

全天球ボールカメラによる撮影動画の視点固定手法の提案

中澤 正和* 小池 英樹*

概要. スポーツのボールにカメラを装着することによるボール視点の新しいスポーツ観戦が実現しようとしている。問題となるのはスポーツ中にボールが回転することで視点が不安定になってしまうことである。本稿では全天球カメラを用いたボールカメラのプロトタイプを作成し、全天球動画の視点を固定するアルゴリズムを提案する。本アルゴリズムではフレーム間で特徴点マッチングを行い、8点アルゴリズムによってフレーム間の姿勢変化を推定する。そして姿勢修正を各フレームに施すことで動画全体を基準フレームと同じ視点に固定することができる。本アルゴリズムは本稿のハードウェアで撮影されたものではない全天球動画にも適用でき、動画を撮影したカメラの種類に依存しない。

1 はじめに

スポーツは人々の生活に深く関わっている。スポーツは体を動かすという欲求を満たしてくれるだけでなく、爽快感や達成感、他者との連帯感を与えてくれる。またスポーツは見て楽しむこともできる。最近ではテレビのようなメディアを通してスポーツ観戦をする機会が多い。普通スポーツの動画は選手の邪魔にならないようにフィールドの外側に設置されたカメラを使って撮影されるので、選手から遠い距離の動画しか見ることができない。しかし Woodman Labs 社の GoPro¹ に代表される小型で頑丈なアクションカメラが登場したことで、スポーツ観戦の形が変化している。選手がアクションカメラを装着してスポーツを行い、選手視点の動画を撮影できるようになった。これらの動画により今まで以上にスポーツの臨場感を味わうことができるようになった。

そして近年、さらに新しい形のスポーツ観戦が実現しようとしている。スポーツのボールに装着したカメラを使ったボール視点のスポーツ観戦である。問題となるのはボールが回転することでカメラの視点が不安定になり、動画中で特定の方向を見続けることが困難なことである。視点の不安定な動画から視点の固定された動画を生成する研究は現在までに行われてきた [5, 3, 2]。これらの研究では特定方向を向いたときのフレームを取り出し合成することで、疑似的に視点の固定された動画が生成された。しかしその視点はカメラの視野の中という限られた範囲にしか固定することはできない。つまりカメラの視野はそのときのカメラの姿勢に依存するので、カメラの視野の外を含む任意の方向に視点を固定することはできない。

この問題を解決してくれる全天球カメラの人気の現在高まっている。全天球カメラは複数のレンズを

使うことで全天球パノラマ動画を撮影することができる。周囲のすべてのシーンを記録するので、まるでそこにいるかのような没入感を体験できる。360度 GoPro マウント² は全天球動画を撮影するために6個の GoPro を使う。最近ではリコー社の RICOH THETA³ や Kodak 社の PIXPRO SP360⁴、Samsung 社の Gear 360⁵ が発売され、誰でも手軽に全天球動画を撮影できるようになった。これからも多くの企業が全天球カメラを発売すると考えられる。

本稿では全天球動画の視点を固定するアルゴリズムを提案する。本アルゴリズムによって全天球カメラの姿勢を変化させながら撮影した動画から、任意の方向に視点を固定した動画を生成することができる。図1は本稿の寄与を示す。本稿では RICOH THETA S を用いたボールカメラのプロトタイプで全天球動画を撮影したが、例えば Brazucam⁶ や GoPro 等の他の種類のカメラで撮影した全天球動画にも本アルゴリズムは適用できる。

本アルゴリズムは3つのステップからなる。最初に全天球画像から特徴点と特徴量を抽出する。次にフレーム間の全天球カメラの姿勢変化を推定する。最後にフレームに姿勢修正を施す。

2 関連研究

2.1 ボールカメラ

Pfeil らは Panoramic Throwable Ball Camera (Panono) を開発した [8]。Panono は 36 個のカメラユニットをもち全天球画像を撮影できる。加速度

