

平成18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： (株)映蔵

平成19年4月

情報通信研究機構

平成18年度 研究開発成果報告書 (一般型)

「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」

目次

1	研究開発課題の背景	3
2	研究開発の全体計画	3
2-1	研究開発課題の概要	3
2-2	研究開発目標	9
2-2-1	最終目標	9
2-2-2	中間目標	9
2-3	研究開発の年度別計画	10
3	研究開発体制	11
3-1	研究開発実施体制	11
4	研究開発実施状況	
4-1	複合センサカメラの開発	12
4-1-1	はじめに	12
4-1-2	複合センサカメラの試作	12
4-1-3	複合センサカメラのキャリブレーション	14
4-1-4	キャリブレーション評価実験	17
4-1-5	まとめ	17
4-2	リアルタイム高解像度動画像の作成	17
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	17
4-2-2	高解像度画像作成のためのPSF推定	20
4-2-3	モーフィングアプローチによる高解像度画像の作成	22
4-2-4	高解像度高フレームレート生成手法の有効性の検証	27
4-2-5	まとめ	30
4-3	高解像度全方位カメラ用光学系の設計	30
4-3-1	双曲面型全方位視覚センサ	31
4-3-2	二枚反射全方位視覚センサ	32
4-3-3	まとめ	34
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	35
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	35
4-4-2	書き込み速度の最適化	35
4-4-3	パルスジェネレータの作成	36
4-4-4	まとめ	36
4-5	高解像度動画配信用ソフトウェアの開発	37
4-5-1	受信クライアントソフトウェアの構築	37
4-5-2	配信サーバの構築	39
4-5-3	まとめ	42

4-6	ライブ入力, 配信システムの開発	42
4-6-1	FPGAによる実時間画像補正	42
4-6-2	まとめ	44
4-7	マルチアクセス時の性能評価	45
4-7-1	マルチアクセス実験環境	45
4-7-2	配信能力の評価	45
4-7-3	まとめ	47
4-8	統合システムの試作と評価	47
4-9	総括	49
5	参考資料・参考文献	
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、通常の高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) 高解像度だが時間的には粗い画像と、b) 通常の高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。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。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）レベルの高解像度で、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、今回提案する高解像度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高解像度を出すためには、通常の高解像度カメラであれば 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する。

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° の全方位の高解像度動画をリアルタイム入力し蓄積できるカメラの開発と、それを高効率で配信するシステムを開発を行う。

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) 通常の高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。

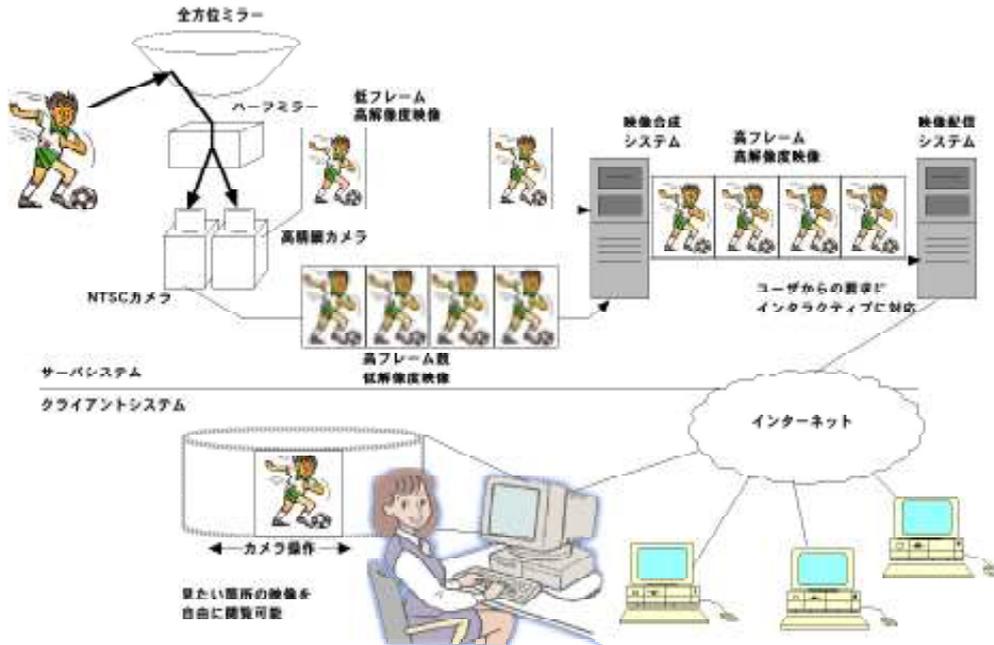


図1 全体システムのイメージ図

(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。

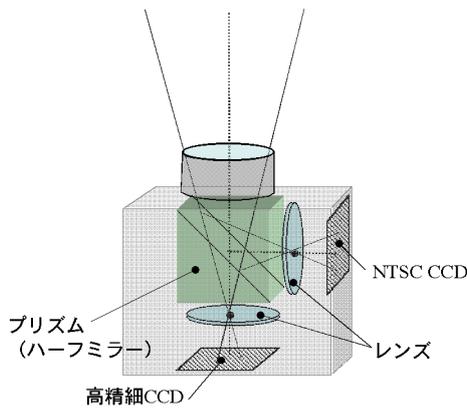


図2. 複合センサカメラ

(2) リアルタイム高解像度動画の作成：

上で試作した複合センサカメラから得られる2つの動画から、空間的にも高解像度で、しかも時間的にも密な動画(4000x4000画素、30フレーム/秒)を作成する技術を確認する。すでに、理論的な解析は終わっているので、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と、通常の解像度だが時間的に密な動画の2つの動画から空間的にも時間的にも密な動画を作成できることを実験的に確かめる。

高解像度動画の作成には、2台のセンサからの動画を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。低解像度カメラと高解像度カメラの2つの映像信号を、空間オーバーサンプリング、ディレイ、高解像度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、低解像度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、高解像度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における高解像度化処理は、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、低解像度カメラと高解像度カメラの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。

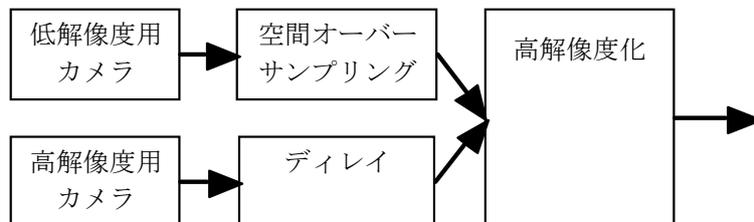


図3 時空間高解像度化処理

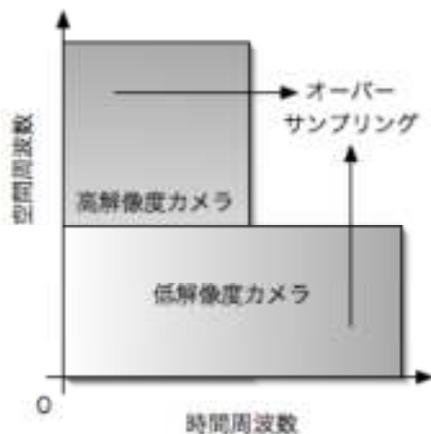


図4 各カメラの周波数帯域

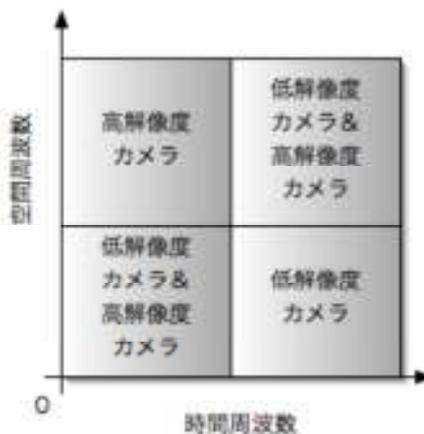


図5 時空間高解像度化後の周波数帯域

(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（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、通常的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、 $9\mu\text{m}$ 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画像 (4000×4000 画素、30 フレーム/秒) を作成する技術を確認する。このように、高解像度の実時間全方位動画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。

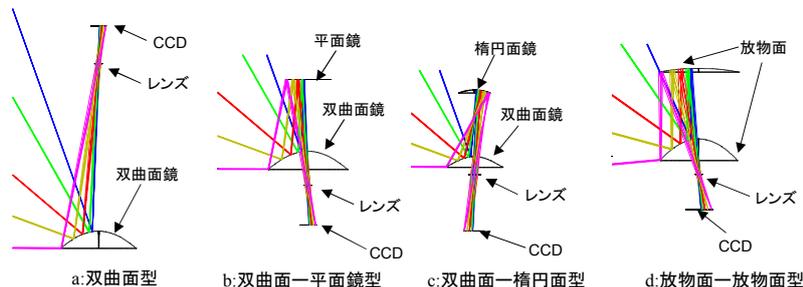


図6：全方位視覚センサの光学系

(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画像表示するシステムを実現する。

2種類の全方位動画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を受け取っているため、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。

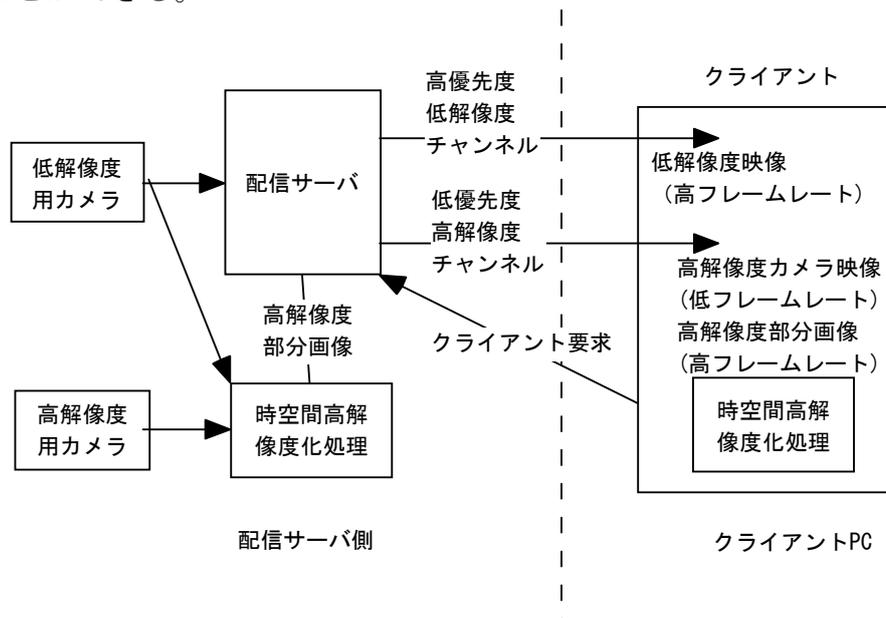


図7. 高解像度動画配信システム

(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ン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の作成 ・高解像度動画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 ・動画配信アルゴリズムの開発 ・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				→			
間接経費							
合計							

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。

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ジタルで行うデ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えない。

一方で、デ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。また、これらのキャリブレーションによる画像補正を実時間でこなうハードウェアの開発を行なった。

4-1-2 複合センサカメラの試作

本研究では、高解像度低フレームレートカメラと低解像度高フレームレートカメラを用いて同一視点、同一視野で時間的、空間的な分解能の異なる2種類の動画を撮像する。そして、高解像度低フレームレート画像のテクスチャ情報、低解像度高フレームレート画像のモーション情報を用いてそれらを統合することによって高解像度高フレームレート画像を得ることを提案する。本研究ではそれらの時間的、空間的な分解能の異なる2種類の動画を撮像するための複合センサカメラシステムを提案する。本システムでは同一視点、同一視野を2台のカメラで撮像するため、プリズムを用いる。具体的にはプリズムの透過光と反射光をそれぞれのカメラで撮像することで同一視点、同一視野を撮像する。

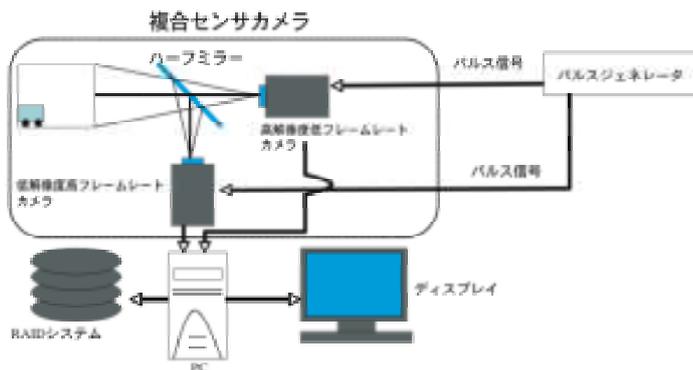


図 8 複合カメラセンサシステムの構成

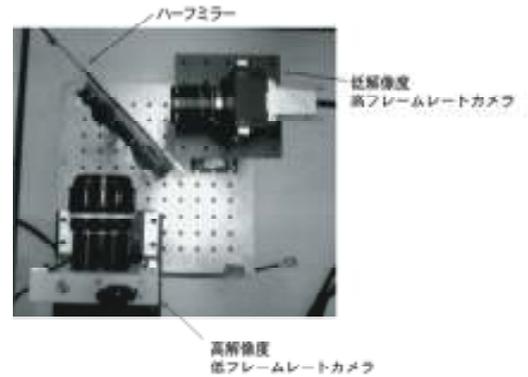


図 9 複合センサカメラ

図 8 に示すように、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るためのパルスジェネレータから構成される。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。

図 9 に示すように、複合センサカメラは高解像度低フレームレートカメラと低解像度高フレームレートカメラの 2 台とシーン情報を分光するためのハーフミラーから構成される。2 台のカメラは 3 軸ステージに固定されており位置の調整ができる。2 台のカメラ間で視点が一致するように、3 軸ステージで位置を調整して配置する。また、ハーフミラーの反射光と透過光の比は 4:6 である。高解像度低フレームレートカメラの方が感度が低いために、光量の多い透過光を高解像度低フレームレートカメラで、反射光を低解像度高フレームレートカメラで撮像する。



