

マニュアル作成支援を目的とした作業動画からの 自動画像抽出システムの提案

村山 築[†] 西垣 貴央[†] 小野田 崇[†]

[†] 青山学院大学 理工学研究科

1 はじめに

近年、インターネット上での動画コンテンツの増加に伴い、様々な動画の利活用が行われている。その一つに、動画の要約を作成するために、動画から特徴的な画像を抽出するものがある。動画から特徴的な画像を抽出する方法として、RGB 表色系の変化を追跡する手法が用いられている [1]。しかし、ブロックの組み立てのような細かい作業を含んだ動画からの画像抽出には、作業対象の変化が小さすぎて適用できない。そこで本研究では、細かい作業を含む動画の特徴に着目して、マニュアルに使用できる画像を自動抽出する方法を提案する。提案手法の有効性を示すために、実験では複数のブロックの組み立て動画を用いて既存手法との比較を行った。

2 ビデオ映像からのマニュアル作成

動画の要約を作成するための研究として、連続するフレーム間の変化量を用いたシーン分割 [2] がある。ここで扱う「シーン」とは、作業動作の一区切りを指すものである。[2] では明瞭な分割点が存在しない未編集動画に対して、ある一定の「まとまり」を持つ動画単位を取り出すことで、未編集動画の汎用的なシーン分割を実現している。また汎用性を持たせるために、ユーザーがシーン分割する動画の時間的長さを制御できるパラメータを設定し、様々な動画への対応を可能にした。しかし、この手法は分割後も動画であるために、紙媒体のマニュアルの作成を支援することができない。そこで本稿では、動画から作業内容を表した画像を抽出し、作業マニュアルの作製を支援可能なシステムを提案した (図 1)。

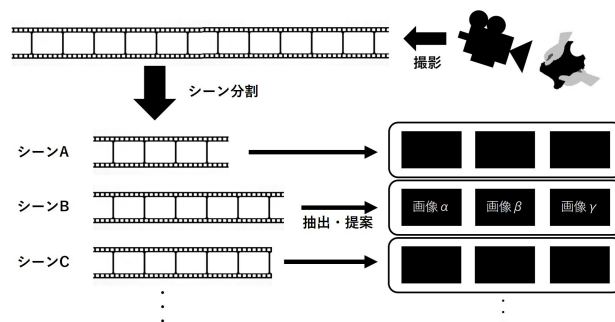


図 1: 提案システムのイメージ

3 提案手法

先行研究 [2] では未編集動画をより短い動画に分割した。しかし、作業マニュアルは紙媒体での需要もあり、動画では紙媒体にそのまま利用することはできない。そこで本研究では、作業の変化を捉えることができる要素として、動画内に出現する作業者の手の動きに着目することで特徴の抽出を行った。

3.1 映像の分割

本研究では、マニュアルを作成するための動画はユーザーの想定するシーン分割がされていることを前提とした。そのため、実験で用いた動画は分割が行いやすい作業間に休止の時間を設け、その休止を検出することによって正確な位置で分割を行った。

3.2 動作検出時に想定される手の動きの分類

今回着目する手の動きは「並行移動」、「形態変化」、「回転」に分類を行うことができ、これらの変化を追跡する必要がある。「並行移動」は画面に対して横方向、もしくは縦方向に手が移動することを指す。画面内の手領域の大きさや形状は変化しないが、領域の座標が変化する。「形態変化」は日常的な手の動作を指や関節の動きから分類したものを指す [4]。「回転」は手がカメラに対して水平方向もしくは垂直方向に回転する場合を指す。動画から手の状態を推定する研究として、手話の自動認識 [5] が行われているが、本研究では手の変化を検出できればよい。したがって手の形状そのも

Development of a Support System for Automatic Image Extraction to Make a Manual

Kizuku MURAYAMA[†], Takahiro NISHIGAKI[†] and Takashi ONODA[†]

[†] Graduate School of Science and Engineering, Aoyama Gakuin University

252-5258, Fuchinobe, Japan

のの推定は行わない。

3.3 分割されたシーンからの手領域に着目した画像抽出

手領域の抽出では照明や作業者固有の肌色の違いによる影響を受けにくくするために、HSV 表色系がよく用いられる [3]。本研究でも HSV 表色系を用いて肌色に相当する色相を持つ領域を抽出した。また、動画に登場する人物が1人であると仮定し、左右それぞれの手領域が抽出できるよう、一定以上の大きさがある領域を最大2つまで抽出した。