² Freedom360 マウント <http://freedom360.us/>

³ RICOH THETA <https://theta360.com/ja/>

⁴ PIXPRO SP360 <https://www.maspro.co.jp/products/pixpro/sp360/>

⁵ Gear 360 <http://www.samsung.com/global/galaxy/gear-360/>

⁶ Brazucam <https://www.youtube.com/watch?v=pZskBajN5oI>

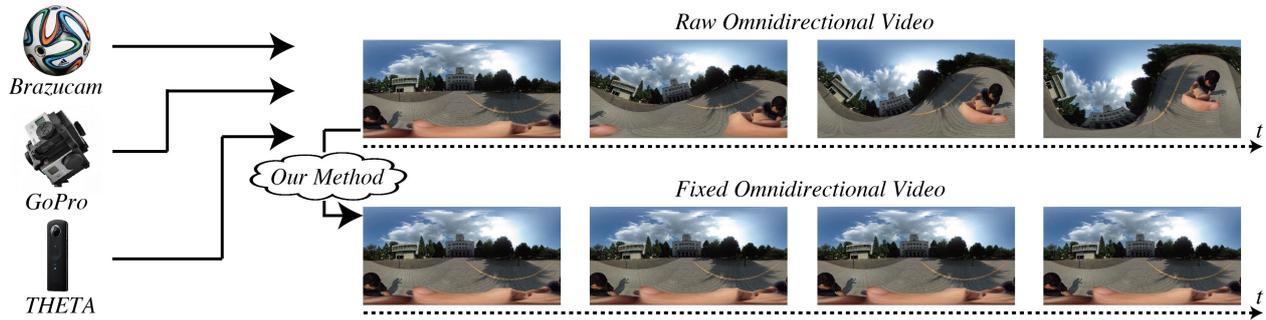


図 1: 全天球動画は様々な種類のカメラで撮影される。本アルゴリズムでその視点を固定することができる。

センサも内蔵しているので、Panono を投げ上げて頂点に達すると自動的にシャッターが切られる。球形であり投げることができるという点で Panono の使用法はスポーツのボールと類似している。しかし Panono は動画を撮影することはできない。

Funakoshi らはボールの表面に複数のカメラを装着してボールを回転させながら撮影した動画から pseudo straight view を合成するアルゴリズムを開発した [2]。ボールの表面にカメラを 6 個均等に装着し、それらが撮影した動画の中から特定方向を向いたときのフレームを取り出し合成することで視点の固定された動画を生成する。しかしこのアルゴリズムの中で実際に使われるカメラは円周上に配置された 4 個のみであり、ボールの回転軸はその円に垂直な方向であると想定されている。スポーツ中ボールの回転軸は変化していくと考えられるので実用的ではない。

2.2 全天球動画の回転の補正

Kasahara らは一人称視点全天球動画の体験共有システム、JackIn Head を開発した [4]。この研究では頭部に装着した 6 個のカメラを使って全天球動画が合成される。ただしその底部はカメラ装着者の体の部分に当たるので欠損している。全天球動画は合成後、頭部の回転の補正処理を施され遠隔ユーザーに送られる。この回転補正処理の問題点として、処理に必要な特徴点の抽出領域が画像内の一部に限定されているため高速な回転に対応できないことや、特徴点がカメラから充分離れていないと機能しないアルゴリズムになっていることがあげられる。

3 全天球カメラ

本稿では全天球カメラとして RICOH THETA S を使用する。動画のフレームサイズは最大で 1920×1080 [pixel]、フレームレートは 30 [fps] である。RICOH THETA S は 2 つの広角レンズからそれぞれ得られた情報を 1 つのフレームに並べて記録する。この動画をそのまま処理するのは困難なので、

