

輝度分布の相関を用いた動画像検索手法の検討

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[†]

[†] 広島大学

[‡] 株式会社 佐竹製作所

〒 739-0046 東広島市鏡山 3-10-31

〒 739-0046 東広島市鏡山 3-13-26-303

TEL : (0824) 21-3646

TEL : (0824) 20-0545

FAX : (0824) 21-3639

FAX : (0824) 20-0503

E-mail : mei@ipc.hiroshima-u.ac.jp

あらまし 本稿では、画像検索として動画像マッチング手法に着目した。輝度値の頻度分布に着目し、動画像情報としての特徴量である、時間的変化量を利用したマッチング手法を提案した。特に、フレーム前後間の分布相関値を求めることにより、前後のフレーム間の特徴量を比較したマッチング手法について検討した。さらに、動画像情報の相関値の偏った頻度特性を利用して、しきい値による一定相関値のグループ化及びその時間的前後の変化量との2次元情報によるマッチングアルゴリズムを提案した。本提案方式についてシミュレーション実験を行い、マッチングの高速化が可能であることを示した。

キーワード 動画像検索手法, 輝度分布, 相関, 高速化アルゴリズム, 検索時間

Video Searching Methods using the Coorelation of the Luminance Variance

Mei KODAMA[†] Hidekazu TAKAHASHI[†] Masayuki TAKEMOTO[‡] Tomoji IKEDA[‡]
Tsuayoshi MASAKI[‡] Kazufumi KANEDA[†] Hideo YAMASHITA[†]

[†]HIROSHIMA University

[‡]SATAKE Corporation

3-10-31, Kagamiyama, Higashi-Hiroshima, 3-13-26-303, Kagamiyama, Higashi-Hiroshima,
739-0046 JAPAN

TEL : +81-824-21-3646

TEL : +81-824-20-0545

FAX : +81-824-21-3639

FAX : +81-824-20-0503

E-mail : mei@ipc.hiroshima-u.ac.jp

Abstract In this paper, we pay attention to video matching methods as image serching system. Especially, frecuenys of luminance signals are used. In this paper, we proposed the video matching methods using temporal variances as image characteristics. And then, we verified these schemes on the matching process and matching pattern. The proposed methods were efficiet by simulation results. From the viewpoint of matching process, the grouping method was available in our mothods.

key words Video Searching Method, Luminance Variance, Correlation, Fast Algorithm, Matching Time

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][2]。その中でも、情報伝達に最も有効であるのは、画像、特に動画の検索システムである。そこで、動画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、輝度値の頻度分布に着目し、特にフレーム前後間の輝度分布の相関値を利用することにより、前後のフレーム間の特徴量を比較したマッチング手法について検討する。さらに、動画の相関値の偏った頻度特性を利用して、しきい値による一定相関値のグループ化及びその時間的前後の変化量との2次元情報によるマッチングアルゴリズムを提案する。また、本手法を用いたシミュレーション実験を行い、検索時間の高速化における有用性を示す。

2. 提案動画検索方式の概要

本検索手法の流れを図1に示す。

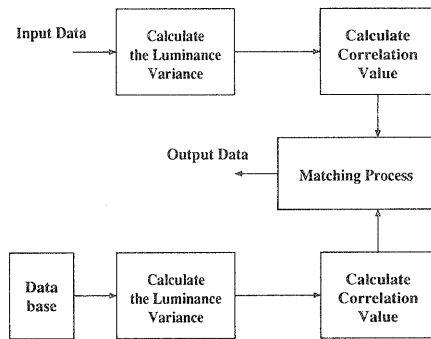


図1: 頻度分布を用いた動画検索の流れ

本検索手法は、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、

$$\bar{x} = \frac{1}{256} \sum_{j=0}^{255} x_j \quad \bar{y} = \frac{1}{256} \sum_{j=0}^{255} y_j$$

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。

3. マッチング手法

マッチング手法として、分布相関値を用いる方法と相関値特性を用いる方法を検討する。

3.1 分布相関値マッチング法

本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とデータベースの輝度分布相関値を1フレームごとに比較していく方法である。具体的には、入力 N フレームとデータベースの先頭から N フレームを式(2)を用いて比較する。

$$|C_{X_i} - C_{Y_{i+s}}| < T_c \quad (2)$$

ただし、

C_{X_i} : 入力画像
 $C_{Y_{i+s}}$: データベース画像
 の相関値である。

式(2)を満たすフレーム数をカウントし、そのカウント数が式(3)を満たすとき、マッチング判定を行う。

$$count > T_f \quad (3)$$

この操作をデータベースの相関値に対して1フレームずつシフトして行う。

この分布相関値マッチング法は、1フレームずつすべてのフレームについてマッチング処理を行っている。

3.2 相関値特性を利用したマッチング法

頻度分布を用いた検索手法の高速化を図るため、相関値特性を利用したマッチング法(以下、相関値特性マッチング法と呼ぶ)を考える。本手法は、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の特徴量とすることにより、動画のマッチングに利用する。具体的には、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。グループ化することにより、分布相関値マッチング法に比べて入力に用いるデータ量が少なくなるため、時間

