

**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**  
-----



**ĐOÀN THỊ PHƯƠNG**

**NGHIÊN CỨU MỘT SỐ PHƯƠNG PHÁP PHÂN  
ĐOẠN ẢNH MÀU**

**Chuyên ngành: Khoa học máy tính**

**Mã số: 60.48.01**

**Người hướng dẫn khoa học: PGS.TS NGÔ QUỐC TẠO**

**TÓM TẮT LUẬN VĂN THẠC SĨ**

**HÀ NỘI – 2012**

## MỞ ĐẦU

Cùng với sự phát triển ngày càng mạnh mẽ của khoa học kĩ thuật trong một vài thập kỷ gần đây, xử lý ảnh tuy là một ngành khoa học còn tương đối mới mẻ so với nhiều ngành khoa học khác nhưng hiện nay nó đang là một trong những lĩnh vực phát triển rất nhanh và thu hút sự quan tâm đặc biệt từ các nhà khoa học, thúc đẩy các trung tâm nghiên cứu, ứng dụng về lĩnh vực hấp dẫn này.

Xử lý ảnh đóng vai trò quan trọng trong nhiều ứng dụng thực tế về khoa học kĩ thuật cũng như trong cuộc sống thường ngày như: sản xuất và kiểm tra chất lượng, sự di chuyển của Robot, các phương tiện đi lại tự trị, công cụ hướng dẫn cho người mù, an ninh và giám sát, nhận dạng đối tượng, nhận dạng mặt, các ứng dụng trong y học, sản xuất, hiệu chỉnh video,...

Để xử lý được một bức ảnh thì phải trải qua nhiều khâu khác nhau tùy theo mục đích của việc xử lý, nhưng khâu quan trọng và khó khăn nhất đó là phân đoạn ảnh. Trong một số lượng lớn các ứng dụng về xử lý ảnh và hiển thị máy tính, phân đoạn đóng vai trò chính yếu như là bước đầu tiên trước khi áp dụng các thao tác xử lý ảnh mức cao hơn như: nhận dạng, giải thích ngữ nghĩa, và biểu diễn ảnh.

Phân đoạn ảnh là một thao tác ở mức thấp trong toàn bộ quá trình xử lý ảnh. Quá trình này thực hiện việc phân vùng ảnh thành các vùng rời rạc và đồng nhất với nhau hay nói cách khác là xác định các biên của các vùng ảnh đó. Các vùng ảnh đồng nhất này thông thường sẽ tương ứng với toàn bộ hay từng phần của các đối tượng thật sự bên trong ảnh. Vì thế, trong hầu hết các ứng dụng của lĩnh vực xử lý ảnh, phân đoạn ảnh luôn đóng một vai trò cơ bản và thường là bước tiền xử lý đầu tiên trong toàn bộ quá trình trước khi thực hiện các thao tác khác ở mức cao hơn như nhận dạng đối tượng, biểu diễn đối tượng, nén ảnh dựa trên đối tượng, hay truy vấn ảnh dựa vào nội dung ... Trước đây, các phương pháp phân vùng ảnh được đưa ra chủ yếu làm việc trên các ảnh mức xám do các hạn chế về phương tiện thu thập và lưu trữ. Ngày nay, cùng với sự phát triển về các phương tiện thu nhận và biểu diễn ảnh, các ảnh màu đã hầu như thay thế hoàn toàn các ảnh mức xám trong việc biểu diễn và lưu trữ thông tin do các ưu thế vượt trội hơn hẳn so với ảnh mức xám. Do đó, các kỹ thuật, thuật giải mới thực hiện việc phân vùng ảnh trên các loại ảnh màu liên tục được phát triển để đáp ứng các nhu cầu mới.

