

映像データベースとその類似検索手法の開発

柳沼 良知

近年、日本オープンコースウェア・コンソーシアムJOCWやインターネット教育支援サービスTIES等で教育映像のWeb配信が行われるようになってきた。このような教育映像のデータベース化を行う場合、特に、投稿型の映像データベースでは、いかに権利の問題がある映像を検出し、排除するかが重要になる。しかし、実際にYouTube等に投稿されている映像を見ると、同じ映像が横長になっていたり、映像が黒い枠に囲まれている場合があるなど、単純な類似検索手法では類似する映像の検索を正しく行えない場合がある。このため、本稿では、類似検索手法として、画像の類似度の計算に用いる色相のヒストグラムを作成する際に、映像に付加された黒枠の影響を受けにくくする処理を加えることで、検索性能を向上させる手法の開発を行った。また、評価実験を行い、提案手法が、映像に付加された黒枠の幅が大きい場合に有効であることを示した。

キーワード

教育映像、権利管理、映像データベース、類似検索、色ヒストグラム

1. はじめに

近年、日本オープンコースウェア・コンソーシアムJOCW (2008) やインターネット教育支援サービスTIES (2008) 等で教育映像のWeb配信が行われるようになってきた。このような教育映像のデータベース化を行う場合、映像の作成者がWebブラウザから直接映像を登録できれば、授業の進行に合わせて新しい映像を追加したり、映像を見せるか見せないかの設定を変えたりといった修正が行いやすいという利点がある。

しかし、実際にサービスが行われているYouTube (2008) のような投稿型の映像データベースでは、管理者を通さずに映像の登録が可能であることから著作権の問題が生じやすいという問題がある。また、特に教育映像の場合、例えば教室内の使用では問題のない映像でもそれをWeb上に公開すれば著作権の問題が生じる場合もある。このため、映像データベースでは、いかに権利の問題がある映像を検出し、排除するかが重要になる。しかし、投稿される映像の数が増えた場合、映像をすべて人が見て権利の問題の有無を判断するには多くの時間が必要となり、すべてを人手で行うのには限界があるため、検出処理の自動化が必要となる。

このような映像の検出に必要な類似画像検索手法としては、意匠・商標画像を対象とし、自動的に抽出さ

れた特徴点や対称軸といった情報の類似度により画像検索を行うTRADEMARK (加藤ほか 1989)、色相のベクトル間の角度により画像の類似度を計算して検索を行うシステム (Yaginuma et al. 1994)、形状、色、輪郭、物体と背景の配置等を組み合わせた検索が可能なQBIC (Flickner et al. 1995) 等がある。

しかし、実際にYouTubeに投稿されている映像を見ると、同じ映像が複数投稿されている場合、横長になっていたり、映像が黒い枠に囲まれている場合があるなど、単純な類似検索手法では検索を正しく行えない場合がある。

このため、映像データベースを実装するとともに、その類似検索手法の開発を行った。提案手法では、画像の類似度の計算に用いる色相のヒストグラムを作成する際に、黒枠の有無によるヒストグラムの変化を小さくする処理を加えることで、黒枠の有無にかかわらず類似した映像を検索することができるようにした。以下、その詳細について述べる。

2. 映像データベースの構築

2.1 データベース機能

映像データベースの本体としては、ユーザによる映像の投稿できること、RSSと呼ばれるメタデータで登録された映像の付加情報が管理されており、iPod等への映像配信にも対応していることから、JavaベースのBlogソフトであるBlojsom (2008) を用いた。

図1がBlojsomを用いて実装した映像データベースで



図1 映像データベース

ある。画面右側には検索窓があり、投稿された内容のキーワード検索を行うことができる。また、画面右側の「新規エントリー」をクリックすることで、投稿画面が開き、タイトルや説明文とともに、映像を投稿することができる。映像を視聴する場合は、登録されたそれぞれの項目の「Podcast:」の部分に書かれているファイル名をクリックすることで、PC等で対応する映像を視聴することができる。

