

佐藤 洋一

東京大学生産技術研究所・教授

「日常生活空間における人の注視の推定と誘導による情報支援基盤の構築」

§1. 研究実施体制

(1) 佐藤グループ

- ① 研究代表者:佐藤 洋一 (東京大学・生産技術研究所、教授)
- ② 研究項目
 - 環境埋め込みカメラによる視線計測技術の開発
 - 不完全な視線情報からの注視推定技術の開発
 - 視覚刺激に対する注視反応モデルの構築
 - 視覚刺激による注視誘導技術の開発

(2) 杉本グループ

- ③ 研究代表者:杉本 晃宏 (国立情報学研究所・コンテンツ科学研究系、教授)
- ④ 研究項目
 - 環境埋め込みカメラによる視線計測技術の開発
 - 視覚刺激に対する注視反応モデルの構築
 - 視覚刺激による注視誘導技術の開発

(3) 久野グループ

- ① 研究代表者:久野 義徳 (埼玉大学・大学院理工学研究科、教授)
- ② 研究項目
 - ロボットの身体動作に対する注視反応モデルの構築
 - ロボット動作に対する注視反応モデルに基づいた注視誘導技術の開発

(4) 小池グループ

- ① 研究代表者:小池 英樹 (電気通信大学・大学院情報システム学研究科、教授)

② 研究項目

- 視覚刺激による注視誘導技術の開発
- 視覚適応型インタフェースの開発

§ 2. 研究実施内容

(文中に番号がある場合は(3-1)に対応する)

環境埋め込みカメラによる視線計測技術の開発

本年度は学習時に未知の姿勢における視線推定を実現するために、任意の頭部姿勢に対応する目画像を疑似的に合成する手法を提案した[16]。すなわち、基準となる頭部姿勢で獲得した学習目画像群の見え方が異なる頭部姿勢下ではどのように変化するかを推定し、これを元に実際には撮影されていない頭部姿勢での学習データを生成する。このような画像変形を実現するための画素フローを、学習時頭部姿勢以外で撮影された4枚の追加画像のみを用いて推定し、頭部視線変動を許容した他の手法と比べ、大幅に少ない学習画像を利用しながら同程度かそれ以上の視線推定精度を実現している。

不完全な視線情報からの注視推定技術の開発

これまで、視覚的顕著性を学習に利用することで明示的なキャリブレーション作業を行うことなく視線推定を行う手法を開発してきたが、本年度は更に推定フローの中で顕著性モデルの性能改善を行う手法を提案した[2]。顕著性マップは色や輝度など様々な基礎特徴から計算されたマップを足し合わせることで算出されているが、一度推定された注視点座標に最も一致するようにこの統合重みを最適化することで、顕著性マップ自体の注視点予測性能が向上し、さらにはこのフィードバックループによって最終的な注視点推定性能も向上することが期待できる。その他、推定全体のパラメータ最適化手法を新たに提案することで、平均誤差3.5度での視線推定を実現している。

視覚刺激に対する注視反応モデルの構築

これまでに提案されてきた視覚的顕著性マップの計算モデルにとって、視線移動順序を予測することの困難さが一つの課題として知られている。研究[11]では、実際に計測した人間の視線移動の大規模なデータベースを用いたデータ駆動型のアプローチによりこれを実証し、視野特性と視覚的顕著性との関わりを明らかにした。一方、人間の注視行動はその時人物がどのような作業に従事しているかにも深く関連している。一人称視点の状況において頭部自己運動と視線運動が得られた場合、これらの情報から得られる注視の動きから逆に、その時対象人物が行っている作業を識別することができる可能性がある。研究[7]ではこの仮説に基づき、自己運動と視線運動を用いた自己動作認識手法を提案した。

ロボットの身体動作に対する注視反応モデルの構築

これまで、注視誘導を行うためには、注視獲得、アイコンタクト、注視移動の3つの行動が必要なことを明らかにしてきた[5]。今年度は、対象者以外の人物も存在する複数人環境での実験で提案モデルの有効性を確認した[12]。また、ロボットと相手の人間の位置関係に応じて注視獲得を行う適切な行動を実験的に明らかにした[14]。さらに、相手は何らかの作業に集中しているときに注視を獲得しようとするのは、相手の作業を妨害することになるので、相手の行動を認識してどのようなタイミングで注視獲得行動を始めると良いかについても検討した[13]。以上により、相手との位置関係や相手の行っている行動などに応じて適切な非言語行動を用いて注視誘導を行う基本的なモデルが構築できた。

ロボット動作に対する注視反応モデルに基づいた注視誘導技術の開発

前項目で構築されたモデルをもとに、実際に注視誘導を行うことのできるロボットを開発する。アイコンタクトや共同注視を実現するためには、ロボットが何を見ているかを読み取れるように、ロボットの視線の読み取りやすさが重要である。また、日常生活で人と共存するロボットでは、外見的な親しみやすさも重要である。そこで、被験者を用いた実験により、視線の読み取りやすさと親しみやすさを兼ね備えたロボットの目のデザインを明らかにした[18]。そして、その結果に基づき、小型プロジェクタにより投影した視線表現を行う目と、超小型カメラによる外界を見る目を組み込んだロボット頭部を開発した。

視覚刺激による注視誘導技術の開発

昨年度までに、画像中の指定領域の視覚的顕著性を高めることで、人の注視を自然に誘導する手法を提案した。そこでは、指定領域内の視覚的顕著性を高め、領域外の視覚的顕著性を下げするために、入力画像の各ピクセルの RGB 値を独立に変更する操作によって画像を加工している。そのため、注視自体は誘導できるものの、加工画像と入力画像との間に違和感が生じるという問題があった。そこで本年度は、画像加工によって生じる違和感を軽減するために、画像の加工に画像の輝度値を補正する操作を取り入れる手法に拡張した。具体的には、入力画像と加工画像との間で色ヒストグラムになるべく差が出ないように輝度値の補正を画像全体で一括して行う。これによって、指定領域の顕著性を高めつつ違和感のない画像加工を実現した。

