

## コメント付きユーザ配信型ライブ放送の 「面白い部分」を抜き出したアーカイブ化

大泉 拓†

相田 仁†

† 東京大学工学系研究科電気系工学専攻

### 1 はじめに

ユーザ配信型のライブ放送はフルサイズでアーカイブされても視聴者への時間的な負担が大きい。番組の面白い部分、話の流れの中で必要と思われる部分のみを抜き出したようなアーカイブを作成することでこれを解決できると考えられる。本研究では動画に対して単位時間における「面白さの値」というものがあると仮定し、これを用いて内容を圧縮したアーカイブを作成する事を提案する。「面白さの値」を計算して動画の編集領域を決定する方法を考案し、「面白さの値」の妥当性について検討した。

### 2 研究の背景

大量に投稿される動画の中から面白い動画を短時間で視聴したいという需要を受けて、コメントを用いたダイジェスト動画作成の研究がなされている [1][2]。一方、ここ数年でユーザ配信型のライブ放送サービスが普及し始めた。ライブ放送は編集された投稿型動画とは違い冗長性が高く、フルアーカイブの視聴は時間的な負担がとても大きい。そこで内容を圧縮したアーカイブを作ることによって、視聴者が見逃した番組を追ったり、新しい番組を探すに当たってその番組の趣向を掴むことが容易になると考えられる。またライブ放送と投稿型の動画との違いは放送主と視聴者のインタラクションがあることである。こういった状況を踏まえ動画のコメント、音声、映像、双方向性を考慮したアーカイブ手法を提案する。

### 3 提案手法

#### 3.1 面白さの値

動画のどの部分を編集するかを決める基準のようなものがあると便利であると考え、各時間における面白さの値  $P$  を定義した。 $P_v, P_s, P_c$  はそれぞれ映像、音声、コメントによる面白さの値である。編集する領域はこ

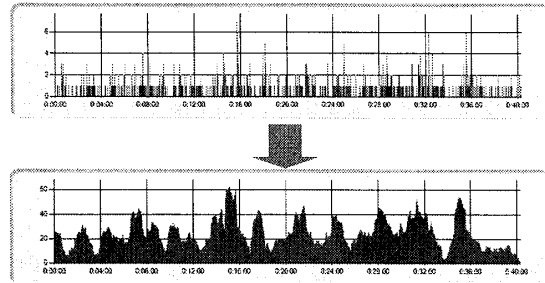


図 1: コメント数の畳み込み

の値に基づいて決定される。

$$P = P_v + P_s + P_c \quad (1)$$

#### 3.2 コメントによる面白さの値

図 1 の上図はある動画の毎秒におけるコメント量 (全 1000 件) をグラフにしたものである。生放送におけるコメントは視聴していた人間のみに限られるため投稿型の動画に比べるとコメント資源に乏しい。そこでコメントによる面白さの値については窓を用いた畳み込み計算によって計算する。ここで  $x$  を放送時間内における時間的位置  $t$ 、 $N_c$  を各時間におけるコメント数とし以下の式で定義する。

$$P_c(x) = \sum_{t=\tau}^{t=\tau+W} G(t) N_c(x+t) \quad (2)$$

$G$  は畳み込みの窓関数であり、矩形窓、ガウス窓などが考えられる。そこでコメントなどのある事象に対する反応の時間推移に対してどのような窓を用いるのが妥当であるのか検討する必要がある。

#### 3.3 コメントの意味解釈

本研究ではライブ放送を対象としているため単純に面白い部分を抽出するだけでなく、双方向性の強い対話部分を抽出することも目的としている。そこでまずコメント内の“w”に注目する。“w”は日本語圏内のインターネットにおけるスラングであり、笑っている状態を表現したものである。図 2 はこの“w”(または全角の

Extraction of “Enjoyable Parts” from Commented User-delivered Live Streaming for Archiving.

