

MSP 情報検索システムとその動画像検索エンジンに関する検討

児玉 明 高橋 秀和

広島大学 地域共同研究センター

〒739-0046 東広島市鏡山 3-10-31

TEL : (0824) 21-3646

FAX : (0824) 21-3639

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

あらまし 我々はメディア統合およびマルチメディアサービスの統合化を目指した構造化データとして、マルチメディアスケーラビリティパッケージを提案している。本稿では、MSP データを利用した検索システムを提案した。また、動画像の時間的特徴量を利用した動画像検索方式を提案した。本方式は、フレーム間の時間相関値を利用し、動的変化構造を検索テーブルとして用いて、各検索パラメタの条件を満たすものを抽出する方式である。本提案方式について、確率モデルの観点から考察する。頻度分布とマッチング数の観点から、検索方式としての可能性について評価する。

キーワード MSP 通信システム, MSP 検索システム, 動画像検索, 情報記述,
マルチメディアスケーラビリティパッケージ

A Study on MSP Search System and its Video Search Method

Mei KODAMA Hidekazu TAKAHASHI

†Center for Technology Research and Development,
HIROSHIMA University

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

TEL : +81-824-21-3646

FAX : +81-824-21-3639

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

Abstract We had already proposed "Multimedia Scalability Package:MSP" as structure data in integrated multimedia services. In this paper, we proposed MSP Search System and the video search method based on temporal characteristics.

MSP data. We considered these system using proceedings data and video contents. In these system, it is realized that both of query data and answer data are treated as MSP data and these are transmitted between client and server terminals. In addition, proposed video search method was evaluated from the viewpoint of probability model.

key words MSP Communication System, MSP Search System, Video Search Method,
Information Description, Multimedia Scalability Packages

1. はじめに

我々は、機能化パッケージデータによる統合的な情報交換方式を用いたマルチメディアサービスの概念を提案している [1]。本サービスにおけるマルチメディア情報はスケーラビリティの概念を基本構造として用いており、このような機能を備えたパッケージデータを MSP (マルチメディアスケーラビリティパッケージ) と呼んでいる。現在、マルチメディアサービスとして、情報検索機能に着目し、その中でも特に、画像検索を中心に検討している [2]。MSP の特徴は、データ自身で動作機能を有することであり、よって、画像検索システムにおいて、画像特徴量抽出、マッチング処理などの検索処理や情報配信などの自動的な処理動作を行うことが可能である。

本稿では、MSP 検索システムにおける時間特徴量を利用した動画像検索方法を提案し、その評価を行い、提案方法の有効性を示す。

2. MSP 検索システム

2.1 概念

MSP 検索システムの概念を図 1 に示す。本検索システムは、利用者の持っている情報から、利用者の意志を受け、情報検索という機能を持った MSP 情報を生成し、データベースへ送信する。また、データベースにおいて、スケーラビリティを利用した情報検索を行い、その結果を、MSP 検索結果データとして返信してもらい、データを解凍した際に、検索結果を知るというシステムである。ただし、検索データの生成に予め利用者側でスケーラビリティを利用することも可能である。

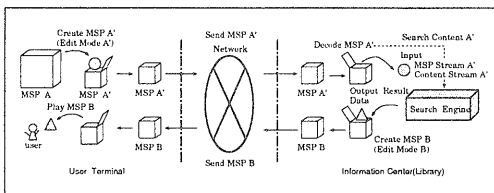


図 1: MSP 検索システムの概念

具体的には、利用者が検索対象とするコンテンツを元 MSP A から部分的な MSP データ (MSP ヘッダ情報とコンテンツ) を抽出して検索用の MSP データを作成する。ここで新たに生成された MSP データ

を検索用の MSP A' とする。MSP A' は検索動作情報と利用者の意図する検索項目、検索数や検索時間などの付加的情報を有する。次に、この MSP A' を検索サーバへ配送する。検索サーバでは、受信した MSP A' を復号して、検索指令のデータであることを認識する。そこで、検索処理の詳細項目及び対象コンテンツデータを MSP A' から抽出し、検索エンジンにかける。検索エンジンからの結果 (出力データや対象情報、あるいはネットワークアドレスなど) を元に、新たな MSP データとして MSP B を生成し、検索サーバから利用者へ返送する。ユーザは記述された MSP B のモードデータに基づき、スケーラビリティ機能を利用して、検索結果あるいは要求するデータを得ることができる。

