

全方位カメラと3D-LIDARを用いた 自由視点画像による隠消現実感

大石 圭^{1,a)} 森 尚平^{1,b)} 齋藤 英雄^{1,c)}

概要：近年、拡張現実感や複合現実感の実利用とともに、隠消現実感の研究も盛んに行われている。遮蔽物の裏側に隠背景観測用カメラを設置し死角を可視化する従来の隠消現実感では、カメラの位置関係上隠背景観測カメラ1台に対する透視範囲が狭まるといった問題が生じた。そこで本研究では、全方位カメラと3D-LIDARを組み合わせることで広範囲の自由視点画像を生成し、それを体験者の視点に重畳させることで死角を透視する隠消現実感を実現した。

キーワード：隠消現実感, 自由視点画像, 全方位カメラ

1. はじめに

近年、現実世界にCGオブジェクトを重畳させる拡張現実感 (Augmented Reality; AR) や複合現実感 (Mixed Reality; MR) の実利用とともに、現実世界から実物体を視覚的に消し去る隠消現実感 (Diminished Reality; DR) 技術の研究も盛んに行われている。隠消現実感を実現させるために主な手法として、現実世界の物体に仮想物体を重畳させ物体を消す手法 [1] や物体に隠背景を重畳させ消す手法がある。また、隠背景を重畳させるの手法にもインペインティング [2] と呼ばれる画像修復の技術を用いて隠背景を推定する手法と、別のカメラで観測した隠背景を体験者の視点に変換し重畳する手法がある。隠消現実の技術を用いて死角等を観測したいと考えた時、後者の手法が最も適当である。

Shen らの研究 [3] では体験者視点の映像のみを使い、時間的に変化する画像から対象物体の隠背景を推定した。Kameda ら [4] は監視カメラを隠背景観測用カメラとして使用し、その映像を建物等に重畳させることによって建物によって生じる死角の可視化を行った。視点変換は十分に距離が離れていることを利用し射影変換が用いられている。この研究ではCADモデルを体験者視点の位置姿勢推定や隠背景に利用しているが、その生成には時間を要するため任意の環境において現実的ではない。Barnum ら [5] は移

動物体に対して体験者視点カメラと隠背景観測用カメラの2次元射影の関係を用いて変換を施すことで、移動物体に対してもより自然な隠背景の可視化を行った。またこれも視点変換には射影変換を用いている。

上記の隠背景観測型の隠消現実感では隠背景に対して2つの問題がある。それは背景が平面に近似できる場合にしかきれいな隠背景が生成できないこと、また、多くの場合隠背景観測用カメラは体験者視点に対し前方に置く必要があるため、カメラ1台に対する透視範囲が狭まる傾向にあることである。前者の問題に対しては、Hosokawa ら [6] のように隠背景の3次元再構成を行い体験者視点に重畳させる手法がある。後者の問題に対しては、透視範囲を拡大させるためにカメラを複数台使用する手法が考えられる。しかし、この手法は得られる効果以上に装置が大掛かりなものになってしまうといった問題がある。またKameda ら [4] のように3次元モデルを用いそれを重畳させることで透視範囲の拡大を狙う手法もあるが、隠背景が変化する場面には対処できないという問題がある。そこで本研究ではカメラ1台に対する透視範囲を拡大させるため、隠背景観測用カメラに全方位カメラを用いる。また非平面へも対応するため広範囲の距離情報を取得できる3D-LIDARと組み合わせ隠背景の3次元再構成を行うことで自由視点画像を生成し、それを体験者視点へ重畳する手法を提案する。また、従来手法と比較することで、その効果を確認する。

2. 提案手法

提案手法では、隠背景観測用カメラは透過する対象に対して固定されているものとする。毎フレームに対し、隠背

¹ 慶應義塾大学
Keio University, Japan
a) oishi@hvrl.ics.keio.ac.jp
b) mori@hvrl.ics.keio.ac.jp
c) saito@hvrl.ics.keio.ac.jp

景の3次元再構成を行い、隠背景観測用カメラに対する体験者視点カメラの位置を推定する。その位置姿勢推定に合わせてレンダリングされた隠背景を、体験者視点の映像に重畳させることで、実物体の透視映像を生成する。また事前処理として隠背景観測用カメラと体験者視点カメラの内部パラメータの推定と、隠背景観測で用いる全方位カメラと3D-LIDARの位置姿勢推定を行う。

