



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0047528

(51)^{2020.01} G06K 9/00

(13) B

(21) 1-2021-07680

(22) 30/11/2021

(45) 25/06/2025 447

(43) 25/02/2022 407A

(73) TẬP ĐOÀN CÔNG NGHIỆP - VIỆN THÔNG QUÂN ĐỘI (VN)

Lô D26 Khu đô thị mới Cầu Giấy, phường Yên Hoà, quận Cầu Giấy, thành phố Hà Nội

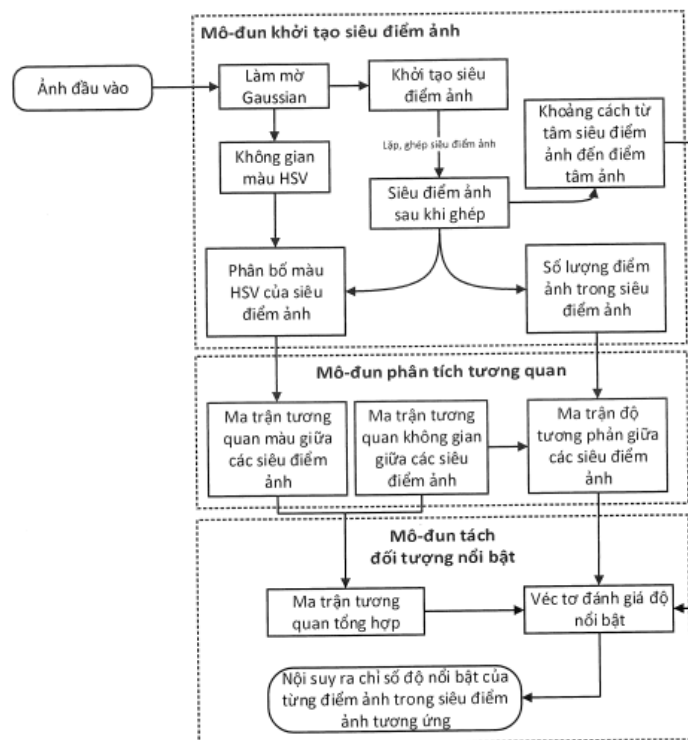
(72) Lê Văn Bằng (VN); Lữ Mạnh Hùng (VN).

(74) Công ty TNHH NACILAW (NACILAW)

(54) PHƯƠNG PHÁP XÁC ĐỊNH ĐỐI TƯỢNG NỔI BẬT TRONG ẢNH

(21) 1-2021-07680

(57) Sáng chế này đề cập đến phương pháp xác định đối tượng nổi bật ở trong ảnh dựa trên phân tích siêu điểm ảnh. Phương pháp thực hiện bao gồm khởi tạo siêu điểm ảnh sử dụng thuật toán SLIC, đồng thời thực hiện ghép các siêu điểm ảnh liền kề mà có phân bố màu sắc tương đồng. Sau đó, thực hiện tính toán tương quan phân bố màu sắc, không gian giữa các siêu điểm ảnh sau khi được ghép. Kết hợp với các chỉ số thống kê về độ lấp đầy, khoảng cách đến tâm ảnh ban đầu và độ tương phản toàn cục của các siêu điểm ảnh, để tính toán ra véc-tơ đánh giá độ nổi bật đối với từng siêu điểm ảnh. Cuối cùng, nội suy độ nổi bật đối với từng điểm ảnh trong siêu điểm ảnh.



Hình 1

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp xác định đối tượng nổi bật trong ảnh. Cụ thể phương pháp xác định đối tượng nổi bật trong ảnh được ứng dụng trong lĩnh vực trí tuệ nhân tạo và thị giác máy tính, giúp xử lý, phân tích các đối tượng có phân bố màu sắc nổi bật hơn.

Tình trạng kỹ thuật của sáng chế

Trong các hệ thống thị giác máy tính nói chung và các hệ thống bóc tách đối tượng nói riêng, việc phân tích các đối tượng có tính chất nổi bật hơn trong ảnh có vai trò quan trọng trong việc xác định những yếu tố cần chú ý trong ảnh, ví dụ như bóc tách sự vật ra khỏi ảnh nền, phát hiện và đánh dấu những đặc điểm, đặc trưng mang mặc nổi bật của người, hoặc xác định những hành vi bất thường trong chuỗi ảnh của video.

Trong lĩnh vực xác định đối tượng nổi bật, siêu điểm ảnh được cho là một trong những kỹ thuật phổ biến để phân tích và gom nhóm những điểm ảnh gần nhau có phân bố màu sắc tương đồng. Tuy nhiên, kỹ thuật này thông thường chỉ có vai trò phân tích toàn cục đối với ảnh, chưa thể xác định được những siêu điểm ảnh nổi bật hơn so với môi trường ảnh xung quanh.

Hiện nay, việc phân tích đối tượng nổi bật trong ảnh cũng có thể sử dụng các mô hình học sâu, tuy nhiên vấn đề của các mô hình này là phải cung cấp dữ liệu huấn luyện, và hiệu quả phân tích đối với các đối tượng chưa xuất hiện trong dữ liệu huấn luyện rất kém. Nếu sử dụng phân tích siêu điểm ảnh, thì tính thích ứng với những sự vật, sự kiện mới rất cao, có thể linh hoạt sử dụng trong nhiều trường hợp ứng dụng khác nhau.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là khắc phục các vấn đề kỹ thuật còn tồn tại, như nêu trên, bằng việc đề xuất phương pháp bóc tách đối tượng nổi bật trong ảnh. Phương pháp đề cập trong sáng chế có mục đích cung cấp mô hình bóc tách các đối tượng được cho là nổi bật, dựa trên phân tích đặc trưng về phân bố màu và tương quan không gian.