## Chương 1 - TỔNG QUAN VỀ XỬ LÝ ẢNH

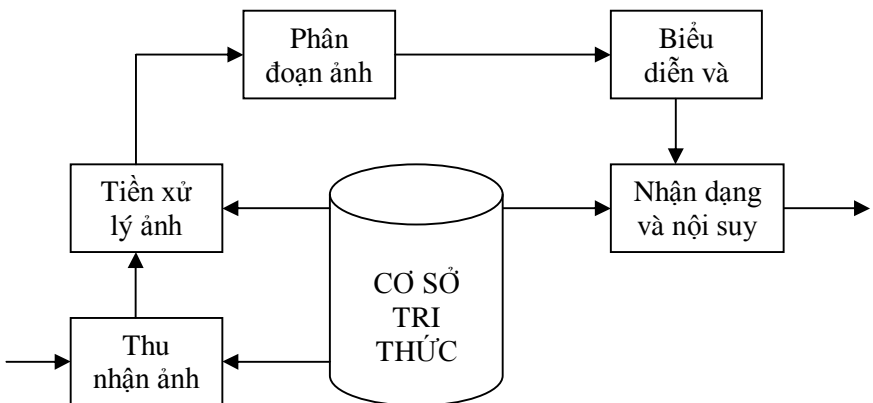
### 1.1. Giới thiệu về xử lý ảnh

Quá trình xử lý ảnh là một quá trình thao tác nhằm phân tích, biến đổi một ảnh đầu vào để đưa ra một kết quả mong muốn. Kết quả của quá trình xử lý ảnh có thể là một ảnh tốt hơn hoặc một kết luận. Điều này tùy thuộc vào mục đích yêu cầu của quá trình.

### 1.2. Quá trình xử lý ảnh

Các phương pháp xử lý ảnh bắt đầu từ các ứng dụng chính như nâng cao chất lượng ảnh và phân tích ảnh. Do vậy, quá trình xử lý ảnh bao giờ cũng bắt đầu bằng công việc thu nhận ảnh và kết thúc là việc nhận dạng ảnh hoặc một phán đoán theo ý nghĩa trên cơ sở nhận dạng ảnh. Cụ thể, các bước cơ bản trong quá trình xử lý ảnh được thể hiện thông qua hình

1.1 sau :



**Hình 1.1 - Các bước cơ bản trong xử lý ảnh.**

### ***1.2.1. Thu nhận ảnh***

Công việc cụ thể ở giai đoạn này là thu ảnh qua một bộ thu ảnh và số hoá những tín hiệu liên tục được sinh ra bởi bộ thu ảnh đó. Bộ thu ảnh có thể là máy chụp ảnh đơn sắc hay màu, máy quét ảnh, máy quay...

### ***1.2.2. Tiền xử lý ảnh***

Công việc cụ thể của bước này là cải thiện độ tương phản của ảnh, khử nhiễu. Mục đích của các công việc này là làm cho chất lượng ảnh trở lên tốt hơn chuẩn bị cho các bước xử lý tiếp theo

### ***1.2.3. Phân đoạn ảnh***

Đây là giai đoạn này tách một ảnh đầu vào thành nhiều vùng khác nhau hay còn gọi là các đối tượng để biểu diễn phân tích, nhận dạng ảnh.

### ***1.2.4. Biểu diễn và mô tả***

Ảnh sau khi số hoá sẽ được lưu vào bộ nhớ hoặc chuyển sang các khâu tiếp theo để phân tích. Nếu lưu trữ ảnh trực tiếp từ các ảnh thô đòi hỏi dung lượng bộ nhớ rất lớn và không hiệu quả theo quan điểm ứng dụng và công nghệ. Thông thường, các ảnh thô đó được biểu diễn lại theo các đặc điểm của ảnh được gọi là các đặc trưng ảnh như: biên ảnh, vùng ảnh. Các thông tin này sẽ được chọn các tính chất đặc trưng để thể hiện gọi là trích chọn đặc trưng.

### ***1.2.5. Nhận dạng và nội suy***

Nhận dạng ảnh là quá trình xác định ảnh bằng cách so sánh ảnh với mẫu chuẩn đã được lưu từ trước. Nội suy là phán đoán theo ý nghĩa trên cơ sở nhận dạng ảnh.

Một số đối tượng nhận dạng khá phổ biến hiện nay đang được áp dụng trong khoa học và công nghệ là: Nhận dạng ký tự (chữ in, chữ viết tay, chữ ký điện tử), nhận dạng văn bản, nhận dạng vân tay, nhận dạng mã vạch, nhận dạng mặt người...

