

爆笑カメラ：笑い声により自然な笑顔を撮影するカメラシステム

伏見 遼平 *1 福嶋 政期 *1 苗村 健 *1

Laughin'Cam: Camera System to Induce Natural Smiles with a Laughter Sound

Ryohei Fushimi*1, Shogo Fukushima*1 and Takeshi Naemura*1

Abstract – When taking a photo, such verbal prompts as “cheese” or “smile” are often used by photographers to get natural smiles from their subjects. However, since the obtained smiles usually look forced or resemble a grimace, capturing a natural smile is difficult. We propose an active camera system called “Laughin’Cam” that obtains natural-looking portraits by eliciting spontaneous smiles by presenting the sound of laughter. The proposed system utilizes an emotional contagion effect, which is the tendency of emotional behavior to spread from person to person. We conducted an evaluation experiment with our proposed system using computer vision and also performed a subjective experiment. Our results suggested that our system could efficiently induce naturally-looking and spontaneous smiles.

Keywords : Photography; Active camera; Infectious laughter; Emotional contagion; Involuntary facial expression; Natural smile.

1. はじめに

カメラを向けられた時、思わず緊張したり顔がこわばったりした経験や、うまく笑顔が作れなかった経験はないだろうか。多くの人は、普段は自然な表情を作られていても、カメラを向けられた時は不自然な笑顔になってしまう。本稿では、この課題の解決を目的とし、シャッターを切る前に笑い声を再生することで自然な笑顔を撮影するアクティブカメラシステムを構築し、パラメータについて評価を行う

近年、写真を撮影したり他人の写真を見る機会は大きく増えた。写真共有サービス Instagram では、サービス開始から4年間で30億枚もの写真が共有されている^[1]。このうち題材として顔を含む写真は全体の30%を占め、他の題材に比べてコメント等のリアクションも多い^[2]。

写真の流通の変化に応じて写真を撮影する手段も変化したが、顔写真撮影において“カメラを向けると表情がこわばる”という問題は未だ存在している。この課題への対処法は主に撮影者による声掛けや指示についてのアドバイスが中心であり^{[3],[4]}、システムによる技術的検討はなされてこなかった。一方で、撮影後に加工を行うことによって表情をより自然に見せる手法^{[5],[6]}はあるものの、表情そのものへの加工に抵抗を持つ人は多い。

そこで著者らは、“つられ笑い”現象に着想を得て、撮影時に笑い声を提示することで自然な笑顔を撮影

するシステムを構築し、その効果やパラメータについて検討を行った。このような撮影時に笑顔を誘うための音声を再生するようなシステムは市販品やスマートフォンアプリなどですでに公開しているものもあるが、本稿のような考察が加えられたことは今までなかったと考えている。笑い声は誰もが理解し慣れ親しむことができるものであり、つられ笑いも広く共有された現象である。よって、笑い声による笑い誘発を利用した本システムは、文化・国籍・年齢を問わず効果を発揮し、なおかつ仕組みを理解することのできる、普遍性の高いシステムであると考えられる。

本稿は提案システムの効果を検証し、メカニズムについて考察を加える。図1に本研究で検証するシステムによって撮影された結果の一例を示した。

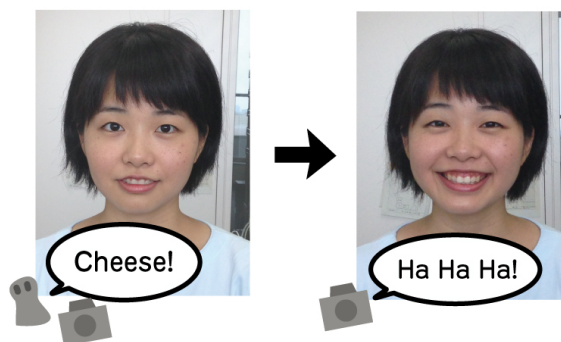


図1 撮影された代表画像。(左: 統制条件, 右: 幼児の笑い声を再生しながら撮影した画像)

Fig.1 Representative photograph from our evaluation experiment. (Left: portrait taken with shutter sound. Right: portrait induced by child's laughter)

*1: 東京大学

*1: The University of Tokyo

本稿では提案システムのプロトタイプを用いて撮影したデータを対象に、笑い声再生開始後シャッターを切るまでの遅延というパラメータについて、つられ笑いが最も顕著になるように評価を行った(実験1)。さらに他者による主観評価実験、被験者自身による主観評価実験を行い、笑顔の度合いや笑顔の自然さ、プロフィール写真としてのふさわしさなどについてシステムの有効性を確認した(実験2, 3)。

2. 関連研究

本研究は、シャッターを切る前に笑い声を再生することで自然な笑顔を撮影するカメラシステムを提案するものである。この章では、6つの節に分けて関連研究を紹介しながら、各節の中で本研究の位置づけを明らかにしていく。

2.1 笑顔の随意性

人の笑顔には2つの異なるタイプがある。(1)大頬骨筋と眼輪筋両方の動きが観察される Duchenne Smile と、(2)大頬骨筋のみの動きしか見られない non-Duchenne Smile である^[7]。このうち眼輪筋は不随意であり、(1)Duchenne Smile が真の笑い、(2)non-Duchenne Smile は愛想笑いの表出であるとされる。ただし近年では Duchenne Smile を意識的に表出することができるとする研究もある^[8]。

本研究で提案するシステムが撮影することを目指すのは Duchenne Smile、すなわち意識して表出できるものではない不随意的眼輪筋の収縮を伴う笑顔である。この不随意的な笑顔を本稿では「自然な笑顔」と呼ぶ。

