

LỜI CẢM ƠN

Trước hết em xin chân thành cảm ơn các thầy cô giáo trong khoa công nghệ thông tin trường đại học dân lập Hải Phòng đã trang bị những kiến thức cơ bản cần thiết để em thực hiện đề tài của mình.

Đặc biệt em xin bày tỏ lòng kính trọng và biết ơn sâu sắc tới thầy giáo hướng dẫn Ths. Ngô Trường Giang đã tận tình hướng dẫn, chỉ bảo và tạo mọi điều kiện thuận lợi giúp em trong quá trình làm đồ án tốt nghiệp.

Trong quá trình thực hiện đồ án tốt nghiệp, mặc dù đã cố gắng hết sức xong do trình độ còn hạn chế, nội dung đề tài còn quá mới mẻ và khó với em nên khó tránh khỏi những sai sót trong quá trình tiếp nhận kiến thức. Vì vậy, em rất mong nhận được sự thông cảm, chỉ dẫn, giúp đỡ của các thầy cô và sự góp ý bạn bè.

Một lần nữa em xin chân thành cảm ơn !

Hải Phòng, ngày.....tháng.....năm.....

Sinh viên

Đỗ Thanh Thủy.

MỤC LỤC

MỞ ĐẦU	4
DANH MỤC HÌNH VẼ	6
CHƯƠNG 1: TỔNG QUAN VỀ XỬ LÝ ẢNH.....	7
1.1 Các khái niệm cơ bản trong xử lý ảnh.....	7
1.1.1 Xử lý ảnh là gì.....	7
1.1.2 Ảnh và điểm ảnh	7
1.1.3 Quan hệ giữa các điểm ảnh.....	8
1.1.4 Mức xám của ảnh.....	8
1.1.5 Độ phân giải	9
1.2 Các phép toán cơ bản trên ảnh nhị phân.....	9
1.2.1 Các phép toán logic.....	9
1.2.2 Các phép toán hình thái học.....	10
1.3 Các giai đoạn trong xử lý ảnh.....	16
1.4 Một số ứng dụng cơ bản.....	18
CHƯƠNG 2: TỔNG QUAN VỀ PHÂN ĐOẠN ẢNH.....	20
2.1 Khái niệm phân đoạn ảnh.....	20
2.2 Các hướng tiếp cận trong phân đoạn ảnh.....	20
2.2.1 Phân đoạn dựa vào ngưỡng.....	20
2.2.2 Phân đoạn dựa theo đường biên.....	22
2.2.3 Phân đoạn theo miền đồng nhất.....	26
CHƯƠNG 3: PHÂN ĐOẠN ẢNH DỰA TRÊN RWR.....	28

3.1	Giới thiệu	28
3.2	Random Walker Restart (RWR).....	30
3.3	Phương pháp phân đoạn dựa trên RWR.....	34
3.3.1	Mô hình đồ thị.....	35
3.3.2	Học	36
3.3.3	Phân đoạn.....	38
CHƯƠNG 4: CÀI ĐẶT THỬ NGHIỆM.....		40
4.1	Môi trường cài đặt	40
4.2	Chương trình thực nghiệm.....	40
4.2.1	Kết quả phân đoạn ảnh sử dụng RWR.....	40
4.2.2	So sánh kết quả phân đoạn bằng RWR với một số phương pháp khác.	41
KẾT LUẬN.....		45
TÀI LIỆU THAM KHẢO		46

MỞ ĐẦU

Xử lý ảnh (XLA) là một trong những chuyên ngành quan trọng và lâu đời của Công nghệ thông tin. XLA được áp dụng trong nhiều lĩnh vực khác nhau như y học, vật lý, hoá học, tìm kiếm tội phạm,... Mục đích chung của việc XLA thường là: (1) xử lý ảnh ban đầu để có được một bức ảnh mới theo một yêu cầu cụ thể; (2) phân tích ảnh để thu được các thông tin đặc trưng trên ảnh nhằm hỗ trợ cho việc phân loại và nhận biết ảnh; (3) phân đoạn ảnh (image segmentation) để nhận diện được các thành phần trong ảnh nhằm hiểu được kết cấu của bức ảnh ở mức độ cao hơn. Để xử lý được một bức ảnh thì phải trải qua nhiều bước, nhưng bước quan trọng và khó khăn nhất đó là phân đoạn ảnh. Nếu bước phân đoạn ảnh không tốt thì dẫn đến việc nhận diện sai lầm về các đối tượng có trong ảnh.

Trong khoảng 30 năm trở lại đây đã có rất nhiều các thuật toán được đề xuất để giải quyết bài toán phân đoạn ảnh. Các thuật toán hầu hết đều dựa vào hai thuộc tính quan trọng của mỗi điểm ảnh so với các điểm lân cận của nó, đó là: sự khác (dissimilarity) và giống nhau (similarity). Các phương pháp dựa trên sự khác nhau của các điểm ảnh được gọi là các phương pháp biên (boundary-based methods), còn các phương pháp dựa trên sự giống nhau của các điểm ảnh được gọi là phương pháp miền (region-based methods). Tuy nhiên, cho đến nay các thuật toán theo cả hai hướng này đều vẫn chưa cho kết quả phân đoạn tốt, vì cả hai loại phương pháp này đều chỉ nắm bắt được các thuộc tính cục bộ (local) của ảnh. Do đó, trong thời gian gần đây, việc tìm ra các thuật toán nắm bắt được các thuộc tính toàn cục (global) của bức ảnh đã trở thành một xu hướng.

Mục đích chính của em là nắm được tổng quan về xử lý ảnh số, nắm được các hướng tiếp cận chính trong phân đoạn ảnh và cài đặt thử nghiệm một vài thuật toán phân đoạn ảnh. Vấn đề mấu chốt trong đề án này là em tập trung tìm hiểu và trình bày thêm một phương pháp được đánh giá là hiệu quả hơn các phương pháp trước đây, khắc phục được hai khó khăn quan trọng trong ảnh tự nhiên là bài toán đường biên yếu và kết cấu yếu. Phương pháp này dựa vào việc coi một bức ảnh như một đồ thị có trọng số. Sau khi tính xác suất trạng thái ổn định của mỗi điểm ảnh bằng cách sử dụng RWR, chúng ta có thể ước lượng khả năng phân tách và cuối cùng gán nhãn vào mỗi điểm ảnh.

Ngoài phần mở đầu và kết luận, đề án được chia làm 4 chương, cụ thể nội dung các chương như sau:

Chương 1: Tổng quan về xử lý ảnh

Chương 2: Phân đoạn ảnh và các hướng tiếp cận trong phân đoạn ảnh.

Chương 3: Tìm hiểu phương pháp phân đoạn ảnh RWR (Random Walker Restart).

Chương 4: Cài đặt thử nghiệm thuật toán phân đoạn ảnh dựa trên RWR.

DANH MỤC HÌNH VẼ

<i>Hình 1.1. Hình minh họa các phép toán trên ảnh nhị phân</i>	<i>10</i>
<i>Hình 1.2. Hiệu quả của thao tác nhị phân đơn giản trên một ảnh nhỏ</i>	<i>11</i>
<i>Hình 1.3. A dẫn bởi B.....</i>	<i>12</i>
<i>Hình 1.4. Dẫn mất điểm ảnh.....</i>	<i>12</i>
<i>Hình 1.5. Dẫn ảnh sử dụng phần tử cấu trúc</i>	<i>13</i>
<i>Hình 1.6. Phép co nhị phân.....</i>	<i>13</i>
<i>Hình 1.7. Sử dụng phép toán mở.....</i>	<i>15</i>
<i>Hình 1.8. Phép đóng</i>	<i>15</i>
<i>Hình 1.9. Phép đóng với độ sâu lớn.....</i>	<i>16</i>
<i>Hình 1.10. Các giai đoạn chính trong xử lý ảnh</i>	<i>16</i>
<i>Hình 2.1. Đường biên lý tưởng</i>	<i>23</i>
<i>Hình 2.2. Đường biên bậc thang.....</i>	<i>23</i>
<i>Hình 2.3. Đường biên thực</i>	<i>24</i>
<i>Hình 3.1 Phân đoạn đơn nhãn.....</i>	<i>30</i>
<i>Hình 3.2. Kết quả phân đoạn</i>	<i>38</i>
<i>Hình 4.1. Một ví dụ về sự thay đổi xác suất trạng thái ổn định r theo xác suất khởi động lại c.....</i>	<i>40</i>
<i>Hình 4.2. Một ví dụ về phân đoạn đối với sự biến đổi của các xác suất khởi động lại c trong ảnh tự nhiên.....</i>	<i>41</i>
<i>Hình 4.3. So sánh thuật toán GC, RW, RWR cho việc tìm kiếm đường biên yếu</i>	<i>42</i>
<i>Hình 4.4. So sánh phân đoạn kết cấu giữa các thuật toán GC, RW, RWR.....</i>	<i>43</i>
<i>Hình 4.5. So sánh thuật toán GC, RW, RWR trên ảnh tự nhiên</i>	<i>44</i>

CHƯƠNG 1: TỔNG QUAN VỀ XỬ LÝ ẢNH

1.1 Các khái niệm cơ bản trong xử lý ảnh

1.1.1 Xử lý ảnh là gì

Quá trình xử lý ảnh được xem như là quá trình thao tác ảnh đầu vào nhằm cho kết quả mong muốn. Kết quả đầu ra của một quá trình xử lý ảnh có thể là một ảnh tốt hơn hoặc một kết luận.

Mục tiêu của xử lý ảnh có thể chia làm ba hướng như sau:

- Xử lý ảnh ban đầu để có được ảnh mới theo một yêu cầu xác định (Ví dụ như ảnh mờ cần xử lý để được ảnh rõ hơn).
- Phân tích ảnh để thu được các thông tin đặc trưng giúp cho việc phân loại, nhận biết ảnh (Ví dụ phân tích ảnh vân tay để trích chọn đặc trưng vân tay).
- Hiểu ảnh đầu vào để có những mô tả về ảnh ở mức cao hơn (Ví dụ từ một ảnh tai nạn giao thông có thể phát họa hiện trường tai nạn).

1.1.2 Ảnh và điểm ảnh

Ảnh tự nhiên là ảnh liên tục về không gian và độ sáng. Để xử lý bằng máy tính (số), ảnh cần phải được số hóa. Số hóa là sự biến đổi gần đúng một ảnh liên tục thành một tập điểm phù hợp với ảnh thật về trí (không gian) và độ sáng (mức xám). Khoảng cách giữa các điểm ảnh được thiết lập sao cho mắt người không phân biệt được ranh giới giữa chúng. Mỗi một điểm như vậy gọi là điểm ảnh (PEL: Picture Element) hay gọi tắt là Pixel. Trong khuôn khổ ảnh hai chiều, mỗi pixel tương ứng với cặp tọa độ (x, y) .

Điểm ảnh (pixel) là một phần tử của ảnh số tại tọa độ (x, y) với độ xám hoặc màu nhất định. Kích thước và khoảng cách giữa các điểm ảnh được chọn thích hợp sao cho mắt người cảm nhận được sự liên tục về không gian và mức xám của ảnh số gần như ảnh thật. Mỗi phần tử trong ma trận được gọi là phần tử ảnh.

Ảnh được xem như tập hợp các điểm ảnh.

1.1.3 Quan hệ giữa các điểm ảnh

1.1.3.1 Các lân cận của điểm ảnh

Giả sử một ảnh số được biểu diễn bằng hàm $f(x, y)$, p và q là cặp điểm ảnh có quan hệ với nhau, điểm ảnh p có tọa độ (x, y) . Định nghĩa các lân cận của điểm ảnh.

- Lân cận 4 của p kí hiệu $N_4(p)$: $N_4(p) = \{(x-1, y); (x, y-1); (x, y+1); (x+1, y)\}$
- Lân cận chéo của p kí hiệu $N_p(p)$: $N_p(p) = \{(x+1, y+1); (x+1, y-1); (x-1, y+1); (x-1, y-1)\}$
- Lân cận 8 của p kí hiệu $N_8(p)$: $N_8(p) = N_4(p) + N_p(p)$

1.1.3.2 Các mối liên kết điểm ảnh

Các mối liên kết được sử dụng để xác định giới hạn của đối tượng hoặc xác định vùng trong một ảnh. Một liên kết được đặc trưng bởi tính liên kề giữa các điểm và mức xám của chúng. Có ba loại liên kết:

- Liên kết 4: Hai điểm ảnh p và q được gọi là liên kết 4 nếu q thuộc $N_4(p)$
- Liên kết 8: Hai điểm ảnh p và q được gọi là liên kết 8 nếu q thuộc $N_8(p)$
- Liên kết m (liên kết hỗn hợp): Hai điểm ảnh p và q được gọi là liên kết hỗn hợp nếu q thuộc $N_4(p)$ hoặc q thuộc $N_8(p)$

1.1.3.3 Đo khoảng cách giữa các điểm ảnh

Khoảng cách $D(p, q)$ giữa hai điểm ảnh p tọa độ (x, y) , q tọa độ (s, t) là hàm khoảng cách (Distance) nếu:

- $D(p, q) \geq 0$ (Với $D(p, q)=0$ khi và chỉ khi $p=q$)
- $D(p, q) = D(q, p)$
- $D(p, z) \leq D(p, q) + D(q, z)$; z là một điểm ảnh khác.

Khoảng cách Euclide giữa hai điểm ảnh $p(x, y)$ và $q(s, t)$ được định nghĩa như sau: $D_e(p, q) = [(x - s)^2 + (y - t)^2]^{1/2}$

1.1.4 Mức xám của ảnh

Mức xám của điểm ảnh là cường độ sáng của nó được gán bằng giá trị số tại điểm đó.

