

# 認知症高齢者の情動機能回復評価を目的とした笑顔画像計測

## Smile Measurement for Estimation of Emotional Recovery in Dementia Elders

○花沢 明俊<sup>1</sup>, 中村 克樹<sup>2</sup>○Akitoshi Hanazawa<sup>1</sup>, Katsuki Nakamura<sup>2</sup><sup>1</sup>九州工業大学<sup>1</sup>Kyushu Institute of Technology<sup>2</sup>京都大学<sup>2</sup>Kyoto University

**Abstract:** Learning therapy is a kind of non-drug therapy for dementia elders. It ameliorates, stops or slows the progress of dementia symptoms. Its effectiveness is observed not only in the functional recovery of intelligence but also in that of emotion. The recovery of intelligence is quantified by indices measured by tests similar to intelligence tests. On the other hand, the recovery of emotion has been a matter of subjective impression shared by staffs in nursing homes. We, thus, attempted to quantify the recovery of emotion by measuring smile intensity. We recorded movies of elders undergoing the learning therapy across a year, and developed an image analysis program that can measure the smile intensity of elder people.

### 1. はじめに

高齢化社会の進行にともない、老人の認知症は発症者数が年々増加しており、大きな社会問題となっている。その対策として、認知症の予防・進行抑制・改善を目的とした非薬物療法である認知的なリハビリテーション手法が複数存在し、その1つは学習療法と呼ばれている。学習療法では、読み書き計算などの比較的単純な課題を行うことにより認知機能を活性化している。また、その長期的効果は、知能テスト的な手法により測定されている。学習療法では、認知機能以外に情動機能についても改善が見られ、笑顔などの感情表出が増えることが知られている。認知症における情動機能回復は、周囲とのコミュニケーション機能の回復につながる重要なものである。しかし、情動機能回復については、笑顔が増えた等の介護者の主観による報告が主で、客観的・定量的な測定データを得る方法が確立されていない。また、情動機能回復は、学習療法で行われる課題遂行そのものの効果と、介護者とのコミュニケーションの効果の2つが指摘されている。後者については、学習療法に対する熟練度など介護者個々人の要因に影響を受けるが、その要因の特定も行われていない。そこで我々は、画像処理手法を用いた笑顔の計測により、情動機能回復の客観的・定量的評価を試みた。特別養護老人ホームの協力のもと、認知症発症者の学習療法開始後の経過を定期的にビデオ撮影し、その画像解析を行なった。

### 2. 学習療法

学習療法は、読み書き計算といった単純な課題が、前頭前野を効率的に活性化させるという脳機能イメージングの知見にもとづき、認知症改善のための非薬物療法として開発された[1][2][3]。現在、療法に用いる教材の開発、介護スタッフ研修等の人材育成、老人福祉施設へのシステム導入は、くもん学習療法センターが行なっている。1回のセッションは、通常10～20分程度の時間、学習者が音読や計算の教材を介護者とともに行う。これを週3日から5日

程度で、継続して行う。認知機能の評価は、月に一度程度、一般的認知機能(MMSE) および前頭葉機能(FAB)と呼ばれる認知テストによって行われ、スコア化される。学習療法では、これらのスコアが、3ヶ月程度の期間で対照群と比較して有意に向上することがわかっている。また、会話ができるようになる、自分でトイレへいけるようになる、など、日常的な機能の改善が同時に起こる。

### 3. 画像処理による表情の定量化

笑顔など、顔の表情の検出およびその度合の定量化は、数多くの研究が行われ、デジタルカメラの笑顔シャッター等へ応用されている[4]。その技術には、Haar-like 特徴やガボール特徴などの特徴量を用いる特徴量ベースの手法[5]、口角や目尻などの特徴点を検出し、その座標を用いる特徴点ベースの手法[6]、顔の2次元または3次元モデルを用い、輝度、色、テクスチャーなどの見え(アピアランス)に基づいて表情変化に伴う口や目などの形状変化を特定するアピアランスベースの手法などがある[7]。これらは、特徴量や顔の部位の座標データを SVM などの何らかの学習識別器にかけることで、表情の認識を達成している。これらの手法では、顔を正面からとらえている条件では良好に機能するが、顔の向きが大きく変化したときに性能が低下することが共通の問題点である。しかし、日常的な状況あるいはカメラによる撮影を意識しない状況では、顔の向きは通常大きく変化するため、実用上はこの問題に対し、可能な限り対処する必要がある。本研究では、顔の位置や角度に対する追従処理には、顔の3Dモデルを用いたアピアランスベースの手法を用い、表情によって大きく変化する目および口の周辺領域の形状変化に対しては、3Dモデルを用いた正面画像化処理を行った後、ガボール特徴による特徴量ベースの手法を用いて、笑顔度の評価を行った。

