

(エリアⅡ) 人を観る

エリア成果概要

【成果の概要】

「人を観る」エリアでの成果は以下の通りである：

1. 装着型視覚センサを用いた人物の運動軌跡推定－杉本晃宏(NII)・松山隆司(京大)
2. 自由変形を考慮した頭部姿勢追跡－佐藤洋一(東大) 松山隆司(京大)
3. 人間の視線検出に基づく高精細情報獲得と人の行為の認識－松山隆司(京大)・杉本晃宏(NII)・鷺見和彦(京大)
4. 複数指先軌跡の実時間計測にもとづくジェスチャ認識－佐藤洋一(東大)・松山隆司(京大)
5. 人間の状態と行動の解析－松山隆司・鷺見和彦(京大)・松山隆司(京大)
6. 複数センサからの観測情報と環境モデルの統合による人物追跡－佐藤洋一(東大)・杉本晃宏(NII)・松山隆司(京大)
7. ユーザインタフェースのための実時間手形状計測－有田大作・谷口倫一郎(九大)
8. ビジョン技術による実時間身体動作解析の高精度化－有田大作・谷口倫一郎(九大)
9. 伝統舞踏の獲得－池内克史(東大)
10. 装着型体験記憶装置の情報を用いた、体験記憶再生方式に関する研究－上岡玲子・廣瀬通孝(東大)
11. 複数話者追跡法の開発とインタラクションへの応用－奥乃博(京大)
12. 混合音分離と音声認識を統合したミッシングフィーチャ理論による音声認識手法の開発－奥乃博(京大)
13. 話し言葉音声認識のモデルと認識技術－河原達也(京大)
14. 音声の音響的普遍構造を用いたパラ言語情報の推定－峯松信明・広瀬啓吉(東大)
15. 音声の音響的特徴を用いた知覚的年齢の推定とその高精度化－峯松信明・広瀬啓吉(東大)
16. 事象関連電位を用いたヒトと非ヒト認知に関する研究－開一夫・國吉康夫(東大)

これらの成果は、人の「何を」観るか、という観点から以下のように分類される(2005年度柱会議)。

- 外部状態－身体動作(5～9)、行為・動作(1～4, 10～16)
- 内部状態－意図・感情(5, 13)、記憶(9)、属性(年齢等、13～15)

また、どのようなメディアを利用するのか、という観点から以下のように分類できる。

- メディア処理－画像処理(1～10)、音響処理(11～15)、脳波(16)

さらに同一メディアであってもマルチモーダル処理を行っているものも少なくない。

各々の成果が対象としている応用は、以下のようにまとめられる。

- 新しい技術に基づいた従来型マルチメディア・モーダルインタフェースの拡張(5, 7, 12～15)
- 体験記憶・ライフログ(1, 2, 6, 8, 10)
- ロボットの挙動生成(9)、追跡(11)
- 脳内処理の解明(16)

さらに、新たな処理アルゴリズムとしては、以下のような特徴があった。

- 複数メディアの情報統合－音響と画像のストリーム統合(11)
- 複数モダリティの情報統合－画像メディア(6)、音声メディア(12～15)
- 統合のレベル手法－パーティクルフィルタ、カルマンフィルタ、画像・音響ストリーム統合
- 実時間処理(6～8, 11)

このように、エリア2では、「人を観る」ための多数の基礎技術の開発が行われた。特に「複数のセンサ情報を統合した実時間複数人物(話者)追跡システムの開発」は人を観る性能を大幅に改善し、デモによりその有効性を確認している。

【今後の展開】

今後、これら個々の成果を生かして、人との共生を追求していくためには、単一メディア、モダリティに着目して、その認識性能を向上させるだけでなく、それらを統合して、お互いの欠点を補完しあい、より高性能化を目指すために、情報統合、および、それらの実時間処理が不可欠と考えられる。さらに、新たな技術を、上記のような応用に加えて新たな分野に展開することにより、着実な進展をはかっていく必要がある。今後の主たる技術的な課題を列挙する：

- 複数人を観る－見え隠れ、登場・退場、交代発話、同時発話
- 五感を使用して観る－画像、音響、脳波、触覚、呼吸、嗅覚・・・
- 場所を超えて観る－カメラ(観察者に装着、被観察者に装着、環境に設置)、マイクロフォン(話者に装着、聞き手に装着、環境に設置)、磁気センサ、脳波、触覚、GPS、・・・
- 時間を超えて観る－アーカイブされたデータ、現時点のデータ、シミュレートされたデータ、等の統合
- 観たものを伝達－動作・状態のシンボル化による情報伝達。体験記録再現のアノテーション。
- 無意識下の行動を観る－脳波、非接触センサ、LifeLog。
- 誤りを前提とした処理－フィードバック処理、階層的でロバストなインタフェース設計論。

【本エリアの代表的な成果】

環境センサと装置センサとをパーティクルフィルタで情報統合した、高精度な複数人物追跡を実時間で実現した(6)