

全天球カメラを用いた視線推定技術の提案

小西 優実^{1,a)} 益子 宗^{2,b)} 落合 裕美^{2,c)} 小池 英樹^{1,d)}

概要：近年，全天球カメラの普及が進んでおり，視覚センサとしての利用もされつつある．また，実世界での視線を用いた対話システムは，その有用性から長らく研究されてきた．これら 2 つを組み合わせることで，簡易な設備による，直感的かつ素早い対話的なシステムを構築することが可能であるということが想定できる．本稿は全天球カメラを利用した外観ベースの手法を用いた視線計測のシステムの提案およびその応用例についての考察を行う．

1. はじめに

近年，仮想現実技術の急速な発展に伴い，全天球パノラマ映像，およびそれを撮影することができる全天球カメラの普及が進んでいる．全天球カメラは，単にコンテンツ映像撮影の用途として使われるのみならず，テレビ会議システム [1]，ロボットのセンサー [2] などに用いられている．全天球カメラがこのように用いられる理由としては，一台のカメラで周囲全体の環境の情報を取得できるという点が挙げられる．

一方，実世界での視線を用いたインタラクティブなシステムについての研究は古くから多く提案されてきた．Pfeuffer らによるタブレットにおける視線およびタッチの組み合わせによるインタラクション [3]，Esteves らによるスマートウォッチの操作における視線情報を利用した Orbits[4] などがその例である．こうした研究はいずれも，視線情報を直感的かつ素早いシステムとのインタラクションの手段として用いている．

これら 2 つを組み合わせることで，簡易な設備による，直感的かつ素早い対話的なシステムを構築することが可能であるということが想定できる．加えて，従来のシステムは視線計測および周囲環境情報の取得を別々に行っていたのに対し，全天球カメラは一台でユーザとユーザの周囲の環境の双方の情報を取得できることから，ユーザの視線方向だけでなく，ユーザが見ている対象物も合わせて推定できるという利点が考えられる (図 1)．

そこで本研究では，全天球カメラを利用した視線計測のシステムを提案するとともに，その応用例についての考察を行う．

2. 先行研究

2.1 全天球カメラを用いたインタラクションシステム

本研究に類似する全天球カメラを用いたインタラクティブシステムは，これまでもいくつか提案されている．具体的には，Li らの，全天球カメラを用いたビデオ会議システム OmniEyeBall[5] や，双方向のジェスチャーを認識できるシステム OmniGesture[6] が挙げられる．いずれのシステムも，全天球カメラによる入力および，複数方向へのフィードバック提示を可能とする球面，円柱形ディスプレイを搭載している．

全天球カメラと視線を用いたインタラクションシステムとしては，Shiro らの OmniGaze がある [7]．このシステムはテレビ会議システムであり，一方のデバイス A は平面ディスプレイおよび視線計測装置，他方のデバイス B は全天球カメラおよび球体ディスプレイからなる．A で計測した視線位置情報を B のディスプレイに提示することで，B 側のユーザ達は A 側のユーザが見ているユーザが誰であるのかを把握することができる．OmniGaze は，全天球カメラの情報ではなく既存の視線計測装置を用いて視線を計測している点で本研究とは異なるものである．

2.2 視線推定技術

Hansen らによれば，視線推定技術は大別すると特徴量ベース，モデルベース，外観ベースの三種類の手法が提案されている [8]．特徴量ベースの手法とは目に関する特徴量から直接視線方向を推定する方法，モデルベースの手法は

¹ 東京工業大学

² 楽天株式会社 楽天技術研究所

a) konishi.y.ae@m.titech.ac.jp

b) so.masuko@rakuten.com

c) hiromi.ochiai@rakuten.com

d) koike@c.titech.ac.jp



図 1 全天球カメラを用いた視線検出システム想定．赤点が過去の視線位置，緑点が現在視線位置，青枠が対象物を表す．

目に関する特徴量を眼球の 3D モデルに当てはめ視線方向を推定する方法，そして外観ベースとはカメラ画像から機械学習技術を用いて直接視線方向を推定する方法である．このうち，近年の深層学習技術の発展およびデータセット蓄積の結果，精度向上が行われているのは外観ベースの技術である．加えて Zhang らは，OpenGaze という外観ベースの手法を用いた視線推定ライブラリを提供している [9]．そこで本研究では，OpenGaze をベースとした全天球映像の視線推定を提案する．

