平成15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」

1 研究開発課題の背景
2 研究開発分野の現状 ······4
 3 研究開発の全体計画 4 3-1 研究開発課題の概要 4 3-1-1 複合センサカメラの試作 5 3-1-2 リアルタイム高解像度動画像の作成 6 3-1-3 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 9 3-1-4 高解像度動画配信用ソフトウェアの開発 9 3-1-5 ライブ入力,配信システムの開発 9 3-1-6 マルチアクセス時の性能評価 10 3-1-7 総合システムの試作と評価 10 3-2 研究開発目標 10 3-2-1 最終目標 10 3-2-2 中間目標 11 3-3 研究開発の年度別計画 12
3-4 研究開発体制 ······13
 4 研究開発の概要(平成15年度まで) 4-1 研究開発実施計画 4-1-1 研究開発の計画内容 4-1-2 研究開発の計画内容 4-1-2 研究開発の実施内容 4-2 研究開発の実施内容 4-2-1 複合センサカメラの試作 4-2-2 リアルタイム高解像度動画像の作成 4-2-3 高解像度全方位カメラ用光学系の設計 16 4-2-4 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・16
5 研究開発実施状況(平成15年度)・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・

目 次

5-2-1	周波数空間アプローチによる高解像度画像の作成38
5-2-2	モーフィングアプローチによる高解像度画像の作成 ・・・・49
5-2-3	まとめ ・・・・・・65
53 葿	高解像度全方位カメラ用光学系の設計 ・・・・・・・・・・66
5-3-1	全方位視覚センサ ・・・・・66
5-3-2	全方位ミラーの設計 ・・・・・68
5-3-3	まとめ
5−4 ≦	き方位高解像度リアルタイム動画像入力記録方式の開発・72
5-4-1	複合センサカメラシステムを利用したリアルタイム
重	カ画像記録 ・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・
5-4-2	書き込み速度の最適化 ・・・・・・・・・・・・・・・・
5-4-3	パルスジェネレータの作成
5-4-4	まとめ
5-5 総括	岳 ··········77

参考資料,参考文献

(添付資料)

1 研究発表,講演,文献等一覧

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(640 x 480 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.

2 研究開発分野の現状

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,多数のカメラを並ベ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.

3 研究開発の全体計画

3-1 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.

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.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

3-1-1 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. 基本設計は研究担当者らが行うが、実際の試作は光学機器の開発に高い技術力を持 つ協力会社に依頼する.



図2 複合センサカメラ

3-1-2 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.



3-1-3 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.なお、複合 双曲面ミラーの設計は研究担当者らが行うが、その試作は協力会社(山田光学)で行う.

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.



図6 全方位視覚センサの光学系

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.

3-1-4 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.

3-1-5 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、配信サーバは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

3-1-6 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.

3-1-7 総合システムの試作と評価

エジプトのピラミッドや日本の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.

3-2 研究開発目標

3-2-1 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.

3-2-2 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.

3-3 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
 (町九開光項日 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 					十皮	ĀI	'''' 方
8.総合システムの試作と評価 ・データ収集,総合評価							
間接経費額(税込み)							
おおお 合い 計 のうちょう おうしょう おうしょう しょうしょう おうしょう しょうしょう ひょうしん ひょう ひょうしん ひょうしん ひょうしん ひょうしん ひょうしん ひょうしん ひょうしん ひょう							

3-4 研究開発体制



4 研究開発の概要

4-1 研究開発実施計画

4-1-1 研究開発の計画内容

(1) 複合センサカメラの試作

カメラ視点(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(640 x 480 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(4000 x 4000 画素, 30 フレーム/秒) を作成する技術を開発する.

まず,その第一歩として,同一の視野を持った2種類の異なった解像度を持つ動画像 を撮影し蓄積する複合センサカメラを試作する.

(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.

(3) 高解像度全方位カメラ用光学系の設計

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.

今回対象とする CCD 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、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.

(4) 方位高解像度リアルタイム動画像入力記録方式の開発

全方位複合センサカメラを用いて,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)を入出力および記録する技術を開発す る.

4-2-1 研究開発実施計画

(金額は非公表)

研究開発項目	第1四半期	第2四半期	第3四半期	第4四半期	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 						
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.). (合計の計は,「3-1の研究開発課題必要概算経費」の総額と一致)

2 備考欄に再委託先機関名を記載

4-2 研究開発の実施内容

4-2-1 複合センサカメラの試作

レンズ主点(射影中心)を合わせることにより,同一の視野を持つ時空間分解能の異なる2種類の動画像 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)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)動画像を同時に撮像するカメラシステムを構築する.

4-2-2 リアルタイム高解像度動画像の作成

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.

4-2-3 高解像度全方位カメラ用光学系の設計

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. この全方位複合センサカメラを用いて,時空間分解能の異なる2種類の全方位動画像を 同時撮像できる.

対象とする CCD は, 9.0×9.0µm のセルサイズの 4000×4000 画素という高解像度センサを想定しているため,従来設計してきたものに比べてスペックが1ランク上となる. 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シミュレータを用いて行う.

4-2-4 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

全方位複合センサカメラを用いて,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)を入出力および記録する技術を開発す る.

5 研究開発実施状況

5-1 複合センサカメラの試作

5-1-1 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.そんな高品質な映像を実現しようと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[1]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,ディジタルシネマが提案されており,映画の撮影から上映までをディジ タルで行う.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.これらのよう な撮像システムや映像配信システムの開発からも世の中の高品質な映像に対するニー ズの重要性がわかる.しかし,これらのシステムは,非常に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困難である.

一方,一般的には従来のテレビカメラは解像度が 640×480 画素でフレームレートが 30fps であり,フレームレートは申し分ないが,高解像度であるとはいい難い.また, 近年解像度を重視したディジタルスチルカメラなどでは解像度が 4000×4000 画素とい うものがあるがフレームレートは 1fps 程度しかなく,滑らかな動画を撮影するには無理がある.このように,画像の吐き出し速度の限界のため高解像度と高フレームレート を両立させるのは困難であった.

そこで,高解像度低フレームレート画像と低解像度高フレームレート画像を撮像し, これら2つの画像から高解像度高フレームレート動画像を生成する.また,その時間的, 空間的に分解能の異なる2種類の動画像を撮像する複合センサカメラシステムを試作 した.

5-1-2 複合センサカメラシステム

従来より, 複数のカメラ(センサ)を組み合わせて高解像度画像を得ようとするアプロ ーチが数多く提案されてきた.ここでは従来の研究を紹介して, 提案システムとの違い について述べていく.

杉田らは 2 台の視野角の異なるカメラを用いて中心視野付近の解像度を重視して撮像するシステム[2]を提案している.このシステムは、45°に傾けたハーフミラーをはさんで光学的に共役な位置に 2 台のカメラを設置する.一方のカメラで視野を中心付近にしぼってその部分を高解像度に撮像し、もう一方のカメラで広視野を撮像する.広視野を撮像した画像に、中心付近を高解像度に撮像した画像を貼り付けて、広視野かつ中心視野付近の解像度を高めた画像を得ることができる(図8).

また,田中らは2台のカメラで撮像視野角を分けて撮像することにより広視野で高解 像度画像を実現するシステム[3]を提案した.このシステムは画角θを有する2台のカ メラと2枚の鏡を用いて,図9のように配置することによって1台のカメラで撮像でき る画角の倍である2θの画角を有する高解像度映像を撮像することができる.また,続 らによるさらに多くのカメラを用いて垂直方向にも視野を分けて高解像度な画像を撮 像するシステム[4]もある.

菅原らは、多板 CCD(charge coupled device)カメラの CCD 間で画素ずらしを利用して

解像度を高める手法[5]を提案した. この手法はGのCCDに対して, R, BのCCDを斜め 方向(水平方向に1/2 画素, 垂直方向に1/2 画素)に画素ずらしを行うことにより高解像 度画像を撮像できる. GとRの画素ずらしを図10に示す.





図10 斜め方向の画素ずらし



低解像度高フレームレートカメラ



これらの研究に対して, 我々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 の動画像を撮像する.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種類の動画像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11).

ここで本研究と複数カメラを用いた撮像システムに関する従来の研究[2][3][4][5] の違いを表1にまとめた.表1からわかるように従来の研究は、いずれも複数の同じフ レームレートのカメラで部分撮像しつなぎ合わせることで空間解像度のみを高めてい る空間統合であるといえる.それに対して、本研究では時間的、空間的な分解能の異な る2種類の動画像を統合することにより、解像度だけでなくフレームレートも高めてい く時空間統合であるといえ、これはまったく新しいアプローチである.

手法	異なる視野角を 撮 像[2]	視野角を分割して 撮像[3][4]	CCD 間の 画素ずらし[5]	本研究
特徴	視野の中心部分の 解像度を高める	広視野で高解像度 な画像が得られる	CCD の画素ずら しを利用して解像 度を高める	解像度, フレームレ ート共に高める

表1 従来の研究と本研究のアプローチの違い

5-1-3 ハードウェア構成

図12に示すように本システムは、高解像度低フレームレートカメラ、低解像度高フレームレートカメラ、2台のカメラで同一視点、同一視野を撮像するためにシーン情報を分光するキューブ型のプリズム、フレームレートの異なる2台のカメラの同期をとるためのパルスジェネレータ、それらを制御するワークステーション、取り込んだ画像を記憶する RAID (Redundant Arrays of Inexpensive Disks)システムのハードディスクから構成される.

2 台のカメラの仕様を表 2 に,画像を記憶するための RAID システムのハードディス クの仕様を表 3 に示す.また,キューブ型のプリズムとしてエドモンド社のキューブ型 ビームスプリッタを,ワークステーションとして(hp 社)hp workstation xw6000 を使用 する.



図12 システム全体の構成図

これらの2台のカメラで同一視点,同一視野を撮像するためにプリズムの透過光を一 方のカメラで,反射光をもう一方のカメラで撮像する.得られた画像をワークステーシ ョン上に取り込んで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 に記憶していく.ここで高解像度低フレームレートカメラは 5fps,低解像度高フレー ムレートカメラは 30fps であるので,パルスジェネレータからそれぞれのカメラにパル ス信号を送り,シャッターを駆動することで2台のカメラを同期させる.図13に示す ように高解像度低フレームレートカメラで撮像するタイミングで低解像度高フレーム レートカメラでも時間的に同じシーンを撮像する.2台のカメラを同期させるために高 解像度低フレームレートカメラには 5Hz,低解像度高フレームレートカメラには 30Hz のパルス信号をそれぞれ送ることにする.

	ES11000(REDLAKE 社)	A201bc(BASLER 社)			
	(高解像度	(低解像度			
	低フレームレート)	高フレームレート)			
解像度	4008×2672	1008×1018			
フレームレート	5fps	30fps			
出力画像	Bayer Color	Bayer Color			
出力フォーマット	8 or 10 or 12bits/pixel	8 or 10bits/pixel			

表2 カメラの仕様

表3 RAIDシステムのハードディスク仕様

	Newtech SCSI-RAID System C	onfidence ATA II Fe
	(ニューテック社)	
最大転送速度	160Mbyte/sec	
最大容量	300GByte	
最大書き込み速度	100Mbyte/sec	



図13 同期のとれた画像列

5-1-4 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、CCD 上 で R、G、B は図14のように配置されている. なお G 成分については後の色補間処理の 説明のため G₁、 G₂と表記することにする.