Trong sáng chế này, phương pháp tách đối tượng nổi bật trong ảnh dựa trên một hệ thống bao gồm các bộ phận cấu thành như sau: mô-đun khởi tạo siêu điểm ảnh đối với ảnh đầu vào, đồng thời ghép các siêu điểm ảnh liên kề có phân bố màu sắc tương đồng, mô-

đơn phân tích sự tương quan không gian và màu sắc giữa các siêu điểm ảnh, mô-đun lựa chọn và bóc tách các siêu điểm ảnh, thực hiện bóc tách đối tượng nổi bật trong ảnh.

Cụ thể, để đạt được mục đích đã nêu ở trên, phương pháp tách đối tượng bao gồm các bước thực hiện như sau:

Bước 1: thực hiện tiền xử lý ảnh đầu vào; bước này sử dụng thuật toán làm mờ Gaussian, chuyển ảnh từ không gian màu đỏ-xanh lá-xanh lam (red-green-blue: RGB) sang tông màu-độ bão hòa-độ sáng (Hue-Saturation-Illumination: HSV), đồng thời sử dụng thuật toán SLIC (Simple Linear Iterative Clustering) để khởi tạo siêu điểm ảnh và thống kê các đặc điểm về phân bố màu sắc và không gian của từng siêu điểm ảnh.

Bước 2: ghép các siêu điểm ảnh có độ tương quan màu sắc tương đồng

Bước 3: thống kê đặc tính màu sắc và không gian của các siêu điểm ảnh; tại bước này, dựa trên các siêu điểm ảnh được ghép sau bước 2, tiến hành thống kê mô tả đặc tính về quy mô của siêu điểm ảnh, khoảng cách đến trung tâm ảnh ban đầu và tính toán độ tương quan về phân bố màu sắc, phân bố không gian, độ tương phản của các siêu điểm ảnh.

Bước 4: phân tích kết quả phán đoán cuối cùng về độ phân biệt giữa các siêu điểm ảnh; tại bước này dựa trên tính toán ma trận tương quan tổng hợp, và véc-tơ đánh giá độ nổi bật của siêu điểm ảnh, và nội suy ra chỉ số độ nổi bật của từng điểm ảnh trong siêu điểm ảnh tương ứng.

Mô tả vắn tắt các hình vẽ

Hình 1 là hình vẽ thể hiện các bước thực hiện phương pháp xác định đối tượng nổi bật trong ảnh dựa trên phân tích siêu điểm ảnh;

Hình 2 là hình vẽ thể hiện phương pháp tính chỉ số lấp đầy của siêu điểm ảnh;

Hình 3 là hình vẽ thể hiện hiệu quả phương pháp đối với ảnh nải chuối;

Hình 4 là hình vẽ thể hiện hiệu quả phương pháp đối với ảnh đồng hồ;

Hình 5 là hình vẽ thể hiện hiệu quả phương pháp đối với ảnh người.

Mô tả chi tiết các phương án thực hiện sáng chế

Tham chiếu Hình 1, là hình vẽ thể hiện các bước thực hiện phương pháp xác định đối tượng nổi bật trong ảnh.

Cụ thể, các bước thực hiện phương pháp như sau:

Bước 1: thực hiện tiền xử lý ảnh đầu vào;

Bước này sử dụng thuật toán làm mờ Gaussian, chuyển ảnh từ không gian màu đỏ-xanh lá-xanh lam (red-green-blue: RGB) sang tông màu-độ bão hòa-độ sáng (Hue-Saturation-Illumination: HSV), đồng thời sử dụng thuật toán lặp phân nhóm màu tuyến tính đơn (Simple Linear Iterative Clustering, SLIC) để khởi tạo siêu điểm ảnh và thống kê các đặc điểm về phân bố màu sắc và không gian của từng siêu điểm ảnh. Cụ thể được thực hiện qua các bước nhỏ sau:

B1.1: ảnh $I(x, y)$ là ảnh RGB, được tích chập với một nhân Gaussian k có kích thước $[3, 3]$, và sigma bằng 0,5, thực hiện thao tác làm mờ ảnh, công thức như sau:

$$I(x, y) \leftarrow I(x, y) \otimes k(i, j)$$

B1.2: khởi tạo siêu điểm ảnh từ ảnh đã làm mờ $I(x, y)$, sử dụng thuật toán SLIC, với tham số n là số lượng siêu điểm ảnh tối đa có thể được khởi tạo. Sau khi khởi tạo, ảnh ban đầu được chia thành nhiều siêu điểm ảnh lớn $S(i)$, trong đó các vùng ảnh có tính chất tương đồng gần nhau sẽ được gom vào cùng một siêu điểm ảnh, và vị trí của siêu điểm ảnh được đánh dấu từ nhỏ đến lớn theo chiều từ trái sang phải, từ trên xuống dưới.

B1.3: chuyển ảnh ban đầu $I(x, y)$ từ không gian màu RGB sang không gian màu HSV, được ảnh $I'(x, y)$.

Bước 2: ghép các siêu điểm ảnh có độ tương quan màu sắc tương đồng;

Cụ thể được thực hiện qua các bước nhỏ như sau:

B2.1: ảnh HSV $I'(x, y)$ tính ở bước B1.3 kết hợp tập hợp siêu điểm ảnh $S(i)$ tính ở B1.2, tính toán phân bố màu sắc trên không gian màu HSV đối với từng siêu điểm ảnh. Cụ thể công thức tính toán như sau:

$$h(i) = \underset{i \in [1, m]}{\text{histogram}}(I'_{S(i)}(x, y), \text{bins} = [8, 8, 8])$$

Trong đó, bins là số nhóm được phân chia từ 256 bậc xám, sau khi tính toán phân bố màu sắc đối với mỗi siêu điểm ảnh, được ma trận $h(i)$ có kích thước $m \times 8 \times 8 \times 8$.

