

科学研究費助成事業 研究成果報告書

令和元年6月11日現在

機関番号：15501

研究種目：基盤研究(C) (一般)

研究期間：2013～2018

課題番号：25330197

研究課題名(和文)高精度両眼立体視手法の高速化および応用に関する研究

研究課題名(英文) development of a fast precise stereo disparity estimation and its applications

研究代表者

水上 嘉樹 (Mizukami, Yoshiki)

山口大学・大学院創成科学研究科・准教授

研究者番号：60322252

交付決定額(研究期間全体)：(直接経費) 3,800,000円

研究成果の概要(和文)：ステレオ画像対に基づいて高速かつ高精度な距離測定を実現することを目的として、本研究ではコストボリュームの傾斜変換を行う視差推定手法を提案した。近年、高精度な視差推定は誤差伝播法またはグラフ理論に基づいた大局アプローチにより可能であることが報告されているが、リアルタイム性の観点から局所並列演算で実装可能な局所アプローチに注目した。一般に、局所アプローチはテクスチャが弱い領域や物体表面がカメラ方向に対して傾いている場合には視差推定精度が劣化する。そこで、提案手法では物体の傾きを想定して、コストボリュームの傾斜変換を行った後に、エッジ保存フィルタ、または、水平および垂直方向の積分フィルタを適用する。

研究成果の学術的意義や社会的意義

交通事故を減らすための手段として安全運転支援機能の充実が望まれている。自動車周囲の距離情報を取得するためには、距離センサを用いる方式とステレオカメラ(画像センサ)を用いる方式の2つが主に検討されているが、後者は得られる距離情報の密度と導入コストにおいて有利である。ステレオ画像対から距離情報を復元するために、視差推定を行う必要があるが、計算負荷と精度の間にはトレードオフがある。本研究では一般に計算負荷が小さいと考えられている局所アプローチに着目して、精度の改善を目的としてコストボリュームの傾斜変換を行うことを特徴とする視差推定手法を提案し、シミュレーションによって有効性を確認した。

研究成果の概要(英文)：This study discussed fast and accurate distance measurement based on a pair of stereo images, and proposed a new disparity estimation method which transforms the slope of the cost volume as a pre-processing for cost aggregation. Global approaches such as belief propagation and graph theory have been applied to obtain accurate disparity, but they require a heavy computational burden. We focused on local approaches which have an advantage of simple and fast computation for real-time applications and disadvantages in dealing with weak texture region or non-frontal region. To overcome these disadvantages, the proposed method transforms the slope of the cost volume and applied edge-preserving smoothing filters such as guided filter or line-wise integral filter on the cost volume. Experimental results showed that the slope transformation improves the accuracy of the estimated disparity.

研究分野：画像処理

キーワード：両眼立体視 視差推定 局所並列演算

様式 C-19、F-19-1、Z-19、CK-19（共通）

1. 研究開始当初の背景

両眼立体視とは、左右に設置したカメラで撮影した画像に含まれる視差に基づいて対象物までの距離を推定する技術である。左カメラで撮影した画像を基準画像、右カメラで撮影した画像を参照画像と呼んでおり、基準画像上の画素を参照画像上の画素に対応付けることによって相対的な位置である視差を求める。しかし、対象物の表面模様が希薄なウィークテクスチャ領域、基準画像には含まれているが前方物体の隠蔽により参照画像には含まれていないオクルージョン領域、物体表面の法線方向がカメラ方向に対して大きく傾いているノンフロンタル領域、照明条件や反射特性によって画像間で画素値が異なる領域などに起因する視差推定の困難性が指摘されており、多くの研究者が視差推定精度の改善のために試行錯誤を行っている。