図14からわかるように、各画素には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.例えば、図14のRの部分ではG成分とB成分がないので近 傍の画素から推定することになり、G成分は隣接する上下左右のGから、B成分は斜め 4方向のBから補間していく.また、一番上と一番下の行、右端と左端の列にある画素 は近傍画素から補間できないので、補間した後の画像は行、列共に2減ることになる. 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.以下に色の補間処理の式を示 す.

G 1	R	G ₁	R
В	G 2	В	G 2
G ₁	R	G ₁	R
В	G ₂	В	G ₂

図14 Bayer 配列

まず,図14のG₁における補間処理の式を以下に示す.

$$R(x_{G_1}, y_{G_1}) = (r(x_{G_1} - 1, y_{G_1}) + r(x_{G_1} + 1, y_{G_1}))/2$$

$$G(x_{G_1}, y_{G_1}) = g(x_{G_1}, y_{G_1})$$

$$B(x_{G_1}, y_{G_1}) = (b(x_{G_1}, y_{G_1} - 1) + b(x_{G_1}, y_{G_1} + 1))/2$$
(1)

ただし, (x_{G1}, y_{G1}) における補間する前の R, G, B の値をそれぞれ $r(x_{G1}, y_{G1})$, $g(x_{G1}, y_{G1})$, $b(x_{G1}, y_{G1})$ とし, (x_{G1}, y_{G1}) における補間した後の R, G, B の値をそれぞれ $R(x_{G1}, y_{G1})$, $G(x_{G1}, y_{G1})$, $B(x_{G1}, y_{G1})$ とする.

同様に図14のRにおける補間処理の式を以下に示す.

$$R(x_{R}, y_{R}) = r(x_{R}, y_{R})$$

$$G(x_{R}, y_{R}) = (g(x_{R} - 1, y_{R}) + g(x_{R} + 1, y_{R}) + g(x_{R}, y_{R} - 1) + g(x_{R}, y_{R} + 1))/4$$

$$B(x_{R}, y_{R}) = (b(x_{R} - 1, y_{R} - 1) + b(x_{R} - 1, y_{R} + 1) + b(x_{R} + 1, y_{R} - 1) + b(x_{R} + 1, y_{R} + 1))/4$$
(2)

ただし, (x_{R}, y_{R}) における補間する前の R, G, B の値をそれぞれ $r(x_{R}, y_{R})$, $g(x_{R}, y_{R})$, $b(x_{R}, y_{R})$ とし, (x_{R}, y_{R}) における補間した後の R, G, B の値をそれぞれ $R(x_{R}, y_{R})$, $G(x_{R}, y_{R})$, $B(x_{R}, y_{R})$ とする.

図14の G_2 における補間処理の式を以下に示す.

$$R(x_{G_2}, y_{G_2}) = (r(x_{G_2}, y_{G_2} - 1) + r(x_{G_2}, y_{G_2} + 1))/2$$

$$G(x_{G_2}, y_{G_2}) = g(x_{G_2}, y_{G_2})$$

$$B(x_{G_2}, y_{G_2}) = (b(x_{G_2} - 1, y_{G_2}) + b(x_{G_2} + 1, y_{G_2}))/2$$
(3)

ただし, (x_{62}, y_{62}) における補間する前の R, G, B の値をそれぞれ r (x_{62}, y_{62}) , g (x_{62}, y_{62}) , b (x_{62}, y_{62}) とし, (x_{62}, y_{62}) における補間した後の R, G, B の値をそれぞれ R (x_{62}, y_{62}) , G (x_{62}, y_{62}) , B (x_{62}, y_{62}) とする.

最後に図14のBにおける補間処理の式を以下に示す.

$$R(x_{B}, y_{B}) = (r(x_{B} - 1, y_{B} - 1) + r(x_{B} - 1, y_{B} + 1) + r(x_{B} + 1, y_{B} - 1) + r(x_{B} + 1, y_{B} + 1))/4$$

$$G(x_{B}, y_{B}) = (g(x_{B} - 1, y_{B}) + g(x_{B} + 1, y_{B}) + g(x_{B}, y_{B} - 1) + g(x_{B}, y_{B} + 1))/4$$

$$B(x_{B}, y_{B}) = b(x_{B}, y_{B})$$
(4)

ただし, (x_B, y_B) における補間する前の R, G, B の値をそれぞれ $r(x_B, y_B)$, $g(x_B, y_B)$, $b(x_B, y_B)$ とし, (x_B, y_B) における補間した後の R, G, B の値をそれぞれ $R(x_B, y_B)$, $G(x_B, y_B)$, $B(x_B, y_B)$ とする.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15に,補間処理を行った後の画像を図16に示す.

図16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16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.

次に、色補間によって得た画像のカラーバランスの補正について述べていく. Gray World Assumption に基づくアルゴリズム[6]を用いてカラーバランスの補正を行なう. この手法は画像中の R, G, B の平均値は全て一定の値に収束してそれは全て等しい値になるという考えを用いている. この手法を数式に表すと以下のようになる.

$$R(x,y) = r(x,y) \times \frac{A_{ave}}{R_{ave}}$$

$$G(x,y) = g(x,y) \times \frac{A_{ave}}{G_{ave}}$$

$$B(x,y) = b(x,y) \times \frac{A_{ave}}{B_{ave}}$$
(5)



図15 入力画像(1008×1018pixel)



図16色の補間処理を行った後の画像(1006×1016pixel)

ただし, (x, y)における補正する前のR, G, Bの値をそれぞれr(x, y), g(x, y), b(x, y) とし, (x, y)における補正した後のR, G, Bの値をそれぞれR(x, y), G(x, y), B(x, y)とす る. また,画像中の全画素から求めたR,G,Bそれぞれの平均値を R_{ave} ,G $_{ave}$,B $_{ave}$ とし, 全画素のR,G,B全てを足したものの平均値を A_{ave} とする.

図17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.



図17 カラーバランス補正後の画像(1006x1016画素)

5-1-5 複合センサカメラのキャリブレーション

複合センサカメラでは2台のカメラの特性が各々異なるため、2枚の画像を一致させ るためには、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。幾何学的キ ャリブレーションとは物体が存在する3次元空間と2次元画像面の幾何学的関係およ びカメラ内部の幾何学的変換特性を求めることであり、光学的キャリブレーションとは レンズによって生じる光学的歪みのモデル化と補正を行なうことである。

今年度は、2枚の画像の幾何学的特性を一致させる手法について開発を行なった. そ れぞれのカメラに対してカメラモデルを設定し、幾何学的キャリブレーションを行う. 推定されたパラメータを用いることで、2枚の画像を3次元空間中の同一基準平面に投 影し、解像度の異なる画像の一致をはかる. さらに、シミュレーションを行うことで、 モデル誤差と観測誤差が与える影響について検討した.

(1) カメラモデル

カメラの投影モデルとしてはピンホールカメラモデルがよく用いられるが,解析の目 的によってはフォーカスなどの影響を考慮した薄凸レンズカメラモデル[7]や厚凸レン ズカメラモデル[8]も使用される.本研究ではボケの影響は考慮しないため,最も単純なピンホールカメラモデルを使用する.

図18に示すように規準となる世界座標における,点 Pの座標を (x_w, y_w, z_w) と表し, カメラ座標における点 Pの座標を(x, y, z)と表す.ここで、0はレンズ中心を表し、z 軸 はレンズの光軸に一致するように設定する.次に、x-y 平面に平行で、z 座標が f の位 置にある画像平面を考える.理想的なピンホールカメラの場合には、点 Pの座標(x, y, z)は画像平面上では (X_u, Y_u) と表すことができるが、レンズの幾何学的な歪みにより実際に は (X_d, Y_d) の位置に対応するものとする.そして、 (X_d, Y_d) のディジタル画像上での離散 化された座標を (X_f, Y_f) と表す.



図18 世界座標と画像座標の関係

これらの座標系の関係を用いると、以下の関係式が得られる.

$$d_{x}X + d_{x}X\kappa_{1}R^{2} = f \frac{r_{1}x_{w} + r_{2}y_{w} + r_{3}z_{w} + T_{x}}{r_{7}x_{w} + r_{8}y_{w} + r_{9}z_{w} + T_{z}}$$

$$d_{x}Y + d_{x}Y\kappa_{1}R^{2} = f \frac{r_{4}x_{w} + r_{5}y_{w} + r_{6}z_{w} + T_{y}}{r_{7}x_{w} + r_{8}y_{w} + r_{9}z_{w} + T_{z}}$$

$$R = \sqrt{(d_{x}X)^{2} + (d_{x}Y)^{2}}$$

$$X = X_{f} - C_{x}$$

$$Y = Y_{f} - C_{y}$$
(6)

ここで、カメラモデルの歪みは radial distortion の第1項のみを考慮している. また、 (C_x, C_y) は画像中心の画像座標、 d_x, d_y は受光素子の水平・垂直方向の中心間距離である.

(2) 画像の一致

ここでは、2枚の画像について幾何学的特性を一致させる手法について述べる.高解

像度カメラと低解像度カメラは焦点を合わせることによりある程度一致した視野を獲得できている.つまり、画像中には3次元空間中のほぼ同じシーンが撮影されているので、3次元空間中にある基準平面を設定すると、その同一平面上に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ことができる.

画像座標(X_f, Y_f)が与えられると,式(6)から世界座標へは式(7)の変換式で変換できる.

 $\begin{bmatrix} x_{w} \\ y_{w} \\ z_{w} \end{bmatrix} = \mathbf{R}^{-1} \begin{bmatrix} \frac{z}{f} \begin{bmatrix} X_{u} \\ Y_{u} \\ f \end{bmatrix} - \mathbf{T} \end{bmatrix}$ (7)

しかし、一般にzが未知であるため、逆投影変換で一意に決めることができない.ここではオブジェクトが一つの平面内にあると仮定する.この平面を基準平面として、2枚の画像を基準平面に投影し基準平面上で画像の一致をはかる.本実験では、まず高解像度側の画像を基準平面に投影し、その投影された画像に一致するように低解像度側のパラメータを最適化する.なお、画像座標を世界座標に変換するには、各カメラの幾何学的なパラメータを推定しておく必要がある.

5-1-6 幾何学的キャリブレーションのアルゴリズム

本節では、各カメラの幾何学的なパラメータである内部パラメータ(焦点距離、画像 中心、レンズ歪み係数)と外部パラメータ(カメラの位置、姿勢)の推定法について述 べる.対応する N 点の世界座標と画像座標からカメラの位置と姿勢を推定する問題は、 いわゆる PnP 問題として知られている.その解法は多数提案および検証がなされており [9]、十分な指標があればこれらの手法によって問題なく解ける場合が多い.ここでは、 カメラモデルのパラメータ推定をその正確さと計算速度により現実的な手法として知 られている Tsai のキャリブレーション手法[9]で行う.

Tsai の手法では,多数の点の世界座標 (x_x , y_x , 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, κ_1 , (C_x , C_y),外部パラメータとし c R と 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.以下,Tsai の手法によるキャリブレーション手順について述べる.

Step 1:

まず,カメラの仕様から受光素子の水平・垂直方向の中心間距離 d_x, d_yを求め,画像 中心(C_x, C_y)の初期値を画像の中央とする.

Step 2:

同一平面上に存在する N 点のキャリブレーション点を撮影した画像から点 P_i の画像 座標 (X_{fi}, Y_{fi}) を求める.