2.2 笑い声提示による笑いや笑顔の誘発

笑い声の伝染現象は古くから研究されている。Provine は、実験協力者に1分おきに18秒間の笑い声を聞かせることを10回繰り返し、初回では笑顔については9割、笑いについては6割の協力者について誘発することに成功したが、繰り返しにより笑顔・笑いを誘発できた割合は減少し、不快感を感じたという報告もあった^[9]。“情動の感染現象 (Emotional Contagion)”と呼ばれるこの伝染現象は学習ではなく、無意識的で生得的な行動とされている^[10]。この現象については社会的に近い集団による笑い声の方が、そうでない集団の笑い声よりも笑いの伝染を引き起こしやすいことが指摘されている^[11]。また、自然な笑いか作り笑いという表出の種別よりも、被験者が笑いを自然だと感じたかが情動の伝染において重要であり、^[12]。誘発に用いる笑い声自体は必ずしも非随意的に起こった笑いによるものである必要はない。

笑い (laugh) と笑顔 (smile) は日常会話においてはよく混同されて用いられるが、実際には異なる概念である。本システムは、笑いに伴って発生する笑い

声 (laughter) が、(しばしば幸福感情や笑いという情動体験を伴いながら) 笑顔を誘発する効果を持つという現象を基盤にしている。この現象の機序に関しては Provine らにより複数の仮説が提案されている^{[10][12]}。

次に笑い誘発効果の応用例について取り上げる。「笑い袋」は、1969年ごろ流行したおもちゃで、ボタンを押すとシュールな笑い声がとめどなく流れるというものである。「くすぐりエルモ (Tickle me Elmo)」^[13]は、子供に人気のキャラクター「エルモ」のぬいぐるみであり、腹部を触るとエルモが笑う。嶋本らは、プレゼンテーション時に聴衆のPCから笑い声を再生し、笑いや拍手を誘発するシステムを提案している^[14]。また笑い声を音声や映像コンテンツへ応用した事例は多く存在し、一般的にラフトラック (Laugh Track) と呼ばれる。ジョーク集にラフトラックを付加することで、人のおかしみが増幅されて感じられたという研究成果が報告されている^[15]。福嶋らはこれらの研究を元に「笑い増幅器」を提案し、実装した^[16]。

本研究で提案するシステムは、このように広く知られ、検証されてきた笑い声の提示により笑顔を誘発する効果を、写真撮影のために応用するものである。

2.3 緊張と緩和による笑顔誘発

18世紀の哲学者カントは「笑いは期待が俄に無に消失することから生じる」と述べ^[17]、落語家である2代目桂枝雀も、笑いの起因について「生理上で最初に緊張があり、それが緩和されると笑いが生じる」という理論を唱えた^[18]ように、笑顔や笑いとは緊張は切っても切り離せない関係にある。

これによれば Provine の初回の結果は情動感染の効果に加えて、講義中に心理学実験を行う緊張が笑い声という音声により唐突に破られたことにもよるものであると考えることができる。Provine の実験^[9]で見られた、提示を繰り返すごとに誘発効果が大きく落ちていった現象は、「笑い声が笑いを誘う」という情動感染の効果だけでは説明できないが、「提示を繰り返すにつれ、緊張感が失われる」とすることで説明ができる。

本研究で提案するシステムも、笑い声より笑顔が誘発される作用を用いたものであるが、笑いを引き起こす効果のメカニズムは情動感染だけでなく、緊張が破られる驚きでも説明できる。そのため、繰り返しにより効果が落ちていたり、システムの仕組みを体験の前に事前に知っているとも効果が薄れる可能性がある。

2.4 カメラシステムから被撮影者への働きかけ

Cheese Cam^[19]、EyeCatcher^[20]、“笑顔がとれるこどもカメラ”^[21]はすべて視覚的な素材を用いたアクティブカメラシステムである。Cheese Cam は表情を模したイラストを提示し、表情模倣のはたらきを用いて自然な笑顔を撮影するシステムである。EyeCatcher

爆笑カメラ: 笑い声により自然な笑顔を撮影するカメラシステム

はカメラのホットシューに視覚刺激を提示するディスプレイを装着し自然な表情を引き出すシステムである。“笑顔が撮れる こどもカメラ”は、幼児の自然な笑顔を写真におさめるためのスマートフォン向けアプリケーションであり、画面にキャラクターを表示させ、これを動かすことで幼児の興味を引き、カメラの方を向かせてシャッターを切る。これらの視覚コンテンツを用いたアクティブカメラシステムは、どれも視線がカメラの光軸とわずかに一致しないという問題があるが、今回提案する聴覚コンテンツを用いたシステムでは、この問題を特別な道具なしに回避することができる。

今回検証するシステムに類似したコンセプトの商品は、すでに発売・配布されているものもある。また市販デジタルカメラのうち、1994年に発売されたPolaroid 636 Talking Camera (ポラロイド社) や、プリセットされた音源を再生した後にシャッターを切るという機能があり、またキヤノン社の一部機種 (IXY DIGITAL 930IS, Powershot G11 など) でも、ユーザが付属ソフトで、撮影時に再生されるシャッター音を設定できるものがある。ガラケーと呼ばれる日本製フィーチャーフォンでもシャッター音を設定できる機能がある。本研究はこのようなシステムについて統制を加えたうえで効果や最適なパラメータについて検証を加えたものである。

3. カメラシステムの設計

この章では、制作したカメラシステムの設計および実装について述べる。

3.1 カメラシステムの要件定義

提案するカメラシステムの満たすべき要件を示す。カメラシステムは撮影機能とそれと同期した音声提示装置を備える必要がある。笑い声の繰り返しによる笑い誘発効果の減少を緩和するために、複数の音声コンテンツを利用でき、ユーザが好きな音声コンテンツを指定したり、システムに音声コンテンツを任せることができることが望ましい。ユーザがシャッターボタンを押すと、まず音声提示装置から音声の再生を始め、一定時間後に静止画を撮影する。もしくは、音声の再生を始める前に動画の録画を始め、一定時間後に録画を終了するものとする。

