

平成 29 年度 特別研究報告書

料理サイトに投稿する動画の自動編集プログラム
の検討

龍谷大学 理工学部 情報メディア学科

T140487 新田 晃基

指導教員 三好 力 教授

内容梗概

近年 1 分程度の料理動画が人気である。またスマートフォンの普及により、通勤時間や休憩時間など少しの時間でも場所を選ばず料理動画を視聴できるようになった。本研究では、自分が作った料理の動画を作りたいという人のために、簡単に 1 分程度の料理動画を作ることができるアルゴリズムを提案する。自分の作った料理を動画にすることにより、より自分のレシピがわかりやすくなり、さらに SNS などに自分のレシピを気軽に投稿することが期待できる。

目次

第 1 章 緒言	1
1.1 研究背景.....	1
第 2 章 関連研究及び既存技術	3
2.1 クックパッド.....	3
2.1.1 クックパッドとは.....	3
2.1.2 クックパッドの特徴.....	4
2.2 自動ハイライト動画作成アプリ.....	4
2.3 問題点.....	4
2.3.1 クックパッド.....	4
2.3.2 自動ハイライト動画作成アプリ.....	4
第 3 章 提案手法	5
3.1 概要.....	5
3.2 ハイライトシーン.....	5
3.3 アルゴリズム.....	6
第 4 章 実験	8
4.1 実験目的.....	8
4.2 実験条件.....	8
4.3 実験内容.....	8
4.3.1 動画撮影.....	8
4.3.2 動画を画像化.....	8
4.3.3 画像の色の取得.....	8
4.4 麻婆豆腐の料理動画での実験結果.....	9
4.5 ラザニアの料理動画での実験結果.....	13
4.6 実験結果.....	17
4.7 考察.....	18
第 5 章 結論	19

第1章 緒言

1.1 研究背景

近年、スマートフォンの普及に伴い、いつでもどこでも手軽に料理のレシピを見ることができるようになった。図 1.1 から、料理をする時に 86.2%の人が「参考をしている」ことがわかる。さらに、図 1.2 から参考をしている人の 56.7%が「レシピ検索(アプリ)」であり、続いて「レシピアプリ」が 38.4%となっている。つまり半数以上の人々が、スマートフォンで料理のレシピを見ていることがわかる。

● 料理をするときに参考しているもの (n=552)

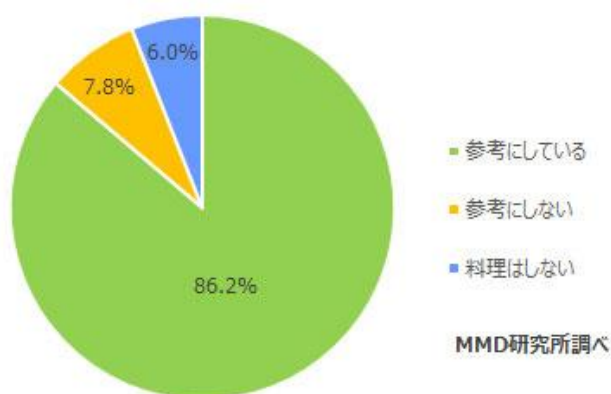


図 1.1 : 料理をするときに参考しているもの[1]

● 料理をするときに参考しているもの (n=476)

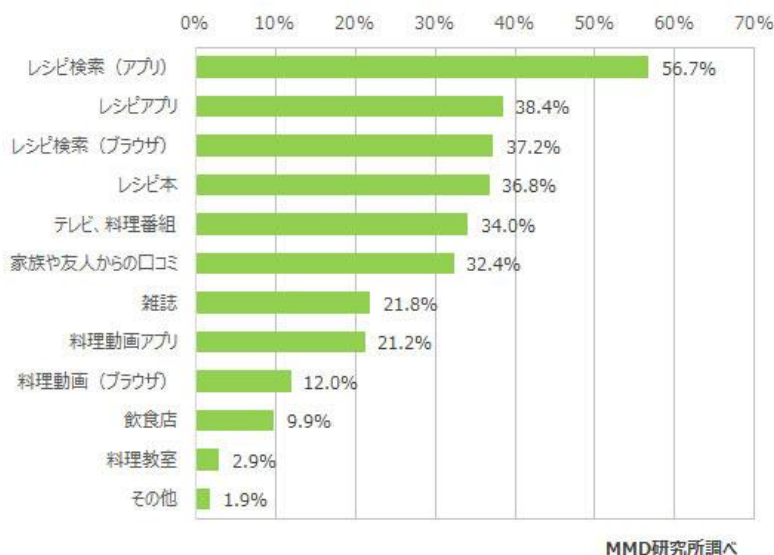


図 1.2 : 料理をするときに参考しているもの

その中でも料理の全ての手順を 1 分程度の動画で説明するレシピ動画を視聴する人が増えている。最近では、検索数が増加しており、図 1.3 より 2016 年の 5 月以降大きく伸びていることがわかる。さらに詳しく、「レシピ」と「動画」の掛け合わせに関する検索と、「レシピ」だけの検索について、検索数の推移を比較すると、2011 年 1 月の検索数を 100 として検索数を比べてみると、2015 年まではいずれも緩やかな増加傾向であるが、2016 年は「レシピ」と「動画」を掛け合わせた検索数が飛躍的に伸びていることが分かる。

レシピ動画が人気の理由は、食材を切ったり炒めたりする様子が動画で見られるので、文章で説明するレシピよりも手順や流れが分かりやすいからである。そのため、自分でも作れるという期待感を抱くことができる。また 1 分程度にまとめられた料理の動画は、SNS などで拡散させやすく、内容も 1 分間の動画の中で完結しているので、見ていだけでも楽しい。このため、普段料理をする人はもちろん、単身赴任で自炊する必要に迫られた男性や、料理をすることに慣れていない若者など、幅広い世代に受け入れられていると考える。また働く人にとっては仕事場への行き帰りの電車内などの少しの時間でも視聴でき、1 分間動画なら、スマートフォンや携帯さえ持っていれば気軽にどこでも視聴が可能である。これらの理由から今後もさらにレシピ動画を視聴する人が増えていくと考える。

本研究では、自分が作った料理の動画を作りたいという人のために、簡単に 1 分程度の料理動画を作ることができるアルゴリズムを提案する。自分の作った料理を動画にすることにより、より自分のレシピがわかりやすくなり、さらに SNS などに自分のレシピを気軽に投稿することが期待できる。