Các thang giá trị mức xám thông thường là: 16, 32, 64, 128, 256 (Mức 256 là mức phổ dụng nhất vì máy tính dùng 1 byte (8 bit) để biểu diễn mức xám. Mức xám dùng 1 byte biểu diễn: $2^8=256$, tức là từ 0 đến 255)

Ảnh đen trắng là ảnh có hai màu đen và trắng. Nếu phân mức đen trắng thành L mức, sử dụng số bit B để mã hóa mức đen trắng (hay mức xám) thì L được xác định: $L=2^B$. Nếu $L=2$, $B=1$ nghĩa là chỉ có 2 mức 0 và 1. Ảnh dùng hai mức 0 và 1 để biểu diễn mức xám gọi là ảnh nhị phân. Mức 1 ứng với màu sáng còn mức 0 ứng với màu tối. Nếu L lớn hơn 2 đó là ảnh đa cấp xám. Như vậy ảnh nhị phân mỗi điểm ảnh được mã hóa trên 1 bit, còn ảnh 256 mức mỗi điểm ảnh được mã hóa trên 8 bit. Ảnh đen trắng nếu dùng 8 bit (1 byte) để biểu diễn mức xám số mỗi mức xám được biểu diễn dưới dạng một số nguyên nằm trong khoảng từ 0 đến 255, mức 0 biểu diễn cho cường độ đen nhất và mức 255 biểu diễn cho cường độ sáng nhất.

Ảnh màu: là ảnh tổ hợp từ 3 màu cơ bản đỏ (Red), lục (Green), lam (Blue). Để biểu diễn cho một điểm ảnh màu dùng 3 byte để mô tả 24 bit màu $2^{8*3}=2^{24} \approx 16,7$ triệu màu.

1.1.5 Độ phân giải

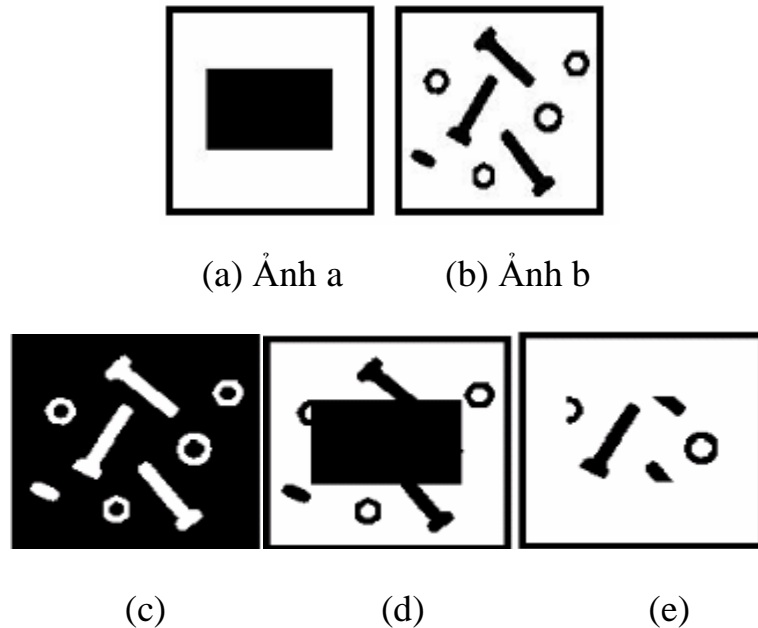
Độ phân giải (Resolution) của ảnh là mật độ điểm ảnh được ấn định trên ảnh số khi hiển thị. Như vậy khoảng cách giữa các điểm ảnh được chọn sao cho mắt người vẫn thấy được sự liên tục của ảnh. Việc lựa chọn khoảng cách thích hợp tạo nên một mật độ phân bố, đó chính là độ phân giải và được phân bố theo trục x và y trong không gian hai chiều.

1.2 Các phép toán cơ bản trên ảnh nhị phân

1.2.1 Các phép toán logic

P	Q	P AND Q	P OR Q	NOT P
0	0	0	0	1
0	1	0	1	1
1	0	0	1	0
1	1	1	1	0

Hình 1.1 dưới đây minh họa những thao tác nói trên với giá trị nhị phân “1” có màu đen, còn giá trị nhị phân “0” có màu trắng.



Hình 1.1. Hình minh họa các phép toán trên ảnh nhị phân

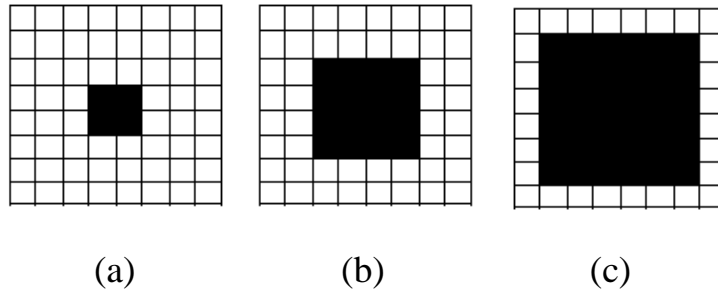
Trong hình 1.1: hình (a) và (b) là ảnh ban đầu; (c) phép NOT (b); (d) phép OR (a,b); (e) phép AND (a,b).

1.2.2 Các phép toán hình thái học

Hình thái (morphology) có nghĩa là “hình thức và cấu trúc của một đối tượng”, hoặc là cách sắp xếp mối quan hệ bên trong giữa các phần của đối tượng. Hình thái có liên quan đến hình dạng, và hình thái số là một cách để mô tả hoặc phân tích hình dạng của một đối tượng số.

Những thao tác hình thái nhị phân được xây dựng trên ảnh chỉ có 2 mức xám 0 và 1, “0” ứng với màu trắng, “1” ứng với màu đen. Trước hết, để bắt đầu, ta hãy xem hình 1.2a. Tập hợp các điểm ảnh đen tạo nên đối tượng ảnh hình vuông và trong hình 1.2b, đối tượng ảnh cũng là hình vuông nhưng là hình vuông lớn hơn so với hình 1.2a một điểm ảnh về mọi phía, nghĩa là thay mọi lân cận trắng của các điểm ảnh trong hình 1.2a thành các điểm ảnh đen. Đối tượng trong hình 1.2b cũng được thao tác tương tự, tức là hình 1.2b được tăng thêm một điểm ảnh về mọi phía. Thao tác đó có thể coi như một phép dẫn đơn giản, phép dẫn một điểm ảnh về mọi phía. Việc dẫn đó có thể được thực hiện cho đến khi toàn bộ ảnh được thay bằng các điểm ảnh đen. Do vậy, đối tượng ảnh trong hình 1.2a có thể được viết lại là $\{(3, 3) (3, 4) (4, 3) (4,4)\}$, với điểm

ảnh phía trên bên trái là $(0, 0)$. Tuy nhiên, việc viết như vậy sẽ rất dài dòng và bất tiện nên ta gọi đơn giản đối tượng ảnh là A , và các phần tử trong đó là các điểm ảnh.



Hình 1.2. Hiệu quả của thao tác nhị phân đơn giản trên một ảnh nhỏ

Trong hình 1.2, hình (a) ảnh ban đầu; (b) ảnh dẫn 1 điểm ảnh; (c) ảnh dẫn 2 điểm ảnh so với ảnh ban đầu.

1.2.2.1 Phép dẫn nhị phân

Bây giờ ta sẽ chỉ ra thao tác tập hợp đơn giản nhằm mục đích định nghĩa phép dẫn nhị phân. Phép dịch A bởi điểm x (hàng, cột), được định nghĩa là một tập:

$$(A)x = \{c \mid c = a + x, a \in A\} \quad (1.1)$$

Chẳng hạn nếu x có tọa độ $(1, 2)$, khi đó điểm ảnh đầu tiên phía trên bên trái của A sẽ dịch đến vị trí: $(3, 3) + (1, 2) = (4, 5)$. Các điểm ảnh khác trong A sẽ dịch chuyển một cách tương ứng, tức ảnh được dịch sang phải (cột) điểm ảnh và xuống phía dưới (hàng) điểm ảnh.

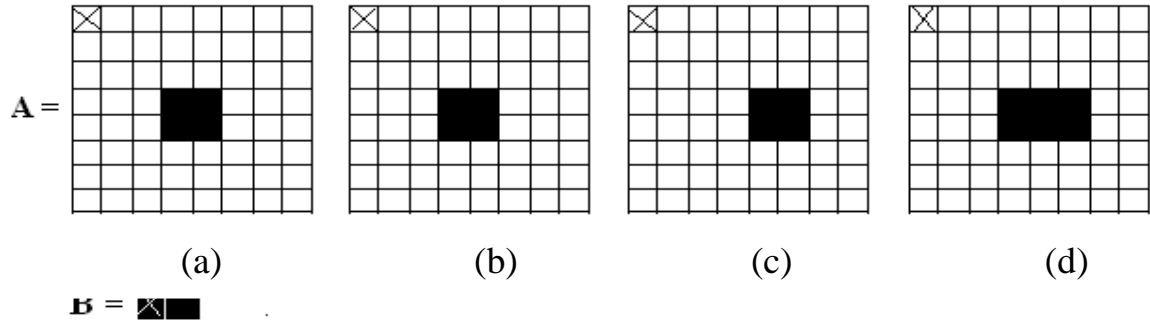
Bây giờ ta có thể định nghĩa phép dẫn (dilation) qua lý thuyết tập hợp như sau: Phép dẫn tập A bởi tập B , đó là tập:

$$A \oplus B = \{c \mid c = a + b, a \in A, b \in B\} \quad (1.2)$$

Để thấy trong toán học, đây là phép tổng trực tiếp A và B . A là đối tượng ảnh được thao tác và B được gọi là phần tử cấu trúc (viết tắt là cấu trúc). Để hiểu kỹ hơn về điều này, ta hãy coi A là đối tượng trong hình 1.2a và $B = \{(0,0), (0, 1)\}$.

Những phần tử trong tập $C = A \oplus B$ được tính dựa trên công thức (1.1), có thể viết lại như sau:

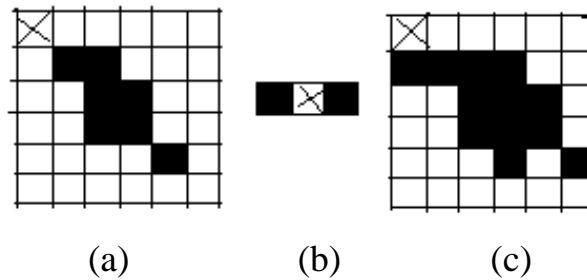
$$A \oplus B = (A + \{(0, 0)\}) \cup (A + \{(0, 1)\}) \quad (1.3)$$



Hình 1.3. A dẫn bởi B

Trong hình 1.3: (a) tập A ban đầu; (b) tập A cộng phần tử $(0, 0)$; (c) tập A cộng phần tử $(0, 1)$; (d) hợp của (b) và (c) (kết quả của phép dẫn).

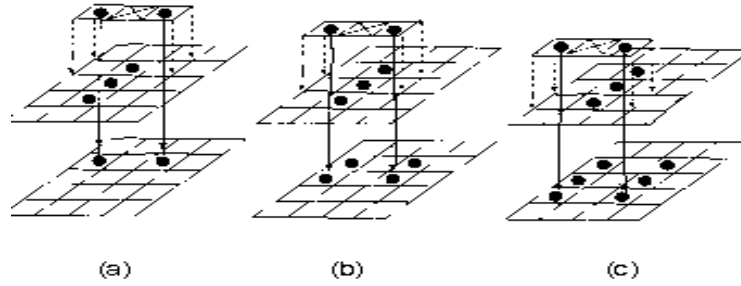
Nhận thấy rằng trong hình 1.4, có một số phần tử của đối tượng ban đầu sẽ không có.



Hình 1.4. Dẫn mất điểm ảnh

Trong hình 1.4. (a) ảnh $A1$; (b) phần tử cấu trúc $B1$; (c) $A1$ được dẫn bởi $B1$.

Từ những điều trên, giúp ta tiếp cận đến một thao tác dẫn ảnh có thể được “máy tính hóa”. Ta hãy coi những phần tử cấu trúc như là một mẫu và dịch nó trên ảnh. Điều này được thể hiện khá rõ trong hình 1.5.



Hình 1.5. Dẫn ảnh sử dụng phần tử cấu trúc

Trong hình 1.5: (a) là góc cấu trúc định vị trên điểm ảnh đen đầu tiên và những điểm đen cấu trúc được chép sang ảnh kết quả ở những vị trí tương ứng; (b) quá trình tương tự với điểm đen tiếp theo; (c) quá trình hình thành.

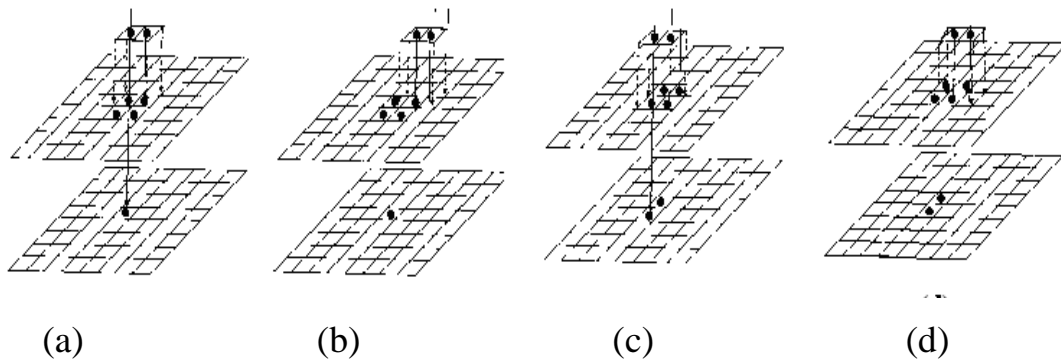
1.2.2.2 Phép co nhị phân

Nếu như phép dẫn có thể nói là thêm điểm ảnh vào trong đối tượng ảnh, làm cho đối tượng ảnh trở nên lớn hơn thì phép co sẽ làm cho đối tượng ảnh trở nên nhỏ hơn, ít điểm ảnh hơn. Trong trường hợp đơn giản nhất, một phép co nhị phân sẽ tách lớp điểm ảnh bao quanh đối tượng ảnh, chẳng hạn hình 1.2b là kết quả của phép co được áp dụng đối với hình 1.2c.

Nhìn chung, phép co một ảnh A bởi cấu trúc B có thể được định nghĩa như là tập:

$$A \ominus B = \{c \mid (B)c \subseteq A\} \quad (1.4)$$

Đầu tiên, ta hãy xét một ví dụ đơn giản sau đây:



Hình 1.6. Phép co nhị phân.