リコー社のソフトウェアを使用して正距円筒図法の形式に変換してから処理を行う。

本稿で対象とする全天球画像は、カメラの周囲 360 度の球面上に得られた情報を正距円筒図法によって投影したものである。正距円筒図法では球面の経緯度 (θ, ϕ) と全天球画像座標 (u, v) に以下のような関係がある。

$$u = \frac{w}{2\pi} (\theta + \pi) \quad (-\pi \leq \theta < \pi) \quad (1)$$

$$v = \frac{h}{\pi} \left(-\phi + \frac{\pi}{2} \right) \quad \left(-\frac{\pi}{2} \leq \phi < \frac{\pi}{2} \right) \quad (2)$$

ただし w と h はそれぞれ全天球画像の幅と高さを表す。図 2a と図 2b はそれぞれ球面上の経緯度と正距円筒図法による全天球画像座標を示す。正距円筒図法による全天球画像には経線と緯線が均等な間隔で配置されるという性質がある。しかし球面上ではそのようなになっていない。したがって正距円筒図法による全天球画像は低緯度領域では歪みが少ないが、高緯度領域では情報が引き伸ばされて大きな歪みが発生する。

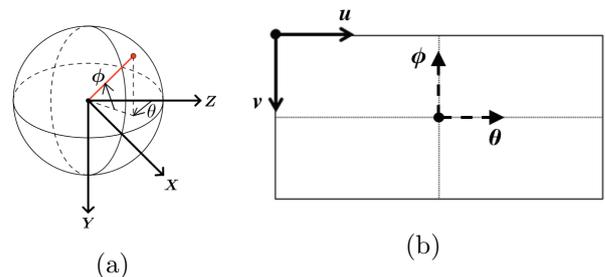


図 2: 正距円筒図法における球面の経緯度と全天球画像座標座標の関係。(a) 球面の経緯度。(b) 全天球画像座標。

4 画像処理アルゴリズム

全天球カメラの姿勢を変化させながら撮影した視点の不安定な動画から視点の固定された動画を生成するアルゴリズムを説明する。本アルゴリズムは画

像処理のみに重点をおき、ジャイロセンサ等の外部センサは使用しない。カメラが高速に回転したときにフレームとセンサ値を同期させることが困難だからである。アルゴリズムは3つのステップからなる。

4.1 全天球画像からの特徴点と特徴量の抽出

3章で述べたように、正距円筒図法による全天球画像は低緯度領域では歪みが少ないが、高緯度領域では歪みが大きい。Kasaharaら[4]は歪みの少ない低緯度領域からのみ特徴点を抽出していたが、これではフレーム間で高緯度領域と低緯度領域に跨って特徴点が移動する場合に対応できない。スポーツのボールは高速に回転するのでそのような状況は頻繁に発生すると考えられる。したがって低緯度領域からだけでなく高緯度領域からも特徴点を抽出することが求められる。

球面上の特徴点を抽出する研究はMorelら[7]やCruz-Mota[1]らなどにより行われてきた。本稿ではTairaら[9]の方法を参考に、特徴量としてはSIFT[6]を用いる。注目領域が高緯度にあるとき、特徴点を抽出する前にカメラ座標系のX軸まわりに球面を回転させることで注目領域を低緯度へ移動させる。図3は極を含む高緯度領域を低緯度へ移動させる例を示す。上が元画像で、色の付いた部分(マゼンタ)が特徴点を抽出しようとする高緯度領域である。下が回転後の画像で、色の付いた部分が元画像のものに対応している。図3aでは $+\pi/3$ 、図3bでは $-\pi/3$ 球面を回転させている。

低緯度へ移動させた注目領域から特徴点と特徴量を抽出する。抽出した特徴点の座標は元画像での座標とは異なるので、低緯度へ移動させるために行った回転と逆の回転の処理を施すことで、元画像における座標に変換する。