2.1 カメラの内部パラメータ推定

多くのカメラの場合いわゆるピンホールカメラモデルを取り扱う場合が多いが、全方位カメラは撮像手法が一般的なピンホールカメラとは異なるためそれに当てはまらない。したがってこの節では、ピンホールカメラの概要については省略し、全方位カメラの内部パラメータの推定手法のみ述べる。多くの全方位カメラの射影方式には画面の中心からの距離と角度が比例する等距離射影が用いられる。等距離射影は画面中心からの距離を y 、焦点距離を f 、入射角を θ としたとき式 (1) のように表される。

$$y = f\theta \quad (1)$$

しかし、実際の全方位カメラは理想的な射影方式には従わない。したがって、本研究では Scaramuzza ら [7] の提案する全方位カメラモデル (図 1) を用いて内部パラメータを推定する。ある3次元ベクトル $\mathbf{P} = [X \ Y \ Z]^T$ と画像座標 $\mathbf{p} = [u \ v]^T$ は式 (2) の関係にある。

$$\mathbf{P} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \approx \begin{bmatrix} u - u_0 \\ v - v_0 \\ f(\rho) \end{bmatrix} \quad (2)$$

ここで、 $\rho = \sqrt{(u - u_0)^2 + (v - v_0)^2}$ は画像座標 $\mathbf{p} = [u \ v]^T$ と画像の中心 $\mathbf{p}_0 = [u_0 \ v_0]^T$ の距離である。また、 $f(\rho)$ は ρ の多項式であり、式 (3) のように表される。

$$f(\rho) = a_0 + a_1\rho + a_2\rho^2 + a_3\rho^3 + a_4\rho^4 + \dots \quad (3)$$

本研究では4次の項までの係数を内部パラメータとして用

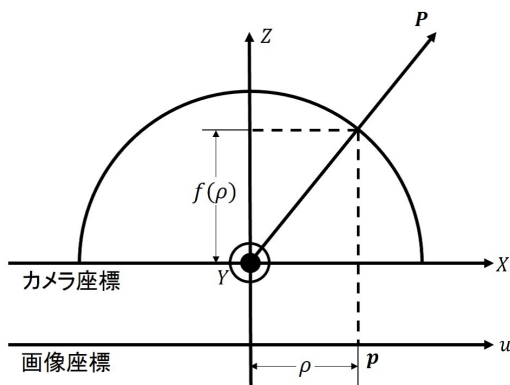


図 1: 全方位カメラのモデル

いる。したがって内部パラメータ \mathbf{A} は式 (4) のように表される。

$$\mathbf{A} = [a_0 \ a_1 \ a_2 \ a_3 \ a_4 \ u_0 \ v_0]^T \quad (4)$$

本研究では、チェスボードの交点の画像座標と3次元位置を用いて、その投影誤差が最小になるような式 (4) で表される内部パラメータを推定した。

2.2 3D-LIDAR に対する全方位カメラの位置姿勢推定

2.2.1 位置姿勢推定手法

本研究で使用した3D-LIDAR(YVT-X002:HOKUYO)はデバイスを中心とした3次元位置を観測するが、その走査範囲に対し取得点群が少ないため、反射強度等点群の3次元位置以外の情報を使うのが難しい。したがって、本研究では平面の法線と平面の点群に注目した位置姿勢推定手法である Pandey ら [8] の手法を用いて3D-LIDAR に対する全方位カメラの位置姿勢推定を行う。

まず、チェスボードを使用し目的平面の法線と距離を算出する。チェスボードの交点の画像位置と3次元位置を用いて PnP 問題を解くことでカメラの目的平面に対する回転行列 \mathbf{R}_c と平行移動ベクトル \mathbf{t}_c が定まる。また世界座標に対し目的平面を $Z = 0$ とすれば、回転行列 $\mathbf{R}_c = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3]$ における \mathbf{r}_3 は Z 軸に対する変換として考えられるため、目的平面の単位法線ベクトルになる。また、移動ベクトル \mathbf{t} の \mathbf{r}_3 に対する正射影はカメラと目的平面までの距離になる。その正射影ベクトル \mathbf{N}_c は内積を用いて式 (5) のように表される。

