

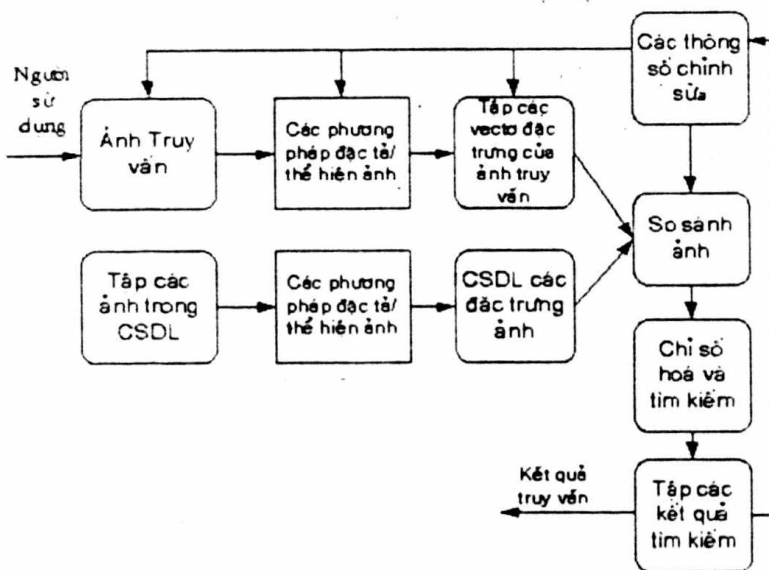
# XÂY DỰNG GIẢI THUẬT VÀ KIẾN TRÚC PHẦN MỀM TÌM KIẾM ẢNH DỰA TRÊN TẬP CÁC ĐẶC TRƯNG

Huỳnh Quyết Thắng

Khoa Công nghệ Thông tin, Đại học Bách Khoa Hà Nội

## I. ĐẶT VẤN ĐỀ

Bài toán nhận dạng và tìm kiếm ảnh dựa trên nội dung được mô tả: với một ảnh có sẵn hoặc đã được mô tả hãy tìm các ảnh tương tự trong cơ sở dữ liệu. Quy trình thực hiện nhận dạng và tìm kiếm ảnh được chỉ ra ở hình 1, trong đó hai nhiệm vụ chính cần giải quyết gồm có: (1)- Trích chọn đặc trưng ảnh, (2) So sánh ảnh [1].



Hình 1: Quy trình thực hiện nhận dạng và tìm kiếm ảnh

Nhiệm vụ 1- Trích chọn đặc trưng ảnh (Feature Extraction): Đặc trưng để nhận dạng là các thuộc tính trực quan của ảnh như màu sắc hình dạng, kết cấu, thuộc tính thống kê và thuộc tính về tần số. Khi nhận dạng có thể sử dụng một hoặc kết hợp nhiều đặc trưng. Tập các đặc trưng của ảnh được gọi là một véc tơ đặc trưng. Véc tơ đặc trưng này chính là một rút gọn của ảnh.

### (2) Nhiệm vụ 2 - So sánh ảnh

Để so sánh ảnh, người ta chuyển về bài toán so sánh độ tương tự của các đặc trưng ảnh: Có hai ảnh  $I_1, I_2$  và  $F_1, F_2$  là tập các đặc trưng tương ứng của hai ảnh này.  $F_1$  có  $n$  phần tử và  $F_2$  có  $m$  phần tử.  $F_1 = \{f_{1i} | i = 0, 1, \dots, n\}$   $F_2 = \{f_{2j} | j = 0, 1, \dots, m\}$ . Thông thường người ta chọn  $m = n$  và giá trị này là xác định. Một phép đo độ tương tự thường được sử dụng nhất là tính tỷ lệ giống nhau của hai đặc trưng tương ứng:  $d_i = |f_{1i} - f_{2i}|$  ( $i = 0, \dots, n$ ) Nếu  $d_i < d_{nguong}$  thì đặc trưng thứ  $i$  của hai ảnh là giống nhau. Để xác định hai ảnh có giống nhau không người ta xét tỷ lệ giống nhau trong tập đặc trưng. Hai ảnh thường được quan niệm là như nhau nếu có lớn hơn 60% đặc trưng giống nhau tùy thuộc và loại ảnh đang xét [1,3]. Giá trị dnguong được xác định dựa trên yêu cầu về độ chính xác, kinh nghiệm thực tiễn của người xây dựng hệ thống và tất nhiên là phụ thuộc vào loại ảnh. Khoảng cách giữa hai

tập đặc trưng được tính như sau:  $D_1 = \sum_{i=0}^n |f_{1i} - f_{2i}|$   $D_2 = \sqrt{\sum_{i=0}^n (f_{1i} - f_{2i})^2}$  Khoảng cách này có

thể xem như là nghịch đảo của độ tương tự giữa hai ảnh. Hai ảnh được coi là giống nhau nếu  $D_1$  và  $D_2$  nhỏ hơn giá trị các giá trị ngưỡng tương ứng.

