

令和元年6月6日現在

機関番号：11301

研究種目：挑戦的萌芽研究

研究期間：2016～2018

課題番号：16K12426

研究課題名(和文)空間浮遊3D点描画を用いた3D映像システム

研究課題名(英文)3D Imaging System using Aerial 3D Display

研究代表者

青木 輝勝(Aoki, Terumasa)

東北大学・未来科学技術共同研究センター・准教授

研究者番号：00302787

交付決定額(研究期間全体)：(直接経費) 2,600,000円

研究成果の概要(和文)：本研究では、ディスプレイやスクリーンが存在しない空間上に表示された点描画とデジタル機器を連携させるための画像認識技術を開発する。これを実現するための中心課題は、(1)発光体、(2)点描画、(3)半透明という従来画像処理/画像理解分野で研究対象とされることの少なかった被写体を対象とした新しい画像認識技術の開発である。この問題を解決するため、形状認識的アプローチと深層学習的アプローチの2つの手法を提案し、それらの手法の有効性を実証している。

研究成果の学術的意義や社会的意義

本研究の学術的特色は、(1)発光体、(2)点描画、(3)半透明(点描画の点と点の隙間に背景が写るため)という従来画像処理/画像理解分野で研究対象とされることの少なかった被写体を対象に新しい画像マッチング技術を構築する点である。本研究の成果はAerial 3D Displayで生成された点描画コンテンツのみならず、従来認識が困難であった(1)～(3)の性質を持つ被写体に対しても認識を可能にするため、その用途も非常に広い。

研究成果の概要(英文)：An image recognition method for digital devices (such as smartphones) to recognize 3D point clouds projected by Aerial 3D Display is developed in this research. The core technology to achieve this is development of a novel image recognition method for light-emitting, point-drawing and semi-transparent objects. To solve this problems, two approaches - a shape-recognition based method and deep-learning based method - are developed. And the effectiveness of these methods are shown in this report.

研究分野：画像処理・画像理解

キーワード：Aerial 3D Display 画像認識 3次元点描画 形状認識 深層学習

1. 研究開始当初の背景

ディスプレイやスクリーンが存在しない空間上に映像コンテンツを映し出すことは長らく無謀な夢であったが、Aerial 3D Display (日本未来科学館空中 3D ディスプレイプロジェクト) が発表され、その夢に一步近づく状況となってきた。Aerial 3D Display は空間上に点描画を映し出す将来性の高いシステムであるが、提示される点描画コンテンツは表現力が限られているため次世代エンターテインメントなどの用途に対しては実用レベルとは言い難く、現在、本システムは試作レベルに留まっている状態である。本研究は、日本未来科学館空中 3D ディスプレイプロジェクトの全面的な協力のもと、Aerial 3D Display によって映し出される点描画コンテンツとデジタル機器 (スマートフォン、タブレット、ヘッドマウントディスプレイなど) を連携させるための新しい認識手法を研究開発することを目的とする。

空間中に描かれた 3 次元の点描画コンテンツを正確に認識することは技術的に容易ではない。一般に、画像処理 / 画像理解分野では、(I) 発光体の認識は困難である、(II) 点描画コンテンツの認識は困難である、(III) 半透明コンテンツの認識は困難である、などが広く知られているのに対し、対象コンテンツ (点描画コンテンツ) はまさにこれらのすべての性質を有するためである。これまで、このような 発光体、点描画、半透明 (点描画の点と点の隙間に背景が写るため) という従来画像処理 / 画像理解分野で研究対象とされることの少なかった被写体を対象とした新しい画像認識技術が研究開発された事例は研究代表者の知る限り存在しない。本研究の成果は Aerial 3D Display で生成された点描画コンテンツのみならず、従来認識が困難であった (I) ~ (III) の 1 つ以上の性質を持つ被写体に対しても認識を可能にするため、その用途も非常に広い。

2. 研究の目的

Aerial 3D Display によって空間中に描かれた 3 次元の点描画コンテンツを正確に認識する技術は現在の技術 (の組み合わせ) のみでは非常に困難である。それは、3 次元点描画コンテンツが (I) 発光体である、(II) 点描画である、(III) 半透明である、という性質を持つのに対し、画像処理 / 画像理解分野においてこのような被写体を対象とした研究事例はこれまで非常に少なく (特にこの 3 つの性質すべてを対象とした研究事例は研究代表者の知る限り存在しない)、その精度も不十分であるためである。

