

PHÂN TÍCH TỰ ĐỘNG DỮ LIỆU VIDEO SỐ DỰA TRÊN MÔ HÌNH PHÂN CẤP DỮ LIỆU

NGUYỄN LÂM¹, LÝ QUỐC NGỌC²

¹ Trường Đại học Lương Thế Vinh, Tp. Nam Định

² Trường Đại học Khoa học tự nhiên - ĐHQG Tp. HCM

Abstract. Nowadays, digital video documents are stored and grown in the number and the size. Therefore, it requires more efficient management techniques that allow retrieving visual information in an efficient way. Manual management is not appropriate for a huge volume data as that of digital video data, so a trend of research is rapidly developing for automatically classifying digital video data based on content.

Following this trend, in this paper, we represent a method for automatically analysing structure of digital video data to generate the table of content (ToC) and the index table. Basing on that structure we can automatically enrich firstly video data with semantic labels.

We solve this problem by studying the decomposition of the video sequence into elementary segments, then merging them into hierarchical diagram based on Binary Partition Tree and finally generating the ToC and the index table from binary partition tree.

We experimented on the sports video, documentary video, and the results showed that ToC and the index table contained many useful information for retrieving visual information more efficiently.

Based on this structure, we can:

- + Retrieve visual information.
- + Browse following the structure of video documents.
- + Filter the content of video documents.

Tóm tắt. Hiện nay dữ liệu video số được lưu trữ và phát triển với số lượng ngày càng tăng. Vì vậy dẫn đến một nhu cầu là cần có một cách thức quản lý hữu hiệu hơn để phục vụ việc truy tìm thông tin. Việc quản lý bằng phương pháp thủ công sẽ không còn thích hợp đối với lượng dữ liệu cực lớn như video số, do đó hiện nay tồn tại một hướng nghiên cứu đang phát triển mạnh mẽ nhằm đáp ứng việc phân loại tự động dữ liệu video số dựa vào nội dung.

Nằm trong xu hướng đó, bài báo trình bày một phương pháp giúp phân tích tự động cấu trúc của dữ liệu video số nhằm tạo ra bản mục lục và chỉ mục cho dữ liệu video số, làm cơ sở để có thể đi xa hơn nữa trong việc tự động làm giàu thêm thông tin bằng cách gán nhãn ngữ nghĩa vào cấu trúc được rút trích.

Bài toán được tiếp cận bằng việc đầu tiên là phân tích tự động video số thành các đoạn cơ sở, sau đó nhóm chúng lại theo một sơ đồ phân cấp dựa vào cây nhị phân và cuối cùng là rút gọn cây để tạo bảng mục lục và chỉ mục.

Kết quả được thử nghiệm trên các đoạn video về thể thao, tin tức thời sự, kết quả cho thấy bảng mục lục và chỉ mục chứa đựng nhiều thông tin có ý nghĩa giúp việc truy tìm thông tin hiệu quả hơn. Dựa trên đặc tả này, chúng ta có thể:

- + Truy tìm thông tin.

- + Duyệt theo cấu trúc của tài liệu.
- + Lọc nội dung tương ứng với các đặc trưng.

1. GIỚI THIỆU

Dữ liệu hình ảnh là một dạng dữ liệu đặc biệt, con người có được may mắn lớn là có thể nhận biết và hiểu chúng mà hầu như không cần diễn dịch. Tuy nhiên họ lại không biết được tại sao họ lại làm được điều này. Khi dạng dữ liệu này được số hóa thì việc thiết kế các giải thuật để máy tính có thể truy tìm chúng dựa vào nội dung là công việc đầy thử thách và thú vị. Một trong những thử thách chính là việc lấp hố ngăn cách giữa một bên là các đặc trưng cấp thấp của ảnh mà cụ thể là giá trị màu tại điểm ảnh và một bên là ngữ nghĩa của ảnh.

Bài báo trình bày một phần việc trong công đoạn tiền xử lý của quá trình truy tìm thông tin thị giác dựa vào nội dung.