2.2 検索処理器

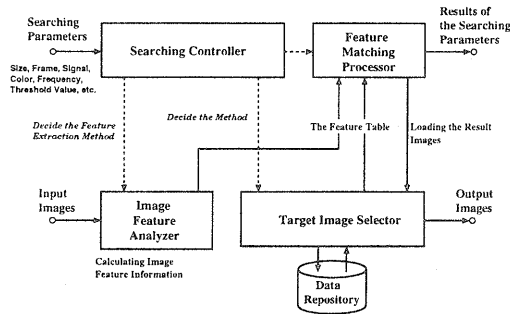


図 2: 検索処理手順

図 2 に MSP 検索システム内の検索処理手順を示す。本処理器は、検索処理制御器、画像特徴量算出器、特徴量マッチング判定器、対象画像選択器の 4 つから構成される。

検索処理制御器は、検索用のパラメータとして入力データのサイズや検索処理を実行する際の閾値などの情報を入力として検索手法を決定する。そして、それらの入力情報をもとに、どの特徴量を抽出するかを決定し、画像特徴量算出器を動作させる。画像特徴量算出器は、検索処理制御器の指令を受けて、入力画像よりその特徴量を算出する。そして、算出結果として特徴量テーブルを出力する。また、検索処理制御器の指令を受けた対象画像選択器は、データベース側からマッチング処理に適した特徴量テーブルを選択し、両特徴量テーブルを特徴量マッチング判定器に引き渡す。特徴量マッチング判定器は、検索処理制御器より閾値等のパラメータを取得し、両

テーブルを利用してマッチング処理を行う。そして検索結果により対象のコンテンツデータを決定し、それをデータベースから出力させる指令を出す。また検索処理の結果、生成されたマッチング数やマッチング処理時間等のパラメータを出力する。

検索処理を行う際、比較する対象として、入力情報と、データベース中の情報との特徴量を用いる。入力情報の特徴量を F_i 、データベース中の特徴量を F_d とし、判定に使用する閾値を Th とすると、例えば判定式は、 $|F_i - F_d| \leq Th$ で表される。

3. 時間的特徴量を利用した動画像検索方式

ここで、動画像におけるフレーム間の時間的変化量として時間相関を利用したマッチング手法を提案する。本マッチング手法は、フレーム前後間の輝度分布の相関値を動的な変化量として利用することにより、その特徴量を比較する手法である。

3.1 概要

図2の画像特徴量算出器における本マッチング手法の処理手順を図3に示す。本マッチング手法にお

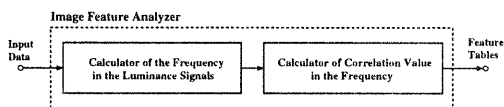


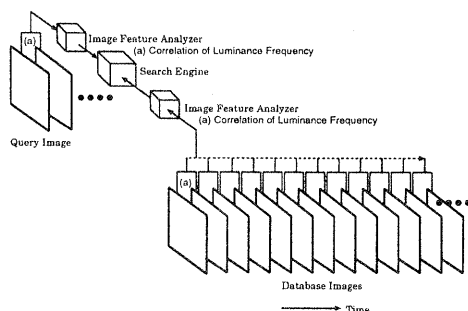
図3: 時間特徴量算出の処理手順

いて、画像特徴量算出器では入力データの輝度分布算出を行い、その算出結果をもとに分布相関値を算出し、これにより特徴量テーブルを作成する。

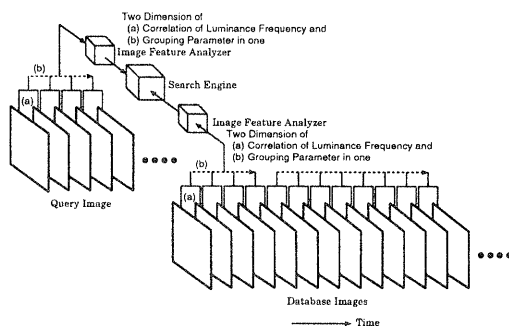
さらに、動画像においてその相関値の偏った頻度特性を利用して、閾値による一定分布相関値のグループ化及びその時間的前後の変化量との2次元情報によるマッチングを行うことで動画像検索を実現する。