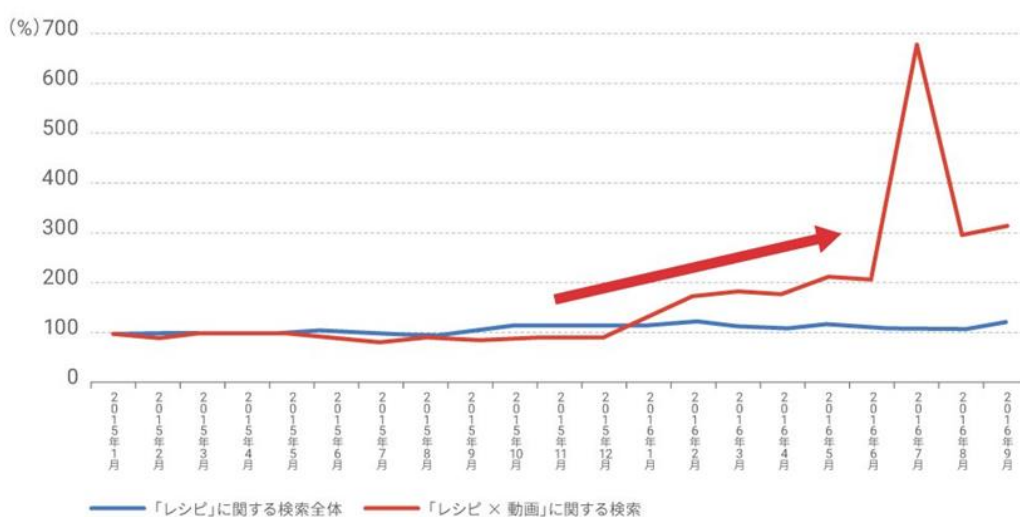


図 1.3 : レシピ × 動画とレシピ全般に関する検索ボリュームの推移 [2]

第2章 関連研究及び既存技術

2.1 クックパッド

2.1.1 クックパッドとは [3]

日本最大の料理レシピのコミュニティウェブサイトであり、278万品を超えるレシピの中から、作り方を検索できる。毎日の食事はもちろん、ハロウィンやクリスマスなどのイベント料理、バレンタインや誕生日ケーキなどのスイーツまで、多数のレシピを見つけることができる。利用者は6000万人の人気の料理サイトである。

図2.1よりレシピ動画アプリの中で1位は「クックパッド」(60.0%)であり、レシピ動画の中で最も人気な料理サイトである。

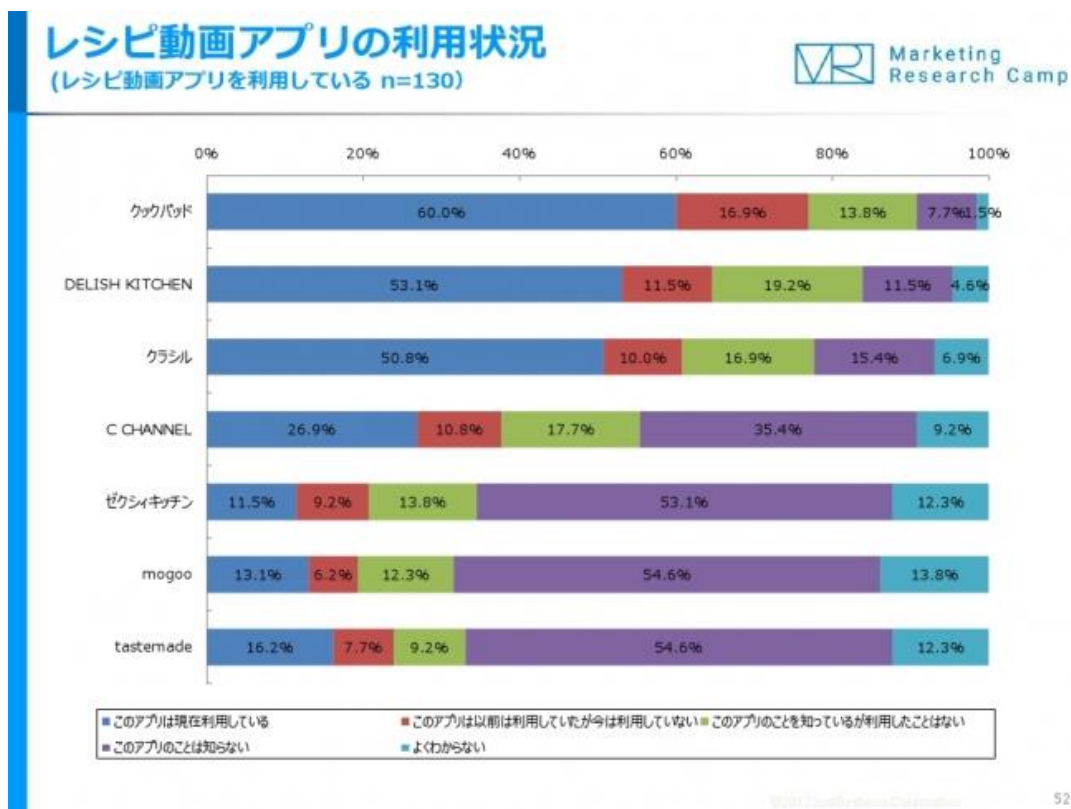


図 2.1 : レシピ動画アプリの利用状況 [4]

2.1.2 クックパッドの特徴

クックパッドの最大の特徴は、自分のレシピを公開することができることであり、家庭の主婦が実際につくった簡単実用レシピが多い。また人気になった料理レシピは、クックパッドのスタッフがその料理レシピを忠実に再現した料理動画を制作してくれるサービスがある。そのほか、自分でも料理動画を投稿することもできる。

2.2 自動ハイライト動画作成アプリ

自動でアクティブなシーンを検出してハイライト動画を生成するアプリなど、動画のハイライトを自動で作れるアプリは多数存在する。例えば、「HaiLyts」というアプリがあるが、これは画像処理技術を利用して、動きの大きなシーンを優先的に選び、逆に動きが少ない箇所はカットするという仕組みである。



図 2.2 : HaiLyts

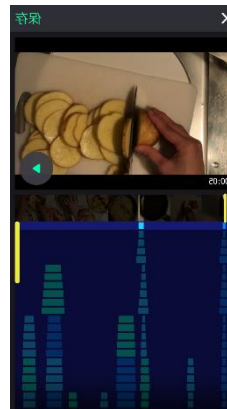


図 2.3: HaiLyts の画面

2.3 問題点

2.3.1 クックパッド

クックパッドのスタッフが料理動画を制作してくれるサービスがあるが、それはごくわずかな人気の料理レシピだけで、すべてのレシピに網羅できていない。またクックパッドにある料理動画は、ほとんどスタッフが制作しており、一般ユーザーが投稿している動画は約 278 万品中 232 品しか投稿されていない。これは自ら撮った料理動画を自分で編集しなければならないことが原因であると考えられる。動画編集というのは大変時間がかかり、普段動画編集をしない人であると敷居が高い。実際、本研究でも動画編集をいくつか行ったところ、動画の時間によるが、30 分の動画なら約 1 時間かかってしまう。