Từ ma trận $h(i)$, tính toán ma trận tương quan về màu sắc giữa các siêu điểm ảnh, sử dụng chỉ số Bhattacharyya (Bhattacharyya coefficient, BC). Công thức tính toán như sau:

$$\text{color}(i, j) = BC(h_i, h_j)_{i, j \in [1, m]}$$

với BC là chỉ số Bhattacharyya, công thức tính BC đối với hai véc-tơ trên cùng không gian như sau:

$$BC(p, q) = -\ln(10^{-1} + \sum \sqrt{p_i q_i})$$

Trong công thức trên, 10^{-1} được thêm vào để tránh trường hợp không tồn tại giá trị từ hàm $\ln(\cdot)$.

B2.2: tiến hành ghép các siêu điểm ảnh gần nhau có phân bố màu sắc tương đồng. Đối với quá trình này, các siêu điểm ảnh có vị trí liên kề là những siêu điểm ảnh liên kết trực tiếp với nhau trên ảnh, do đó, nếu trên ma trận tương quan màu sắc các vị trí $(i, i+1)$ có chỉ số BC thấp hơn ngưỡng epsilon được định nghĩa từ trước, thì có thể gộp hai siêu điểm ảnh i và $i+1$ thành một siêu điểm ảnh mới.

Sau khi ghép siêu điểm ảnh, các siêu điểm ảnh được cập nhật, thành các siêu điểm ảnh mới, cụ thể:

$$S_{updated}^{(i)} \leftarrow S^{(i)}_{i \in [1, n]}$$

B2.3: tiếp tục thực hiện hai bước B2.1 và B2.2, đến lúc số lượng siêu điểm ảnh trước và sau cập nhật không đổi, kết thúc quá trình ghép siêu điểm ảnh, và được siêu điểm ảnh cuối cùng $\{S(i)|i=1, 2, \dots, n\}$.

Bước 3: thống kê đặc tính màu sắc và không gian của các siêu điểm ảnh; dựa trên các siêu điểm ảnh $\{S(i)|i=1, 2, \dots, n\}$ được ghép sau bước 2, tiến hành thống kê mô tả đặc tính về quy mô của siêu điểm ảnh, khoảng cách đến trung tâm ảnh ban đầu và tính toán độ tương quan về phân bố màu sắc, phân bố không gian, độ tương phản của các siêu điểm ảnh. Cụ thể được thực hiện qua các bước nhỏ sau:

B3.1: thống kê khoảng cách từ tâm siêu điểm ảnh đến tâm ảnh và kê số lượng điểm ảnh trong siêu điểm ảnh. Thông thường, ảnh cần tính toán khu vực nổi bật thường tập trung vào khu vực chính giữa, do đó các siêu điểm ảnh gần trung tâm ảnh hơn sẽ có khả năng là khu vực nổi bật hơn.

Khoảng cách từ siêu điểm ảnh đến điểm trung tâm của ảnh được định nghĩa bởi khoảng cách Euclid của tâm siêu điểm ảnh đến điểm chính giữa của ảnh ban đầu. Công thức tính như sau:

$$d(i) = \left\| \frac{\sum_{S(i)} (x, y)}{\sum_{S(i)} |(x, y)|}, (cx, cy) \right\|$$

Trong đó (cx, cy) là điểm trung tâm của ảnh ban đầu, giá trị lần lượt bằng một nửa của chiều cao và chiều rộng của ảnh, và $\sum_{S(i)} |(x, y)|$ là tổng số điểm ảnh trong siêu điểm ảnh $S(i)$.

B3.2: Tính toán ma trận tương quan màu sắc $color(i, j)$, cụ thể phương pháp tương tự như bước B2.1.

B3.3: Từ tập hợp siêu điểm ảnh $S(i)$, tính chỉ số lấp đầy hình tròn ngoại tiếp của từng siêu điểm ảnh, quá trình được mô tả như trong Hình 2. Cụ thể được thực hiện qua các bước nhỏ như sau:

B3.3.1: gọi điểm trung tâm của siêu điểm ảnh thứ i là c_i , lúc này bán kính đường tròn ngoại tiếp của siêu điểm ảnh này được tính như sau:

$$r(i) = \max_{(x, y) \in S(i)} (\|(x, y), c_i\|)$$

B3.3.2: chỉ số lấp đầy hình tròn ngoại tiếp $compact(i)$ của siêu điểm ảnh được tính bởi công thức:

$$compact(i) = \frac{\sum_{(x, y) \in S(i)} |(x, y)|}{\pi \times r^2(i)}$$

B3.4: từ tập hợp siêu điểm ảnh $S(i)$ tính ở B2.3 và chỉ số lấp đầy hình tròn ngoại tiếp $compact(i)$ tính được ở B3.3.2, tính được ma trận tương quan không gian giữa các siêu điểm ảnh. Công thức như sau:

$$spatial(i, j) = \left(1 - \frac{\|c_i, c_j\|}{\sqrt{cx^2 + cy^2}}\right) \times (1 - compact(i))$$

Trong đó c_i, c_j là điểm trung tâm của siêu điểm ảnh thứ i và j .

