

アニメ作品映像要約におけるキーフレーム検出の一手法

高木 麻妃[†] 石川 知一[†] 柿本 正憲[†]東京工科大学 メディア学部[†]

1. はじめに

近年、インターネットによるアニメ作品の動画配信が盛んとなっており、放送中の作品の無料配信や過去の作品の一举放送などが行われている。しかしながら、動画であるアニメ作品の情報消費には多くの時間が必要であり、すべての人が手軽に楽しめるとは限らない。

そのような背景から、映像要約に関する研究も盛んに行われており、必要最低限の情報を抜き出す手法として漫画形式の映像要約も注目されている[1][2]。

本研究ではフィルムコミックという書籍メディアに着目し、漫画形式の映像要約を目的とする。そのために、アニメ作品内の静止画から計算される特徴量を利用した最適なキーフレームの検出を提案する。ここでキーフレームとは漫画形式のレイアウトに使用されるべき見栄えのよいフレームと定義する。

2. 提案手法

提案手法の流れを示す(図1)。

2.1 シーン変更判定

動画を連続した静止画として抽出し、連番画像の中のひとまとまりをシーンとする。シーンが切り替わったかどうかを決めるためにシーン変更判定を行う。入力とする動画はカラーのアニメ作品であり、登場するキャラクターはそれぞれ色による個性が付けられているため、岡寄らの手法[3]に基づき色差によるシーン変更判定を行った。入力画像を $M \times N$ 個のブロックに分割しダウンサンプリングを行う。各ブロックから得られた RGB の値を $L^*a^*b^*$ 表色系に変換し、隣接フレームのユークリッド距離を計算することにより色差を求める。色差が閾値以上の箇所

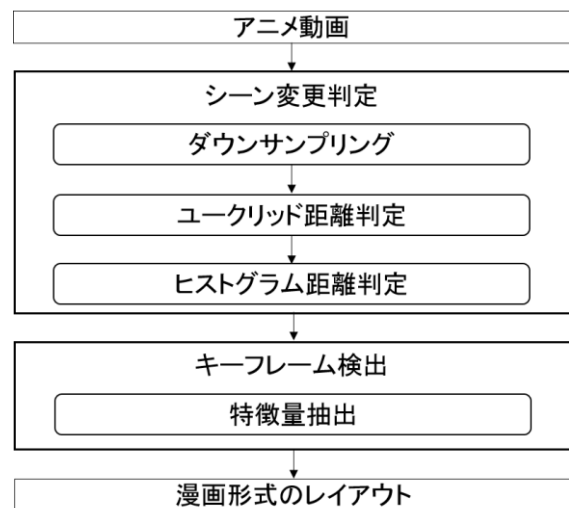


図1 処理全体の流れ

をシーン変更点とする。ここで、1シーンの中で重要なフレームが2つ以上となる場合のために、ヒストグラム距離判定を用いてさらにシーン変更判定を行う。これにより色差の判定では1つのシーンと判定されたものが再分割され、1シーンから複数フレームを選ぶことができる。

2.2 キーフレーム検出

シーン変更点を求めた後、漫画形式のレイアウトを行う際にコマとして使用するフレーム(キーフレーム)を選択する必要がある。キーフレームを決定するために、各フレームにおける特徴量を以下の要素に基づいて計算する。

(1) エッジピクセル差分

まばたきや口を動かすことによりエッジのピクセル数が変化する。まず、キャニー法によりエッジを抽出し、エッジと判定されたピクセル数を数える。連続した2つのフレームのエッジピクセル数の差分を後方差分法により求める。差分の大きい部分はシーン内で何か変化があったと考えられるため、キーフレームと判定される要素として用いる。

(2) 肌色領域抽出

キャラクターやカメラの動きにより肌色の領

A Keyframe Extraction Method for Anime Movie Content Summarization

Maki TAKAGI[†], Tomokazu ISHIKAWA[†], Masanori KAKIMOTO[†]

[†]School of Media Science, Tokyo University of Technology

域は変化する。岡崙らの手法に基づき、入力する画像の色を HSV 表色系に変換し、H 色相が $0^\circ \sim 44^\circ$ の範囲を肌色領域として求め、そのピクセル数を数える [3]。

(3) BGM・SE

アニメ動画から音声ファイルを抽出し、元の音ファイルに逆波形の音声ファイルを重ね合わせることによって BGM と効果音 (SE) を求める。効果音を付けたり、音量が大きくなったりする部分は感情の高まりや視聴者を引き付ける効果があると考えられるため、音量の強弱を用いてキーフレーム検出の要素とする。