3. 提案システム概要

本研究の提案システムの概要を図 2 に示す．本システムは全天球映像から顔部分を抽出し，ライブラリを利用して視線方向推定を行う．推定された視線方向を元にユーザーのしている位置の推定を行う．本システムは，ゲームエンジン Unity を利用して実装を行った．

まず，全天球カメラからの映像に対し Dlib Facemark Detector を利用した顔検出を行う．検出結果をもとに顔抽出を行い，抽出された顔画像から OpenGaze を用いて視線方向を推定する．顔抽出は，全天球映像から直接抽出するのではなく，仮想空間上の天球に投影した全天球カメラ映像を，天球中心に設置した仮想カメラで撮影することで行う (図 3)．これにより全天球映像を用いつつ，パースペクティブカメラによる撮影を擬似的に再現することができるため，全天球映像を用いた学習データを用意する必要がないと考えた．

全天球映像上でのユーザの視線位置は，天球上の顔位置 \vec{F} および推定された視線方向ベクトル \vec{d} にて表される半直線と，天球との交点 \vec{P} を求めることで推定する (図 4)． \vec{P} は以下のように導出される：

$$\vec{P} = \vec{F} + t\vec{d} \quad (1)$$

$$t = \frac{-\vec{F} \cdot \vec{d} + \sqrt{\vec{F} \cdot \vec{d} - (\vec{d} \cdot \vec{d})(\vec{F} \cdot \vec{F} - R^2)}}{\vec{d} \cdot \vec{d}} \quad (2)$$

ただし， R は天球の半径を表す．

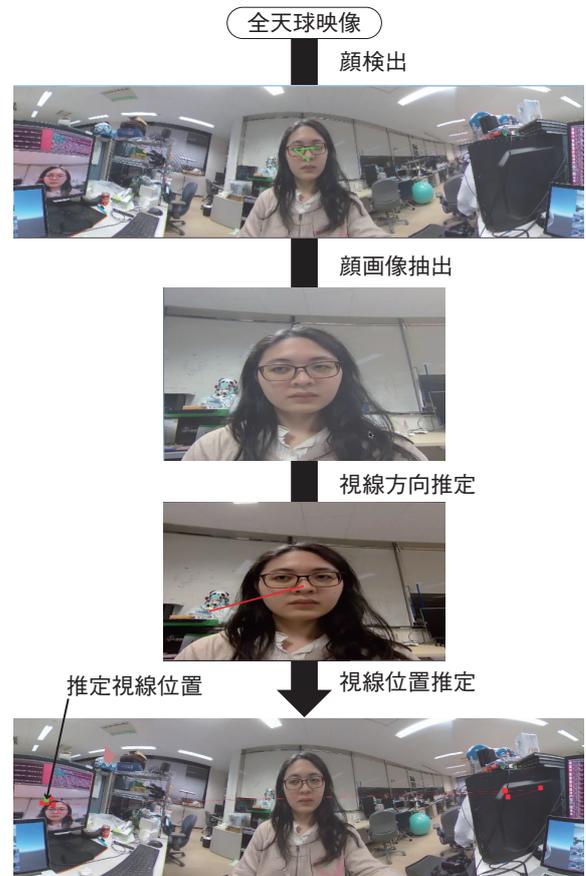


図 2 システム概要図

4. 応用例

具体的な応用例としては，まずは先述のテレビ会議システムへの組み込みが考えられる．また，ユーザの視線を，対象物への注目とみなすことで，その情報を利用したフィードバックシステムが考えうる．具体例としては，システムを店舗へ設置し，商品の注目度に応じたサイネージ等のデジタルコンテンツの動的変化などが挙げられる．