(I) 発光体の問題に関しては、困難要因のひとつは描画システム (Aerial 3D Display) とビデオカメラの周波数の差である。この両者の周波数がきちんと同期していない場合、あるビデオフレームには発光体が写るが、次のビデオフレームにはその部分が映らない、という現象が生じてしまう。このため、認識以前の問題として被写体がどのような描画点の集合なのかをきちんと復元しなければならない。また、別の問題として、発光体によって生成された点は画像内ではいわば孤立点に相当するため、雑音との区別がつきにくいという問題も発生してしまう。

(II) 点描画の問題に関しては、仮にすべての描画点が取得できたとしても、「図形としてどの点とどの点を接続すべきなのか？」という問題がある。これは画像処理分野でも広く要求される技術 (例 : (閉ループの) 輪郭線検出など) であるが、汎用的に高精度な性能を示す手法は現時点では提案されていない。また、近年、3 次元計測機器 (RGB-D カメラ等) の著しい低コスト化に伴い、実在の 3D オブジェクトから Point Cloud (奥行きを含めた 3 次元上の計測点の座標の集合) を生成し、その情報から高精細な 3 次元モデルを生成するなど様々な Point Cloud 処理を試みる研究が盛んに行われている。しかしながら、発光体点描画の集合に対しては、そもそも既存機器で奥行き情報を取得することが不可能であることに加え、単眼画像の情報をベースとした 3 次元復元処理 (Shape from X 等) によって奥行き情報を推定することも困難である。

(III) 半透明の問題に関しては、主として雑音除去の問題に帰着する。特に、描画点と描画点の間に写る背景に発光体が含まれる場合や鏡面反射の強いオブジェクトが含まれる場合、それらの背景情報を雑音として取り除く必要がある。

本研究では、以上 (I) ~ (III) の問題を同時に解決し、その結果として点描画コンテンツを高速高精度に認識するための技術を研究開発することを目的とする。

3. 研究の方法

(1) 解決すべき課題の実証的な集約

本研究は、広い意味で形状認識に属する内容である。形状認識は、様々な画像認識・理解技術の基盤となる重要な技術であるが、既存の画像認識技術のほとんどは実在の物体を対象としており、対象の輪郭が線状に存在していることが前提となっている。このため、2. で述べた通り、点描画コンテンツに対してそのまま利用することができない。

そこで、はじめに簡易試作システムを構築して基礎的な実験を行うことにより、具体的な課題を集約することとした。紙面の都合上、この詳細は割愛するが、この作業の結論として、2. で述べた (II) (III) については既存技術の応用である程度の解決が図れることが明らかになる

反面（ただし改良の余地は大いにある）(1)に関連して以下3つの課題が抽出された。

人間の目には残像効果があるが、カメラには残像効果がない。このため、Aerial 3D Display から出力された点描画は、人間には点描画全体が“ある図形”として捉えることができるものの、カメラの各フレームにはその一部分しか映らない(図1)。

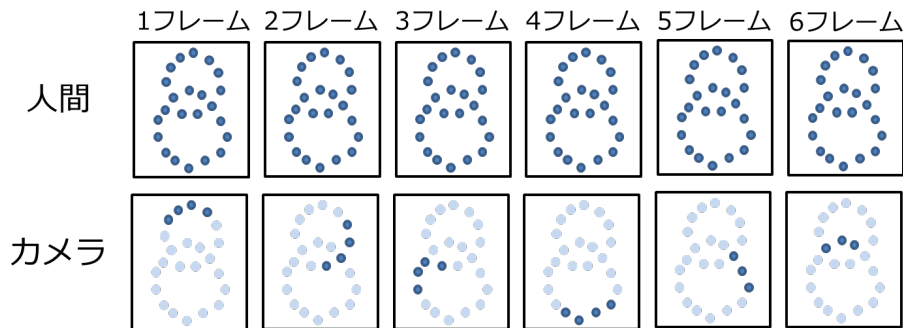


図1 カメラには残像効果がないことによる影響

点描画とカメラの周波数差により、点描画の「飛び」が発生してしまう(図2)。ここで言う「飛び」とは、正しく空中に描画された点描画コンテンツの点のいくつか、カメラで撮影されたビデオのどのフレームにも映らない現象を表す。どの点が「飛び」となるのかは毎回(周期毎に)変わるため、「飛び」問題を解決する最も簡単な方法は、多数のフレームを重畳し、点描画全体が映る画像を作り出すことである。しかしながら、この手法はカメラの手振れの影響により困難であることが明らかになっている(図3)。

