

ステレオ全天球カメラによる距離推定と「音の AR」への活用

池田輝政^{†1} 遠藤正隆^{†2} 中嶋裕一^{†2} 松井瑠偉人^{†2} 菱田隆彰^{†1}

概要：魚眼カメラを前後両面に配置することで、360度の範囲を同時に撮影することが可能な全天球カメラが普及し始めている。このカメラで撮影される画像は、撮影範囲の広さから様々な用途での利用が期待される。もし、カメラから画像中の被写体までの距離が取得できれば更に活用の幅も広がるが、現行の全天球カメラには距離を測定する機構は搭載されていない。本研究ではこれまで、全天球カメラのステレオ撮影による距離推定の手法を提案し、手法の妥当性を検証してきた。本稿では、改良したシステムを用い取得した距離情報の活用事例として、被写体の3次元位置を音声の定位として表現することで、被写体の位置を音で知覚するシステムを製作した。

1. 全天球カメラによる距離推定

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ており、カメラを中心とした360度の範囲を一度に撮影することが可能である。現状ではVRコンテンツの素材として用いられることが多いが、画角が有限である一般的な広角レンズのカメラと比べて撮影範囲が広く、一度に取得できる情報が増えるため、娯楽目的以外にも監視カメラなど様々な用途への活用が期待できる。

全天球カメラを有効に活用できる場面を増やすためには、画像以外にも情報が取得できた方が望ましい。例えば、カメラに映る被写体をリアルタイムで検出し、その被写体との間の距離を取得できれば、それを反映したVRコンテンツやシステムの製作が可能となる。しかし、現行の全天球カメラには被写体の検出や距離を測定する機能は搭載されていない。また、一般的な深度センサに用いられるパターン照射型や飛行時間型の測定手法は、全天球カメラの画角に対応させることが容易ではない。したがって、これらの手法を用いることなく、現行の全天球カメラで取得できる撮影画像から距離を推定する手法を検討する必要がある。

我々はこれまでに、ベースライン上に並べた2台の全天球カメラによるステレオ撮影を行い、その視差から三角測量の要領でベースラインから被写体までの距離を推定する手法を提案した。そして、動画像に対してリアルタイムで被写体の検出と測定点の特定、距離の推定を行うシステムを製作し、手法の妥当性を検証した[1]。本稿では提案システムに対して改良を行い、精度、速度の向上を目指す。

2. システムの概要とその実証

本システムでは全天球カメラとしてRICOHの「THETA」を採用している。THETAは前、後面に設置された2枚の魚眼レンズで撮影された画像に対して、画質の調整、全天球



図1 ステレオ全天球カメラ

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枚の正距円筒図として出力する[2]。このTHETAをベースライン上に平行に並べたステレオ全天球カメラ(図1)を用意し、ステレオ撮影された動画像の各フレームを対象に被写体の検出、測定点の特定、距離推定を行う。

本システムでは次のような手順で距離推定を行う。まず、片方の画像(A)と事前に撮影しておいた背景画像との差分を取り、被写体を検出する。次に、被写体を含む矩形領域ともう片方の画像(B)との間でA-KAZEアルゴリズムによる特徴点抽出とマッチングを行う。この時、画像(B)に被写体が出現する位置は限定されるので、検索範囲を狭めて処理を減らすことができる。マッチングされた特徴点の中から距離が中央値になる組み合わせを左右の測定点とする。そして、カメラと測定点を結んだ線とベースラインとの間の角度を算出する。THETAの出力画像は正距円筒図なので、単純な計算で画像上の座標をレンズ球面に写る被写体の緯度、経度に変換することが可能である。両画像から得られた被写体との角度により、三角測量の要領でベースラインからの距離を算出する。

今回は既存のシステムのプログラムを見直し、処理の最

^{†1} 愛知工業大学

^{†2} 株式会社リオ

適化を行い実行速度の向上を図った。また、背景差分生成時の下処理で行っていたコントラスト調整などに用いるパラメータの適正化を行い、被写体の検出精度向上を試みた。表1は改良後のシステムにおける測定精度の検証をすべく、両カメラの中央で被写体を前後させつつ距離推定を行った結果の一例である。測定点の特定が適切に行われた場合、どの距離でも誤差率3%以内で推定することができた。

