

笑い声呈示により自然な笑顔を撮影するカメラの提案

伏見 遼平^{1,a)} 福嶋 政期^{1,b)} 苗村 健^{1,c)}

概要：記念撮影で自然な表情を撮影するのは難しい．本研究では被撮影者に自然な笑顔の表出を促すカメラを検討している．本稿では笑い声が笑顔や笑いを誘発する現象に着目し，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カメラシステムを提案する．効果を検証するために，シャッター音・笑い声（青年）・笑い声（幼児）の 3 条件で表情の変化を比較した．さらに笑い声呈示から表情表出までの時間差に着目しシャッターを切るタイミングを検討した．

How to Prepare Your Paper for IPSJ Journal (version EC2014)

RYOHEI FUSHIMI^{1,a)} SHOGO FUKUSHIMA^{1,b)} TAKESHI NAEMURA^{1,c)}

Abstract: This document is a guide to prepare a draft for submitting to IPSJ Journal, and the final camera-ready manuscript of a paper to appear in IPSJ Journal, using L^AT_EX and special style files. Since this document itself is produced with the style files, it will help you to refer its source file which is distributed with the style files.

1. はじめに

カメラは被写体のありのままの姿を記録するためにデザインされている．ストロボなどを除けば、基本的にはカメラや付属する機器は被写体に対して働きかけを行うものは少ない．しかし、ポートレートなどの人物写真については、被写体となる人間とのコミュニケーションも重要である．カメラマンからは「ハイチーズ」「笑って」などの声掛けがなされるが、言語でコミュニケーションを取ったり、笑顔を作るように誘導したとしても、随意的な笑顔しか撮影することができない．自然な笑顔

本研究では、被写体にカメラそのものが働きかけ、笑顔を誘発するようなカメラシステムを研究する．笑顔を誘発するために、つられ笑いや思い出し笑いなどの非随意的なメカニズムを使えば、自然な笑顔が撮影できるはずである．中でも本論文では、笑い声が笑いを誘発する効果を利用し

て、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．また、システムをスマートフォン上で動作するアプリケーションとして実装し、実際に効果を検討する．

まず、予備実験として笑顔を誘発しやすい笑い声音を検討する．次に実験室の統制された環境で、2 種類の笑い声音と通常のシャッター音で撮影された笑顔にどのような違いがあるかを調べる．最後に、スマートフォンアプリとして配布し、実際の記念撮影のシチュエーションでどのような写真が撮影できたかを調べる．

2. 関連研究

2.1 カメラ側からのインタラクション（自然な表情を取るための試み）

本研究は、写真撮影のシーンにおける、カメラと被撮影者の間の非随意的・非言語的インタラクションとして位置づけることができる．

まず、被写体への光学的な働きかけとして、照明を用いて被写体を明るくみせるストロボが代表的である．キセノンランプや LED を用いたストロボは多くのカメラデバ

¹ 東京大学
UTokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo
a) fushimi@nae-lab.org
b) shogo@nae-lab.org
c) naemura@nae-lab.org

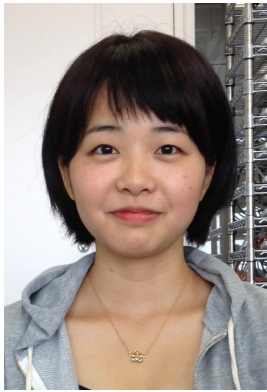


図 1 「笑って」の笑顔

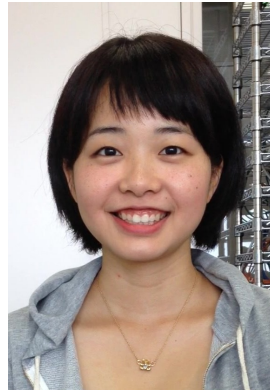


図 2 自然な笑顔

スに搭載されている．レフ板を用いて環境光を反射させて被写体にあてることも写真撮影には一般的である．人物写真においてはこれに加えて、瞳が表情の印象に影響を与えることから、ポートレート写真撮影のための書籍や Web サイトでは、被写体のまわりに照明や白いレフ板や布を置いてこれらの光を瞳に写しこむキャッチライトというテクニックが紹介されている．

