

## 映像検索エージェントの検討

4U-3

伏木田 勝信 樋渡 良継 脇 英世<sup>†</sup>

通信・放送機構 東京臨海部リサーチセンター †東京電機大学

### 1. はじめに

インターネット上の映像（ビデオ）データベースが急激に増加しつつあり、今後、これら映像コンテンツの検索に対する要求も大きくなってくると予想される。これまで、曖昧画像検索（Content-based image retrieval）方式の検討を行ってきたが [1~5]、今回、ネットワーク上に分散している映像データベースを検索する方式として映像検索エージェント（VQA：Video Query Agent）についての提案を行い、映像検索エンジンの特性等の検討を行ったので報告する。

### 2. 映像検索エージェント（VQA）の概念

ビデオの検索においては、標準的な映像検索方式が確立されていないことや通信容量の制約などが大きな問題となる。ここで提案するVQAの使用環境は、例えば、図1に示すように、インターネット上のビデオデータベースサイト（VDBS）、VQAを配信するビデオサーチサイト（VSS）、そして検索要求を出すクライアントから構成される。

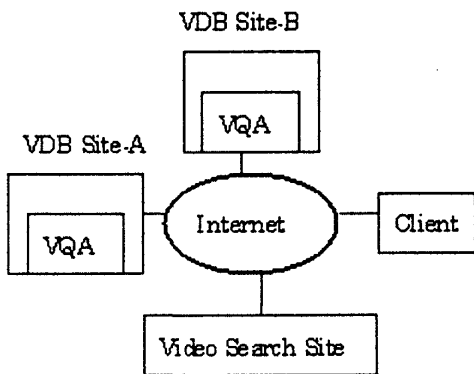


図1 映像検索エージェント動作概念図

VQAを用いた映像検索を行う手順例を以下に示す。

- (1) クライアントからVSSへの検索要求。
- (2) VSSによるVQAのVDBSへの移送。

(3) VQAの起動、インデックシング、映像検索実行と検索結果のクライアントへの転送。

(4) クライアント-VQA（VDBS上）間のインタラクティブ映像検索とブラウジング。

(5) VQAの自身による自動消去。

また、複数のVDBSに対してマルチエージェントによる協調的な検索系も構成することができる。

### 3. VQAのアーキテクチャー

VQAのアーキテクチャーは、図2に示すように、自動インデックス作成部、映像検索部、クライアントとの通信制御部等を内臓しており、インテリジェントコントローラーにより全体が制御される。

VQAに要求される特性としては、(1) プラットフォームに依存しないこと、(2) VDBSのマシンの性能・利用可能メモリ量、通信容量等によって処理の精度を変更できる柔軟性をもつこと、

(3) さまざまな種類のビデオに適應できる汎用性を持ちロバストであること、などがあげられる。

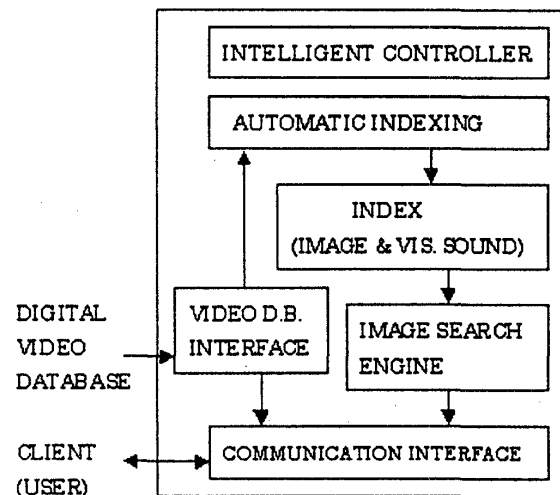


図2 VQAのアーキテクチャー

### 4. インデックス作成方式

上記のVQAに対する要求にしたがい、インデックシング方式としては、N個のビデオフレームを含むブロックに対して、各ビデオフレームの特徴ベクトル系列を最適に矩形波近似するI個の代表フレームをダイナミック・プログラミング法（DP法）

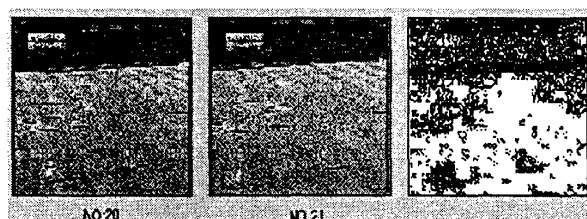
Studies toward video query agent / Katsunobu Fushikida, Yoshitsugu Hiwatari and †Hideyo Waki / Tokyo Waterfront Research Center, TAO Japan  
 †Tokyo Denki University

