

顔及び手領域による身振り認識システム

1 L-7

鶴田 泰士[†] 植田 訓弘 小川 均

立命館大学大学院理工学研究科

1 はじめに

日常生活において、人間は無意識のうちに様々な身振りをしている。ある場合は、会話の補助として意図や雰囲気伝えることができる。手話においては、対話の中心的手段として情報や意志を伝える役目を果たしている。従って、身振り認識は、ヒューマンインタフェースには不可欠な技術である。

近年の身振り認識の研究では、データグローブや磁気センサ、マーカ等の装着に頼らない非接触な認識手法が研究されている。

大和ら [1] は、隠れマルコフモデル (Hidden Markov Model : HMM) を用いた認識手法を提案している。HMM は最尤推定の枠組で認識が可能である。しかしながら、動作を行っている時間区間のセグメンテーションを手で行わなければならない。

西村ら [2] は白黒動画画像から動き特徴及び形状特徴を抽出し、連続 DP (Dynamic Programming) により認識を行っている。この方法は、画像全体を一括して処理を行うため、標準パターンと比較して、動作の大きさの違い、身振りと関係ない部分の動作等に対応する仕組みを持っていない。即ち、身振り認識が動作位置や、直接関係ない動作に影響される。また、動作の回数が関係する身振りも考慮されていない。

本研究では、上記の問題を解決し、身振りの意味を認識するため、人物の画像を顔・左手・右手の3つの領域に分割し、連続 DP と HMM を組み合わせて認識する手法を提案する。

なお、本研究では「動作」を各領域における動き、「身振り」を全体の動きと定義する。

2 システムの概要

本研究では利用者の使いやすさを考慮して、マーカ等に頼らず、非接触な手法としてカメラから得られた動画画像を用いる。画像から肌色をもとに顔・左手・右手の3つの領域を検出し、各々の領域から基本的な動

作を認識し、その集合により全体の身振りの認識を行う。3つの領域に分けることにより、各領域の動作位置にとらわれない、柔軟な身振り認識ができる。図1にシステムの概要を示す。

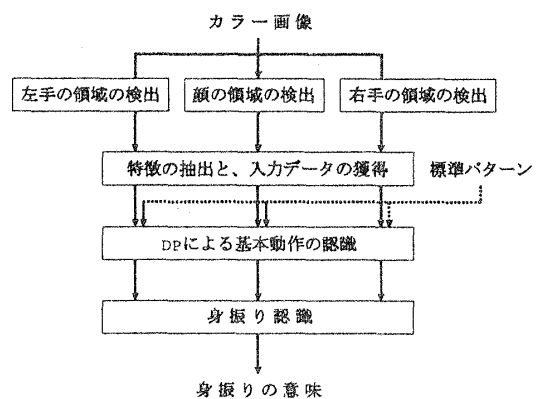


図1: システムの概要

本システムでは、身振りを認識するに当たり、認識対象の前提条件として以下のものを設定している。

- 条件1 画像中には1人の人物がカメラに正対している。
- 条件2 被験者は長袖の服を着用し、服や背景の色は肌色とは異なる。
- 条件3 顔・左手・右手領域は常に画像中に存在し、互いに交差することはない。
- 条件4 追跡開始時は、画像上において左手は顔の右側に、右手は顔の左側にある。

3 顔・手領域の検出と追跡

顔領域は、取り込んだ画像のすべてを、肌色領域、暗色領域、その他の領域に分割し、その中から肌色領域の大きさやその中にある暗色領域の位置関係により検出する [3]。左手領域は顔領域より右側にある最大の肌色領域とし、右手領域は顔領域より左側にある最大の肌色領域とする。

