平成17年度 研究開発成果報告書

全方位高解像リアルタイム動画入力と その配信システムに関する研究開発

委託先: ㈱映蔵

平成18年4月

情報通信研究機構

平成17年度 研究開発成果報告書

「全方位高解像リアルタイム動画入力とその配信システムに関する研究開発」

日	1/1-
	- 1八

1	研究開発課題の背景	. 2
2	研究開発の全体計画 2-1 研究開発課題の概要 2-2 研究開発目標 2-2-1 最終目標 2-2-2 中間目標 2-3 研究開発の年度別計画	. 2 . 8 . 8 . 8 . 8
3	研究開発体制 3-1 研究開発実施体制	1 0 1 0
$ \begin{array}{c} 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\ 4 \\$	 研究開発実施状況 -1 複合センサカメラの研究開発 4-1-1 はじめに 4-1-2 複合センサカメラの試作 4-1-3 複合センサカメラのさ作 4-1-4 評価実験 -2 リアルタイム高解像度動画像の作成 4-2-1 周波数空間アプローチによる高解像度動画像の作成 4-2-2 モーフィングアプローチによる高解像度動画像の作成 -3 高解像度全方位カメラ用光学系の設計 -4 全方位高解像度リアルタイム動画像入力記録方式の開発 4-4-1 複合センサカメラシステムを利用したリアルタイム動画像記録 4-4-2 書き込み速度の最適化 -5 高解像度動画配信用ソフトウェアの開発 4-5-1 受信クライアントソフトウェアの開発 4-5-2 配信サーバの開発 -6 ライブ入力,配信システムの開発 -7 マルチアクセス時の性能評価 	$1 1 \\ 1 1 \\ 1 1 \\ 1 2 \\ 1 4 \\ 1 5 \\ 1 6 \\ 1 9 \\ 2 7 \\ 2 8 \\ 2 9 \\ 2 9 \\ 3 0 \\ 3 6 \\ 4 1 \\ 4 3$
4 4 5	 4-7-1 マルチアクセス実験環境 4-7-2 配信能力の評価 -8 統合システムの試作と評価 -9 総括 参考資料・参考文献 	43 44 47 48

5-1 研究発表・講演等一覧

1 研究開発課題の背景

ブロードバンド化の急速な伸びにより、高品質な映像配信への需要は急速に拡大してき ている。研究担当者らが世界に先駆け開発した全方位カメラを用いると、周囲 360°(全 方位と呼ぶ)のシームレスな動画像がリアルタイムで撮影でき、非常に臨場感の高い映像 を得ることができる。この全方位映像をインターネット配信すれば、閲覧者(クライアン ト)は全方位映像を見ることができると同時に、各ユーザが各自の見たい方向の視野を見 ることができるので、臨場感、現実感の高い映像を見ることができる。このため、全方位 映像の入力と処理技術は世界的にも研究が活発化し、市販されるようにもなってきた。し かし、全方位カメラの欠点はある部分に着目したときの解像度が劣るという点である。す なわち、通常のカメラが水平画角 40 度前後であるのに対し、全方位カメラでは1台で周囲 360 度を撮影していることからある部分に着目したときの空間分解能は 1/9 程度しかない。 このため、通常の解像度の全方位カメラではその用途が限られてくる。たとえば、遠隔監 視を例にとってみても、人がいることは分かってもその人の顔を判別できるまでの解像度 がない。この解像度の低さが実用化の妨げとなっていた。

本研究では、全く同一の視野を持った2種類の全方位動画像、a)高解像度(8000 x 8000 画素)だが時間的には粗い(1フレーム/秒)画像と、b)通常の解像度(640x480 画素)だ が時間的に密な(30 フレーム/秒)画像を同時に撮影できる全方位カメラを開発し、これ ら2つの全方位動画像から、空間的に高解像度で、しかも時間的にも密な全方位動画像 (8000×8000 画素、30 フレーム/秒)を作成する技術を開発する。また、これら2種類の 全方位画像を圧縮してインターネットにより複数の閲覧装置に配信するサーバシステムと 配信された2種類の全方位動画像から閲覧者側で高解像度(8000x8000 画素)で、かつ実 時間(30 フレーム/秒)の動画像を作成する技術を開発する。さらに、エジプトのピラミ ッドや日本の古墳等の高精細なディジタルアーカイブ化とその配信による実証実験を行い、 本提案技術の有効性の検証とその周知を行う。現在、実用化されている競合技術として(株) 立山科学、iPIX 社、BeHere 社の全方位入力装置がある。しかし、いずれもNTSC(40 万画 素)レベルの解像度で、現行の全方位カメラでは、画質の点で問題があり商品力に欠ける。 また、研究開発段階のものもハイビジョン(200 万画素)や130 万画素カメラ対応のもの などが最上位で、今回提案する 6400 万画素カメラとは全く画質が異なり、競合の対象とは ならない。

一方、多数のカメラを並べリアルタイムで全方位の画像を取り込むシステムも提案され ているが、カメラ間のキャリブレーションと映像のつなぎ目のずれが問題とされている。 また提案技術である 6400 万画素相当の解像度を出すためには、通常の 40 万画素のカメラ であれば 160 台必要となり、システム規模が著しく大きくなるという問題もある。従って、 携帯性も意識した提案システムとは競合しないものと考える。

2 研究開発の全体計画

2-1 研究開発課題の概要

ブロードバンド化の急速な伸びにより、高品質な映像配信への需要は急速に拡大してきている。高品質として広視野、高精細、高速が考えられるが、本研究では、広視野としてその究極である 360°の全方位、高精細として 8000x8000 の高解像度、高速としてリアルタイム(30 フレーム/秒)の動画を入力し蓄積できるカメラの開発と、それを高能率で配信するシステムの開発を行う。

研究担当者らが世界に先駆け開発した全方位カメラを用いると、周囲 360°のシームレス な動画像がリアルタイムで撮影でき、臨場感の高い映像を得ることができる。しかし、従 来の全方位カメラは、リアルタイムで撮像することが可能であるが空間解像度が低いのが 欠点であった。一方、ディジタルカメラの発展により高解像度カメラも市販されるように なってきたが、サンプリング間隔が長くリアルタイムで撮像することはできないという欠 点がある。そこで、本研究では、同一の視野を持った2種類の全方位動画像;a)高解像度 だが時間的には粗い画像と、b)通常の解像度だが時間的に密な画像を同時に撮影できる全 方位カメラを開発し、これら2つの全方位動画像から、空間的に高解像度で、しかも時間 的にも密な全方位動画像を作成する技術を開発する。これにより、全方位カメラの欠点で あった解像度の低さの問題が解決され、全方位カメラの真の実用化が可能となる。 また、これら2種類の全方位画像を圧縮してインターネットにより複数の閲覧装置に配信 するサーバシステムと、配信された2種類の全方位動画像から閲覧者側で高解像度で、か つ実時間の動画像を作成する技術を開発する。これにより、遠隔からのモニタリング、セ キュリティ、遠隔会議、遠隔医療、遠隔教育、コンサートやスポーツなどのインタラクテ

ィブ放送などへの応用が実現可能となる。図1に全体システムのイメージ図を示す。



図1 全体システムのイメージ図

(1) 複合センサカメラの試作:

従来、動画入力のできるビデオカメラとしては、NTSC クラス(640x480 画素、30 フレーム /秒)のカメラが長年使われてきた。しかし、NTSC クラスの画像サイズ(640x480 画素)で は、広視野で画像を撮像すると解像度が不足するという欠点があった。一方、近年のディ ジタルカメラの急速な発展と高機能化により、高解像度のカメラが次々と現れ、現状で 4000x4000 画素を持つ高解像度のカメラも市販されている。しかし、高解像度になると、 フレームレートは低下し、リアルタイムで撮像することはできないという欠点がある。た とえば、4000x4000 画素のセンサの場合、1秒間に1フレームの速度である。実時間入力 (30 フレーム/秒) できるものは NTSC(640x480 画素) クラスの画像サイズがほとんどであ る。本研究では、焦点(レンズ中心)を合わせることにより、同一の視野を持った2種類 の動画像:a)高解像度(4000 x 4000 画素) だが時間的には粗い動画像(1フレーム/秒) と、b) 通常の解像度(640x480 画素) だが時間的に密な動画像(30 フレーム/秒) から、空 間的にも高解像度で、しかも時間的にも密な動画像(4000x4000 画素、30 フレーム/秒)を 作成する技術を開発する。このように、高解像度の実時間全方位動画像を撮影できるカメ ラは他に例がなく本研究独自のものである。まず、その第一歩として、図2に示すように、 焦点(レンズ中心)を合わせることにより、同一の視野の動画像を2種類の異なった解像 度で撮影し蓄積するカメラを試作する。



(2) リアルタイム高解像度動画像の作成:

上で試作した複合センサカメラから得られる2つの動画像から、空間的にも高解像度で、 しかも時間的にも密な動画像(4000x4000 画素、30 フレーム/秒)を作成する技術を確立す る。すでに、理論的な解析は終えているので、そのソフトウェアを試作し、実際に、高解 像度だが時間的には粗い動画像と、通常の解像度だが時間的に密な動画像の2つの動画像 から空間的にも時間的にも密な動画像を作成できることを実験的に確かめる。

高解像度動画像の作成には、2台のセンサからの動画像を2つの3次元時空間データとし て扱い、3次元時空間で統合することにより高解像度化を図る。具体的な時空間高解像度 処理ブロックを図3に示す。低解像用カメラと高解像用カメラの2つの映像信号を、空間 オーバーサンプリング、ディレイ、高解像化処理を行うことにより時空間周波数の広帯域 化を行う。図4に示すように、低解像カメラは空間周波数は低いが、フレームレートは高 いので時間周波数は高いという特徴がある。また、高解像カメラは空間周波数は高いがフ レームレートは低いので時間周波数は低いという特徴がある。高帯域化を行うには、図4 における右上領域まで有効信号成分を含ませることである。通常、これらの領域には信号 の折り返し成分(エイリアス成分、ノイズ)が含まれているので、オーバーサンプリング によりエイリアス成分をより高周波に移動させ、有効信号成分を含むことができるように する。この処理は、図3における空間オーバーサンプリングとディレイ(時間方向のオー バーサンプリング)で実現する。図3における高解像化処理は、図5のような周波数成分 を持つ映像信号を混成する処理となる。左上と右下の周波数領域は、それぞれのカメラの 信号成分を利用する。左下の信号成分は、低解像カメラと高解像カメラの2つの有効信号 成分を本来含んでいるので、この信号にノイズ低減処理を施し、高品質映像信号を生成す る。右上の領域はオーバーサンプリングによる疑似信号成分であるが、この部分にもノイ ズ低減処理を行うことで、最終的に時空間高周波映像を生成する。





図4 各カメラの周波数帯域 図5 時

図5 時空間高解像度化後の周波数帯域

(3) 高解像度全方位カメラ用光学系の設計と全方位高解像度リアルタイム動画像入力記録 方式の開発:

上で作成した複合センサカメラ用に全方位ミラーを設計し、全方位複合センサカメラを試 作する。また、このカメラから得られる2つの全方位動画像から空間的にも高解像度で、 しかも時間的にも密な全方位動画像(4000x4000 画素、30 フレーム/秒)を作成する技術を 開発する。研究担当者らが世界に先駆け開発したリアルタイム全方位カメラを用いると、 周囲 360°のシームレスな動画像がリアルタイムで撮影でき、臨場感の高い映像を得るこ とができる。また、この全方位画像入力に研究担当者らの提案した双曲面ミラーを用いる と、双曲面は焦点を持つことから、パノラマ画像や透視変換画像に精度良くかつ容易に変 換できる。このため、世界的にこれを用いた応用研究がなされ、すでにNTSC(640x480 画素) クラスの画像を対象とした全方位カメラは市販されている。しかし、全方位カメラの欠点 は、ある部分に着目したときの解像度が劣るという点である。すなわち、通常のカメラが 水平画角 40°前後であるのに対し、全方位視では1台のカメラで周囲 360°を撮影してい ることから、ある部分に着目したときの空間分解能は 1/9 程度しかないため、通常の全方 位カメラではその用途が限られてくる。たとえば、遠隔監視を例にとってみても、人がい ることは分かってもその人の顔を判別できるまでの解像度がない。この解像度の低さが実 用化の妨げとなっていた。

この問題を解決するための高解像度化には高解像度のセンサを用いることが当然考えられる。しかし、すでに述べたように、非常に高価格なものを除いては高解像度になるとフレームレートは低下し、実時間入力できるものは NTSC (640x480 画素) クラスの画像サイズがほとんどである。このため、本研究では、上で開発した複合センサカメラを入力カメラとして用いることにより、全方位の高解像度リアルタイム動画像入力記録システムを開発する。

全方位画像の解像度を上げるには、センサの解像度を上げると同時に全方位画像を撮影す るための光学系(主にミラー)の解像度も上げる必要がある。しかし、我々の提案した双 曲面で代表される単一のミラーを用いた全方位カメラの場合、非点収差、コマ収差、球面 収差などの影響を受け、像にぼけが生じることが原因で、解像度を上げるにはミラーを大 型化する必要があり、カメラとしての携帯性に問題が生じる。この問題を研究担当者らが 新たに考案した二回反射全方位光学系(特許)を用いることで解決する。二回反射光学系 の例を図6-b~6-dに示す。図6-bは双曲面鏡と平面鏡を用いた構造で、平面鏡により 光路を折り曲げることで、図6-aの一枚反射光学系に対して双曲面鏡とカメラ間の距離を 約1/2にできるため小型化に有利である。また,図6-c,dはそれぞれ双曲面-楕円面と 放物面- 放物面の組み合わせの凹凸面鏡ペアによる光学系である。二枚の曲面を用いるこ とで設計自由度が上がるため、更なる小型化が期待できる。図6-b~dの二回反射光学系 は、いずれも図6-a同様に単一視点を保持しており、従来の双曲面鏡を用いた全方位セン サと同様に、自由視点の画像やパノラマ画像に無ひずみで変換可能である。これら二回反 射光学系は、これまでの試作ならびにシミュレーション実験から、単一ミラーを用いた場 合の数分の1の大きさでほぼ同程度の光学系が実現できている。これにより、実用性とい う点では重要となる、小型軽量で携帯性のある高解像度の全方位カメラが初めて実現可能 となる。

また、今回対象とする CCD センサは、9µmのセルサイズの 4000×4000 画素を想定してい るため、従来設計してきたものに比べ、スペックが1ランク上となる。また、ミラーが2 枚になることで設計自由度が上がるので、最適解の計算には、専用のシミュレーションプ ログラムを作成し、並列計算による高速化が必要である。このシミュレーションは、研究 担当者らの所属機関で開発したシミュレータを利用し行う。また試作した全方位カメラの 解像度分布などの特性評価もこのシミュレータを用いて行う。

