

画像検索およびその高速化技術の放送局における実用

Application of Image Search and Its Speed-up Technology in Broadcast Station

望月 貴裕† 住吉 英樹† 佐野 雅規† 藤井 真人†
Takahiro MOCHIZUKI Hideki SUMIYOSHI Masanori SANNO Mahito FUJII

1. まえがき

ソフトウェアとハードウェアの高度化や、通信環境の整備により、通信網上の広い範囲に蓄積されたさまざまな画像・映像データにユーザーが容易にアクセスできるようになった。放送の分野においても、放送アーカイブスなどの大規模な映像資産を活用する需要が高まっており、NHKも風景や動物などを撮影した素材映像を無償で提供するWebサイト「NHK クリエイティブ・ライブラリー（以下CL）」の運用を開始している[1]。

そのような流れの中で、大量の映像データから所望のシーンを効率的に探し出す検索技術が求められている。本稿では、我々が研究を進めてきた基本技術である、クエリー画像の描画による画像検索手法および、画像の「構図」を用いた検索処理の高速化手法について解説するとともに、本技術を適用したシステムの例として、CLに適用したクリップ検索システムと、アーカイブスのような大量の番組映像を検索対象とした検索システムを紹介する。

2. クエリー画像描画による画像検索

本章では、クエリー画像の描画による画像検索手法[2]の概要について述べる。この手法は、検索対象画像集合を予め解析して抽出した、出現頻度の高い主要な色およびテクスチャ（ブロック画像）のリストを「パレット」として持ち、このパレットを使ってクエリーを描画することで、「見た目がこんな画像」という検索意図と検索対象の画像特徴が合致した、納得性の高い画像検索を実現するものである。

2.1. 特徴ベクトルとパレット生成

特徴ベクトルは、画像を分割したブロック領域単位での色特徴およびテクスチャ特徴により構成される。色特徴ベクトルは、RGBそれぞれの平均値および色相成分のヒストグラムにより構成する。テクスチャ特徴ベクトルは、テクスチャの粗密および複雑性などの多様な情報を含むフラクタルシーケンス[3]と、テクスチャの方向性に関する情報を補強するためのエッジ方向ヒストグラムにより構成する。

パレットの色リストは、検索対象画像集合の色ヒストグラム解析による度数の高い色で形成する。また、テクスチャリストは、検索対象画像集合の全ブロック領域画像の、特徴ベクトルの類似性に基づいて生成した各クラスタの代表画像で構成する。

2.2. 検索処理

ユーザーは2.1節で生成したパレットを用いて、ブロック領域や簡単な図形の塗りつぶしによりクエリー画像を描画する。次に、描画した画像のブロック領域ごとの特徴ベクトルを計算する。この特徴ベクトルと検索対象の各画像の特徴ベクトルとの類似度を求め、類似度の高い画像を一



図1 クエリー画像の描画例と検索結果

定の数だけ検索結果として抽出する。図1にクエリー画像の描画例とそれに対する検索結果を示す。

3. 主要構図データを用いた検索処理の高速化

一般的な画像は、一面の桜、空と山、ビルに囲まれた道路など、ある典型的な構図で構成されている。本章では、[4]の技術を基にした、検索対象の画像から典型的な構図のバリエーション（以下、主要構図データ）を生成し、主要構図データであらかじめ検索対象の画像を分類して検索処理を高速化する手法について述べる。

3.1. 主要構図データの自動生成

主要構図データは検索対象からサンプリングして生成した画像集合（以下、部分画像集合）の各画像から取得した構図データをクラスタリング処理して生成する。各画像の構図データは次のように取得する。まず、画像を $M \times N$ 個のブロック領域に分割し、各ブロック領域の色特徴ベクトルの類似度に基づいて最大2グループに分類する。そして、同じグループに属するブロック領域に同じフラグ（1または-1）を付与する。このフラグ系列が構図データとなる。

部分画像集合のすべての画像の構図データを求めた後、構図データの類似度に基づいてクラスタリングする。なお、2つの構図データ $B_1 = (b_1^{(1,1)}, \dots, b_1^{(M,N)})$ と $B_2 = (b_2^{(1,1)}, \dots, b_2^{(M,N)})$ の類似度 $S_S(B_1, B_2)$ は(1)式で与えられる。

$$S_S(B_1, B_2) = \frac{\sum_{i=1}^M \sum_{j=1}^N (1 - 0.5 \times |b_1^{(i,j)} - b_2^{(i,j)}|)}{MN} \quad (1)$$

生成されたクラスタのうち、属する画像数が設定値以上のすべてのクラスタについて構図データの平均データを求め、小数値である平均データの各フラグ値を、正値を1、負値を-1に量子化したものを主要構図データとする。

3.2. 主要構図データによる検索対象画像の分類

検索対象のすべての画像をあらかじめ主要構図データに対応した構図グループに分類する。まず、すべての画像の構図データを求める。次に、画像 I の構図データ B_I と各主要構図データ B_T ($T=1, \dots, N_S$; N_S は主要構図データの数) との類似度 $S'_S(B_I, B_T)$ を(2)式で計算する。

$$S'_S(B_I, B_T) = \max\{S_S(B_I, B_T), S_S(B_I, \tilde{B}_T)\} \quad (2)$$

† NHK 放送技術研究所

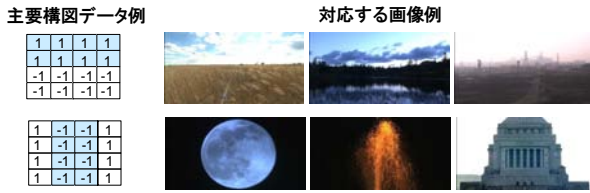


図2 主要構図データ例と分類された画像例

ここで、 \tilde{B}_T は B_T の符号をすべて反転した構図データである。類似度 $S_s(B_i, B_T)$ と $S_s(B_i, \tilde{B}_T)$ を計算する理由は、フラグの符号が全て反転した構図データ同士は同じ構図を意味するためである。そして、最も類似度 $S_s(B_i, B_T)$ の高い主要構図データに対応した構図グループへ画像 I を分類する。図2に主要構図データの例とそれに対応した構図グループへ分類された画像の例を示す ($M=N=4$)。

