

色領域の位置情報に基づく動画像からの類似シーンの自動獲得

沼口義和[†] 高野茂[‡] 新島耕一[‡]九州大学大学院 システム情報科学府情報理学専攻[†]九州大学大学院 システム情報科学研究院情報理学部門[‡]

1. はじめに

近年のデジタルビデオ機器の急速な普及により、動画像中からユーザが見たいシーンを検索するシステムの開発が望まれている。

本稿では、質問画像の色領域の位置情報が類似したシーンを、動画像中から自動的に獲得する手法を提案する。本手法では、まず動画像中のフレーム画像に対してダイアディックウェーブレット分解を施す。次に得られた低周波画像から、その代表となる色を抽出し、色に基づく領域分割を行う。その際、代表色の類似するフレーム画像毎に動画像をクラスタリングしておく。さらに、クラスタリングされたフレーム画像群から、色領域の配置の情報を抽出し、データベース化する。こうして計算された色の位置情報をもった特徴データベースを用いて、ユーザが入力した質問画像に類似するシーンを動画像中から高速に検索するシステムを構築する。

2. 画像の特徴抽出

2.1. ダイアディックウェーブレット変換

色情報に基づく領域分割を行う前に、入力画像に対してダイアディックウェーブレット変換[1]を施し、サイズを変えずともとの画像を滑らかにした低周波画像を得る。この低周波画像から次節で説明する方法により代表色を抽出する。

2.2. パレットの作成

色情報による領域分割が画像の明度の変化に依存しないように、2.1節で得られた低周波画像の色空間を赤(R)、緑(G)、青(B)から色相(H)、彩度(S)、明度(V)に変換し、画像の特徴として色相と彩度のみを使用する。その際、図1のように、色相と彩度からなる色分布のマップを作成する。マップの値は次の3値からなるベクトル $\vec{x} = (h, s, \alpha)$ で表現する。ただし、 h と s はそれぞれ色相と彩度を表し、 α はその頻度である。このマップから16個のピークを検出し、それらをフレーム画像の代表色（以下パレットとよぶ）とする。パレット \vec{P} は、抽出された16色を要素にもつベクトルである。

$$\vec{P} = (\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{16}).$$

抽出されたパレットだけを用いて低周波画像を表示すると、図2のように色領域に分割できる。

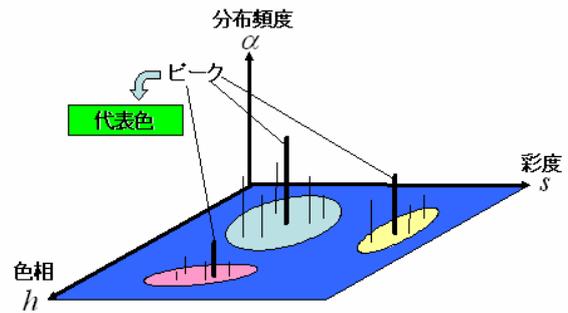


図1. 分布ピークの計算

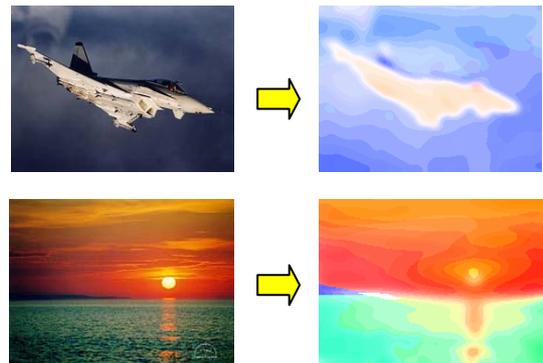


図2. 色情報に基づく領域分割

2.3. クラスタリング

動画像を効率よく検索するために、あらかじめパレットの類似するフレーム画像をクラスごとに振り分ける。まず、それぞれのフレーム画像のパレットを計算し、距離の近いものを同じクラスに分類する。さらに各クラスにおいて、パレットの重心をとり、クラスパレットを計算する。各クラスにおいて、このクラスパレットを用いて、ふたたび色領域に分割する。この操作は、次節で説明する色領域の並びの特徴を抽出するとき必要となる。

2.4. CRTの作成

本節では、画像の特徴を表す合成領域テンプレート[2] (Composite Region Template, 以下CRTで表す)のデータベースを作成する方法について説明する。まず、クラスパレットにより色領域に分割された画像を水平方向に5等分し、その各画像列に対し、上から下に色の並びを調べる。この色領域の並びを文字列化したものを