Phần tử cấu trúc được dịch chuyển đến vị trí một điểm đen trong ảnh. Trong trường hợp này, các thành viên của cấu trúc đều phù hợp với những điểm đen của ảnh cho nên cho kết quả điểm đen.

Phần tử cấu trúc dịch chuyển tới điểm ảnh tiếp theo trong ảnh, và có một điểm không phù hợp và kết quả là điểm trắng.

Ở lần dịch chuyển tiếp theo, các thành viên của cấu trúc lại phù hợp nên kết quả là điểm đen.

Tương tự được kết quả cuối cùng là điểm trắng.

Ta nhận thấy một điều quan trọng là: Phép co và phép dẫn không phải là những thao tác ngược nhau. Có thể trong một số trường hợp đúng là phép co sẽ giải hoạt hiệu quả của phép dẫn. Nhưng nhìn chung thì điều đó là không đúng, ta sẽ quan sát chúng một cách cụ thể hơn ở sau. Tuy nhiên, giữa phép co và phép dẫn có mối quan hệ qua biểu thức sau đây:

$$(B \ominus A)_c = Bc \oplus \hat{A} \quad (1.5)$$

Tức là phần bù của phép co ảnh A bởi B được coi như phép dẫn phần bù của A bởi tập đối của B. Nếu như cấu trúc B là đối xứng (ở đây ta quan niệm đối xứng theo toạ độ) thì tập đối của B không thay đổi, nghĩa là $\hat{A} = A$

Khi đó:

$$(B \ominus A)_c = Bc \oplus A \quad (1.6)$$

Hay, phần bù của phép co A bởi B được coi như phép dẫn nền của ảnh A (ta quy ước trong ảnh nhị phân rằng: đối tượng ảnh là những điểm đen quan sát, ảnh A là bao gồm cả điểm đen và nền).

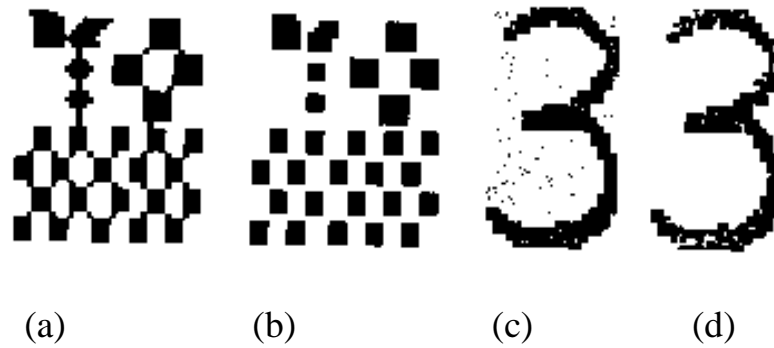
1.2.2.3 Phép mở (Opening)

Nếu như ta áp dụng phép co ảnh đối với một ảnh và sau đó lại áp dụng tiếp phép dẫn ảnh đối với kết quả trước thì thao tác đó được gọi là phép mở ảnh, hay với I là ảnh, D là Dilation (dãn) và E là Erosion (co).

$$\text{Opening}(I) = D(E(I)) \quad (1.7)$$

Tên của phép toán “mở” ảnh dường như đã phản ánh rõ tác dụng của nó. Tác dụng của nó chính là “mở” những khoảng trống nhỏ giữa các phần tiếp xúc trong đối

tượng ảnh, làm cho ảnh dường như bớt “gai”. Hiệu quả này dễ quan sát nhất khi sử dụng cấu trúc đơn giản. Hình 1.7 trình bày ảnh có những phần của nó tiếp xúc nhau. Sau thao tác mở đơn giản đối tượng ảnh đã dễ nhận hơn so với ban đầu.



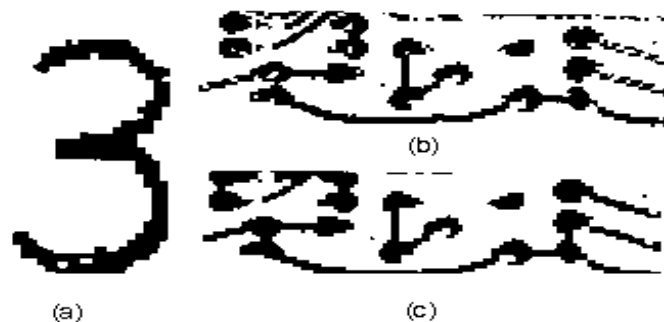
Hình 1.7. Sử dụng phép toán mở

Trong hình 1.7: (a) một ảnh có nhiều vật thể được liên kết; (b) các vật thể được cách ly bởi phép mở với cấu trúc đơn giản; (c) một ảnh có nhiễu; (d) ảnh nhiễu sau khi sử dụng phép mở, các điểm nhiễu.

1.2.2.4 Phép đóng (Closing)

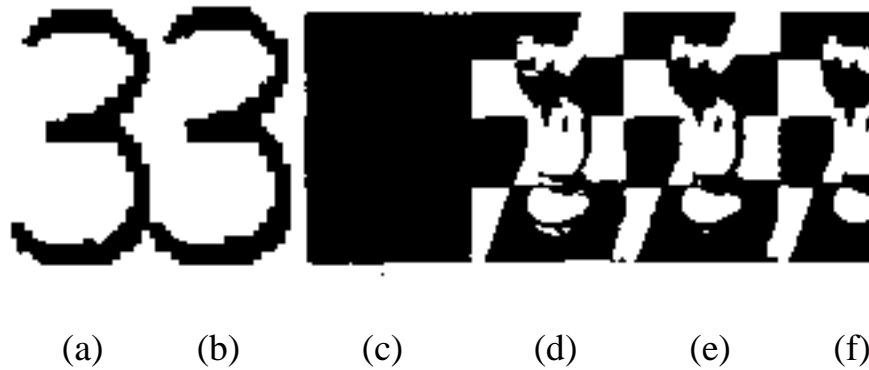
Tương tự phép mở ảnh nhưng trong phép đóng ảnh, thao tác dẫn ảnh được thực hiện trước, sau đó mới đến thao tác co ảnh và cùng làm việc trên cùng một phần tử cấu trúc.

$$\text{Close}(I) = E(D(I)) \quad (1.8)$$



Hình 1.8. Phép đóng

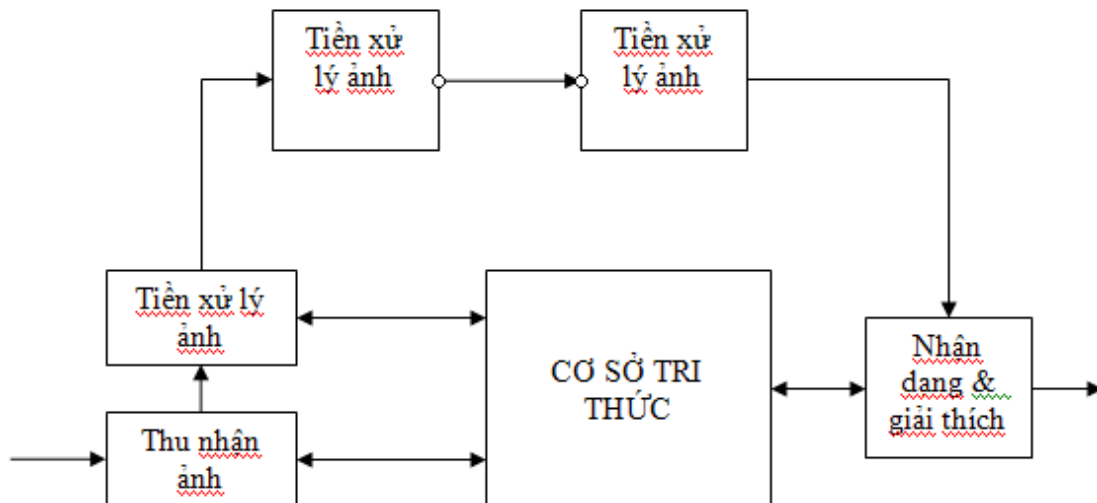
Trong hình 1.8: (a) kết quả đóng sử dụng cấu trúc đơn giản; (b) ảnh của một bảng mạch được phân ngưỡng và có các vết đứt; (c) ảnh tương tự sau khi đóng nhưng những nét đứt đã được nối liền.



Hình 1.9. Phép đóng với độ sâu lớn

Trong hình 1.9: (a) từ hình 1.8a, sử dụng phép đóng với độ sâu 2; (b) phép đóng với độ sâu 3; (c) một vùng bàn cờ; (d) vùng bàn cờ được phân ngưỡng thể hiện những điểm bất quy tắc và một vài lỗ; (e) sau khi thực hiện phép đóng với độ sâu 1; (f) Sau khi thực hiện phép đóng với độ sâu 2.

1.3 Các giai đoạn trong xử lý ảnh



Hình 1.10. Các giai đoạn chính trong xử lý ảnh

Bước 1: Thu nhận ảnh. Để thực hiện bước này chúng ta cần có 1 bộ cảm biến lấy ảnh và khả năng số hóa những tín hiệu liên tục được sinh ra bởi bộ cảm biến đó. Bộ cảm biến ở đây có thể là 1 máy chụp ảnh đơn sắc hay màu hoặc 1 máy chụp ảnh kiểu quét dòng cho ra 1 dòng ảnh ở một thời điểm cụ thể. Mặc dù đây chỉ là bước đầu tiên nhưng kết quả của nó có thể ảnh hưởng rất nhiều đến công đoạn kế tiếp tùy theo loại hình ứng dụng, chất lượng và chủng loại của thiết bị lấy ảnh.

Bước 2: Tiền xử lý ảnh. Ở bước này, ảnh sẽ được cải thiện về độ tương phản, khử nhiễu, khử bóng, khử độ lệch,... với mục đích làm cho chất lượng ảnh trở lên tốt hơn nữa chuẩn bị cho các bước xử lý phức tạp hơn về sau trong quá trình xử lý ảnh.

Bước 3: Phân đoạn ảnh. Trong bước này, ảnh đầu vào được chia thành nhiều phần nhỏ khác nhau hay còn gọi là các đối tượng. Việc phân đoạn ảnh thành tập những đối tượng khác nhau là nhiệm vụ phức tạp nhất trong xử lý ảnh số hóa. Nếu kết quả phân đoạn ảnh chỉ dừng lại ở mức thô thiển thì toàn bộ những bước xử lý tiếp theo sẽ không cho kết quả tốt. Mặt khác, các thuật toán phân đoạn ảnh không đủ mạnh, hoạt động không ổn định cũng là nguồn gốc dẫn đến sự thất bại của một giải pháp xử lý ảnh. Kết quả của bước phân đoạn ảnh thường được cho dưới dạng dữ liệu thô, trong đó hàm chứa biên của 1 vùng ảnh, hoặc tập hợp tất cả những điểm ảnh thuộc về chính vùng ảnh đó. Trong cả 2 trường hợp, sự chuyển đổi dữ liệu thô này thành 1 dạng thích hợp hơn cho việc xử lý trong máy tính là hết sức cần thiết. Để chuyển đổi chúng, câu hỏi đầu tiên cần phải trả lời là nên biểu diễn một vùng ảnh dưới dạng biên hay dưới dạng 1 vùng hoàn chỉnh gồm tất cả những điểm ảnh thuộc về nó. Biểu diễn dạng biên cho 1 vùng phù hợp với ứng dụng chỉ quan tâm chủ yếu đến các đặc trưng hình dạng bên ngoài của đối tượng. Còn biểu diễn dạng vùng lại thích hợp cho những ứng dụng khai thác các tính chất bên trong của đối tượng ví dụ như vân ảnh hoặc cấu trúc xương của ảnh.

Bước 4: Biểu diễn và mô tả. Bước này đề cập đến sự rút trích từ ảnh những đặc trưng cần thiết dẫn đến sự hình thành các thông tin định lượng giúp chúng ta có thể phân biệt các lớp đối tượng khác nhau trong ảnh.

Bước 5: Nhận dạng và giải thích. Nhận dạng là công đoạn gán nhãn cho đối tượng dựa trên thông tin do bộ mô tả của đối tượng đó cung cấp. Giải thích là công đoạn gán nghĩa cho 1 tập các đối tượng đã được nhận biết.

Cơ sở tri thức: Tri thức được đề cập đến có thể chỉ đơn giản là sự chi tiết hóa các vùng ảnh, nơi được biết trước là sẽ có những thông tin đáng quan tâm để tìm ra lời giải cho bài toán. Ngoài mục đích hướng dẫn cách thức làm việc phù hợp cho mỗi bước xử lý ảnh, nó còn giúp điều khiển môi tương tác giữa các bước xử lý với nhau.

1.4 Một số ứng dụng cơ bản

Kỹ thuật xử lý ảnh trước đây chủ yếu được sử dụng để nâng cao chất lượng hình ảnh, chính xác hơn là tạo cảm giác về sự gia tăng chất lượng ảnh quang học trong mắt người quan sát. Thời gian gần đây, phạm vi ứng dụng xử lý ảnh mở rộng không ngừng, có thể nói hiện không có lĩnh vực khoa học nào không sử dụng các thành tựu của công nghệ xử lý ảnh số.

Trong y học các thuật toán xử lý ảnh cho phép biến đổi hình ảnh được tạo ra từ nguồn bức xạ X-ray hay nguồn bức xạ siêu âm thành hình ảnh quang học trên bề mặt film x-quang hoặc trực tiếp trên bề mặt màn hình hiển thị. Hình ảnh các cơ quan chức năng của con người sau đó có thể được xử lý tiếp để nâng cao độ tương phản, lọc, tách các thành phần cần thiết (chụp cắt lớp) hoặc tạo ra hình ảnh trong không gian ba chiều (siêu âm 3 chiều).

Trong lĩnh vực địa chất, hình ảnh nhận được từ vệ tinh có thể được phân tích để xác định cấu trúc bề mặt trái đất. Kỹ thuật làm nổi đường biên (image enhancement) và khôi phục hình ảnh (image restoration) cho phép nâng cao chất lượng ảnh vệ tinh và tạo ra các bản đồ địa hình 3-D với độ chính xác cao.

