

平成30年6月19日現在

機関番号：14603

研究種目：挑戦的萌芽研究

研究期間：2016～2017

課題番号：16K12407

研究課題名（和文）4D超臨場TV向け16Kライトフィールド映像処理アクセラレーション

研究課題名（英文）16K Lightfield Acceleration for Super-realistic 4D-TV

研究代表者

中島 康彦（Nakashima, Yasuhiko）

奈良先端科学技術大学院大学・情報科学研究科・教授

研究者番号：00314170

交付決定額（研究期間全体）：（直接経費） 2,500,000円

研究成果の概要（和文）：ライトフィールド動画の伝送高速化、および、処理高速化を目的として、まず動画圧縮専用ハードウェアの検討を行った。圧縮率優先時には、先行8K動画圧縮ハードに対し50%のハード増加と34%の圧縮率向上、小型化優先時には、先行研究に対し24%のハード削減と6%の圧縮率向上を確認した。処理高速化にはCGRAを導入し、組込GPUの1/3の計算資源でレンダリング89%、距離計測4倍の性能を達成した。また、動画認識へのCGRA導入も検討し、ARM Coretex-A9の40倍、Vivante GC2000+の11倍、Xilinx Zynq(Z-7020)高位合成の6倍の性能を確認した。

研究成果の概要（英文）：For the purpose of speeding up transmission of light field movie and speeding up processing, we first studied dedicated video compression hardware. At the time of compression priority, we confirmed hardware increases in 50% and a compression ratio is improved by 34% compared with the preceding 8K motion picture compression hardware. At the time of miniaturization priority, we confirmed hardware is reduced by 24% and a compression ratio is improved by 6% compared with the previous research. For high-speed processing, CGRA was introduced, achieving 89% of the performance in rendering and 4 times faster in distance extraction with 1/3 of computing resources in embedded GPU. We also examined the employment of CGRA to motion picture recognition, confirmed the performance of 40x of ARM Coretex-A9, 11x of Vivante GC 2000+, 6x of Xilinx Zynq (Z-7020) high-level synthesis.

研究分野：計算機アーキテクチャ

キーワード：ライトフィールド アクセラレータ CGRA GPU DCNN

## 1. 研究開始当初の背景

CCD 撮像素子の高解像度化および動画圧縮復元技術の高性能化により、4Kテレビが実用化されて久しい。また、ステレオ立体視を応用した3Dテレビも一時期注目を集めた。しかし、臨場感向上を掲げた高解像度化や3D化は、消費者が求める質的向上からほど遠いものとなり、テレビ離れと相まって、テレビ業界および映像電子産業全体の衰退を招いている。要因の1つは、メモリの微細化競争と同様に、単に高解像度化することが価値を生み出すとの誤解が長く業界を支配したことであり、もう1つは、従来の3D映像では3軸方向の視点情報が全て固定されており、Z軸方向の視覚情報があるのに自由度がない点が視聴を不自然にし、長時間の視聴を困難にしたことにある。総じて、これまで、消費者の真の臨場感演出要求に応える先端画像処理技術の追求がなされてこなかったことが敗因であると言える。

そもそも臨場感とは、単に解像度が高いとか、立体的に飛び出して見えることではない。例えば、360度全周映像のリアルタイム3D表示技術を究極のゴールとする考え方がある。しかし、あまりに多い画像情報はいかに高効率の圧縮アルゴリズムを適用しても十分な速度で配信することができない。また、視野の全てを消費者が自由に変更できる状況は、映像製作側の意図が消費者に正しく伝わらない不具合を生じさせる。すなわち、テレビに求められる次世代臨場感とは、あらかじめ決められた構図やストーリーから大幅に外れることなく、しかし、消費者がある程度の自由度を与えられて映像の変化を能動的に制御できる極自然な環境(完全フリーではなくツアーガイド付きの旅行)であると考えている。以降、このような環境を4Dと呼ぶ。

