

PHƯƠNG PHÁP CHỌN NGUỖNG TỪ BIỂU ĐỒ MỨC XÁM OTSU'S METHOD

NGUYỄN HOÀI NAM

Ngày 10 tháng 12 năm 2023

Tóm tắt nội dung

Phương pháp chọn ngưỡng Otsu (Otsu's Method) - Một phương pháp chọn ngưỡng một cách tự động không cần tham số và không giám sát để phân đoạn ảnh. Ngưỡng tối ưu được chọn theo tiêu chí phân biệt, cụ thể là để tối đa hóa khả năng phân tách của các lớp kết quả ở mức xám. Quy trình của phương pháp này sử dụng mômen tích lũy bậc 0 và bậc 1 của biểu đồ mức xám. Từ Otsu's Method, thật đơn giản để mở rộng phương pháp này cho các bài toán đa ngưỡng. Một minh họa cũng được trình bày để làm rõ cách thức tìm ngưỡng Otsu (Otsu Thresholding), đồng thời, giải thích ý nghĩa của công thức có trong Otsu's Method.

1 Giới thiệu

Phương pháp chọn ngưỡng Otsu, hay còn được gọi là phân loại Otsu hoặc Otsu Thresholding, là một kỹ thuật quan trọng trong xử lý ảnh và thị giác máy tính. Phương pháp này được phát triển bởi Nobuyuki Otsu [1] vào năm 1979 và thường được sử dụng để tự động xác định ngưỡng tốt nhất để phân tách hai lớp hoặc nhóm pixel trong ảnh.

Điều quan trọng trong xử lý ảnh là chọn ngưỡng mức xám thích hợp để trích xuất các đối tượng từ nền của chúng. Trong trường hợp lý tưởng, biểu đồ mức xám có một thung lũng sâu và sắc nét giữa hai đỉnh tương ứng đại diện cho vật thể và nền, do đó có thể chọn ngưỡng ở đáy thung lũng này [2]. Tuy nhiên, đối với hầu hết các bức ảnh thực, thường khó phát hiện chính xác đáy thung lũng, đặc biệt trong những trường hợp như khi thung lũng bằng phẳng và rộng, nhiều nhiễu hoặc khi hai đỉnh có chiều cao cực kỳ chênh lệch, thường không thể theo dõi được như hình 1. Đã có một số kỹ thuật được đề xuất nhằm khắc phục những khó khăn này, trong đó có phương pháp chọn ngưỡng Otsu.

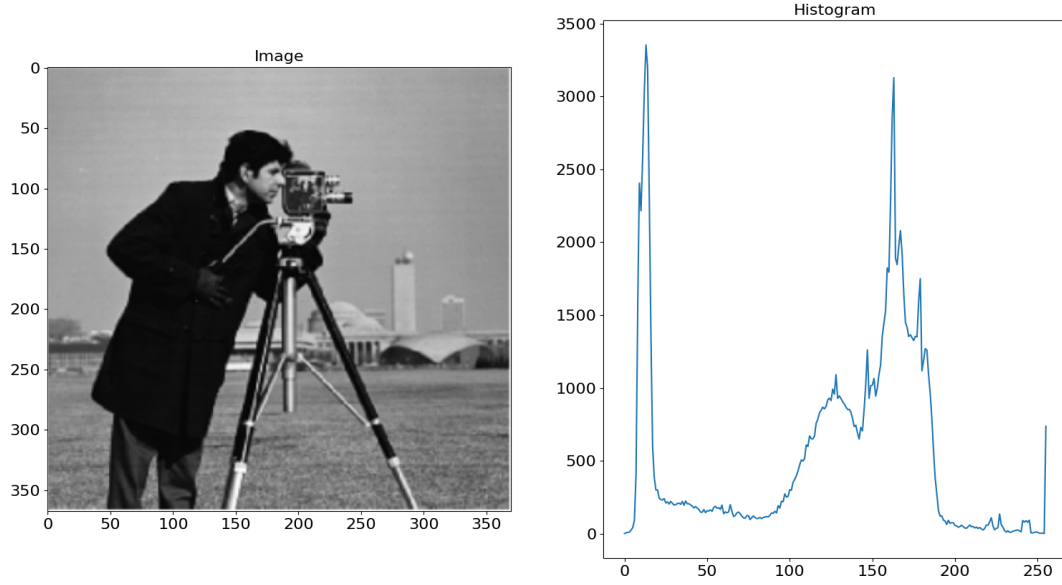
Phương pháp lựa chọn ngưỡng Otsu sẽ giới hạn trong trường hợp cơ bản về lựa chọn ngưỡng trong đó chỉ có biểu đồ mức xám là đủ. Nó không chỉ quan trọng như một kỹ thuật tiêu chuẩn trong xử lý ảnh mà còn cần thiết cho các vấn đề quyết định không giám sát trong nhận dạng mẫu. Otsu's Method là một phương pháp mới được đề xuất từ quan điểm phân tích phân biệt; nó trực tiếp tiếp cận tính khả thi của việc đánh giá mức độ "tốt" (khả năng phân tách) của ngưỡng và tự động chọn ngưỡng tối ưu.