2.3.2 自動ハイライト動画作成アプリ

料理動画はカメラが固定された状態で撮影され、料理動画自体に動きがほとんどないため、無駄な動きもハイライトしてしまう。自動ハイライト動画を作成するアプリは、動きの激しいスポーツに特化しているものが多く、動きの少ない料理動画には適していない。

第3章 提案手法

3.1 概要

第2章で述べた問題点を解決するために、本研究では誰でも簡単に自動で1分程度の料理動画を作ることができるアルゴリズムを提案する。1分程度のハイライト動画を作成するには、実際撮影した1本の動画からハイライトシーン以外のシーンをカットしていく必要がある。そこでハイライトシーンを定義することで、ハイライトシーンを抽出し、1分程度の動画を作成することを考える。

3.2 ハイライトシーン

料理動画において、どのシーンが必要であるか不必要であるかを検討した。

現在の投稿動画を分析すると、ハイライトシーンとなるシーンは「食材を切り始めるシーンと切り終わりのシーン」や、「調味料を入れ始めるシーンと入れ終わりのシーン」などであり、これらのシーンはハイライトシーン以外のシーンに比べて、動きや色の変化が激しいことがわかった。例えば食材のピーマンが出てきたら、緑の成分が増えるのでそこで色が変わる。そしてピーマンを切り終わりピーマンが画面から消えるとき、また色の成分が変わる。この色が変わったシーンを繋ぐと、ピーマンを切り始めるシーンからピーマンを切り終わるシーンとなると考える。

これらのようにハイライトシーンは色の変化が比較的大きく、動きよりも色の方が単純なアルゴリズムで変化を取り出せると考えられるので、本研究では色の変化に着目した。したがって本研究では、ハイライトシーンを色がある一定以上変化するシーンと定義する。

3.3 アルゴリズム

まず撮影した動画から 1 秒置きに静止画を抽出する。次に抽出された順に静止画に番号をつけ、前から 5 秒ずつの間隔で静止画を比較する。比較には色の変化を使うので、まず画像に使われている各ピクセルの RGB を取得する。次に各 RGB を平均化しその平均化された RGB をすべて足す。比較内容は、その平均化された RGB の合計の差分を取り、値を絶対値にすることで色の変化量を調べる。平均化された RGB の合計の差分値が設定した閾値以上となった場合、比較した静止画を色が変わったと判定して、2 秒間使い、そのシーンはハイライトシーンとして用いる。逆に閾値以下になった場合その静止画をカットし、またさらに 5 秒後の静止画を比較する。その後 4 秒カットし、これを繰り返し行い、ハイライト動画を作成していく。最後に動画を 2 倍速にし、さらに動画の時間を 1 分に近づける。

最後に 2 倍速にする理由は、1 倍速で 1 秒間だけの情報であると短すぎるため、2 秒間の情報を得たほうがより、レシピがわかりやすくなるため。また 2 倍速にすることでテンポよく動画を見ることができる。