顔領域の追跡は、前フレームで検出された領域の周辺のみで顔検出を行い、検出を繰り返すことによって

実行する。左手・右手領域は、前フレーム検出された各領域の周辺で、最も大きい肌色領域を、その領域として検出する。

4 基本動作の認識

本研究では、「右手を上動かす」のような、身振りを構成する基本的な動作を基本動作と呼ぶ。全体の身振りを認識する前に、まずこの基本動作を認識する。

各領域の動きの特徴を、追跡過程で得られた各領域の中心座標を連続するフレーム間で差分を行うことによって得たベクトル列で表す。この時系列データと事前に作成した基本動作のモデルとなる標準パターンを連続 DP (Dynamic Programming) を用いることにより認識する。連続 DP を用いることにより、時系列データの時間的な伸縮を補正して標準パターンとのマッチングが行える。また、時間的セグメンテーションの問題も解決できる。

現在では、基本動作として上下左右 8 方向への動きと静止の 9 種類の標準パターンを用いている。

5 身振り認識

5.1 各領域の基本動作の関連付け

各領域で得られた基本動作の時間的な関連付けをタイムテーブルに記述し、顔と左右の手の動きの組みを抽出する。このとき、次の方針で行う。

1. 各領域の動作の同期を合わせるため、各領域の基本動作の始点、終点のタイミングがずれていても、全体に対して微小であれば無視し、ずれを吸収する (図 2)。
2. 動作途中の一時停止など全体の動作に比べ微小なものは無視する (図 3)。

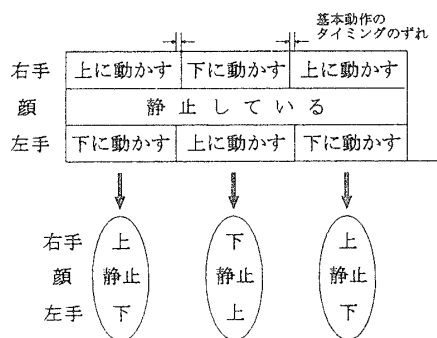


図 2: タイムテーブルの例 1

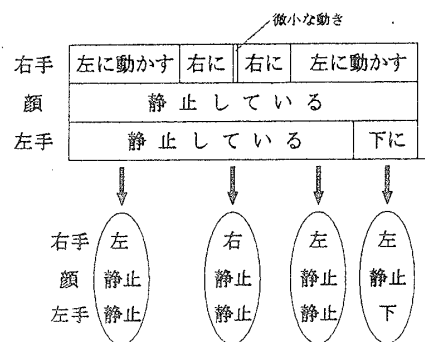


図 3: タイムテーブルの例 2

5.2 一連の基本動作からの身振りの認識

前節のタイムテーブルから得られた基本動作列から、必要な情報を用いて HMM により全体の身振りの認識を行う。例えば、図 3 の例では最後に左手が動いているが、両手の動作から身振りは認識されない。右手のみを使用して「byebye」が認識される (図 4)。

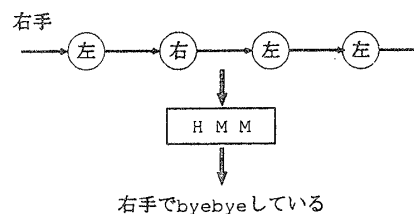


図 4: 身振りの認識の例

6 おわりに

本研究では次のような身振りを認識するために、基本動作とその結果に基づく身振り認識手法を提案した。

1. 動作位置に影響されない身振り認識
2. 身振りに直接関係ない動作がある場合の認識
3. 動作の回数が関係する身振りの認識

参考文献

- [1] 大和淳司, 大谷淳, 石井健一郎: “隠れマルコフモデルを用いた動画像からの人物の行動認識”, 電子情報通信学会論文誌 (D-II), vol.J76-D-II, No.12, pp.2556-2563. (1993).
- [2] 西村拓一, 向井理朗, 岡隆一: “白黒動画像からの形状特徴を用いた動作認識”, 電子情報通信学会論文誌 (D-II), vol.J81-D-II, No.8, pp.1812-1820. (1998).
- [3] 徐剛, 杉本岳夫: “カラー画像を用いた実時間顔検出追跡システム”, 電子情報通信学会技術研究報告, PRMU97-130, pp.9-16(1997).