的特徴量を効果的に利用でき、また高速化が図れると考えられる。

3.2.1 テーブル化

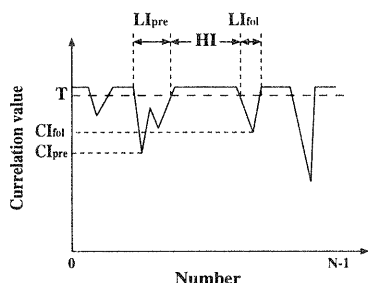


図 2: しきい値によるテーブル化 (入力データ)

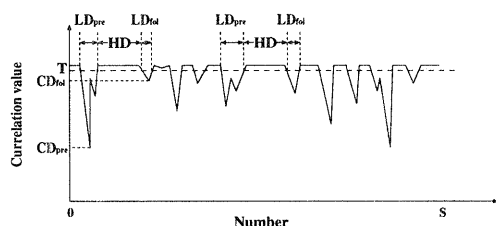


図 3: しきい値によるテーブル化 (データベース)

まず、テーブル化の方法について述べる。

図 2, 図 3 に示すように、分布相関値があるしきい値 T より高いフレーム数、低いフレーム数をカウントする。カウントしたフレーム数を、フレーム番号の小さい方から順にテーブルに格納する。また、しきい値より低い場合は、その領域における相関値の極小値 CI もテーブル化しておく。これを、入力データ、データベースのデータ両方について行う。

ここで、相関値がしきい値より高い領域を high 領域、しきい値より低い領域を low 領域と呼ぶことにする。

3.2.2 マッチング方法

マッチング方法は次の 1 から 15 の過程からなる。

1. 入力画像のテーブル (table_in とする) の high 領域の中で、グループフレーム数が一番大きいもの (図 2: HI) を選択する。
2. データベース画像のテーブル (table_db とする) の中から、グループフレーム数の HI 以上の大

さを持つもの (図 3: HD) を選択する。

3. 入力テーブルにおいて、 HI の直前のテーブル (図 2: LI_{pre}) を選択する。
4. データベースについても同様に HD の直前のテーブル (図 3: LD_{pre}) を選択する。
5. 入力テーブルとデータベーステーブルで選択されたテーブル (LI_{pre} と LD_{pre}) を比較し、等しいものだけを選択する。
6. 5 で選択されたものについて、入力テーブルの HI の直後のテーブル (図 2: LI_{fol}) を選択する。
7. データベースについてもまた同様に HD の直後のテーブル (図 3: LD_{fol}) を選択する。
8. 選択されたテーブル (LI_{fol} と LD_{fol}) を比較し、等しいものだけを選択する。
9. 8 で選択されたものについて、3 で選択された入力テーブルの領域 (LI_{pre}) における相関値の極小値 (図 2: CI_{pre}) を選択する。
10. データベースについても 4 で選択されたデータベーステーブルの領域 (LD_{pre}) における相関値の極小値 (図 3: CD_{pre}) を選択する。
11. それらを比較し、式 (4) を満たすものだけを選択する。

$$|CI_{pre} - CD_{pre}| < T_c \quad (4)$$

12. 11 で選択されたものについて、6 で選択された入力テーブルのテーブル (LI_{fol}) の領域における相関値の極小値 (図 2: CI_{fol}) を選択する。
13. データベースについてもデータベーステーブルのテーブル (LD_{fol}) の領域における相関値の極小値 (図 3: CD_{fol}) を選択する。
14. それらを比較し、式 (5) を満たすものだけを選択する。

$$|CI_{fol} - CD_{fol}| < T_c \quad (5)$$

15. これらの操作の結果、選択された部分が入力データにマッチしたものとして出力される。ただし、各比較選択の操作において、選択された部分が 1 つになった場合は、その時点で操作を終了する。

4. シミュレーション実験

4.1 実験方法

今回用いた入力データとデータベースのデータを表 1 に示す。

表 1 の条件の下で入力データ、データベースデータそれぞれについて頻度分布を求め、それを用いてフレーム前後間の分布の相関値を求める。データ

表 1: 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
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]