Trong ngành khí tượng học, ảnh nhận được từ hệ thống vệ tinh theo dõi thời tiết cũng được xử lý, nâng cao chất lượng và ghép hình để tạo ra ảnh bề mặt trái đất trên một vùng rộng lớn, qua đó có thể thực hiện việc dự báo thời tiết một cách chính xác hơn. Dựa trên các kết quả phân tích ảnh vệ tinh tại các khu vực đông dân cư còn có thể dự đoán quá trình tăng trưởng dân số, tốc độ ô nhiễm môi trường cũng như các yếu tố ảnh hưởng tới môi trường sinh thái.

Xử lý ảnh được sử dụng nhiều trong các hệ thống quản lý chất lượng và số lượng hàng hóa trong các dây chuyền tự động, ví dụ như hệ thống phân tích ảnh để phát hiện bọt khí bên vật thể đúc bằng nhựa, phát hiện các linh kiện không đạt tiêu chuẩn (bị biến dạng) trong quá trình sản xuất hoặc hệ thống đếm sản phẩm thông qua hình ảnh nhận được từ camera quan sát.

Xử lý ảnh còn được sử dụng rộng rãi trong lĩnh vực hình sự và các hệ thống bảo mật hoặc kiểm soát truy cập: quá trình xử lý ảnh với mục đích nhận dạng vân tay hay khuôn mặt cho phép phát hiện nhanh các đối tượng nghi vấn cũng như nâng cao hiệu quả hệ thống bảo mật cá nhân cũng như kiểm soát ra vào. Ngoài ra, có thể kể đến các ứng dụng quan trọng khác của kỹ thuật xử lý ảnh tĩnh cũng như ảnh động trong đời sống như tự động nhận dạng, nhận dạng mục tiêu quân sự, máy nhìn công nghiệp trong các hệ thống điều khiển tự động, nén ảnh tĩnh, ảnh động để lưu và truyền trong mạng viễn thông v. v...

CHƯƠNG 2: TỔNG QUAN VỀ PHÂN ĐOẠN ẢNH

2.1 Khái niệm phân đoạn ảnh

Phân đoạn ảnh là một vấn đề quan trọng trong thị giác máy. Nói một cách dễ hiểu, phân đoạn ảnh có nghĩa là chia một ảnh đầu vào thành các vùng có cùng tính chất hay còn gọi là các đối tượng.

2.2 Các hướng tiếp cận trong phân đoạn ảnh

Phân đoạn ảnh là chia ảnh thành các vùng không trùng lặp. Mỗi vùng gồm một nhóm pixel liên thông và đồng nhất theo một tiêu chí nào đó[1]. Tiêu chí này phụ thuộc vào mục tiêu của quá trình phân đoạn. Ví dụ như đồng nhất về màu sắc, mức xám, kết cấu, độ sâu của các layer... Sau khi phân đoạn mỗi pixel chỉ thuộc về một vùng duy nhất. Để đánh giá chất lượng của quá trình phân đoạn là rất khó. Vì vậy trước khi phân đoạn ảnh cần xác định rõ mục tiêu của quá trình phân đoạn là gì. Xét một cách tổng quát, ta có thể chia các hướng tiếp cận phân đoạn ảnh thành ba nhóm chính như sau:

- Phân đoạn dựa vào ngưỡng.
- Phân đoạn dựa theo đường biên.
- Phân đoạn dựa theo miền đồng nhất.

2.2.1 Phân đoạn dựa vào ngưỡng

2.2.1.1 Giới thiệu chung

Biên độ của các tính chất vật lý của ảnh (như là độ phản xạ, độ truyền sáng, màu sắc ...) là một đặc tính đơn giản và rất hữu ích. Nếu biên độ đủ lớn đặc trưng cho ảnh thì chúng ta có thể dùng ngưỡng biên độ để phân đoạn ảnh. Thí dụ, biên độ trong bộ cảm biến hồng ngoại có thể phản ánh vùng có nhiệt độ thấp hay vùng có nhiệt độ cao. Đặc biệt, kỹ thuật phân ngưỡng theo biên độ rất có ích đối với ảnh nhị phân như văn bản in, đồ họa, ảnh màu hay ảnh X-quang.

Việc chọn ngưỡng trong kỹ thuật này là một bước vô cùng quan trọng, thông thường người ta tiến hành theo các bước chung như sau:

- Xem xét lược đồ xám của ảnh để xác định và khe. Nếu ảnh có nhiều đỉnh và khe thì các khe có thể sử dụng để chọn ngưỡng.
- Chọn ngưỡng T sao cho một phần xác định trước η của toàn bộ số mẫu là thấp hơn T.
- Điều chỉnh ngưỡng dựa trên lược đồ xám của các điểm lân cận.
- Chọn ngưỡng bằng cách xem xét lược đồ xám của những điểm thoả tiêu chuẩn đã chọn.

Một thuật toán đơn giản trong kỹ thuật này là: giả sử rằng chúng ta đang quan tâm đến các đối tượng sáng (object) trên nền tối (background), một tham số T - gọi là ngưỡng độ sáng, sẽ được chọn cho một ảnh $f[x,y]$ theo cách:

$$\text{If } f[x,y] \geq T \text{ } f[x,y] = \text{object} = 1$$

$$\text{Else } f[x,y] = \text{Background} = 0.$$

Ngược lại, đối với các đối tượng tối trên nền sáng chúng ta có thuật toán sau:

$$\text{If } f[x,y] < T \text{ } f[x,y] = \text{object} = 1$$

$$\text{Else } f[x,y] = \text{Background} = 0.$$

Vấn đề chính là chúng ta nên chọn ngưỡng T như thế nào để việc phân vùng đạt được kết quả cao nhất. Có rất nhiều thuật toán chọn ngưỡng: ngưỡng cố định, dựa trên lược đồ, sử dụng Entropy, sử dụng tập mờ, chọn ngưỡng thông qua sự không ổn định của lớp và tính thuần nhất của vùng vv... Ở đây chúng tôi đề cập đến hai thuật toán chọn ngưỡng đó là chọn ngưỡng cố định và chọn ngưỡng dựa trên lược đồ.

2.2.1.2 Chọn ngưỡng cố định

Đây là phương pháp chọn ngưỡng độc lập với dữ liệu ảnh. Nếu chúng ta biết trước là chương trình ứng dụng sẽ làm việc với các ảnh có độ tương phản rất cao, trong đó các đối tượng quan tâm rất tối còn nền gần như là đồng nhất và rất sáng thì việc chọn ngưỡng $T = 128$ (xét trên thang độ sáng từ 0 đến 255) là một giá trị chọn khá chính xác. Chính xác ở đây hiểu theo nghĩa là số các điểm ảnh bị phân lớp sai là cực tiểu.

2.2.1.3 Chọn ngưỡng dựa trên lược đồ

Trong hầu hết các trường hợp, ngưỡng được chọn từ lược đồ độ sáng của vùng hay ảnh cần phân đoạn. Có rất nhiều kỹ thuật chọn ngưỡng tự động xuất phát từ lược đồ xám $\{h[b] \mid b = 0, 1, \dots, 2^B - 1\}$ đã được đưa ra. Những kỹ thuật phổ biến sẽ được trình bày dưới đây. Những kỹ thuật này có thể tận dụng những lợi thế do sự làm trơn dữ liệu lược đồ ban đầu mang lại nhằm loại bỏ những dao động nhỏ về độ sáng. Tuy nhiên các thuật toán làm trơn cần phải cẩn thận, không được làm dịch chuyển các vị trí đỉnh của lược đồ. Nhận xét này dẫn đến thuật toán làm trơn dưới đây:

$$h_{smooth}[b] = \frac{1}{W} \sum_{w=-(W-1)/2}^{(W-1)/2} h_{raw}[b-w] \quad W \text{ odd} \quad (2.1)$$

Trong đó, W thường được chọn là 3 hoặc 5.

Chọn ngưỡng dựa theo lược đồ có các thuật toán như:

- Thuật toán đẳng liệu.
- Thuật toán đối xứng nền.
- Thuật toán tam giác.

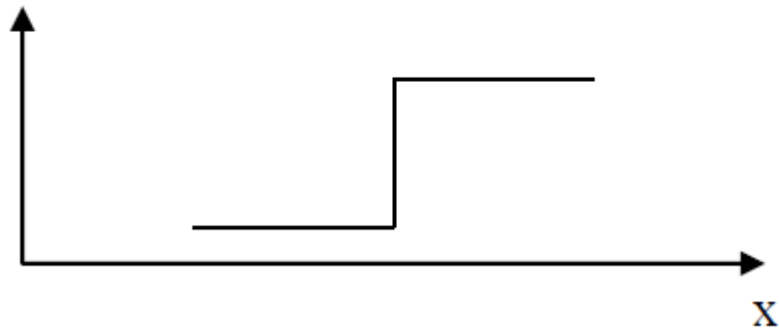
2.2.2 Phân đoạn dựa theo đường biên

2.2.2.1 Giới thiệu chung

Như chúng ta đã biết, Biên là một đặc tính rất quan trọng để phân vùng các đối tượng. Có thể hình dung tầm qua trọng của biên thông qua ví dụ sau: Khi một người họa sĩ vẽ một cái bàn gỗ, chỉ cần phác thảo vài nét về hình dáng như cái mặt bàn, cái chân bàn mà không cần thêm các chi tiết khác, người xem đã có thể nhận ra đó là cái bàn. Vài nét phác thảo của người họa sĩ chính là đường biên bao quanh đối tượng. Nếu ứng dụng của ta là phân lớp nhận diện các đối tượng thì coi như nhiệm vụ đã hoàn thành. Tuy nhiên, nếu đòi hỏi thêm các chi tiết khác như vân gỗ, màu sắc, kích thước vv ... thì chừng ấy thông tin là chưa đầy đủ.

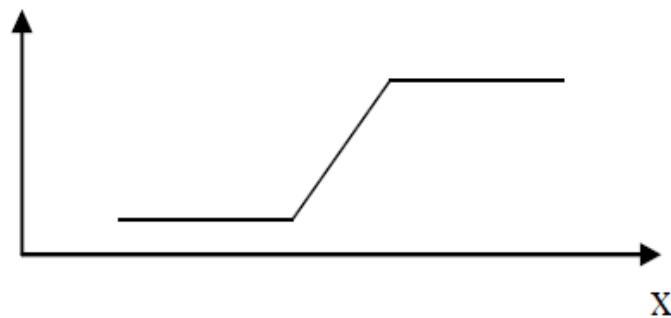
Trong toán học, người ta đưa ra khái niệm đường biên lý tưởng như sau: Đường biên lý tưởng là sự thay đổi giá trị cấp xám tại một vị trí xác định. Vị trí của đường biên chính là vị trí thay đổi cấp xám. Thể hiện của định nghĩa là hình 2.1

Mức xám

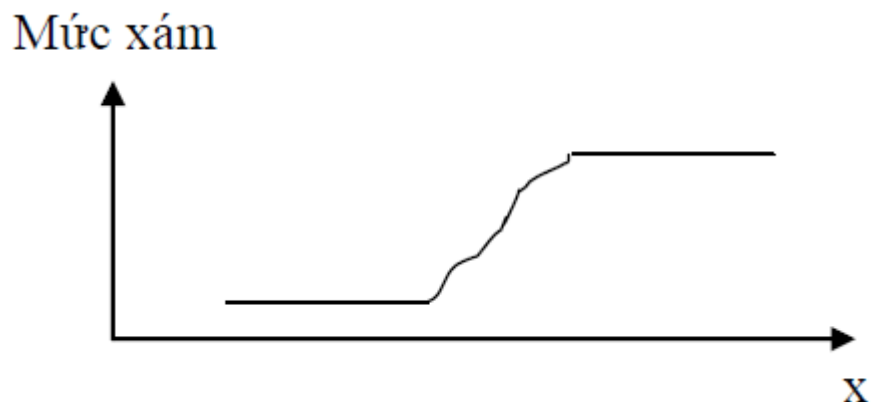
*Hình 2.1. Đường biên lý tưởng*

Một loại đường biên nữa - được gọi là đường biên bậc thang: Đường biên bậc thang xuất hiện khi sự thay đổi cấp xám trải rộng qua nhiều điểm ảnh. Vị trí của đường biên được xem như vị trí chính giữa của đường nối giữa cấp xám thấp và cấp xám cao.

Mức xám

*Hình 2.2. Đường biên bậc thang*

Trong thực tế đường biên của chúng ta thường có dạng như sau:



Hình 2.3. Đường biên thực

Như đã nói ở trên, biên là một trong những đặc trưng quan trọng của ảnh, chính vì vậy mà trong nhiều ứng dụng người ta sử dụng cách phân đoạn dựa theo biên. Việc phân đoạn ảnh dựa vào biên được tiến hành qua các bước:

- Phát hiện biên và làm nổi biên
- Làm mảnh biên
- Nhị phân hoá đường biên
- Mô tả biên

2.2.2.2 Phát hiện biên

Phát hiện biên một cách lý tưởng là xác định được tất cả các đường bao trong các đối tượng. Có nhiều phương pháp phát hiện biên, thông thường chúng ta sử dụng phương pháp phát hiện biên trực tiếp. Phương pháp này nhằm làm nổi biên dựa vào sự biến thiên về giá trị độ sáng của điểm ảnh. Kỹ thuật chủ yếu dùng ở đây là kỹ thuật đạo hàm. Nếu lấy đạo hàm bậc nhất của ảnh ta có phương pháp Gradient, nếu lấy đạo hàm bậc hai ta có kỹ thuật Laplace. Phương pháp này có ưu điểm là ít chịu ảnh hưởng của nhiễu, song nếu sự biến thiên của độ sáng không đột ngột thì hiệu quả đạt được là rất kém.

Một số kỹ thuật sử dụng trong phát hiện biên:

- Kỹ thuật Gradient
- Kỹ thuật Laplace
- Kỹ thuật la bàn

2.2.2.3 Làm mảnh biên

Làm mảnh biên thực chất là làm nổi biên với độ rộng chỉ 1 pixel. Chúng ta cũng đã biết rằng chỉ có kỹ thuật Laplace mới cho biên có độ rộng 1 pixel trong khi các kỹ thuật khác thì không hoàn toàn như thế. Vấn đề đặt ra là sau khi thu được bản đồ biên của ảnh chúng ta cần phải làm mảnh biên.