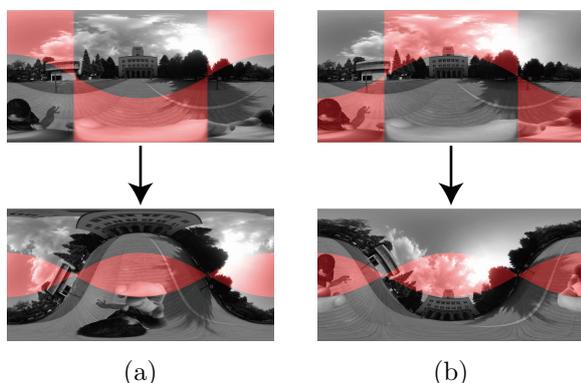


図3: 球面の回転による高緯度領域の低緯度への移動。(a)X軸まわりに $+\frac{\pi}{3}$ 回転。(b)X軸まわりに $-\frac{\pi}{3}$ 回転

4.2 フレーム間の全天球カメラの姿勢変化の推定

4.1節で抽出した特徴量を使ってフレーム間で特徴点マッチングを行い、マッチングに含まれる特徴点の座標を正規化画像座標に変換する。マッチングと正規化画像座標をもとに8点アルゴリズムによりフレーム間の基本行列を推定する。ロバストに推定するためにRANSACやLMedsを利用する。推定した基本行列を分解することで回転行列と歪対称行列が得られるので、前者を全天球カメラの姿勢変化とする。

4.3 フレームの姿勢修正

基準フレームから現フレームまでのすべての隣接するフレーム間で回転行列を推定する。これらの総乗は基準フレームと現フレーム間の全天球カメラの姿勢変化とみなせる。この行列を現フレームに施し姿勢修正を行うことで、現フレームの視点を基準フレームと同じものに修正することができる。図4は基準フレームから数えてn番目の現フレームを姿勢修正の様子を示す。

本アルゴリズムでは基準フレームと現フレーム間の全天球カメラの姿勢変化を直接推定しない。以下の理由により特徴点マッチングの精度を向上させることができるからである。

- 隣接するフレーム間の時間は微小であるので、特徴点の移動量も微小である。
- 隣接するフレーム間の時間は微小であるので、シーンは静的である。
- 隣接するフレーム間の時間は微小であるので、全天球カメラの並進による特徴点の見え方の変化はない。

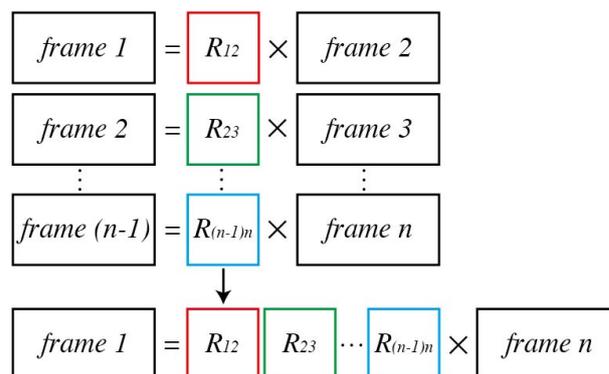


図4: n番目のフレームの姿勢修正。

5 実験

実際に全天球カメラの姿勢を変化させながら動画を撮影し、その視点を固定する実験を行った。

5.1 ボールカメラのプロトタイプ

全天球カメラの姿勢を変化させながら動画を撮影するために、カメラを透明なアクリルボールの中心に固定したボールカメラを考えた。図 5a はボールカメラの構造を示す。全天球カメラをボールの中心に固定するために円形板をカメラの形に繰り返し抜き、カメラをそこにはめ込む。それらを2つのアクリルドームで包むことでボールカメラができる。

図 5b は本稿で使用したボールカメラのプロトタイプを示す。アクリルドームの直径は約 25cm で厚さは約 2.5mm である。本稿ではこのようなボールカメラを使って全天球動画を撮影したが、本アルゴリズムは他の種類のカメラで撮影された全天球動画にも適用できることに留意してほしい。