B3.5: tính toán ma trận độ tương phản toàn cục giữa các siêu điểm ảnh, trong đó, các siêu điểm ảnh có số lượng điểm ảnh lớn hơn sẽ có ảnh hưởng lớn hơn trong quá trình đánh giá tính nổi bật của siêu điểm ảnh.

$$contrast(i, j) = \underset{i, j \in [1, n]}{dot} \left(|color(i, j)|, \left\{ npixel_i = \sum_{S(i)} |(x, y)| \right\} \right)$$

Trong đó $npixel_i$ là số lượng điểm ảnh trong siêu điểm ảnh i . Đồng thời $contrast(i,j)$ được chuẩn hóa về $[0,1]$ dựa trên khoảng $[min(contrast), max(contrast)]$. Ngoài ra $dot(.)$ mô tả phép nhân hai ma trận.

Bước 4: phân tích kết quả phán đoán cuối cùng về độ phân biệt giữa các siêu điểm ảnh; bước này tiến hành tính toán ma trận tương quan tổng hợp, và véc-tơ đánh giá độ nổi bật của siêu điểm ảnh, và nội suy ra chỉ số độ nổi bật của từng điểm ảnh trong siêu điểm ảnh tương ứng. Bước này được thực hiện theo các bước nhỏ như sau:

B4.1: ma trận tương quan tổng hợp $similar(i,j)$, là sự kết hợp của các yếu tố: ma trận tương quan màu sắc, ma trận tương quan không gian, cụ thể:

$$similar(i, j) = color(i, j) \times spatial(i, j)$$

Ma trận tương quan tổng hợp được coi là ma trận tương quan với căn cứ phán đoán là ma trận tương quan màu sắc dưới tác động của ma trận tương quan không gian như một ma trận hệ số. Các siêu điểm ảnh có kích thước lớn (số lượng điểm ảnh lớn) sẽ có ảnh hưởng nhiều hơn.

B4.2: véc-tơ đánh giá độ nổi bật của siêu điểm ảnh $decision(i)$ được tính toán từ ma trận tương quan điểm ảnh từ bước B4.1, ma trận tương phản toàn cục từ bước B3.5 và kết quả thống kê khoảng cách từ tâm các siêu điểm ảnh đến điểm trung tâm ảnh ban đầu từ bước B3.1. Cụ thể được thực hiện theo các bước nhỏ sau:

B4.2.1: tính toán ma trận không gian mở rộng $spread(i,j)$, sử dụng ma trận tương quan tổng hợp $similar(i,j)$ và khoảng cách $d(i)$. Cụ thể công thức tính như sau:

$$spread(i, j) = \frac{dot(d(i), similar(i, j))}{\sum_{axis=0} similar(i, j)}$$

Sau đó ma trận $spread(i,j)$ được chuẩn hóa về $[0,1]$ dựa trên khoảng $[min(spread), max(spread)]$.

Bước này chủ yếu bổ sung hệ số hội tụ về trung tâm ảnh ban đầu của các siêu điểm ảnh. Các siêu điểm ảnh gần trung tâm hơn sẽ được chú ý hơn, và có tiềm năng là một phần trong đối tượng nổi bật cao hơn.

B4.2.2: véc-tơ đánh giá độ nổi bật $decision(i)$ là một chuỗi các chỉ số đánh giá đối với các siêu điểm ảnh, được tính như sau:

$$decision(i) = 1 - \frac{dot(contrast(i, j), similar(i, j)) \times dot(spread(i, j), similar(i, j))}{\sum_{axis=0} similar(i, j)^2}$$

Véc-tơ $decision(i)$ có chiều dài bằng số lượng siêu điểm ảnh đã khởi tạo, giá trị càng cao càng nổi bật ở trong ảnh.

B4.3: tính toán chỉ số nổi bật của từng điểm ảnh $p(x, y)$ trong ảnh đã chuyển hóa sang không gian màu HSV ở bước B1.3 thông qua véc-tơ đánh giá độ nổi bật cho từng siêu điểm ảnh. Gọi $N(p)$ là 8 điểm ảnh lân cận của $p(x, y)$.

$$S(p) = \frac{\sum_{i \in N(p)} decision(i) \times histogram(i) [bin(p)]}{\sum_{i \in N(p)} histogram(i) [bin(p)]}$$

Trong công thức trên, các điểm ảnh có giá trị màu càng phổ biến trong siêu điểm ảnh, thì chỉ số nổi bật của điểm ảnh đó càng cao. Trong đó $bin(p)$ là giá trị màu (h, s, v) ở không gian HSV của điểm ảnh $p(x, y)$ sau khi được nén lại thành 3 bit, cụ thể:

$$bin(p) = \left[round\left(\frac{h}{8}\right), round\left(\frac{s}{8}\right), round\left(\frac{v}{8}\right) \right]$$

Cuối cùng, từ ảnh RGB ban đầu, đã được chuyển hóa thành dạng ma trận có kích thước tương đương, giá trị của mỗi phần tử trong đó thể hiện độ nổi bật của điểm ảnh tương ứng.

Ví dụ thực hiện sáng chế

Nhóm nghiên cứu đã tiến hành thử nghiệm và đánh giá hiệu quả của phương pháp được đề cập trong sáng chế này với nhiều dạng dữ liệu khác nhau.

Như thể hiện trong Hình 3, 4 và 5, lần lượt là hiệu quả phân tích độ nổi bật ở mức siêu điểm ảnh đối với đối tượng là nải chuối, đồng hồ và người. Trong đó, các Hình 3a, 4a và 5a lần lượt là ảnh ban đầu cần phân tích, Hình 3b, 4b, 5b là kết quả phân tích độ nổi bật (màu càng tối càng nổi bật).