### ***1.2.6. Cơ sở tri thức***

Trong nhiều khâu xử lý và phân tích ảnh, ngoài việc đơn giản hoá các phương pháp toán học đảm bảo tiện lợi cho xử lý thì chúng ta luôn hướng đến việc xây dựng hệ thống tự động tiếp nhận và xử lý theo cách của con người. Vì vậy, nhiều khâu hiện nay đã được xử lý theo các phương pháp trí tuệ nhân tạo, sử dụng cơ sở tri thức của con người.

## **1.3. Tổng quan về phân đoạn ảnh**

Phân đoạn ảnh là một thao tác ở mức thấp và là bước then chốt trong quá trình xử lý ảnh. Giai đoạn này nhằm phân tích ảnh thành những vùng rời rạc có cùng tính chất nào đó dựa vào việc xác định biên và các vùng liên thông cho từng vùng. Tiêu chuẩn để xác định các vùng liên thông có thể là cùng mức xám, cùng màu hay cùng độ nhám... Các vùng ảnh này thông

thường sẽ tương ứng với toàn bộ hay từng phần của đối tượng thật bên trong ảnh.

## **1.4. Một số khái niệm cơ bản**

### ***1.4.1. Điểm ảnh***

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 đó được chọn thích hợp sao cho mắt người cảm nhận sự liên tục về không gian và mức xám (hoặc màu) của ảnh số gần như ảnh thật.

### ***1.4.2. Độ phân giải của ảnh***

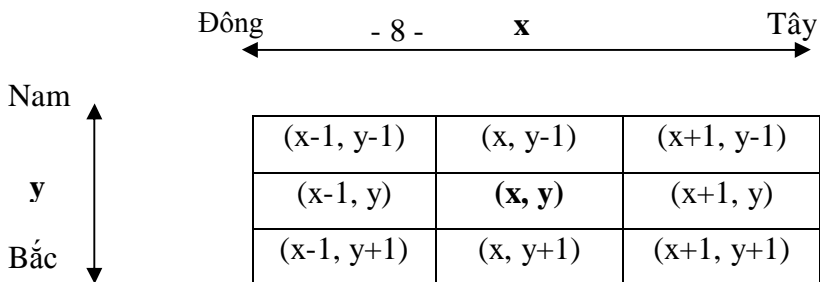
Độ phân giải (Resolution) của ảnh là mật độ điểm ảnh được ấn định trên một ảnh số được hiển thị.

### ***1.4.3. Mức xám của ảnh***

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

### ***1.4.4. Quan hệ giữa các điểm ảnh***

#### **1.4.4.1. Các lân cận của điểm ảnh (Image Neighbors)**



**Hình 1.2 – Lân cận các điểm ảnh của tọa độ  $(x,y)$**

#### 1.4.4.2. Khoảng cách giữa các điểm ảnh

- Khoảng cách Euclide
- Khoảng cách khối
- Khoảng cách bàn cờ

#### **1.4.6. Nén ảnh**

Ảnh dù ở dạng nào vẫn chiếm không gian nhớ rất lớn. Vì vậy, khi mô tả ảnh có thể sử dụng kỹ thuật nén ảnh để thu gọn dung lượng nhớ dành cho ảnh.

#### **1.5. Các định dạng cơ bản trong xử lý ảnh**

Hình ảnh khi lưu trữ dưới dạng tệp tin sẽ được số hóa. Một số dạng ảnh đã được chuẩn hóa như: ảnh GIF, BMP, PCX, IMG,...



## **Chương 2 - MỘT SỐ PHƯƠNG PHÁP PHÂN ĐOẠN ẢNH MÀU**

### **2.1. Phương pháp phân đoạn dựa trên ngưỡng cục bộ thích nghi**

Ý tưởng chính của phương pháp này là chia một ảnh ban đầu ra làm nhiều vùng nhỏ. Sau đó, dựa vào ngưỡng cục bộ của từng vùng để trộn các vùng nhỏ thành các vùng có cường độ xám nằm trong một ngưỡng cho phép.

#### **2.1.1. Phân đoạn sơ khởi bằng Watershed**

Dữ liệu đầu vào của giải thuật Watershed là một ảnh xám. Vì vậy, trước tiên phải biến đổi ảnh đầu vào  $I$  thành ảnh xám. Sau đó, dùng giải thuật tìm cạnh Canny để lấy cường độ gradient, kí hiệu là  $I_G$ . Cuối cùng áp dụng giải thuật Watershed phân đoạn ảnh thành  $n$  vùng không trùng lặp nhau.

