

# Mục lục

<b>1</b>	<b>Xử lý ảnh và các vấn đề cơ bản của xử lý ảnh</b>	<b>3</b>
1.1	Xử lý ảnh là gì? . . . . .	3
1.2	Một số vấn đề trong xử lý ảnh . . . . .	6
1.2.1	Một số khái niệm trong xử lý ảnh . . . . .	6
1.2.2	Tăng cường khôi phục ảnh . . . . .	6
1.2.3	Biến đổi ảnh . . . . .	6
1.2.4	Phân tích ảnh . . . . .	6
1.2.5	Nhận dạng ảnh . . . . .	6
1.2.6	Nén ảnh . . . . .	6
<b>2</b>	<b>Một số phương pháp tiếp cận phân vùng ảnh</b>	<b>7</b>
2.1	Phân vùng ảnh theo ngưỡng biên độ . . . . .	7
2.2	Phân vùng ảnh theo miền đồng nhất . . . . .	8
2.2.1	Phương pháp tách cây tứ phân . . . . .	9
2.2.2	Phân vùng ảnh dựa vào phát triển vùng cục bộ . . . . .	10
2.2.3	Phân vùng ảnh dựa trên hợp và tách vùng . . . . .	11
2.3	Phân vùng ảnh dựa trên phân tích kết cấu . . . . .	11
2.4	Phân vùng ảnh dựa trên sự phân lớp điểm ảnh . . . . .	11
2.5	Phân vùng ảnh dựa vào lý thuyết đồ thị . . . . .	13
2.6	Phân vùng ảnh dựa trên biểu diễn và xử lý đa phân giải . . . . .	13

<b>3</b>	<b>Phương pháp phân vùng ảnh sử dụng biên động</b>	<b>14</b>
3.1	Phương pháp Active contour without Edege . . . . .	14
3.1.1	Mô hình . . . . .	14
3.1.2	Giải quyết vấn đề sử dụng phương pháp tập mức . . . . .	16
3.1.3	Ưu nhược điểm của mô hình . . . . .	19
3.2	Phương pháp Local binary fitting energy . . . . .	19
3.2.1	Mô hình . . . . .	19
3.2.2	Giải quyết vấn đề sử dụng phương pháp tập mức . . . . .	20
3.2.3	Ưu nhược điểm của mô hình . . . . .	22
3.3	Phương pháp Local image fitting . . . . .	22
3.3.1	Mô hình và phương trình với tập mức . . . . .	22
3.3.2	Ưu nhược điểm của mô hình . . . . .	23
3.4	Mô hình kết hợp Global và Local . . . . .	24
3.4.1	Mô hình . . . . .	24
3.4.2	Giải số . . . . .	26
<b>4</b>	<b>Áp dụng mô hình GLIF trong tìm biên đối tượng trong video</b>	<b>27</b>

# Chương 1

## Xử lý ảnh và các vấn đề cơ bản của xử lý ảnh

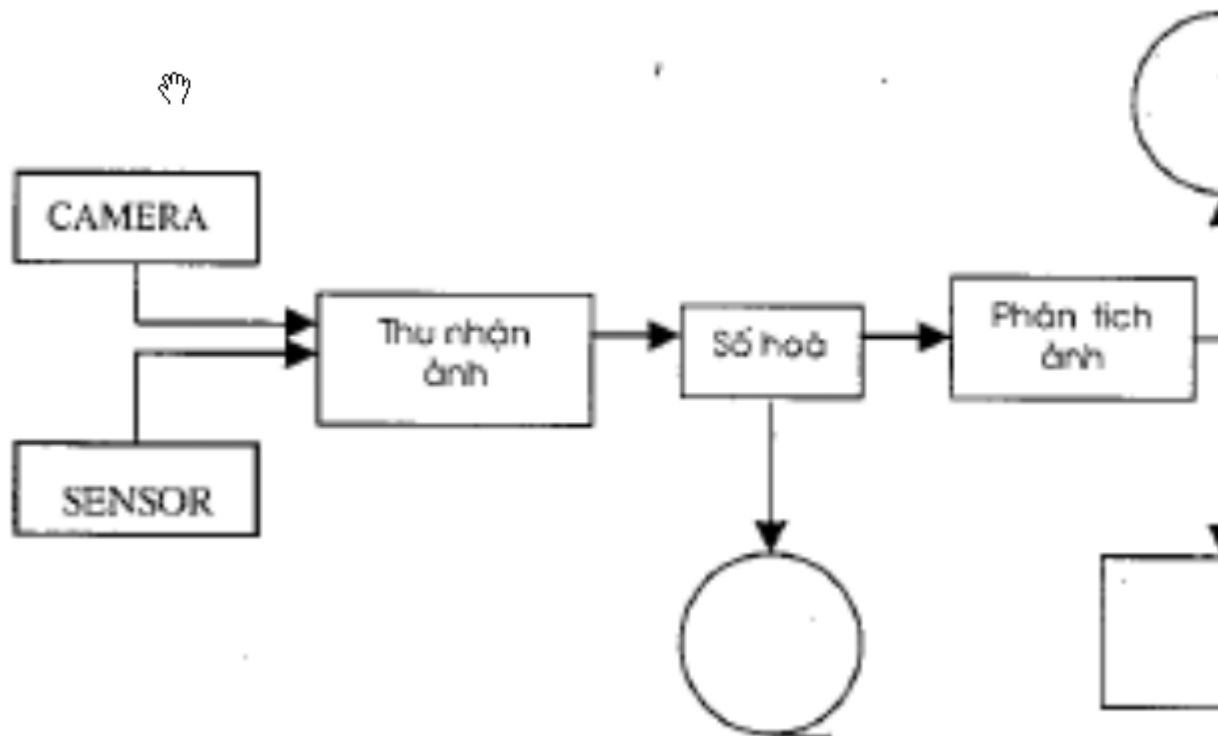
### 1.1 Xử lý ảnh là gì?

Một ảnh có thể được đại diện bởi một hàm 2 chiều  $f(x, y)$ .  $(x, y)$  là tọa độ của các điểm trong mặt phẳng, giá trị  $f$  tại các cặp tọa độ  $(x, y)$  được gọi là cường độ hay mức xám của ảnh tại điểm đó. Ảnh số là tập hợp hữu hạn các phần tử, mỗi phần tử có một tọa độ riêng và một giá trị  $f$  riêng. Các giá trị này gọi là *pixel* hay *imageelements*. Một trong các ứng dụng của ảnh số là trong ngành công nghiệp báo chí, khi ảnh lần đầu tiên được truyền qua cáp biển giữa London và NewYork. Việc truyền ảnh qua cáp biển này đã rút ngắn quá trình truyền ảnh qua Đại Tây Dương từ hơn một tuần xuống còn chưa đến 3 giờ. Một máy in đặc biệt mã hóa bức ảnh để truyền tải đi và khôi phục lại ảnh gốc ở đầu cuối. Sự phát triển của phần cứng máy tính đã giúp cho lĩnh vực đồ họa và xử lý ảnh phát triển một cách mạnh mẽ và ngày càng có nhiều ứng dụng trong cuộc sống.