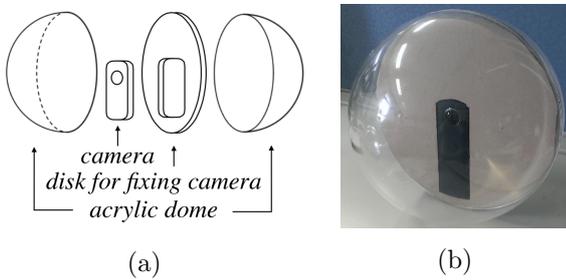


図 5: 全天球カメラを用いたボールカメラ。(a) ボールカメラの構造。(b) ボールカメラのプロトタイプ。

5.2 結果

5.1 節のボールカメラのプロトタイプを使って全天球動画を撮影し、本アルゴリズムで視点を固定する処理を行った。撮影時にボールを回転させる必要があるが、ボールの素材がアクリルであるため、サッカーのキックやバスケットボールのドリブルのようなボールに強い衝撃を与えるプレイはできない。そこでボウリングとパスをサンプルシナリオに選んだ。

図 6 は本アルゴリズムの処理の前後の全天球画像列の比較を示す。図 6a はボウリングを、図 6b はパスを示す。それぞれのシナリオで左側が処理前の画像列、右側が処理後の視点を固定した画像列である。処理後の画像列は視点が一番上の画像のものに固定されていることがわかる。

6 考察

本アルゴリズムによって全天球カメラの姿勢を変化させながら撮影した動画の視点を固定することができた。しかしまだ解決すべき問題がある。

1 つめはフレーム中のブレの問題である。RICOH THETA S はフレームレートが 30 [fps] なので、ボールが高速に回転するとブレが発生し特徴点の抽出が困難になる。抽出される特徴点の数が少ないと全天球カメラの姿勢変化の推定の精度が落ち、最悪の場合推定に失敗してしまう。この問題は RICOH THETA