3.2 カメラシステムの設計・実装

本研究では、簡便に同期を実現できる、音声再生から撮影までの遅延を自由に制御できる、さらに簡単にセットアップ・配布することができるという利点を持つスマートフォン向けのアプリケーションとして実装することとした。

制作したプロトタイプはスマートフォン (iPhone 6),

スピーカ (Logitech Mini Boombox, Logitech Corp.) によって構成されている。事前にインストールした音声を選択し、ボタンを押した時刻から音声を再生しながら動画の録画を始め、タイマーで音声を再生・録画をストップするアプリケーションを Objective-C で記述し、スマートフォンにインストールした。

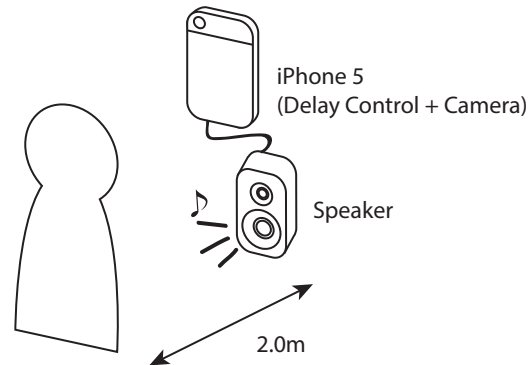


図2 システムの構成および撮影状況
Fig. 2 System setup and shooting situation

4. 実験

この章では、提案システムのプロトタイプを用いて撮影したデータを対象に、コンピュータビジョンによる笑顔尺度を用いて、笑い声再生後シャッターを切り撮影するまでの遅延というパラメータについて評価する (実験1)。また、撮影された表情の自然さを他者や被撮影者自身による主観評価を用いて確かめる (実験2, 3)。

4.1 音声コンテンツの選定

実験で用いる音声コンテンツを選定するために笑顔を誘発しやすい音声コンテンツを選定する実験を行った。ロイヤリティフリーの音声素材を提供する Web サイト audioblocks^[22] から笑顔を誘発する効果を持つと感じた笑い声音声素材を選定し、下記に示す5種類の笑い声 (いずれも10秒程度) を用意した。これらを3人の協力者 (男性2名, 女性1名) に聴かせて、表情を観察し感想を聞いた。結果として、笑顔を誘発する効果の大きかった (1) 幼児の笑い声 (3) 男性の笑い声を本実験に用いることにした。

1. 幼児の笑い声
2. 少年の笑い声
3. 男性の笑い声
4. 男性の3名の笑い声
5. 多人数の笑い声 (ラフトラック)

4.2 撮影の手続きおよび条件

SNS および口頭により実験への協力を募り、集まった21-36歳の撮影実験協力者19名 (以下、被撮影者) に対して実験を行った。なお、被撮影者は東京大学・東

京外国語大学の学生および東京大学の大学職員であった(男性8名, 女性11名, 平均23.1歳).

実験は静かな実験室で, 撮影者と机を挟んで対面する状況で行われた. カメラデバイスをスタンドに設置し, 撮影者はスタンドの右斜め後ろで被撮影者に教示を与えたのち, そのままデバイスを操作した.

実験条件として, 音声を4.1節の予備実験で選定した「男性の笑い声」「幼児の笑い声」に対照条件として「一眼レフカメラのオートフォーカス合焦音+シャッター音(対照条件)」を加えた3条件(表1)で実験を行った.

被撮影者には, この実験は写真を撮る際のシャッター音の表情への影響を調べる実験であることが伝えられ, 「いつも写真に撮られるときの笑顔で映ってください」という教示が与えられた. すなわち3条件とも, 音声の提示がある前は作り笑顔を作った状態で音声提示された. さらに注意としてシャッター音には長いものも短いものもあること, 撮影中はなるべくカメラレンズを見つめるようにすること, 合図があったのちにシャッター音が再生されることを伝えたのち, 3条件で合図を出してから撮影を行った. 順序効果を相殺するため実験条件の順序はラテン方格法により割り当てられた. 3条件すべての撮影後, それぞれの体験についてどういう感想を持ったかを聞いた.

表1 実験条件
Table 1 Experimental Conditions

条件	音声の内容	継続時間
実験条件1	男性の笑い声	10秒
実験条件2	幼児の笑い声	10秒
統制条件	合焦音+シャッター音(対照)	2秒

女性2名から, 男性の笑い声は不快だという意見があった. また女性2名から, 男性の笑い声よりも赤ちゃんのほうが笑いやすいという意見があった. そのうち1名は, それは自分が女性だからではないかという意見を付した.

4名の被撮影者は, 動画の撮影中に笑いを我慢し, 撮影が終了したあと吹き出すように, またこらえていた笑いを解放するように笑った. 実験室という環境の緊張や, 羞恥心などの抑制的な感情から, 動画の撮影中に笑うことを我慢していたと報告した男性がいた.

最後に, 代表的な画像を図3, 4に示す. いずれも音声の流れ始めてから1.5秒後の写真である.

4.3 実験1: コンピュータビジョンを用いた評価

本研究では, コンピュータビジョンによる笑顔尺度の推定を行うことで最適なシャッター遅延の長さを決定し, 得られたシャッター遅延を用いて撮影した画像についてあらためて主観評価を行うという順序で評価実験を行った. 実験1では, まず表情認識ソフトウェ



図3 実験条件2
Fig.3 Cond. 2



図4 統制条件
Fig.4 Control.