図 3.1 にハイライト動画抽出のフローチャートを示す。

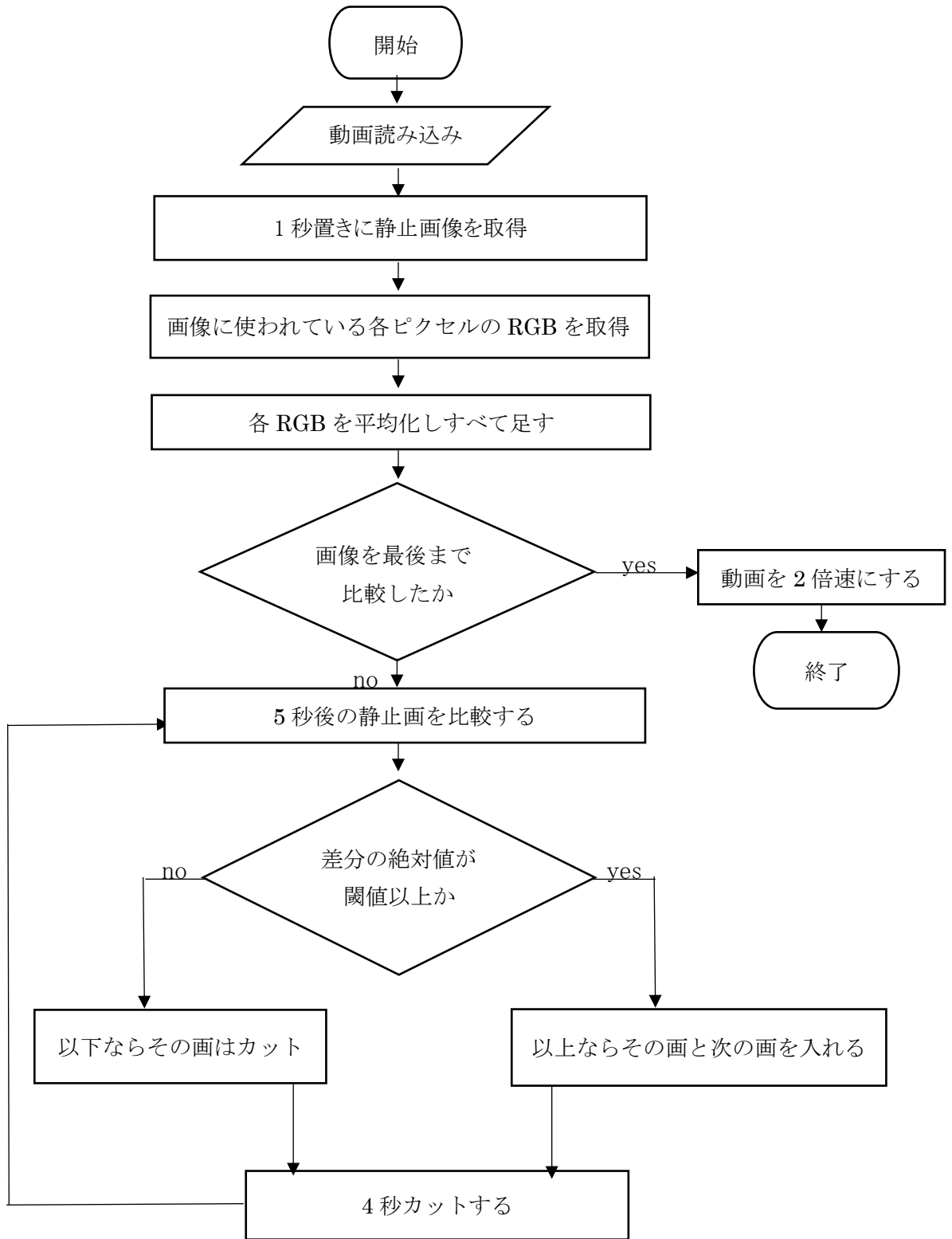


図 3.1 料理動画を作成するアルゴリズム

第4章 実験

4.1 実験目的

本実験は、第3章で述べたアルゴリズムの通りに動画編集を行うと、正確にハイライトシーンを含む1分程度の動画に編集できるのかを確認するために行った。そのために、平均化されたRGBの閾値が適当であるか確かめるために、実際に自分で撮影した「麻婆豆腐」と「ラザニア」の料理動画を用いて実験を行う。

4.2 実験条件

実験条件を以下に示す。

実験場所(撮影場所)	自宅
使用機器	スマートフォン(iPhone)、三脚
実験機器	PC(Windows)
使用ツール	Free Video to JPG Converter 5.0.92.607 [5] 画像色平均計算ツール Version 1.0.1 [6]

4.3 実験内容

4.3.1 動画撮影

本実験では三脚を用いてスマートフォンを固定して設置し、真上の位置になるよう撮影をした。撮影した動画のファイル情報を以下に示す。

拡張子	MOV ファイル (Apple での標準動画形式)
解像度	1280×720
フレーム率	29.97 フレーム/秒
長さ	麻婆豆腐は 42 分、ラザニアは 24 分
ビットレート	約 10000kbps

4.3.2 動画を画像化

今回は、動画内のフレームを、JPEG 画像として抜き出してくれるソフト「Free Video to JPG Converter」というツールを用いて、動画を1秒置きにJPEG形式で抽出した。

4.3.3 画像の色の取得

今回は、画像中のすべての画素の色を調べ、すべてを加算し、画素数で割る作業を自動でしてくれるソフト「画像色平均計算ツール」というツールを用いて、画像に使われている各ピクセルのRGBの平均を求める。

4.4 麻婆豆腐の料理動画での実験結果

平均 RGB の合計の差分がどのくらいあればハイライトシーンとなるのかを実験した。

麻婆豆腐の 42 分中食材を炒めるシーン 1 分 11 秒から 6 分 41 秒間において、画像中の平均 RGB の合計のグラフを以下に示す。

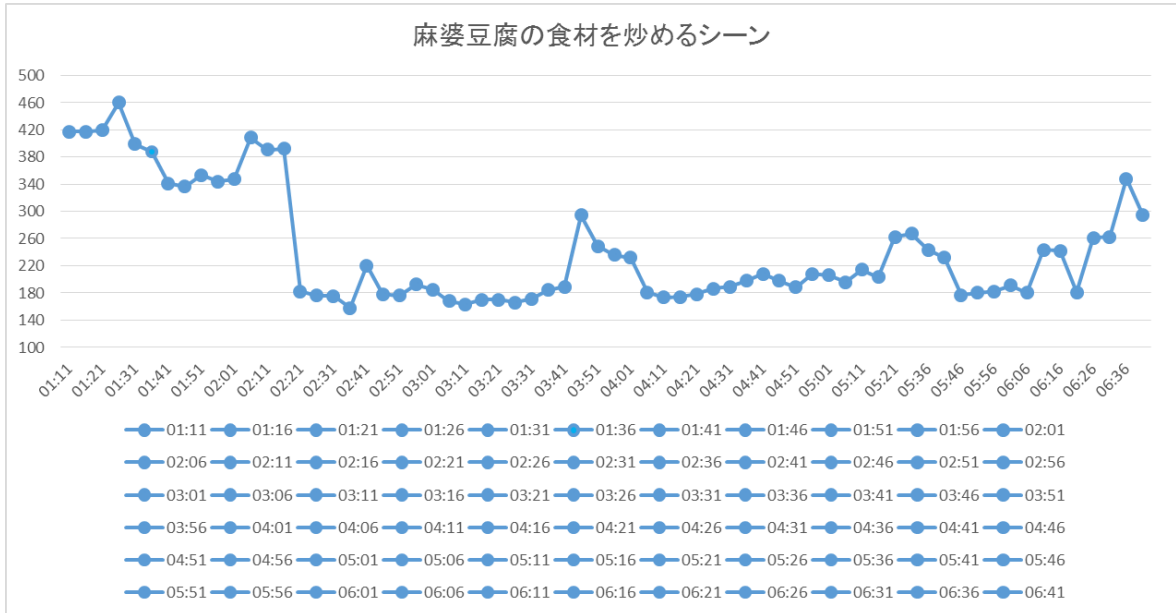


図 4.1 : 麻婆豆腐の食材を炒めるシーン

色の変化が特に大きい、1 分 6 秒の画と 1 分 11 秒の画の画像と平均色を以下の図 4.2 と図 4.3 に示す。図 4.2 から図 4.3 へは、画が大きく変わっている。図 4.4 の色を見ると、オレンジに近い茶色であり、図 4.5 の色は濃い茶色であるので、平均的な色の違いもはっきり違う。こういった色に変化したシーンをハイライトシーンとする。