Taku OIZUMI† Hitoshi AIDA†

†Graduate School of Engineering, The University of Tokyo

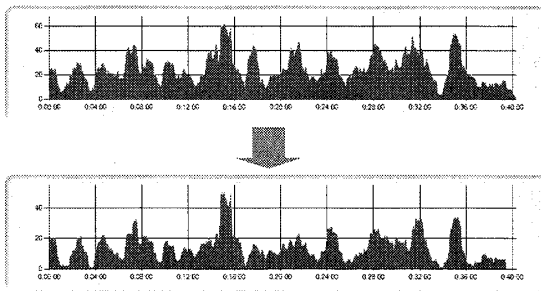


図 2: “w”のコメント数の畳み込み

“w”) の数を同じ窓幅と窓の開始位置で畳み込みしたものである。このように特定の単語の抽出を行うことによって単純なコメント数では得られない、特定の内容を含んだ結果を抽出することが可能である。

### 3.4 映像、音声による面白さの値の計算

映像、音声から動画のダイジェスト部分を抜き出す手法は様々に研究されている [3]。本研究では音声からは音量を用いて面白さの値を計算する。またそれだけではなく音声をフーリエ変換することによって会話部分の抽出を行いクリッピングの位置計算にも利用する。

映像から意味を解釈し面白さの値を計算することは非常に困難である。そこで本研究では場面の変化度 [3] を面白さの値の計算とクリッピングに利用する。

## 4 実装

上記で定義された面白さの値を計算し自動編集するアプリケーション“NicoCommentAnalyzer”(NCA) 図 3 を作成し実験を行った。NCA は主に三つの部分に分けられ、まずはコメントを表示する部分、そして動画を表示する部分、最後に面白さの値をグラフ化しその上に編集領域を表示する編集画面である。NCA では圧縮時間を入力し許容範囲内の圧縮時間の編集結果を返す。任意の編集も可能であり不適切だと思ふ箇所は視聴しながら編集領域を変更することが可能である。

実験ではあらかじめ人間の判断によって編集された領域と NCA によって編集された領域を比較した。人間が編集する際は元の動画を視聴する必要があるため、NCA による編集領域計算においては編集時間が大幅に短縮された。しかし、コメント数が全体的にまばらに存在する動画においては閾値を用いる方法ではうまくいかない場合があり、面白さの値の山の部分を抜き出す手法の必要性を確認した。また NCA ではクリッピングの位置が不適切であることがしばしば発見された。

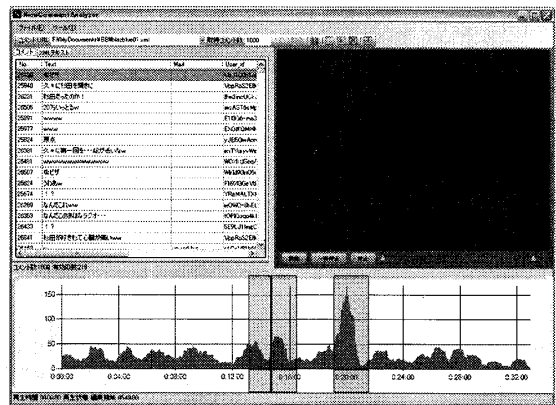


図 3: 製作したアプリケーション

## 5 考察

コメントによる面白さの値の定義については解釈しやすいテキスト形式から計算されるが、音声、映像に関しては未だ意味の解釈が十分ではない。これは画像認識、音声認識分野における自動翻訳技術の発達によって解決できる可能性があるが、その場合はコメントの意味と動画の意味を対応づける必要があると考えられる。

また本研究では動画の単位時間における面白さの値を定義したが、実験を進めるに従いこの面白さの値が実際の動画の面白さを示すには不十分であると感じた。これは面白さの値が畳み込みで計算されていることと、単純なスカラー値で表現していることに起因していると考えられる。この面白さの値をより厳密にするためには擬似人格のモデルが必要ではないかと考えた。観測した事象に対する人間の反応はそれまでの経験の積み重ねによって変化するものである。経験を保持できる人格のモデルによって編集後の動画の検出も容易になるのではないかと考えられる。

## 参考文献

- [1] 青木秀憲, 宮下芳明. “ニコニコ動画における映像要約とサビ検出の試み”, 情処研報 2008-HCI-128/2008-MUS-75, Vol.2008, No.50, pp.37-42, 2008.
- [2] 磯貝佳輝, 斎藤義仰, 村山優子. “視聴者からのコメント情報を用いたダイジェスト動画作成支援システムの実装と評価”, DICOMO2009, 2009.
- [3] 鹿嶋 雅之, 押川修士, 鶴田伸一, 佐藤公則, 渡邊睦. “映像と音声特徴に基づいた映画のダイジェスト生成”, 平成 21 年度 (第 62 回) 電気関係学会九州支部連合大会, 2009.