

スマートフォンの加速度センサを用いた 微小不随意運動検出による動画視聴時の笑い評価手法

L-PoD: evaluation method of micro involuntary muscle movement of
laughing in video watching using acceleration sensor in smartphone

北田大樹¹⁾, 鈴木伸之介¹⁾, 白井暁彦¹⁾

Taiki KITADA, Shinnosuke SUZUKI and Akihiko SHIRAI

1) 神奈川工科大学

(〒 243-0211 神奈川県厚木市下荻野 1030 K1-501, kitada@shirai.la)

概要: 動画視聴者の随意・不随意の笑い動作に着目し、映像視聴時の被験者の不随意身体動作をスマートフォンの高精度な加速度センサで分類するアルゴリズム「L-PoD」およびアプリケーションを試作し、動画視聴者の不随意運動から笑いの開始タイミングの検出と統計処理による視聴者全体の時系列ごとの笑いの度合いの分類をおこない、動画視聴体験を物理評価する手法について提案・報告する。

キーワード: 笑い検出, 非言語評価, 加速度センサ, スマートフォン

1. はじめに

Youtube やニコニコ動画といった動画共有サイトは、バーチャルに時間や空間といった制約を超えて、動画コンテンツを共有する感覚を味わえる VR プラットフォームともみることができる。また、初音ミクや MikuMikuDance といったユーザー生成コンテンツの一般化や、Oculus Rift などの家庭用 HMD の一般化により、今後より多様な感情や表現力を支える VR コンテンツが、幅広いユーザによって創作されることが予測できる。しかしながら、従来のニコニコ動画のような、コメント文による言語的な共有方法だけでは、動画コンテンツ制作者にとって有益なフィードバックであるクオリティの向上や今後の創作活動のモチベーション維持に繋がる具体的なフィードバックを得られないことがある。例えばエンタテインメントコンテンツにとって基本的な情動である「笑い」について、「どこで笑ったのか?」といった単純な情報でさえ、エンドユーザにとって自然でストレスの無い方法で、リアルタイムで取得する方法は多くは存在しない。

本稿では、このような背景から、非言語による動画評価のフィードバックの実現として、笑いを狙った面白い動画コンテンツを視聴した際に生じる笑い動作に着目し、映像再生と同時に動画視聴者の笑いの情動による不随意運動で生じる加速度情報を分類するアルゴリズム「L-PoD」およびアプリケーションの開発と実験から得た加速度データの分析を通して、笑いの開始タイミングの検出と統計処理による視聴者全体の時系列ごとの笑いの度合いの分類による評価手法について、報告をする。

2. 関連研究

2.1 笑い増幅器

関連研究として、動画視聴者の笑い検出と笑いの増幅を行うことを目的とした福嶋らの「笑い増幅器」がある [1]。

笑い増幅器は、映像視聴者の胸部に貼付した電極から得られる筋電位から笑いの検出とその程度を算出し、映像視聴者の笑いに同調した笑いを増幅させる効果があるとされているラフトラック（他者の笑い声）を周囲に配置した人形から再生することで、笑う群衆の中に身を置いているような没入感を与えることにより、動画視聴者に笑いを誘発させることで、笑いを増幅できるシステムである。

2.2 スマートフォンやタブレット端末を用いた動画 視聴時のユーザー分析

倉野らは、iPad に搭載されているマイクロフォン、前面カメラ、加速度センサの 3 つのセンサを複合して用いるマルチモーダルセンシングによって、記録された動画視聴者の発話やしぐさ、動画閲覧時の状況から閲覧している動画に対する視聴者の興味変化の推測をおこない、閲覧者の興味変化を推定できる簡易的なシーンの発見を報告している [2]。しかしながら、スマートフォンに搭載されたカメラやマイクロフォンによる記録は、プライバシーやリソース使用に対するエンドユーザの利益や実用性を考えると、最良の方法とは言えないだろう。

我々は、エンドユーザの動画視聴というエンタテインメントエクスペリエンスを阻害しない方法で、笑いのような情動を抽出する方法を検討している。

加藤らは、抽象的なアニメーション作品視聴時における

瞬間的な加速度を計測・分析することで、非言語による視聴情報を取得する手法を提案した [3]。提案手法によって、視聴者から作家に視聴情報を還元するという課題に対して、iPhone3G に搭載された加速度センサによって取得した連続値に注目することで、意識的な入力を強いることなく、視聴者の身体的反応を計測することができるようになった。