**Hình 1.1:** Một bức ảnh được tạo ra bởi máy điện ấn năm 1921

Quá trình xử lý ảnh được xem là quá trình tho tác với ảnh đầu vào để đưa ra một kết quả mong muốn. Quá trình xử lý ảnh có thể được mô tả theo sơ đồ sau:

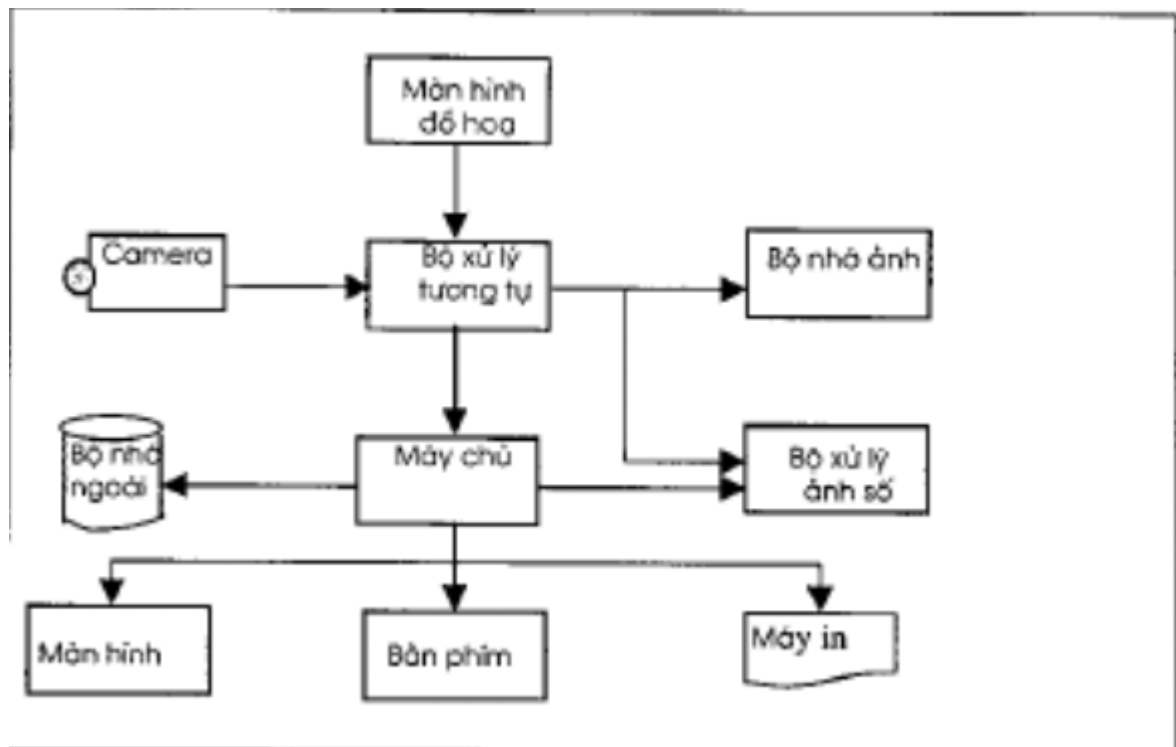


**Hình 1.2:** Các quá trình của xử lý ảnh

Đầu tiên là quá trình thu nhận ảnh. Ảnh có thể được thu nhận qua camera. Đầu ra có

thể là dạng tín hiệu tương tự hoặc cũng có thể là tín hiệu số hóa. Ngoài ra ảnh còn có thể được thu nhận qua bộ cảm ứng sensor ,tranh, ảnh được quét thông qua máy quét. Quá trình số hóa thực hiện biến đổi tín hiệu tương tự thành tín hiệu rời rạc sau đó được đưa vào quá trình phân tích,xử lý hoặc thực hiện lưu trữ ảnh.

Quá trình phân tích ảnh bao gồm nhiều bước nhỏ. Với một số ảnh, do thiết bị thu nhận ảnh, nguồn sáng hay nhiễu mà ảnh có thể bị suy biến. Do vậy chúng ta cần thực hiện quá trình tăng cường ảnh nhằm tăng chất lượng hoặc làm nổi bật một số đặc tính chính của ảnh. Tiếp theo là quá trình phát hiện các đặc tính như biên ảnh (Edge Detection), phân vùng ảnh (Segmentation)... Cuối cùng tùy theo ứng dụng mà có thể thực hiện quá trình nhận dạng, phân lớp hay ra quyết định Một hệ xử lý ảnh có thể được mô tả gồm các thành phần như bên dưới



**Hình 1.3:** Các thành phần chính của hệ thống xử lý ảnh

## **1.2 Một số vấn đề trong xử lý ảnh**

### **1.2.1 Một số khái niệm trong xử lý ảnh**

- Điểm ảnh: là một phần tử nhỏ nhất của ảnh. Một điểm ảnh có thể xem là tọa độ và cường độ sáng tại một điểm trong không gian của đối tượng.
- Mức xám, màu: là kết quả của sự mã hóa cường độ sáng của một điểm ảnh với một giá trị số.

### **1.2.2 Tăng cường khôi phục ảnh**

### **1.2.3 Biến đổi ảnh**

### **1.2.4 Phân tích ảnh**

### **1.2.5 Nhận dạng ảnh**

### **1.2.6 Nén ảnh**

## Chương 2

# Một số phương pháp tiếp cận phân vùng ảnh

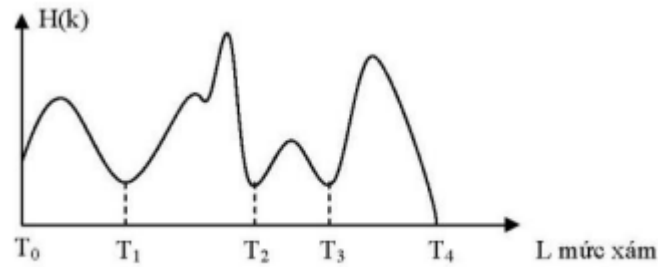
Phân vùng ảnh là một bước quan trọng trong xử lý ảnh, quá trình này nhằm phân tích ảnh thành các thành phần có cùng tính chất nào đó dựa theo biên hay các vùng liên thông. Vùng ảnh là tập hợp các điểm ảnh có cùng hay gần cùng một đặc điểm nào đó. Ví dụ như mức xám, mức màu .. Đường bao quanh vùng ảnh được gọi là biên ảnh

### 2.1 Phân vùng ảnh theo ngưỡng biên độ

Một ảnh được đặc trưng bởi các tính chất như: mức xám, độ tương phản, màu sắc,.... Ta có thể dùng ngưỡng biên độ với các đặc trưng trên để phân đoạn ảnh trong trường hợp ngưỡng biên độ đủ lớn đặc trưng cho ảnh. Kỹ thuật phân ngưỡng theo biên độ có lợi thế với ảnh nhị phân như bản in, đồ họa, ảnh màu Việc chọn ngưỡng sẽ bao gồm các bước sau:

- Xác định các đỉnh và khe của ảnh dựa vào việc phân tích lược đồ xám. Các khe có thể dùng để chọn ngưỡng
- Chọn ngưỡng  $t$  sao cho một phần xác định trước của toàn bộ mẫu là thấp hơn  $t$

- Điều chỉnh ngưỡng dựa trên lược đồ xám của các điểm lân cận
- Chọn ngưỡng theo lược đồ xám của các điểm thỏa mãn tiêu chuẩn chọn
- Khi có mô hình phân lớp xác suất, ta thực hiện xác định ngưỡng dựa vào tiêu chuẩn xác suất nhằm cực tiểu hóa xác suất sai
- Ta có ví dụ minh họa cho việc phân vùng dựa trên ngưỡng như sau:



**Hình 2.1:** Ví dụ minh họa cách chọn ngưỡng

Ở đây 5 ngưỡng được chọn và ảnh sẽ được phân vùng thành 4 vùng. Ta ký hiệu các vùng đó là  $C_k$  trong đó  $k = 1, 2, 3, 4$ . Các vùng sẽ được phân như sau:

$$P(m, n) \in C_k \text{ nu } T_{k-1} < P(m, n) < T_k; k = 1, 2, 3, 4 \quad (2.1)$$