#### **2.1.2. Trộn các vùng**

Đồ thị các vùng lân cận (Region Adjacency Graph - RAG) là cấu trúc dữ liệu được sử dụng để thể hiện quá trình phân vùng ảnh. Đồ thị RAG biểu diễn tập các vùng  $R_i^0$ ,  $i=1, \dots, n$  dưới dạng đồ thị vô hướng  $G=(V,E)$ . Trong đó,  $V=\{1,2, \dots, n\}$ , mỗi node thuộc  $V$  tương ứng với một vùng. Mỗi cạnh  $e(i,j) \in E$  nếu  $i, j \in V$  và miền  $R_i^{m_i}$  với  $R_j^{m_j}$  là lân cận

nhau. Mỗi cạnh  $e(i,j)$  có một trọng số bằng giá trị của hàm  $f(R_i^{m_i}, R_j^{m_j})$ .

Quá trình trộn dựa vào giải thuật Kruskal tìm cây khung nhỏ nhất MST (Minimum Spanning Tree) cho đồ thị các vùng lân cận RAG.

### ***2.1.3. Tìm ngưỡng cục bộ thích nghi***

Quá trình trộn đã được mô tả như giải thuật tìm cây khung MST trong phần trên. Tuy nhiên, vẫn chưa biết cách xác định vùng nào không trộn được và thời điểm nào thì không trộn. Như vậy, cần có cơ chế tự động rút trích thông tin về ngưỡng cục bộ thông qua việc theo dõi sự thay đổi của mỗi vùng trong quá trình trộn. Các ngưỡng này sẽ cho biết có thể trộn một vùng hay không. Như thế, các ngưỡng này giúp hình thành phân vùng hoàn chỉnh cuối cùng.

#### **Cách tính ngưỡng cục bộ thích nghi**

Để xác định được ngưỡng cục bộ thích nghi cần dùng một phương thức động. Phương thức này dựa trên các thuộc tính riêng của các vùng trong suốt quá trình trộn.

#### **Quá trình hồi quy**

## **2.2. Phương pháp phân đoạn dựa theo đường biên**

### **2.3.1. Giới thiệu**

Biên là một đường viền giữa hai miền đồng nhất. Tách biên chính là quá trình nhận dạng và định vị các điểm gián đoạn rõ ràng trong một ảnh.

Tách biên là một công cụ cơ bản sử dụng trong hầu hết các ứng dụng xử lý ảnh. Nó là bước tiền xử lý cho việc rút trích thông tin và phân đoạn đối tượng. Đây là quá trình trích các đường bao của một đối tượng và biên giữa các đối tượng với nền của ảnh. Bộ lọc tách biên có thể cũng được sử dụng để cải thiện ảnh mờ và nhiều ứng dụng khác nữa.

### **2.3.2. Phát hiện điểm gián đoạn**

2.3.2.1. Phát hiện điểm

2.3.2.2. Phát hiện đường

2.3.2.3. Phát hiện biên

2.3.2.4. Toán tử Gradient

### **2.3.3. Kỹ thuật phát hiện biên**

2.3.3.1. Các toán tử Sobel

Việc tính toán đạo hàm từng phần của gradient có thể gần đúng với các ảnh kỹ thuật số bằng việc sử dụng các toán tử Sobel.

### 2.3.3.2. Roberts Cross

Toán tử Roberts Cross thực hiện tính toán độ đo gradient của một ảnh trong không gian hai chiều một cách đơn giản và rất nhanh. Đầu vào của toán tử là một ảnh mức xám. Các giá trị điểm ảnh tại mỗi điểm ở đầu ra biểu thị cho độ lớn của gradient của ảnh đầu vào tại điểm đó.

### 2.3.3.3. Toán tử Laplacian

Laplacian thường được sử dụng để thiết lập một điểm ảnh trên phía bên tối hoặc sáng của một cạnh.