ここでは、時間特徴量としてフレーム相関値を用い、その相関値をある特定の閾値でグループ化する方法を提案する。

以上で求めたフレーム前後間の相関値を用いて、特徴量マッチング判定器においてマッチング処理を行う。ここで本稿では、マッチング手法として分布相関値を用いる方法(分布相関値マッチング法)と相関値特性を用いる方法(分布相関群マッチング法)として、後者を中心に話を進める。(図4)



(i) Temporal Approach with Each Two Frames



(ii) Temporal Extension Approach to Grouping Parameter

図4: 時間的アプローチ

3.2 分布相関値について

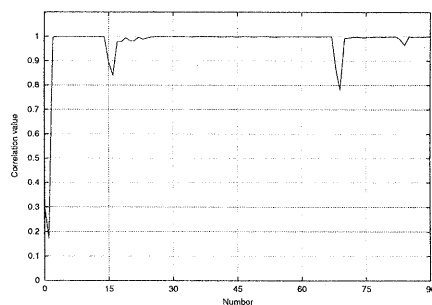


図5: 分布相関値(入力データ: アニメーション)

フレーム前後間の分布相関値は、図5のような波形で示される。図5に、入力データのフレーム前後間(i 番目のフレームと $i+1$ 番目のフレーム)の輝度分布相関値を示す。ここで図の横軸は相関値算出の基準フレーム番号を指し、縦軸は分布相関値を表す。分布相関値の変動が、一つの時間的な特徴量を

示している。

分布相関値マッチング手法は、これらの波形を比較することにより判定を行う。つまり、入力データ、データベースデータそれぞれのフレーム前後間の分布相関値を1つずつ比較する方法である。

この処理手順を図6に示す。

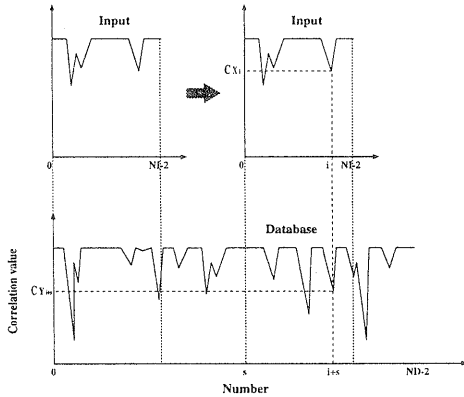


図 6: 分布相関値マッチング法の処理手順

図6に示すように、入力データの分布相関値の波形と、そのフレーム数分のデータベースデータの分布相関値の波形と順番に比較していく。

3.3 分布相関群マッチング法

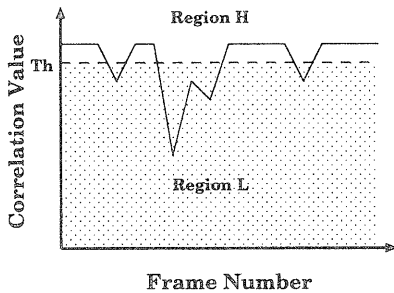


図 7: 閾値によるテーブル化

本研究では分布相関群マッチング法を提案する。本マッチング手法は、動画像の分布相関値が変化が激しいところを除き、1に近い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特徴量とすることにより、動画像データのマッチングに利用する。具体的には、分布相関値がある

閾値より高いか、または低いかで相関値をグループ化し、そのグループ化の数で判定する方式である。

入力映像情報に対して、時間方向に対して前述のように相関値を求め、図7に示すように、時間方向に対して分布相関値がある閾値 Th より連続して高いフレーム数 $Length_{H_n}$ 、連続して低いフレーム数 $Length_{L_n}$ をカウントし、特徴量テーブルとする。

検索用入力動画像情報に対しても同様に、構造テーブルを算出し、両者の間で判定処理を行う。

4. 評価実験

4.1 実験条件

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[pe]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。なお検索処理には、Fujitsu GP7000SM25 400MHz UltraSPARC-2 を用いた。

4.2 実験結果

4.2.1 データベース相関値の頻度分布

データベースの相関値テーブルに対し、相関値の精度を変化させ、相関値の頻度分布の平均値と分散値を求めた結果を表2に示す。相関値の頻度分布は、その平均値約 0.993 付近に分布していることが分かる。