図 10 入力画像 (1008 × 1018 画素)



図 11 色の補間処理を行なった後の画像 (1006 × 1016 画素)

(1) 色の補間

ここでは、カメラから取り込んだ Bayer 画像を RGB カラー画像にするための色の補間処理について説明する。カメラの CCD は単板カラーで Bayer 配列となっており、各画素に



図 12 カラーバランス補正後の画像(1006×1016 画素)

は 1 つの色成分しかない．そこで残りの二つの色成分を近傍の画素から推定して補間することにより RGB カラー画像に変換する．補間方法には線形補間を用いる．本システムで用いる低解像度高フレームレートカメラの画素数は 1008×1018 なので，色補間した後は画素数が 1006×1016 に減ることになる．補間処理を行った結果を示す．カメラからの入力画像を図 10 に，補間処理を行った後の画像を図 11 に示す．

図 11 に示すように，色の補間処理を行なうことにより Bayer 画像を RGB カラー画像にすることができた．しかし，図 11 は全体的に黄色の強い画像になっている．これは CCD の感度の問題で B 成分の感度が弱いためである．図 12 に実際にカラーバランス補正を行なった結果を示す．カラーバランスの補正を行なうと白い部分は正しく白色になっており，全体的に黄色が強かったと言う点も改善されたのが分かる．

4-1-3 複合センサカメラのキャリブレーション

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間で幾何学的，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．ここでは，このような複合センサカメラを対象としたキャリブレーション手法を提案し，ハードウェア化に向けたキャリブレーションの改良法と精度評価を行なった．

(1) 幾何学的キャリブレーション

高解像度カメラと低解像度カメラそれぞれに対してパラメータの推定を行う．パラメータ推定法として Tsai の手法[1]を用いた．Tsai の手法では，多数の点の世界座標 (x_w, y_w, z_w) とそれらに対応する画像座標 (X_f, Y_f) の組が与えられたときに，内部パラメータとして $f, \kappa_1, (C_x, C_y)$ ，外部パラメータとして \mathbf{R} と \mathbf{T} を求めることができる．ただし，実験の容易さから平面のキャリブレーションパターンを使用する．

高解像度カメラは画像座標 (X_f^{high}, Y_f^{high}) が与えられると，高解像度カメラパラメータにより高解像度歪みなし平面 (X_u^{high}, Y_u^{high}) に画像を変換できる．低解像度カメラも同様に画像座標 (X_f^{low}, Y_f^{low}) が与えられると，低解像度カメラパラメータと高解像度カメラパラメータを用いて高解像度歪みなし平面 (X_u^{high}, Y_u^{high}) に変換できる．つまり，高解像度歪みなし平面上に 2 枚の画像を投影することで幾何学的特性の異なる画像の一致をはかる．

(2) 光学的キャリブレーション

a. 周辺明度低下の補正

レンズを用いた撮像系では、輝度が一様なシーンを撮影しても画像中心から離れるに従って明度が低下することが知られている。2 台のカメラでこの明度低下の度合いが異なるため、2 枚の画像の輝度差を正確に補正するためには、事前に各カメラで周辺明度低下の補正を行う必要がある。画像周辺の明度低下の原因としては $\cos^4 \theta$ 現象と口径蝕現象がある。

ここでは、 $\cos^4 \theta$ 現象や口径蝕現象による明度低下に加え、その他の要因による明度低下や画素間の感度のバラツキによる固定パターンノイズも考慮に入れるため、そのようなモデルをもたずに単純な補正法を利用して明度低下を補正する。具体的には、放射輝度一様な白色シーンを撮影し、画像全体で一様な輝度を持つと仮定し補正マップを作成することで周辺明度低下を補正する。

b. 画像間の輝度差の補正

周辺明度低下の補正後、画像間で異なる濃淡や色調を一致させるため、一方の画像に対して輝度変換を行う。2 枚の画像間の輝度変換を行う際には、対応する点の輝度値を比較して最小自乗法により変換式を求めたり、変換テーブルを作成するのが一般的である。ここでは、そのような位置の制約を用いず、任意の非線形な変換に対応するために、2 枚の画像の輝度ヒストグラムが一致するように変換を行う。

ヒストグラムの変換には Yang らの手法[2]を用いた。2 枚の画像の輝度ヒストグラムを正規化し、その累積ヒストグラムをそれぞれ H_i^1, H_j^2 で表す。このとき、カメラ 1 の輝度 i に対応するカメラ 2 の輝度 j を次式のように求める。

$$j = \arg \min_j |H_i^1 - H_j^2| \quad (1)$$

この変換を RGB すべての輝度レベルに対して行う。一度、変換テーブルを作成しておくと、以後は異なる画像に対しても、同一の変換テーブルを用いて補正が行える。

4-1-4 キャリブレーション評価実験

(1) 実画像による光学キャリブレーション評価

キャリブレーションパターンとして半径 10mm の円を中心間距離 25mm 間隔で $10 \times 7 = 70$ 個、平面に配置したパターンを用いる。ただし、撮影に使用した高解像度カメラの画像サイズは 4006×2670 画素、低解像度カメラの画像サイズは 1006×1016 画素である。特徴点として円の重心を検出し、世界座標とそれに対応する画像座標の組を得る。さらに、幾何学的キャリブレーションを行いパラメータを推定する。

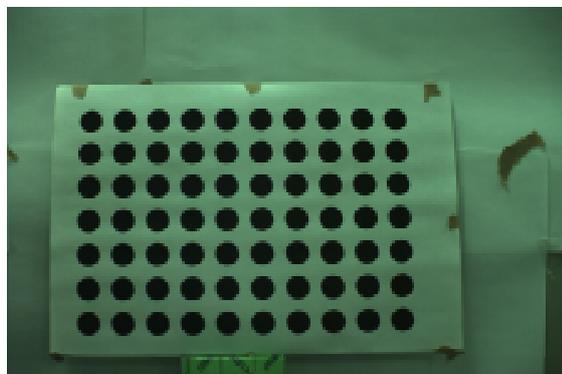


図 13 高解像度入力画像



図 14 低解像度入力画像

高解像度入力画像と低解像度入力画像をそれぞれ図 13、図 14 に、それぞれのヒストグラムを図 15、図 16 に示す。基準平面に投影された特徴点の誤差は平均 0.624pixel、標準

偏差 0.410, 最大 2.342pixel であった. 誤差の平均はサブピクセルとなり十分な精度で幾何補正されたことがわかる. 変換後の画像を図 17, 図 18 に, それぞれのヒストグラムを図 19, 図 20 に示す. また, その差分画像を図 21 に, 差分画像のヒストグラムを図 22 に示す. ただし差分画像はログスケールで示す.

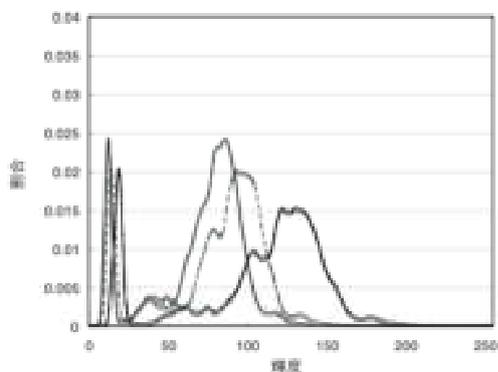


図 15 高解像度入力画像ヒストグラム

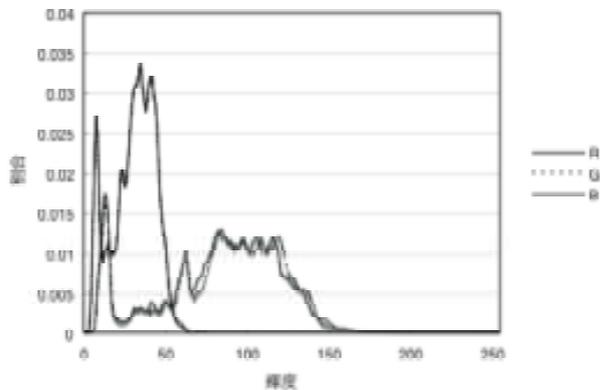


図 16 低解像度入力画像ヒストグラム

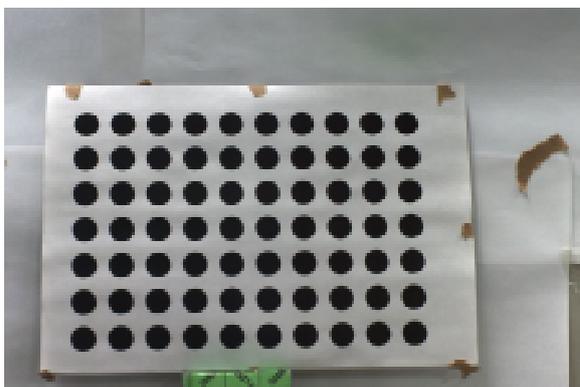


図 17 高解像度補正画像

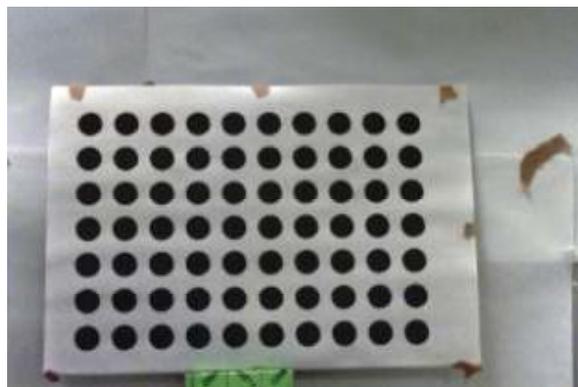


図 18 低解像度補正画像

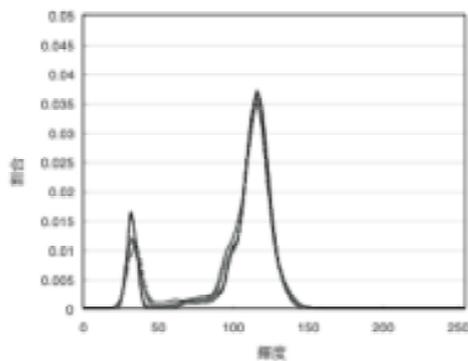


図 19 高解像度補正画像ヒストグラム

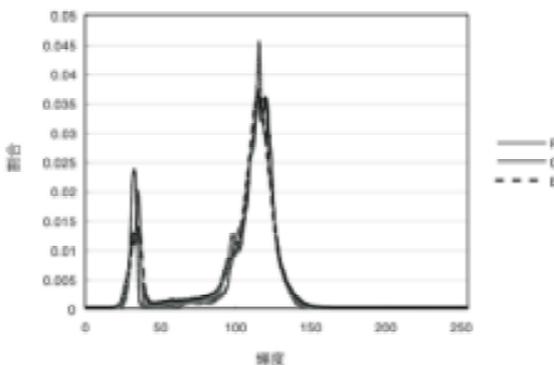


図 20 低解像度補正画像ヒストグラム

図 21 よりエッジの部分で輝度に差があるのがわかる. しかし, 差分画像のヒストグラムを見てみるとその差は小さいものであった.

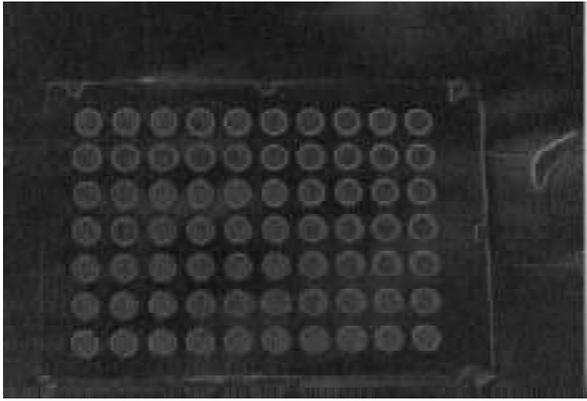


図 21 差分画像 (ログスケール)

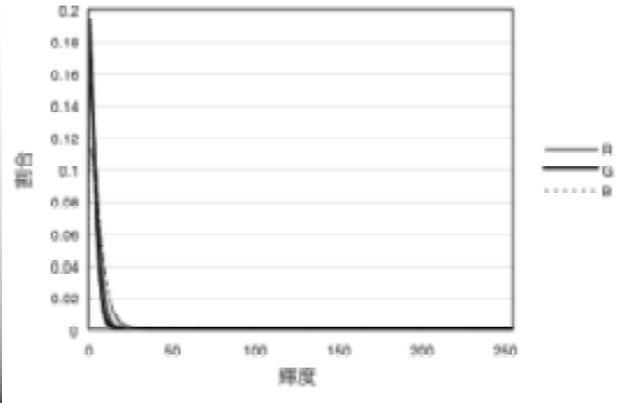


図 22 差分画像ヒストグラム

4-1-5 まとめ

本サブテーマでは、複合視覚センサを試作し、評価を行った。提案手法によって複合センサカメラのキャリブレーションがうまく行われたのを確認した。

4-2 リアルタイム高解像度動画の作成

上で試作した複合センサカメラから得られる2つの動画から、空間的にも高解像度で、しかも時間的にも密な動画を作成する。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、低解像度高フレームレート動画)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。さらに、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提案アルゴリズムの改良を行った。

周波数空間アプローチによる高解像度画像の作成では、複合センサカメラから得られる2種類の動画から、画像処理によって高解像度・高フレームレート動画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。

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

(1) 高解像度動画の生成

本研究では離散コサイン変換(DCT)を用いる方法、および離散ウェーブレット変換(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。