を用いて選択しインデックスとして用いる方式を採用した [1]。この方式は、ロバストであり、I を与えることによりインデックスDBの規模を容易に制御することができる。

## 5. VQA用映像検索エンジン

検索エンジンとしては、処理量の制御が容易で簡易な Feature-base 型の画像検索エンジンを用いた。特徴ベクトルとしては、画面を  $L \times M$  分割し、各分割領域での RGB ヒストグラムを用いた。従って、特徴ベクトルの次元数は  $L \times M \times 3$  である。

また、ビデオの動的な特徴を用いる簡易な検索方式として、隣接するビデオフレーム間の各ピクセルの相違パターンを用いる方式についても検討を行った。相違パターンの例を図3に示す。相違パターン間のマッチングは上記の検索方式を用いて行われた。本方式は、色ずれによる影響を受け難いという特徴も持っている。



先行フレーム 後続フレーム 相違パターン

図3 相違パターン例 (サッカービデオ)

## 6. 映像検索実験と結果

### ・検索エンジンの特性評価

$L \times M$  を変化させたときの音像パターン [3,5] (300 枚の音声と鳥の鳴声等を用いた) による検索実験を行った。検索率を表1に示す。これらの結果では、 $L \times M = 64$  程度の分割数でほぼ飽和している。

表1 分割数 ( $L \times M$ ) を変えた時の音像検索結果

$L \times M$	1 x 1	2 x 2	4 x 4	8 x 8	16 x 16
音声	86	94	97	98	98
鳥	91	95	96	96	97

$$\text{検索率} = ((\text{Recall rank}) / (\text{DB size})) \times 100(\%)$$

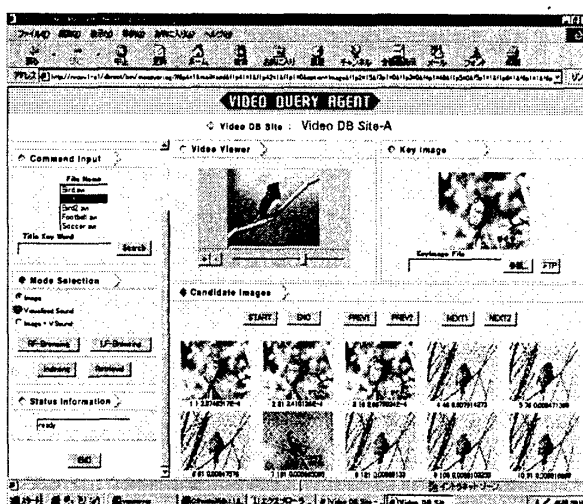
### ・フレーム間相違パターンによる検索方式

前記のインデックス作成方式で選択された、相違パターン系列の代表フレームをターゲット画像とし、キー画像としてビデオからランダムに選ばれたフレームを用いて検索実験を行った。映像DBとしては、スポーツビデオ (サッカー、野球)、野鳥ビデオな

ど約100秒の (10フレーム/秒) のビデオサンプルを用いた。平均検索率は、約92%であった。

### ・VQAによる検索実験

Javaを用いて (一部、映像データ転送部分等はC++を用いた) VQAを作成し、イントラネット (上に映像検索実験システムを構築し (PC (266MHzCPU)・Windows NT/95を使用)、検索実験を行った。VQAの起動に際してはRMI (Remote Method Invocation) を用いた。検索された候補画像や検索モード選択メニュー等をクライアント側で表示したVQAの画面例を図4に示す。インタラクティブな検索実験の結果、映像検索処理はJavaによるシステムでも、実時間で効果的な検索が可能であることが確かめられた。



左側は検索モード等の選択メニュー、右下は候補画像  
図4 VQA画面例

## 7. まとめ

ネットワーク上の分散映像データベースの検索効率良く行う映像検索エージェント (VQA) を提案し、実験システムによる検討を行い有効性を示すとともに、VQAに適した映像検索エンジンについての検索実験を行い、その効果を確認した。

参考文献: [1] 伏木田、最適矩形波近似による代表ビデオフレーム選択、情報処理学会57全大、5P-1, 1998-9. [2] 伏木田、樋渡、イントラネットを利用した対話型分散画像検索システム、電子情報通信学会総合大会'97, 1997-10. [3] 樋渡、伏木田、マルチモーダルインデックスによるパラレル画像検索方式、DICOM '98, 1998-7. [4] K. Fushikida, Y. Hiwatari, H. Waki, Content-based Image Query Method using Parallel Retrieval Scheme, ICCIMA '98, 1998-2. [5] K. Fushikida, Y. Hiwatari, H. Waki, Content-Based Video Retrieval Method using a Visualized Sound Pattern, IFIP VDB-4, 1998-5.