

2 台の全天球カメラを用いた距離推定手法

池田輝政¹ 遠藤正隆² 中嶋裕一² 三浦哲郎² 菱田隆彰¹

概要：近年，2つの魚眼レンズを前後両面に配置した，360度撮影可能な全天球カメラが普及し始めている．全天球カメラは一般的な広角レンズのカメラと比較して撮影範囲が広いため，監視カメラなどの用途においてより多くの情報を得ることが出来る．そのような用途において，被写体との距離が測定出来れば更に活用範囲が広がるが，現行の全天球カメラには距離測定の機構は搭載されていない．本研究では，2台の全天球カメラを用いたステレオ撮影によって，被写体までの距離を推定する手法を提案する．そして，その実証として2台の全天球カメラから取得した動画像からリアルタイムで距離を推定するシステムを製作し，その有用性について検討する．

A method of distance estimation with two omnidirectional cameras

TERUMASA IKEDA¹ MASATAKA ENDO² YUICHI NAKASHIMA²
TETSURO MIURA² TAKA AKI HISHIDA¹

1. はじめに

ここ数年で普及し始めた全天球カメラは，2つの魚眼レンズが前後両面に配置され，360度の範囲を一度に撮影することが可能である．画角が有限である一般的な広角レンズのカメラに比べて，本体の周囲全てが撮影範囲となるため，取得出来る情報量が多くなる．このカメラを監視カメラのような用途に使用出来れば，据え置きの状態であっても，より死角が少ないシステムの構築が期待出来る．

そのような用途を想定したとき，カメラから特定の被写体までの距離が測定出来ると，より活用範囲が広がることが期待出来る．しかし，現状の全天球カメラには被写体の検出や特定，及びその被写体までの距離を測定する機構は搭載されていない．本研究では，全天球カメラとして RICOH の「THETA V」(図1)を2台用意し，それらのステレオ撮影により特定の被写体までの距離を推定する手法を提案する．そして，その手法による距離推定の精度についての検証を行う．また，動画像によるリアルタイムでの距離推定についても検証する．

2. THETA V の仕様と出力画像の特徴

THETA V は前面，後面に設置された2枚の魚眼レンズで撮影された撮像に対して，画質調整，全天球イメージ形式への変換，パターンマッチングによる繋ぎ処理などを内部処理として行い，本体の周囲360度をカバーする1枚のパ

ノラマ画像として出力することが出来る．また，USB 経由で4K (3840×1920) , 2K (1920×960) 画質によるライブストリーミングが可能である[1]．魚眼カメラ画像からパノラマ画像を生成する工程をシステム側で行う必要がない点，高解像度でのライブストリーミングが可能なのは，リアルタイムでの距離推定を行うのに適したスペックを有していると言える．

THETA V が出力するパノラマ画像は，魚眼レンズの撮像から正距円筒図法によって生成される．正距円筒図法とは，球面上の緯度，経度をそれぞれ平面上の縦，横にそのまま読み替えたもので，緯線，経線が平面上で直角，かつ等間隔に交差するように投影される．球体の半径を R ，標準緯線を φ_0 ，経度を λ ，緯度を φ とした時，平面上の点 (x, y) への投影式は式1となる．標準緯線上と縦方向に対しては正距となるが，標準緯線から離れるにつれて横方向に拡大される

$$\begin{cases} x = R\lambda \cos \varphi_0 \\ y = R\varphi \end{cases} \quad (\text{式 1})$$



図1 THETA V

1 愛知工業大学
Aichi Institute of Technology
2 株式会社リオ
RIO CORPORATION



図 2 THETA V の出力画像

ため、画像に歪みが生じるという特徴がある。正距円筒図法は平面上の座標と球面上の緯度、経度との変換処理が簡単のため、VR デバイスなどに背景として投影する画像に適している。

THETA V に搭載されている魚眼レンズは 180 度を超える画角を持つが、出力画像は、前、後面の魚眼レンズ画像それぞれについて、水平方向に 180 度ずつの画角になるように重なる部分をブレンドして繋ぎ処理してあるため、出力画像の幅が本体を中心とした球における円周に対応する。また、画像の高さは本体を中心とした球における半球分の円周に対応する。したがって、画像の幅と高さは 2:1 の割合となる。画像の中心が前面レンズの中心であり、そこから左右に幅の 1/4 ずつの範囲が前面レンズの画像となる。後面レンズの画像は中心で分割され、前面レンズの画像の左右にそれぞれ繋ぎ処理されている (図 2)。