表 2: 相関値頻度分布の平均と分散

精度	平均	分散
1.0×10^{-3}	9.929873×10^{-1}	2.310268×10^{-3}
1.0×10^{-4}	9.929810×10^{-1}	2.310074×10^{-3}
1.0×10^{-5}	9.929812×10^{-1}	2.310032×10^{-3}
1.0×10^{-6}	9.929807×10^{-1}	2.310027×10^{-3}

データベース内の閾値を変化させた場合のグループ H の占有率とその閾値の関係を表3に示す。但し、閾値を 1.0 から低い方向に考えてゆく。また、閾値と占有率の関係を図8に示す。

表 3: グループ H の占有率に対する精度と閾値との関係

占有率	相関値精度			
	1.0×10^{-3}	1.0×10^{-4}	1.0×10^{-5}	1.0×10^{-6}
10%	1.000000	0.999700	0.999670	0.999665
20%	1.000000	0.999500	0.999530	0.999525
30%	0.999000	0.999400	0.999390	0.999394
40%	0.999000	0.999200	0.999240	0.999242
50%	0.999000	0.999100	0.999050	0.999053
60%	0.999000	0.998800	0.998790	0.998786
70%	0.998000	0.998400	0.998350	0.998352
80%	0.998000	0.997700	0.997660	0.997661
90%	0.996000	0.995600	0.995650	0.995648

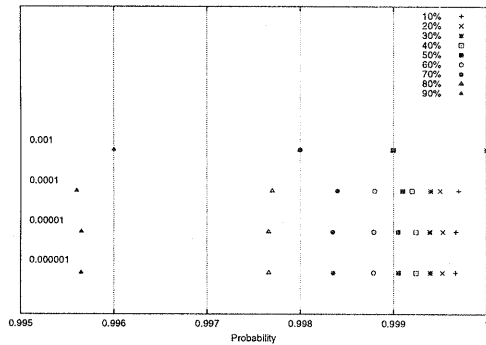


図 8: H の占有率と閾値との関係

図8より、約 40 ~ 60 % で閾値間の差が狭く、相関値頻度が最も多くなっていることが分かる。以下、固定閾値として、H の占有率を 50 % とする閾値を用いて評価を行う。

データベースのフレーム前後間相関値に対して、固定閾値を用いてグループ化を行い、その結果得られたグループ H とグループ L の長さについて、頻度分布を算出した。ここで固定閾値は、精度幅 1.0×10^{-6} の頻度分布より算出したものを用いた。H の確率を 50 % としたときの H と L の長さの頻度分布を、図 9 に示し、また、確率を 10 %, 30 %, 50 %, 70 %, 90 % と変化させて求めた H と L の長さの頻度分布を、それぞれ図 10 に示す。さらに、隣り合う HL についての、各 H, L の長さの頻度分布を 3 次元グラフとして、図 11 に示す。

4.2.2 H と L の長さの頻度分布における分散

データベースの頻度分布から求めた固定閾値を用いて、グループ H, L によるマッチングを行った。表、に、入力動画像の構造とそのテーブルを

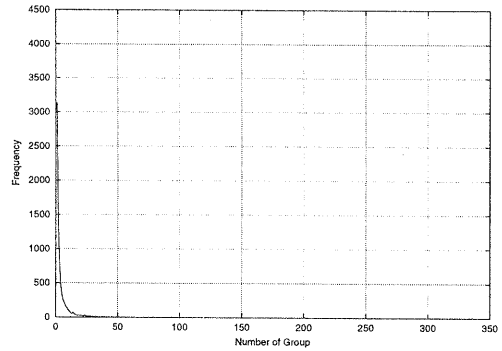


図 9: グループ H の長さの頻度分布 (H の占有率 50%)

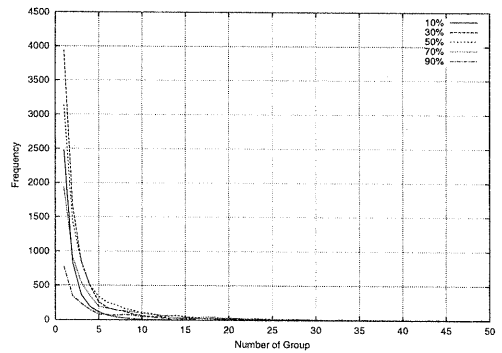


図 10: グループ H の長さの頻度分布 (占有率による推移)