アを用いて要因の相互作用を検討した.

本研究では, 公開されている顔認識・特徴抽出APIである“ReKognition API”を用いた. ReKognition APIは画像を送信すると顔を検出し, その位置や推定年齢, 笑顔尺度などを返す画像解析サービスである. 笑顔尺度は0~1の浮動小数点値で示されており, この値を顔全体の笑顔の度合いとみなすことができる. ReKognition APIを公開しているOrbeus社は, 機械学習を用いた顔認識エンジン専門の企業で, 多数の企業に顔認識システムを提供している^[23].

分析の科学的基礎づけは公開されておらず, この値の信頼性に対して評価を加えることは難しい. 連続した表情の変化に対して, ほぼ連続した値が得られていることから, ある程度信頼できるものと考えたが, 笑顔尺度として公開されている尺度が, 実際には幸福, 驚き等の尺度も含んでいるという可能性もある. 実際に自然な笑顔になっているかどうかは実験2,3で主観評価を行うこと, この実験は実験2,3で評価に使うタイミングを決定するという目的を持っていることから, 十分であると考えた.

4.3.1 評価方法

撮影された10秒間の動画から, 15[fps]で静止画像を切り出し, 条件間の表情の差異, および表情の時間変化について分析を行った. その後, ReKognition APIを用いて, 各画像の笑顔尺度の測定を行った. ただし, 目を閉じている場合, 笑顔尺度は目を閉じていない前後のフレームよりも大きく値が低下する. このことから, 目を閉じていると判定されたフレームとその前後2フレームについては, 3フレーム前の笑顔尺度の値と3フレーム後の笑顔尺度の値の平均値を採用した.

4.3.2 結果と考察

笑顔尺度の時間変化について, 条件ごとに全被撮影者の尺度を平均した値をプロットしたものを図5に示す. 表1の通り, 実験条件1と2では0秒から9秒まで音声の流れていたが, 統制条件では0秒から2秒の区間のみで音声 flowed.

爆笑カメラ: 笑い声により自然な笑顔を撮影するカメラシステム

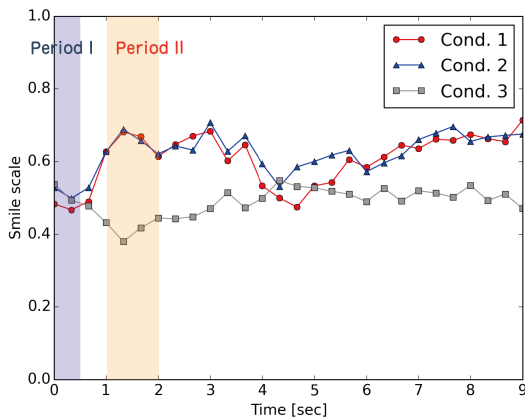


図5 笑顔尺度の時間変化
Fig. 5 Plot between smile scale vs. time

4.2.1 節で述べたように、音声流れる前には被撮影者は作り笑いを作っている。図5からは、音声流れ始めてすぐの0.0-0.6秒(Period I)は条件間に差がほとんどなく、1.0-2.0秒(Period II)の条件間の差が最も大きいことが認められた。

再生した音声(男性・幼児・統制)とタイミング(区間1: 0.0-0.6秒, 区間2: 1.0-2.0秒)を独立変数, 得られた笑顔尺度の値の区間中の平均値を従属変数とした3x2の分散分析を行った。結果として, タイミング要因の主効果が認められ, シャッター音要因の主効果は認められなかった。タイミング要因とシャッター音要因の交互作用が認められたため, さらに単純主効果の検定を行ったところ, シャッター音要因が幼児笑い声・青年笑い声の場合で, タイミング要因の単純主効果が認められた($F(1, 18) = 36.7, F(1, 18) = 35.5$)。統制条件下においては, 主効果は認められなかった($F(1, 18) = 3.41$)。

笑顔の誘発には遅延が存在し, 誘発された笑顔が最も顕著になるのは実験条件1, 実験条件2ともに音声再生からおおよそ1.0秒から2.0秒後であることがわかった。さらに分散分析の結果, 実験条件1(男性笑い声), 実験条件2(幼児笑い声)で笑顔を誘発できていること, さらに対照とした統制条件ではその効果が現れないことが検証できた。シャッター音が鳴り終わるのが音声再生からおおよそ1.6秒程度であることを考慮し, 以後の主観評価実験は, このうち2.0秒を, 笑い声音声が最も顕著に笑顔を誘発できるシャッター遅延時間の代表値として用いた。実際の写真撮影システムを構築する場合は, 別の音声を扱う場合には今回のような実験を繰り返す必要がある。

4.4 実験2: 他者による主観評価

次にコンピュータビジョンを用いた実験で得られた知見を元に, 被撮影者とは無関係の評価実験協力者によ

表2 実験2において評価者の評価した画像数
Table 2 Number of images participants evaluated in expr. 2

日次	1日目	2日目
被撮影者数	19	6
音声条件数	3	3
タイミング条件数	3	3
顔部位条件数	1	3
合計画像数	171	162

る主観評価実験を行った。ここで用いた画像は, 音声再生から0.0秒, 2.0秒, 6.0秒の時点での画像である。

4.4.1 評価方法

主観評価実験のためのシステムはWebアプリケーションとして構築され, インターネット上で実施された。評価者はWebブラウザ上で, 下記の質問について, それぞれ7段階のリッカート法で報告した。(1: 全くそう思わない, 7: とてもそう思う)

- 設問1: この人を自然な笑顔だと思いますか。(自然さ)
- 設問2: この人は作り笑いをしていると思いますか。(つくり笑いらしさ)