次に、被写体への言語的なインタラクションについて整理する．記念撮影のシーンにおいてカメラマンによる被写体への語りかけや掛け声が多くみられ、これには位置やポーズ、表情についての指示だけではなく、被撮影者をリラックスさせようという意図の語りかけも多く見られる．言語的アプローチは非撮影者に表情の指示を伝えることができる点で有効だが、これは随意的な笑顔しか撮影できない．トークなどを通じて非随意的な笑いを意識して引き出すことのできるカメラマンも存在するが、簡単ではない．「チーズ」「ミッキー」などの語末に /i/ 音がある単語を発音させ、口角を上げさせる撮影の手法もある．またブリラ撮影機では、具体的なポーズの指示について、事前収録された音声再生されるが、インタラクティブ性はない．



図 3 プロカメラマンの写真撮影の例

最後に、非言語的アプローチについてまとめる．D2C のリリースしたスマートフォン向けアプリ「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アプリケーションである．スマートフォンの画面にキャ

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、シャッターを切る．これは用途を子供向けに限定しており、視覚的アプローチを取るものであるが、本研究には最も近い事例である．

2.2 笑い声呈示による笑い誘発

笑い声の伝染現象は古くから研究されている．Provine は、心理学の講義を履修した 128 名の学生に対し、18 秒間の笑い声を聞かせ、42 秒間の間を空けることを 10 回繰り返し、それぞれの試行について笑顔になったか否かおよび笑ったか否かを報告させた．繰り返しにより笑顔・笑いを誘発できた割合は大きく下がり、繰り返しの後半では不快感を感じたという報告も多かった．[Provine, 1992]

笑い声は音声・映像コンテンツに対する刺激の重畳としても用いられる（ラフトラック）．映像コンテンツに対して ... らは、多い方は 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を ... らは指摘している．最後に、笑い誘発を実世界のデバイスに応用した例について取り上げる．「くすぐりエルモ (Tickle me Elmo)」は、子供に人気のキャラクター「エルモ」のぬいぐるみであるが、腹部を触るとエルモが笑い転げる．

3. システム

本システムでは、笑い声を非撮影者に呈示しながら写真を撮影する必要がある．一眼レフなどのデジタルカメラのシャッターと同期して音声を再生する装置を検討したが、同期のための機構を特別に必要になってしまう．同期を簡単に実現でき、音声再生から撮影までの遅延を自由に制御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．

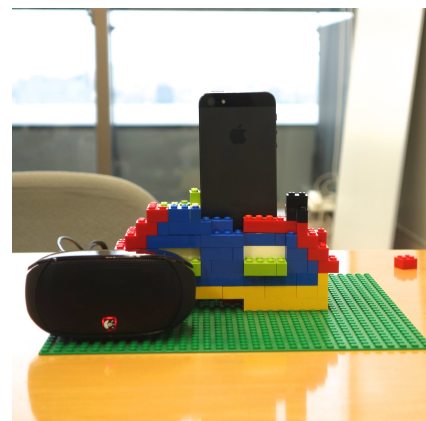


図 4 実装したシステム

本システムはスマートフォン、スピーカによって構成されている．ボタンを押した時刻から録画を始め、タイマーで、音声を再生し、録画をストップするアプリケーション

を Objective-C で記述し、iOS のスマートフォンにインストールした。アプリケーションは通常の写真撮影モードと、動画撮影するモードがある。また、著者以外でも実験ができるように、教示をアプリケーションの画面の中に埋め込んだ。

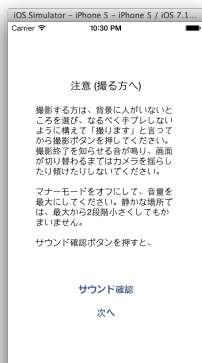


図 5 スクリーンショット

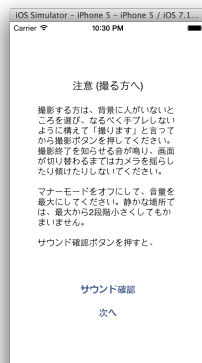


図 6 スクリーンショット

4. 予備実験

予備実験として、下記に示す 5 種類の笑い声を用意し、3 人の被験者（男性 2 名、女性 1 名）に聴かせた。笑顔を誘発する効果の大きかった (1) 幼児の笑い声 (3) 青年の笑い声を本実験に用いることにした。音声データはロイヤリティフリーの音声素材を提供する Web サイト audioblocks から取得した。

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

感想として「5. は笑えない」

5. 実験 1

5.1 目的

本実験では、実験室において本システムが有効に

5.2 実験条件および実験設備環境

21-36 歳の被験者 19 名に対して実験を行った。実験は静かな実験室で、実験者と二人の状況で行われた。実験条件として、音声を予備実験で選定した「青年の笑い声」「幼児の笑い声」「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（対象条件）」に設定した 3 条件で実験を行った。カメラデバイスをスタンドに設置し、実験者はスタンドの右斜め後ろで被験者に教示を与えたのち、デバイスを操作した。