### 2.3.3.5. Toán tử Kiresch

### 2.3.3.6. Kỹ thuật phát hiện biên Canny

Kỹ thuật Canny là phương thức rất quan trọng để tìm các biên bằng việc tách nhiễu từ ảnh trước khi tìm các biên của ảnh mà không ảnh hưởng đến các đặc điểm của biên trong ảnh, sau đó áp dụng việc tìm các biên và giá trị tới hạn cho ngưỡng.

### 2.3.3.7. Kỹ thuật EMT

Trong các ảnh khi có nhiều hơn một miền đồng nhất (ví dụ một ảnh có nhiều đối tượng với các mức xám khác nhau) hoặc có sự thay đổi về độ sáng giữa các đối tượng và nền của nó. Trong trường hợp này, một phần của các đối tượng có thể được trộn với nền hoặc các phần của nền có thể như một đối tượng.

Từ thực tế như trên, các kỹ thuật chọn ngưỡng động trở lên tốt hơn cho các ảnh có các miền đồng nhất lớn.

### ***2.3.4. Mô phỏng thực nghiệm***

2.3.4.1. Thủ tục kiểm tra

2.3.4.1. Các kết quả mô phỏng

## **2.3. Phương pháp phân đoạn dựa theo miền đồng nhất**

### ***2.3.1. Giới thiệu***

Phương pháp phân đoạn dựa theo miền đồng nhất này xác định các miền đồng nhất có ưu thế trong một ảnh màu bằng giải thuật phân cụm Fuzzy C mean. Với phương thức này, số các cụm sẽ quyết định số quy luật. Vì vậy, từ thông tin của cụm xác định luật IF – THEN (nếu – thì). Mỗi điểm ảnh sau khi được đánh giá bởi luật này thì kết quả cuối cùng được lưu lại. Cuối cùng, ngưỡng có hiệu suất cao và giải thuật PDVE được áp dụng trên các kết quả và các điểm ảnh được phân lớp vào các lớp thích hợp.

### ***2.3.2. Các bước của phương pháp***

2.3.2.1. Giải thuật Fuzzy C Mean và ứng dụng của nó với phương pháp

2.3.2.2. Xác định luật

2.3.2.3. Xử lý ngưỡng cho việc phân lớp cuối cùng

2.3.2.4. Các bước của giải thuật PDVE cho việc phân tích biểu đồ

## **2.4. Phương pháp phân đoạn dựa vào đồ thị**

### **2.4.1. Giới thiệu**

Phân đoạn ảnh dựa vào đồ thị là một phương pháp tiếp cận khá hiện đại dựa trên thuộc tính non-local của ảnh đầu vào. Phương pháp này phát hiện ra biên giữa hai vùng của ảnh bằng cách so sánh sự khác nhau giữa nội vùng (inter-component) với các vùng khác. Thuật toán phân đoạn dựa vào đồ thị tuân theo chiến lược tham lam, có thời gian chạy gần như tuyến tính, nhưng vẫn đảm bảo được việc phân đoạn chính xác và hiệu quả.

### **2.4.2. Phân đoạn dựa vào đồ thị**

Cho  $G = (V, E)$  là một đồ thị vô hướng với các đỉnh  $v_i \in V$ , là tập hợp các phần tử cần được phân đoạn và các cạnh  $(v_i, v_j) \in E$ , tương ứng với các cặp đỉnh lân cận nhau. Mỗi cạnh  $(v_i, v_j) \in E$  có một trọng số tương ứng, trọng số là một số không âm đo sự khác nhau giữa hai phần tử lân cận  $v_i$  và  $v_j$ , ký hiệu  $w(v_i, v_j)$ . Ở đây trọng số của các cạnh đo sự khác nhau giữa hai điểm nối bởi cạnh đó (có nhiều mức độ khác nhau: màu sắc, vị trí, sự vận động hoặc các thuộc tính khác).

Như vậy phân đoạn một bức ảnh là việc phân chia  $V$  thành các thành phần, mà mỗi thành phần (hoặc miền)  $C \in V$

tương đương với một thành phần liên thông trong đồ thị  $G' = \langle V, E' \rangle$ , với  $E' \subseteq E$ .