2 Phương pháp

Ý tưởng cơ bản của phương pháp chọn ngưỡng Otsu là tìm ngưỡng k sao cho tổng phương sai của các mức (the total variance of levels) là lớn nhất. Phương sai càng lớn, sự phân tách giữa hai lớp càng tốt.

Đầu tiên, hãy để các pixel của hình ảnh được thể hiện dưới dạng L mức xám $[1, 2, \dots, L]$. Số lượng các pixel ở mức i được kí hiệu là n_i và tổng số các pixel là $N = n_1 + n_2 + \dots + n_L$. Để đơn giản hóa, biểu đồ mức xám được chuẩn hóa và được coi là phân bố xác suất:

$$p_i = n_i/N, \quad p_i \geq 0, \quad \sum_{i=1}^L p_i = 1 \quad (1)$$



Hình 1: Ví dụ biểu đồ Histogram của một bức ảnh thực tế .

Giả sử rằng chúng ta phân đôi các pixel thành hai lớp C_0 và C_1 (nền và đối tượng, hoặc ngược lại) theo một ngưỡng k ; C_0 biểu thị các pixel có mức độ $[1, \cdot, k]$, và C_1 biểu thị các pixel có mức độ $[k + 1, \dots, L]$. Khi đó xác suất xuất hiện của lớp C_0 và lớp C_1 lần lượt được tính bởi công thức 2 và 3:

$$\omega_0 = \Pr(C_0) = \sum_{i=1}^k p_i = \omega(k) \quad (2)$$

$$\omega_1 = \Pr(C_1) = \sum_{i=k+1}^L p_i = 1 - \omega(k) \quad (3)$$

và mức trung bình của lớp tương ứng được tính bởi công thức 4 và 5:

$$\mu_0 = \sum_{i=1}^k i \Pr(i | C_0) = \sum_{i=1}^k i p_i / \omega_0 = \mu(k) / \omega(k) \quad (4)$$

$$\mu_1 = \sum_{i=k+1}^L i \Pr(i | C_1) = \sum_{i=k+1}^L i p_i / \omega_1 = \frac{\mu_T - \mu(k)}{1 - \omega(k)} \quad (5)$$

khi

$$\omega(k) = \sum_{i=1}^k p_i \quad (6)$$

và

$$\mu(k) = \sum_{i=1}^k i p_i \quad (7)$$

lần lượt là moment tích lũy bậc 0 (công thức 6) và bậc 1 (công thức 7) của biểu đồ lên đến mức thứ k , và công thức 8

$$\mu_T = \mu(L) = \sum_{i=1}^L i p_i \quad (8)$$

là tổng mức trung bình (moment tích lũy bậc 1) của ảnh gốc. Chúng ta có thể dễ dàng xác minh mối quan hệ sau đây (công thức 9) với mọi k :

$$\omega_0\mu_0 + \omega_1\mu_1 = \mu_T, \quad \omega_0 + \omega_1 = 1 \quad (9)$$