こうして試作された複合双曲面ミラーと複合センサカメラを用いて全方位高解像度リアル タイム動画像入力記録カメラを作成する。また、このカメラから得られる2つの全方位動 画像から、空間的にも高解像度で、しかも時間的にも密な全方位動画像(4000×4000 画素、 30 フレーム/秒)を作成する技術を確立する。このように、高解像度の実時間全方位動画 像を撮影できるカメラは世界的に見て他に例がない。なお、開発当初は、現状で入手でき る 4000×4000 画素の解像度を持つセンサで開発を行い中間目標とするが、2、3年後には 8000×8000 画素(6400 万画素)のセンサが開発されると思われるので、最終年度までには 8000×8000 画素の解像度を持つ全方位のリアルタイム動画入力記録装置を開発する予定で ある。





(4) 高解像度動画配信用ソフトウェアの開発:

蓄積された2種類の全方位画像を圧縮してインターネットにより複数の閲覧装置(クライ アント)に配信するサーバシステムと、配信された2種類の全方位動画像から閲覧者側で 高解像度(8000×8000 画素)かつリアルタイム(30 フレーム/秒)の動画像を作成する技 術を開発する。また、全方位画像のうち、閲覧者の要求する視野の部分画像のみを切り出 しそれらを能率良く配信し、クライアント側で動画像表示するシステムを実現する。

2種類の全方位動画像を伝送するため、圧縮した映像をネットワーク上に2チャンネル 分の帯域幅を使用して伝送する。1 チャンネルは全方位の低解像度動画像でこれは全クラ イアントにブロードキャストすることでネットワーク全体のトラフィックを少なく抑える。 また、もう1 チャンネルは、各クライアントの要求に応じた解像度の部分画像を転送する ピアツーピア通信に利用する。この各クライアントに専用の通信チャンネルを設けること で、各クライアントとインタラクティブにデータをやり取りすることができる。特に、部 分画像の要求がないときは、全方位高解像度画像を低フレームレートで伝送することで、 ネットワークのトラフィックを抑えつつユーザの要求にリアルタイムに応答できるクライ アントシステムを構築する。さらに、伝送信号に優先度を設定し通信チャンネルの輻輳を 抑える。具体的には、ブロードキャストが行われる全方位低解像度動画像のチャンネルは 高優先度で配信し、各クライアント毎の映像通信チャンネルは低優先度で伝送する。この 仕組みにより、回線が混んでいる状態でも低解像度ながら画像を途切れることなく見るこ とができ、回線に余裕があるときはクライアントがインタラクティブかつ高解像度で映像 を見ることができる仕組みも達成することができるという利点がある。

(5) ライブ入力、配信システムの開発:

以上のシステムを、蓄積された全方位画像に対してだけでなく、撮影しつつ配信するライ ブ入力、配信についても同様の機能を実現する。具体的に本システムは、図7に示すよう な各クライアントの要求を受け付ける配信サーバと時空間高解像度化処理を行う画像入力 及び処理システムから構成される。クライアントの要求の量によって、配信サーバの数を 自由に増減できるようにするために、画像入力及び処理システムからは分離する。この分 離により、クライアント数が増加した場合には配信サーバを増強したり、ネットワークの 通信帯域を広げたり、故障した配信サーバを除去したりすることで、柔軟で対故障性能が よい頑強なシステムを作り上げることができる。ネットワークサービスにおいてシステム が頑強であるということは重要なことである、配信サーバは各クライアントの部分画像切 り出しも担当するが、実際には画像入力及び処理システムから高解像度動画像を受け取っ ているので、複数のクライアントの部分画像の切り出し要求を、配信サーバがまとめて処 理を行うことができる。



図7. 高解像度動画配信システム

(6) マルチアクセス時の性能評価:

複数人に同時配信するときに、複数人が複数の部分画像を同時要求した場合の性能評価 を行う。通常の映像配信の場合は、高優先度映像だけが送られることになるが、部分画像 だけで良い場合は、部分画像だけを転送すれば良いので帯域圧縮が期待できる。そこで、 部分画像を要求したクライアントには、低優先度高空間周波数画像データチャネルを高解 像度部分画像データに差し替えて伝送することにより、配信サーバは個人個人の要求にイ ンタラクティブに応答することができる。ここでは、ネットワーク帯域と応答性能、CPU 能力、システムバス能力とを比較して、配信サーバが何台までのクライアントの要求に応 えられるかを検証する。

2-2 研究開発目標

2-2-1 最終目標(平成19年3月末)

1) 焦点 (レンズ中心) を合わせることにより、同一の視野を持った2種類の全方位動画像、 a) 高解像度(画像サイズ 8000x8000 画素) だが時間的には粗い(1 フレーム/秒)全方位 動画像と、b) 通常の解像度(640x480 画素) だが時間的に密な(30 フレーム/秒) 全方位 動画像を同時に撮影し蓄積するカメラを開発する。

2) 上の2つの全方位動画像から、空間的にも高解像度で、しかも時間的にも密な全方位動 画像(8000x8000 画素、30 フレーム/秒)を作成する技術を開発する。

3) 蓄積された2種類の全方位画像を圧縮してインターネットにより複数の閲覧装置(クラ イアント)に配信するサーバシステムと配信された2種類の全方位動画像から閲覧者側で 高解像度(8000x8000 画素)かつリアルタイム(30 フレーム/秒)の動画像を作成する技 術を開発する。また、閲覧者の要求する視野の部分画像のみを切り出しそれらを能率良く 配信し、クライアント側でその部分画像のみを高解像度で動画像表示するシステムも実現 する。

4) 以上のシステムを蓄積された全方位画像に対してだけでなく、撮影しつつ配信するライ ブ配信についても同様の機能を実現する。

5) エジプトのピラミッドや日本の古墳等の高精細なディジタルアーカイブ化とその配信 による実証実験を行い、本提案技術の有効性の検証とその周知を行う。

2-2-2 中間目標(平成17年1月末)

1) 焦点 (レンズ中心) を合わせることにより、同一の視野を持った2種類の全方位動画像、 a) 高解像度(画像サイズ 4000x4000 画素) だが時間的には粗い(1フレーム/秒)全方位 動画像と、b) 通常の解像度(640x480 画素) だが時間的に密な(30 フレーム/秒) 全方位 動画像を同時に撮影し蓄積するカメラを開発する。

2)上の2つの全方位動画像から、空間的にも高解像度で、しかも時間的にも密な全方位動 画像(4000x4000 画素、30 フレーム/秒)を作成する技術を確立する。

3) 蓄積された2種類の全方位画像を圧縮してインターネットにより複数の閲覧装置(クラ イアント)に配信するサーバシステムと配信された2種類の全方位動画像から閲覧者側で 高解像度(4000x4000 画素)でかつリアルタイム(30 フレーム/秒)の動画像を作成する 技術を開発する。また、閲覧者の要求する視野の部分画像のみを切り出しそれらを能率良 く配信し、クライアント側でその部分画像のみを高解像度で動画像表示するシステムも実 現する。

2-3 研究開発の年度別計画

							(金額は非公表)
研究開発項目	15 年度	16年度	17 年度	18 年度	年度	計	備考
 全方位高解像度動画入力とその配信システム に関する研究開発 1. 複合センサカメラの試作 ・複合センサ設計・試作,センサ特性評価 2. リアルタイム高解像度動画像の作成 高解像度動画像生成アルゴリズムの開発 シミュレーション評価 3.高解像度全方位カメラ用光学系の設計 光学系設計・試作,センサ特性評価 4.全方位高解像度リアルタイム動画像入力記録方式の開発 動画入力フォーマットの設計・試作 5.高解像度動画配信用ソフトウェアの開発 シミュレーション評価 6.ライブ入力,配信システムの開発 ライブ配信システムの試作 7.マルチアクセス時の性能評価 マルチアクセスプロトコル開発 8.総合システムの試作と評価 データ収集,総合評価 							
谷 計							

注) 1 経費は研究開発項目毎に消費税を含めた額で計上。また、間接経費は直接経費の30%を上限として計上(消費税を含む。)。

2 備考欄に再委託先機関名を記載

3 年度の欄は研究開発期間の当初年度から記載。

3 研究開発体制

3-1 研究開発実施体制



4 研究開発実施状況

4-1 複合センサカメラの開発

4-1-1 はじめに

カメラ技術の発達により様々なカメラが市場に出回っている.また,それに伴って高品質な映像 に対する世の中のニーズも高まっている.高品質な映像の条件としては高臨場感を与える高解像度 や,スムーズな動画像再生を実現する高フレームレートが挙げられる.そのような高品質な映像を 実現するために多くの研究が行われてきた.その例として,日本放送協会(NHK)が開発した超高精 細カメラが挙げられる.暫定的な仕様ではあるがこのカメラは走査線 4000 本,フレームレートが 60fps とかなり高品質な映像を撮像できる.また,映画の撮影から上映までをディジタルで行うデ ィジタルシネマが提案されている.これはフィルムに匹敵する品質の高精細映像を電子的手法によ り撮影,編集,保存することができ,またディジタルであるため画質の劣化がない.現在,4Kフォ ーマットと呼ばれる 4096×2160pixel,24fps の超高精細の映像規格が裁定され,これに対応する 機器の開発や,配信実験が行われている.しかし,これらのシステムは,高価な機材を利用するこ とで高品質な映像を実現しており,一般的な利用に結びつけるのはコストの面からも困難である. 一般的なテレビ放送映像方式である NTSC 方式では、フレームレートはスムーズな動画像を表現で きる 30fps であるが,解像度は 640×480pixel で,高解像度であるとはいい難い.

一方で、ディジタルスチルカメラには4000×4000pixelという高解像度のものが存在するが、フレームレートは低く滑らかな動画を撮影するには無理がある.このように、単位時間あたりの画像 データレートは CCD (Charge Coupled Device)の画素数とフレームレートの積で表され、カメラの画 像掃き出し速度の制限のために解像度とフレームレートはトレードオフの関係にある.つまり、一般的なカメラを用いての高解像度と高フレームレートを両立した撮影を行うのは困難である.

そこで、本研究では高解像度と高フレームレートを両立した撮影を実現するために特性の異なる2台のカメラを複合して用いることを提案する.この複合センサカメラでは解像度を重視した高 解像度低フレームレートのカメラと、フレームレートを重視した低解像度高フレームレートのカメ ラを複合して用いる.シーンからの入射光をハーフミラーで分光し、時空間周波数の異なる2台の カメラでそれぞれ撮像することにより同一視点、同一視野で高解像度と高フレームレートの撮像を 行うことができるカメラシステムを構築した.さらに、2台のカメラの特性が各々異なるため、レ ンズ歪み等の幾何学的特性を一致させる幾何学的キャリブレーションと、画素値の光学的特性を一 致させる光学的キャリブレーションを行う必要がある.このような2枚の画像を一致させるための キャリブレーション手法を提案し、改良を行なった.

4-1-2 複合センサカメラの試作

今年度は,前年度試作した複合センサカメラの評価と,改良を行なった.まず,複合センサについて簡単に説明し,その後評価結果を示す.

本システムは複合センサカメラと、画像を取り込むための PC、画像記録用の RAID (Redundant Arrays of Inexpensive Disks)システム、画像表示用のディスプレイ、フレームレートの異なる 2 台のカメラを同期させて撮像するためのパルスジェネレータから構成される(図 8). 複合センサカメラで撮像シーンの入射光をハーフミラーを用いて分光して透過光を一方のカメラで、反射光をもう一方のカメラで撮像する. このようにして得られた画像を PC 上に取り込んでディスプレイに表示させつつ、RAID システムに保存していく.

現状の2台のカメラは CCD の受光面積が異なるため水平画角と垂直画角の両方を一致させることができない.そこで、水平画角が40°で一致するようにそれぞれのレンズを選択する.ここで、 画角 θ と焦点距離 f の関係は、以下の式のようになる.x は CCD の縦もしくは横のサイズを表している.

$$\theta = \frac{360}{\pi} \tan^{-1} \frac{x}{2f} [\text{deg}]$$

(1)

この式より、様々なレンズで画角を計算し、同一視野を撮像できるようにレンズを選択する.レ



図8 複合カメラセンサシステムの構成

	X1 ·			
レンズの	高解像度低フレーム	ムレートカメラ	低解像度高フレーム	ムレートカメラ
焦点距離(mm)	水平画角(deg)	垂直画角(deg)	水平画角(deg)	垂直画角 (deg)
12.5	111	88	40	40
20	84	62	26	26
28	66	47	18	19
35	54	38	15	15
45	44	30	12	12

表1 レンズの焦点距離と画角の関係

表1より,高解像度低フレームレートカメラが焦点距離 50mm のレンズ,低解像度高フレームレートカメラが焦点距離 12.5mm のレンズを選んだ時に水平画角が 40°で一致しているのが分かる. これより,高解像度低フレームレートカメラには焦点距離 50mm のレンズ (ニコン社の Ai Nikkor 50mm F1.2S)を,低解像度高フレームレートカメラには焦点距離 12.5mm のレンズ (PENTAX 社の B1214D-2)を用いることにした.

27

10

10

4-1-3 複合センサカメラのキャリブレーション

40

複合センサカメラには2つのカメラ間での CCD 特性やレンズ特性の違いにより画像に幾何学的, 光学的な差が生じる. 複合センサカメラにより得られる2つの画像列情報を統合するためには,キ ーフレームでの画像一致を前提としている. そのため,幾何学的特性および光学的特性の差を2つ のカメラ間で補正する必要がある. 今年度は,ハードウェア化に向けたキャリブレーション手法の 改良と,精度評価を行なった.

(1) カメラモデル

50

高解像度画像と低解像度画像を一致させるためには何らかの基準平面上で整合を取る必要がある. 前年度までは,仮想的な基準平面を考え高解像度画像と低解像度画像の両方を幾何変換を施していた.今年度は,この基準平面を高解像度画像平面に設定することで,高解像度画像については補正する必要をなくし,計算量の削減を行なうことでハードウェア化に向けて改良を行なった.高解像 度歪み無し平面から高解像度画像への変換は,以下の式で行うことができる.

$$X_d = X_u + X_u \kappa_1 R_u$$

$$Y_d = Y_u + Y_u \kappa_1 R_u$$
(2)

これにより、高解像度歪み無し平面から高解像度画像への変換が四則演算だけで可能となり、ハードウェア化が容易となる.