視線適応型インタフェースの開発

ディスプレイに表示される映像やスクリーンに投影される映像など、視聴者の位置に関することなく一方的に表示される映像には、表示領域やコンテンツのサイズに応じて最適な視聴距離がある。本年度は、モザイク画像の特徴を応用することによって、遠近 2 つの領域にいる視聴者に対してそれぞれ異なった情報を同時に提供できる画像の生成方法と、当該画像の視認性向上処理の提

案・実装を行った[17]。

§3. 成果発表等

(3-1) 原著論文発表

● 論文詳細情報

1. M. Yousuf, Y. Kobayashi, Y. Kuno, K. Yamazaki and A. Yamazaki, "A Mobile Guide Robot Capable of Establishing Appropriate Spatial Formations," IEEJ Transactions on Electronics, Information and Systems, vol.133-C, no.1, January 2013. (10.1541/ieejieiss.133.28)
2. Y. Sugano, Y. Matsushita, and Y. Sato, "Appearance-based Gaze Estimation using Visual Saliency," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), vol. 35, no. 2, pp. 329-341, February 2013. (10.1109/TPAMI.2012.101)
3. Tam T. Le , Yousun Kang, Akihiro Sugimoto: Image Categorization Using Hierarchical Spatial Matching Kernel, The Journal of the Institute of Image Electronics Engineer of Japan, Vol. 42, No. 2, pp.214-221, March 2013.
4. D. Thomas and A. Sugimoto: Range Image Registration using a Photometric Metric under Unknown Lighting, IEEE Trans. on Pattern Analysis and Machine Intelligence (accepted)
5. M. Moshiri, T. Onuki, Y. Kobayashi and Y. Kuno, "Effect of Robot's Gaze Behaviors for Attracting and Controlling Human Attention," Advanced Robotics (accepted).
6. Y. Sugano, Y. Matsushita and Y. Sato, "Graph-based Joint Clustering of Fixations and Visual Entities", in ACM Transactions on Applied Perception (TAP), (accepted).
7. K. Ogaki, K. Kitani, Y. Sugano, and Y. Sato, "Coupling Eye-Motion and Ego-Motion Features for First-Person Activity Recognition", Proc. IEEE Workshop on Egocentric Vision (in conjunction with CVPR2012), pp. 1-7, June 2012. (10.1109/CVPRW.2012.6239188)
8. D. Thomas, Y. Matsushita and A. Sugimoto: Robust Simultaneous 3D Registration by Rank Minimization, Proc. Symposium on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT2012), pp. 33-40, June 2012. (10.1109/3DIMPVT.2012.15)
9. M. Hoque, D. Das, T. Onuki, Y. Kobayashi and Y. Kuno, "Robotic System Controlling Target Human's Attention," International Conference on Intelligent Computing (ICIC2012), Lecture Notes in Computer Science, vol.7390, pp.534-544, July 2012. (10.1007/978-3-642-31576-3_68)

10. M. Yousuf, Y. Kobayashi, Y. Kuno, A. Yamazaki and K. Yamazaki, "Development of a Mobile Museum Guide Robot That Can Configure Spatial Formation with Visitors," *Intelligent Computing Technology, International Conference on Intelligent Computing (ICIC2012)*, Lecture Notes in Computer Science, vol.7389, pp.423-432, July 2012. (10.1007/978-3-642-31588-6_55)
11. 窪田秀行、菅野裕介、岡部孝弘、佐藤洋一、杉本晃宏、開一夫、"人間の視野特性を考慮した学習に基づく視覚的顕著性モデル"、画像の認識・理解シンポジウム(MIRU2012), August 2012.
12. M. Hoque, D. Das, T. Onuki, Y. Kobayashi and Y. Kuno, "Model for Controlling a Target Human's Attention in Multi-Party Settings," *Proc. Ro-Man2012*, September 2012. (10.1109/ROMAN.2012.6343797)
13. D. Das, M. Hoque, T. Onuki, Y. Kobayashi and Y. Kuno, "Vision-Based Attention Control System for Socially Interactive Robots," *Proc. Ro-Man2012*, September 2012. (10.1109/ROMAN.2012.6343800)
14. M. Hoque, D. Das, T. Onuki, Y. Kobayashi and Y. Kuno, "An Integrated Approach of Attention Control of Target Human by Nonverbal Behaviors of Robots in Different Viewing Situations," *Proc. IROS2012*, pp.1399-1406, October 2012. (10.1109/IROS.2012.6385480)
15. Y. Sugano, K. Harada and Y. Sato, "Touch-consistent perspective for direct interaction under motion parallax", in *ACM ITS2012 (Interactive Tabletops and Surfaces Conference)*, November 2012. (10.1145/2396636.2396695)
16. F. Lu, Y. Sugano, T. Okabe, and Y. Sato, "Head Pose-Free Appearance-Based Gaze Sensing Via Eye Image Synthesis", in *21st International Conference on Pattern Recognition (ICPR2012)*, November 2012.
17. 鈴木裕太, 小池英樹、モザイク画像を利用した遠近画像の同時提示手法、日本ソフトウェア科学会インタラクティブシステムとソフトウェア(WISS2012)、December 2012
18. T. Onuki, T. Ishinoda, Y. Kobayashi and Y. Kuno, "Designing Robot Eyes for Gaze Communication", *Frontiers of Computer Vision (FCV2013) Workshop*, January 2013.