Step 3:

N点の (Xdi, Ydi) を求める.

$$X_{di} = d_x (X_{fi} - C_x)$$

 $Y_{di} = d_x (Y_{fi} - C_x)$
(8)

Step 4:

 (x_{wi}, y_{wi}, z_{wi}) と (X_{di}, Y_{di}) の組から T⁻¹_yr₁, T⁻¹_yr₂, T⁻¹_yT_x, T⁻¹_yr₄, T⁻¹_yr₅ を未知数とす る線形方程式を解く.

Step 5:

 $T^{-1}_{v}r_1$, $T^{-1}_{v}r_2$, $T^{-1}_{v}T_x$, $T^{-1}_{v}r_4$, $T^{-1}_{v}r_5$ から T^2_{v} を求める.

Step 6:

画像中心 (C_x , C_y) から十分離れた点 (X_{fi} , Y_{fi}) をその対応する点の世界座標 (x_{wi} , y_{wi} , z_{wi}) から T_v の符号を決定する.まず, T_v の符号を正として

$$r_{1} = (T_{y}^{-1}r_{1})T_{y}, r_{2} = (T_{y}^{-1}r_{2})T_{y}$$

$$r_{4} = (T_{y}^{-1}r_{4})T_{y}, r_{5} = (T_{y}^{-1}r_{5})T_{y}$$

$$T_{z} = (T_{y}^{-1}T_{x})T_{y}$$

$$x = r_{1}x_{w} + r_{2}y_{w} + T_{x}$$

$$y = r_{4}x_{w} + r_{5}y_{w} + T_{y}$$
(9)

を求め、 $x \ge X$ の符号が等しくかつ $y \ge Y$ の符号が等しいなら T_y の符号を正、それ以外では T_y の符号を負とする.

Step 7:

$$r_1, r_2, r_4, r_5$$
から回転行列Rを決定する.

$$R = \begin{bmatrix} r_1 & r_2 & \sqrt{1 - r_1^2 - r_2^2} \\ r_4 & r_5 & \sqrt{1 - r_4^2 - r_5^2} \\ r_7 & r_8 & r_9 \end{bmatrix}$$
(10)

ただし, s=-sgn($r_1r_4+r_2r_5$)である. r_7 , r_8 , r_9 は R が直交行列であるという性質を用いて求める.

Step 8:

レンズ歪みを無視してfとT_zの初期値を求める.N点のキャリブレーション点を用いてfとT_zを未知数とする線形方程式を解く.

Step 9:

Step8 で求めた f と T_z および κ_1 =0 を初期値として, Marquart 法を用いて非線形最適 化問題を解き, f, T_z , κ_1 を求める.

Step 10:

Step9 までで求めたパラメータ(f, κ , **T**, **R**, C_x, C_y)を初期値として, Marquart 法を 用いて非線形最適化問題を解き, f, κ , **T**, **R**, C_x, C_yを求める.

5-1-7 評価実験

(1) キャリブレーション結果

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11×7=77 個を平面に配置したパターンを用いる.

このキャリブレーションパターンを撮影した画像を図19に示す.ただし撮影に使用 したカメラは BASLER 社の A201bc で,画像サイズは1008×1018 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. 特徴点を検出した結果を図20に示す.



図19 キャリブレーション 図20 特徴点抽出結果パターン

キャリブレーションにより推定されたパラメータを以下に示す.

 $\begin{array}{l} f = 25.\ 692552 \quad [mm] \\ \kappa = 6.\ 180576 \times 10^{-4} [1/mm^2] \\ T_x = -142.\ 769652, \quad T_y = -109.\ 329657, \quad T_z = 822.\ 028049 \quad [mm] \\ R_x = -1.\ 840933, \quad R_y = 17.\ 614088, \quad R_z = -0.\ 435503 \quad [deg] \\ C_x = 547.\ 990521, \quad C_y = 525.\ 191173 \quad [pixels] \end{array}$

(2) レンズの異なる画像の一致

1 台のカメラでレンズを変えて撮影した 2 枚の画像,図21および図22を用意し, それぞれについて特徴点検出,キャリブレーションを行いパラメータを求める.さらに, 求めたパラメータを用いて基準平面に投影したものを図23および図24に示す.

基準平面上でのそれぞれの特徴点の位置を図25に示す.特徴点の誤差の二乗平均の 平方根(RMSE:Root Mean Squared Error)は 0.0640mm,標準偏差 0.0338,最大のズレ 0.174mm であった.これは画像上に再投影したとき,サブピクセルの誤差になり,本手 法により十分な精度で画像の一致が行われたことがわかる.



図21 レンズ1

図22 レンズ2



図23 図21を規準平面に投影 図24 図22を規準平面に投影



図25 特徴点の位置

(3) 解像度の異なる画像の一致

次に解像度の異なるカメラで撮影した2枚の画像,図26および図27を用意し,そ れぞれについて特徴点検出,キャリブレーションを行い,パラメータを求める.さらに, 求めたパラメータを用いて基準平面に投影したものを図28および図29に示す.基準 平面上でのそれぞれの特徴点の位置を図30に示す.特徴点のRMSEは0.153mm,標準 偏差0.118,最大のズレ0.601mmであった.レンズだけを変えた時に比べて特徴点の誤 差が大きくなっているが,画像上に再投影したときにはサブピクセルの誤差になり,本 手法により十分な精度で画像の一致が行われたことがわかる.



図26 1008×1018画素

図27 1280×980画素



図28 図26を規準平面に投影

図29 図27を規準平面に投影



(4) シミュレーション実験

シミュレーション実験を行いモデル誤差の与える影響と観測誤差の与える影響について考察する.シミュレーション用のデータとして世界座標で(10,10,0)から(190,190,0)まで20mm刻みの格子状の100点のデータWを用意する. CCDの素子の大きさは9.0µm×9.0µm,素子の個数は1000×1000とする.カメラモデルFは以下のように設定する.

また、W,Fより生成される画像上での理想的なピクセル位置をDとする.

カメラモデル F に誤差を与えたモデルを F+ Δ とし、W と F+ Δ によって生成される画 像上での位置を D+ Δ 'とする.内部パラメータと外部パラメータを変化させ、100 点の D と D+ Δ 'の誤差を図31から図40に示す.ただし、x 軸 y 軸はそれぞれ世界座標系 における点の位置を示し、z 軸はその点の画像上での誤差を示す.また、パラメータの 真値とのズレと RMSE の関係を図41に示す.図41より、f のパラメータが与える影 響が大きいことがわかる.このため、f については精度良く求める必要がある.また、 ょが与える影響が少ないことより、レンズ歪みを考慮しないという手法も考えられる.



図33 Txを0.1mm 変化させたときの誤差











図36 Rxを0.1[deg]変化させたときの誤差



図39 Cxを0.1画素変化させたときの誤差



図41 モデルを変化させたときに与える影響

次に,観測誤差が与える影響について考察する.実際のカメラで撮影を行う場合,特 徴点の検出に観測誤差や量子化誤差が含まれると考えられる.そこで理想的なピクセル 位置DにガウスノイズをのせたデータD+Nを用意する.WとD+Nを用いてキャリブレー ションを行うことによりカメラパラメータFを推定する.WとFを用いることにより画 像上でのピクセル位置D+Nを生成した.標準偏差1.5程度の観測誤差なら,理想的なピ クセル位置との誤差が0.5pixel以内に押さえられることがわかる.これより特徴点検 出誤差が与える影響は少ないと言える.

本論文では複合センサカメラにおいて幾何学的キャリブレーションを行い2種類の 異なった解像度画像を3次元空間中の同一基準平面に投影し,一致をはかる手法を提案 した.そして,実験により本手法の有効性を確認した.これにより全く同一の視野を持 った2種類の動画像,高解像度・低フレームレート動画像と,低解像度・高フレームレ ート動画像を同時に撮影できるカメラを開発した.今後の課題として,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[10]や,カラーへの対応が挙げられる.

5-1-8 まとめ
高解像度低フレームレート画像と低解像度高フレームレート画像から高解像度高フレームレート動画像を生成することを提案した.具体的には,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生成する.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種類の動画像を撮像するシステムを提案した.

また, 複合センサカメラにおいて幾何学的キャリブレーションを行い2種類の異なった解像度画像を3次元空間中の同一基準平面に投影し,一致をはかる手法を提案した. そして,実験により本手法の有効性を確認した.これにより全く同一の視野を持った2 種類の動画像,高解像度・低フレームレート動画像と,低解像度・高フレームレート動 画像を同時に撮影できるカメラを開発し,これら2つの動画像から高解像度・高フレー ムレート動画像を作成する技術の開発を目指す.

今後は、高解像度低フレームカメラについても同様に画像の取り込み、保存に関する ソフトウェアを作成する予定である.また、プリズムやステージを用いてカメラの光学 的な配置決めを行って同一視点、同一視野を撮像できるようにする予定である.さらに、 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が挙げられる.また、パル スジェネレータからのパルス信号により同期のとれた画像を撮像できるようにし、以上 のことに関する評価を行なっていく予定である.

5-2 リアルタイム高解像度動画像の作成

時空間分解能の異なる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,周波数空間的アプローチとモ ーフィングにより欠落している高解像度画像を生成するアプローチを提案し,アルゴリ ズムの検討を行った.

5-2-1 周波数空間アプローチによる高解像度画像の作成

画像処理によって高解像度の画像を得る単純な方法として補間法(内挿法)がある.代表的な補間法として最近傍法,双一次補間法,3次畳み込み補間法がある[11].これらの手法は処理が簡単であるため計算コストは低いが,ぼけた画像が生成されるため画質は良くない.

信号処理の立場からのアプローチによって高解像度画像を得るための手法として超 解像があり、古くから研究されている.超解像では一般に、同一のシーンから複数の低 解像度画像が得られるという仮定を設ける.そして、これらの低解像度画像の間にサブ ピクセルレベルの微小な変位が存在すれば、超解像による画像復元が可能となる.超解 像の処理は、まず低解像度画像と高解像度画像の間の関係を表す画像劣化過程をモデル 化し、その後モデルに従って複数の低解像度画像から高解像度画像を推定するという流 れとなる.劣化モデルは静止画像に対するものと動画像に対するものに大別される.

静止画像に対する超解像の従来研究は[12]にまとめられている.また,動画像に対す る超解像の研究として[13]などがある.[14]では圧縮された低解像度動画像に対する超 解像の従来研究がまとめられている. MPEG 符号化された動画像を用い,符号領域で高 解像度画像を復元する研究などがある.超解像は複雑な処理を必要とするため,オフラ イン処理が前提となっている.そのため,超解像は主に衛星写真の高解像度化や軍事な どの特殊な分野で利用されている.

一方,動画像のフレーム数を向上させる処理はフレームレート変換を行うことに相当 する.フレームレート変換はフレーム補間とも呼ばれ、テレビ信号規格の相互変換やス ローモーション映像生成などの目的で昔から研究されている.また、ハードウェアによ る変換器も実用化されている[15].フレームレートを向上させるための単純な手法とし て、フレーム反復、時間方向フィルタリングによる線形補間がある.しかし、フレーム 反復により生成された動画像は物体の動きが滑らかでない.また、線形補間によって生 成された動画像は、シーン中に含まれる動物体がぼけてしまう.これらの問題を解決す るために動き補償付きフレームレート変換が提案されており、従来研究として[16]など がある.

(1) 複合カメラセンサの設定

本研究で想定する複合センサカメラは、高解像度・低フレームレートの動画像を撮影できる高解像度 CCD と、低解像度・高フレームレートの動画像を撮影できる低解像度 CCD の 2 つの CCD センサを有する.カメラに入射した光はプリズムによって分割され、高解像度 CCD と低解像度 CCD に到達する.複合センサカメラについて以下の 2 つの仮定を設ける(図 4 2).