### 4. 方法

#### 4.1. データ画像の撮影

データ画像の撮影は、特別養護老人ホームいぶき苑（岐阜県）にて行った（図1）。撮影には、ビデオカメラを用いた。データ画像の時空間解像度は、1280×720ピクセル、60フレーム/秒（プログレッシブ）である。18名の学習療法未経験の被験者に対し、学習療法開始直後から1年間、3ヶ月間隔で4回のビデオ撮影を行った。撮影は、学習療法のセッション時に行い、10分から20分程度を記録した。撮影時に、画像解析用のマーカーとなる、色の異なるカラープレートが5枚ついたバンドを被験者の頭部に装着した。



図1：学習療法セッション中のビデオ撮影

#### 4.2. 顔画像の追従・正面画像化処理

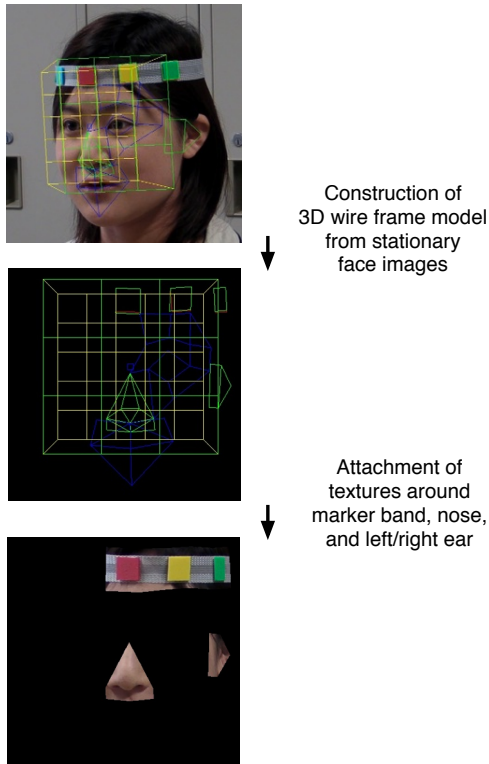


図2：顔の3Dモデル作成

笑顔認識の事前準備として、各被験者について、ビデオ動画から抜き出した3枚の顔の角度が異なる静止画を用い、マーカーバンド部、鼻周辺部、耳周

辺部、目周辺部、口周辺部の5部位からなる頭部の3Dモデルを構成した（図2）。3Dモデルは、ワイヤーフレームで作成し、マーカーバンド部、鼻周辺部、耳周辺部の3部位について基準画像のテクスチャーを貼り付け、テクスチャーを持つ3次元モデルとした。