本研究開始当初、既に本研究グループでは画素単位で推定された視差に基づいてサブピクセル精度の視差を推定するための技術を確立していた。当時、提案されていた両眼立体視手法の多くが画素単位の視差を推定することを目的としていたこともあり、提案手法によって推定されたサブピクセル視差は当時のベンチマークであったミドルベリー大学の評価データベース V2 にて 0.5 画素の許容誤差基準で首位にランクしていた。一方で、提案手法によるサブピクセル視差は既存の手法で得られた画素単位の視差の精度に依存するという性質があったために、更なる高精度化のためには画素単位の視差計算の精度を向上させる必要があった。

2. 研究の目的

本研究では、既に提案していたサブピクセル視差の推定手法を補強するための画素単位視差の推定手法の開発に取り組むこととした。特に、ウィークテクスチャ領域やノンフロンタル領域での視差精度を向上するために、物体表面の傾きを考慮した画像間の対応付けについて検討することとした。また、両眼立体視に関連する技術として、特徴点に基づいて平面部分の形状を推定するための手法、および、推定されたセンサからの距離情報に基づいて環境地図を構築するための手法についても開発および検討を行った。

3. 研究の方法

両眼立体視は、コスト計算、コスト平滑化、視差推定、後処理の 4 段階に分けて実施されることが多い。コスト計算では両画像に含まれる視差を想定した場合のコスト分布を水平、垂直、視差からなる 3 次元の直方体（コストボリューム）として表現する。コストは基準画像上の画素値に対応する参照画像上の画素値との差異から計算されるが、従来、水平垂直方向に局所的な積分計算を行うことで耐雑音性を高めていた。しかし、2015 年に J. Zbontar らが提案した畳み込みニューラルネットワークを用いたコスト算出手法は秀逸であり、その後の多くの研究者が同グループによって提案されたコスト算出手法を取り入れるようになった。

コスト平滑化においてはより積極的な雑音除去が試みられる。複数の形状の平滑化フィルタを用いることが検討されていたが、2005 年に K. J. Yoon らによってエッジ保存フィルタであるバイラテラルフィルタが適用され、さらには 2010 年に K. He らによって提案されたガイドドフィルタの適用によって画像中の物体境界を保存したコストボリューム平滑化の有効性が広く知られるようになった。

視差推定は、大局アプローチと局所アプローチに大きく分類される。大局アプローチは適切な評価関数を設定し、これを最小化するような視差関数を求めようとする。評価関数は画像間の対応誤差を表すデータ項と物体境界以外での不連続視差にペナルティを与える制約項によって構成される。代表的な大局アプローチは確率伝播法とグラフ理論である。前者は全ての状態確率を保持するために多くの計算機資源を要することと、反復処理のために計算時間が長くなる傾向がある。後者は、枝切りによって効率的な計算を実現することが可能である一方、離散状態しか扱えないためにサブピクセルでの視差推定ができないという欠点があった。しかし、2017 年に T. Tanai らの研究グループによって連続的な視差推定が実装可能であることが示されて以降高い注目を集めている。局所アプローチの長所はその高速性にある。平滑化されたコストボリュームの視差方向に最小のコストを与える視差を検出する所謂ウィナー・テーク・オール手法(WTA)だけで視差推定を完了するために簡潔に実装することができる。なお、大局アプローチの利点である広い範囲での整合性の配慮するために、上下、左右（さらには斜め）方向の不連続性にペナルティを課したコスト積分を行うセミ・グローバルマッチング手法(SGM)が H. Hirschmuller によって提案されている（2007 年）。SGM は大局的な情報を配慮しながら精度を向上させるだけでなく、反復計算を必要としないために優れた高速性を有している。そのために、効率性を重視する産業的な用途でも使われることが多い。

本研究では、高速性の観点から局所アプローチに基づいた両眼立体視手法の開発に取り組んだ。局所アプローチは本質的にノンフロンタル領域およびウィークテクスチャ領域での視差推定が苦手である。その主な理由は、物体表面の法線が視線方向に直交していることを仮定して、コストボリュームの水平垂直方向でコストの平滑化を行うことが挙げられる。解決策の一つは、水平垂直方向から傾いた平面上で平滑化計算を行うことであるが、複雑性および効率性の観点から実装上の問題がある。そこで、本研究では、視差の傾斜を許容する簡潔な平滑化処理の導