3.3. 主要構図データを用いた検索処理の高速化

ユーザーが描画したクエリー画像 Q の構図データ B_Q を算出し、各主要構図データ B_T との類似度 $S'_s(B_Q, B_T)$ を計算する。 $S'_s(B_Q, B_T)$ がしきい値以上となる B_T に対応した構図グループに属する画像のみを検索対象とする。このように、主要構図データを用いた絞り込みを行うことにより、画像の構図の類似性を重視した検索の高速化が可能となる。

4. 実用化に向けたシステムの例

4.1. 関連クリップ検索システム EN-Vision

EN-Vision[5]は、CLにおける「関連クリップ」検索サイトとして一般公開されており、あらかじめ付与されているキーワードやクリップの内容記述文を用いた言語的な検索および、「見た目」の類似性による検索機能を備えている。

見た目の類似性によるクリップ検索は、クリップの代表サムネイル画像を検索対象としたものであり、2章で述べたクエリー画像の描画による検索機能を含み、3章で述べた構図による高速化処理も適用している。

図3にEN-Vision画面の概観を示す。図3(a)は中央に表示された閲覧中のクリップに対する関連クリップ表示画面であり、赤枠で囲まれた部分がサムネイル画像の類似による関連クリップである。また、図3(b)は、クエリー画像の描画例である。

4.2. 番組映像検索システム

放送アーカイブを対象として、まず放送局内向けの実用化を目指し試作した多機能検索システムであり、キーワード検索以外に、様々なメディア解析技術に基づく以下の機能を有する。

- A) 簡単な画像描画によるショット検索 (図4(a))
- B) 顔の位置、カメラの動きの指定によるショット検索
- C) スーパー表示場面および顔アップ場面の一覧表示
- D) 番組の概要文の類似性による関連番組の表示
- E) 指定したショットとサムネイル画像の類似した関連ショットの表示 (図4(b))

A)には2章で述べた技術を適用している。放送アーカイブでは映像のジャンルが多種にわたるので、クエリー画像の描画画面においては、幾つかのジャンルごとのパレットを用意し、それらを切り替えて画像を描画できるようにした。

また、本システムでは現在のEN-Visionのサイトよりもさらに大量の映像データを扱うことを前提としている。そ

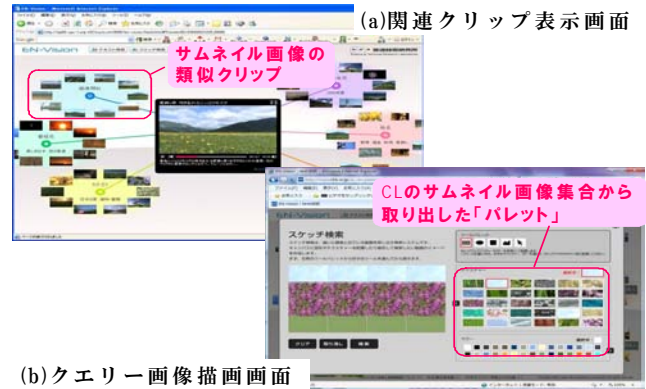


図3 EN-Visionの外観



図4 番組映像検索システムの外観

こで、A) および E) のサムネイル画像の検索処理においては、3章で述べた構図グループへの分類だけでなく、各構図グループに対して画像全体の色特徴による分類を行い更なる検索処理の高速化を図った。

5. あとがき

検索対象画像集合の主要な色やテクスチャを用いてクエリー画像を描画する画像検索手法と、主要構図データを用いて画像をあらかじめ分類することで検索処理を高速化する手法について述べた。また、実用化を目指し、本稿で述べた技術をシステムに導入した例を紹介した。

本稿で述べた検索技術は、サムネイル画像などの静止画像を対象としたものであるが、カメラの動きなどの動画特徴を用いた検索手法への拡張も必要である。更に、ショットなどの短い映像ではなく、番組全体などの長い映像の特徴を記述する手法を確立し、ジャンルやタイトルにとらわれない、「画作り」の似た“意外な”番組に出会える検索システムの実現も目指す。

参考文献

- [1] NHKクワイティブライブラリー <http://www.nhk.or.jp/creative/>
- [2] 望月, 蓼沼, 藤井, 伊藤: “データベース中の代表的なテクスチャと色を用いて作成した問合せ画像による画像検索,” 信学論, Vol. J88-D-II No.8, pp.1729-1739 (2005)
- [3] 望月, 藤井, 伊藤: “新しいフラクタル特徴とロバストな構図情報を用いた画像検索,” 映像誌, Vol.57 No.6, pp.719-728 (2003).
- [4] 望月, 住吉, 藤井: “クエリー画像描画による画像検索の構図テンプレートを用いた高速化,” 情報科学技術フォーラム(FIT2010) 講演論文集, 9(3), pp.219-220 (2010)
- [5] EN-Vision <http://www.nhk.or.jp/en-vision/>