著者と同じ研究室に属する9名(男性4名, 女性5名)が評価実験に参加した(以下, 評価者。被撮影者とは異なる)。課題の実施は2日間に分割され, それぞれ約160枚の画像について評価を行った。課題実施の間には48時間以上の間を置いた。実験中には2度, 3分間の休憩を取るよう指示し, 全員が60分以内に評価を終えた。1日目・2日目ともに, 音声再生から0.0秒, 2.0秒, 6.0秒の時点での画像を取り出したものを用いた。ただし, 2日目は評価者の負担を抑える観点から, 実験時間を60分以内に収めるために被撮影者のうち男女3名ずつ合計6名について行われた。結果として評価者の評価した画像数は表2の通りである。

“もし得られた表情がDuchenne Smileであれば, 目元の画像の自然さや, 作り笑いしにくさは, 口元の画像や顔全体の画像に比べて大きく判断されるはずである”という仮説を立て, この仮説に基づいて1日目の課題では顔全体の画像について評価させ, 2日目の課題には目元のみ, 口元のみ, 顔全体についての画像を評価させた。これは得られた笑顔がDuchenne Smileであるかどうかを確かめるためである。画像の切り抜きについては, ReKognition APIを用いた顔の各要素の位置の推定情報を用いた(図6)。

4.4.2 結果と考察

各設問・部位・音声コンテンツ条件について, (再生開始から2.0秒の画像の評価値 - 0.0秒の画像の評価値)の値は, それぞれの音声コンテンツが評価値に与

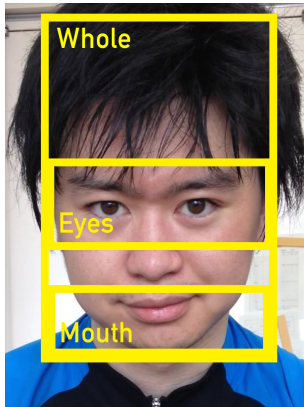


図6 実験2に用いられた切り抜き部位
Fig.6 Cropped regions for expr. 2

えたの持つ効果の量であると考えられる。各部位・条件ごとの誘発量を、被撮影者ごとに平均した値の分布を箱ひげ図を用いてプロットしたものを、図7(設問1)および図8(設問2)に示す。

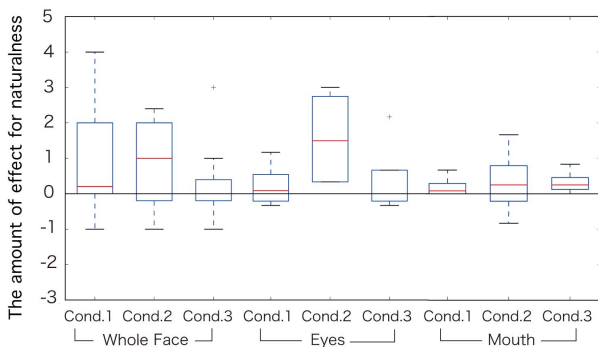


図7 自然さに対する誘発量の分布
Fig.7 Distribution of the amount of effect for naturalness

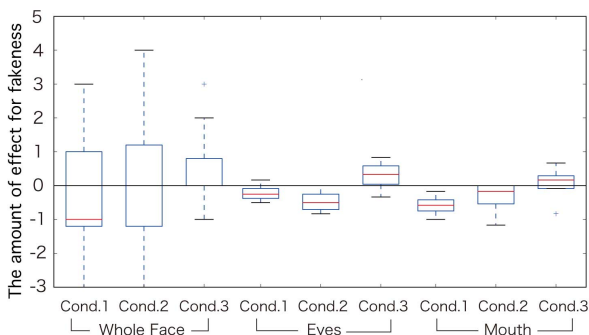


図8 つくり笑いらしさに対する誘発量の分布
Fig.8 Distribution of the amount of effect for contrivance

再生音声(男性・幼児・シャッター)と、タイミング(時点1: 0.0秒, 時点2: 2.0秒)を独立変数(いずれも被験者内)、各部位(全体・目元・口元)について得られた得点の値(自然さ得点)を従属変数とした3x2の分

散分析を行った。結果として、全体・目元の自然さ得点に関してはタイミング要因の主効果が認められ、音声要因の主効果は認められなかった。口元の自然さ得点に関しては、いずれの主効果も認められなかった。全体・目元の自然さ得点に関しては音声要因とタイミング要因の交互作用が認められたため、さらに単純主効果の検定を行ったところ、再生音声および部位が(幼児笑い声, 顔全体)(青年笑い声, 顔全体)(幼児笑い声, 目元)の組み合わせの場合において、タイミング要因の単純主効果が認められた。(青年笑い声, 目元)の組み合わせに関しては有意傾向があった。

設問1の結果は顔全体や目元の笑顔の自然さについて、それぞれシステムが有効に寄与していることを示唆している。ただし、青年笑い声条件では、目元の笑顔の自然さに関しては有効に寄与されていることは示されなかった。また、本実験では口元では自然な笑顔を誘発する効果がなく、全体や目元に関しては効果があることが示され、提案システムが有効に Duchenne Smile を誘発できているという仮説を支持する結果となった。

4.5 実験3: 被撮影者本人による主観評価

本実験では、被撮影者が、撮影された自分自身の画像について評価した際の主観評価を行わせて、その結果について、実験2の結果と比較しながら検討する。

4.5.1 評価方法

撮影実験の6ヶ月後、19名の被撮影者に追加実験への参加を要請した。実験への参加の意志を表明した12名(男性6名, 女性6名)に、男性笑い声・幼児笑い声・シャッター音の3つの音声条件, 0.0秒・2.0秒の2つのタイミング条件の計6枚の画像を用意し、ランダムに並べ替えた上で送付し、全員から回答を得た。