領域列とよび、1枚の画像に対して5つの領域列を作成する。次に、これらの領域列を用いて、画像の色領域の上下関係を図3のように、16×16サイズの行列で表現する。この行列が画像の色領域の位置情報を表すCRTである。CRTの行成分は先に現れる色領域のラベルを、列成分は後に現れる色領域のラベルをそれぞれ表しており、行列の要素はその領域の並びの頻度を表す。

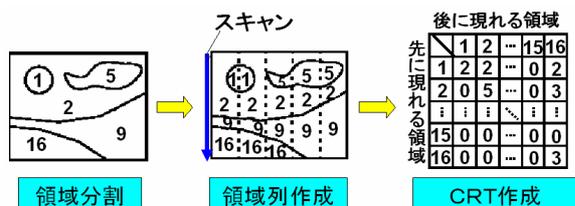


図3. CRTの作成

各クラスにおいて、クラスパレットを用いて、それぞれのフレーム画像のCRTを作成し、それらをデータベース化する。

3. 動画からの類似シーンの獲得

本節では、質問画像と類似するシーンを獲得する方法について述べる。まず質問画像のパレットを作成し、最も距離が近いクラスパレットを見つける。次に、このクラスパレットを用いて、質問画像を色領域に分割して、その領域列からCRTを作成する。最後に、クラスに所属するフレーム画像の中から、質問画像のCRTに類似するCRTをもつ画像を検索する。CRTの類似度を測るための判定式は次のように定義する。

$$|Q - T_k| = \sum_{i=1}^{16} \sum_{j=1}^{16} (Q[i, j] - T_k[i, j])^2 \quad (1)$$

ただし、 Q は質問画像のCRTであり、 T_k は k 番目のフレーム画像のCRTである。クラス内において、(1)が最大となるCRTをもつフレーム画像が類似シーンとして出力される。

図4に画像の入力から類似シーンの獲得までのフローチャートを示しておく。

4. シミュレーション

実験では、画像サイズが128×85で1秒間に30フレームからなる5分間の動画から、図5に示す質問画像に類似するシーンを検索した。その結果を図6に示す。

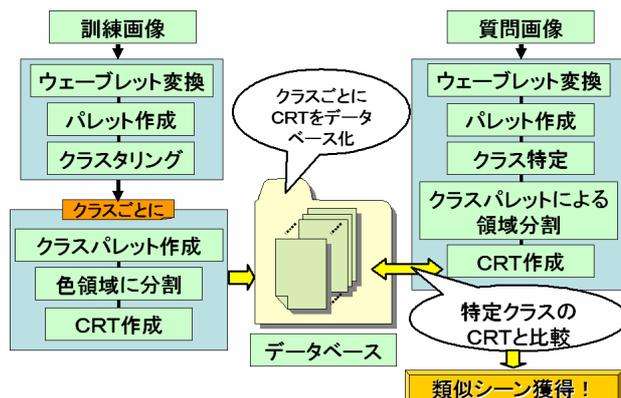


図4. システムの概要



図5. 質問画像

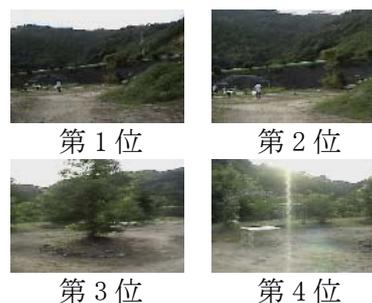


図6. 獲得された類似シーン

5. まとめと今後の課題

本稿では、色領域の位置情報をもとに、質問画像に類似するシーンを動画データから自動的に検索するシステムを構築した。フレーム画像を色領域に分割するために、画像の代表色であるパレットを作成する方法を提案した。各フレーム画像間のパレットの類似度を測ることで、フレーム画像をクラスターリングし、類似シーンを効率よく検索することができた。今後の研究課題のひとつは、ダイアディックウェーブレット変換による高周波画像を用いて、画像の形状の特徴を抽出し、それを利用した画像検索システムを構築することである。

参考文献

- [1]T. Abdulkirim, K. Nijima and S. Takano, "Lifting dyadic wavelets", DOI Technical report, No.212, Department of Informatics, Kyushu University, 2002.
- [2]John R. Smith and Chung-Sheng Li, "Image Classification and Querying using Composite Region Templates", Computer Vision and Image Understanding, Vol.75, No.1/2, pp.165-174, 1999.