

非線形写像学習アルゴリズム PaLM-tree を用いた顔方向推定

Face Direction Estimation using Nonlinear Map Learning Algorithm "PaLM-tree"

佐藤 哲[†]Tetsu Satoh[†]和田俊和^{††}Toshikazu Wada^{††}中村恭之^{††}Takayuki Nakamura^{††}

1 はじめに

情報通信研究機構が推進している「ゆかりプロジェクト (UKARI: Universal Knowledgeable Architecture for Real-Life appliances)」[1] の目標の一つは、住居内で人間の行動を観測・学習して行動意図を推測し、システムから積極的にサービスを提供する能動的な生活支援パートナーを開発することである。そのため本研究では、人間は関心が高い方向に顔を向けると仮定し、行動意図推測の手がかりとして顔方向を推定するシステムを提案する。提案手法では、人物を撮影している複数の画像センサにより得られる情報と、被験者の顔の向きの情報を、非線形写像学習アルゴリズム PaLM-tree を用いて学習することで顔方向を推定する。

2 複数カメラを用いた顔方向推定

画像センサを用いて顔方向を推定するには、目や鼻などの部位から特徴点を抽出し、位置関係を考慮して方向を推定する手法や、画像と方向を対応付けて学習することで方向を推定する手法など多くの研究がある [2]。これらの手法は、モデルを準備する必要があったり、大量の顔画像を準備する必要があるなど、知識ベース処理に特有の汎用性に欠ける問題がある。そこで本研究では、人間の周囲のカメラ画像から顔領域を抽出し、顔領域の面積を学習させることで顔方向を推定するシステムを構築する。目的は屋内で人間の行動意図の推測のための一情報を得ることなので、おおよその方向を高速に求めることとする。また、方向は水平方向のみを対象とする。

2.1 顔領域の抽出

本研究はモデルを持たずに顔方向を推定する試みなので、顔領域の抽出には輪郭の形や目鼻の部位といった形状特徴は用いず、単純に肌色領域を検出することで顔領域を抽出する。肌色領域の検出には、LUT(Look Up Table) を導入した最近傍識別器 [3] を用いる。この手法により高速に類似色と肌色を分離して検出することが可能となる。ただし、頭部領域は画像中に十分な大きさで写っていると仮定する。この仮定は人物の頭部をカメラ制御により追跡することで満たされるが、現時点では実装が終了していないので、今回は静止人物に対し実験する。

2.2 顔方向の推定

顔領域の面積から顔方向を推定するために、ある顔方向と顔領域面積の対応を学習させる。あらゆる方向に対応するために、人間を囲むように配置されている4台のカメラを用いる。従って学習は、4つの面積の値を入力とし、顔方向を出力するネットワークを形成する問題に帰着する。学習アルゴリズムにはニューラルネットワークを始めとして莫大な研究があるが、本研究では、高速性、記憶領域消費量、設定パラメータの量、出力精度保証など多くの利点を有する PaLM-tree アルゴリズム [4] を採用する。

PaLM-tree は対応付けが不明な写像について、既知の対応の集合を入力することで局所的に線形な写像で近似し、その近似線形写像の集合として非線形写像を構成する。具体的には、本研究での写像は以下ようになる：

$$(S_1, S_2, S_3, S_4) \mapsto \theta \quad (1)$$

ここで S_i はカメラ i で検出された顔領域の面積、 θ は設定された座標系の下での顔方向を示す角度である。学習により写像が形成されると、PaLM-tree の基本的なデータ構造は KD-tree と同様であり、ノード数に対し対数時間で入力に対する出力を検索することができ、高速な応答が期待できる。

3 顔方向推定実験結果

本節では、顔方向推定の精度と速度について、計測結果を述べる。まず、肌色領域を対話的に指定し、最近傍識別器により肌色領域を検出した結果を図1に示す。左図がカメラにより得られた画像で、右図の黄緑色の部分が肌色領域として検出された部分を示している。ソファの表面の色や腕の色は、顔領域の肌色に近い色をしているにも関わらず、識別されている。首は顎の下のある部分を除いて顔領域だと誤認識されているが、顔の方向を推定する問題の場合、顔方向に関わらず一定領域で首が画像中に存在することが多いので、首の誤認識が顔方向推定に与える影響は少ない。

次に、顔領域の抽出に基づき顔方向の学習をする。学習は、被験者の体の向きに対し左方向を0度、正面を90度、右方向を180度とする座標系で、45度、90度、135度の3種類の方向を対象とし行った。各角度については約30秒の顔画像撮影・顔領域抽出結果を入力として与えた。本システムは秒間10フレームの速度で画像を更新しており、肌色領域抽出はほぼ実時間で行えているので、入力は約300種類の面積値を4台のカメラ分だけ与えていることになる。

