

令和6年3月31日

令和5年度 学生自主研究成果報告書

教 育 本 部 長 様

学生自主研究グループ名	切り抜き隊	
研究課題名	映像中の音声から切り抜く箇所を見つける研究	
研究代表者 (学生)	学籍番号	B25N046
	氏 名	根平 了輔
指導教員	学 科	情報工学科
	氏 名	安倍 幸治

学生自主研究の報告書を別紙のとおり提出します。

映像中の音声から切り抜く箇所を見つける研究

システム科学技術学部 知能メカトロニクス学科

1年 根平 了輔

1年 坂本 侑也

1年 三澤 直生

1年 小松 頼人

1年 佐々木 陽向

指導教員 システム科学技術学部 情報工学科

助教 安倍 幸治

1. 目的

本研究はYouTubeやTikTokなど、多くの動画配信サービスで投稿されている切り抜き形式の動画を手軽に作成できるように、適切な切り抜き箇所を選定できるシステムを作成・評価することを目的とする。今回は、MATLABを使用して動画の音量に着目して切り抜きポイントを解析的に抽出することを試みる。

2. 研究準備

今回、研究に参加したメンバーはMATLABに触れるのが初めてなので、MATLABの習得から行った。解析ソフトを作った経験が無かったため、いろいろと調べても独力で解析ソフトのプロトタイプを作ることはできなかったが、そのソフトの改良を行うことができる能力は習得することができた。

3. 切り抜きソフトの作成

I. 解析ソフトの作成・改良

安倍先生に協力いただき、wavファイルの解析を行う基本となるソフトを作成した。具体的には、はじめに読み込んだwavファイルの音量の値を2乗してパワーに準ずる形に変換する。その後、計算負荷を軽減するために、一定の時間ごとにその値の間引く作業を行っている。さらに、音声等の音信号は瞬時の増減が大きいため、平均的な音量の変化を観察できるように、間引いた値を設定した一定の時間長で平均を取った。最後に、しきい値を設定し、一定の値以上の音量が出ている区間を切り抜き区間としてピックアップしている。この解析は、動画の面白い部分では、演者の声であったり、効果音であったりが、一定以上の音量になっているのではないかという我々の仮説に基づいている。このソフトはこのままでも十分研究に使用できるような完成度の高いものだった。

しかし、解析プログラムの初期型は数値計算に多く時間がかかってしまうため計算を効率化し、計算にかかる時間を少し短縮した(1時間の動画の計算時間が10分から5分程度に減少)。

さらに、今回の研究では実験段階で同じ動画に対して複数回切り抜きを行うので、実験段階で切り抜き区間が分かりやすくなる様に区間の表示形式を「秒」から「時:分:秒」に変更するようなプログラムを追加した。

II. システムの評価実験

本研究メンバーである根平が独断で選んだYouTubeに投稿されている下記の3つの動画について、作成した解析ソフトを使用して実験を行った。下記の動画は本当に音を分析することで面白い部分を切り抜くことができるのかを試すために、音量と面白さが対応していなようなものから対応していそうなものまで①～③まで選んだ。選出した動画の概要を以下に示す。

