

輝度分布の相関と色空間情報を利用した動画像検索方式の検討

高橋 秀和 児玉 明 金田 和文 山下 英生

広島大学

〒739-0046 東広島市鏡山 3-10-31

TEL : (0824) 21-3646

FAX : (0824) 21-3639

E-mail : hide@eml.hiroshima-u.ac.jp

あらまし 本稿では、画像検索としての動画像マッチング手法において、時間特徴量と空間特徴量の2つに着目した。時間特徴量を利用したマッチング手法として、輝度値の頻度分布に着目し、フレーム前後間の輝度分布の相関値を用いた。また、空間特徴量を利用したマッチング手法として、XYZ色空間を用い、さらに、縮小画像を用いることによって高速化を図った。最後に、それら2つの手法を組み合わせることでマッチングを行うことにより、輝度分布の相関を用いた手法の空間方向への拡張を行った。本提案手法について、検索速度と精度の両方の面からの有用性について検討を行った。

キーワード 動画像検索手法、輝度分布、相関、色空間、高速化アルゴリズム

A Study on Video Matching Methods using the Correlation of the Luminance Variance and Color Space

Hidekazu TAKAHASHI Mei KODAMA Kazufumi KANEDA Hideo YAMASHITA

HIROSHIMA University

3-10-31, Kagamiyama, Higashi-Hiroshima,
739-0046 JAPAN

TEL : +81-824-21-3646

FAX : +81-824-21-3639

E-mail : hide@eml.hiroshima-u.ac.jp

Abstract In this paper, we pay attention to temporal characteristics and spatial characteristics in video matching methods as image searching system. As temporal characteristics, we used the luminance variance and correlations of the variance. The other side, as spatial characteristics, we made use of XYZ color space and we aimed to fast matching algorithm using downsampled images. Last, we proposed method combined these two schemes. The proposed method were efficient by simulation results.

key words Video Searching Method, Luminance Variance, Correlation, Color Space, Fast Algorithm

1. はじめに

現在、膨大なコンテンツ情報の中からユーザが自分の望む情報を手に入れるためには、検索システムが必要である^[1]。また、近年のコンピュータ処理能力の進歩により、特に言葉では表現しきれない情報に対して、動画像を扱うことが増えてきている。そこで、動画像検索システムにおける画像マッチングの手法について考える。

本稿では、時間特徴量を利用したマッチング手法として、輝度値の頻度分布に着目し、フレーム前後間の輝度分布の相関値を用いる。また、空間特徴量を利用したマッチング手法として、XYZ色空間を用い、さらに、縮小画像を用いることにより高速化を図る。最後に、それら2つの手法を組み合わせることでマッチングを行うことにより、輝度分布の相関を用いた手法の空間方向への拡張を行う。そして、提案手法について、検索速度と精度の両方の面からの有用性を示す。

2. 時間特徴量を利用したマッチング手法

時間特徴量を用いたマッチング手法として、輝度値の頻度分布とフレーム前後間の輝度分布の相関値を利用する^[2]。

2.1 検索方式の概要

本手法は、次の各処理からなる。

- 輝度分布の算出
まず、入力データ、データベースのデータそれぞれに対して輝度値 Y についての頻度分布を求める。頻度分布は各フレーム全体の特徴を表す。ここで、輝度値は 8bit 精度で扱う。
- 相関値の算出
次に、求めた頻度分布を用いて、フレーム前後間の分布の相関値を求める。 i 番目のフレームの頻度分布を x 、 $i+1$ 番目のフレームの頻度分布を y とすると、相関値 C は式 (1) で求まる。

$$C = \frac{\sum_{j=0}^{255} (x_j - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_{j=0}^{255} (x_j - \bar{x})^2 \sum_{j=0}^{255} (y_j - \bar{y})^2}} \quad (1)$$

ここで、 \bar{x} , \bar{y} は、平均値である。

以上で求めた相関値を用いて、マッチング処理を行う。そのマッチング手法について次節で述べる。

2.1.1 分布相関値マッチング法

分布相関値マッチング法は、入力データの輝度分布相関値の波形のうち、入力フレーム数分の波形と比較する方法で、この操作を、入力相関値の波形を1つずつシフトさせて、データベースのすべての相関値の波形に対して行う。つまり、データベースの相関値の波形の中で、入力相関値の波形と似たものを探す処理である。