入が必要との視点に立ち、対象物表面の上下方向、または、左右方向の傾斜を仮定して、コストボリュームの傾斜変換を行った上で平滑化計算を行うことの妥当性について検討することとした。

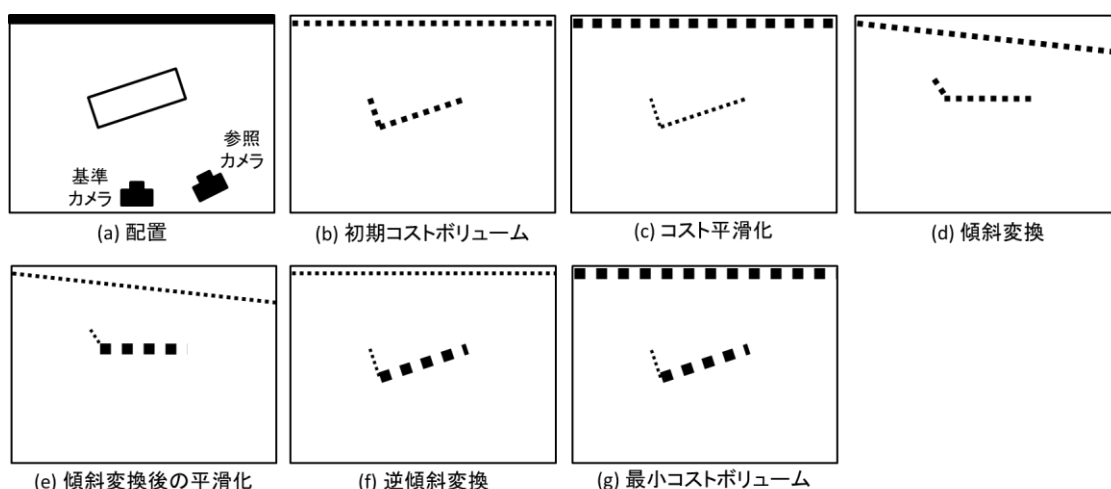


図 1. 傾斜変換を行うコストボリューム平滑化

図 1 に傾斜変換を行うコストボリューム平滑化の概要を示す。図(a)に被写体である情景と基準カメラ、参照カメラの位置を垂直方向から記述する。ここでは、背景の法線は視線方向と一致しているが、配置された物体表面の法線は傾いている。図(b)は基準画像の水平垂直座標に基づいて、奥行方向を視差とする初期コストボリュームである。なお、コストは小さいほど望ましいが、分かりやすさのために図の説明においてはコストが小さくなるほど点線を太くしている。前述の背景に相当する視差は水平垂直平面に分布する一方で、物体表面に相当する視差は水平垂直平面から傾いて分布している。コストボリュームに対して平滑化処理を行った場合には、背景に相当する視差のコストはさらに小さくなるために、ノイズを多く含む領域やウィークテクスチャ領域に対してより頑強な視差推定が可能になる。一方、傾斜を有する物体表面の視差コストは平滑化方向に合致しないために、間違った視差の有するコストの影響で大きくなってしまいう危険性がある(図(c))。そこで、水平方向の視差の傾斜を仮定して図(d)に示されるようにコストの傾斜変換を行う。この時、コストボリュームの全体形状が直方体ではなくなるが、以後の計算効率を配慮して部分的な転移を行うことで直方体形状を維持する。傾斜変換を行った後のコストボリュームに対して平滑化処理を行い(図(e))、平滑化後のコストボリュームを逆傾斜変換する(図(f))。本来、物体表面の傾斜は水平成分および垂直成分を同時に含むはずであるが、計算負荷の観点から水平成分のみを考慮した傾斜変換、または、垂直成分のみを考慮した傾斜変換を行うこととする。なお、傾斜の度合いを表す傾斜角の範囲や解像度についても実験的に調整する。異なる傾斜方向および傾斜角で得られた逆傾斜変換後のコストボリュームの同じ座標に位置するコストの最小値を求めることで最小コストボリュームを生成する(図(g))。得られた最小コストボリュームに基づいて、視差推定段階では視差方向に最小コストを有する視差が推定結果として選出される。

