TẠP CHÍ KHOA HỌC TÍNH TOÁN VÀ ĐIỀU KHIỂN TẬP VI (1990), Số 4 (1-5)

THUẬT TOÁN NHẬN DẠNG THÍCH NGHI CÁC ĐỐI TƯỢNG TRONG HỆ THỊ GIÁC MÁY

NGÔ QUỐC TẠO, HOÀNG KIẾM

Viện Tin học, Viện KHVN

Nhiệm vụ của các hệ thị giác máy công nghiệp đối với các quá trình công nghệ là nhận được các thông tin tối ưu về ảnh theo thời gian thực. Vấn đề đặt ra là nâng cao tốc độ tính toán của hệ dựa trên một số phương pháp heuristique.

Cách tiếp cận là tính sơ bộ ma trận dấu hiệu mẫu C_{ik}, ở đây $i \leq n$ - số lớp, $k \leq m$ - số dấu hiệu. Các dấu hiệu C_{ik} dao động trong miền nào đó ta tìm được trong giai đoạn học, H là véc tơ cần nhận dạng. Nhiệm vụ đặt ra cần phải giải quyết hai bài toán sau đây:

1. Tìm số $l < m$ đảm bảo tách các lớp đã cho.

2. Xây dựng tự động đồ thị giải dựa trên ma trận dấu hiệu ban đầu với $l < m$. Thuật toán đặc biệt tiện lợi trong trường hợp các dấu hiệu nhận dạng được tính theo công thức truy hồi: $C_{ik} = f(C_{ik-1})$, chẳng hạn các mô men bất biến [1] và các hệ số khai triển hàm chu tuyến thành chuỗi furie [2]. Các dấu hiệu heuristique như hệ số compac (độ dài chu tuyến/căn bậc hai của diện tích), tỉ số các bán kính cực tiểu và cực đại cũng như các dấu hiệu hình học [3].

Bài này đưa ra tiêu chuẩn đánh giá chất lượng của các dấu hiệu được chọn và xác định xác suất tiên nghiệm của thuật toán nhận dạng dựa trên bộ mẫu có số dấu hiệu $< l$.

1. THUẬT TOÁN XÁC ĐỊNH SỐ DẤU HIỆU TỐI THIỂU

Bộ dấu hiệu ban đầu của tập mẫu học là ma trận $C_{m \times 2n}$. Ở đây m là số dấu hiệu còn n là số lớp. Vì ảnh của các đối tượng rời rạc nên ở giai đoạn học lớp thứ i bằng biểu diễn các đối tượng của lớp này với các hướng, vị trí và tỉ lệ khác nhau, chúng ta nhận được dấu hiệu thứ k nằm trong khoảng nào đó $C_{ik}^{min} \leq C_{ik} \leq C_{ik}^{max}$, với $i \leq n, k \leq m$. Như vậy mỗi lớp được biểu diễn bởi hai hàng của ma trận xác định cận biến đổi của các dấu hiệu. Sau khi hình thành ma trận C, để giảm thời gian nhận dạng cần tìm số l nhỏ nhất các dấu hiệu đủ để nhận dạng đối tượng bất kỳ trong bộ dấu hiệu có chứa n lớp.

Nếu các dấu hiệu C được tính theo hệ số truy hồi thì ta chỉ cần tính l dấu hiệu đầu tiên của C.

Để giải bài toán này chúng ta tính ma trận truy hồi sau:

$$B_{j+1} \equiv (B_{ik}^{j}) \equiv B_j \wedge M_{j+1} \tag{1}$$

$$R_{j+1} \equiv (r_{ik}^{j+1}) \equiv R_j + B_{j+1} \tag{2}$$

Ở đây $j = 0, ..., m-1$; B và R là ma trận kích thước $n \times n$ gồm các phần tử 0 hoặc 1; $b^0_{ik} = r^0_{ik} = 1$ đối với $i, k = 1, ..., n$; $\wedge$ là phép nhân logich.