3. 提案と開発

3.1 笑い評価アルゴリズム「L-PoD」の提案

加藤らの方法では、近年のスマートフォンに実装されている 16bit 深度以上の分解能をもった加速度センサを用い、適応型フィルタリングを実施し、動画視聴時の定在的な重力をキャンセルした加速度、つまり人間の動作の中でも高周波成分にあたる動作に注目し、0.2[G] 近辺の運動に注目する事で、情動に関係のある不随意運動が検出可能であることが報告された。

しかしながら、加藤らの研究はエンドユーザの不随意運動を、映像作家の仕掛けにフィードバックする視点で研究されており、より明確で利用価値の高い情動、例えば「笑い」に注目した実験をおこなっていない。

本報告では、加藤らのフィルタリング手法をそのまま利用し、エンドユーザの「笑い」に注目して、同一のスマートフォンデバイスでのコンテンツの再生時における、ユーザ間の高周波加速度センシングの「違いの特性」(PoD; Property of Difference) として取得できる簡易なアルゴリズムを提案する。このアルゴリズムを以下、「L-PoD」と呼ぶ。

まず、人間の運動には随意・不随意があるが、笑いにおける不随意運動に注目する。また、人間は完全に静止した状態で姿勢を取り続ける事は不可能であるため、加速度値取得時にローパスフィルタとハイパスフィルタによるフィルタリングをおこなうことで、微小不随意運動を検出することができる (図 1)。

(1) ローパスフィルタによるフィルタリング

```

{
  gx = acceleration.x * 0.1 + gx * (1.0 - 0.1)
  gy = acceleration.y * 0.1 + gy * (1.0 - 0.1)
  gz = acceleration.z * 0.1 + gz * (1.0 - 0.1)
}

```

(2) ハイパスフィルタによるフィルタリング

```

{
  ax = acceleration.x - gx
  ay = acceleration.y - gy
  az = acceleration.z - gz
}

```

図 1: 使用した 2 つのフィルタ

ローパスフィルタによるフィルタリングでは、低い値のフィルタ係数 (0.1) を使用し、新しく得られた加速度 (フィルタリングされていない加速度データ) の 10% と前回取得した値 (直前にフィルタリングした値) の 90% を加算する。

これにより、ユーザの 3 軸 (x, y, z) の重力加速度や緩やかな加速度運動を分離することができる。

ハイパスフィルタによるフィルタリングでは、取得した加速度から重力加速度 (g_x, g_y, g_z) を引くことで、瞬間的な加速度を取得することができ、瞬間的な身体の反応 (微小不随意運動) を取得することができる。

次に重力加速度 (g_x, g_y, g_z) から式 (1) を用いて、重力成分のマグニチュード (以降、 G_{Mag}) として利用し、瞬間的な加速度 (a_x, a_y, a_z) も式 (2) を用いて、微小不随意運動のマグニチュード (以降、 A_{Mag}) として利用する。本報告では、この A_{Mag} を使って、笑いの検出と分類をおこなう。

$$G_{mag} = \sqrt{g_x^2 + g_y^2 + g_z^2} \quad (1)$$

$$A_{mag} = \sqrt{a_x^2 + a_y^2 + a_z^2} \quad (2)$$

最後に、集成的特徴 (PoD) を得るため、各ユーザの A_{Mag} からユーザごとに 1 秒ごとの $A_{Mag(i)}$ の平均値を求める。これを用いて、ユーザ全体の平均加速度 $\bar{\mu}$ を求め、分散および標準偏差を式 (3) から算出し、表 1 の分類表により、ユーザ全体の動画の時系列 (1 秒) ごとの笑いの度合いを 4 つのステータスに分類する。

$$\sigma = \sqrt{\frac{(A_{Mag(n)} - \bar{\mu})^2 + \dots + (A_{Mag(n)} - \bar{\mu})^2}{n - 1}} \quad (3)$$

$n = \text{users}$

表 1: 笑いの分類アルゴリズム

	平均加速度 > 0.02[G]	
	高い (動的)	低い (静的)
標準偏差 > 0.02[G]	(P) 人によって笑いが起こる	(N) 該当なし
	(A) 全員笑う	(S) 安定

表 1 の各ステータスについて説明する。S (Stable) は、すべてのユーザにおいて、平均値が低く安定しており、重力以外の加速度が検出できていない状態で誰も笑っていないことを表す。P (Particular) は、人によっては笑っている状態で、加速度センサに大きな値が記録されている。A (All) は、すべてのユーザにおいて、一様に加速度センサに大きな値が入っている。N (Noise) は、平均値が低く、個々のユーザによって異なる行動をしていることを表している。

