

MSP 検索システムにおける検索手法の検討

児玉 明[†] 竹本 正行[‡] 池田 朋二[‡] 真崎 剛[‡]

[†] 広島大学 地域共同研究センター

[‡] 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 我々はマルチメディアスケーラビリティパッケージ(MSP)という機能化データのパッケージ化による、マルチメディアサービスの概念を提案している。そして、MSPを用いた交換型画像検索システムの実現を目指している。本稿では、画像の検索手法として、空間方向と時間方向の特徴量を抽出し、その結果をマッチング処理するという手法の効果を実証した。そこで、MSPを用いることにより、これらの行程をクライアント側とサーバ側に分散させて処理を行い、参照テーブルを用いた効果的な画像検索の方式が実現できることを述べた。

キーワード マルチメディアスケーラビリティパッケージ, 動画画像検索システム, 画像特徴量抽出
画像通信システム, スケーラビリティ

A Study on the Image Matching Methods in MSP Video Searching System

Mei KODAMA[†] Masayuki TAKEMOTO[‡] Tomoji IKEDA[‡] Tsuyoshi MASAKI[‡]

[†]Center for Technology Research and Development,
HIROSHIMA University

[‡]SATAKE Corporation

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

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 We had indicated communication services using multimedia scalability packages(MSP), which can provide multimedia information according to functionality and availability for users. In this paper, we propose the MSP video searching system as one of the MSP multimedia services. And then, we inspected effects of methods, which have image feature sampling as first step and matching process as second step in video searching system. Especially, we study on effective video searching methods by two dispersion process from the client side and the server side using the MSP in reference tables.

key words Multimedia Scalability Packages, Video Searching System, Image feature sampling,
Image Communication System, Scalability

1. はじめに

近年、コンピューターの処理能力の進歩にともない動画像データを利用するケースが増えてきている。そこで蓄積された動画像に対してユーザーが必要なものを効率良く検索できる技術が求められている。また、我々の提案しているマルチメディアスケーラビリティパッケージ(MSP)通信システムでは、MSPデータによる相互運用性のあるマルチメディアサービスの実現を目的としている。MSPデータは、コンテンツ情報に対して情報利用時のサーバ側処理量並びに機能性を考慮し、情報作成時に前処理を行いスケーラビリティ構造化したデータである[1]。

本稿では、動画像検索を行う際の、効果的な検索手法を実証するとともに、MSP画像検索システムへ応用するための方式について提案する。

2. 画像検索処理手順

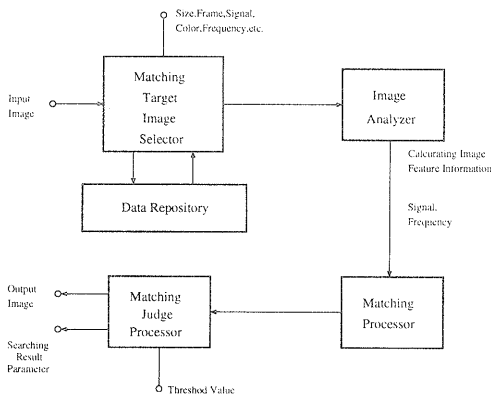


図 1: 画像検索処理手順

画像検索処理手順について図1に示す。まず、入力画像と入力パラメータ(手法の選択、画像サイズ、閾値など)が共に入力され、対象画像選択器(Matching Target Image Selector)より、対象のデータベース上の画像が選択される。対象データベース画像、入力画像、入力パラメータと共に、画像解析器(Image Analyzer)で特徴量を抽出し、次に検索処理器(Matching Processor)で特徴量の比較による検索処理を行う。最後に検索判定器(Matching Judge Processor)で閾値判定を行い、選択されたデータベース画像と共に検索結果付加情報(検索所用時間、検索枚数など)を出力する。こ

のように、画像の検索を行う際には、画像の特徴量を抽出し、それをもとにマッチングを行うという手順を踏むことになる。

3. 動画像検索手法

本研究の動画像検索手法においては、画像の空間方向及び時間方向の特徴量に着目する。空間方向での検索には、入力フレームとデータベースのフレーム間の特徴量の誤差を判定に利用し、時間方向での検索には、入力とデータベース双方の連続したフレーム間の特徴の変化量の誤差を判定に利用する。本研究ではフレーム単位に特徴量の平均値を算出してテーブル化し、検索時に利用する事で検索時間の短縮化を計っている。画像の特徴情報に着目し3.1に一般的な特徴量抽出の手法を示す。