## II. TRÍCH CHỌN ĐẶC TRƯNG ẢNH DỰA TRÊN MÀU SẮC VÀ LƯỢC ĐỒ MÀU

### II. 1. Biểu đồ phân bố màu sắc tổng thể - Histogram và phân bố màu sắc cục bộ trên từng vùng lưới ảnh (Grid)

Biểu đồ phân bố màu sắc tổng thể (GCH Global Color Histogram) là cách tiếp cận đơn giản và thông dụng để mã hoá thông tin về màu sắc phục vụ nhận dạng và tìm kiếm dựa trên nội dung. Một GCH được mô tả bằng một tập các số (mỗi số ứng với một màu), mỗi số có giá trị cho bởi hàm  $p(ck) = nk/n$  trong đó  $ck$  là màu thứ  $k$  trong không gian màu đã được lượng tử hoá,  $nk$  là số điểm có màu  $ck$  và  $n$  là tổng số điểm trong ảnh. Phân bố màu sắc của ảnh được biểu diễn không kèm theo một thông tin về vị trí không gian hay hình dạng của các vùng cùng màu, các ảnh hoàn toàn khác nhau về hình dạng có thể có cùng GCH. Do đó nhận dạng tìm kiếm ảnh sử dụng GCH thường cho về các kết quả sai đặc biệt là khi thực hiện trên cơ sở dữ liệu lớn.

Nếu chỉ xét phân bố màu sắc trên toàn bộ ảnh thì không có thông tin về vị trí xuất hiện của các màu trên ảnh. Do đó ảnh được chia thành từng ô nhỏ và tính histogram trên từng ô đó. Lợi điểm của cách tiếp cận này là mã hoá nhiều thông tin hơn so với GCH. Ngoài sự phân bố về màu sắc của ảnh, thêm vào đó còn có vị trí không gian của màu sắc. ảnh được phân tích thành nhiều ô có kích thước cố định (thường là không được chèn lên nhau) và với mỗi ô đó, tách ra một histogram màu sắc cục bộ của nó. Cách tính mỗi histogram màu sắc cục bộ giống như là GCH, nhưng chỉ tính cho một vùng cục bộ của một ô thay vì toàn bộ ảnh. Một ảnh được chia làm 64 ô và được biểu diễn bằng 64 histogram màu cục bộ cho mỗi ô. Khi nhận dạng, hai histogram của hai ảnh nhỏ tương ứng sẽ được so sánh với nhau. Kiểu phân tích ảnh này có thể được xem như là một trong các cách cơ bản để phân đoạn ảnh, tuy rằng có nhiều phương án phân lưới khác nhau [3].

### II.2. Biểu đồ phân bố theo không gian của từng màu - Color Shape Histogram

Cách tiếp cận này có thể xem như là mở rộng của của cách tiếp cận phân lưới. Cách này làm giảm không gian của lưới qua việc lợi dụng đặc điểm chỉ có một số tương đối nhỏ các màu xuất hiện trong mỗi ảnh. Sử dụng ý tưởng phân bố theo không gian của từng màu, màu sắc không xuất hiện trong ảnh thì không cần phải biểu diễn, do đó đặc tính của ảnh được tách càng gọn. Xét một ảnh được chia thành các ô không đè lên nhau như cách tiếp cận lưới. Một CSH cho một màu  $C$  là một tập các số (mỗi số cho một ô) và giá trị của nó được tính  $p(\text{cell}k) = nk/n$ . Trong hàm này,  $\text{cell}k$  là ô thứ  $k$  của ảnh,

nk là số điểm của ô thứ k với màu C và n là tổng số điểm trong ảnh. Một ảnh được tạo thành từ N màu sẽ được mô tả bởi N CSHs, mỗi cái mô tả phân bố vị trí của một màu. Theo cách phân tích này, nếu một màu không xuất hiện trong ảnh, thì toàn bộ histogram không cần phải biểu diễn hay lưu trữ. Số màu xuất hiện trong ảnh ít hơn rất nhiều so với tổng số màu trong không gian ảnh. Hơn nữa cách tiếp cận CS có thể biểu diễn trực tiếp vị trí xuất hiện của mỗi màu sắc trong ảnh, trong khi đó cách tiếp cận lưới cần bước xử lý nữa mới lấy được các thông tin này.

