

MỘT PHƯƠNG PHÁP PHÁT HIỆN ĐỐI TƯỢNG ỨNG DỤNG TRONG HỆ THỐNG TỰ ĐỘNG BẮM MỤC TIÊU

Nguyễn Văn Hùng¹, Nguyễn Văn Xuất², Lê Mạnh Cường³

¹ Viện Vũ khí, Tổng cục Công nghiệp Quốc phòng, Bộ Quốc phòng

² Khoa Công nghệ Thông tin, Học viện Kỹ thuật Quân sự, Bộ Quốc phòng

³ Trung tâm 80, Cục Tác chiến Điện tử, Bộ Quốc phòng

hungitd@yahoo.com, xuatnguyenvan@yahoo.com, lemanhcuong76@gmail.com

TÓM TẮT - Phát hiện đối tượng là một thành phần rất quan trọng trong các hệ thống tự động giám sát và bám mục tiêu. Bài báo này đề xuất một phương pháp dựa trên kỹ thuật xử lý ảnh và các thuật toán máy học để phát hiện các mục tiêu quân sự. Phương pháp đề xuất dựa trên các đặc trưng hình dạng và màu sắc để tìm đối tượng trong ảnh cảnh được chụp từ một camera. Trong phương pháp của chúng tôi, ảnh đầu vào trước tiên được phân thành các vùng nhỏ đồng màu sử dụng thuật toán dựa vào lý thuyết đồ thị. Sau đó, các vùng có khả năng thuộc về đối tượng thấp sẽ được loại bỏ bằng việc sử dụng các thông tin về màu sắc. Cuối cùng, đối tượng cần tìm được xác định từ một tập các vùng đồng màu kết nối với nhau, sử dụng một mô hình xác suất với các đặc trưng hình dạng và màu sắc. Đặc trưng hình dạng của một vùng ảnh được biểu diễn bằng các bộ đặc tả hình dạng (shape contexts). Các bộ đặc tả về hình dạng là bất biến với sự méo hình cục bộ, sự dịch chuyển và xoay... Phương pháp đề xuất được đánh giá trên các tập dữ liệu lớn khác nhau. Mỗi tập dữ liệu bao gồm các ảnh của một kiểu đối tượng được thu thập từ các cảnh khác nhau dưới các điều kiện ánh sáng khác nhau. Các kết quả thực nghiệm đã chỉ ra tính hiệu quả của phương pháp đề xuất.

Từ khóa - Phân vùng ảnh; dò đối tượng; đặc trưng màu, đặc trưng hình dạng, mô hình xác suất.

I. GIỚI THIỆU

Phát hiện đối tượng tự động có vai trò quan trọng trong các hệ thống trình sát phát hiện mục tiêu, các hệ thống giám sát, nhận dạng và khảo sát. Việc sử dụng thị giác máy tính để phát hiện các đối tượng là một giải pháp được sử dụng rộng rãi trên toàn thế giới. Phương pháp này ứng dụng các kỹ thuật xử lý ảnh và các thuật toán máy học để tìm các đối tượng trong các hình ảnh. Tuy nhiên, việc phát hiện đối tượng dựa trên thị giác máy tính là một vấn đề khó bởi vì hệ thống phải giải quyết với sự thay đổi về điều kiện tạo ảnh (ví dụ như sự thay đổi về các điều kiện ánh sáng và thời tiết) và sự thay đổi của các cảnh và môi trường.

Nhiều phương pháp phát hiện đối tượng khác nhau sử dụng thị giác máy tính đã được phát triển và ứng dụng rộng rãi trong đời sống thực tiễn. Các phương pháp này phát hiện đối tượng với ba bước chính. Bước thứ nhất là dựa vào các thuộc tính của đối tượng như màu sắc, kết cấu bề mặt và hình dạng để trích chọn các đặc trưng ảnh. Bước thứ hai là sử dụng tập dữ liệu mẫu để xác định các tham số cho các bộ nhận dạng đối tượng trong ảnh. Bước thứ 3 là sử dụng bộ nhận dạng để xác định đối tượng trong các ảnh đầu vào bất kỳ.