Nếu sau khi phân vùng theo ngưỡng vừa mà chọn ảnh rõ nét thì kết thúc, ngược lại phải điều chỉnh ngưỡng và phân vùng lại cho đến khi đạt được kết quả mong muốn

## 2.2 Phân vùng ảnh theo miền đồng nhất

Kỹ thuật này sử dụng tính chất đồng nhất của các đặc trưng nào đó của ảnh. Cách thức phân vùng sẽ phụ thuộc vào việc lựa chọn đặc trưng dùng để thực hiện phân vùng. Các đặc trưng thường được dùng là mức xám với ảnh đen trắng và màu với ảnh màu. Với cách tiếp cận này, các phương pháp thường được sử dụng là :

- Phương pháp tách cây tứ phân



- Phương pháp cục bộ
- Phương pháp tổng hợp

### 2.2.1 Phương pháp tách cây tứ phân

Một tiêu chuẩn đồng nhất sẽ được chọn. Quá trình phân vùng sẽ thực hiện như sau: Đầu tiên ta thực hiện kiểm tra tính chất đồng nhất trên toàn bộ miền ảnh. Nếu tính đồng nhất thỏa mãn thì kết thúc thuật toán. Ngược lại miền ảnh sẽ được chia làm 4 miền con. Tiếp tục kiểm tra tính chất đồng nhất của các miền con và thực hiện phân tách khi tính chất đồng nhất không thỏa mãn. Quá trình kết thúc khi tất cả các miền con trong miền ảnh ban đầu thỏa mãn tính đồng nhất Phương pháp này được mô tả như sau

---

```

Func Phanvung(MienAnh)
{
    dsmiencon.Push(MienAnh)
    while(dsmiencon.Count()!=0)
    {
        miendangxet=dsmiencon.Pop();
        if(!ktdongnhat(miendangxet))
        {
            m1,m2,m3,m4=miendangxet.chiamien();
            for(int i=0;i<4;i++)
            {
                dsmiencon.Push(m_i)
            }
        }
    }
}

```

---

Tiêu chuẩn đồng nhất có thể dựa vào mức xám. Một cách đơn giản ta có thể chọn giá

trị chênh lệch giữa giá trị mức xám lớn nhất và giá trị mức xám nhỏ nhất trong miền đang xét. Hàm kiểm tra mức xám có thể được viết như sau: Giả sử  $(m_1, n_1), (m_2, n_2)$  là tọa độ điểm đầu và điểm cuối của miền đang xét

---

```
Func ktdongnhatch(miendangxet)
{
    int min=0;
    int max=255
    for(i=n1;i<n_2;i++)
    for(j=m1;j<m_2;j++)
    {
        if(I(i,j)<min)
            I(i,j)=min;
        if(I(i,j)>max)
            I(i,j)=max;
    }
    if((max-min)<nguong) return 1;
    return 0;
}
```

---

### 2.2.2 Phân vùng ảnh dựa vào phát triển vùng cục bộ

Ý tưởng của phương pháp này ngược lại với phương pháp cây tứ phân. Phương pháp thực hiện xét từ các miền ảnh nhỏ nhất của ảnh rồi nối chúng lại nếu chúng thỏa mãn điều kiện đồng nhất. Thực hiện tiếp tục khi không thể tiếp tục nối các miền lại với nhau nữa. Trong cách này, hai vùng sẽ được nối lại với nhau nếu chúng thỏa mãn 2 điều kiện :

- Kế cận nhau
- Có mức xám đồng nhất

Để xác định tính kế cận giữa 2 vùng người ta sử dụng tính liên thông. Có 2 quan niệm

về liên thông là 4 liên thông và 8 liên thông. Với 4 liên thông, một điểm ảnh sẽ có 4 điểm ảnh kế cận theo 2 hướng  $x, y$ . Còn với 8 liên thông, một điểm ảnh sẽ có 4 điểm ảnh theo 2 hướng  $x, y$  là 4 điểm ảnh theo các hướng chéo 45 độ

### 2.2.3 Phân vùng ảnh dựa trên hợp và tách vùng

Để khắc phục được nhược điểm của 2 phương pháp cây tứ phân và phương pháp phát triển cục bộ người ta đưa ra phương pháp kết hợp ý tưởng hợp và tách vùng của 2 phương pháp trên. Với phương pháp tách, việc thực hiện chia quá chi tiết còn với phương pháp hợp mặc dù giảm được tối thiểu số vùng sau khi chia tuy nhiên không cho ta thấy rõ được mối liên hệ giữa các miền. Phương pháp kết hợp giữa hợp và tách thực hiện như sau. Đầu tiên, ta thực hiện tách miền ảnh theo cây tứ phân, thực hiện phân đoạn từ gốc đến lá, tiếp theo tiến hành duyệt cây theo chiều ngược lại để hợp các vùng đồng nhất. Trong thao tác hợp các vùng đồng nhất, có thể có nhiều vùng thỏa mãn điều kiện đồng nhất. Vì vậy ta phải xây dựng một hàm đánh giá giá trị đồng nhất của các vùng với nhau trả về các giá trị trong đoạn  $[0, 1]$  trong đó 0 là không đồng nhất, và 1 là hoàn toàn đồng nhất. Với hàm đánh giá kiểu này, trong quá trình thực hiện thao tác hợp, trong trường hợp có nhiều vùng thỏa mãn điều kiện đồng nhất, ta sẽ thực hiện chọn vùng mà giá trị hàm đồng nhất trả về lớn nhất để thực hiện hợp vùng.

## 2.3 Phân vùng ảnh dựa trên phân tích kết cấu

### 2.4 Phân vùng ảnh dựa trên sự phân lớp điểm ảnh

Cho một ảnh với các điểm ảnh:  $p_i$  với  $i = [1, N * M]$ ,  $M, N$  lần lượt là chiều rộng và chiều cao của ảnh. Với mỗi điểm ảnh ta chọn một thuộc tính để phục vụ cho việc phân vùng. Giả sử giá trị thuộc tính là  $A(p_i)$  với điểm ảnh thứ  $i$ . Thuật toán thực hiện chia tập hợp các điểm ảnh thành vào cá lớp và từ các lớp này tiến hành xây dựng vùng tương ứng. Tiêu chí dùng để phân vùng là độ tương đồng giữa các điểm ảnh trên thuộc tính phân

vùng được chọn. Ta hoàn toàn có thể sử dụng thuộc tính phân vùng nhiều chiều tùy vào yêu cầu của việc phân vùng. Ta có ví dụ về việc phân vùng với thuộc tính được chọn là một chiều, cụ thể là mức xám của điểm ảnh:

- Bước 1: Khởi tạo  $t = 0$ :

- Thực hiện phân tích lược đồ xám để dự đoán số lớp cần phân
- Giả sử số lớp chọn được là  $K$ , ta khởi tạo các ngưỡng đều nhau. Cụ thể là:

$$T_i(t) = i \frac{l_{max} - l_{min}}{K} + l_{min}; i = 0, \dots, K. \quad (2.2)$$

Với  $l_{max}, l_{min}$  lần lượt là giá trị mức xám lớn nhất và giá trị mức xám nhỏ nhất của các điểm ảnh trong ảnh.