Một số kỹ thuật làm mảnh biên:

- Kỹ thuật loại bỏ các điểm không cực đại.
- Kỹ thuật do Sherman đề xuất.

2.2.2.4 Nhị phân hóa đường biên

Nhị phân hóa đường biên là giai đoạn then chốt trong quá trình trích chọn vì nó xác định đường bao nào thực sự cần và đường bao nào có thể loại bỏ. Nói chung, người ta thường nhị phân hóa đường biên theo cách thức làm giảm nhiễu hoặc tránh hiện tượng kéo sợi trên ảnh. Điều này cũng giải thích tại sao phân đoạn dựa theo biên có hiệu quả khi ảnh có độ tương phản tốt. Trong trường hợp ngược lại, có thể sẽ bị mất một phần đường bao hay đường bao có chân, không khép kín, v.v..., do đó sẽ bất lợi cho biểu diễn sau này. Một phương pháp hay được dùng là chọn ngưỡng thích nghi. Với cách chọn này, ngưỡng sẽ phụ thuộc vào hướng của gradient nhằm làm giảm sự xoắn của biên. Đầu tiên, người ta định ra một ngưỡng nào đó và sau đó sử dụng một hệ số sinh thích nghi thông.

2.2.2.5 Mô tả biên

Khi đã có bản đồ biên ảnh, ta cần phải biểu diễn nó dưới dạng thích hợp phục vụ cho việc phân tích và làm giảm lượng thông tin dùng để miêu tả, lưu trữ đối tượng. Người ta thường thực hiện theo nguyên tắc: tách riêng từng biên và gán cho mỗi biên một mã.

Có rất nhiều phương pháp miêu tả biên, mỗi phương pháp thích hợp với một loại ứng dụng riêng. Tuy nhiên, nhìn chung các biên sẽ được làm rõ hơn thông qua các thao tác: loại bỏ đường biên hở, khép kín đường biên, loại bỏ các chân rết bám theo đường biên v.v...

Thông thường, các cấu trúc cơ sở mã hoá đường biên bao gồm 4 loại: điểm, đoạn thẳng, cung và đường cong. Tuy nhiên, nếu ta biểu diễn đường biên bởi các điểm

thì rất đơn giản về mặt tính toán nhưng lại nghèo nàn về mặt cấu trúc và không cô đọng. Ngược lại, nếu biểu diễn biên bởi đường cong đa thức bậc cao thì cấu trúc dữ liệu rất cô đọng nhưng độ phức tạp tính toán lại khá lớn. Do đó, tùy từng loại ứng dụng cụ thể và từng bài toán cụ thể mà chúng ta có thể chọn cách mã hoá đường biên theo kiểu nào.

Một số phương pháp mã hoá đường biên hay dùng:

- Mã hóa theo tọa độ Đê-các.
- Mã hóa Freeman.
- Xấp xỉ bởi đoạn thẳng.

2.2.3 Phân đoạn theo miền đồng nhất

2.2.3.1 Giới thiệu

Giả sử rằng một miền ảnh X phải được phân thành N vùng khác nhau: R_1, \dots, R_N và nguyên tắc phân đoạn là một vị từ của công thức $P(R)$. Việc phân đoạn ảnh chia tập X thành các tập con $R_i, i = 1..N$ phải thoả mãn:

- Các vùng $R_i, i=1..N$ phải lấp kín hoàn toàn ảnh:

$$X = \bigcup_{i=1}^N R_i \quad (2.2)$$

- Hai vùng khác nhau phải là những tập hợp rời nhau:

$$R_i \cap R_j = \emptyset \text{ với } i \neq j \quad (2.3)$$

- Mỗi vùng R_i phải có tính đồng nhất:

$$P(R_i) = \text{TRUE} \text{ với } i = 1..N \quad (2.4)$$

- Nếu R_i, R_j là hai vùng rời nhau thì $(R_i \cup R_j)$ phải là một vùng ảnh không đồng nhất:

$$P(R_i \cup R_j) = \text{FALSE} \text{ với } i \neq j \quad (2.5)$$

Kết quả của việc phân vùng ảnh phụ thuộc vào dạng của vị từ P và các đặc trưng được biểu diễn bởi vectơ đặc trưng. Thường thì vị từ P có dạng $P(R, X, t)$, trong đó X là vectơ đặc trưng gắn với một điểm ảnh và t là một tập hợp các tham số (thường là các ngưỡng). Trong trường hợp đơn giản nhất, vectơ đặc trưng X chỉ chứa giá trị

mức xám của ảnh $I(k,l)$ và vector ngưỡng chỉ gồm một ngưỡng T . Một nguyên tắc phân đoạn đơn giản có công thức:

$$P(R): f(k,l) < T \quad (2.6)$$

Trong trường hợp các ảnh màu, vector đặc trưng X có thể là ba thành phần ảnh RGB $[fR(k,l), fG(k,l), fB(k,l)]T$. Lúc đó luật phân ngưỡng có dạng:

$$P(R,x,t): ((fR(k,l)<TR)\&\& (fG(k,l)<TR)\&\&(fB(k,l)<TR)) \quad (2.7)$$

2.2.3.2 Một số phương pháp phân đoạn ảnh theo miền đồng nhất

- Phương pháp tách cây tứ phân.
- Phương pháp phân vùng bởi hợp.
- Phương pháp tổng hợp.

CHƯƠNG 3: PHÂN ĐOẠN ẢNH DỰA TRÊN RWR

3.1 Giới thiệu

Phân đoạn ảnh là một vấn đề quan trọng trong thị giác máy tính. Đặc biệt, phân đoạn ảnh tự nhiên là một trong những vấn đề khó khăn nhất. Hai khó khăn quan trọng của phân đoạn trong ảnh tự nhiên là bài toán đường biên yếu và bài toán kết cấu yếu. Bài toán đầu tiên là tìm đường biên yếu khi chúng là những phần đường biên thích hợp. Bài toán thứ hai là phân tách vùng kết cấu trong ảnh có tính phức tạp. Trong thực tế, những khó khăn này thường phát sinh trong ảnh tự nhiên. Trong những trường hợp này, các vùng thường rất nhập nhằng, và do đó phương pháp tiếp cận phân đoạn ảnh có giám sát thường được ưu tiên.

Gần đây, một số phương pháp tiếp cận phân đoạn ảnh có giám sát đã được đưa ra. Có ba loại thuật toán phân đoạn ảnh giám sát theo đầu vào người sử dụng:

- Loại 1: Phân đoạn thu được dựa trên những mảnh biên được yêu cầu, chẳng hạn như intelligent scissors.
- Loại 2: Đưa ra một biên ban đầu sát với biên được yêu cầu, chẳng hạn như Active Contour và Set Level.
- Loại 3: người sử dụng cung cấp một nhãn ban đầu của một số điểm ảnh.

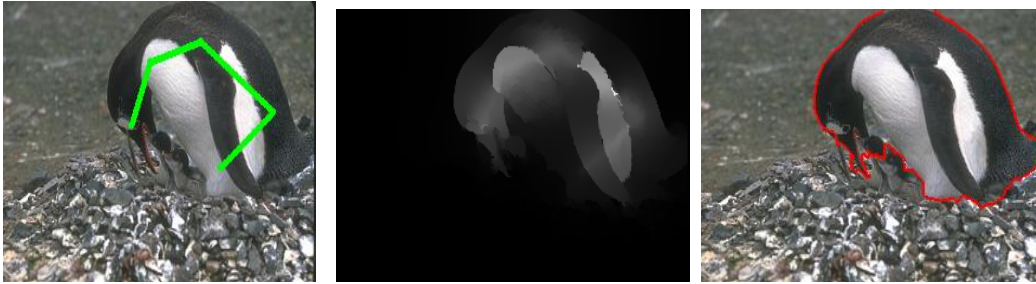
Một trong các phương pháp tiếp cận phổ biến là phương pháp đồ thị cắt (GC) dựa trên hàm năng lượng được giảm thiểu thông qua kỹ thuật tối ưu hóa rời rạc. Thiết lập của các cạnh có tổng trọng số nhỏ nhất thu được thông qua dòng tối đa (maxflow) / cắt giảm năng lượng tối thiểu (min-cut). Kể từ khi GC xử lý tiêu chí cắt giảm tối thiểu này, nó thường gây ra vấn đề cắt nhỏ khi tương phản thấp hoặc số lượng seed của điểm ảnh là nhỏ.

Khoảng cách đo đạc từ seed được sử dụng cho phân đoạn ảnh. Bằng cách gán cho mỗi điểm ảnh 1 nhãn với khoảng cách tối thiểu, ta thu được phân đoạn. Khoảng cách giữa hai điểm ảnh được định nghĩa là tách rời nhỏ nhất của một thành phần trọng số trên tất cả các đường dẫn. Tuy nhiên, vì nó không xem xét các mối quan hệ bao trùm giữa hai điểm ảnh, nó không đáng tin cậy để sử dụng khoảng cách đo đạc đơn giản như là sự đo lường chính xác giữa 2 điểm ảnh.

Một cách khác là thuật toán phân đoạn ảnh RandomWalker (RW) đề xuất bởi Grady. Sau khi xác suất tiên nghiệm RW bắt đầu từ một điểm ảnh đầu tiên đạt đến một trong những hạt giống với mỗi nhãn được tính, điểm ảnh đó được gán nhãn với xác suất tối đa. Nó thể hiện rằng RW có hiệu suất tốt hơn GC trong điều kiện khó khăn. Tuy nhiên, xác suất tiên nghiệm có một số hạn chế. Vì 1 RW bắt đầu từ một điểm ảnh đầu tiên phải đến khu vực biên được gán trước, nó chỉ xem xét mối quan hệ giữa các điểm ảnh và biên đó. Do đó, các thông tin của các hạt giống bên trong khu vực gán trước được bỏ qua mà không có tương tác cao hơn. Ngoài ra, xác suất này phụ thuộc vào số lượng seed. Nếu seed chỉ với 1 nhãn phát triển dưới vấn đề biên ảnh yếu, xác suất tiên nghiệm của nhãn đó được tăng lên mà không có liên quan đến toàn bộ mối quan hệ giữa một điểm ảnh và seed. Những hạn chế này giải thích lý do tại sao RW vẫn bị hai vấn đề: vấn đề biên ảnh yếu và vấn đề kết cấu yếu. Gần đây nhất, cách tiếp cận phân đoạn được định nghĩa bởi một định mức l_∞ . Vì RW với các ràng buộc đã được sử dụng như một phương pháp qui tắc cho năng suất một giải pháp duy nhất, phương pháp này vẫn còn có những hạn chế của RW.

Hầu hết các thuật toán phân đoạn giám sát trước đều tập trung vào phân biệt giữa các nhãn, không tìm thấy mô hình sinh cho mỗi nhãn. Mặc dù họ đã cố gắng để giải quyết vấn đề biên ảnh yếu và vấn đề kết cấu yếu, hai vấn đề khó khăn nhất trong phân đoạn ảnh. Thuật toán phân đoạn ảnh RWR được đề xuất bởi nhóm tác giả thuộc trường đại học Quốc gia Seoul Hàn Quốc có thể giải quyết vấn đề nói trên. Đóng góp mang tính mâu chốt của thuật toán như sau:

- Thuật toán giới thiệu một mô hình sinh cho phân đoạn ảnh. Từ lý thuyết quyết định cơ bản, có thể biết rằng các phương pháp sinh là tốt hơn suy luận thuật toán. Ngược lại với mô hình hiện có trong đó tập trung trên đa nhãn phân biệt, thuật toán này giải quyết vấn đề của việc tìm kiếm mô hình sinh cho mỗi nhãn. Ví dụ, chúng ta có thể chỉ xem xét vấn đề phân đoạn đơn nhãn như trong hình 3.1. Thuật toán này có thể tạo ra kết quả phân đoạn với một mức ngưỡng tối ưu như trong hình 3.1 (c). Điều này là có thể vì xác suất likelihood có thể được tạo ra bằng cách sử dụng mô hình sinh như mô tả trong hình 3.1 (b). Vì mô hình sinh của mỗi nhãn được xây dựng độc lập, nó cũng có thể thêm một nhãn mới mà không cần thay đổi các mô hình của các nhãn trước.



(a) ảnh gốc (b) khả năng xảy ra (c) Kết quả phân đoạn

Hình 3.1 Phân đoạn đơn nhân

- Thuật toán phân đoạn ảnh sử dụng xác suất trạng thái ổn định của RWR như là một phần của khả năng có thể xảy ra. Vì khả năng của một điểm ảnh được định nghĩa là mức trung bình của tất cả các xác suất trạng thái ổn định giữa điểm ảnh đó và seed với cùng một nhãn, thuật toán này có thể làm giảm sự phụ thuộc vào số lượng hạt giống trong vấn đề biên ảnh yếu. RWR tương tự như đồ thị học bán giám sát, là một kỹ thuật rất thành công để xác định mối quan hệ liên quan giữa hai nút trong đồ thị khai thác. Nó có hiệu suất tốt trên nhiều ứng dụng khác: phát hiện tương quan mô hình chéo, phát hiện đồ thị con phân mảnh trung tâm, tra cứu ảnh dựa trên nội dung, xây dựng vùng lân cận, vv. Vì xác suất trạng thái ổn định này của RWR xem như mối quan hệ toàn bộ giữa hai điểm ảnh, nó phản ánh những tác động của kết cấu một cách tự nhiên.
- Trong hai vấn đề khó khăn là biên ảnh yếu và kết cấu yếu, thuật toán mới này cho các kết quả phân đoạn rất tốt trên ảnh tự nhiên.

3.2 Random Walker Restart (RWR)

Một vấn đề đặt ra là làm thế nào có thể liên kết chặt chẽ hai nút trong một đồ thị? Làm thế nào để tính toán số điểm này một cách nhanh chóng trên đồ thị thực? Và Random Walk with restart (RWR) được đưa ra là giải pháp nhanh chóng cho vấn đề nói trên. RWR cung cấp một số điểm liên quan giữa hai nút trong một đồ thị có trọng số, và nó đã được sử dụng thành công trong nhiều cài đặt như: phụ đề tự động của hình ảnh, khái quát kết nối đồ thị con,...RWR khai thác hai thuộc tính quan trọng được chia sẻ bởi nhiều đồ thị thực đó là: mối tương quan tuyến tính và khối thông minh, cấu trúc liên kết chặt chẽ.