続いて、本研究で採用したコストボリュームの平滑化手法について言及する。ここでは、外デッドフィルタとラインスキャンの2つの方法について検討した。前者のガイドドフィルタはエッジ保存フィルタの一種であるために、物体の輪郭(エッジ)を考慮した平滑化が可能である。後者のラインスキャンとは、水平方向または垂直方向にコストを積分していく計算を表しており、閾値より大きな画素値の差異がある場所はエッジと見做して積分計算をリセットすることで、エッジを考慮した平滑化を実現する。また、積分されたコストを積分距離で割ることでコストの平均値を平滑化後の結果としてコストボリュームに保存している。水平方向にラインスキャンを行って得られた結果に対して、垂直方向にラインスキャンを行うことで平滑化を完了することとする。

最後に、両眼立体視に関連する研究として実施した3次元再構成手法、および、距離情報に基づいて環境地図を構築するための手法についてそれぞれ段落を分けて記述する。

[3次元再構成手法] 近年、複数画像に基づいた3次元再構成について多くの研究が行われている。特に、都市空間や屋内空間においては、被写体が平面で構成されることが多いことを踏まえると、空間内の平面領域を検出することは詳細な距離分布の推定に有用である。本研究では、Structure from Motion(SfM)で求めた疎らな3次元特徴点間の色分布に基づいて、基準カメラ画像上の複数の平面領域を選出する。次に、平面領域毎に平面方程式を解き、法線方向に平面スイープを行う。得られた視差空間画像からコストが最小となる視差候補を選出して視差画像を作成する。最後に、複数の視差画像をコストに基づいて合成し、最終的な視差画像を得る。

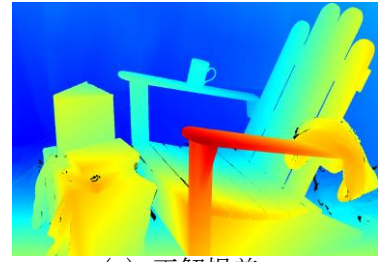
[環境地図構築手法] SLAMとは自己位置推定と環境地図作成を同時に行うことができる技術である。作成された環境地図に基づいて、地形や建築物の形状評価、自動車やロボットの経路探