2.1.2 相関値特性マッチング法

相関値特性マッチング法は、分布相関値が1に近づく頻度が高いという傾向を利用した手法である。したがって、相関値が1から離れる部分を時間的特徴量とすることにより、動画のマッチングに利用する。図1に示すように、分布相関値があるしきい値より高いか低いかで相関値をグループ化し、しきい値より低い場合には、さらにそのグループの代表極小値を用いてマッチングを行う方法である。具体的には、まず、分布相関値がしきい値より連続して高い部分のフレーム数 (H)、連続して低い部分のフレーム数 (L)、しきい値より低い領域における相関値の極小値 (C) をテーブル化し、そのテーブルの比較を、入力テーブルにおける最大の H とその前後の L 、そしてその L の領域における C の5段階で行う。

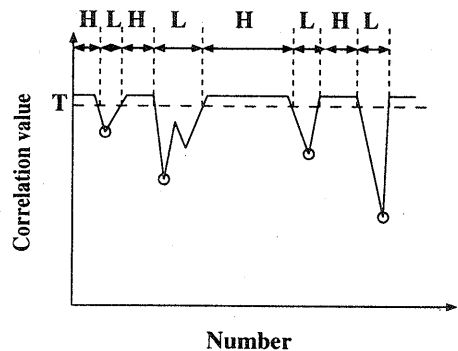


図 1: テーブル化

3. 空間特徴量を利用したマッチング手法

時間特徴量を利用したマッチング手法では、検索時間の高速化は図れるが、フレーム前後間の輝度分布の相関値だけを用いてマッチング処理を行って

るため、精度の面で問題点がある [2]。そこで、精度の向上を目的として、空間特徴量を利用したマッチング手法について考える。

3.1 XYZ 色空間

空間特徴量として、XYZ 色空間を用いた。YCbCr 色空間から XYZ 色空間への変換は、式 (2)、(3) による。

$$\begin{aligned} R &= \frac{Y-16}{219} + \frac{Cr-128}{160} \\ G &= \frac{Y-16}{219} - \frac{0.114}{0.587} \frac{Cb-128}{126} - \frac{0.299}{0.587} \frac{Cr-128}{160} \\ B &= \frac{Y-16}{219} + \frac{Cb-128}{126} \end{aligned} \quad (2)$$

$$\begin{aligned} X &= 2.7689R + 1.7517G + 1.1302B \\ Y &= 1.0000R + 4.5907G + 0.0601B \\ Z &= 0.0565G + 5.5943B \end{aligned} \quad (3)$$

入力画像とデータベース画像それぞれに対して、YCbCr 空間から XYZ 空間への変換を行い、その特徴量を式 (4) で比較する。

$$(X_{in} - X_{db})^2 + (Y_{in} - Y_{db})^2 + (Z_{in} - Z_{db})^2 < Th \quad (4)$$

ただし、 X_{in}, Y_{in}, Z_{in} は、入力画像の XYZ の各成分、 X_{db}, Y_{db}, Z_{db} は、データベース画像の XYZ の各成分である。式 (4) を満たすものをカウントし、そのカウント数が式 (5) を満たすとき、マッチング判定を行う。

$$count > Th_{count} \quad (5)$$

3.2 縮小画像を用いたマッチング

空間特徴量を用いたマッチング手法は、各画素の成分について比較を行うので、時間特徴量を用いたマッチング手法に比べてマッチング時間がかかると考えられる。そこで、縮小画像を用いることによって、マッチングに用いる空間特徴量のデータ量を削減し、マッチング時間の高速化を図る。

今回、縮小画像を作成する際に用いたダウンフィルターには、ラグランジェの補間公式 (式 (6)) より導き出した 4 タップ $[-1, 9, 9, -1]$ を利用している。

$$\begin{aligned} L(t) &= a_0(t-t_1)(t-t_2)\cdots(t-t_{n-1}) \\ &\quad + a_1(t-t_0)(t-t_1)\cdots(t-t_{n-1}) + \cdots \\ &\quad + a_i(t-t_0)\cdots(t-t_{i-1})(t-t_{i+1})\cdots(t-t_{n-1}) \\ &\quad + \cdots + a_{n-1}(t-t_0)(t-t_1)\cdots(t-t_{n-2}) \\ &= \sum_{k=0}^{n-1} \left\{ a_k \prod_{\substack{i=0 \\ i \neq k}}^{n-1} (t-t_i) \right\} \end{aligned} \quad (6)$$