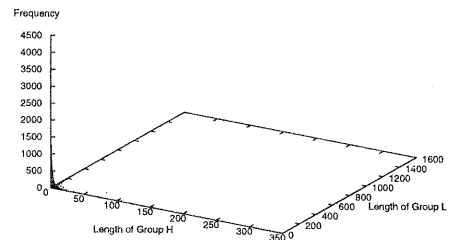


図 11: グループ HL の長さの頻度分布 (H の確率 50%)

利用した際の、マッチング数を示す。なお、用いた固定閾値は、データベースの頻度分布算出の際の精度: 1.0×10^{-6} 、頻度分布における H の占有率が 50 % のものである。また、表中で (-) となっている部分は、入力グループ数が 2 以下であり、本マッチング手法ではマッチングの際に最初と最後のグルー

表 4: 入力映像の HL 構造とマッチング数: アニメ

フレーム数	総構造数 (H,L)									
150	31(15,16) 1									
75	21(10,11) 1					11(5,6) 1				
30	4(2,2) 1		13(7,6) 1		5(2,3) 63		7(3,4) 6		4(2,2) 7	
15	2(1,1) -	2(1,1) -	9(5,4) 1	4(2,2) 79	5(2,3) 63	1(0,1) -	2(1,1) -	5(2,3) 65	2(1,1) -	3(1,2) 3134

表 5: 入力映像の HL 構造とマッチング数: スポーツ 3

フレーム数	総構造数 (H,L)									
150	58(29,29) 1									
75	26(13,13) 1					33(16,17) 1				
30	9(5,4) 1		15(7,8) 1		11(5,6) 1		12(6,6) 1		11(5,6) 1	
15	5(3,2) 91	5(3,2) 278	10(5,5) 4	6(3,3) 16	3(1,2) 1444	9(4,5) 6	6(3,3) 5	6(3,3) 27	8(4,4) 3	4(2,2) 10

但し、上段括弧外は、総 HL 構造数、括弧内は (H,L) の内訳を示す。下段は、マッチング数。

プを省いたためマッチングが行えないことを表している。

表より、構造数の増加と、マッチング数の減少できることがわかる。入力フレーム別に見ると、150 フレーム、75 フレームの時は、構造数が多いため、マッチング数が少ないが、入力フレームを 30、15 とするにつれて構造数が減少し、それに伴って、マッチング数が増加している。よって、H の長さ L の長さの分散を大きくすることで、本方式を効果的に利用できることが示された。但し、特徴的な動的構造が表現できないシーケンスに対しては課題である。また、入力動画シーケンスに適応的に切り替える閾値算出方法、テーブル作成方法も課題である。

4.2.3 H と L の長さの頻度分布における分散

グループを時間方向に拡張させる方法として、HLHL... の構造に分けて考察する。隣り合うグループ H とグループ L、隣り合う 3 つ HLH、隣り合う 4 つの HLHL について、各々共分散を求めた。その結果を表 6 に示す。表 6 より、グループの構造が増えるにしたがって共分散の絶対値が大きくなっており、値が全体に散らばるため、構造数を増やすほどマッチング精度が向上することを示し、本時間方向への拡張方法の有効性を示した。但し、H の占有率と分散との関係について考慮すると、閾値の設定方法に課題がある。

表 6: グループ構造による共分散

占有率	HL	HLH	HLHL
10%	-4.429703	-47.002135	404.666598
20%	-4.400504	-34.345996	438.301228
30%	-4.584205	-19.080808	581.002253
40%	-7.559226	-67.736800	1304.699734
50%	-9.038803	-126.033382	2861.160377
60%	-11.180389	-348.203568	2513.700159
70%	-13.441557	-366.967778	5153.716540
80%	-16.161928	-655.604857	-348.482203
90%	-15.890356	-568.180316	405.444337

5. まとめ

本稿では、MSP 検索システムの提案を行い、また、動画の時間的特徴量を利用した動画検索方法を提案した。評価実験により本提案手法の有効性を示した。動的変化の構造を表現できないシーケンスに対しての評価は今後の課題である。

参考文献

- [1] 児玉 明, 笠井 裕之, 村井 正人, 富永 英義: “マルチメディアプロシーディングとその情報構成に関する検討”, 信学技報, IN 96-122, OFS 96-10 (1997).
- [2] 児玉 明, 竹本 正行, 池田 朋二, 真崎 剛: “MSP 画像検索システムにおける情報構成法の考察”, PCSJ 99, P-3.08, pp. 53-54 (1999).