Các phương sai của lớp C_0 và C_1 lần lượt được tính bởi công thức 10 và 11:

$$\sigma_0^2 = \sum_{i=1}^k (i - \mu_0)^2 \Pr(i | C_0) = \sum_{i=1}^k (i - \mu_0)^2 p_i / \omega_0 \quad (10)$$

$$\sigma_1^2 = \sum_{i=k+1}^L (i - \mu_1)^2 \Pr(i | C_1) = \sum_{i=k+1}^L (i - \mu_1)^2 p_i / \omega_1 \quad (11)$$

Hai công thức trên yêu cầu moment tích lũy bậc 2 (thống kê).

Để đánh giá mức độ "tốt" của ngưỡng (ở mức k), sau đây là các độ đo được sử dụng trong phân tích phân biệt (the discriminant analysis):

$$\lambda = \sigma_B^2 / \sigma_W^2, \quad \kappa = \sigma_T^2 / \sigma_W^2, \quad \eta = \sigma_B^2 / \sigma_T^2 \quad (12)$$

khi

$$\sigma_W^2 = \omega_0\sigma_0^2 + \omega_1\sigma_1^2 \quad (13)$$

$$\begin{aligned} \sigma_B^2 &= \omega_0 (\mu_0 - \mu_T)^2 + \omega_1 (\mu_1 - \mu_T)^2 \\ &= \omega_0\omega_1 (\mu_1 - \mu_0)^2 \end{aligned} \quad (14)$$

(dựa theo công thức 9) và

$$\sigma_T^2 = \sum_{i=1}^L (i - \mu_T)^2 p_i \quad (15)$$

lần lượt là phương sai trong lớp (the within-class variance - λ), phương sai giữa các lớp (the between-class variance - κ) và tổng phương sai của các mức (the total variance of levels - η). Khi đó, bài toán được rút gọn thành bài toán tối ưu hóa để tìm kiếm ngưỡng k làm cực đại hóa một trong các độ đo (hàm đối tượng) trong công thức 12.

Tuy nhiên, λ, κ và η tương ứng với k là tương đương với nhau.

ví dụ: $\kappa = \lambda + 1$ và $\eta = \lambda / (\lambda + 1)$ xét theo λ , vì mối quan hệ cơ bản sau (công thức 16) luôn đúng:

$$\sigma_W^2 + \sigma_B^2 = \sigma_T^2 \quad (16)$$

Cần lưu ý rằng σ_W^2 và σ_B^2 là các hàm của mức ngưỡng k , nhưng σ_T^2 là độc lập với k . Cũng cần lưu ý rằng σ_W^2 dựa trên moment tích lũy bậc 2 (phương sai lớp), trong khi σ_B^2 dựa trên moment tích lũy bậc 1 (trung bình lớp). Do đó, η là độ đo đơn giản nhất đối với k . Do đó, phương pháp chọn ngưỡng Otsu sử dụng η làm độ đo để đánh giá mức độ "tốt" (hoặc khả năng phân tách) của ngưỡng ở mức k .

Ngưỡng tối ưu k^* giúp tối đa hóa η hoặc tương đương với tối đa hóa σ_B^2 (dựa theo công thức 12), được chọn bằng tìm kiếm tuần tự sau bằng cách sử dụng các đại lượng tích lũy đơn giản ở công thức 6, 7 và 8:

$$\eta(k) = \sigma_B^2(k) / \sigma_T^2 \quad (17)$$

$$\sigma_B^2(k) = \frac{[\mu_T \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \quad (18)$$

và ngưỡng tối ưu k^* là (công thức 19):

$$\sigma_B^2(k^*) = \max_{1 \leq k < L} \sigma_B^2(k) \quad (19)$$

Từ bài toán, phạm vi k mà giá trị cực đại được tìm kiếm có thể được giới hạn ở công thức 20:

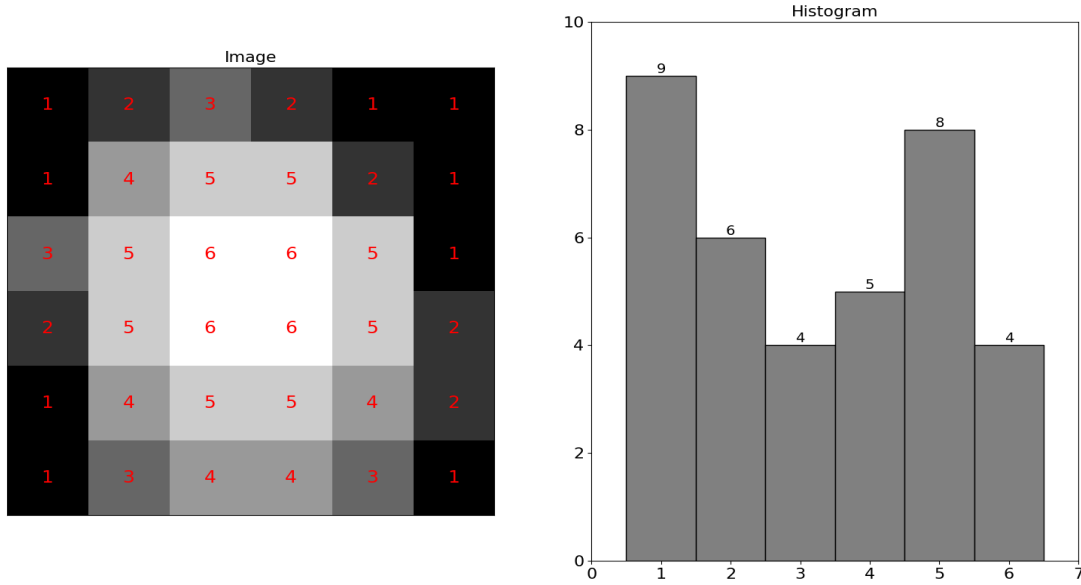
$$S^* = \{k; \omega_0\omega_1 = \omega(k)[1 - \omega(k)] > 0, \quad \text{or } 0 < \omega(k) < 1\}. \quad (20)$$

Nó được gọi là phạm vi hiệu quả của biểu đồ mức xám. Dựa theo công thức 14, độ đo σ_B^2 (hoặc η) nhận giá trị tối thiểu bằng 0 đối với k như $k \in S - S^* = \{k; \omega(k) = 0 \text{ hoặc } 1\}$ và nhận giá trị dương và giới hạn đối với $k \in S^*$. Do đó, hiển nhiên là cực đại luôn tồn tại.

3 Minh hoạ và giải thích ý nghĩa công thức

Cho một ảnh I được biểu diễn bởi ma trận số liệu như bên trái hình 2.

Thực hiện theo phương pháp chọn ngưỡng Otsu, bước đầu tiên, ta tính biểu đồ mức xám (histogram) của ảnh I như hình bên phải hình 2:



Hình 2: Ảnh I và biểu đồ Histogram của ảnh I.

Dựa vào hình 2, ta dễ dàng nhận thấy ảnh I được thể hiện dưới dạng 6 mức xám $[1, 2, \dots, 6]$ ($L = 6$). Tổng các pixel là $N = n_1 + n_2 + n_3 + n_4 + n_5 + n_6 = 9 + 6 + 4 + 5 + 8 + 4 = 36$ pixel. Áp dụng công thức 1:

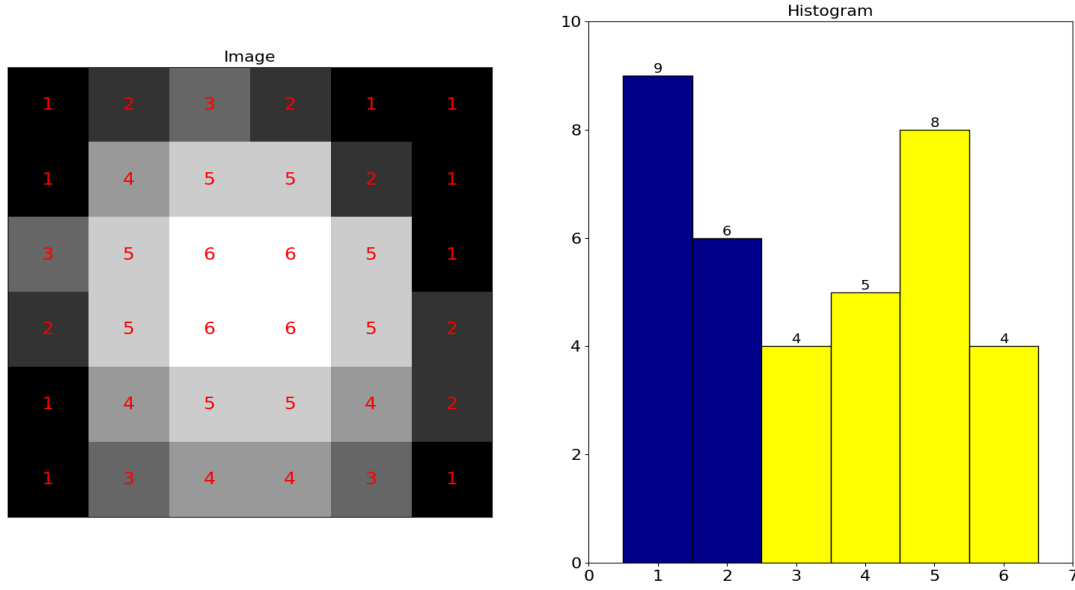
$$p_i = n_i/36, \quad p_i \geq 0, \quad \sum_{i=1}^6 p_i = 1$$

Như vậy, ta được

$$p_1 = 0.25; p_2 = 0.17; p_3 = 0.11; p_4 = 0.14; p_5 = 0.22; p_6 = 0.11$$

Từ đó, có thể thấy, ý nghĩa của công thức 1 chính là để làm đơn giản hoá bài toán, chuẩn hoá biểu đồ xám và được xem là phân bố xác suất.

Bước thứ hai, chọn ngưỡng $k = 2$, chia ảnh I thành hai lớp C_0 (màu xanh đậm) và C_1 (màu vàng). Lớp C_0 biểu thị các pixel có mức độ $[1, 2]$ và lớp C_1 biểu thị các pixel có mức độ $[3, 4, 5, 6]$ như hình 3 dưới đây.



Hình 3: Biểu đồ Histogram sau khi được chia lớp theo ngưỡng $k = 2$.

Áp dụng công thức 2 và 3, ý nghĩa của hai công thức này là để tính được xác suất xuất hiện của hai lớp C_0 và lớp C_1 .

$$\omega_0 = \Pr(C_0) = \sum_{i=1}^2 p_i = p_1 + p_2 = 0.25 + 0.17 = 0.42 = \omega(2)$$

$$\omega_1 = \Pr(C_1) = \sum_{i=3}^6 p_i = p_3 + p_4 + p_5 + p_6 = 0.11 + 0.14 + 0.22 + 0.11 = 0.58 = 1 - \omega(2)$$

Áp dụng công thức 4 và 5, ý nghĩa của hai công thức này là để tính được mức trung bình của hai lớp C_0 và lớp C_1 .

$$\mu_0 = \sum_{i=1}^2 i \Pr(i | C_0) = \sum_{i=1}^2 i p_i / \omega_0 = \frac{1 * p_1 + 2 * p_2}{0.42} = \frac{0.59}{0.42} = \frac{\mu(2)}{\omega(2)} = 1.404$$

$$\mu_1 = \sum_{i=3}^6 i \Pr(i | C_1) = \sum_{i=3}^6 i p_i / \omega_1 = \frac{3 * p_3 + 4 * p_4 + 5 * p_5 + 6 * p_6}{0.58} = \frac{2.65}{0.58} = \frac{\mu_T - \mu(2)}{1 - \omega(2)} = 4.57$$

