

## THIẾT KẾ VÀ CÀI ĐẶT HỆ QUẢN TRỊ VÀ XỬ LÝ DỮ LIỆU-ẢNH: DIMS

ĐOÀN VĂN BAN, GIANG VŨ THẮNG,  
NGÔ QUỐC TẠO

*Viện Tin học, Viện KHAVN*

### GIỚI THIỆU

Dựa vào mô hình dữ liệu quan hệ của codd [2], nhiều hệ quản trị cơ sở dữ liệu đã được xây dựng như Dbase, Foxbaset, Foxprolan, Oracle v.v..., các hệ này đang được sử dụng rộng rãi và tỏ ra khá hữu hiệu trong việc ứng dụng máy tính vào công tác quản lý và xử lý dữ liệu.

Trong quá trình phát triển của xã hội loài người, nhu cầu xã hội ngày một cao, bên cạnh những thông tin dạng dữ liệu, thực tế ngày càng đòi hỏi phải xử lý cả những dạng thông tin phức tạp gồm cả hình ảnh, âm thanh, tiếng nói, v.v... Do đó xuất hiện những nhu cầu cần có những hệ thống tương đối thông minh, xử lý được những thông tin đa dạng, nhận biết và có khả năng suy luận để tìm ra những giải pháp hợp lý nhất cho bài toán đặt ra trên cơ sở khai thác triệt để những thông tin thu nhận được.

Hệ DIMS được thiết kế và cài đặt nhằm phục vụ cho quản lý, lưu trữ và xử lý các thông tin cả văn bản (text) lẫn hình ảnh (image).

DIMS được phát triển dựa trên nguyên tắc kế thừa những ưu điểm, những chức năng tiện dụng của các hệ quản trị CSDL phổ dụng, bổ sung thêm một số chức năng mới để xử lý các thuộc tính ảnh về những đối tượng quản lý.

### I - MỘT SỐ KHÁI NIỆM CƠ SỞ

Thông tin về các đối tượng cần quản lý trong các cơ sở dữ liệu (CSDL) được tổ chức thành tập cơ sở dữ liệu. Tập CSDL là tập thông tin có cùng bản chất được tổ chức theo một nguyên tắc nhất định nhằm phản ánh về một lớp các đối tượng nào đó. Các thông tin được lưu thành các bản ghi (record); mỗi bản ghi chứa một bộ giá trị mô tả thuộc tính của một đối tượng cụ thể.

Trong tập cơ sở dữ liệu, mỗi thuộc tính phản ánh một khía cạnh của một đối tượng gọi là trường, các trường phân biệt nhau bởi tên trường, kiểu và kích thước (chi tiết xem [2], [5]).

DIMS tuân theo các qui định chung của hệ quản trị CSDL phổ dụng về tên, kiểu và kích thước của dữ liệu. Ngoài kiểu trường thông dụng như ký tự (Character), số (Numeric), ngày (Date), Logic (Logic), và ký ức (Memo), DIMS còn quản lý thêm về thuộc tính hình ảnh của đối tượng.

Ảnh được đọc vào qua scanner hoặc được tạo ra bởi các hệ EDITOR, đồ họa phổ dụng khác như MS-Paint, Dr. Halo, v.v... được lưu trữ và nén ở dạng .PCX chuẩn. DIMS quản lý tất cả các thuộc tính đã khai báo và thông tin ảnh số [4].

Ảnh số là ma trận  $A = \{a_{ij} (1 \leq i \leq M; 1 \leq j \leq N)\}$ ,  $N$  là chiều rộng của ảnh,  $M$  là chiều cao của ảnh.  $a_{ij}$  (hoặc  $A(i, j)$ ) là một phần tử ảnh (Pixel) tại tọa độ  $(i, j)$  và

$$a_{ij} \in L = \{0, 1, \dots, l\}$$

Người ta có thể chia ảnh thành hai loại, ảnh nhị phân ( $l = 2$ ) và ảnh đa mức  $l > 2$ . Trong một số trường hợp, người ta có thể dùng phương pháp phân đoạn ảnh thích hợp để đưa ảnh đa mức về ảnh nhị phân.

