



題目：温度画像処理を用いた頷きコミュニケーションシステム・表情分析に関する研究

CG キャラクタと人間との頷きによる身体的コミュニケーションシステム、および、表情強度を用いて人間の表情分析および情動認識を可能にする手法を提案した。本研究では、温度画像処理を用いることにより、人間に環境的・物理的制約を課す必要がないシステムおよび手法を実現している。

第1章 序論

高齢化と高度情報化が同時進行で進みつつある社会において、「人間の視点から機械と人間の接点のあり方を考える」という観点に立って「共生」という概念を考える上で、“人間とコンピュータ(ロボット)との協調動作”という「パラダイム」が着目されている。ここでいう協調動作とは、「人間の状態を認識してコンピュータ(ロボット)側が応答する機能」をベースとした相互協力的または相互補完的動作を指す。

コンピュータ(ロボット)と人間とが、円滑なコミュニケーションをとるため、さまざまな観点から研究が進められてきた。対面コミュニケーションのシステムでは、相手の言葉を理解するための音声・言語解析などの研究が進められてきた。一方、人間と人間との対面コミュニケーションでは、単に言葉の理解だけではなく、身振り・手振りといったジェスチャーや表情によるノンバーバル情報の果たす役割も大きい。このため、人間とコンピュータ(ロボット)との対面コミュニケーションにおいても、ノンバーバル情報が重要だと考えられている。そこで、人間とコミュニケーションをとるロボットやCG キャラクタたちに、人間との身体的インタラクションを行わせる研究が進められている。しかし、種々のコミュニケーションシステムにおいて、可視光画像を用いる場合には、光環境(太陽光、照明、影、反射)の変化が、ノンバーバル情報のコンピュータでの認識精度に大きな影響を及ぼす。そこで、原理的に光環境の変化に影響を受けない温度画像を用いた表情認識などの研究が行われてきた。

本研究では、人間の状態として、ヒューマンインタフェースの観点で近年注目されている「顔向き」に関連深く、身振り・手振りのうち、視覚情報としてわかりやすい頭部の動きである「頷き」に着目した。そして、温度画像を入力情報として、発声により処理を開始し、人間の「頷き」を予測してCG キャラクタに同調的「頷き」を行わせるシステムを、ファジィ推論を用いて提案した。また、人間は話しかける際に情動表現をする傾向があるので、本研究では、表情強度を用いて、温度画像を入力情報として、発声時における表情分析および情動認識を行う手法を提案した。

第2章 温度画像処理を用いたCG キャラクタと人間との同調的頷きシステム

人間とCG キャラクタとのコミュニケーションシステムを提案した。本システムでは、顔と首領域の温度画像を取得できる赤外線カメラを用いて、人間の頷きを予測し、CG キャラクタによる同調的頷きを実現している。本研究では、人工知能の手法として、ファジィルール、メンバーシップ関数、など一連の処理の調整の柔軟性が高いファジィ推論を用いた。人間の頷き角度の予測には

MA モデルを用いた。本システムにおいて予測した傾き角度の平均誤差は、約 5°であった。本システムでは、頭部の動きを検出するためのセンサーをユーザに装着してもらうという物理的制約もなく、また、システム使用に際して光環境を制御するという環境的制約もない。

また、人間と CG キャラクタによるコミュニケーションをより自然なものにするために、上記システムに、人間の発声を検出する機能を加えた、CG キャラクタと話者との同調的傾きシステムを提案した。そして、CG キャラクタとコミュニケーションを行うユーザの印象を改善するために必要な CG キャラクタの様相、動きの速さなどの具備条件を明らかにした。

第 3 章 温度画像での表情強度を用いた表情分析と情動認識

コンピュータビジョンの研究分野で、種々の表情認識手法が研究されてきたが、現在でも実用化が試みられている段階に留まっている。可視光を用いた表情認識では、光環境の変動による入力画像の濃度変化が表情認識精度の低下を引き起こすため、実用化を難しくしていると考えられる。このため、光環境の変化にロバストな表情認識手法を開発するために、温度画像を用いた表情認識の研究が進められてきた。本研究では、温度画像での表情強度を用い、「テレビ電話中の表情分析手法」および「発声時における情動認識手法」を提案した。

まず、テレビ電話(Skype)で会話中の人間の表情を分析する手法において、記録した映像を、温度画像処理および 2D-DCT 処理を施された口・顎領域から求めることができる表情強度による表情分析に用いた。無感情の際の特徴ベクトルと対象となる特徴ベクトルの差分ベクトルのノルムを表情強度と定義し、表情変化の分析に用いた。発声の有無の区別では、音声データの強度の閾値を用いた。そして、表情の変化を、表情強度の変化として定量的に計測できることを実験で実証した。

また、発声時における表情強度を測定し、規格化した表情強度の平均値と、規格化した発声時間を、情動認識に用いる手法を提案した。そして、被験者に中立的な単語である日本名「taro」(最初と最後の母音はそれぞれ a,o)と「tsubasa」(最初と最後の母音はそれぞれ u,a)を 5 情動(「怒り」、「喜び」、「中立」、「悲しみ」、「驚き」)で、発声してもらい、平均 80%の認識率を得た。

第 4 章 総括

本研究では、温度画像処理を用いることにより、人間に環境的・物理的制約を課す必要のない、CG キャラクタと人間との傾きによる身体的コミュニケーションシステム、および、表情強度を用いた人間の表情分析および情動認識を可能にする手法を提案した。

人間と CG キャラクタにおけるコミュニケーションシステムでは、人間の傾きを予測し、CG キャラクタによる人間との同調的傾きを実現している。また、CG キャラクタの外観および動作速度は、被験者の印象に大きな影響を及ぼすことを明らかにした。また、CG キャラクタの動作に、傾き以外の動作を加えることもユーザの印象の改善に貢献すると考えられる。

次に、温度画像での表情強度を用いて人間の表情分析および情動認識を可能にする手法を提案した。温度画像処理および 2D-DCT 処理を施された口・顎領域から表情強度を求め、表情分析および情動認識に用いた。

第 2, 3 章の結果を基に、温度画像処理を用いた傾きコミュニケーションシステム、表情分析および情動認識に関して、本法の有効性を実証した。