

## "First Person Vision"への挑戦

亀田 能成<sup>†1</sup> 藤吉 弘亘<sup>†2</sup> 羽下 哲司<sup>†3</sup> 山口 晃一郎<sup>†4</sup> 中村 裕一<sup>†5</sup>

<sup>†1</sup> 筑波大学 大学院工学研究科 〒305-8573 茨城県つくば市天王台 1-1-1 3M304

<sup>†2</sup> 中部大学 工学部情報工学科 〒487-8501 愛知県春日井市松本町 1200

<sup>†3</sup> 三菱電機(株) 自動車機器開発センター 〒669-1513 兵庫県三田市三輪 2-3-33

<sup>†4</sup> (株)豊田中央研究所 〒480-1192 愛知県愛知郡長久手町長湫横道 41-1

<sup>†5</sup> 京都大学 学術情報メディアセンター 〒606-8501 京都府京都市左京区吉田本町

E-mail: <sup>†1</sup> kameda@iit.tsukuba.ac.jp, <sup>†2</sup> hf@cs.chubu.ac.jp <sup>†3</sup> Haga.Tetsuji@cj.mitsubishielectric.co.jp

<sup>†4</sup> yamaguchi@mosk.tytlabs.co.jp <sup>†5</sup> yuichi@media.kyoto-u.ac.jp

**あらまし** 近年、カメラセンサを人の頭部等に装着し、その人の行動意図を認識しより精密に行動支援を実現しようとする取り組みが、人間中心型のIT技術 First Person Vision として改めて注目を集めている。国際会議 CVPR2009 では、Egocentric Vision という第一回目のワークショップが開催され、そこで First Person Vision に関する問題提起から様々な取り組みまで紹介された。本セッションでは、改めて First Person Vision の本質を見つめ直し、問題の定義からその応用先について議論する。また、それを支える画像生成・計測・認識技術や、ウェアラブルカメラを用いた新しい方式・システムとその応用事例についてアイデア段階から実用化の手前の研究まで幅広く取り上げる。

**キーワード** Egocentric Vision, 意図理解, 行動認識, ウェアラブルカメラ

## Challenge to “First Person Vision”

Yoshinari KAMEDA<sup>†1</sup> Hironobu FUJIYOSHI<sup>†2</sup> Tetsuji HAGA<sup>†3</sup> Koichiro YAMAGUCHI<sup>†4</sup>  
and Yuichi NAKAMURA<sup>†5</sup>

<sup>†1</sup> University of Tsukuba, 3M304, 1-1-1 Tennoudai, Tsukuba, Ibaraki, 305-8573

<sup>†2</sup> Chubu University, 1200 Matsumoto-cho, Kasugai, Aichi, 487-8501

<sup>†3</sup> Mitsubishi Electric Corp., Automotive Electronics Development Center, 2-3-33 Miwa, Sanda, Hyogo, 669-1513

<sup>†4</sup> Toyota Central R&D Labs., Inc., 41-1 Yokomichi, Nagakute, Nagakute-cho, Aichi-gun, Aich, 480-1192

<sup>†5</sup> Kyoto University, Yoshidahonmachi, Sakyo-ku, Kyoto, 305-8573

E-mail: <sup>†1</sup> kameda@iit.tsukuba.ac.jp, <sup>†2</sup> hf@cs.chubu.ac.jp <sup>†3</sup> Haga.Tetsuji@cj.mitsubishielectric.co.jp

<sup>†4</sup> yamaguchi@mosk.tytlabs.co.jp <sup>†5</sup> yuichi@media.kyoto-u.ac.jp

**Abstract** A head mounted camera or a human borne camera could be served for estimating a user's action and for realizing better life by supporting his/her action. This research domain is coming to be in focus of our society because it is one of the key technologies of human-centered IT, and it is called “First Person Vision” recently. The First Workshop of Egocentric Vision was held in conjunction with CVPR2009 and wide research topics including new problem statement and novel methodologies were presented. In this session, we revisit the concept of “First Person Vision” and discuss it from wide viewpoints, including its definition and its applications. In addition, new computer graphics, computer vision, image understanding that are suitable for First Person Vision will be introduced and newly developed systems that exploit a head-mounted camera will be explored.

