

# PHÂN TÍCH TỰ ĐỘNG DỮ LIỆU VIDEO SỐ HỖ TRỢ TRUY TÌM THÔNG TIN THỊ GIÁC DỰA VÀO NỘI DUNG

Nguyễn Lâm<sup>(1)</sup>, Lý Quốc Ngọc<sup>(2)</sup>

Phan Vĩnh Phước<sup>(2)</sup>, Nguyễn Văn Kỹ Cang<sup>(2)</sup>, Nguyễn Quốc Tuấn<sup>(2)</sup>

<sup>(1)</sup>Trường Đại học Lương Thế Vinh, Nam Định

<sup>(2)</sup>Trường Đại học Khoa học Tự nhiên – ĐHQG-HCM

(Bài nhận ngày 16 tháng 11 năm 2004, hoàn chỉnh sửa chữa ngày 03 tháng 02 năm 2005)

**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 gán nhãn thủ công cùng với việc lưu trữ lượng dữ liệu thô sẽ không còn thích hợp đối với lượng dữ liệu cực lớn như video số, do đó trong bài báo này chúng tôi trình bày một phương pháp giúp phân tích tự động cấu trúc của video số nhằm tạo ra bản mục lục và chỉ mục, giúp lưu trữ nội dung đoạn video số với các đặc trưng cấp cao, kích thước lưu trữ bé, 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 phương pháp phân lớp phân cấp và cuối cùng là rút gọn cấu trúc phân cấp 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

Để thiết kế hệ thống truy tìm dữ liệu video số dựa vào nội dung, việc đầu tiên cần giải quyết là phải có phương pháp phân tích dữ liệu video số vốn rất lớn về kích thước thành các đặc trưng cấp cao với kích thước bé để tiện lưu trữ và truy tìm.

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.

Chúng tôi 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 2 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. Các phương pháp truyền thống 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 lược đồ màu toàn cục.

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.

*Phương pháp so sánh lược đồ màu toàn cục [4]*

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

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

$h_t[m]$  là lược đồ màu tại ô màu thứ  $m$  của khung hình thứ  $t$ ,  $M$  là số ô màu trong lược đồ màu.

Giải thuật này phát hiện chuyển cảnh khi  $HD(t)$  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 bé của đối tượng và sự dịch chuyển bé của camera.

Khuyết điểm phương pháp này là giá trị của đường cong sai biệt thường có giá trị rất bé tại các vị trí xảy ra chuyển cảnh tiệm tiến.

Vì vậy chúng tôi sẽ sử dụng lược đồ tự tương tự màu nhằm làm lộ rõ độ sai biệt tại vị trí xảy ra chuyển cảnh trên đường cong sai biệt.

## 2.3. Phương pháp đề xuất

Phương pháp kết hợp lược đồ tự tương tự màu và giải thuật Watershed.

Phương pháp giúp tăng độ chính xác (mức độ tìm đúng) và độ trung thực (mức độ tìm sót).

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.

Chúng tôi đề xuất giải pháp khắc phục vấn đề trên dựa trên 3 bước cơ bản.

Trong bước 1, sử dụng lược đồ tự tương quan màu để xây dựng ĐCSB.

Phương pháp này giúp khắc phục khuyết điểm của lược đồ màu vì có xét đến yếu tố không gian trong phân bố màu, đối phó được với sự dịch chuyển của đối tượng và sự dịch chuyển của camera, làm lộ rõ độ sai biệt giữa các chuyển cảnh).

Trong bước 2, tiến hành phân đoạn ĐCSB, sử dụng toán tử phi tuyến hình thái học nhằm loại bỏ 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 phân đoạn đường cong dựa trên 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.

Chúng tôi giải quyết 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.

Trong bước 3, tiến hành phát hiện các đoạn cơ sở bằng giải thuật Watershed [6] 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.

### 2.3.1. Xây dựng đường cong sai biệt dựa trên lược đồ tự tương quan màu

Tạo ĐCSB với độ đo tương tự giữa các khung hình theo lược đồ tự tương quan màu [2].