(2) 処理の流れ

キャリブレーション処理の流れを図9に示す.現在のシステムではカメラから送られてきたデー タに対して色補間を行ってからキャリブレーションを行っていた.しかし,それでは扱うデータ量 が3倍になり計算コストがかかってしまう.そこで,ハードウェア実装するために,図10のよう に,Bayer 配列をG1,G2,R,B成分に分離し,それぞれのデータに対してキャリブレーションを行う. なお,輝度変換テーブルは従来どおりの方法で成分毎に作成する.

複合センサカメラを対象としたキャリブレーション処理のブロック図を図11に示す.カメラか ら送られてきたデータはそれぞれ Bayer 配列の成分毎に分けられ,対応する輝度変換テーブルによ り輝度補正が行なわれる.さらにその画素がどのモデルに対応するかをモデルに従って計算し,対 応するバッファに書き込む.





4-1-4 評価実験

ハードウェアでの画像補正が実装して上手く性能が出るか,ソフトウェアでシミュレーション実験を行った.

カラー補間した画像を図12に示す.このカラー補間された図に幾何学的,光学的補正を行った 結果を図13に示す.図13に示した高解像度画像を低解像度画像と同じサイズに縮小し,PSNR を計算した結果R成分22.32,G成分23.72,B成分23.28となった.

次にハードウェアによる変換をソフトウェアでシミュレーションする.低解像度画像を G1, R, B, G2 成分に分けそれぞれに対して幾何学的,光学的補正を行った結果を図14に示す.カラ ー補間を行い PSNR を計算した結果 R 成分 23.77, G 成分 23.85, B 成分 23.25 となった. PSNR はど



図12 カラー補間した入力画像

ちらも同じような値となった.これより、キャリブレーションをハードウェア化したとしても精度に問題がないことを確認した.



図13 ソフトウェアによるキャリブレーション



図14 ハードウェアによるキャリブレーションのシミュレーション結果

4-2 リアルタイム高解像度動画像の作成

上で試作した複合センサカメラから得られる2つの動画像から,空間的にも高解像度で,しかも時間的に も密な動画像を作成する.前年度,時空間解像度の異なる2つの画像系列(高解像度低フレームレート動 画像、低解像度高フレームレート動画像)の情報を統合する手法として周波数アプローチとモーフィングア プローチの2つの手法を考案し,それらのソフトウェアを実装した.今年度は,様々なシーンの実画像を用 いて実験を行い,手法の適応範囲や問題点の評価を行った.また,他の周波数変換手法や特徴点追跡 手法を利用するなど現在のアルゴリズムに改良を行った.

4-2-1 周波数空間アプローチによる高解像度画像の作成

複合センサカメラから得られる2種類の動画像から,画像処理によって高解像度・高フレームレ

ート動画像を生成するアルゴリズムを開発した.開発手法の有効性を確かめるため,MPEGテストシ ーケンスと試作した複合センサカメラから得られた実画像に対して各手法を適用し,高解像度動画 像を生成する実験を行った.また,生成画像のピーク信号対雑音比(PSNR)を測定し手法の性能を比 較した.

(1) 高解像度動画像の生成

本研究では複合センサカメラから得られた動画像をインターネット配信することを想定している.動画像のもつデータ量は非常に大きいため,通常は動画像をインターネット配信する場合に符号化するが,各フレームは符号化処理の過程で何らかの周波数変換により周波数領域へ移される.したがって,動画像をインターネット配信するという前提の下では,高解像度画像の生成処理を画像空間領域と周波数領域の両方で行うことが可能である.画像を周波数空間に移す周波数変換はこれまでに様々な方法が提案され,実用化されているが,本研究ではDCTを用いる方法、およびDWTを用いる方法を提案する.

本研究で提案する2種類のアルゴリズムは、いずれも以下のアイデアに基づいている.

- 1. 低解像度・高フレームレート動画像を用いてシーン中の動きを推定する.
- 2. 推定された動き情報を高解像度画像(キーフレーム)に適用することにより、動き補償を行う.
- 3. 動き補償によって生成された高解像度画像と、これに時間的に対応する低解像度画像を、周波 数空間で合成する.

提案手法は,解像度の異なる2種類の画像を周波数空間で合成することにより,キーフレーム間で のシーンの変化(画像全体の明るさの変化など)を生成画像に反映させることができるのが特徴で ある.

a. ウェーブレット変換を用いた高解像度動画像生成

離散ウェーブレット変換(DWT)は周波数変換の一種で,静止画像符号化の最新の国際標準規格である JPEG2000 にも採用されている. DWT は離散 Fourier 変換や DCT などの他の多くの周波数変換とは異なり,周波数領域においても画像空間情報を保持しているという特徴がある.また,一枚の画像に対して DWT を実行すると,異なる解像度を有する階層的な表現が得られる(オクターブ分割).提案手法では, DWT がもつこれらの特徴を利用する.

DWT は周波数領域においても画像空間情報を保持しているのが特徴であるが, shift-variant であることが知られている.このため,DWT 係数に対して動き補償を行うと誤差を生じる場合がある.そこで,提案 手法では,DWT を拡張した変換である冗長ウェーブレット変換(RDWT; redundant DWT)を採用している. RDWT を用いた高解像度画像生成処理の流れを図15に示す.ウェーブレット変換を用いた高解像度画像 生成のアイデアは以下の通りである.



- 1. 動き推定により、低解像度画像の各画素に対して動きベクトルを求める.
- 2. 高解像度画像に対して, RDWT をレベルαまで繰り返す. ただし, レベルαの低周波成分 LL^αはこの 後の処理に使用しないので求める必要はない.
- 3.1 で求めた動きベクトルを用いて,高解像度画像の高周波成分に対して動き補償を行う.ただし,動き

ベクトルが推定できなかった部分については、高周波成分に0を代入する.

- 4. 動き補償された高周波成分をダウンサンプリングし, RDWT 係数を DWT 係数に変換する.
- 5. レベル α の低周波成分として低解像度画像を統合する. すなわち LL^α=I_kとする.

6.5 で得られる DWT 係数に対して IDWT を実行し, 高解像度画像を得る.

ここで, RDWT に使用するウェーブレット関数は任意である. 画像統合によって解像度を向上させるという 目的に最も適合するウェーブレット関数は,実験を通して明らかにする.また,動画像のネットワーク配信を 考える場合は,上の1,2がサーバ側の処理,3から6がクライアント側の処理となる.

ウェーブレット合成法は、動き補償および2種類の画像の合成の両方の処理を周波数空間(ウェーブレット領域)で実行できる.また、低解像度画像に対する周波数変換が必要ないため、DCT 合成法と比較すると単純なアルゴリズムとなっている.しかし、RDWT による冗長な係数を扱う必要があるため、計算コストは必ずしも小さくならない.また、画像生成に必要なメモリ量が大きくなるという欠点がある.1次元 RDWT とその逆変換のフィルタバンクによる実現を図16に示す.DWT と異なるのは、スケールの増加に合わせてフィルタ係数をアップサンプリングする点である.レベル j+1 におけるフィルタはレベル jのフィルタを2倍アップサンプリングしたものとなる.画像のような2次元信号に対して RDWT を適用する場合、まず画像の各行について横方向に1次元 RDWT を実行し、2つの帯域に分割する.次に、2つの帯域それぞれに対し、画像の各列について縦方向に1次元 RDWT を実行し、2つの帯域に分割する.最終的に、1レベルの RDWT 分解により画像は4つの帯域に分割される.



図16 2レベルの1次元 RDWT 分解・合成フィルタバンク

(2) 評価実験

提案手法の有効性を検証するため, MPEG テストシーケンスを用いたシミュレーション実験を行った. 実験には,表2に示す MPEG テストシーケンスを使用した.

シーケンス名称	画像サイズ	使用フレーム
Coast guard	352×288	No. 0 \sim No. 294
Football	352×240	No. 1 \sim No. 120
Foreman	352×288	No. 0 \sim No. 294
Hall monitor	352×288	No. 0 \sim No. 294

表2 実験に使用した MPEG テストシーケンス

a. 空間アップサンプリングとの比較

各手法について、テストシーケンス全体の PSNR を測定した結果を表3に示す. すべてのテスト シーケンスについて、DCT 合成法の PSNR が最大となった. Daubechies 4-tap フィルタを使用した 場合の生成画像を見ると、生成画像の画質が視覚的には悪くないにもかかわらず、PSNR の結果は全 てのテストシーケンスについて低くなった. これは、Daubechies 4-tap フィルタのインパルス応答 (フィルタの数列)が対称でないため、直線位相特性をもたないからである. 直線位相特性をもたな いフィルタによりサブバンド分割を繰り返すと、画像中のエッジがずれていく. したがって、生成 画像にひずみが生じ、結果として PSNR が下がったと考えられる.

シーケンス	Coast guard	Football	Foreman	Hall monitor
最近傍法	22.03	20.69	25.29	22.82
双3次スプライン	22.03	20.76	25.73	22.89
frame rate 変換	22.58	16.44	20.46	29.18
整数 2/6	24.04	20.91	26.93	29.39
整数 5/3	22.27	19.72	24.78	24.82
Haar	23.95	20.53	26.70	29.51
Daubechies 4-tap	23.47	20.51	25.66	27.12
DCT 合成法	24.62	21.15	27.14	30.31

表3 PSNR の測定結果

b. 時間アップサンプリングとの比較

次に,提案手法と時間アップサンプリング法(フレームレート変換)の結果の比較を行う.ここでは、フレームレート変換アルゴリズムとして位相相関法を用いる手法を用いた.前節で述べた実験において、入力として用いた高解像度・低フレームレート動画像(4.29 [fps])をフレームレート変換アルゴリズムに適用し、高フレームレート動画像(30 [fps])を生成した.フレームレート変換アルゴリズムによって生成された画像を図17に示す.また、PSNRの結果を表3に示す.





(a) 生成画像(b) (a)を4倍拡大図17 フレームレート変換により生成された画像

本実験で入力として用いた動画像のフレームレートは、4.29 [fps]とかなり低い.このため、シ ーンによっては、連続する2フレーム間でかなり急激な動きが発生する.したがって、動き推定が 困難な場合が多く発生する.表3に示されているように、テストシーケンス Football および Foreman に対する PSNR の値は、時間アップサンプリング(フレームレート変換)よりも空間アップサ ンプリング(双3次スプライン補間)の方が上回った.これらのシーケンスは、シーン中の広い範囲 を動領域が占めている.このため、シーン中の動き推定が困難となり、PSNR が下がった.一方、テ ストシーケンス Coast guard および Hall monitor に対する PSNR の値は、双3次スプライン補間よ りもフレームレート変換の方が上回った.これらのシーケンスは、シーン中のほとんどが静止領域 か、あるいは単純な並進移動が占めている.したがって、動き推定が正確に行え、結果として高い PSNR が得られた

以上より,最近傍法や双3次スプライン補間といった空間アップサンプリング法は,シーン中に 占める動領域が大きい場合に有効であると言える.また,時間アップサンプリング法(フレームレ ート変換)はシーン中に含まれる動きが少ない場合に向いている.しかし,提案手法はすべてのテ ストシーケンスに対し,PSNRの点で空間アップサンプリング法および時間アップサンプリング法を 上回った.よって,提案手法はどのようなシーンに対しても,より高品質な動画像を生成できる性 能を有していると結論付けられる.

c. 実画像からの高解像度画像生成

複合センサカメラの試作機から得られた2種類の動画像を提案手法に適用し,サイズ4008×2672

画素,フレームレート30 [fps]の動画像を生成した.DCT 合成法を用いて生成されたフレームの例 を図18(a)に示す.(b)は,生成画像(a)のうち静止領域の一部を拡大したものである.比較のた め,同様の実験条件で整数2/6,整数5/3ウェーブレットを用いて生成したフレームを拡大したも のをそれぞれ(c)と(d)に示す.また,(a)から(d)の画像に時間的に対応する低解像度入力画像を(e) に示す.提案手法により生成された画像(図18(b)(c)(d))を低解像度画像(図18(e))と比較する と,提案手法により生成された画像では文字が明確に判読できるため,提案手法により解像度の向 上が実現できたと言える.この結果から,実画像においてもシミュレーション同様に高解像度・高 フレームレート動画像が提案手法によって生成できることが示された.



(a) 生成画像(DCT スペクトル合成法)



4-2-2 モーフィングアプローチによる高解像度画像の作成

高解像度低フレームレート動画像からはテクスチャ情報を利用し、低解像度高フレームレート 動画像からはモーション情報を利用することで、それぞれ優れた特徴から高品質な高解像度高フレ ームレート動画像の獲得を提案する.そのために、高解像度だが低フレームレートの動画像を撮影 するカメラと、低解像度だが高フレームレートの動画像を獲得するカメラからなる複合センサカメ ラを構築し、同一視点・同一視野を持つ時間的・空間的なサンプリングの異なる動画像を取得する. この2種類の動画像を用いて、高解像度で高フレームレートの動画像を得ることを目指す.本提案 手法では、モーフィングというコンピュータグラフィックス(CG)分野でよく用いられている技法 によって、高解像度低フレームレート動画像の中間フレームを補間し、高解像度高フレームレート 動画像を生成する.

(1) 高解像度高フレームレート動画像の生成

a. 動画像生成アルゴリズム

低解像度高フレームレート動画像は時間的サンプリングの面で優れ,高解像度低フレームレート 動画像は空間的サンプリングの面で優れている.提案手法では,これら両者の長所を活かして,高 解像度高フレームレート動画像を生成する.提案手法の概略を図19に示す.一般的に入力の高解 像度低フレームレート動画像のようにフレームレートが低い場合,特徴点のトラッキングが難しい. また,モーフィングでの特徴点の位置変化は線形移動と仮定することになる.そこで,本提案手法 では,フレーム数の多い低解像度画像から特徴点を検出し,特徴点のトラッキングにより求めた位 置関係変化を高解像度画像のモーフィングに用いる.このように低解像度高フレームレート動画像 を用いることで,トラッキングを比較的に容易に行える.さらには,三角パッチによりフレームを 部分分割し,モーフィングによりテクスチャを生成することで,近似的に動領域の非線形的な変化 を再現できる点が本手法の特徴である.また,キーフレームは,両画像において時間的に対応がと れているため,高解像度画像と低解像度画像の特徴点の対応付けも容易である.



提案アルゴリズムのフローチャートを図20に示す.

b. モーフィングによる高解像度テクスチャの生成

本手法では、モーフィングにより高解像度低フレームレート動画像の中間フレームを生成する. 前節で述べたように、ポリゴン単位で処理することで、形と色の両方を補間できるワーピング法を 用いる. 三角パッチ毎でアフィン変換により高解像度キーフレームのテクスチャから生成フレーム のテクスチャを算出する. これにより動領域の非線形なテクスチャ変化も近似的に再現できる. モ ーフィングに用いるテクスチャ情報は生成フレーム近傍の高解像度キーフレームから得られる.