以上の処理をした後、式(1)により特徴量 f を求める。

$$f(E, S, V) = \alpha E + \beta S + \gamma V \quad \dots(1)$$

ここで、 E はエッジのピクセル差分、 S は肌色領域のピクセル数、 V は BGM・SE の音量を示す。これらの要素はシーン毎の最大値、最小値を用いて正規化された値を使用する。 α 、 β 、 γ は重み係数を示す。これらの重み係数は、いくつかのシーンに対して、フィルムコミック作品中に使用されているフレーム (正解フレーム) と、特徴量の最大値となるフレームの誤差が少なくなるように決定する。すなわち、以下の誤差が最小になるように係数を決定する。

$$error = \sqrt{\frac{1}{N} \sum_{i=1}^N (n_i^* - \hat{n}_i)^2} \quad \dots(2)$$

ここで、 N は重み係数を決定するために使用するサンプリングシーンの数で、本研究では 16 とした。 n_i^* は特徴量が最大となるフレーム番号で、 \hat{n}_i は実際にフィルムコミックで使用されているコマのフレーム番号である。

3. 処理結果

事前に重み係数を決定するために使用したシーン以外に対して各フレームで特徴量を計算し、最大となったものをキーフレームとして検出した。

実際にフィルムコミックで使用されているフレームを正解とし、検出された結果と比較する。図 2 に提案法によって抽出されたキーフレームと正解フレームの結果例を示す。完全に正解フレーム(1-b)と同じフレームを検出できた結果を 1-a に示す。正解フレーム(2-b)と近いフレームを検出した結果を 2-a、正解フレーム(3-b)から遠い

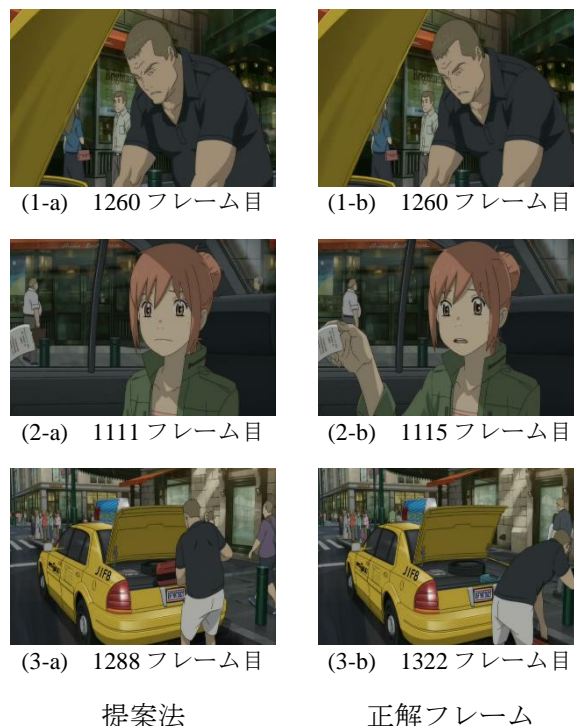


図 2 キーフレーム検出を行った結果

(©プロダクション I.G. 東のエデン劇場版 I The King of Eden. アスミック/フジテレビ. 2010.)

フレームを検出した結果を 3-a に示す。

同じフレームや見た目に大差無いフレームを選択することができた。

4. おわりに

アニメ作品の最適なフレームの選択を目的とし、シーン変更判定や画像の特徴量を用いてキーフレーム検出を行った。

今後の課題として、作品全体に対して複数フレームのためのシーン変更判定を行うと枚数が増え過ぎてしまうため、ユーザーが所望する要約枚数に応じてシーン変更の閾値を自動設定する方法を提案する予定である。

参考文献

- [1] Meng Wang, Richang Hong, Xiao-Tong Yuan, Shuicheng Yan, "Movie2Comics: Towards a Lively Video Content Presentation", IEEE Transactions on Multimedia Vol.14 No.3, June 2012.
- [2] Ying Cao, Rynson W.H. Lau, Antoni B.Chan, "Look Over Here: Attention-Directing Composition of Manga Elements", ACM SIGGRAPH 2014 Vol.33, July 2014.
- [3] 岡崙堅仁, 石川知一, 柿本正憲, 西田友是, "フィルムコミック形式によるアニメ作品の自動映像要約手法", VC/GCAD 合同シンポジウム, 2014.