3.1 画像特徴量抽出手法

平均値、分散値、最小値、最大値、ダイナミックレンジ、標準偏差を利用した画像特徴量の抽出手法を挙げる。そして、これらの特徴量抽出手法を利用して、シミュレーション実験を行った。

3.1.1 平均値

フレーム内の信号の平均値を a とすると、 a は式(1)で表される。

$$a = \frac{\sum_{m=1}^m \sum_{n=1}^n X_{mn}}{M \times N} \quad (1)$$

但し、 M は、画像の縦サイズ、 N は、画像の横サイズを表し、 X_{mn} は、座標 (m, n) の信号値を示す。

3.1.2 分散値

フレーム内の信号の分散値を b とすると、 b は次の式で表される。

$$b = \frac{\sum_{m=1}^m \sum_{n=1}^n (X_{mn} - a)^2}{M \times N} \quad (2)$$

3.1.3 最小値、最大値、ダイナミックレンジ

画像1フレームにおける信号の最小値、最大値、ダイナミックレンジを算出し、画像の特徴量として扱う。

3.2 マッチング手法

画像の空間的特徴量を利用したマッチング手法と時間的特徴量を利用したマッチング手法について検討する。本稿では、画像信号として、輝度値に着目して、検討を進めてゆく。

3.2.1 特徴量の平均値誤差

データベース画像と入力画像の空間的特徴量の誤差を利用する。その判定式は式(3)を利用し、閾値 th_1 で判定する。

$$|M_i - N_k| < th_1 \quad (3)$$

但し、 M_i は、データベース画像の特徴量を表し、 i は、データベース内のフレーム番号を示す。 N_k は、入力画像の特徴量を表し、 k は、入力画像内のフレーム番号を示す。

3.2.2 特徴量の平均値累積誤差

データベース画像と入力画像において、入力フレーム数方向に跨った空間方向の特徴量誤差の累積は、式(4)で表される。その際の、誤差を閾値 th_2 で判定する。

$$\frac{\sum_{k=1}^n (M_{i+k} - N_k)^2}{k} < th_2 \quad (4)$$

3.2.3 フレーム間の特徴量の平均値の変動誤差

データベース画像と入力画像において、それぞれの時間方向に隣り合うフレーム間の特徴量を求め、その特徴量の誤差は式(5)で表される。その際の誤差を閾値 th_3 で判定する。

$$|(M_{i+k} - M_{i+k+1}) - (N_k - N_{k+1})| < th_3 \quad (5)$$

3.2.4 フレーム間の特徴量の平均値の累積変動誤差

3.2.3における入力フレーム数方向に跨った特徴量誤差の累積は、式(6)で表される。その際の、誤差を閾値 th_4 で判定する。

$$\frac{\sum_{k=1}^n \{(M_{i+k} - M_{i+k+1}) - (N_k - N_{k+1})\}^2}{k} < th_4 \quad (6)$$

3.3 マッチング用特徴量の算出

3.3.1 算出条件

画像中の輝度信号に着目し、各特徴量の算出を行った。今回用いたデータベース用の動画の条件及び特徴を表1に示す。但し、画像フォーマットは、それぞれ ITU-R Rec.601 4:2:0 704[pe]x480[line] である。

表 1: データベース画像内容

フレーム番号	種別	特色
0 ~ 8999	アニメーション	全体的に暗め、モノクロシーン有り 動きは少ない、ズームイン ズームアウト有り
9000 ~ 17999	スポーツ1	テロップ有り、背景固定 シーンチェンジ少、対象移動
18000 ~ 26999	スポーツ2	シーンチェンジ少 背景固定、人物移動
27000 ~ 35999	音楽1	背景固定、全体的に暗め ディゾルフ領域有り、人物ほぼ固定
36000 ~ 44999	音楽2	シーンチェンジ多、動きが激しい 照明を多用
45000 ~ 53999	映画	夜間、爆発のシーン有り 照明を多用
54000 ~ 62999	ニュース	人物背景固定、テロップ多い シーンチェンジ少
63000 ~ 71999	スポーツ3	屋外、背景ほぼ固定 人物が多く、細かい動き
72000 ~ 80999	スポーツ4	屋内、人物アップ ズームイン、ズームアウト有り
81000 ~ 89999	スポーツ5	屋外、シーンチェンジ少ない 背景固定、人物移動