$$\mathbf{N}_c = (\mathbf{r}_3 \cdot \mathbf{t}_c)\mathbf{r}_3 \quad (5)$$

また目的平面からカメラまでの距離は式 (6) のように表される。

$$\|\mathbf{N}_c\| = \mathbf{r}_3 \cdot \mathbf{t}_c \quad (6)$$

次に上で求められた目的平面の法線とカメラと目的平面までの距離と、3D-LIDAR から得られる目的平面の点群から3D-LIDAR に対する全方位カメラの位置姿勢推定を行う。3D-LIDAR に対する全方位カメラの回転行列 \mathbf{R} と平行移動ベクトル \mathbf{t} は常に固定されており、各場面において変化しないものとする。目的平面 i 上に存在する3次元点 \mathbf{P}_l^j は、 \mathbf{R} と \mathbf{t} を用いて式 (7) によってカメラ視点の3次元点 \mathbf{P}_c^j に変換される。

$$\mathbf{P}_c^j = \mathbf{R}\mathbf{P}_l^j + \mathbf{t} \quad (7)$$

このとき、目的平面の上の点 \mathbf{P}_c^j に対する単位法線ベクトルへの射影と、カメラと目的平面までの距離の差が最小となればよい。したがって式 (8) を最小にする \mathbf{R} と \mathbf{t} を求める。

$$F = \sum_{i=1}^m \sum_{j=1}^n \left(\frac{\mathbf{N}_c^i}{\|\mathbf{N}_c^i\|} \cdot (\mathbf{R}\mathbf{P}_l^j + \mathbf{t}) - \|\mathbf{N}_c^i\| \right)^2 \quad (8)$$

また、本研究では Levenberg Marquardt algorithm [9] を用いて式 (8) を最小化した。

2.2.2 平面抽出の仕方と必要枚数

目的平面の点群の抽出方法について述べる、まず固定された場面にに対して 3D-LIDAR を用いて点群を取得する。次に平面を設置して再度点群を取得し、その差を比較することで平面と思われる点群を取得する。精度を高めるために RANSAC アルゴリズムを用いて得られた点群のなかから正しいと思われる点群だけを取得する。具体的な手法としては、ランダムに選んだ 3 点から平面の式を算出し、すべての点に対しその平面との距離を比較する。そしてその距離がある閾値以下の数を記録し、もっともそのスコアが高かった時の点群を平面の点群とした。

次に平面の必要枚数について述べる。本研究で用いた手法では 3 枚以上の平面を必要とする。仮に 1 枚だけの平面を用いた場合、位置姿勢推定に平面の法線と距離の情報しか用いらなかったため、ある平面上でカメラ位置が収束しない。また 2 枚の場合においてもある 3 次元位置で平面が交わる線上においてカメラ位置が収束しない。したがってカメラ位置を収束させるためには 3 枚以上必要である。

2.3 隠背景観測用カメラに対する体験者視点カメラの位置姿勢推定

隠背景観測用カメラと体験者視点カメラの位置姿勢推定を行うための模式図を図 2 に示す。M は位置姿勢を表す 4 × 4 の行列である。

隠背景観測用カメラ k と体験者視点カメラ t は 共通の視野を持たない場合がある。したがって、あらかじめ両方の視野を観測できるカメラ a を一時的に用いることで 2 台のカメラの位置姿勢推定を行う。

求めたい行列は隠背景観測用カメラ k に対する体験者視点カメラ t の位置姿勢行列 M_{kt} である。そこでまず両方の視野を観測できるカメラ a を用いて、隠背景観測用カメ