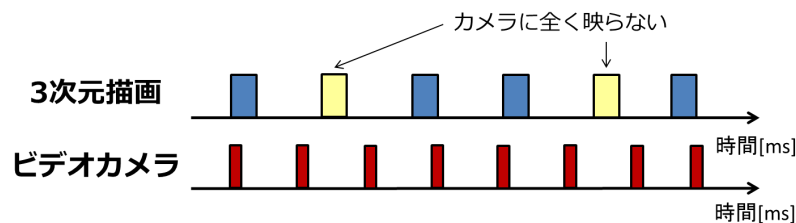


図2 「飛び」の発生



図3 手振れの影響

カメラ撮影時に焦点ぼけが発生してしまう。これは、Aerial 3D Display によって生成された点描画は、「点はあるが面はない」かつ「光はあるがモノ(物理的実体)はない」という性質を持つため、既存カメラでは焦点合わせがうまく機能しないためである(図4)。

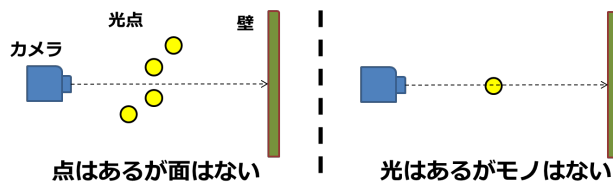


図4 焦点ぼけの発生

以下、これらの問題を解決するための2つの手法を提案する。

(2)提案手法1(画像処理的手法)

提案手法1は、描画周波数の推定、線描画の生成、線描画のマッチングの3段階から構成される。具体的には、動画入力に対して最初の数フレームからまず図形全体描画の周波数を求め、その後1フレームごとに点描画の最小全域木から線描画を作り、図形が一周するフレーム分の線描画を足し合わせることで図形全体の線描画を完成させる。そして最後にCRN(Cross Ratio Number)を用いて線描画のマッチングを行う。CRNとは研究代表者らが開発したピクトグラムマッチングの手法であり、非常に大きな射影変化に対しても高いマッチ率を実現することが可能な手法である。以下、各段階での処理について順を追って説明する。

描画周波数の推定

第 1 段階として、図形全体が十分映る、任意の数のフレームを取り出し、描画された図形全体を得るのに最低限必要なフレーム数を推定する。まず、1 フレームごとの画像を高い輝度値で二値化し光点を抽出する。これにより図 5(b)のような画像が得られる。この時、光点の面積が一定値以上ならば焦点ぼけによって複数の点が連結しているものと判定する。この場合、連結している光点の凸包をとって最も距離の遠い 2 つの頂点と重心の 3 点を 3 つの光点として扱う。光点抽出によって 1 フレームにつき複数の光点が多くなるので光点の座標の平均値を求める。この値を時間軸上にプロットし、フーリエ変換を行うことでピーク周波数を求め、ここから線描画を生成するために必要なフレーム数を決定する。

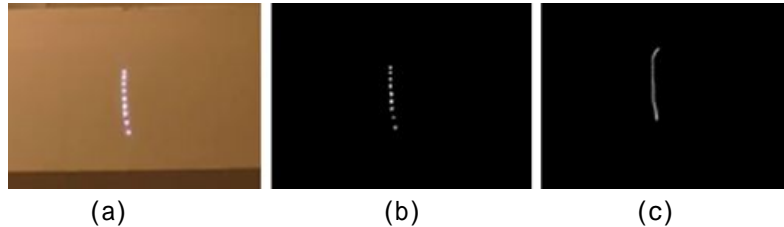


図 5 フレームごとの線描画生成

(a)入力画像 (b)抽出した光点 (c)最小全域木による線分

線描画の生成

第 2 段階として、(2-1)で得られたフレーム数をもとに線描画を生成する。まず、(2-1)と同様に焦点ぼけを考慮した発光点抽出を行う。次に発光点を結ぶ最小全域木を生成する。発光点は、多くの場合、線状に並ぶため 1 フレームごとの最小全域木は図 5(c)に示すような線分の形になる。そして、(2-1)で得たフレーム数だけ最小全域木を重ね合わせる。この時、隣接するフレームの発光点のうち最も距離の近い 2 点を選んで接続する。こうして、描画の欠損を補間した線描画が生成される。

線描画のマッチング

最後に、第 3 段階として生成した線描画を使ってマッチングを行う。これには CRN を用いる。CRN は、画像内の情報が少なく SIFT 等で特徴点の取りづらいピクトグラム等向けの高速高精度なマッチングのために提案された射影変換にも頑強な手法である。本手法で点描画から生成される線描画もまた画像内の情報が少なく、更に空中の点描画を地上から撮影する都合上射影変換に強いマッチング手法が必要となるため、CRN は本システムに適していると言える。