図 4.2 : 1 分 6 秒の画



図 4.3 : 1 分 11 秒の画



図 4.4 : 図 4.2 の平均色



図 4.5 : 図 4.3 の平均色

次に、差分の絶対値の閾値を 10 と設定した場合のグラフを以下に示す。

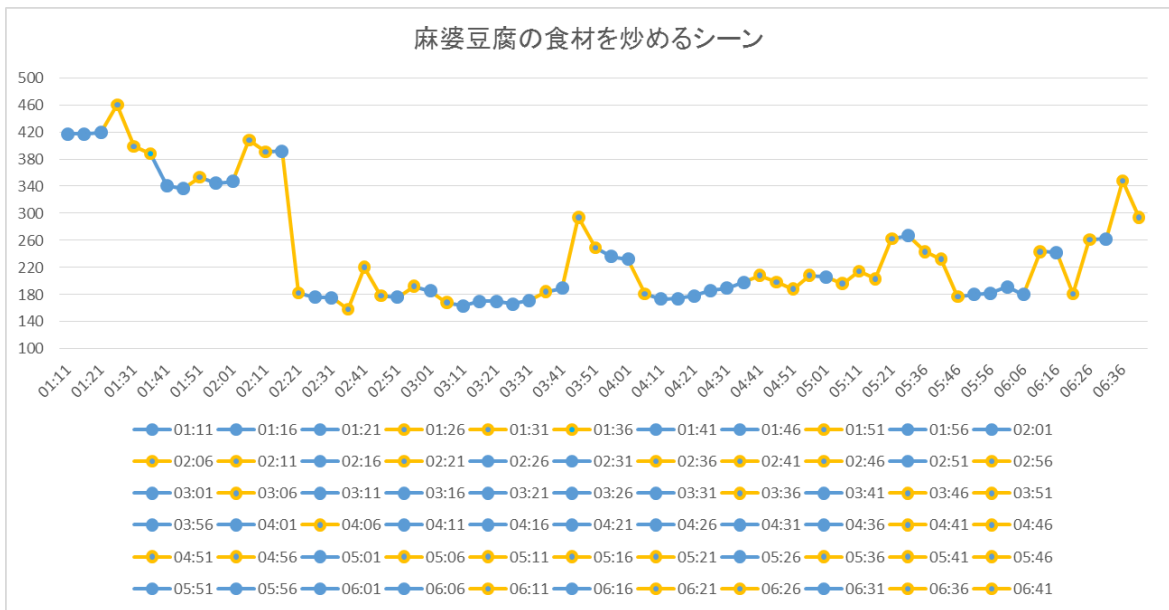


図 4.6 : 差分の絶対値の閾値が 10 の場合

枠線が黄色となっているのが差分の絶対値の閾値を 10 と設定した場合を表している。この場合ハイライトシーンとする箇所が多すぎて不適切と考える。またあまり色の変化のないシーンでもハイライトとしてしまうので、実用性は期待できない。図 4.7～図 4.9 の画像は、それぞれ RGB の差が 10 であるが、実際画像を見てもあまり変化がないのがわかる。このような同じような色が続く箇所がいくつもあるため、同じようなハイライトシーンが続き、その分動画の時間が長くなる。



図 4.7: 3 分 31 秒の画



図 4.8: 3 分 36 秒の画



図 4.9: 3 分 41 秒の画

次に、差分の絶対値の閾値を 20 と設定した場合のグラフを以下に示す。

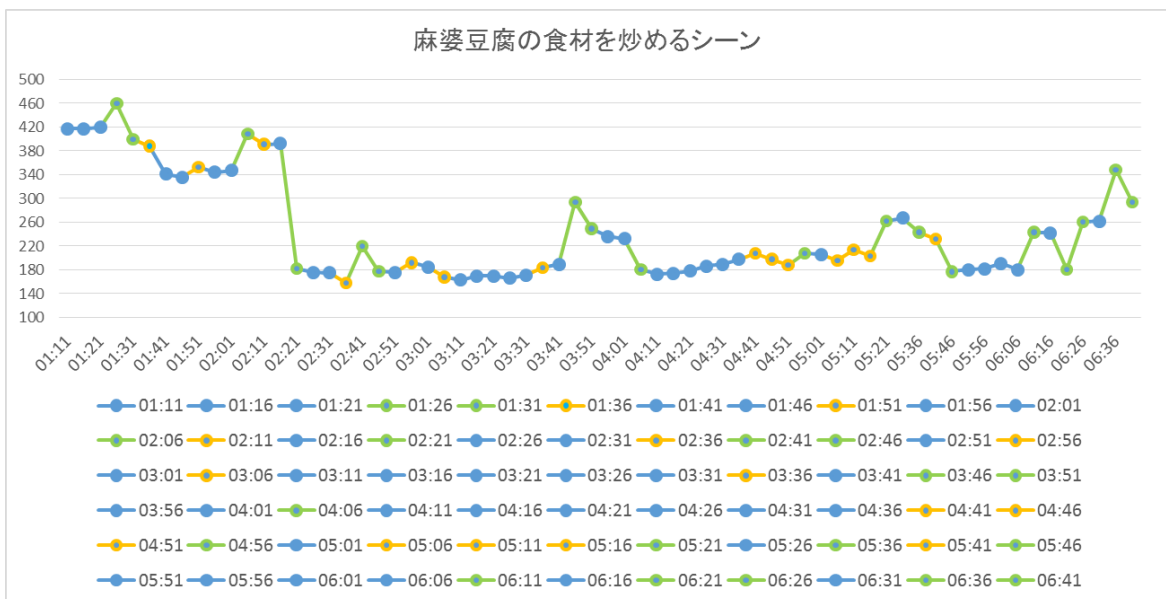


図 4.10: 差分の絶対値の閾値が 20 の場合

枠線が緑色となっているのが差分の絶対値の閾値を 20 と設定した場合を表している。前述で述べた、同じような色が続く箇所は少なくなり、図 4.11～図 4.12 の画像は同じようなシーンであり、RGB の差が 24 であったが、この一か所だけであった。