被撮影者はそれぞれの画像について4問のアンケートに答えた。前半の2問は実験2と共通であり、実験2の結果と比較するために用意された。後半の2問は、撮影された画像について、他者に向けた自分のプロフィール写真として好ましい写真が撮影できているかを評価するために用意された。それぞれの設問について最も当てはまるものを、実験2と同じく7段階のリッカート法を用いて報告させた(1: 全くそう思わない, 7: とてもそう思う)。

設問1(自然さ) この写真のあなたは自然に笑っていると思いますか。

設問2(作り笑いらしさ) この写真のあなたは作り笑いをしていると思いますか。

設問3(友人プロフ) この写真は友人に見せるプロフィール写真にふさわしいと思いますか。

設問4(他人プロフ) この写真は知らない人に見せるプロフィール写真としてふさわしいと思いま

すか。

4.5.2 結果と考察

実験 2 と同じく、(再生開始から 2.0 秒の画像の評価値 - 再生開始から 0.0 秒の画像の評価値) の値の分布および平均値を、各部位・条件ごとに箱ひげ図でプロットしたものを、図 9 に示す。

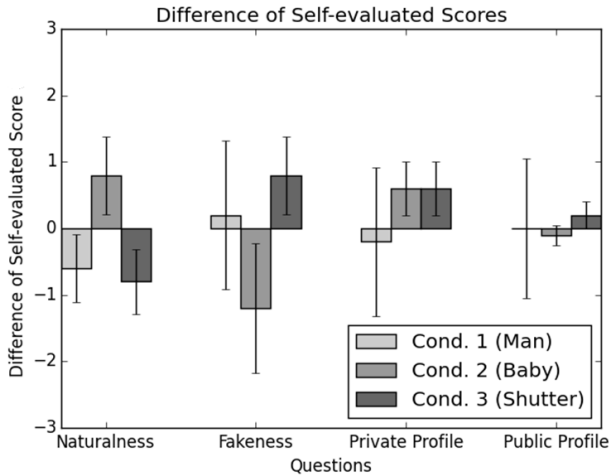


図 9 評価値の差の平均値 (エラーバー: 標準偏差)

Fig. 9 Average of the amount of effect (Error bar: standard dev.)

再生音声 (男性・幼児・シャッター) と、タイミング (時点 1: 0.0 秒, 時点 2: 2.0 秒) を独立変数 (いずれも被験者内), 各設問 (自然さ・つくり笑いらしさ・友人プロフ・他人プロフ) について得られた得点の値 (自然さ得点) を従属変数とした 3x2 の分散分析を行った。

結果, 設問 1~4 を通じて再生音声要因・タイミング要因の主効果は認められなかったが, 設問 1~3 に関して (自然さ・つくり笑いらしさ・友人プロフ) 再生音声要因とタイミング要因の交互作用が認められた。

自然さ設問, つくり笑いらしさ設問では, 幼児声再生時にタイミング要因による主効果が認められた ($p < .05$) が, 青年声・シャッター音再生時には認められなかった。これは, 被撮影者自身にとっての自然さ・作り笑いらしくさについて, 幼児声再生時にはそれぞれシステムが有効に寄与していることを示唆している。

全体として, プロフィール写真としてのふさわしさに関しては, 有効に寄与しているかどうかは不明であった。

実験 2,3 を通じて, 幼児声再生時に関してシステムの寄与が示された場合でも青年声再生時には示されない場合がある傾向がみられた。実際に撮影された写真を見ても, 幼児笑い声のほうが顔を崩して自然に笑っているように見える写真が多かった。実験 2,3 の結果から即座に「笑い声」自体が笑顔を誘発したと

結論付けるのは尚早であり, 「幼児の声」そのものが, その幼児図式に対する反応として微笑を誘発したと考えることも可能である。

ただ, 被験者数の不足が原因で検定力が不足していること, また実験 1 では尺度に大きな差が出ていなかったこととあわせると, 本実験の結果を踏まえても, 笑い声の主によらず, 笑い声そのものが笑顔を誘発しているという仮説はやはり否定されないものと考えられる。

最後に, 設問 1,2 について実験 2(他者評価)・実験 3(自己評価) の誘発量を比較したプロットを図 10 に示した。自己評価のサンプル数が少ないため検定は行わなかったが, 男性笑い声・シャッター音に関しては, システムの効果による表情の変化を, 他者評価に比べて自己評価のほうがよりネガティブに (すなわち, より自然でない, 作り笑いらしいと) 捉える傾向があった。この結果からは, 表情の自然さについて過小評価する傾向にあることが示唆される。

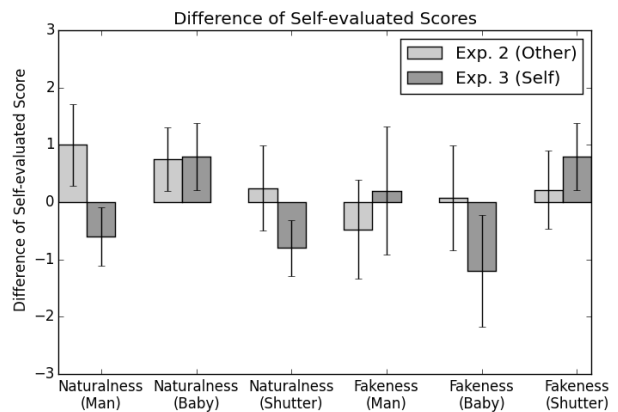


図 10 誘発量の平均値 (他者評価と自己評価の違い, エラーバー: 標準偏差)

Fig. 10 Average of the amount of effect (Error bar: standard dev.)