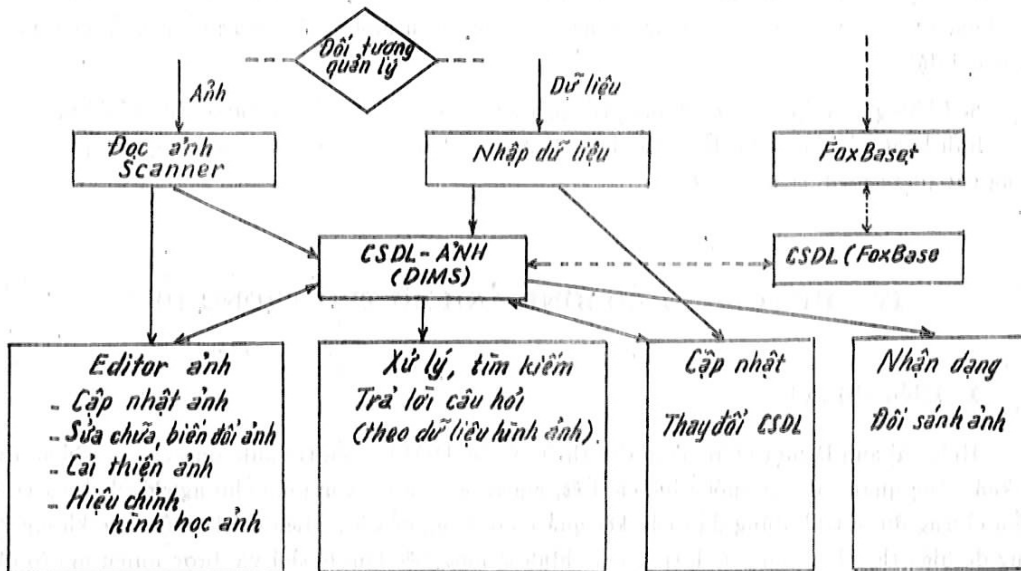
Ảnh nhị phân gồm hai thành phần là đối tượng ảnh và nền ảnh. Một đối tượng  $\theta$  trên ảnh  $A$  được xác định như sau:

$$\theta = \{a_{ij} | a_{ij} \in A \ \& \ a_{ij} = 1\} \text{ và nền ảnh } B = A \setminus \theta$$

Ở đây chúng tôi sử dụng phương pháp so sánh ngưỡng theo độ sáng và gradient để phân loại ảnh. Phương pháp này được mô tả tóm tắt như sau: Cho trước ngưỡng độ sáng  $T$  và ngưỡng gradient  $E$ . Từ ảnh đa mức có thể tạo ra ảnh nhị phân  $B$ . Các phần tử ảnh này với mức ánh sáng nhỏ hơn  $T$  coi như là nền. Ngoài ra cần tách các phần tử tăng gradient đột ngột và lớn hơn  $E$ . Giá trị gradient được đánh giá theo toán tử Robert.

$$B(i, j) = \begin{cases} 1, & A(i, j) \geq T \ \& \ |A(i, j) - G(i + 1, j + 1)| + |A(i + 1, j) - A(i, j + 1)| \leq E \\ 0 & \text{trọng trường hợp ngược lại} \end{cases}$$

## II - SƠ ĐỒ THIẾT KẾ



Trong đó  $\leftarrow - \rightarrow$  thể hiện mối quan hệ tương thích của DIMS với hệ quản trị CSDL phổ dụng mà FoxBase+ tạo ra, đơn giản là bổ sung thêm thuộc tính ảnh để thành CSDL-ảnh. Ngược lại từ CSDL ảnh do DIMS quản lý, khi lược bỏ thuộc tính ảnh thì cho lại hệ CSDL thông thường mà các hệ quản trị CSDL phổ dụng khác vẫn quản lý.

### III - CÁC CHỨC NĂNG CHÍNH

Mục đích khi thiết kế DIMS là xây dựng hệ mềm phục vụ cho việc lưu trữ quản lý các thông tin dữ liệu lẫn thông tin hình ảnh. Hệ được cài đặt phải tiện lợi cho người sử dụng và tương thích được với những hệ phổ dụng hiện nay như Foxbase+ [5]. DIMS đã được cài đặt bằng ngôn ngữ lập trình C chạy trên các máy PC XT/AT và tương thích, bộ nhớ  $\geq 360$  KB, ổ đĩa cứng  $\geq 10$  MB.