- Bước 2: Thực hiện bước lặp theo thời gian:

- Phân lớp theo các ngưỡng  $T_i(k-1)$  với tiêu chí

$$p_i \in C_k \text{ nu } T_{k-1}(t-1) < A(p_i) < T_k(t-1), i = 1 \dots K \quad (2.3)$$

- Tính giá trị trung bình mức xám của các lớp

$$m_k(t-1) = \text{avarage}_{p_i \in C_k} A(p_i) \quad (2.4)$$

- Tính lại giá trị ngưỡng dựa trên giá trị trung bình mức xám của các lớp như sau:

$$T_i(t) = \frac{m_{i+1}(t-1) - m_i(t)}{2} \quad (2.5)$$

- Kiểm tra điều kiện dừng Nếu  $T_i(t) \approx T_i(t-1)$  hoặc  $m_i(t-1) \approx m_i(t)$  Thì dừng

- Thực hiện xây dựng vùng theo các lớp vừa phân

**2.5    Phân vùng ảnh dựa vào lý thuyết đồ thị**

**2.6    Phân vùng ảnh dựa trên biểu diễn và xử lý đa  
phân giải**

## Chương 3

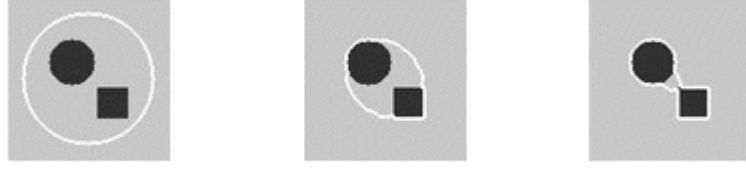
# Phương pháp phân vùng ảnh sử dụng biên động

Ý tưởng của các phương pháp theo hướng tiếp cận này là xuất phát từ đường cong ban đầu, ta thực hiện biến đổi đường biên ban đầu theo thời gian và dừng lại ở biên của vật. Việc biến đổi đường biên được thực hiện thông qua xây dựng một vòng lặp nhằm đạt được giá trị cuối cùng của đường cong thỏa mãn cực tiểu hoặc cực đại một hàm năng lượng (energy function) nào đó.

### 3.1 Phương pháp Active contour without Edege

#### 3.1.1 Mô hình

Chúng ta định nghĩa đường cong  $C \in \Omega$  là biên của một tập con mở  $\omega$  của  $\Omega$ . Giả sử rằng ảnh được chia làm hai vùng được xấp xỉ bởi các hằng số cường độ khác nhau  $u_0^i$  và  $u_0^o$ . Đối tượng cần xác định trong ảnh được đại diện bởi vùng có cường độ  $u_0^i$ . Giả sử biên của nó là  $C_0$ . Vì vậy chúng ta có  $u_0 \approx u_0^i$  trong đối tượng, và  $u_0 \approx u_0^o$ .



Mô hình bao gồm 2 phần chính Fitting term và Regularization term:

- Fitting term: Đặt

$$F_1 = \iint_{inside(C)} |u_0 - c_1|^2 dx dy \quad (3.1)$$

$$F_2 = \iint_{outside(C)} |u_0 - c_2|^2 dx dy \quad (3.2)$$

Trong đó

$$c_1 = average(u_0 inside(C)) \quad (3.3)$$

$$c_2 = average(u_0 outside(C)) \quad (3.4)$$

Trong trường hợp đường cong  $C$  nằm ngoài đối tượng, ta có  $F_1 > 0, F_2 \approx 0$ . Nếu đường cong  $C$  nằm trong đối tượng ta có  $F_1 \approx 0, F_2 > 0$ . Nếu đường cong  $C$  một phần nằm trong và một phần nằm ngoài đối tượng ta có  $F_1 > 0, F_2 > 0$ . Nếu đường cong  $C$  khớp với đối tượng  $F_1 \approx 0, F_2 \approx 0$ :



- Regularization term:

$$\mu \cdot |C| + \nu \cdot Area(inside(C)) \quad (3.5)$$

Trong đó  $C$  là độ dài đường cong  $C$ ,  $Area(inside(C))$  là độ lớn phần diện tích giới hạn bởi đường cong. Trong hầu hết bài toán ta chọn  $\nu = 0$  và  $\mu$  là tham số thay đổi tùy theo đầu vào để được kết quả tốt. Tham số  $\mu$  được chọn càng nhỏ thì mô hình càng có khả năng phát hiện các đối tượng nhỏ. Trong trường hợp chỉ cần phát hiện các đối tượng có kích thước lớn,  $\mu$  được chọn lớn hơn.

Tổng kết lại ta có hàm năng lượng

$$\begin{aligned} inf_{c_1, c_2, C} F(c_1, c_2, C) &= \mu \cdot |C| + \nu \cdot Area(inside(C)) \\ &+ \lambda_1 \cdot \iint_{inside(C)} |u_0 - c_1|^2 dx dy + \lambda_2 \cdot \iint_{outside(C)} |u_0 - c_2|^2 dx dy \end{aligned} \quad (3.6)$$

Bài toán cực tiểu đặt ra là

$$inf_{c_1, c_2, C} F(c_1, c_2, C) \quad (3.7)$$

### 3.1.2 Giải quyết vấn đề sử dụng phương pháp tập mức

Ta coi biên  $C$  được đại diện bởi tập mức 0 của một hàm  $\phi : \omega \rightarrow \mathbb{R}$ . Đặt

$$\begin{cases} C = \omega = (x, y) \in \Omega : \Phi(x, y) = 0 \\ inside(C) = \partial\omega = (x, y) \in \Omega : \Phi(x, y) > 0 \\ outside(C) = \Omega \setminus \omega = (x, y) \in \Omega : \Phi(x, y) < 0 \end{cases} \quad (3.8)$$

Thay biến chưa biết  $C$  bởi biến chưa biết  $\Phi$ . Trong quá trình xây dựng lại phương trình theo  $\Phi$  ta sử dụng hàm Heviside  $\Phi$  và Hàm Delta Dirac  $\delta_0$

$$H(z) = \begin{cases} 1 & \text{nếu } z > 0 \\ 0 & \text{nếu } z < 0 \end{cases} \quad (3.9)$$



Hàm năng lượng sẽ được viết lại thành

$$\begin{aligned} \inf_{c_1, c_2, \Phi} F(c_1, c_2, C) &= \mu \int_{\Omega} \delta(x, y) |\nabla \Phi(x, y)| dx dy + \nu \int_{\Omega} H(\Phi(x, y)) dx dy \\ &+ \lambda_1 \cdot \int_{\Omega} |u_0 - c_1|^2 H(\Phi(x, y)) dx dy \\ &+ \lambda_2 \cdot \iint_{\Omega} |u_0 - c_2|^2 (1 - H(\Phi(x, y))) dx dy \end{aligned} \quad (3.10)$$

Trong đó  $c_1, c_2$  có thể được tính theo  $\Phi$  theo công thức:

$$c_1(\Phi) = \frac{\int_{\Omega} u_0(x, y) H(\Phi(x, y)) dx dy}{\int_{\Omega} H(\Phi(x, y)) dx dy} \quad (3.11)$$

$$c_2(\Phi) = \frac{\int_{\Omega} u_0(x, y) (1 - H(\Phi(x, y))) dx dy}{\int_{\Omega} H(\Phi(x, y)) (1 - H(\Phi(x, y))) dx dy} \quad (3.12)$$

Để tính toán phương trình Euler-Lagrange cho hàm  $\Phi$  ta thực hiện chuẩn hóa  $H$  và  $\delta_0$  tương ứng thành  $H_{\epsilon}$  và  $\delta_{\epsilon}$ ,  $\epsilon \rightarrow 0$ .  $H_{\epsilon}$  là mở rộng của  $H$  trên  $C^2(\bar{\Omega})$  và  $\delta_{\epsilon} = H'_{\epsilon}$ . Hàm mở rộng của  $F(c_1, c_2, C)$  thành

$$\begin{aligned} F(c_1, c_2, \Phi) &= \mu \int_{\Omega} \delta_0(x, y) |\nabla \Phi(x, y)| dx dy + \nu \int_{\Omega} H(\Phi(x, y)) dx dy \\ &+ \lambda_1 \cdot \int_{\Omega} |u_0 - c_1|^2 H(\Phi(x, y)) dx dy \\ &+ \lambda_2 \cdot \iint_{\Omega} |u_0 - c_2|^2 (1 - H(\Phi(x, y))) dx dy \end{aligned} \quad (3.13)$$

