

**科学研究費助成事業 研究成果報告書**

平成 29 年 6 月 21 日現在

機関番号：14303

研究種目：若手研究(B)

研究期間：2015～2016

課題番号：15K15993

研究課題名（和文）内発的表情の認識に基づく感情推定によるライフログ映像検索に関する研究

研究課題名（英文）Lifelog video retrieval using emotion estimation based on recognition of spontaneous facial expressions

研究代表者

野宮 浩揮（Nomiya, Hiroki）

京都工芸繊維大学・情報工学・人間科学系・准教授

研究者番号：80533116

交付決定額（研究期間全体）：（直接経費） 1,600,000円

研究成果の概要（和文）：感情推定に基づいて、ライフログ映像から印象的なシーンを効率的に検索・提示するシステムを構築した。印象的なシーンでは、映像中の人物に様々な表情が表出している可能性が高いため、顔面上の特徴点を用いて、表情の種類ならびに表情の強さを推定する手法、および映像中から特定の種類・強さの表情が表出しているシーンを検出する手法を確立した。これらを用いて、ウェブブラウザ上で動作し、ライフログ映像中から様々な表情が表出しているシーンを容易に検索・閲覧できる映像検索システムを構築した。

研究成果の概要（英文）：A video retrieval system to efficiently retrieve and display impressive scenes from lifelog videos is developed on the basis of emotion estimation. In most of impressive scenes, a person in a video often shows a variety of facial expressions. Therefore, this system retrieves impressive scenes by means of facial expression recognition and the estimation of the intensity of facial expression using facial feature points. In addition, a method to detect emotional scenes in which a certain facial expression is observed is introduced. By using them, a lifelog video retrieval system is developed. This system works on a web browser and has a user-friendly interface so that a user easily finds impressive scenes.

研究分野：マルチメディアデータ工学

キーワード：ライフログ 映像検索 感情推定 表情認識 表情強度

### 1. 研究開始当初の背景

スマートフォンや動画撮影機器、データ記録媒体などの発展・普及により、手軽に動画を撮影できるようになってきたため、日常の行動や体験を映像として記録することが広く行われるようになってきている。これらは、ライフログ映像として、個人の記憶の補助や、過去の出来事の振り返り、行動分析などに幅広く活用できる。

しかし、ライフログ映像は主に個人で作成・所有するため、映像データの管理が難しいという問題がある。公開できる映像は、SNS や動画投稿サイトなどを通じて管理・検索することができるが、プライベートな映像は個人で管理しなければならず、大量の映像データから、見返したいシーンを検索することは困難である。タグ付けなどによる従来の検索手法は、映像に手作業でタグをつける手間が煩わしいといった問題から、利用することが難しいと考えられる。

このような背景により、ライフログ映像から、タグ付けなどの作業を介さず、有用なシーンを的確かつ効率的に取得する手法の確立が望まれていた。

### 2. 研究の目的

本研究では、人物を主体としたライフログ映像を対象に、重要なシーンを効率的に検索・提示するシステムの構築を目的とする。ここで、重要なシーンとは、印象的な出来事が起こったシーンであり、従来では写真やホームビデオなどの形で残しておくような場面を指す。

重要なシーンでは、人物の感情が変化し、それが表情の変化に現れると考えられることから、映像中の人物の表情から重要なシーンを推定できると考えられる。しかし、人の表出する表情は様々であり、感情の強さによっても表情が変化するため、表情の微妙な違いを捉える必要がある。

そこで、表情表出の有無に加えて、表出している表情の強度を適切に推定し、どのような表情がどの程度の強さで表出しているかを分かりやすく提示することにより、容易に重要なシーンを検索・閲覧できるシステムを構築する。

### 3. 研究の方法

(1)まず、シーンの重要性と感情・表情の関係を、十分なライフログ映像データを作成、分析することにより実現する。

(2)表情の識別・強度推定に有効な特徴量の明確化を行う。(1)で重要と判断された表情を正しく識別し、その強度を正確に推定するために、様々な特徴量を組み合わせ、最適な特徴量を明らかにする。

(3)表情識別・表情推定モデルの構築法を確立する。これには、アンサンブル学習等の機

械学習法を用いる。

(4)(3)で構築したモデルをもとに、推定された表情強度に基づいて、表情表出シーンの検出手法を確立する。

(5)ライフログ映像検索システムの構築と評価を行う。ライフログ映像作成に携わった実験協力者によりシステムの評価を行う。

### 4. 研究成果

(1)これまでにやってきた研究では、ライフログ映像検索手法の評価を行うために使用する映像データセットが不十分であったため、新たにライフログ映像データセットを作成した。

