

令和 2 年 6 月 16 日現在

機関番号：14303

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K00156

研究課題名(和文) 表情強度を用いた多様な感情の推定によるライフログ映像検索システムの構築

研究課題名(英文) Development of Lifelog Video Retrieval System Using Estimation of Various Emotions Based on Facial Expression Intensity

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) 3,000,000円

研究成果の概要(和文)：ライフログ映像から、映像中の人物に感情が現れているシーンを重要なシーンとして検索する手法、ならびに、容易にシーンの検索と閲覧ができるシステムの構築を行った。人物の感情は、表情の種類と強度を認識することにより推定した。顔画像から得られる顔特徴点の位置関係をもとに生成した特徴量を用いて、機械学習により表情の識別・強度推定モデルを作成し、映像中の重要なシーンの検出に用いた。また、どのような特徴量が有用であるかの分析を行った。その結果、喜びや驚き等、様々な表情について、認識に有用な特徴量を明らかにした。

研究成果の学術的意義や社会的意義

スマートフォンの普及等により、ライフログ映像に相当する映像データが多く作成されている。従来、映像検索を行う際には、注釈をつける等の手間が必要であったが、本研究では、映像中の人物の表情のみを手がかりにシーン検索ができるため、映像検索を容易に行うことができるようになる点で社会的意義があるといえる。また、表情認識に関する多くの研究では、どのような特徴量が認識に有効かという点は考慮されていない。本研究では、いくつかの特徴量に対してこの点を明らかにしたことにより、学術的意義があるといえる。

研究成果の概要(英文)：Focusing on lifelog videos, we developed a video retrieval system to retrieve and display important scenes in lifelog videos. An important scene is defined as a scene in which a person in the video shows a certain emotion. The emotion of a person is estimated by recognizing the type and intensity of a facial expression of that person. Our video retrieval system finds important scenes using a facial expression recognition model developed by means of a machine learning technique and several facial features computed from the positional relationships of facial feature points. In addition, we analyzed the effectiveness of facial features. As a result of the analysis, we found a number of facial features which were effective to recognize various facial expressions such as happiness and surprise.

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 映像検索 表情認識 表情強度推定

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

近年のスマートフォン等の普及により、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、記憶の補助や過去の振り返り等に活用できる。しかし、ライフログ映像は個人で作成・所有するため、検索を行う際には手作業で注釈等をつける必要があり、映像データの量が増えるにつれて、必要な映像を検索することが困難になる。そこで、ライフログ映像から、手作業を介さずに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。

本研究を開始する以前に、ライフログ映像検索に関する研究を既に開始していたが、研究を始めてからの期間が短かったため、検索の精度が不十分であり、検索システムも使いやすいものではなく、多くの改善の余地があった。

2. 研究の目的

ライフログ映像から、映像中の人物に何らかの感情が現れているシーンを重要なシーンとして検索・提示するシステムの構築を目的とする。感情は表情に現れやすいことから、表情認識を用いて感情を推定する。既存の表情認識法には、表出している表情の種類を見分けるものが多いが、人の感情は多様であり、感情の程度も様々であることから、表情の種類を識別するだけでなく、表情の強度も推定することにより、多様な感情が現れているシーンを的確かつ高速に検索・提示するシステムを確立する。

3. 研究の方法

- (1) 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、シーンの重要性和感情・表情の関係を明確化する。
- (2) 重要と判断された感情を表す表情を正しく識別し、その強度を正確に推定するために、表情の識別・強度推定に有効な特徴量の明確化を行う。
- (3) 得られた特徴量をもとに、機械学習法を用いて、表情識別・表情推定モデルを構築する。
- (4) 構築されたモデルを用いて、映像中から特定の表情が表出しているシーンを検出する手法を確立する。
- (5) ライフログ映像の検索・提示システムを構築し、システムの有用性の評価を行う。