(3)提案手法 2 (深層学習を用いた手法)

(2)の提案手法 1 では画像処理的手法での問題解決を試みた。この提案手法 1 によってある程度の性能が実現できることは次の 4. で示すが、さらなる性能向上のため、深層学習を用いた手法についても検討を行った。

深層学習を用いるにあたっては、

- ・入力データを何とするか？(前処理を含む)
- ・どのような深層学習ネットワークを用いるか？

などの議論が必要となるが、紙面の都合上、結論のみを記載する。

まず、入力データは以下の手順で作成することとした。

手順 1: 図 5(b)と同様に撮影された点描画ビデオの各フレームを二値化し、光点を抽出する。
手順 2: 二値化された各フレームを n 枚重畳した重畳画像を作る。この重畳画像を深層学習の入力データとする(n の値は 5, 10, 15 とした)。

また、使用した深層学習ネットワークであるが、現在画像認識(画像分類)分野で広く使用されている Inception3 をそのまま利用することとした。

4. 研究成果

3. で述べた提案手法1(画像処理的手法)と提案手法2(深層学習を用いた手法)の性能評価結果を以下に示す。

(1)評価手法

Aerial 3D Display によって出力された点描画を iPhone 6 で撮影した動画を実験データとして用いた。描画される図形は図6の4種類(曇、雪、晴れ、雨)とし、それぞれの図形に対し(約)180秒の動画を100本ずつ用意した。

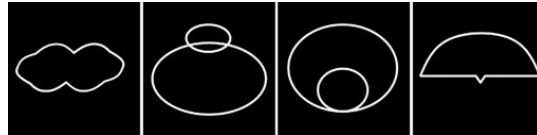


図6 実験で用いた図形

(2)評価結果

提案手法1によって再構成された図形の一列を図7に示す。この結果より、おおむね正しく元の図形が復元できていることがわかる。

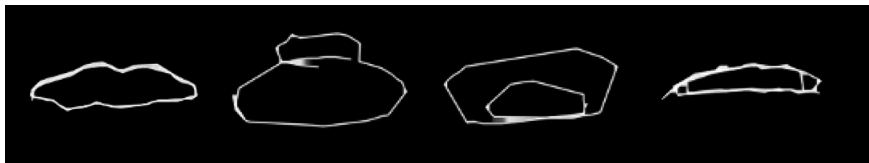


図7 提案手法1により再構成された図形の一列

しかしながら、実際にCRNを用いて形状認識を行ってみると、特に曇りと雨の誤認識が多く発生してしまい、全体としての認識精度は68%止まりであった。人間の感覚的にはおおむね正しく図形が復元できていることから、CRNに代わるマッチング手法を導入することにより認識性能が向上する可能性はあるのかもしれない。このあたりは今後の検討課題である。

また、提案手法2については、 $n = 5, 10, 15$ の3種類の重畳画像を作成し、これらを統合することにより各図形毎に(約)4400枚ずつの実験用画像を準備した。このうち、4000枚を学習データ、400枚をテストデータとして使用して機械学習を行った。この結果、認識性能は92.0%まで向上させることができた。この性能はほとんどのアプリケーションに対して実用レベルの性能であると言える。

5. 主な発表論文等

[雑誌論文](計7件)

Vicky Sintunata, Kurumi KAMINISHI and Terumasa AOKI, “Skewness Map: Estimating Object Orientation for High Speed 3D Object Retrieval System”, International Journal of Intelligent Engineering Informatics, vol.6, issue.1-2, pp.44-60, 2018. (査読有)

Terumasa AOKI and Van.Ng, “Global Distribution Adjustment and Nonlinear Feature Transformation for Automatic Colorization”, Advances in Multimedia, vol.2018, pp.1 – 15, 2018. (査読有)

Vicky Sintunata, Kurumi Kaminishi and Terumasa Aoki, “Skewness Map: Estimating Object Orientation for High Speed 3D Object Retrieval System”, International Journal of Intelligent Engineering Informatics, no.1, 2018. (査読有)

青木 輝勝, “8Kスーパーハイビジョン放送を支えるメディア伝送技術 -8K時代の伝送と信号処理- : 6. 8K/MMT時代のプライバシー保護技術 -ブラー不変マップと定点観測カメラシステムへの応用-”, 情報処理,58(2),pp.120-123, 2018. (査読有)

