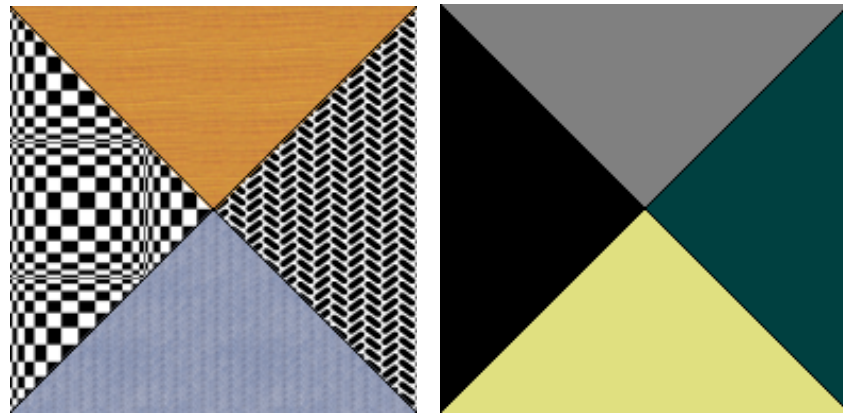


CHƯƠNG 3: CÁC THUẬT GIẢI PHÂN ĐOẠN ẢNH

3.1. Giới thiệu

Phân đoạn ảnh là công đoạn rất quan trọng trong các bài toán xử lý ảnh và thị giác máy tính. Mục đích của phân đoạn ảnh là chia ảnh đầu vào thành các vùng đồng nhất. Trong đó, mỗi vùng đồng nhất là tập các điểm ảnh có chung một số tính chất như màu sắc, kiểu tô, v.v (hình 3.1). Các thuật giải phân đoạn ảnh có thể được xếp vào một trong các loại:



Hình 3.1-Minh họa ảnh gốc và ảnh kết quả mong muốn sau khi phân đoạn

- **Phân ngưỡng (thresholding), gom nhóm (clustering), phân lớp (classifier).**
 - **Thresholding:** chia ngưỡng cường độ và gán điểm ảnh vào từng lớp khác nhau. Cách giải quyết dựa trên histogram của ảnh. Ưu điểm của các phương pháp này là đơn giản, và được dùng cho phân đoạn sơ bộ. Tuy nhiên, phương pháp này lại không thích hợp cho ảnh màu, ảnh xám, nhiễu.
 - **Clustering:** k-Mean, Expectation Maximization (EM): dựa trên cường độ điểm ảnh nhưng xét thêm quan hệ giữa các điểm ảnh lân cận khi phân nhóm, và cập nhật giá trị đại diện của mỗi nhóm.

- **Classifier:** được xem như mở rộng của phương pháp phân ngưỡng với các thuật giải phổ biến k-NN, maximum likelihood, SVM.
- **Loang**
 - **Loang vùng:** từ mỗi điểm ảnh nhân (xác định trước) của mỗi vùng, tìm các điểm ảnh liên thông dựa trên một số tiêu chuẩn được xác định trước và loang dần để tạo thành vùng.
 - **Watershed:** phương pháp heuristic với việc xem ảnh là bề mặt ba chiều và thực hiện loang dần từ đáy lên để hình thành nên các “hồ chứa” thể hiện các vùng.
- **Dựa trên biên:** sử dụng các thuật giải dò cạnh như Canny, Robert, Sobel, v.v. hay các mô hình biến dạng (deformable) Snake, Gradient Vector Flow.
- **Dựa trên phân hoạch đồ thị:** Max Cut, Normalized Cut là hai đại diện phổ biến của phương pháp. Ảnh được xem như đồ thị, và cách giải quyết là xác định tiêu chuẩn phân hoạch đồ thị “tốt nhất”, tạo ra mỗi đồ thị con tương ứng với lớp cần phân đoạn.
- **Dựa trên mô hình xác suất:** sử dụng các mô hình xác suất như trường ngẫu nhiên Markov, Gibbs, Gaussian, và Conditional. Ảnh được xem như trường ngẫu nhiên với hàm phân phối xác định, và mục tiêu là tìm lời giải tối ưu thể hiện trạng thái nhãn của từng điểm ảnh.

Trong [48], các tác giả đã trình bày một thuật giải phân đoạn ảnh dựa trên phân hoạch đồ thị. Phương pháp được gọi là Normalized Cut với độ phức tạp tính toán được xác định là NP-đầy đủ. Đặc điểm chính của thuật giải này là xem ảnh như một đồ thị $G = (V, E)$. Trong đó, V là tất cả điểm ảnh trong ảnh ứng với các nút trong đồ thị, E là tập các cạnh nối giữa mọi cặp nút. Dựa vào tiêu chuẩn Normalized Cut (NCut-độ đo thể hiện trọng số khi tách đồ thị) chia đồ thị thành hai (hoặc K-ứng với số lớp) đồ thị con và đệ quy tiếp tục trên các đồ thị con. Trong thuật giải này, các đồ thị con được tách theo tiêu chuẩn NCut chính là các vùng ảnh cần phân đoạn.

Nhận xét rằng tính đúng đắn của thuật giải phụ thuộc vào ứng dụng, loại ảnh, và thực tế cho thấy rằng không tồn tại phương pháp phân đoạn tốt nhất cho tất cả các loại ảnh.

Phương pháp dựa trên mô hình xác suất, cụ thể là trường ngẫu nhiên Markov là trọng tâm nghiên cứu của chúng tôi trong chương này. Đối với mô hình MRF, chúng tôi khai thác thuật giải SA với các đóng góp về xác định nhiệt độ khởi đầu, và sử dụng hàm giảm nhiệt độ thích nghi, cũng như kết hợp MRF với các mô hình phân rã nhằm thực hiện phân đoạn ảnh đa tỉ lệ. Bên cạnh đó chúng tôi cũng trình bày phương pháp watershed ứng dụng trong phân đoạn ảnh. Phương pháp watershed được trình bày ở đây vì các lý do:

- Đây là ứng dụng thực tiễn - phân tích dữ liệu ảnh wafer bán dẫn với nhu cầu phân đoạn các hạt (grain) - mà chúng tôi đã sử dụng watershed với các cải tiến.
- Chúng tôi minh họa ý tưởng việc sử dụng công cụ hay thuật giải có xét đến yếu tố tốc độ thực hiện trong điều kiện công nghiệp.

3.2. Phân đoạn với watershed cải tiến

3.2.1. Phân đoạn với watershed phân cấp

Thuật giải watershed do Digabel và Lantéjoul [20] giới thiệu và sau đó được Beucher, Vincent [7] bổ sung. Thuật giải này được nghiên cứu và ứng dụng để xác định đối tượng trong ảnh. Biến đổi watershed là phương pháp phân đoạn phổ biến xuất phát từ các phương pháp hình thái học. Biến đổi này đơn giản và có thể mô tả trực quan như sau: Xem ảnh là vùng địa hình 3 chiều, trong đó mỗi điểm ảnh đều có độ cao thể hiện tương ứng giá trị mức xám của nó. Cho mưa rơi trên vùng địa hình (ảnh xám đang xét) thì watershed là đường ngăn cách tạo thành hồ chứa nước (catchment basin) tương đương với đối tượng được phân đoạn trên ảnh. Thuật giải watershed được tóm tắt như sau

Sử dụng phương pháp loang dần từ đáy lên [36], [53], nghĩa là xác định tất cả mức xám có trong ảnh, sau đó chọn điểm ảnh hay nhóm điểm ảnh có giá trị cường độ xám nhỏ nhất, loang dần từ các điểm ảnh này ra để xác định các điểm ảnh khác (thuộc cùng mức xám hay mức xám lớn hơn gần nhất) thuộc về vùng đang xét. Thuật giải thực hiện tương tự cho các mức xám nhỏ nhất kế tiếp chưa được xét. Một điểm ảnh được cho là nằm trên biên nếu nó thuộc về ít nhất hai vùng.

Tiêu chuẩn để xác định một điểm ảnh thuộc về vùng đang xét là quan trọng. Quan hệ giữa hai điểm ảnh có thể xác định dựa trên khoảng cách hình học, hay dựa trên khác biệt về cường độ màu. Việc thực hiện phân đoạn ảnh màu với thuật giải watershed cũng tương tự như ảnh xám khi xét tuần tự từng kênh màu. Tuy nhiên, dữ liệu ảnh với không gian màu RGB cần được chuyển về không gian màu HSV hay $I_1I_2I_3$ [<http://www.couleur.org/index.php?page=transformations>] trước khi phân đoạn.

Thuật giải watershed dựa trên khoảng cách được thực hiện như sau

- B1. Chuyển ảnh I từ sắc xám về đen trắng I_b với ngưỡng I_r cho trước.
Theo đó
$$I_{i,j} = \begin{cases} 1 & I_{i,j} \geq I_r \\ 0 & \text{khác} \end{cases}$$
- B2. Tính giá trị mỗi điểm ảnh trong ảnh là khoảng cách Euclide. Trong đó giá trị mỗi điểm ảnh là khoảng cách Euclide gần nhất đến giá trị điểm ảnh khác zero trong ảnh I_b .
- B3. Sắp xếp các điểm ảnh trong ảnh theo giá trị khoảng cách tăng dần.
- B4. Tất cả điểm ảnh có giá trị khoảng cách lớn nhất D_{\max} được gán nhãn duy nhất L_{\max} và xem là các marker (có thể là một điểm ảnh hay là nhóm các điểm ảnh kề nhau) và được lưu trong ảnh marker.
- B1. Xét các điểm ảnh có khoảng cách D_{next} (giá trị lớn kế sau D_{\max}). Nếu có các điểm ảnh kề L_{\max} thì được xấp nhập vào L_{\max} . Các điểm ảnh

(hay nhóm các điểm ảnh kề nhau) chưa được xét nhập sẽ tạo thành các marker mới với nhãn mới.

B2. $D_{\max} = D_{\text{next}}$.

B3. D_{next} = Giá trị khoảng cách lớn kế tiếp.

B4. Lặp lại các bước B5, B6, B7 cho đến khi D_{\max} bằng D_{\min} .

Tương tự, watershed dựa trên biến đổi mức xám cũng theo các bước trên, nhưng thay vì sử dụng khoảng cách hình học, biến đổi mức xám được sử dụng trên ảnh thông qua ma trận biến thiên mức xám G , với G_{\max} là biến đổi mức xám lớn nhất, G_{\min} là biến đổi mức xám nhỏ nhất.

Qua kết quả thực nghiệm thuật giải watershed dựa trên khoảng cách hình học với tập vi ảnh chụp wafer bán dẫn, chúng tôi có những nhận xét sau:

Nhận xét 1: biến đổi watershed dựa trên ma trận D chứa giá trị khoảng cách của các điểm ảnh trên ảnh chỉ hiệu quả trong trường hợp các đối tượng (ở đây là các grain) tương đối tách rời nhau.

Nhận xét 2: biến đổi watershed dựa trên ma trận G chứa biến đổi mức xám của các điểm ảnh trên ảnh bị ảnh hưởng bởi nhiễu và dễ gây ra tình trạng phân đoạn quá nát (over-segmentation).

Từ các nhận xét trên, chúng tôi đề nghị cả hai yếu tố khoảng cách và biến đổi mức xám sẽ được kết hợp cho phân đoạn watershed nhằm làm rõ các biên grain ở những chỗ mà biến đổi khoảng cách với ma trận D không xác định được.

Gọi D là ma trận (có kích thước bằng kích thước ảnh tính theo đơn vị điểm ảnh) chứa khoảng cách hình học của từng điểm ảnh trong ảnh I đang xét, G là ma trận (có kích thước bằng D) chứa biến thiên mức xám của các điểm ảnh trong ảnh. Khi đó D' là ma trận biến đổi kết hợp khoảng cách và biến thiên mức xám được xác định như sau:

$$D' = D * \exp\left(1 - \frac{G - G_{\min}}{G_{\max} - G_{\min}}\right)$$

Với G_{max} và G_{min} là giá trị lớn nhất và nhỏ nhất của G .