## 2. 研究の目的

人間が景色や対象物を観察する時に重要な情報は、固定された視点位置から見える静止画と距離情報(従来の3D映像)だけではない。意識的あるいは無意識的な視点移動から得られる、対象物の重なり具合の能動的な微小変化(網膜像のリフレッシュに必要な眼球の微小運動とは異なる)が重要である。テレビが真の臨場感を演出するためには、消費者に届ける情報自体に、消費者が視点位置を微小変化させた場合の様々なシーンを全て包含し、消費者が能動的に制御できる十分なデータが含まれている必要があり、かつ、360度全周映像のような過剰な情報を必要としない適切なデータ構造でなければならない。本研究では【1】既存技術では到達不可能な16K解像度4D映像処理に必要な基盤技術(圧縮・復元・能動表示)をコンピューティング基盤の視点から明らかにし、【2】Light Field 原画像をベースとして次世代テレビに相応しい臨場感演出に最適なハニカム構造画像形

式と準専用ハードウェア構成方式を突き止める。

## 予想される結果と意義

2D動画やステレオ動画には8Kを超える解像度の撮像素子は過剰スペックであり、今後の発展は望めない。しかし、4D動画への新たな展開と、次世代コンピューティング基盤の組み合わせは、真に消費者の要求に応える映像製品の実現や産業利用だけでなく、8Kを超える次世代撮像素子(センサ)および超高速計算基盤(SoC)の両面から最先端技術開発を牽引できる。新興国の追い上げに右往左往することなく、明確な目標を掲げて新技術を開発する道標となる。

## 3. 研究の方法

平成28年度は、市販最大解像度(8K)のLight Field Cameraを購入し、まず、メニコア、GPGPU、FPGAなどの既存計算基盤上において圧縮・復元・能動表示の基本検討に取り組む。

### 【1】16K解像度4D映像処理に必要な圧縮・復元技術の開発



図1 Light Field 原画像(一部)

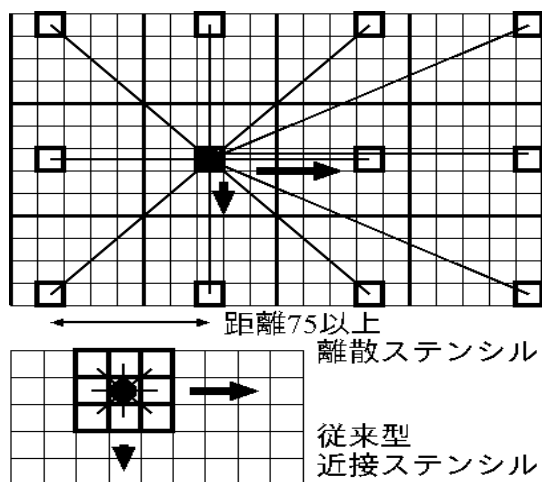


図2 Light Field 画像処理の難しさ

最新の H.265 規格でも、動画の圧縮には、フレーム間の類似性に着目した動き予測に計算時間の大部分を費やしている。一方、Light Field 映像には、図1に示したように、フレーム間の類似性に加えて、同一フレーム内の隣接微小画像間に類似部分が多い。ただし、微小画像間の差分として奥行き情報が表現されているため、僅かな差分を全て冗長情報として除去してしまうと奥行き情報が完全に失われる。75×75程度の微小画像内において輪郭抽出を行い、除去可能な情

報を選別するなどの高度な圧縮手法が必要であると予想する。あるいは、微小画像の辺縁部分は歪が大きいので圧縮対象から除外するほうが最終的には良い結果を得られることも考えられる。

また、フレーム内の類似性に基づき第1段階の圧縮を行った後、さらに圧縮に用いた情報をフレーム間の類似性検出に適用して第2段階の圧縮を行う技術が考えられる。圧縮時に階層符号化を行うことで、復元時にスケーラブルな選択が可能となり、様々なデバイスへの最適表示が可能とすることができる。特に、空間解像度と画質別のスケーラビリティは既存映像でも実装例がない。また、データ復元時の高並列処理は一般に難しい。復元時に高並列度が得られるとともに、圧縮率が損なわれない圧縮・復元手法の開発が必要である。