4. 研究成果

- (1) 作成したライフログ映像データから、ポジティブな感情が現れているシーンの重要性が特に高いと考えられたことから、基本6表情(怒り、嫌悪、恐怖、幸福、悲しみ、驚き)の中から、特に幸福(笑顔)の表情に重点を置いて、重要なシーンの検索を行うこととした。しかし、単一の表情に限定してしまうと、ライフログ映像検索システムとしての汎用性が低くなると考えられることから、幸福の表情に重点を置きつつ、基本6表情の全ての表情をシーン検索の対象とすることとした。
- (2) 表情の識別や強度推定の精度と速度のトレードオフを考慮して、図1に示す顔特徴点(眉、目、鼻、口の周辺の点49点)の位置関係に基づく特徴量を用いることとした。特徴量として、幸福の表情の表出の有無や、表情の強度を精度よく見分けられるよう、幸福の表情が表出した際に特徴的な変化が見られると考えられる、口角やほうれい線に関する特徴量を定めた。また、幸福以外の表情の識別や強度推定もできるだけ精度よく行えるように、様々な表情の表出に関連していると考えられる、汎用的な特徴量も定めた。特徴量の一覧を表1に示す。

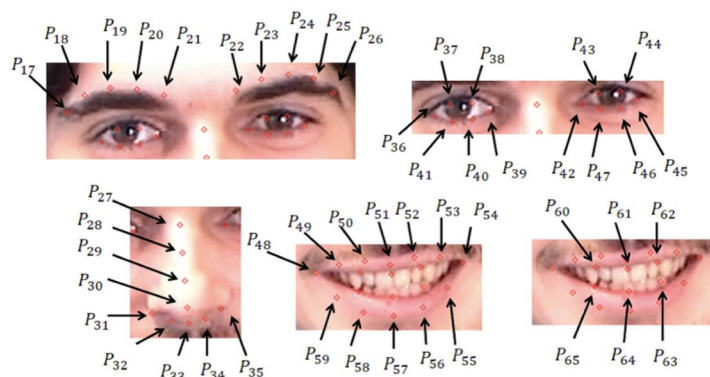


図1 顔特徴点

表 1 顔特徴点の位置関係から得られる特徴量 ($f_1 \sim f_{16}$ は特徴量の略称を示す)

f_1 : 眉の角度	f_9 : 口の外側の縦横比
f_2 : 眉と目の距離	f_{10} : 口角の上がっている度合い
f_3 : 眉間の面積	f_{11} : 口角の角度
f_4 : 目の面積	f_{12} : ほうれい線の角度
f_5 : 目の縦横の比	f_{13} : 口の両端と鼻の頂点を結んだ三角形の面積
f_6 : 口の内側の面積	f_{14} : 口と鼻の間の距離
f_7 : 口の外側の面積	f_{15} : ほうれい線と口の両端を結んだ四角形の面積
f_8 : 口の内側の縦横比	f_{16} : 眉と目の面積