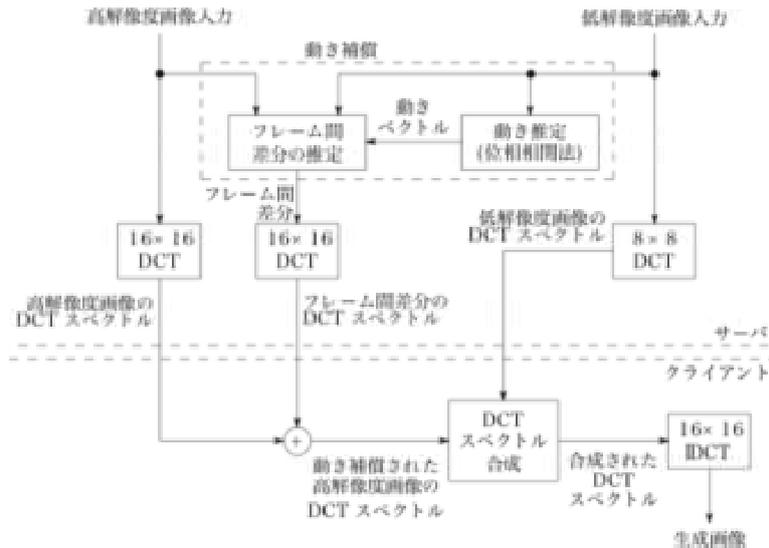


図 23 DCT を用いた高解像度画像生成処理の流れ

a. DCT を用いた高解像度動画画像生成アルゴリズム

DCT を用いた高解像度画像生成手法では、生成する高解像度画像の高周波成分と低周波成分を異なる方法によって推定する。

- ・ 生成画像の高周波成分は、画像空間上で動き補償を行った高解像度画像のスペクトルをそのまま利用する。シーン中の動きは低解像度画像を用いて推定する。ただし、補償できない部分については低解像度画像から補間する。
- ・ 生成画像の低周波成分は、画像空間上で動き補償された高解像度画像のスペクトルに低解像度画像のスペクトルを合成することによって生成する。

DCT を用いた高解像度画像生成処理の流れを図 23 に示す。ただし、2 種類の動画画像の画像サイズ比を 2:1 とする。

b. ウェーブレット変換を用いた高解像度動画画像生成

次に、直交周波数変換として、離散ウェーブレットを利用した方法について述べる。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。なお、shift-variant とは、周波数変換の係数の値が画像の位置に依存することであり、平行移動を行うとその値が変化する。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。

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。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(拡大)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。

(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 1 に示す MPEG テストシーケンスを使用した。

表 1 実験に使用した MPEG テストシーケンス

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

a. DCT を用いた高解像度画像生成手法のシミュレーション評価

図 24 は“Flower garden”の 29 番目のフレームの生成結果である。図 24 の右側は、左側の四角で囲んだ一部分を拡大したものである。



(a) 原画像



(b) 低解像度画像を 2 倍拡大



(c) 生成画像

図 24 Flower garden フレーム No. 29

b. 空間アップサンプリング手法との比較

空間アップサンプリング手法（最近傍法，双 3 次スプライン，frame rate 変換）と提案手法（DCT 合成法，RDWT 合成法）との比較のため，テストシーケンス全体の PSNR を測定した結果を表 2 に示す。なお，RDWT 合成法で使用したウェーブレットは，整数 2/6，整数 5/3，

Haar, Daubechies 4-tap である。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。

表 2 PSNR の測定結果

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	
DCT 合成法	24.62	21.15	27.14	30.31	
RDWT 合成法	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

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 種類の動画画像を提案手法に適用し、サイズ 4000×2600 画素、フレームレート 30 [fps]の動画画像を生成した。DCT 合成法を用いて生成されたフレームの例を図 25(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 25(b)(c)(d))を低解像度画像(図 25(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画像が提案手法によって生成できることが示された。

4-2-2 高解像度画像作成のための PSF 推定

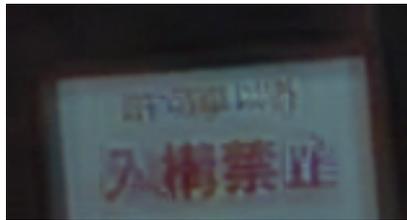
提案手法は、高フレームレート動画画像から得られる動きの情報によって高解像度画像に動き補正を行い、高解像度・高フレームレート動画画像を生成した。しかし、動画画像の撮影には 2 組の異なった CCD・レンズを有するカメラを使用しているため、2つの動画画像ではぼけが異なる。前節の提案手法ではこれらのぼけの影響を考慮していなかったために、低解像度画像のぼけが生成画像に伝播して、生成画像の画質が低下していた。そこで、本節では画像のぼけを表す関数である PSF(Point Spread Function)を用いて、低解像度画像のぼけを補正し、高解像度高フレームレート画像を生成する手法について提案する。

(1) PSF の推定法

一般にカメラで撮影された画像は、レンズの光学的収差や回折によりぼけが生じる。このぼけを表す関数として PSF があり、これはオブジェクト領域の点光源による放射照度分布のことである。撮影するシーンを原画像、撮影したぼけのある画像を劣化画像とすると、一般に劣化画像は原画像と PSF との畳み込み演算で得られる。特に PSF が位置不変(画像中の各点でぼけが一定)の場合、劣化画像 $g(x,y)$ は、原画像を $f(x,y)$ 、PSF を $h(x,y)$ とすると



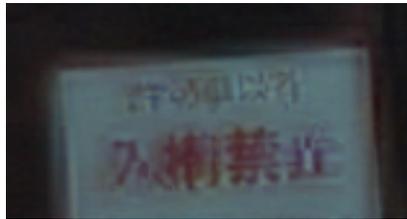
(a) 生成画像 (DCT スペクトル合成法)



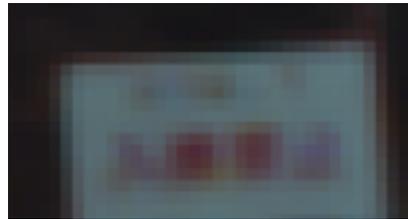
(b) (a)を拡大した画像



(c) 整数 2/6



(d) 整数 5/3



(e) 対応する低解像度画像

図 25 実画像より生成された高解像度画像

$$g(x,y) = h(x,y) * f(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\alpha,\beta) f(x-\alpha, y-\beta) d\alpha d\beta \quad (2)$$

と表される．ここで*は畳み込み演算を意味する．また，空間領域 (x, y) での畳み込み演算は，周波数領域 (u, v) の乗算に対応するので， $g(x, y), h(x, y), f(x, y)$ のフーリエ変換をそれぞれ $G(u, v), H(u, v), F(u, v)$ とすると

$$G(u, v) = H(u, v) F(u, v) \quad (3)$$

と表す事ができる．これに基づくと，図 26 に示すように，高解像度画像 G_h と低解像度画像 G_l は，理想画像 F から，それぞれのカメラの撮像過程 H_h, H_l を経て撮像される．ここで，図 26 に示す $H_x(u, v)$ を新たな PSF として導入する． $H_x(u, v)$ は以下の式を満たす．

$$G_l(u, v) = H_x(u, v) G_h(u, v) \quad (4)$$

この式を変形すると

$$G_h(u, v) = H_x^{-1}(u, v) G_l(u, v) \quad (5)$$

となる．上式から， $H_x^{-1}(u, v)$ は，低解像度画像のぼけレベルを高解像度画像のぼけレベルに補正する関数であるといえる．なお，この $H_x^{-1}(u, v)$ はキーフレームの高解像度画像と低解像度画像において計算する．

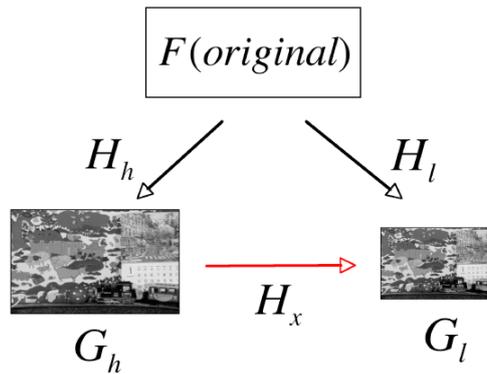


図 26 PSF と各画像間の関係

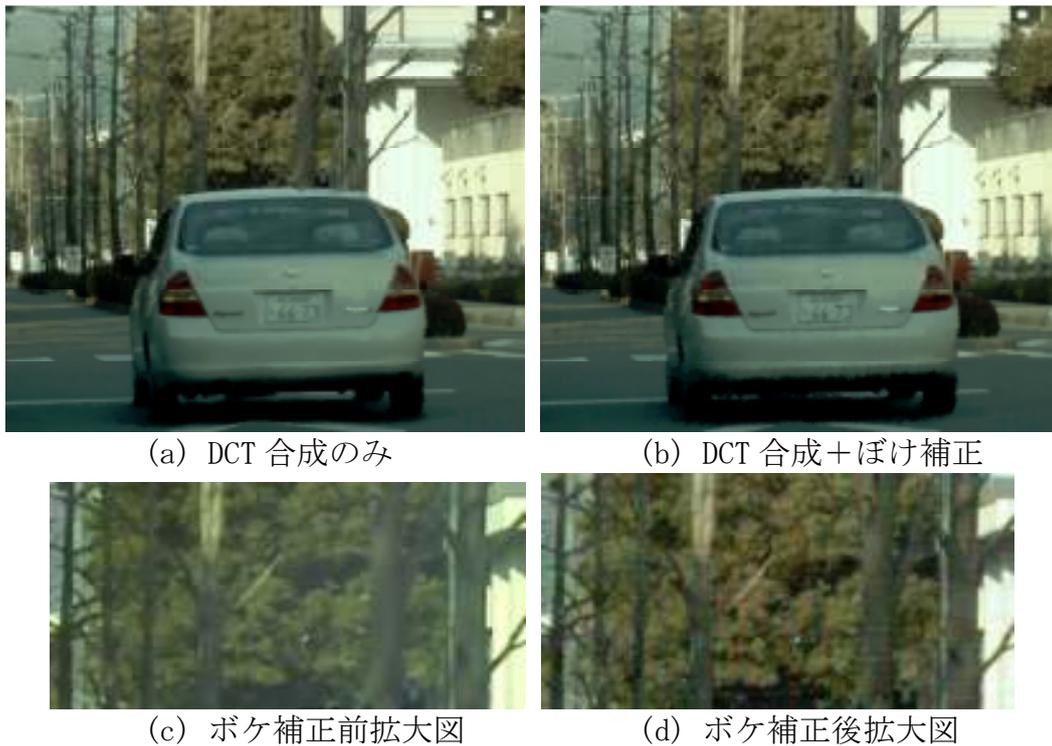


図 27 生成画像の比較 (実画像)

(2) 評価実験

実際のカメラで撮影した 2 種類の動画像（低解像度高フレームレート画像・高解像度低フレームレート画像）を用いて，前節の実験と同じく 2 パターンの処理を行いその結果を比較した．結果を図 27 に示す．図 27 (c), (d) より，ぼけ補正を加えることにより，画像中の木が鮮明になっているのがわかる．しかし，補正した画像（図 27 (b)）は全体的にノイズが加わっている．これは，PSF の推定誤差などによるノイズと考えられる．今後は，キーフレームを多数使うなどの方法により，PSF の推定精度をより高める必要がある．

4-2-3 モーフィングアプローチによる高解像度画像の作成

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

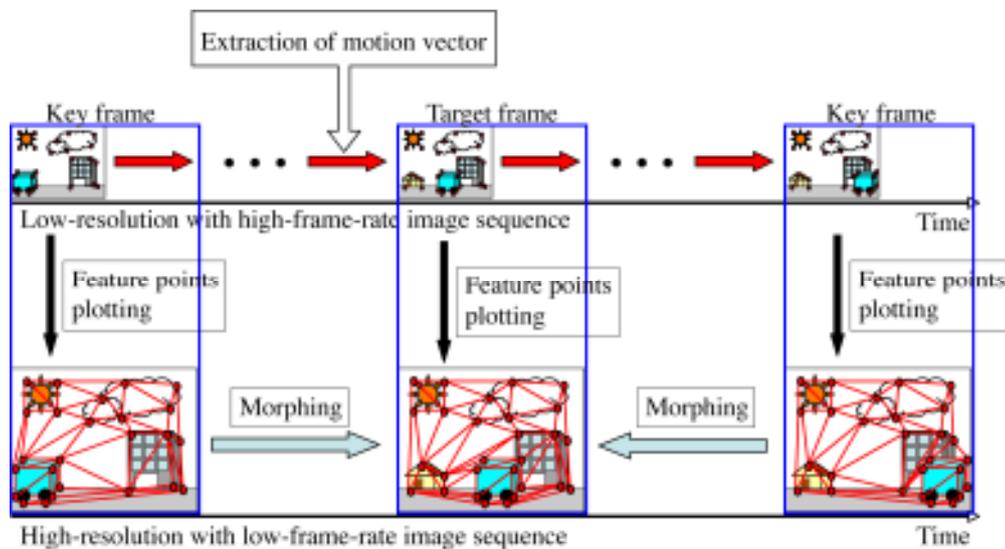


図 28 高解像度動画の生成

低解像度高フレームレート動画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を生成する。提案手法の概略を図 28 に示す。一般的に入力の高解像度低フレームレート動画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を用いることで、トラッキングを比較的に行える。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 29 に示し、以下に処理の概略を示す。

1. 特徴点追跡 (Feature tracking) : 低解像度高フレームレート動画のキーフレームにおいて特徴点を抽出する。この特徴点をキーフレーム間で追跡することで、特徴点の動きベクトルを検出しモーション情報を得る。特徴点追跡には、KLT tracker (Kanade-Lucas-Tomasi tracker) [3]を用いた。KLT tracker は時空間の輝度勾配を用いた特徴点追跡法であり、サブピクセル単位での追跡が可能である。また、KLT tracker に適した特徴点を抽出する手法として Tomasi らの手法[4]を用いた。
2. 特徴点の適用 (Feature points plotting) : 特徴点追跡により低解像度動画の各フレーム間で対応づけされた特徴点位置を、高解像度動画の同一時刻のフレームに適用し、高解像度動画のフレーム間で特徴点の対応付けを行う。特徴点追跡に用いた KLT tracker は、各低解像度画像においてサブピクセル単位で特徴点位置情報が得られる。したがって、高解像度画像への適用は、低解像度画像の特徴点位置と縦横それぞれの解像度比の積算により行う。つまり、解像度比が $1 : MN$ (横 $1 : M$, 縦 $1 : N$) の場合、低解像度画像における特徴点位置 (x, y) は、高解像度画像の座標 (Mx, Ny) に適用される。
3. ポリゴン分割 (Polygon partitioning) : (2) で得られた特徴点位置を頂点とした三角パッチに高解像度画像をポリゴン分割する。ポリゴン分割にはドロネー分割 (Delaunay triangulation) [5]を用いる。
4. モーフィング (Image morphing) : (1) で得られた動きベクトルと (3) で得られた高解像度ポリゴンを用いて、補間フレームの高解像度ポリゴンをモーフィング (Image morphing) により生成する。モーフィングは、ポリゴン単位で処理するこ