2.2 映像検索機能

映像データベースの機能として、色特徴による検索機能を実現した。色特徴としては、画像のRGBを色相、明度、彩度に変換し、色相について32段階のヒストグラムを作成した(図2)。色相のヒストグラムは、左から右に赤、黄、緑、青、シアン、マゼンタに対応する。例えば、図2の上側の図は、草木が多いことからヒストグラムでは緑色に対応する部分にピークがある。図2下側の図では、青い背景に紫色のテロップがあるため、ヒストグラムには大きなピークが2つあり、右端のピークは紫色のテロップ、その左側のピークが青い背景に対応する。このような色相のヒストグラムの類似度により類似する画像の検索を行う。

画像間の類似度の尺度としては、2つの画像のヒストグラムを $h1(n)$ 、 $h2(n)$ とした場合、

$$\sum_{n=1}^N \frac{(h2(n) - h1(n))^2}{h1(n)}$$

を計算し、この値が小さいほど画像は類似しているとした。Nはヒストグラムの段数である。この式は映像からの場面の変わり目を検出するシーンチェンジ検出(長坂ほか 1992)等で用いられる式である。

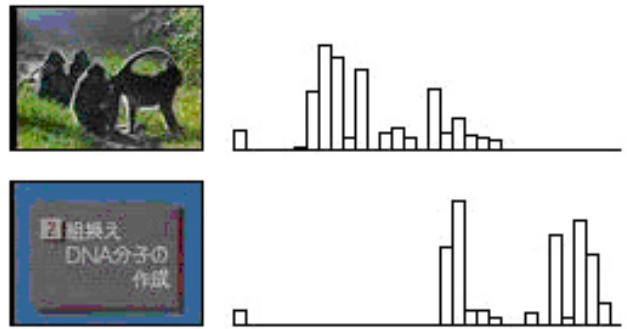


図2 色相のヒストグラムの例

類似度の計算結果をもとに、それぞれの画像に対する類似検索結果のHTMLファイルを予めすべて作成することで、画像検索機能を実装した(柳沼ほか 2007)。図3は、左上の3次元CG画像に類似する画像を検索した結果であり、上位20枚のうち17枚が同じ映像に含まれる画像であった。表示されている画像をクリックすることで、クリックした画像の類似検索結果を表示することができる。また、それぞれの画像の下の「Open HTML」のアイコンをクリックすることで、対応するBlojsomのページを表示することができる。

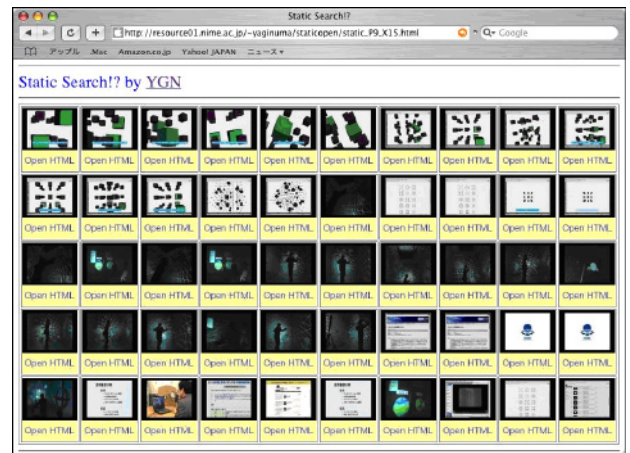


図3 色特徴による検索

3. 類似検索手法

3.1 類似検索手法

本章では、2章で述べた映像データベースの一機能として開発を行った類似検索手法について述べる。類似検索手法の開発をするにあたり、まず、実際に稼働している投稿型映像データベース1つであるYouTube上の映像について、その特徴の解析を行った。

YouTubeに投稿されている映像を見ると、同じ映像が複数投稿されている場合、必ずしも、すべて画面が同じとは限らず、同じ映像でも、図4(a)のように横長になっていたり、図4(b)のように映像が黒い枠に囲まれてい

る場合があった。また、枠の幅も必ずしも一定ではなかった。これは、もとの映像の表示方法や取り込み装置の特性によるものと考えられ、映像の権利管理を考える場合には、このような画像の縦横比の変更や黒枠が付加されることを考慮した類似検索ができる必要がある。