Với trình độ khoa học kỹ thuật hiện nay, chúng ta không nên đòi hỏi một phương pháp vạn năng áp dụng được cho mọi đoạn video với đủ loại nội dung, vì vậy chúng tôi mạnh dạn đề xuất một phương pháp giúp tự động rút trích cấu trúc đoạn video dưới dạng bảng mục lục và chỉ mục dựa trên đặc trưng màu và đặc trưng chuyển động, điều này có sự mô phỏng gần giống với bảng mục lục và chỉ mục của một cuốn sách. Với dữ liệu nhập là đoạn video thì chúng tôi cho kết xuất là bảng mục lục và chỉ mục của đoạn video đó, điều này cũng phù hợp với xu hướng lưu trữ đặc tả nội dung của chuẩn lưu trữ MPEG-7.

Sơ đồ tổng quát của quá trình tự động tạo bảng mục lục và chỉ mục: Phân đoạn tuần tự → Tạo cấu trúc phân cấp → Tạo bảng mục lục và chỉ mục.

2. PHÂN ĐOẠN TUẦN TỰ

2.1. Định nghĩa về đoạn cơ sở

Đoạn cơ sở là đoạn bao gồm các khung hình liên tiếp nhau theo thời gian được giới hạn giữa hai chuyển cảnh.

Hai dạng chuyển cảnh thường được khảo sát trong đoạn video: Chuyển cảnh tức thì và chuyển cảnh tiệm tiến.

Với định nghĩa này đoạn cơ sở thường mang một nội dung riêng biệt, làm cơ sở cho các bước phân tích cấu trúc đoạn video số ở cấp độ cao hơn.

2.2. Phương pháp tìm đoạn cơ sở

Việc tìm biên giới hạn của đoạn cơ sở dựa vào đặc tính về độ biến thiên về màu sắc không lớn giữa các khung hình trong đoạn cơ sở và độ biến thiên đáng kể tại các biên của đoạn cơ sở.

Các phương pháp tìm đoạn cơ sở thường dùng:

- + Phương pháp so sánh điểm ảnh.
- + Phương pháp so sánh khối điểm ảnh.
- + Phương pháp so sánh lược đồ màu toàn cục.
- + Phương pháp so sánh lược đồ màu cục bộ.

Các phương pháp này dựa vào độ đo tương tự giữa các khung hình kề nhau, khi độ đo này vượt một ngưỡng cho trước thì biên của đoạn cơ sở sẽ được xác lập.

2.2.1. Phương pháp dựa vào so sánh điểm ảnh

Đại lượng được dùng để phát hiện biên của đoạn cơ sở là độ đo sự sai biệt giữa 2 khung hình liên tiếp nhau tại thời điểm t , gọi là $FD(t)$ (đối với ảnh độ xám) và $FD^{mau}(t)$ (đối với ảnh màu).

$$FD_t = I_{t+1} - I_t, \quad FD(t) = \frac{1}{n_x n_y} \sum_{x=0}^{n_x} \sum_{y=0}^{n_y} |FD_t(x, y)|, \quad FD^{mau}(t) = \frac{1}{3} \sum_{i=1}^3 |FD^{C_i}(t)|$$

Giải thuật này phát hiện chuyển cảnh khi độ đo sai biệt giữa hai khung hình vượt quá một ngưỡng cho trước.

Khuyết điểm của phương pháp này là không đối phó được với hai hiện tượng thường gặp trong đoạn video là sự dịch chuyển của đối tượng và sự dịch chuyển của camera.

2.2.2. Phương pháp dựa vào so sánh lược đồ màu toàn cục

Đại lượng được dùng để phát hiện biên của đoạn cơ sở là độ đo sự sai khác giữa 2 lược đồ màu trong 2 khung hình liên tiếp nhau tại thời điểm t và $t + 1$, gọi là $HD^{mau}(t)$:

$$HD^{mau}(t) = \frac{1}{M} \sum_{m=1}^M |h_{t+1}^{mau}[m] - h_t^{mau}[m]|$$

Giải thuật này phát hiện chuyển cảnh khi $HD^{mau}(t)$ vượt quá một ngưỡng cho trước. Ưu điểm của phương pháp này là đối phó được với 2 hiện tượng thường gặp trong đoạn video là sự dịch chuyển của đối tượng và sự dịch chuyển của camera.