ここでは、ライフログ映像検索において、検索対象となりやすいと考えられる、笑顔と驚きの表情に重点を置き、様々な強度の笑顔と驚きの表情(自然に表出した表情)が含まれるように、カードゲーム等を行っているシーンを撮影し、ライフログ映像データセットとした。

また、撮影の参加者全員に、撮影したシーンの中から、後で見返したいシーンと、その時に抱いていた感情を記述してもらった。

(2)(1)で重要と判断した表情(笑顔と驚き)に、比較的ネガティブな感情を表す、怒り、嫌悪、恐怖、悲しみを含め、それぞれの識別に有効な特徴量を明らかにした。なお、これらの表情は基本6表情として知られているものである。

特徴量として、比較的高速に求めることのできる、顔特徴点(図1に示すような、眉・目・口の端点など、表情表出に関係すると考えられる点)の位置関係を用いた。

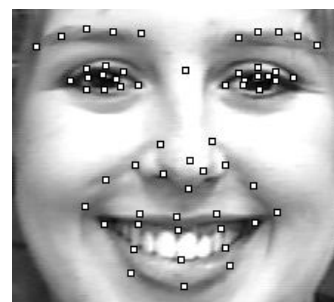


図1. 顔特徴点

表情の識別、ならびに表情強度の推定に有効な特徴量を明らかにするため、2通りの手法を構成した。

一方の手法では、複数の顔特徴点の位置関係から得られる値を特徴量とし、すべての可能な特徴量の組み合わせの中から、機械学習法を用いて、学習用データに対して高い表情識別性能を示すものを選択することにより、有用な特徴量を明らかにしている(この手法は、雑誌論文、ならびに図書に掲載されている。また、学会発表にて発表してい

る) 図 2 に、実際に有用と判断された特徴量の例を示す。なお、以後この手法を「手法 A」と記述する。

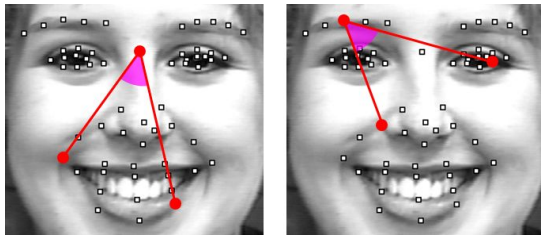


図 2. 有用と判断された特徴量の例 (ここでは、3 点の顔特徴点により構成される 2 本の線分のなす角を特徴量としている)

他方の手法では、無表情時と表情表出時の顔特徴点の位置を目視で比較し、表情ごとに、表情の識別や表情強度の推定に有効と考えられる特徴量を人手で定めている。例えば、左右のほうれい線の角度の平均値や、鼻と口の間の平均距離などを特徴量として定めている。(この手法は、雑誌論文 に掲載されている。また、学会発表 にて発表している) なお、以後この手法を「手法 B」と記述する。

(3)手法 A にて得られた有効な特徴量を用いて、複数の識別器を統合することにより、通常の識別手法に比べて高い認識性能を実現することのできる、アンサンブル学習法に基づいて、表情識別モデルならびに表情強度推定モデルを構築した。

ライフログ映像データを用いて表情認識性能の評価を行ったところ、平均して 60%程度の表情認識精度が得られた。これは、Support Vector Machine など、広く用いられている学習法と比べて高い認識精度であった。同じ映像データを用いて、表情強度推定の評価を行ったところ、人手により作成した正解データと、本手法での推定値との間に高い相関が見られたことから、適切に表情強度の推定ができていたことを確認した(この研究成果は、雑誌論文 に掲載されている)。

また、手法 B にて定めた特徴量の値を組み合わせることで、表情強度の値を定めた。表情強度の推定に有効となる特徴量は、表情によって異なることから、基本 6 表情のそれぞれについて、表情強度の推定に有効な特徴量を定め、表情ごとに個別の表情強度算出式を構成した。

ライフログ映像データならびに、表情認識性能の評価に広く用いられている表情データセットを用いて提案手法の表情強度推定性能を評価したところ、人手で判断した表情の強さと表情強度の推定値がおおむね比例しており、表情強度の推定が適切にできていることを確認した。(この研究成果は、雑誌論文 に掲載されている。また、学会発表 にて発表している)。