RWR đòi hỏi phải có ma trận nghịch đảo. Có hai giải pháp đơn giản, nhưng không thể áp dụng cho các đồ thị lớn: đầu tiên là tính toán trước và lưu trữ các ma trận nghịch đảo (phương pháp PreCompute). Thứ hai là tính toán nghịch đảo ma trận (phương pháp OnTheFly). Phương pháp đầu tiên thì thời gian truy vấn nhanh nhưng không gian hạn chế (về số lượng các nút trên đồ thị), trong khi phương pháp thứ hai thì chậm về thời gian truy vấn. Một giải pháp mới, nhanh chóng, và thực tế được đưa ra đó là B_LIN có lợi thế của hai thuộc tính được chia sẻ bởi nhiều đồ thị thực: khối thông minh, cấu trúc liên kết chặt chẽ và tương quan tuyến tính giữa các hàng và cột của ma trận kề.

RWR được xác định dựa vào phương trình sau:

$$\vec{r}_i = c\mathbf{W}\vec{r}_i + (1 - c)\vec{e}_i \quad (3.1)$$

Xem xét một phân tử ngẫu nhiên bắt đầu từ node i . Nó truyền lặp đi lặp lại đến các node bên với xác suất tỷ lệ thuận với trọng số của cạnh. Cũng tại mỗi bước, nó có xác suất c là khả năng trở lại node i . Kết quả liên quan của node j với node i được định nghĩa là trạng thái ổn định xác suất $r_{i,j}$ mà phân tử đó cuối cùng sẽ ở lại tại node j . Phương trình (3.1) định nghĩa một vấn đề về hệ thống tuyến tính mà \vec{r}_i được xác định bởi phương trình (3.2):

$$\begin{aligned} \vec{r}_i &= (1 - c)(I - c\mathbf{W})^{-1}\vec{e}_i \\ &= (1 - c)\mathbf{Q}^{-1}\vec{e}_i \end{aligned} \quad (3.2)$$

Kết quả liên quan được xác định bởi RWR có nhiều thuộc tính tốt: nắm bắt toàn bộ cấu trúc của đồ thị, so sánh với những khoảng cách đồ thị truyền thống (tìm đường đi ngắn nhất,...), nắm bắt các khía cạnh trong mối quan hệ giữa hai nút.

Một trong những cách được sử dụng rộng rãi nhất để giải quyết RWR là phương pháp lặp, lặp lại phương trình (3.1) cho đến khi hội tụ, nghĩa là cho đến khi chỉ tiêu kế tiếp L2 ước tính của \vec{r}_i là dưới ngưỡng ξ , hoặc vòng lặp tối đa đạt được m bước. Mặt khác, có thể được nhìn thấy từ phương trình (3.2), ma trận hệ thống \mathbf{Q} xác định tất cả xác suất trạng thái ổn định của RWR. Vì vậy, nếu chúng ta có thể tính toán

trước và lưu trữ Q^{-1} , chúng ta có thể nhận được \vec{r}_i thời gian thực. Tuy nhiên, tính toán và lưu trữ Q^{-1} là không thực tế khi các tập dữ liệu lớn. Mặt khác, mối tương quan tuyến tính tồn tại trong nhiều đồ thị thực, có nghĩa là chúng ta có thể tính gần đúng \tilde{W} . Tính chất này cho phép chúng ta tính gần đúng Q^{-1} rất hiệu quả. Thuật toán được đưa ra đó là B_LIN:

$$\tilde{W}_1 = \begin{pmatrix} \tilde{W}_{1,1} & 0 & \dots & 0 \\ 0 & \tilde{W}_{1,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \tilde{W}_{1,k} \end{pmatrix} \quad (3.3)$$

$$Q_1^{-1} = \begin{pmatrix} Q_{1,1}^{-1} & 0 & \dots & 0 \\ 0 & Q_{1,2}^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & Q_{1,k}^{-1} \end{pmatrix} \quad (3.4)$$

Đầu vào: Ma trận trọng số bình thường \tilde{W} và vector bắt đầu \vec{e}_i

Đầu ra: vector xếp hạng \vec{r}_i

Giai đoạn tính toán trước:

p1. Phân vùng đồ thị vào phân vùng k bởi METIS.

p2. $\tilde{W} = \tilde{W}_1 + \tilde{W}_2$ theo kết quả phân vùng, nơi \tilde{W}_1 có chứa tất cả các liên kết trong phân vùng và \tilde{W}_2 chứa tất cả liên kết chéo phân vùng.

p3. Để $\tilde{W}_{1,i}$ là phân vùng thứ i, áp dụng \tilde{W}_1 vào phương trình (3.3).

p4. Tính toán và lưu trữ $Q_{1,i}^{-1} = (I - c\tilde{W}_{1,i})^{-1}$ cho mỗi phân vùng i.

p5. Tính xấp xỉ bậc thấp cho $\tilde{W}_2 = USV$.

p6. Xác định Q_1^{-1} như phương trình (3.4). Tính toán và lưu trữ

$$\tilde{\Lambda} = (S^{-1} - cVQ_1^{-1}U)^{-1}$$

Giai đoạn truy vấn:

$$q1.Đầu ra \vec{r}_i = (1 - c)(Q_1^{-1}\vec{e}_i + cQ_1^{-1}U\Lambda VQ_1^{-1})$$

Bảng 3.1. B_LIN

Kí hiệu	Định nghĩa
$W = [w_{i,j}]$	Đồ có thị trọng số, $1 \leq i, j \leq n$
\bar{W}	Ma trận trọng số bình thường liên quan đến W
\bar{W}_1	Ma trận phân vùng kết hợp với \bar{W}
\bar{W}_2	Ma trận chéo phân vùng liên quan với \bar{W}
Q	Hệ thống ma trận liên quan đến $W : Q=I - c\bar{W}$
U	Ma trận $n \times t$ node- khái niệm
S	Ma trận $t \times t$ khái niệm - khái niệm
V	Ma trận $t \times n$ node - khái niệm
0	Ma trận khối có tất cả các thành phần bằng 0
\vec{e}_i	Vector bắt đầu $n \times 1$, phần tử thứ $i = 1$, các phần tử còn lại bằng 0
$\vec{r}_i = [r_{i,j}]$	Vector xếp hạng $n \times 1$, $r_{i,j}$ là số điểm liên quan của node j với node i
c	Xác suất khởi động lại, $0 \leq c \leq 1$
n	Tổng số node trong đồ thị
k	Số lượng phân vùng
m	Số lần lặp tối đa
ξ_1	Ngưỡng dừng quá trình lặp đi lặp lại
ξ_2	Ngưỡng thừa thớt ma trận

Bảng 3.2. Một số ký hiệu

3.3 Phương pháp phân đoạn dựa trên RWR

Xem xét phân đoạn ảnh như là một vấn đề gán nhãn, trong đó mỗi điểm ảnh $x_i \in X = \{x_1, \dots, x_n\}$ được gán vào một nhãn $l_k \in L = \{l_1, \dots, l_k\}$. Từ lý thuyết quyết định cơ bản, chúng ta biết rằng sự mô tả hoàn thiện nhất của giải pháp được thể hiện trong các điều khoản của tập hợp các xác suất sau $p(l_k | x_i)$. Một khi chúng ta biết các xác suất này, không khó để gán x_i vào nhãn có xác suất lớn nhất. Trong tiếp cận di truyền, thuật toán mô hình hóa phân phối chung $p(l_k, x_i)$ của các điểm ảnh và nhãn. Điều này có thể được thực hiện bằng cách tính riêng xác suất tiên nghiệm nhãn $p(l_k)$ và khả năng $p(x_k | l_i)$. Các yêu cầu xác suất hậu nghiệm thu được bằng cách sử dụng các quy tắc Bayesian:

$$p(l_k | x_i) = \frac{p(x_i | l_k) p(l_k)}{\sum_{n=1}^K p(x_i | l_n) p(l_n)} \quad (3.5)$$

Trong đó, tổng trong các mẫu số được thực hiện trên tất cả các nhãn

$X^{l_k} = \{x_1^{l_k}, \dots, x_{M_k}^{l_k}\}$ ($X^{l_k} \subset X$) là 1 tập hợp của seed M_k với nhãn l_k . Sau đó,

các khả năng $p(x_i | l_k)$ có thể thu được bằng cách:

$$\begin{aligned} p(x_i | l_k) &= \frac{1}{Z} \sum_{m=1}^{M_k} p(x_i | x_m^{l_k}, l_k) p(x_m^{l_k} | l_k) \\ &= \frac{1}{Z \times M_k} \sum_{m=1}^{M_k} p(x_i | x_m^{l_k}, l_k), \end{aligned} \quad (3.6)$$

Trong đó Z là một hằng số bình thường hóa. Mỗi điểm ảnh được mô hình hóa bởi một phân phối $p(x_i | x_m^{l_k}, l_k)$ từ mỗi hạt giống $x_m^{l_k}$ mà có một $p(x_m^{l_k} | l_k)$. Sự phân bố điểm ảnh $p(x_i | x_m^{l_k}, l_k)$ cho thấy số điểm liên quan giữa điểm ảnh x_i và một seed $x_m^{l_k}$. So với khoảng cách đồ thị truyền thống (Chẳng hạn như đường ngắn nhất, maximum flow), trạng thái ổn định này có thể nắm bắt toàn bộ mối quan hệ giữa hai điểm ảnh. Sự phân bố seed $p(x_m^{l_k} | l_k)$ được định nghĩa bởi một phân bố đồng đều $1 / M_k$. Vì khả

năng $p(x_i | l_k)$ được tính = trung bình phân phối điểm ảnh của tất cả các hạt giống với các nhãn l_k , phương pháp này ít phụ thuộc vào số lượng seed.

Quá trình phân đoạn ảnh dựa trên RWR gồm các bước:

- Xây dựng đồ thị trọng số trong hình ảnh.
- Tính xác suất trạng thái ổn định $p(x_i | x_m^{l_k}, l_k)$ mà 1 RW bắt đầu tại một hạt giống $x_m^{l_k}$ mở trong 1 điểm ảnh x_i trong đồ thị.
- Ước tính khả năng $p(x_i | l_k)$.
- Gán nhãn với xác suất tối đa ở (1) vào điểm ảnh.

3.3.1 Mô hình đồ thị

Cho một ảnh I , hãy xây dựng một đồ thị vô hướng $G = (V, E)$ với nút $v \in V$, cạnh $e \in E$. Mỗi nút v_i trong V xác định duy nhất một điểm ảnh x_i . Cạnh E ở giữa hai nút được xác định bởi hệ thống láng giềng. Trọng số $w_{ij} \in W$ được gán cho cạnh $e_{ij} \in E$ kéo dài giữa các nút $v, v \in V$. Nó tính khả năng các nút lân cận v_i, v_j có cùng nhãn.

Sự thay đổi trọng số mã hóa màu ảnh được sử dụng trong nhiều thuật toán phân đoạn đồ thị. Ở đây, trọng số w_{ij} được định nghĩa là trọng số chức năng Gaussian điển hình đưa ra bởi:

$$w_{ij} = \exp\left(-\frac{\|g_i - g_j\|^2}{\sigma}\right) \quad (3.7)$$

Trong đó, g_i và g_j cho thấy màu ảnh ở hai nút v_i và v_j trong dải màu thí nghiệm. Nó cung cấp cho chúng ta một thông số đo lường, một số giữa 0 và 1, cho sự giống nhau giữa một cặp điểm ảnh. Chức năng gaussian có bản chất của đo đặc khoảng cách. Ví dụ, phép nhân giữa hai trọng số w_{ij}, w_{jk} :

$$w_{ij}w_{jk} = \exp\left(-\frac{\|g_i - g_j\|^2 + \|g_j - g_k\|^2}{\sigma}\right),$$

v_i, v_k . Tính chất này phù hợp với thuật toán này.

3.3.2 Học

Giả sử một RW bắt đầu từ seed thứ m điếm ảnh $x_m^{l_k}$ của nhãn l_k trong đồ thị G . RW truyền qua lại tới các cạnh bên với xác suất tỷ lệ thuận với trọng số cạnh giữa chúng. Cũng tại mỗi bước, có một xác suất khởi động lại c để trở về hạt giống $x_m^{l_k}$. Sau khi hội tụ, thu được xác suất khả năng $r_{im}^{l_k}$ mà RW sẽ ở lại cuối cùng tại điếm ảnh x_i . Ở đây, thuật toán sử dụng xác suất trạng thái như phân phối $p(x_i | x_m^{l_k}, l_k)$ ở (2):

$$p(x_i | x_m^{l_k}, l_k) \approx r_{im}^{l_k} \quad (3.8)$$

Bằng cách biểu thị $r_{im}^{l_k}, i = 1, \dots, N$ trong các điều khoản của một vector N chiều $r_{im}^{l_k} = [r_{im}^{l_k}]_{N \times 1}$ và xác định một ma trận kề $W = [w_{ij}]_{N \times N}$ sử dụng (3), RWR có thể được xây dựng như sau.

$$\begin{aligned} r_m^{l_k} &= (1 - c)\mathbf{P}r_m^{l_k} + c\mathbf{b}_m^{l_k} \\ &= c(\mathbf{I} - (1 - c)\mathbf{P})^{-1}\mathbf{b}_m^{l_k} \\ &= \mathbf{Q}\mathbf{b}_m^{l_k} \end{aligned} \quad (3.9)$$

Trong đó $\mathbf{b}_m^{l_k} = [b_i]_{N \times 1}$ là vector N -chiều với $b_i = 1$ nếu $x_i = x_m^{l_k}$ và $b_i = 0$, và chuyển đổi ma trận $\mathbf{P} = [p_{ij}]_{N \times N}$ là ma trận kề W chuẩn hóa hàng:

$$\mathbf{P} = \mathbf{D}^{-1} \times \mathbf{W}, \quad (3.10)$$

Trong đó $D = \text{diag}(D_1, \dots, D_N)$, $D_i = \sum_{j=1}^N w_{ij}$. Nếu những xác suất trạng thái $r_{im}^{l_k}$ này được chèn vào (3.2) theo định nghĩa, những khả năng $p(x_i | l_k)$ ($i = 1, \dots, N$) được tính bằng:

$$[p(x_i | l_k)]_{N \times 1} = \frac{1}{Z \times M_k} \mathbf{Q} \tilde{\mathbf{b}}^{l_k} \quad (3.11)$$

Trong đó $\tilde{\mathbf{b}}^{l_k} = [\tilde{b}_i] N \times 1$ là vector N chiều với $\tilde{b}_i = 1$ if $x_i \in \mathbf{X}^{l_k}$ và $\tilde{b}_i = 0$ nếu ngược lại.