以上のような様々な可能性を探索しつつ技術開発を推進し、有用な距離情報を残しつつ十分な圧縮率を達成し、現実的な通信インフラにより配信できる、Light Field 動画の基本圧縮・復元アルゴリズムを完成させる。

## 【2】能動的に鑑賞できる自然な4D表示環境の実現

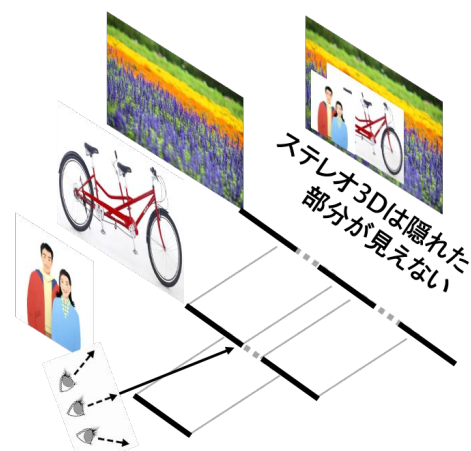


図3 視点移動により隠れた部分が見える

復元したデータから距離情報を含む映像情報を取り出して高速動画表示しつつ、消費者の視点位置に追従して能動的に視点位置を変化させる技術開発に取り組む。

まず、視点位置は消費者の頭部の位置（上下左右前後の3軸）により決定できる。表示モニタに取り付けた簡便なカメラにより十分な精度で計測できる。次に、Light Field 映像から、距離情報により打ち抜いた画像を取り出して重ね合わせる。図3に示すように、従来のステレオ3D画像では、視点をどのように移動しても隠れた部分を見ることができない不自然さがあるのに対し、Light Field 画像では、1枚の静止画であっても視点位置に応じて重なり具合を動的に変化させることが可能である特長を最大限に利用する。さらに動作再生中に本機能を活用して能動表示を実現する。このような映像は、十分な解像度と視点移動の自由度があれば、専用メガネによる強制的な立体視を不要とできる可能性がある。3Dテレビが普及しなかった理由の1つは、メガネ人

口が多い日本において、メガネの上にメガネをかけることの不自由さにあった。メガネ不要のシステムも存在したものの、視野角の狭さから同様に視聴は不自由であった。消費者の負担を圧倒的な演算能力でカバーできる超高速計算基盤を追及し、解像度一辺倒であったこれまでの映像表示装置の常識を高価な全周スクリーンを投入することなく根底から変える4D映像処理システムの基本技術を確立する。

平成29年度は、基本検討結果をベースに、プロトタイプの実装および評価を行う。また、研究過程において判明したボトルネックやオーバヘッドを解消する効果がある、原画像形式の最適化に取り組み、最高性能の4Dシステムとして総括する。最終的には、最近登場したハニカム構造 Light Field 映像を最適利用する高精度な4Dシステムのプロトタイプ完成を目指す。

## 【3】超高速能動表示のための準専用ハードウェアの策定

図2に示した離散ステンシル計算を圧倒的な絶対性能で高速に行う専用ハードウェアを策定し、プロトタイプシステムを開発する。汎用 CPU や GPGPU は離散ステンシル計算との相性が悪く、絶対性能を達成するには大量の計算資源を投入しなければならない。また、プログラマビリティが全くない真正専用ハードウェアでは設計後の微調整ができず、前述のように多数の FIFO を利用する一般的構成では FIFO が巨大化するため効果が期待できない。FPGA では動作周波数やメモリバンド幅を十分に確保することができない。圧縮・復元・能動表示のアルゴリズムを以上の既存計算基盤上に実装し、ボトルネックを明らかにしたうえで、アルゴリズムの主要なパラメタを設計後に調整可能な準専用ハードウェアの策定を開始する。