図21のようにモーフィングは三角パッチ単位で処理を行う.特徴点追跡,ドロネー分割により近傍の高解像度キーフレームと生成フレームの三角パッチは対応付けされている.モーフィングでは,対応している高解像度キーフレーム中の対応パッチ内の輝度情報を補間することにより生成する.



c. 中間フレーム生成

これまで述べてきたように,本手法における中間フレームのテクスチャ生成はモーフィングによ り行うことを基本としている.しかし、撮像したシーンの激しい動きにより、急激に画素値が変化 し特徴点追跡が失敗した場合、モーフィングによる高解像度生成ポリゴンにおいて画像の歪みが発 生し、画質が著しく低下する場合がある.また、本手法で使用する特徴点追跡法では、キーフレー ム間の全てのフレームで対応付けられた特徴点を三角パッチの頂点とする. このため、フレームの 外縁領域ではモーフィングに使用する三角パッチを作成することはできず、モーフィングのみでフ レーム内の全画素の画素値を算出することは不可能であり、テクスチャの欠落領域が生じる. そこ で、中間フレームのテクスチャを生成する際は以上の2点を考慮する必要がある。中間フレーム生 成は以下の手順で行う.まず,モーフィングにより高解像度ポリゴンを生成すると同時にその歪み 判定を行う. 歪んでいないと判定されたポリゴンには、モーフィングにより生成したテクスチャを 貼付ける.一方,歪んでいると判定されたポリゴンには、テクスチャを貼付けない.次に、モーフ ィングによりテクスチャが生成されなかった領域に対して、テクスチャ補間処理を施す.テクスチ ャ補間には、生成フレームと同一時刻の低解像度画像、および近傍高解像度キーフレームを使用す る.低解像度画像は双線形補間により、近傍高解像度キーフレームと画像サイズをそろえて両者の 差分をとる.差分が大きな画素に対しては、低解像度画像を双線形補間したテクスチャにより内挿 を行う.一方,差分が小さな画素に対しては,近傍高解像度キーフレームのテクスチャにより内挿 を行う.

歪み判定には、同一時刻の低解像度画像を双線形補間したものとの比較により生成画像の破綻部 位を検出し、補正を行うことで画像劣化を回避する.判定は画素毎にポリゴンサイズで差分をとり、 このポリゴン内の平均、つまり1画素当たりの差分を評価することで、大域的に比較し生成画像の 破綻ポリゴンの検出を行う.生成画像と低解像度画像の比較にポリゴン内の画素平均を用いている のは、ノイズや提案手法による解像度向上を破綻部位であると誤検出することを防ぐためである. このように、生成画像内のポリゴン毎に補正を行うことで、高解像度処理が成功した領域を保った まま、破綻領域のみを補正するため、画像全体としての画質向上を図ることができる.

(2) 評価実験

撮像システムのシミュレートにより作成した 2 種類の動画像を入力として実験を行った. MPEG テストシーケンスを加工することで,同一視点・同一視野を持つ低解像度高フレームレート動画像 と高解像度低フレームレート動画像を作成し,この2つの動画像を入力として高解像度高フレーム レート動画像の生成を行った. 画質評価は Peak Signal to Noise Ratio (PSNR)により行った.

a. 高解像度低フレームレート入力動画像の有効性

提案手法は、解像度とフレームレートのいずれかをそれぞれ優先サンプリングした2種類の動画 像を入力として高解像度高フレームレート動画像の生成を行う.一方、低解像度高フレームレート 動画像のみを使用することで高解像度高フレームレート動画像を生成するアプローチに低解像度 画像補間手法がある.代表的な補間手法としては、最近傍法 (Nearest Neighbor),双線形補間法 (Bilinear Interpolation),バイキュービック法 (Bicubic Interpolation)が挙げられる.

上記の3手法,および提案手法により高解像度高フレームレート動画像を生成し,それぞれの生成画像について検証を行った.ここでは,試作された複合センサカメラの仕様に合わせ,2種類の入力動画像の解像度比を1:16(縦横比は共に1:4),フレームレート比を7:1として実験を行った.

図22に理想画像である MPEG テストシーケンス,提案手法による生成画像,および3種類の補 間手法による生成画像の例として "Container"の第18フレームとその一部の拡大図を示す.こ の結果より,提案手法,および補間手法のそれぞれにより高解像度低フレームレート動画像には存 在しないフレームが生成できていることが確認できた.また,低解像度高フレームレート動画像同 様 30[fps]の動画像が生成でき,同様の動きが再現できていることが確認できた.補間手法による 生成画像を主観的に評価する場合,バイキュービック法が最も自然な画像を得られる手法として知 られている.図22の "Container"についても補間手法の中でバイキュービック法が最も自然な 画像を生成できたが,そのバイキュービック法と比較しても,本手法による生成画像の解像度が視 覚的に明らかに向上していることが分かる.

シーケンス	提案手法	最近傍法	双線形補間法	バイキュービック法
Container	34.26	22.31	22.52	21.09
Hall monitor	32.21	23.46	23.83	22.03
Coast guard	25.89	22.48	22.68	21.22
Foreman	27.69	25.72	26.72	25.08
Mobile	20.04	17.13	17.21	16.33
Garden	19.09	17.41	17.39	16.53

表4 提案手法と低解像度補間手法の PSNR 比較

表4に6種類のMPEGテストシーケンスについて各手法により生成した動画像のPSNRを示す.補間手法の生成画像を主観的に評価した場合,バイキュービック法が最も自然な画像を生成できたが, PSNRにより画質を客観評価した場合は,ほとんどのテストシーケンスで双線形補間法が他の2手法よりもPSNRが高い数値を示した.一方,提案手法は,補間手法の最も高い数値と比較して全ての テストシーケンスで高いPSNRが得られた.この中で特に "Container"と "Hall monitor"が 補間手法に対する PSNR の向上が著しかった. "Container"は、固定カメラで撮像され、シーン中の動領域・動物体は単純な平行移動を行っている.また、 "Hall monitor"も同様に固定カメラで撮像され、シーンのほとんどが静止領域である.このような固定カメラで撮像したシーンでは、大部分が静止領域である.よって、ポリゴン形状を変形させることなく高解像度キーフレームのテクスチャから直接、静止領域を補填できるため容易に高解像度画像を生成できる.また、動物体の単純な平行移動は、特徴点のトラッキングが容易であり、さらにキーフレームからのテクスチャの変化が少ないため、本手法での高解像度画像の再現性が高いものと考える.

一方, "Coastguard"は、移動カメラで撮像されているため、シーンのほとんどが動領域であ



図22 "Container" Frame No. 18, 上から順に理想画像(MPEG テストシーケンス),提案 手法,最近傍法,双線形補間法,バイキュービック法 り、さらにオクルージョンが数多く出現している.また、 "Foreman"は、カメラのパンによるシ ーンの急激な変化があり、人間の表情などテクスチャの変化が複雑であるが、これらにおいても補 間手法より PSNR が向上した. "Mobile"や "Garden"において他の画像に比べて相対的に PSNR が低いのは、これらのシーンが全体的にテクスチャが複雑であり、オクルージョンが存在するため、 動領域のトラッキングが困難であることが原因と考えられる.ただこれらシーンにおいてもでも、 若干ながら補間手法に対して PSNR が高い値を示した.このように、本手法はシーンに依存して性 能は異なるものの、全てのシーンにおいて解像度が向上できることが示された.

b. 低解像度高フレームレート入力動画像の有効性

前節では、高解像度低フレームレート動画像を入力画像として使用する有効性を検証した.こ こでは、低解像度高フレームレート動画像を入力画像として使用する有効性を検証する.高解像度 高フレームレート動画像生成には、低解像度画像補間手法の他に、高解像度低フレームレート動画 像のみを使用するアプローチがある.ここでは、図23のように高解像度キーフレーム間において KLT tracker により特徴点を対応付け、キーフレーム間の動きを線形移動と仮定して高解像度高フ レームレート動画像の生成を行い、提案手法と比較した.それぞれの手法で生成した動画像の PSNR を表5に示す.全てのテストシーケンスで歪み判定を省いた生成画像の方が高い PSNR の数値を示 している.





図24 左: 歪み判定を省いた提案手法による生成画像,右: フレーム間線形移動手法による生成画像

表5 PSNR 比較,手法A: 歪み判定を省いた提案手法,手法B: キーフレーム間特徴点対応・ 動きベクトル線形移動

手法 A [dB]	手法 B [dB]
34.26	25.15
32.21	24. 53
25.89	19.48
27.69	22.48
20.04	16.98
19.09	15.03
	手法 A [dB] 34. 26 32. 21 25. 89 27. 69 20. 04 19. 09

図24に "Foreman" について今回の比較実験と提案手法による生成画像の一例を示す. キー

フレーム間線形移動手法により中間フレームを生成する場合,動きが大きく各画素で激しい輝度変化が起きている"Foreman"のような動画像では、ゴーストが生じ、PSNR が急激に低下する.以上により、時間的サンプリングを優先した低解像度高フレームレート動画像で動きベクトルを検出し、モーフィングに用いる提案方法が非常に有効であることを確認した.

c. 解像度比,フレームレート比の検証

提案手法では、2種類の入力画像の解像度比やフレームレート比が大きいほどセンサの低コスト化や映像データ量の削減が可能となり、意義が大きくなる.しかしながら、これらの比が大きいほど、生成画像の画質改善が困難となる.そこで、解像度比やフレームレート比を様々に組み合わせた動画像を入力とした生成画像の PSNR による客観的画質の評価を行った.解像度比は1:4,1: 16,1:64(縦横比はそれぞれ1:2,1:4,1:8)とし、フレームレートは低解像度高フレームレート動画像が 30fps、高解像度低フレームレーと動画像が1,2,3,4.29,5,6,10,15[fps]で



サンプルをとった. 6 種類の MPEG テストシーケンスにおける結果を図25 に示す.

フレームレートに関してはいずれも同様の挙動を示しており,4.29~6[fps]よりも低いフレーム レートでは急激に PSNR が減少する.一方,解像度比に関してはシーケンスによって結果が大きく 異なっている.これは低解像度画像で補正を行っている領域の大きさが関係していると考えられる.

"Container"のように固定カメラで撮像され特徴点追跡が容易なシーケンスでは、画質破綻が少なく歪み判定の結果により高解像度ポリゴンのテクスチャ領域が多くなるため、解像度比を変えても画質の差異が比較的小さくなる.一方、移動カメラで撮像されシーンの急激な変化があり、人間の表情などテクスチャの変化が複雑な"Foreman"では画像のトラッキングが失敗し、歪み判定の結果により低解像度画像で補正する領域が多くなるため、解像度の差が大きくなると著しく画質が劣化する.これらの結果から、全てのシーンを総括して考えると、解像度比1:16、フレームレート比5:1あたりが、画質とデータ量削減のバランスポイントであると考えられる.



図26 (左)低解像度高フレームレート入力画像,(中央)高解像度低フレームレート入力画像, (右)高解像度高フレームレート生成画像

d. 実画像による実験

試作した複合センサカメラにより撮像した動画像を入力として,評価実験を行った. 高解像度 高フレームレート動画像の生成結果を図26に示す.この結果により,高解像度低フレームレート 動画像に存在しないフレームが,本手法により生成できることが確認できた.したがって,中間フ レームが補間され,低解像度高フレームレート入力動画像と同様に30[fps]の動画像が生成でき, 動物体・動領域の滑らかな動きが再現されていることが確認できた.図27に低解像度高フレーム レート入力画像と高解像度高フレームレート生成画像の一部を拡大して示す.高解像度高フレーム レート生成画像は,低解像度高フレームレート入力画像と比べ解像度が改善していることが確認で きた.



図27 (上段)静止領域の拡大図,左:低解像度高フレームレート入力画像,右:高解像度高フレームレート生成画像,(下段)動領域の拡大図,左:低解像度高フレームレート入力画像,右: 高解像度高フレームレート生成画像

(3) まとめ

本研究では、同一視点・同一視野で、解像度とフレームレートの異なる2種類の動画像から高 解像度高フレームレート動画像を生成するためのアルゴリズムを提案した.提案手法では、低解像 度高フレームレート動画像からモーション情報を抽出し、高解像度低フレームレート動画像からデ クスチャ情報を抽出し、このモーション情報とテクスチャ情報をモーフィングにより統合すること で高解像度高フレームレート動画像を生成する.また、提案手法の有効性を確認するために、シミ ュレーション実験と実画像を用いた実験を行った.シミュレーション実験では、複合センサカメラ を想定した2種類の入力動画像を作成し、高解像度高フレームレート動画像を生成した.この実験 では、解像度とフレームレートの異なる2種類の動画像を用いる提案手法に対して、いずれか一方 を入力とする手法と比較検討することで、提案手法の有効性を確認した.実画像を用いた実験では、 実際に複合センサカメラで撮像した2種類の動画像を入力として高解像度高フレームレート動画 像を生成した.この実験では提案手法により高解像度キーフレーム間の中間フレームが高解像度画 像として生成されることを確認した.

4-3 高解像度全方位カメラ用光学系の設計

提案した複合視覚センサ用の全方位ミラーの設計試作を行った.全方位ミラーと組み合わせるこ とにより,複合視覚センサを全方位の広視野角で利用することができる.試作した複合センサに用 いた2台のカメラ(画像サイズ4008×2672画素,セルサイズ9.0×9.0µmと画像サイズ1008×1018 画素,セルサイズ9.0×9.0µm)のCCDを用いている.従来の単一双曲面ミラーによる全方位視覚 センサはその形状に起因する像面湾曲が問題になっている.像面湾曲とは,入射角に対して焦点距 離が異なることから像点位置が湾曲する収差で,中心部にピントを合わせた場合は周辺部がぼける といったように,全体にピントの合った画像が得られなくなる現象である.この像面湾曲の影響か らミラーの高解像度化やセンササイズの小型化が制限されていた.本試作では,この像面湾曲を抑 制して複合センサカメラの4008×2672画素という高解像度に対応するため,2枚の双曲面反射型ミ ラーを試作した.前年度に導出した2枚反射全方位光学系の幾何や光学シミュレーションの結果を 基に,本年度は試作機の具体的な光学設計を行い,試作を行なった.図28に設計した全方位ミラ ーの設計図を試作した図29に概要を示す.



