

確率的 ARG マッチングを用いたスケッチ画像を問い合わせとする 動画検索システム

中村 直人[†] 高野 茂[‡] 岡田 義広[‡]

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

近年におけるグラフィックデバイスの急速な性能向上により、一般ユーザが動画像などのマルチメディアデータを扱うことが容易になった。そのため、動画を扱うサービスが増加しコンテンツの総量が爆発的に増加している。再利用などの観点からも、これらを効率よく管理するために精度のよい検索システムの開発が求められている。

本論文では、属性付き関係グラフ (ARG) を動画フレームの特徴とした検索システムの提案を行う。ARG は内容物の構造を表現することに適しているため、これを用いることで、精度のよい内容物ベースの検索を行うことができる[3]。また、ユーザがコンテンツの検索を効率よく行うためには、直感的な入力インターフェイスが必要である。そこで、本システムではスケッチインターフェイスを質問の入力に用いる。システムの検索速度に関する利便性を向上させるため、動画像のキーフレーム抽出を行う。これには分割カイ二乗検定法を用いた手法を採用する[1]。

2. システムの概要

本システムの概要図を図 1 に示す。本システムは質問入力のためのスケッチダイアログ、動画像を表示するための動画像ウィンドウ、様々な操作を行う操作ダイアログ、そしてキーフレームや検索結果を表示するためのサムネイルダイアログで構成される。ユーザはスケッチインターフェイスを用いて直感的に質問画像を入力する。検索時には質問画像から ARG を構築し、キーフレームからあらかじめ構築された ARG 群とマッチングを行い

類似度を計算する。

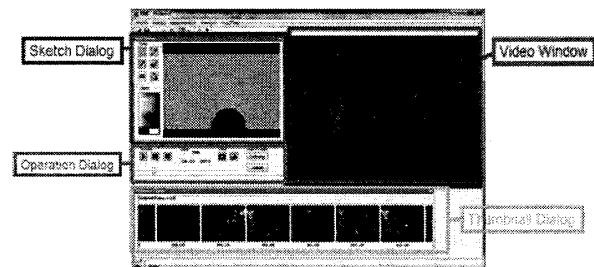


図 1: システムの概要図

3. キーフレーム抽出

本システムでは、検索速度向上のために前処理としてキーフレーム抽出を行う。先行研究として、分割カイ二乗検定法を用いた手法があるが、これは色の強度情報のみを使用したものであった[1]。本システムでは色相情報をコンテンツ抽出の全行程において活用するため、色相情報を考慮した分割カイ二乗検定法を提案する。本手法において使用される評価値 V は次式であらわされる。

$$V = \sum_{j=0}^{63} \frac{\{H_c(f_1, j) - H_c(f_2, j)\}^2}{H_c(f_1, j)}$$

ここで f_1, f_2 は時間的に隣接するフレームであり、 $H_c(f_i, j)$ はフレーム f_i における色相値のヒストグラムの、量子化階級 j の値である。

4. 確率的 ARG マッチング法

ARG はコンテンツの構造的な特徴を効率よく表す。ARG は頂点集合 V 、辺集合 E 、そして属性集合 A からなる。通常のグラフ構造は頂点と辺からなるが、ARG では各頂点と辺に属性が付加される。確率的 ARG マッチングでは、対象となる 2 つの ARG のうち片方を、VCP, ATP という 2 つの過程を経てランダムに変形させた際に、もう片方の ARG に変形する確率を、類似度の計算に用いる

Video scene search system based on stochastic ARG matching with sketch input interface.

[†]N. Nakamura, Information science and Electrical Engineering, Kyushu University.

[‡]S. Takano and Y. Okada, Information science and Electrical Engineering, Kyushu University.

[3]。このマッチングの概要を図2に示す。

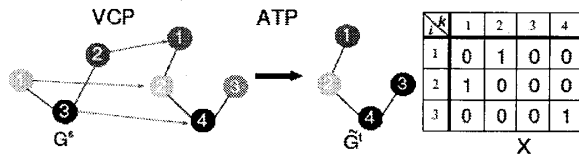


図2：確率的 ARG マッチング

マッチングを行う ARG をそれぞれ G^s , G^t とすると、類似度 $S(G^s, G^t)$ は以下のように定義される。

$$S(G^s, G^t) = \frac{p(Y^t | Y^s, H_p = 1)}{p(Y^t | Y^s, H_p = 0)}$$

ここで Y^s , Y^t はそれぞれ G^s , G^t の属性であり、 $H_p = 0$ の際には VCP を行わず、 $H_p = 1$ の際には VCP を行うものとする。