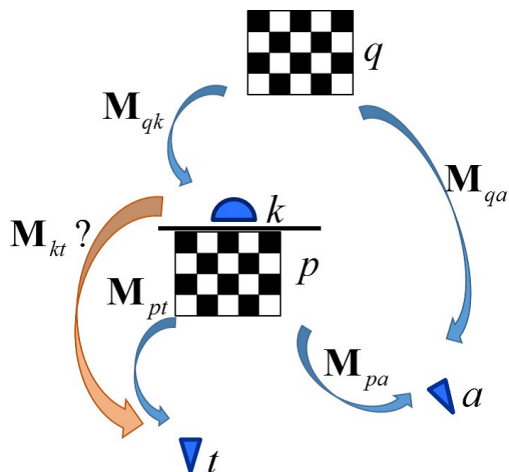


図 2: 各カメラとチェスボードの位置姿勢関係

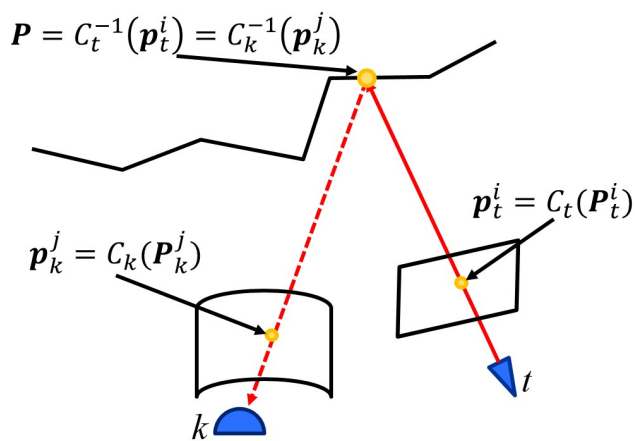


図 3: レンダリング

ラ k に対するチェスボード p の位置姿勢行列 M_{kp} を求め、その行列を用いて M_{kt} を求める。式 (9) (10) にその関係をまとめる。

$$M_{kt} = M_{kp}M_{pt} \quad (9)$$

$$= M_{qk}^{-1}M_{qa}M_{pa}^{-1}M_{pt} \quad (10)$$

このとき体験者視点カメラ以外は固定であるため、行列 M_{kp} は毎フレームで変化しない。したがって一度だけ求めれば十分である。

2.4 自由視点画像の生成

3D-LIDAR から得られる 3 次元点群は大変疎であるために、体験者視点の隠背景もまた疎になってしまう。そこで、点群から三角メッシュを構成することで隠背景の 3 次元再構成を行い、体験者視点の隠背景をレンダリングする。また物体のエッジ等で三角形が歪むことを避けるために、三角形の辺の長さが 3D-LIDAR から三角形重心位置までの距離と、ある定数の乗算によって定まる閾値より大きい場合には三角メッシュを構成しない。

レンダリングの模式図を図 3 に示す。体験者視点カメラのある画像位置 p_t^i と隠背景観測用カメラのある画像座標 p_k^j が 3 次元空間において同じ座標を示していたとすると、その画素値 I は等しくなり式 (11) の関係式が得られる。

$$I_t(p_t^i) = I_k(p_k^j) \quad (11)$$

先に述べた内部パラメータ推定と位置姿勢推定によって求められる 3 次元位置の画像座標への投影を C, また 3 次元再構成によって求められる投影の逆変換を C^{-1} とすると、式 (11) は以下のように変形できる。

$$\begin{aligned} I_t(p_t^i) &= I_k(p_k^j) \\ &= I_k(C_k(P_t^i)) \end{aligned} \quad (12)$$

$$= I_k(C_k(C_t^{-1}(p_t^i))) \quad (13)$$

したがって、体験者視点カメラの画素を隠背景観測用カメラから決定することができ、視点変換が可能となった。

2.5 透視画像の生成

体験者視点の隠背景を生成する手法は 2.4 節で述べた。この節では、生成された体験者視点の画像を体験者視点に重畳する手法について述べる。今回チェスボードを遮蔽物体に取り付け体験者視点のカメラ位置を算出したため、遮蔽物の 3 次元位置をユーザが指定しやすい。したがって本研究では遮蔽物の 3 次元位置をユーザが指定するものとする。また、体験者が隠背景を遮蔽物の奥側にあることが識別しやすいよう Barnum ら [5] の手法を参考に重畳画像を生成する。任意の重み付けで体験者視点の画像と視点変換された隠背景の画像を合成するほかに、遮蔽物のエッジの情報も重畳させることで遮蔽物を認識しやすいように透視画像を重畳した。

3. 評価実験

実験 1 では複数平面を用いた 3D-LIDAR と全方位カメラのキャリブレーションの実験を行いその精度の評価を検証する。実験 2 では隠消現実感の実験を行い従来手法と比較することで透視範囲が拡大したことを確認する。

3.1 実験環境

実験 1

- 全方位カメラ: SP360 4K Kodak PIXPRO (画素数:1440 × 1440, 画角:214 × 214)
- 3D-LIDAR: YVT-X002 HOKUYO (最大取得点数:10360, 走査範囲:210 × 40)

実験 2

- RGB カメラ: C905 Logicool (画素数:640 × 480, 画角:61 × 46)
- 全方位カメラ: SP360 4K Kodak PIXPRO (画素数:1440 × 1440, 画角:214 × 214)
- 3D-LIDAR: YVT-X002 HOKUYO (最大取得点数:10360 点, 走査範囲:210 × 40)