Trong bước này chúng tôi áp dụng lược đồ tự tương quan màu thay thế lược đồ màu nhằm khắc phục nhược điểm về tính cục bộ trong phân bố màu của lược đồ màu (2 ảnh có lược đồ màu tương tự nhưng có thể rất khác nhau) và làm lộ rõ độ sai biệt tại vị trí xảy ra chuyển cảnh, nhằm phát hiện chính xác hơn các đoạn cơ sở.

Gọi  $[D]$  là tập gồm  $D$  khoảng cách  $d_1, \dots, d_D$  được đo bằng độ đo  $L_\infty$ ,

Lược đồ tự tương quan màu của ảnh  $I$  được xác định với cặp màu  $(c_i, c_j)$  và khoảng cách  $d$  như sau :

$$\gamma_{c_i, c_j}^{(d)}(I) = \Pr_{p_1 \in I_{c_i}, p_2 \in I_{c_j}} [p_2 \in I_{c_j} \mid |p_1 - p_2| = d],$$

trong đó :

$I$  là ảnh kích thước  $M \times N$  (điểm ảnh),  $I_c = \{p \in I \mid I(p) = c\}$

Lược đồ tương quan màu thể hiện xác suất cặp điểm ảnh bất kỳ  $p_1$  và  $p_2$  chịu sự ràng buộc về màu ( $p_1$  có màu  $c_i$ ,  $p_2$  có màu  $c_j$ ) và vị trí ( $|p_1 - p_2| = d$ ).

Để giảm kích thước lược đồ tương quan màu mà không ảnh hưởng đến kết quả khảo sát, chúng tôi sử dụng tập con của lược đồ này là lược đồ tự tương quan màu.

Lược đồ tự tương quan màu được xác định như sau:

$$\alpha_c^{(d)}(I) = \gamma_{c,c}^{(d)}(I)$$

Sử dụng độ đo  $L_1$  đo độ sai biệt giữa khung hình  $t$  và  $t+1$  dựa vào lược đồ tự tương quan màu

$$CD(t) = \sum_{c \in [C], d \in [D]} |\alpha_c^{(d)}(I_{t+1}) - \alpha_c^{(d)}(I_t)|,$$

trong đó:

$[C]$  là tập các màu được sử dụng để định lượng màu,

$[D]$  là tập các khoảng cách được dùng để đánh giá mức độ phân bố màu cục bộ,

$\alpha_c^{(d)}(I_t)$  là lược đồ tự tương quan màu của khung hình thứ  $t$  ứng với màu  $c$  và khoảng cách  $d$

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.

### 2.3.2. Lọc đường đồ thị bằng các toán tử hình thái học

Nhằm mở rộng vùng ứng với các đỉnh âm và đỉnh dương có giá trị bé.

Gồm các bước:

- Lọc các đỉnh âm bằng toán tử phi tuyến hình thái học nhưng vẫn giữ được các đỉnh dương. Chúng tôi chọn toán tử đóng (Closing operator) [6] để thực hiện việc này.

$$CD_{fp}(t) = \varphi_w(CD(t)),$$

trong đó:

$$\varphi_w(f(x)) = \varepsilon_w(\delta_w(f(x))), \delta_w(f(x)) = \max_{b \in W} f(x+b), \varepsilon_w(f(x)) = \min_{b \in W} f(x+b)$$

$w$  là cửa sổ lân cận quanh vị trí đang xét  $x$ ,

$CD(t), CD_{fp}(t)$  là đường cong sai biệt trước và sau khi lọc đỉnh âm

- Loại bớt một số đỉnh dương có giá trị bé bằng toán tử phi tuyến hình thái học. Chúng tôi chọn toán tử tái cấu trúc mở (Opening reconstruction operator) [6] để thực hiện việc này

$$CD_{fp}(t) = \gamma_h(CD_{fp}(t), CD_{fp}(t) - h)$$

trong đó:

$$\gamma_w(f(x), f(x) - h) = (\delta_w(f(x) - h) \wedge f(x)), \delta_w(f(x)) = \max_{b \in W} f(x+b),$$