図 4.11 : 4 分 16 秒の画



図 4.12 : 4 分 21 秒の画

次に、差分の絶対値の閾値を 30 と設定した場合のグラフを以下に示す。

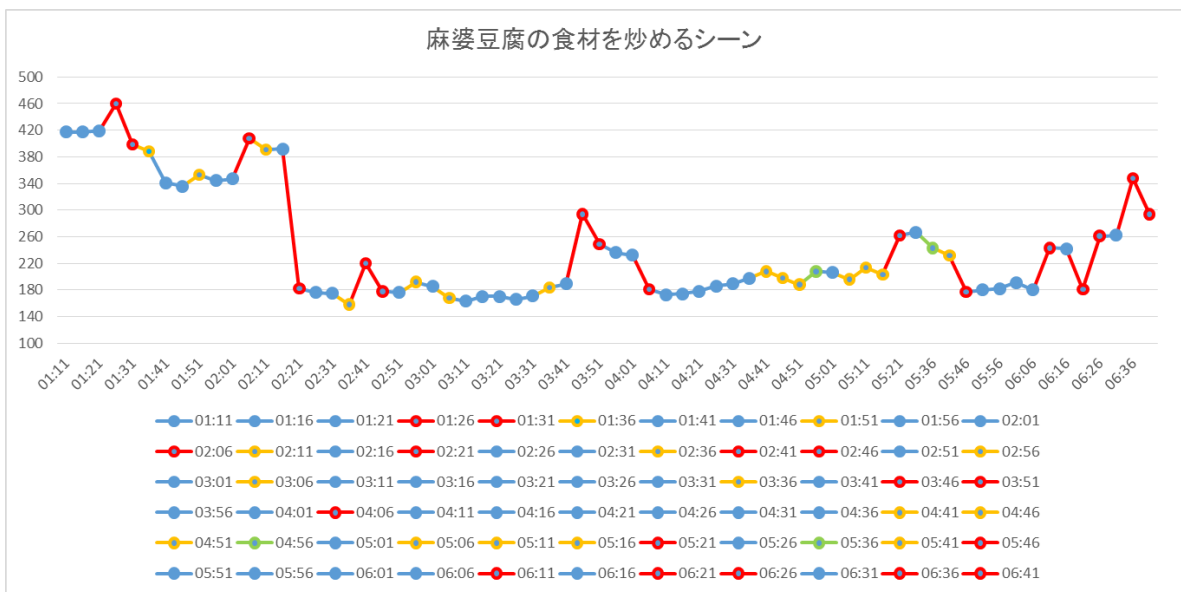


図 4.13 : 差分の絶対値の閾値が 30 の場合

枠線が赤色となっているのが差分の絶対値の閾値を 30 と設定した場合を表している。比較的色の変化が大きい箇所が目立つので、ハイライトシーンとして必要なシーンを判別している。

4.5 ラザニアの料理動画での実験結果

ラザニアの 24 分中食材を炒めるシーン 7 分 56 秒から 12 分 46 秒間において、画像中の平均 RGB の合計のグラフを以下に示す。

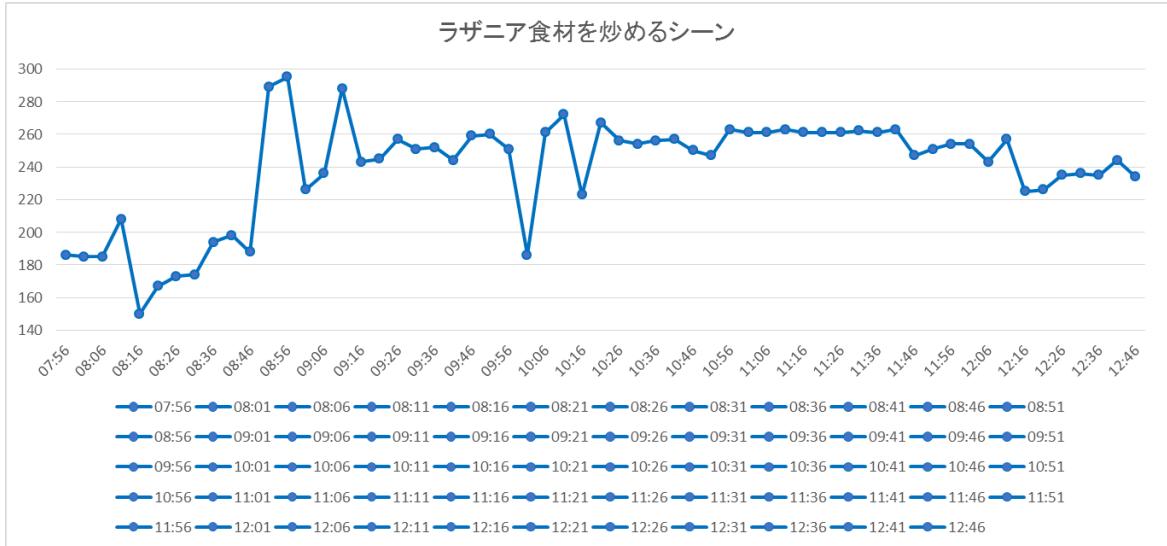


図 4.14 : ラザニアの食材を炒めるシーン

次は実験 4.4 とは逆に、色の変化があまりない箇所を見る。図 4.14 のグラフを見ると 10 分 56 秒から 11 分 46 秒の間は、平均 RGB の合計がほとんど変わらない。実際図 4.15 ~ 図 4.16 の画像を見ても画がほとんど変わらない。図 4.17 ~ 図 4.18 の平均色を見ても多少の色の濃さはるもののあまり違いがない。つまり色の変化がない箇所は 1 分動画として不必要なシーンであることがわかる。



図 4.15 : 10 分 56 秒の画



図 4.16 : 11 分 46 秒の画



図 4.17 : 図 4.15 の平均色



図 4.18 : 図 4.16 の平均色

次に、差分の絶対値の閾値を 10 と設定した場合のグラフを以下に示す。

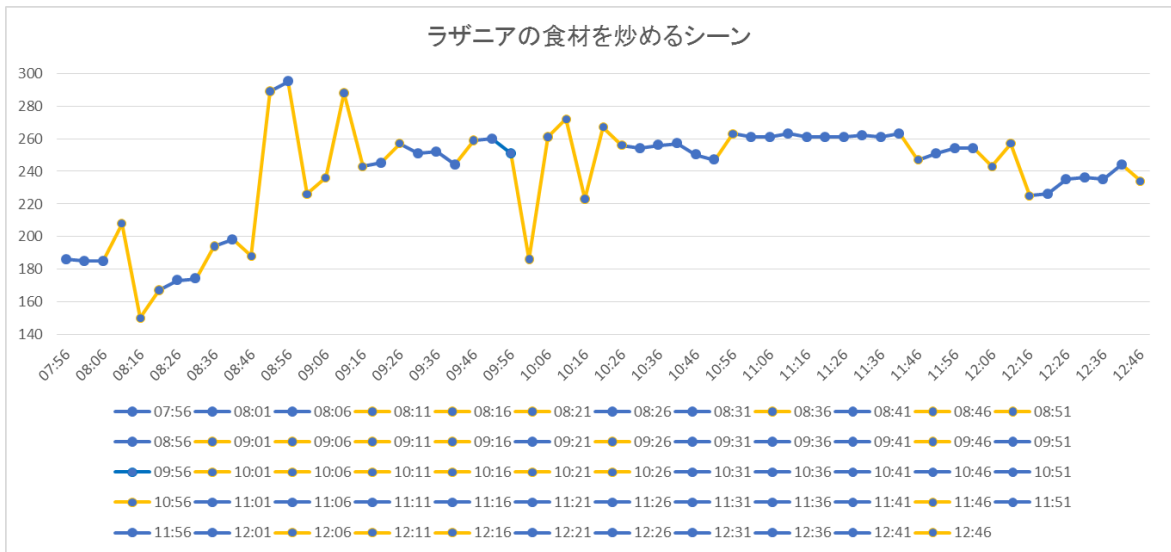


図 4.19 : 差分の絶対値の閾値が 10 の場合

同じく枠線が黄色となっているのが差分の絶対値の閾値を 10 と設定した場合を表している。この場合も図 4.20～図 4.21 の画像を見ると、RGB の差が 11 であるが、実際画像を見てもあまり変化がないのがわかる。この料理動画に関しても図 4.20～図 4.21 のようなハイライトシーンではないシーンでもハイライトしてしまう箇所が多い。