- (3) 表 1 に示した特徴量を用いて、統計的手法により、表情の表出強度を求める手法を考案した。具体的には、無表情の学習用顔画像を用いて、予め無表情時の各特徴量の基準値を求めておき、その基準値からの差分をもとにして、各表情の表情強度を算出する。差分が大きいほど、強く表情が表出していることを表す。
また、各表情が表出している学習用の顔画像から、各特徴量に対して、特徴量の値と表情強度の相関係数を求め、それぞれの表情について、どの特徴量が表情強度と相関があるかを求めた。例えば、幸福の表情では、 $f_4 \sim f_{10}$ と f_{12} 、 f_{14} の 9 つの特徴量が、幸福の表情強度と相関があることが明らかになった。不要な特徴量を用いて表情強度を算出すると、表情強度の推定精度が低下する可能性があるため、相関がある特徴量のみを用いて表情強度を算出することとした。
- (4) 主に笑顔が表出しているライフログ映像に対して、映像中の各フレームの画像から、(3) に示した手法により求めた表情強度と、実際の表情強度を比較する実験を行った。4 名のライフログ映像について実験を行ったところ、4 名中 3 名については、提案手法により求めた表情強度と、実際の表情強度がおおむね対応しているという結果が得られた。残り 1 名については、笑顔が強く表出しているにもかかわらず、提案手法により推定された表情強度が低くなるという結果となった。この 1 名については、顔特徴点が正しく抽出できていないフレームが散見されていたことが、精度の低下につながった可能性が考えられる。また、表情の個人差が精度に影響している可能性も考えられる。人によって表出する表情にそれぞれ特徴があり、表情の強度も異なることから、個人差の影響を低減し、多くの人に対して精度よく表情認識や表情表出シーンの検出を行うことが今後の課題である。
- (5) 基本 6 表情が表出しているシーンをライフログ映像中から検索・提示するシステムを構築した。ライフログ映像の検索結果の画面の一例を図 2 に示す。左の図は、特定の表情（この図では幸福の表情）が表出しているシーンの一覧を表示している画面であり、右側の図は、これらの中から 1 つのシーンを選択した際に表示される、より詳細な検索結果を示している画面である。図の下側には、映像のどの部分でどれくらいの強さの表情が表出しているかを目視で確認できるよう、ヒートマップの形式で表情強度を表示している。また、表情強度は、映像中に含まれる人物それぞれについて確認できるようになっている。なお、本研究を開始する以前にも、ライフログ映像検索システムを構築していたが、機能や使いやすさの点に関して改善の余地があったため、本研究において構築したライフログ映像検索システムは、これらの点を改善したものとなっている。

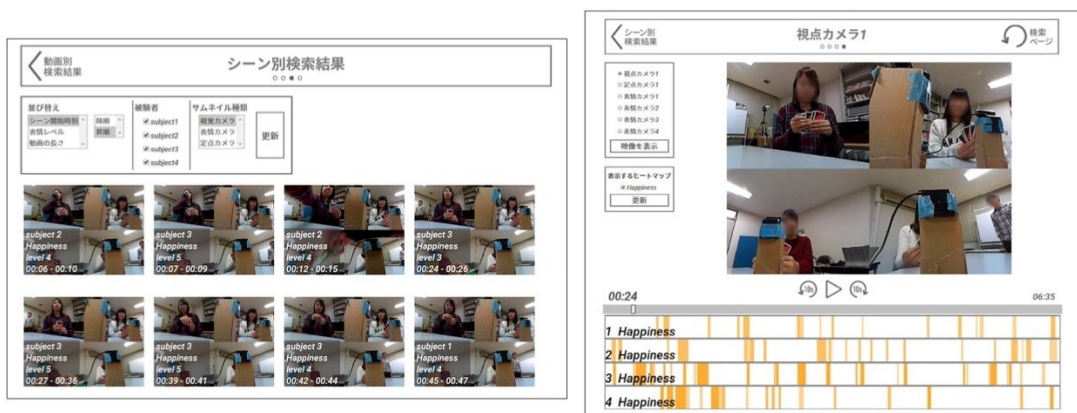


図 2 ライフログ映像検索結果画面

- (6) (5)までで、ライフログ映像検索システムを構築することができたが、表情の認識精度が十分ではないという点に問題が残った。そこで、表情認識に有効な特徴量を発見し、表情認識精度を向上することを目的として、表1に挙げた16種類の特徴量について、どの特徴量がどの表情の認識に寄与しているかを明らかにすることを試みた。各特徴量の表情認識精度への貢献度を求めるため、ニューラルネットワークによる表情識別器を作成し、deep Taylor decompositionに基づく手法により、特徴量の貢献度を求める手法を構築した。その結果、各表情の識別に有用な特徴量として、表2に示す特徴量が得られた。

表2 各表情の識別に有用な特徴量

表情	識別に有用な特徴量
怒り	f_2, f_5, f_{10}, f_{11}
嫌悪	$f_2, f_5, f_{10}, f_{11}, f_{13}, f_{15}$
恐怖	$f_1, f_2, f_5, f_{10}, f_{11}$
幸福	$f_2, f_3, f_4, f_{10}, f_{11}, f_{13}$
悲しみ	$f_2, f_5, f_{10}, f_{15}, f_{16}$
驚き	$f_2, f_4, f_{10}, f_{11}, f_{12}, f_{15}$