3.3 シミュレーション実験

3.3.1 実験方法

表 1: データベース画像データ

シーケンス	アニメ, スポーツ 1, スポーツ 2, 音楽 1, 音楽 2, 映画, ニュース, スポーツ 3, スポーツ 4, スポーツ 5
データベースシーケンス数	10 シーケンス
データベースフレーム数	各シーケンス 9000 枚 計 90000 枚
画像フォーマット	ITU-R Rec.601 4:2:0 704[pel]x480[line]

今回用いたデータベースのデータを表 1 に示す。データベースには、表 1 にあるシーケンスが、アニメ、スポーツ 1、...、スポーツ 5 の順で入っている。

入力はデータベース内の静止画 1 枚 (スポーツ 1 の中の 1 枚) を用い、704x480、352x240(1/4)、176x120(1/16)、88x60(1/64)、44x30(1/256) の 5 種類の画像サイズのものを用いて、データベースの全ての画像とマッチング処理を行った。

なお使用計算機は、Fujitsu GP7000SM25 400MHz UltraSPARC-2 である。

3.3.2 実験結果

実験結果を、時間的、精度的側面から評価する。なお、式 (4)、(5) にあるしきい値 Th 、 Th_{count} は、それぞれ 1.0、全画素数の 80% とする。

表 2: 空間特徴量を用いたマッチング手法

size	time [s]	matching pattern
704x480	34631.54	11
352x240	9129.53	11
176x120	2484.12	10
88x60	1034.11	23
44x30	476.85	55

5 種類の各画像サイズにおけるマッチング処理に要した時間とマッチングパターンの結果を表 2 に示す。

表 2 から分かるように、704x480 の画像サイズは 44x30 の画像サイズに比べて約 70 倍マッチング時間に要している。精度の面から見ると、44x30

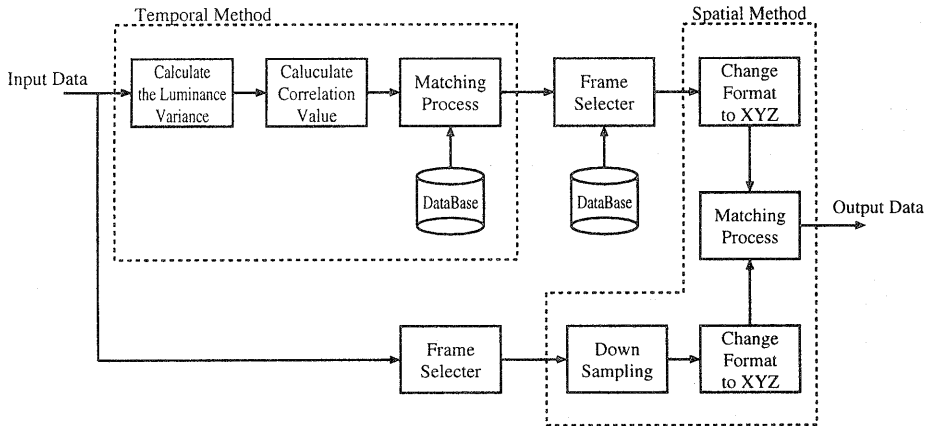


図 2: 輝度分布の相関と色空間情報を用いたマッチング手法の概要

の画像サイズでも 90000 枚中の 55 枚であり、マッチング時間が 70 倍高速化されたのに対し、マッチング枚数は 5 倍にしか増えていない。352x240 や 176x120 では、704x480 とほとんど変わらない精度が得られている。以上のことから、空間特徴量を利用したマッチング手法では、縮小画像を用いることが有効であると考えられる。

4. 時間・空間特徴量を用いたマッチング手法

時間特徴量を用いた手法、空間特徴量を用いた手法それぞれについて考えてきたが、時間特徴量を用いた手法では精度の面で、また、本稿での空間特徴量を用いた手法では速度の面でそれぞれ問題点があった。これらの問題点を解決する方法として、以上の 2 つの手法を組み合わせることによって、速度、精度両方の良い部分を活用することができるのではないかと考えられる。

4.1 提案マッチング手法の概要

輝度分布の相関と色空間情報を用いたマッチング手法の概要を図 2 に示す。本マッチング手法は、次の各処理からなる。

- 時間特徴量を利用したマッチング
 2. 節で述べたように、入力 N 枚の静止画に対して、次の各処理を行う。
 1. 輝度分布の算出
 2. フレーム前後間の相関値の算出