図28 設計図



図29 試作2枚反射型全方位ミラー

試作した2枚反射全方位ミラーの光学性能を確認するために、テストチャートを用いた解像度比 較実験を行った.本実験では、図30に示すように、全方位ミラーに対してチャートを設置し撮像 した.撮像された画像中のチャートを視認により調べることにより、入射角に対する解像度をもと めた.図31に実験により得られた入射角に対する解像度特性を示す.図31に比較として、前年 度に試作した単一双曲面全方位ミラーによる結果も同様に示した.ただし、この図は各角度で得ら れた解像度の値を曲線近似により表している.図31より、レンズ焦点の合っている入射角0度付 近では差はないが、従来ミラーは入射視野角が焦点位置から外れるのに対して著しく解像度が低下 しているのにたいして、今回試作した2枚反射型全方位ミラーでは像面湾曲が抑制され、全域において高い解像度が保持されている事が示されている.この結果から、試作した2枚反射型全方位ミラーは光学シミュレーション同様、実験においても高解像度に適応できることが示された.

次に,実際に複合センサカメラに試作ミラーを接続し,撮像した結果を図32に示した.図32 左図より実際の撮像画像により,試作全方位ミラーが複合センサカメラの高解像度側の解像度 (4008x2672 画素)に十分対応できている事を確認した.

従来の全方位ミラー(双曲面ミラー1枚) 14 カメラ 12 Ê 10 解像度(line/ 8 6 第15号 (双曲面) 4 2 第2ミラー 0 (双曲面) -25 -20 -15 -10 -5 0 5 10 15 20 25 ミラー支持具 入射視野角(deg) 図30 解像度測定実験 図31 入射角に対する解像度の変化

- 双曲面2枚ミラーシステム



図32 全方位複合センサ画像 (左:高解像度例フレームレート画像,右:低解像度侯フレームレート画像)

4-4 全方位高解像度リアルタイム動画像入力記録方式の開発

4-4-1 複合センサカメラシステムを利用したリアルタイム動画像記録

前年度,各々のカメラで撮像した画像データをワークステーション上に取り込んで,ディスプレイに表示させつつ RAID システムのハードディスクに記憶させるソフトウェアを開発した.今年度は、これらのソフトウェアドライバの修正と、データ取得タイミングの最適化による高速化を行った.リアルタイムの保存,表示を実現するために、画像の取り込みから保存までの一連の処理を「ダブルバッファ処理」を用いて行う.ここでダブルバッファ処理について簡単に説明する.

まず,画像データ2フレーム分のメモリを用意しておく.その二つのメモリを M1, M2 という名称 で呼ぶことにする.画像取り込みの命令を出して M1 に画像データを格納する.その処理の裏で M2 に格納された画像データをディスプレイ上に表示させつつ, RAID システムのハードディスクに書き 込んでいく.これらの処理が終了したら,フレームバッファを切り替える.同様に画像取り込みの 命令を出して、今度は M2 に画像データを格納していく. 裏では前述の処理で M1 に格納した画像デ ータをディスプレイ上に表示させつつ、RAID システムのハードディスクに書き込んでいく. ダブル バッファを用いる方が画像の取り込みと、画像の表示、保存を並列に処理するため高速に処理でき る.

4-4-2 書き込み速度の最適化

それぞれのカメラの解像度とフレームレートにより、低解像度高フレームレートカメラは1秒間 に約 30Mbyte の画像データを、高解像度低フレームレートカメラは1秒間に約 47Mbyte の画像デー タを同時に実時間で書き込まなければならない.実時間書き込みを実現するためには、書き込みブ ロックサイズとオープンしているファイル数を考慮する必要がある.

まず,書き込みブロックサイズについて考える.書き込みをする際にデータの転送速度等の関係 上,データの書き込みブロックサイズを変えると書き込み速度も変わる.そこで,書き込み速度の 最適化を行うために実験を行った.実験内容は 30Mbyte のデータを RAID システムのハードディス クに書き込む際に,書き込みブロックサイズを変化させていき,その時の書き込み所要時間を計測 した.この結果を図33に示す.図33より,書き込みブロックサイズが約 500KByte の時に書き 込み所要時間が短いことがわかる.この結果より書き込みブロックサイズを約 500KByte にして保 存していくことにした.同様に高解像低フレームレートカメラについても同様の実験を行った.図 34に結果を示す.高解像度画像の時も低解像度画像の時と同様に 100KByte から 500KByte の範囲 で所要時間が短くなっているので,書き込みブロックサイズを 500KByte にして書き込むようにし た.

次に、ファイルオープンの回数について考える.それぞれのカメラで、1枚の画像を保存するたびに書き込み先のファイルをオープンすると、その分無駄な時間を消費してしまう.実際に上の実験で求めた書き込みブロックサイズで画像1フレームずつファイルオープンすると、書き込みが追いつかずコマ落ちが生じて実時間での書き込みができない.そこで、ファイルオープンの回数を減らして一つのファイルに1000フレーム分のデータを保存することにした.

4-5 高解像度動画配信用ソフトウェアの開発

本研究で目標とするシステムの概略図を図35に示す.全方位視覚センサによって撮像された高 解像度低フレームレートの動画像と低解像度高フレームレートの動画像の二つの動画像から,高解 像度高フレームレートの動画像を生成する.生成された高解像度高フレームレートの動画像はサー バに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において,クライアン トが見たい領域のみを分割して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画像に変換してクライアントが望む領域を出力する.クライアントが別の 視点を見たい場合は、視点変更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 チャンネルで配信することによってクライアントが見たい部分を見られるようなインタラクティ ブな動画像配信システムを構築する.

上で述べたシステムを実現する際に、全方位画像全体を配信せずに部分画像を多チャンネルで配 信する理由は、既存のプロトコルや伝送路には次のような制約があるためである.



図33 低解像度高フレームレートカメラの 書き込みブロックサイズと書き込み所要時間



図34 高解像度低フレームレートカメラの 書き込みブロックサイズと書き込み所要時間



図35 高解像度動画配信システム

1. 現存するプロトコルではHDTV (1920×1080) までの解像度の動画像しか配信できない

2. 配信する伝送路には帯域制限がある

よって高解像度動画像を配信する為に、画像を分割する事によってデータ量を落として配信する. また、分割して配信する事により各伝送路の帯域幅を節約でき、さらに、クライアントが見たい部 分だけを配信することによってチャンネル数も節約できる.動画像を分割して多チャンネルでスト リーミング配信するために、既存の RTSP(Real Time Streaming Protocol)を DSS(Darwin Streaming Server)を用いて拡張したプロトコルを使用する.

4-5-1 受信クライアントソフトウェアの開発

この節では本研究の目的であるストリーミング再生プレイヤの開発において,ベースプログラムとした MPlayer と目標システムの実装について述べる.

(1) MPlayer の概略

MPlayer は動画や音声の再生ソフトで,動画は MPEG-1 や MPEG-2 だけでなく,ビデオ CD や DVD な どを直接再生することも可能である.また DivX や Windows Media Video にも対応しており.音声 も MP3 だけでなく, Windows Media Audio 等,様々なフォーマットに対応している.また MPlayer はフリーソフトウェアかつオープンソースであり,OpenGL の出力モジュールをサポートしているの で本研究においてベースプログラムとした.

(2) MPlayer の解析と目標システム

この節では動画像を再生する際に, MPlayer がどのような処理をしているか解析し,目的とする特性を持つクライアントソフトウェアの作成のために, MPlayer に加えた改良点について述べる.

a. 出力モジュール選択

MPlayer はコンソールアプリケーションなので、ターミナルからコマンドを入力して実行する. OpenGL の出力モジュールを使用して動画像を再生するには

mplayer -vo gl filename.mov

のように入力する.オプション -vo は映像出力先を指定しており,この場合は OpenGL のモジュールを選択している.すなわち,動画像の各フレームをテクスチャにし,それを画面に表示することになる. OpenGL のモジュールには

LIBVO_EXTERN(g1)

の様な命令が書かれており、これが実行されるとモジュール内に、モジュール内の関数へのポイン タをメンバにもった video_out_gl という構造体ができる.これは関数マクロであり、これが定義 されたヘッダファイルを読み込むことにより関数ポインタ群をもつ構造体を作っている. OpenGL のモジュールに限らず、各モジュールには初期化、テクスチャ作成、描画や制御といった共通の関 数が用意されていてこの関数ポインタを用いて、main モジュールから映像出力(ここでは OpenGL) のモジュールの関数にアクセスし実行することができる.

b. 座標変換

全方位視覚センサによって得られた画像は、単一視点を持つという特徴から透視投影画像に変換で きる.全方位視覚センサには様々なものが考えられているが、本研究では双曲面ミラーが1枚の一 回反射タイプの全方位視覚センサと、双曲面ミラーが2枚の2回反射タイプの全方位視覚センサに ついて変換を考える.2回反射タイプの全方位視覚センサは1回反射タイプの全方位視覚センサの 問題点であった周辺部の解像度低下が軽減されている.この二つのセンサは透視投影画像への変換 式が異なるので、それぞれの変換式をこの章で示し、全方位画像から透視投影画像への幾何変換を 実現する.

一回反射タイプ

図36に一回反射タイプの全方位視覚センサの幾何特性を示す.ここで世界座標系(X, Y, Z)と画像の座標系(x, y)には

$$x = X \times f \times \frac{(b^2 + c^2)}{(b^2 - c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}}$$

$$y = Y \times f \times \frac{(b^2 + c^2)}{(b^2 - c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}}$$
(3)

の関係があり、この対応関係に基づいてテクスチャマッピングを行う事によって透視投影変換を行う事ができる.変換結果を図38に示す.

二回反射タイプ

二回反射タイプの幾何特性の図を図37に示す.世界座標系は左右の回転方向 θ と、上下の角度を図37の α で表し、画像座標系を(x, y)で表すと、世界座標系と画像座標系には以下の関係が成り立つ.この対応関係に基づいてテクスチャマッピングを行うと図39のようになる.

$$x = r \cos \theta$$

$$y = r \sin \theta$$

$$t = \frac{(b_1^2 + c_1^2) \tan \alpha + 2b_1 c_1 \sqrt{\tan^2 \alpha + 1}}{(c_1^2 - b_1^2)}$$

$$r = \arctan \frac{f(c_2^2 - b_2^2)}{(b_2^2 + c_2^2)t + 2b_2 c_2 \sqrt{t + 1}}$$

(4)



MPlayer でのフレームの描画は各映像出力モジュールの中の関数で行われている. OpenGL のモジュ ールにはフレームからテクスチャを作成する draw_frame という関数と,テクスチャマッピングを 行って画像を表示する flip_page という関数がある.本研究では後者の関数を変更することによっ てテクスチャ化されたフレームに幾何変換を加え,全方位画像を透視画像に変換する.サーバから 得られたカメラパラメータは構造体に格納されており、そのメンバに全方位視覚センサのミラータ イプを加えることによって、センサタイプに応じて変換式を変更する.





図38 一回反射タイプの変換結果





図39 2回反射タイプの変換結果

(3) 評価実験

この節では動画像の画像サイズ,テクスチャマッピングを行う際の分割格子数,そして透視投影変換に関してフレームレートの観点から評価実験を行った.

X 0 11	
OS	Fedora Core 4
CPU	Intel Xeon 2.40 GHz
Compiler	gcc version 3.2.3
GPU	NVIDIA Quardo4 280NVS
cache memory	512KB
main memory	1GB

表6 評価実験環境

a. 実験方法

フレームレートを計測する為に、画面に表示する関数、flip_pageの中で時間を取得する.本研究 では 30fps の動画像において 100 フレーム表示する間にかかったミリ秒を計測し、フレームレート を推定した.1 フレーム目はマッピングの為の計算が行われるため、ここでは 50 フレーム目から 150 フレーム目までの 100 フレーム間を計測することにした.また分散を小さくするために 3 回の 平均を計測値とし、画像サイズ、分割格子数、変換式を変更してフレームレートを求めた.ここで 分割格子数が n とは、テクスチャを縦横方向に n 分割するという意味である.画像サイズは 720× 480 を基準サイズとし、その 1/4 倍~3 倍までで変更した. ビットレートは 90kbps を基準ビットレートとし、45kbps~360kbps の範囲で評価した. また実験環境を表 6 に示す.

b. 画像サイズとフレームレートに関する評価実験

実験結果より、ビットレートが90kbpsの時の画像サイズとフレームレートの関係を図40に示 す.一般的なNTSC 方式のテレビのフレームレートは30fps であるが、図40より画像サイズが2 倍までは十分なフレームレートが得られている事が分かる.画像サイズが2倍を超えたあたりから フレームレートは減少し始め、サイズが3倍になるとフレームレートが一桁になってしまい、動画 像としては成り立たなくなる.2倍のサイズが1440×960のサイズなので、このサイズの解像度な ら問題なくソフトウェアで再生することができる.





c. 分割格子数とフレームレートに関する評価実験

テクスチャマッピングにおいてテクスチャを切り出す格子数を変化させた時の, 描画にかかる時間について評価実験を行った. MPlayer が起動すると視点情報の初期値によって格子点のマッピン グの計算が行われる.ここでは格子間のテクスチャの線形補間にかかる計算量とフレームレートの 関係についてのみ言及するために, 視点変更を行わない 50 フレーム目から 150 フレーム目までの 経過時間を測定する.ビットレートは 90kbps とする.実験結果より分割格子数とフレームレート の関係を図41に示す.図41より分割格子数が 32,64,128 の時は,フレームレートはサイズが 2 倍の時まで維持できている事が分かる.しかし,分割格子数が 256 になるとフレームレートはサイズが 1 倍の時から下がり始め,サイズが 2 倍になると 20fps を下回り動画像としてはかなり不連 続になってしまう.このことから,描画面積は同じでも小さいテクスチャを多く変形する方が時間 がかかり,サイズが 2 倍までを 30fps で再生するには分割格子数は 128 以下が望ましい事が分かる. また図40と比較してみると,分割格子数が 128 まででは,透視投影変換をしてもフレームレート に与える影響はほとんどないと言える.

d. 透視投影変換とフレームレートに関する評価実験

実際に本研究で提案するシステムを用いて配信を行うと、当然視点変更が発生すると予想される ので、その際の格子点のマッピングの再計算にかかる時間とフレームレートについて評価する.プ ログラムを視点変更が行われなくても、フレーム毎にマッピングのための格子点の計算を行うよう に変更し、透視投影変換を行ったときのフレームレートを計算する.ビットレートは 90kbps で、 変換式は前節で述べた二つの式について検討し、分割格子数を変えて実験する.図42に1回反射 タイプの透視投影変換とフレームレートの関係を、図43に2回反射タイプの透視投影変換とフレ ームレートの関係を示す.図42、43より分割格子数が32、64の時はどちらの変換式もサイズ が2倍までは30fpsを維持しており、減少の様子もほぼ一致していると言える.また分割格子数が 128のときはサイズが2倍の時は30fpsを下回っているものの、1.5倍のサイズまでは30fpsを維 持している.一方、分割格子数が256の時はどちらの変換式も著しくフレームレートが減少してい くが、2回反射タイプの変換式の方が低い値をとる.これは前節での議論からも分かるように、2 回反射タイプの変換式に上へて大きいため であると考えられる.これらの議論より、分割格子数が32、64の時は計算量の違いがフレームレ ートに与える影響は少ないと言え、分割格子数が128の時も1.5倍のサイズまでならフレームレートに影響を与えないと言える.

e. プロセス実行時間の計測

1 枚のフレームを描画する為には、デコード、テクスチャメモリへの転送、格子点の計算、テクス チャマッピングの処理を行わなければならない.かつ複数のファイルを読み込むときは、デコード とテクスチャメモリへの転送はファイルの数だけ行われる. ここではサイズとビットレートの観点 からデコード、テクスチャメモリへの転送、テクスチャマッピングにかかる時間を計測し、分割格 子数の観点から格子点の計算時間を計測する。計測方法は、プログラム中で各処理が行われる直前 と直後で時間を計測し、その差から各処理の実行時間を求め、3回の平均を計測値とする. 前節ま ではビットレートを 90kbps に固定したが、実際のシステムで部分画像を生成すると元の画像に対 してビットレートは減少する. ビットレート 180kbps で 720×480 の画像を 4 分割すると, 360×240 の部分画像のビットレートは 45kbps となる. このように部分画像を生成しても、単位ピクセルあ たりのビットレートは変化しない. そこで 720×480 を単位サイズとして, これらの画像を単位サ イズあたりのビットレートが 180kbps と呼ぶことにする. 単位サイズあたりのビットレートを 45kbpsとしたときの画像サイズと処理時間の関係を図44に,画像サイズが1倍の時のビットレー トと処理時間の関係を図45に、分割格子数と処理時間の関係を表7に示す.図44より、テクス チャマッピング時間はサイズに依存しないが、デコード時間とテクスチャ転送時間はサイズが大き くなると増加することが分かる、図45からは、テクスチャ転送時間とテクスチャマッピング時間 はビットレートに依存しないが、 デコード時間はビットレートが増加すると共に増加することが分 かる.また表7より格子点の計算にかかる時間は、1回反射タイプの変換式と2回反射タイプの変 換式でそれほど大差ではないということが分かる.



