

内容情報を扱う映像構造化・検索手法

5D-4

佐伯 剛幸 元木 誠

NEC C&C 研究所

1 はじめに

映像素材が溢れている現在においては、映像の内容情報をデータベース化し、すべてを見なくても必要な映像を探せるような機構を提供する必要がある。このため、我々は、内容情報を扱う映像構造化・検索モデルを提案している [1][2]。提案したモデルでは、内容情報として、映像中の特定の内容に関する部分を指すシーンや物体の領域を指すホットスポットを扱っている。だが、実際にこのような構造に基づいて内容情報を扱うためには、動画エディタなどを使用して情報を手動で作成する必要があり非常に工数がかかる。

このために、本稿では、構造化を画像処理技術を用いて行なう方法を提案する。従来、画像処理技術を用いて映像から内容情報を抽出する手法は多数提案されているが、それらの手法は適用範囲を限定したものが多く、また、処理の重いものが多かった。だが、映像はデータ量が膨大であるためにできるだけ処理を軽くする必要がある。このようなことから、本稿では、映像からキーフレームのみを取り出し、キーフレーム間で物体の動きを抽出して物体追跡を行ないホットスポットを検出する手法を提案する。

2 内容情報に基づく映像構造化モデル

[1]のモデルでは、映像情報を構造化する際に、映像の内容として本質的に存在する物理的な構造と人間が映像の意味に基づいて付加する論理的な構造とに分離している (図1)。構造を二つに分けることによりそれぞれの構造での情報の変更が局所化され、また、固定した情報である物理構造に対して、論理構造をつなぎ換えることによりさまざまな視点から映像に対して意味情報を与えることが可能になる。

3 映像自動構造化手法

3.1 概要

本稿で提案する映像構造化手法は、映像から図1に示した物理構造のホットスポットを半自動的に構築することを目的とするものである。

映像からホットスポットを抽出する際には、(1)物体

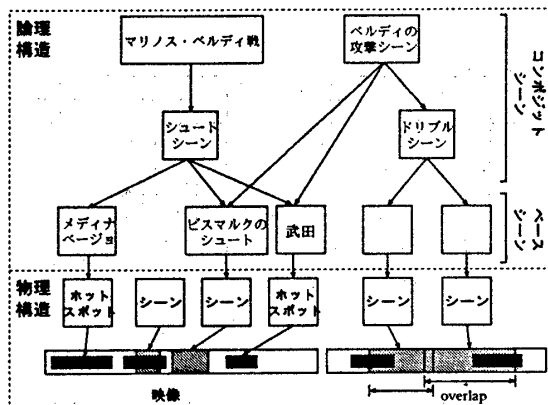


図1: 内容情報を扱う映像構造化モデル

が存在することを認識できることが重要で領域抽出の精度は低くてもよい、(2)物体の動きの細かい追跡は不要であり間は補間すれば十分である、などを考慮して処理を軽くすることできる。このようなことから、提案手法では、映像からキーフレームを取り出し、物体抽出・追跡処理はキーフレームに対してのみ行なう。また、これにより誤差が多く生じるために、統計的手法などによる誤差除去処理を行なう。

提案手法の処理の流れを以下に示す。以降では、このうちの動き抽出について説明する。

1. キーフレーム取得
色ヒストグラム、フレーム間差分により画面が大きく変化したフレームをキーフレームとして取り出す。
2. 物体領域抽出
キーフレームを縮小化し、色、輪郭情報に基づいて領域分割した画像 (24×24 画素) を作成する。
3. 動き抽出
キーフレームからサンプリング画像 (96×96 画素) を作成し、時間的に連続した2枚のサンプリング画像から4×4画素のブロックによるブロックマッチング法により動きの検出を行なう。さらに、領域情報を基に物体全体としての動きを導出する。
4. 領域の対応付け
動きベクトル使用して物体の一連の動きを追跡する。