3.2 スマートフォンアプリケーションの試作

3.1 節で提案した L-PoD アルゴリズムによる笑いの評価の実験をおこなうために、映像再生と同時に加速度値を取得する iOS および Android デバイス向けの実験用スマートフォンアプリケーションの開発をおこなった。iOS, Android

版ともに基本的な機能は同じであるが、本稿では、試作した iOS アプリケーションの基本機能と分析のために使用したサーバー側のプログラムの仕組みを以下に述べる (図 2)。

実験に使用した iOS デバイスは、2010 年に販売された 3.5 インチディスプレイの iPod Touch であり、システムのバージョンは iOS6.1.3 である。試作した iOS アプリケーション上で映像再生を開始すると、再生と同時に現在時刻、3 軸の加速度値、端末の向き、現在の動画再生時刻の取得を 0.1 秒間隔でおこない、取得した加速度から重力加速度 (g_x, g_y, g_z) と瞬間的な加速度 (a_x, a_y, a_z)、 G_{Mag} 、 A_{Mag} を算出し、全部で 14 つのパラメータをデバイス内に CSV ファイルとして保存する。保存された CSV ファイルは、動画視聴終了後にサーバーの役割をおこなう Mac OSX ver.10.8.4 上で動作している Ruby 言語 (ruby 2.0.0dev) と Web アプリケーションフレームワーク Sinatra (ver. 1.4.3) で開発したサーバー側のプログラムに HTTP 通信経由で自動でファイル送信がおこなわれる。

サーバー側のプログラムは、受信した CSV ファイルをサーバー内に保存し、統計解析ソフト GNU R に CSV ファイルのデータを渡すことで、 G_{Mag} 、 A_{Mag} をはじめとする各種グラフを自動生成する仕組みとなっている。現段階では、実験後の分析に必要な全てのグラフを自動生成することができないが、将来的には実時間処理が可能な演算である。

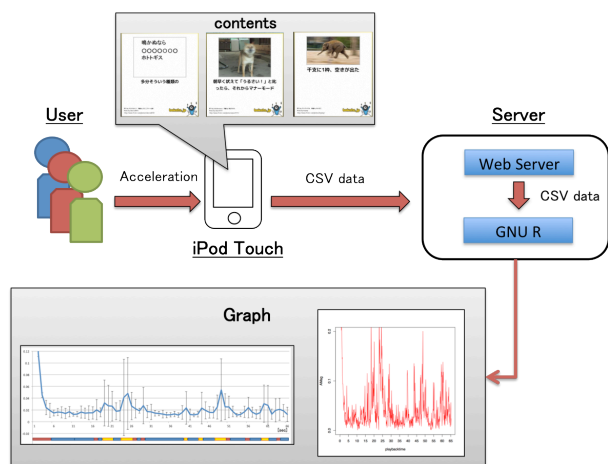


図 2: システム構成

4. 実験

4.1 実験方法

3.1 節で提案したアルゴリズムを実装した iOS 向けアプリケーションを使って、 A_{Mag} による笑いの検出と提案したアルゴリズムによる動画コンテンツの笑いの評価の実験をおこなった。

被験者は、日常的にスマートフォンを使っている大学生男性 8 名である。4 名ずつのグループ A と B に分け、椅子に座った状態 (座位) で、デバイス (iPod Touch) の持ち方を含め、自然な形で自由にデバイスを持ち、グループご

とに割り当てられた画面上の再生ボタンを押すことで動画コンテンツの視聴を開始し、映像終了まで視聴する (図 3)。



図 3: 実験風景：自由な持ち方・座位で笑い誘発刺激をするコンテンツを視聴する

実験で使用する笑い誘発刺激をおこなう動画コンテンツは、写真に一言 (ボケ) を書き込んで投稿・共有する Web サービス bokete (<http://bokete.jp/>) で公開されている「3 秒で笑える殿堂入りボケ」の画像から 10 枚選び、冒頭に実験を説明するための画像を追加した説明文の画像 1 枚 (表示時間 10 秒) と bokete 画像 10 枚、ダミーとして記号のみ表示される画像 1 枚 (表示時間 5 秒) の静止画で構成された約 65 秒の MPEG 動画を使用する。コントロール条件として、グループによって画像が表示される順番を入れ替えた 2 種類の動画 A, B をグループ別に分けて視聴する。また、実際に笑っているかどうか表情を記録するために、テーブルを挟んで被験者の 1m ほど前にビデオカメラをセットし、録画をおこなった。

