

漫画技法を考慮したアニメ作品の映像要約手法

岡寄 堅仁[†] 石川 知一[†] 柿本 正憲[†]

[†] 東京工科大学 メディア学部 メディア学科

1. はじめに

近年、アニメ作品の動画配信が盛んになっており、放送中のアニメ作品の無料配信や、過去のアニメ作品の一回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、時間が必要であり全ての人が手軽に楽しむことはできない。

このような背景から、映像要約に関する研究が盛んであり、漫画形式の映像要約も注目されている。漫画形式の映像要約に関しては、画像から得られる特徴量からコマのサイズを決定する[1][2]などの研究がある。

本研究では、画像から得られる特徴量を利用するとともに、フィルムコミックという漫画形式の書籍メディアに着目し、漫画技法を考慮した映像要約を行うことを目的とする。

2. 提案手法

2.1 全体の構成

本研究の提案手法の全体構成を図 1に示す。

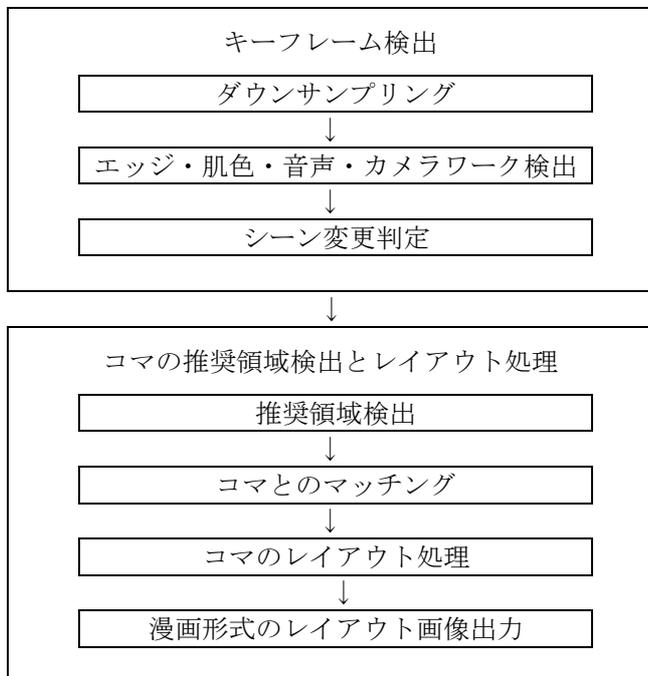


図 1 全体の構成

“A Method for Summarizing an Animation Film Taking into Account a Comic Technique”

Kento OKAZAKI, Tomokazu ISHIKAWA,

Masanori KAKIMOTO

Tokyo University of Technology, 1404-1 Katakura-machi, Hachioji-shi, Tokyo 192-0982 Japan

2.2 シーン変更判定

アニメ作品のキャラクターは特徴的な色を持っている。それらを比較するために、隣接フレームごとにブロックに分けて、各ブロックの RGB 平均を計算し、ダウンサンプリングを行う。ブロックごとに得られた RGB を CIE L*a*b*表色系に変換し、隣接フレームにおけるユークリッド距離を求める。ユークリッド距離が閾値よりも大きいとき、距離から求める特徴量の累計を取り、その累計が閾値を超えたときにシーンの変更を判定する。

上記に加えて、平井ら[3]の手法にならない、隣接フレームでの輝度値のヒストグラムの距離を評価の一つとして利用し、本研究の手法と平井ら[3]の手法を組み合わせ、シーンの変更判定を行う。

2.3 キーフレーム判定

漫画のコマとして選ばれるフレームは見栄えの良いフレームである必要がある。見栄えの良いフレームの判定のためにキーフレーム検出を行う。

はじめに、エッジ検出を行う。目を閉じているシーンよりも、目を開いているシーンは基本的にエッジのピクセル数が増加するため、これを評価の一つとする。しかしながら、服などが複雑なシーンにおいては、顔が映るシーンよりも、エッジの複雑な対象があるシーンが選ばれてしまう。そのため、肌色領域のピクセル数も考慮する。

また、音声も映像における重要な要素であるためこれも考慮する。隣接フレーム間で音量を比較し、音量の大きさを評価の一つとする。

同一ショット中のフレーム番号 f の各評価値から算出した特徴量が最大となるフレームを、そのショットのキーフレームとする。特徴量 S_f は次式から求められる。

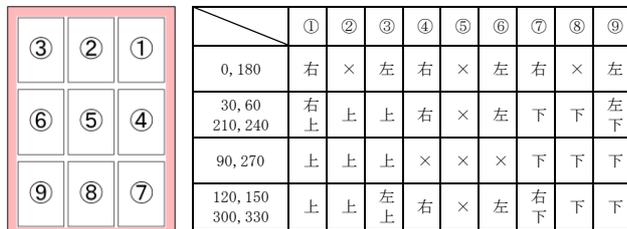
$$S_f = \alpha \cdot E_f + \beta \cdot V_f + \gamma \cdot F_f \quad \dots(1)$$