Ma trận $M_j \equiv (m_{ik})$ xác định như sau:

$$m^j_{ik} = \begin{cases} 0 & \text{nếu} [C^{min}_{ij}, C^{max}_{ij}] \cap [C^{min}_{ik}, C^{max}_{ik}] \equiv \emptyset \\ 1 & \text{nếu ngược lại} \end{cases} \tag{3}$$

Biểu thức (3) xác định các dấu hiệu giao nhau của lớp thứ i và lớp thứ k theo dấu hiệu thứ j. B_j xác định khả năng nhận dạng của đối tượng trong mẫu n lớp. Dù chỉ có một phần tử $b^j_{ik} = 1$; $i = k$ chỉ ra rằng các lớp i và k rời nhau theo dấu hiệu j. Hiển nhiên rằng điều kiện cần và đủ để nhận dạng là $B_j = 1$ nếu $B_l = 1$ với l là chỉ số nhỏ nhất $< m$ đủ để nhận dạng. Nếu $B_m \neq 1$ thì dấu hiệu mẫu chưa đủ tách mẫu các lớp đã cho, cần tăng m. Tại thời điểm thỏa mãn $B_j = 1$ tại ma trận R tích lũy các thông tin về dấu hiệu tách hai lớp. Các phần tử $r_{ik}, 1 \le r_{ik} \le l (i \neq k)$ là chỉ số dấu hiệu tách các lớp i và k. Ma trận R được dùng để giải các bài toán nhận dạng.

2. THUẬT TOÁN HÌNH THÀNH ĐỒ THỊ GIẢI BÀI TOÁN NHẬN DẠNG

Để xây dựng chiến thuật nhận dạng từ tập l dấu hiệu ta chọn ra dấu hiệu có nhiều thông tin nhất là f trong đó số các lớp giao nhau ít nhất.

Từ các ma trận M_j ta xác định

$$\min \sum_{i=k}^{n} \sum_{k=1}^{n} m^j_{ik}, \quad l \ge j \ge 1 \tag{4}$$

Theo dấu hiệu f ta sửa đổi R_1 thành R^*_1 như sau:

Đối với hai lớp bất kỳ được tách nhau theo dấu hiệu f tương ứng với phần tử r_{ik} ta sẽ gán cho $r^{*1}_{ij} = 0$; sau đó tách từ ma trận C_{ik} cột f chứa vùng biến đổi của dấu hiệu f đối với mỗi lớp chúng ta sắp xếp C_{ir}, $i = 1, 2, ..., 2n$ như sau:

$$C_{p_1 r} \le C_{p_2 r} \le C_{p_3 r} \le ... \le C_{p_{2n} r} \tag{5}$$

Khi đó ta nhận được (2n+1) khoảng. Phân tích các khoảng này cho phép xây dựng lời giải bài toán nhận dạng.

Ta xây dựng ma trận chiến thuật S $2 \times (2n+1)$. Hàng đầu tiên chứa các giá trị chỉ số (S) : $p_1, p_2, ..., p_{2n}$. Hàng thứ hai của ma trận S chứa các nhãn đặc trưng cho trạng thái mỗi khoảng theo qui tắc sau:

1) $S_{2i} = Q$ nếu trong khoảng thứ i không chứa dấu hiệu f.

2) $S_{2i} = k$ nếu khoảng thứ i chứa dấu hiệu f chỉ có lớp k.

3) $S_{2i} = Q + k$ nếu có nhiều lớp có dấu hiệu f rơi vào với k là số nhỏ nhất của danh sách các lớp giao nhau.

3. NHẬN DẠNG

Ta có C_{ik}, R, S, véc tơ cần nhận dạng $H(h_j)$ $1 \leq j \leq l$ chiến thuật nhận dạng theo phương pháp lưỡng phân theo dãy số tự nhiên từ 1 tới $(2n+1)$. Qua chỉ số S_{1i} tìm được khoảng thứ i chứa h_j. Hàng thứ hai của S chứa nhãn xác định trong khoảng xác định một trong ba tình huống nêu trên. Phân tích S_{2i} quan hệ với Q.