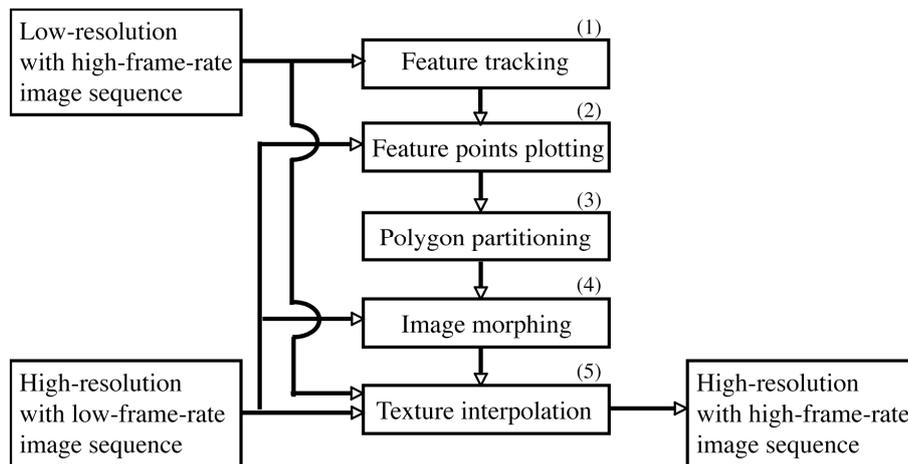


図 29 高解像度動画像生成の流れ

とで、形と色の両方を補間できるワーピング法を用いる。ただし、特徴点抽出・追跡精度の問題で動領域のテクスチャの歪みが生じることがある。よって、生成した高解像度ポリゴンに対し歪み判定を行い、歪んでいないと判定されたポリゴンのみ中間フレームのテクスチャとして使用する。

5. テクスチャ補間 (Texture interpolation) : (4)までの処理では、歪み判定によるテクスチャの欠落領域、あるいはポリゴン自体が生成されない領域が存在する可能性がある。以上のことを起因とするテクスチャ欠落領域に対してテクスチャ補間 (Texture interpolation) を行う。

(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。

MPEG テストシーケンスの $M \times N$ pixels (横 : M pixels, 縦 : N pixels) のブロック毎に平均値を計算し、その値を各画素の画素値とする画像を作成する。これにより、MPEG テストシーケンスに対し、解像度比が $1 : MN$ 、フレームレート比が $1 : 1$ の動画像が得られる。この動画像を低解像度高フレームレート動画像とする。一方、MPEG テストシーケンスの F_l/F_h フレームにつき1フレームを残し、その $F_l/F_h - 1$ 枚の中間フレームを間引き処理する。これにより、MPEG テストシーケンスに対し、解像度比が $1 : 1$ 、フレームレート比が $F_h : F_l$ の動画像が得られる。この動画像を高解像度低フレームレート動画像とする。シミュレーション実験では、加工前の MPEG テストシーケンスは、元画像であると同時に高解像度高フレームレート動画の生成において、目指すべき理想画像であるとも言える。

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.

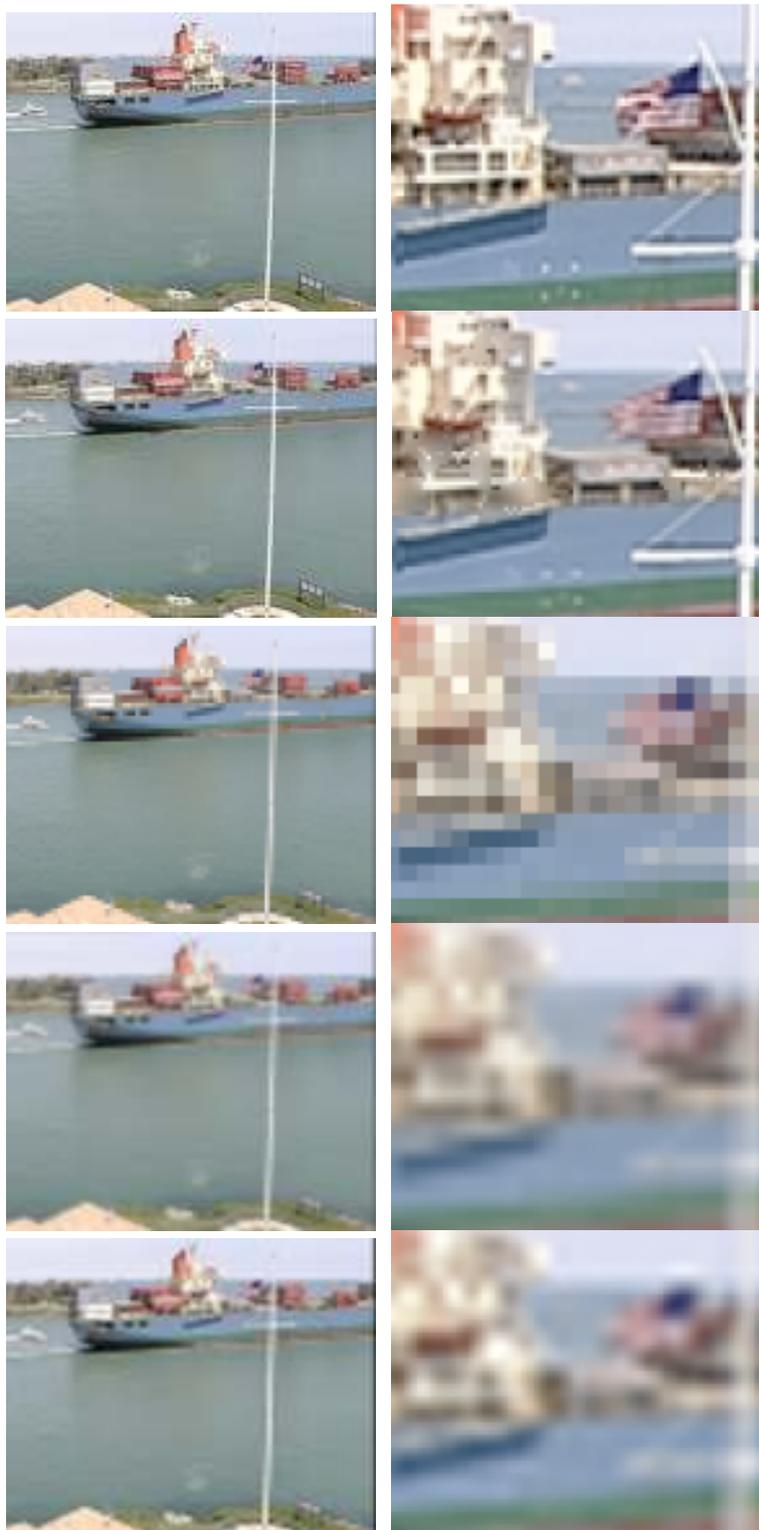


図 30 “Container” Frame No. 18, 上から順に理想画像(MPEG テストシーケンス), 提案手法, 最近傍法, 双線形補間法, バイキュービック法

図 30 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同様 30[fps]の動画が生成でき, 同様の動きが再現でき



図 31 (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

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 30 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.

表 3 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する.

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. 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.

表 3 提案手法と低解像度補間手法の PSNR 比較

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

b. 全方位画像への適用

提案手法を全方位画像に適用して評価を行なった. 前節の実験で使用した複合センサカメラに, 全方位ミラーを取り付け, 撮像を行った. ここで得られた動画像に対し, 幾何学的, 光学的キャリブレーションを行い CCD センサカメラ間の特性差を補正し, 解像度 1002

×668 pixels, フレームレート 30 fps の低解像度高フレームレート動画像と, 解像度 4008 ×2672 pixels, フレームレート 4.29 fps の高解像度低フレームレート動画像を得た. これら2種類の動画像を入力とした高解像度高フレームレート動画像の生成結果を図31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てい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動画像同様に, 30 fps の動画像が生成でき, 動物体・動領域の滑らかな動きを再現できた. 図32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分かる. 以上により全方位画像に対しても提案手法を適用できることを確認した.



図32 (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

4-2-4 高解像度高フレームレート生成手法の有効性の検証

本節では新たに追加実験を行うことで提案手法の検証を補い, 改めて提案手法の有効性を検証する.

(1)データ量圧縮効果の検討手法

本研究の提案手法は, CCD の画素数とサンプリング速度の物理的制約を超える目的で, 高解像度・高フレームレート映像の合成を行っている. しかし, 高解像度低フレームレートと低解像度高フレームレートに分離撮像し合成を行うことは, 同時に画像データ圧縮の効果を生むとも考えられる. また, 高品質, つまり高解像度高フレームレート動画像を生成する際, 提案手法では入力画像として高解像度低フレームレートと低解像度高フレームレートをを用いているが, 低解像度高フレームレート動画像のみを入力とした従来手法も提案されており, それらも同様にデータ圧縮効果があると考えられる. そのような従来手法として最近傍法, 双線形補間法, バイキュービック法などがあるが, 提案手法とそれら従来手法を比較することで, データ圧縮効果の検証を行う.

まず, 高解像度高フレームレート実動画像を加工することで, 同一視点・同一視野を持

表 4 同一入力画像における PSNR 比較

Name	Proposed [dB]	Nearest Neighbor [dB]	Linear Interpolation [dB]	Bi-cubic [db]
Container	34.26	22.58	22.70	21.91
Hall monitor	32.21	23.61	23.94	24.47
Coast guard	25.89	22.57	22.72	23.01
Foreman	28.73	26.53	27.32	27.97
Mobile	20.04	17.10	17.17	17.39
Garden	19.09	17.35	17.33	17.48

つ低解像度高フレームレート動画画像と高解像度低フレームレート動画画像を作成する。提案手法ではそれら両方を入力として高解像度高フレームレート動画画像の生成を行うが、従来手法では低解像度高フレームレート動画画像のみを入力として用いる。

検証には、従来手法と提案手法に同等の入力データ量を与えた時に生成される、高解像度高フレームレート出力画像の画質を比較する、という方法を取った。入力データ量を増やした時に画質が良くなると仮定すると、「同じ入力データ量で画質が良い」とことと「同じ画質を実現するための入力データ量が少ない」とことは同等であるためである。この仮定の正当性については実験により検証を行う。

(2)評価実験

a. データ圧縮効果の検証

前節で述べた検証を実際に行った。提案手法には、評価実験により最適であると結論付けられたフレームレート比 7:1、解像度比 1:16 の入力画像の組み合わせを用いた。一方、比較画像として、提案手法とデータ量が同一になるように MPEG テストシーケンスを縦横 0.44 倍縮小した映像(30[fps])を入力とし、従来手法(最近傍法, 双線形法, バイキュービック法)を用いて元の解像度に補間した画像を用いた。

表 4, 表 5 に各 MPEG テストシーンごとの PSNR の結果を示す。ただし表 4 の結果は、従来手法の入力低解像度高フレームレート動画画像のデータ量を提案手法の低解像度高フレームレート入力画像と同じ 1 とした時の結果であり、これは 4-2-3 節で行なった検証と同じ方法で実験を行ったものである。それに対し表 5 中では、入力低解像度高フレームレート動画画像のデータ量を 22/7 とした時の結果である。

この同一データ量での比較では、比較の補間手法では入力画像データ量が表 4 の条件より大きくなっているため、表 5 の PSNR は表 4 のそれより向上している。それでもなお、“Container”, “Hall Monitor” や “Mobile” においては、依然として提案手法の方が高い PSNR を示している。しかし、“Foreman” など他のシーケンスに関しては、データ量同一の条件では、提案手法は単純な補間手法と同等または若干劣る結果となった。“Foreman” などの動きの激しいシーンでは、多くの動領域部分で特徴点追跡に失敗し、その結果、低解像度画像からの内挿補間による画像生成部分が大半を占めるようになる。すなわち、このようなケースでは、低解像度入力画像の低い解像度が全体の画質に大きな影響を与えるため、同一データ量での比較では入力解像度の大きい比較手法に対して劣る結果となった。

全体をまとめると、提案手法では同一データ量の比較においてもシーン依存ではあるが PSNR が向上している結果があるため、限定的ではあるがデータの削減効果があることが示された。この特徴の更なる適用範囲の拡大には、特徴点追跡手法の高精度化が求められる。また、現状では追跡を失敗した部位は低解像度入力画像からの双線形補間により補正を行っているが、表 4, 表 5 の結果より、補間手法ではバイキュービック法の方が性能が高いた

表5 同一データ量における PSNR 比較

Name	Proposed [dB]	Nearest Neighbor [dB]	Linear Interpolation [dB]	Bi-cubic [db]
Container	34.26	23.95	25.37	26.11
Hall monitor	32.21	25.27	27.10	27.94
Coast guard	25.89	24.09	25.53	26.30
Foreman	28.73	28.11	30.00	30.93
Mobile	20.04	17.89	18.79	19.18
Garden	19.09	17.99	18.95	19.50

め、補正にバイキュービックを用いる事で画質向上が期待出来る。

b. 超高解像度実画像を用いた実験

これまでのシミュレーション実験では解像度 400×300 pixels 程度の画像を生成し、提案手法の有効性を検証してきた。この節では、本研究目標である解像度 8000×8000 pixels の実画像を扱い、提案手法の有効性について検証していく。