DIMS thực hiện các chức năng chính sau:

- 1- Thực hiện các chức năng chính của hệ quản trị CSDL như: Nhập, xử lý, sửa đổi và lưu trữ dữ liệu ở dạng .DBF.
- 2- DIMS đọc, quản lý được những CSDL được tạo lập bởi những hệ quản trị phổ dụng (FoxBase+) và còn có các khối về xử lý, cập nhật và lưu trữ ảnh số.
- 3- Ảnh của đối tượng quản lý được đưa vào qua Scanner hoặc được tạo ra từ những hệ EDITOR đồ họa như MS-Paint, Dr. Halo và được lưu trữ ở dạng .PCX. Thông qua khối EDITOR, ảnh số của đối tượng quản lý có thể được cắt xóa, hoặc thêm bớt, làm rõ hơn những thông tin về ảnh phục vụ cho việc lưu trữ, tìm kiếm và nhận dạng.
- 4- DIMS cho phép thực hiện đối sánh ảnh của đối tượng mẫu (đối tượng dạng quan tâm) để tìm trong CSDL đối tượng có thuộc tính ảnh gần giống nhất với ảnh mẫu. Trong DIMS, một vài thuật toán nhận dạng đối tượng thông minh dựa trên một số đặc trưng bất biến của ảnh đã được cài đặt.
- 5- Phần giao tiếp với người dùng rất tiện dụng thông qua việc sử dụng các chế độ menu đơn giản, linh hoạt cả bàn phím lẫn con chuột theo các chuẩn mực đã được nghiên cứu và sử dụng trong các phần mềm tiên tiến [8].

### IV - HỆ SOẠN THẢO HÌNH ẢNH BỔ TRỢ TRONG DIMS

#### 1. Hiển thị ảnh

Hiển thị ảnh là một chức năng cần thiết của hệ DIMS. Hiển thị ảnh giúp cho người sử dụng có khả năng quan sát ảnh một cách chi tiết, ngoài ra hiển thị còn giúp cho người sử dụng có thể kiểm chứng được tính đúng đắn của kết quả nhận dạng của hệ. Hiển thị là một việc không khó, song để hiển thị nhanh một ảnh trên màn hình là một việc làm lý thú và được nhiều người quan tâm.

Giá sử ảnh có kích thước  $M \times N$ , bao gồm  $N$  dòng và  $M$  cột. Các dòng ảnh lần lượt được chuyển vào bộ nhớ VIDEO giúp cho việc hiển thị ảnh được nhanh và tạo cho người quan sát có cảm giác ảnh xuất hiện một cách tức thời và mịn.

## 2. Hiệu chỉnh hình học ảnh

Hiệu chỉnh hình học là một kỹ thuật rất phức tạp trong khâu tiền xử lý ảnh. Song ở đây chúng ta chỉ xét hiệu chỉnh hình học đơn giản, đó là hiệu chỉnh mật độ điểm sáng trên màn hình hiển thị.

Ở các chế độ màn hình khác nhau, khoảng cách giữa hai pixels kế cận treo chiều ngang và chiều thẳng đứng là khác nhau. Sự khác nhau này sẽ gây cảm giác sai lệch khi quan sát ảnh đối tượng trên màn hình.

Gọi  $w$  là chiều rộng của màn hình,  $h$  ký hiệu chiều cao của màn hình. Giá sử màn hình được khởi tạo ở một chế độ graphic nào đó. Ở đây  $x_{max}$  là số pixel cực đại theo chiều rộng và  $y_{max}$  là số pixel cực đại theo chiều cao. Màn hình được gọi là đang ở chế độ đồ họa (graphic) hiển thị chuẩn, nếu

$$\frac{w}{x_{max}} = \frac{h}{y_{max}}$$

Chúng ta ký hiệu

$$Q = \frac{w * y_{max}}{h * x_{max}}$$

Trong thực tế,  $Q$  được gán giá trị nguyên của biểu thức bên phải.