Nếu tình huống 1 huặc 2 xảy ra thì chúng ta nhận được lời giải. Trong tình huống thứ 3 cần phải phân tích thêm ma trận R. Các phần tử $r_{N_1N_2} = 0$ nếu N_1 và N_2 tách nhau theo dấu hiệu f. Ngược lại nó chứa các chỉ số dấu hiệu tách các lớpN_1, N_2.

Giai đoạn phân tích như sau:

1. Đặt $N_1 = S_{2i} - Q$;

2. Nếu $R_{N_1N_2} = 0$ thì $N_2 = N_2 + 1$; lặp lại quá trình này.

3. $r_{N_1N_2} \neq 0$ kiểm tra xem lớp nào không phải là nghiệm

$$T_{1,2} = \frac{1}{2}(C_{vu} \mp C_{v-1,u}),\ T_1 - |T_2 - h_u| < 0 \tag{6}$$

Ở đây chỉ số $u = r_{N_1N_2}$ còn $v = 2N_1$ hoặc $v = 2N_2$ phụ thuộc vào chỉ số lớp cần phân tích. Nếu lớp nào thỏa mãn (6) thì bị loại khỏi nghiệm. Nếu $r_{N_1N_2} \neq 0$ vàN_2 thỏa mãn (6) thì chuyển tới giai đoạn 2. Nếu (6) không thỏa mãn với N_2 thì $N_2 = N_2 + 1$ chuyển tới giai đoạn 2. Thực hiện chu trình cho tới khi $N_2 > n$, khi đó N_1 sẽ là nghiệm, Tức là H được xếp vào lớp N_1.

Độ phức tạp của thuật toán trong giai đoạn nhận dạng là $log_2(2n+1) + n(1 - 1/m)$ phép toán.

4. TIẾP CẬN THỐNG KÊ

Hiệu quả của thuật toán thích nghi ở chỗ xác định được số l dấu hiệu đảm bảo tách các đối tượng thành các lớp. Ta coi rằng thuật toán thích nghi có hiệu quả trong trường hợp xác suất nhận dạng các đối tượng từ mẫu với số dấu hiệu nhận dạng là l bằng 1. Như vậy chúng ta đã xác định với các dấu hiệu $j \leq l$ thì nhận dạng với xác suất tin cậy nào. Điều kiện cần và đủ để thuật toán hiệu quả là $B_l = 1$ với xác suất 1.

Bài toán đặt ra được giải quyết sau khi xác định số tối thiểu các dấu hiệu nhận dạng, thông tin được sử dụng chứa trong ma trận C.

Giả sử $A^j_{ik} \equiv (m^j_{ik} = 1)$ là biến cố ngẫu nhiên biểu thị sự giao nhau giữa hai lớp i và k theo dấu hiệu j (3), còn $\overline{A}^j_{ik}$ là biến cố bù trừ của A^j_{ik} (tức là $m^j_{ik} = 0$).

Như vậy chúng ta có biểu thức

$$P(m^j_{ik} = 1) = \varepsilon^j_{ik}$$

xác suất giao nhau của các lớp i và k theo dấu hiệu thứ j

$$P(m^j_{ik} = 0) = 1 - \varepsilon^j_{ik}$$

Đối với hai lớp bất kỳ i và k, sau khi sử dụng ma trận thông tin C_{ik} có thể xác định C^j_{ik} dưới dạng

$$C^j_{ik} = \frac{\max\{\min(C^{max}_{ij} - C^{min}_{ij}, C^{max}_{ik} - C^{min}_{ik}, C^{max}_{ik} - C^{min}_{ij}, C^{max}_{ij} - C^{min}_{ik}), 0\}}{\max(C^{max}_{ij} - C^{min}_{ij}, C^{max}_{ik} - C^{min}_{ik}, C^{max}_{ik} - C^{min}_{ij}, C^{max}_{ij} - C^{min}_{ik})}$$

Xác suất để hai lớp i và k tách nhau với j dấu hiệu là

$$P(B^j_{ik} = 0) = 1 - P(B^j_{ik} = 1) = 1 - \prod P(m^a_{ij} = 1) = 1 - \prod_{a<j} c^a_{ik}$$