4.2 A_{Mag} による笑い検出

実験終了後にビデオカメラで取得した映像を確認し、両グループの被験者が笑ったかどうかを検証作業をおこなった。その結果、被験者グループ A は実験中に笑った被験者が約 2 名であり、グループ B は 3 名の被験者であった。図 4 は両グループの中で、一番笑った被験者の A_{Mag} のグラフである。

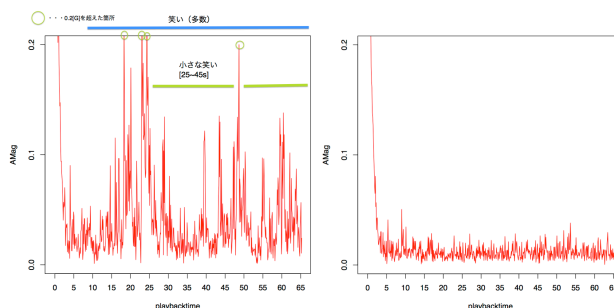


図 4: 笑いの確認できた被験者の A_{Mag} [G] (左) と笑いを確認できなかった被験者の A_{Mag} [G] (右)

この被験者は、ビデオ記録では声を出して笑っているのを確認しており、グラフ上でも 15～25 秒と 45～50 秒の間に 0.2[G] を超える A_{Mag} を記録している。

ユーザー単独の A_{Mag} のグラフを見た場合、図 4 左のように笑いやすい被験者は 0.2[G] を超えている。しかし、0.1[G] 以下の A_{Mag} の値でも、映像から笑っていることが判断できる被験者も存在する。閾値の設定によって、幅広い笑いを対象とするか、より確実な笑いを対象にするかをコントロールすることができる。

4.3 評価アルゴリズムによる分類

より多くの笑った被験者が確認できた「グループ B」の集会的笑い評価をおこなった (図 5)。

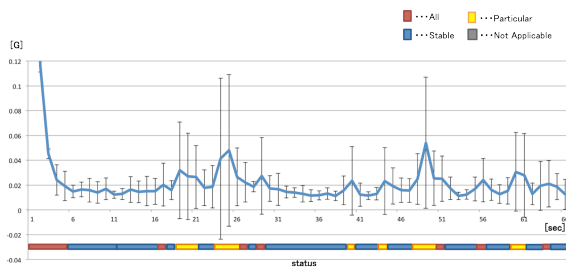


図 5: グループ B の笑いの評価

おおまかではあるが、Stable と Particular の判定には成功している。All に関しては、フィルタの特性上、冒頭の画面操作時の動作の部分に課題が残るが、75% (4 人中 3 人) の反応で、All と判定されている。

5. 課題と今後の展望

本実験により、今後より一般化しつつあるスマートフォンデバイスでの動画コンテンツ視聴において、デバイス非装着でユーザーにとって自然な視聴環境で自然な笑いを検出できた意義は大きい。特にデバイスの持ち方を指定せず、加速度センサにとって不利な座位での検出に成功した点は有用である。

本手法を発展させ、インターネット動画共有サイトや再生アプリ側に実現することで、動画視聴時の「笑いと楽しさ」という主観的な現象を、自然な動画視聴体験と統合した簡易な方法で物理評価できるようになり、VR エンタテインメントシステムでの高精度なコンテンツの評価にも寄与可能になると考える。具体的には、笑いの傾向を分類し、類似度によるおすすめの動画の紹介、ネットワークを活用したテレグジスタンス、VR コミュニケーションの評価手法としても利用可能性は大きいと考える。

課題として、このシンプルなアルゴリズムをリアルタイム化することが挙げられる。例えば、動画コンテンツ視聴中に中間データをサーバー側に送信することで、PoD の算出を現在の動画コンテンツ視聴後ではなく、毎秒ごとに算出することができ、分析速度を向上させることができる可

能性がある。

2 つ目により多くのストーリー性のある動画コンテンツを使った実験を多くの被験者で実施し、L-PoD アルゴリズムの有効性の検証をするべきであろう。より正確で自然な形で評価可能にするためのアルゴリズムの改良につなげたい。

参考文献

- [1] 福嶋政期, 橋本悠希, 野澤考司, 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, Vol.12, No.3, pp.199-207, 2010.
- [2] 倉野大二郎, 松村耕平, 角康之: マルチモーダルデータを用いた映像閲覧者の興味推定, 情報処理学会 インタラクション 2013, 2013.
- [3] 加藤匠, 山下泰介, 田中健司, 早川貴泰, 白井暁彦: 加速度センサを用いたエンタテインメントシステムの非言語評価手法の提案, エンタテインメントコンピューティング 2011, 2011.