Xét trong 1 RWR, $\mathbf{Q} = [q_{ij}] N \times N$ được sử dụng để tính toán số điểm giữa 2 pixel. Nói cách khác, q_{ij} là khả năng q_i có cùng một nhãn được gán cho q_j . Nó có thể được cải tiến công thức như sau:

$$\begin{aligned} \mathbf{Q} &= c(\mathbf{I} - (1 - c)\mathbf{P})^{-1} \\ &= c \sum_{t=0}^{\infty} (1 - c)^t \mathbf{P}^t \end{aligned} \quad (3.12)$$

\mathbf{Q} được định nghĩa là tổng trọng số của tất cả các ma trận \mathbf{P}^t , $t = 0, \dots, \infty$. Lưu ý rằng \mathbf{P}^t là ma trận chuyển đổi thứ t , mà các thành phần p_{ij}^t của nó có thể được hiểu là tổng xác suất cho RW bắt đầu tại v_i đến kết thúc tại v_j với t lặp đi lặp lại, xem xét tất cả các đường đi có thể có giữa hai điểm ảnh. Bằng cách thay đổi số lần của vòng lặp t , tác giả thấy rõ ràng mối quan hệ ở các quy mô khác nhau trong ảnh, và khi tăng t , tác giả hy vọng tìm ra cấu trúc thô (coarser). Vì vậy, RWR các tác động của kết cấu bằng cách xem xét tất cả các đường đi giữa hai điểm ảnh ở tất cả các quy mô (bất kỳ số vòng lặp ($t = 0, \dots, \infty$)) trong hình ảnh. Vì điểm ảnh kẻ có giá trị tương đồng cao, nếu tăng t , \mathbf{P}^t có trọng số thấp hơn, $(1-c)^t$ ($0 < c < 1$). Ma trận kết quả \mathbf{Q} có thể được giải quyết bằng cách sử dụng một phương pháp tuyến tính ma trận đảo ngược. Mặc dù nó đòi hỏi nhiều không gian bộ nhớ hơn, có thể tính nhanh nếu ma trận thưa. Vì hệ thống láng giềng các điểm ảnh nhỏ được sử dụng trong bài này, ma trận chuẩn hóa \mathbf{P} là rất thưa. Do đó, \mathbf{Q} có thể được tính nhanh. Tuy nhiên, nếu số lượng các lân cận gần nhất là chọn được một số lượng lớn cố định, sự phức tạp của ma trận đảo ngược là rất cao.

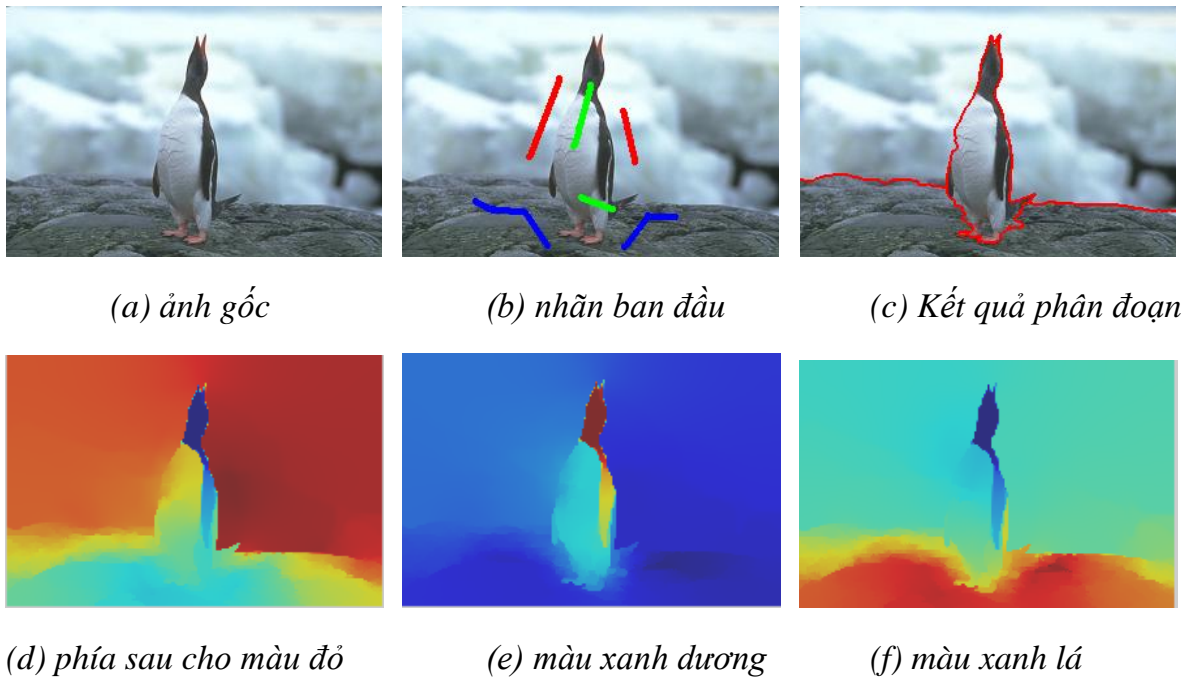
Trong trường hợp này, một số phương pháp tính xấp xỉ như phương pháp RWR nhanh được sử dụng.

3.3.3 Phân đoạn

Giả sử rằng xác suất trước $p(l_k)$ ở (1) là thống nhất. Sử dụng khả năng $p(x_i | l_k)$ ở (7), các quy tắc quyết định của mỗi điểm ảnh x_i cho phân đoạn ảnh như sau:

$$R_i = \arg \max_{l_k} p(l_k | x_i) = \arg \max_{l_k} p(x_i | l_k) \quad (3.13)$$

Gán nhãn R_i cho mỗi điểm ảnh x_i , thu được phân đoạn.



Hình 3.2. Kết quả phân đoạn

Hình 3.2 cho thấy quá trình tổng thể của thuật toán từ hạt giống để tính Xác suất hậu nghiệm $p(l_k | x_i)$ của mỗi nhãn và kết quả phân đoạn. Nó bắt đầu với ba nhãn hạt giống ban đầu: Red, Green và Blue như trong hình 3.2(b). Sau khi tính toán khả năng cho mỗi nhãn, chúng ta tạo ra xác suất sau như trong hình 3.2(d), (e) và (f). Bởi một quy tắc quyết định, mỗi điểm ảnh được gán nhãn có xác suất lớn nhất. Cuối cùng,

chúng ta có được kết quả phân đoạn trong hình 3.2(c), đối tượng biên ảnh được vẽ bằng màu đỏ phủ lên trên ảnh gốc.

CHƯƠNG 4: CÀI ĐẶT THỬ NGHIỆM

4.1 Môi trường cài đặt

Chương trình được cài đặt trên Môi trường Windows XP, sử dụng ngôn ngữ Matlab 2008 với máy tính có cấu hình như sau:

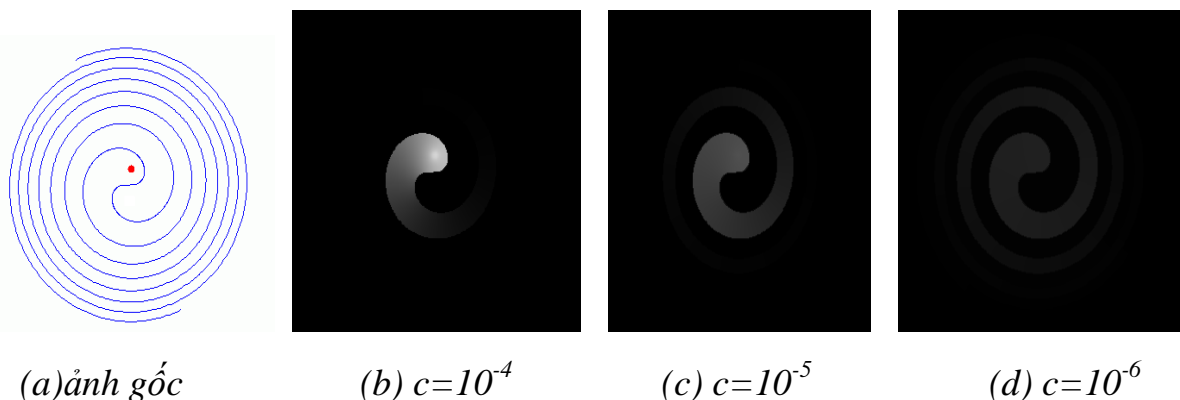
- CPU: Intel® Pentium E5500(2.8 GHz)
- HDD: 160 GB
- Memory: 2GB

4.2 Chương trình thực nghiệm

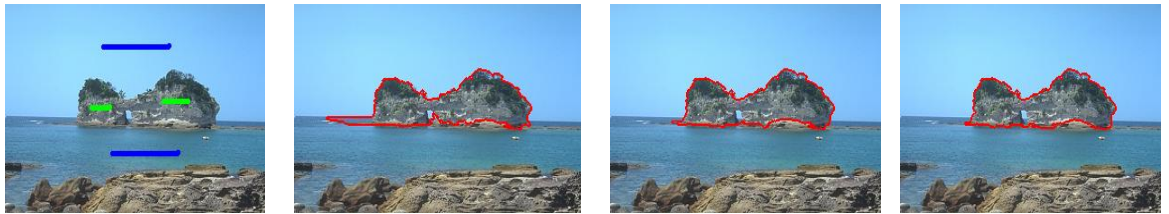
4.2.1 Kết quả phân đoạn ảnh sử dụng RWR

4.2.1.1 Thiết lập thông số

RWR cần một tham số, xác suất c . Theo thông số đó, phạm vi tuyên truyền của một random walker từ nút bắt đầu bị biến đổi như trong hình 4.1. Nếu c giảm, xác suất mà random walker đi qua một diện tích lớn hơn sẽ tăng lên. Điều này có nghĩa là bằng cách thay đổi c , chúng ta có thể kiểm soát mức độ thông tin nhận của một seed ở quy mô khác nhau trong ảnh. Hình 4.2 cho thấy một ví dụ khác của phân đoạn đối với sự biến đổi của xác suất c trong một ảnh thường. Theo xác suất c , kết quả phân đoạn bị thay đổi. Do đó, quan trọng là tìm xác suất c theo số lượng (hay chất lượng) của seeds và kích thước ảnh. Ở đây, c được lựa chọn theo kinh nghiệm, và thuật toán thiết lập $c = 4 \times 10^{-4}$ cho tất cả các hình thường được thử nghiệm.



Hình 4.1. Một ví dụ về sự thay đổi xác suất trạng thái ổn định r theo xác suất khởi động lại c

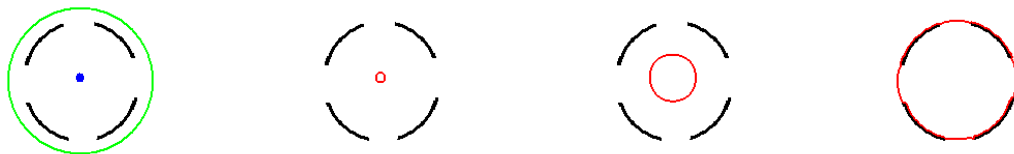


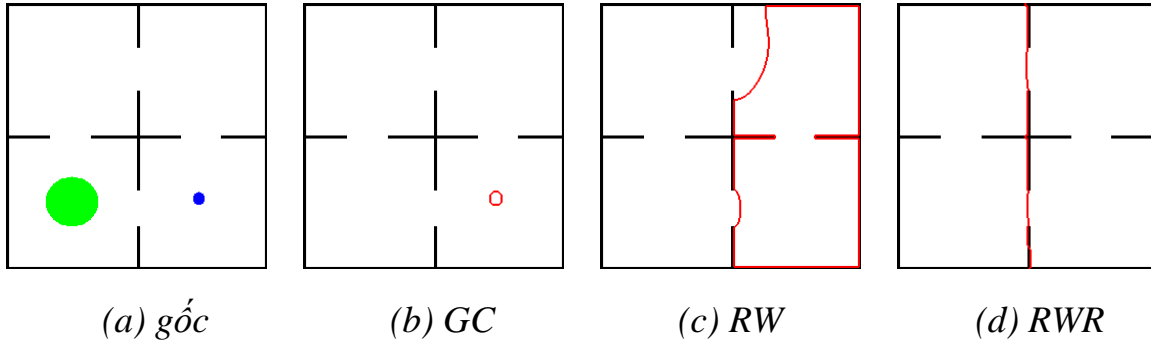
(a) Ảnh gốc $a_o=0.826165$ (b) $c = 10^{-3}$ $a_o=0.850946$ (c) $c = 10^{-4}$ $a_o=0.877542$ (d) $c = 10^{-5}$

Hình 4.2. Một ví dụ về phân đoạn đối với sự biến đổi của các xác suất khởi động lại c trong ảnh tự nhiên