この問題を解決するため、提案手法では、黒枠がある可能性のある範囲を設定し、もしそこに含まれる画素が黒である場合に、その画素を用いないで色相のヒストグラムの作成を行う。これにより、黒枠がある場合でも、ない場合でも、色相のヒストグラムがほぼ同じになり、黒枠の有無にかかわらず類似した映像を検索することができるようになる。

実際に、YouTubeに投稿された映像100本について調べたところ、黒枠の幅の分布は図5のようになった。映像については、音楽ビデオは同じ映像が複数投稿される場合が多いと考え、5名のアーティスト名で検索を行い、それぞれの検索結果の上位20件を選ぶことで100本の映像を選択した。320×240の大きさの映像に対して、横方向の黒枠の幅(W)が0であった映像が65本、縦方向の黒枠の幅(H)が0であった映像が54本であった。これらのうち、横方向にも縦方向にも黒枠が存在しない映像は40本あった。黒枠が無い映像の割合が多いため、横方向の黒枠の幅(W)の平均は1.6ピクセル、縦方向の黒枠の幅(H)の平均は11.3ピクセルとなったが、3.3で述べるように提案手法が従来手法に対して検索性能が高くなる黒枠の幅が24ピクセル以上に該当する映像は29本(全体の29%)あった。また、縦方向の黒枠の幅(H)の大きさは最大で55ピクセルであった。また、横方向の黒枠の幅(W)の最大値は25ピクセルであった。このため、黒枠がある可能性のある範囲は、 $H \leq 55$, $W \leq 25$ と設定した。

具体的な色相のヒストグラム作成の手順は、以下のようになる。

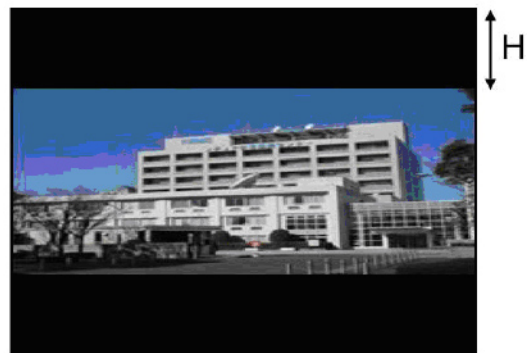
- (1) 処理時間を短縮するため、画像を一定の大きさに縮小する。(以下の実験では、画像を80×60に変換し処理を行った。)
- (2) それぞれの画素のRGBを色相、明度、彩度に変換し、色相のN段階のヒストグラムを作成する。(ヒストグラムの段数は32とした。)ただし、黒枠がある可能性のある範囲 ($H \leq 55$, $W \leq 25$) の画素については、明度を計算し、明度が256階調で5未満ならその画素をヒストグラムの作成に用いないこととした。
- (3) 黒枠がある場合、ヒストグラムの作成には黒枠に含まれる画素は使われないため、ヒストグラム全体の画素数が変わってしまう。このため、全体の画素数が一定になるようにヒストグラムの値を規格化する。具体的には、作成したヒストグラムを $h1_{org}(n)$ 、規格化したヒストグラムを $h1(n)$ として、