Trong hệ thống phát hiện mục tiêu quân sự thì việc trích chọn đặc trưng ảnh đóng một vai trò quan trọng. Có hai xu hướng chính trong việc trích chọn đặc trưng: 1) sử dụng các đặc trưng cạnh, 2) sử dụng các đặc trưng màu và kết cấu bề mặt của các vùng ảnh. Các phương pháp dựa trên các đặc trưng cạnh xác định đối tượng trong ảnh bằng việc đi tìm các đường bao của đối tượng [1-5]. Điểm mạnh của các phương pháp này là bất biến với các điều kiện ánh sáng và sự dịch chuyển của đối tượng. Tuy nhiên, các phương pháp dựa vào cạnh lại nhạy cảm với các cạnh nhiễu và không có hiệu quả khi trong ảnh xuất hiện nhiễu cạnh nhiễu. Mặt khác, hiệu quả của các phương pháp này lại phụ thuộc vào việc dò tìm các điểm cạnh.

Các phương pháp sử dụng các đặc trưng màu và kết cấu bề mặt của các vùng ảnh có hiệu quả cao trong việc dò tìm các đối tượng mà chúng có màu sắc hoặc kết cấu bề mặt rất khác so với các đối tượng nền khác trong ảnh [6, 7]. Điểm yếu của các phương pháp này là chúng rất nhạy cảm với sự thay đổi của các điều kiện ánh sáng và thường lỗi trong trường hợp đối tượng bị ảnh hưởng bởi các vùng rất sáng và các bóng râm.

Trong bài viết này chúng tôi đề xuất một phương pháp mới để phát hiện các mục tiêu quân sự. Phương pháp của chúng tôi sử dụng điểm mạnh của phương pháp dựa vào cạnh và phương pháp dựa vào màu và kết cấu bề mặt. Phương pháp đề xuất dựa trên sự kết hợp giữa đặc trưng màu sắc và đặc trưng hình dạng để xây dựng một mô hình xác suất cho việc phát hiện các mục tiêu trong ảnh. Có hai bước chính trong phương pháp của chúng tôi. Bước thứ nhất là phân mảnh ảnh đầu vào thành các vùng đồng màu khác nhau. Bước thứ hai là xác định mục tiêu trong ảnh từ các vùng đồng màu sử dụng các đặc trưng màu và hình dạng. Mục tiêu được tìm như là một tập con các vùng đồng màu kết nối với nhau (các vùng này đều thuộc về một vùng lớn) sao cho xác suất hậu nghiệm của tập này là lớn nhất.

Cấu trúc của bài báo với các phần còn lại như sau: Phần II trình bày phương pháp đề xuất; Phần III trình bày về thực nghiệm và các kết quả đạt được; Phần IV đưa ra các kết luận chính.

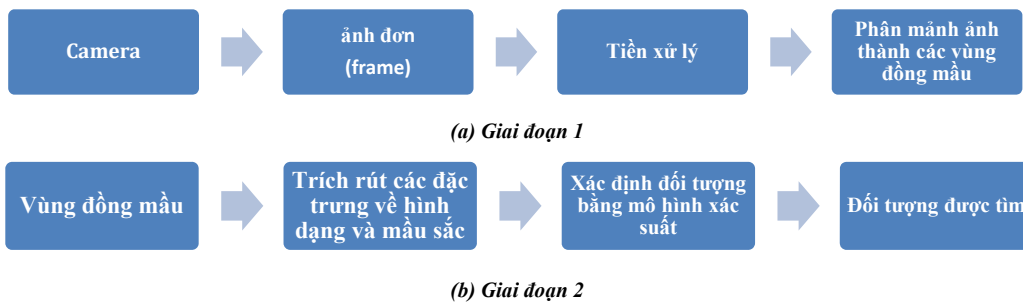
II. PHƯƠNG PHÁP ĐỀ XUẤT

Các mục tiêu quân sự thường phải được quan sát từ xa và các ảnh thu thập được từ camera sẽ bao gồm nhiều đối tượng nền. Do vậy việc sử dụng các phương pháp dựa vào cạnh sẽ không hiệu quả. Mặt khác nếu chúng ta sử dụng các phương pháp dựa vào đặc trưng màu sắc sẽ rất nhạy cảm với điều kiện ánh sáng. Từ việc phân tích các kết quả thực nghiệm phân vùng ảnh bằng phương pháp đồ thị trong [9], chúng tôi tìm thấy rằng thuật toán này phân vùng ảnh nhanh, có độ chính xác và tin cậy cao. Hơn nữa, đặc trưng hình dạng (shape context) được đề xuất trong [10] là một đặc trưng quan trọng trong việc xác định đối tượng, nó có nhiều ưu điểm như là bất biến với sự méo hình cục bộ, sự dịch chuyển và xoay. Thay vì chỉ sử dụng đặc trưng màu, việc kết hợp đặc trưng hình dạng với đặc trưng màu sắc sẽ tăng đáng kể độ chính xác và độ tin cậy trong việc xác định đối tượng dưới các điều kiện môi trường và ánh sáng khác nhau.