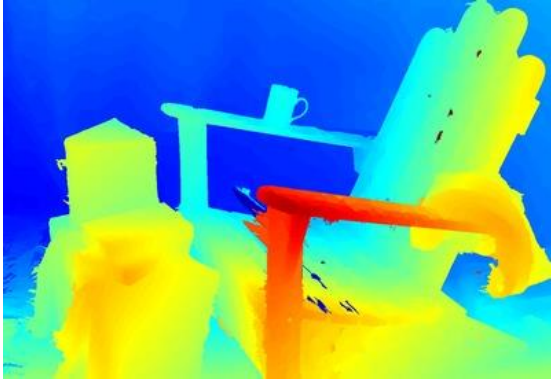
(a) 基準画像



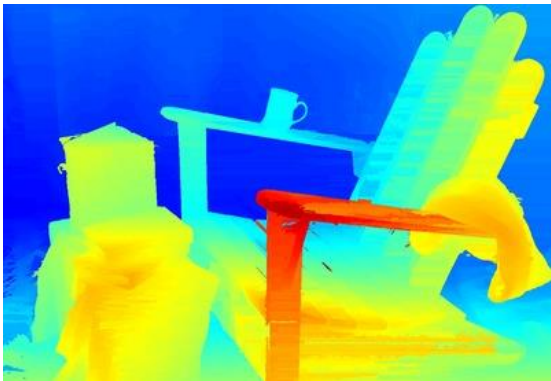
(b) 参照画像



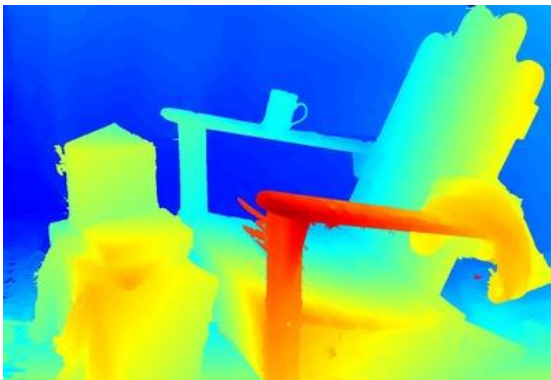
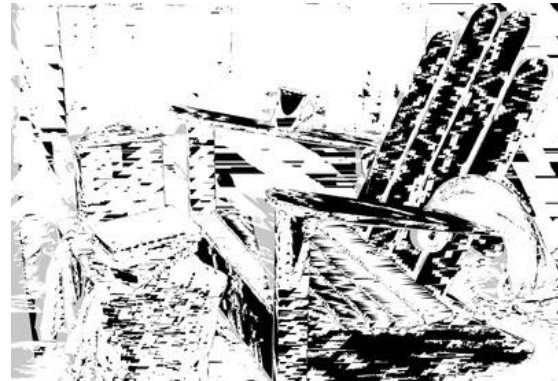
(c) 正解視差



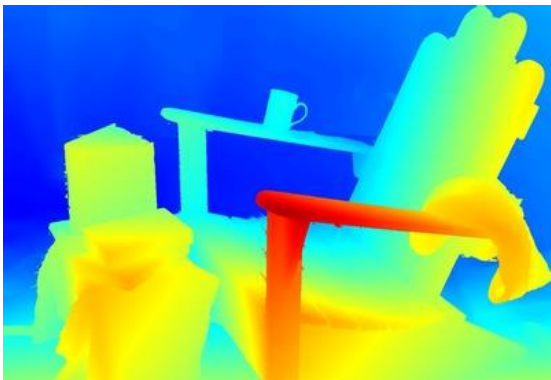
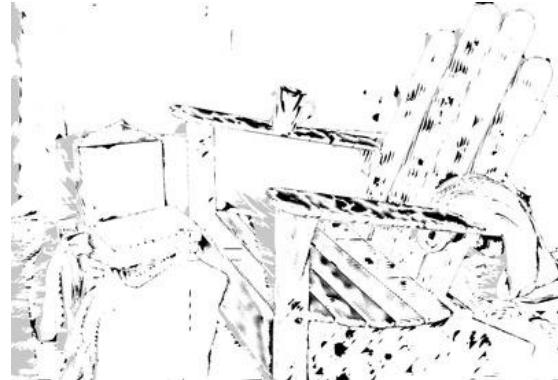
(d) 傾斜変換を用いない GF (推定視差と誤り視差領域)



(e) 傾斜変換を用いない LS (推定視差と誤り視差領域)



(f) 傾斜変換を用いた GF (推定視差と誤り視差領域)



(g) 傾斜変換を用いた LS (推定視差と誤り視差領域)