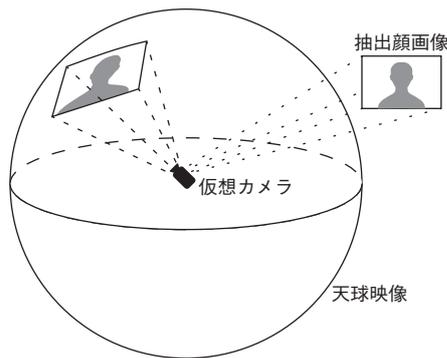


図 3 顔抽出

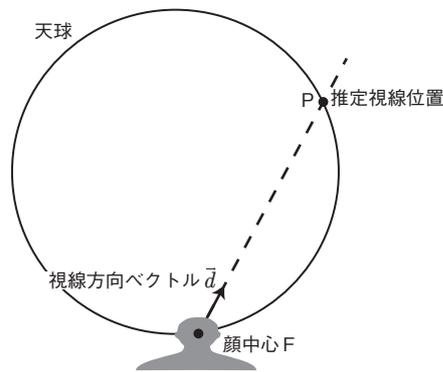


図 4 視線位置推定

5. 今後の課題と展望

本稿地点での実装は、1 人のみの視線推定を行っている。しかしながら本稿で用いられている技術は全て複数人に対応できるものであるため、理論上は複数人の視線推定を同時に行うことができる。今後、仮想カメラの数を増やすことで、多人数対応を実現したい。

また、提案システムをアプリケーションに組み込むことに際して、推定視線の精度評価は不可欠である。ユーザの顔と対象物との位置関係、および顔とカメラとの位置関係の 2 つの実験条件に応じた精度の評価を行いたい。

精度によっては、その向上のために学習済みモデルの再トレーニングを行うことが考えられる。OpenGaze は学習データとして MPIIGaze を用いているが、よりデータ量および視野角の広いデータセットはいくつか公開されている [10][11]。こうしたデータセットを用いた再学習、あるいはアプリケーションに応じた最適化を施すことが重要であると考えられる。

参考文献

- [1] OWLLabs(online), <https://www.owllabs.com> (Accessed on 2019.12.12).
- [2] Libor Spacek and Christopher Burbridge, "Instanta-

- neous robot self-localization and motion estimation with omnidirectional vision," *Robotics and Autonomous Systems*, vol. 55, no. 9, pp. 667-674, 2007.
- [3] Pfeuffer, Ken and Alexander, Jason and Chong, Ming Ki and Gellersen, Hans, "Gaze-touch: Combining Gaze with Multi-touch for Interaction on the Same Surface," *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, pp. 509-518, 2014.
- [4] Esteves, Augusto and Velloso, Eduardo and Bulling, Andreas and Gellersen, Hans, "Orbits: Gaze Interaction for Smart Watches Using Smooth Pursuit Eye Movements," *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, pp. 457-466, 2015.
- [5] Li, Zhengqing and Miyafuji, Shio and Sato, Toshiki and Koike, Hideki, "OmniEyeball: Spherical Display Embedded With Omnidirectional Camera Using Dynamic Spherical Mapping," *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 193-194, 2016.
- [6] Kim, Heesun and Lee, Dongeon and Kim, Min Gyeong and Jang, Hyejin and Park, Ji-Hyung, "Omni-Gesture: A Hand Gesture Interface for Omnidirectional Devices," *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*, pp. 44:1-44:2, 2018.
- [7] Shiro, Keisuke and Okada, Atsushi and Miyaki, Takashi and Rekimoto, Jun, "OmniGaze: A Display-covered Omnidirectional Camera for Conveying Remote User's Presence," *Proceedings of the 6th International Conference on Human-Agent Interaction*, pp. 176-183, 2018.
- [8] D. W. Hansen and Q. Ji, "In the Eye of the Beholder: A Survey of Models for Eyes and Gaze," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 478-500, 2010.
- [9] Xucong Zhang and Yusuke Sugano and Andreas Bulling, "Evaluation of Appearance-Based Methods and Implications for Gaze-Based Applications," *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2019.
- [10] Funes Mora, Kenneth and Monay, Florent and Odobez, Jean-Marc, "EYEDIAP: a database for the development and evaluation of gaze estimation algorithms from RGB and RGB-D cameras", *Proceedings of the Symposium on Eye Tracking Research and Applications*, pp. 255-258, 2014.
- [11] P. Kellnhofer*, A. Recasens*, S. Stent, W. Matusik and A. Torralba, "Gaze360: Physically Unconstrained Gaze Estimation in the Wild," *IEEE International Conference on Computer Vision (ICCV)*, 2019