分割格子数	計算時	間[ms]
	1回反射タイプの変換式	2回反射タイプの変換式
32	0.4	0.4
64	1.8	1.8
128	7.0	7.1
256	28.3	28.2

表7 格子点の計算時間

f. ビットレートと分割ファイル数における考察

理想的なフレームレートは 30fps なので,各プロセスの実行時間の総和が 33ms を超えなければ 遅延なく再生可能であり,各プロセスの実行時間の総和が 33ms を超えてしまうと遅延が発生する. ここでデコード,テクスチャメモリへの転送,格子点の計算,テクスチャマッピンにかかる時間を それぞれ,Td,Tt,Tk,Tmとする.ファイル数を n とすると,視点変更が行われない場合の1フレー ムあたりの時間 T は

$$T = nT_d + nT_t + T_m$$

となる. 30fps を維持するためにはこれが 33ms を超えなければいいので

(6)

(5)

という関係式が得られる.この式を満たせば複数ファイルの再生は可能であると言える.また視点 変更が行われると格子点の計算が発生するので

$$nT_d + nT_t + T_k + T_m \le 33$$

 $nT_d + nT_t + T_m \leq 33$

(7)

という関係式を満たせば、視点移動中でもフレームレートは低下せずに再生を行える.

前節より各プロセスの実行時間が分かったので、これより、実際に複数の部分画像を同時再生したときの推定フレームレートfを求める. 導出式は

 $f = \frac{1}{nT_d + nT_t + T_k}$ (8) となる、単位サイズあたりのビットレートを 45kbps, 90kbps, 180kbps, 360kbps とし, 1440×960

2/26. 単位サイスあたりのビットレートを45kops, 90kbps, 180kbps, 360kbps とし, 1440×960 の動画像を 4, 9, 16 のファイルから生成したときの推定フレームレートを導出した結果を表 8 に 示す. これらの議論より, 部分画像を多チャンネルで配信する際には画像サイズ, ファイル数 (チ ャンネル数), 部分画像のビットレートの選択が重要となる事が分かる.

		=		
	45kbps	90kbps	180kbps	360kbps
4分割	30.6	29.5	27.2	24.5
9分割	29.2	27.0	24.2	21.1
16分割	25.7	24.1	21.0	17.3

表8 複数画像再生時の推定フレームレート(単位:fps)

以上,画像サイズ,分割格子数,変換式,ビットレートに対するフレームレートについて検証を行い,各処理の実行時間を求めた.これらの議論から,本実験で用いた環境においては,1440×960サイズの動画像は問題なくソフトウェアで再生することができると考えられる.分割格子数は128以下であればフレームレートが減少することなく再生ができた.透視投影変換の計算も,分割格子数が128までならば問題なく行えた.ここで,二つの変換式において画像サイズが2倍のときは,それぞれフレームレートが28.3fps,25.1fpsとなったが,視点変更中はフレームレートよりも,クライアントが望む視野の発見が目的であるのでそれほど影響はないと考える.また,多チャンネルで部分画像を配信する際には,サイズとチャンネル数とビットレートの選択が重要となる.

(4) まとめ

本研究では視点情報に基づいて,高解像度高精細の動画像を配信するシステムとそれを支援する ストリーミング再生クライアントを構築した.フリーウェアである MP1ayer を改良して,サーバと 視点情報をやりとりするプレイヤを作成する事ができた.また,OpenGL を用いて,全方位画像を透 視投影変換してクライアントに提示する事ができた.そして,多チャンネル再生の実現のための機能として,複数のファイルを読み込み,画面を分割して同時に再生する機能を実装した.さらに,再生する画像サイズとフレームレートの関係,テクスチャマッピングを行うときの分割格子数とフレームレートの関係,そして透視投影変換がフレームレートに与える関係について評価実験を行い,実際のシステムにおけるストリーミング再生の実現可能性について検討した.その結果,1440×960のサイズの画像を分割格子数128で透視投影変換しながらストリーミング再生可能であると判断できた.また,各処理にかかる時間をビットレートを考慮して計測し,多チャンネル配信の実現可能性について検討した.今後は多チャンネル配信に対応したプロトコルの作成と,実際にネットワークを介して多チャンネルで配信された複数ストリームに対して,部分画像を合成して再生できるようにする予定である.

4-5-2 配信サーバの開発

低解像度全方位動画像配信なら、従来のストリーミングサーバーは困難なく実現できるが、解像度 とフレームレートが高くなると、通信帯域の需要も増大する.インターネットの通信帯域は有限 であるため、通信帯域の限界を超える高解像度動画像を配信するには転送データを削減する必要 がある.本研究で提案する配信システムでは、クライアントが見たい方向の高解像度動画像だけ をクライアントに配信することで、解像度とフレームレートを確保できる.また、視点(見たい 方向)の指定より、サーバー・クライアント間のインタラクティブ配信を実現できる.

(1) 本研究の目的と目標

本研究で提案する配信システムでは、全方位カメラから撮られた動画像はサーバーを通し、イン ターネット上にある各クライアントに転送される.クライアントが指定する視点情報により、サー バーは指定された方位の動画像をクライアントに転送する.そして各クライアントは同時に異なる 方位の動画像を見ることができる.図46は本研究のシステム概略図を示す. 本研究の目標は以下の3つである:

- 1. クライアントからの要求を受付
- 2. 動的に要求した動画像データを多チャンネルで配信
- 3. 強固なサーバーと配信品質を保証



(2) 多チャンネル配信の提案手法

全方位カメラから撮られた動画像は映像処理サーバーで複数の小さな低解像度動画像に分割して、ストリーミング配信サーバーの動画像データベースに保存される.そして配信サーバーを介し、クライアントに転送される.基本的なアイデアはストリーミングサーバーがクライアントに要求された視点情報に対応する多数の低解像度動画像をクライアントに転送するというものである.

本研究では全方位動画像を多チャンネルで配信するため、サーバー・クライアント間の通信プロト コル(RTSP)の設計とストリーミングサーバーの配信モジュールの構築を行う.通信の概略を図4 7に示す.



図47 リーハー・クノイノント通信慨略回

(3) 多チャンネル配信のためのプロトコル設計

本節では、多チャンネル配信するため、サーバー・クライアント間のセッションコントロールプロトコルの設計について述べる.本研究のセッションコントロールプロトコルは、既存の RTSP Method をベースとし、多チャンネル配信に必要なサーバー・クライアント間の情報やり取りへ RTSP を拡張することで設計する.

a. 本研究の RTSP の流れ

一般的なストリーミングサーバーとクライアント間の RTSP の流れは前節で述べた. OPTION Method から PLAY Method まで、クライアントとサーバーの間は使用可能な Method 交換、メディア の記述情報の請求と応答、セッションのセットアップ、メディア再生一連のやり取りを行ってい



図48 本研究のRTSPの流れ

る.本研究の多チャンネル配信を実現する RTSP プロトコルも基本的にこのように考えられている が、全方位動画像を多チャンネルで配信するため、視点情報を載せる RTSP Method の SET_PARAMETER を追加し、そして各段階の RTSP Method の役割を多少増加させる. RTSP プロトコルの流れを図48 に示す.

1. OPTION

一般的なストリーミングサーバーはクライアントの OPTION 請求により,サーバーが対応する RTSP Method をクライアントに応答する.本研究では,さらにサーバーが配信できる全方位ミラー の動画像の種類をクライアントに返す.そして,クライアントは受信したい全画像を選択し,次の RTSP Method へ進める.

2. DESCRIBE(全方位動画像に対して)

一般的なストリーミングサーバーはクライアントの 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応答する.本研究では、クライアントの DESCRIBE 請求に含まれている全方位動画像タイプから、クライアントが受信したい全方位動画像の 全体的な記述情報をクライアントに返す.記述情報の中には、全方位カメラのミラーパラメータ、 低解像度動画像への分割情報、全方位動画像を構成したすべての低解像度動画像のファイルリス ト、各低解像度動画像ファイルの位置とサイズ情報が記述されている.

3. SET_PARAMETER

SET_PARAMETER Method はクライアントからサーバーにある変数とその値を渡す時とセッション を維持するためによく使われている.本研究では、SET_PARAMETER を介して、クライアントからサ ーバーに視点情報を送る.サーバーは視点情報より、対応するファイルのリストをクライアントに 返す.

4. DESCRIBE(単一の低解像度動画像に対して)

単一の低解像度動画像の記述情報を取得するため、本研究の DESCRIBE Method は一般的なストリーミングサーバーの DESCRIBE Method と変わらない.ただし、SET_PARAMETER 請求で受信したすべてのファイルリストの各ファイルに対し、複数の DESCRIBE 請求を行うことになっている.

5. SETUP

本研究も同じように SETUP Method でセッションをセットアップする.ただし,SET_PARAMETER 請求で受信したすべてのファイルリストの各ファイルに対し,複数の SETUP で同じセッションに複数のストリームをセットアップする.

6. **PLAY**

本研究では PLAY 請求で, セッションに含まれているすべてのストリームを同時に配信する.

クライアントは視点を変えたい時に,SET_PARAMETER → DESCRIBE(数回で単一の低解像度動画像に対する) → SETUP(数回で単一の低解像度動画像に対する) → PLAY の順番で視点を変え,新たな 視野範囲に対応するメディアファイルを受信することができる.

(4) ストリーミングサーバーの実装

ストリーミング技術が発明されてから、多くのストリーミングプラットフォームが提出された. それらの中でも、Microsoft社のWindows Media Real Networks社のReal MediaとApple社のQuickTime は業界で最も有名な 3 つのストリーミングプラットフォームである.本研究で用いる Darwin Streaming ServerはApple社のQuickTime Streaming Serverの開発者向けオープンソースバージ ョンである. Darwin Streaming Serverはサーバーの機能を自由に拡張でき、本研究においては一 番大きなメリットである.他のストリーミングサーバーと比べると、無料である、サポート対象で あるプラットフォームが多い、配信品質が高い、業界標準プロトコル(RTSP, RTP, RTCP, SDP)を サポートしている、オープンソースであるなどの利点がある.以上の理由により、Darwin Streaming Server を本研究の配信サーバーとして用いる.

a. Server アーキテクチャ

DSS は1つのペアレントプロセスといくつかのチャイルドプロセスから Core Server を構成して いる. Core Server はネットワークとクライアントのインタフェースであり, RTP と RTSP を介して, 請求と応答を行う. Server Module はクライアントの請求を処理し, クライアントにパケットを転 送する. 4つのタイプのスレッドで Core Server を動かす. 図49はサーバーのアーキテクチャを 示す. 多プラットフォームのサポート, そしてオープンソースである DSS は開発者向けであり, モ ジュール二次開発のインタフェースにより, ストリーミングサーバーの機能を拡張できる. このイ ンタフェースを利用し, プロトコルの拡張ができ, また特定な配信方式も実現できる. DSS の多く の重要な機能もモジュールとしてサーバーのソースに記述されている. モジュールにはそのコンパ イル方式により, スタティックモジュールとダイナミックモジュールがある. スタティックモジュ ールはサーバーをコンパイルする時に同時にサーバーにコンパイルされるモジュールである. ダイ ナミックモジュールはサーバーが起動する時,サーバーにロードされるモジュールである.サーバ ーが起動する時,ダイナミックモジュールの方が先にサーバーにロードされる.モジュールの二次 開発に対し,ダイナミックモジュールの方が良いと考えられている.



図 4 9 Darwin Streaming Server Architecture

サーバーは最初にダイナミックモジュールをサーバーにロードする、そしてスタティックモジュ ールをロードする.サーバーの機能を拡張したい場合、ダイナミックモジュールにコンパイルする ことが薦められている.よって、本研究はモジュールの二次開発により、RTSPの拡張と多チャンネ ル配信機能を実現するダイナミックモジュールを開発する.

b. モジュールの二次開発

本研究では、ダイナミックモジュールの二次開発により、サーバーの機能を拡張する.本研究で 目的とする機能は主にモジュールに配置した QTSS_RTSPRequest_Role と QTSS_RTPSendPackets_Role により実現する. QTSS_RTSPRequest_Role は視点情報の解析, RTSP Response, RTSP Session Setup, RTSP Session Update などを行う. QTSS_RTPSendPackets_Role は多チャンネル配信, RTP Session

Setup, RTP Session Update などを行う.以下にモジュールにある各 Role の説明を述べる.