Kết quả cho thấy khả năng phân tích độ nổi bật ở nhiều cấp độ khác nhau, ví dụ như đồng hồ nổi bật hơn so với nền, còn trong đồng hồ thì kim đồng hồ nổi bật hơn các phần còn lại, điều này cho thấy tính tự thích ứng tốt của phương pháp.

Phương pháp phân tích độ nổi bật đối tượng được đề cập trong sáng chế này phù hợp cho nhiều dạng đối tượng khác nhau, tính thông dụng cao và có thể là tiền đề cho các ứng dụng phân tích điểm bất thường trong các sản phẩm công nghiệp, và các ứng dụng

phân tích hành vi trong chuỗi ảnh của video, cũng như ứng dụng bóc tách đối tượng nổi bật ra khỏi nền, giúp ích cho quá trình xử lý dữ liệu tự động.

Lợi ích (hiệu quả) có thể đạt được bởi sáng chế

Phương pháp được đề cập trong sáng chế đã được sử dụng trong các sản phẩm phân tích đặc trưng trực quan của người và phương tiện, triển khai cho các hệ thống giám sát an ninh, tìm kiếm đối tượng của Tổng công ty công nghiệp công nghệ cao Viettel. Trong đó, những đối tượng là người và phương tiện giao thông được bóc tách ra khỏi phần nền ảnh, thực hiện phân tích đặc điểm về màu sắc, bên cạnh các yếu tố nổi bật về ngoại hình, mang mặt, phụ kiện, lô-gô. Từ đó là căn cứ để tìm kiếm đối tượng người và phương tiện trong hệ thống giám sát qua video.

Ngoài ra, phương pháp này còn được sử dụng một cách hữu hiệu trong các công cụ xử lý, phân tích dữ liệu và chuẩn hóa dữ liệu video từ camera. Trong đó, lượng dữ liệu lớn sẽ được tự động phân tích và phân mảnh dựa vào đặc trưng mang mặt, đồng thời có thể thực hiện bóc tách nhiều loại đối tượng khác nhau từ phần nền của ảnh.

Bên cạnh đó, về mặt triển khai thực tế, phương pháp như được đề cập trong sáng chế này không yêu cầu cao đối với phần cứng vận hành, do đó công nghệ này có tính ứng dụng cao trong các sản phẩm công nghiệp.

Yêu cầu bảo hộ

1. Phương pháp xác định đối tượng nổi bật trong ảnh, bao gồm:

bước 1: thực hiện tiền xử lý ảnh đầu vào, sử dụng thuật toán làm mờ Gaussian, chuyển ảnh từ không gian màu đỏ-xanh lá-xanh lam (red-green-blue: RGB) sang tông màu-độ bão hòa-độ sáng (hue-saturation-ilumination: HSV), đồng thời sử dụng thuật toán SLIC (simple linear iterative clustering) để khởi tạo siêu điểm ảnh và thống kê các đặc điểm về phân bố màu sắc và không gian của từng siêu điểm ảnh; cụ thể được thực hiện như sau:

ảnh $I(x, y)$ là ảnh RGB, được tích chập với một nhân gaussian k có kích thước $[3, 3]$, và sigma bằng 0,5, thực hiện thao tác làm mờ ảnh, công thức như sau:

$$I(x, y) \leftarrow I(x, y) \otimes k(i, j)$$

khởi tạo siêu điểm ảnh từ ảnh đã làm mờ $I(x, y)$, sử dụng thuật toán SLIC, với tham số n là số lượng siêu điểm ảnh tối đa có thể được khởi tạo; sau khi khởi tạo, ảnh ban đầu được chia thành nhiều siêu điểm ảnh lớn $S(i)$, trong đó các vùng ảnh có tính chất tương đồng gần nhau sẽ được gom vào cùng một siêu điểm ảnh, và vị trí của siêu điểm ảnh được đánh dấu từ nhỏ đến lớn theo chiều từ trái sang phải, từ trên xuống dưới;

chuyển ảnh ban đầu $I(x, y)$ từ không gian màu RGB sang không gian màu HSV, được ảnh $I'(x, y)$;

bước 2: ghép các siêu điểm ảnh có độ tương quan màu sắc tương đồng; cụ thể bước này được thực hiện như sau:

ảnh HSV $I'(x, y)$ kết hợp tập hợp siêu điểm ảnh $S(i)$ tính ở bước 1, tính toán phân bố màu sắc trên không gian màu HSV đối với từng siêu điểm ảnh; cụ thể công thức tính toán như sau:

$$h(i) = \underset{i \in [1, m]}{\text{histogram}}(I'_{S(i)}(x, y), \text{bins} = [8, 8, 8])$$

trong đó, bins là số nhóm được phân chia từ 256 bậc xám, sau khi tính toán phân bố màu sắc đối với mỗi siêu điểm ảnh, được ma trận $h(i)$ có kích thước $m \times 8 \times 8 \times 8$;

từ ma trận $h(i)$, tính toán ma trận tương quan về màu sắc giữa các siêu điểm ảnh, sử dụng chỉ số bhattacharyya (bhattacharyya coefficient, BC); công thức tính toán như sau:

$$color(i, j) = BC(h_i, h_j) \\ i, j \in [1, m]$$

với BC là chỉ số bhattacharyya, công thức tính BC đối với hai véc-tơ trên cùng không gian như sau:

$$BC(p, q) = -\ln(10^{-1} + \sum \sqrt{p_i q_i})$$

trong công thức trên, 10^{-1} được thêm vào để tránh trường hợp không tồn tại giá trị từ hàm $\ln(\cdot)$;