Trong phần này, chúng tôi trình bày một phương pháp phát hiện đối tượng mới từ các vùng ảnh cục bộ, sử dụng phương pháp xác suất với sự kết hợp giữa đặc trưng hình dạng với đặc trưng màu sắc. Đầu tiên, ảnh đầu vào được phân mảnh thành các vùng đồng màu sử dụng thuật toán trong [9]. Sau đó, đối tượng cần tìm được xác định trong ảnh như là một tập con các vùng đồng màu kết nối với nhau. Tập con này có xác suất hậu nghiệm của đặc trưng màu sắc và hình dạng là cực đại. Theo một cách khác, phương pháp đề xuất phát hiện đối tượng quân sự trong ảnh bao gồm có hai giai đoạn chính như trong Hình 1.

Giai đoạn 1: Phân mảnh ảnh đầu vào thành các vùng đồng màu.

Giai đoạn 2: Tìm đối tượng từ các vùng đồng màu sử dụng mô hình xác suất với đặc trưng màu sắc và hình dạng.



Hình 1. Sơ đồ khối phương pháp đề xuất và ảnh minh họa

Các giai đoạn này được trình bày chi tiết trong các phần sau đây.

A. Phân vùng ảnh

Có nhiều thuật toán phân vùng ảnh khác nhau như các thuật toán chia và trộn vùng, các thuật toán tăng trưởng vùng, các thuật toán áp dụng lý thuyết đồ thị, các thuật toán sử dụng các kỹ thuật máy học, các thuật toán dựa vào cạnh... Tuy nhiên phương pháp dựa vào lý thuyết đồ thị được sử dụng phổ biến nhất. Phương pháp này thể hiện được tính tổng quát hóa trong phân vùng ảnh và có thể áp dụng cho nhiều trường hợp thực tế. Trong công việc này, chúng tôi sử dụng thuật toán dựa vào lý thuyết đồ thị được đề xuất trong [9] để phân mảnh ảnh đầu vào thành các vùng đồng nhất về màu sắc. Thuật toán này độ chính xác cao và thời gian tính toán nhanh.

Từ một ảnh đầu vào, một đồ thị vô hướng $G = \{V, E\}$ được thiết lập với V là tập đỉnh và E là tập cạnh. Các đỉnh $v \in V$ là các điểm ảnh (pixels) và các cạnh (v_i, v_j) nối hai đỉnh liền kề nhau v_i and v_j . Mỗi cạnh (v_i, v_j) có một trọng số $w(v_i, v_j)$ được đo bằng sự khác nhau về màu sắc giữa hai điểm ảnh v_i and v_j như sau:

$$w(v_i, v_j) = \sqrt{(C_1^i - C_1^j)^2 + (C_2^i - C_2^j)^2 + (C_3^i - C_3^j)^2} \quad (1)$$

Trong công thức (1) C_k^i là thành phần màu thứ k của điểm ảnh v_i . Thuật toán phân vùng ảnh được trình bày cụ thể trong Thuật toán 1.

Thuật toán 1: Phân vùng ảnh dựa vào lý thuyết đồ thị

Đầu vào: Một đồ thị vô hướng $G = \{V, E\}$, với n đỉnh, và m cạnh

Đầu ra: Các vùng ảnh đồng màu $S_1, S_2, S_3 \dots$

Bước 1: Sắp xếp các cạnh $e \in E$ theo thứ tự có trọng số $w(e)$ tăng dần.

Bước 2: Khởi gán các vùng ảnh S_i là các đỉnh v_i .

Bước 3: Lặp lại bước 4 với $q = 1..m$ (m là tổng số cạnh trong E).