3. ステレオ全天球カメラによる距離推定

図 3 のように、ベースラインと直交する直線上にある被写体との距離は、三角測量によって推定することが出来る。ベースライン上に置いた 2 点の観測点と被写体とを結んだ各ベクトルとベースラインとがなす三角形の内角をそれぞれ a 、 b とし、観測点間の距離を L とした時、被写体までの距離 D は式 2 によって求められる。

図 3 の両観測点に一般的な広角レンズのカメラをベースラインと直交する向きに置き、ステレオ撮影した撮像から

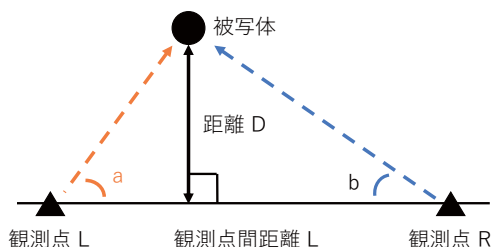


図 3 三角測量による距離推定

$$D = \frac{L}{\left(\frac{1}{\tan a} + \frac{1}{\tan b}\right)} \quad (\text{式 2})$$

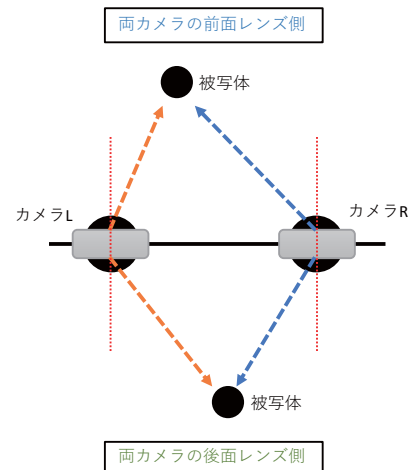


図 4 THETA V による全方位距離推定

距離を推定する手法は一般的に行われているが、本研究ではそこに THETA V を用いることで、カメラの画角に影響されない距離推定を行う (図 4)。ベースラインに THETA V を設置する際に両カメラの前面、後面の方向を合わせておくことで、図中でベースラインの上部にある被写体は必ず両カメラの前面レンズに、下部にある被写体は後面レンズに写る。こうすることで、両カメラの全周囲において距離推定を行うことが出来る。

被写体までのベクトルとベースラインとがなす三角形の内角は、魚眼レンズに射影された被写体のレンズ球面上における経度から算出可能である。2 節で述べた通り、画像内における被写体の経度は式 1 から逆算して求めることが出来る。したがって、THETA V を用いることで三角測量による距離推定を簡単に行うことが出来る。

4. 被写体検出と座標特定における問題点

THETA V の出力画像中から距離推定の対象となる被写体を検出し、その座標を特定する処理は自動化されていることが望ましいが、実現するために検討すべき点はいくつか存在する。まず、被写体の検出手法について検討しなければならない。被写体が人間の場合、AI や分類器による人物検出、顔検出が候補に挙がるが、THETA V の出力画像は正距円筒図法による投影の影響で歪んでいる箇所があるため、一般的な画像向けの検出手法では精度が上がらない可能性がある。また、検出精度の高い手法は計算量も多くなる傾向があるため、リアルタイムでの処理に支障が出ることも考えられる。全天球カメラにおける出力画像の歪みを考慮した物体検出手法については井上ら [2] によって提案されているが、やはり処理速度の低下が課題として挙げられている。被写体検出の精度をある程度保ちつつ、計算量を抑えられるような手法を選択する必要がある。

次に、左右のカメラ画像で同じ被写体を検出する手法についても検討する必要がある。両カメラ画像から被写体の

検出を行った時に、違う被写体が検出されてしまうと距離推定が成り立たない。両画像に対して単純に被写体検出を行うのではなく、片方で検出された被写体をもう片方の画像から探索するなど、同じ被写体が検出されるような工夫をしなければならない。

経度の算出に用いる被写体の座標の特定手法についても検討が必要である。被写体までの距離推定を行うには、何らかの手法で出力画像内から被写体を検出した後、その被写体の範囲内から測定点を選定し、その点の座標から経度を算出することになる。この時、左右のカメラ画像で被写体の同じ箇所を選定出来なければならない。例えば、対象が人物であり左右のカメラで同じ被写体を検出していても、片方のカメラ画像では被写体の左肩、もう片方のカメラ画像では被写体の右肩を測定点として選定してしまうと、内角の算出に誤差が生じる。距離推定を正確に行うには、どちらの画像からも右肩が選出されるのが望ましい。測定点の選定方法として、検出された被写体の重心や外接矩形の中心点を使うことが考えられるが、カメラに映る被写体の角度が左右で違うため、これらも正確に同じ点を選定出来るとは限らない。また、被写体の位置によっては歪みの影響も受けてしまう。これらの問題を考慮して測定点を選定する必要がある。