[†] 独立行政法人 情報通信研究機構^{††} 国立大学法人 和歌山大学

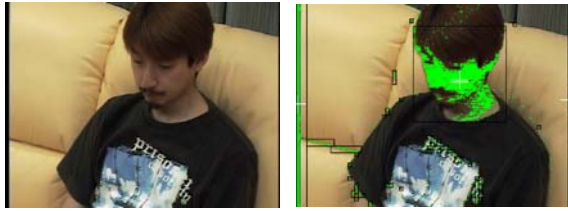


図 1: 顔領域抽出

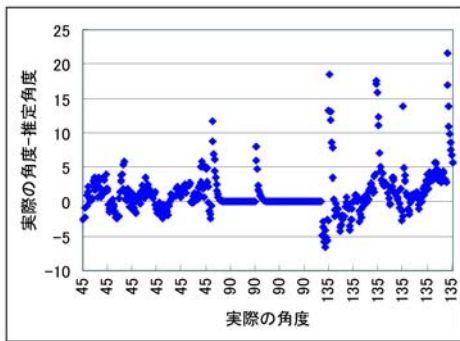


図 2: 顔方向を 45 度, 90 度, 135 度としたときの推定誤差

学習の間は被験者は顔方向を固定させているが、カメラのノイズにより検出される肌色領域が変動するので、約 300 種類の面積値は一定ではない。同様の理由で学習時のみならず顔方向推定時にもカメラのノイズは影響してくるので、長時間の学習を行うほど安定した顔方向推定には有効である。図 3 に、学習後に改めて顔方向を 45 度, 90 度, 135 度とし、4 つのカメラより得られる入力画像から顔方向を推定した場合の、各角度と推定値の差を示す。最大で 20 度程度の誤差があることが分かるが、正面と左方向, 右方向の区別は確実にされている。精密な角度検出が目的ではないので、このような 45 度程度の角度検出ができれば本研究の目標としては達せられている。また、顔方向を固定させずに左右に振った場合の方向推定結果を図 3 に示す。この場合は、被験者は 45 度から 135 度の間を、40 秒間で 5 往復の速度で左右に顔を振っている。図 3 より、顔振りに追従して、45 度から 135 度の間を約 40 秒間で 5 往復している様子が検出できていることが分かる。本研究では実時間性を重視していることを踏まえ、図 3 に顔振りの速度を速くした場合の推定結果を示す。この場合は、10 秒間で 45 度から 135 度の間を 5 往復する日常生活の動作としては速いと思われる動作を想定した。図 3 より、角度が左方向で 45 度, 右方向で 135 度に達していないなど誤推定があるが、左方向, 正面, 右方向という向きに対しては、ほぼ追従していることが分かる。誤推定が生じている原因は、一時遅れ系のフィルタにより推定結果を平滑化していることによる。

以上の実験では、画像センサは日立国際電気社のドームカメラ HC-300 を、キャプチャボードは

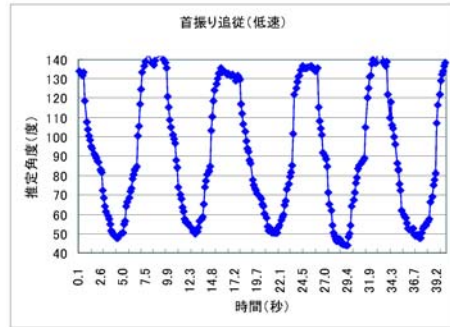


図 3: 40 秒間の 5 往復首振りに対する方向推定

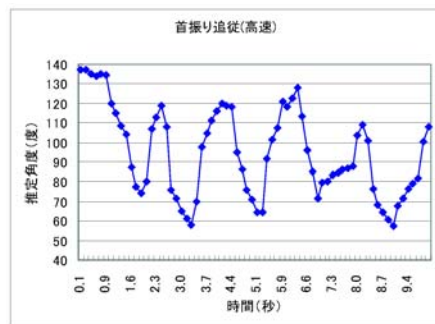


図 4: 10 秒間の 5 往復首振りに対する方向推定

ViewCast 社の Osprey-100 を用いた。画像処理には Celeron 2.50GHz (メインメモリ 512Mbyte) の PC を、画像キャプチャと最近傍識別器による顔領域検出用に 4 台, PaLM-tree による顔方向推定用に 1 台用いた。我々のプロジェクトでは、提案手法の応用例としてロボット (東芝製 ApriAlpha) の顔を人の顔向きと連動させて動かす実験をしている[†]。

4 おわりに

最近傍識別器により肌色情報から顔領域を抽出し、複数の顔領域面積と顔方向を非線形写像学習アルゴリズム PaLM-tree により学習することで、顔方向を推定できることを示した。今後、人物の動き追跡処理を加えることで、人物の移動に対応した実用的なシステムを開発する予定である。

参考文献

- [1] 美濃導彦: ゆかりプロジェクトの目的と概要 - UKARI プロジェクト報告 No.1 -, 第 66 回情処全大 (2004).
- [2] Xu, M. and Akatsuka, T.: Detecting Head Pose from Stereo Image Sequence for Active Face Recognition, *Proc. International Conference on Face and Gesture Recognition*, pp. 82-87 (1998).
- [3] 和田俊和: 最近傍識別器を用いた色ターゲット検出, 情報処理学会論文誌, Vol. 41, No. 6, pp. 1234-1242 (2000).
- [4] 中村恭之, 加藤英介, 和田俊和: 非線形写像学習のための PaLM-tree の提案とその応用, 第 9 回ロボティクスシンポジウム予稿集, pp. 360-366 (2004).

[†]ApriAlpha の動作実装を担当してくれた東芝 山本大介氏に感謝する。