5. おわりに

本研究では, 音声を用いて笑顔を引き出す簡便な装置により実装できるアクティブカメラシステムを提案し, これを実装し自然な笑顔を撮影するという目的が達成できているかを評価した。

カメラシステムのプロトタイプとして, 音声コンテンツを再生しながらムービーを撮影するシステムを実装した。さらに予備実験で選定した 2 種類の笑い音声に統制条件を加えた 3 種類の音声を再生しながら撮影を行い, システムの有効性を確かめるため, コンピュータビジョン・他者による主観評価・被撮影者自身による主観評価による 3 つの評価を行った。

すべての実験を通して, 提案システムが他とともに自然な笑顔と認められる表情を引き出し撮影できてい

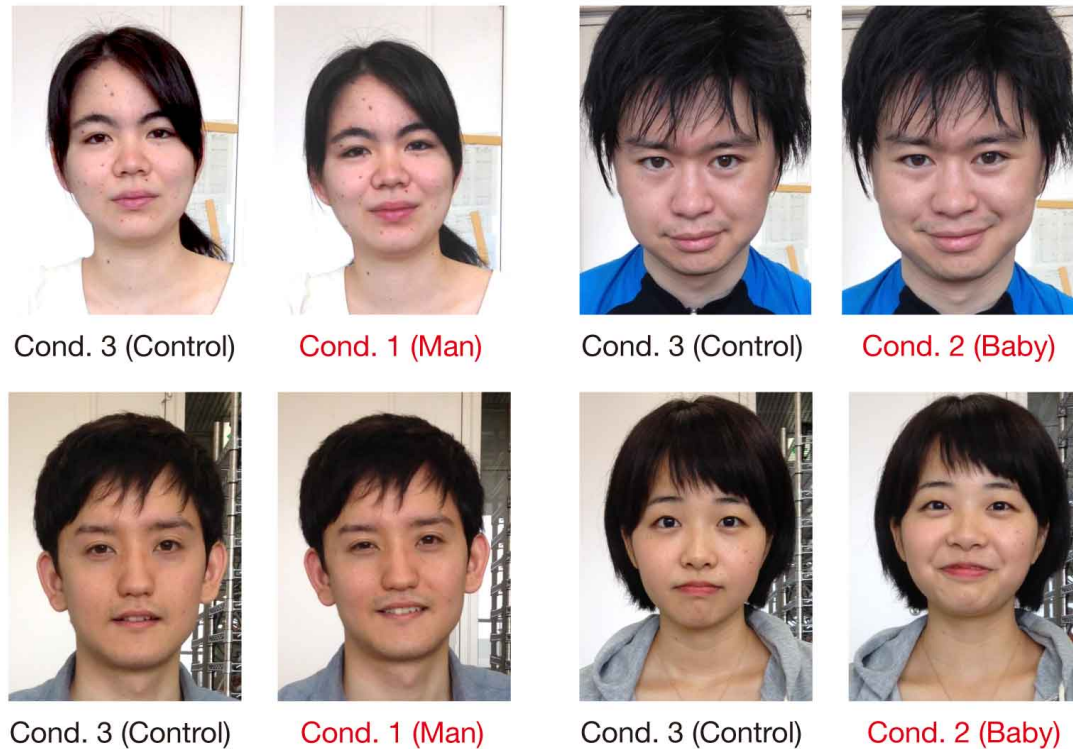


図 11 提案したシステムによって撮影出来た画像の例 (左: 統制条件, 右: 笑い声提示条件)

Fig. 11 Examples of obtained images (Left: Forced smiles, Right: Induced smiles)

るといふ仮説は支持されたが、コンテンツの種類と被撮影者の性別は効果の大きさに交互作用を及ぼすため、より注意深く選定する必要がある。システムによって撮影された画像の代表例を図 11 に示す。本研究で提案したシステムの効果は、個人差があり、実験協力者の中には全く効果のない者もいた。男性笑い声よりも幼児笑い声のほうが頑強な効果を発揮できたように、コンテンツによる違いも大きいと考えられる。今回のシステムを実用に供するためには、より個人差が小さく頑強な効果を持つコンテンツを選定すべきである点は課題である。

また本研究では、繰り返し使用した際の評価は行わなかった。実用のためには、多くの種類のコンテンツを用意してランダムな順序で流すなど、効果を減衰させないための工夫が必要である。

本研究でおこなった実験では、撮影時には撮影者本人の表情の変化が視界に入らないように工夫していたが、実験以外でシステムを利用した際、研究者本人が撮影した場合や、撮影者がシステムについて知らなかった場合に、撮影者自身も笑うことがあった。撮影者が笑っていたほうが被撮影者の笑いを引き出す効果は大きい。笑い声により、被撮影者・撮影者が同時に笑っていた場合の効果について測定する実験を行うこ

とで、この効果を測定することができると考えられる。また、被撮影者・撮影者にコンテキストに応じた視覚コンテンツを非対称に提示し、より効果の高いアクティブカメラの実装方法を模索することもできるだろう。

2.3 節でも見たように、笑いは緊張と深い関係にある。今回の撮影実験では、音声コンテンツ再生の際には合図があったので、2回目以降は笑い声などが再生されることや、そのタイミングは予測ができた。ボタンを押したあとランダムな時間をおいて音声再生を開始するなど、被撮影者が再生されるタイミングを分からなくすることで、撮影前の緊張感を高めることによって、より大きな効果が得られると考えられる。

今も様々な形で新しい撮影体験が求められ、日々新しい写真や映像の撮影方法について探求が行われている。本研究をさらに推し進めアクティブカメラの可能性を探求することは、写真撮影という文化の発展に貢献するものとする。

6. 謝辞

本研究は、JST, CREST の支援を受けたものである。

参考文献

[1] Instagram. Press news (2015). <http://instagram.com/press/>.