既存技術のうち最も近いハードウェア構成は、CGRA ( Coarse Grained Reconfigurable Architecture) であるものの、CGRA のままでは演算器間ネットワークの自由度が過剰であり動作周波数が上がらない。そこで、本研究では、CGRA よりもさらに再構成粒度を粗くし、ハードウェアの複雑度を大幅に低下させる準ハードウェア化を最終目標とする。演算器とメモリブロックの間を離散ステンシルアルゴリズムの実装に必要な最低限のネットワークにより接続することで、特に問題となるネットワーク遅延を大幅に低減させることを狙う。また、4D映像処理は、ある程度の画面分割により単純な高速化が可能と考える。各ユニットの複雑さを極限まで抑えることで、最小限のプログラマビリティを維持しつつ、CGRA よりも高速かつコストのかからない高並列構成を目指す。シミュレータや現有 FPGA システムを利用して性能見積もりを行い、最終的には ASIC 化容易な基本設計を完了する。

## 【4】臨場感演出に最適なハニカム構造画像形式の最適利用技術

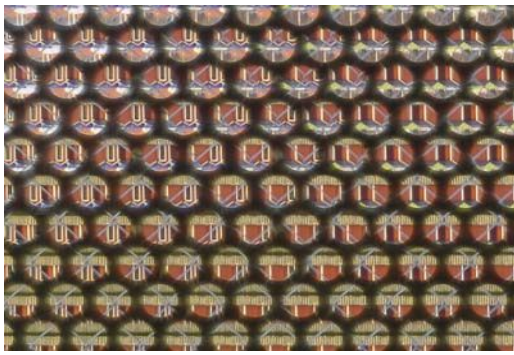


図4 ハニカム構造の微小画像形式

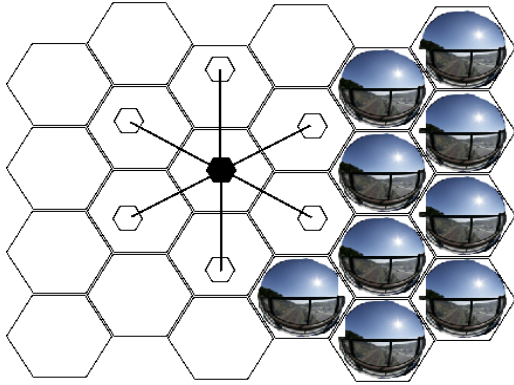


図5 ハニカム構造の離散ステンシル計算

ここまで、図1に示すような矩形の微小画像から構成される原画像を想定して説明を行った。矩形の場合、多くの静止画像が入手可能であり、特定の Camera に固有の画質や解像度に縛られない普遍的なアルゴリズム開発を行うことが可能である。また、矩形は、既存の圧縮・復元アルゴリズムとの相性が良いため、研究を早期に立ち上げることができるメリットがある。

しかし、最近登場した高解像度 Light Field Camera(購入予定の RAYTRIX 製 R29)は、図4に示すようにハニカム構造の微小画像を採用している。この形状は、精度の高い球形微小レンズの最密配置に対応させることで、微小画像の辺縁部分の歪を格段に減らせるメリットがあり、画像全体に占める有用なデータの比率を向上させ、全体4D映像再生時の臨場感向上に大きく貢献できると予想される。しかし、ハニカム構造は、映像をメモリに格納し取り出す際のアドレス計算(座標計算)を極めて複雑にするため、離散ステンシル計算の計算量をさらに増大させる。ハニカム構造の表現と参照に適した斬新なメモリ構成方法や演算器間ネットワークを考案することができれば、計算量を増大させることなく高性能を達成できる。本研究では、最終的に、このようなハニカム構造の高精度原画像に対して超高速に4D映像処理を行うことができるプロトタイプシステムを完成させ、4D映像による真の臨場感演出の効果を実証する。

#### 4. 研究成果

[28年度]