図 4.20 : 6 分 56 秒の画



図 4.21 : 7 分 1 秒の画

次に、差分の絶対値の閾値を 20 と設定した場合のグラフを以下に示す。

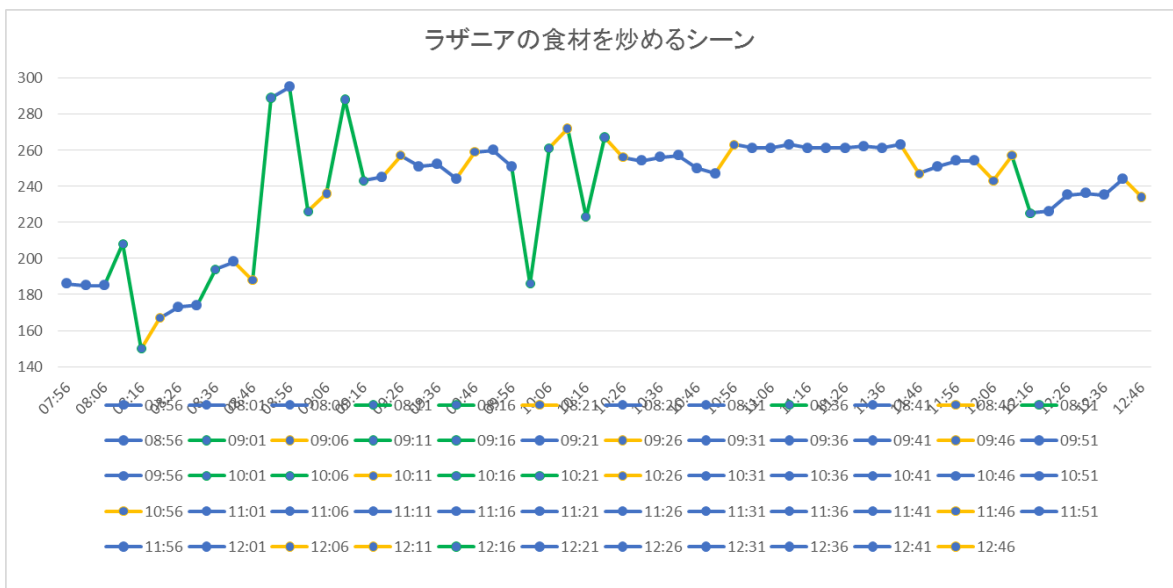


図 4.22 : 差分の絶対値の閾値が 20 の場合

同じく枠線が緑色となっているのが差分の絶対値の閾値を 20 と設定した場合を表している。図 4.23～図 4.24 の場合、RGB の差が 23 であり、油を入れるシーンは、ハイライトシーンとなるので非常に重要なシーンといえる。また図 4.25～図 4.26 の場合も、RGB の差が 21 であるので、油を入れるシーンやにんにくを入れるシーンなど画が大きくは変化していない場合のハイライトシーンでも判別できる。



図 4.23 : 4 分 41 秒の画



図 4.24 : 4 分 51 秒の画



図 4.25 : 5 分 6 秒の画



図 4.26 : 5 分 11 秒の画

次に、差分の絶対値の閾値を 30 と設定した場合のグラフを以下に示す。

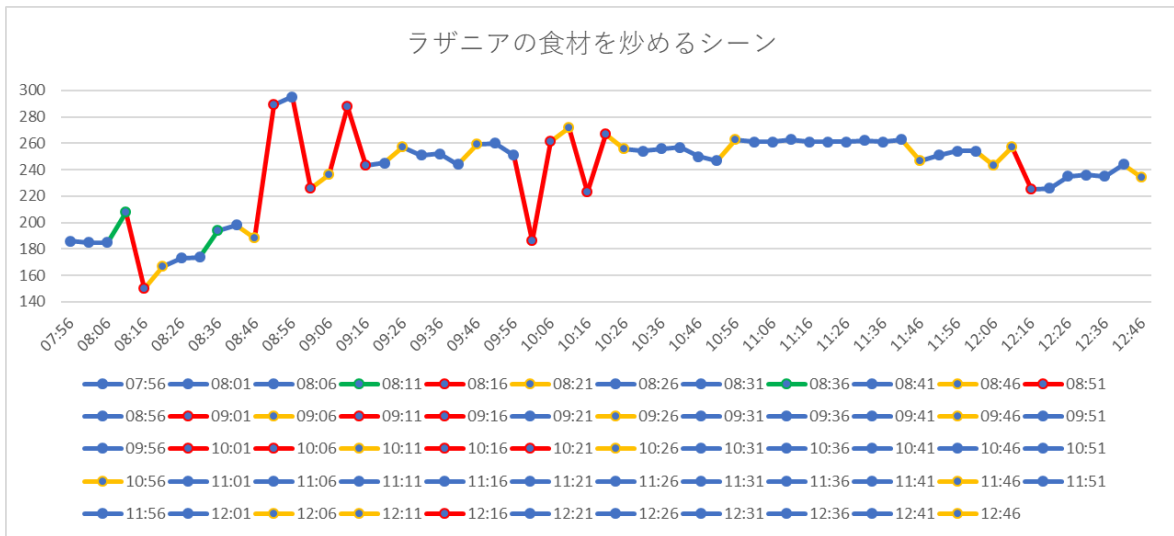


図 4.27：差分の絶対値の閾値が 30 の場合

同じく枠線が赤色となっているのが差分の絶対値の閾値を 20 と設定した場合を表している。閾値を 30 と設定すると、図 4.23～図 4.24 と図 4.25～図 4.26 はハイライトシーンとして重要なシーンであるが、カットされてしまう。

4.6 実験結果

差分の絶対値の閾値を 20 と設定して、実際撮影した動画をアルゴリズム通りに編集を行い、その動画の平均 RGB の合計のグラフを以下の図 4.28、図 4.29 に示す。

