

一人称視点映像と3人称視点情報を融合した生活行動ナビゲーション

Life behavior navigation that integrates third and first-person point of view

梁 泳成*¹
Yongsong Ryang

土井 健輔*¹
Kensuke Doi

大井 翔*²
Sho Ooi

佐野 睦夫*¹
Mutsuo Sano

*¹ 大阪工業大学情報科学部

Faculty of Information Science and technology, Osaka Institute of technology

*² 大阪工業大学大学院情報科学専攻科

Mayor in Information Science and technology, Graduate School of Osaka Institute of technology

This paper proposes a method to navigate activities for daily livings (ADL) safely for the elder and cognitive disabled patients using both a first-person vision and a third-person vision. The proposed method can detect dangerous ADL elements from an ADL 3D map at hand (first-person vision) generated by a Structure from Motion (SfM) approach using wearable camera, which is complemented with an overall ADL 3D map (third-person vision) obtained by a Visual SLAM using RGB-D sensor mounted on an autonomous robot. We verify the effectiveness of the proposed method through the experiments of safely cooking navigation.

1. はじめに

日常生活行動のナビゲーションにおいて、3次元生活空間情報はユーザの生活を安心・安全な生活にするうえで必要不可欠である。生活空間のうち、キッチンなどの料理空間では、調理中にももの配置が変化するため、位置が安定せず、調理者は足元や火元での不注意を起こす危険性がある。したがって、調理者は常に料理空間全体の状態を把握する必要がある。

3次元空間情報の取得には、多視点画像より自己位置姿勢の推定および3次元空間を復元することが可能である SfM (Structure from Motion) や自己位置推定と環境地図作成を同時に行うことが可能である SLAM (Simultaneous Localization and Mapping) などのアルゴリズムが挙げられる [満上 11][SatoshiRobatoFujimoto 15]。そこで、今回3次元の生活空間の情報を再現するために、頭部搭載型のウェアラブルカメラを用いることで実現することが考えられる。特に、調理中の料理空間においては、一般的に手元での作業に注意し、調理を行う。そのため、ウェアラブルカメラなどの1人称視点では、手元の作業を3人称視点に比べ近くで観測することが可能であるので、3人称視点より手元の3次元空間情報を高解像度で取得することができる。

一方で、1人称視点では手元の作業に着目し、3次元空間の取得を行うため、1人称視点により料理空間全体を把握することは困難である。

そこで、本研究では、1人称視点を3人称視点により補完することで、調理者に調理空間における危険領域を検出し、認識させることが可能であると仮定した。3人称視点による3次元空間情報の取得では、1人称視点でのオクルージョンなどによる観測することができない空間情報に対して、ロボットを用いて補完する。1人称視点では観測できない空間情報を、ロボットが能動的に移動することで詳細な料理空間全体の状態を取得することが可能である。

1人称視点では、ウェアラブルカメラを用いて手元の3次元空間情報を取得することを想定するため、SfMアルゴリズムを採用し、3人称視点では、SLAM (Simultaneous Localization and Mapping) 技術を採用した。近年、SLAMでは、KinectなどのRGB-Dカメラを用いることでリアルタイムに自己位置推定と環境地図作成が行えるほか、PCL (Point Cloud Library) などの3次元点群を扱うライブラリを併用することが可能であるため、1人称視点映像より得られた位置情報をリアルタイムに更新することが可能である。しかし、1人称視点ではRGB-Dカメラを頭部に搭載することは困難であるので、SLAMではなくSfMアルゴリズムを採用した。

本研究では、1人称視点映像と3人称視点情報を統合することで、より容易かつ高精度な3次元空間情報の取得が行えると仮定し、調理環境における危険領域検出のためのセグメンテーションおよび特定物体認識、1人称視点映像による3次元空間情報の対応付けを行うための方式を提案する。

2. 提案方式

2.1 料理空間情報対応付けの枠組み

本研究では、1人称視点映像によって得られたデータと3人称視点映像によって得られたデータがあり、それぞれに対して3次元点群のデータを作成し、統合する必要がある。以下に各映像データから3次元点群の取得方法について説明する。

(1) 1人称視点映像

ウェアラブルカメラにより取得した映像は2次元の画像である。そこで、多視点から画像を構成した多視点画像系列を Visual SfM というツールを用いて対象空間の3次元点群データを作成する。この作成した3次元点群データをターゲットと定義した。

Visual SfMによる多視点画像系列より、3次元点群を取得する手順を以下に示した [山口 15]。

STEP1 画像群の入力

・「File>Open+Multi Images」を選択して、対象画像群を入力する。

STEP2 画像のマッチング

連絡先: 梁泳成, 大阪工業大学情報科学部, 大阪府枚方市北
山 1-79-1, 072-866-5189, elc13116@st.oit.ac.jp

・「SfM>Pairwise Matching>Compute Missing Match」を選択し、特徴点の抽出および画像間の対応点の探索を行う。なお、特徴点の抽出には、SIFT 特徴を用いる。