3.2 実験 1

2.2 節で述べた 3D-LIDAR と全方位カメラのキャリブレーションについて実験を行う。正確な位置が計測できないため、キャリブレーション後自由視点画像を生成し視覚的に評価を行う。

図 4 は実験に使用した入力シーンの例である。図 5 から入力するシーンが多ければ多いほどその精度が向上することがわかる。しかし、実験の過程でどれだけ平面を増やしてもある程度以上は精度が向上しなかった。これには複数の原因が考えられる。最も原因としてあげられるのは取得点群の誤差である。今回使用した 3D-LIDAR は得られる点群の数が少ないため、平面の点群をより多く得られるようにできるだけ大きな平面を用いようとしたが、それによってたわみが生じ正確な平面の点群が得られなかった可能性



図 4: 入力シーン例

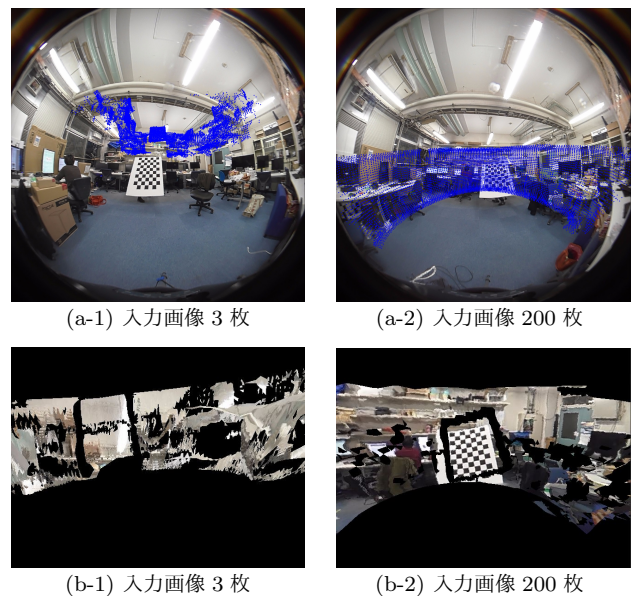


図 5: 実験 1 ((a) 3 次元点群の投影 (b) 自由視点画像)

がある。また、平面までの距離や色、材質によっても平面点群に誤差が生じた可能性がある。今後は使用する平面の検討や、十分な平面枚数についても研究していきたい。

3.3 実験 2

隠背景観測用カメラに RGB カメラを用い、視点変換を射影変換によって求めた従来手法と提案手法を目視によって比較することで、提案手法がそれと比べ透視範囲が拡大していることを確認する。また、静止画で比較がしやすいようあらかじめ隠背景観測用カメラに対する体験者視点カメラの位置姿勢を推定した。

実験結果を図 6 に示す。提案手法の方がカメラ一台に対

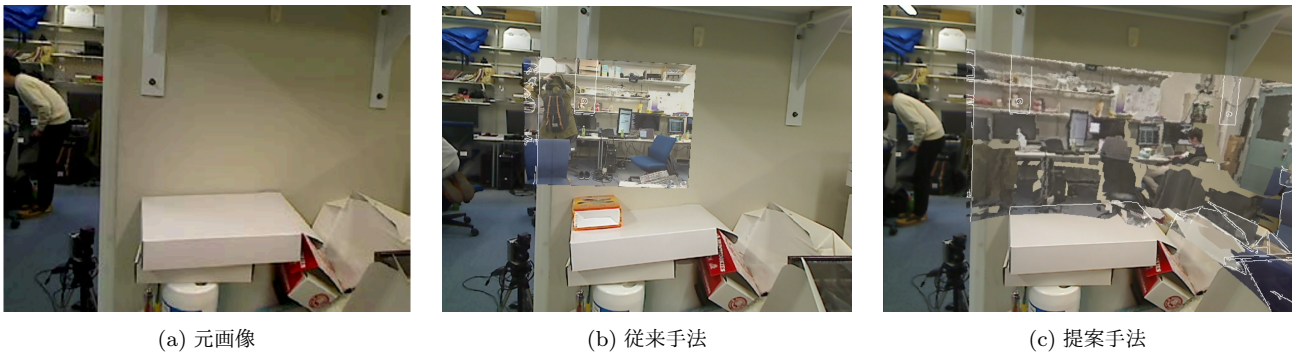


図 6: 実験 2

する透視範囲が拡大したことがわかる。今回透視できた範囲は 3D-LIDAR の走査範囲とほぼ等しい。また、全方位カメラの画角は 3D-LIDAR の走査範囲よりも広いので、今後の 3D-LIDAR の発展によって走査角度が広がれば、それに伴って本研究と同様の手法でも透視範囲がさらに広がることが期待される。