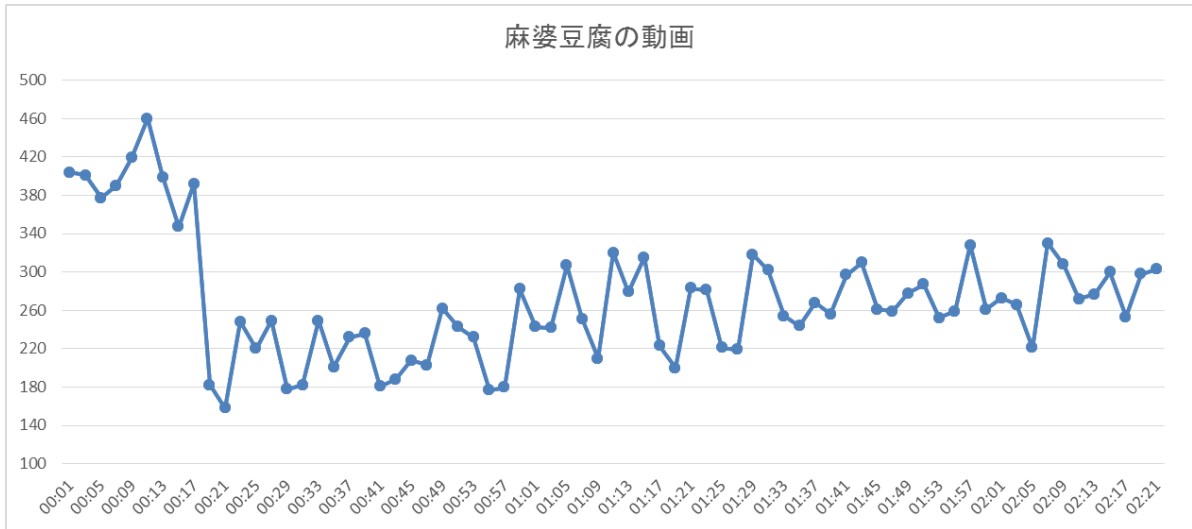


図 4.28 : 動画編集後の麻婆豆腐の動画のグラフ

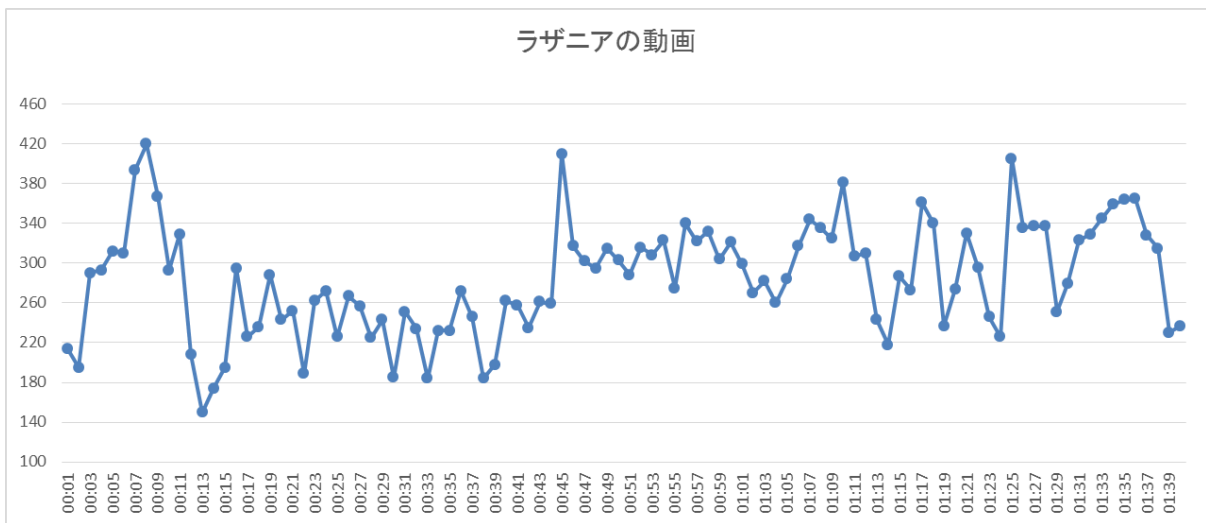


図 4.29 : 動画編集後のラザニアの動画のグラフ

編集前の動画のグラフと比べると、横ばいになっている箇所は編集後のグラフにはあまり目立たない。

4.7 考察

実験結果から RGB の差分の絶対値の閾値を 10 と設定した場合は、色の変化がない箇所でもハイライトとして判定してしまうため、閾値が 10 の設定では不適切だといえる。閾値を 20 と設定した場合は、色が大きくは変化していない箇所でもハイライトシーンとして判定できる。最後に、閾値を 30 と設定した場合は、色の変化が大きい箇所は判別できるが、上記のような箇所では判別できない。よって閾値は 20 が妥当といえる。

実際、麻婆豆腐の動画を閾値 20 と設定してアルゴリズム通りに動画編集を行うと、42 分の動画から 2 分 22 秒になり、さらに 2 倍速にして 1 分 11 秒と約 1 分の動画になった。またラザニアの動画も、同じようにすると 24 分の動画から 1 分 40 秒になり、さらに 2 倍速にしても 50 秒となり、こちらも約 1 分の動画になった。2 倍速にしても何を行っているか判断できるため支障はない。

第5章 結論

本研究では、自動で1分程度の料理動画を作るため、三脚を用いて動画撮影を行い、約1分ハイライト動画を作るためのアルゴリズムを提案した。そのため、色がどのくらい変化したらハイライトとなるのかを研究した。実験結果からハイライトとなるシーンを判別でき、1分程度の料理動画を作成することができた。作成された1分動画に手動で文字や音楽を入れると、クックパッドに投稿されているような動画になる。

本実験では動画編集を手動で行っているため、今後はアプリなどを使ってハイライトシーンの作成や、カットした部分を認識して文字を自動的に入力するなど、自動で動画編集が行えるようにするのが、今後の課題である。

謝辞

本論文を作成するにあたり、ご指導頂きました三好力教授に深謝いたします。
また、多忙の中日頃の議論にご協力くださった同三好研究室の皆様や、学友の皆様に心より感謝いたします。

参考文献

- [1] https://mmdlabo.jp/investigation/detail_1646.html
- [2] http://apac.thinkwithgoogle.com/intl/ja_ALL/articles/videotrend-recipes2016.html
- [3] <https://cookpad.com/>
- [4] <https://prtimes.jp/main/html/rd/p/000000259.000007597.html>
- [5] Free Video to JPG Converter 5.0.92.607
<https://free-video-to-jpg-converter.jp.uptodown.com/windows>
- [6] 画像色平均計算ツール Version 1.0.1
<http://www.vector.co.jp/soft/winnt/art/se511230.html>