ここで、 E_f はエッジのピクセル数、 V_f は音量、 F_f は肌色領域のピクセル数、 α 、 β 、 γ は重み係数を示す。

2.4 カメラワーク判定

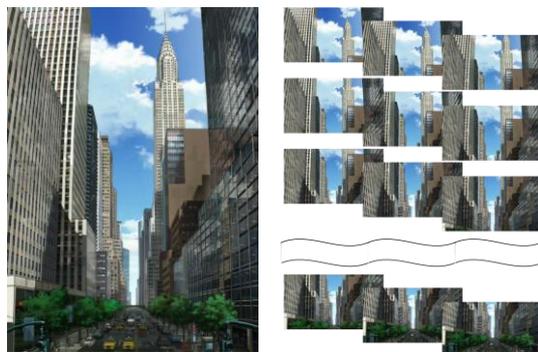
漫画の技法に、印刷範囲限界まで描写を行う断り切りゴマという描写がある。断り切り方向の判定のために、隣接フレーム間におけるオプティカルフローを利用した。画像を 4×4 の全 16 ブロックに分割し、各ブロックのオプティカルフローから計算した

画像の x 軸に対するベクトルのなす角度を 30 度ずつの 12 方向に量子化し、最頻出のベクトルの角度を求める。断ち切りの方向はベクトル角度とコマの位置により変化し、その対応関係を図 2 の (a)、(b) に示す。



(a) コマ番号 (b) コマ番号による断ち切り方向
 図 2 ベクトル角度による断ち切り方向の判定

アニメのシーンには、図 3 のように、本来縦長のシーンをカメラワークとしてティルトを行って表現することがある。このようなシーンは、フィルムコミック、漫画ともに大きなコマで表現される。ティルトシーンの判定のために、上記で求めたベクトル角度から 90、-90 度の場合、隣接フレームでテンプレートマッチングを行い、連結部分候補を求める。連結部分候補の HSV 表色系における色相 H を比較し、閾値以下のとき、ティルトシーンと判定する。



(a) 本来のシーン (b) 連番画像

図 3 ティルトによる縦長のシーン

2.5 コマの推奨領域検出

コマの推奨領域検出では、コマのサイズや形を決定するために必要な候補領域検出行う。

漫画、フィルムコミックともにコマ中心はキャラクターの顔が配置されることが多いため、顔検出によってキャラクターの位置を検出する。顔検出には高山ら[4]の手法の一部である肌色領域の検出と、輪郭の二次曲線近似を用いた。得られた各顔領域候補の全体の領域を推奨領域とする。

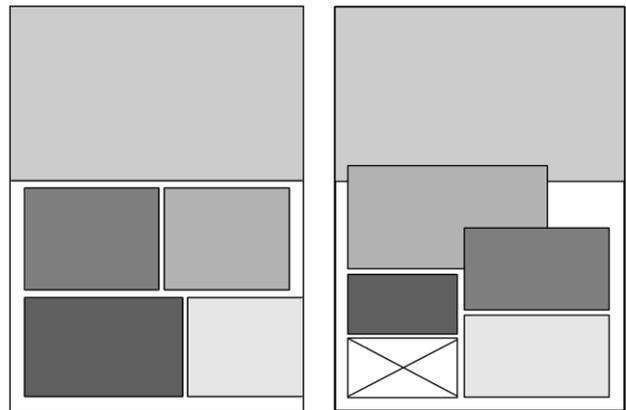
2.6 コマとのマッチング・レイアウト処理

コマとのマッチングにおいては、推奨領域検出で得られた、顔領域を利用する。顔領域が画像全体もしくは、小さい、検出なしの場合、背景や重要なシーンとして判断し大きなコマとする、その他のフレ

ームにおいては、コマの領域幅の最適化処理を行う。コマ間、推奨領域の各領域幅を、座標と領域幅を変数とした二次式として捉え、最小二乗法により、各領域幅を決定する。

3. 結果

同一ショットのキーフレームを同色で示したレイアウトの出力画像とフィルムコミック[5]の比較を図 4、キーフレームの要約フレーム数と同一フレーム数の一致率を図 5 に示す。



(a) 本手法 (b) フィルムコミック

図 4 レイアウト比較

	本手法	フィルムコミック	一致率
要約フレーム数	836 枚	862 枚	96.9%
同一フレーム数	419 枚	862 枚	48.6%

図 5 キーフレーム一致率

4. おわりに

画像から得られる、色特徴量やエッジ、オプティカルフローに加え音声に着目することで、漫画技法を考慮したキーフレームの検出と、コマとのマッチングを行い、フィルムコミックに近い枚数の映像要約、漫画形式のレイアウト画像の出力に成功した。

今後は、フィルムコミックと一致するキーフレームの検出率向上と、全時系列を考慮した重要シーンの検出や、吹き出しのような構成要素を考慮したレイアウトの最適化が望まれる。

参考文献

- [1] 福里他：アニメ作品のキーフレーム検出における漫画形式の映像要約手法、VC/GCAD 合同シンポジウム, 2013.
- [2] Calic, et al: Efficient Layout of Comic-Like Video Summaries. IEEE Transactions on Circuits and Systems for Video Technology, vol.17, No.7, pp931-936, 2007.
- [3] 平井他：シーンの連続性と顔類似度に基づく動画のコンテンツ中の同一人物登場シーンの同定, 映像情報メディア学会誌 Vol.66, No.7, ppJ251-J259, 2012.
- [4] 高山他:形状および対称性を用いたアニメキャラクターの顔検出, 情報処理学会第 73 回全国大会, 2011.
- [5] 株式会社扶桑社：フィルムコミック東のエデン劇場版.