2.3. Phương pháp đề xuất

Tuy nhiên trong thực tế, các chuyển cảnh tiệm tiến thay đổi từ từ nội dung của đoạn video và do đó có thể xuất hiện dao động bé trên đường cong thể hiện sai biệt giữa các khung hình (ĐCSB). Ngoài ra nội dung của đoạn cơ sở có thể thay đổi nhanh chóng với một số lý do như sự di chuyển nhanh camera, sự dịch chuyển các đối tượng lớn, đèn chớp... và do đó có thể xuất hiện các đỉnh có giá trị lớn trên ĐCSB không tương ứng với chuyển cảnh.

Vì vậy nếu chỉ các đỉnh có giá trị lớn trên đường cong được ưu tiên phát hiện để tránh phát hiện dư thì các chuyển cảnh tiệm tiến sẽ bị bỏ qua. Ngược lại, ưu tiên phát hiện chuyển cảnh tiệm tiến sẽ dẫn đến phát hiện dư. Bài báo đã khắc phục vấn đề này dựa trên 3 bước cơ bản.

Bước 1, dùng phương pháp lược đồ tự tương quan màu để xây dựng ĐCSB. (Phương pháp này giúp khắc phục khuyết điểm của hai phương pháp trên trong việc có xét đến yếu tố không gian trong phân bố màu và đối phó được với sự dịch chuyển của đối tượng và sự dịch chuyển của camera).

Bước 2, tiến hành phân đoạn ĐCSB, sử dụng toán tử phi tuyến hình thái học nhằm loại bớt các đỉnh ảnh hưởng đến việc phân đoạn dư nhưng vẫn bảo toàn được các đỉnh thực cho chuyển cảnh tức thì và tiệm tiến. Trong bước này, chấp nhận phát hiện dư. Mục đích là tạo tập các đoạn gần với đoạn cơ sở.

Đa số các phương pháp truyền thống thử phân đoạn đường cong từ quan điểm chuyển cảnh. Đường cong được dò tìm để xác định các đỉnh dựa vào ngưỡng (tức là dò tìm theo chiều dọc của ĐCSB và đoạn cơ sở được xác định như là phần giữa các đỉnh này).

Khuyết điểm xảy ra là: nếu ngưỡng quá bé thì xảy ra tình trạng phát hiện dư và nếu ngưỡng quá lớn thì xảy ra tình trạng phát hiện thiếu. Giải quyết vấn đề này bằng cách dùng các toán tử hình thái học tác động vào ĐCSB có tác dụng như lọc phi tuyến nhằm bám sát hình dáng đường cong để dò tìm đoạn cơ sở theo chiều ngang kết hợp với chiều dọc trên ĐCSB.

Bước 3, tiến hành phát hiện các đoạn cơ sở bằng giải thuật Watershed và ghép nối các đoạn cơ sở có độ tương đồng cao và độ dị biệt thấp nhằm loại bớt các phát hiện dư ở Bước 2.

Các bước của quá trình tìm đoạn cơ sở:

Bước 1: Tạo đường đồ thị phân biệt khung hình bằng cách tính độ đo tương tự giữa các khung hình theo lược đồ tự tương quan màu:

$$CD(t) = \frac{1}{M} \sum_{m=1}^M \sum_{d=1}^D |C_{t+1}[m, d] - C_t[m, d]|,$$

$$C_t[c, d] =$$

$$\frac{1}{I_x I_y} \sum_{x=0}^{I_x-1} \sum_{y=0}^{I_y-1} \left\{ \sum_{x'=0}^{I_x-1} \sum_{y'=0}^{I_y-1} \sum_{d=1}^D \left\{ \begin{array}{l} 1 \quad \text{nếu } (\exists(x', y') : Q(I(x', y')) = c) \\ \quad \text{và } \|(x', y') - (x, y)\| = d \end{array} \right\} \right\} \left. \begin{array}{l} \text{nếu } Q(I(x, y)) = c \\ \\ \text{nếu khác} \end{array} \right\},$$

Q là hàm định lượng màu.