Ma trận biến thiên mức xám G được tính theo đạo hàm bậc nhất giá trị mức xám giữa các điểm ảnh kề nhau duyệt lần lượt từ trái sang phải và từ trên xuống dưới trên ảnh xám I đã qua các bước tiền xử lý (khử nhiễu, lọc morphological). Toán tử gradient rời rạc được xác định như sau:

$$\nabla f(m, n) = f_1(m, n)i_m + f_2(m, n)i_n \quad (3.1)$$

Trong đó (m, n) là vị trí điểm ảnh đang xét, i_m, i_n là vector đơn vị theo hướng ngang và dọc,

$$f_1(m, n) = f(m, n) \otimes h_1(m, n)$$

$$f_2(m, n) = f(m, n) \otimes h_2(m, n)$$

Với \otimes là phép toán chập được xác định bởi:

$$c(m, n) = a(m, n) \otimes b(m, n) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a(j, k)b(m-j, n-k)$$

Và h_1 và h_2 là các nhân tử hàng và cột được xác định như sau:

$$h_1 = h_2^T = (-1 \quad 0 \quad 1)$$

Trong thực tế, phải dùng phần đảo của D' cho watershed. Nghĩa là chọn

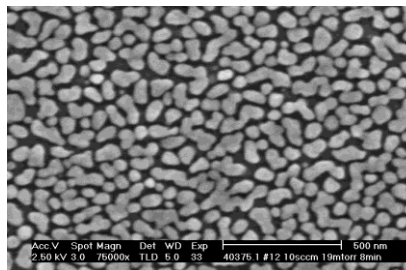
$$D'' = \max(D') - D'$$

làm ma trận biến đổi khoảng cách và biến thiên mức xám của ảnh I cho watershed. Kết quả thực nghiệm cho thấy, phần lớn các biên grain đều được tách chính xác. Tuy nhiên tình trạng phân đoạn quá nát vẫn còn, do đó bước hậu xử lý phân đoạn ảnh là cần thiết. Trong phần 3.2.2, chúng tôi trình bày ứng dụng của thuật giải watershed cùng với việc giải quyết vấn đề phân đoạn quá nát vào một vấn đề công nghệ cụ thể.

3.2.2. Ứng dụng thuật giải watershed trong phân tích vi ảnh grain

Trong phần này, chúng tôi đề xuất một thuật giải hai pha và ứng dụng để tách tự động các grain trong vi ảnh chụp quá trình phủ vật chất nhằm phục vụ các yêu cầu thống kê hay phân tích một số tính chất định lượng trong quá trình sản xuất. Trong pha thứ nhất, biến đổi watershed với hàm phân đoạn sử dụng khoảng cách hình học kết hợp với biến thiên mức xám. Kết quả phân đoạn của pha này có thể còn hiệu ứng phân đoạn quá nát của watershed sẽ được khắc phục trong pha thứ hai. Pha thứ hai sẽ thực hiện vector hóa trên ảnh kết quả của pha thứ nhất và dựa vào tính chất topo của dữ liệu vector để giải quyết tình trạng phân đoạn quá nát của watershed. Bằng cách sử dụng thuật giải phân đoạn hai pha này, các tính toán như chu vi, diện tích, phân bố của grain được chính xác hơn. Cuối phần này, chúng tôi cũng nêu rõ vai trò của chương trình phân tích tính chất của các grain trong vi ảnh nhằm ứng dụng đánh giá chất lượng phủ vật chất trong công nghệ bán dẫn cũng như hỗ trợ xác định được hiệu suất tối ưu của quy trình sản xuất.

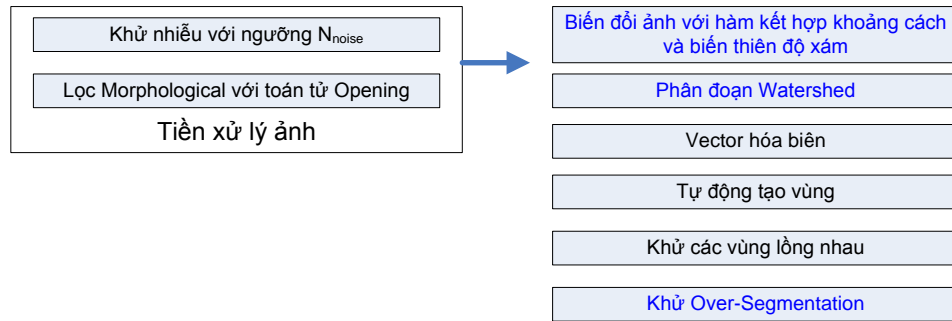
Việc phân đoạn ảnh nhằm trích các grain của vi ảnh chụp lớp phủ phân tử trong quá trình phủ vật chất trong ngành bán dẫn phục vụ thống kê về số lượng, tính chất (diện tích, chu vi, phân bố, v.v.) của chúng (như hình 3.2) là cần thiết cho các kỹ sư đánh giá được phương thức nào là hiệu quả và nhờ đó xác định được hiệu suất tối ưu.



Hình 3.2-Ảnh chụp lớp phủ phân tử (nguồn công ty Vigilant).

Việc trích grain đầy đủ về số lượng và độ chính xác cao là quan trọng vì sẽ ảnh hưởng đến các số liệu cần cho phân tích ở bước sau. Các ảnh dùng trong

thực nghiệm là ảnh Scanning Electron Microscopy (SEM) chụp ngang bề mặt lớp phủ phân tử và được chuyển về định dạng TIFF. Quy trình phân đoạn được thực hiện qua các bước như trong hình 3.3.



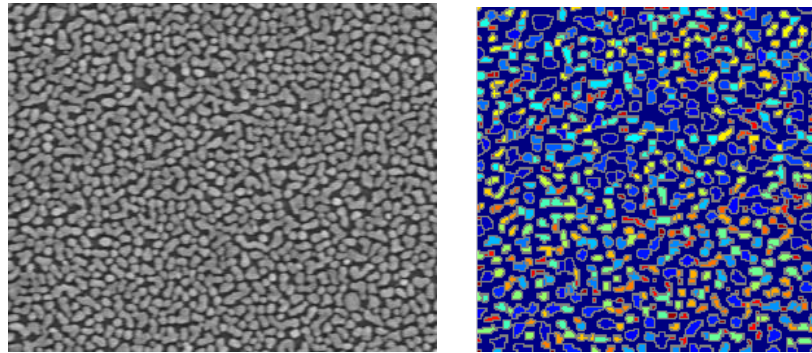
Hình 3.3-Các bước chính trong phân tích ảnh grain

Xem ảnh là ma trận hai chiều I bằng kích thước ảnh (kích thước ảnh trong các thực nghiệm là 712×484), trong đó $I_{i,j}$ là mức xám tại điểm ảnh (i, j) . Chúng tôi sử dụng bộ lọc Median với kích thước nhân tử N_{noise} kết hợp với việc loại bỏ các nhóm điểm ảnh cùng sắc xám nhỏ được xem là nhiễu (có đường kính nhỏ hơn ngưỡng N_{noise}) để làm giảm bớt tình trạng phân đoạn quá nát sau này. Tiếp bộ lọc morphological với toán tử “Opening” được thực hiện nhằm làm nổi rõ các biên grain theo công thức.

$$I_{i,j} \circ K = I_{i,j} \ominus K \oplus K$$

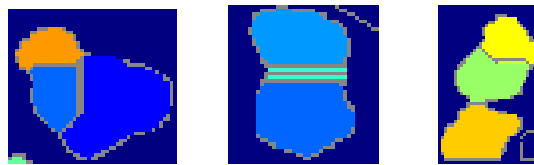
Với \circ là toán tử Opening, \ominus là toán tử erode và \oplus là Dilate, K là ma trận nhân tử.

Hình 3.4 cho thấy một kết quả phân đoạn ảnh SEM tỉ lệ cao khi có rất nhiều grain trong ảnh. Phần lớn các biên grain đều trích đúng, tuy nhiên một số biên mới được tạo ra do tình trạng phân đoạn quá nát.



Hình 3.4-Minh hoạ kết quả watershed với khoảng cách và biến thiên mức xám.

Việc thực hiện watershed như trên có thể dẫn đến tình trạng phân đoạn quá nát. Nghĩa là một grain bị chia thành nhiều grain nhỏ không mong muốn như hình 3.5 dưới đây.



Hình 3.5-Minh hoạ grain bị chia nhỏ do phân đoạn quá nát.

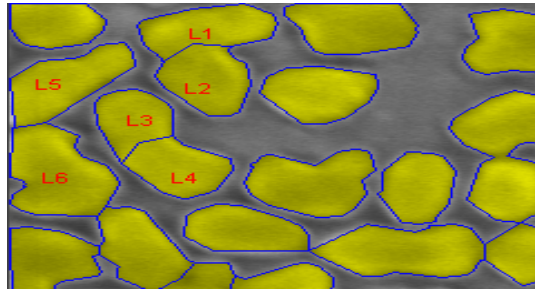
Có nhiều thuật giải đã được đưa ra nhằm giải quyết tình trạng này. Ví dụ như với phương pháp phân đoạn với nhãn đánh dấu có giám sát. Trong đó, các nhãn đánh dấu được xác định trước và áp dụng như là các cực tiểu trong quá trình thực hiện watershed [36]. Mỗi vùng chỉ được phép chứa một nhãn đánh dấu. Tuy nhiên việc xác định các nhãn đánh dấu khá phức tạp và đòi hỏi tri thức trước về loại ảnh đang xử lý. Cách tiếp cận thứ hai gọi là watershed phân cấp, nghĩa là thực hiện watershed trên ảnh ở các cấp phân rã khác nhau. Phương pháp phân rã thường dùng là Laplacian Pyramid. Tại mỗi cấp phân rã, biến đổi watershed được thực hiện trên ảnh đã được giảm kích thước. Các vùng của ảnh phân đoạn tại cấp phân rã mịn hơn được xem xét nhằm kết hợp lại với nhau dựa trên kết quả của các bước phân đoạn ở cấp thô hơn. Phương pháp khác là thêm bước hậu xử lý để tìm các đường biên đúng của đối tượng. Trong thuật giải đã đề nghị, chúng tôi sử dụng cách hậu xử lý dữ liệu vector có chứa thông tin topo thể hiện quan hệ giữa các đối tượng vùng kề nhau nhằm kết hợp lại các vùng bị rã không

hợp lý khi thực hiện watershed. Kỹ thuật hậu xử lý dùng thông tin topo được mô tả trong phần sau

Khử phân đoạn quá nát

Các biên grain sau bước biến đổi watershed sẽ được vector hoá. Quá trình vector hóa được thực hiện sao cho các dữ liệu quan hệ giữa các node và cung được giữ nguyên nhằm làm dữ liệu đầu vào cho bước tạo vùng tự động và khử phân đoạn quá nát sau này.

Bước khử các sub-polygon (đa giác nhỏ nằm trong đa giác lớn) có thể giải quyết bằng cách thực hiện toán tử UNION [23] giữa hai đa giác sau khi kiểm tra đa giác lồng nhau. Việc xác định đa giác nằm trong đa giác khác có thể dễ dàng thực hiện thông qua việc kiểm tra hình chữ nhật bao nhỏ nhất và điểm bất kỳ của đa giác nhỏ nằm trong đa giác lớn hơn. Bước khử phân đoạn quá nát có thể xử lý trên các đa giác (sau khi đã khử sub-polygon) kết hợp với mức xám của các điểm ảnh trong đối tượng nhằm loại bỏ các đường ngăn không hợp lý. Hình 3.6 minh họa một số trường hợp cần khử phân đoạn quá nát.



L_1, L_2 cần kết hợp thành một grain, trong khi L_5, L_6 thuộc về hai grain khác nhau.

Hình 3.6-Minh họa các trường hợp phát sinh khi khử phân đoạn quá nát.

Trước tiên, độ đo biến thiên mức xám được dùng để khử phân đoạn quá nát. Giá trị biến thiên mức xám trung bình của mỗi đối tượng (cung và đa giác) được sử dụng nhằm xác định khả năng kết hợp hai đối tượng kề nhau.