### II.3. Phân tích Wavelet

Phân tích wavelet tương tự như phân tích Fourier theo nghĩa là chia tín hiệu thành phần tử để phân tích. Biến đổi wavelet tách tín hiệu thành các “wavelet” của nó, được co dãn, dịch từ “wavelet gốc” (mother wavelet). Điều kiện để một hàm có thể thực hiện phân tích Wavelet là tích phân bình phương của wavelet hữu hạn, nghĩa là mọi hàm  $\psi(t)$  thỏa mãn:  $\int_R |\psi(t)|^2 dt < \infty$  có thể được xem như một wavelet gốc. Họ của hàm

wavelet  $\psi_{a,b}(t)$  tạo thành một cơ sở cho không gian Lebesgue  $L^2(\mathbb{R})$  có thể tạo thành từ hàm wavelet gốc  $\psi(t)$  như sau: 
$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad a > 0, b \in \mathbb{R}$$

Lợi thế của đặc trưng này là có thể xây dựng một không gian đo tính toán nhanh, yêu cầu không gian lưu trữ nhỏ cho mỗi ảnh dữ liệu. Phân tích wavelet hai chiều của ảnh sẽ cung cấp một cơ sở tốt để xây dựng không gian này. Phân tích wavelet cho phép sự xấp xỉ ảnh rất tốt với rất ít hệ số. Thông thường theo mô hình này, chỉ các hệ số wavelet có giá trị cao nhất mới được sử dụng. Phân tích wavelet có thể được sử dụng để tách và mã thông tin về biên ảnh. Các hệ số phân tích wavelet cung cấp thông tin phụ thuộc vào độ phân giải của ảnh gốc. Do đó mô hình dựa trên wavelet cho phép ảnh truy vấn và ảnh dữ liệu được tách rất hiệu quả. Phân tích wavelet rất nhanh và dễ tính toán, yêu cầu thời gian là tuyến tính với kích thước ảnh [10].

Trong [4, 6] đã đề xuất thuật toán: Trong bước tách đặc trưng, mỗi ảnh trong cơ sở dữ liệu được phân tích bằng phép biểu diễn wavelet chuẩn hai chiều với hàm gốc Haar, và chỉ ghi lại trung bình màu tổng thể và chỉ số và giá trị của m hệ số wavelet có biên độ lớn nhất. Các chỉ số cho toàn bộ cơ sở dữ liệu ảnh sau đó được tổ chức thành một cấu trúc dữ liệu đơn riêng trong chương trình để tối ưu quá trình tìm kiếm. Sau đó, mỗi ảnh truy vấn cũng được thực hiện cùng một phân tích wavelet như vậy, và bỏ qua tất cả các hệ số trừ m hệ số trung bình màu lớn nhất.

Một phân tích wavelet hai chiều chuẩn theo hàm gốc Haar rất đơn giản để cài đặt. Nó bao gồm phân tích một chiều trên mỗi hàng của ảnh, tiếp theo phân tích một chiều theo các cột của kết quả. Phân tích được trình bày trong giải thuật 1 [4].

<pre> Proc WaveletTransform(A:array[0..h-1] of color) A = A/ sqrt(h) While h&gt;1 do: ` h = h/2 for i=0 to h-1 do: A'[i] =(A[2i]+A[2i+1])/sqrt(2) A'[h+i] =(A[2i]-A[2i+1])/sqrt(2) end for A = A' End while End proc </pre>	<pre> Proc DecomposeImage(T:array[0..r-1,0..r-1] of color) For row = 1 to r do WaveletTransform(T[row,0..r-1]) End for For col = 1 to r do WaveletTransform(T[0..r-1,col]) End for End proc </pre>
<p><b>Giải thuật 1. Phân tích wavelet theo Haar</b></p>	<p><b>Giải thuật 2. Phân tích ảnh sử dụng wavelet</b></p>

Trong đoạn giả mã trên ở giải thuật 1, đầu vào A được giả sử là các thành phần màu ba chiều, mỗi chiều chạy trong khoảng [0,1]. Số phép tính số học được thực hiện trên mỗi thành phần màu là khác nhau. Toàn bộ ảnh T kích thước r x r có thể được phân tích như ở giải thuật 2. Sau quá trình phân tích, T[0,0] là tỷ lệ với trung bình màu của toàn ảnh, trong các phần khác của T chứa hệ số wavelet. Cuối cùng, chỉ T[0,0] và chỉ số và dấu của m hệ số wavelet lớn nhất của T được giữ lại. Để tối ưu tiến trình tìm kiếm, m hệ số wavelet còn lại được tổ chức thành một tập 6 mảng, gọi là mảng tìm kiếm, mỗi mảng là tổ hợp của một dấu (+ hoặc -) và kênh màu (R,G,B).

