

イラストレーションメイキング動画における注目箇所の効率的な探索システム

高橋 リサ†

† 東京農工大学院 工学府 情報工学専攻

藤波香織 ‡

‡ 東京農工大学 工学研究院 先端情報科学部門

1 はじめに

インターネットの普及により動画投稿サイトではあらゆる動画が投稿され、イラストレーション等の創作活動の制作過程を動画として発信する「メイキング動画」の投稿も一般的になりつつある。メイキング動画は初学者にとって優れた学習教材である一方で、全過程を視聴するために多くの時間を要し、ユーザが自身の視聴目的に沿った映像を探索することが困難であるという問題点が存在する。この問題はイラストレーションのメイキング動画に限らず多くのジャンルの動画視聴時に生じており、その対策として動画の要約に関する研究が広く行われている [1][2][3]。しかし、これらの要約手法の多くは対象をフレーム情報が大きく変化する実世界の動画やアニメーションに限定しているため、フレーム情報が緩やかに変化するイラストレーションのメイキング動画を同手法で要約することは困難であると考えられる。

この解決のために本稿ではメイキング動画の完成形イラストを利用する。完成イラストとはメイキング動画内の最終的な制作物のことであり、この画像からユーザは自身の興味ある項目の有無を確認することが可能である。したがって、予め提示されたメイキング動画の完成形のイラストから自身の興味のある箇所を選択及び視聴可能な動画再生システムは、メイキング動画の項目探索に有効であると考えられる。そこで本研究ではイラストレーションのメイキング動画に特有の要素である完成形イラストを用いた新たな動画探索システムの提案及びその実装を行う。

2 提案システムの概要と実装

2.1 システム概要

本システムは完成イラストから選択された部分的領域をクエリとして類似部分が描かれている時間を動画内から探索し、その時点のフレームを一覧として表示する。更に、フレームを選択すると生成されたフレームの時間まで動画を遷移させる処理を付与することでユーザが注目した項目の動画内探索を支援する。本システムは図 1 に示すように、1) 完成イラスト取得部、2) 描画座標生成部、3) 動画再生部といった主要機能で構成される。

2.2 完成イラスト取得部

一般的なメイキング動画の画面は図 2 に示すようにツール領域とキャンバス、そのキャンバスを取り囲む領

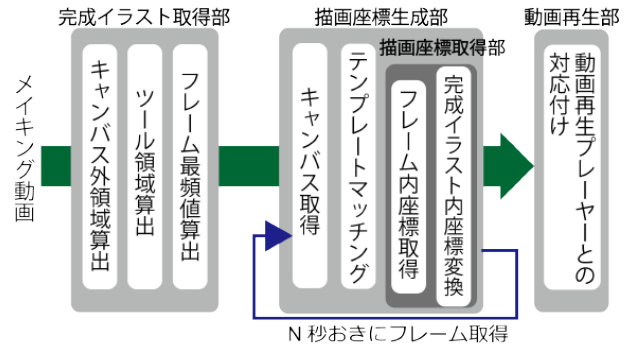


図 1 システム概要図

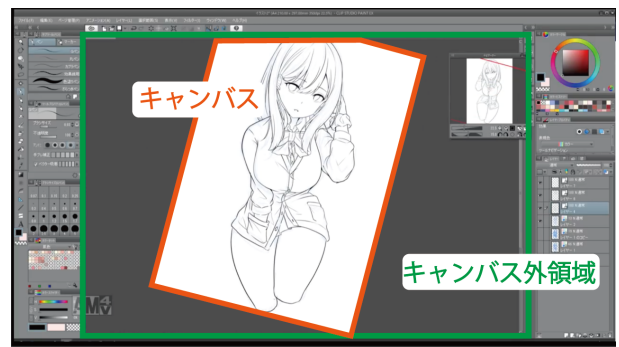


図 2 メイキング動画の画面例

域（以降キャンバス外領域とする）からなる。キャンバス外領域は表示するキャンバスの拡大や縮小に応じて増減し、キャンバスが少ないほど広くなる。そのため、キャンバス外領域が狭いフレームほどキャンバスの縮小率が高く、この領域に囲まれる部分は完成イラストの全体像である可能性が高いと考えられる。したがって、各動画において一定時間おきにキャンバス外領域を算出し、その領域が最大となる時間のキャンバスを完成イラストとして取得する。キャンバス外領域は複数枚の動画フレームの HSV 空間画素値の最頻値を閾値として算出した。この時、ツール領域が最頻値になってしまう可能性を考慮するために予めツール領域の画素値も推定し、その値は最頻値算出時に除外する。ツール領域は動画内において移動せず固定されている場合がほとんどであるため、移動物体の検出方法の 1 つであるフレーム差分法を利用し、反転した出力結果を逐次論理積合成することで固定領域であるツール領域の画素値を取得する。最頻値を探索する時間は本システムでは大まかなイラストが完成していると推測される動画時間全体の 9 割以降のフレームとする。キャンバス外領域が最大である時間の決定後は輪郭検出を利用し完成イラストを取得する。