Qing Tong and Terumasa AOKI, “A Blur-Invariant Local Feature Descriptor for Gaussian and Motion Blurred Image Matching”, International Journal of Computer Science (IJCS), Vol.5, Issue 10, pp.150 – 161, 2017. (査読有)

Vicky Sintunata and Terumasa AOKI, “Skeletonization in Natural Image using Delaunay Triangulation”, Advances in Science, Technology and Engineering Systems Journal (ASTESJ), vol.2, no.3, pp.1013-1018, 2017. (査読有)

Qing Tong and Terumasa AOKI, “A Blur-Invariant Interest Point Detector Based on Moment Symmetry for Gaussian and Motion Blurred Image Matching”, International Journal of Signal Processing vol.2, pp.96-106, 2017. (査読有)

[学会発表](計21件)

Terumasa AOKI and Kurumi KAMINISHI, “A Local Descriptor for High-speed and

High-performance Pictogram Matching”, IEEE International Conference on Image Processing (IEEE ICIP2017), Sept., 2017. (査読有)

Junjie Hu and Terumasa AOKI, “Non-rigid Structure from Motion via Sparse Self-Expressive Representation”, IEEE International Conference on Image Processing (IEEE ICIP2017), Sept., 2017. (査読有)

Wang Tianqi and Terumasa AOKI, “Multicolor removal based on Color Lines and Improved Hough Transform for SFS”, IEEE International Conference on Image Processing (IEEE ICIP2017), Sept., 2017. (査読有)

Qiang Tong and Terumasa AOKI, “A Novel Blur-Invariant Local Feature Scheme for Image Matching”, IEEE International Conference on Consumer Electronics Taiwan (IEEE ICCE-TW2017), June, 2017. (査読有)

Vicky Sintunata and Terumasa AOKI, “Grey-Scale Skeletonization using Delaunay Triangulation”, IEEE International Conference on Consumer Electronics Taiwan (IEEE ICCE-TW2017), June, 2017. (査読有)

Qiang Tong and Terumasa AOKI, “A Blur-Invariant Local Feature for Motion Blurred Image Matching”, International Conference on Digital Image Processing, May, 2017. (Best Presentation Award) (査読有)

Junjie Hu and Terumasa AOKI, “Improving the Performance of Non-rigid 3D Shape Recovery by Points Classification”, IAPR International Conference on Machine Vision Applications (IAPR MVA2017), May, 2017. (査読有)

Junjie Hu and Terumasa Aoki, “A Convex Approach for Non-rigid Structure from Motion via Sparse Representation”, International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP2017), Feb., 2017. (査読有)

Vicky Sintunata and Terumasa Aoki, “Skeleton Extraction in Cluttered Image Based on Delaunay Triangulation”, IEEE International Symposium on Multimedia (ISM), Dec., 2016. (査読有)

Wang Tianqi and Terumasa Aoki. “Multicolor removal based on Color Lines for SFS”, IEEE International Conference on Image Processing (IEEE ICIP2016), Sept., 2016. (査読有)

Junjie Hu and Terumasa Aoki, “Points classification for non-rigid structure from motion”, IEEE International Conference on Image Processing (IEEE ICIP2016), Sept., 2016. (査読有)

Ossi Hirvola, Timo Viitanen, Vicky Sintunata and Terumasa Aoki, “Improved Image Quality in Fast Inpainting with Omnidirectional Filling”, IEEE International Conference on Image, Vision and Computing (IEEE ICIVC2016), Aug., 2016. (査読有)

Qiang Tong and Terumasa Aoki, “Moment Symmetry: A novel method for interest point detection to match blurred and non-blurred images”, IEEE International Conference on Image, Vision and Computing (IEEE ICIVC2016), Aug., 2016. (査読有)

Nguyen Thi Huyen Van and Terumasa AOKI, “Enhanced Hemisphere Concept for Color Pixel Classification”, International Conference on Multimedia Systems and Signal Processing (ICMSSP2016), Aug., 2016. (査読有)

Vicky Sintunata and Terumasa Aoki, “High speed 3D object retrieval using skewness value”, IEEE Region 10 Symposium (IEEE TENSYPMP 2016), June, 2016. (査読有)

Vicky Sintunata and Terumasa Aoki, “3D object retrieval system using skewness database”, IEEE International Conference on Control, Decision and Information Technologies (IEEE CoDIT 2016), May, 2016. (査読有)

(他 5 件)

6 . 研究組織

(1)研究分担者

研究分担者氏名：

ローマ字氏名：

所属研究機関名：

部局名：

職名：

研究者番号 (8 桁):

(2)研究協力者

研究協力者氏名：

ローマ字氏名：