STEP3 3次元点群の復元

・「SfM>Reconstruct Sparse」を選択し、位置姿勢の推定および3次元点群の復元を行う。

STEP4 バンドルの調整

・「Run Bundle Adjustment」を選択し、バンドルを調整する。

STEP5 3次元点群結果の表示

・Visual SfM にて復元した3次元点群を ply 形式で保存し、保存した ply データを開く。

(2)3人称視点映像

3人称視点では RGB-D カメラから得られた映像データに対して、Visual SLAM を用いて3次元点群データを取得する。この作成した3次元点群データをソースと定義した。

また、49枚の画像から図5の結果を導くため、使用したPCのスペックは windows8.1/8GHz に対して処理時間は約420sである。

2.2 位置合わせ

2.1節で説明した2種類の3次元点を対応付けさせる必要がある。そこで、今回対応付けには、PCL(Point Cloud Library)を使用した。図1に対応付けのフローを示す。

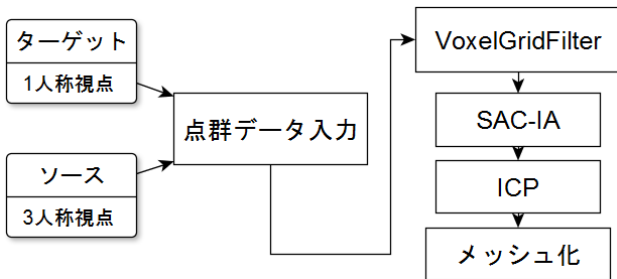


図1 各点群の対応付けの流れ

Visual SfM と Visual SLAM により取得したターゲットおよびソースの点群データを入力データとする。入力された各点群データに対して、VoxelGridFilter を適用する。初期の入力された点群データの密が高く、点の数は多いため、処理が重たくなってしまいう問題がある。各点群データに VoxelGridFilter を適用することで処理負荷の軽減を図る。フィルタを適用した各点群をまず SAC-IA(Sampled Consensus-Initial Alignment)アルゴリズムを用いて初期位置合わせを行う。SAC-IA により、おおまかな位置合わせを行うことで、ICP(Iterative Closest Point)による高精度な位置合わせを行うことが可能となる。ICP は最近傍点の計算、位置合わせの計算、位置合わせの結果で構成される [増田 9]。そのため、初期の最近傍点が遠くに存在する場合、位置合わせが正しく行われないので、おおまかな初期位置合わせを行う必要がある。ICP による位置合わせの結果をメッシュ化する。

2.3 危険領域検出の枠組み

調理中の料理空間での危険領域には、火元、割れ物や包丁の置き場所、足元の置物などが挙げられる。3人称視点による危険領域の検出には KinectV2 を用いる。本研究では、割れ物や包丁の置き場所、足元の置物に着目し、これらの危険領域を検出するための料理空間におけるセグメンテーションを

行う方式を検討した。具体的には、3人称視点によりリアルタイムに取得した点群データに対して、平面を検出する。検出した平面領域は削除し、平面以外の領域では、検出された平面と高さが異なる場合、点の赤色情報の値を最大にする。これにより床に落ちている危険な物体に対してセグメンテーションすることが可能である。図2は床に落ちている危険なオブジェクトに対してセグメンテーションを行っている結果であり、図2が(aセグメンテーションした結果)、図2(b)が元のキッチン空間である。