従来手法の入力画像は解像度 2000×2000 pixels, フレームレート 30fps の低解像度高フレームレート実画像列とした。また、提案手法の入力画像は、解像度 8000×8000 pixels, フレームレート 4.29fps の高解像度低フレームレート実画像列と、解像度 2000×2000 pixels, フレームレート 30fps の低解像度高フレームレート実画像列とした。さらに、提案手法の各種パラメータには、特徴点検出時の最大特徴点数, 最小近傍特徴点間距離, トラッキング時の特徴窓サイズ, 探索領域, 最大反復計算回数などがある。パラメータ変更前は解像度 400×300 pixels 程度の画像を用いた時のパラメータと同じだが、変更後は解像度 8000×8000 pixels 画像に最適なパラメータを用いた。

図 33 に 8000×8000 pixels 画像での生成結果を示す。解像度の違いがわかりやすいように、入力高解像度, 低解像度, 生成画像の一部を拡大したものを図 34 に示した。図 34-(b) と (c) との比較により、入力低解像度に対して解像度が向上し、高解像度入力キーフレームである (a) の画像の解像度に近づいていることが確認できた。

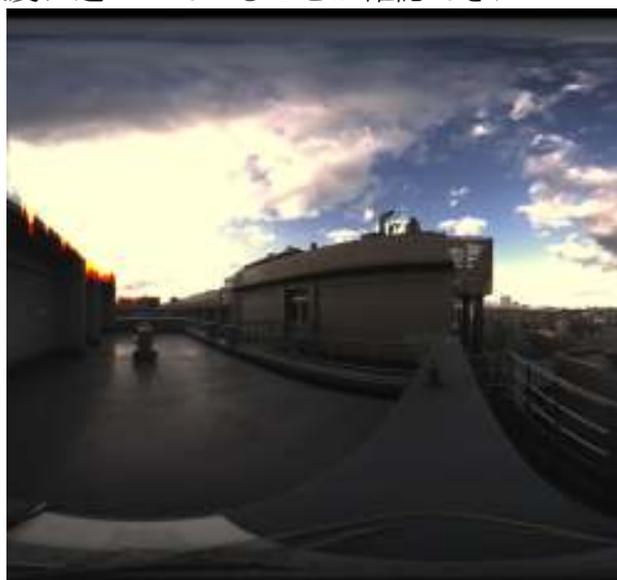


図 33: 高解像度生成画像 (8000×8000 pixels)

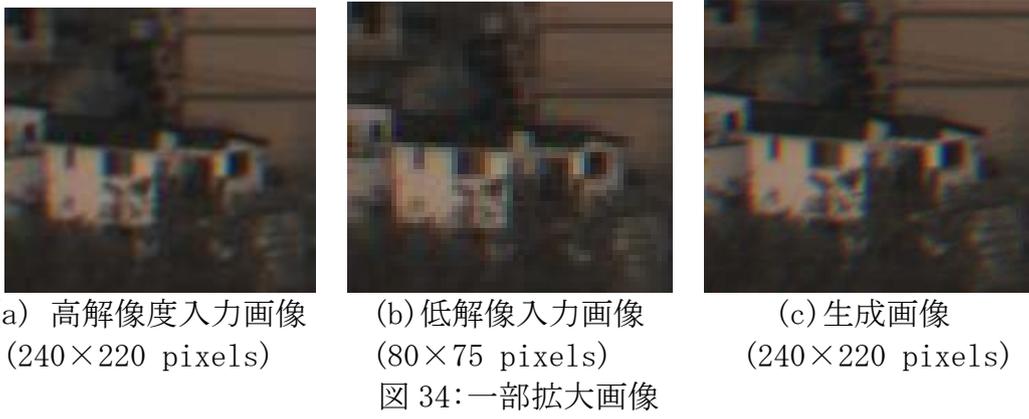


図 35 に解像度 8000×8000 pixels の動画生成時の時系列 PSNR データを示す．図 35 より超高解像度画像生成において，現段階で提案手法ではパラメータ変更後は変更前よりかなり改善したものの，従来手法より依然劣っている事が分かる．これは，図 34 に示すように高精細化される部位もあるが，画像サイズの拡大に伴いトラッキングミスやパッチ分割数の不足により画質向上が制限されたものであると考える．ただし，トラッキングの高精度化，ドローン分割時に移動ベクトル情報や色情報などを用いるなどの方法で PSNR の向上が予想される．また，提案手法はパラメータによって PSNR が大きく変化しており，現段階ではパラメータの設定は使用者の経験に依っているが，今後は，パラメータ設定を自動で最適化することにより，PSNR 向上を目指す．

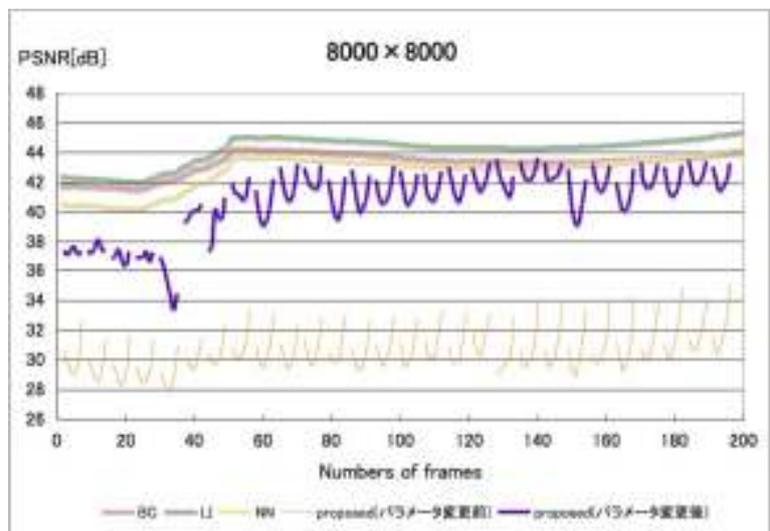


図 35 8000×8000 pixels 画像での PSNR 比較

4-2-5 まとめ

本サブテーマでは，試作した複合センサから得られる 2 つの動画像から空間的にも高解像度で，しかも時間的に密な動画像を作成する手法として，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，評価を行った．試作した複合センサによる実画像実験により本提案手法が有効であることを確認した．

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を設計，作成する．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．本研究では，実時間撮像と無歪みの仮想視野画像変換の利点より，双曲面ミラーを用いた全方位光学系を設計，作成した．試作した複合センサは，高解像度 CCD に画像サイズ 4008×2672 画

素, セルサイズ $9.0 \times 9.0 \mu\text{m}$, 低解像度 CCD に画像サイズ 1008×1018 画素, セルサイズ $9.0 \times 9.0 \mu\text{m}$ を用いている. 全方位ミラーは, これらのカメラを想定して光学設計を行った. また, 更に高解像度化に向く 2 枚反射全方位光学系のシミュレーションによる光学特性の評価とミラーの評価と試作を行った.

4-3-1 双曲面型全方位視覚センサ

双曲面を用いた全方位視覚センサは, 図 36 に示すように, 鉛直下向き of 双曲面鏡と上向き of カメラから構成される. 双曲面鏡は, 二つの焦点を持ち, カメラは双曲面の中心軸と同軸に主点が双曲面鏡のその一方の焦点にくるよう配置されている. このように配置することで, センサの周囲 360° を一度に撮像することができる. また, 図 36 に示すように双曲面鏡のミラー焦点に向かって入ってくる光は, 双曲面の特性によりミラーに反射後, もう一方の焦点に向かう. この関係から, 全方位入力画像を図 37 に示すようにミラー焦点を視点とした透視変換画像やパノラマ画像に容易に変換できるという特性を持つ.

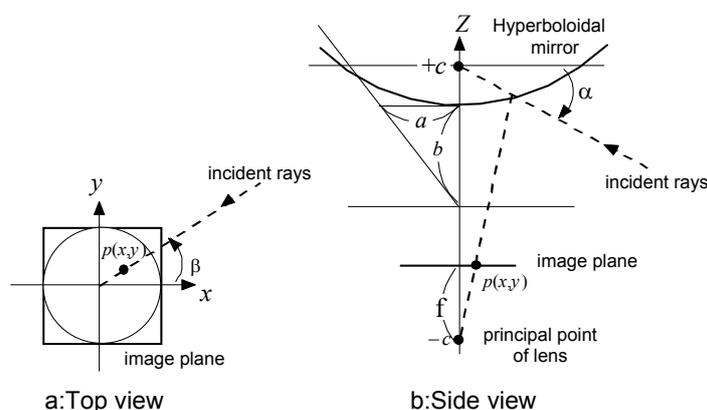


図 36 全方位視覚センサの構造

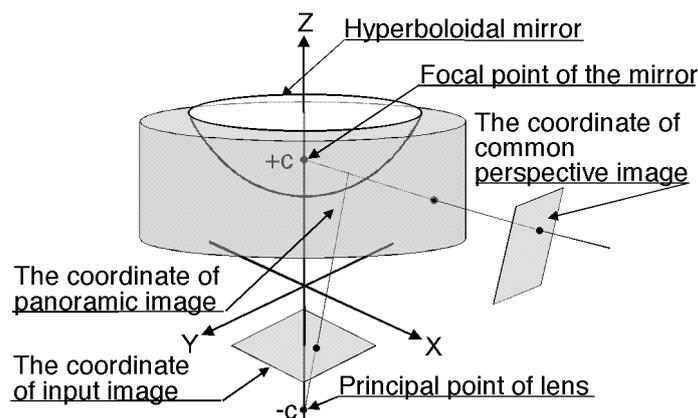


図 37 変換画像の幾何学的関係

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 $9.0 \times 9.0 \mu\text{m}$ と画像サイズ 1008×1018 画素, セルサイズ $9.0 \times 9.0 \mu\text{m}$) の CCD を想定した全方位ミラーの設計を行った. 高解像度化に際しての問題点を評価するために, まず, 従来の双曲面 1 枚ミラーにて設計・試作を行った.

ミラー形状は, 全方位視覚センサを構成するカメラの画角と対象とする垂直視野角によ

り決定される. 今回は画角が43度のカメラを用いて, 全方位センサの垂直視野角が88度(上方25度, 下方63度)になるように設定した. 図38に設計したミラーの概形を示す.

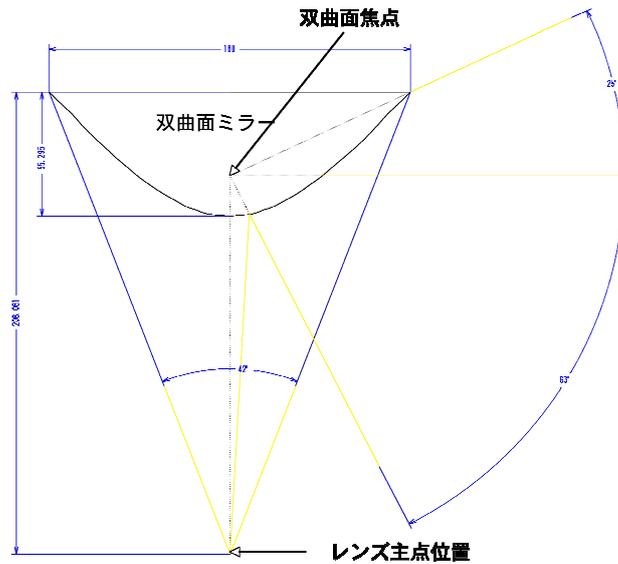


図 38 試作ミラー概形と視野角

試作した複合センサカメラ単体の光学的評価により得られた特性を想定して, 全方位ミラーの評価を行った. この結果を用いて, 複合センサに最適な単一双曲面形状の全方位ミラーの設計を行い, 試作を行った. 図39に試作した全方位ミラーを示す.



図 39 試作した双曲面全方位ミラー

4-3-2 二枚反射全方位視覚センサ

二枚反射全方位視覚センサの光学系では, 反射曲面として2葉双曲面を用いた. 2葉双曲面とは双曲線をZ軸周りに回転させることで得られる曲面である. 2枚反射光学系は, 図40に示すように, 鉛直下向きの双曲面ミラー(第一ミラー)と鉛直上向きの双曲面ミラー(第二ミラー), そして第一ミラーの内側にカメラを設置した構成をとる. 双曲面は図41に示すように, 焦点F1に向かって入射する光線は他方の焦点F1'に向かって反射するという幾何的な性質をもつ. そこで2枚ミラーシステムでは, 第一ミラーの一方の焦点F1'と第二ミラーの一方の焦点F2を一致させ, 第二ミラーの他方の焦点F2'とレンズ中心Pを一致させる構成をとる. この構造によって, 第一ミラーの焦点F1に向かって入射した光線がカメラで撮像されるため, 2枚ミラーシステムにおいても単一の投影中心が実現される. つまり, 光学特性が透視投影となるので, 2枚ミラーシステムは双曲面ミラーを用いた一枚ミラーの全方位視覚センサ同様, 様々な画像処理に適している.