3. 分布相関値マッチング法または、相関値特性マッチング法を用いたマッチング処理

- 空間特徴量を利用したマッチング

時間特徴量を用いたマッチング処理で得られる結果は、入力 N 枚の静止画に対して、静止画 N 枚単位の動画である。この結果に対して、次の各処理を行う。

1. 代表フレームの選択

入力画像 N 枚の最初と最後のフレームを選択し、また得られた結果 N 枚の最初と最後のフレームの縮小画像をデータベースから選択する。
2. 縮小画像の作成

選択された入力画像の 2 枚のフレームについて縮小画像を作成する。空間特徴量を利用したマッチング処理では、高速化のため、これらの 2 枚に対してのみマッチング処理を行う。
3. XYZ 色空間への変換

それぞれの縮小画像に対して、色空間を YCbCr から XYZ へ変換する。
4. マッチング処理

3. 節で述べたマッチング方法によりマッチング処理を行う。

以上が、提案マッチング手法の処理の流れである。

時間特徴量を利用したマッチング処理においては、高速なマッチングは可能である [2]。その時間特

微量を利用したマッチング処理で得られた結果の、最初と最後の2枚のフレームだけに空間特徴量を利用したマッチング処理を行うことにより、時間特徴量を利用したマッチング手法だけでは得ることができなかった高精度な結果を得ることができると考えられる。また、処理に時間を要する空間特徴量を利用したマッチング処理は2フレームだけであるので、マッチング時間の面から見ても、この手法は有効であると考えられる。

4.2 シミュレーション実験

4.2.1 実験方法

表 3: 入力画像データ

シーケンス	アニメ, スポーツ 1, スポーツ 2, 音楽 1, 音楽 2, 映画, ニュース, スポーツ 3, スポーツ 4, スポーツ 5
入力データ数	10 シーケンス
入力フレーム数	90 枚
画像フォーマット	ITU-R Rec.601 4:2:0 704[pe]x480[line]

用いたデータベースの画像は、表 1 に示すものと同じであり、また用いた入力画像は、表 3 に示す。これらの条件のもと、時間特徴量を利用したマッチング手法で、分布相関値マッチング法と相関値特性マッチング法を用い、また、空間特徴量を利用したマッチング手法では、それぞれ 704x480 と 44x30 の画像サイズを用いたときについてシミュレーション実験を行った。そして、時間特徴量を利用したマッチング手法のみを用いた場合と、組み合わせによる手法を用いた場合について、マッチングに要した時間とマッチングパターンをそれぞれ測定した。

4.2.2 実験結果

時間特徴量を利用したマッチング手法で、分布相関値マッチング法を用いた実験結果を表 4 に、また、相関値特性マッチング法を用いた実験結果を表 5 にそれぞれ示す。

4.2.3 考察

マッチング手法において時間・空間といった複数方向の特徴量を用いることは、マッチング時間、マッチング精度の面から見て有効である。本稿では、輝度分布の相関を用いた手法によるマッチング

処理を、さらに色空間情報を用いることにより、空間方向へ拡張した。

時間特徴量を用いた手法では、表 4, 5 に示すように、分布相関値マッチング法で約 1.6[s], 相関値特性マッチング法で約 0.4[s] の結果が得られており、表 2 に示す空間特徴量を用いる場合と比べて、格段に高速なマッチングが行える。しかし、時間の特徴としてフレーム前後間の相関値のみを用いているため、精度の面では、図 3 に示すような時間の特徴の少ない入力に対して、良い結果が得られていない。この傾向は、特に相関値マッチング法に現れている。このように、時間特徴量を用いた手法では、入力に依存するという問題点があった [2]。

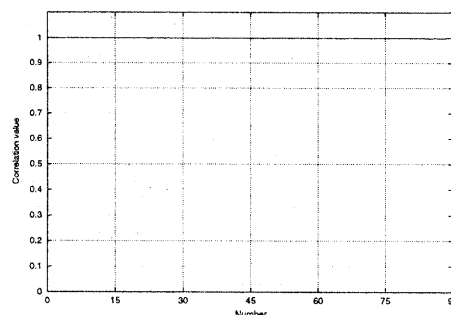


図 3: フレーム前後間相関値 (入力データ: スポーツ 1)