図2 オブジェクトのセグメンテーション

3. 実験・結果と考察

実験として、今回研究室内にあるキッチン空間を対象として3次元空間の作成と危険領域の検出を行った。

3.1 1人称視点映像による空間情報作成

1人称視点による手元の料理空間作成には、Visual SfM を用いて、3次元点群データを復元した。図3に元のキッチン空間の画像を示す。



図3 対象としたキッチン空間

対象としたキッチン空間に対し、図4に示すように49枚の多視点画像系列の映像情報より作成し、図5に Mesh Lab にて表示した結果を示す。



図4 キッチン空間の多視点画像の一部

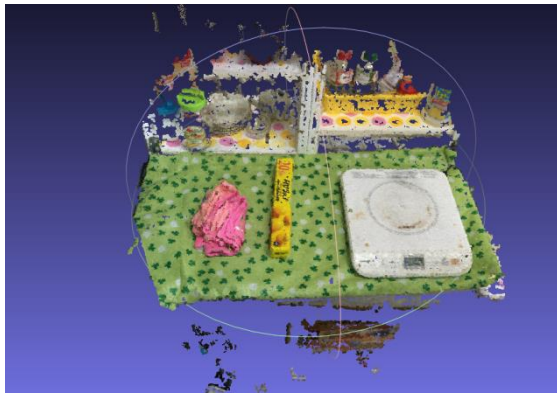


図 5 1 人称視点によって作成された料理空間

図 5 では、手元の料理作業空間以外の足元領域などの危険領域が再現されておらず、3 人称視点により補完する必要があることが確認できる。SfM では、1 人称視点映像に基づき料理空間の作成を行っているため、調理中では、手元の調理空間に集中しているため、足元などの空間映像情報を多数取得することが困難である。

3.2 足元領域の危険領域の検出

2.3 節で示したように、セグメンテーションし、危険領域を検出した画像と実際の画像を比較してみる。図 6 に足元危険領域を検出した例を示す。



図 6 3 人称視点による足元危険領域

ここで、図 6(b)に示した赤い領域に着目する。赤い領域に対して図 6(a)の結果をみると、物体の半分領域は危険領域の検出が行われているが、もう半分領域は検出が行われていないことが確認できる。赤枠で囲まれた赤枠で囲まれた対象物体の半分領域のみ検出された原因として、対象物体とカメラの距離が挙げられる。

次に、図 6(b)で示した黄色の領域に着目する。黄色の領域に対して図 6(a)の結果をみると、危険領域として検出されていないことが確認できる。赤枠と黄枠の対象物体が正しく危険領域として検出されなかった問題は、今後、ロボットによる能動視点により解決が行えることを検証する予定である。

3.3 キッチン上での危険領域

キッチン上での危険領域検出も 2.3 節で示した方式で検出した。図 7 にキッチン上の危険領域を検出した例を示す。



図 6 3 人称視点によるキッチン上の危険領域

図 6(a)は、セグメンテーションを行った結果であり、図 6(b)はセグメンテーションを行った画像の視点を右側から見た視点の対象空間の元画像である。キッチン上の割れ物である瓶や包丁に対して、足元領域の危険領域と同様に、危険領域として赤色で検出されていることが確認できる。一方、図 6(b)の緑枠で囲まれた鍋の物体に着目する。鍋の取っ手となる領域では正しく危険領域として検出が行えていることが確認できるが、鍋の金属部分の領域では大半の領域が危険領域として検出されていないことが確認できる。RGB-D カメラで金属光沢を持つ物体を対象とする場合、金属は赤外線レーザを反射するため、点群の取得が困難である。また、同画像の緑枠で囲まれた雑誌に着目する。この対象物体も足元領域の危険領域と同様に対象物体領域の半分領域のみ危険領域として検出されていることが確認できる。

4. まとめと今後の課題

本稿では、日常生活空間における、1 人称視点映像により取得した空間情報を 3 人称視点情報で補完することで、危険領域を検出し、人の生活を安心・安全な生活とすることに着目した。1 人称視点情報のみでは観測できない情報を 3 人称で補完する方式を提案した。今後の課題としては、危険領域をセグメンテーションのみで検出するのではなく、一般または特定物体認識を行う必要がある。また、3 人称視点を補完するうえでロボットが能動的に視点を変更する必要がある。

謝辞

本研究の一部は文部科学省研究費補助金(基盤 C15K00368)の支援を受けた。

参考文献

- [中村 15] 中村薫, 杉浦司, 高田智広, 上田智章: KINECT for Windows SDK プログラミング V2 センサー対応版, 株式会社 秀和システム, (2015).
- [満上 11] 満上 育久: Instruction of "Bundler", http://www.ite.or.jp/data/journal/passed_issues/tool1104/ (accessed 2016.3.29).
- [増田 9] 増田 健: ICP アルゴリズム, 情報処理学会研究報告 IPSJ SIG Technical Report Vol.2009-CVIM-168 No.23, 情報処理学会(IPSJ) (2009).
- [山口 15] 山口 悠志: VisualSfM と CloudCompare によるドローン撮影画像を用いたデジタル地図作成, http://www.slideshare.net/HiroshiYamaguchi5/visualsfmmes_hlabcloudcompare-ccbby (accessed 2016.3.29).
- [SatoshiRobatoFujimoto 15] SatoshiRobatoFujimoto: TUM の RGB-D SLAM についてまとめてみる, <http://qiita.com/SatoshiRobatoFujimoto/items/1e4cf330b92d-f43afafe> (accessed 2016.3.31).
- [Jochen Sprickerhof 12] Jochen Sprikerhof: PCL::Registration, <http://www.pointclouds.org/assets/icra2012/registration.pdf> (accessed 2016.3.31).