Điều chúng ta mong muốn là đường đồ thị có giá trị bé trong đoạn cơ sở và giá trị lớn ở biên đoạn cơ sở. Nhưng thực tế lại phức tạp hơn, đường đồ thị thường có giá trị bé trong đoạn chuyển cảnh từ từ và có giá trị lớn tại giữa đoạn cơ sở tại các cảnh có sự dịch chuyển nhanh của camera, của đối tượng kích thước lớn.

Bước 2: Lọc đường đồ thị để loại các đỉnh âm và các đỉnh dương có giá trị bé nhưng vẫn bảo toàn các đỉnh dương có giá trị lớn. Gồm các bước:

+ Lọc các đỉnh thấp (đỉnh âm) bằng toán tử phi tuyến hình thái học nhưng vẫn giữ được các đỉnh cao.

+ Loại bớt một số đỉnh (đỉnh dương) có giá trị bé bằng toán tử phi tuyến hình thái học.

Bước 3: Tìm các đoạn cơ sở và nối ghép. Gồm các bước:

+ Xác định vị trí điểm mầm trên ĐCSB.

+ Từ điểm mầm, lan rộng để loại bớt các đỉnh dương thấp theo giải thuật phân đoạn Watershed nhằm nối ghép các đoạn cơ sở bị phân cách các đỉnh ảo.

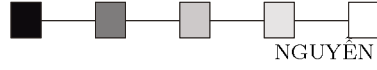
+ Nối ghép các đoạn cơ sở có độ tương đồng cao và độ dị biệt thấp.

3. XÂY DỰNG CÂY NHỊ PHÂN

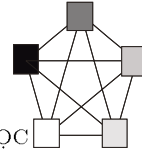
Đây là giai đoạn chuẩn bị cho giai đoạn xây dựng bảng mục lục và chỉ mục, giai đoạn này gồm các bước sau:

Bước 1: Tạo liên kết lân cận.

Có hai dạng liên kết lân cận là liên kết cho bảng mục lục và liên kết cho bảng chỉ mục. Liên kết lân cận cho bảng mục lục chịu sự ràng buộc theo chiều thời gian và liên kết lân cận cho bảng chỉ mục không chịu sự ràng buộc này.



NGUYỄN LÂM, LÝ QUỐC NGỌC



Liên kết lân cận cho bảng mục lục

Liên kết lân cận cho bảng chỉ mục

Khoảng cách giữa các nút lá trong liên kết được xác định như sau.

Chọn các đặc trưng màu và đặc trưng chuyển động:

Đặc trưng màu:

+ Lượng đồ màu trung bình của đoạn cơ sở

$$\bar{h}_s^{mau} = (\bar{h}_s^{mau}[0], \dots, \bar{h}_s^{mau}[M-1]), \quad \bar{h}_s^{mau}[m] = \frac{1}{N} \sum_{t=0}^{N-1} \bar{h}_t[m],$$

N là số khung hình trong đoạn cơ sở, M số ô màu định lượng.

+ Ảnh trung bình của đoạn cơ sở

$$\bar{I}_s^{mau} = (\bar{I}_s^h, \bar{I}_s^s, \bar{I}_s^i), \quad \bar{I}_s^h = \frac{1}{N} \sum_{t=0}^{N-1} I_t^h, \quad \bar{I}_s^s = \frac{1}{N} \sum_{t=0}^{N-1} I_t^s, \quad \bar{I}_s^i = \frac{1}{N} \sum_{t=0}^{N-1} I_t^i.$$

Đặc trưng chuyển động:

+ Lượng đồ chuyển động trung bình của đoạn cơ sở

$$\bar{h}_s^{cdong} = (\bar{h}_s^{cdong}[0], \dots, \bar{h}_s^{cdong}[D-1]), \quad \bar{h}_s^{cdong}[d] = \frac{1}{N-1} \sum_{t=0}^{N-2} \bar{h}_t^{cdong}[d],$$

N là số khung hình trong đoạn cơ sở, D là số vectơ dịch chuyển.