Bước 4: Hợp 2 vùng S_i và S_j chứa hai đỉnh v_i and v_j ($v_i \in S_i, v_j \in S_j$) được nối với nhau bởi cạnh thứ q nếu thỏa mãn 2 điều kiện sau:

$$\begin{cases} S_i \neq S_j \\ w(v_i, v_j) < MInt(S_i, S_j) \end{cases} \quad (2)$$

Trong công thức (2) $MInt(S_i, S_j)$ là hàm đo sự khác nhau giữa 2 vùng S_i và S_j . Hàm này được tính dựa trên sự khác nhau nội bộ giữa các điểm ảnh trong mỗi vùng S_i và giá trị ngưỡng của S_i . Sự khác nhau nội bộ giữa các điểm ảnh trong một vùng S là tổng số lớn nhất của các cạnh nối các điểm ảnh trong S :

$$Int(S) = \max_{e \in S} w(e) \quad (3)$$

Giá trị ngưỡng của một vùng S được tính như sau:

$$\tau(S) = \frac{K}{|S|} \quad (4)$$

Trong công thức (4) K là một hằng số dương, $|S|$ là tổng số điểm ảnh trong vùng S .

Sự khác nhau giữa hai vùng S_i và S_j được tính như sau:

$$MInt(S_i, S_j) = \min[Int(S_i) + \tau(S_i), Int(S_j) + \tau(S_j)] \quad (5)$$

Bảng 3 minh họa các kết quả phân vùng của một số loại ảnh áp dụng thuật toán trên.

B. Xác định đối tượng

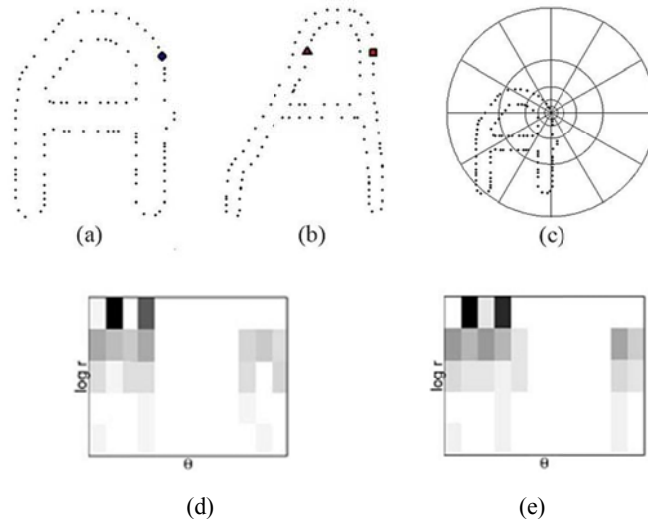
Gọi $\mathcal{S} = \{S_1, S_2, S_3, S_4, \dots\}$ là tập các vùng đồng nhất về màu sắc trong ảnh đầu vào. Các vùng đồng màu này giành được bằng việc áp dụng thuật toán phân vùng ảnh ở phần trên. Đối tượng trong ảnh đầu vào được tìm như là một tập con $\mathcal{S}^* \subseteq \mathcal{S}$. Tất cả các thành phần $S_i \in \mathcal{S}^*$ là các thành phần kết nối và đều thuộc về một vùng ảnh lớn Z . Để xác định tập con \mathcal{S}^* từ \mathcal{S} chúng tôi sử dụng hai đặc trưng nổi bật của đối tượng: *màu sắc* và *hình dạng*. Các đặc trưng của đối tượng được tính toán bằng việc sử dụng các hàm mật độ xác suất trên tập dữ liệu ảnh mẫu (training).

Trên tập dữ liệu ảnh mẫu (training), mỗi ảnh được phân vùng thành các vùng đồng màu S_i , các vùng thuộc đối tượng được xác định bằng tay như là các vùng S_k^o . Như vậy, từ tập mẫu chúng ta sẽ có một tập các vùng đồng màu khác nhau thuộc về đối tượng $\mathcal{S}^o = \{S_1^o, S_2^o, S_3^o, S_4^o, \dots\}$. Đối với mỗi vùng S_k^o , đặc trưng màu $\mathbf{c} = (c_1, c_2, c_3)$ của nó là một vector mà mỗi thành phần c_i ($i = 1, 2, 3$) là giá trị trung bình của thành phần màu thứ i của tất cả các điểm ảnh trong S_k^o . Đặt \mathbf{O} biểu thị cho lớp đối tượng cần tìm. Hàm mật độ xác suất của đối tượng đối với đặc trưng màu $p(\mathbf{c}|\mathbf{O})$ được tính bằng phương pháp histogram màu 3D của tất cả các màu \mathbf{c} của các vùng S_k^o . Mỗi thành phần màu được lượng tử hóa thành N bin.