Từ công thức 6, ta có:

$$\omega(2) = \sum_{i=1}^2 p_i = p_1 + p_2 = 0.25 + 0.17 = 0.42; \mu(2) = \sum_{i=1}^2 i p_i = 1 * p_1 + 2 * p_2 = 1 * 0.25 + 2 * 0.17 = 0.59$$

Và từ công thức 7, ta có:

$$\mu_T = \mu(6) = \sum_{i=1}^6 i p_i = 1 * p_1 + 2 * p_2 + 3 * p_3 + 4 * p_4 + 5 * p_5 + 6 * p_6 = 3.24$$

Như vậy, chúng ta có hai cách tính mức trung bình của hai lớp C_0 và lớp C_1 : cách đầu tiên, ta sử dụng hai công thức 4 và 5; cách thứ hai, ta sử dụng hai công thức 6 và 7 lần lượt là moment tích lũy bậc 0 và 1, và công thức 8 là tổng mức trung bình của ảnh I.

Dựa vào các giá trị đã tính được, ta dễ dàng thấy được mối quan hệ giữa xác suất xuất hiện và mức trung bình của các lớp (dựa theo công thức 9):

$$\omega_0\mu_0 + \omega_1\mu_1 = 0.42 * 1.404 + 0.58 * 4.57 = 3.24 = \mu_T, \quad \omega_0 + \omega_1 = 0.42 + 0.58 = 1$$

Ta có thể tính phương sai của hai lớp C_0 và lớp C_1 theo công thức 10 và 11. Tuy nhiên, hai công thức này yêu cầu moment tích lũy bậc 2, để tìm ra ngưỡng tối ưu k^* dễ dàng hơn, ta tối đa một giá trị dựa trên moment tích lũy bậc 1 là giá trị σ_B^2 (dựa theo công thức 12).

Bước thứ ba, áp dụng công thức 14:

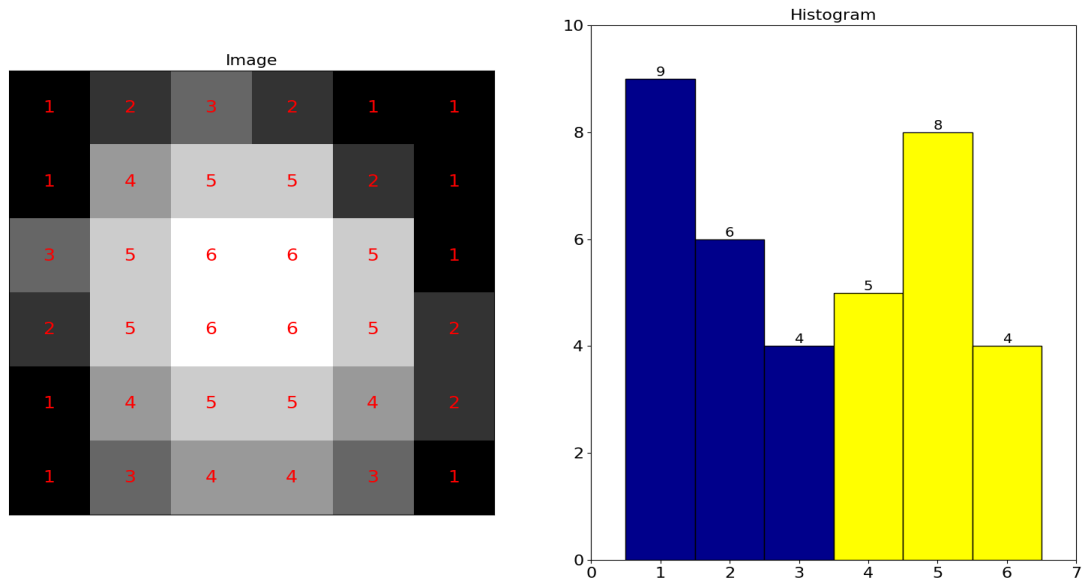
$$\sigma_B^2 = \omega_0\omega_1 (\mu_1 - \mu_0)^2 = 0.42 * 0.58 * (4.57 - 1.404)^2 = 2.442$$