Cố định  $c_1, c_2$ , thực hiện cực tiểu hàm  $F_{\epsilon}$  theo  $\Phi$ , ta thu được phương trình Euler-Lagrange theo  $\Phi$ . Tham số hóa  $\Phi$  theo thời gian và áp dụng phương pháp hướng giảm ta được:

$$\frac{\partial \Phi}{\partial t} = \delta_{\epsilon}(\Phi) [\mu - \text{div}(\frac{\nabla \Phi}{|\nabla \Phi|}) - \nu - \lambda_1(u_0 - c_1)^2 - \lambda_2(u_0 - c_2)^2] \text{trong}(0, \infty) \times \Omega \quad (3.14)$$

$$\Phi(0, x, y) = \Phi_0(x, y) \text{trong} \Omega \quad (3.15)$$

$$\frac{\delta_{\epsilon}(\Phi)}{|\nabla \Phi|} \frac{\partial \Phi}{\partial \vec{n}} = 0 \quad (3.16)$$

Trong đó  $\vec{n}$  là vector pháp tuyến hướng ra ngoài của biên  $\Omega$ , và  $\frac{\partial \Phi}{\partial \vec{n}}$  là đạo hàm theo vector pháp tuyến của  $\Phi$ . Trong quá trình tính toán xấp xỉ các giá trị  $\Phi_x, \Phi_y$  ở các điểm gần biên, ta cần biết được giá trị  $\Phi(x, y)$  tại các điểm  $(x, y)$  nằm ngoài biên  $\partial\Omega$  trong khi

ta chỉ có giá trị tại các điểm nằm trong và trên biên của  $\Omega$ , một cách tự nhiên người ta có thể chọn  $\Phi(x, y)$  tại các điểm  $(x, y)$  nằm ngoài biên bằng với giá trị của  $\Phi(x, y)$  tại điểm gần  $(x, y)$  nhất trên biên, hay  $\frac{\partial \Phi}{\partial \vec{n}} = 0$ . Thuật toán theo mô hình này được mô tả cụ thể như sau:

- Bước 1: Khởi tạo  $\Phi^0 = \Phi_0, n = 0$
- Bước 2: Tính toán các giá trị  $c_1(\Phi^n), c_2(\Phi^n)$
- Bước 3: Giải phương trình đạo hàm riêng (1.14) để có  $\Phi^{n+1}$
- Bước 4: Khởi tạo lại hàm  $\Phi$  bằng hàm dấu khoảng cách tới đường cong
- Bước 5: Kiểm tra điều kiện dừng, nếu không thỏa mãn thì gán  $n = n + 1$  và quay lại bước 2

Các tính toán xấp xỉ được dùng trong quá trình giải số hệ phương trình đạo hàm riêng (1.14) như sau:

$$\begin{aligned}
 k &= \text{div}\left(\frac{\nabla \Phi}{|\Phi|}\right) = \frac{\Phi_{xx}\Phi_y^2 - 2\Phi_{xy}\Phi_x\Phi_y + \Phi_{yy}\Phi_x^2}{(\Phi_x^2 + \Phi_y^2)^{3/2}} \\
 \Phi_x &= \frac{1}{2h}(\Phi_{i+1,j} - \Phi_{i-1,j}) \\
 \Phi_x &= \frac{1}{2h}(\Phi_{i,j+1} - \Phi_{i,j-1}) \\
 \Phi_{xx} &= \frac{1}{h^2}(\Phi_{i+1,j} - \Phi_{i,j} + \Phi_{i-1,j}) \\
 \Phi_{yy} &= \frac{1}{h^2}(\Phi_{i,j+1} - \Phi_{i,j} + \Phi_{i,j-1}) \\
 \Phi_{xy} &= \frac{1}{h^2}(\Phi_{i+1,j+1} - \Phi_{i-1,j+1} - \Phi_{i,j+1} + \Phi_{i,j-1})
 \end{aligned} \tag{3.17}$$

trong đó  $h$  là khoảng cách lưới. Khi đó (7) có thể được rời rạc hóa thành thành

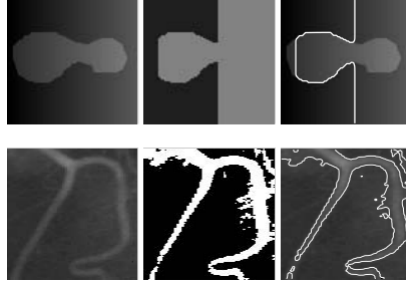
$$\frac{\Phi_{n+1} - \Phi_n}{\Delta t} = \delta_\epsilon(\Phi_{n+1})[\mu k - \lambda_1(u_{i,j} - c_1(\Phi^n))^2 + \lambda_2(u_{i,j} - c_2(\Phi^n))^2] \tag{3.18}$$

Từ đó ta có thể tính được  $\Phi^{n+1}$  theo  $\Phi_n$

### 3.1.3 Ưu nhược điểm của mô hình

Ưu điểm của mô hình là có thể phát hiện được biên trơn, xử lý tốt với ảnh nhiễu, mô hình có thể phát hiện được biên từ duy nhất một đường cong khởi tạo ban đầu. Đường cong này có thể ở mọi vị trí trên ảnh mà không nhất thiết phải bao quanh đối tượng trong ảnh.

Nhược điểm của mô hình là độ phức tạp tính toán còn lớn do thao tác re-initialization và kết quả còn thiếu chính xác trong trường hợp ảnh không đồng nhất cường độ.



## 3.2 Phương pháp Local binary fitting energy

### 3.2.1 Mô hình

Cho ảnh  $I : \Omega \rightarrow \mathbb{R}^d$ . Trong đó  $\Omega \subset \mathbb{R}^2$  là miền ảnh,  $d > 1$  là bậc của vector  $I(x)$ . Với ảnh xám  $d=1$ , với ảnh màu  $d=3$ . Gọi  $C$  là biên của ảnh trong  $\Omega$ , với mỗi  $x$  ta định nghĩa một hàm năng lượng sau:

$$\epsilon_x^L BF(C, f_1(x), f_2(x)) = \lambda_1 \int_{in(C)} K(x-y) |I(y) - f_1(x)|^2 dy + \lambda_2 \int_{out(C)} K(x-y) |I(y) - f_2(x)|^2 dy \quad (3.19)$$

trong đó  $\lambda_1, \lambda_2$  là các hằng số dương,  $K$  là hàm nhân với thuộc tính địa phương  $K(u)$  giảm và dần về 0 khi  $|u|$  giảm,  $f_1(x), f_2(x)$  là hai số khớp với cường độ ảnh tại các điểm gần  $x$ . Ta gọi  $x$  là điểm trung tâm của hàm năng lượng trên. Trong mô hình này  $K$  được chọn là hàm nhân Gaussian

$$K_\sigma(x) = \frac{1}{(2\pi)^{n/2} \sigma^n} e^{-|x|^2/2\sigma^2} \quad (3.20)$$

với  $\sigma$  là tham số dương có thể tùy chỉnh được. Nhấn mạnh rằng  $f_1(x), f_2(x)$  thay đổi theo  $x$ .  $f_1(x), f_2(x)$  làm phương pháp này trở nên khác các phương pháp khác.