提案手法は従来手法より透視範囲が拡大したものの、従来手法より画質が荒くなっている。これの最も大きな原因は 3D-LIDAR の取得点群の少なさにあると考えられる。本手法では得られる点群から三角メッシュを構成することで 3 次元再構成を行い自由視点画像を生成しているが、3D-LIDAR の取得点群は最大でも 10360 点とその走査範囲に対し少ないために、構成された三角メッシュの一つ一つが大きくなってしまふ。そのため、各画素値のレンダリングを行う際、実際とはずれた位置で画素値を決定してしまふぼやけが生じてしまったと考えられる。

4. むすび

本論文では隠背景観測用カメラに 3D-LIDAR と全方位カメラを用いることで広範囲な隠背景の 3 次元再構成を行い、それを体験者視点に重畳させることで広範囲の透視を行う隠消現実感を提案した。

また目的の異なるに 2 つの実験を行った。1 つは 3D-LIDAR と全方位カメラのキャリブレーションである。この実験では本研究で用いたキャリブレーション手法が正確な位置姿勢推定ができるかどうか確認した。本研究で用いたキャリブレーション手法ではある程度の位置合わせはできるものの正確な位置姿勢推定を行うことができなかったことを確認した。したがって、さらに研究の余地があることが分かった。

もう一つは、実画像を用いた透視画像の生成である。この実験では提案手法が従来手法よりカメラ 1 台に対する透視範囲が拡大したことを確認した。提案手法では 3D-LIDAR の走査範囲と同等の透視が可能となり、透視範囲が拡大したことを確認した。また、3D-LIDAR の発展によってその走査範囲が広がれば、本研究と同様の手法でもより広範囲

の透視が可能となると考えられる。

謝辞 本研究の一部は、科研費・基盤研究 (S) (課題番号 24220004) と学振・特別研究員奨励費 (課題番号 16J05114) による。

参考文献

- [1] S. Mann and J. Fung, "VideoOrbits on eye tap devices for deliberately diminished reality or altering the visual perception of rigid planar patches of a real world scene," Int. Symp. on Mixed and Augmented Reality (ISMAR), pp. 48 - 55, 2001.
- [2] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," ACM Conf. Comp. Graphics (SIGGRAPH), pp.417 - 424, 2000.
- [3] Y. Shen, F. Lu, X. Cao, and H. Foroosh, "Video completion for perspective camera under constrained motion," International Conference on Computer Vision (ICPR), pp.63 - 66, 2006.
- [4] Y. Kameda, T. Takemasa, Y. Ohta, "Outdoor Seethrough Vision Utilizing Surveillance Cameras," Int. Symp. on Mixed and Augmented Reality (ISMAR), pp.151 - 160, 2004.
- [5] P. Barnum, Y. Sheikh, A. Datta, T. Kanade, "Dynamic Seethroughs: Synthesizing Hidden Views of Moving Objects," Int. Symp. on Mixed and Augmented Reality (ISMAR), pp.111 - 114, 2009.
- [6] T. Hosokawa, S. Jarusirisawad, and H. Saito, "On-line video synthesis for removing occluding objects using multiple uncalibrated cameras via plane sweep algorithm," ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC), pp1 - 8, 2009.
- [7] D. Scaramuzza, A. Martinelli and R. Siegwart, "A Toolbox for Easy Calibrating Omnidirectional Cameras, Intelligent Robots and Systems (IROS), pp.5695 - 5701, 2006.
- [8] G.Pandey, J.McBride, S.Savarese, R.Eustice, "Extrinsic calibration of a 3D laser scanner and an omnidirectional camera," 7th IFAC Symposium on Intelligent Autonomous Vehicles, pp. 336 - 341, 2010.
- [9] K.Levenberg, "A method for the solution of certain problems in least squares," The Quarterly of Applied Mathematics, pp. 164 - 168, 1944.