3.3.2 各特徴量の結果

表1に基づいて、各特徴量を抽出した結果を図2~図4に示す。ここで、各図の横軸は、データベース上のフレーム番号を示し、縦軸は各特徴量の値を示している。

3.3.3 各特徴量の考察

図2より、各フレームの輝度値の平均値を見ると、それぞれのシーケンスにより変動が大きく現れているところもある。また、シーケンスごとに変動の中心が異なっていることが分かる。ニュース番組のように画像に大きな変化がない場合には、平均値の変動も小さくなっている。よって、各シーケンスによる違いが現れている。

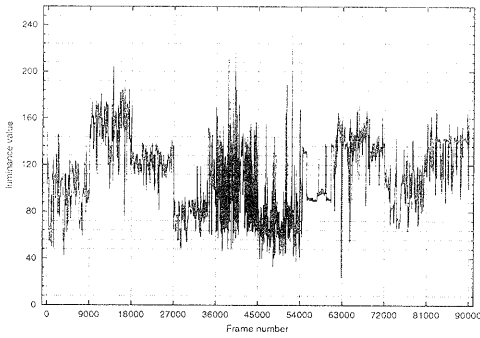


図 2: 平均値

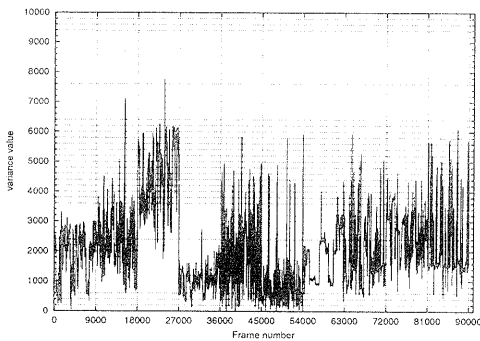


図 3: 分散値

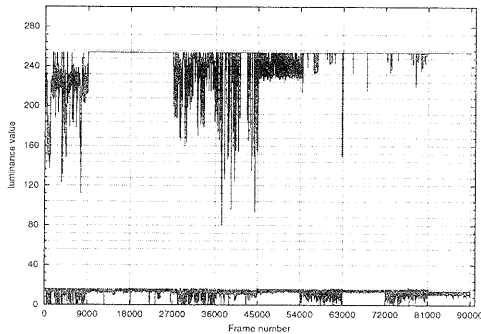


図 4: 最大値, 最小値

図3より、分散値については、各シーケンスともフレーム単位で変動が大きく現れている。特に明るいシーンが多いスポーツ1やスポーツ2、スポーツ3やスポーツ4、また音楽2のようなシーケンスはその変動が大きく現れている。また、振幅がシーケンスごとに偏っている傾向があり、各シーケンス

による特徴が現れている。

最大値、最小値、ダイナミックレンジでは、振幅差があまり見られず、輝度値が集中しているシーケンスがあるために区別するには、不十分と言える。

以上の結果より、本稿ではマッチングに用いる特徴量としてフレームの平均値と分散値について着目し検討する。

3.3.4 マッチング手法の考察

式(3)~式(6)の検証を行うため、実際にそれぞれの手法で各特徴量を利用し、マッチング処理を行った。入力画像は、式(3)では90枚のうち先頭1フレームを使用し、式(5)では、先頭2フレームを使用した。式(4)、式(6)では90フレーム全てを使用した。マッチング条件は、入力画像1フレーム、2フレームあるいは、90フレームの特徴量の平均値に対し、全データベース画像90000フレームの平均特徴量テーブルとの差分マッチングを行った。結果を表2に示す。但し、表2中のThは閾値を示す。なお、特徴量の平均値によるマッチング処理で、入力画像と同一フレームが、データベース上に存在しない場合を考え、閾値を変化させた場合のマッチング枚数の変化を見るため3パターン of 閾値を用いた。