#### **2.4.2. Tính chất của so sánh cặp miền**

Để đánh giá dễ dàng về một đường biên giữa hai thành phần trong một phân đoạn cần định nghĩa một thuộc tính D. Tính chất này dựa vào độ đo sự khác nhau giữa các phần tử dọc theo một đường biên của hai thành phần liên quan nhằm đo sự khác nhau giữa các phần tử lân cận trong mỗi thành phần. Kết quả là so sánh sự khác nhau giữa nội vùng (inter-component) với sự khác nhau với các vùng khác.

#### **2.4.3. Giải thuật và các thuộc tính**

Phần này sẽ mô tả và phân tích một giải thuật cho việc phân đoạn ảnh sử dụng tiêu chuẩn D được trình bày trong phần trên.

#### **2.4.4. Độ phức tạp tính toán**

Thời gian thực hiện của thuật toán này được chia làm hai phần:

Một là thời gian cần thiết để sắp xếp dãy trọng số theo chiều không giảm.

Hai là thời gian thực hiện bước 1-3. Để kiểm tra được hai đỉnh có cùng chung trong một thành phần hay không cần sử dụng biến set-find trên mỗi đỉnh nhằm lưu lại số hiệu thành phần mà đỉnh đó đang phụ thuộc vào.

#### ***2.4.5. Kết quả cho các đồ thị lân cận gần nhất***

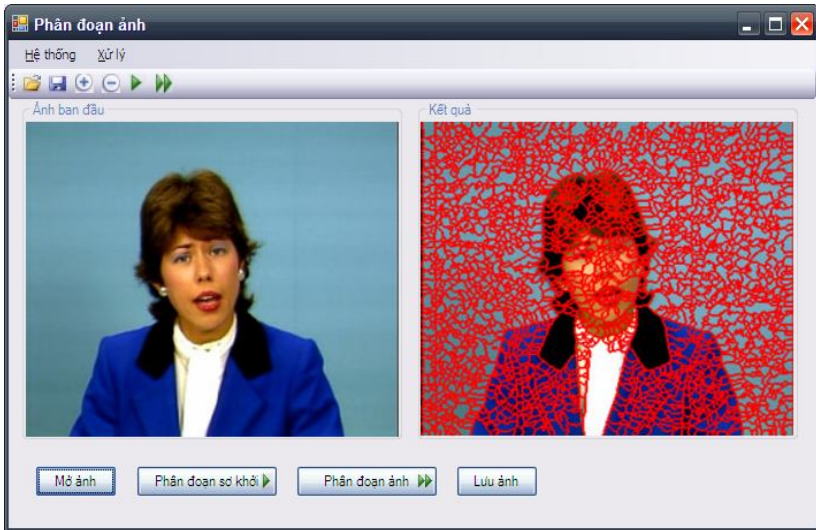
Một giải pháp chung cho phân đoạn ảnh là dựa vào đồ thị mỗi điểm ảnh và sau khi tìm các cụm của các điểm giống nhau. Phần này sẽ kiểm tra bằng việc sử dụng giải thuật phân đoạn dựa vào đồ thị để tìm các cụm các điểm giống nhau.



## Chương 3 - CÀI ĐẶT THỬ NGHIỆM

### 3.1. Phân đoạn sơ khởi áp dụng giải thuật Watershed

Giải thuật Watershed đã được trình bày cụ thể trong chương 2. Dưới đây là chương trình thử nghiệm cho quá trình thực hiện và kết quả sau khi áp dụng giải thuật:



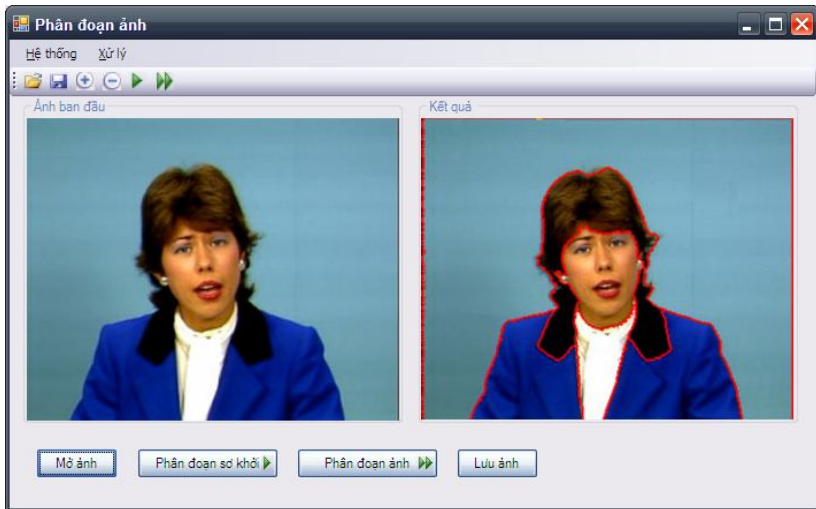
**Hình 3.1 – Giao diện thực hiện chức năng phân đoạn sơ khởi**