+ Ảnh chuyển động trung bình của đoạn cơ sở

$$\bar{M}_s = (\bar{M}_s^X, \bar{M}_s^Y), \quad \bar{M}_s^X = \frac{1}{N-1} \sum_{t=0}^{N-2} |M_t^X|, \quad \bar{M}_s^Y = \frac{1}{N-1} \sum_{t=0}^{N-2} |M_t^Y|.$$

Đặc trưng của nút lá:

$$F(l) = \{\overline{h}_s^{mau}, \overline{I}_s^{mau}, \overline{h}_s^{cdong}, \overline{M}_s\}.$$

Khoảng cách giữa hai nút lá:

Giả sử nút lá gồm các N đặc trưng:

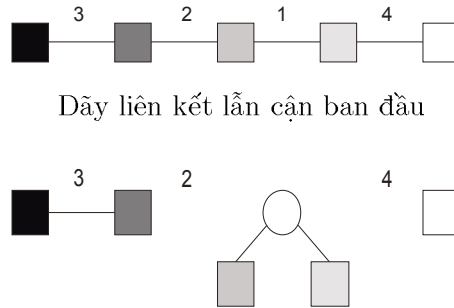
$$F(l) = \{F_1(l), \dots, F_N(l)\},$$

$$d(l_a, l_b) = \sum_{i=0}^{N-1} w_i d(F_i(l_a), F_i(l_b)),$$

$$\sum_{i=0}^{N-1} w_i = 1.$$

Bước 2: Xây dựng cây nhị phân.

Chọn 2 nút có khoảng cách bé nhất trong dãy liên kết để tạo nên nút mới. Sơ đồ liên kết:



Kết quả sau khi liên kết hai nút có khoảng cách bé nhất

Bước 3: Cập nhật liên kết và khoảng cách giữa nút mới và cũ trong liên kết.

Giả sử xét cặp nút không là nút lá n_a và n_b ,

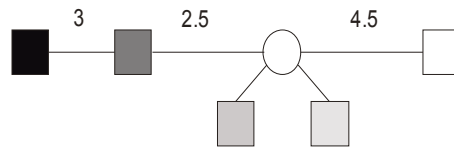
$$n_a = \{l_0, \dots, l_{p-1}\}, n_b = \{l_0, \dots, l_{q-1}\}$$

Khoảng cách giữa hai nút không là nút lá được xác định như sau:

$$d(n_a, n_b) = \{d_{\min}(n_a, n_b), d_{\max}(n_a, n_b)\},$$

$$d_{\max}(n_a, n_b) = \max\{d(l_i, l_j), l_i \in n_a, l_j \in n_b, i \in [0 \dots p-1], j \in [0 \dots q-1]\},$$

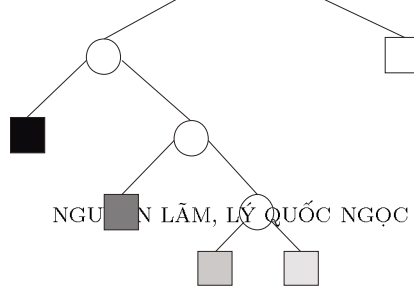
$$d_{\min}(n_a, n_b) = \min\{d(l_i, l_j), l_i \in n_a, l_j \in n_b, i \in [0 \dots p-1], j \in [0 \dots q-1]\},$$



Kết quả sau khi cập nhật liên kết và khoảng giữa nút mới và cũ

Bước 4: Lặp lại bước 2 cho đến khi tất cả các nút đều được xử lý.

Sau cùng chúng ta có được cây nhị phân liên kết các đoạn cơ sở dựa vào mức độ tương tự giữa các đoạn cơ sở.