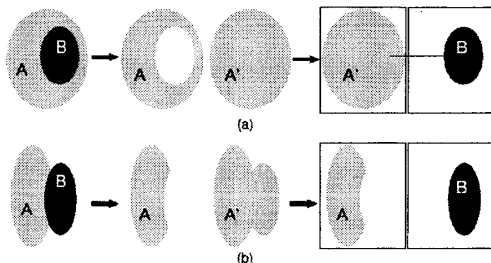
変形確率を計算する際に多次元正規分布を使用するが、この正規分布の分散共分散行列はテストデータを用いた学習により決定する。よって、各属性間の重みを考慮した類似度となる。

5. ARG の構築

通常、動画のフレームには様々な内容物が含まれており、それぞれの内容物は隣接していたり前後していたりといった関係を持つ。本システムでは、動画フレームの内容物の構造を特徴とした検索を行うために、画像を層状に分解する[2]。この分解をレイヤー化と呼ぶ。これは円形度という指標を用いて、分解後の領域が単純な形状になるよう分解を行う。円形度 e は次のように定義される。

$$e = 4\pi \frac{S}{l^2}, \quad 0 \leq e \leq 1,$$

ここで、 S は領域の面積、 l は領域の周囲長を表す。レイヤー化の概要図を図3に示す。



(a) A は下方レイヤーの候補領域であり、B は上方レイヤーの候補領域、 A' はそれらの和となる領域である。 A' の円形度が A のものより大きい場合、A, B 間にレイヤー関係を結び、領域 A を A' に更新する。
(b) A の円形度が A' のものより大きい場合、A, B はそのままとする。

図3：レイヤー化の概要

6. 実験

前述の検索システムの実装を行い、検索実験を行った。検索対象には約 2200 秒の長さの動画像を用い、これから 1 秒ごとにフレームを切り出し、さらにキーフレーム抽出を行った結果、約 300 枚のキーフレームを抽出した。ARG の特徴としては円形度、面積、色相値などを使用した。検索結果を図4に示す。検索されたフレームは、質問画像と内容物が構造的に類似している。

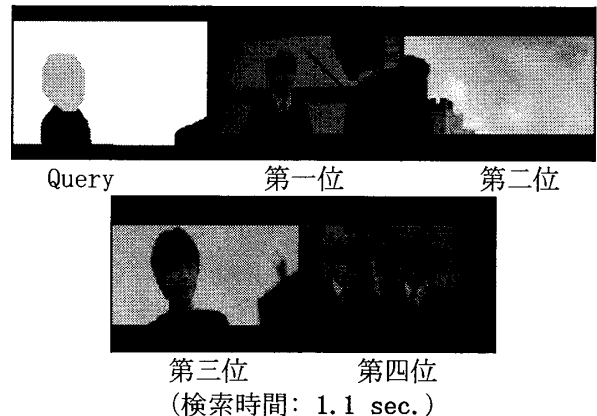


図4:動画シーンの検索結果

7. まとめ

本論文では確率的 ARG マッチングを用いた動画像シーンの検索システムを提案、実装し、検索実験を行った。本システムでは入力システムにスケッチインターフェイスを採用しているため、直感的な入力が可能となっている。実験では質問画像と構造的に類似した検索結果を得ることができた。

本システムでは精度の高い検索を直感的な操作で行うことが可能だが、検索時間が長いという問題がある。これはグラフマッチングの計算コストが高いことに起因する。そこで今後の課題としては、システムの高速化が挙げられる。

参考文献

- [1] Nagasaka, A. and Tanaka, Y., "Automatic Video Indexing and Full Video Search for Object Appearances", *Visual Database System II*, pp. 113 – 127, 1992.
- [2] Nakamura, N., Takano, S., and Nijima, K.: Video Scene Retrieval Based on The Layerization of Images and The Matching of Layer-Trees. In: *Proceedings of IVCNZ 2005 Image and Vision Computing New Zealand 2005*, pp. 449-454 (2005).
- [3] Zhang, D., and Chang, S.: Stochastic Attributed Relational Graph Matching for Image Near-Duplicate. Columbia University ADVENT Technical Report 206-2004-6 Columbia University (2004).