1. 複合センサカメラの2つのCCDから得られる画像は,解像度は異なるが同一の視 野をもつとする.厳密には,図43のようなサンプル点の位置関係を仮定する (ただし、図は画像サイズ比2:1の場合).

2. 複合センサカメラは2つのCCDの外部同期を取ることができる. つまり,外部からパルスを入力することにより,時間的なずれのない高解像度画像と低解像度 画像の組が得られるとする. なお,本論文ではこのようなフレームのことをキ ーフレームと呼ぶこととする.

本研究では、以上の仮定の下で、同一の視野をもつ既知の高解像度・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から未知の高解像度・高フレームレートの動画像を生成するアルゴリズムの開発に取り組む.ただし、本研究では 256 階調(8 ビット)のモノクロ動画像に限定する.



(2) 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、この処理では、動画像は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.提案手法はこのことを利用する.提案手法では、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.

なお、本研究では周波数変換として DCT (離散コサイン変換)を用いる.DCT を用いる理 由は、低周波成分への電力集中度が高いために符号化効率が非常に高く、さらに高速変 換アルゴリズムが存在するからである.

(a) 画像生成アルゴリズム

高解像度画像生成処理の流れを図44に示す.ただし、2種類の動画像の画像サイズ比を2:1とする.高解像度画像の生成は次の手順で行う.

- 1. 低解像度画像を用いて,位相相関法[15]により動きベクトルを求める.
- 1で求めた動きベクトルを用いて高解像度画像のフレーム間差分を推定する.
 ただし、動きベクトルが推定できなかった部分は、図39の仮定に基づき、双一 次補間法により低解像度画像から画素値を補間する.
- 3. 高解像度画像とフレーム間差分に対してサイズ16×16のブロックDCTを実行する.また,低解像度画像に対してサイズ8×8のブロックDCTを実行する.
- 4. 高解像度画像とフレーム間差分のそれぞれのDCTスペクトルをDCT領域で足し 合わせることにより、動き補償された高解像度画像のDCTスペクトルを求める.
- 5. 動き補償された高解像度画像のDCTスペクトルに対し、低解像度画像の対応す る位置のDCTスペクトルを合成する.
- 6. 合成したスペクトルに対してIDCT(逆離散コサイン変換)を実行することにより,高解像度画像を生成する.

なお,動画像のインターネット配信を考える場合は,上の1から3までがサーバ側の処理,4から6までがクライアント側の処理となる.



図44 高解像度動画像生成の流れ

(b) フレーム間差分を用いた動き補償

動画像をインターネット配信する場合,動き補償処理をクライアントで行うとクラ イアント側の負担が大きくなる.そこで,フレーム間差分を用いて周波数領域で動き 補償を行う.この原理について述べる.

$$I_{k+1} = I_k + S_{k,k+1} \tag{11}$$

の関係がある.すると,DCTの線形性より

$$C[I_{k+1}] = C[I_k] + C[S_{k,k+1}]$$
(12)

が成り立つ. ここで C[] は DCT スペクトルを表す. すなわち, 動き補償された画像 の DCT スペクトルを得るためには, フレーム間差分の DCT スペクトルを前フレームの DCT スペクトルと足し合わせるだけでよいことになる.

提案手法では位相相関法により動きベクトルを推定し、これを用いて高解像度画像の動き補償を行い、フレーム間差分を推定する.

(c) 位相相関法による動き推定

連続する 2 つのフレーム I_k , I_{k+1} の間で単純な並進移動 (d_x, d_y) が生じていると仮定する. すなわち

$$I_{k}(x,y) = I_{k+1}(x+d_{y},y+d_{y})$$
(13)

が成り立つとする.上式を2次元 Fourier 変換すると

$$F_{k}(u,v) = F_{k+1}(u,v)e^{j2\pi(ud_{x}=vd_{y})}$$
(14)

となる.ここで $F_k = F[I_k]$ とする.すなわち,画像空間領域の並進移動は,Fourier 領域において位相差として現れる.位相相関法はこの性質を利用して動き推定を行う. 連続するフレーム I_k と I_{kt} の間の正規化相互パワースペクトルは次式で定義される.

$$\Phi_{k,k+1}(u,v) = \frac{F_{k+1}(u,v)F_k^*(u,v)}{\left|F_{k+1}(u,v)F_k^*(u,v)\right|}$$
(15)

ここで*は複素共役を表す.式14と式15から

$$\Phi_{k\,k+1}(u,v) = e^{-j2\pi(ud_x + vd_y)} \tag{16}$$

が成り立つ.上式の2次元逆 Fourier 変換は

$$c_{k,k+1}(x,y) = F^{-1}[\Phi_{k,k+1}(u,v)] = \delta(x - d_x, y - d_y)$$
(17)

となる. ここで δ は Dirac のデルタ関数を表す. c_{kk+1} は位相相関関数と呼ばれ,座

標 (d_x, d_y) にインパルスが現れる. c_{kk+1} に現れるインパルスの座標を調べることによ

り,画像中に含まれる並進移動量と方向,すなわち動きベクトルを知ることができる. 図45に位相相関関数の例を示す.なお,方向や移動量の異なる複数の動きが2フレ ーム間に含まれている場合,位相相関関数のインパルスはそれらの動きの方向・移動 量に対応する座標に複数現れる.しかも,位相相関関数*c*₄₄₄の点(x, y)における値は,

ブロック中で(x,y)だけ移動した物体の面積(画素数)を近似する.

位相相関関数のインパルスの座標を調べることにより,画像中に含まれる動きベクトルを知ることができる.しかし,動きベクトルが画像中のどこに対応するかは,位相相関関数からはわからない.そこで,位相相関法による動き推定は図46に示すような処理の流れで画像の各ピクセルに動きベクトルを割り当てる.

以下,図46の各ブロックの処理を説明する.

前処理

まず,画像を24×24 画素のブロックに分割する.次に,24×24 ブロックを縦に2つ, 横に2 つ並べた48×48 ブロックを取る.その後,ブロックの境界で連続になるよう にするため,48×48 ブロックを64×64 ブロックに拡張する.以降の処理はすべて この64×64 ブロックごとに行う.そして,ブロックを拡張した部分に対して,ブロ ックの周囲で白と黒の中間色(mid-grey)になるようなぼかし処理を施す.

FFT

前処理で得られた 64×64 ブロックに対して Fourier 変換を実行する.

正規化相互パワースペクトルの計算

2 つの画像の対応する位置の 64×64 ブロック同士で正規化相互パワースペクトルを 計算する.



図46 位相相関法による動き推定

IFFT

正規化相互パワースペクトルに対して逆 Fourier 変換を実行し, 位相相関関数を求める.

ピーク検出

計算された位相相関関数の値を調べて、ピーク検出を行う.値の大きい方から10点 を、動きベクトルの候補としてリストVにする.本研究では位相相関関数に関するし きい値を T=300 とした.

ブロックマッチング

画像の各画素に対して動きベクトル v を割り当てる.動きベクトルの候補リスト V の 各要素について,動きベクトルを求めたい画素を中心とした 3×3 ブロックによるブ ロックマッチングを行う. V の要素の中で輝度値の絶対値差分が最も小さいものを, その画素の動きベクトルとする.

(3) DCTスペクトル合成

一般的な画像信号は低周波成分を多くもつという性質がある.つまり,画像のもつ情報の大部分は低周波成分に集中しているということが言える.したがって,動き補償を行った高解像度画像の低周波成分に低解像度画像のスペクトルを合成することによって,生成画像を真の高解像度画像に近づけることができる.

今,動き補償された高解像度画像の DCT スペクトル(サイズ $N_h \times N_h$)を $C_h(u, v)$ とし, それに対応する位置の低解像度画像の DCT スペクトル(サイズ $N_1 \times N_1$)を $C_1(u, v)$ とする (ただし $N_h = \Sigma N_1$). このとき,合成された DCT スペクトル C(u, v)を

$$C(u,v) = \begin{cases} w_h(u,v)C_h(u,v) + \frac{N_h}{N_l}w_l(u,v)C_l(u,v) & \text{if } 0 \le u, v \le N_1 \\ C_h(u,v) & \text{otherwise} \end{cases}$$
(18)

により求める(サイズ $N_h \times N_h$). サイズの異なるスペクトルを合成するため,低解像度画像のスペクトル C_1 にエネルギー補正係数 N_h / N_1 をかける.また, w_h , w_1 はスペクトル合成の重みである.

(4) 評価実験

本節では画像生成の実験方法とその結果を示す.最初に,画像評価の指標であるピーク信号対雑音比の定義を与える.

(a) ピーク信号対雑音比の定義

画像サイズが M×N,フレーム数が K である 2 つの映像信号 $I_k(x, y) \ge \hat{I}_k(x, y)$ の間の 平均二乗誤差 (MSE) は次式で定義される.

$$MSE = \sigma_e^2 = \frac{1}{KMN} \sum_{k=1}^{K} \sum_{x=1}^{M} \sum_{y=1}^{N} [I_k(x,y) - \hat{I}_k(x,y)]^2$$
(19)

ピーク信号対雑音比(PSNR)は平均二乗誤差を用いて次式のように定義される.単位は デシベル(dB)である.

$$PSNR = 10\log_{10}\frac{I_{max}^2}{\sigma_e^2}$$
(20)

ここで I_{max} は信号強度の最大値で、一般的な8ビット信号の場合は I_{max} =255である.

(b) 実験結果

複合センサカメラから得られる 2 種類の動画像をシミュレートするため、MPEG テストシーケンス (30[~][fps])を用いて高解像度画像生成実験を行った.2 種類の動画像の画素位置に関する仮定に基づき、サイズ M×N のテスト画像 $\tilde{I}(x,y)$ に対して次式に

基づいて 50%の画像縮小を行い(σ =2), サイズ $\frac{M}{2} \times \frac{N}{2}$ の低解像度画像 I(x, y)を生成 する. そして, 原画像を 5 [fps]の高解像度画像, 縮小した画像を 30 [fps]の低解像 度画像と想定して(R=6)画像生成を行った. すなわち, 高解像度画像(原画像)の 1 番 日, 7 番目, 13 番目, …のフレームのみと, 低解像度画像(縮小した画像)のすべてを 使用して高解像度画像を生成した. 本研究の実験に使用した MPEG テストシーケンス を表4に示す.



(a) 原画像



(b) 低解像度画像を2倍拡大



(c)生成画像 図47 Flower garden フレームNo. 29

シーケンス名称	画像サイズ	使用フレーム	
Flower garden	352 × 240	No. 1 ~ No. 115	
Football	352 × 240	No. 1 ~ No. 121	
Foreman	352 × 288	No. 1 ~ No. 295	
Hall monitor	352 × 288	No. 1 ~ No. 295	

表4 使用したMPEGテストシーケンス



(a) 原画像



(b) 低解像度画像を2倍拡大



(c)生成画像 図48"Football"フレーム No. 57

図47は"Flower garden"の29番目のフレームの生成結果である.また,

"Football"の57番目のフレームの生成結果を図48に示す. "Flower garden"の フレームごとのピーク信号対雑音比の変化を図49に示す.ただし、キーフレームは 原画像と同一であるため、ピーク信号対雑音比が計算できない.このため、図49に おいて、キーフレームについては生成画像のピーク信号対雑音比(実線)を描画してい ない.つまり、実線の途切れている部分がキーフレームである.また、2倍拡大した 低解像度画像のピーク信号対雑音比を破線で示す.各テストシーケンス全体のピーク 信号対雑音比を表5に示す.表5において「キーフレームあり」は、動画像の総フレ ーム数にキーフレームを含めて評価する場合の数値である.表には、拡大した低解像 度画像(LRI)のピーク信号対雑音比も併記してある.