動画①

動画タイトル：そ・・・？

投稿チャンネル：Siro Channel

動画概要：視聴者から送られたプロフィール帳を読みながら一人で雑談する

URL：https://www.youtube.com/live/3szkhHT0m_k?si=9JbLndwLCr1pSTVN

選出理由：音量と面白さはあまり対応していなそうだと予想



動画②

動画タイトル：🎃Halloween Party 2021🎃 ～ブラックちえりからの招待状～

投稿チャンネル：どっとライブ

動画概要：司会進行役1人と演者4人でオリジナルのすごろくを楽しむ

URL：<https://www.youtube.com/live/3uX7nkmCoyo?si=FNQSmivikL6ir8m0>

選出理由：動画全体で複数人がワイワイしているため、それなりに切り抜き箇所が選べると予想



動画③

動画タイトル：【夜間警備】ホラゲにビビったら先輩からお尻にお仕置き👊👹!!スパンキングホラー配信👁👁【十六夜ちはや/ぶいばい】

投稿チャンネル：十六夜ちはや / Izayoi Chihaya

動画概要：演者2人でお尻などを叩きながらホラーゲームを行う

URL : <https://www.youtube.com/live/ZhTopVfWqrk?si=0BsomHwABsNkXUI7>

選出理由：ホラーゲームは怖がって叫んでいる箇所が見どころになるため、今回のソフトでいい箇所が切り抜けると予想

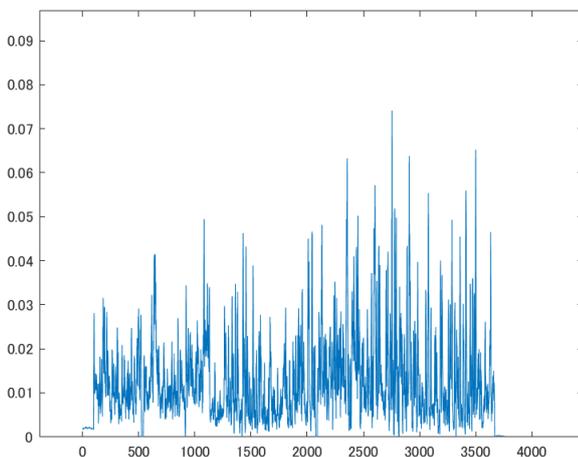


切り抜き動画を作成するための手順を以下にします。基本的なルールは下記の2項目となる。

- ・解析プログラムを実行した際に0.1秒ごとの音量データを取り除き、処理するデータ量を減らす
- ・音データは5秒間の平均を取る

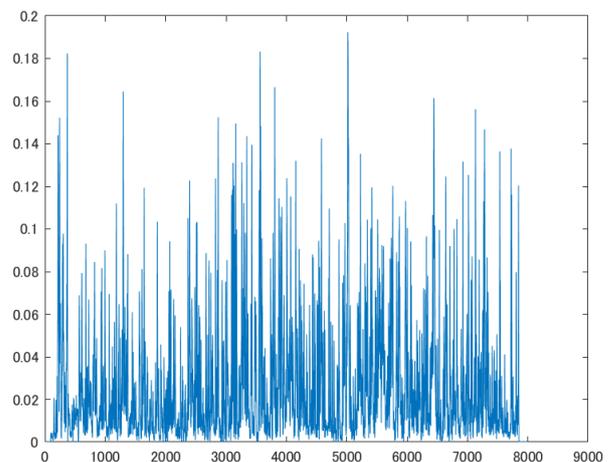
その後、しきい値だけを変更して切り抜き区間がおよそ10区間程度になるよう調整した。

最後にその切り抜きポイントに対して、前後20秒程度を参考にしながら動画視聴することで、そのポイントが実際に切り抜き動画に適するのかを○(切り抜きに適する)・△(編集方法によっては切り抜ける)・×(切り抜きに適さない)の3段階で評価した。実験結果は以下のようになった。



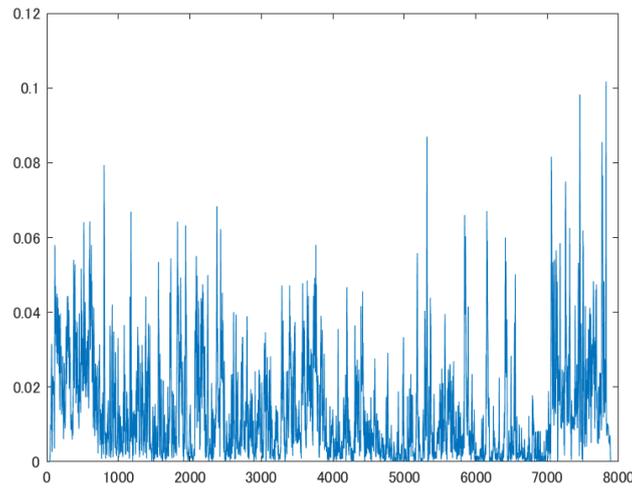
動画①

しきい値 0.05
切り抜き箇所 10



動画②

しきい値 0.15
切り抜き箇所 9



動画③

しきい値 0.065

切り抜き箇所 11

上の実験で選んだ切り抜き箇所について、評価は下のようになった。

表1. 実験結果の評価

	○	△	×
動画①	0	3	7
動画②	2	3	4
動画③	5	4	2

4. 実験結果の考察

・動画①について

現在のシステムでは音量による盛り上がりしか検知できないため、ピックアップされた箇所は不思議な箇所ばかりで、切り抜きに適した箇所を検知できているとは言えなかった。

・動画②について

動画上で盛り上がっている箇所はある程度検知できていたが、この動画ではそれまでの流れが影響した盛り上がりの箇所が多かった為、×評価が多く出た。

・動画③について

動画上で盛り上がっている箇所を多く検知できていた。口上の挨拶やネタバレになる箇所には△や×の評価を付けたが、ほとんど切り抜きに適した箇所を切り抜けていたと言える。

5. 考察

雑談動画が顕著だったが、動画上の盛り上がりにはそれまでの流れが必要な場合が多く、今のシステムではそれは考慮できない。そのため、現在のシステムでは切り抜きに適さない箇所が多く選ばれることがある。しかし、ホラーなど動画内での盛り上がりが音量に反映されやすいものならばなかなかの精度で検出することができる。

音量意外のアプローチを行ったり、音量のデータの取り方をもう少し工夫したりすることで万人に面白いと思ってもらえるような箇所を切り抜けるようになると考える。

6. 結論

現在のシステムはどんな動画に対しても使用できるようなものではない。しかし、相性のいいジャンルの動画に対してはある程度精度の高い切り抜きポイントを出すことができた。今後さらなる改良を行っていき、切り抜きポイントを選ぶ精度を上げていきたい。