Nếu  $M \times N$  là kích thước của ảnh ban đầu ( $N$  dòng,  $M$  cột). Gọi  $M' \times N'$  là kích thước của ảnh hiển thị trên màn hình thì phép hiệu chỉnh hình học đơn giản có thể thực hiện như sau:

$$M' = M, N' = N * Q \quad \text{nếu } Q > 1,$$

$$N' = N, m' = M/Q \quad \text{nếu } Q < 1,$$

$$m' = m \& N' = N \quad \text{nếu } Q = 1.$$

## 3. Quay ảnh

Trong hệ DIMS, đa số các ảnh cần lưu trữ và xử lý là được đọc qua Scanner. Tuy nhiên, đôi lúc do sơ xuất, chúng ta có thể quét ảnh ở vị trí không chuẩn (ví dụ lệch một góc  $\alpha$  so với trục thẳng đứng). Việc thu nạp ảnh lệch sẽ gây trở ngại trong việc quan sát, lưu trữ và xử lý sau này. Do vậy quay ảnh là một thủ tục không thể thiếu trong hệ.

Giá sử  $a_{ij}$  là giá trị của pixel ảnh tại tọa độ  $(i, j)$ , phép quay ảnh đi một góc  $\alpha$  là một ánh xạ  $R: (i, j) \mapsto (i', j')$ , sao cho

$$i' = i * \cos \alpha - j * \sin \alpha$$

$$j' = i * \sin \alpha + j * \cos \alpha$$

Trong cài đặt,  $i'$  và  $j'$  sẽ lấy giá trị nguyên của biểu thức bên phải.

#### 4. Vấn đề cắt và xóa một phần của ảnh

Trong quá trình xử lý và nhận dạng, chúng ta chỉ tập chung xử lý những phần ảnh quan trọng, thể hiện đúng đối tượng mà chúng ta đang quan tâm. Ảnh đọc vào qua Scanner thường còn có những phần ảnh dư thừa cần phải được loại bỏ trước khi lưu trữ vào cơ sở ảnh. Cắt ảnh là một thủ tục nhằm tách ra từ ảnh toàn cảnh một ảnh con đủ để mô tả đối tượng mà chúng ta cần lưu trữ và xử lý trong hệ.

Khác với chức năng cắt ảnh, chức năng xóa một phần ảnh là xóa đi trên ảnh các vùng quá nhiều, bản do quá trình quét ảnh, với việc giúp đỡ của người sử dụng. Chức năng đơn giản này góp phần tăng cường chất lượng của ảnh và độ tin cậy của thuật toán nhận dạng được sử dụng trong hệ DIMS.

#### 5. Nén và giải mã ảnh

Thông thường ảnh có kích thước khá lớn và chiếm nhiều không gian nhớ để lưu trữ. Do vậy, vấn đề quan trọng trong khi xây dựng hệ cơ sở dữ liệu ảnh là vấn đề nén ảnh, nhằm giảm tối đa khoảng không gian lưu trữ ảnh, nhất là đối với cơ sở dữ liệu ảnh lớn. Như chúng ta đã biết, hiện nay có nhiều thuật toán nén và giải mã ảnh ở dạng nén. Hai chỉ tiêu quan trọng quyết định chất lượng của thuật toán nén và giải mã ảnh là lượng không gian tiết kiệm được khi lưu trữ ảnh nén và thời gian chi phí để giải mã mà nén thành ảnh gốc bit-map. Thiếu một trong hai chỉ tiêu trên thì thuật toán không thể coi là có hiệu quả. Đã tồn tại những thuật toán nén, đảm bảo giảm không gian lưu trữ ảnh tới hàng vài chục lần, thậm chí hàng trăm lần, nhưng thời gian giải mã thì quá lớn. Sự chậm chễ này sẽ kéo dài thời gian xử lý ảnh và làm ngưng trệ cả thời gian truy nhập tới cơ sở dữ liệu ảnh.

Trong hệ DIMS, chúng tôi đã chọn mã nén dạng .PCX. Đó là một kiểu mã đơn giản, không gian chứa ảnh nén giảm nhiều so với ảnh gốc và thời gian giải mã nhanh. Trên tệp nén, việc giải mã được tiến hành từ đầu đến cuối. Tại byte thứ  $i$  của ảnh nén, nếu hai bit cao nhất của nó có giá trị 11, thì đó là byte không nén và được sao nguyên vẹn sang tệp ảnh. Trong trường hợp khác, tệp ảnh được nạp thêm liên tiếp  $n$  byte có giá trị bằng giá trị của byte thứ  $i + 1$  trong tệp nén. Ở đây  $n$  là giá trị được biểu diễn bằng 6 bit thấp của byte thứ  $i$  trong tệp ảnh nén.