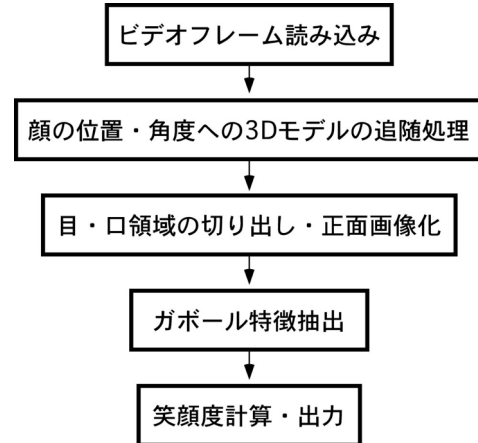


図3：笑顔度計測処理の流れ

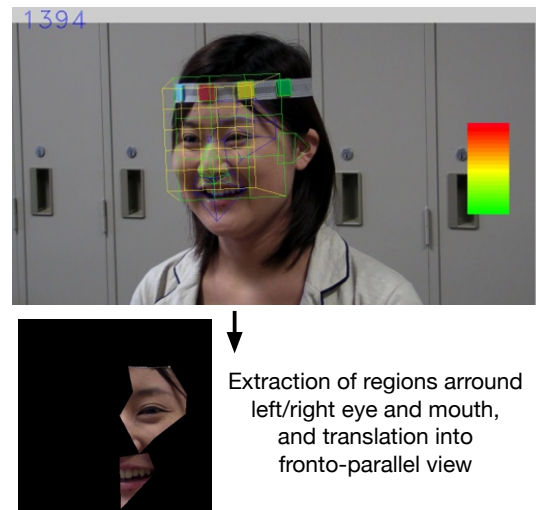


図4：各フレームにおける画像処理

この3Dモデルを用いた被験者の笑顔評価を図3のような流れで行った。まず第1段階として、ビデオ動画の各フレームを静止画として読み込んだ。次の第2段階では、顔の動きへの追従処理を行うため、各データフレームで3Dモデルの顔画像へのフィッティングを行い、顔の位置・角度を求めた。各フレームでのモデルのフィッティングは、最初のビデオフレームに対して初期値を手動で与え、その後は、前画像の位置からパーティクルフィルタおよび再急勾配法によって、フィッティングを行った（図4上部）。第3段階では、顔画像にフィッティングしたワイヤーフレームモデルの目領域、口領域に対して、該当フレームのテクスチャーを貼り付け、更にワイヤーフレームモデルを回転させることにより、目・口領域について正面から見た状態を模擬した画像を作成した（図4下部）。この目・口領域の正面画像において、ガボール特徴抽出および笑顔度の評価を行った。

### 4.3. 特徴抽出と笑顔度計測

目領域、口領域の正面化後の画像に対し、ガボールフィルタによる特徴抽出を行った(図5)。用いたガボールフィルタのサイズは、16x16ピクセルで、1周期が8ピクセルの正弦波・余弦波に半値幅8ピクセルのガウス関数を乗じて作成した。4方位のフィルターを用い、画像の各ピクセル位置でフィルターとの畳み込み演算を行った。正弦波フィルターと余弦波フィルターの出力の2乗和平方根を最終出力とし、20x20ピクセルを1ブロックとして出力値を各方位で足し合わせ、ヒストグラム化した。400x400ピクセルの画像に対し、ブロック数は20x20、さらに4方位のヒストグラム値で、20x20x4=1600特徴で特徴ベクトルを構成し、笑顔の認識処理を行った。

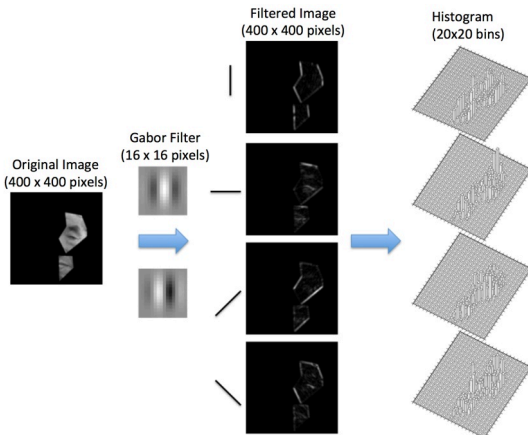


図5：ガボール特徴ヒストグラムによる顔画像のベクトル化

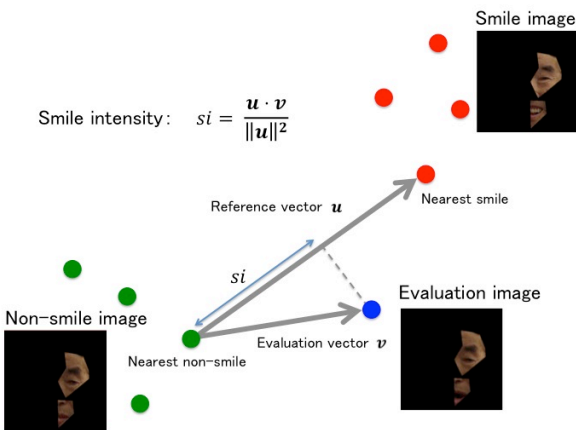


図6：笑顔・非笑顔プロトタイプに基づいたデータ画像の笑顔度評価

笑顔度の計算には、ビデオデータ中から選んだ典型的な非笑顔画像5~10枚と笑顔画像5~10枚をプロトタイプとし、各画像について特徴ベクトルを算出した。処理対象画像の特徴ベクトルに対し、特徴ベクトル間のユークリッド距離から求めた最近傍の非笑顔画像1枚と笑顔画像1枚を選定した。非笑顔画像を原点としたときの笑顔画像ベクトルと処理画像ベクトルの内積を笑顔画像ベクトルのユーク