5. 距離推定手法の提案と処理手順

本研究では、4 節で述べた問題点に対処した距離推定手法を提案する。図 5 にその処理手順を示す。まず、被写体の検出手法については、事前に撮影しておいた背景画像との差分を用いた動体検出を用いて行うこととした。この手法では被写体は何であるかを認識することは出来ないが、歪みに関係なく被写体の位置を検出する事が可能であり、処理が単純で全体的な処理速度への影響が少なくなる。具体的には、1 フレーム毎に背景画像と撮影した画像との差分画像を取得し、その画像に対して輪郭抽出を行う。検出された輪郭の中から最も大きいものを被写体とし、その外接矩形を取得する。被写体の検出は図 4 におけるカメラ L の出力画像に対してのみ行い、検出した被写体の外接矩形を探索領域として特徴点の抽出を行う。カメラ R の出力画像においても同じように特徴点の抽出を行い、両画像から取得した特徴点に対して総当たり法によるマッチングを行う。マッチングされた全ての特徴点に対して、左右の特徴点同士が持つ x 座標の差分を計算し、その中央値を持つ特徴点同士を測定対象とする。そして、それらの座標から 3 節で述べた手法にて距離推定を行う。この手法であれば、1 フレーム毎に被写体の検出を 1 度行えばいいため、計算量を軽減出来る。また、特徴点マッチングによって、左右の画像から適切な測定点を選定出来る可能性が増える。

図 4 のようにカメラを設置し、その高さが揃えてある場

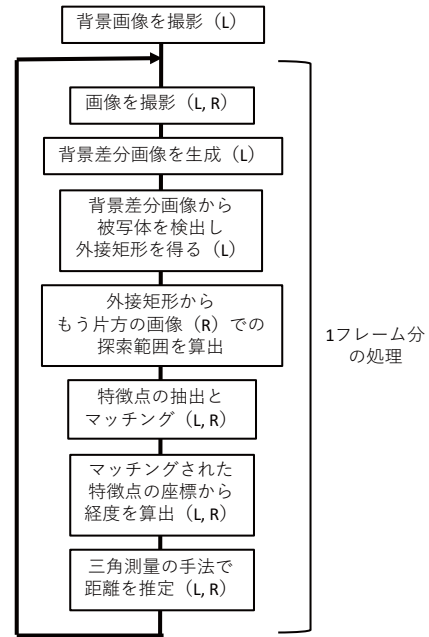


図 5 距離推定の処理手順

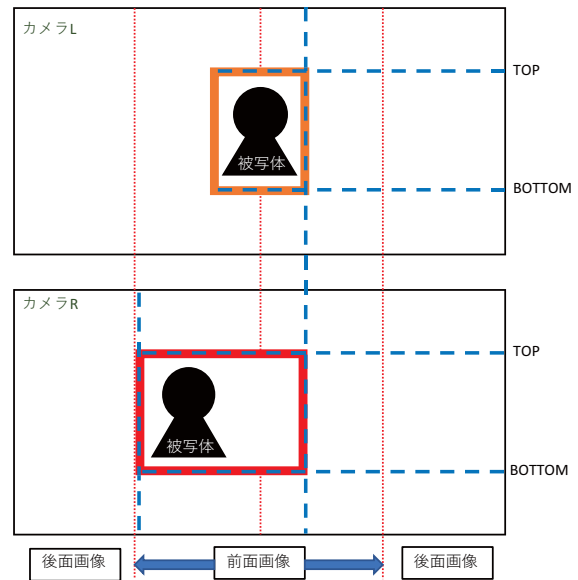


図 6 カメラ R に対する探索範囲の限定

合、両カメラに映る被写体には次のような関係性がある。

- 両カメラとも同じ面のレンズに写る
- 両カメラともほぼ同じ高さに写る
- カメラ R に写る被写体の位置はカメラ L に写る被写体の位置よりも左にくる

この関係性を考慮すると、カメラ R の出力画像から特徴点を抽出する際の探索範囲は、図 6 のように限定することが可能である。探索範囲が限定されることで計算量が軽減出来る。また、特徴点マッチングの精度が向上することも期待出来る。