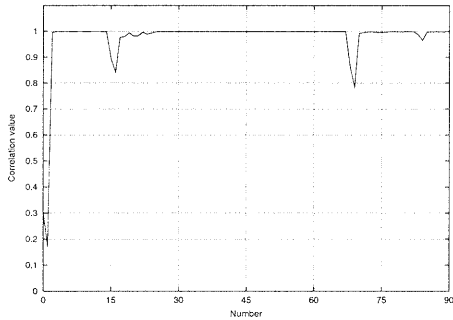


図 4: フレーム前後間の相関値 (入力データ: アニメ)

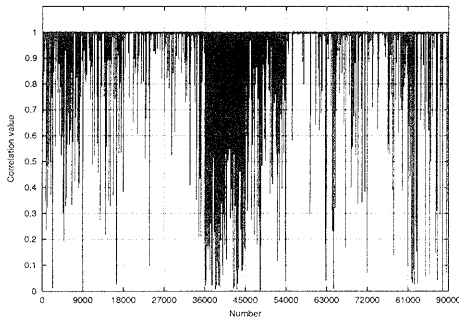


図 5: フレーム前後間の相関値 (データベース)

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。

図 4, 図 5 に入力データ 1 つと、データベースデータについて、フレーム前後間 (i 番目のフレームと $i+1$ 番目のフレーム) の輝度分布相関値のグラフを示す。グラフの横軸は相関値算出の基準フレーム番

号を指し、縦軸は相関値を表す。

これらの相関値を用いて、分布相関値マッチング法、相関値特性マッチング法の両方についてシミュレーション実験を行った。

4.2 実験結果

実験結果を、以下のマッチングの各処理における時間的側面から評価する。

4.2.1 相関値算出

まず、フレーム前後間の相関値の算出にかかった時間を、入力データを表 2、データベースを表 3 に示す。

表 2: 相関値算出時間 (入力)

name	time [s]
アニメ	0.308
スポーツ 1	0.235
スポーツ 2	0.327
音楽 1	0.283
音楽 2	0.306
映画	0.283
ニュース	0.305
スポーツ 3	0.300
スポーツ 4	0.341
スポーツ 5	0.287

表 3: 相関値算出時間 (データベース)

name	time [s]
データベース	3900.140

データベースの相関値算出時間は、入力データに比べてフレーム数分処理に要する。

4.2.2 分布相関値マッチング法

次に、分布相関値マッチング法を用いた検索の実験結果について述べる。表 4,5 に、各入力に対するマッチング処理に要した時間とマッチングパターンを示す。ただし、式 (2), 式 (3) のしきい値は、表中に定めるものを用いた。

表 4 から分かるように、しきい値 T_f を下げると、マッチングパターンが多くなる。しかし、 $T_f = 80$ の時には、かなり絞り込まれており、この条件においては、動画検索に有効であると考えられる。ま

表 4: マッチング時間とマッチングパターン ($T_c = 0.0005$)

name	$T_f = 70$		$T_f = 79$		$T_f = 80$	
	time [s]	matching pattern	time [s]	matching pattern	time [s]	matching pattern
アニメ	1.719	1	1.676	1	1.683	1
スポーツ 1	1.702	5397	1.654	123	1.647	46
スポーツ 2	1.697	7405	1.620	11	1.621	4
音楽 1	1.641	1	1.644	1	1.640	1
音楽 2	1.707	1	1.697	1	1.703	1
映画	1.710	4846	1.619	6	1.621	1
ニュース	1.707	1	1.693	1	1.693	1
スポーツ 3	1.674	11	1.662	1	1.664	1
スポーツ 4	1.714	406	1.691	11	1.689	8
スポーツ 5	1.688	1	1.659	1	1.662	1

表 5: マッチング時間とマッチングパターン ($T_c = 0.00005$)

name	$T_f = 70$		$T_f = 79$		$T_f = 80$	
	time [s]	matching pattern	time [s]	matching pattern	time [s]	matching pattern
アニメ	1.693	1	1.649	1	1.649	1
スポーツ 1	1.614	1	1.602	1	1.603	1
スポーツ 2	1.561	1	1.575	1	1.568	1
音楽 1	1.650	1	1.612	1	1.622	1
音楽 2	1.733	1	1.695	1	1.698	1
映画	1.616	1	1.579	1	1.575	1
ニュース	1.684	1	1.653	1	1.658	1
スポーツ 3	1.607	1	1.610	1	1.614	1
スポーツ 4	1.662	1	1.659	1	1.659	1
スポーツ 5	1.624	1	1.613	1	1.608	1

た、データベースの 90000 枚のすべてのフレームとマッチングしているにもかかわらず、マッチング時間は約 1.6[s] の結果が得られている。

という特徴を利用して、前後のフレーム間の分布相関値を用いてマッチングを行った。

4.2.3 相関値特性マッチング法

最後に、相関値特性マッチング法を用いた検索の実験結果について述べる。

表 6,7 に、マッチングにかかった時間とマッチングパターンを示す。ただし、テーブル化する基準となるしきい値、及び、式 (4)、式 (5) のしきい値は、表中に定めるものを用いた。