### V - MỘT SỐ THUẬT TOÁN NHẬN DẠNG HÌNH ẢNH HIỆU QUẢ TRONG DIMS

Nhận dạng ảnh là một ngành khoa học quan trọng và có nhiều ứng dụng thực tiễn trong các lĩnh vực khác nhau như: nhận dạng sản phẩm trong công nghiệp, nhận biết nhãn hàng, nhận dạng chân dung đối tượng, chữ viết, vân tay, v.v...

Hiện nay tồn tại nhiều thuật toán nhận dạng khác nhau. Những thuật toán này dựa trên các nguyên lý nhận dạng khác nhau như nguyên tắc hàm thế, nguyên lý thống kê, nguyên tắc cấu trúc, v.v... [1], [6], [7].

Trong phần này của bài, chúng tôi sẽ trình bày một số thuật toán hiệu quả đã được cài đặt để xử lý thông tin hình ảnh trong hệ DIMS. Các ảnh được đối sánh với nhau dựa trên những đặc trưng bất biến của chúng. Trong ảnh hai chiều chúng ta cần tìm những đặc trưng thỏa mãn các tính chất sau đây:

- Bất biến đối với phép quay.
- Bất biến đối với phép tịnh tiến.
- Bất biến đối với phép biến đổi tỉ lệ.

Trong số các đặc trưng bất biến của ảnh, những bất biến hình học sau đây có vai trò quan trọng đặc biệt [1].

- x1 - Diện tích của hình.
- x2 - Mômen quán tính cực tiểu.
- x3 - Mômen quán tính cực đại.
- x4 - Độ dài chu tuyến.
- x5 - Khoảng cách nhỏ nhất từ tâm tới chu tuyến.
- x6 - Khoảng cách lớn nhất từ tâm tới chu tuyến.
- x7 - Khoảng cách trung bình từ tâm tới chu tuyến.
- x8 - Số lỗ hổng.
- x9 - Tổng số diện tích các lỗ hổng.
- x10 - Tổng độ dài các véc tơ từ tâm vật thể đến tâm lỗ hổng.

Ngoài ra có thể xác định thêm các bất biến bằng cách chia vật thể ra làm 4 phần dựa vào các trục momen quán tính cực đại và cực tiểu. Sau đó tính khoảng cách từ tâm mỗi phần tới tâm vật thể. Các dấu hiệu trên là bất biến đối với phép quay, tịnh tiến. Để thu được các bất biến theo phép biến đổi tỉ lệ chúng ta cần tính các tham số là tổ hợp của các bất biến trên, ví dụ: tỉ số giữa độ dài chu tuyến và căn bậc hai diện tích. Trong hệ DIMS chúng tôi sử dụng các đặc trưng thống kê để xác định mômen bất biến để phục vụ cho việc đối sánh nhanh hình ảnh.

Gọi  $f(i, j)$  là mật độ sáng tại điểm  $(i, j)$ ,  $1 \leq i \leq M$ ;  $1 \leq j \leq N$ ,  $M$  và  $N$  là kích thước của ảnh. Mômen bậc  $(p + q)$  của ảnh  $e$ , ký hiệu là  $m_{pq}$ , được xác định như sau [4], [9]:

$$m_{pq} = \sum_{i=1}^N \sum_{j=1}^M i^p j^q f(i, j)$$

Momen tám bậc  $(p + q)$ , ký hiệu  $\mu_{pq}$  được tính như sau:

$$\mu_{pq} = \sum_{i=1}^N \sum_{j=1}^M (i - \bar{i})^p (j - \bar{j})^q f(i, j)$$

Trong đó  $\bar{i} = m_{10}/m_{00}$ ,  $\bar{j} = m_{01}/m_{00}$ , còn mômen tâm chuẩn hóa, ký hiệu là

$$\nu_{pq} = \mu_{pq} / (\mu_{00}^{(p+q)/2} + 1)$$

Chúng ta dựa vào 7 mômen bất biến sau đây để đối sánh:

$$\varphi_1 = \eta_{20} + \eta_{02}$$

$$\varphi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\varphi_4 = (\eta_{30} + 3\eta_{12})^2 (\eta_{21} + \eta_{03})^2$$

$$\varphi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]$$

$$+ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$\varphi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$\varphi_7 = (3\eta_{12} - \eta_{30})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]$$

$$+ (3\eta_{21} - \eta_{03})(\eta_{03} + \eta_{21})[(\eta_{03} + \eta_{21})^2 - 3(\eta_{12} + \eta_{30})^2]$$

Mỗi khi đưa một đối tượng vào cơ sở dữ liệu ảnh, hệ thống tiến hành xem xét và tính toán một số đặc trưng bất biến cơ bản và lưu vào hệ thống. Khi thực hiện chức năng nhận dạng, phương pháp người láng giềng gần nhất được áp dụng.