表2に示す結果より、 f_2 (眉と目の距離) や f_{10} (口角の上がっている度合い)、 f_{11} (口角の角度) といった特徴量は、多くの表情の識別に有用であると判定された。一方で、 f_6 (口の内側の面積)、 f_7 (口の外側の面積)、 f_8 (口の内側の縦横比)、 f_9 (口の外側の縦横比)、 f_{14} (口と鼻の間の距離) は、いずれの表情においても有用であると判定されなかった。このことから、眉と目の位置関係や、口角の変化が、様々な表情の認識に有用である可能性が示唆された。また、口の大きさや形状などは、表情の識別にはあまり有効に働かない可能性が示唆された。表情認識に関する研究では、どのような特徴量が表情の識別に有効であるかを示しているものは少ないため、本研究から、有用な知見が得られたものと考えられる。これらの知見をもとに、より精度の高い表情認識・表情強度推定法を確立することが今後の課題である。

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 6件/うち国際共著 0件/うちオープンアクセス 0件）

1. 著者名 Shota Sakaue, Hiroki Nomiya, Teruhisa Hochin	4. 巻 721
2. 論文標題 Estimation of Emotional Scene from Lifelog Videos in Consideration of Intensity of Various Facial Expressions	5. 発行年 2017年
3. 雑誌名 Studies in Computational Intelligence	6. 最初と最後の頁 121-136
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-319-62048-0_9	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Mitsuaki Maeda, Hiroki Nomiya, Shota Sakaue, Teruhisa Hochin, Yukiko Nishizaki	4. 巻 -
2. 論文標題 Emotional Video Scene Retrieval System for Lifelog Videos Based on Facial Expression Intensity	5. 発行年 2017年
3. 雑誌名 Proc. of 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing	6. 最初と最後の頁 551-556
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/SNPD.2017.8022777	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kazuya Sugawara, Hiroki Nomiya, Teruhisa Hochin	4. 巻 -
2. 論文標題 Improvement of Emotional Video Scene Retrieval System for Lifelog Videos Based on Facial Expression Intensity	5. 発行年 2018年
3. 雑誌名 Proc. of 5th International Confefence on Computational Science/Intelligence and Applied Informatics	6. 最初と最後の頁 108-113
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/CSII.2018.00026	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Yamato Shinohara, Hiroki Nomiya, Teruhisa Hochin	4. 巻 -
2. 論文標題 Estimation of Facial Expression Intensity for Lifelog Videos Retrieval	5. 発行年 2018年
3. 雑誌名 Proc. of 5th International Confefence on Computational Science/Intelligence and Applied Informatics	6. 最初と最後の頁 133-138
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/CSII.2018.00030	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Naoki Imamura, Hiroki Nomiya, Teruhisa Hochin	4. 巻 -
2. 論文標題 Extraction of Useful Features from Neural Network for Facial Expression Recognition	5. 発行年 2019年
3. 雑誌名 Proc. of 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing	6. 最初と最後の頁 221-226
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/SNPD.2019.8935652	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Naoki Imamura, Hiroki Nomiya, Teruhisa Hochin	4. 巻 8
2. 論文標題 Finding Useful Features for Facial Expression Recognition and Intensity Estimation by Neural Network	5. 発行年 2020年
3. 雑誌名 International Journal of Software Innovation	6. 最初と最後の頁 68-84
掲載論文のDOI (デジタルオブジェクト識別子) 10.4018/IJSI.2020040105	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計1件 (うち招待講演 0件 / うち国際学会 0件)

1. 発表者名 今村 直樹
2. 発表標題 表情認識におけるニューラルネットワークからの有用な特徴量の抽出
3. 学会等名 第24回電子情報通信学会関西支部学生会研究発表講演会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	篠原 大和 (Shinohara Yamato)		

6. 研究組織（つづき）

	氏名 (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	今村 直樹 (Imamura Naoki)		
研究協力者	菅原 和弥 (Sugawara Kazuya)		
研究協力者	前田 光晶 (Maeda Mitsuaki)		