信するために、複数の RTP ポートを各分割された低解像度に対して確保し、サーバから受け取った情報を映像生成モジュールに渡す。また、視点の変更がある場合は、映像生成モジュールからその視点情報を受け取り、サーバに送信するというものである。

また RTSP にはコンテンツを制御するためにセッションという概念が提案されており、この通信モジュールでは1セッションに1ファイルの情報を格納し、サーバから受信した各動画をこのセッションによって管理する。また図6にあるサブセッションとは、セッション中の音声と映像のことを指しているだけである。つまり、セッションの削除、追加を行なうことで複数の RTP ポートを破棄したり、確保する。これにより、必要な動画ファイルに対して、セッション単位でセットアップ、再生を繰り返すことで多チャンネル受信を実現することができる。

3.2.2 映像生成モジュール

MPlayer は OpenGL[11] のモジュールをサポートしているので、GL のテクスチャマッピング機能を用いて全方位画像を透視投影変換して表示するようにした。テクスチャマッピングとは画像を格子状に切り出し格子点のみを変換に応じてマッピングを行い、格子点間は線形補間によって近似変換する手法であり、全方位画像を透視投影画像に変換するような非線形変換に対して有効である。

また多チャンネル配信によって配信された複数ファイルを再生するために、複数のファイルを読み込んで同時に再生できるように改良した。この流れは図7のようになっている。モニタに出力する前に受信した複数の動画を順番に1フレームずつデコードし、各動画の描画位置をずらしてテクスチャマッピングを行い、1枚のテクスチャに合成し表示することにより複数ファイルの同時再生を可能にした。

視点情報の変更に関しては、キーボードの矢印キーに対応させる。ここでいう視点情報とは、図3に示すように水平方向角度 α_1 と画角 α_2 で表される視点方向と、画角 α_3 と α_4 で表される視野のことである。MPlayer ではキーボードイベントが発生すると、イベントチェック関数が呼び出される。本研究では、MPlayer にキーボードイベントに対応する関数が用意されているので、これを変更することで視点情報を書き換える。そして、通信モジュールと描画モジュールはルーチンの中でこの値を外部変数としてチェックし、値の変更があれば通信モジュールは SET_PARAMETER でサーバに視点情報に対応した部分画像をリクエストし、新たに必要になった動画ファイルの RTP ポートが確保され、OpenGL のモジュールは再びデコード・テクスチャマッピングを行ない、クライアントが望む領域の部分画像を提供する。この処理の流れを図8に示す。プログラム変更によって動画の早送り、早戻しをなくした。本研究では動画ファイルの再生ではなくリアルタイムのストリーミング再生を想定しているので、時間の操作は必要ないと考え、クライアントが直感的に視点変更操作を行えるようなキーに割り当てた。

3.3 時間同期

視点変更されると、サーバとクライアントは SET_PARAMETER によって新たに必要になる動画と不要となる動画を確認し、不要な動画は破棄 (TEARDOWN) され、新たに必要な部分画像が DESCRIBE, SETUP, PLAY される。このとき、今まで再生されていた動画と新しく再生される動画は時間同期をとる必要がある。これはサーバ側で送信するパケットの番号を記憶しておき、新たな動画を送信する時はそのパケット番号に合わせて途中からパケットを送信することで、クライアント側においても各チャンネルから同じ番号の RTP パケットを同時に受信することができる。つまり、サンプル時間が同じであれば、同時刻の動画を復元でき、同期が保証される。