Đặt $B_C = (\sum_{i \in C} B_i) / n$ là biến thiên mức xám của đối tượng C . Trong đó giá trị n là tổng số điểm ảnh trong đối tượng C , và B_i là biến thiên mức xám của điểm

ảnh i tính theo công thức (3.1). Hai đối tượng L_i, L_k và cung chung A_{ik} đang xét có giá trị cường độ trung bình là $B_{L_i}, B_{L_k}, B_{A_{ik}}$. Từ nhận xét, nếu A_{ik} là đường ngăn cách thật sự (nghĩa là L_i, L_k là hai grain khác nhau) thì giá trị của $B_{A_{ik}}$ sẽ lớn hơn B_{L_i} , hay B_{L_k} . *Như vậy điều kiện để kết hợp hai đa giác đang xét dựa trên độ đo biến thiên mức xám là*

$$U_{A_{ik}} = \frac{2.B_{A_{ik}}}{(B_{L_i} + B_{L_k})} \geq U_{merge} \quad (3.2)$$

Với U_{merge} là ngưỡng được xác định thông qua thực nghiệm. Độ đo thứ hai liên quan đến tính chất hình học của các đối tượng vector. Gọi $L = \{\text{tập hợp các đa giác}\}$ của ảnh đang xét. Với mỗi đa giác L_i , ta dễ dàng tính được chu vi P_i , đường kính D_i theo dữ liệu vector. Một số tính chất khác cũng được xét cho quá trình kết hợp đa giác, bao gồm:

Diện tích: số lượng điểm ảnh nằm trong đối tượng đang xét. Nếu là cung thì là các điểm ảnh nằm trên cung. Chúng tôi không tính diện tích theo dữ liệu vector vì các cung chung không phải là đối tượng dạng vùng.

Độ giãn (Elongation): tỉ lệ giữa đường kính phụ và đường kính chính của đối tượng. Trong đó đường kính chính là khoảng cách giữa hai điểm xa nhau nhất trên đối tượng, và đường kính phụ là đường vuông góc với đường kính chính và nối hai điểm gần nhau nhất so với các điểm còn lại.

Gọi X_i là vector đặc trưng đối tượng đa giác thứ i trong tập L : $X_i = (S_i, P_i, E_i)$. Trong đó: S : diện tích, P : chu vi, E : độ giãn.

Với hai đa giác bất kỳ L_i, L_k kề nhau, ta có thể lấy được cung chung gọi là A_{ik} . Vector đặc trưng của cung A_{ik} tương tự như vector đặc trưng của các đa giác trong tập L , nhưng giá trị độ giãn luôn bằng zero.

Gọi \bar{X} là vector trung bình của các vector của cung và tập đa giác mẫu. Gọi C là ma trận hiệp phương sai trên các vector X mẫu.

Độ đo của mỗi đối tượng theo phân bố Gauss là giá trị xác định bởi :

$$H_x = \frac{1}{2\pi\sqrt{|\det(C)|}} \exp\left(-\frac{1}{2}\sqrt{(X - \bar{X})^T C^{-1} (X - \bar{X})}\right)$$

Trong thực tế ta chỉ xét hợp các đa giác có cung chung dài hơn ngưỡng cho trước. Nghĩa là chỉ xét tập $A = \{A_{ik} | P_{ik} \geq T_A\}$ chứa các cung có độ dài hơn ngưỡng T_A xác định trước. Ngưỡng T_A có thể xác định theo $\alpha \bar{T}_A$, với α là tỉ lệ có giá trị từ 0 đến 1, và \bar{T}_A độ dài trung bình của các cung.

Đặt :

$$R_{A_{ik}} = \frac{2 * H_{ik}}{(H_i + S_k)} \quad (3.3)$$

là giá trị đo quan hệ giữa L_i và L_k và cung chung A_{ik} . Hai đa giác L_i và L_k được hợp lại khi $R_{A_{ik}}$ trong (3.3) lớn hơn ngưỡng T_{merge} xác định.

Hai độ đo trong (3.2) và (3.3) có thể liên kết thành một công thức nhằm xác định tiêu chuẩn hợp của hai đa giác L_i, L_k là.

$$T_{A_{ik}} = (U_{A_{ik}} + R_{A_{ik}}) \geq (T_{merge} + U_{merge}) = V_{merge} \quad (3.4)$$

Các ngưỡng T_{merge} , U_{merge} có thể xác định được sau quá trình thực nghiệm trên tập mẫu chứa các đa giác trên các ảnh theo mức phân giải khác nhau. Mức phân giải có thể xác định tùy theo tỉ lệ nanomet trên điểm ảnh, kích cỡ grain, và số lượng grain trên ảnh và thông thường có thể chia làm ba loại: là thấp, trung bình và cao.

Thuật giải khử phân đoạn quá nát

- B1. Sắp xếp các cung trong A theo giá trị độ dài P tăng dần.
- B2. Xét cung chung A_0 có giá trị P nhỏ nhất, gọi là P_0 .
- B3. Tìm hai đa giác L_0 và L_1 có cung chung A_0 . Ta có thể dễ dàng xác định đa giác L_1 kề với L_0 nhờ thông tin topography trong khi vector hoá. Gọi cung chung giữa L_0, L_1 là A_{01} .
- B4. Tính $T_{A_{01}}$ theo công thức (3.4).

- B5. Nếu $T_{A_0} \geq V_{merge}$ thì kết hợp L_0, L_1 . Loại L_0 cũ ra khỏi tập L . Đưa L_0 mới được kết hợp vào tập L . Đồng thời loại bỏ cung A_0 ra khỏi tập A .
- B6. Sắp xếp lại quan hệ topo nếu có quá trình kết hợp xảy ra ở B5.
- B7. Thực hiện B2 với cung chung có giá trị P lớn kế tiếp cho đến khi mọi cung trong A được duyệt hết.

Phân tích đặc trưng ảnh chụp grain

Mỗi grain trên ảnh đã trở thành một đối tượng vùng dạng vector, do đó việc tính toán các tính chất định lượng trở nên dễ dàng. Trong thực tế ta có thể chọn ngẫu nhiên một số khu vực chứa các grain hay toàn bộ grain trên ảnh cho phân tích. Các biểu đồ (Gauss hay histogram) thể hiện các giá trị như diện tích, chu vi, đường kính, phân bố mật độ grain trên các trục X, Y mang lại thông tin hữu ích phục vụ cho quá trình sản xuất wafer.

Từ nhu cầu thực tế cần cho phân tích dữ liệu có trong ảnh grain, chúng tôi đề xuất các đặc trưng hình học của mỗi grain như trong bảng 3.1 nhằm phục vụ các yêu cầu dự đoán và phân loại chất lượng lớp phủ vật chất.

Bảng 3.1-Một số đặc trưng hình học của grain.

Tham số	Ý nghĩa
Diện tích (Area)	Diện tích của đối tượng. Có thể đo theo số điểm ảnh trong đối tượng grain dạng đa giác hay đo theo nano mét nếu tỉ lệ đã được xác định.
Chu vi (perimeter)	Chu vi của đối tượng đa giác tương ứng với grain.
Mức độ tròn (roundness)	Xác định bởi $\frac{4\pi * S}{P^2}$

	Giá trị thay đổi từ 0 đến 1. Giá trị càng lớn thì đối tượng càng tròn. Nếu giá trị là 1, đối tượng gần như hình tròn và ngược lại đối tượng sẽ dẹt khi giá trị tiến về 0.
Chiều dài trục chính	Được xác định bởi khoảng cách giữa hai điểm xa nhau nhất của đối tượng grain.
Góc trục chính	Góc giữa trục ngang và trục chính của đối tượng grain.
Chiều dài trục phụ	Chiều dài của đường dài nhất vuông góc với trục chính nối hai điểm của đối tượng.
Góc trục phụ	Góc giữa trục ngang với trục phụ của đối tượng.
Mức độ giãn (Elongation)	Tỉ lệ giữa độ dài trục phụ và trục chính của đối tượng. Giá trị dao động từ 0 đến 1. Nếu giá trị là 1, đối tượng gần như vuông hay tròn và ngược lại đối tượng sẽ càng dẹt khi giá trị tiến về 0.
Đường kính đường tròn bao.	Đường kính của hình tròn có cùng diện tích với đối tượng. Giá trị được tính là $\sqrt{\frac{4 * S}{\pi}}$.
Độ nén (compact)	$\sqrt{\frac{4 * S / \pi}{D_{major}}}$. Trong đó D_{major} là chiều dài trục chính. Đặc trưng này thể hiện tính tròn của đối tượng grain. Giá trị thay đổi từ 0 đến 1. Đây là tỉ số giữa đường kính đường tròn bao với độ dài trục chính của đối tượng. Nếu giá trị là 1, đối tượng sẽ là tròn, ngược lại đối tượng sẽ ít tròn hơn khi giá trị tiến về 0.

Với các đặc trưng trên kết hợp với các mô hình phân loại đơn giản như k-NN [50] hay phức tạp hơn như mạng nơ-ron [61] hoặc SVM, chúng tôi đã xây dựng được mô hình phân loại dựa trên vi ảnh SEM.

Kết quả thực nghiệm với tham số $N_{noise} = 3$ trên khoảng 200 ảnh chụp ngang bề mặt lớp phủ phân tử (nguồn công ty Vigilant Video). Thực nghiệm được viết trên MATLAB và C++. Giá trị ngưỡng T_A trong thực nghiệm trong khoảng 1-3 điểm ảnh. Nghĩa là hai đa giác có ít hơn 2 hay 3 điểm ảnh kề nhau thì không được xét để kết hợp lại. Kết quả thực nghiệm cho thấy nhiều hơn 90% các grain được xác định chính xác so với ảnh gốc. Quá trình tự động tạo vùng cũng được thực hiện dễ dàng nhờ các quan hệ giữa node và cung sau khi được vector hoá. Quá trình hậu xử lý rất tiện lợi vì tất cả dữ liệu đã được chuyển sang dạng vector có kèm theo thông tin thể hiện quan hệ giữa các cung, đa giác.

Các giá trị đặc trưng trong bảng 3.1 thể hiện tốt nhiệm vụ phân tích và dự báo chất lượng lớp phủ phân tử. Trong thực tế, có thể dùng tất cả 11 đặc trưng hay chỉ kết hợp một số đặc trưng cho phân tích.

Nhận xét rằng phương pháp trích grain kết hợp giữa watershed và khử phân đoạn quá nát trên dữ liệu vector và độ biến thiên cường độ màu đã trình bày trong phần này hiệu quả cho lớp vi ảnh chụp lớp phủ phân tử. Các đặc trưng trong phân loại và phương pháp phân tích đã được kiểm nghiệm trong thực tế và cho kết quả tốt. Tuy nhiên, trong quá trình khảo sát và thực hiện bài toán phân đoạn trên một số lớp vi ảnh với tính chất về dải màu trải đều cũng như các khó khăn đã nêu trong chương 1, chúng tôi nhận thấy việc dựa trên mô hình ngẫu nhiên có thể nâng cao tính tổng quát và chất lượng phân đoạn như sẽ được trình bày chi tiết trong phần 3.3 sau đây.

3.3. Phân đoạn ảnh với mô hình trường ngẫu nhiên Markov

Từ nhận xét là trong một ảnh bất kỳ, trạng thái của mỗi điểm ảnh không là một đối tượng độc lập mà có tương tác với các điểm ảnh khác, nhất là các điểm ảnh lân cận. Như vậy, các mô hình thống kê rất thích hợp để biểu diễn các tương tác này. Thời gian gần đây, các nghiên cứu để giải quyết bài toán phân đoạn tập trung vào mô hình trường ngẫu nhiên Markov (Markov Random Field-MRF) [6], [16]. Trong phần này chúng tôi giải quyết bài toán phân đoạn ảnh dựa trên mô hình MRF.

Trường ngẫu nhiên Markov có thể xem là trường hợp đặc biệt của trường ngẫu nhiên [66], trong đó trạng thái của một biến ngẫu nhiên chỉ phụ thuộc vào trạng thái của các biến ngẫu nhiên “lân cận”. Ưu điểm của mô hình MRF thể hiện ở các tính chất:

- Dựa trên một nền tảng lý thuyết thống kê vững chắc, và có tính hệ thống để giải quyết.
- Có thể đánh giá sự hội tụ tại bất kỳ thời điểm nào.