tiến hành ghép các siêu điểm ảnh gần nhau có phân bố màu sắc tương đồng; đối với quá trình này, các siêu điểm ảnh có vị trí liền kề là những siêu điểm ảnh liên kết trực tiếp với nhau trên ảnh, do đó, nếu trên ma trận tương quan màu sắc các vị trí $(i, i+1)$ có chỉ số BC thấp hơn ngưỡng epsilon được định nghĩa từ trước, thì có thể gộp hai siêu điểm ảnh i và $i+1$ thành một siêu điểm ảnh mới;

sau khi ghép siêu điểm ảnh, các siêu điểm ảnh được cập nhật, thành các siêu điểm ảnh mới, cụ thể:

$$S_{updated}^{i \in [1, n]}(i) \leftarrow S^{i \in [1, n]}(i)$$

tiếp tục thực hiện hai bước 2, đến lúc số lượng siêu điểm ảnh trước và sau cập nhật không đổi, kết thúc quá trình ghép siêu điểm ảnh, và được siêu điểm ảnh cuối cùng $\{S(i) | i=1, 2, \dots, n\}$;

bước 3: thống kê đặc tính màu sắc và không gian của các siêu điểm ảnh; dựa trên các siêu điểm ảnh $\{S(i) | i=1, 2, \dots, n\}$ được ghép sau bước 2, tiến hành thống kê mô tả đặc tính về quy mô của siêu điểm ảnh, khoảng cách đến trung tâm ảnh ban đầu và tính toán độ tương quan về phân bố màu sắc, phân bố không gian, độ tương phản của các siêu điểm ảnh; cụ thể bước này được thực hiện như sau:

thống kê khoảng cách từ tâm siêu điểm ảnh đến tâm ảnh và kê số lượng điểm ảnh trong siêu điểm ảnh; thông thường, ảnh cần tính toán khu vực nổi bật thường tập trung vào khu vực chính giữa, do đó các siêu điểm ảnh gần trung tâm ảnh hơn sẽ có khả năng là khu vực nổi bật hơn;

khoảng cách từ siêu điểm ảnh đến điểm trung tâm của ảnh được định nghĩa bởi khoảng cách euclid của tâm siêu điểm ảnh đến điểm chính giữa của ảnh ban đầu; công thức tính như sau:

$$d(i) = \left\| \frac{\sum_{S(i)} (x, y)}{\sum_{S(i)} |(x, y)|}, (cx, cy) \right\|$$

trong đó (cx, cy) là điểm trung tâm của ảnh ban đầu, giá trị lần lượt bằng một nửa của chiều cao và chiều rộng của ảnh, và $\sum_{S(i)} |(x, y)|$

là tổng số điểm ảnh trong siêu điểm ảnh $S(i)$;

tính toán ma trận tương quan màu sắc $\text{color}(i, j)$, cụ thể phương pháp tương tự như bước 2;

từ tập hợp siêu điểm ảnh $S(i)$, tính chỉ số lấp đầy hình tròn ngoại tiếp của từng siêu điểm ảnh; cụ thể được thực hiện như sau:

gọi điểm trung tâm của siêu điểm ảnh thứ i là c_i , lúc này bán kính đường tròn ngoại tiếp của siêu điểm ảnh này được tính như sau:

$$r(i) = \max_{(x, y) \in S(i)} \|(x, y), c_i\|$$

chỉ số lấp đầy hình tròn ngoại tiếp $\text{compact}(i)$ của siêu điểm ảnh được tính bởi công thức:

$$\text{compact}(i) = \frac{\sum_{(x, y) \in S(i)} |(x, y)|}{\pi \times r^2(i)}$$

từ tập hợp siêu điểm ảnh $S(i)$ tính ở bước 2 và chỉ số lấp đầy hình tròn ngoại tiếp $\text{compact}(i)$, tính được ma trận tương quan không gian giữa các siêu điểm ảnh theo công thức như sau:

$$\text{spatial}(i, j) = \left(1 - \frac{\|c_i, c_j\|}{\sqrt{cx^2 + cy^2}} \right) \times (1 - \text{compact}(i))$$

trong đó c_i, c_j là điểm trung tâm của siêu điểm ảnh thứ i và j ;

tính toán ma trận độ tương phản toàn cục giữa các siêu điểm ảnh, trong đó, các siêu điểm ảnh có số lượng điểm ảnh lớn hơn sẽ có ảnh hưởng lớn hơn trong quá trình đánh giá tính nổi bật của siêu điểm ảnh; công thức tính toán như sau:

$$contrast(i, j) = \underset{i, j \in [1, n]}{dot} \left(\left| color(i, j) \right|, \left\{ npixel_i = \sum_{S(i)} |(x, y)| \right\} \right)$$

trong đó $npixel_i$ là số lượng điểm ảnh trong siêu điểm ảnh I ; đồng thời $contrast(i, j)$ được chuẩn hóa về $[0, 1]$ dựa trên khoảng $[\min(contrast), \max(contrast)]$; ngoài ra $dot(.)$ mô tả phép nhân hai ma trận.