#### II.4. Một số kỹ thuật đề xuất nhằm cải tiến thuật toán

Các nhận xét:

(1) Nhận dạng bằng color shape khá chính xác nhưng khối lượng tính toán gấp nhiều lần nhận dạng bằng histogram.

(2) Một đặc điểm quan trọng của nhận dạng bằng histogram như đã nói là không bỏ sót. Nếu ảnh dữ liệu giống ảnh truy vấn sẽ chắc chắn có trong kết quả.

**Kết luận:** Nếu dùng histogram để nhận dạng thô thì sẽ cho phép thu hẹp không gian tìm kiếm rất nhiều. Trên kết quả của tìm kiếm bằng histogram sẽ thực hiện tìm kiếm bằng Color Shape hoặc Wavelet.

Như vậy quá trình thực hiện tìm kiếm được thực hiện theo mô hình phân cấp như sau:

**Bước 1.** Tìm kiếm thô: áp dụng các bước

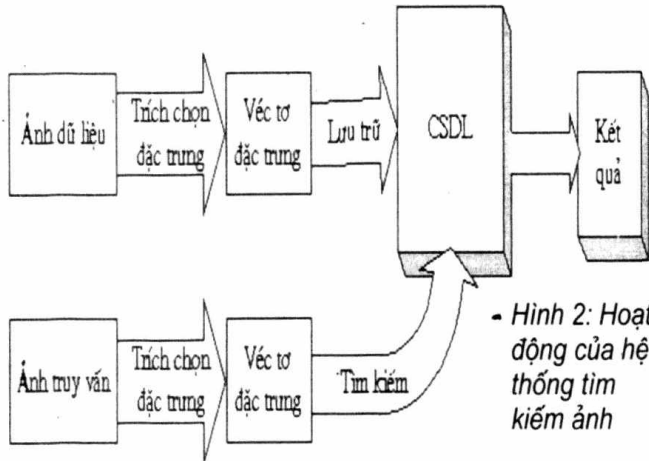
- 1) Tìm kiếm bằng histogram tổng thể của ảnh
- 2) Sắp xếp kết quả theo độ giống nhau.

**Bước 2.** Tìm kiếm chính xác: áp dụng các phương pháp

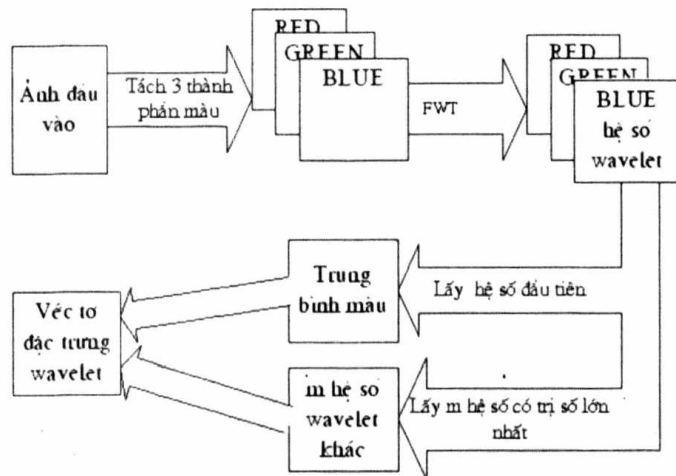
- 1) Thuật toán nhận dạng dựa trên color shape
- 2) So sánh nhận dựa trên các hệ số wavelet.

Trên cơ sở mô hình này chúng ta có thể tổng hợp thành quy trình tìm kiếm ảnh dựa trên hai bước. Chi tiết của thuật toán sẽ được mô tả trong mục III.