図 4 9 "Flower garden"のピーク信号対雑音比

シーケンス	ピーク信号対雑音比 [dB]		
名称	キーフレームあり	キーフレーム無し	LRIを2倍拡大
Flower garden	21.37	20. 54	18. 82
Football	25. 25	24. 42	23. 38
Foreman	32. 19	31. 38	27.87
Hall monitor	35. 39	34. 58	25. 19

表5 生成画像のピーク信号対雑音比

図49から分かるように、生成画像のピーク信号対雑音比のグラフがV字型となっている.これは、フレーム間差分を推定する際に、生成済みの高解像度画像から画素値を推定しているからである.このため、生成画像が時間的にキーフレームから離れるほど誤差が蓄積する.本研究のシミュレーション実験では2種類の動画像のフレームレート比を1:6としたが、この条件が厳しくなった場合、キーフレームの中間フレームの画質がさらに落ちると考えられる.

なお,全てのテストシーケンスについて,生成画像のピーク信号対雑音比が,拡大

した低解像度画像よりも大きくなった.したがって,提案手法により真の高解像度画 像に近い画像が得られたということができる.また,テストシーケンス "Flower garden"と "Football" ではピーク信号対雑音比の値が 25 [dB]以下と,他の 2 つ よりも結果が悪かった.これは動きが激しいため,動きベクトルの推定に失敗し,低 解像度画像から補間を行った画素が,生成画像中に多く含まれていたためであると考 えられる.より精度の高い動き推定法については今後の課題とする.

(c) スペクトル合成の有効性

提案手法におけるスペクトル合成の有効性を検証するため、テストシーケンス "Foreman"を用いて、スペクトル合成処理を行わずに画像を生成した. "Foreman" の82番目のフレームの生成結果を図50に示す.

スペクトル合成を行わなかった場合,シーケンス全体のピーク信号対雑音比は 29.92 [dB]となった(ただしキーフレームを評価に含めない).これを表5の結果と 比較すると,スペクトル合成によって+1.46 [dB]のゲインが得られていることになる. したがって,スペクトル合成により,生成画像が真の高解像度画像に近づいていると いうことができる.このため,提案手法によって得られる生成画像の画質は動き推定 の性能に大きく依存するということがわかる.



(a) 生成画像(合成あり)



(b) 生成画像(合成無し)図50 "Foreman" フレーム No. 82

(d) 画像の位置合わせの正確性と画質の関係

高解像度画像と低解像度画像間の位置関係の仮定が成立しない場合を考える.ここでは、高解像度画像の(0.5,0)画素分ずれた場合、高解像度画像の(0.5,0.5)画素分ずれた場合のそれぞれについて、テストシーケンス"Foreman"を用いて画像を生成し、 ピーク信号対雑音比を測定した.画素がずれた場合のシーケンス全体のピーク信号対 雑音比は、ずれ量が(0.5,0)の場合は29.85 [dB]、(0.5,0.5)の場合は28.41 [dB]と なった(ただし、いずれもキーフレームを評価に含めない).これらは表5の結果の 31.38 [dB]に比べて小さい.したがって、生成画像が真の高解像度画像を近似するた めには、2種類の画像の相対的位置関係が仮定に近いほどよいということが言える.

今後の課題として、カラー画像への対応、処理の高速化および生成画像の高画質化 が挙げられる.提案手法のスペクトル合成が高画質化に寄与していることが明らかに なったので、生成画像の画質は動き推定の結果に依るといえる.今後、より精度の高 い動き推定を実現することを目指す.

5-2-2 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、それぞれの優れた特徴から高品質で高解像度な高フレームレート動画像の獲得手法を提案する.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[17-19]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(図51).



(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.本手法では,両者の長所を生かして,高フレームレート動画像からモーション情報を高解像度動画像からはテクスチャ情報を抽出して,モーフィングを用いて高解像度高フレームレート動画像の生成を行う.高解像度高フレーム動画像生成のフロー図を図52に示す.

まず,低解像度高フレームレート動画像を用いて,動領域を検出する(図52(1)). フレーム間差分を用いて動領域を検出し,動領域と静止領域を二値画像で表現したマ スク画像を作成する.1枚のマスク画像を作成するのに,隣接した3枚の低解像度フ レームを使用する.作成したマスク画像と高解像度低フレームレート動画像を使用し て,背景画像を生成する(図52(2)).背景画像1枚を生成するのに,隣接した3 枚の高解像度フレーム使用する.マスク画像より得た高解像度フレームの静止領域が 背景画像のテクスチャとなる.背景画像が得られると,中間フレーム補間における静 止領域処理は完了である.

動領域マスクと低解像度高フレームレート動画像を用いて,特徴点の抽出・追跡を 行う(図52(3)).低解像度高フレームレート動画像の動領域において,ハリスオ ペレータ出力が局所最大となる画素を特徴点として抽出し,この特徴点の近傍を探索 することにより,低解像度高フレーム動画像のフレーム中の特徴点を追跡する.この 追跡結果より点の各フレームにおけるモーションベクトルを求める.また,特徴点抽 出と同様に検出した動領域に対してのみ追跡を行う.これにより特徴点の動領域誤追 跡の低減や計算コストの削減を行うことができる.

低解像度高フレームレートにおける特徴点の位置を高解像度低フレーム動画像に 反映することで、フレーム間の対応点をとることができる.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の同期がとれているキーフレームでは、 同じシーンが撮像されているため、高解像度画像での特徴点対応づけは容易に行える. この対応点を用いて動領域を三角形のパッチに分割する(図52(4)). 三角パッチ 分割にはドロネー分割を用いる.これにより、高解像度フレーム間で対応する三角パ ッチが得られる.





対応する三角パッチと特徴点の追跡結果を使用して、中間フレームをモーフィング 補間する(図52(5)).モーフィングは、クロスディゾルブ法とワーピング法の2 つがある.本手法は、画像をパッチに分け、パッチ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採用する.モーフィングにより生成した中間フレーム の動領域を、背景画像と合成することで中間フレームが完成する(図52(6)).こ れにより、高解像度高フレームレート動画像が生成できる.

(b) 動領域検出と背景画像

動画像は、動領域と静止領域に区別できる.動領域は、移動物体の存在や光の影響 により、画像輝度値が変化している領域である.静止領域は、画像輝度値の変化がな い領域で背景を指す.静止領域は、動画像において複数フレームで共通して使用でき るので、動領域と分けて処理した方が、計算時の負担をメモリ的にも時間的にも削減 できる.ここでは、動領域と静止領域を分けた二値画像であるマスク画像の作成手法 を説明する.マスク画像作成過程を図53に示す.

1枚のマスク画像を生成するには連続する3枚のフレームが必要である.動画像に おいて,動領域・動物体は隣接フレーム間で輝度値の差として現れる.動領域を検出 したい注目フレームkに対し、その前後のフレーム(隣接過去フレームk-1、隣接未 来フレーム k+1) それぞれと差分をとり(図53(1)),任意のしきい値より二値化 する(図53(2)).この際、輝度値の差をとるため、動領域内にある画素でもフレー ム間で差がないものが存在する. そのため, 単純にフレーム間差分をとった二値画像 は動領域・動物体の輪郭部分が現れるだけである. それに対処するため, 二値画像の 動領域を横方向,縦方向それぞれで連結する(図53(3)). しかし,依然として同 一の動領域にあっても、横方向で連結した画像は縦方向に分断が残っている.また、 縦方向連結についても同様である.よって,それぞれ動領域に対し拡大法を行い,同 一の動領域をつなぎ合わせる(図53(4)).ここまでの処理で横方向連結画像は横 に並んでいる別の動領域・動物体、縦方向連結画像は縦方向に並んでいる別の動領 域・動物体が一つの領域に統合されている.ここで,横方向連結画像と縦方向連結画 像の AND をとれば, 異なる動領域が別々の領域として現れる (図53 (5)). この AND 画像は,注目フレームkと隣接過去フレームk-1から生成したものと,注目フレーム kと隣接未来フレーム k+1 から生成したものの2枚がある. さらにこの2枚の画像の AND をとれば,動領域と静止領域が区分できたマスク画像が生成できる(図53(6)). この動領域は、特徴点抽出・追跡、三角パッチ分割、モーフィングなどの処理に利用 し,静止領域は,背景画像の生成に利用する.

背景画像を生成する際,注目フレーム k とタイミングが同じ高解像度フレームのテ クスチャ情報を主に使用することになる.その前に,各フレームに対応するマスク画 像は低解像度フレームよりとっているため,高解像度フレームとは空間的サンプリン グが異なるという問題に対処しなければならない.まず,マスク画像を背景画像に合 わせて空間サンプリングを拡大する(54(1)).その拡大マスク画像から高解像度 フレームの静止領域が分かるので,その静止領域は背景画像にコピーされる.これに より使用時間に背景となるほとんどの領域が得られたことになる.さらに前後の高解 像度フレーム k-1,フレーム k+1 の静止領域情報を使用して,まだ背景画像の情報が ない画素を埋めること(図54(2))で,背景画像が生成できる.







図54 背景画像の生成

(c) 特徴点抽出と追跡

低解像度高フレームレート動画像で抽出した特徴点は、中間画像を補間するために 追跡しなければならない(図55).よって、追跡するために有効な特徴点を抽出す る必要がある.特徴点追跡で単純な手法であるテンプレートマッチングは、テンプレ ート内の輝度値を照合する手法であり、特徴点が線形的な移動をする場合にしか使用 できない.特徴点周りのテクスチャが回転しただけでも追跡が破綻してしまう.一方、 ハリスオペレータは、フレームの輝度勾配により算出した応答関数の値が局所最大と なる画素を検出する.このため、周りのテクスチャの回転や輝度値変化が生じても、 応答関数値が局所最大に保たれていれば、対応点として追跡が可能となる.したがっ て、特徴点抽出には、ハリスオペレータ[20]を使用する.

ハリスオペレータ処理の際,隣接特徴点との間隔・特徴点数の上限をパラメータとして与えており,追跡精度を確認しながら特徴点抽出を調節できるようにした.

ここで、ハリスオペレータについて簡単に説明しておく.まず、式21に示すように、画像の輝度値勾配より求めた行列Aとする.

$$A = \sum_{(x,y)\in N_0} \nabla I(x,y) \nabla I(x,y)^T$$
(21)

 N_0 : neighborhood of a pixel (x_0, y_0)

$$\nabla I(x,y) = \left(\frac{\partial I(x,y)}{\partial x}, \frac{\partial I(x,y)}{\partial y}\right)^T$$

I(x,y) : pixel value (intensity)

行列Aより応答関数Rを求める.

```
R = \det A - \kappa^* trace^2 A
det A
A(1,1)* A(2,2) - A(1,2)<sup>2</sup>
traceA : A(1,1) + A(2,2)
```

式22に示すようにRの値が局所最大の画素が特徴点として抽出される.抽出した 特徴点は、高解像度低フレームレート動画像のフレーム間隔時間だけ追跡する.追跡 には、フレーム間で対応点が近傍画素に存在するという拘束条件とハリスオペレータ 出力値を使用する.前者については、低解像度高フレームレート動画像は、隣接フレ ーム間での動領域の移動が小さいため、対応点は近傍画素に存在するという仮定が成 り立つからである.後者については、特徴点抽出にハリスオペレータを使用したので、 追跡フレームの特徴点の対応が精度よく得られるからである.