Trường ngẫu nhiên Markov có thể được biểu diễn bởi xích Markov và đã được chứng minh là luôn tìm được lời giải (còn gọi là trạng thái) tại bất kỳ thời điểm nào [6]. Tương tự tính chất của xích Markov, mô hình MRF cũng được xác định bởi tập các tham số. Cụ thể là ma trận xác suất chuyển giữa các trạng thái, các hàm xác suất, v.v. Nếu gọi X là trường ngẫu nhiên Markov, θ là tham số của mô hình, x_n là trạng thái ở lần chuyển thứ n , thì $P(x_n)$ hoàn toàn được xác định dựa trên $P(x_0)$ và θ , với $P(x_0)$ là xác suất của trạng thái ban đầu:

$$P(x_n | \theta) = P(x_0 | \theta) \cdot \sum_i P(x_i, x_{i-1} | \theta) \quad (3.5)$$

Với xích Markov một chiều, xác suất của lần chuyển thứ n chỉ phụ thuộc vào các lần chuyển trước đó (lân cận gần nhất). Mô hình MRF khi áp dụng cho phân đoạn ảnh hai hoặc ba chiều thì khái niệm lân cận được mở rộng hơn khi xét các hướng khác nhau thay vì chỉ xét các trạng thái liền kề trước.

3.3.1. Các khái niệm cơ bản về trường ngẫu nhiên Markov

Trường ngẫu nhiên

Cho $S = \{X_1, \dots, X_n\}$, với $X_i \in \{0, 1, \dots, G-1\}$, là tập các biến ngẫu nhiên trên không gian mẫu $\Omega = \{0, 1, \dots, G-1\}^n$. Độ đo xác suất p là một trường ngẫu nhiên nếu thoả điều kiện

$$0 < p(w) \leq 1, \forall w \in \Omega \quad (3.6)$$

Có nhiều loại trường ngẫu nhiên, trong số đó có thể kể đến là trường ngẫu nhiên Markov (MRF), Gibbs (GRF), Gaussian (GMRF), hay trường ngẫu nhiên có điều kiện (Conditional Random Field-CRF).

Trường ngẫu nhiên Markov

Trường ngẫu nhiên Markov là trường ngẫu nhiên được xác định trên cấu trúc lân cận cho trước gọi là clique, và thỏa tính chất Markov (lân cận cục bộ) xác định bởi:

$$P(X_i = x_i | X_k = x_k, k \neq i) = P(X_i = x_i | X_k = x_k, k \in N_i) \quad (3.7)$$

Từ đây, chúng tôi sử dụng song song hai thuật ngữ clique và lân cận. Như vậy, xác suất của một biến ngẫu nhiên lấy giá trị xác định chỉ phụ thuộc vào giá trị của các biến ngẫu nhiên lân cận. Trong trường hợp một chiều, MRF được chứng minh là xích Markov [14].

Các mô hình dựa trên MRF khác nhau ở một số tính chất: hàm phân phối xác suất, cấu trúc lân cận, tham số tương tác, hàm thế, v.v. Các mô hình này đều có chung đặc điểm là hàm năng lượng (còn gọi là hàm chi phí, hay cường độ) cho một lời giải của hệ thống trên dữ liệu hai chiều có kích thước $M \times N$ được xác định bởi

$$U(x) = \sum_i x_i G_i(x_i) + \sum_{0 \leq i \leq N-1} \sum_{0 \leq j \leq M-1} \beta_{i,j} x_i x_j \quad (3.8)$$

Trong đó, G_i là hàm thể hiện năng lượng tại vị trí thứ i , phần thứ hai trong phép cộng của (3.8) thể hiện sự tương tác giữa các vị trí, trong đó $\beta_{i,j}$ biểu diễn độ tương tác giữa hai vị trí i và j .

3.3.2. Biểu diễn ảnh dựa trên MRF

Xét ảnh Y kích thước $M \times N$. Ký hiệu $s = (i, j)$ là vị trí điểm ảnh, $S = \{s\}$ là tập tất cả các vị trí. Gọi x là một kết quả phân đoạn trên ảnh Y . Nghĩa là với mỗi vị trí s , tồn tại một giá trị x_s xác định trạng thái của điểm ảnh, thể hiện điểm ảnh tại vị trí s thuộc về vùng nào. Trong phần dưới đây, không mất tính tổng quát, các công thức xác suất không bao gồm tham số mô hình θ .

Gọi:

- X_s là trạng thái của biến ngẫu nhiên X tại vị trí s , và \mathbf{x}_s là giá trị của trạng thái tại vị trí s .
- $N_{s,d}$ ký hiệu các vị trí lân cận (theo nghĩa thông thường của xử lý ảnh) bậc d của s , viết tắt là N_s .

$$N_{s,d} = \{(k,l) : 0 < (k-i)^2 + (l-j)^2 < d; (k,l) \neq (i,j)\}$$

- L là tập hữu hạn các trạng thái (các nhãn) của biến ngẫu nhiên X .
 $L = \{L_1, L_2, \dots, L_G\}$.

Như vậy, giả sử một vùng có nhãn g , quan hệ giữa cường độ tập các điểm ảnh được cho là có nhãn g được thể hiện bởi hàm phân phối xác suất: $P(y_s, s | x_s = g)$. Xét trên toàn bộ ảnh, thì quan hệ này được xác định bởi:

$$\prod_{g \in L} P(y_s, s | x_s = g) \quad (3.9)$$

X được gọi là MRF nếu thỏa hai tính chất sau:

- $(P(X = x) = P(x)) > 0$.
- Tính chất Markov: $P(x_s | x_{S-s}) = P(x_s | x_{N_s})$.

Trong đó, x_{S-s} biểu diễn tất cả các nhãn của ảnh Y tại các vị trí khác s , x_{N_s} là các nhãn của các điểm ảnh thuộc lân cận của vị trí s .

Như vậy, phân đoạn ảnh là tìm lời giải “gán nhãn” tối ưu x^* cho toàn bộ ảnh. Nghĩa là tìm cực đại xác suất hậu nghiệm (Maximum a Posteriori - MAP). MAP được xác định bởi

$$x_{MAP} = \arg \max_x P(x | Y) \quad (3.10)$$

Ta có:

$$P(x | Y) = \frac{P(x) \times (Y | x)}{P(Y)} \text{ (quy tắc Bayes).}$$

Nhận xét rằng, $P(Y)$ thể hiện xác suất xuất hiện của một ảnh, và không tham gia vào công thức tìm MAP. Vì vậy (3.10) được đưa về xác định

$$\arg \max_x P(x).P(Y | x),$$

tương đương với $\arg \max_x \log(P(x).P(Y | x))$

dẫn đến việc xác định

$$\arg \max_x P(x) + P(Y | x) \quad (3.11)$$

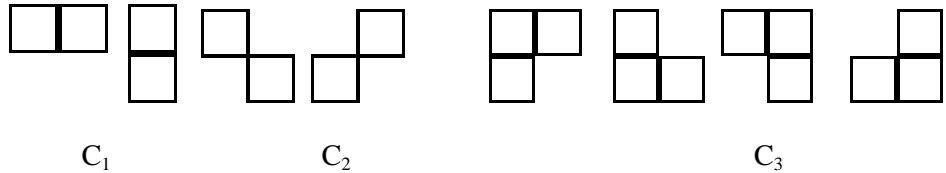
MRF có thể được biểu diễn bởi trường ngẫu nhiên Gibbs (định lý Hammersley-Clifford). Trường ngẫu nhiên X được xem là GRF (trường ngẫu nhiên Gibbs) nếu hàm phân bố xác suất có dạng:

$$P(X = x) = \frac{1}{Z} e^{\left(\frac{-U(x)}{T}\right)} \quad (3.12)$$

Trong đó, Z là hằng xác định trước, gọi là hằng chuẩn hóa, T được gọi là nhiệt độ, U là hàm năng lượng hay còn gọi là hàm Halmiton của phân bố Gibbs và được xác định qua một hàm thế năng theo công thức:

$$U(x) = \sum_{c \in C} V_c(x), \quad Z = \sum_x e^{\left(\frac{-U(x)}{T}\right)} \quad (3.13)$$

Trong đó c là một cụm (clique), và C là tập tất cả các clique. Hàm thế V_c chỉ phụ thuộc vào giá trị của các biến ngẫu nhiên tại các vị trí lân cận đặc biệt của x , gọi là clique c . Tập con $c \in S$ được gọi là clique nếu hai vị trí bất kỳ thuộc c đều là lân cận của nhau. Một số các clique có dạng sau (hình 3.7) thường được dùng trong phân đoạn ảnh.



Hình 3.7-Các cấu trúc clique dùng trong phân đoạn ảnh

Hàm V_c được dùng để thể hiện sự tương tác giữa các biến ngẫu nhiên. Chẳng hạn, dạng đơn giản nhất trên C_I được xác định bởi:

$$V_{(s,r) \in C_I} = \begin{cases} -\beta & s = r \\ +\beta & s \neq r \end{cases},$$

trong đó $\beta \geq 0$ là tham số thể hiện cường độ tương tác giữa các biến ngẫu nhiên trong mô hình.

Một số dạng hàm thể dựa trên sự khác biệt giữa các trạng thái của các biến ngẫu nhiên có thể dùng cho phân đoạn ảnh.

Đặt $\Delta = |x_r - x_s|$, thể hiện sự khác biệt giữa trạng thái của hai biến ngẫu nhiên tại các vị trí r và s , một số dạng các hàm thể sau có thể dùng trong (3.13).

- $V_c(\Delta) = |\Delta|$ [45].
- $V_c(\Delta) = \log \cosh(\Delta)$ [62], với $\cosh(x) = \frac{e^x + e^{-x}}{2} = \cos ix$.
- $V_c(\Delta) = \min(|\Delta|^2, 2|\Delta| - 1)$ (hàm Huber) [68].
- $V_c(\Delta) = \log(1 + |\Delta|^2)$ [77].

3.4. Thuật giải nung luyện mô phỏng thích nghi trong phân đoạn ảnh

Khi sử dụng MRF với hàm phân bố xác suất Gibbs để mô hình hoá ảnh F , thì $P(x)$ được xác định hoàn toàn theo công thức (3.12). Vì vậy, bài toán tìm lời giải tối ưu trở thành xác định cực tiểu hàm năng lượng $U(x)$, nghĩa là tìm:

$$\arg \min_x U(x) \quad (3.14)$$

Trong đó, năng lượng $U(x)$ là tổng năng lượng của các điểm ảnh. Năng lượng trạng thái của điểm ảnh được tính bao gồm năng lượng của chính điểm ảnh đó và sự

tương tác với các điểm ảnh lân cận theo một clique cho trước. Năng lượng của điểm ảnh s được xác định bởi:

$$U_s(X_s = x_s, F) = \ln(\sigma_{x_s} \sqrt{2\pi}) + \frac{f_s - \mu_{x_s}}{2 \times \sigma_{x_s}^2} + \sum_{r \in C_2(s)} V_c(x_s, x_r) \quad (3.15)$$

trong đó

f_s là cường độ của điểm ảnh tại vị trí s .

μ_{x_s} giá trị trung bình của các pixel được cho là thuộc lớp xác định.

σ_{x_s} phương sai của các điểm ảnh được cho là thuộc lớp xác định.

Và hàm năng lượng $U(x)$ của một lời giải được xác định bởi:

$$U(x) = \sum_s U_s(X_s = x_s | F) \quad (3.16)$$

Có nhiều cách để giải bài toán tối ưu theo (3.14), bao gồm các phương pháp cục bộ như cơ chế điều kiện lặp (Iterated Conditional Modes – ICM) [4], [55] hay các phương pháp toàn cục GA, SA. Các thuật giải tìm tối ưu cục bộ có ưu điểm về tốc độ nhưng lời giải sau cùng là chỉ mang tính cục bộ. Thuật giải GA cho phép tìm lời giải tối ưu toàn cục, tuy nhiên tính dừng của nó khó kiểm soát được. Một số phương pháp ngẫu nhiên như Metropolis [70], Gibbs [72] cũng được nhiều nhà nghiên cứu quan tâm để tìm lời giải tối ưu. Một lớp thuật giải tìm kiếm khác được xem như định hướng quan trọng trong tìm lời giải tối ưu là thuật giải nung luyện mô phỏng (Simulated Annealing-SA). Một ưu thế nổi bật của loại thuật giải này là tìm được quy luật toán học để điều khiển quá trình hội tụ của bài toán. Chúng tôi tiếp cận theo hướng xác suất với hạt nhân là thuật giải SA. Hơn nữa, chúng tôi sử dụng SA để tạo quá trình Markov cho mô hình MRF áp dụng trên ảnh. Trong phần sau, chúng tôi trình bày thuật giải SA với hai đóng góp chính về mặt lý thuyết liên quan đến tham số nhiệt độ để giải bài toán phân đoạn ảnh.