- 1. QTSS_Register_Role:モジュールがサポートする Roles をサーバーに登録する. 多チャンネル 配信に必要なオブジェクトもここで定義する.
- 2. QTSS_Initialize_Role: モジュールが支持する RTSP Method をサーバーに登録する. 本研究の クライアント視点情報を指定するための SET_PARAMETER Method を追加する.
- 3. QTSS_RereadPrefs_Role:モジュールの内部パラメータを初期化する.全方位カメラのミラーパ ラメータと全方位動画像の記述情報を外部ファイルから読み込む.
- 4. QTSS_RTSPRequest_Role: 拡張した RTSP をクライアントから受け取る. そして対応する応答を クライアントに返す. 主にサーバーは SET_PARAMETER Method の中の視点情報に対応するファイ ルリストをクライアントに返し、複数の SETUP Method で複数のストリームを1つのセッション にセットアップし、PLAY Method でセッションに含まれているすべてのストリームを配信する.
- 5. QTSS_RTPSendPackets_Role: 多チャンネルに対応する配信スケジュールで同時に複数のストリ ームをクライアントに配信する.
- 6. QTSS_ClientSessionClosing_Role: サーバーとクライアントのセッションを断ち切る.

RTSP 請求の処理は前節で述べた流れの通りに行う.ただし,一般的なメディアファイルも配信 できるようにするため,クライアントの請求に含まれるキーワードより,一般的なクライアント か,全方位動画像を受信したいクライアントかを判断する必要がある.全方位動画像を受信したい クライアントとして判断したら、QTSS_RTSPRequest_Role は図50のように RTSP 請求を処理する. OPTION からサーバーは請求が全方位動画像に対するかどうかを判断する.全方位動画像に対するな ら,サーバーはサポートする全方位動画像タイプをクライアントに返す.次に DESCRIBE も同じよ うに、全方位動画像に対するなら、対応する全方位動画像の記述情報をクライアントに返す.一般 のメディアファイルに対するなら、メディアファイルの記述情報を返す.次に SET_PARAMETER に含 まれている視点情報より、対応する方向のファイルリストをクライアントに返す.不正な視点情報



図50 拡張した QTSS_RTSPRequest_Role

なら、エラーを返す.そして、各ファイルに対し、DESCRIBE と SETUP を行う、単一ファイルについては、一般的なメディアファイルと同じように、DESCRIBE と SETUP 請求を処理する.最後に全方位動画像に対する PLAY 請求により、多チャンネル配信を開始する.

c. 多チャンネルストリーム配信

RTP の RFC により,多重配信は複数のストリームをクライアントの複数のポートに送信すること が薦められている. クライアントは異なる SSRC より各ストリームを識別する,各ストリームが対応する RTP チャンネルと RTCP チャンネルにより,パケット受信と通信状況の報告を行う.

クライアントは SETUP で、各ファイルの受信 RTP 番号と RTCP 番号をサーバーに通知する.そし て、PLAY Method よりサーバーはクライアントが設定した受信番号にその対応する各ストリームパ ケットを順番に送信する.各ストリームはそれぞれの異なる SSRC 同期源番号を持っている.クラ イアントはその SSRC 番号により、各ストリームを区別する.また、サーバーは PLAY に対するの応 答の中に、各ファイルのはじめの RTP PACKET 番号と RTP PACKET のサンプル時間情報を含めてクラ イアントに返す.この情報によって、クライアントはパケットから動画像に復元する事ができる.

各ストリームの配信スケジュールはストリーム1の最初のパケットから,順番に各ストリームの パケットをクライアントに転送する.つまり,サンプル時間が同じ様な各ストリームのパケットを 一緒にクライアントに転送する.このように配信することで,クライアント側は同時刻の低解像度 動画像を同期化することができる.

d. 視点情報からファイルリストへの計算

視点情報の4つのパラメータ(水平方向角度,垂直方向角度,水平視野範囲角度,垂直視野範囲 角度)から,実空間の四角形状の透視投影画像領域を表示することができる.図51に視点と各パ ラメータの関係を示す.透視投影面を全方位画像に投影することで,透視投影画像作成に必要な透 視投影領域を求めることができる.図52に透視投影面と全方位画像中の透視投影領域の関係を示 す.視点情報が対応する透視投影領域を計算できれば、転送するファイルリストを計算できる.分 割された格子状全方位動画像の各動画像が透視投影領域に含まれるなら、転送リストに追加され る.図53に転送リストの生成を示す.各低解像度メディアファイルの命名方法は一番左上のファ



イルを 0_0 として,右に行くと,0_1,0_2,0_3 のように,下に行くと 1_0,2_0,3_0 のように各 ファイルの名前をつける.そして RTSP の DESCRIBE と SET_PARAMETER でファイルリストをクライア ントに返す,全方位動画像の分割方法より,各ファイルの位置情報とサイズ情報もつける.

(4) まとめ

本研究では、全方位、高精細高解像度、高速としてリアルタイムの動画像を高能率でストリーミング配信するシステムの構築を目的として、Darwin Streaming Server のモジュール二次開発に基づくサーバー・クライアント通信プロトコルの設計、多チャンネルでクライアントの視点情報に対応する方向の動画像をクライアントに配信する方法を開発した。一般的なストリーミングサーバーにとって、1つのクライアントに対する配信できるチャンネル数は制限されるため、全方位高精細動画像を1 つチャンネルで全部クライアントに転送するのは通信帯域が足りないという問題がある。この問題に対して、本研究では、高解像度全方位動画像を複数の格子状低解像度に分割し、多チャンネルストリーミング配信によって、クライアントからの視点情報に対応する方向の複数の低解像度動画像を同時にクライアントに転送するという配信手法を構築した。

4-6 ライブ入力,配信システムの開発

現在ソフトウェアで行っている高解像度画像と低解像度画像の幾何・光学補正をハードウェアで 行うために, FPGA 画像処理ボード(Tsunami: SBS Technologies)を導入した. FPGA 画像処理ボー ドは Altera 社の FPGA を搭載し, PCI を通して PC と Cameralink インタフェイスを通してカメラと 通信する形をとる.本ボードを用いて,低解像度カメラの取り込みを部分の回路を実装し,画像取 り込みのテストを行った.

(1) FPGA による実時間画像補正

画像補正処理は、ドットクロックの低い低解像度画像のみに加える.低解像度画像は画像処理ボ ードによって取り込まれ、ボード上にある FPGA で補正処理が加えられる.このボードにはカメラ インターフェース、FPGA、メモリが備えられている.FPGA として ALTERA 社の EP1S40 が搭載されて いる.画像処理ボードの仕様を表9に、FPGA の仕様を表10に示す.また、ハードウェア記述言語 として VHDL (Very high speed integrated circuit Hardware Description Language)を使用した. 得られた高解像度画像と補正された低解像度画像を PC に取り込んでディスプレイに表示させる. 処理の流れを図54に示す.FPGA 内での処理はカメラ入出力制御、画像補正、PCI データ転送制御 の3つのステップに分けられる.FPGA 内での処理の流れを図55に示す.カメラ入出力制御ブロッ クではカメラからのデータ入力および画像補正ブロックへのデータ出力の制御を行う.画像補正ブ ロックではソフトウェアで求めたパラメータを用いて輝度変換や幾何変換といった画像補正を行 う.PCI データ転送制御ブロックでは PC へのデータ転送の制御を行う.カメラ入出力制御や PCI データ転送制御は既存のコントローラで行えるため、画像補正処理の実装手法について述べる.



	表 9	画像処理ボー	ドの仕様
--	-----	--------	------

	Tsunami PCI A40
PCI bus interface (bit range/speed)	32bits/33MHz
local bus (bit range/speed)	32bits/50MHz
SRAM Memory	2MB per bank, 4MB total
SDRAM Memory	512MB per bank, 1024 MB total

表 I O FPGA の仕様			
	EP1S40		
Total Logic Elements	41250		
TriMatrix memory, total RAM bits	3423744		
PLLs	8 (4 Enhanced, 4 Fast)		
DSP blocks	14		
Embedded multipliers	112		

カメラの CCD は Bayer 配列となっており,1 画素ごとに RGB のいずれかのデータのみが入力とし て与えられる.なお,後の説明のためG成分はG1,G2 と表記することにする.アドレスジェネレー タでは入力データの数からその座標を計算する.ただし,画像左上隅を原点とし,水平方向右向き に x 軸,垂直方向下向きに y 軸をとるものと考える.入力データは原点から水平方向右向きに順に 与えられる.セレクタで取り込んだ入力データをその座標から B,G1,G2,Rのいずれのデータで あるかを判別し,対応する出力信号線に送る.そのデータに輝度変換ブロックで輝度補正を加える. 輝度補正後のデータはその座標とともに B,G1,G2,Rごとに用意したバッファに保存される.ま た一方で,出力データの座標に幾何変換ブロックで幾何補正を加え,その座標に対応する入力デー タの座標を求める.求められた座標のデータをバッファから取り出し出力する.出力データの座標 は高解像度観測画像平面に投影された低解像度画像の座標であり,幾何変換により投影前の元の低 解像度観測画像上の点の座標を求める.

試作された複合センサカメラでは2つのカメラの位置調整を手動で行っている.そのため,2つ のカメラ間で光軸方向や光軸垂直方向にずれが生じる.このずれにより高解像度観測画像平面上の 直線を低解像度観測画像平面に投影した際,その2直線の傾きは必ずしも一致しない.さらに,レ ンズ歪みの度合いの差から高解像度観測画像平面上の直線は糸巻き型に歪曲して低解像度観測画 像平面に投影される.これらのことから,高解像度観測画像平面上の水平な直線,すなわち y 座標 値を一定とした直線の座標から対応する低解像度観測画像上の点の座標を求めると,その y 座標値 は一定とはならず,ある変動幅を持つ.従って,バッファを用いることでこの幾何変換におけるず れ幅を許容する.必要となるバッファの大きさは複合センサカメラの位置調整の精度によって異な る.本研究では FPGA 内の RAM の大きさの制約から B, G1, G2, R それぞれに 32 ライン分のデータ が保存できるようにした.

輝度変換は B, G1, G2, Rごとに先に求めた変換テーブルを用意し,そのテーブルによって変換 を行う. 幾何変換は高解像度観測画像平面に投影された低解像度画像の座標から投影前の元の低 解像度観測画像上の点の座標を演算によって求めることで行われる.この演算におけるそれぞれの 変数の値は bit vector 型で表現され,固定小数点数として表される.固定小数点数で表現するこ との利点として,情報落ちが起こらないことや演算が高速に行えることが挙げられる.VHDLでは, 除算は integer 型同士でのみ定義されているため,変数のデータタイプを bit vector 型から integer 型に変換してから行う.integer 型は符合付き 32 ビット長の整数値である.除算を行う際 は桁落ちの影響を少なくするため,integer 型への変換前に被除数に対して MSB (Most Significant Bit)側へのシフト演算を行う.しかし,先に述べた integer 型のビット長の制限から十分な有効桁 数が得られるだけのシフト演算が行えず,除数が大きな値である場合は除算の結果大きく桁落ちす る.この桁落ちにより幾何変換後の座標に大きな誤差が生じる.そこで,除算による桁落ちの影響 を少なくするため,先に数式をまとめて除算回数を減らす.また,演算速度を向上させるために, 定数はまとめて,あらかじめ求めておく.

(2) まとめ

本研究では、FPGA を用いることで複合センサ画像の補正処理を高速に行うことで、ライブ入力 を実現する手法を提案した.FPGA 内での画像補正処理の流れについて述べ、輝度補正や幾何学的補 正の方法について述べた.

今後は、実際に FPGA を用いて補正処理を行い、ソフトウェアを用いた場合と比較しての処理の 速度や精度といった性能の評価を行う予定である.高解像度高フレームレート生成手法を FPGA を 用いて行うことで、実時間での高解像度かつ高フレームレートな動画像の生成を目指す.

4-7 マルチアクセス時の性能評価

本章では、本研究で提案した全方位高精細動画像配信システムの実際配信能力の評価を行う.

4-7-1 マルチアクセス実験環境

本研究が使うストリーミングサーバー(ハードウェア)はHewlett-Packard(HP)社のProLiant DL560 である.データ解析など高度な計算能力を実現するために設計され,高いプロセッサパワーを持 ち,本研究のストリーミングサーバーとして使われている.詳しいスペックは表11に示す.オペ レーティングシステムについては,本研究はRed Hat Enterprise Linux を用いる.サーバーのパ フォーマンスやスケーラビリティを向上させ,Red Hat Enterprise Linux は,ハードウェアとソフ トウェアに幅広く対応する最新のLinux環境となる.Linuxの2.6.9カーネルを基に強力で安定し たサーバー環境を提供できる.また,本研究ではApple社のQuickTime Streaming Serverのオー プンソースバージョンであるDarwin Streaming Serverを配信サーバーとして用いる.本研究で用 いたバージョンはDSS-v5_0_3_2である.テスト段階で,十分な通信帯域を確保するため,ネット ワーク環境については,LAN (Local Area Network)において,テストを行う.ネットワーク構成 図を図56に示す.

System unit	ProLinant DL550		
CPU	Intel Xeon MP 3.0GHz		
Number of CPUs	2 CPUs		
Cache memory	4MB L3 cache		
Main memory	1GB 200MHz DDR SDRAM		
HDD	80GC		
Interface	Serial, Keyboard, Mouse, Monitor, RJ-45, USB		
Network Interface	NC7781 PCI-X Gigabit server adaptor		
	10/100/1000 (onboard) x 2,		
	10BaseT/100BaseT/1000Base-T		

表11 マルチアクセス実験に使用した計算機の仕様

4-7-2 配信能力の評価

実験用の全方位動画像については、全方位画像のサイズは 2560 画素×2560 画素とし、格子状の 1024 (32×32) 個の低解像度動画像へ分割されているものとする. 各低解像度動画像の画面サイズ は 80 画素×80 画素, MPEG4 フォーマット形式で圧縮されている. 単一ファイルの転送ビットレートは 50Kbps 程度,各ファイルの再生時間は 60 秒とする. 全方位画像の特徴により、視野範囲の大きさは同じ(水平視野範囲角度が同じ,かつ垂直視野範囲角度も同じの異なる透視投影面)でも、全方位画像に対応するファイル数は異なる. 全方位画像の中心に近づくほど対応するファイル数は少なくなる. 逆に全方位画像の周辺に近づくほど対応するファイル数は多くなる. この性質から,実験は視点の仰角が-30,0,30 度の 3 つの場合について行った. 各視点の視野角は水平 30 度,垂直 30 度とした.

以上の条件においてクライアント数,各クライアントごとのストリーム数,サーバーの配信ビットレート,配信の遅延時間について,レイヤーなしとレイヤーありの2つの画面分割方法で,サーバーの配信能力の評価を行った.実験結果を以下の表に示す.