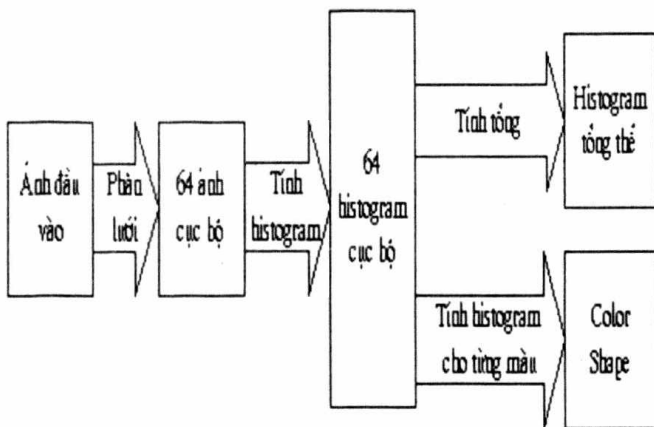
### III. MÔ TẢ THUẬT TOÁN TỔNG THỂ TÍCH HỢP VỚI CÁC CẢI TIẾN ĐỀ XUẤT



- Hình 2: Hoạt động của hệ thống tìm kiếm ảnh



Hình 3: Quá trình tách đặc trưng wavelet của ảnh



Hình 4: Quá trình tính đặc trưng về màu sắc: Histogram và ColorShape

Quy trình tách đặc trưng trong hai giai đoạn tạo cơ sở dữ liệu và trước khi truy vấn là như nhau. Đặc trưng sử dụng bao gồm đặc trưng về màu sắc, và biểu diễn đa phân giải sử dụng biến đổi wavelet (Xem hình 2). Hình 3 minh họa quá trình tách đặc trưng wavelet và hình 4 minh họa quá trình tách đặc trưng về màu sắc chúng tôi đã sử dụng.

Đặc trưng về histogram tổng thể có thể được đi qua một bước lọc nhiễu để tăng độ chính xác khi nhận dạng. Các đặc trưng nhận dạng được thiết kế thành lớp. Mỗi lớp bao gồm các thuộc tính là giá trị các đặc trưng, các hàm để tách đặc trưng từ ảnh và hàm để so sánh hai ảnh. Đặc trưng về màu sắc bao gồm histogram và colorshape histogram được lưu trữ và xử lý trong lớp CSH. Cách tính histogram theo lưới không sử dụng để nhận dạng vì khối lượng tính toán lớn và khi lưu trữ trong cơ sở dữ liệu sẽ rất tốn không gian. Các hệ số biến đổi wavelet và trung bình màu được lưu và xử lý trong lớp CWavelet. Đặc trưng ảnh được lưu trữ trong tệp CSDL. Trong phần mềm, chúng tôi sử dụng thử nghiệm cơ sở dữ liệu đặc trưng ảnh trên SQL Server. ảnh đầu vào được tách lấy đặc trưng và đặc trưng này sẽ sử dụng để tìm kiếm. Đặc trưng về màu sắc như histogram tổng thể và colorshape được nhận dạng bằng cách khôi phục từng véc tơ đặc trưng của mỗi ảnh trong cơ sở dữ

liệu và so sánh với véc tơ đặc trưng của ảnh truy vấn. Nếu kết quả so sánh lớn hơn ngưỡng (chúng tôi thử nghiệm với giá trị mặc định 60%) thì sẽ giữ lại trong tệp kết quả.

Các ảnh trong cơ sở dữ liệu được so sánh với ảnh truy vấn và kết quả được sắp xếp theo thứ tự giảm dần. Những ảnh có kết quả lớn hơn ngưỡng hoặc  $n$  ảnh có kết quả cao nhất được cho vào tập tìm thấy. Việc truy vấn tùy thuộc vào lựa chọn cơ chế tìm kiếm của người sử dụng. Có thể chọn để thực hiện tìm kiếm phân cấp theo hai bước: tìm kiếm thô và tìm kiếm chính xác. Các ngưỡng cũng được để cho người sử dụng lựa chọn phù hợp với loại dữ liệu. Giá trị các ngưỡng trong khoảng từ 50-90%. Chương trình phần mềm được xây dựng trên VC6.0, các thử nghiệm được tiến hành trên máy tính Pentium III, tốc độ 700 Mhz, RAM 256 MB. Cơ sở dữ liệu ảnh (3000 ảnh) được sử dụng từ [www.ece.gatech.edu/users/gt0371a/link.html](http://www.ece.gatech.edu/users/gt0371a/link.html) và các ảnh tải về từ Internet từ các Website [vnexpress.net](http://vnexpress.net) và [www.vnn.vn](http://www.vnn.vn). Các thử nghiệm cho thấy:

+ Nếu ảnh có trong cơ sở dữ liệu thì chắc chắn tìm thấy bằng cách sử dụng bất kỳ đặc trưng nào.