しかし、照明の影響などで被写体の検出が適切でない場合は測定点を的確に特定することができず、精度が落ちてしまう。被写体検出の手法の変更も含め今後の課題となる。処理速度については、各部の計算量が減少するよう改善を行ったが、既存のシステムからの課題となっている、被写体の大きさによって必要な処理量の変動が大きいという点を解消するには至らず、結果として表示のフレームレートが安定しないため目に見える高速化は得られなかった。

3. 音声の定位による被写体位置の知覚

ステレオ全天球カメラによって得られる被写体までの距離情報を活用する一例として、被写体の3次元位置を音声の定位によって知覚する機能をシステムに追加した。昨今では、視覚だけでなく聴覚から得る情報で現実を拡張する「音のAR」と呼ばれる仕組みが提唱されており、音のARを利用することを前提としたウェアラブルデバイスも登場し始めている。図2の「BOSE FRAMES」はサングラスにスピーカ、マイク、モーションセンサが内蔵されており、スマートフォンなどの外部デバイスと連携することで、音のARを利用することができる[3]。本システムではカメラで得た視覚情報を聴覚情報に置き換えることで、音のARによる新たなユーザ体験を提供する。

具体的には、ステレオ全天球カメラで得られたベースラインから被写体までの距離、及びカメラから見た被写体の緯度、経度より、カメラと被写体を結ぶ3次元ベクトルの距離を求める。これらの情報を元に、カメラをユーザ、被写体を点音源とした時にユーザに聞こえる音声に生じるであろう両耳間時間差や両耳間強度差を計算し、それを反映した音声データを生成する。また、3次元ベクトルの距離に応じて音量を調節し、音源の遠近を表現する。これらの処理をリアルタイム距離推定と並行して行い再生すること

表1 実際の距離と推定値の例

実際の距離[m]	推定値[m]	誤差率[%]
1.0	1.01	1.0
2.0	1.97	1.5
3.0	2.98	0.6
4.0	3.91	2.3



図2 BOSE FRAMES

で、カメラ内での被写体の動きを音声の定位の動きに置き換えてユーザに提供できる。前述のBOSE FRAMESを装着してシステムに接続してみると、実際に被写体の動きと音声の定位が連動することが確認できた。

4. まとめ

ステレオ全天球カメラによる被写体までの距離推定手法について提案し、改良を加えた距離推定システムによって推定値の妥当性を検証した。また、取得した距離情報を用いて被写体の3次元位置を割り出し、その位置に応じて音声の定位を変化させることで、視覚情報を聴覚情報に置き換える音のAR機能を追加し、その動作を確認した。

新しいシステムによる距離推定については、測定点の特定が上手くいった場合、実用に足る精度で距離を推定することが分かった。今後も処理の高速化と測定点特定の精度向上が課題となる。また、ステレオ動画像による距離推定で本来重要となるカメラ間のフレーム同期[4]についても導入を検討していきたい。

音のAR機能については、ひとまず目的であった視覚情報を聴覚情報に置き換えたユーザ体験の実現は達成できた。今後は実際の位置情報と音声の定位がより妥当に感じるようにチューニングをしていく必要がある。

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] 池田輝政, 遠藤正隆, 中嶋裕一, 三浦哲郎, 菱田隆彰. 2台の全天球カメラを用いた距離推定手法. マルチメディア、分散、協調とモバイル(DICOMO2019)シンポジウム論文集, pp.147-150, 2019.
- [2] “製品紹介 | RICOH THETA”. <https://theta360.com/ja/about/theta/technology.html>, (参照 2019-12-15).
- [3] “ボーズ、革新的な「BOSE FRAMES」の国内販売を本日より開始”. <https://globalpressroom.bose.com/jp-jp/pressrelease/view/2024>, (参照 2019-12-15).
- [4] 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, vol. 49, pp.22-34, 2018.