Đặc trưng hình dạng của đối tượng được xác định bằng việc sử dụng các bộ đặc tả về hình dạng (shape context) [10]. Bộ đặc tả hình dạng này bất biến khi đối tượng bị xoay, dịch chuyển, méo, biến đổi theo tỷ lệ. Đặc trưng hình dạng \mathbf{s} của một đối tượng bao gồm các đặc tả hình dạng của các điểm ảnh nằm trên đường bao của đối tượng. Xét một đối tượng có K các điểm mẫu $p_1, p_2, p_3, \dots, p_k$ trên đường viền (xem Hình 2). Đặc tả hình dạng của một điểm p_i là một histogram của các cực tọa độ tương đối giữa p_i và $K - 1$ điểm còn lại trên đường viền của đối tượng như sau:

$$h_i^k = \#\{q \neq p_i: (q - p_i) \in \text{bin}(k)\}. \quad (6)$$

Các bin là đồng nhất trong không gian \log .



Hình 2. Đặc tả hình dạng: (a) và (b) là các điểm biên của mẫu; (c) là sơ đồ các bin được sử dụng để tính đặc trưng hình dạng; (d) là đặc trưng hình dạng của điểm mẫu trên (a); (e) là đặc trưng hình dạng của điểm mẫu trên (b).

Đối với hai điểm khác nhau trên cùng một đối tượng thì các đặc tả hình dạng của chúng sẽ khác nhau. Đối với các đối tượng có hình dạng giống nhau thì đặc tả hình dạng của hai điểm tương ứng nhau sẽ là giống nhau. Sự khác nhau giữa hai bộ đặc tả hình dạng của hai điểm p và q được tính như sau:

$$C(p, q) = 0.5 \sum_{i=1}^M \frac{(h_p^i - h_q^i)^2}{h_p^i + h_q^i} \tag{7}$$

Trong (7), M là số bin của các histogram tọa độ cực.



Hình 3. Một số hình dạng mẫu của một đối tượng máy bay.

Gọi $T = \{T_1, T_2, T_3, \dots\}$ là tập các hình dạng mẫu của đối tượng cần tìm. Trong ứng dụng của chúng tôi, mỗi loại đối tượng bao gồm 9 hình dạng mẫu. Hình 3 trình bày một hình dạng mẫu của đối tượng máy bay F22. Đối với mỗi hình dạng mẫu, chúng tôi lấy các điểm mẫu p nằm trên đường bao của đối tượng sao cho hai điểm liền kề nhau cách nhau τ điểm ảnh. Các bộ đặc tả hình dạng của các điểm này là đặc trưng hình dạng của mỗi hình dạng mẫu. Đối với mỗi một vùng ảnh Z (nó có thể bao gồm nhiều vùng đồng màu), thì đặc trưng hình dạng s_z của Z là các đặc tả hình dạng của các điểm mẫu nằm trên đường bao ngoài của Z . Sự khác nhau về hình dạng giữa một vùng ảnh Z và một hình dạng mẫu T của đối tượng được tính như sau:

$$D(s_z, T) = \frac{1}{|s_z|} \sum_{p \in s_z} \min_{q \in T} C(p, q) \tag{8}$$

Trong công thức (8), $|s_z|$ biểu thị cho tổng số điểm mẫu trong Z . Hàm mật độ xác suất về đặc trưng hình dạng của một vùng ảnh Z thuộc về đối tượng được tính như là:

$$p(s_z | \mathbf{O}) = \exp[-\beta \min_{T \in T} D(s_z, T)] \tag{9}$$

Trong (9), β là tham số tỷ lệ được xác định thông qua tập dữ liệu mẫu. Công thức (9) có nghĩa rằng khi hình dạng của vùng ảnh Z tương tự với một hình dạng mẫu T của đối tượng thì giá trị $D(s_z, T)$ là nhỏ và do vậy giá trị $p(s_z | \mathbf{O})$ là cao. Ngược lại $p(s_z | \mathbf{O})$ có giá trị thấp khi hình dạng của vùng ảnh Z khác với các hình dạng mẫu T của đối tượng.