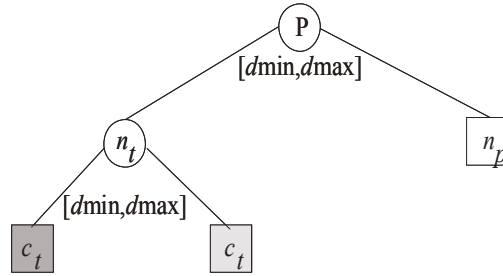
Kết quả của giai đoạn xây dựng cây nhị phân

4. XÂY DỰNG BẢNG MỤC LỤC VÀ CHỈ THỊ

Xây dựng bảng mục lục và chỉ mục từ cây phân hoạch nhị phân. Nút cần loại bỏ dựa trên tiêu chuẩn xét độ tương đồng và dị biệt với nút cha. Trong quá trình tạo cây nhị phân, ta có:

Nút p_i được tạo lập và được liên kết 2 giá trị: $d_{\min}(n_t, n_p), d_{\max}(n_t, n_p)$.

Nút n_i được tạo lập và được liên kết 2 giá trị: $d_{\min}(c_t, c_p), d_{\max}(c_t, c_p)$.



Cây nhị phân

Giả sử nút n là con của nút p , xét hai đại lượng:

$$\Delta_{\min}(p, n) = \frac{d_{\min}^p - d_{\min}^n}{d_{\min}^n}, \text{ thể hiện sự tương đồng giữa nút } p \text{ và nút } n.$$

$$\Delta_{\max}(p, n) = \frac{d_{\max}^p - d_{\max}^n}{d_{\max}^n}, \text{ thể hiện sự dị biệt giữa nút } p \text{ và nút } n.$$

Nút n được loại bỏ nếu: $\Delta_{\min} < \varepsilon$ và $\Delta_{\max} < \varepsilon$.

Khi nút n bị loại bỏ, nút con của nút n là nút c không bị mất đi mà được liên kết trở lại với nút p . Sau bước này, chiều sâu của cây nhị phân ban đầu được giảm đáng kể và ta có một cấu trúc cây mới gọi là cây phân hoạch nhị phân, với các nút không lá chứa các đoạn cơ sở với độ tương đồng cao và độ dị biệt thấp.

5. KẾT QUẢ THỰC NGHIỆM

Chúng tôi chọn thử nghiệm trên các đoạn video trong giải bóng đá Euro 2004, với 10 đoạn video, mỗi đoạn 45 phút.

Mô hình màu được chọn là mô hình màu HSI , với 12 thành phần H , 3 thành phần S , 3 thành phần I , tập màu được dùng gồm 108 màu.

Tập $[D] = \{1, 3, 5, 7\}$ là tập các khoảng cách được dùng trong trích đặc trưng tự tương quan màu của các khung hình.

Chúng tôi sử dụng hai đại lượng: độ chính xác và độ trung thực để đánh giá tính hiệu quả của hệ thống.

Độ chính xác = Số đoạn cơ sở tìm được đúng / Số đoạn cơ sở tìm được.

Độ trung thực = Số đoạn cơ sở tìm được đúng / Tổng số đoạn cơ sở thực có.

Bảng 1

Đoạn video	Phương pháp lược đồ màu toàn cục + ngưỡng toàn cục			Phương pháp lược đồ màu toàn cục + giải thuật Watershed			Phương pháp lược đồ tự tương quan màu + giải thuật Watershed		
	Số đoạn cơ sở tìm được	Độ chính xác (%)	Độ trung thực (%)	Số đoạn cơ sở tìm được	Độ chính xác (%)	Độ trung thực (%)	Số đoạn cơ sở tìm được	Độ chính xác (%)	Độ trung thực (%)
Video 1	450	83	76	450	89	82	473	92	89
Video 2	412	85	72	463	88	84	480	91	90
Video 3	389	86	70	434	90	82	460	90	87
Video 4	426	82	77	442	87	85	457	89	90
Video 5	465	87	75	478	91	81	519	92	89
Video 6	372	85	80	367	89	83	390	90	89
Video 7	378	82	75	374	90	82	402	92	90
Video 8	415	84	80	401	90	83	429	91	90
Video 9	470	87	79	486	88	83	510	90	89
Video 10	421	85	81	410	89	83	430	92	90