6. 提案手法による距離推定の検証

5 節で述べた手法による距離推定の有用性を確かめるべく、プログラムを作成して動作を検証した。プログラミング言語には Python を用い、背景差分の生成、被写体の輪郭抽出と外接矩形の取得、特徴点の抽出とマッチングには OpenCV ライブラリを使用した。特徴点抽出のアルゴリズムには A-KAZE を採用した。A-KAZE は拡大、縮小や回転に対してロバストであり、比較的フラットな部分からも特徴点が抽出出来ることが知られている。A-KAZE の特徴は、左右のカメラで被写体が写る角度が違うことや、THETA V の出力画像が歪みを持つことに対して効果的に働くことが期待出来る。

作成したプログラムを用いて、まずは距離推定が実際に可能かどうか、実用に足る精度が出るのかどうかについて、1 フレーム分の静止画を撮影して検証した。事前準備としてベースライン上に THETA V を 2 台、カメラ間の距離を 1[m] として設置し、手作業でキャリブレーションを施した。画像の解像度は 2K (1920×960) とした。その後、ベースラインから 0.5, 1.0, 2.0, 3.0, 4.0 [m] の各計測ライン上において、複数のポイントで被写体を撮影し、距離推定を試みた。例として、被写体を両カメラ間の中央で前後に移動させた時の、実際の距離と計測値を表 1 に示す。どの計測ラインにおいても、誤差率が 5[%] 内に収まっていることが分かる。全体的に実際の距離よりも短めに測定されているのは、手作業でキャリブレーションを行ったことによる誤差が影響していると考えられる。より正確なキャリブレーションを行うことで、更に精度を高められる可能性がある。それ以外の計測位置でも、特徴点が正確にマッチングされた場合には同程度の誤差で計測することが出来た。しかし、マッチングが上手くいかずに両カメラで別々の測定点が選定されてしまい、精度が落ちることもあった。

次に、ライブストリーミングされる動画像に対してリアルタイムでの距離推定を行い、その動作速度について検証した。2K (1920×960)、及び 1K (960×480) の動画像に対し、それぞれ距離推定を行った上で PC の画面上に画像を表示させた際の FPS を算出した。THETA V は約 30[fps] でライブストリーミングを行う能力を持っているが、プロ

表 1 実際の距離と計測値の例

実際の距離[m]	計測値[m]	誤差率[%]
0.5	0.489	2.2
1.0	0.990	1.0
2.0	1.946	2.7
3.0	2.951	1.6
4.0	3.803	4.9

グラムでは 1 フレームで 2 台のカメラから画像を取得していること、被写体が存在しない場合でも背景差分の取得は行われていることなどから、被写体を検出していない状態で 2K で 3~5[fps] 程度、1K で 9~13[fps] 程度での動作となった。被写体を検出している状態では特徴点抽出とマッチングの処理が加わるため、確実に速度が低下するはずであるが、FPS が全く安定せず正確にどの程度低下するかは判別出来なかった。これは、背景差分画像が照明などの影響を受けることで、安定した被写体検出が出来ていないことや、同じ理由で特徴点の検出量もフレーム毎に安定していないことが原因と考える。

7. まとめ

全天球カメラのステレオ撮影による被写体までの距離推定について、THETA V が出力する画像の特徴を活かした推定手法を提案し、その有用性について検証した。結果として、特徴点マッチングによる測定点の選定が上手くいった場合には実用に足る精度で距離推定が出来ることが分かった。動作速度についても検証したが、安定した速度で動作させることが出来ず、こちらは有用性を示すには至らなかった。よって、提案手法は距離推定については有用であるが、リアルタイムでの安定した動作については未だ改善の余地があると言える。

今後の課題としては、より安定した動作を目指して様々な被写体の検出手法や特徴点マッチングの手法を試すことが挙げられる。被写体の検出と測定点の選定が安定して行えば、より実用的な距離推定システムに近づけることが出来る。また、ステレオ動画像による距離推定では本来、カメラ間のフレーム同期を考慮する必要がある[3]。フレーム同期に対応させることで、距離推定の精度がより高まることが期待出来る。これらの課題に取り組み、様々な場面で活用出来る距離推定システムを目指す。

謝辞

本研究は JSPS 科研費 JP19K12073 の助成を受けたものです。

参考文献

- [1] 製品紹介 | RICOH THETA : <https://theta360.com/ja/about/theta/technology.html>
- [2] 井上 慶彦, 岩村 雅一, 黄瀬 浩一: 全方位カメラを用いた物体検出とトラッキング - 視覚障害者支援システムの実現に向けて-, 情報処理学会研究報告コンピュータビジョンとイメージメディア 20 号, pp.1-6 (2018)
- [3] 関 晃仁, 奥富 正敏: 非同期ステレオ動画像を用いた同時最適化による位置とモーションの推定, 情報処理学会論文誌コンピュータビジョンとイメージメディア 49 号, pp.22-34 (2008)