$w$  là lân cận quanh vị trí đang xét  $x$ ,  $\wedge$  là toán tử min,

$CD_{fp}(t)$  là ĐCSB sau khi lọc đỉnh âm,  $CD_{fp}(t)$  là ĐCSB sau khi lọc đỉnh dương,

$h$  là đại lượng sẽ được tính trong thực nghiệm, thể hiện độ cao của đỉnh cần loại bỏ

### 2.3.3. Tìm các đoạn cơ sở và nối ghép.

Trong bước này chúng tôi áp dụng giải thuật phân đoạn Watershed [6] cho ĐCSB sau khi lọc đỉnh dương. Tiến hành xác định vị trí các điểm mầm trên ĐCSB  $CD_{fp}(t)$  dựa vào các đoạn có độ biến thiên thấp, các điểm mầm này là chính là các khung hình đại diện trong đoạn cơ sở cần tìm.

Từ các điểm mầm, lan rộng sang 2 phía, 2 điểm mầm liên tiếp nhau sẽ gặp nhau tại vị trí có giá trị lớn nhất giữa 2 điểm mầm trên đường cong sai biệt  $CD_{fp}(t)$ , trong quá trình lan rộng theo nguyên tắc trên đã loại được các đỉnh ảo. Đoạn cơ sở chính là đoạn ở giữa các vị trí này.

Sau đó cần thực hiện nối ghép các đoạn cơ sở có độ tương đồng cao và độ dị biệt thấp nhằm loại bớt các đoạn cơ sở thừa ở bước trước.

## 3. Xây dựng phân lớp phân cấp

Đây là giai đoạn chuẩn bị cho giai đoạn xây dựng bảng mục lục và chỉ mục.

Giai đoạn này thực chất là giai đoạn phân lớp các đoạn cơ sở. Chúng tôi chọn phương pháp phân lớp phân cấp [5] vì nó rất phù hợp với đặc tính phân cấp về nội dung của video số.

### 3.1 Giải thuật phân lớp phân cấp

Giả sử  $X = \{x_i, i = 1, 2, \dots, N\}$ , với  $x_i$  là các đoạn cơ sở.

*Giai đoạn khởi động:*

B1. Tạo phân cấp  $R_0 = \{x_i, i = 1, 2, \dots, N\}$ , với các lớp  $C_i = x_i$

B2. Đặt  $P_0 = P(X)$

Trong đó  $P(X)$  là ma trận kích thước  $N \times N$  với  $P(i, j) = d(x_i, x_j)$ ,

$d$  là độ đo sự khác biệt giữa 2 đoạn cơ sở.

**B3.  $t=0$**

*Giai đoạn phân lớp:*

B1.  $t=t+1$

B2. Chọn cặp  $C_i, C_j$  sao cho  $d(C_i, C_j) = \min d(C_r, C_s), r, s = 1, 2, \dots, N, r \neq s$

B3. Kết nạp  $C_i, C_j$  vào lớp  $C_q$  và tạo một phân cấp  $R_t$ ,

$$R_t = (R_{t-1} - \{C_i, C_j\}) \cup \{C_q\}$$

B4. Cập nhật  $P(t)$  từ  $P(t-1)$  bằng 2 bước:

- Xoá các dòng và cột tương ứng với 2 lớp vừa được kết nạp.

- Thêm dòng mới và cột mới chứa khoảng cách giữa lớp mới tạo và các lớp cũ.
- B5. Lập lại B1 khi nào các  $x_i$  đều chưa thuộc cùng một lớp.

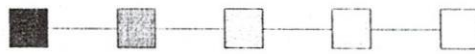
**Tạo liên kết lân cận**

Trong bước 2 tại giai đoạn phân lớp của giải thuật phân lớp phân cấp, việc chọn cặp  $C_i, C_j$  để đối sánh phụ thuộc vào dạng liên kết.

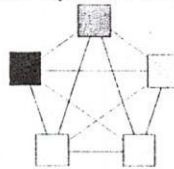
Chúng tôi khảo sát 2 dạng liên kết lân cận :

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.