2枚ミラーシステムは, 従来の単一ミラーの場合に比べ, カメラに対する入射視野角が小さくなり, 像面湾曲が抑制されるという特性がある. 図40の $\angle\beta$ と $\angle\gamma$ の比較から, カ

メラに対する入射視野角が小さくなっていることが確認できる。この像面湾曲の抑制により、2枚ミラーシステムは全方位視覚センサの高解像度化や小型化を実現できる。

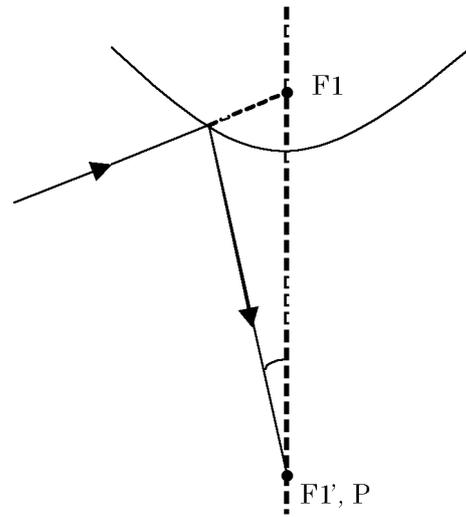
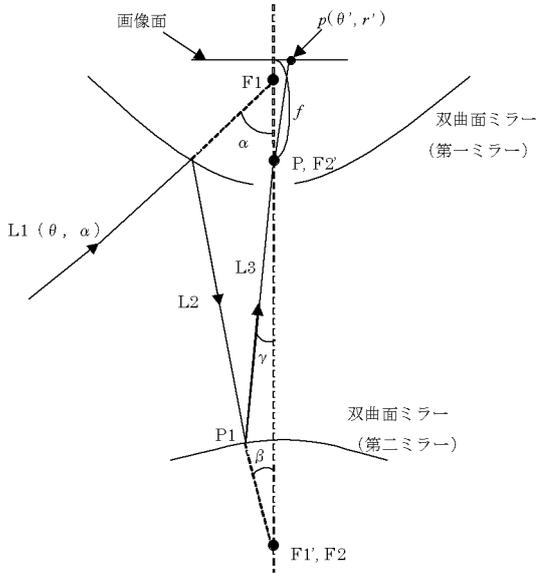


図 40 2枚ミラーシステムの幾何学特性 図 41 双曲面ミラーの幾何学特性

また、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。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。図 42 に設計した全方位ミラーの設計図を、図 43 に試作ミラー概要を示す。さらに複合センサカメラと接続した写真を図 44 に示す。

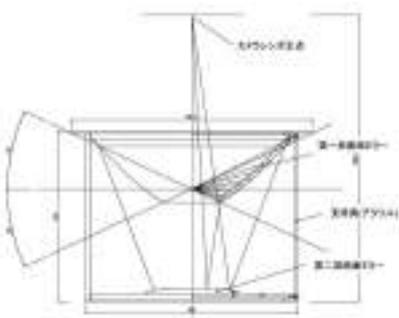


図42 設計図



図43 試作2枚反射型全方位ミラー



図44 全方位複合カメラ

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45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で、入射角

に対する解像度をもとめた。図46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46に比較として、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46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、単一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対して、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47に示した。図47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。

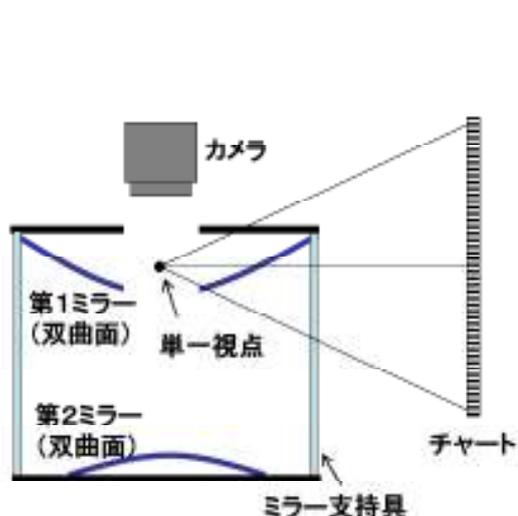


図 45 解像度測定実験

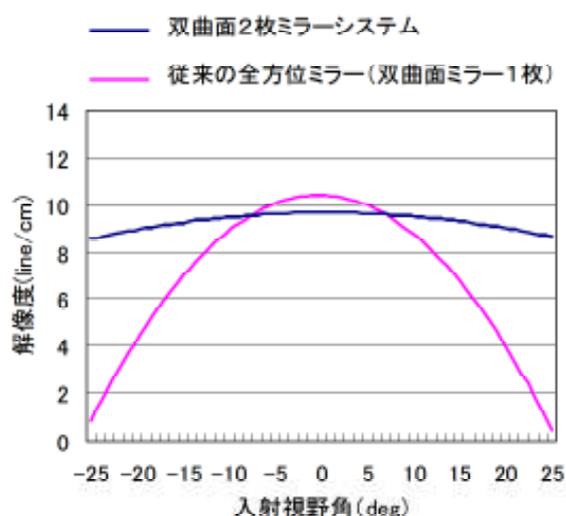


図 46 入射角に対する解像度の変化



図 47 全方位複合センサ画像
(左：高解像度低フレームレート画像，右：低解像度高フレームレート画像)

4-3-3 まとめ

本サブテーマでは、複合センサカメラ用に新たに考案した2枚反射系全方位ミラーを設計し、高解像度複合センサカメラを試作した。小型化しても十分な性能が得られることを

実験により確認した。

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

ここでは、複合センサカメラで撮像した画像データを実時間で PC 上に取り込んで、ディスプレイに表示させつつ RAID システムに記憶させるソフトウェアを実装した。

画像を取り込んで保存する際に、取り込みの終了を待ってから画像の表示、保存を行うと画像取り込みの終了を待つ時間が無駄である。そこで、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。ダブルバッファ処理は、画像の取り込みと、画像の表示、保存を並列に行うものであり、ダブルバッファを用いた時の処理と用いない時の画像 2 フレーム分の取り込み、保存処理の流れを図 48 に示す。具体的には、画像を取り込むためのメモリを 2 つ用意しておき、メモリ 1 で画像の取り込みを行い、それに並行してメモリ 2 に格納された画像データの表示、保存を行う。それぞれの処理が終了した後に、メモリ 2 で画像の取り込みを行い、それに並行してメモリ 1 に格納された画像データの表示、保存を行う。これらの処理を繰り返して行う。

ダブルバッファ処理を用いることで、次の画像の取り込みまでの待ちの時間が大幅に減り高速に取り込み処理が実現できるようになった。

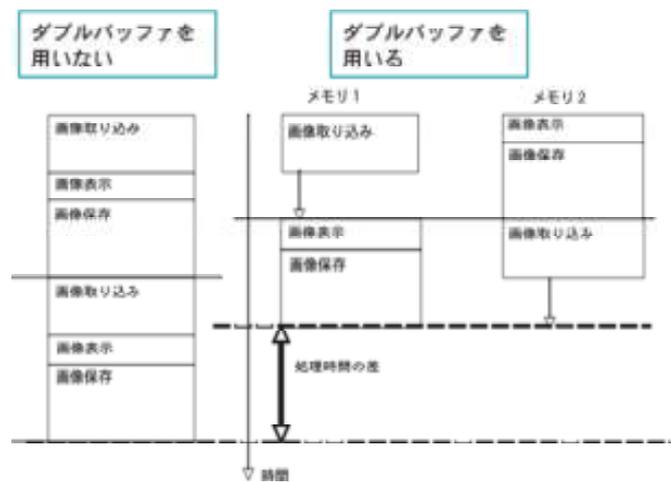


図 48 ダブルバッファ処理

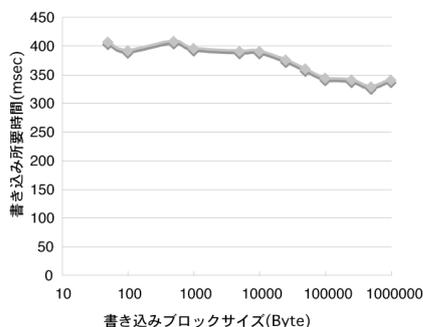


図 49 低解像度高フレームレートカメラの書き込みブロックサイズと書き込み所要時間

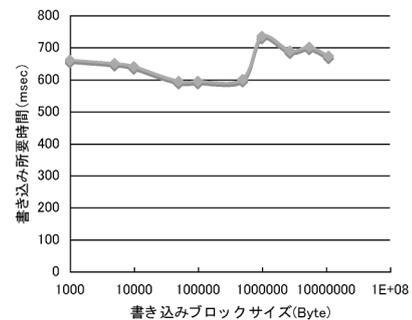


図 50 高解像度低フレームレートカメラの書き込みブロックサイズと書き込み所要時間

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 49 に示す。図 49 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。同様に高解像度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 50 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。

4-4-3 パルスジェネレータの作成

本システムでは高解像度低フレームレートカメラと低解像度高フレームレートカメラという 2 台のフレームレートの異なるカメラを同期させて撮像する。高解像度低フレームレートカメラは 4.29fps、低解像度高フレームレートカメラは 30fps である。そのため、1:7 の割合で高解像度低フレームレート動画画像と低解像度高フレームレート動画画像を同期して撮像する。

2 台のカメラの同期をとるために、パルスジェネレータを作成しそこから 2 台のカメラにパルス信号を送る。本システムで用いるカメラは 4.29fps と 30fps であるので高解像度低フレームレートカメラに 4.29Hz、低解像度高フレームレートカメラに 30Hz のパルス信号をそれぞれ送ることで 2 台のカメラの同期をとる。

4.29Hz と 30Hz のパルス信号を出力するためのパルスジェネレータを Max+Plus II Baseline (Altera 社) というソフトを用いて設計した。設計した回路は、ヒューマンデータ社の CSP-001 という FPGA (Field Programmable Gate Array) 評価ボードを用いて実装した。

パルスジェネレータからは 30Hz と 4.29Hz のパルス信号を出力するので、まず FPGA 評価ボードのベースクロック信号を 30Hz にまで分周して、高フレームレートカメラにはそのまま 30Hz の信号を、低フレームレートカメラには 30Hz の信号を 7 進カウンタに通して 4.29Hz に分周したものを出力する。それぞれのカメラに送られたパルス信号は、カメラの取り込みボードに送られ、そこから取り込みトリガが生成され、カメラがトリガを検知して画像を取り込んでいく。これにより、高解像度低フレームレート動画画像と低解像度高フレームレート動画画像を 1:7 の割合で同期して撮像できた。

4-4-4 高解像度動画配信用ソフトウェアの開発

本サブテーマでは、複合センサカメラから得られる 2 つの全方位動画から、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画像を入力および記録する技術を開発した。試作したシステムは、複合センサカメラの出力をとりこぼしなく実時間で記録できること

を確認した。

4-5 高解像度動画配信用ソフトウェアの開発

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。

1. 現存プロトコルでは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は、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必要がある。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、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。また、映像配信には、従来の MPEG 規格を利用した。

4-5-1 受信クライアントソフトウェアの開発

ここでは本研究の目的であるストリーミング再生プレイヤーの開発において、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。図 51 に多チャンネル動画配信の 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示す。

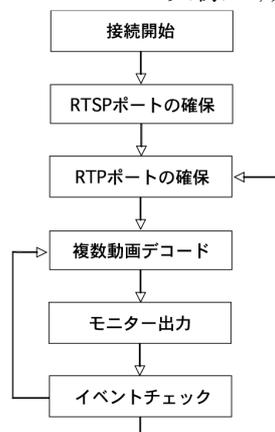


図 51 MPlayer の処理の流れ

a. 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 種類の通信モジュールが担っている。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とされる通信モジュールの役割は、サーバから受け取った情報を映像生成モジュールに渡し、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。

b. 映像出力モジュール

サーバから受信した動画ファイルのデコードや表示は decode_video, flip_page, vo_check_event という映像生成モジュールが担っている。本研究で必要とされる映像生成モジュールの役割は、多チャンネルによって配信された複数のファイルを読み込んで一つに合成し、ユーザに透視投影画像を提示し、キーボード操作によって視点変更を可能にするというものである。

(2) 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。

表 6 評価実験環境

OS	Fedora Core 4
CPU	Intel Xeon 2.40 GHz
Compiler	gcc version 3.2.3
GPU	NVIDIA Quadro4 280NVS
cache memory	512KB
main memory	1GB

a. 実験方法

フレームレートを計測する為に、画面を更新する関数の呼び出し回数をカウントした。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に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。

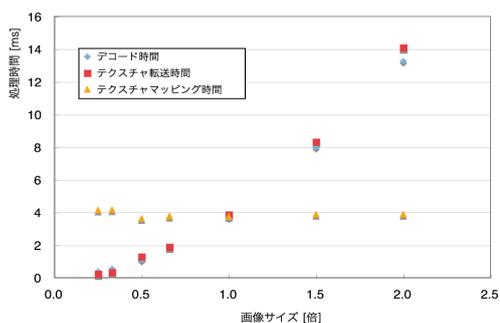


図 52 画像サイズと処理時間
(ビットレート 180kbps)

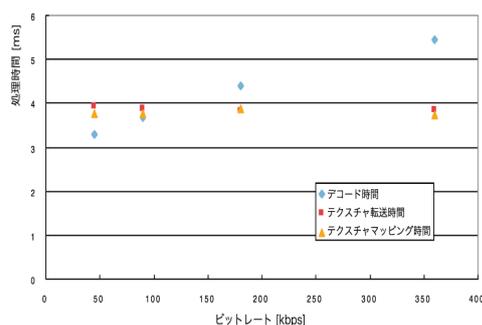


図 53 ビットレートと処理時間
(画像サイズ 1 倍)

b. 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52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53に、分割格子数と処理時間の関係を表7に示す。図52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53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がないということが分かる。

表7 格子点の計算時間

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

以上、画像サイズ、分割格子数、変換式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。

4-5-2 配信サーバの開発

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する。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する。

(1) 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は映像処理サーバーで複数の小さな低解像度動画に分割して、ストリーミング配信サーバーの動画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をクライアントに転送する。通信の概略を図 54 に示す。

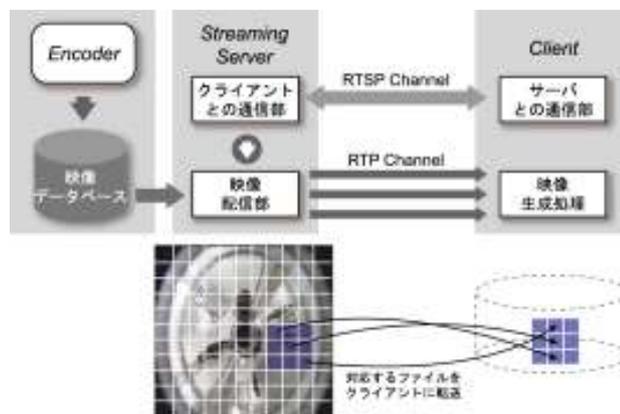


図 54 サーバー・クライアント通信概略図

(2) ストリーミングサーバーの実装

a. 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を同期化することができる。

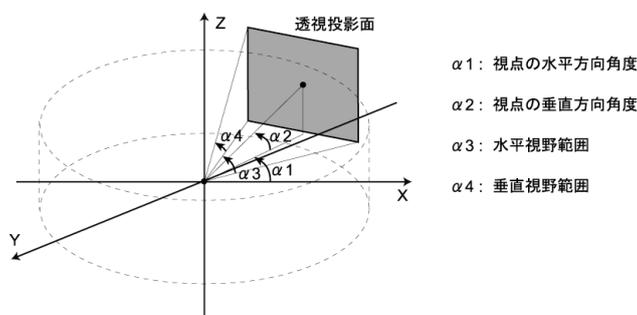


図 55 視点情報の 4 つのパラメータ

b. 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55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56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の各動画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57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．