3.4.1. Giới thiệu thuật giải nung luyện mô phỏng

Thuật giải nung luyện mô phỏng được giới thiệu lần đầu trong [71], và được trình bày chi tiết trong [57], cũng như ứng dụng tìm lời giải tối ưu trong [43], [44]. Ý tưởng chính xuất phát từ phương pháp nung chảy kim loại làm cho các phân tử rơi vào trạng thái hỗn độn sau đó làm nguội dần cho đến khi các phân tử đạt đến trạng thái năng lượng nhỏ nhất. Về mặt toán học, thuật giải SA là thiết kế quá trình Markov trên không gian trạng thái với hàm nhiệt độ để quá trình Markov hội tụ đến cực tiểu toàn cục. Quy trình SA bắt đầu đặt toàn bộ trạng thái ở “nhiệt độ cao” gọi là trạng thái nung chảy. Sau đó “nhiệt độ” được giảm dần cho đến khi đạt đến trạng thái ổn định (còn gọi là trạng thái đông cứng), và không còn sự thay đổi đáng kể nữa. SA được xem như một trong vài giải thuật heuristic hiệu quả nhất (cùng với thuật giải genetic) để tìm lời giải tối ưu. Ở “nhiệt độ cao”, sự thay đổi giữa các trạng thái rất nhiều, trong khi ở “nhiệt độ” thấp thì sự thay đổi trạng thái ít diễn ra. Trong khi GA mô phỏng quá trình tiến hoá tự nhiên, thì SA mô phỏng quá trình nung luyện. Điều đáng chú ý là cả hai đều mô phỏng quá trình tự nhiên và đạt được thành công đáng kể.

Thuật giải 3.1 tóm tắt SA tổng quát. Nhận xét rằng, nhiệt độ là tham số quan trọng điều khiển quá trình hội tụ cũng như ảnh hưởng trực tiếp đến chất lượng lời giải. Vấn đề đặt ra cho các thuật giải SA là dung hoà giữa hai mục tiêu:

- Mong muốn hội tụ nhanh đến lời giải tối ưu bằng cách thay đổi nhiệt độ sao cho các độ đo Gibbs (năng lượng của lời giải) tập trung quanh giá trị cực tiểu U mong muốn.
- Tuy nhiên nếu giảm nhanh quá nhiệt độ sẽ ảnh hưởng đến kết quả của thuật giải.

Nhận xét là khả năng hội tụ của thuật giải SA phụ thuộc rất nhiều vào nhiệt độ khởi đầu T_0 và quá trình hạ nhiệt. Thông thường, nhiệt độ ban đầu T_0 được xác định thủ công (dựa trên kinh nghiệm) và quá trình hạ nhiệt theo một hàm

xác định trước (hàm giảm theo bước lặp). Chúng tôi nghiên cứu SA và đề xuất các giải pháp:

- Xác định T_0 từ ảnh đầu vào.
- Một hàm giảm thích nghi cho quá trình cập nhật nhiệt độ.

Thuật giải 3.1-Phân đoạn ảnh dựa trên SA

- B1. Tính nhiệt độ khởi đầu T_0 . Gán $T=T_0$.
- B2. Chọn ngẫu nhiên một lời giải làm lời giải ban đầu, gọi là X_0 (chứa tất cả các nhãn gán ngẫu nhiên cho từng điểm ảnh). Tính chi phí (còn gọi là năng lượng) $U(X_0)$.
- B3. Thay đổi ngẫu nhiên trạng thái cho mọi điểm ảnh để xác định lời giải mới
- B4. Tính chi phí cho lời giải X , gọi là $U(X)$. Tính $\Delta U = U(X) - U(X_0)$
- B5. Nếu $\Delta U < 0$, thì xem X là lời giải được chọn với xác suất bằng 1. Ngược lại, chọn X nhưng với xác suất $\frac{\Delta U}{T}$
- B6. Giảm T với theo phương pháp xác định trước. Cách giảm T phổ biến nhất là $T = T * rate$, với giá trị $rate$ từ 0.95 đến 0.99.
- B7. Lặp lại B3 đến B6 cho đến khi sự thay đổi ít hơn ngưỡng cho trước hay giá trị T đủ nhỏ.

3.4.2. Xác định giá trị nhiệt độ khởi đầu

Trong [24], giá trị ban đầu của tham số điều khiển nhiệt độ T_0 được chọn theo heuristic: “ T_0 được chọn sao cho tại giá trị này, tất cả các bước chuyển từ kết quả phân đoạn này sang phân đoạn khác đều được chấp nhận”. Chúng tôi sử dụng các phương pháp lấy mẫu Gibbs và Metropolis tổng quát để xác định lời giải sau mỗi bước chuyển. Sau đó cần xác định giá trị cụ thể cho T_0 thỏa mãn heuristic này.

Giả sử dãy các phép thử (kết quả phân đoạn ngẫu nhiên) được sinh ra tại giá trị T , và tỉ số chấp nhận $\chi(T)$ là:

$$\frac{\text{Số các bước chuyển được chấp nhận}}{\text{Số các bước chuyển phát sinh ngẫu nhiên}}$$

Đặt:

m_1 : số các bước chuyển được đề nghị $x \rightarrow y$ sao cho $U(y) \leq U(x)$ (chấp nhận).

m_2 : số các bước chuyển được đề nghị $x \rightarrow y$ sao cho $U(y) > U(x)$ (phát sinh ngẫu nhiên).

$\bar{\Delta}_U$: trung bình của hiệu chi phí trên m_2 các bước chuyển có chi phí tăng. Khi đó tỉ số chấp nhận có thể được xấp xỉ bởi biểu thức sau:

$$\chi(T) = \frac{m_1 + m_2 \cdot \exp(-T \cdot \bar{\Delta}_U)}{m_1 + m_2} \quad (3.17)$$

$$\text{Suy ra: } T = \frac{\ln\left(\frac{m_2}{m_2 \chi(T) - m_1(1 - \chi(T))}\right)}{\bar{\Delta}_U} \quad (3.18)$$

Giá trị T_0 có thể xác định được theo công thức (3.18).

Bây giờ áp dụng vào bài toán phân đoạn ảnh. Với ảnh cần phân đoạn bất kỳ, giá trị T_0 được xác định thông qua thuật giải 3.2.

Thuật giải 3.2-Xác định nhiệt độ khởi đầu cho thuật giải SA

- B1. Phát sinh ngẫu nhiên M lời giải phân đoạn từ ảnh ban đầu.
- B2. Chọn một lời giải làm lời giải ban đầu, gọi là X_0 . Tính chi phí (là năng lượng) $U(X_0)$.
- B3. Lặp với $M-1$ lời giải còn lại. Với mỗi lời giải X_i ,
Tính $U(X_i)$.

So sánh $U(X_i)$ và $U(X_0)$. Tăng m_1 nếu $U(X_i) \leq U(X_0)$. Ngược lại tăng m_2 .

B4. Tính $\bar{\Delta}_U$ trên các $U(X_i) > U(X_0)$.

B5. Tính $\chi(T) = \frac{m_1}{m_1 + m_2}$

B6. Tính T_0 theo công thức (3.18)

3.4.3. Hàm giảm nhiệt độ thích nghi

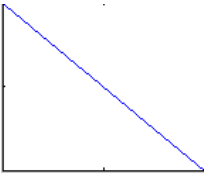
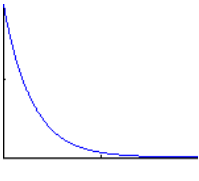
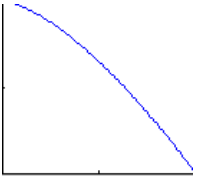
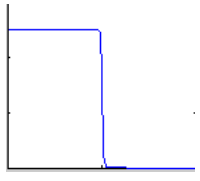
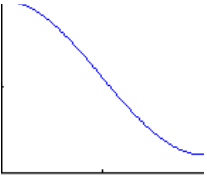
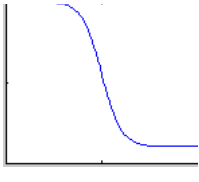
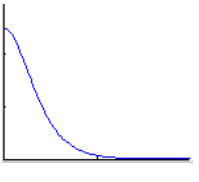
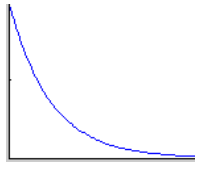
Trong SA, nhiệt độ T được giảm dần sau mỗi bước lặp theo một hàm (bảng 3.2 giới thiệu một số hàm giảm nhiệt độ phổ biến) phụ thuộc 3 tham số:

T_0 Nhiệt độ ban đầu của thuật giải SA.

T_N Nhiệt độ lạnh cuối cùng của thuật giải SA. Giá trị xác định bằng thực nghiệm

N Số vòng lặp dự kiến của thuật giải SA.

Bảng 3.2-Các hàm giảm nhiệt độ cho thuật giải SA

$T_i = T_0 - i \cdot \frac{T_0 - T_N}{N}$ 	$T_i = \left(\frac{T_N}{T_0} \right)^{\frac{i}{N}}$ 	$A = \frac{\ln(T_0 - T_N)}{\ln(N)}$ $T_i = T_0 - i^A$ 	$T_i = \frac{T_0 - T_N}{1 + e^{3(i-N/2)}} + T_N$ 
$T_i = 0.5 * (T_0 - T_N) * (1 + \cos(i * \pi)) + T_N$ 	$T_i = 0.5 * (T_0 - T_N) * (1 - \tanh(10 * i / N - 5)) + T_N$ 	$T_i = \frac{(T_0 - T_N)}{\cosh(10 * i / N)} + T_N$ 	$A = \frac{1}{N} \ln \left(\frac{T_0}{T_N} \right)$ $T_i = T_0 e^{-A \cdot i}$ 

Lưu ý là với các hàm trong bảng 3.2, nhiệt độ được giảm một cách đều đặn qua từng bước lặp bất chấp trạng thái dữ liệu hiện hành. Chúng tôi đã nghiên cứu và đề nghị một hàm giảm thích nghi theo lời giải của mỗi bước như sau:

Ký hiệu S là nhiệt độ nghịch đảo, $S = 1/T$ (không mất tính tổng quát khi bỏ đi chỉ số của lần lặp thứ i). Khi đó

$$S_{i+1} = S_i + \frac{\ln(1+\delta)}{3\sigma_i}, \text{ và } T_{i+1} = \frac{1}{S_{i+1}} \quad (3.19)$$

Với $\delta > 0$ là tham số điều khiển, σ_i là độ lệch chuẩn của các chi phí lời giải tại “nhiệt độ” T_i . Ta cũng có thể chọn σ_i là độ lệch chuẩn của các chi phí lời giải tại các nhiệt độ từ T_0 đến T_i thay vì chỉ tại “nhiệt độ” T_i .

Tiêu chuẩn dừng

Thuật giải SA dừng khi một trong các điều kiện sau xảy ra:

- Giá trị T nhỏ hơn ngưỡng cho trước (lân cận zero).
- Số lần “nung luyện” đủ nhiều. Nghĩa là số lần lặp (B3-B6 trong thuật giải 3.1) nhiều hơn giá trị K cho trước.
- $\frac{1}{T_i \cdot U(x_{i+1})} * (U(x_{i+1}) - U(x_i)) < \varepsilon$, với ngưỡng ε cho trước, $U(x_i)$ là chi phí của lời giải, T_i là “nhiệt độ” ở bước lặp thứ i . Hay có thể ngừng với điều kiện đơn giản hơn: $(U(x_{i+1}) - U(x_i)) < \varepsilon$.

Nhằm cải tiến chất lượng cũng như tốc độ thực hiện của thuật giải, chúng tôi sử dụng MRF phân cấp theo các mô hình đa phân giải đã trình bày. Khi đó, giá trị β trong hàm thế tuyến tính cũng được thay đổi tùy theo cấp cần phân đoạn, và được xác định bởi

$$\beta_k = \beta_{k-1} + \frac{1}{1 + \ln(L-k)}, 0 < k < L \quad (3.20)$$

Trong đó, β_0 là tham số điều khiển tương tác tại lớp phân rã thô nhất β_k là tham số điều khiển tương tác tại cấp phân rã thứ k , L là số cấp cần phân rã (chúng tôi sử dụng $L=6$ cho ảnh 512x512).