$$h1(n) = h1_{org}(n) \cdot \frac{80 \times 60}{\sum_{n=1}^N h1_{org}(n)}$$

と規格化した。Nはヒストグラムの段数である。

上の手順のうち、(2)の「ただし、」以下の部分が提案手法で追加した点であり、この部分を省き、すべての画素を使ってヒストグラムの作成を行うのが従来手法となる。

このようにして作成された色相のヒストグラムを用いた検索については、従来手法、および、提案手法のいずれも2.2の検索手法と同様に行うことができる。



(a) 縦横比の変化



(b) 枠の付与

図4 映像の変化

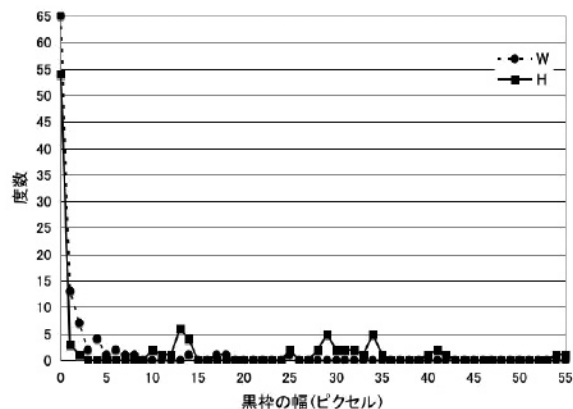


図5 黒枠の幅の分布

3.2 シーンチェンジ検出

提案手法の性能評価のため、検索実験を行った。実験対象とした映像は、生物に関するビデオ教材25本、新素材に関するビデオ教材22本の計47本、全体で約23時間の映像である。映像の大きさは、320×240である。この映像から1秒ごとに、全体で83811枚の画像を抜き出した。しかし、画像のデータ量は大きく、それらをそのまま扱うことは効率的ではない。また、これらの画像の中で、例えば、講師が説明をしているような場面では、ほとんど画像が変わらないため、単純にこれらすべての画像を対象に類似検索したのでは、検索結果に類似した画像が多くなり過ぎ、必要な画像を効率的に探すことが難しくなる。このため、場面の変わり目（シーンチェンジ）の検出を行い、画像枚数を1/10の8381枚に減少させた。シーンチェンジ検出は、2.2の式により連続する2枚の画像の類似度を計算し、類似度が小さい画像順に8381枚の画像を選択した。

処理時間については、CPU 2.5GHz Dual, メモリ 1.5GBの環境で測定を行った。なお、320×240の大きさの画像で処理を行った場合、処理時間が長くなるため、処理プログラムの内部で画像を80×60の大きさに縮小し、処理を行っている。色相のヒストグラムを作成する際、提案手法では、黒枠の部分をヒストグラムの作成から省く処理が加わるものの、明度のみを計算して黒枠と判断された場合にはその画素のRGBを色相、明度、彩度に変換する処理が必要なくなるため、表1に示すように、従来手法、提案手法のいずれも1986秒程度で色相のヒストグラムの作成を行うことができた。シーンチェンジ検出には、従来手法、提案手法のいずれも88秒程度で処理を行うことができた。このように、検索の前処理には全体で30分程度の時間がかかっているが、これらの処理は最初に1度だけ実行すれば良い。

3.3 検索性能

検索性能を評価するため、8381枚の画像の中から検索キーとなる画像を選び、その画像に黒枠を加えて検索実験を行った。黒枠は図4(a)のように画像の上下に付加し、その中央にもとの画像を横長にして貼付けた。図6が、黒枠の幅を変えた場合の検索性能である。横軸は、320×240の大きさの画像に対する黒枠の幅（ピクセル）である。検索性能の指標としては、黒枠を加える前の画像が何番目に検索されたかの平均値を用いた。図6の縦軸は、100枚の画像に対して検索を行った場合の平均順

表1 処理時間

	従来法	提案法
色相のヒストグラムの作成 (秒)	1987.3	1985.0
シーンチェンジ検出 (秒)	88.0	87.0

位である。

図6を見ると、黒枠の幅が24ピクセル、すなわち画像の縦方向の1割までの場合は、従来手法も提案手法もほぼ1番目に黒枠を付加する前の画像を検索することができている。しかし、黒枠の幅がそれ以上大きくなった場合、従来手法では、平均順位が大きく増加していることが分かる。例えば、黒枠の幅が56ピクセルの場合、従来手法では平均順位が89.3であるが、データ数が今回の実験の10倍の規模とした場合、893程度の検索結果を見なければもとの画像を見つけ出すことができず、類似検索に従来手法を利用するのは現実的ではないことが分かる。一方、提案手法では、黒枠の幅が56ピクセルの場合でも平均順位が3.1と性能が悪くならず、提案手法は、黒枠が大きい場合に有効な手法であることが分かる。