Trong mô hình này, hàm năng lượng có tính đại phương với  $x$  hay  $f_1(x), f_2(x)$  chỉ khớp với cường độ ảnh tại các điểm gần  $x$ . Điều này có được là do tính chất của hàm  $K$  là  $K(x - y)$  có giá trị lớn hơn khi  $y$  gần  $x$ . Vì vậy cường độ ảnh tại các điểm  $y$  gần  $x$  ảnh hưởng nhiều hơn đến giá trị  $f_1, f_2$  làm cực tiểu hàm năng lượng  $\epsilon_x^{LBF}(C, f_1, f_2)$ , trong khi cường độ ảnh tại các điểm  $y$  xa  $x$  hầu như không ảnh hưởng tới các giá trị  $f_1, f_2$

Tại mỗi điểm trung tâm  $x$ , hàm năng lượng  $\epsilon_x^{LBF}(C, f_1, f_2)$  có thể đạt cực tiểu khi đường con  $C$  khớp với biên của đối tượng và các giá trị  $f_1, f_2$  được chọn sao tối ưu. Tuy nhiên hàm năng lượng  $\epsilon_x^{LBF}(C, f_1, f_2)$  được định nghĩa một cách cục bộ theo điểm trung tâm  $x$ . Để tìm được toàn bộ biên của vật, ta cần cực tiểu hàm  $\epsilon_x^{LBF}$  với mọi  $x$  trong miền ảnh  $\Omega$ . Điều này có thể làm được bằng cách định nghĩa hàm năng lượng mới sau

$$\epsilon(C, f_1, f_2) = \int_{\Omega} \epsilon_x^{LBF}(C, f_1(x), f_2(x)) dx \quad (3.21)$$

### 3.2.2 Giải quyết vấn đề sử dụng phương pháp tập mức

Đường con  $C \subset \Omega$  được đại diện bởi tập mức 0 của một hàm Lipschit  $\Phi \rightarrow \mathbb{R}$  Viết lại hàm năng lượng  $\epsilon_x^{LBF}(C, f_1, f_2)$  theo  $\Phi$  ta được

$$\begin{aligned} \epsilon_x^{LBF}(\Phi, f_1(x), f_2(x)) = & \lambda_1 \int_{\Omega} K_{\sigma}(x - y) |I(y) - f_1(x)|^2 H(\Phi(y)) dy \\ & + \lambda_2 \int_{\Omega} K_{\sigma}(x - y) |I(y) - f_2(x)|^2 (1 - H(\Phi(y))) dy \end{aligned} \quad (3.22)$$

trong đó  $H$  là hàm Heaviside. Theo đó hàm năng lượng  $\epsilon^{LBF}$  có thể được viết lại thành

$$\begin{aligned} \epsilon(C, f_1, f_2) = & \int_{\Omega} \epsilon_x^{LBF}(C, f_1(x), f_2(x)) dx \\ = & \lambda_1 \int_{\Omega} \int_{\Omega} K_{\sigma}(x - y) |I(y) - f_1(x)|^2 H(\Phi(y)) dy dx \\ & + \lambda_2 \int_{\Omega} \int_{\Omega} K_{\sigma}(x - y) |I(y) - f_2(x)|^2 (1 - H(\Phi(y))) dy dx \end{aligned} \quad (3.23)$$

Để đảm bảo rằng  $\Phi$  ổn định, chúng ta thêm hàm độ lệch giữa hàm tập mức  $\Phi$  và hàm dấu khoảng cách.

$$P(\Phi) = \int_{\Omega} \frac{1}{2} (|\nabla \Phi| - 1)^2 dx \quad (3.24)$$

$$\mathcal{L}(\Phi) = \int_{\Omega} \delta(\Phi(x)) |\nabla \Phi(x)| dx \quad (3.25)$$

Bây giờ ta có hàm energy cuối cùng

$$\mathcal{F}(\Phi, f_1, f_2) = \epsilon(C, f_1, f_2) + \mu \mathcal{P}(\Phi) + \nu \mathcal{L}(\Phi) \quad (3.26)$$

trong đó  $\mu$  và  $\nu$  là các hằng số. Trong thực tế, các  $H$  được xấp xỉ bởi một hàm trơn  $H_{\epsilon}$

$$H_{\epsilon}(x) = \frac{1}{2} \left[ 1 + \frac{2}{\pi} \arctan\left(\frac{x}{\epsilon}\right) \right] \quad (3.27)$$

và  $\delta_{\epsilon}(x)$  được chọn là

$$\delta_{\epsilon}(x) = H'_{\epsilon}(x) = \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2} \quad (3.28)$$

Hàm năng lượng này sẽ được cực tiểu hóa để tìm biên Với  $\Phi$  cố định ta thực hiện cực tiểu hàm năng lượng  $\epsilon(C, f_1, f_2)$  theo  $f_1, f_2$  ta tìm được  $f_1, f_2$  như sau:

$$f_1(x) = \frac{K_{\sigma} * [(H_{\epsilon}(\Phi(x))) * I(x)]}{K_{\sigma} * (H_{\epsilon}(\Phi(x)))} \quad (3.29)$$

$$f_2(x) = \frac{K_{\sigma} * [(1 - H_{\epsilon}(\Phi(x))) * I(x)]}{K_{\sigma} * (1 - H_{\epsilon}(\Phi(x)))} \quad (3.30)$$

Theo công thức  $f_1, f_2$  ở trên luôn dương do  $H_{\epsilon}$  và  $1 - H_{\epsilon}$  luôn dương. Cố định  $f_1, f_2$  và cực tiểu hàm  $\epsilon^{LBF}(\Phi, f_1(x), f_2(x))$  theo  $\Phi$ , sử dụng phương pháp hướng giảm ta được:

$$\begin{aligned} \frac{\partial \Phi}{\partial t} = & -\delta_{\epsilon}(\Phi)(\lambda_1 e_1 - \lambda_2 e_2) + \nu \delta_{\epsilon}(\Phi) \operatorname{div}\left(\frac{\nabla \Phi}{|\nabla \Phi|}\right) \\ & + \mu(\nabla^2 \Phi - \operatorname{div}\left(\frac{\nabla \Phi}{|\nabla \Phi|}\right)) \end{aligned} \quad (3.31)$$

trong đó  $\delta_{\epsilon}(x)$  là hàm trơn Dirac được cho như công thức trên và  $e_1(x), e_2(x)$  được tính theo công thức sau:

$$e_1(x) = \int_{\Omega} K_{\sigma}(y - x) |I(x) - f_1(y)|^2 dy \quad (3.32)$$

$$e_2(x) = \int_{\Omega} K_{\sigma}(y - x) |I(x) - f_2(y)|^2 dy \quad (3.33)$$

### 3.2.3 Ưu nhược điểm của mô hình

Ưu điểm của mô hình này là không cần thiết phải chuẩn hóa  $f_1, f_2$ . Thực tế,  $f_1, f_2$  làm cực tiểu hàm năng lượng có thể được cho bởi công thức (14)(15) và là các hàm trơn. Hơn nữa mô hình này cũng không cần phải mở rộng  $f_1, f_2$  vì nó được định nghĩa trên toàn miền ảnh  $\Omega$ . Một ưu điểm nữa của mô hình này là việc khởi tạo lại hàm  $\Phi$  là không cần thiết do việc thêm hàm chuẩn hóa khoảng cách. Nhờ việc chuẩn hóa khoảng cách này mà việc khởi tạo hàm  $\Phi$  lúc đầu trở nên linh hoạt hơn. Tuy nhiên việc tính toán  $\lambda_1 e_1 - \lambda_2 e_2$  còn tốn nhiều tài nguyên