Việc tìm nhiệt độ khởi đầu và lượng giảm nhiệt độ thích nghi cho thuật giải SA đóng góp nhiều vào bài toán phân đoạn ảnh. Cùng với các cấu trúc clique và dạng của hàm thế, trong phần này chúng tôi đã trình bày chi tiết cách thức tính toán hai thành phần quan trọng của thuật giải. Tuy nhiên, bản chất của SA là dựa trên thủ tục tính MAP, nên việc tính toán khá nhiều. Vì vậy, chúng tôi xây dựng MRF phân cấp nhằm cải tiến tốc độ cũng như chất lượng phân đoạn ảnh. Ở mỗi cấp, giá trị β (nếu sử dụng hàm thế tuyến tính) cũng được thay đổi theo quy luật tương tự như lượng giảm nhiệt độ nhằm làm cho chất lượng phân đoạn tốt hơn so với sử dụng duy nhất một giá trị β cho mọi cấp. Phần kế tiếp chúng tôi trình bày mô hình MRF đa phân giải dựa trên các phân rã khác nhau.

3.5. Phân đoạn ảnh dựa trên MRF đa phân giải

Mô hình MRF như trình bày ở mục 3.3.2 biểu diễn các tương tác ngang cấp giữa các điểm ảnh. Nhận xét rằng, chi phí tính toán tương tác giữa các điểm ảnh trên ảnh gốc quá cao ảnh hưởng đến tốc độ hội tụ. Để khắc phục, ảnh sẽ được phân cấp nhằm giảm kích thước khi thực hiện tiến trình phân đoạn dùng MRF. Hơn nữa, việc sử dụng MRF phân cấp cũng giúp cho việc cập nhật lại tham số mô hình cho các lần phân đoạn ở cấp mịn hơn và có thể làm tăng chất lượng phân đoạn.

Nguyên lý của mô hình này như sau: kết quả phân đoạn với MRF ở cấp thô được xem như đầu vào hay là ràng buộc cho phân đoạn ở cấp mịn kế tiếp. Thuật giải tiếp tục cho đến khi ảnh được phân đoạn với kích thước ban đầu (cấp 1 – cấp mịn nhất). Trong trường hợp đơn giản, các bộ tham số của MRF giữa các tỉ lệ trong quá trình phân đoạn không được xét đến. Trong các phần từ 3.5.1 đến 3.5.4, chúng tôi sẽ trình bày các mô hình MRF phân cấp dựa trên phân rã LP, WT, UWT, và AWP.

3.5.1. Phân đoạn ảnh dựa trên kết hợp của MRF và phân rã LP

Một mô hình cho MRF phân cấp là thực hiện phân rã ảnh với Laplacian Pyramid [63]-chúng tôi gọi là LPMRF-trong đó phân rã ảnh LP dựa trên phép toán chập với một trong các hàm biến đổi như: Gaussian, Fourier hay Cosine rời rạc, sau đó thực hiện MRF trên các cấp. Thuật giải 3.3 trình bày tóm tắt các bước thực hiện LPMRF. Trong đó, MRF thực hiện trực tiếp trên dữ liệu ảnh ở các cấp khác nhau, bắt đầu từ cấp thô nhất. Ở mỗi cấp phân rã với LP, kết quả sẽ là ảnh lowpass với kích thước giảm một nửa so với ảnh ở cấp mịn hơn và ảnh còn lại lưu trữ sự khác biệt giữa ảnh phân rã và ảnh dự đoán, được gọi là ảnh bandpass.

Đặt g_l là dữ liệu tại cấp phân rã thứ l , w là ma trận nhân tử hai chiều, \hat{w} là ma trận nhân tử một chiều. Các giá trị trong w có thể xác định dựa trên \hat{w} bởi:

$$w(m, n) = \hat{w}(m)\hat{w}(n); \sum_{m=-2}^2 \hat{w}(m) = 1; \hat{w}(t) = \hat{w}(-t) \quad (3.21)$$

$$g_l(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) g_{l-1}(2i + m, 2j + n)$$

Thuật giải 3.3 mô tả các bước chính thực hiện phân đoạn ảnh dựa trên LPMRF. Bước ánh xạ kết quả phân đoạn từ cấp thô lên cấp mịn để có kết quả phân đoạn khởi đầu được thực hiện theo nguyên tắc phân rã của LP.

Thuật giải 3.3-Phân đoạn ảnh dựa trên LPMRF

- B1. Phân rã ảnh với LP rời rạc L cấp.
- B2. Thực hiện phân đoạn với thuật giải SA đã trình bày trong thuật giải 3.1 trên dữ liệu *lowpass* ở cấp thô nhất. Gán kết quả phân đoạn cho ảnh ở cấp thô.
- B3. Lời giải ban đầu của cấp mịn hơn chính là quá trình tự ánh xạ giữa từng điểm ảnh đã phân đoạn ở cấp thô đến các điểm ảnh tương ứng.
- B4. Cập nhật lại tập tham số mô hình θ cho cấp mịn hơn.

B5. Xác định lại giá trị “nhiệt độ khởi đầu” khi sử dụng phân đoạn dựa trên SA theo thuật giải 3.2.

B6. Thực hiện bước B2 cho đến khi phân đoạn xong cấp mịn nhất.

Chúng tôi sử dụng nhân tử Gaussian cho phân rã LP. Nhân tử Gaussian rời rạc có thể được xác định bởi ma trận

$$\frac{1}{273} \begin{pmatrix} 1 & 4 & 7 & 4 & 1 \\ 4 & 16 & 26 & 16 & 4 \\ 7 & 26 & 41 & 26 & 7 \\ 4 & 16 & 26 & 16 & 4 \\ 1 & 4 & 7 & 4 & 1 \end{pmatrix}$$

ứng với giá trị $\sigma = 1.0$, và trung bình $(0,0)$ của hàm Gaussian

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Phân rã Gaussian LP trên dữ liệu hai chiều có thể thực

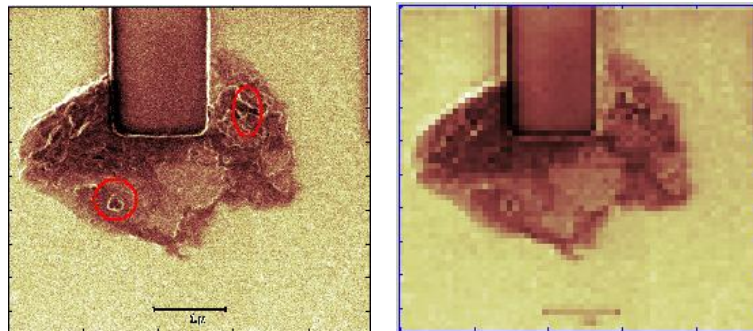
hiện tuần tự với giá trị x trên từng hàng và giá trị y trên từng cột, với ma trận

nhân xác định bởi $\begin{pmatrix} \frac{1}{4} & \left(\frac{1}{4} - \frac{a}{2}\right) & a & \left(\frac{1}{4} - \frac{a}{2}\right) & \frac{1}{4} \end{pmatrix}$, trong đó giá trị a thích hợp

thuộc khoảng từ 0.3 đến 0.6.

3.5.2. Phân đoạn ảnh dựa trên kết hợp giữa MRF và WT

Một số nhu cầu của phân đoạn vi ảnh đòi hỏi phân biệt được đối tượng lớn có nghĩa và bỏ qua các đối tượng nhỏ (hay nhiễu) nằm trong vùng lớn hơn. Hình 3.8 minh họa ảnh chụp lỗi có trong vi ảnh, trong đó các “đối tượng nhỏ” được khoanh tròn không cần được phân tích và cần bị loại bỏ trước khi thực hiện bước phân đoạn.



Hình 3.8-Minh họa nhu cầu phân đoạn ảnh đa tỉ lệ.

Nếu sử dụng MRF đơn cấp, dẫn đến các đối tượng nhỏ không có nghĩa cũng sẽ được xem như các đối tượng lớn, và như vậy cũng là một phần trong kết quả phân đoạn. Các đối tượng không có nghĩa này có thể làm ảnh hưởng đến quá trình phân tích của các công đoạn sau. Ngoài việc khử hiệu quả các tín hiệu nhiễu (bằng cách loại bỏ các hệ số giá trị nhỏ theo tiêu chuẩn cho trước) hay “đối tượng nhỏ”, MRF đa phân giải trên hệ số wavelet còn cải tiến tốc độ thực hiện thuật giải so với MRF một cấp. Các bước thực hiện tương tự như thuật giải 3.3, nhưng thay bước “phân rã LP” bằng biến đổi wavelet rời rạc và được tóm tắt trong thuật giải 3.4 như sau:

Thuật giải 3.4-Phân đoạn ảnh dựa trên WTMRF

- B1. Phân rã ảnh với biến đổi wavelet rời rạc L cấp.
- B2. Thực hiện phân đoạn với thuật giải SA đã trình bày trong thuật giải 3.1 trên dữ liệu LL (dữ liệu biến đổi qua bộ lọc tỉ lệ) ở cấp thô nhất. Gán kết quả phân đoạn cho ảnh ở cấp thô.
- B3. Lời giải ban đầu của cấp mịn hơn chính là quá trình tự ánh xạ giữa từng điểm ảnh đã phân đoạn ở cấp thô đến 4 điểm ảnh tương ứng.
- B4. Cập nhật lại tập tham số mô hình θ cho cấp mịn hơn.
- B5. Xác định lại giá trị “nhiệt độ khởi đầu” khi sử dụng phân đoạn dựa trên SA theo thuật giải 3.2.
- B6. Thực hiện bước B2 cho đến khi phân đoạn xong cấp mịn nhất.

Nhận xét ưu thế của mô hình này là có thể loại bỏ những phần nhỏ hay tích hợp luôn các phần nhỏ vào phần lớn hơn trong quá trình phân đoạn. Nguyên nhân là vì các "đối tượng nhỏ" sẽ bị mất đi trong quá trình phân rã WT ở các cấp thô.

Tuy nhiên trong một số trường hợp, ưu thế này có thể trở thành nhược điểm khi “đối tượng nhỏ” được xem là quan trọng trong ảnh, vì vậy chúng tôi đề nghị thuật giải trong phần 3.5.3 dưới đây nhằm giải quyết nhu cầu này.

3.5.3. Phân đoạn ảnh dựa trên kết hợp giữa MRF và UWT

Chúng tôi đề nghị MRF đa phân giải là sự kết hợp MRF với phân rã dạng UWT hoặc IUWT nhằm giải quyết vấn đề đã nêu trong 3.5.2. Điều khác biệt của lược đồ này so với WTMRf là thay vì dùng phân rã WT, chúng tôi đề nghị sử dụng UWT hoặc IUWT với thuật giải *a trous* như đã trình bày trong 2.3.4. Ưu thế là nhiều dạng Gaussian đều bị khử với UWT, vì vậy kết quả phân đoạn sẽ tốt hơn. Mặc dù thời gian thực hiện thuật giải này chậm hơn đáng kể so với phân đoạn dựa trên MRF một cấp (vì tất cả ảnh cần phân đoạn ở các cấp đều bằng kích thước ảnh gốc), tuy nhiên vẫn giữ được ưu thế của MRF một cấp và không làm mất quá nhiều đối tượng nhỏ như trong WTMRf. Các bước thực hiện tương tự như thuật giải 3.3, tuy nhiên với hai thay đổi: *thay bước B1 bằng phân rã UWT hoặc IUWT, và bước B3 không cần ánh xạ từ một điểm ảnh ở cấp thô đến 4 ở cấp mịn hơn, thay vì vậy ánh xạ 1-1 kết quả phân đoạn từ cấp thô đến cấp mịn hơn (vì ảnh ở các tỉ lệ vẫn giữ nguyên kích thước)*. Thuật giải 3.5 minh họa các bước chính thực hiện UWTMRf.