**Keyword** Egocentric Vision, Intent Understanding, Action Recognition, Wearable Camera

### 1. はじめに

コンピュータビジョンの分野の中でも、人間の頭部や胸部等にカメラを搭載して、その映像からその人間

の行動を理解したり周囲の状況を認識したりしようという研究はこれまでも様々な取り組みが行われてきた。本稿では1人称視点のカメラを利用した画像処理および

びそれに基づく人間行動支援の枠組みを **First Person Vision** と呼ぶ。

社会的にみても、高齢化社会やバリアフリー社会を目指す上で、人間中心型のIT技術はさらに重要性を増していくものと考えられる。CVPR2009 併催の **First Workshop on Egocentric Vision**[1]が開催されたことから、この分野が改めて注目を集めていることがわかる。

これまでこの研究分野は、行動認識、状況理解、Lifelog、障害者支援など様々な分野に分かれていた。そこで、本セッションでは、1人称視点を扱う研究という観点からこれらをまとめて概観し、**First Person Vision** の定義や共通項目、さらには期待される未来像や応用について公開討論形式で議論する。

## 2. First Person Vision の研究が狙うもの

### 2.1. 上層

1人称カメラを装着している人間の認知や行動、さらにはその社会的影響に関する項目をここでは列挙する。

- 利用者の視界を常にカバーすることで利用者が見ている情報を完全に把握する。
- 利用者の生活行動モデルを学習し、それに基づいて行動支援を行う。
- 視覚障害者に対して、視覚機能を何らかの形で提供する。
- 歩行障害者に対して、移動装置として位置推定と周囲の安全確認を行い、移動ナビゲーションを提供する。
- 利用者の生活プライバシーへの配慮。
- 1人称カメラに写る周囲の被写体のプライバシーへの配慮。
- 支援情報の提示方法。一つは複合現実感などを利用した視覚的提示が考えられ、また一つにはそれ以外の聴覚・触覚に対する提示が考えられる。
- 画像認識から行動理解、およびその支援までを含めた被験者実験評価による人間の視覚心理の解明。

### 2.2. 下層

ここでいう下層とは、装着している人間の認知や行動に踏み込まないという概念である。

- 日常生活中で常に装着することが負担にならない超軽量カメラの開発。
- 利用者が装着を肉体的・心理的に負担に感じない人体上の装着位置。
- 生活行動支援における計算機処理能力の削減。ウェアラブルデバイス上で完結することが望ましい。
- 人間向けに表示された標識や文字の認識。
- 1人称映像として視覚的に無意味な時間区間の高

速判定。

- 眼球画像観測による視線方向追跡。
- 1人称カメラのリアルタイムレジストレーション。

## 3. First Person Vision の問題点

- 1人称カメラシステムでは、現時点では、人間の眼ほどのハイダイナミックレンジ・広視野角・立体視・リアルタイムを同時に実現できない。そのため用途を限定する必要がある。
- 1人称カメラは人間に搭載されるため、遥動対策が不可避である。暗所ではぶれが頻発することが想定される。
- 生活の場で映像が撮影されるため、想定外の物体に遭遇する可能性が常に存在する。学習データに完全性を期待することはできない。
- 1人称カメラで得られる日常生活の映像を構造化していくときに、明確なルールが存在しない。これはスポーツ撮影など目的をもって収集された映像に対する検索と大きく異なる点である。
- 行動理解のためには、他のウェアラブルセンサとの連携が必須である。ウェアラブルセンサは、人間の内部状態を直接図ろうとするものと、1人称カメラに代わって外界と利用者との関係を観測するものとに大別される。

## 4. First Person Vision を支える技術

- 全般に映像検索技術は有用である。
- SIFT, SURF のような光量変動に強い特徴量。画像平面上だけでなく空間方向にも展開するコンパクトで低計算量の局所特徴量の出現が待たれる。
- 長時間大量のデータを扱うためのデータマネジメント技術。
- 大量データに対して高速に動作する Hash 法や ANN, KD ツリーなどのカテゴリ認識・データ検索技術。

## 5. First Person Vision が導く未来

**First Person Vision** は、カメラを装着した自身の行動支援を目的とし、人間に近い、人間に寄り添う技術である。それだけに、研究としてどこまでも性能・効率を追求するだけでなく、私達人間にとって心地よく使える技術を実現するよう留意すべきであろう。

## 文 献

- [1] First Workshop on Egocentric Vision (in conjunction with CVPR2009), 2009.  
<http://www.seattle.intel-research.net/egovision09/>