Từ một tập các vùng ảnh đồng màu \mathbf{S} , đối tượng cần tìm được xác định như một tập con \mathbf{S}^* như sau:

$$\mathbf{S}^* = \arg \max_Z s p(Z | \mathbf{O}) \tag{10}$$

với

$$p(Z | \mathbf{O}) = p(c_z | \mathbf{O}) p(s_z | \mathbf{O}) \tag{11}$$

Tập \mathbf{S}^* trong (10) có thể được xác định bằng phương pháp vét cạn với việc tìm tất cả các tập con trong \mathbf{S} . Tuy nhiên phương pháp này sẽ mất nhiều thời gian tính toán, độ phức tạp tính toán của nó là $O(2^{|\mathbf{S}|})$ với $|\mathbf{S}|$ là tổng số thành phần của \mathbf{S} . Để giảm thời gian tính toán, chúng tôi sử dụng thuật toán thêm bớt các vùng trong [11]. Thuật toán này hiệu quả hơn khi ta thêm một số điều kiện lọc như: loại bỏ vùng đồng màu lớn (phần lớn là các vùng nền) bằng việc thống kê tiết diện đối tượng so với khung hình. Thuật toán này được mô tả trong Thuật toán 2.

Thuật toán 2: Tìm đối tượng bằng việc thêm bớt vùng
$\mathbf{S}' \leftarrow \{S \in \mathbf{S} p(c_s \mathbf{O}) \geq \tau\}$ $\mathbf{S}^* \leftarrow \arg \max_{S \in \mathbf{S}'} p(c_s \mathbf{O})$ <p><i>Temp</i> \leftarrow Đúng</p> <p>While (<i>Temp</i>) do</p> <div style="margin-left: 20px;"> $\mathbf{S}_a \leftarrow \{S_i \in \{\mathbf{S}' - \mathbf{S}^*\}, \text{với } \mathbf{S}^* \cup S_i \text{ là một vùng liên thông} \}$ $\mathbf{S}^+ \leftarrow \arg \max_{S \in \mathbf{S}_a} p(\{S \cup \mathbf{S}^*\} \mathbf{O})$ $\mathbf{S}_r \leftarrow \{S_i \in \mathbf{S}^*, \text{với } \{\mathbf{S}^* - S_i\} \text{ là một vùng liên thông} \}$ $\mathbf{S}^- \leftarrow \arg \max_{S \in \mathbf{S}_r} p(\{\mathbf{S}^* - S\} \mathbf{O})$ </div> <p>If $p(\{\mathbf{S}^+ \cup \mathbf{S}^*\} \mathbf{O}) \geq p(\{\mathbf{S}^- \cup \mathbf{S}^*\} \mathbf{O})$ and $p(\{\mathbf{S}^+ \cup \mathbf{S}^*\} \mathbf{O}) > p(\mathbf{S}^* \mathbf{O})$ then</p>

```


$$S^* \leftarrow S^* \cup S^+$$

Else if  $p(\{S^* - S^-\} | \mathbf{O}) > p(S^* | \mathbf{O})$  then

$$S^* \leftarrow S^* - S^-$$

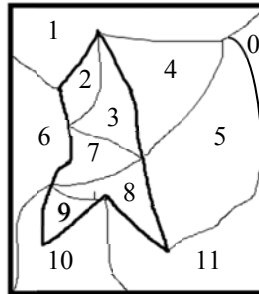
Else

$$Temp \leftarrow Sai$$

End if
End while

```

Trong Thuật toán 2, ở mỗi bước lặp khi thêm hoặc bớt một vùng S_i vào S^* thì tính kết nối của tập $\{S^* \cup S_i\}$ và $\{S^* - S_i\}$ được kiểm tra. Một tập các thành phần được xem là kết nối nếu như tất cả các thành phần của nó đều thuộc về một vùng ảnh chung. Ví dụ trong Hình 4, giả sử S^* bao gồm các thành phần là các vùng ảnh 2, 3, 7, 8 và 9 thì vùng ảnh 7 không thể bớt được từ S^* bởi vì nó sẽ bẻ gãy tính kết nối của S^* . Tương tự thì vùng số 0 không được thêm vào S^* . Bởi vì Thuật toán 2, có số các thành phần và thao tác là xác định nên thuật toán sẽ hội tụ.



Hình 4. Minh họa quá trình thêm và bớt vùng.