(1)レイヤーなし分割方法
 View Point: 0_30_30_30
 各クライアントのストリーム数:63
 各クライアントのデータ転送量:58.181MB



		-		
クライアント数	ストリーム数	転送ビットレート	平均転送時間	平均遅延時間
1	63	0.969 MB/sec	60.25 sec	0.25 sec
4	252	3.878 MB/sec	60.35 sec	0.35 sec
10	630	9.696 MB/sec	60.33 sec	0.33 sec
14	882	13.58 MB/sec	60.35 sec	0.35 sec
17	1071	16.48 MB/sec	60.33 sec	0.33 sec
18	Too much streams			

表12 View Point 0_30_30_30 におけるサーバ配信能力

View Point: 0_0_30_30

各クライアントのストリーム数:22

各クライアントのデータ転送量: 15.637MB

表13 View Point 0_0_30_30 におけるサーバ配信能力

クライアント数	ストリーム数	転送ビットレート	平均転送時間	平均遅延時間
1	22	0.260 MB/sec	60.51 sec	0.51 sec
4	88	1.042 MB/sec	60.27 sec	0.27 sec
10	220	2.606 MB/sec	60.24 sec	0.24 sec
30	660	7.818 MB/sec	60.22 sec	0.22 sec
50	1100	13.03 MB/sec	60.23 sec	0.23 sec
51	Too much streams			

View Point: 0_-30_30_30

各クライアントのストリーム数:10

各クライアントのデータ転送量: 7.48MB

表14 View Point 0_-30_30_30 におけるサーバ配信能力

		· ·		
クライアント数	ストリーム数	転送ビットレート	平均転送時間	平均遅延時間
1	10	0.124 MB/sec	60.08 sec	0.08 sec
20	200	2.493 MB/sec	60.18 sec	0.18 sec
40	400	4.986 MB/sec	60.18 sec	0.18 sec
80	800	9.973 MB/sec	60.18 sec	0.18 sec
105	1050	13.09 MB/sec	60.19 sec	0.19 sec
106	Too much streams			

(2) レイヤーあり分割方法

各クライアントのス	ストリーム数:15			
各クライアントの	データ転送量: 13	.648MB		
	表15 View Pe	oint 0_30_30_30 におり	ナるサーバ配信能力	
クライアント数	ストリーム数	転送ビットレート	平均転送時間	平均遅延時間
1	15	0.227 MB/sec	60.05 sec	0.05 sec
16	240	3.639 MB/sec	60.07 sec	0.07 sec
36	540	8.189 MB/sec	60.09 sec	0.09 sec
64	960	14.53 MB/sec	60.09 sec	0.09 sec
70	1050	15.92 MB/sec	60.09 sec	0.09 sec
71	Too much streams			

View Point: 0_30_30_30

View Point: 0_0_30_30

各クライアントのストリーム数:8

各クライアントのデータ転送量: 7.142MB

表16 View Point 0_0_30_30 におけるサーバ配信能力

クライアント数	ストリーム数	転送ビットレート	平均転送時間	平均遅延時間
1	8	0.119 MB/sec	60.04 sec	0.04 sec
20	160	2.380 MB/sec	60.05 sec	0.05 sec
40	320	4.760 MB/sec	60.06 sec	0.06 sec
80	640	9.520 MB/sec	60.06 sec	0.06 sec
126	1008	14.99 MB/sec	60.06 sec	0.06 sec
127	Too much streams			

View Point: 0_-30_30_30 各クライアントのストリーム数:8

各クライアントのデータ転送量: 6.819MB

表17 View Point 0-30 30 30 におけるサーバ配信能力

クライアント数	ストリーム数	転送ビットレート	平均転送時間	平均遅延時間
1	8	0.114 MB/sec	60.03 sec	0.03 sec
20	160	2.280 MB/sec	60.05 sec	0.05 sec
40	320	4.560 MB/sec	60.06 sec	0.06 sec
80	640	9.120 MB/sec	60.05 sec	0.05 sec
128	1024	14.59 MB/sec	60.05 sec	0.05 sec
130	Too much streams			

実験の結果により、サーバーの配信能力が示されている.サーバーとクライアント間の RTSP の やり取りも確認した.レイヤーありとレイヤーなしの画面分割方法の2つの場合に、サーバーの配 信能力の比較により、レイヤーありの画面分割方法は、サポートする最大クライアント数がより多 い、転送遅延時間がより短いという利点がある.無駄の配信領域が生ずるが、クライアントの受信 帯域に越えないを条件として、ストリーム数を減らすことで、サーバーの配信能力を高めることが できる.今回行った実験では、ストリーシングサーバーの最大転送ビットレートに超えてないが、 サーバーが配信できる最大ストリーシ数は1000 であることが結論として得られた.レ イヤーあり分割方法とレイヤーなし分割方法のサーバーパフォーマンスの比較を図57に示す.た だし、本研究のストリーミングサーバーを実際に応用すれば、配信環境はもっと複雑になる.例え ば、各クライアントが持っている視点情報は同じではないので、各クライアントに対するストリー ム数は異なる.またインターネットにある各クライアントはサーバーに接続する経路が違うなどの 要素がある.本研究に対する実験評価は最も理想的な配信環境でストリーミングサーバーの最大配 信能力の計測にあたる.

以上,本研究が提案したストリーミングサーバーの実際配信能力の評価を行った,サーバー・クライアント間のRTSPやり取りとサーバーの最大配信能力を確認した.

4-8 統合システムの試作と評価

日本の神社や仏閣、古墳等を高精細にディジタルアーカイブ化し、それらを実際に配信サービス を行う実証実験を行い、本提案技術の有効性の検証とその一般への周知を行う。平成16年度に試 作した複合センサカメラとリアルタイム画像入力記録システムを用いて、近畿周辺の神社や仏閣な どで高精細画像コンテンツの撮影を行う。これらの撮影データを用いて高解像度画像生成実験や画 像配信実験を行う。また、学会や映像機器の展示会などに本システム持ち込み、実機デモンストレ ーションを行うことで、本提案技術の一般への周知を行う。

本年度は、試作した複合カメラとリアルタイム画像入力記録装置を、CEATECの展示、セキュリティ展などの企業、一般向けのデモ展示した.また、商業誌へ本システムの解説記事を載せ、一般向



けの周知を行った.さらに、電子情報通信学会主催の画像と認識理解シンポジウム(MIRU2005)において、デモ展示とポスター展示を行い、画像処理技術者向けに本提案技術の周知を行った.また、 淡路島などの近畿周辺において撮影テストを行った.

4-9 総括

各サブテーマ毎に、今年度の研究成果と今後の課題および最終目標に対する達成状況を示す。

1) 複合センサカメラの開発

今年度は、市販の高解像度低フレームレートカメラ(4008×2672 画素、4.29 フレーム/秒)と低解 像度高フレームレートカメラ(1008×1018 画素 30 フレーム/秒)を用いた複合センサカメラの評価 を行ない最適なパラメータを推定した。本試作センサは 1:7 の割合で2 種類の動画像を同期撮像で きる。この試作したカメラを実際に利用して屋内外での撮像実験を行った。現状においては、最終 目標である 8000×8000 画素の CCD が出回っておらず、4008×2672 画素が最大性能である。そ こで、今後の課題としては、高解像度の CCD が利用できるようになったときに、スムーズに移行 できるように、複合センサカメラのキャリブレーションの精度を上げることである。

2) リアルタイム高解像度動画像の作成

リアルタイム高解像度動画像の作成では、試作した複合センサカメラから得られる2つの動画像 から、空間的にも高解像度で、しかも時間的にも密な動画像(4000x4000 画素、30 フレーム/秒) を作成する技術を確立した。時空間解像度の異なる2つの画像系列(高解像度低フレームレート動 画像、低解像度高フレームレート動画像)の情報を統合する手法として周波数アプローチとモーフ ィングアプローチの2つの手法を考案し、それらのソフトウェアを実装し、MPEG動画像による PSNR評価と、試作した複合センサによる実画像の高解像度化実験を行った。これにより、最終目 標である高解像度高フレームレート画像を生成する手法を完成することが出来た。今後の課題とし ては、8000x8000 画素の高解像度シミュレーション画像を用意し、最終的な性能評価を行なうこと である。

3) 高解像度全方位カメラ用光学系の設計

高解像度全方位カメラ用光学系の設計では、複合センサカメラ用に全方位ミラーを設計し、全方 位複合センサカメラを試作するのが目標である。今年度は、試作した複合センサカメラと全方位ミ ラーを用いて、全方位複合センサカメラを構成した。これにより複合センサカメラを全方位撮像す ることが出来、最終目標に到達した。

4) 全方位高解像度リアルタイム動画像入力記録方式の開発

全方位高解像度リアルタイム動画像入力記録方式の開発では、複合センサカメラから得られる2 つの全方位動画像から、空間的にも高解像度で、しかも時間的にも密な全方位動画像(4008×2672 画素、30 フレーム/秒)を、入力および記録する技術を開発している。前年度試作したシステムで は数秒に1フレーム程度コマ落ちが発生したので、今年度は、複合センサカメラの2つの出力を、 取りこぼしなくリアルタイムで同期記録できるようにシステムを改良した。これにより、現状のお いては目標に到達した。今後の課題としては、8000x8000 画素の高解像度の CCD が利用できるよ うになったときに、スムーズに移行できるように、ソフトウェアの性能向上を行なう事などが挙げ られる。

5) 高解像度動画配信用ソフトウェアの開発

高解像度動画配信用ソフトウェアの開発では、蓄積された2種類の全方位画像を圧縮してインタ ーネットにより複数の閲覧装置(クライアント)に配信するサーバシステムと、配信された2種類 の全方位動画像から閲覧者側において高解像度でかつリアルタイム(30 フレーム/秒)の動画像を 作成する技術を開発する。また、全方位画像のうち、閲覧者の要求する視野の部分画像のみを切り 出しそれらを能率良く配信し、クライアント側で動画像表示するシステムを実現する。今年度は、 Streaming Server を利用した配信サーバの構築、および OpenGL を利用した全方位提示クライア ントの試作を行った。今後は、試作クライアントと配信サーバの相互接続実験を繰り返して、性能 評価を行なう予定である。また、最終目標に対しては、高解像度で配信する技術を確立したので、 8000x8000 画素の高解像度シミュレーション画像を用意し、最終的な性能評価を行なう。

6) ライブ入力、配信システムの開発

ライブ入力、配信システムの開発では、ライブ入力のために現在ソフトウェアで行っている画像 補正のハードウェア化による高速化を行った。今後は、ハードウェア回路の改良を重ねて安定して 画像補正を行なえるようにする。最終目標に対して計画どおり順調に進んでいる。

7) マルチアクセス時の性能評価

マルチアクセスの時の評価では、前年度構築した配信システムと、今年度作成したソフトウェア を用いて配信サーバの評価を行った。これにより、配信システムの最大性能が明らかになった。今 後は、8000x8000 画素の高解像度シミュレーション画像を用意し、最終的な性能評価を行なう。最 終目標に対して計画どおり順調に進んでいる。

8) 統合システムの試作と評価

統合システムの試作と評価では、配信サーバと試作クライアントの結合評価を行なった。今後は、 さらに実機デモンストレーションを行い、一般への周知を行う。また、最終目標に対しては、平成 18年度に、コンテンツ撮影を行いディジタルアーカイブを構築する予定である。

5 参考資料・参考文献

5-1 研究発表・講演等一覧

[1]松延徹、星川章、重本倫宏、渡辺清高、長原一、岩井儀雄、谷内田正彦、田中紘幸「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・理解シンポジュウム (MIRU2005) JULY. 2005"

[2]渡辺清高、岩井儀雄、長原一、谷内田正彦、田中紘幸、「ウェーブレット領域での動き補償と画 像統合による高解像度高フレームレート動画像の生成」"画像の認識・理解シンポジュウム (MIRU2005) JULY. 2005"

[3]Akira Hoshikawa, Tomohiro Shigemoto, Hajime Nagahara, Yoshio Iwai, Masahiko Yachida, Hiroyuki Tanaka 「Dual Sensor Camera System with Different Spatio-Temporal Resolution」"SICE Annual Conference 2005" Aug. 2005

[4] Kiyotaka Watanabe, Yoshio Iwai, Hajime Nagahara, Masahiko Yachida, Hiroaki Tanaka 「Video Synthesis wiwh High Spatio-Temporal Resolution Using Motion Compensation and Spectral Fusion」"SICE Annual Conference 2005" Aug. 2005

[5]T. Matsunobu、H. Nagahara、Y. Iwai、M. Yachida、H. Tanaka 「Generation of High Resolution Video Using Morphing」"SICE Annual Conference 2005" Aug. 2005 [6] 間下以大,岩井儀雄,谷内田正彦 "円錐曲線を用いた全方位視覚センサのキャリブレーショ ン",情報処理学会 CVIM 研究会論文誌 (採録決定)

[7] Tomohiro Mashita, Yoshio Iwai, and Masahiko Yachida, "Calibration Method for Misaligned Catadioptric Camera", IEICE Trans. on Information & Systems, (採録決定)

[8] Hajime Nagahara, Akira Hoshikawa, Tomohiro Shigemoto, Yoshio Iwai, Masahiko Yachida, Hiroyuki Tanaka, "Dual-Sensor Camera for Acquiring Image Sequences with Different Spatio-temporal Resolution", Proc. IEEE Int. Conf. Advanced Video and Signal based

Surveillance, pp. 450-455, Como, Italy, Sep. 2005.

[9] 重本 倫宏, 星川 章, 長原 一, 岩井 儀雄, 谷内田 正彦, 田中 紘幸, "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", 情報処理学会論文誌: コンピュータビジョンと イメージメディア(採録決定)

[10] 渡邊清高,岩井儀雄,長原一,谷内田正彦,鈴木俊哉, "ウェーブレット領域での動き補償 と画像統合による高解像度高フレームレート動画像の生成",情報処理学会論文誌:コンピュータ ビジョンとイメージメディア(投稿中・条件付き採録)

[11] Kiyotaka Watanabe, Yoshio Iwai, Hajime Nagahara, Masahiko Yachida, Toshiya Suzuki, ``Video Synthesis with High Spatio-Temporal Resolution Using Motion Compensation and Spectral Fusion'', IEICE Trans. Inf. & Syst. (採録決定)

[12] Kiyotaka Watanabe, Yoshio Iwai, Hajime Nagahara, Masahiko Yachida, Toshiya Suzuki, ``Video Synthesis with High Spatio-Temporal Resolution Using Motion Compensation and Image Fusion in Wavelet Domain'', Proc. Asian Conf. Computer Vision, pp. 480-489, Jan. 2006. [13] 岩井儀雄, 長原一, 鈴木俊哉, "全方位高解像度動画像配信システムの開発", 画像ラボ(日 本工業出版), Vol. 17, No. 3, pp. 60-65, Mar, 2006.