4. 再生結果

本節では、実際にサーバと接続し、ストリーミング再生した結果を示す。図9は、左図がサーバ側に格納して

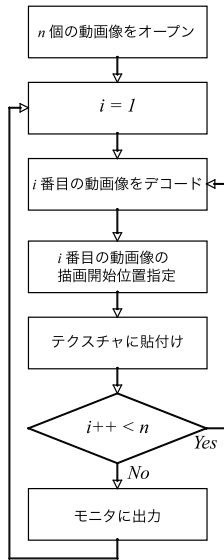


図 7: 複数ファイル再生

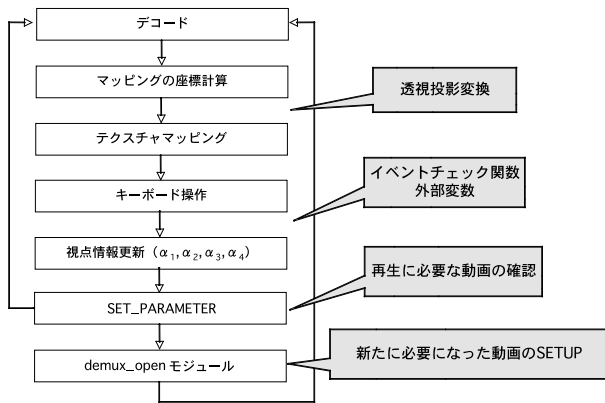


図 8: 視点変更

ある全方位動画像で、右図は viewPoint: 1_35_9_9 をサーバに請求し、9 チャンネル受信して透視投影画像をモニターに出力した結果である。

5. おわりに

本研究では視点情報に基づいて高解像度動画を部分的に配信するシステムと、それを実現するための拡張プロトコルと分割された複数の動画を合成し、再生できる多チャンネルストリーミング再生プレイヤーを提案した。今後の課題として、クライアントの性能評価、拡張プロトコルの改良、サーバ・クライアントのインタラクティブ性の向上などが挙げられる。

謝辞

本研究は、独立行政法人情報通信機構「民間基盤技術研究促進制度」および、科学研究費補助金基盤研究(A)「3次元高精細全方位動画像の獲得と提示」(課題番号 18200012)の援助を受けた。

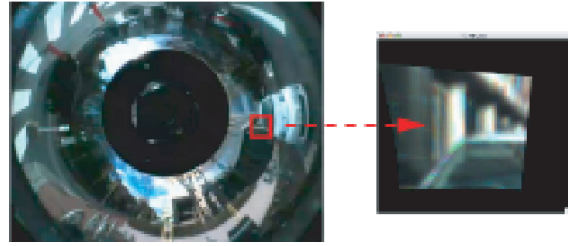


図 9: 9 チャンネル再生

参考文献

- [1] 山澤, 八木, 谷内田: “移動ロボットのナビゲーションのための全方位視覚センサ hyperomni vision の提案”, 電子情報通信学会論文誌 D-II, Vol. J79D-II, No. 5, pp. 698-707 (1996).
- [2] 重本, 星川, 長原, 岩井, 谷内田, 田中: “時間的・空間的分解能の異なる複合センサカメラシステム”, CVIM-147, pp. 88-90 (2005).
- [3] 松延, 星川, 重本, 渡邊, 長原, 岩井, 谷内田, 田中: “複合センサカメラを用いた高解像度動画像の撮像・提示システム”, 画像の認識・理解シンポジウム 2005, Vol. 2005, No. 7, pp. 1602-1603 (2005).
- [4] 渡邊, 岩井, 長原, 谷内田: “時空間周波数の異なる画像列からの高解像度動画像の合成”, 情報科学技術レターズ (FIT2004), Vol. 3, No. L1-004, pp. 169-172 (2004).
- [5] 松延, 長原, 岩井, 谷内田, 田中: “モーフィングによる高解像度高フレームレート動画像の生成”, 電子情報通信学会技術研究報告, Vol. PRMU-2004, No.178, pp. 85-90 (2005).
- [6] 山澤, 石川, 中村, 藤川, 横矢, 砂原: “Web ブラウザと全方位動画像を用いたテレプレゼンスシステム”, 電子情報通信学会論文誌 (D-II), Vol. J88-D-II, No. 8, pp. 1750-1753 (2005).
- [7] 藤原, 安田: 標準ブロードバンド+モバイル MPEG 教科書, アスキー (2003).
- [8] F. Pereira and Touradj Ebrahimi: THE MPEG BOOK, Prentice Hall (2002).
- [9] H. Schulzrinne, A. Rao and R. Lanphier, “Real Time Streaming Protocol (RTSP)”, RFC 2326, Internet Engineering Task Force (1998).
- [10] MPlayer - The Movie Player, <<http://www.mplayerhq.hu/DOCS/HTML-single/en/MPlayer.html>>
- [11] Manson Woo, Jackie Neider, Tom Davis: OpenGL プログラミングガイド, アジソン・ウェスレイ・パブリッシャーズ・ジャパン株式会社, 1997