(22)

まず,フレームkの特徴点座標をフレームk+1にとる.次に,その近傍画素をサー チェリアとし,特徴点を探索する(図56).サーチェリア内でハリスオペレータ出力 が最も大きな画素を対応する特徴点とする.この処理を繰り返し行い,特徴点を追跡 する.



図55 特徴点抽出·追跡



図56 特徴点追跡

(d) 3角パッチ分割

動領域テクスチャのモーフィングにはワーピング法を使用するので,動領域をパッ チ分割する必要がある.パッチの形状は,四角形と三角形が考えられる.フレーム全 体ではなく,形状が不規則な動領域を処理する場合は,三角の方が柔軟に対処できる ので,ドロネー分割によって動領域を三角パッチに分割する.

特徴点追跡結果を同時刻の高解像度低フレーム動画像に移すことで,高解像度フレーム間で対応点がとれる.この対応点が三角パッチの頂点となる(図57).

ドロネー分割は、特徴点の座標情報のみを使用する.以下では、ドロネー分割の処 理手順について説明する.ドロネー分割には(図58(1))の7つの特徴点を使用す る.計算に使用するのは特徴点の座標情報のみ.分割したい全ての特徴点を囲む十分 大きな三角形を仮に置く(図58(2)).その大きな三角形の中に特徴点を追加してい き、三角形を分割していく.特徴点が三角形内部にあるときはその三角形が3つの三 角形に分割される(図58(3)).追加する特徴点が三角形の辺上にあるときは2つ の三角形がそれぞれ2つの三角形に分割、つまり5つの三角形になる(図58(4)). (3)または(4)の操作を繰り返しながら全ての特徴点を追加する.最後に(2)で仮にお いた3つの点とそれを頂点に持つ三角形を削除する.

分割した三角パッチ単位でモーフィング処理を行うため、テクスチャがつぶれないように、できるだけ正三角形に近い三角形に分割したい.しかし、三角パッチ分割の 過程において図59左フレームのように平たく薄い三角形(△ABD)が現れる場合が ある.図59のケースでは、△ABDと△CBDの共有辺BDをACに換えることで共に正 三角形に近い△ABCと△ADCに修正できる.



図57 3角パッチ分割



図58 ドロネー分割の手順



図59 三角形の修正

(e) モーフィング

モーフィングは,対象領域をパッチ分割することで色と形の両方を補間できるワー ピング法により行う.よって,図60のように三角パッチ単位の処理になる.

特徴点追跡により得られた目的パッチ内の画素 X の対応画素 X', X', を開始パ ッチ1,開始パッチ2からそれぞれ探し出す(step 1).そして,X',X',の輝度 値から X の輝度値を計算する(step 2).図60を例にすると,低解像度高フレーム 6枚に対し高解像度フレーム1枚なので,中間フレームを5枚補間する.中間フレー ムは,特徴点情報を得た低解像度高フレームに同期して生成される.また,特徴点追 跡結果から頂点座標が決まる目的パッチのテクスチャには開始パッチ1と開始パッ チ2のテクスチャを用いる.



図61 パッチ間の対応付け

以下, step1, step2 についてそれぞれ説明する.

(step 1)

開始フレームと目的フレームの画素の対応づけは、基本的なモーフィング手法である複数対応線を用いて行った.まず、図61左のように三角パッチの3つの辺と注目 画素 X との位置関係を v_{i}, u_{i} (*i* = 1,2,3)で表す.

$$u_{i} = \frac{(X-P) \cdot (Q-P)}{\left\|Q-P\right\|^{2}}$$

$$v_{i} = \frac{(X-P) \cdot Perpendicular(Q-P)}{\left\|Q-P\right\|} \qquad (i = 1, 2, 3)$$
(23)

 u_i は線分 P_iQ_i とそれに画素 X から下ろした垂線の交点の位置を表し、 v_i は線分 P_iQ_i と画素 X との距離を表す.また、関数 Perpendicular()は、ベクトルを引数として与えると、その線分に大きさが等しく垂直なベクトルを返す.

この手法では、座標変換において各対応線分の重みづけを行う.各対応線分においてXに対応する位置X'を計算する.

$$X_{i} = P_{i}' + u_{i} \cdot (Q_{i}' - P_{i}') + \frac{v \cdot Perpendicular(Q_{i}' - P_{i}')}{\|Q_{i}' - P_{i}'\|} \qquad (i = 1, 2, 3)$$
(24)

この計算により図61において画素 X に対応する画素は $\Delta X_1'X_2'X_3'$ の内部にまで 絞られる.この中から対応する1つの画素を選ぶために、 X_i' と X の差を $D_i = X_i' - X$ として、3つ線分について求めた D_i より重みづけ平均差を計算する.その重みは各 線分と X との距離によって決まる.

 $weight = \left[\frac{length^{p}}{(a+dist)}\right]^{b}$ (25)

変数 a は十分大きな値を持つか、線分から画素までの距離(dist)が0 ならば重み は無限大になる. a に適切な値を与えることで線分上の画素を思ったところへ正確に 移動できる. 変数 b によって各線分の相対的な重みを調整できる. b が大きければ全 ての画素が最も近い線分からの影響のみを受ける. また b が0 ならば全ての線分に等 しく影響を受ける. b については $0.5 \le b \le 2$ という範囲の値が最も効果的である. 変 数 p は $0 \le p \le 1$ という範囲で、もし値が0 ならば全ての線分が同じ重みを持ち、1 な らば長い線分が大きな重みを持つ.

重みづけ平均差をXに加えることで目的フレームにおけるXの対応点X'が決まる. 複数対応ベクトルのアルゴリズムを下に示す.

For each X in the destination image DSUM =(0,0) weightsum=0 For each line P_iQ_i calculate u,v based on P_iQ_i calculate X_i' based on $P_i'Q_i'$ calculate displacement $D_i = X_i' - X$ for this line dist = shortest distance from X to P_iQ_i

$$weight = \left\lfloor \frac{length^{p}}{(a+dist)} \right\rfloor^{p}$$

$$DSUM += D_{i}*weight$$

$$weightsum += weight$$

$$X' = X+DSUM/weightsum$$

$$distinationImage(X)=sourceImage(X')$$

(step 2)

基本的に画素 X'と X''の輝度値は近い値をとるため、下の式のように内分値を 線形的に求めれば良い.

$$I(X) = \frac{1}{n} ((n-i)*I(X') + i*I(X')) \qquad (i = 1, ..., n-1)$$
(26)

5枚のフレームを補間する場合は n=6 として計算を行う.ただし,照明条件の変化 やオクルージョンにより輝度値が急激に変化する画素がある.この画素を考慮すると, 輝度値計算では低解像度フレームの輝度値を参考にする必要が生じる.

(2) 評価実験

今回は,各段階の処理の動作確認を行った.使用動画像は,720×428 画素,30fps である.

(a) 動領域検出と背景画像の生成

動領域検出結果(図63)と背景画像の生成結果(図64)を以下に示す.それぞれの使用フレーム(図62,63)中のフレームkに対応する動領域と背景画像である.

背景画像(図64)は背景テクスチャのみで生成できているのに対し,動領域検出 (図63)では動物体の周りの背景テクスチャが含まれている.3枚のフレームを使 用してマスク画像を作成する場合,動領域と静止領域を正確に切り分けることはでき ないため,どちらかに領域を広くとる必要がある.



図62 動領域検出使用フレーム



図63 動領域検出の結果



図64 背景画像の生成結果

多くのフレームを使用することで、より正確な動領域検出を行うこともできるが、計 算量が増大してしまう.今回は、動領域を広くとることで、動領域が静止領域に含ま れることを防いだ.5-2-2章(1)-(b)節で説明した拡大法は同一動領域の統合の他に、 動領域が静止領域にはみ出ることを防ぐ処理も目的としている.よって、動領域とし て検出されている背景テクスチャも動領域と同様の処理が行われる.使用する画像に よっては、背景テクスチャを動領域と同じように処理すると、画質の劣化を引き起こ す.今回の手法では、背景画像に動領域のテクスチャが現れていないため、高解像度 高フレーム動画像の生成に適用可能である.

なお,背景画像の白い領域では,使用した3枚のフレームのマスク画像が常に動領 域であるため,背景情報は存在しない.この領域については,この背景画像を使用し ている間,常にモーフィング処理された動領域が合成されることになるので問題ない. つまり,高フレーム高解像度動画像を生成して段階で情報が欠落した画素が存在する ことはない.

(b) 特徵点抽出·追跡

図65は、ハリスオペレータによる特徴点抽出結果である.図65の手前の自動車 やモノレールに顕著に現れているように、テクスチャが鮮明である動領域や背景が青 空などの単純なテクスチャである場合には、背景と対象の境界の特徴点が抽出できる. このことは、この後のモーフィング処理においても非常に有効である.

図66では,前の自動車のボンネットで反射している光の周りに特徴点が集まって

いる. ハリスオペレータは輝度値勾配により特徴点を抽出する. 光の反射を示してい る画素では周りの画素より輝度値が高くなるため,高い輝度値勾配が生じる. これに より自動車のコーナーを示す特徴点が抽出できなくなる. また,道路の横断歩道に特 徴点が抽出されている. 自動車の移動とともにその影が常に動いているからである. 屋外で撮影を行う際,影が動領域として検出されるのは当然であり,自動車本体と同 様に処理しなければならない.



図65 特徴点抽出









図67 5フレーム間追跡使用動画像



図68 5フレーム追跡結果

図67に示す5フレームの入力画像を用いて,特徴点追跡を行った結果を図68に 示す.ここでは,特徴点間距離を20画素以上とり,5フレーム間追跡を行った.図 66での考察で述べたように,自動車のボンネットが光で反射した時点で追跡が破綻 している.本手法では,抽出している近傍にハリスオペレータ出力値の近い画素が存 在する場合,隣接フレームでは出力値が逆転することがある.そして,サーチエリア 内に出力値の高い別の特徴点が入ってくる可能性があり,特徴点間隔を広く設定しな ければならない.以上述べてきたハリスオペレータ欠点を考慮して,その他の拘束条 件を加える必要がある.

(c) 3角パッチ分割

図68で求められた特徴点を用いて三角パッチ分割を行った結果を図69に示す. 図68において,手前の自動車やモノレールなど特徴点抽出で良い結果が得られた動 領域については,ドロネー分割でも特に輪郭がはっきりととらえられている.これに より本提案手法において,特徴点抽出・追跡の精度が大きな影響を持っていることが 分かる.

ドロネー分割では、5-2-2 章(1)-(d)節で説明したように大きな三角形から分割を 開始する.そのため特徴点情報のみを使用したドロネー分割では、動領域が凸形状に なる.すると、図69右図のような背景のみの三角形が生じる.これは前節図65の 影とは異なり本来ならモーフィングの必要のない領域である.現時点では、このパッ チもモーフィング処理し,背景画像と合成するといった動領域と同じ処理を行うこと にしている.もし支障が生じた場合は,マスク画像の利用など静止領域のパッチを削 除する方法について検討する必要がある.