## 3.3 Phương pháp Local image fitting

### 3.3.1 Mô hình và phương trình với tập mức

Một hàm khớp địa phương (Local fitted image fomulation) được định nghĩa như sau:

$$I^{LFI} = m_1 H_\epsilon(\Phi) + m_2 (1 - H_\epsilon(\Phi)) \quad (3.34)$$

Trong đó  $m_1, m_2$  được định nghĩa như sau

$$\begin{cases} m_1 = \text{mean}(I \in (\{x \in \Omega | \Phi(x) < 0\} \cap W_k(x))) \\ m_2 = \text{mean}(I \in (\{x \in \Omega | \Phi(x) > 0\} \cap W_k(x))) \end{cases} \quad (3.35)$$

trong đó  $W_k(x)$  là một hàm cửa sổ hình chữ nhật. Ở đây ta chọn  $W_k(x)$  là hàm cửa sổ Gaussian  $K_\sigma(x)$  với độ lệch chuẩn  $\sigma$  và với kích thước  $4k+1$  và  $4k+1$  với  $k$  là số nguyên lớn nhất không vượt quá  $\sigma$ . Mô hình này sử dụng hàm năng lượng khớp địa phương (local image fitting energy functional) thể hiện độ lệch giữa ảnh sau khi khớp và ảnh gốc:

$$E^{LIF}(\Phi) = \frac{1}{2} \int_{\Omega} |I(x) - I^{LFI}(x)|^2 dx \quad (3.36)$$

Sử dụng phương pháp hướng giảm ta có

$$\frac{\partial \Phi}{\partial t} = (I - I^{LFI})(m_1 - m_2) \delta_\epsilon(\Phi) \quad (3.37)$$

Trong quá trình làm việc theo phương pháp với tập mức, để đảm bảo tính hội tụ của  $\Phi$ , ta cần thực hiện thao tác khởi tạo lại  $\Phi$ . Tuy nhiên việc này đòi hỏi khá nhiều tài nguyên. Li et al đề xuất việc thêm hàm phạt độ lệch giữa  $\Phi$  và hàm dấu khoảng cách(SDF).

Các bước của thuật toán như sau:

- Bước 1: Khởi tạo hàm  $\Phi$  là hàm nhị phân như sau:

$$\Phi(x, t = 0) = \begin{cases} -\rho, x \in \Omega_0 - \partial\Omega_0 \\ 0, x \in \partial\Omega_0 \\ \rho, x \in \Omega - \Omega_0 \end{cases} \quad (3.38)$$

trong đó  $\rho > 0$  là hằng số,  $\Omega_0$  là tập con của miền ảnh  $\Omega$  và  $\partial\Omega_0$  là biên của  $\Omega_0$

- Bước 2: Tính toán  $\Phi$  sử dụng phương trình (3.4)
- Bước 3: Chuẩn hóa hàm tập mức  $\Phi$  bởi hàm nhân Gaussian. Ví dụ  $\Phi = \mathcal{G}_\psi * \Phi$  trong đó  $\psi$  là độ lệch chuẩn thỏa mãn lớn hơn  $\sqrt{t}$  để tăng cường khả năng làm mịn
- Bước 4: Kiểm tra điều kiện dừng, nếu không thỏa mãn quay lại bước 2

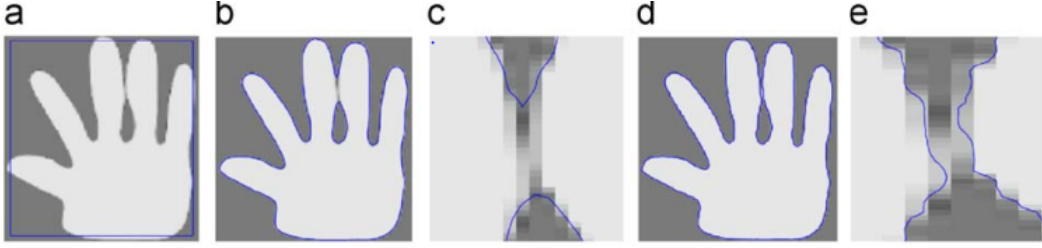
Qua quá trình thử nghiệm, tham số  $\psi$  ở bước 3 thường được chọn trong khoảng 0.45 đến

1. Nếu nhiễu lớn thì  $\psi$  nên chọn lớn hơn

### 3.3.2 Ưu nhược điểm của mô hình

So với mô hình LBF, tuy mô hình LBF không cần thao tác re-initialization không cần thiết, tuy nhiên độ phức tạp tính toán còn tương đối cao. Chi phí tính toán của mô hình LBF chủ yếu nằm ở thao tác tính toán  $\lambda_1 e_1 - \lambda_2 e_2$ . Mô hình LIF có tốc độ tính toán tốt hơn nhiều so với LBF. Đồng thời thực nghiệm cho thấy mô hình LIF có thể đạt tới việc phân chia nhỏ hơn pixel. Hình dưới cho thấy với phương pháp C-V kết quả cuối cùng thu được 2 ngón tay ở giữa bị dính vào nhau trong khi với phương pháp LIF thì chúng tách rời hoàn toàn.

**Hình 3.1:** Hình b và c thể hiện kết quả của phương pháp C-V, hình d và e thể hiện kết quả của phương pháp LIF



## 3.4 Mô hình kết hợp Global và Local

### 3.4.1 Mô hình

Cả 3 phương pháp phân vùng sử dụng biên động như đã trình bày ở trên đều có ưu nhược điểm của riêng mình. Với mô hình CV, nó hoạt động tốt với các ảnh mà cường độ xám trong và ngoài đối tượng trong ảnh đồng nhất, trong trường hợp ảnh có cường độ sáng không đồng nhất thì mô hình này thường cho ra kết quả sai. Ngược lại, mô hình LBF, LIF trình bày ở phần 2,3 lại hoạt động tốt với ảnh không đồng nhất mức xám và tỏ ra không hiệu quả với ảnh đồng nhất mức xám. Trong phần này tôi đề xuất một mô hình kết hợp mô hình Global và Local để có thể xử lý được với nhiều loại ảnh khác nhau, giảm thời gian tính toán, đồng thời tăng tính hội tụ do việc giảm bớt các thành phần dễ bị ảnh hưởng bởi đường biên ban đầu. Trong mô hình CV ta lấy thành phần fitting term và bỏ đi thành phần regularize term :

$$E^{(GIF)} = \int_{\Omega} |u_0 - c_1|^2 H(\Phi(x, y)) + |u_0 - c_2|^2 (1 - H(\Phi(x, y))) dx dy \quad (3.39)$$

Chúng ta gọi nó là GIF (Global image fitting). Ngoài ra trong mô hình này sử dụng thêm thành phần  $E^{LIF}$  trong mô hình LIF:

$$E^{LIF}(\Phi) = \frac{1}{2} \int_{\Omega} |I(x) - m_1 H_{\epsilon}(\Phi) - m_2 (1 - H_{\epsilon}(\Phi))|^2 dx \quad (3.40)$$



trong đó

$$\Phi(x, t = 0) = \begin{cases} -\rho, x \in \Omega_0 - \partial\Omega_0 \\ 0, x \in \partial\Omega_0 \\ \rho, x \in \Omega - \Omega_0 \end{cases} \quad (3.41)$$

Hàm năng lượng được đề xuất chứa thành phần global và thành phần local :

$$E^{GLIF} = \alpha E^{LIF} + (1 - \alpha) E^{GIF} \quad (3.42)$$

trong đó  $\alpha$  là số không âm nằm trong đoạn  $[0,1]$ . Khi xử lý với ảnh có cường độ ảnh đồng nhất ta sử dụng hệ số alpha cao. Trong khi xử lý với ảnh có cường độ không đồng nhất ta sử dụng hệ số alpha thấp. Thành phần LIF chứa lực địa phương (local force) để trích biên và dừng nó ở biên của vật. Thành phần này cho phép mô hình có thể xử lý với vấn đề cường độ không đồng nhất. Thành phần GIF chứa lực toàn cục (global force) cho phép dịch chuyển biên khi biên tạm ở xa biên của vật. Điều này cho phép mô hình xử lý linh động hơn với các đường cong khởi tạo ban đầu. Ảnh hưởng của các lực địa phương và lực toàn cục bổ sung cho nhau. Lực địa phương có hiệu quả khi ở gần biên của đối tượng trong khi lực toàn cục có hiệu quả khi ở xa biên của đối tượng.