なお、明度5未満の画素をヒストグラムの作成に用いないことから、例えば、全体的に明度が低く黒っぽい画像を扱う場合、ヒストグラムの作成がうまくいかず結果として類似度の判定が難しくなることも考えられるが、提案手法では、

- 1) 画像の周辺部分の予め設定した範囲に含まれる明度5未満の画素のみをヒストグラムの作成から除外し、周辺部分に比べて重要と思われる画像の中央部分については、明度5未満の画素についてもヒストグラムの作成に用いている
- 2) 黒っぽい画像については、類似していれば同じように周辺部分の画素がヒストグラムの作成に用いられず、類似した画像のヒストグラムの変更が同じように行われるため、結果として類似した画像のヒストグラムは類似したものになる

ことから、黒っぽい画像であっても検索性能が大きく低下することは無いと考えられる。

実際に検索性能の評価に用いた8381枚の画像のうち、明るさを黒が0、白が255の256階調で表した場合に、明るさの平均が白の1/4以下となる黒っぽいと考えられる画像の枚数が1877枚 (22.4%)、検索キーとして選んだ100枚の画像のうち同様に黒っぽいと考えられる画像の枚数が22枚 (22%)あり、これらを含んだ検索結果である図6を見ると、提案手法の検索性能が従来手法に比べて悪くなる傾向は見られず、提案手法は、黒っぽい画像が含まれている場合でも適用できると考えられる。

処理時間については、CPU 2.5GHz Dual, メモリ 1.5GBの環境で測定を行った。表2に示すように、100回検索を行った場合の検索時間は従来法、提案手法のいずれも19秒程度であった。これは、1回の検索あたり0.2秒以下であり、十分高速に処理が行えていることが分かる。

3.4 ヒストグラムの段数の最適化

提案手法では、類似度の計算に色相のヒストグラムを用いている。このヒストグラムの段数が小さい場合には、

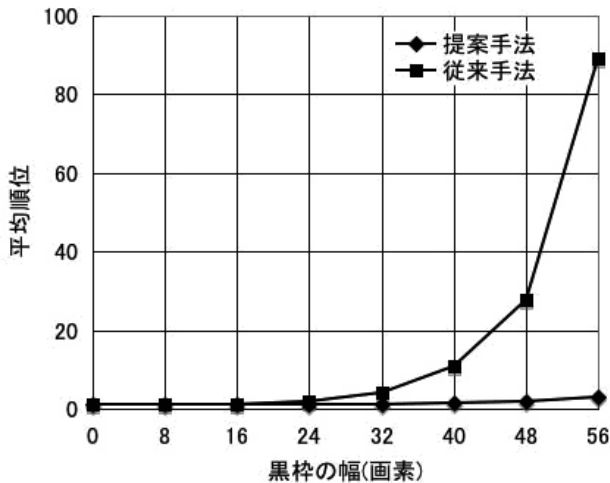


図6 黒枠の幅を変えた場合の検索性能

表2 検索時間

	従来法	提案手法
100回の検索時間 (秒)	19.3	19.0

画像ごとのヒストグラムの差があまり出ないため、画像の分類精度が悪くなると考えられる。一方で、ヒストグラムの段数を大きくしすぎると、色ムラなどによりヒストグラムの境目近くの色が隣のヒストグラムに加算されることで、本来類似している画像の類似度が小さく計算されてしまうという問題が生じやすくなる。このため、実験で用いた映像を対象にヒストグラムの段数の最適値を求めた。

図7が黒枠の幅を実際の映像の最大値であった55ピクセルに設定し、ヒストグラムの段数を変えて類似検索を行った場合の検索性能である。(なお、処理プログラムでは、画像を80×60に縮小して処理を行っているため、黒枠の幅が55ピクセルの場合の平均順位は、黒枠の幅が52ピクセルの場合の平均順位と同じになる。)図7の横軸はヒストグラムの段数である。縦軸は、検索性能の指標として、100回の検索を行った場合の、もとの画像の平均順位を示している。この結果、ヒストグラムの段数が32の場合に最も平均順位が小さくなっており、検索性能が高くなっていることが分かる。このため3.2, 3.3での評価実験では、ヒストグラムの段数を32として評価実験を行った。