Giá trị trên là giá trị σ_B^2 ở ngưỡng $k = 2$, để tìm được ngưỡng k^* tối ưu, ta áp dụng công thức 19 để tìm được giá trị k tốt nhất, kết quả tìm được như bảng 1.

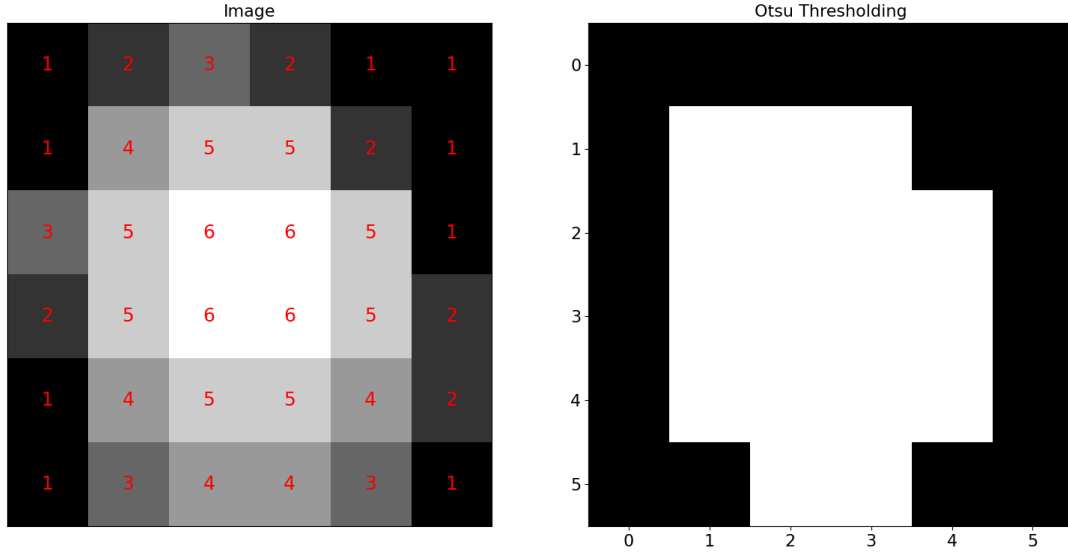
Bảng 1: Bảng số liệu quá trình tìm ngưỡng tối ưu k^* .

Mức xám	1	2	3	4	5
ω_0	0.25	0.42	0.53	0.67	0.89
μ_0	1	1.40	1.74	2.21	2.91
ω_1	0.75	0.58	0.47	0.33	0.11
μ_1	3.97	4.57	4.94	5.33	6
σ_B^2	1.65	2.44	2.55	2.16	0.93

Dựa vào kết quả ở bảng 1, nhận thấy với $k = 3$, giá trị σ_B^2 là lớn nhất là 2.55, như vậy, ngưỡng tối ưu k^* chính là 3. Với $k = 3$, ta có kết quả như sau:



Hình 4: Kết quả phân lớp ảnh I dựa trên biểu đồ Histogram (với $k = 3$).



Hình 5: Kết quả cuối cùng phân lớp ảnh I bằng phương pháp chọn ngưỡng Otsu (với $k = 3$).

4 Thảo luận và nhận xét

4.1 Phân tích các khía cạnh quan trọng

Otsu Thersholding cung cấp thêm phương tiện để phân tích các khía cạnh quan trọng ngoài việc lựa chọn ngưỡng tối ưu.