Bước 4: phân tích kết quả phán đoán cuối cùng về độ phân biệt giữa các siêu điểm ảnh; bước này tiến hành tính toán ma trận tương quan tổng hợp, và véc-tơ đánh giá độ nổi bật của siêu điểm ảnh, và nội suy ra chỉ số độ nổi bật của từng điểm ảnh trong siêu điểm ảnh tương ứng; bước này được thực hiện như sau:

ma trận tương quan tổng hợp $similar(i, j)$, là sự kết hợp của các yếu tố: ma trận tương quan màu sắc, ma trận tương quan không gian, cụ thể:

$$similar(i, j) = color(i, j) \times spatial(i, j)$$

ma trận tương quan tổng hợp được coi là ma trận tương quan với căn cứ phán đoán là ma trận tương quan màu sắc dưới tác động của ma trận tương quan không gian như một ma trận hệ số; các siêu điểm ảnh có kích thước lớn (số lượng điểm ảnh lớn) sẽ có ảnh hưởng nhiều hơn;

véc-tơ đánh giá độ nổi bật của siêu điểm ảnh $decision(i)$ được tính toán từ ma trận tương quan điểm ảnh từ bước 4, ma trận tương phản toàn cục và kết quả thống kê khoảng cách từ tâm các siêu điểm ảnh đến điểm trung tâm ảnh ban đầu từ bước 3; cụ thể được thực hiện như sau:

tính toán ma trận không gian mở rộng $spread(i, j)$, sử dụng ma trận tương quan tổng hợp $similar(i, j)$ và khoảng cách $d(i)$. Cụ thể công thức tính như sau:

$$spread(i, j) = \frac{dot(d(i), similar(i, j))}{\sum_{axis=0} similar(i, j)}$$

sau đó ma trận $spread(i, j)$ được chuẩn hóa về $[0, 1]$ dựa trên khoảng $[\min(spread), \max(spread)]$;

bước này chủ yếu bổ sung hệ số hội tụ về trung tâm ảnh ban đầu của các siêu điểm ảnh; các siêu điểm ảnh gần trung tâm hơn sẽ được chú ý hơn, và có tiềm năng là một phần trong đối tượng nổi bật cao hơn;

véc-tơ đánh giá độ nổi bật $decision(i)$ là một chuỗi các chỉ số đánh giá đối với các siêu điểm ảnh, được tính như sau:

$$decision(i) = 1 - \frac{dot(contrast(i, j), similar(i, j)) \times dot(spread(i, j), similar(i, j))}{\sum_{axis=0} similar(i, j)^2}$$

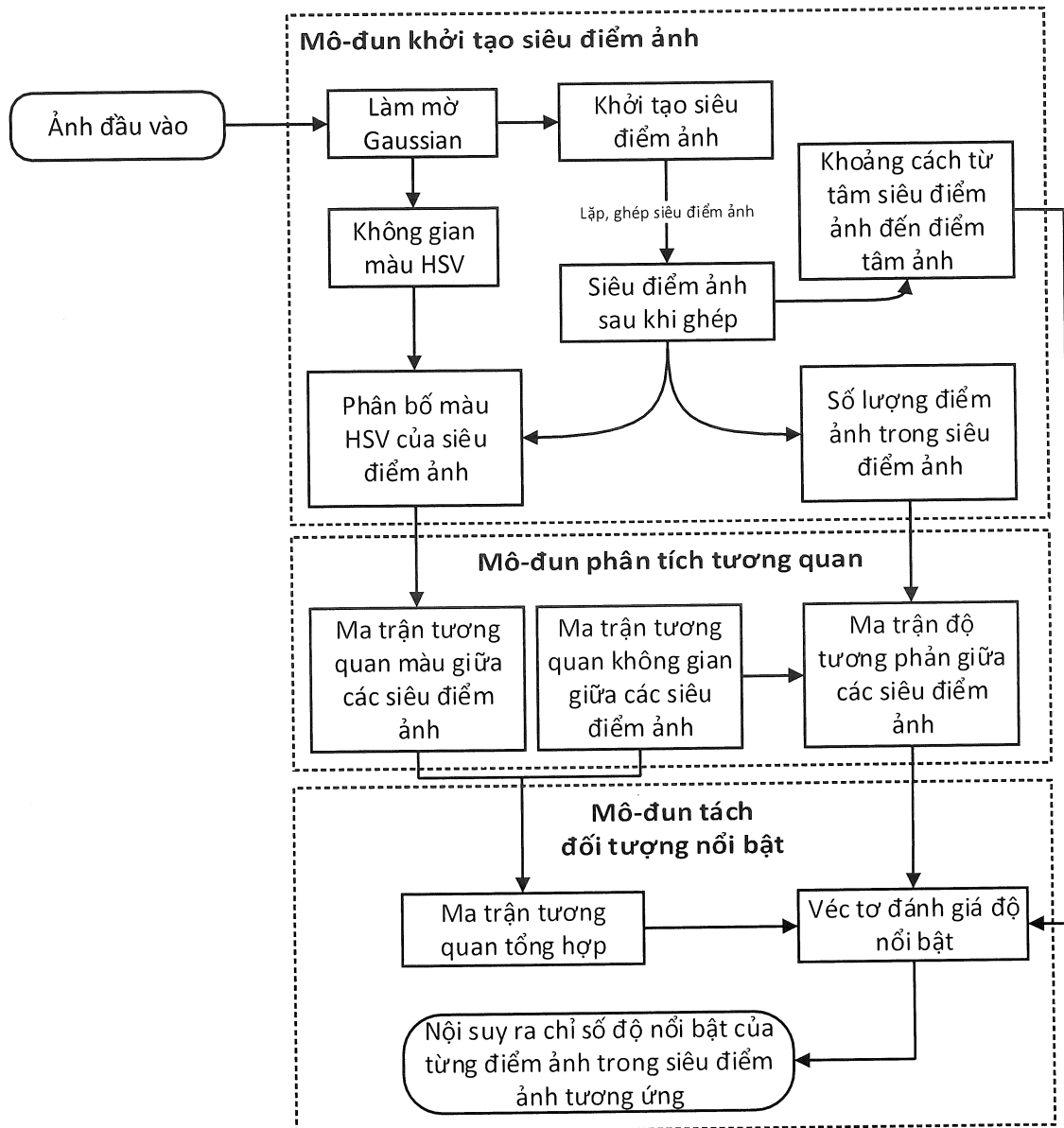
véc-tơ $decision(i)$ có chiều dài bằng số lượng siêu điểm ảnh đã khởi tạo, giá trị càng cao càng nổi bật ở trong ảnh;