Liên kết lân cận cho bảng mục lục



Liên kết lân cận cho bảng chỉ mục

**3.2. Đặc trưng và khoảng cách giữa các đoạn cơ sở trong liên kết**

Chọn các đặc trưng màu và đặc trưng chuyển động.

**Đặc trưng màu:**

- Lược đồ tự tương tự màu trung bình của đoạn cơ sở

$$\overline{cd}_s = (\overline{cd}_s[0], \dots, \overline{cd}_s[CD-1]), \quad \overline{cd}_s[m] = \frac{1}{N} \sum_{t=0}^{N-1} cd_t[m], m \in [0..CD-1],$$

N là số khung hình trong đoạn cơ sở, C là số ô màu định lượng, D là số lượng khoảng cách được dùng trong bước tính lược đồ tự tương tự màu.

$cd_t$  là lược đồ tự tương tự màu của khung hình thứ t trong đoạn cơ sở s,

$\overline{cd}_s$  là lược đồ tự tương tự màu trung bình của đoạn cơ sở s,

- Ảnh trung bình của đoạn cơ sở có N khung hình

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

$\overline{I}_s^h, \overline{I}_s^s, \overline{I}_s^i$  là ảnh trung bình đối với thành phần H,S,I của đoạn cơ sở

**Đặc trưng chuyển động:**

- Lược đồ chuyển động trung bình của đoạn cơ sở [1].

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

$h_t^m[d]$  là lược đồ chuyển động của khung hình thứ t ứng với chuyển động thứ d.

$\overline{h}_s^m[d]$  là lược đồ chuyển động trung bình của đoạn cơ sở ứng với chuyển động thứ d

N là số khung hình trong đoạn cơ sở, D là số vector dịch chuyển đã được định lượng

- Ảnh chuyển động trung bình của đoạn cơ sở [1]

$$\overline{M}_s = (\overline{M}_s^x, \overline{M}_s^y), \quad \overline{M}_s^x = \frac{1}{N-1} \sum_{t=0}^{N-2} |M_t^x|, \quad \overline{M}_s^y = \frac{1}{N-1} \sum_{t=0}^{N-2} |M_t^y|$$

$\overline{M}_s^X$  là ảnh chuyển động trung bình theo phương ngang của đoạn cơ sở

$M_i^X$  là ảnh chuyển động theo phương ngang của khung hình thứ  $t$

$\overline{M}_s^Y$  là ảnh chuyển động trung bình theo phương dọc của đoạn cơ sở

$M_i^Y$  là ảnh chuyển động theo phương dọc của khung hình thứ  $t$

- Đặc trưng của đoạn cơ sở  $x_i$ :

$$F(x_i) = \{ \overline{cd}_s, \overline{I}_s, \overline{h}_s^m, \overline{M}_s \}$$

Khoảng cách giữa 2 đoạn cơ sở:

Giả sử đoạn cơ sở gồm các  $N$  đặc trưng:

$$F(x_i) = \{ F_0(x_i), \dots, F_{N-1}(x_i) \}, F_j(x_i), j = 0..N-1 \text{ là đặc trưng thứ } j \text{ của } x_i$$

$$d(l_a, l_b) = \sum_{i=0}^{N-1} w_i d(F_i(l_a), F_i(l_b)), w_i \text{ là trọng số xác định mức độ quan trọng của đặc trưng thứ } i.$$

$d$  được chọn là khoảng cách  $L_1$

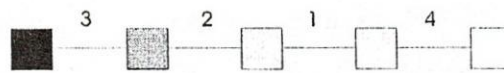
$$\sum_{i=0}^{N-1} w_i = 1,$$

### 3.3 Giai đoạn phân lớp

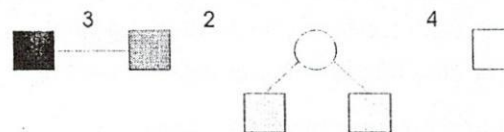
Cụ thể hoá các bước ở mục 3.1.

B1. Chọn 2 đoạn cơ sở có khoảng cách bé nhất trong dãy liên kết để tạo nên nút mới.