Đối với ngưỡng đã chọn k^* , xác suất của lớp C_0 và C_1 ở công thức 2 và 3 chỉ ra các phần diện tích được chiếm giữ bởi các lớp C_0 và C_1 trong hình được phân ngưỡng. Công thức 4 và 5 đóng vai trò ước tính mức trung bình của các lớp trong ảnh cấp độ xám ban đầu.

Giá trị tối đa $\eta(k^*)$, được ký hiệu đơn giản là η^* , có thể được sử dụng làm độ đo để đánh giá khả năng phân tách của các lớp đối với hình ảnh gốc hoặc tính lưỡng cực của biểu đồ. Đây là một độ đo quan trọng, vì nó bất biến dưới các phép biến đổi affine của thang mức xám (tức là, đối với bất kỳ sự dịch chuyển và giãn nở nào, $g'_i = ag_i + b$). Nó được xác định duy nhất trong phạm vi

$$0 \leq \eta^* \leq 1$$

Giới hạn dưới (zero) có thể đạt được bởi và chỉ bởi các ảnh có một mức xám không đổi duy nhất, và giới hạn trên (unity) có thể đạt được bởi và chỉ bởi các ảnh có hai giá trị.

4.2. Mở rộng ra bài toán Đa ngưỡng (Multithresholding)

Việc mở rộng phương pháp này cho các bài toán đa ngưỡng là đơn giản nhờ tiêu chí phân biệt. Ví dụ: trong trường hợp ba ngưỡng, Otsu giả sử hai ngưỡng: $1 \leq k_1 < k_2 < L$, để phân tách ra ba lớp, C_0 cho $[1, \dots, k_1]$, C_1 cho $[k_1 + 1, \dots, k_2]$ và C_2 cho $[k_2 + 1, \dots, L]$. Khi đó, độ đo σ_B^2 (cũng là η) là hàm của hai biến k_1 và k_2 và một tập hợp các ngưỡng tối ưu k_1^* và k_2^* được chọn bằng cách tối đa hóa σ_B^2 :

$$\sigma_B^2(k_1^*, k_2^*) = \max_{1 \leq k_1 < k_2 < L} \sigma_B^2(k_1, k_2).$$

Cần lưu ý rằng các ngưỡng được chọn thường trở nên kém tin cậy hơn khi số lượng lớp được tách ra tăng lên. Điều này là do độ đo (σ_B^2), được xác định theo thang đo một chiều (mức xám), có thể dần mất đi ý nghĩa khi số lượng lớp tăng lên. Biểu thức σ_B^2 và quy trình cực đại hóa cũng ngày càng trở nên phức tạp hơn. Tuy nhiên, chúng rất đơn giản với $M = 2$ và 3 , bao gồm hầu hết tất cả các ứng dụng thực tế, do đó hầu như không cần đến một phương pháp đặc biệt để giảm bớt quá trình tìm kiếm. Cần

lưu ý rằng các tham số được yêu cầu trong phương pháp hiện tại cho ngưỡng M là các ngưỡng riêng biệt $M - 1$, trong khi phương pháp tham số, trong đó biểu đồ mức xám được tính gần đúng bằng tổng phân bố Gaussian, yêu cầu $3M - 1$ tham số liên tục.

5 Kết luận

Phương pháp chọn ngưỡng Otsu là phương pháp chọn ngưỡng tự động từ biểu đồ mức xám được bắt nguồn từ quan điểm phân tích phân biệt. Điều này trực tiếp giải quyết vấn đề đánh giá mức độ tốt của các ngưỡng. Ngưỡng tối ưu được chọn theo tiêu chí phân biệt; cụ thể là bằng cách tối đa hóa độ đo phân biệt η (hoặc độ đo khả năng phân tách của các lớp kết quả ở mức xám).

Tài liệu

- [1] Nobuyuki Otsu. A threshold selection method from gray-level histograms. 1979.
- [2] R. W. Sittler. An optimal data association problem in surveillance theory. IEEE Trans, Mil. Elect., vol. MIL-8, pp. 125-139, 1964.