Thuật giải 3.5-Phân đoạn ảnh dựa trên UWTMRf

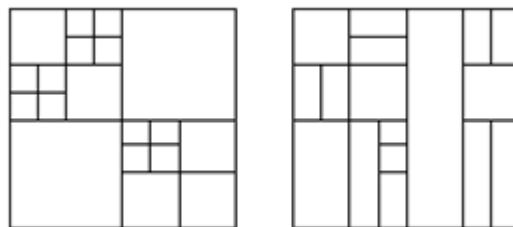
- B1. Phân rã ảnh với UWT hoặc IUWT rồi rạc L cấp với thuật giải *a trous*.
- B2. Thực hiện phân đoạn với thuật giải SA đã trình bày trong thuật giải 3.1 trên dữ liệu LL (dữ liệu biến đổi qua bộ lọc tỉ lệ) ở cấp thô nhất. Gán kết quả phân đoạn cho ảnh ở cấp thô.
- B3. Lờ giải ban đầu của cấp mịn hơn chính là quá trình tự ánh xạ giữa từng điểm ảnh đã phân đoạn ở cấp thô đến 1 điểm ảnh tương ứng.
- B4. Cập nhật lại tập tham số mô hình θ .
- B5. Xác định lại giá trị “nhiệt độ khởi đầu” khi sử dụng phân đoạn dựa trên SA theo thuật giải 3.2.
- B6. Thực hiện bước B2 cho đến khi phân đoạn xong cấp mịn nhất.

Tương tự như WTMRf, thuật giải phân đoạn ở mỗi cấp thực hiện trên hệ số đã biến đổi. Sau khi ảnh được phân rã qua UWT với L cấp, MRF được thực hiện trên các ảnh LL . Do LL chỉ chứa các dữ liệu tần số thấp, nên kết quả phân đoạn sẽ loại bỏ các yếu tố nhiễu trên ảnh. Các kết quả thực nghiệm trình bày trong phần 3.6 dưới đây thể hiện tính đúng đắn của mô hình.

3.5.4. Phân đoạn ảnh dựa trên kết hợp giữa MRF và WP bất đẳng hướng

Ý tưởng của mô hình này là thực hiện phân đoạn với MRF trên các hệ số của biến đổi wavelet bất đẳng hướng (Anisotropic Wavelet Packet-AWP) giới thiệu trong [15]. Chúng tôi đề xuất và gọi thuật giải phân đoạn này là AWPMRF. Trong đó, bước phân rã AWT được tóm tắt từ [15] như sau:

Thay vì thực hiện phân rã ảnh hai chiều thành quad-tree (theo biến đổi wavelet packet đẳng hướng) ở mỗi cấp phân rã, AWP thực hiện phân rã theo các cửa sổ chữ nhật kích thước khác nhau dựa trên độ đo hàm chi phí. Trong AWP, mỗi phân rã từ cấp mịn lên cấp thô kế tiếp theo dạng là phân rã hàng (hay cột) không nhất thiết phải thực hiện ngay phân rã cột (hay hàng) như minh họa trong hình 3.9. Thay vào đó, chỉ thực hiện phân rã wavelet một chiều trên hàng hoặc cột để tách thành hoặc thành hai khối ngang, hoặc thành hai khối dọc, hoặc không phân rã (tùy theo giá trị của hàm chi phí) ở mỗi cấp. Như vậy WP chỉ là trường hợp đặc biệt của AWP. Hình 3.9 minh họa sự khác biệt giữa phân rã AWP và WP, trong đó WP luôn phân rã dạng 4 bốn phần vuông, trong khi AWP theo các hình chữ nhật.



Hình 3.9-Sự khác biệt giữa wavelet packet và AWP

Mỗi bước phân rã wavelet hai chiều cho hàng (hay cột) thực chất là thực hiện DWT một chiều trên từng hàng (hay cột). Cho ảnh kích thước $M \times N$, thì

phân rã hàng sẽ sinh ra ma trận hệ số (tỉ lệ và wavelet) có kích thước $M \times (N/2)$ khi phân rã theo hàng, hay $(M/2) \times N$ cho phân rã cột. Trong [15] đã trình bày hai thuật giải: tham lam, và cây kép để xây dựng AWP. Tiêu chuẩn chọn phân rã tối ưu tại mỗi bước là chọn lời giải với hàm chi phí có giá trị nhỏ nhất trong ba trường hợp: không phân rã tiếp; phân rã theo hàng hoặc phân rã theo cột thỏa mãn tiêu chuẩn (17) trong [15]. AWP đã được chứng minh là có tính tái tạo đầy đủ. Các độ đo entropy sử dụng trong phân rã wavelet packet như Shannon, Log, hay chuẩn P có thể sử dụng như hàm chi phí. Tại mỗi cấp phân rã, nếu ảnh P được tách (theo cột hoặc theo dòng) thành P_A (phần xấp xỉ) và P_D (chi tiết) thì log entropy xác định bởi:

$$e(P') = e(P_A^c) + e(P_D^c), \text{ theo cột}$$

$$e(P'') = e(P_A^r) + e(P_D^r), \text{ theo hàng}$$

và dễ dàng chọn được dạng phân rã dựa trên tiêu chuẩn $\min(e(P), e(P'), e(P''))$, với P là ảnh gốc khi chưa biến đổi.

Thực nghiệm dưới đây (hình 3.10) so sánh kết quả giữa phân rã WT và AWT. Nhận xét rằng, số lượng khối của AWT ít hơn so với WT, đồng thời độ lợi thông entropy tin cũng ổn định hơn.



Theo thứ tự từ trái sang: ảnh gốc, phân rã WT ($\log \text{entropy} = -1.51024 E+06$), AWT ($\log \text{entropy} = -2.03765 E+06$) sử dụng hàm Db4 với 3 cấp phân rã.

Hình 3.10-So sánh phân rã WT và AWT.

Thuật giải hoàn chỉnh sử dụng quy hoạch động để tìm cách phân rã bất đối xứng tối ưu được trình bày trong [11].

Phân đoạn ảnh dựa trên mô hình này tương tự như WTMRF, và chúng tôi gọi là Anisotropic Wavelet Packet Markov Random Field (AWPMRF). AWPMRF cũng thực hiện phân đoạn với MRF trên dữ liệu ảnh LL (qua bộ lọc tỉ lệ) ở mỗi cấp phân rã, và không xét đến HH (dữ liệu biến đổi qua bộ lọc wavelet). Các bước thực hiện của AWPWRF tương tự như thuật giải 3.5, nhưng bước B1 được thay thế bởi AWP; và bước B3 không lấy kết quả ban đầu của phân đoạn cho tỉ lệ mịn hơn dựa vào ánh xạ từ 1 điểm ảnh đến 4 điểm ảnh. Thay vì vậy ánh xạ điểm ảnh sẽ thực hiện theo cấu trúc phân rã cây AWP [15]. Cụ thể hơn, kết quả khởi đầu cho phân đoạn dựa trên MRF của bước mịn hơn xác định bởi. Trong đó j thể hiện cấp phân rã, (p, q) vị trí của điểm ảnh:

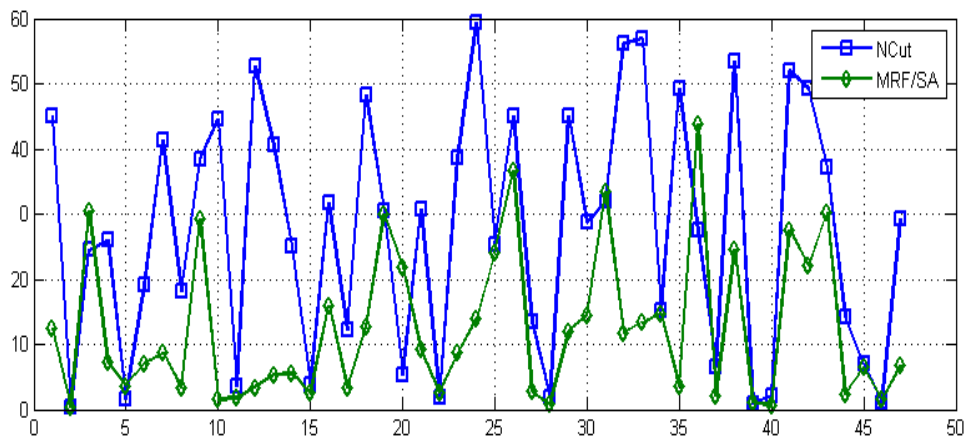
$$C_j(p, q) = \begin{cases} C_{j+1}(p, q) & null \\ C_{j+1}(\frac{p}{2}, q) & row \\ C_{j+1}(p, \frac{q}{2}) & column \end{cases} \quad (3.22)$$

Thuật giải 3.6-Phân đoạn ảnh dựa trên AWPMRF

- B1. Phân rã ảnh dựa trên AWP rời rạc L cấp.
- B2. Thực hiện phân đoạn với thuật giải SA đã trình bày trong thuật giải 3.1 trên dữ liệu LL (dữ liệu biến đổi qua bộ lọc tỉ lệ) ở cấp thô nhất. Gán kết quả phân đoạn cho ảnh ở cấp thô.
- B3. Lời giải ban đầu của cấp mịn hơn chính là quá trình tự ánh xạ giữa từng điểm ảnh đã phân đoạn ở cấp thô đến các điểm ảnh tương ứng theo công thức (3.22).
- B4. Cập nhật lại tập tham số mô hình θ .
- B5. Xác định lại giá trị “nhiệt độ khởi đầu” khi sử dụng phân đoạn dựa trên SA theo thuật giải 3.2.
- B6. Thực hiện bước B2 cho đến khi phân đoạn xong cấp mịn nhất.

3.6. Thực nghiệm trong phân đoạn vi ảnh

Thực nghiệm phân đoạn được thực hiện trên tập chứa 400 vi ảnh bán dẫn (nguồn công ty SiGlaz, Vigilant), và khoảng 100 ảnh vi sinh học trích từ phim quay quá trình phân bào (nguồn khoa sinh học, ĐHKHTN). Thuật giải SA được trình bày ở trên với các cấu trúc clique C_2 và C_3 được thực hiện trên các vi ảnh. Chúng tôi phân đoạn vi ảnh với các hàm thế V_c theo dạng tuyến tính và một số dạng như đã mô tả trong phần 3.3.1. Mỗi ảnh được phân đoạn với nhiệt độ khởi đầu T xác định bởi thuật giải 3.2, và áp dụng một số các hàm giảm nhiệt độ trong bảng 3.2, cũng như hàm giảm nhiệt độ thích nghi theo công thức (3.19). Các thí thực nghiệm phân đoạn với MRF một cấp và các mô hình MRF phân cấp đã trình bày ở trên. Ngoài ra, chúng tôi so sánh kết quả phân đoạn dựa trên MRF một cấp và thuật giải dựa trên Normalized Cut. Tiêu chuẩn so sánh dựa trên so khớp độ sai lệch giữa ảnh phân đoạn và ảnh phân đoạn lý tưởng (có được do vẽ thủ công). Kết quả được cho là tốt hơn nếu tỉ lệ lỗi nhỏ hơn. Bảng 3.3 và biểu đồ trong hình 3.11 thể hiện kết quả so sánh giữa thuật giải MRF/SA và NCut thực hiện trên một số ảnh ngẫu nhiên trong tập ảnh thực nghiệm. Hình 3.12 minh họa một số ảnh được phân đoạn thủ công, và kết quả dựa trên các thuật giải NCut, MRF/SA.