被験者の属性は東京大学・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）

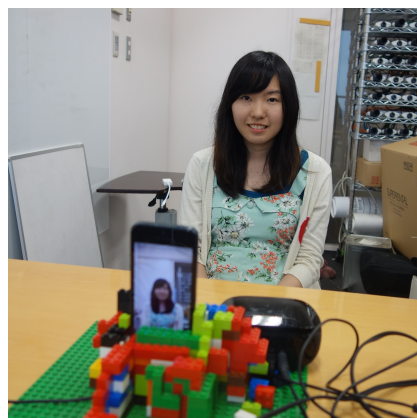


図 7 実験の様子

表 1 実験条件

条件 1	青年の笑い声
条件 2	幼児の笑い声
条件 3	一眼レフカメラの AF 合焦音＋シャッター音（対照）

5.3 手続き

被験者には、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を伝えたのち、3 条件で撮影を行った。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。

5.4 評価方法

撮影された 8 秒間の動画から、15fps 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。

まず、Rekognition API を用いて、各画像について笑顔尺度の測定を行った。Rekognition API は Orbeus 社による表情認識 API で、画像を送信すると映っている表情について笑顔尺度を解析し、0～100 の値で結果を返す。この値を顔全体の笑顔の度合いとみなし、値の時間変化と条件間の差異を分析した。

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。

5.5 結果

笑顔尺度の時間変化を条件ごとに平均した値をプロットしたものを図 8 に示す。音声の流れ始めてすぐは全く変化がなく、1.0 秒ごろから笑顔度に条件間の差が認められる。

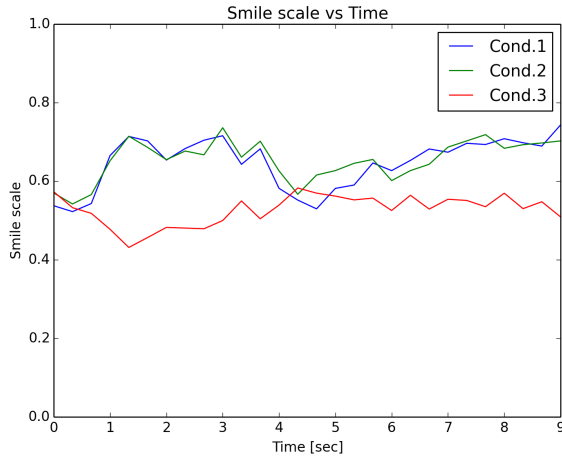


図 8 笑顔尺度の値の変化

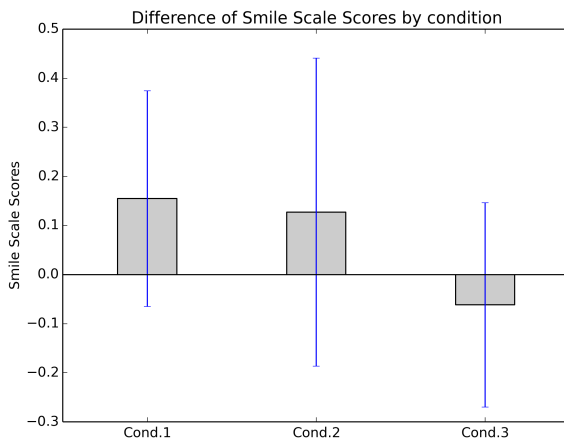


図 9 笑顔尺度の差の平均値

実際に効果が出ていることを確かめるため、0.0 秒～1.0 秒の笑顔度の値を作り笑いの笑顔の値、1.0～2.0 秒の笑顔度の値の平均値を音声によって誘発された笑顔の値の代表値として、条件内で作り笑い-誘発笑いの間に差があるかどうかウィルコクソンの符号順位検定を用いて検定した結果、作り笑い-誘発笑いの笑顔度の差は条件 1,2 で有意 ($p < .02$) だが、条件 3 では有意ではなかった。

さらに (誘発笑いの笑顔度) - (作り笑いの笑顔度) の差について、各条件群について対応あり 3 群の分散分析を行った結果、 $p < 0.05$ で有意に差があるという結果が出た。2 条件間で検定を行ったところ、条件 1-条件 3、条件 2-条件 3 について有意に差があり、Bonferroni の補正を考慮して

表 2 対応あり分散分析の結果 (p 値, カッコ内は Bonferroni の補正)

3 群	0.025
条件 1-2	0.760(2.280)
条件 2-3	0.004(0.012)
条件 1-3	0.041(0.123)

