

# 平成14年度 情報工学コース卒業研究報告要旨

末永	研究室	氏名	根木 大輔
卒業研究題目		人物像を用いた計算機ユーザインタフェースにおける頭部姿勢利用に関する検討	

本研究では、単眼カメラ映像から顔方向を推定し、うなずき動作を検出する手法を提案する。

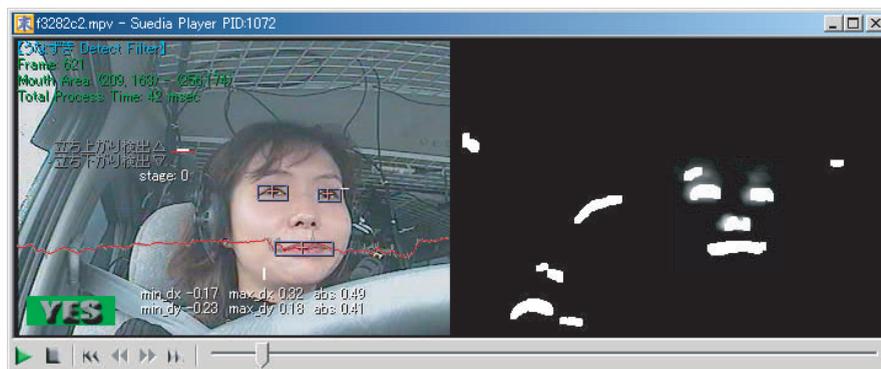
カーナビゲーションシステムを始めとする、音声認識を利用した計算機ユーザインタフェースの普及には目を見張るものがある。しかし、自動車内や街頭ではロードノイズなど雑音が多いため、音声認識による認識率は必ずしも高くない。そこで、画像を用いて音声認識を補うインタフェースがあれば非常に有用である。カーナビゲーションシステムにおけるユーザとの会話タスクでは、目的地の設定などの重要な項目の設定の際に必ず「はい、いいえ」の確認対話が行われる。「はい」という単語は、音声の観点から見ると音圧が小さく、ノイズに埋もれがちで認識しにくい。そこで、画像から「うなずき」を検出することにより、雑音の多い環境下でも安定して計算機に意思を伝えることができると思われる。本研究では、「うなずき」を画像から頭部動作を認識することにより検出する。

単眼カメラ映像から両眼と口領域を抽出するために、入力画像をグレースケール変換し、暗部領域を抽出して横長の領域を強調するフィルタを適用する。2値化してから顔部品として適当な面積をもつ領域のみを残す。あらかじめ用意された顔テンプレートとの相関を取り、両眼、口領域を決定する。

3次元空間における原点からカメラの投影面までの焦点距離が与えられると、原点から両眼と口の3点の重心位置へ向かう3つの3次元ベクトルが求められる。実際の3次元空間における両眼、口は、得られた3次元ベクトルの延長線上に存在する。通常2次元投影座標から3次元座標を復元するのは不可能であるが、3点間の距離が既知であると仮定することにより、連立非線形方程式を立てることができる。ニュートン・ラフソン法を用いて方程式を解くことにより両眼、口の3次元位置が求まり、3点を含む平面の法線方向を顔方向として決定する。

求められた顔方向ベクトルから、顔の上下方向の回転(ピッチ)成分を求め、その時系列変化を見ることでうなずきを検出する。移動平均をとることで雑音を除去し、ピッチが一定フレーム以上連続で下がり続けた場合を立ち上がり、一定フレーム以上連続で上がり続けた場合を立ち上がりとする。立ち下がりを検出してから一定時間以内に立ち上りを検出し、さらに各顔部品の2次元的な並進移動がしきい値以下である場合に、「うなずき」と判定する。

顔部品の並進成分だけではなく顔方向を用いることで、路面状況や体のゆすりなどによる顔の揺れとうなずき動作を分離することができる。リアルタイム動作するシステムを構築し、対話映像や自動車内映像で実験を行い、本手法の有効性を確認した。



「うなずき」動作リアルタイム検出システム動作例

(図右: 顔部品候補領域抽出結果, 図左: 両眼, 口領域抽出結果, ピッチ成分グラフ)