市販最大解像度の8K-Light Field Cameraを購入し、一次画像データを用いて、メモリ、GPGPUなどの既存計算基盤上における圧縮・復

元・能動表示の性能評価、および、ハードウェアによるアクセラレーション効果の見積りを行った。

[1] 圧縮・復元については、8K動画像向け最新HEVC規格に準拠したエンコーダ(Kvazaar)を使用し、ライトフィールド画像特有の多数微小レンズ画像とHEVCエンコーダの親和性について調査を行った。評価の結果、ライトフィールド動画を圧縮する場合、一般的な8K動画像向けに実装されている多くのHEVC圧縮手順が省略できること、また、不要な機能を削除して専用ハードウェアを小型化できる見通しが得られた。また、HEVCハードウェアエンコーダに関する先行研究と比較しながら、ライトフィールド動画圧縮専用ハードウェアの規模を見積もった結果、圧縮率優先構成では先行研究の8K動画圧縮ハードに対し、50%のハードウェア増加により平均34%の圧縮率向上が可能であること、また、小型化優先構成では先行研究に対し、24%のハードウェア削減と平均6%の圧縮率向上が可能であることを明らかにした。

[2] 能動的に鑑賞できる4D表示環境の実現に関しては、復元した一次画像を可視化するレンダリングと、一次画像から奥行き情報を取り出す距離計測を高速に行う計算基盤について、GPGPUとCGRAによる性能評価を行った。GPGPUの特性を生かすチューニング手法を考案し、GTX-780ではレンダリングと距離計測を各々30%、82%高速化することに成功した。

[3] 超高速能動表示のための準専用ハードウェアの策定に関しては、さらにCGRAによる高速化に取り組み、ライトフィールド画像処理に特化したCGRAの詳細シミュレータを開発して評価した結果、提案CGRAは、組み込みGPUの3分の1の計算資源でも、レンダリングで89%、距離計測では4倍の性能を達成できる見通しを得た。この際、[4] 臨場感演出に最適なハニカム構造画像形式の最適利用技術も同時に組み入れ、ハニカム画像についても同等の高速化が可能である見通しを得た。

[29年度]

順調な研究進捗により、ライトフィールド画像処理そのものを高速化する目処が立ったため、さらに、注目を集めるようになった動画認識技術への応用を試みた。特に、カラーアプリケーションとして有望視されはじめた動画認識においては、認識精度向上のために背景除去が役立ち、このために、ライトフィールド画像から得られる奥行き情報を利用できる。また、動画認識処理では、畳み込み層と全結合層の連携により計算を行う。この際問題となるのが、動画を取得するデバイスからクラウドへ全ての動画を送信すると膨大なデータ量のために通信路がパンクすることである。デバイスあるいはエッジにおいて、動画認識処理を行い、結果のみをクラウドに送信することが望ましいものの、認識処理には相当の計算パワーを必要とするため、計算の共有化による計算量の削減、および、中間結果を送信する際に必

要となるデータ圧縮技術が重要となる。そこで、分散 CNN システムのモデルを構築し、圧縮率と認識精度の関係について研究を行った。

複数のデバイスに採用されている DCNN (Deep Convolutional Neural Network) を使用した並列処理モデルと、デバイス間のネットワークラフィックのサイズの縮小に取り組み、中間データを圧縮し、動画認識のために AlexNet の共通計算を集約する手法を提案した。実験の結果、Zip の無損失圧縮はデータ量を最大 1/24 まで削減でき、HEV 非可逆圧縮では 1/208 までデータを削減でき、わずか 3.5% の認識精度の低下で済むことが明らかになった。さらに、共通計算の集約により、30 個の DCNN の計算量を 90% 削減できた。