tính toán chỉ số nổi bật của từng điểm ảnh $p(x, y)$ trong ảnh đã chuyển hóa sang không gian màu HSV ở bước 1 thông qua véc-tơ đánh giá độ nổi bật cho từng siêu điểm ảnh; gọi $N(p)$ là tám điểm ảnh lân cận của $p(x, y)$:

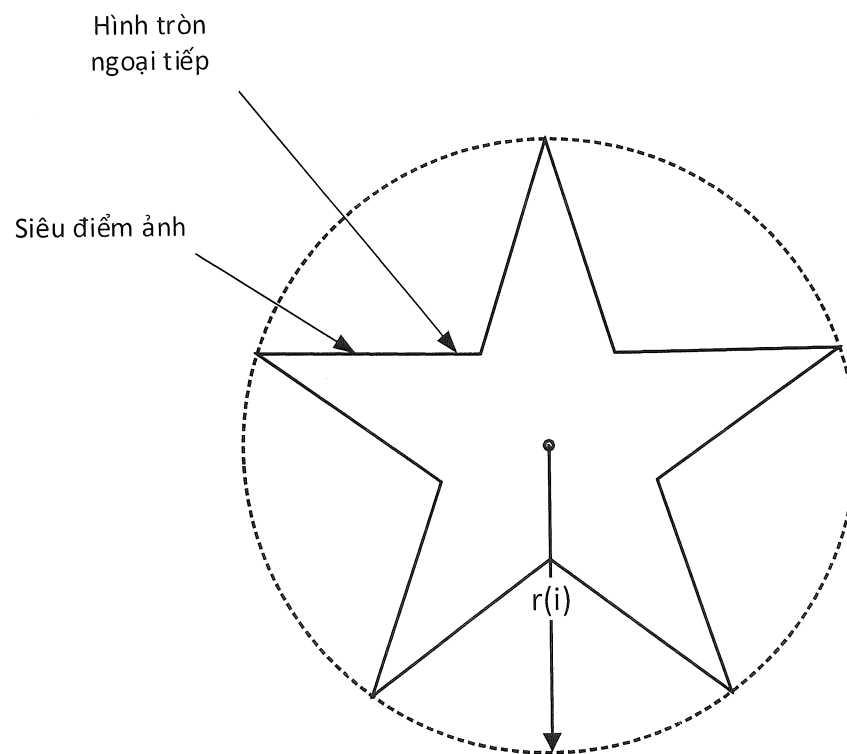
$$S(p) = \frac{\sum_{S(i) \in N(p)} decision(i) \times histogram(i)[bin(p)]}{\sum_{S(i) \in N(p)} histogram(i)[bin(p)]}$$

trong công thức trên, các điểm ảnh có giá trị màu càng phổ biến trong siêu điểm ảnh, thì chỉ số nổi bật của điểm ảnh đó càng cao; trong đó $bin(p)$ là giá trị màu (h, s, v) ở không gian HSV của điểm ảnh $p(x, y)$ sau khi được nén lại thành ba bit, cụ thể:

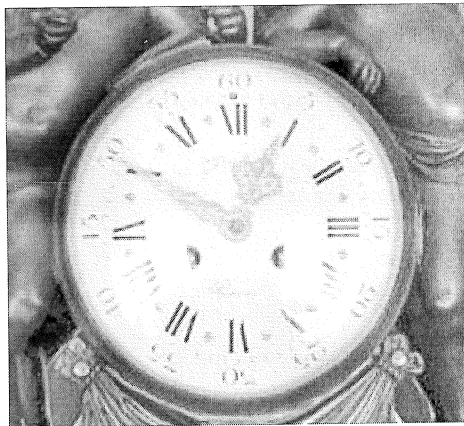
$$bin(p) = \left[round\left(\frac{h}{8}\right), round\left(\frac{s}{8}\right), round\left(\frac{v}{8}\right) \right].$$



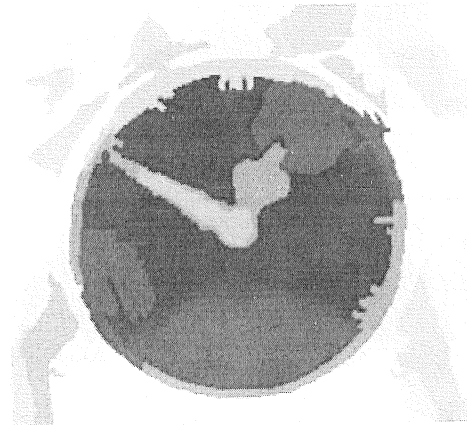
Hình 1



Hình 2

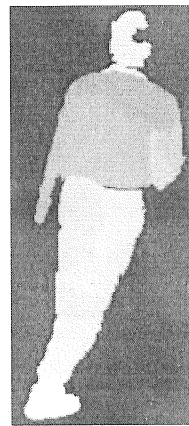


Hình 3a



Hình 3b

Hình 3

**Hình 4a****Hình 4b****Hình 4****Hình 5a****Hình 5b****Hình 5**