表 2: 各手法におけるマッチング数

手法	マッチング数 [Frame]					
	Th = 0		Th = 1		Th = 2	
	a	b	a	b	a	b
式(3)	925	1	2449	53	3848	100
式(4)	1	1	4	2	7	4
式(5)	46401	1	73936	3372	77778	6756
式(6)	1	1	29975	143	41504	1927

空間的特徴量によるマッチング手法について考える。式(3)より、平均値誤差を用いて検索を行うと、平均値の場合、しきい値を0にしても900枚以上のフレームを抽出している。一方、分散値の場合、同様にしきい値を0にするとマッチング数を低減できる。

式(5)より、フレーム間の平均値変動誤差を用いて検索を行うと、平均値の場合、式(3)の値に比べて多くのフレームを抽出している。これは動画像の場合、フレームの前後間での相関が高いからである。逆に分散値の場合は、フレーム間の前後の相関

が高いことが多いことにより、式(3)の値に比べてるとマッチング数が多くなり、また処理量の面からも、式(5)より、式(3)の方が単純且つよい。

本シミュレーション実験では、データベースとして90000枚のフレームを用いたが、データベースのフレーム数がさらに多くなると、同一の値を持つフレームの確率が高くなることが予想できる。さらに、どちらの空間方向の手法も、閾値をこれ以上下げることができないことを考慮すると、空間方向の検索手法における複数のパラメタを用いた多次元検を行う必要がある。

次に、時間的特徴量によるマッチング手法について考える。式(4)より、平均値の累積誤差を用いて検索を行うと、空間方向における検索手法に比べて、抽出したフレーム枚数が少なくなっていることが分かる。これは、式(6)のフレーム間の平均値変動の累積誤差を用いて検索を行った場合についても、 $Th=0$ のときは同様である。しかし、式(4)と式(6)を比較すると、閾値を上げた場合のマッチング数が、式(6)の方が著しく増加している。これは、相関の高いフレーム間で、差分をとってマッチングしたことにより、閾値の範囲に入ったフレームが多くなったと考えることができる。よって、処理面から見ると、式(4)と式(6)では、式(4)の方が単純であり、また、マッチング率から見ても、優れていると言える。

ただ、時間方向における検索手法は、いずれの手法も、空間方向における検索手法に比べて、マッチング数を抑えることができると言える。しかし、一方、式(3)から式(6)に示したように、時間方向の計算は、空間方向の計算と比較して処理に時間がかかるという問題がある。

以上をまとめると、今回扱ったマッチング手法では、空間方向の検索には式(3)がよく、時間方向の検索には式(4)がよい。

そこで、動画の検索を空間方向の検索手法と、時間方向の検索手法とを段階的に行うことにより効率的な動画検索が可能であるかを検討する。

4. 複合型画像検索手法

4.1 提案方法

本稿では、空間方向での検索を行い、フレーム数がある程度に絞り込み、その後、時間方向での検索を行うことにより、効果的な画像検索を行うという検索手順を提案し、シミュレーション実験を行うことにより、その有効性を示す。

図5に、提案する複合検索手法の概念図を示す。

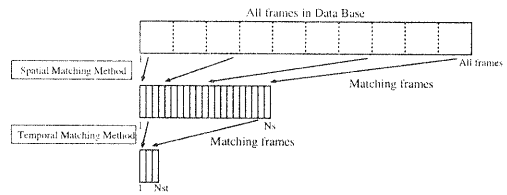


図 5: 空間と時間の複合方式

まず、データベース全体のフレームに対して、空間方向の検索を行い、マッチングしたフレームを N_s 枚までに減らしておく。次に、時間方向の検索を行い、マッチングするフレーム数を N_s 枚にして検索を行う。

4.2 シミュレーション実験

ここで、各手法に要する処理時間についてシミュレーション実験を行う。本稿におけるシミュレーション条件として、入力画像とデータベース側の両方で、フレーム単位の特徴量情報テーブルを事前に作成しておき、そのテーブルに基づいた検索を行う。また、今回の前提条件として入力画像と同一の画像がデータベース上に存在していることとする。シミュレーション実験には、Fujitsu GP7000SM25 400MHz UltraSPARC-2 上で行った。