+ Với một ảnh thu nhỏ cũng có thể sử dụng để tìm thấy ảnh thật trong cơ sở dữ liệu.



Hình 5: Giao diện của chương trình

+ Tốc độ tìm kiếm là chấp nhận được nếu so sánh với các phần mềm miễn phí mà chúng tôi có được [1,3,7] (với CSDL là 400 ảnh, trung bình thời gian tìm kiếm là 2.125 giây, 1000 ảnh – 5.235 giây, 2000 ảnh 11.234 giây).

#### IV. KẾT LUẬN

Xây dựng quá trình tìm kiếm ảnh theo nội dung theo hai bước: tìm kiếm thô và tìm kiếm chính xác làm tăng đáng kể tốc độ tìm kiếm ảnh. Vì bản chất thực hiện theo

thuật toán nhận dạng phân cấp hai bước nên chỉ có thời gian tìm kiếm thô là tuyến tính theo kích thước cơ sở dữ liệu. Còn nhận dạng chính xác phụ thuộc vào kết quả tìm kiếm thô. Các cài đặt thử nghiệm cho thấy với mô hình tìm kiếm này chúng ta có thể áp dụng các phát triển trên cơ sở xây dựng các giải thuật khác như sử dụng các đặc trưng Entropy, Color Correlogram. Hiện tại hướng nghiên cứu tiếp theo của chúng tôi tập trung vào sắp xếp ảnh theo chỉ số và các giải thuật sử dụng đặc trưng về kết cấu ảnh (texture) có nhiều tính chất tốt cho việc phân lớp ảnh và so sánh ảnh. Ngoài ra ứng dụng của mô hình phân cấp hai bước này với các giải thuật thích hợp trong nhận dạng theo nội dung của ảnh y học DICOM [8], các ảnh chụp thiên tai, chụp vệ tinh cũng đang được nghiên cứu thử nghiệm.

## TÀI LIỆU THAM KHẢO

1. Paul W. Fieguth, Riyin Wan, Fast retrieval methods for images with significant variations, *International Conference on Image Processing (ICIP 2000)*, Vancouver, BC, Canada, September, 2000, pp.10-13.
2. Jiri Walder, Using 2-D wavelet analysis for matching two images, *Technical Report, University of Ostrava Czech Republic*, N.3, 2001, pp.15-21.
3. Charles E. Jacobs Adam Finkelstein David H. Salesin, Fast multiresolution image querying, Department of Computer Science and Engineering University of Washington, <http://grail.cs.washington.edu/projects/query/mrquery.pdf>.
4. John R. Smith, Shih-Fu Chang, Quad-tree segmentation for texture-based image query, *Proceedings of the second ACM international conference on Multimedia*, San Francisco, California, United States, 1994, pp. 279 – 286.
5. Y. Rubner and C. Tomasi, *Perceptual Metrics for Image Database Navigation*. Kluwer Academic Publishers, Boston, December 2000, 137p.
6. V. Gudivada and V. Raghavan, Content-based image retrieval systems, *IEEE Computer*, vol. 28, 1995, pp. 18-22.
7. Huỳnh Quyết Thắng, Lê Tấn Hùng, Ứng dụng thư viện BKGRAPHICS trong xây dựng phần mềm xử lý ảnh y học V-doctor theo chuẩn DICOM”, Báo cáo tại FAIR 2003 “*Hội thảo Khoa học Quốc gia lần thứ nhất – Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin*”, Hà Nội, 4-5 / 10 / 2003, 7tr.

VNU. JOURNAL OF SCIENCE, Nat., Sci., & Tech., T.XX, N<sub>0</sub>2, 2004

## DEVELOPING ALGORITHM AND SOFTWARE ARCHITECTURE FOR MULTI FEATURE-BASED IMAGE SEARCHING

**Huynh Quyet Thang**

*Faculty of Information Technology*

*Hanoi University of Technology*

Content-based image retrieval, which provides convenient ways to retrieve images from large image databases, has been studied actively. We presented a software architecture which performs image presentation by the main features: color, histogram and color shape. An algorithm to measure similarity of two image features based on the wavelet transform of the Histogram and Color shape features and implementation results also presented.

**Keyword:** content-based image retrieval, wavelet transformation, image feature, image shape, histogram