分布相関値マッチング法と比べて、さらに高速化が実現できた。

4.3 考察

マッチング手法において頻度分布を用いることは、マッチングのための情報量を少なくできるので、処理面で有効である。さらに本手法においては、動画像では前後のフレームが似ていることが多

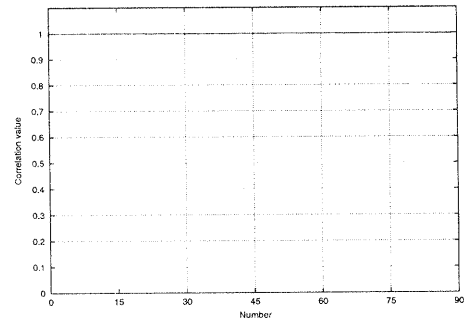


図 6: フレーム前後間相関値 (入力データ: スポーツ 1)

相関値特性マッチング法においては、図 6 のように特徴の少ない、つまり前後のフレームが似ている時間が長い入力画像に対して、どれも同じ結果を出

表 6: マッチング時間とマッチングパターン ($T_c = 0.0005$)

name	$T = 0.995$		$T = 0.996$		$T = 0.997$	
	time [s]	matching pattern	time [s]	matching pattern	time [s]	matching pattern
アニメ	0.364	1	0.366	1	0.371	1
スポーツ 1	0.407	242	0.403	230	0.409	205
スポーツ 2	0.367	1	0.357	1	0.343	1
音楽 1	0.367	7	0.365	4	0.289	2
音楽 2	0.381	96	0.372	26	0.310	2
映画	0.368	1	0.370	1	0.369	1
ニュース	0.398	242	0.408	230	0.411	205
スポーツ 3	0.405	242	0.401	230	0.376	43
スポーツ 4	0.370	1	0.371	10	0.373	11
スポーツ 5	0.370	1	0.368	1	0.366	1

表 7: マッチング時間とマッチングパターン ($T_c = 0.00005$)

name	$T = 0.995$		$T = 0.996$		$T = 0.997$	
	time [s]	matching pattern	time [s]	matching pattern	time [s]	matching pattern
アニメ	0.362	1	0.369	1	0.373	1
スポーツ 1	0.397	242	0.412	230	0.415	205
スポーツ 2	0.366	1	0.361	1	0.364	1
音楽 1	0.360	1	0.365	1	0.365	1
音楽 2	0.374	14	0.373	8	0.378	2
映画	0.368	1	0.363	1	0.369	1
ニュース	0.410	242	0.411	230	0.413	205
スポーツ 3	0.403	242	0.405	230	0.366	7
スポーツ 4	0.368	1	0.359	1	0.368	2
スポーツ 5	0.371	1	0.360	1	0.368	1

力するという問題点があり、入力画像に依存するところが大きい。このような入力画像に対しては、相関値が high 領域と low 領域に分かれるようにしきい値 T を変動するとよいと考えられる。表 6 と表 7 を見ると、しきい値 T の高い方が検索結果が絞り込まれていることが分かる。しかし、時間特徴量が少ないため、相関値特性マッチング法だけの絞り込みでは不十分であり、時間方向以外の特徴量との組み合わせによるマッチング手法が必要であると考えられる。

また、マッチング時間の点から見ると、相関値特性マッチング法は分布相関値マッチング法に比べても高速化できている。このマッチング時間には、相関値の算出時間などは含まれていないが、表 2 にあるように、入力データの相関値の計算には約 0.3[s] 要する。よって、入力データの前処理として頻度分布を求めておき、またデータベース側はフレーム前後間の分布相関値をデータとして与えておけば、この相関値特性マッチング法で、さらに高速なマッチング処理が可能となる。

5. まとめ

頻度分布を用いたマッチング手法としてのフレーム前後間の分布相関値を利用した手法について述べた。また、相関値特性マッチング法により、時間特徴量を効果的に利用することで処理面の改善方法を示し、高速性を検証した。今後の課題としては、相関値特性マッチング法における入力に依存しない方法の確立が挙げられる。

参考文献

- [1] 是津 耕司, 上原 邦昭, 田中 克己: “時刻印付オーサリンググラフによるビデオ映像のシーン検索”, 情処論誌, **39**, 4, pp. 923-932 (1998).
- [2] 高橋克直, 寺島信義, 富永英義: “画紋情報を用いた動画検索手法に関する検討”, 映像学技報, BCS98-61, pp. 1-8 (1998).