4.2.2 Một số so sánh.

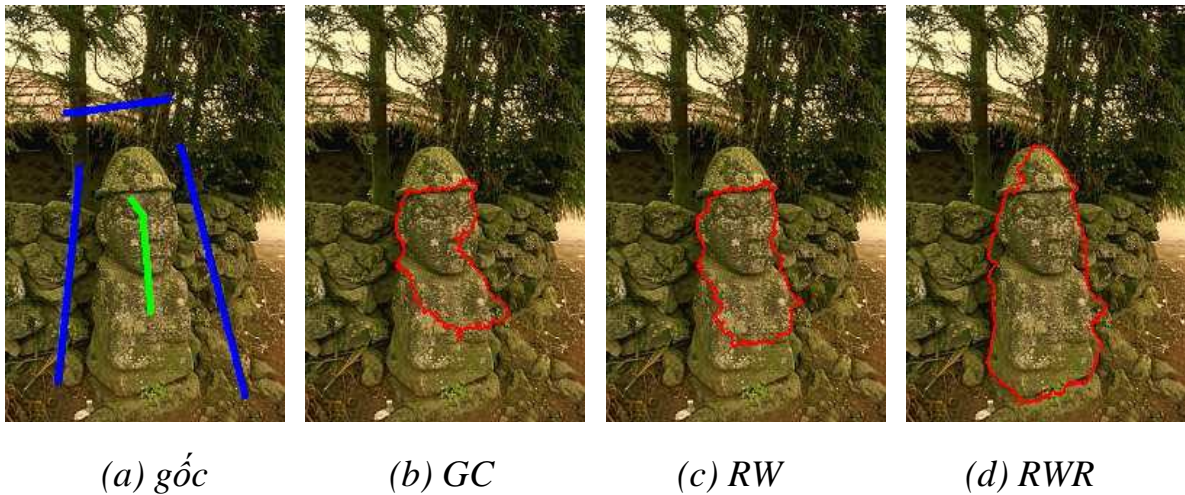
Vấn đề biên ảnh yếu: Các vấn đề biên ảnh yếu là để tìm thấy những biên ảnh yếu khi chúng là một phần của một biên ảnh phù hợp. RW cho thấy kết quả phân đoạn tốt hơn GC trong sự tương phản thấp với số lượng seed nhỏ. Mặc dù GC và RW có khả năng tìm kiếm biên ảnh yếu, thuật toán RWR cung cấp các kết quả đầu ra trực quan hơn. Trong hình 4.3, thuật toán RWR được so sánh với GC và RW trong vấn đề biên ảnh yếu. Hai ví dụ tổng hợp: vòng tròn và lưới 3×3 với bốn phần bị xóa. Với seeds (màu xanh lá cây và xanh dương) trong hình 4.3 (a), các phân đoạn đã thu được trong hình 4.3 (b) - (d). Hình 4.3 (b) cho thấy rõ ràng rằng GC có vấn đề cắt nhỏ. Trong hình 4.3 (c), chúng ta có thể xác định rằng các phân đoạn của RW bị ảnh hưởng đáng kể bởi sự khác biệt giữa số lượng của Green và Blue seeds. Cụ thể, RW nhạy cảm với số lượng seeds. Ngược lại, hình 4.3 (d) cho thấy rằng thuật toán của phân đoạn ảnh dựa trên RWR là ít phụ thuộc vào số lượng seeds và tạo ra phân đoạn tốt hơn, bởi vì khả năng được tính bằng trung bình của số lượng phù hợp của tất cả seeds.





Hình 4.3. So sánh thuật toán GC, RW, RWR cho việc tìm kiếm đường biên yếu

Vấn đề kết cấu: Trong GC và RW, rất khó để tách biệt khu vực kết cấu mà không cần xem xét các kết nối bậc cao hơn, bởi vì chúng có quan hệ với tiêu chuẩn cắt giảm tối thiểu và xác suất khởi động lại tương ứng. Vì hai thuật toán này không xem xét các thông tin của seeds bên trong khu vực nhãn trước, không dễ dàng để đưa vào tính toán các tác động của kết cấu. Mặt khác, thuật toán này có thể phản ánh các thông tin kết cấu bằng cách sử dụng xác suất trạng thái ổn định của RWR, bởi vì RWR xem xét tất cả các đường đi có thể giữa hai nút trong một hệ thống láng giềng nhỏ. Mặc dù sử dụng một hệ thống láng giềng nhỏ, nhưng nó nắm bắt cấu trúc kết cấu tốt và thu được chi tiết đối tượng. Hình 4.4 sử dụng những hình ảnh tổng hợp bao gồm bốn hoặc năm loại kết cấu khác nhau. Đây là vấn đề phân đoạn kết cấu mà một kết cấu được rút ra trong số đó. Các kết quả phân đoạn trong hình 4.4 (d) cho thấy rằng thuật toán RWR tạo ra phân đoạn kết cấu đáng tin cậy hơn GC và RW.



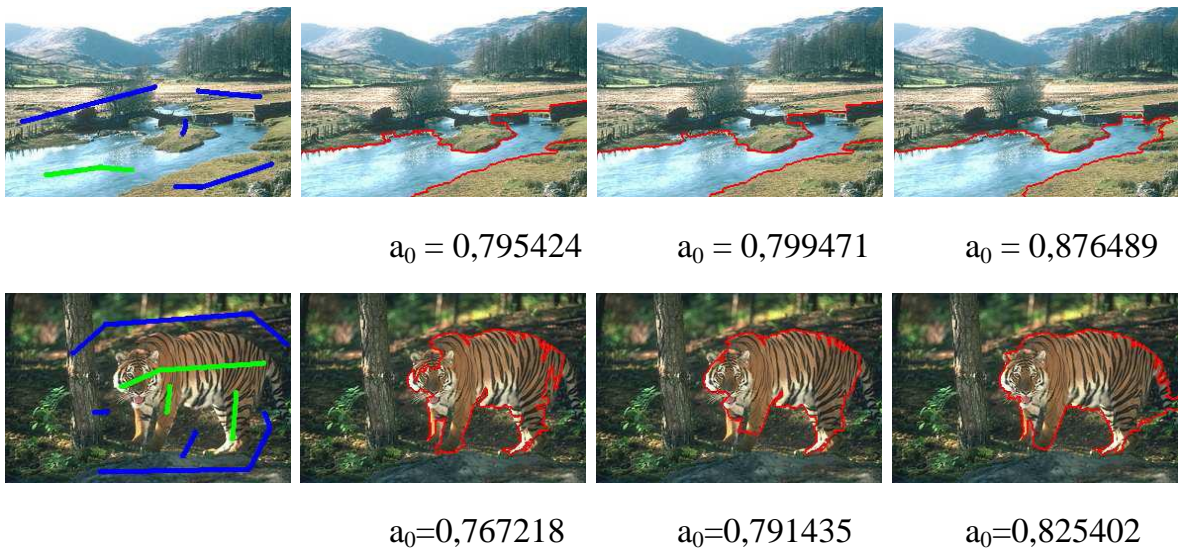
Hình 4.4. So sánh phân đoạn kết cấu giữa các thuật toán GC, RW, RWR

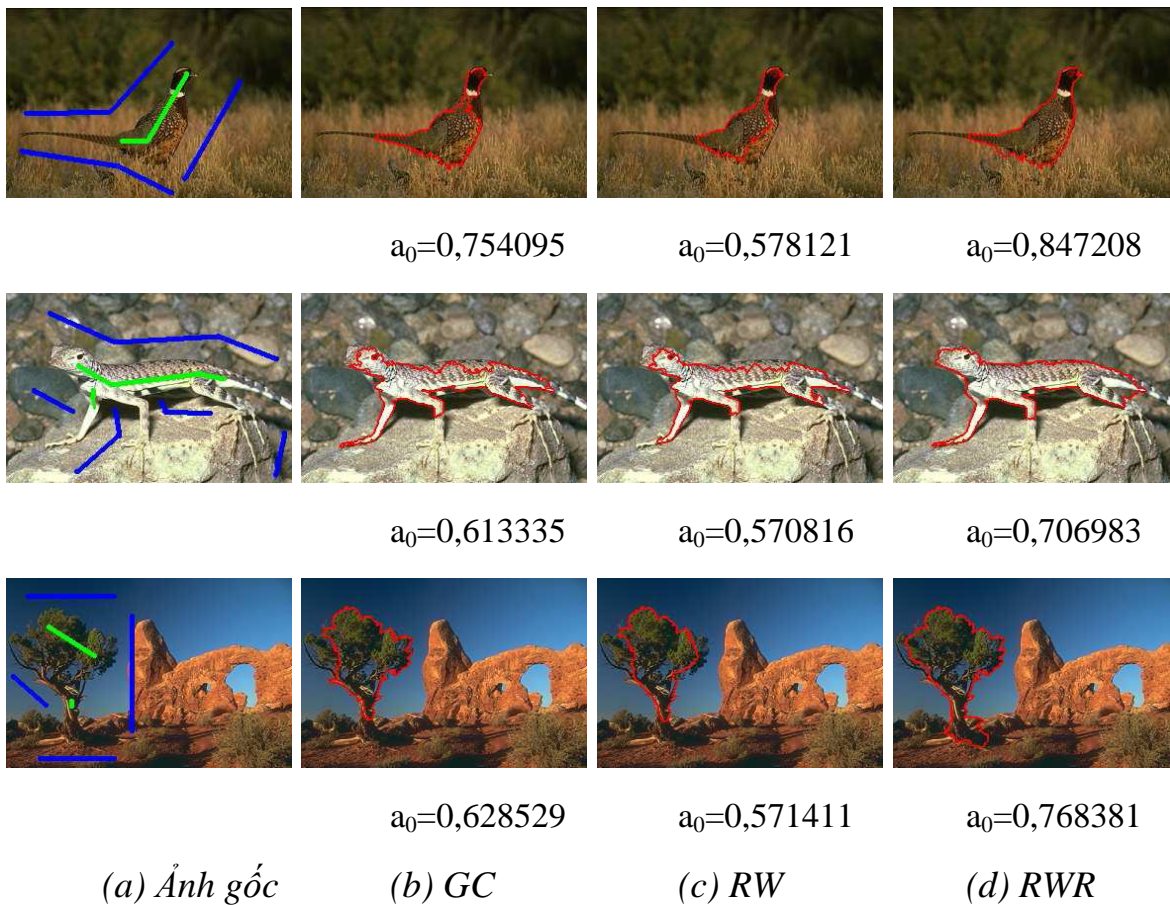
Định lượng so sánh: Hai trường hợp trước đó thường phát sinh trong ảnh tự nhiên. Bây giờ, so sánh các phân đoạn thu được từ GC, RW, và thuật toán RWR trên ảnh tự nhiên. Thuật toán sử dụng một tập dữ liệu của hình ảnh tự nhiên mà đối tượng người cung cấp một vị trí nổi bật / nhãn nền: cảnh thật. Để so sánh định lượng, sự giống nhau giữa kết quả phân đoạn và cài đặt sẵn cảnh thật được đo bằng cách sử dụng chuẩn hóa chồng:

$$a_0 = \frac{|R \cap G|}{|R \cup G|} \quad (4.1)$$

trong đó R là tập hợp các điểm ảnh được gán như cận cảnh từ các kết quả phân ảnh và G là ground truth. Ở đây, nó được sử dụng như là đo độ chính xác của các phân đoạn ảnh. Đối với thí nghiệm này, chọn hình ảnh tự nhiên với vùng kết cấu cao (cluttered) hoặc với có sự phân phối màu sắc tương tự giữa cận cảnh và nền cảnh.

Trong hình 4.5, phân đoạn đã được tạo ra từ ba thuật toán khác nhau trên ảnh tự nhiên. So với các phân đoạn từ GC và RW trong hình 4.5, thuật toán RWR có chất lượng và số lượng tốt hơn về phân đoạn. Với mỗi một ảnh khác nhau ta thu được kết quả phân đoạn khác nhau





Hình 4.5. So sánh thuật toán GC, RW, RWR trên ảnh tự nhiên

KẾT LUẬN

Kết quả đạt được

Trong quá trình nghiên cứu tài liệu và thực hiện đồ án dưới sự định hướng của thầy hướng dẫn em thấy bản thân đã đạt được một số kết quả như sau:

- Tìm hiểu được một cách tổng quan các vấn đề về XLA, phân đoạn ảnh và cách hướng tiếp cận trong phân đoạn ảnh.
- Trong chương 3 và 4 em đã tìm hiểu và cài đặt được phương pháp phân đoạn ảnh dựa trên RWR. Phương pháp này tối ưu hơn rất nhiều so với các phương pháp phân đoạn ảnh trước đó vì nó khắc phục được hai khó khăn quan trọng trong ảnh tự nhiên là đường biên yếu và kết cấu yếu. Phương pháp này dựa vào việc coi một bức ảnh như một đồ thị có trọng số. Sau khi tính xác suất trạng thái ổn định của mỗi điểm ảnh bằng cách sử dụng RWR, chúng ta có thể ước lượng khả năng phân tách và cuối cùng gán nhãn vào mỗi điểm ảnh.
- Ngoài ra, trong quá trình nghiên cứu em cũng tự tích lũy thêm cho mình các kiến thức về toán học, về kỹ thuật lập trình,... Và quan trọng là rèn luyện kỹ năng để thực hiện một nghiên cứu khoa học.

Một số hạn chế cần khắc phục

Bên cạnh những kết quả đạt được em tự thấy bản luận văn vẫn còn một số hạn chế:

- Chưa đưa ra được một phương pháp phân đoạn mới hoàn toàn. Trong khuôn khổ một đồ án tốt nghiệp, em mới chỉ trình bày lại các kiến thức tìm hiểu được chứ chưa đề xuất được một phương pháp hoàn toàn mới.
- Do thời gian có hạn, nên việc trình bày các thuật toán phân đoạn cũng chưa được hệ thống và khoa học. Có nhiều thuật toán được trình bày sơ lược.

TÀI LIỆU THAM KHẢO**Tài liệu Tiếng Việt**

[1]. Đỗ Năng Toàn, Phạm Việt Bình (2007). *Giáo trình xử lý ảnh*, Nhà xuất bản Đại học Thái Nguyên.

[2]. Ths. Võ Đức Khánh, GS.TSKH. Hoàng Kiếm (2007). *Giáo trình xử lý ảnh*. Nhà xuất bản Đại học Quốc Gia TP Hồ Chí Minh.

[4]. Nguyễn Thị Anh Thư (2008), *Đồ án tốt nghiệp*. Trường ĐHDL Hải Phòng.

Tài liệu Tiếng Anh

[5]. Hanghang Tong, Christos Faloutsos, Jia-Yu Pan (2007). *Fast RandomWalk with Restart and Its Applications*

[6]. Leo Grady (2006). *Random Walks for Image Segmentation*.

[7]. Tae Hoon Kim, Kyoung Mu Lee, and Sang Uk Lee (2008). *Generative Image Segmentation Using Random Walks with Restart*. Dept. of EECS, ASRI, Seoul National University, 151-742, Seoul, Korea