図69 ドロネー分割による動領域の三角パッチ分

(d) モーフィング

モーフィング実験では、720×428 画素、30fps の動画像から連続する7枚のフレームを抜き出し、最初と最後の2枚のフレームを使用フレーム(図70)として中間フレームを補間した.なお、2枚の開始フレームの対応点は手入力により与えた座標であり、目的画像の座標は開始フレーム間の対応点から内分点を求めたものである.

図71は、動領域の一部について、図70の使用フレームをモーフィングにより補間した結果を示している.2種類の結果は異なるパッチ数でモーフィングしたもので、 パッチ数によるモーフィングの滑らかさを比較した.また、これらと図70のフレームk+1、フレームk+3、フレームk+5のそれぞれを比較した.

パッチ数の多い方は、動領域のコーナーを密に、少ない方はその半分の割合でとっ たものである.図71 a)と d),g),および c)と f),i)を見て分かるように、使用 フレームから時間的に近いフレームではほぼ実フレームのテクスチャを表現できて いる.なお、実フレーム a)と b)の光の反射が補間フレームに現れていないのは、使 用フレームに光の反射情報がないからであり、当然の結果と言える.実画像両方の補 間フレームにおいて、使用フレームから時間的に遠いフレームである f)と h)ではテ クスチャの歪みが確認できた.ただし、パッチ数を多くとった h)は歪みが小さく、 1フレームあたり使用時間が短い高フレームレート動画像として扱うには十分であ る.一方、モーフィング処理の対象画素の対応点が2つの使用フレーム内で一致して ないため、パッチ数の少ない方 f)では、テクスチャにタイヤが二重に映り込んでい



図70 使用フレーム



る.

a) フレームk+1



b) フレームk+3 実フレーム



c) フレームk+5



d) フレームk+1



- - f) フレームk+5
- 補間フレーム (パッチ数:7 特徴点数:9)



レームK+1 n) フレームK+3 補間フレーム(パッチ数:22 特徴点:18) 図71 モーフィング実験結果 以上より,モーフィングには動領域のコーナーを密にとった方が良いことは明らか である.また,実フレームから遠い補間フレームで生じる歪みは,対応点のズレや重 みづけパラメータ設定が問題として挙げられる.対応点に関しては精度の良い特徴点 追跡が,パラメータ設定に関しては微調整が必要である.

5-2-3 まとめ

本研究では,高解像度 CCD と低解像度 CCD をもつ複合センサカメラを考え,このカメ ラから得られる同一視野の2種類の動画像から,高解像度・高フレームレートの動画像 を画像処理によって生成する2つのアルゴリズムを提案した.

そして、周波数を利用した手法では、MPEG テストシーケンスを用いたシミュレーション実験と生成画像の数値的評価を行い、2 種類の画像のサイズ比が 2:1、フレームレート比が 1:6 の場合には、真の高解像度画像に近い画像が提案手法によって得られることを確認した.周波数を利用した手法の今後の課題として、カラー画像への対応、処理の高速化および生成画像の高画質化が挙げられる.提案手法のスペクトル合成が高画質化に寄与していることが明らかになったので、生成画像の画質は動き推定の結果に依るといえる.今後、より精度の高い動き推定を実現することを目指す.

一方、モーフィングを利用した手法では、処理の流れを4段階に分けて、各段階にお ける画像処理の実装を行った.特徴点追跡では、拘束条件が近傍画素とハリスオペレー タだけでは、精度の良い追跡ができなかった.特徴点追跡については、文献[21]に詳し い. 今後、従来提案されている特徴点抽出手法である、IPAN tracker[22]、KLT tracker[23]、GOA tracker[24]などを比較検証する.その結果より、本応用に適切な 特徴点追跡手法を開発して追跡の安定化を計る予定である.モーフィングでは、動領 域・動物体においてコーナーなどのテクスチャの変化が大きなポイントの数だけ特徴点 を抽出した方が、質の高い中間フレームが生成できることが分かった.今後の課題は、 より精度の高い特徴点抽出・追跡を行うための拘束条件を検討および実装を行うことで ある.そして、中間画像を自動的に生成するアプリケーションの完成を目指す.

5-3 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.従来か ら、全方位視覚センサは、その広視野角を活かした監視やイメージメディア、ヴァーチ ャルリアリティなどの応用に用いられてきた.しかしながら、一般のビデオカメラの解 像度は制限されており、全方位ミラーと組み合わせて利用するには、解像度が低くこの 様な人に呈示する目的においての問題点となってきた.本提案の高解像度で高フレーム レートのビデオセンサを用いることにより、従来の全方位視覚センサの低解像度の問題 を解決することができ、全方位視覚センサの応用が広がるものと期待される.

従来より,全方位 360 度の広視野角を撮像するセンサが数多く提案されてきた[25]. これらの中で,凸面ミラーとビデオカメラを組み合わせたものは,全方位を実時間で計 測できる特徴を持つ.これら凸面ミラーを用いたセンサには,円錐面[26],球面[27], 双曲面[28],放物面[29]など,様々なミラー形状が提案されてきた.特に,双曲面と放 物面ミラーを用いた全方位視覚光学系は,さらにミラー局面中の焦点を仮想視点とした 無歪みの仮想視野画像に容易に変換できる特徴をもつため,人に呈示する目的に利用で 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.試作した全方位ミラーは,これらカメラ を想定して光学設計を行った.

5-3-1 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72に示すように、鉛直下向きの双曲面鏡と 上向きのカメラから構成される.双曲面鏡は(26)、(27)式のように表され、二つの焦 点(0;0;c)、(0;0;c)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(0;0;c)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(図74-(a)).また、図72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.すなわち、画像上での任意の写像点 p(x;y)と俯 角、方位角の関係は、(3)(5)式から一意に決まる.この式から、全方位入力画像(図 74-(a))を図73に示すようにミラー焦点を視点とした透視変換画像(図74-(b)) やパノラマ画像(図74-(c))に容易に変換できるという特性を持つ.

$$\frac{X^2 - Y^2}{a^2} - \frac{Z^2}{b^2} = -1, Z > 0$$
(26)

$$c = \sqrt{a^2 + b^2} \tag{27}$$

$$\beta = \tan^{-1}(y/x) \tag{28}$$

$$\alpha = \tan^{-1} \frac{(b^2 + c^2) \sin \gamma - 2bc}{(b^2 + c^2) \cos \gamma}$$
(29)

$$\gamma = \tan^{-1} \frac{f}{\sqrt{x^2 + y^2}}$$
(30)







図73 変換画像の幾何学的関係



a:全方位入力画像 b:透視投影画像

c: パノラマ画像 図74 全方位画像の画像変換

5-3-2 全方位ミラーの設計

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.今回はカメラの視野角 43 度で,上方垂直視野角 25 度,下方視野角 63 度に設定した.表6に式26に示される双曲面形状バラメータを示す.また,図75 に設計したミラーの概形を示す.

次に試作ミラー形状を用いて集光の評価を行った.この評価において,光学シュミレ ーションソフト ZEMAX (Focus software inc.)を用いた.図76にシミュレーションに より得られたスポットダイアグラムを示す.図中では-60,-40,-20,0,20の5つの俯角 におけるスポットダイアグラムを示している.スポットグラムから,コマ収差や色収差, 非点収差などが起こっておらず,良好な結像を示していることが分かる.また,図77 に各俯角に対するエンサークルドエネルギーを示す.図77から,俯角20度以外は, 70%以上のエネルギーが,10µm以内に収まっており,CCDセルサイズに対して十分な集 光が得られていることが示されている.この結果,試作した全方位ミラーは,幾何的に も光学的にも提案する複合視覚センサのスペックを十分満たすものであると考えられ る.今後は,実際の機器を用いて撮像を行い,画質や解像度を評価していく予定である.

	双曲面パラメータ[mm]	
а	52.0802	
b	66.3876	
С	84.3780	

表6 試作全方位ミラーの双曲面形状パラメータ





2.	• 0.4993
DB1 + 8.00, -50.00 B15 (B2 + 0.00, -10	1866.5 · 2.5581
807: 0.06, -58.80 000	
374 (5.000, 3.448 Ht 398 (6.680, 3	1.791.FM
•	
2521 e.465, e.465 060 2531 e.465, 38	A 26 115
280 8.400, 9.207 (H)	01
•	
A STOCK	p.
TH: 0.080, 6.853 HI 2H: 0.000, 0	0.72% HW
SPOT DIAGRAM	
UBIS HAS NO TITLE THE HAY 30 2001 LATTS HAE MICHONS.	
BELEVIELE : 12:27 7:46 1:51 1:21 1:21 504 6 999 : 45 999 CON	HYPER.ZMX FICURATION 1 OF 1

図76 スポットダイアグラム



図77 エンサークルドエネルギー



図78:全方位ミラーの支持具形状

また,従来は、ミラーを支持するのに、図78-aに示すような透明なドーム型支持 具を用いていた.このドーム型支持具は強度が高いが、透明ドームを高い面精度で球形 状に削り出しにより製作しなくてはならないため、コストがかかる.また、光線がドー ムを通り抜けるとき屈折が起こるため、画像に歪みが生じる.これらの問題を解決する 低コストで光学的歪の影響が無い支持具の構造として、図78-bに示すようなミラー とカメラレンズの間を金属棒により支えるミラー取り付け支持具を設計し試作を行っ た.試作した支持具により、十分な強度を持つ構造であることを確認し、従来の取り付 け支持具の意図しない屈折による歪みなどの光学的な問題やドーム曲面の切削、研磨と いう製作上の問題点を改善した.

5-3-3 まとめ

試作した複合センサに対応した全方位カメラ用光学系を設計し,試作を行った.試作 した全方位ミラーは,複合センサの CCD サイズを想定して光学設計を行った.また,従 来のドーム型支持具の問題点であった,屈折や加工コストを解決する棒型支持具を設計 試作し、十分に実用に耐える強度を持つことを確認した.

今後は、試作したミラーとプロトタイプの複合センサカメラを用いて、ミラーやセン サ光学系の評価を行い、高解像度化による問題点を検証していく予定である.双曲面鏡 による全方位ミラーは、カメラ幾何的に単一視点を持つという長所を持つが、画像の分 解能が仰角に応じて変化するという欠点を持つ.これを改善するために、仰角に対して 画像の分解能が均一(線型変化)になるような、自由曲面形状の全方位ミラー[30]を新 たに設計した.この様な、画像分解能が均一な全方位ミラーは、人がモニタリングする ような応用に適していると考える.今後、この自由曲面ミラーを試作し、光学特性を評 価する.さらに、双曲面、自由曲面ともに、小型化のために二枚反射全方位ミラー[31,32] に拡張設計していく.

5-4 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

5-4-1 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,低解像度高フレームレート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ための手法について説明する.また,1 画素あたりのデータは 8bit である.







図80 画像の取り込み、保存アルゴリズム


(a) 撮像画像(1008×1018画素)



(b) 一部を拡大した図(122×98画素)

図81 撮像画像

画像を取り込み,取り込みの終了を待ってからの画像の表示,保存を行うと画像取り 込みの終了を待つ時間が無駄になり,処理全体が遅くなってしまう.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.

ここでダブルバッファ処理について説明しておく.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79に示す.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.図79に示すようにダブルバッフ 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.また,ダブルバッファを用いた取り込みから保存までのアルゴリズムを図80 に示す.

実際にカメラで撮像した画像を図80(a)に示す.図81(b)は撮像画像を一部拡大したものであるが、CCDが単板カラーのBayer 配列となっているので各画素にR,G,Bの1つの情報しかない.この画像を後に述べる色の補間処理によってRGBカラー画像に変換する.