Vùng ảnh được tạo ra từ các thành phần S^* được xem như là đối tượng cần tìm nếu điều kiện sau đây được thỏa mãn:

$$p(S^* | \mathbf{O}) \geq \tau_c \quad (12)$$

Trong (12), tham số τ_c là một ngưỡng được xác định trước bằng việc học từ tập dữ liệu mẫu. Bảng 3 trình bày một số ví dụ về xác định đối tượng trong các ảnh đầu vào.

III. PHƯƠNG PHÁP THỰC NGHIỆM VÀ KẾT QUẢ

Trong phần này chúng tôi trình bày các dữ liệu hình ảnh mà chúng tôi thử nghiệm, chi tiết các bước thực hiện thuật toán, phương pháp đánh giá, các tham số và kết quả thử nghiệm phương pháp đề xuất.

A. Dữ liệu ảnh

Để đánh giá phương pháp phát hiện đối tượng, chúng tôi đã thu thập được một cơ sở dữ liệu gồm 5 loại đối tượng và mỗi loại đối tượng có 2000 ảnh ở các môi trường tự nhiên khác nhau và khoảng cách từ camera tới đối tượng khoảng 1000 m. Trong tập dữ liệu của chúng tôi, mỗi ảnh chứa một đối tượng tìm kiếm trong các cảnh thay đổi liên tục (các ảnh thu nhận khi đối tượng di chuyển). Trong nhiều trường hợp đối tượng bị mất một phần do ở đó đường lồi lõm hoặc bị che khuất bởi đối tượng khác, đặc biệt hơn nữa là điều kiện chiếu sáng hoàn toàn khác nhau. Với mỗi ảnh chúng tôi dò bằng tay vùng ảnh của đối tượng. Mỗi đối tượng chúng tôi lấy 500 ảnh mẫu dùng để học và 1500 ảnh dùng để thử nghiệm.

B. Phương pháp đánh giá

Tiêu chí để đánh giá việc phát hiện đối tượng được dựa trên việc so sánh hai vùng ảnh: vùng ảnh đối tượng được phát hiện bằng máy và vùng ảnh được xác định bằng tay. Gọi R_d là một vùng được phát hiện và R_g là vùng được đánh dấu bởi cách làm bằng tay. Tỷ lệ so khớp giữa R_d và R_g được tính như sau:

$$\gamma(R_g, R_d) = \frac{|R_g \cap R_d|}{|R_g \cup R_d|}$$

Trong đó $|R|$ là diện tích của vùng R , \cap và \cup là giao và hợp của hai vùng R_g, R_d . Một vùng R_d được cho là phát hiện đúng khi và chỉ khi tồn tại một vùng được xác định bằng tay R_g sao cho $\gamma(R_g, R_d)$ lớn hơn hoặc bằng một giá trị ngưỡng τ . Tương tự như việc đánh giá của một số hệ thống phát hiện đối tượng, giá trị ngưỡng này được thiết lập là 0.5.

Để đánh giá hiệu quả của một phương pháp phát hiện đối tượng, chúng tôi sử dụng hai phép đo: tỷ lệ phát hiện đúng và độ chính xác phát hiện. Tỷ lệ phát hiện đúng là tỷ lệ phần trăm của số đối tượng được phát hiện đúng trên tổng

số ảnh được kiểm tra. *Độ chính xác phát hiện* là tỷ lệ phần trăm của số đối tượng được phát hiện đúng trên tổng số các đối tượng được dò. Ví dụ chúng ta kiểm tra 100 ảnh, chúng ta phát hiện được 90 đối tượng, có 80 đối tượng được phát hiện đúng (cho $\gamma(R_g, R_d) > \tau$) thì tỷ lệ phát hiện đúng là 90% và độ chính xác là $80/90 = 89\%$.

C. Kết quả thực nghiệm

Kết quả thực nghiệm của phương pháp đề xuất cho 05 loại đối tượng trên tập ảnh kiểm tra được trình bày trong Bảng 1.

Bảng 1. Kết quả đánh giá trên tập ảnh kiểm tra cho 05 đối tượng

STT	Loại ảnh đối tượng	Tỷ lệ phát hiện đúng	Độ chính xác
1	Xe u oát quân sự	78.93%	95.60%
2	Máy bay	95.87%	99.87%
3	Xe tăng	81.27%	94.54%
4	Ca nô	78.7%	95.76%
5	Xe tải quân sự	73.32%	97.63%

Bảng 2 trình bày một số kết quả hình ảnh phát hiện đối tượng của phương pháp đề xuất. Các kết quả này chứng minh rằng phương pháp đề xuất có thể phát hiện được nhiều loại đối tượng khác nhau với độ chính xác cao.