抽出した手領域から変化を追跡するために「面積」, 「方向」, 「中心座標」の3つを特徴量として利用した。これら3つの特徴量について前フレームとの差分を計算することで追跡を行った。「面積」は手領域の面積を表した特徴量である。「方向」は手領域の広がる方向を表す特徴量である。手領域の境界に対して主成分分析を行い、第一主成分軸を手領域の広がる方向とした。「中心座標」は手領域を構成する境界の座標の平均値である。また中心座標は手領域の形状によっては手領域外に存在する可能性がある。これらの特徴量を各フレームで計算し抽出に利用した。また、画像の抽出は分割された区間内で各特徴量が大きいものを対象として行った。これにより、1シーンあたり3枚の画像を抽出することができる。

4 実験

抽出の実験は、2種類の異なるおもちゃのブロック(図2)の組み立てを撮影し行った。また比較手法として、シーン分割に用いられる RGB 表色系の変化度を追跡する手法 [2] を利用した。これにより、既存手法と提案手法では抽出できる画像にどのような差異があるかを検証した。提案手法は、作業を撮影した動画に出現する手に着目することで、より細かな内容の変化に対応した抽出が行えることを確認した(図3)。ブロックの組み立てマニュアルを考えた場合、必要となる画像は主にブロックを組付けている様子を表した画像となる。図3では、先行研究 [2] はパーツを組付ける前の移動部分を抽出している。それに対し、提案手法は組付けた時点の画像を抽出している。これらを比較した場合、組み立てマニュアルにより適した画像は、提案手法が抽出した画像だといえる。

先行研究 [2] は、比較的動きの小さい、部品を組み付ける瞬間を捉えられないという問題点があった。これは、人間の手の動きが一定ではなく、緩急が存在する場合に、“移動量の変化=作業内容の変化”が成り立たないからである。しかし、提案手法では、物体の移動以外を捉える特徴量として、手領域の面積と方向を用いることで、この問題を解決した。同様に、組付け前

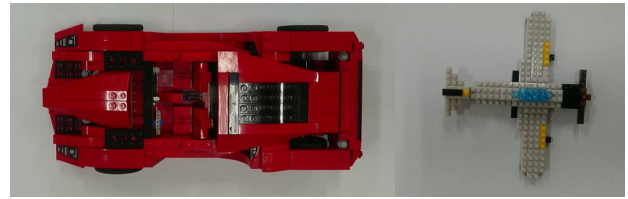


図 2: 実験に用いた2つのブロック

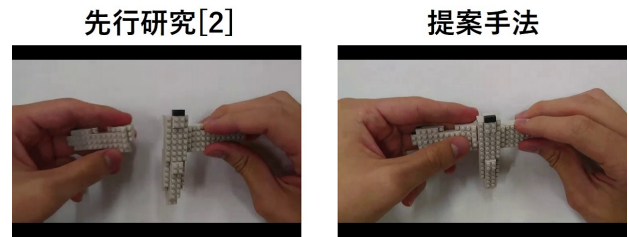


図 3: 抽出された画像の比較例

に部品を裏返す動作を含むシーンでも、提案手法は先行研究 [2] よりも良い結果を得ることができた。

5 まとめと今後の課題

提案手法で用いたのは「面積」, 「方向」, 「中心座標」の3つを特徴量であったが、手領域から得られる特徴量は他にも存在する。今回の実験では特に「形態変化」の検出が困難であった。よって、多様な作業動作に対応した画像の提案を行うために、手領域の周囲長や手領域の多角形近似など、本手法で取り上げなかった特徴量についても比較・検討を行いたい。

参考文献

- [1] 精廬幹人, 橋本周司 (2003) 「統計的な特徴量に基づく動画画像検索手法」, 情報処理学会第 65 回全国大会講演論文集
- [2] 堤富士雄 (2001) 「ウェアラブルな視覚記憶補助装置のための映像分割手法」, WISS2001
- [3] 松橋聡, 藤本 研司, 中村 納, 南 敏 (1995) 「顔領域抽出に有効な修正 HSV 表色系の提案」, テレビジョン学会誌 Vol. 49, No. 6, pp. 787-797
- [4] 鎌倉矩子, 三星文子, 浅海奈津美, 中田真由美 (1986) 「物体の操作における健常手の動きのパターン」, リハビリテーション医学 Vol. 23 no. 2, 59-67
- [5] 喜安千弥, 藤村貞夫 (1999) 「動画像による連続手話の認識方法」, 計測自動制御学会論文集 Vol.35, No.8, 1099-1104