(4)(3)で得られた表情強度をもとに、特定の

表情が特定の強度で表出しているシーンを検出する手法を確立した。表情強度は実数値で求められるが、人が表情の強さを判断する際には、それほど細かく表情の強さを見分けているわけではないと考えられるため、表情強度に閾値を設け、表情の強さを 5 段階に分類し、それぞれの強さ(表情強度レベルと呼ぶ)の表情が表出しているシーンを検出することとした。

表情表出シーンの検出には、シーンの分割・統合を繰り返す手法を採用した。まず、すべての表情表出フレームをそれぞれ独立したシーンとし、図 3(a)に示すように、互いに隣接したシーン  $S_{i-1}$  と  $S_i$  を統合して、単一のシーン  $S_i$  とする。ここで、シーン  $S$  の添字はシーンの番号を示す。また、二つのシーンが隣接しているとは、シーン間の距離(シーン間にあるフレームの数)  $D(S_{i-1}, S_i)$  が閾値  $\Theta$  以下であることを意味する。さらに、図 3(b)に示すように、孤立した短いシーン  $S_i$  を削除する。ここで、孤立した短いシーンとは、シーンの長さが  $\Theta$  以下であり、かつ両隣のシーンとの距離がいずれも  $\Theta$  以上であるようなシーンを指す。この操作を、統合・削除できるシーンがなくなるまで繰り返し、最終的に得られたシーンを表情表出シーンとしている(この手法は、図書 に掲載されている)。

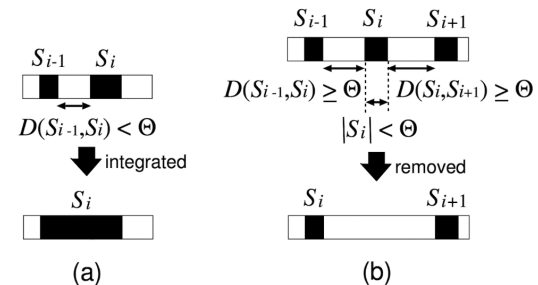


図 3. 表情表出シーンの検出法

(5)(4)の表情表出シーン検出法を用いて、ライフログ映像からの印象的なシーンの検索システムを構築した。検索の利便性を考慮し、一般的なウェブブラウザ上で動作するシステムとした。本システムは、図 4 に示すように、検索モジュールとデータベース管理モジュール(RDBMS モジュール)から構成される。

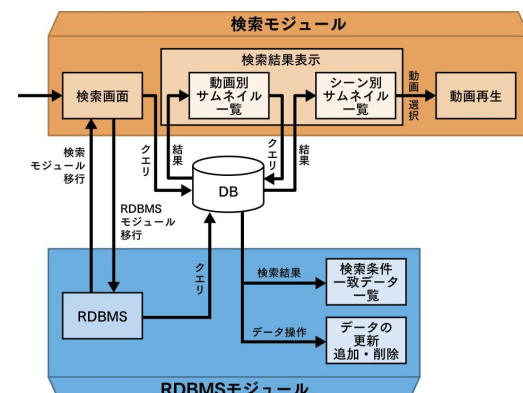


図 4. ライフログ映像検索システムの構成

検索モジュールは、ユーザから検索要求を受け付け、映像データベースを介して検索結果をウェブブラウザ上に提示するモジュールであり、RDBMSモジュールは、データの追加、変更などの機能を提供する、主にライフログ映像の管理者向けのモジュールである。

シーン検索を行う際には、まず検索条件指定画面にて、表情の種類、表情強度レベル、シーンの長さ、映像の撮影日などの条件を指定し、検索を実行する。検索が完了すると、検索の結果得られたシーンがサムネイルとして一覧表示されるので、それらの中から閲覧したいシーンを選択する。選択されたシーンは、図5に示すように、検索結果提示画面に表示される。



図5. 検索結果提示画面

検索結果提示画面では、シーン中で基本6表情が表出しているシーンがヒートマップとして表示される(図5下部参照)。ヒートマップには、表情ごとに異なる色が割り当てられており、表情強度レベルが高いほど濃い色が表示される。なお、ヒートマップは、映像中の人物一人一人に対して提示される。これにより、ユーザは特定の表情が特定の強さで表出しているシーンを容易に閲覧することができる。

また、検索結果提示画面には、シークバーが設けられており、ユーザはこれ进行操作することにより、任意のシーンを閲覧することができる。

ライフログ映像作成に携わった実験協力者により、本検索システムの使いやすさの評価を行った。評価はアンケート形式で、反応の良さ、見やすさ、構成のわかりやすさ、好感度のカテゴリに関する33項目について、1~5点(5点が最も良いことを表す)で評価する形式で行った。各カテゴリの平均評点を図6に示す。