Đoạn video 1: Trận khai mạc Hy Lạp - Bồ Đào Nha (Hiệp 1)

Đoạn video 2: Trận khai mạc Hy Lạp - Bồ Đào Nha (Hiệp 2)

Đoạn video 3: Trận chung kết Hy Lạp - Bồ Đào Nha (Hiệp 1)

Đoạn video 4: Trận chung kết Hy Lạp - Bồ Đào Nha (Hiệp 2)

Đoạn video 5: Trận bán kết Bồ Đào Nha - Hà Lan (Hiệp 1)

Đoạn video 6: Trận bán kết Bồ Đào Nha - Hà Lan (Hiệp 2)

Đoạn video 7: Trận bán kết Hy Lạp - Tiệp Khắc (Hiệp 1)

Đoạn video 8: Trận bán kết Hy Lạp - Tiệp Khắc (Hiệp 2)

Đoạn video 9: Trận vòng loại Pháp - Anh (Hiệp 1)

Đoạn video 10: Trận vòng loại Pháp - Anh (Hiệp 2)

Kết quả phân đoạn theo thời gian như bảng 1.

6. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Mục đích chính của bài báo này nhằm trình bày phương pháp phân tích cấu trúc nội dung đoạn video số với kết xuất là cây phân hoạch nhị phân cho bảng mục lục và bảng chỉ mục. Từ một đoạn video số với lượng dữ liệu thô rất lớn, chúng ta rút ra được một cấu trúc gọn nhẹ (mục lục và chỉ mục), trên cơ sở đó có thể chủ động hỗ trợ việc tìm kiếm thông tin hiệu quả hơn với các thao tác tìm kiếm, lọc, duyệt.

Với trình độ khoa học kỹ thuật hiện tại, chúng ta chưa thể có được một công cụ phân tích cấu trúc video số phù hợp với mọi thể loại. Tuy nhiên trong một số dạng nội dung đặc

thù thì có thể thích ứng được, chẳng hạn như các đoạn video về thể thao, tin tức thời sự,...

Trong các bài báo tiếp theo, chúng tôi sẽ phát triển thêm ở các điểm sau:

- + Tìm kiếm giải thuật phân đoạn video thành các đoạn cơ sở với độ chính xác cao hơn.
- + Tìm kiếm các đặc trưng giúp cho việc tạo cây phân hoạch nhị phân hữu hiệu hơn.
- + Chuyển đổi dạng kết xuất của cây phân hoạch nhị phân sang dạng đặc tả hình thức tiện lợi hơn cho việc tra cứu, chẳng hạn như dưới dạng XML hoặc lưu trữ dưới dạng cây thư mục trên thiết bị lưu trữ.
- + Gán nhãn ngữ nghĩa vào kết xuất cuối.

TÀI LIỆU THAM KHẢO

- [1] N. Adami, A. Bugatti, A. Corghi, R. Leonardi, P. Migliorati, L. A. Rossi, and C. Sarceno, TOCAI: a framework to indexing and retrieval of multimedia documents, *IEEE Proceedings of the International Conference on Image Analysis and Processing, ICIAP'99*, Venezia, Italy, September 1999.
- [2] D. Bimbo, *Visual Information Retrieval*, Morgan Kaufmann, 1999.
- [3] I. Koprinska, S. Carrato, Temporal video segmentation: A survey. *Signal Processing: Image Communications* **16** (5) (2001) 477–500.
- [4] “MPEG, Overview of the MPEG - 7 standard”, Technical report ISO/IEC/JTC1/SC29/WG11 N4031, MPEG, Singapore, March 2001.
- [5] Lý Quốc Ngọc, Truy tìm thông tin thị giác dựa vào nội dung bằng phương pháp gán nhãn ngữ nghĩa cho ảnh, *Tạp chí Phát triển khoa học công nghệ*, tập 7 (tháng 4&5, năm 2004).
- [6] M. Schmitt, J. Mattioli, *Morphologie Mathématique*, Masson, Paris, 1994.

Nhận bài ngày 5 - 8 - 2004