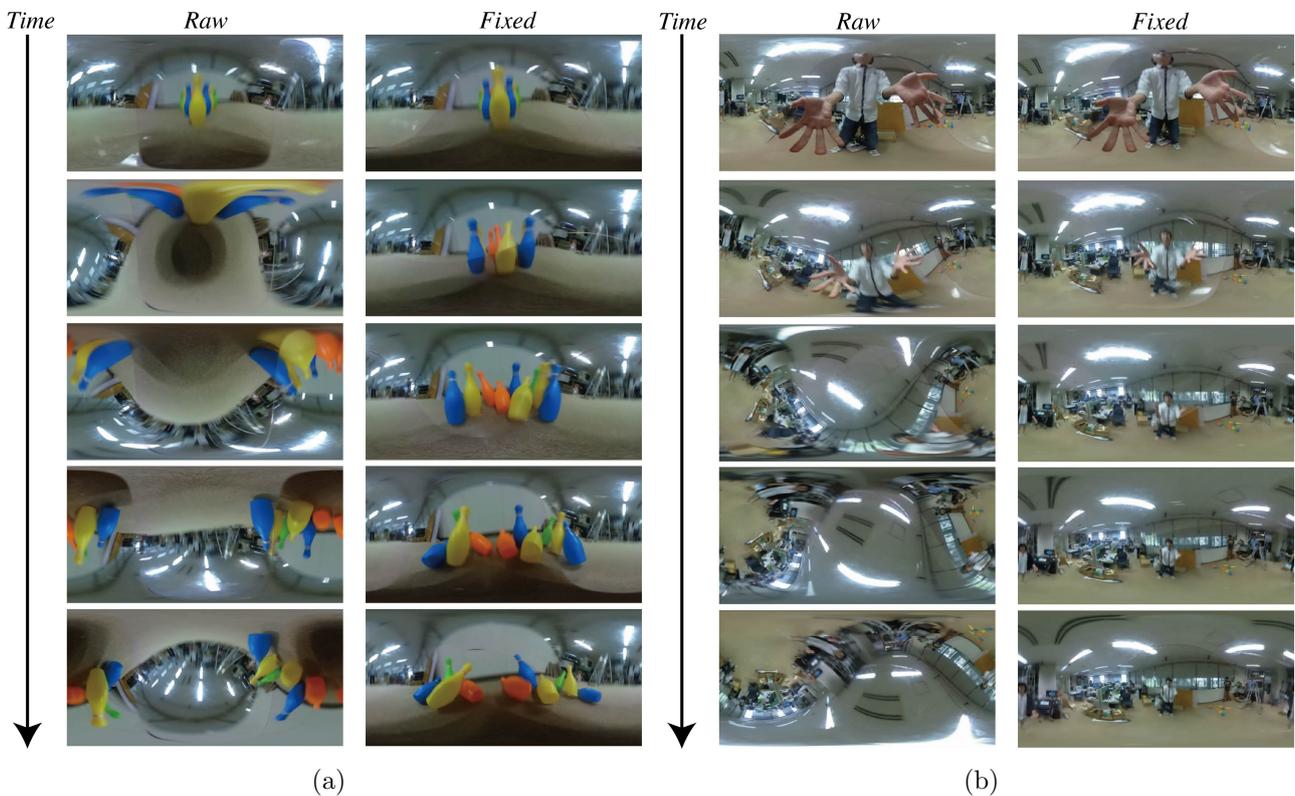


図 6: 本アルゴリズムによる処理の前後の全天球画像列の比較。(a) ボウリング。(b) パス。

Sを高フレームレートのカメラで代替することで解決できると考えられるので、今後はPIXPRO SP360 4Kを用いることを検討している。PIXPRO SP360 4Kは2つで容易に全天球動画を合成することができる。フレームレートは全天球動画の撮影モードで最大120 [fps]なので、ブレの軽減が期待できる。

2つめはフレームの姿勢修正に用いられる行列に誤差が蓄積されてしまう問題である。本アルゴリズムでは基準フレームから数えて n 番目のフレームの姿勢修正に、それまでに推定した $(n-1)$ 個の回転行列が必要である。つまり基準フレームに近いフレームの姿勢修正に必要な回転行列は少数であるが、フレームが後方になればなるほど必要な回転行列の数は増加する。それぞれの回転行列には推定誤差が含まれているので、その数が増加すればそれらの総乗である行列はより多くの誤差を含んでしまうと考えられる。この問題はバンドル調整を利用することで解決できると考えられる。バンドル調整はフレームの処理を進めるたびにそれまでに推定したパラメータを最適化するので、後方のフレームでも精度よく姿勢修正が行える。

7 結論

本稿では全天球カメラを用いたボールカメラのプロトタイプを作成し、全天球動画の視点を固定するアルゴリズムを提案した。最初に球面を回転させることで全天球画像全体から特徴点と特徴量を抽出した。次に8点アルゴリズムによりフレーム間の全天球カメラの姿勢変化を推定した。そして基準フレームから現フレームまでのすべての隣接するフレーム間で推定した回転行列の総乗を用い現フレームの姿勢修正を行うことで、現フレームの視点を基準フレームのものに修正することができた。

本稿の提案は既存のハードウェアとソフトウェアを組み合わせたものであるが、ボールカメラを用いた新しいスポーツ観戦の方向性と全天球カメラの利用法を示し、その中でも重要な問題を解決したものである。