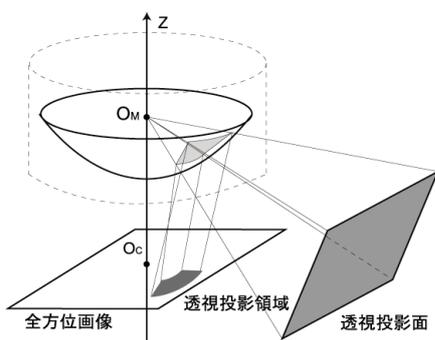


図 56 視点情報の対応領域の透視投影

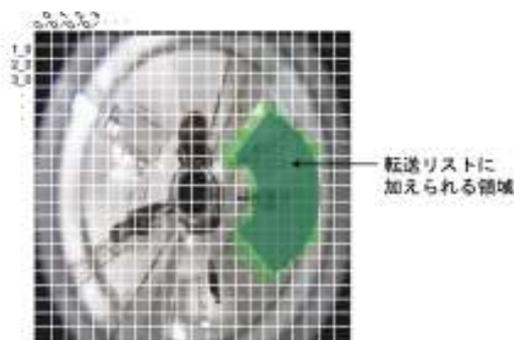


図 57 転送ファイルリストの生成

c. 画面のレイヤー分割

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．

クライアントが視野を画面の周辺に指定する場合，対応するファイル数は非常に増加する．ストリームが増加すると，サーバの負担が増大すると考えられる．サーバの負担を軽減するため，画面の分割については，周辺のところを中心のところより大きな格子で分割する．これにより，クライアントは視野を全方位動画の周辺のところ指定しても，対応するストリーム数はそれほど多くなならない．図58のように，画面を周辺から中心へ4つのレイヤーで分割し，更に各レイヤーの同じ面積に含まれるファイル数の比を1:4:4:16として，画面の分割を行う．

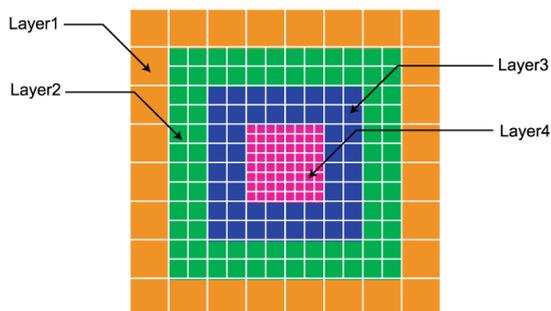


図 58 画面のレイヤー分割

画面のレイヤー分割方法は，ストリーム数を減らすことができるが，単一低解像度画像

の面積が大きくなるため、透視投影領域よりもっと大きな無駄な配信領域が生ずる、従って、配信帯域は増大する。レイヤーあり分割方法とレイヤーなし分割方法の2つの方法におけるサーバ配信能力の比較は評価実験のところで述べる。

4-5-3 まとめ

本サブテーマでは、蓄積された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、配信された全方位動画像を閲覧するクライアントソフトウェアを開発した。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、それらを効率よく配信し、動画表示するシステムを開発した。

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部分の回路を実装し、画像取り込みのテストを行った。

4-6-1 FPGA による画像補正

(1) ハードウェア化のためのカメラモデル

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。4-1-3 節では、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。本節では、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。

実装では冗長な幾何補正計算を避けるため、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元の低解像度観測画像上の点の座標を求める。前節で示したカメラモデルでは、世界座標系の点 $P(x_w, y_w, z_w)$ の画像平面への透視投影点 $P_u(X_u^{low}, Y_u^{low}, Z_u^{low})$ から、レンズ歪みを考慮した実際の画像平面への投影点 $P_d(X_d^{low}, Y_d^{low}, Z_d^{low})$ を求める際に3次方程式を解く必要が出てくる。しかし、ハードウェアで3次方程式を解くには多くのクロックが必要となる。そこで、座標計算の高速化をはかるため低解像度カメラの歪みモデルを変更した。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った。

表8 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	2 MB/bank, 4MB total
SDRAM	512MB/bank, 1024MB total

(2) 画像補正処理システムの構成

画像補正処理は前節で述べたように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。画像処理ボ

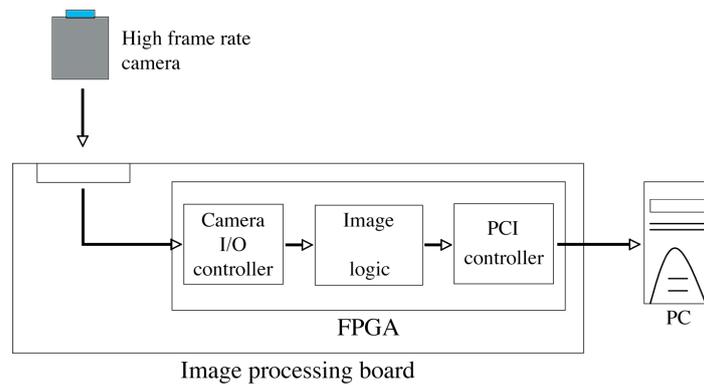


図 59 FPGA 内での処理の流れ

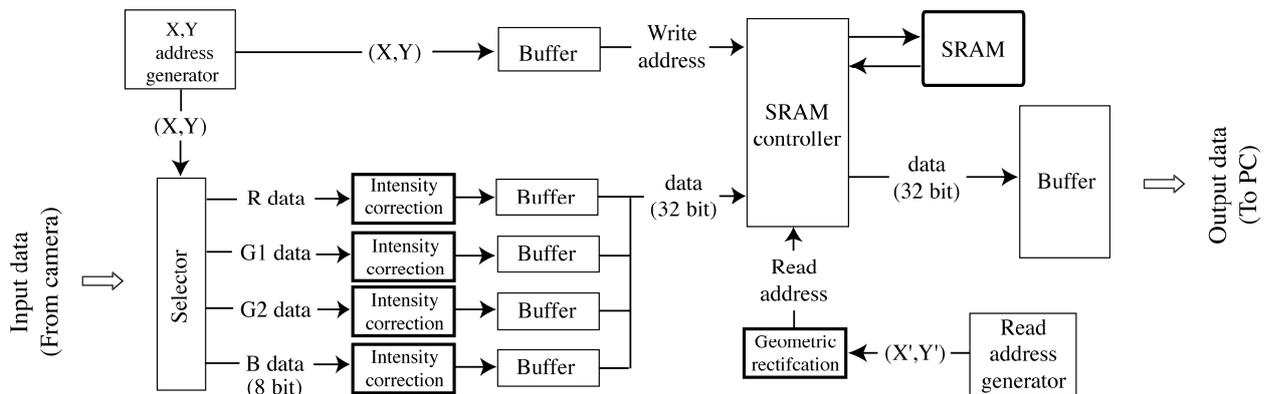


図 60 画像補正の流れ

ードとして soliton 社の Tsunami PCI A40 を用いた. 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 8 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 Discription 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. 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 59 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, ここでは画像補正処理の実装手法について述べる.

(3) 画像補正処理

FPGA 内における画像補正ブロックでの処理を図 60 に示す. カメラの CCD は Bayer 配列となっているため, 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. VHDL では多ビット同士の除算には多くの時間がかかるが, 本研究では入力データ量は 30MB/s で一定のため 1 つの処理にかけることのできる時間には制限がある. そこで, 本研究では多ビット同士の除算を行う際は複数の処理に分割して演算を行うことで処理時間の制限を満たす. また, 除算による桁落ちの影響を少なくするため, 先に前節で述べた数式をまとめて除算回数を減らす.

(4) ハードウェアによる画像補正の精度評価

次に、画像補正処理を FPGA に実装し、精度検証を行った。前節で述べた方法を FPGA に実装し、低解像度カメラで撮像を行うことで、画像補正を行いながら 30fps で画像取得ができることが示された。高解像度画像および補正後の低解像度画像をそれぞれ図 61 (a), 図 61 (b) に示す。図 62 に示した高解像度画像を補正後の低解像度画像と同じサイズに縮小し、PSNR を計算した結果を表 9 に示す。また、差分画像を図 63 に示す。ただし、差分画像は絶対値で表現している。PSNR の比較や差分画像からハードウェアでもソフトウェアと同等の精度で補正ができていることが確認できた。



図 61 入力画像

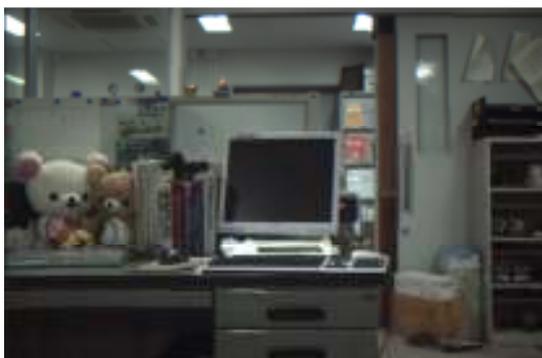


図 62 ハードウェアによる補正後の低解像度画像



図 63 差分画像

表 9 PSNR 比較

	ハードウェア [dB]	ソフトウェア [dB]
R	25.13	25.20
G1	25.42	25.81
G2	26.48	25.82
B	26.68	26.42
平均	25.93	25.81

4-6-2 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 により、実時間でソフトウェア処理と遜色ない精度で変換できることを確認した。

4-7 マルチアクセス時の性能評価

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つので、本研究でストリーミングサーバーとして用いた。詳しいスペックも表 10 に示す。

表 10 マルチアクセス実験に使用した計算機の仕様

System unit	ProLiant 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

オペレーティングシステムについては、本研究では Red Hat Enterprise Linux を用いた。Red Hat Enterprise Linux は、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、Linux の 2.6.9 カーネルを基に強力に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た。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、テストを行った。ネットワーク構成図を図 64 に示す。

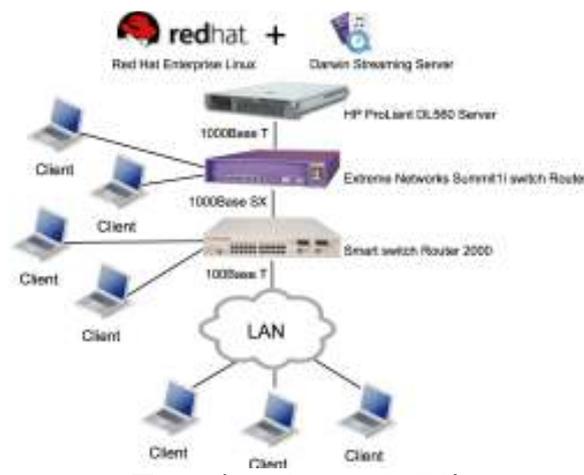


図 64 ネットワーク環境

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。全方位動画像の特徴により、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. 実験結果の一部を以下の表に示す.

(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

表 11 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	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			

(2) レイヤーあり分割方法

View Point: 0_30_30_30

各クライアントのストリーム数: 15

各クライアントのデータ転送量: 13.648MB

表 12 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	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			

実験の結果により, 転送ビットレートが約 15MB/sec 程度であれば, サーバーが遅延なく配信できることが示された. また, 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. 無駄な配信領域が生ずるが, クライアントの受信帯域を越えないことを条件として, 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, ストリーミングサーバーの最大転送ビットレートを超えてないが, サーバーが配信できる最大ストリーム数の限界に至った. 実験データより, 規定された条件において, サーバーが安定に配信できる最大ストリーム数は 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 65 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。

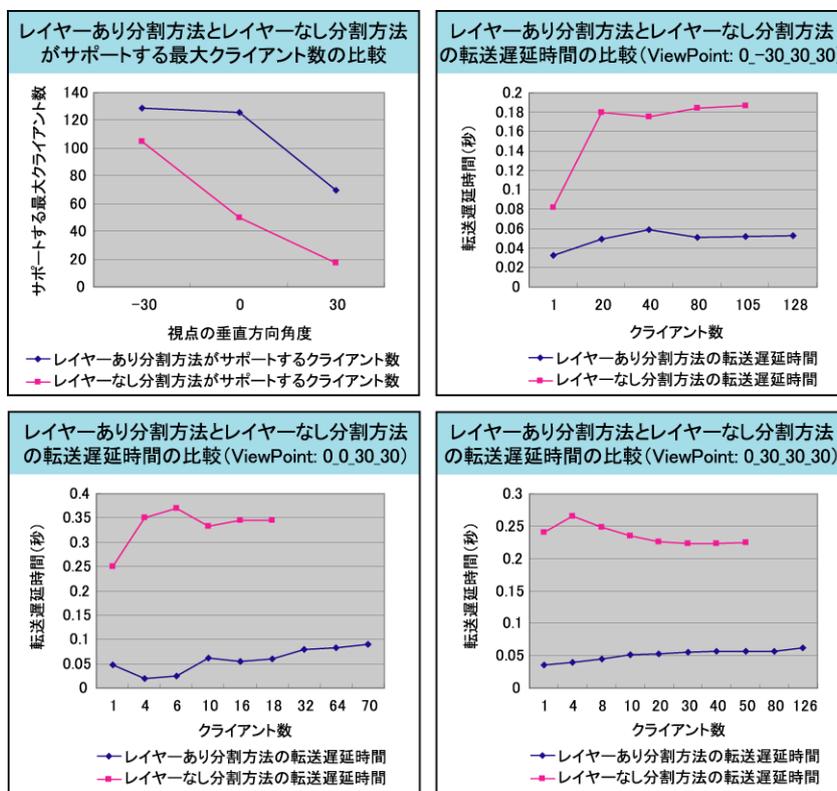


図 65 サーバパフォーマンスの比較

以上、本研究が提案したストリーミングサーバーの実際の配信能力の評価を行った。また、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。