も条件 2-条件 3 間では $p < 0.05$ で優位に差があった。

また、笑い誘発の男女差を確認するため、各条件の (誘発笑いの笑顔度) - (作り笑いの笑顔度) の値についての男女差を、マン・ホイットニーの U 検定を用いて検定した結果を表 3 に示す。条件 1 (青年の笑い声) については女性のほうが、条件 3 (シャッター) については男性のほうが、それぞれ有意に笑顔が誘発されやすい (いずれも $p < .05$) という結果となった。

表 3 マン・ホイットニーの U 検定の結果 (p 値)

条件 1 (青年)	0.037	(男性) < (女性)
条件 2 (幼児)	0.156	(男性) < (女性)
条件 3 (シャッター)	0.046	(男性) > (女性)

さらに、音声の種類の違いが笑顔誘発度合いに及ぼす影響の男女差を調べるため、各被験者について 2 条件間の誘発笑いの笑顔度の差について、男女差があるかを確認した。条件 1 (青年) の誘発笑いの笑顔度と、条件 2 (幼児) の誘発笑いの笑顔度の符号を考慮した差は、女性のほうが有意に大きい ($p < .05$) がほかの 2 条件間では男女差は見られなかった。これは、女性は比較的青年よりも幼児の笑い声に笑いを誘発され、男性はその逆の傾向があることを示唆する。

最後に、代表的な画像を下記に示す。

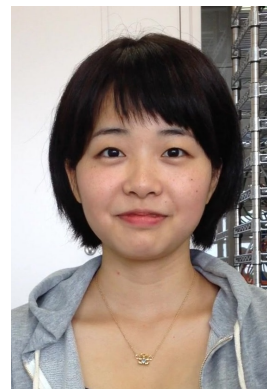


図 10 条件 2

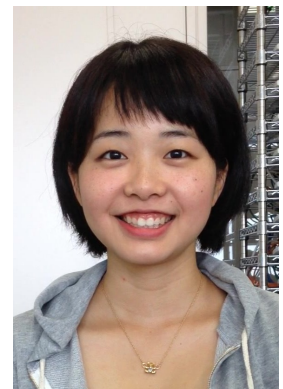


図 11 条件 3

5.6 その他

女性被験者 3 名から、男性の笑い声は不快だという意見があった。また女性の被験者 2 名から、男性青年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．これらの報告は，実際に女性では青年より幼児の方が笑顔度の誘発度合いが大きかったことと符合する．

また男性のうち2名は，「いつも写真に撮られるくらいの笑顔で映ってください」という教示に対し，「いつも写真には笑顔では映らない」と応えたため，「それでは，いつも写真に取られる表情で写ってください」と指示した．今回の実験では，このような被験者に対する対応は準備していなかった．

5.7 まとめ

本システムにおいて，条件1（青年笑い声），条件2（幼児笑い声）で笑顔を誘発できていること，さらに対照とした条件3ではその効果が現れないことを検証した．さらに男女差を分析した．また，男性青年の笑い声よりも幼児の笑い声のほうが笑顔を誘発しやすい傾向が女性では強かった．

また，笑顔の誘発にはディレイが存在し，誘発された笑顔が最も顕著になるのは条件1，条件2ともに音声再生からおおよそ1.5秒後であることがわかった．写真撮影システムを構築する場合は，この結果を元にシャッターを切る時間を設定すれば良い．ただし，この値は再生する音声によって大きく異なるだろうから，異なる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．

比較可能な尺度として Rekognition API の笑顔尺度を利用した．Rekognition を公開している Orbeus 社は，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．ただし，笑顔度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．しかし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られる．

6. 実験2

6.1 目的

実験1では，3種類の音声を呈示しながら撮影した写真についてコンピュータビジョンの技術を使って笑顔尺度を測定し，この値を比較することで，我々のシステムが有効に笑顔を誘発できていることが分かった．しかし実験1だけでは，今回構築したカメラシステムが我々の考える「自然な笑顔」であることはわからない．

本実験では，非撮影者とは異なる評価者に撮影された画像を見せ，笑顔の自然さを評価させることで，我々のシステムが自然な笑顔を誘発できているかどうかを検証することを目的とする．

6.2 実験条件および実験設備環境

node.js を用いて Web サーバを構築し，評価者
評価者には

6.3 手続き

被験者は，割り当てられた写真2枚の組について，どちらが自然な笑顔であるかを判定した．割り当てられた写真は同じ人物の，顔全体・鼻より上・鼻より下の3パターンがあった．

参考文献