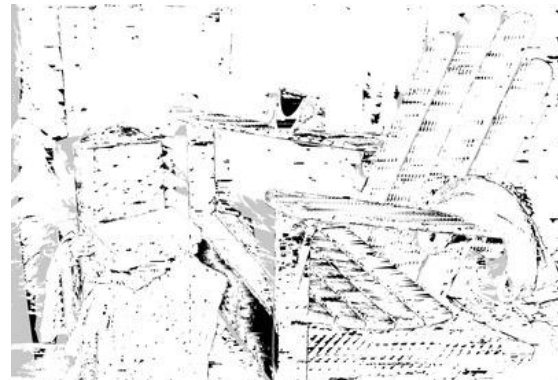


図 2 傾斜変換を前処理に採用した場合の推定視差の精度改善効果

索および自律移動を支援することができる。既に多くの SLAM 手法が提案されているが、大きく 2 次元 SLAM と 3 次元 SLAM に分類される。ここでは、Pomerleau ら (2013 年) によって開発された 3 次元 SLAM パッケージ ETHZASL-ICP-Mapper における計算効率性の改善に着目した。処理経過とともに効率性が低下する要因である全体点群の肥大化、および、点群密度制御における k-d tree の計算負荷を鑑み、前者に対して点群のグリッド管理、後者に対してボクセルグリッドの導入を行うこととする。

4. 研究成果

最初に、傾斜変換を行わなかった場合の視差推定結果を図 2 に提示する。ここでは、図 (a) に基準画像、図 (b) に参照画像、図 (c) に正解視差を示している。さらに、図 (d) または図 (e) ではガイドッドフィルタまたはラインスキャンで平滑化処理を行った視差推定結果をそれぞれ示している。図 (f) および (g) は傾斜変換を行った結果を示している。ガイドッドフィルタ (GF) およびラインスキャン (LS) を用いた結果のどちらにおいても被写体表面の傾斜部が良好に表現されているのが分かる。続いて、提案手法による視差推定精度を数値的に検証する。ミドルベリー大学の評価データベース V3 に含まれる訓練用データセット (画像セット数 15) に基づいて許容誤差基準を 2 画素とした場合の誤り領域の割合 (誤り領域比率) について述べる。ガイドッドフィルタ採用時、傾斜変換を用いない場合 20.5%であったのに対して、傾斜変換を用いた場合は 13.0%であった。ラインスキャン採用時、傾斜変換を用いない場合 22.4%であったのに対して、傾斜変換を用いた場合は 12.9%であった。

以上の結果から、本研究で提案した傾斜変換はコスト平滑化の前処理として有効であることが目視比較および数値に基づいた精度比較によって確認された。提案手法は局所アプローチに属するものであり、また導入した処理手順は全て局所並列演算で実装されるために比較的短い時間で実行することが可能であることも長所といえる。しかし、計算効率性を重視した SGM が提示している誤り領域比率 10.1%と比べると、提案手法によって得られた誤り領域比率には引き続きの改善の余地が認められる。

最後に、3次元再構成手法、および、環境地図構築手法について得られた研究成果についても概観する。

[3次元再構成手法] 本研究では、複数画像に基づいた 3次元再構成について検討した。対象物体が平面から構成されることを仮定し、平面を考慮した距離画像の生成手法の提案を行い実装した。最初に複数の視点画像から SfM 計算を行い特徴点の 3次元座標とカメラパラメータを導出した。次に特徴線分 (Key Line Segment; KLS) を導入することで、シーン中の複数の平面領域の抽出を行った。続いて、平面領域ごとに平面方程式を求め、その法線方向に平面スイープを行うことで視差候補を取得し、さらに抽出された平面領域以外の物体の視差を求めるために基準カメラの視線方向にも平面スイープを行った。最後に、複数の視差候補を統合することで最終的な視差画像を生成した。屋内ならびに屋外の風景に提案手法を適用し、対象物体の平面領域上で勾配に沿った滑らかな視差が取得できることを確認した。用いた撮影画像および視差推定結果の一例をそれぞれ図 3 および図 4 に示す。