また、開発した CGRA を動画認識へ応用する研究も行った。Deep Learning を用いた画像認識で標準的に使用される DCNN は、非常に計算量が多いため、高性能な計算機が必要である。高性能な計算機が搭載できない組み込み機器へ、この画像認識技術を搭載するために、低消費電力で高性能を狙った CGRA 型のアクセラレータである EMAXV をベースに、DCNN 最適な CGRA を探索し、予備評価を行った。評価によって得られた結果から、さらに探索を進め、ARM Coretex-A9 のデュアルコアに比べて 40 倍、Vivante GC2000+ に比べて 11 倍、Xilinx Zynq(Z-7020) の FPGA 部に高位合成を使用し、演算処理回路を実装したものに比べて 6 倍の性能が出せることを見積もりにより確認した。

## 5. 主な発表論文等

(雑誌論文) (計 3 件すべて査読有り)

Takamasa Mitani, Hisakazu Fukuoka, Yuria Hiraga, Takashi Nakada and Yasuhiko Nakashima: "Compression and aggregation for optimizing information transmission in distributed CNN", CANDAR'17, REGULAR PAPER, pp.112-118, Nov. (2017)

Yuttakon YUTTAKONKIT, Shinya TAKAMAEDA-YAMAZAKI and Yasuhiko NAKASHIMA, "Performance Optimization of Light-field Applications on GPU", IEICE Trans., E99-D, pp.3072-3081, <http://doi.org/10.1587/transinf.2016EDP7090> (2016)

Yuttakon YUTTAKONKIT, Shinya TAKAMAEDA-YAMAZAKI and Yasuhiko NAKASHIMA, "Performance Comparison of CGRA and Mobile GPU for Light-field Image Processing", CANDAR'16, REGULAR PAPER, CANDAR'16, pp.174-180, 10.1109/CANDAR.2016.0040 (2016)

(学会発表) (計 8 件すべて査読無し)

福岡久和, 平賀由利亜, 三谷剛正, 中田尚,

中島康彦: "分散CNNにおける圧縮と集約による情報転送の最適化", 信学技報, vol.117, no.314, CPSY2017-59, pp.51-54, 2017年

福岡久和, 三谷剛正, 平賀由利亜, 中田尚, 中島康彦: "動画圧縮技術を利用した分散機械学習における情報伝達効率化", 信学技報, vol.117, no.153, CPSY2017-30, pp.151-155, 2017年

一倉孝宏, 山野龍佑, 福岡久和, 中島康彦, "DCNNに最適なCGRAの探索と予備評価", 電子情報通信学会コンピュータシステム研究会, 2017年

嶋谷知, 三谷剛正, 中島康彦, "ライトフィールド動画の符号化配信ソリューションの検討", 電子情報通信学会コンピュータシステム研究会, 2016年

三谷剛正, 嶋谷知, ユッタコンキット ユッタコン, 中島康彦, "レンズ境界部除去によるライトフィールド動画のHEVC圧縮高速化検討", 電子情報通信学会コンピュータシステム研究会, 2016年

中島康彦, "EMAXVにおける複数バースト転送と複数ベクトル演算のオーバーラップ手法", 電子情報通信学会コンピュータシステム研究会, 2016年

嶋谷知, 三谷剛正, ユッタコンキット ユッタコン, 中島康彦, "GPUを用いたライトフィールド動画のHEVC圧縮高速化手法の検討", 電子情報通信学会コンピュータシステム研究会, 2016年

三谷剛正, 嶋谷知, ユッタコンキット ユッタコン, 中島康彦, "ライトフィールド動画に最適なHEVC圧縮アルゴリズムの提案", 電子情報通信学会コンピュータシステム研究会, 2016年

[産業財産権]

出願状況(計1件)

名称: データ処理装置(メモリ内蔵アクセラレータの構成方法)

発明者: 中島康彦, 高前田伸也

権利者: 奈良先端大

種類: 特許, PCT/JP2016/061302

出願年月日: 2016年4月6日

国内外の別: 海外

[その他]

ホームページ等

<http://arch.naist.jp/>

## 6. 研究組織

(1) 研究代表者

中島 康彦 (NAKASHIMA YASUHIKO)

奈良先端科学技術大学院大学・情報科学研究科・教授

研究者番号: 00314170