3.2 動き抽出処理

処理の概要を図2に、処理の流れを図3に示す。

3.2.1 ブロックの動き抽出

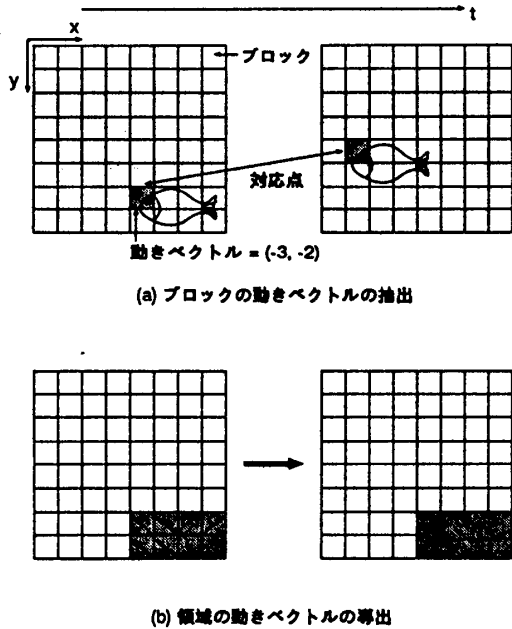


図2: 動き情報の抽出の概要

図2は、時間的に連続する2枚のサンプリング画像(図3.(2))をブロック分割した状態(図3.(3))を示している。まず、1枚目の画像について、コントラストが低いブロックを除去する(図3.(4))。これは、コントラストが低いブロックは特徴が少ないために対応点に誤差が入りやすいためである。次に、除去されなかった1枚目のすべてのブロックの色情報を2枚目のブロックの色情報と比較して色の違いが最小となるブロックを対応点とし、ブロックの位置ベクトルの変化をブロックの動きベクトルとする(図3.(5))。この動きベクトル抽出処理の概要を図2.(a)に示す。さらに、ブロックマッチングの際の色の違いの大きいブロックを除去する(図3.(6))。これは、処理するフレームの間隔が空いているために対応点が消えてしまう場合に発生する不正ベクトルの影響を除去するためである。以上により、ブロックの動きベクトルを導出する。

3.2.2 領域の動き抽出

物体領域情報を使用して領域内に存在するすべての動きベクトル値のヒストグラムを作成し、出現頻度最大の値を領域の動きベクトルとする(図3.(9))。この処理の概要を図2.(b)に示す。ヒストグラムを用いることにより、ブロック単位で除去できなかった不正ベクトルを取り除くことができる。以上により、領域全体としての動

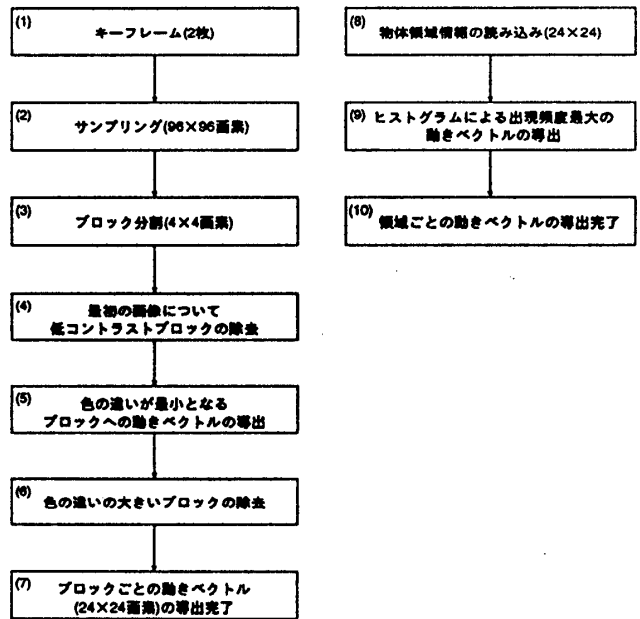


図3: 動き情報の抽出処理流れ図

きベクトルを算出する。

4 評価

2つのサンプル映像に対して処理を行なった評価結果を示す。サンプル映像1はニューヨークの町並みを映した映像、サンプル映像2は花を映した映像である。ここで、総領域数とは、すべてのキーフレームに含まれる物体領域の合計であり、認識率は動きが正常に取れた領域の割合を示す。

	キーフレーム数	総領域数	認識率 (%)
サンプル映像1	6	40	33
サンプル映像2	4	9	67

5 まとめ

提案手法は、低コストでも一定の動き抽出に成功することを確認した。今後は、サンプルを増やしてさらに方式を改良していく必要がある。

謝辞

本研究を進めるに当たって御指導頂いた NEC 北米 C&C 研究所 平田 恭二氏に感謝します。

参考文献

[1] 佐伯, “映像の構造化と操作方式”, 信学会, DEWS(1995)
 [2] 元木, 佐伯, “内容に基づく映像 DB 検索のためのオブジェクト指向フレームワーク”, 情処研報, DBS104-18(1995)