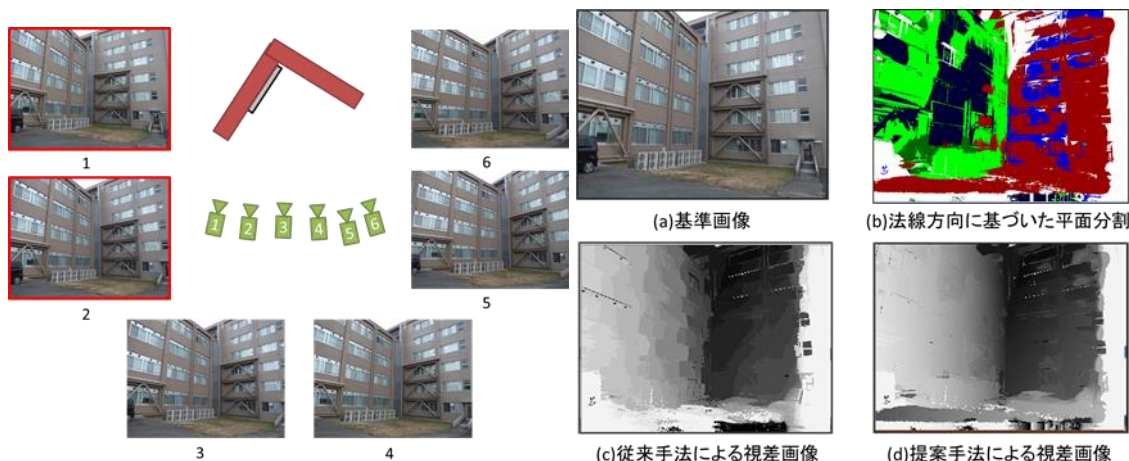
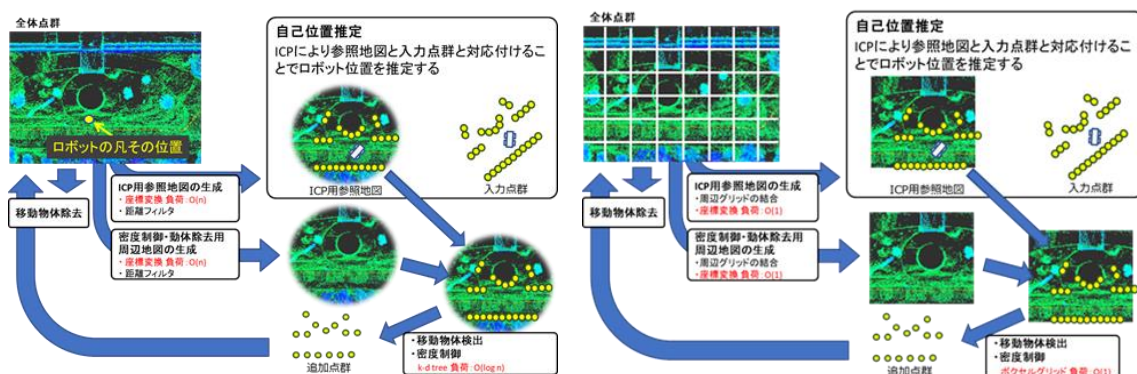


図 3 撮影画像 (建築物)

図 4 視差推定結果

[環境地図構築手法] 本研究では、チューリッヒ工科大学の研究グループによって開発された 3 次元 SLAM パッケージ ETHZASL-ICP-Mapper に着目して、計算効率、および、実行時の安定性に関する改善を行った。具体的には、環境地図を構成する点群をグリッド管理することによる点群追加効率の維持 (改良点 1) と密度制御における計算効率の改善 (改良点 2) について提案を行った。改善の効果を議論するために、作成される環境地図の点群規模を数値的な指標とし、実環境で得られた入力点群およびエンコード値の履歴に基づいて提案手法の有効性を議論した。その結果、改良点 1 によって、点群規模の増大に伴う追加点群の低減を回避することができる

ことを確認した.これにより SLAM の時間経過に依存しない安定な 3 次元 SLAM が可能となった.さらに,改良点 2 によって,点群密度制御を低負荷かつ効率的に実施できることを確認した.ETHZASL-ICP-Mapper の処理概要ならびに提案手法の処理概要を図 5 に示す.



(a) 従来手法 (b) 提案手法
 図 5 環境地図構築手法における従来手法と提案手法の処理概要

5. 主な発表論文等

[雑誌論文] (計 0 件)

[学会発表] (計 2 件)

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。