Bây giờ chúng ta xác định xác suất để n đối tượng từng cặp tách nhau là

$$P(B_1 = I) = \prod_{i<k\leq n} P(B^j_{ik} = 0) - \prod_{i<k\leq n} (1 - \prod_{j\leq l} c^j_{ik}) \tag{8}$$

Ở giai đoạn học thực hiện các thủ tục sau:

1. Theo hệ thức (3) đối với các cột của ma trận C các ma trận được xét M biểu thị các sự kiện giao của vùng biến đổi các dấu hiệu và ánh xạ đồ thị

$$M = \begin{vmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{vmatrix} \qquad M_2 = \begin{vmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{vmatrix}$$

$$M_3 = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} \qquad M_4 = \begin{vmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{vmatrix}$$

2. Theo (1) tính các ma trận B: Ma trận B biến thành I khi $j = 4$.

3. Sử dụng (4), trong số l dấu hiệu tìm được dấu hiệu chứa thông tin nhất là $f = 3$.

4. Theo hệ thức (2) tính R_l và R^*_l với ma trận thông tin $M_{f=3}$:

$$R_l = \begin{vmatrix} 5 & 3 & 1 & 1 & 1 \\ 3 & 5 & 1 & 1 & 1 \\ 1 & 1 & 5 & 4 & 2 \\ 1 & 1 & 4 & 5 & 3 \\ 1 & 1 & 2 & 3 & 5 \end{vmatrix} \qquad R^*_l = \begin{vmatrix} 5 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 1 & 0 \\ 0 & 1 & 5 & 4 & 0 \\ 0 & 1 & 4 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{vmatrix}$$

5. Ứng với (5) theo thứ tự cột f các ma trận C và hình thành ma trận chiến thuật S, đặt $Q = 20$:

$$S = \begin{vmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 20 & 1 & 20 & 2 & 22 & 22 & 23 & 4 & 20 & 5 & 20 \end{vmatrix}$$

6. Chúng ta dựa vào đánh giá thống kê thuật toán đã trình bày đối với ví dụ này. Theo công thức (7) và (8) chúng ta nhận được xác suất tách các đối tượng P_j, Ở đây j là số dấu hiệu

$$P_1 = 0,298;\ P_2 = 0,842;\ P_3 = 0,991;\ P_4 = 1$$

Kết quả này chỉ ra rằng xác suất nhận được ma trận B_l bằng 1 với số dấu hiệu $l = 4$ trùng với kết quả của thuật toán đưa ra ở trên. Tính toán cũng chỉ ra rằng xác suất tách của thuật toán gần 1 khi $l = 3$ ($P_3 = 0,991$). Điều này cho phép chúng ta rút ngắn số dấu hiệu nhận dạng còn 3, nó sẽ nâng cao tốc độ đoán nhận. Ở giai đoạn học chúng ta có thể thay đổi như sau: nếu như $b^j_{ik} < \varepsilon$ nào đó ta sẽ đồng nhất $b^j_{ik} = 0$ sẽ nâng cao tốc độ nhận dạng.

TÀI LIỆU THAM KHẢO

1. Hu M. K. – Visual Pattern Recognition by moment Invariants IRE transaction on Information Theory, vol. II-8, 1962.

2. Hoifordt J. R., Hakonsen H. – SIMAGE - SI's image recognition system. Proc. 11th International Symposium on Industrial Robots. Oct. 1981, Tokio, Japan, pp. 137–144.

3. Bạch Hưng Khang, Hoàng Kiếm – AI - Các phương pháp và ứng dụng. Nhà xuất bản KHKT, Hà Nội, 1989.

4. Ngô Quốc Tạo, Hoàng Kiếm, Thuật toán và chương trình cho hệ nhận dạng thị giác máy TV1. Báo cáo kết quả nghiên cứu đề tài 48A-04-06 (1990).

Nhận ngày 20-7-1990

ABSTRACT

An adaptive recognition algorithm for vision systems

In this paper, we present adaptive recognition algorithm for visition systems.

Our methode is divided into following principal steps:

- Determine the minimum number of feature for object recognition.
- Graphical Matching.
- Object recognition.
- Statical evalution.

The experimental results are presented also.