[2] Bakhshi, S., Shamma, D. A., , Gilbert, E. Faces engage us: Photos with faces attract more likes and comments on instagram. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14*, pages 965–974. ACM, New York, NY, USA (2014).

[3] nanapi. 写真を撮る時に自然な笑顔を引き出すコツ (2013). <http://nanapi.jp/17883>.

[4] 藪田織也. プロが伝授「自然な笑顔の作り方」 (2005). http://aska-sg.net/ht_photo2/029-20051031.html.

[5] Aradhye, H., Toderici, G., , Yagnik, J. Adaptive, selective, automatic tonal enhancement of faces. In *Proceedings of the 17th ACM International Conference on Multimedia, MM '09*, pages 677–680. ACM, New York, NY, USA (2009).

[6] Fujishiro, H., Suzuki, T., Nakano, S., Mejima, A., , Morishima, S. A natural smile synthesis from an artificial smile. In *SIGGRAPH '09: Posters, SIGGRAPH '09*, pages 59:1–59:1. ACM, New York, NY, USA (2009).

[7] de Boulogne, G.-B. D. , Cuthbertson, R. *The mechanism of human facial expression*. Cambridge university press (1990).

[8] Krumhuber, E. G. , Manstead, A. S. R. Can duchenne smiles be feigned? new evidence on felt and false smiles. *Emotion*, 9(6):807 (2009).

[9] Provine, R. Contagious laughter: Laughter is a sufficient stimulus for laughs and smiles. *Bulletin of the Psychonomic Society*, 30(1):1–4 (1992).

[10] Provine, R. Laughter. *American scientist*, pages 38–45 (1996).

[11] Platow, M., Haslam, S., Both, A., Chew, I., , Cud-don, M. “it’s not funny if they’re laughing”: Self-categorization, social influence, and responses to canned laughter. *Journal of Experimental Social Psychology*, 41(5):542–550 (2005.).

[12] 蔭山洋介, 新. 自然な笑いと作り笑いにおける情動の伝染: facial emg 及び内省報告による検討 (社会心理とコミュニケーション及び一般). 電子情報通信学会技術研究報告. HCS, ヒューマンコミュニケーション基礎, 105(306):49–54 (2005).

[13] Time.com all-time 100 greatest toys: Tickle me elmo (2011). <http://content.time.com/time/specials/packages/article/0\,28804\,2049243\,2048661\,2049231\,00.html>.

[14] 嶋本諒太, 宮下芳明. 笑いや拍手を誘発するプレゼンテーションシステム. In *インタラクション 2013*. 情報処理学会 (2013).

[15] Chapman, A. J. Funniness of jokes, canned laughter and recall performance. *Sociometry*, pages 569–578 (1973).

[16] 福嶋政期, 橋本悠希, 野澤孝司, , 梶本裕之. 笑い増幅器: 笑い増幅効果の検証. *ヒューマンインタフェース学会論文誌*, 12(3):199–207 (2010).

[17] Kant, I. *Critique of Judgment*. The Cambridge Edition of the Works of Immanuel Kant. Cambridge University Press (1790).

[18] 長島平洋. 生理的に見た笑いの分布: 桂枝雀の「緊張の緩和」論を検証するために (i). *笑い学研究*, (14):3–11 (2007).

[19] Lee, B. , Lee, W. Cheese cam: Unconscious interaction between humans and a digital camera. In *CHI '09 Extended Abstracts on Human Factors*

in Computing Systems, CHI EA '09, pages 4285–4290. ACM, New York, NY, USA (2009).

[20] Tsukada, K. , Oki, M. Eyecatcher: a digital camera for capturing a variety of natural looking facial expressions in daily snapshots. In *Pervasive Computing*, pages 112–129. Springer (2010).

[21] D2C, Inc. こどもカメラ (2012). <http://www.d2c.co.jp/news/2012/20130118-1609.html>.

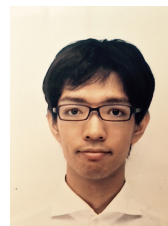
[22] Audioblocks. <http://audioblocks.com>.

[23] Orbeus, Inc. <http://orbe.us/>.

(2016年1月30日受付, 5月8日再受付)

著者紹介

伏見 遼平



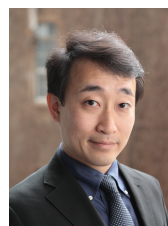
2015年東京大学工学部電子情報工学科卒業。同大学大学院学際情報学府修士課程に在籍中。音声を利用したヒューマンインタフェースの研究に従事。

福嶋 政期



2011年, 日本学術振興会特別研究員(DC2)。博士課程在学中に米国MITへ留学。2013年, 電気通信大学大学院博士課程修了。博士(工学)。2013年, 東京大学大学院情報理工学系研究科特任研究員。触覚や情動を誘発するヒューマンインタフェース, 実世界指向メディアなどの研究に従事。

苗村 健



1997年, 東京大学大学院工学系研究科電子工学専攻博士課程修了。米国スタンフォード大学客員助教授(日本学術振興会海外特別研究員)を経て, 2002年, 東京大学大学院情報学環 助教授。同情報理工学系研究科電子情報学専攻准教授を経て, 2013年, 同情報学環 教授, 現在に至る。メディア+コンテンツ, 実写に基づく映像合成, 複合現実感, 実世界指向情報環境, アート&エンタテインメントなどの研究に従事。博士(工学)。