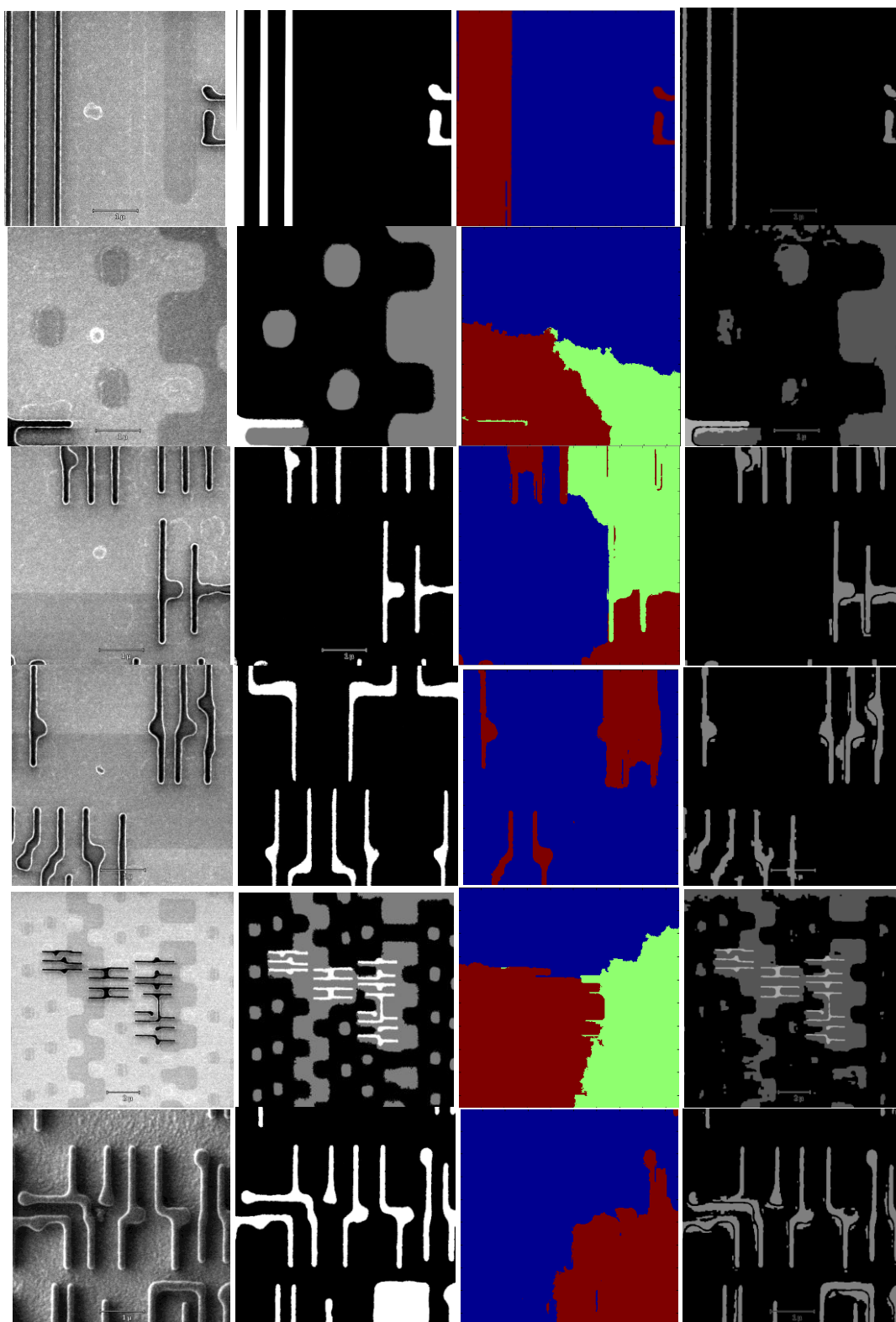


Hình 3.11-Biểu đồ so sánh tỉ lệ lỗi của NCut và MRF/SA

Bảng 3.3-So sánh tỉ lệ lỗi phân đoạn của NCut và MRF/SA

Stt	Tỉ lệ lỗi NCut trên toàn bộ ảnh (%)	Tỉ lệ lỗi của MRF/SA (%)
1	45.185	12.477
2	0.461	0.395
3	24.526	30.411
4	25.962	7.236
5	1.504	3.679
6	19.221	7.123
7	61.401	8.700
8	18.073	3.233
9	58.426	29.318
10	64.495	1.492
11	3.586	1.751
12	52.700	3.371
13	40.650	5.401
14	25.030	5.551
15	3.795	2.499
16	31.715	15.825
17	12.181	3.361
17	68.368	12.612
19	30.590	29.898
20	5.421	21.871
21	30.697	9.372
22	1.898	2.631
23	38.635	8.761
24	79.375	13.928
25	25.398	24.030
26	45.143	36.596

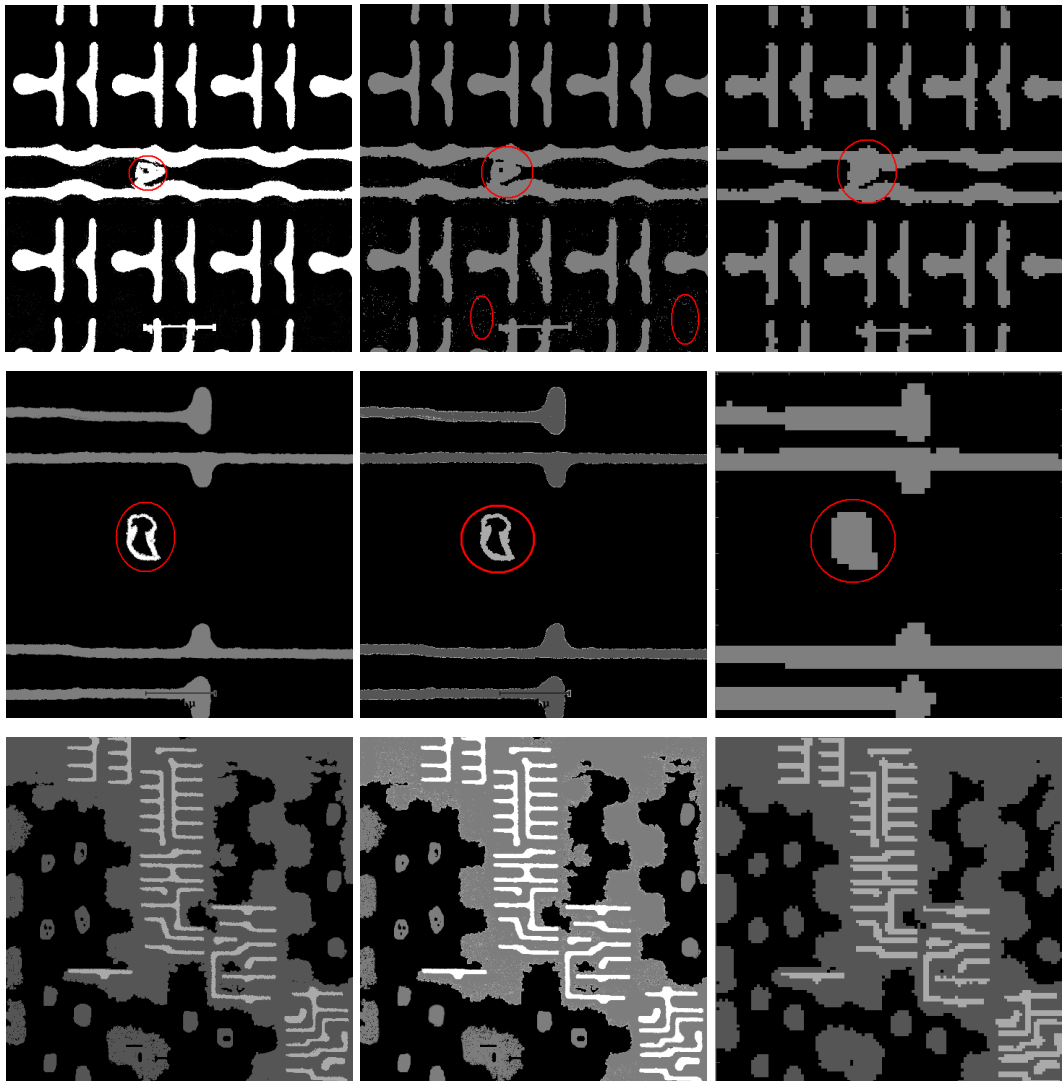
27	13.428	2.730
28	1.746	0.822
29	45.120	12.041
30	28.678	14.467
31	31.872	33.510
32	56.104	11.783
33	56.939	13.446
34	15.196	14.746
35	49.176	3.655
36	27.567	43.866
37	6.442	2.074
38	53.478	24.577
39	0.967	1.172
40	2.145	0.562
41	52.082	27.599
42	69.218	22.076
43	37.267	30.161
44	14.159	2.304
45	6.918	6.438
46	1.171	1.533
47	29.351	6.693



Theo thứ tự từ trái sang: hình gốc (nguồn công ty Vigilant), phân đoạn thủ công, Normalized Cut, MRF/SA.

Hình 3.12-Một số ảnh so sánh kết quả phân đoạn dựa trên MRF/SA và NCut

Hình 3.13 dưới đây thể hiện một số kết quả phân đoạn với MRF phân cấp dựa trên kết hợp của MRF/SA và biến đổi wavelet. Kết quả cho thấy rằng, các phần “nhiều” và các “đối tượng nhỏ” bị loại bỏ, trong khi các “đối tượng” lớn hơn được phân đoạn trong ảnh kết quả.



Theo thứ tự từ trái sang: hình gốc, MRF một cấp, UWTMRF

Hình 3.13-Một số kết quả phân đoạn dựa trên MRF một cấp, UWTMRF.

3.7. Tóm tắt

Phân tích dữ liệu đặc trưng không gian của ảnh SEM chụp lớp phủ phân tử giúp ích nhiều cho tăng hiệu suất quy trình sản xuất mạch bán dẫn. Chúng tôi đã mô tả phương pháp hai pha trích grain rất hiệu quả kết hợp giữa watershed dựa trên khoảng cách và cường độ màu và khử phân đoạn quá nát trên cả dữ liệu raster và vector. Các đặc trưng trong phân loại ảnh chụp bề mặt lớp phủ phân tử và phương pháp phân tích đã được kiểm nghiệm trong thực tế và cho kết quả tốt.

Với mô hình dựa trên MRF, việc tìm nhiệt độ khởi đầu và lượng giảm nhiệt độ thích nghi cho thuật giải SA đóng góp nhiều vào bài toán phân đoạn ảnh. Trong chương này, cùng với các cấu trúc clique và dạng của hàm thế, chi tiết cách thức tính toán hai thành phần quan trọng (nhiệt độ khởi đầu và quy trình giảm nhiệt độ thích nghi) của thuật giải SA đã được nghiên cứu và cải tiến. Tuy nhiên, bản chất của SA là dựa trên thủ tục tính MAP, nên việc tính toán khá nhiều. Vì vậy mô hình MRF phân cấp được sử dụng nhằm cải tiến tốc độ cũng như chất lượng phân đoạn ảnh. Trong mỗi cấp, giá trị β thể hiện sự tương tác giữa các điểm ảnh lân cận (nếu sử dụng hàm thế tuyến tính đẳng hướng) cũng được thay đổi theo quy luật tương tự như lượng giảm nhiệt độ nhằm làm cho chất lượng phân đoạn tốt hơn so với sử dụng duy nhất một giá trị β cho mọi cấp. Ngoài ra, với mô hình MRF đa phân giải, chúng tôi có thể sử dụng hàm giảm nhiệt độ thích nghi ở các cấp phân rã với nguyên tắc nhiệt độ khởi đầu cao ở các cấp phân rã thô là điều kiện thuận lợi cho khả năng hội tụ nhanh của thuật giải, và nhiệt độ khởi đầu nhỏ ở các tỉ lệ mịn. Mô hình WTMRF thể hiện ưu thế khi cần tích hợp các phần nhỏ trong quá trình phân đoạn. Kết quả thực nghiệm phân đoạn vi ảnh dựa trên UWTMRF/IUWTMRF cho thấy chất lượng tốt hơn, và đặc biệt hiệu quả khi được thực hiện trên vi ảnh có độ nhiễu lớn. Một điểm nổi bật của chương này là dựa trên kết hợp giữa phân rã AWP và MRF nhằm đạt được chất lượng phân đoạn tốt hơn khi quá trình phân rã wavelet bất đẳng hướng đã xét đến yếu tố độ lợi thông tin, do đó các ảnh con phân rã ở các tỉ lệ hầu như đã được “phân đoạn” thô trước. Kết quả thực nghiệm cho thấy các mô

hình này đều cho kết quả phân đoạn tốt hơn (chất lượng, tốc độ thực hiện) so với MRF một cấp. Việc kết hợp MRF với phân rã giảm kích thước mẫu ngoài việc cải tiến tốc độ hội tụ, còn mang lại lợi ích khi phân đoạn ảnh ở các tỉ lệ khác nhau nhằm trích được dạng tổng quát của “đối tượng”. Trong khi đó với các phân rã giữ nguyên kích thước mẫu, chúng tôi phân đoạn được cả đối tượng nhỏ trong khi các nhiễu bị loại trừ. Kết quả thực nghiệm được so sánh với thuật giải Normalized Cut cho thấy ưu điểm của các phương pháp đã đề nghị.