Sơ đồ liên kết:



Dãy liên kết lạn cận ban đầu



Kết quả sau khi liên kết 2 đoạn cơ sở có khoảng cách bé nhất

B2. Cập nhật liên kết và khoảng cách giữa nút mới và cũ trong liên kết

Giả sử xét cặp nút không là nút cơ sở  $n_a$  và  $n_b$ ,

$$n_a = \{ l_0, \dots, l_{p-1} \}, l_i \text{ là nút cơ sở, } i \in [0..p],$$

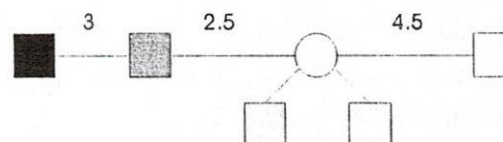
$$n_b = \{ l_0, \dots, l_{q-1} \}, l_j \text{ là nút cơ sở, } j \in [0..q]$$

Khoảng cách giữa 2 nút không là nút cơ sở được xác định như sau:

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

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

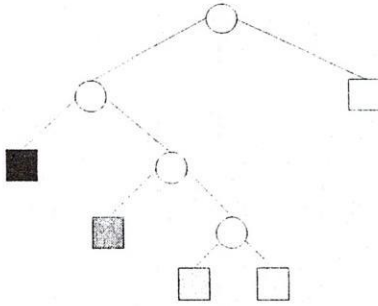
$$d_{\min}(n_a, n_b) = \min \{ d(l_i, l_j), l_i \in n_a, l_j \in n_b, i \in [0..p-1], j \in [0..q-1] \}$$



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

B3. Tiếp tục B1 cho đến khi tất cả các nút đều được xử lý

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



Kết quả của giai đoạn phân lớp phân cấp

**4. Xây dựng bảng mục lục và chỉ mục**

Trong giai đoạn này, chúng tôi có một số cải tiến đối với giải thuật phân lớp phân cấp giúp giảm chiều cao của cấu trúc phân cấp.

Nút cần loại bỏ dựa trên tiêu chuẩn xét độ tương đồng và dị biệt với nút cha.

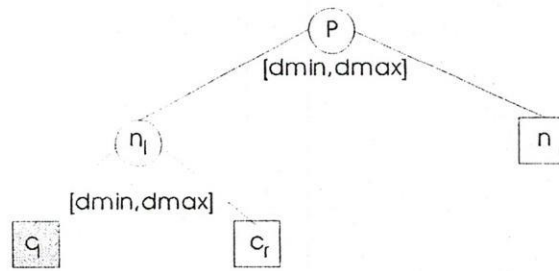
Trong quá trình tạo cấu trúc phân cấp, ta có:

Nút  $p_l$  được tạo lập và được liên kết 2 giá trị:

$$d_{\min}(n_l, n_r), d_{\max}(n_l, n_r), n_l, n_r \text{ là nút con trái và phải của nút } p_l$$

Nút  $n_l$  được tạo lập và được liên kết 2 giá trị:

$$d_{\min}(c_l, c_r), d_{\max}(c_l, c_r), c_l, c_r \text{ là nút con trái và phải của nút } n_l$$



**Cấu trúc phân cấp**

Giả sử nút  $n$  là con của nút  $p$ , xét 2 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:

$$\Delta_{\min} < \varepsilon \text{ và } |\Delta_{\max}| < \varepsilon$$

Khi nút  $n$  bị loại bỏ, nút con của nút  $n$  là nút  $c$  không bị mất đi mà được liên kết với trở lại với nút  $p$ .

Sau bước này, chiều cao của cấu trúc phân cấp ban đầu được giảm đáng kể và ta có một cấu trúc phân cấp mới, với các nút không cơ sở chứa các đoạn cơ sở với độ tương đồng cao và độ dị biệt thấp.

Tùy theo việc chọn liên kết lân cận cho bảng mục lục hoặc liên kết lân cận cho bảng chỉ mục mà ta có bảng mục lục và chỉ mục cho đoạn video.