実験結果を表3に示す。ここで用いた特徴量は分散値を用い、空間方向、時間方向におけるマッチング処理の平均処理時間を計測した。

表 3: 各手法のマッチング時間

手法	時間 (sec)
式(3)	0.262
式(5)	0.266
式(4)	1.133
式(6)	1.653

表3より、空間方向の検索の方が時間方向の検索より、マッチング処理時間を短縮することができる。また、表2より、空間方向の方がマッチングヒット率を低減できることを考慮すると、総検索時間を短縮するには、まず、空間方向の検索を行い、マッチング数を低減してから、時間方向の検索を行うことで対処でき、提案方法が有効であると言え

る。

5. MSP 画像検索システムへの応用

以上の結果を踏まえて、MSP 画像検索システムにおける検索処理手順について考察する。

MSP 画像検索システムは、機能化パッケージ交換型の検索システムである [1]。ので、クライアント側とサーバ側とで処理を分散することが可能である。よって、クライアント側とサーバ側とで処理を分散できることが特徴である。

この検索システムにおいて、複合画像検索手法を利用する際に、特徴量の抽出と、マッチング処理の両工程をサーバ側で行う場合と、クライアント側で行う場合を分けて考えることができる。例えば、前処理として特徴量をクライアント側で処理しておき、その後、サーバ側で、その結果をマッチング処理に利用する場合が考えられる。ここで、両工程をクライアント側で処理することは、サーバを介さないでクライアント端末内でローカル検索を行う場合である。

分散処理を行う際の処理の概略を、図6に示す。

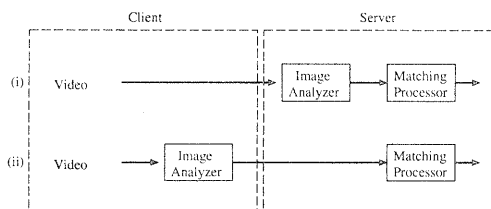


図6: 分散処理の概略図

(i)の方式は、画像の特徴量抽出と、マッチング処理の両方を、サーバ側で行っている。(ii)の方法は、画像の特徴量の抽出をクライアント側で処理させている。本稿では、画像の特徴量を抽出する処理をクライアント側で行うことにより、処理の分散を実現する。(ii)の方式場合について考えてゆく。

○ クライアント側の処理

クライアント側で画像の特徴量抽出を行うことにより、画像の特徴量を記述した参照テーブルを作成する。これにより、サーバ側に受け渡すコンテンツデータは、画像データそのものではなく、参照テーブルを利用することが可能となる。参照テーブルをコンテンツデータとして作

成したMSPデータをサーバ側に送信することにより、送受信の負荷を低減することが期待できる。

○ サーバ側の処理

次に、参照テーブルを受信した、サーバ側の処理について検討する。サーバ側では、MSPデータを受信すると、MSPマネージャによって、サポートしているモード(振舞い)に従い、参照テーブルを用いてマッチング処理を行う [1]。

マッチング処理の結果、得られた検索結果をMSPデータとして作成し、クライアント側へ返信するが、そのとき、受信した参照テーブルの情報をサーバ側の情報として反映する。これにより、サーバ側は、クライアント側で処理した、画像の特徴量抽出の結果をもとに、常に参照テーブルを最新のものに更新することができる。そして、更新された参照テーブルの情報を、返信するMSPデータの中を含むことにより、クライアント側の参照テーブルも、常に最新の情報に更新することができる。

このように、最新の参照テーブルをクライアント側とサーバ側で共有することにより、処理の負荷は低減でき、最新の情報は共有するという効果的な検索システムが実現できる。

6. まとめ

本稿では、空間方向の検索と、時間方向の検索を複合的に行う場合の組み合わせと、その順序について、シミュレーション実験を行い、本稿で提案する複合画像検索手法の有効性を示した。そして、この手法をMSP画像検索システムに応用した場合、サーバ側とクライアント側で検索処理を分散することで、その情報を両方で共有することが可能であることを示した。今後は、MSP検索システムへ複合型検索手法を実装し、評価してゆく予定である。

参考文献

- [1] 児玉, 竹本, 池田, 真崎: “MSP 画像検索システムにおける情報構成法の考察”, PCSJ'99, pp. 53-54 (1999).