Bảng 2. Một số hình ảnh kết quả của một số đối tượng

STT	Ảnh đầu vào	Ảnh phân mảnh	Ảnh kết quả	Xác định bằng tay
1				
2				
3				
4				
5				
6				
7				
8				

IV. KẾT LUẬN

Bài báo này đề xuất một phương pháp phát hiện đối tượng dựa trên kỹ thuật xử lý ảnh và các thuật toán máy học để phát hiện các mục tiêu quân sự. Đối tượng trong ảnh được xác định bằng một phương pháp xác suất sử dụng kết hợp

các đặc trưng màu sắc và hình dạng. Phương pháp đề xuất được đánh giá trên các tập dữ liệu lớn khác nhau. Mỗi tập dữ liệu bao gồm các ảnh của một kiểu đối tượng được thu thập từ các cảnh khác nhau dưới các điều kiện ánh sáng khác nhau. Các kết quả thực nghiệm đã chỉ ra rằng phương pháp đề xuất có thể phát hiện được các đối tượng khác nhau với độ chính xác cao.

V. TÀI LIỆU THAM KHẢO

- [1] A. Opelt, A. Pinz, and A. Zisserman, "Learning an alphabet of shape and appearance for multi-class object detection," *International Journal of Computer Vision*, vol. 80, pp. 16-44, 2008.
- [2] Z. Si, H. Gong, Y. N. Wu, and S. C. Zhu, "Learning mixed templates for object recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 272-279.
- [3] J. Shotton, A. Blake, and R. Cipolla, "Multiscale categorical object recognition using contour fragments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, pp. 1270-1281, 2008.
- [4] L. Szumilas and H. Wildenauer, "Spatial configuration of local shape features for discriminative object detection," in *Lecture Notes in Computer Science*, vol. 5875, ed, 2009, pp. 22-33.
- [5] L. Szumilas, H. Wildenauer, and A. Hanbury, "Invariant shape matching for detection of semi-local image structures," in *Lecture Notes in Computer Science*, vol. 5627, ed, 2009, pp. 551-562.
- [6] Dilip K. Prasad, "Survey of The Problem of Object Detection In Real Images" *International Journal of Image Processing (IJIP)*, vol. 6, 2012.
- [7] A. R. Pope and D. G. Lowe, "Probabilistic models of appearance for 3-D object recognition," *International Journal of Computer Vision*, vol. 40, pp. 149-167, 2000.
- [8] M. Andriluka, S. Roth, and B. Schiele, "Discriminative Appearance Models for Pictorial Structures," *International Journal of Computer Vision*, vol. 99, pp. 259-280, 2012.
- [9] Pedro . Felzenszwalb, "Efficient Graph-Based Image Segmentation," *International Journal of Computer Vision* 59(2), 167–181, 2004.
- [10] Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape context. *Pattern Analysis and Machine Intelligence*, *IEEE Transactions*, vol. 24, pp.509-522, 2002.
- [11] M. C. Le, S.L. Phung, A. Bouzerdoum, "Lane Detection in Unstructured Environments for Autonomous Navigation Systems," in *Proceedings of the Asian Conference on Computer Vision*, pp.414-429, 2014.

A OBJECT DETECTION METHOD FOR AUTONOMOUS TRACKING AND SURVEILLANCE SYSTEMS

Nguyen Van Hung, Nguyen Van Xuat, Le Manh Cuong

ABSTRACT- *Object detection is a crucial component in autonomous surveillance and tracking systems. This paper proposes a vision-based method for military object detection. The proposed method detects the object in each scene image captured by a camera, based on shape and color features. In our method, the image is first segmented into color homogeneous sub-regions using the graph-based algorithm. Then, the background sub-regions are removed by employing color information. Finally, the object is determined from connected sub-regions using a probabilistic model that integrates shape and color features. The shape feature of an image region is represented by shape contexts. Shape contexts are known for their invariance to illumination, local shape deformation, scale and rotation. The proposed method is evaluated on various large datasets in which each dataset includes images of an object type collected from various scenes under vary illumination conditions. The experimental results show the efficiency and robustness of the proposed method.*