リッドノルムで正規化することにより、笑顔度を算出した(図6)。

算出された笑顔度を数値出力し(図7)、同時に入力動画に笑顔度のスケール表示を加えた動画を出力した。

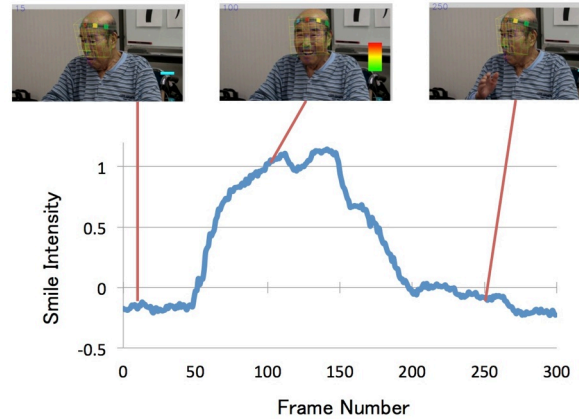


図7：笑顔度の数値出力と画像表示

## 5. 結果

ある1名の被験者の1セッションの笑顔度計測結果を図8に示す。計測の妥当性を評価するため、人間による4段階の笑顔度評定の結果も重ねてプロットした。人間による評価が0の非笑顔の場合は、画像計測による笑顔度の値が0.5以下に分布している。人間による評価値1が2カ所あり、笑顔度の数値は0.5近傍となっている。人間による評価が3以上の場合は、笑顔度の数値は1近傍である。これらの結果から、算出された笑顔度の数値は、実用上妥当性が高いと考えられる。

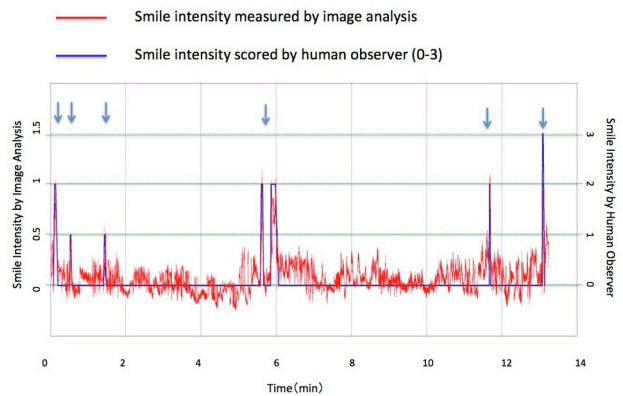


図8：1セッションの笑顔度計測結果

## 6. まとめ

学習療法による情動機能回復の定量化のため、特別養護老人ホームにおいて、学習療法時のビデオ撮影を行い、笑顔表出について画像解析を行った。顔の3Dモデルによる鼻・耳領域およびマーカーバンドを用いたアピランスペースの、位置・角度に対する追従処理を構築し、良好に動作した。各フレームから切り出し、正面化処理を行った目・口領域を用いた特徴量ベースの笑顔度評価を行い、人間によ

る判定との良好な一致を確認した。現在、ビデオデータの解析，特に顔の追従処理に長時間を要しているため，撮影ビデオ全体について解析の途上であり，今後，笑顔度評価の妥当性の検証を進めるとともに，学習療法の効果として，笑顔の頻度や強度の長期的変化の解析を行う。

## 7. 謝辞

本研究を遂行するにあたり，多大なご協力をいただいた特別養護老人ホームいぶき苑，くもん学習療法センターの方々に深く感謝いたします。

## 参考文献

- [1] 川島隆太，脳機能のブレインイメージング，医学書院，2002.
- [2] 川島隆太，山崎律美，痴呆に挑む-学習療法の基礎知識，くもん出版，2004.
- [3] くもん学習療法センター，山崎律美，川島隆太，学習療法の秘密—認知症に挑む，くもん出版，2007.
- [4] 勞世広，山口修，“顔画像処理の応用事例”，情報処理，vol.50，pp.436-443，2009.
- [5] 小西嘉典，木下航一，勞世広，川出雅人，“リアルタイム笑顔度推定”，情報処理学会インタラクシオン 2008 予稿集，2008.
- [6] A. Gee and R. Cipolla, Fast visual tracking by temporal consensus, Image and Vision Computing, vol.14, pp.105-114, 1996.
- [7] T. F. Cootes, G. J. Edwards, and C. J. Taylor, Active appearance models, Proc. European Conf. on Computer Vision (ECCV 1998), pp.484-498, 1998.

## 連絡先

花沢 明俊  
九州工業大学工学研究院基礎科学研究系  
Email: hanazawa@mns.kyutech.ac.jp