**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 tự màu của các khung hình.

Chúng tôi sử dụng 2 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

Đoạn video 1, video 2 : Trận khai mạc Hy Lạp-Bồ Đào Nha (Hiệp 1, Hiệp 2)

Đoạn video 3, video 4 : Trận chung kết Hy Lạp-Bồ Đào Nha (Hiệp 1, Hiệp 2)

Đoạn video 5, video 6 : Trận bán kết Bồ Đào Nha-Hà Lan (Hiệp 1, Hiệp 2)

Đoạn video 7, video 8 : Trận bán kết Hy Lạp-Tiệp Khắc (Hiệp 1, Hiệp 2)

Đoạn video 9, video 10 : Trận vòng loại Pháp-Anh (Hiệp 1, Hiệp 2)

**Kết quả phân đoạn theo thời gian:**

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 tự 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

**Kết quả phân lớp phân cấp:**

Đoạn video	Số đoạn cơ sở (Phương pháp lược đồ tự tương tự màu + giải thuật Watershed)	Phân lớp phân cấp tạo cấu trúc mục lục (số lớp)	Phân lớp phân cấp tạo cấu trúc chỉ mục (số lớp)
Video 1	473	78	46
Video 2	480	80	42
Video 3	460	67	38
Video 4	457	72	49



Video 5	519	73	40
Video 6	390	65	32
Video 7	402	59	45
Video 8	429	69	50
Video 9	510	72	39
Video 10	430	66	43

Kết quả thực nghiệm cho thấy phương pháp kết hợp lược đồ tự tương tự màu và giải thuật Watershed giúp tăng độ chính xác và độ trung thực trong giai đoạn phân đoạn tuần tự theo thời gian so với các phương pháp truyền thống. Mỗi đoạn video được phân tích thành cấu trúc phân cấp mục lục và chỉ mục với số lớp bé, rất tiện cho việc truy tìm sau này.

## 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 cấu trúc phân cấp cho bảng mục lục và bảng chỉ mục. Từ một đoạn video số với lượng dữ liệu thô rất lớn, chúng tôi 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.

Chúng tôi thấy cần 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 cấu trúc phân cấp hữu hiệu hơn.
- Chuyển đổi dạng kết xuất của cấu trúc phân cấp 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.

## AUTOMATIC ANALYSIS DIGITAL VIDEO DATA TO CONTENT-BASED VISUAL INFORMATION RETRIEVAL

Nguyen Lam <sup>(1)</sup>, Ly Quoc Ngoc <sup>(2)</sup>

Phan Vinh Phuoc <sup>(2)</sup>, Nguyen Van Ky Cang <sup>(2)</sup>, Nguyen Quoc Tuan <sup>(2)</sup>

<sup>(1)</sup> Luong The Vinh University, Nam Dinh City, <sup>(2)</sup> University of Natural Sciences – VNU-HCM

**ABSTRACT:** Nowadays, digital video documents are stored and are grown in the number and the size. Therefore, it requires efficient management techniques that allow retrieving visual information in an efficient way. Traditional management by manual annotation and storing raw data is not appropriate for a huge volume data as digital video data, so in this paper we represent method for automatic structural analysis digital video to generate the table of content (ToC) and the index table, storing digital video data by high-level features in small size, based on that structure we can enriched video data with semantic labels.

We approach this problem by studying the decomposition of the video sequence into elementary segments, after that they are classified by hierarchical clustering algorithm and finally reducing hierarchical structure to generate ToC and the Index Table.

We experimented on the sports video, documentary video, the results showed that ToC and the Index Table contained 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.*

### TÀI LIỆU THAM KHẢO

- [1] Al Bovik. *Handbook of Image and Video Processing*. Academic Press, 2000
- [2] D.Bimbo. *Visual Information Retrieval*. Morgan Kaufmann, 1999
- [3] I.Koprinska and S.Carrato. *Temporal video segmentation: A survey*. *Signal Processing: Image communications*, 16(5) : 477-500, January 2001.
- [4] 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.
- [5] Sergios Theodoridis, Konstantinos Koutroumbas, *Pattern Recognition*, Academic Press, 1999
- [6] Rafael C.Gonzalez, Richard E.Woods, *Digital Image Processing. Second Edition*, Prentice - Hall, Inc. 2002.