2.3 描画座標生成部の実装

2.3.1 キャンバス取得部

完成イラスト取得部で算出したキャンバス外領域とツール領域の画素値を元にキャンバスを取得する。キャンバス外領域とツール領域の画素値を0とするようにフレームをマスク化し、キャンバスの領域のみを抽出した画像に輪郭検出を適用することで、各フレームからキャンバスを取得する。

2.3.2 テンプレートマッチング

メイキング動画内のキャンバスは拡大や回転操作が実施された場合フレームから取得した画像は完成イラストの部分的な領域である可能性がある。そこで、2.2節で取得したキャンバス内のイラストをテンプレート画像、完成イラストを入力画像としてテンプレートマッチングを利用し完成イラスト内から該当する領域を探索する。

まず、拡大処理による両画像のサイズ差に対応するためにテンプレート画像のリサイズを行う。完成イラストの幅を基準として1,2,4...100%とリサイズし、リサイズ後の画像をテンプレートマッチングで処理する。テンプレートマッチングの類似度は零平均相互相関から評価する。この手法は本来照明条件に強固になる手法であるが、イラストレーションにおける使用用途の場合、各画像の線の太さの相違が類似度に影響しにくくなることが見込まれる。したがって、同一の箇所を描いている場合でもラフ段階の薄いイラストや線画段階のはっきりしたイラストなどの違いによる類似度の低下を抑制することができると考え、採用した。

2.3.3 描画座標取得部

フレーム間差分法を利用してフレーム内における描画部分を取得し、2.3.1項で実施したテンプレートマッチングの結果を元に完成イラスト内において対応する座標を導出する。キャンバス部分以外の差分取得を防ぐために、差分画像からキャンバス外領域及びツール領域をマスク化し実行後の画像の重心座標を描画部分の座標とする。描画部分の座標を (x, y) 、取得したキャンバスの左上頂点の座標を (P_x, P_y) 、テンプレートマッチング実行時のリサイズ比率を R 、テンプレートマッチングで検出したキャンバスに対応する矩形座標の左上頂点の座標を (Q_x, Q_y) とすると、変換座標 (x', y') は式 (1) から導出する。

$$\begin{aligned} x' &= (x - P_x) \times R + Q_x \\ y' &= (y - P_y) \times R + Q_y \end{aligned} \quad (1)$$

対応座標と先頭からの動画の経過時間は各フレーム毎に CSV ファイルに保存し、動画再生システムにおいて利用する。

2.4 動画再生部

被験者が操作する動画再生システムを実装では、Python 用 GUI ツールキットである Tkinter を使用する。本システムは動画再生プレーヤー領域、完成イラスト領域、検索候補フレーム表示領域からなる。動画再

生プレーヤーは再生や一時停止、動画時間遷移といった一般的な動画再生プレーヤーの機能を保持する。右側の完成イラストからユーザが描画過程を視聴したい部位を選択すると、動画プレーヤー下部に選択した項目が描画されていると推測される動画のフレームを候補として表示する。この候補の中からユーザが更に選択すると、そのフレームの先頭からの経過時間まで動画の時間を遷移させる処理を行う。この候補は完成イラストがクリックされた際その座標を取得し、2.3.2 項で作成した CSV ファイル内の描画座標と距離に近い座標を抽出し、その動画時間のフレームを表示することで実現する。候補には類似した内容のフレームが含まれることが予想されるので、この対策としてクラスタリングを使用する。用いるクラスタリング手法の候補は k-means 法と階層的クラスタリング、DBSCAN が挙げられる。これらの中から本システムに適した手法を今後調査する予定である。



図3 動画再生部

3 おわりに

本稿では、イラストレーションのメイキング動画に特有な完成イラストを活用した新たな動画探索システムを提案し、主要機能についての設計と実装方法を述べた。

今後は動画再生部におけるフレーム候補の生成手法を変数とした実験を実施した上で、効果的であった手法を用いた本システムと一般的な動画プレーヤーを模したシステムとの間で動画時間の短縮性、操作性に関しての比較評価実験を実施する予定である。

参考文献

- [1] D. Zhong, S.-F. Chang. Structure analysis of sports video using domain models. In *Proc. of IEEE ICME 2001*, pp. 713–716.
- [2] S. E. F. De Avila *et al.* VSUMM: A mechanism designed to produce static video summaries and a novel evaluation method. *Pattern Recognition Letters*, Vol. 32(1), pp. 56–68, 2011.
- [3] S. Uchihashi *et al.* Video manga: generating semantically meaningful video summaries. In *Proc. of ACM Multimedia 1999*, pp. 383–392.