5-4-2 書き込み速度の最適化

低解像度高フレームレートカメラのフレームレートは 30fps, 解像度が 1008× 1018pixel であり画像1フレームのデータサイズは約 1Mbyte であるので,1秒間に約 30Mbyte の画像データを実時間で書き込まなければならない.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次に示す予備実験を行った.

実験内容は 30Mbyte のデータを書き込む際に, 書き込みブロックサイズを変化させて いき, その時の書き込み速度を計測するというものである. この結果を図82に示す.



図82 書き込みブロックサイズと書き込み速度

図82より,書き込みブロックサイズが500Byteから50000Byteの範囲では書き込み 速度が速いことがわかる.この結果より書き込みブロックサイズを1008Byte,つまり 画像データ1ライン分にして保存していくことにした.書き込みブロックサイズを 1008Byteにした結果,30fpsの画像データを実時間で取り込み保存することができた.

5-4-3 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5fps,低解像度高フレームレートカメラは30fpsであるため,高 フレームレート画像列の1フレーム目と低フレームレート画像列の1フレーム目,高フ レームレート画像列の7フレーム目と低フレームレート画像列の2フレーム目が同期す るように撮像しなければならない.

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5 fps と 30 fps であるので、高 解像度低フレームレートカメラに 5 Hz,低解像度高フレームレートカメラに 30 Hz 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.

図83に示すパルス信号を出力するようなプログラムを PeakFPGA Design Suite (Protel 社) というソフトを用いて VHDL 言語 (Very high-speed integrated circuit Hard-ware Description Language) で記述し論理シミュレーションを行なった. シミュ レーション結果を図84に示す. 図84に示すように 5Hz のパルス信号と 30Hz のパル ス信号を同期させながら出力できた. 今後は,前述した VHDL のプログラムを論理合成 して Max+Plus II Baseline (Altera 社) というソフトを用いて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シミュ レーションを行う予定である.



図83 各カメラへのパルス信号

VHDL Simulator - [TEST_TAO1]								
🚟 Eile <u>V</u> iew <u>S</u> imulation <u>W</u> indow	<u>H</u> elp							
i 🔁 🔒 😥 😥 😓 🕒	•		•B) 🔑	1	TEST TA	01.VHD	-	
-			100ms		200ms	3	00ms	400ms
CLOCK='0' C_30='0' C_5='0'						ſ		
	►	•						

図84 パルス信号出力の論理シミュレーション結果

5-4-4 まとめ

提案システムのうち、低解像度高フレームレートカメラを用いた画像の取り込み、保存の流れを示した.画像データの書き込み速度の最適化を行なうことにより、30fpsの画像データをコマ落ちすることなく保存できることを示した.また、取り込んだ Bayer 配列の画像を色の補間処理を行なって RGB カラー画像にし、その RGB カラー画像のカラーバランスの補正ができたことを示した.さらに、2台のカメラの同期をとるためのパルスジェネレータの論理シミュレーションを行い、適切なパルス信号が出力されていることを示した.

5-5 総括

- 1. 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種類の動画像を撮像するシス テムを試作した.
- 2. 複合センサカメラにおいて幾何学的キャリブレーションを行い2種類の異なった 解像度画像を3次元空間中の同一基準平面に投影し,一致をはかる手法を開発した.
- 3. 同一視野の2種類の動画像から、高解像度・高フレームレートの動画像を画像処 理によって生成する周波数変換を利用したアルゴリズムを提案した.
- 4. 提案した周波数変換を用いた高精細画像生成アルゴリズムの実装した.
- 5. 周波数変換を利用した手法において、MPEG テストシーケンスを用いたシミュレー ション実験と生成画像の数値的評価を行い、2 種類の画像のサイズ比が 2:1、フレー ムレート比が 1:6 の場合には、真の高解像度画像に近い画像が提案手法によって得ら れることを確認した.
- 6. 同一視野の2種類の動画像から,高解像度・高フレームレートの動画像を画像処 理によって生成するモーフィングを利用したアルゴリズムを提案した.
- 7. モーフィングを利用した手法において,処理の流れを4段階に分けて,各段階に おける画像処理の実装を行った.
- 8. 実装されたモーフィングによる処理を、各実装モジュールごとに実験により評価 検証を行った.
- 9. 複合センサの CCD サイズを想定して全方位ミラーの光学設計を行った.
- 10. 光学設計に基づいて、全方位ミラーの試作を行った.
- 11. 従来のドーム型支持具の問題点であった,屈折や加工コストを解決する棒型支持 具を設計試作した.
- 12. 低解像度高フレームレートカメラを用いた画像の取り込み保存を行うソフトウ ェアを作成し、画像データの書き込み速度の最適化を行なうことにより、30fpsの画 像データをコマ落ちすることなく保存できることを示した.
- 13. 取り込んだ Bayer 配列の画像を色の補間処理を行なって RGB カラー画像にし, その RGB カラー画像のカラーバランスの補正を行うソフトウェアを作成した.
- 14. 2 台のカメラの同期をとるためのパルスジェネレータの論理シミュレーションを 行い, FPGA 回路のプログラムを作成した.

参考資料、参考文献

- [1] 金澤勝, 菅原正幸, 三谷公二, 島本洋, 山下誉行, 浜田宏一: 超高精細映像システムの 開発, 第9回 VMA 研究会
- [2] 杉田馨, 高野孝英, 苗村健, 原島博:4 眼撮像系を用いた広視野角ステレオ画像通信, 3 次元画像コンファレンス'99, 1-7, pp. 37-42 (1996)
- [3] 田中健二, 鈴木保成, 荒川佳樹, 田中英史, 佐藤正人: 800 万画素超高精細カメラ, 第 13回画像入力シンポジウム, pp. 10-14 (2001)
- [4] 続元宏,岩田洋夫:旋回式高解像度実画像ディスプレイ,日本バーチャルリアリティ 学会第6回大会論文集(2001)
- [5] 菅原正幸, 三谷公二, 齋藤敏紀, 藤田欣裕, 末次圭介:4 板撮像方式における画素ずらし効果についての検討, テレビジョン学会誌, Vol. 49, No. 2, pp. 212-218 (1995)
- [6] Barnard. K, Cardei. V, Funt. B:A Comparison of Computational Color Constancy Algorithm-Part 1:Theory and Experiments with Synthetic Data, IEEE Transactions on Image Processing, Vol. 11, No. 9, pp972-984 (2002)
- [7] R. A. Jarvis, ``A perspective on range finding techniques for computer vision'', IEEE Transaction on Pattern Analysis and Machine intelligence. Vol. 5, pp. 122-139, 1993.
- [8] M. Subbarao and N. Gurumoorthy, `Depth recovery from blurred edges', Proceedings of Computer Vision and Pattern Recognition, pp. 253-259, 1982.
- [9] R. Y. Tsai, ``A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-self TV cameras and lenses'', IEEE J. Robot. Automat. RA-3, No. 4, pp. 323-344, 1987.
- [10] 天野晃,浅田尚紀,馬場雅志, ``ズームレンズの光学的ひずみの解析と補正:可 変円筒モデルを用いた口径蝕現象の解析と画像周辺明度低下の補正',電子情報 通信学会論文誌 D-II, Vol. J80-D-II, pp. 1458-1465, 1997.
- [11] 高木幹雄,下田陽久 監修, ``画像解析ハンドブック'',東京大学出版会,1991.
- [12] S. C. Park, M. K. Park, and M. G. Kang, `Super-resolution image reconstruction: A technical overview', IEEE Signal Processing Magazine, Vol. 20, Issue 3, pp. 21-36, 2003.
- [13] H. Shekarforoush and R. Chellappa, `Data-driven multi-channel super-resolution with application to video sequences', Journal of the Optical Society of America A, Vol. 16, No. 3, pp. 481-492, 1999.
- [14] C. A. Segall, R. Molina, and A. K. Katsaggelos, ``High-resolution images from low-resolution compressed video'', IEEE Signal Processing Magazine, Vol. 20, Issue 3, pp. 37-48, 2003.
- [15] J. Watkinson, ``The Engineer's Guide to Motion Compensation'', Snell & Wilcox, 1994.
- [16] Y. K. Chen, A. Vetro, H. Sun, and S. Y. Kung, ``Frame-rate up-conversion using transmitted true motion vectors', In Proceedings of IEEE Second Workshop on Multimedia Signal Processing, Los Angeles, CA, pp. 622-627, 1998.
- [17] T.Beier, S.Neely, 'Feature-Based Image Metamorphosis'' In SIGGRAPH 92 Conference Proceedings (1992), ACM Press, pp. 35-42.
- [18] Robert Szewczyk, Andras Ferencz, Henry Andrews, Brian C.Smith, "Motion and Feature-Based Video Metamorphosis", ACM Multimedia 97 - Electronic Proceedings, Nov 1997.
- [19] Kwai Hung Chan, Rynson W. H. Lau, "Contour-Based Warping", CVGIP: Graphical

Models and Image Processing, Academic Press, 60(5), pp331-348, Sep 1998.

- [20] Harris C., M. Stephens, "A combined corner and edge detector", Proc. 4th Alvey Vision Conf., pp. 147-151, 1998.
- [21] D. Chetverikov, J. Verestoy, "feature point tracking a new algorithm", In Proc. International Conference on Pattern Recognition, pp1436-1438, 1998.
- [22] D. Chetverikov, J. Verestoy, "Feature Point Tracking for Incomplete Trajectories", Computing, Devoted Issue on Digital Image Processing, 62:pp321-338, 1999.
- [23] A.Fusiello, E.Trucco, T.Tommcini, V.Roberto, "Improving Feature Tracking with Robust Statistics", Pattern Analysis & Applications, pp312-320, Feb., 1999.
- [24] C. J. Veenman, M. J. T. Reinders, E. Backer, "Resolving Motion Correspondence for Densely Moving Points", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 1, pp54-72, January 2001.
- [25] Y. Yagi, "Omnidirectional Sensing and Its Applications", IEICE Trans. Information and Systems, Vol E82-D, No. 3, pp. 568-579, Mar, 1999.
- [26] Y. Yagi, S. Kawato and S. Tsuji, "Real-time omnidirectional image sensor (COPIS) for vision guided navigation", IEEE Trans. Robotics & Automation, Vol. 10, No. 1, pp. 11-22, 1994.
- [27] J. Hong, X. Tan, B. Pinette, R. Weiss and E. M. Riseman, "Image-based Homing", Proc. IEEE Int. Conf. Robotics and Automation, pp. 620-625, 1991.
- [28] 山澤,八木,谷内田:移動ロボットのナビゲーションのための全方位視覚系 HyperOmni Visionの提案,電子情報通信学会論文誌 D-II, Vol. J79-D-II, No. 5, pp. 698-707, 1996.
- [29] Nayar, S.K., "Catadioptric Omnidirectional Camera", Proc. Computer Vision and Pattern Recognition, pp. 482-488, 1997.
- [30] J. S. Chahl and M. V. Srinivasan," Reflective surface for panoramic imaging", Applied Optics, Vol. 36, No. 31, pp. 8275-8285, Nov. 1997.
- [31] S. K. Nayar and V. Peri , "Folded Catadioptric Cameras", Proc. of IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, June 1999.
- [32] Y. Yagi, M. Yachida, "Development of A Tiny Omnidirectional Image Sensor", In Proceedings of the Fourth Asian Conference on Computer Vision, vol.1, no. 23-28, 2000.

研究発表、講演、文献等一覧

なし