参考文献

- [1] J. Cruz-Mota, I. Bogdanova, B. Paqueler, M. Bierlaire, and J. Thiran. Scale Invariant Feature Transform on the Sphere: Theory and Applications. In *International Journal of Computer Vision*, Vol.98, No.2, pp. 217–241, 2012.
- [2] R. Funakoshi, Y. Okudera, and H. Koike. Synthesizing Pseudo Straight View from A Spinning Camera Ball. In *Proceedings of the 7th Augmented Human International Conference*, No.30, 2015.
- [3] K. Horita, H. Sasaki, H. Koike, and K. M. Kitani. Experiencing the Ball’s POV for Ballistic Sports. In *Proceedings of the 4th Augmented Human International Conference*, pp. 128–133, 2013.
- [4] S. Kasahara, S. Nagai, and J. Rekimoto. First Person Omnidirectional Video: System Design and Implications for Immersive Experience. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, pp. 33–42, 2015.
- [5] K. M. Kitani, K. Horita, and H. Koike. ”Ballcam!: dynamic view synthesis from spinning cameras. In *Adjunct proceedings of the 25th annual ACM symposium on User interface software and technology*, pp. 87–88, 2012.
- [6] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. In *International Journal of Computer Vision*, Vol.60, No.2, pp. 91–110, 2004.
- [7] J. M. Morel and G. Yu. ASIFT: A new framework for fully affine invariant image comparison. In *SIAM Journal on Imaging Sciences*, Vol.2, No.2, pp. 438–469, 2009.
- [8] J. Pfeil, K. Hildebrand, C. Gremzow, and M. Alexa. Throwable panoramic ball camera. In *SIGGRAPH Asia 2011 Emerging Technologies*, No.4, 2011.
- [9] H. Taira, Y. Inoue, A. Torii, and M. Okutomi. Robust Feature Matching for Distorted Projection by Spherical Cameras. In *IPSJ Transactions on Computer Vision and Applications*, Vol.7, pp. 84–88, 2015.

WISS2016 採録判定時のコメント

採録区分：ロング採録

判断理由：

移動&回転しているボールから撮影された全天球動画の視点固定のためのアルゴリズムの提案です。スポーツやエンタテインメント等の分野において今後高い応用可能性を秘めた研究であると考えられます。特徴点のロバストな抽出という技術的新規性に加え、しっかりした実装が成されており、ロング採録となりました。一方、アルゴリズムの説明で理解が難しい個所があるので、より丁寧な記述が望まれます。

レビューサマリ：

全体の構成：

ボールに内蔵した全天球カメラ映像を活用することで、球技などに新しい視聴体験をもたらす可能性があります。研究は、回転するボールにおいて、どのようにして視点を固定するかという問題に取り組んでいます。

評価すべき点：

- ・新規性の主なものは、特徴点のロバストな抽出方法の提案にあります。全天球動画の回転の補正手法について正距円筒図法においては高緯度領域の画像に歪みが生じることから、緯度変換を行って高緯度領域の画像を一度低緯度に移動させて画像処理（特徴点抽出）を行い、その後再度、高緯度領域に移動させることで歪みの影響を抑えるものです。このアルゴリズムでは、ボールの回転軸の変化に加え、高速な縦方向の回転にも対応できる回転補正処理が実現できている点が、先行研究に対する新規性の点で評価できます。
- ・論文では THETA を用いて撮影された動画を用いていましたが、提案アルゴリズムは他の種類のカメラで撮影した全天球動画にも対応できるという点で、汎用性に点からも評価できます。

改良のためのコメント等：

- ・「隣接するフレーム間の時間は微小であるので、全天球カメラの並進による特徴点の見え方の変化はない」と述べていますが、「特徴点の見え方」が具体的には何を表しているのか不明です。
- ・4 節の画像処理アルゴリズムに関して、もう少し丁寧に（数式等を入れて）説明されると良いと思います。
- ・図 4 の行列 R, frame n について、文章中に説明を追記したほうが良いと思います。
- ・概要に述べられている「8 点アルゴリズム」については説明が無く、理解が困難です。こちらもオリジナルな方法であるとすればその利点や欠点を含めて説明、議論してください。
- ・図 3 の処理による効果分かる動画があれば、提案手法の効果がより明確になります。
- ・おそらく将来的にはボールの視点固定はリアルタイムで行えるべきと思います。提案法と従来法について計算量の比較や、高速化の可能性などについても論じていただけるとありがたいです。
- ・「任意視点ならでは」の面白いアプリケーション例が欲しいです。

本論文に対する各査読者の詳しいコメントは以下のページを参照のこと：

<http://www.wiss.org/WISS2016Proceedings/oral/21.html>

***本ページは論文本体ではありません**

【全天球ボールカメラによる撮影動画の視点固定手法の提案】