一方、空間特徴量を用いた手法では、表 2 に示されているように、精度は良いけれども、マッチングにかなりの時間を要するため、この手法のみでは検索に用いるのは不適であると考えられる。

これら 2 つの短所を、時間特徴量を用いた手法を空間方向へ拡張することによって補うことができる。表 4, 5 に示すように、時間特徴量だけを用いたときには、入力によっては多くのマッチングパターンが得られていたが、組み合わせの手法を用いた結果、マッチングパターンはほとんど 1 になっている。したがって、入力に依存しないマッチング結果を得ることができる。マッチング時間に関しては、空間特徴量を用いた手法でマッチングを行う枚数が多いものは、その分処理に時間を要している。

しかし、空間特徴量の比較を行う際に 44x30 の縮小画像を用いた場合は、704x480 のサイズを用いた場合に比べて、精度を保ったままで全体的に高速化が成されている。またこのマッチング時間には、比較に用いる入力画像 2 枚のダウンサンプリング

表 4: 分布相関値マッチング法の結果

name	時間特徴量のみ		組み合わせ			
			704x480		44x30	
	time [s]	matching pattern	time [s]	matching pattern	time [s]	matching pattern
アニメ	1.679	1	2.440	1	2.270	1
スポーツ 1	1.638	130	49.180	1	2.990	1
スポーツ 2	1.576	11	5.880	1	2.290	1
音楽 1	1.614	1	2.330	1	2.230	1
音楽 2	1.719	1	2.430	1	2.320	1
映画	1.589	6	4.090	1	2.210	1
ニュース	1.693	1	2.430	1	2.300	1
スポーツ 3	1.645	1	2.400	1	2.250	1
スポーツ 4	1.693	11	6.360	2	2.390	2
スポーツ 5	1.638	1	2.410	1	2.260	1

表 5: 相関値特性マッチング法の結果

name	時間特徴量のみ		組み合わせ			
			704x480		44x30	
	time [s]	matching pattern	time [s]	matching pattern	time [s]	matching pattern
アニメ	0.362	1	1.100	1	0.970	1
スポーツ 1	0.414	207	75.020	1	2.240	1
スポーツ 2	0.370	1	1.040	1	0.960	1
音楽 1	0.366	2	1.450	1	0.990	1
音楽 2	0.367	2	1.430	1	1.050	1
映画	0.367	1	1.150	1	1.000	1
ニュース	0.411	207	75.560	1	2.210	1
スポーツ 3	0.372	43	16.210	1	1.280	1
スポーツ 4	0.374	11	4.720	1	1.010	1
スポーツ 5	0.372	1	1.140	1	0.980	1

グに要した時間も含まれており、縮小画像を用いることの有用性を示している。ただ、今回の縮小画像での比較は、データベースに同じく縮小画像が存在していると仮定して、縮小画像と縮小画像で行ったものであり、そのために704x480を用いた場合と比べて、精度に変化が見られなかったと考えられる。したがって、データベースに704x480のサイズの画像しかない場合に、どのように縮小画像を利用するかという点は今後の課題である。

また、今回の手法では、動画像を静止画像の連続として考えており、もし入力符号化されたデータであった場合は、復号して空間情報に戻す必要がある。符号化されたデータをそのまま扱うにはどのようにしたらよいかという点も、今後の課題である。

5. まとめ

動画像マッチング手法において、時間特徴量を利用した手法としてフレーム前後間の相関値を用いた

手法、空間特徴量を利用した手法としてXYZ色空間を用いた手法、またそれら2つの方法を組み合わせた手法について述べた。時間特徴量を用いた手法を空間方向へ拡張することによって、マッチング時間を、マッチング精度両方に効果的であることを示した。今後の課題としては、縮小画像の効果的な利用法と、符号化されたデータを扱う場合への拡張が挙げられる。

参考文献

- [1] 高橋 克直, 寺島 信義, 富永 英義: “画紋情報を用いた動画像検索手法に関する検討”, 映像学技報, BCS98-61, pp. 1-8 (1998).
- [2] 児玉 明, 高橋 秀和, 竹本 正行, 池田 朋二, 真崎 剛, 金田 和文, 山下 英生: “輝度分布の相関を用いた動画像検索手法の検討”, 信学技報, IE99-117, CS99-137, pp. 73-78 (1999).