### 3.2. Giải thuật tìm MST của Kruskal

Việc trộn  $n$  vùng đã được phân đoạn trong bước phân đoạn sơ khởi bằng Watershed thành một số vùng nhất định sẽ dựa trên việc tìm ra cây khung nhỏ nhất MST của đồ thị vô hướng  $G=(V,E)$ . Trong đó mỗi node của đồ thị tương ứng với

một vùng. Mỗi cạnh nối hai đỉnh thể hiện cho hai vùng lân cận. Mỗi cạnh có một trọng số bằng giá trị của hàm sai khác  $f(R_i^{m_i}, R_j^{m_j})$  thể hiện cho sự khác nhau của giá trị Hue và độ gradient giữa hai vùng. Quá trình tạo cây khung nhỏ nhất MST chính là trình tự trộn các vùng trong ảnh.

### 3.3. Kết quả thực nghiệm



**Hình 3.2 – Giao diện thực hiện chức năng phân đoạn ảnh.**

### 3.4. Kết quả đạt được và hướng phát triển

Trong phần cài đặt thử nghiệm này, tôi đã hệ thống lại các bước trong phương pháp phân đoạn dựa trên ngưỡng cục

bộ thích nghi, đồng thời cài đặt thử nghiệm được một số thuật toán cơ bản trong phương pháp này.

Thời gian tới, tôi sẽ tiếp tục cài đặt hoàn thiện các công việc trong phương pháp và cài đặt thử nghiệm một số phương pháp khác đã được trình bày trong đề tài.

## **KẾT LUẬN**

### **KẾT LUẬN:**

Luận văn nghiên cứu về một số phương pháp phân đoạn ảnh màu để thực hiện việc phân đoạn trong quá trình xử lý ảnh. Mỗi một bức ảnh với các đặc điểm khác nhau và các miền trong ảnh theo quan sát của con người cũng có những đặc trưng khác nhau trong việc phân biệt chúng. Mỗi một phương pháp phân đoạn ảnh có sự phù hợp nhất định với một số ảnh nào đó. Vì vậy, khi muốn phân đoạn ảnh, việc lựa chọn một phương pháp phân đoạn phù hợp có ảnh hưởng lớn đến kết quả thu được.

Trong luận văn này, tôi đã trình bày về bốn phương pháp phân đoạn khác nhau. Trong đó, tôi đã nghiên cứu và trình bày chi tiết về phương pháp phân đoạn dựa trên ngưỡng cục bộ thích nghi. Các phương pháp phân đoạn còn lại tôi đã đưa ra và nắm được tư tưởng chính của từng phương pháp.

### **KIẾN NGHỊ VÀ HƯỚNG PHÁT TRIỂN:**

Với bước đầu nghiên cứu cài đặt thử nghiệm chương trình phân đoạn ảnh màu tôi đã nhận thấy nghiên cứu của mình còn có nhiều thiếu sót. Trong thời gian tới, tôi sẽ tiếp tục phát triển đề tài với phương hướng cụ thể như sau:

Nghiên cứu hoàn chỉnh lý thuyết về các phương pháp đã

được đưa ra trong chương 2 và có sự đánh giá, so sánh các phương pháp đó.

Nghiên cứu mở rộng một số phương pháp khác trong việc phân đoạn ảnh màu.

Cài đặt một chương trình hoàn chỉnh cho việc phân đoạn ảnh có tích hợp các phương pháp đã được nghiên cứu. Mỗi một ảnh đầu vào muốn được phân đoạn sẽ được đánh giá và lựa chọn một phương pháp phân đoạn phù hợp để thực hiện.