4. まとめ

本稿では、映像の類似検索手法として、画像の類似度の計算に用いる色相のヒストグラムを作成する際に、映像に付加された黒枠の影響を受けにくくする処理を加えることで、検索性能を向上させる手法の開発を行った。また、評価実験を行い、320×240の画像に対して上下

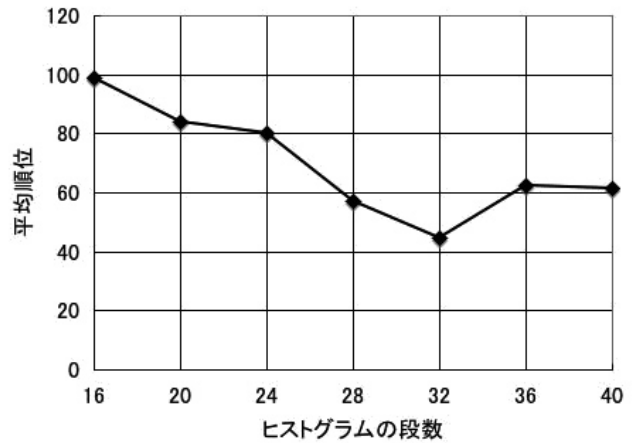


図7 ヒストグラムの段数の最適化

に付加された黒枠の幅が56ピクセルの場合、従来手法では、類似画像が検索される平均順位が89.3であるのに対し、提案手法では類似画像を平均順位3.1で検索できることを示し、提案手法が、映像に付加された黒枠の幅が大きい場合に有効な手法であることを示した。

今後は、登録された映像の色特徴だけではなく、登録の際に付与された関連する文字情報等を統合して処理を行うなど、検索精度を更に向上させる手法について検討を行っていく予定である。

引用文献

- Blojsom (2008). <http://www.blojsom.com/>
- Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D. and Yanker, P. (1995). Query by Image and Video Content: The QBIC System, IEEE Computer, pp.23-32.
- インターネット教育支援サービスTIES (2008). <http://www.tiesnet.jp/>
- 加藤俊一, 下垣弘行, 藤村是明 (1989). 画像対話型商標・意匠データベース TRADEMARK, 電子情報通信学会論文誌, **J72-D-II(4)**, 535-544.
- 長坂晃朗, 田中 譲 (1992). カラービデオ映像における自動索引付け法と物体探索法, 情報処理学会誌, **33-4**, 543-550.
- 日本オープンコースウェア・コンソーシアムJOCW (2008). <http://www.jocw.jp/>
- Yaginuma, Y. and Sakauchi, M. (1994). Multi-Purpose Interface for Still/Moving Image Retrieval, **SPIE94**, 260-267.
- 柳沼良知, 鈴木一史, 清水康敬 (2007). 色特徴とキーワードに基づく教育画像コンテンツの検索, 電子情報通信学会技術研究報告, **106-507**, 111-116.
- YouTube (2008). <http://www.youtube.com/>



やぎぬま よしとも
柳沼 良知

昭63東大・工・計数卒。平2同大大学院修士課程（理・物理）了。同年東大生産技術研究所技官。平5同助手。平12メディア教育開発センター助教授。平13より総合研究大学院大学助教授併任。工博。マルチメディア処理, マルチメディアデータベースなどの研究に従事。

Development of Movie Database and its Similarity Retrieval Method

Yoshitomo Yaginuma

University consortiums such as JOCW and TIES have started Web delivery of educational movies. In such movie databases, right management of movies has to be regarded. When movies are posted to movie databases, aspect ratios of some movies are modified, and black borders are added to some movies. Simple retrieval method using color histograms is not enough to retrieve such movies. Therefore, in this paper, a new similarity retrieval method of movies is proposed. The proposed method reduces the effect of black borders added to movies when color histograms are calculated. Experimental results revealed that the proposed method is effective to retrieve similar movies especially when the black borders added to movies are large.

Keywords

Educational Movie, Right Management, Movie Database, Similarity Retrieval, Color Histogram