Tham số độ lệch chuẩn  $\sigma$ , và tham số chuẩn hóa  $\psi$  đóng vai trò quan trọng trong mô hình. Tham số  $\sigma$  cho phép điều khiển khả năng mở rộng khu phức (region-scalability) từ một lân cận nhỏ cho đến toàn bộ miền ảnh. Tùy thuộc vào chất lượng ảnh và nội dung của ảnh mà ta chọn tham số  $\sigma$  cho phù hợp. Thực tế cho thấy với các ảnh nhiễu nhiễu hoặc độ tương phản thấp thì tham số  $\sigma$  cần được chọn lớn. Điều này sẽ gây ảnh hưởng đến tốc độ chạy thuật toán. Cùng với đó, nếu chọn  $\sigma$  quá nhỏ sẽ làm ảnh hưởng đến kết quả dẫn đến kết quả không mong muốn.

### 3.4.2 Giải số

Tiếp theo chúng ta sẽ xây dựng các bước giải số cho mô hình. Giá trị  $c_1, c_2$  làm cực tiểu hàm năng lượng trên được xác định như sau:

$$c_1(\Phi) = \frac{\int_{\Omega} u_0(x, y) H(\Phi(x, y)) dx dy}{\int_{\Omega} H(\Phi(x, y)) dx dy} \quad (3.43)$$

$$c_2(\Phi) = \frac{\int_{\Omega} u_0(x, y) (1 - H(\Phi(x, y))) dx dy}{\int_{\Omega} H(\Phi(x, y)) (1 - H(\Phi(x, y))) dx dy} \quad (3.44)$$

Cố định  $c_1, c_2$ . Đạo hàm theo  $\Phi$  và cho  $\epsilon$  tiến dần tới 0 ta có

$$\begin{aligned} \frac{\delta E^{GLIF}}{\delta \bar{\Phi}} &= \lim_{\epsilon \rightarrow 0} \frac{d}{d\epsilon} \left( \frac{1}{2} \int_{\Omega} |I(x) - m_1 H_{\epsilon}(\bar{\Phi}) - m_2 (1 - H_{\epsilon}(\bar{\Phi}))|^2 dx \right) \\ &+ \int_{\Omega} |u_0 - c_1|^2 H(\bar{\Phi}) + |u_0 - c_2|^2 (1 - H(\bar{\Phi})) dx \\ &= \lim_{\epsilon \rightarrow 0} \frac{d}{d\epsilon} \left( - \int_{\Omega} -\delta_{\epsilon}(\Phi) (I(x) - m_1 H_{\epsilon}(\Phi) - m_2 (1 - H_{\epsilon}(\Phi))) (m_1 - m_2) \psi dx \right) \\ &+ \alpha \delta_{\epsilon}(\bar{\Phi}) \int_{\Omega} (I - c_1)^2 + (I_0 - c_2)^2 \psi dx \\ &= - \left( \int_{\Omega} -\delta_{\epsilon}(\Phi) (I(x) - m_1 H_{\epsilon}(\Phi) - m_2 (1 - H_{\epsilon}(\Phi))) (m_1 - m_2) \psi dx \right) \\ &+ \alpha \delta_{\epsilon}(\Phi) \int_{\Omega} ((I - c_1)^2 + (I_0 - c_2)^2) \psi dx \end{aligned} \quad (3.45)$$

Theo đó ta có phương trình Euler-Lagarange:

$$\sigma_{\epsilon}(\Phi) \{ (I - I^{LIF})(m_1 - m_2) + \alpha((I - c_1)^2 + (I_0 - c_2)^2) \} \quad (3.46)$$

Cụ thể các bước của thuật toán sẽ như sau:

- Bước 1: Khởi tạo đường cong  $\Phi_0$  ban đầu
- Bước 2: Tính  $c_1, c_2$  theo công thức
- Bước 3: Tính giá trị  $\Phi_{n+1}$  theo  $\Phi_n$  theo công thức
- Bước 4: Chuẩn hóa hàm tập mức  $\Phi$  sử dụng hàm nhân Gauss  $\Phi = G_{\psi}\Phi$ , trong đó  $\psi$  là độ lệch chuẩn

- Bước 5: Kiểm tra điều kiện dừng. Nếu thỏa mãn quay lại bước 2

## Chương 4

# Áp dụng mô hình GLIF trong theo dõi đối tượng trong video

### 4.1 Theo dõi đối tượng trong video là gì? Tại sao phải theo dõi đối tượng trong video

Theo vết đối tượng là một vấn đề quan trọng trong lĩnh vực thị giác máy tính. Sự xuất hiện của các máy tính cấu hình mạnh và các camera chất lượng cao và sự cần thiết của việc phân tích video là điều kiện để thúc đẩy sự phát triển của các thuật toán theo vết. Có 3 bước chính trong phân tích video là: Phát hiện các đối tượng chuyển động trong video, theo vết các đối đó theo từng frame, phân tích vết để phát hiện hành động của đối tượng. Do vậy, theo vết đối tượng phù hợp với:

- Nhận dạng dựa trên chuyển động
- Giám sát tự động, hay là theo dõi một cảnh để phát hiện các hoạt động đáng ngờ hoặc các sự kiện bất thường.
- Lập chỉ mục video, tạo chú thích tự động và tìm iến video trong cơ sở dữ liệu

- Tương tác giữa con người và máy tính, hay là nhận diện cử chỉ, theo dõi mắt làm đầu vào cho máy tính.
- Theo dõi giao thông, thu thập thông kê dòng phương tiện

Ở dạng đơn giản nhất, theo vết có thể được định nghĩa là ước lượng chuyển động của đối tượng trong miền ảnh khi nó chuyển động xung quanh miền ảnh. Một cách khác, người theo dõi thực hiện gán nhãn các đối tượng được theo dõi trong các khung hình của video. Ngoài ra, tùy thuộc vào miền ảnh, việc theo dõi có thể cung cấp các đặc tính của đối tượng như: hình dạng, vị trí, hướng. Việc theo vết đối tượng có thể gặp khó khăn do một số ẩn đề như:

- Thiếu thông tin do các phép chiếu từ khung cảnh 3D ngoài đời thực vào ảnh 2D
- Ảnh nhiễu
- Những chuyển động phức tạp của đối tượng
- Thay đổi khung cảnh
- Yêu cầu xử lý thời gian thực

Điều này có thể đơn giản hơn nhờ việc đưa vào các ràng buộc về chuyển động cũng như là bề ngoài của đối tượng. Ví dụ, hầu hết các thuật toán theo vết đối tượng đều đưa vào ràng buộc tốc độ và gia tốc là hằng số.

Có nhiều phương pháp theo vết đối tượng được đề xuất. Các hướng này khác nhau ở cách tiếp cận để trả lời các câu hỏi: Đối tượng nào thích hợp để theo dõi? Các đặc tính nào của ảnh có thể sử dụng? Làm sao để mô hình hóa hình dạng, chuyển động của đối tượng trong video? Việc trả lời các câu hỏi này còn tùy thuộc vào hoàn cảnh, điều kiện