Đối với mỗi đối tượng, ta xây dựng khoảng cách tới các mẫu. Tập mẫu được biến đổi trong không gian  $l$  chiều. Một đối tượng được nhận dạng tìm đối sánh của nó trong CSDL ảnh, khi chỉ tiêu sau đây đạt cực trị nhỏ hơn giá trị đã qui định

$$D_j = \sum_{i=1}^l \omega_i \frac{|x_i - x_{ji}|}{\bar{x}_{ji}}$$

Ở đây  $x_i$ ,  $1 \leq i \leq l$  là tham số của đối tượng được nhận dạng,

$\bar{x}_{ji}$  với  $1 \leq i \leq l$  là dấu hiệu thứ  $i$  của mẫu thứ  $j$ ,

$\omega_i$ ,  $1 \leq i \leq l$  là trọng số của dấu hiệu thứ  $i$ .

## VII - KẾT LUẬN

Hệ DIMS Ver. 1.1 đã được cài đặt với CSDL demo là hệ quản lý nhân sự. Ảnh của mỗi đối tượng quản lý được nén dưới dạng .PCX chiếm từ 1 đến 3 KB. Thông tin về từng đối tượng gồm cả thuộc tính văn bản, số lần hình ảnh được lưu trữ và xử lý đồng thời. Chúng ta dễ dàng tìm kiếm, khai thác CSDL theo thuộc tính thông thường như các mục trong lý lịch, hồ sơ hoặc từ thuộc tính ảnh; khi đối sánh thành công thì tất cả các thông tin đó sẽ được hiện lên màn hình, cung cấp cho chúng ta đầy đủ và sinh động những thông tin cần thiết về đối tượng mà chúng ta quan tâm.

Hệ DIMS sẽ rất thích hợp cho những ứng dụng mà trong đó cần quản lý những thông tin hình ảnh như: xây dựng hệ CSDL quản lý mẫu, nhận dạng kỹ các mặt hàng sản xuất, các chi tiết thiết bị máy móc, các mặt hàng quảng cáo, về bảo tồn, bảo tàng các di tích lịch sử, du lịch, v.v...

## TÀI LIỆU THAM KHẢO

1. Agin G. J., An Experimental vision System for Industrial Application Proc. 5th International simposium on Industrial Robot, Chicago USA (Sept. 1975), p. 135-148.
2. Codd E. F., A Relational Model for Large Shared Banks, CACM, 6(1970), 1(13), p. 377-381.
3. George R. M., Computer graphic in Application, Prentice-Hall, Inc., 1987.
4. Hall E. L., Computer Image Processing and Recognition, New York - Academic Press, 1979.
5. Hệ quản trị cơ sở dữ liệu Foxbase+, Tài liệu VIE/88/035, Viện Tin học, Hà Nội, 1990.
6. Pavlidis T., Algorithms for Graphic and Image Processing, Springer - Verlag Berlin Heideberg, New York, 1978.
7. Pratt W. K., Digital Image Processing John Willey and Sons, 1978.
8. Shneiderman B., Designing the User Interface strategies for Effective Human-Computer Interaction, ADDISON-WESLEY Pub. Co., 1987.
9. Zakaria M. F. et al., Fast Algorithm for the Computation of Moment Invariants, Pattern Recognition, Vol. 20, No. 6, pp. 639 - 643, 1987. Research Centre for Intelligent Machines, Mc Gill University, Montreal, Qubec, Canada.

## ABSTRACT

## ON THE DESIGN AND IMPLEMENTATION OF DATA-IMAGE MANAGEMENT SYSTEM

Based on relational model of data base we give design and implementation of data-image base management system, DIMS. This system has been devised to allow process, manage data and image base, is compatible with other management systems such as Dbase, FoxBase+, etc.