4-7-3 まとめ

本サブテーマでは、構築した配信システムを用いて、配信サーバの性能評価を行った。また、サーバの配信性能を向上させるため RTP を改良し、多チャンネル映像を効率的に配信する手法を提案した。また、シミュレーション実験によりその改良手法の有効性を確認した。

4-8 統合システムの試作と評価

統合システムの全体図を図 66 に示す。統合システムは、複合センサシステム、映像合成システム、映像配信システム、表示クライアントシステムからなる。複合視覚センサシステムの全体図を図 67 に示す。また、表示クライアントの出力例を図 68 に示す。これらのシステムを利用して、近畿周辺の神社や仏閣、景勝地などで高精細画像コンテンツの撮影を行った。明石海峡大橋で行った撮影風景を図 69 に示す。これらの撮影データを用いて高解像度画像生成実験や統合システムにおける画像配信実験を行った。(統合システムの性能評価に関しては、映像配信システムの性能評価と同一のため 4-7 節を参照していただきたい)。また、学会や映像機器の展示会などに本システムを持ち込み、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った。

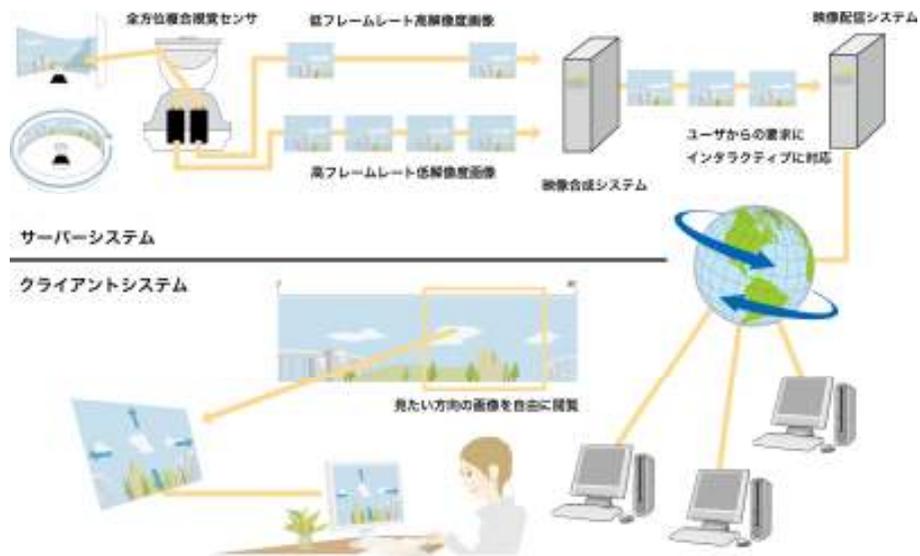


図 66 統合システムの全体像

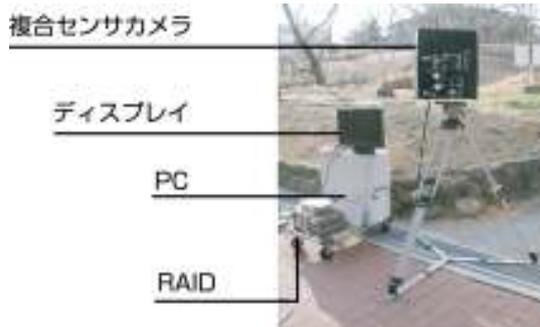


図 67 試作した複合カメラシステム



図 68 クライアントシステムの表示例



図 69 明石海峡大橋での撮影実験



図 70 MIRU2005 におけるデモ

特に、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、図 70 に示すように、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, MIRU2006)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。

4-9 総括

本研究では、全方位、高精細、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、それを高効率で配信するシステムの開発を行った。

1. まず、複合視覚センサの開発では、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を試作し、その評価を行った。また、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。
2. リアルタイム高解像度動画の作成では、試作した複合センサカメラから得られる 2 つの動画から、空間的にも高解像度で、しかも時間的にも密な動画を作成する技術を確認した。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、低解像度高フレームレート動画) 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、MPEG 動画による 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。また、8000x8000 画素の超高解像度画像における提案手法の評価実験を行い本提案手法の有効性を検証した。
3. 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した。また、ミラーの小型化のために、新たに考案した 2 枚反射系全方位ミラーを試作し、小型化を行っても十分な性能が得られることを確認した。
4. 全方位高解像度リアルタイム動画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を、入力および記録する技術を開発した。試作したシステムは、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ことを確認した。
5. 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 種類の全方位動画から閲覧者側において高解像度かつリアルタイム (30 フレーム/秒) の動画を作成する技術を開発した。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表示するシステムを実現した。特に、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い、本提案手法の有効性を検証した。
6. 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ソフトウェアで行った画像補正をハードウェア化により高速化を行った。FPGA により実時間でソフトウェア処理と遜色ない精度で変換出来ることを確認した。
7. マルチアクセスの時の評価では、構築した配信システムを用いて配信サーバの性能評価を行った。また、サーバの配信性能を向上させるために RTP を改良し、多チャンネル映像を効率的に配信する手法を提案した。また、シミュレーション実験によりその改良手法の有効性を確認した。
8. 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い、それぞれの個別研究が統合しても正しく動作し、期待した性能を発揮することを確認した。

以上のように、提案時に計画した目標はほぼ達成できたが、本研究は新規性の高い基盤技術であるため、論文・学会発表にも注力した。今後、さらに成果を発表していく予定であるが、すでに国際発表 11 件を含む 26 件という数多くの論文、学会発表を行うことができ、学術的にも満足のいく成果が得られたと自負している。

「高解像度だが時間的には粗い画像と低解像度だが時間的に密な画像から高解像度でかつ時間的に密な画像を作成する」という着想の新しさから、多くの方々が本研究に興味を持たれ、活発な議論が行われた。また、この複合センサの考え方は、その後、各種の異なるセンサを組み合わせた複合センサが研究されるようになったという波及効果を持った。

この2種類の動画画像から高解像度でかつ時間的に密な画像が作成できる方法は、複合センサとしてのみでなく、圧縮率の高い効率的な記録技術、配信技術としても注目されるようになってきた。これは、提案書の時点では強く意識していなかったが、今後、波及していく可能性の強い技術と思われる。

また、論文学会発表のほか知的財産の獲得にも注力した。国際特許2件を含む8件の特許を申請した。そのうちの1件は、請求項目が40項目の多くから構成され、本研究をほぼ網羅したもので知的財産から見た技術優位性も高いと自負している。

5 参考資料・参考文献

- [1] R. Y. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shell TV cameras and lenses, *IEEE J. Robot. Automat.* RA-3, No. 4, pp. 323-344 (1987)
- [2] X.-D. Yang, Q. Xiao, and H. Raafat: Direct mapping between histograms: An improved interactive image enhancement method, In *IEEE Intl. Conf. on Systems, Man and Cybernetics*, pp. 243-247 (1991)
- [3] C. Tomasi, T. Kanade, “Detection and Tracking of Point Features,” Carnegie Mellon University Technical Report, No. CMU-CS-91-13, April, 1991.
- [4] J. Shi, C. Tomasi, “Good Features to Track,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 593-600, 1994.
- [5] Shewchuk, J.R., “Triangle: Engineering a 2D Quality Mesh Generator and Delaunay Triangulator,” *First Workshop on APplied Computational Geometry, Association for Computing Machinery*, pp. 124-133, 1996.

5-1 研究発表・講演等一覧

注：番号毎末尾の表記内容、例（H.16-1）は5-3外部発表等の状況の外部発表年度と内容番号を示す。

- [1] 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”, 画像の認識・理解シンポジウム(MIRU), pp. I315-I320, 函館, Jul, 2004. (H.16-3)
- [2] 渡邊清高, 岩井儀雄, 長原一, 谷内田正彦: “時空間周波数の異なる画像列からの高解像度動画の合成”, 情報科学技術レターズ, Vol. 3, No. LI-004, pp. 169-172, Sep, 2004. (H.16-1)
- [3] 間下以大, 岩井儀雄, 谷内田正彦: “ミラーの外縁を利用した全方位視覚センサのキャリブレーション”, 情報処理学会研究報告, No. CVIM147, pp. 21-28, 京都, Jan, 2005. (H.16-3)
- [4] 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”, 情報処理学会研究報告, No. CVIM147, pp. 83-90, 京都, Jan, 2005. (H.16-3)
- [5] 松延徹, 長原一, 岩井儀雄, 谷内田正彦, 田中紘幸, “モーフィングによる高解像度高フレームレート動画の生成”, 電子情報通信学会技術報告, Vol. PRMU104, No. 573, pp. 85-90, Jan, 2005. (H.16-3)
- [6] 渡邊清高, 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の生成”, 画像の認識・理解シンポジウム(MIRU), Jul, 2005. (H.17-3)
- [7] 松延徹, 星川章, 重本倫宏, 渡邊清高, 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の撮像・提示システム”, 画像の認識・理解シンポジウム(MIRU), Jul, 2005. (H.17-3)
- [8] Akira Hoshikawa, Tomohiro Shigemoto, Hajime Nagahara, Yoshio Iwai, Masahiko Yachida, Hiroyuki Tanaka: “Dual Sensor Camera System with Different Spatio-Temporal Resolution”, *SICE Annual Conference*, Aug, 2005. (H.17-2)
- [9] Kiyotaka Watanabe, Yoshio Iwai, Hajime Nagahara, Masahiko Yachida, Hiroaki Tanaka: “Video Synthesis with High Spatio-Temporal Resolution Using Motion Compensation and Spectral Fusion”, *SICE Annual Conference*, Aug, 2005.

(H. 17-2)

- [10] Toru Matsunobu, Hajime Nagahara, Yoshio Iwai, Masahiko Yachida, Hiroyuki Tanaka: "Generation of High Resolution Video Using Morphing", SICE Annual Conference, Aug, 2005. (H. 17-2)
- [11] 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. (H. 17-2)
- [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. (H. 17-2)
- [13] 重本 倫宏, 星川 章, 長原 一, 岩井 儀雄, 谷内田 正彦, 鈴木俊哉: "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, Vol. 47, No. SIG5(CVIM13), pp. 35-44, Mar, 2006. (H. 17-1)
- [14] 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", 情報処理学会論文誌: コンピュータビジョンとイメージメディア, Vol. 47, No. SIG5, pp. 45-58, Mar, 2006. (H. 17-1)
- [15] 岩井儀雄, 長原一, 鈴木俊哉: "全方位高解像度動画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. (H. 17-9)
- [16] Kiyotaka Watanabe, Yoshio Iwai, Hajime Nagahara, Masahiko Yachida and Toshiya Suzuki: "Video Synthesis with High Spatio-Temporal Resolution Using Motion Compensation and Spectral Fusion", IEICE Transactions on Information and Systems, Vol. E89-D, No. 7, pp. 2186-2196, Jul, 2006. (H. 17-1)
- [17] 渡邊清高, 岩井儀雄, 長原一, 谷内田正彦, 鈴木俊哉: "ウェ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, Vol. 47, No. SIG 10(CVIM 15), pp. 195-207, Jul, 2006. (H. 17-1)
- [18] Tomohiro Mashita, Yoshio Iwai, and Masahiko Yachida: "Calibration Method for Misaligned Catadioptric Camera", IEICE Trans. on Information & Systems, Vol. E-89-D, No. 7, pp. 1984-1992, Jul, 2006. (H. 17-1)
- [19] 鈴木俊哉, 谷内田正嗣 "全方位視覚センサとその応用" 映像情報 2006. 4. 17 発行 (H. 18-9)
- [20] 中村在宏, 長原一, 岩井儀雄, 谷内田正彦, 鈴木俊哉: "複合センサ画像の補正処理の高速化 情報処理学会研究報告", Vol. CVIM-154, No. 1, pp. 1-8, 奈良, May, 2006. (H. 18-3)
- [21] 中村在宏, 長原一, 岩井儀雄, 谷内田正彦, 鈴木俊哉: "複合センサカメラのための実時間補正処理", 画像の認識・理解シンポジウム(MIRU), No. DS-11, pp. 1391-1392, 仙台, July, 2006. (H. 18-3)
- [22] Hajime Nagahara, Toru Matsunobu, Yoshio Iwai, Masahiko Yachida, Toshiya Suzuki: "High-resolution Video Generation Using Morphing", Proc. IAPR Int. Conf. Pattern Recognition(ICPR), Hong Kong, Aug, 2006. (H. 18-2)
- [23] 上谷巧, 西尾佳祐, 岩井儀雄, 長原一, 谷内田正彦, 鈴木俊哉: "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", 第5回情報科学技術フォーラム講演論文集(FIT), No. L-062, pp. 151-154, 福岡, Sep, 2006. (H. 18-3)

- [24] Yang Haoran, Wang Xiaoqian, Yoshio Iwai, Hajime Nagahara, Masahiko Yachida, Toshiya Suzuki: “ Multiple-channel Streaming Delivery for HyperOmni Vision System” , 第5回情報科学技術フォーラム講演論文集(FIT), No. L-063, pp. 155-158, Sep, 福岡, 2006. (H.18-3)
- [25] Kiyotaka Watanabe, Yoshio Iwai, Hajime Nagahara, Masahiko Yachida, Toshiya Suzuki: “Video Synthesis with High Spatio-Temporal Resolution Using Spectral Fusion” , Proc. Int. Workshop Multimedia Content Representation, Classification and Security(MRCS), Vol. LNCS4105, pp. 683-690, Istanbul, Turkey, Sep, 2006. (H.18-2)
- [26] 松延 徹, 長原 一, 岩井 儀雄, 谷内田 正彦, 鈴木 俊哉: “モーフィングによる高解像度高フレームレート動画の生成” , 電子情報学会論文誌, 2007, 採録決定. (H.18-1)