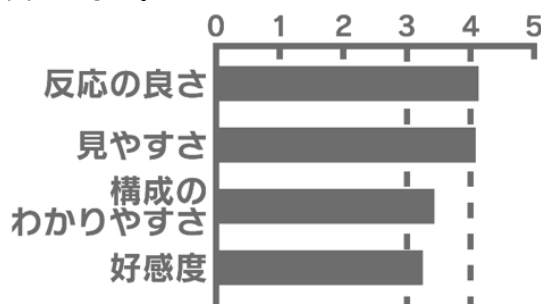


図6. 各カテゴリの平均評点

図6より、すべてのカテゴリについて、3点以上の評価が得られた。特に、反応の良さや見やすさは、4点以上の高評価であった。

一方、シークバーが使いづらい、ヒートマップが画面下部に表示されるため見づらいなどの問題点も明らかになった。これらは、今後改善を検討している(以上の研究成果は、雑誌論文に掲載されている)。

本評価実験では、主に検索システムの使いやすさを評価しているが、これとは別に、(1)で事前に得ていた、撮影した映像中にある、後で見返したいシーンと、その時に抱いていた感情が適切に検出できているかどうかの検討を行った。

その結果、見返したいシーンでは笑顔・驚きの表情レベルが比較的高くなることが多く、これは実際に抱いていた感情とおおむね合致していたため、ある程度見返したいシーンを検索することは可能であると考えられる。

しかし、表情レベルが低いと判断されているが見返したいシーンや、見返したいシーンではないが表情レベルが高いと判断されたシーンが散見されたため、表情認識や表情強度レベルの推定にはまだ改善の余地があると考えられる。これらについては、今後特微量や学習手法の改良などにより、改善を行っていく予定である。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計4件)

Hiroki Nomiya and Teruhisa Hochin, "Emotional Scene Retrieval from Lifelog Videos Using Evolutionary Feature Creation," *Studies in Computational Intelligence*, Vol.612, pp.61-75, 2015, 査読有.

DOI: 10.1007/978-3-319-23509-7\_5

Hiroki Nomiya, Shota Sakaue, and Teruhisa Hochin, "Recognition and Intensity Estimation of Facial Expression Using Ensemble Classifiers," *Proc. of 15<sup>th</sup> International Conference on Computer and Information Science*, pp.825-830, 2016, 査読有.

DOI: 10.1109/ICIS.2016.7550861

Shota Sakaue, Hiroki Nomiya, and Teruhisa Hochin, "Estimation of Emotional Scene from Lifelog Videos in Consideration of Intensity of Various Facial Expressions," *Proc. of 18<sup>th</sup> IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, 2017, 査読有(掲載確定)。

Mitsuaki Maeda, Hiroki Nomiya, Shota Sakae, Teruhisa Hochin, and Yukiko Nishizaki, "Emotional Video Scene Retrieval System for Lifelog Videos Based on Facial Expression Intensity," Proc. of 18<sup>th</sup> IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 2017, 査読有 (掲載確定) .

〔学会発表〕(計2件)

野宮 浩揮、阪上 翔太、宝珍 輝尚、決定木アンサンブルを用いた表情認識と表情強度の推定、第8回データ工学と情報マネジメントに関するフォーラム(DEIM2016)、2016年3月1日、ヒルトン福岡シーホーク(福岡県福岡市) .

阪上 翔太、野宮 浩揮、宝珍 輝尚、顔特徴点の位置変化による特徴量を用いた表情強度の推定、情報処理学会関西支部 支部大会、2016年9月26日、大阪大学中之島センター(大阪府大阪市) .

〔図書〕(計1件)

Hiroki Nomiya, Teruhisa Hochin, Andrew Cummings, Jennifer Rennels, Denise Davidson, Sandra B. Vanegas, Elizabeth Hilvert, Elizabeth Tuminello, Theano Kokkinaki 他, "Emotional and Facial Expressions: Recognition, Developmental Differences and Social Importance," Nova Science Publishers, pp.131-151 (189 pages), 2015.

## 6. 研究組織

### (1)研究代表者

野宮 浩揮 (NOMIYA, Hiroki)  
京都工芸繊維大学・情報工学・人間科学系・准教授  
研究者番号：80533116

### (2)研究協力者

阪上 翔太 (SAKAUE, Shota)  
前田 光晶 (MAEDA, Mitsuaki)