



Báo cáo bài tập lớn về

CHƯƠNG TRÌNH DÒ BIÊN ẢNH

Môn: XỬ LÝ TÍN HIỆU SỐ

Năm học: 2020 – 2021

SINH VIÊN THỰC HIỆN:

Nguyễn Thị Thu Duyên – 1711091

Lê Thị Tuyết Ngọc – 18110163

Dưới sự hướng dẫn của:

ThS. Nguyễn Ngọc Long

LỜI CẢM ƠN

Báo cáo này sẽ không thể thực hiện được nếu không có sự giúp đỡ của nhiều người. Rất cảm ơn giảng viên môn Xử lý tín hiệu số, thầy Nguyễn Ngọc Long, đã tận tình hướng dẫn và giúp chúng em cải thiện bài báo cáo này. Cảm ơn các bạn cùng lớp đã cùng chúng tôi thảo luận về các thiếu sót trong lúc làm việc. Cảm ơn trường Đại học Khoa học Tự nhiên, đặc biệt là khoa Toán – Tin học, đã tạo điều kiện cho chúng em được học tập và thực hành môn học bổ ích này.

MỤC LỤC

Mở đầu	7
Danh sách hình ảnh	5
Danh sách bảng	6
CHƯƠNG 1: GIỚI THIỆU	8
1.1 Đề tài	8
1.2 Mô tả vấn đề	8
1.3 Giới thiệu chương	9
 CHƯƠNG 2: KIẾN THỨC TỔNG QUAN	 10
2.1 Xử lý tín hiệu và Xử lý tín hiệu số	10
2.2 Xử lý hình ảnh kỹ thuật số	10
2.3 Mô hình màu RGB	13
 CHƯƠNG 3: CHƯƠNG TRÌNH XÁC ĐỊNH BIÊN	 14
3.1 Khái niệm biên	14
3.2 Quy trình và Phương pháp xác định biên	14
3.3 Phương pháp xác định biên Sobel	15
3.4 Phương pháp xác định biên Prewitt	17
3.5 Xác định biên ảnh màu	18
3.6 Ngưỡng	19
3.7 Quy trình xác định biên ảnh xám	21
3.8 Quy trình xác định biên ảnh màu	22

3.9 Quy trình xác định biên video và webcam	23
CHƯƠNG 4: KẾT QUẢ THỰC NGHIỆM	25
4.1 Giới thiệu dữ liệu	25
4.2 Kết quả trên tập dữ liệu ảnh xám	27
4.3 Kết quả trên tập dữ liệu ảnh màu	29
4.4 Kết quả trên tập dữ liệu video và webcam	31
CHƯƠNG 5: THẢO LUẬN	32
5.1 Ưu điểm của chương trình	32
5.2 Hạn chế của chương trình	32
5.3 Hướng phát triển	32
KẾT LUẬN	33
NGUỒN THAM KHẢO	34

Danh sách hình ảnh

Hình 1. Hệ thống xử lý ảnh kỹ thuật số	12
Hình 2. Biểu diễn mô hình màu RGB trên hệ trục tọa độ	14
Hình 3. Một số dạng đường biên	15
Hình 4. Quy trình xác định biên	17
Hình 5. Giá trị cường độ của các pixel trong ảnh	20
Hình 6. Quy trình xác định biên ảnh xám	23
Hình 7. Quy trình xác định biên ảnh màu	23
Hình 8. Quy trình xác định biên dành cho video hoặc webcam	24
Hình 9. Một số ảnh từ bộ dữ liệu hình ảnh	27

Danh sách bảng

Bảng 1. Thông số dữ liệu hình ảnh	26
Bảng 2. Thông số dữ liệu video	27
Bảng 3. Thông số video được ghi lại từ webcam	28
Bảng 4. Một số kết quả của quy trình xác định biên ảnh xám	29
Bảng 5. Một số kết quả của quy trình xác định biên ảnh màu	31

Mở đầu

Xác định đường biên là một nhiệm vụ cơ bản trong xử lý hình ảnh, phân tích ảnh, nhận dạng mẫu và thị giác máy tính, đặc biệt là trong các lĩnh vực phát hiện và trích xuất đối tượng. Do đó, một số thuật toán đã được đề xuất nhằm giải quyết bài toán này đã được đưa ra. Trong báo cáo này, chúng tôi thực hiện khảo sát một số phương pháp xác định biên trên tập dữ liệu hình ảnh, video và webcam, từ đó đưa ra những so sánh, đánh giá tổng quan nhất về từng phương pháp đã áp dụng.

Từ khóa - xử lý tín hiệu số ; xử lý ảnh ; đường biên .

CHƯƠNG 1: GIỚI THIỆU

1.1 Đề tài

Viết chương trình dò biên ảnh

Yêu cầu chính:

- Dò biên ảnh độ xám (grayscale) ảnh có kích thước $2^n \times 2^n$

Ví dụ: Ảnh có kích thước 256x256

Yêu cầu phụ và mở rộng:

- Dò biên ảnh độ màu (24bit màu) ảnh có kích thước hình chữ nhật bất kỳ
- Có thể tùy chọn ngưỡng dò biên

1.2 Mô tả vấn đề

Có hai dạng thông tin cơ bản mà con người nhận được hằng ngày, đó là âm thanh và hình ảnh, từ hai nguồn thông tin ấy, con người có thể đưa ra những đánh giá và quyết định cho các vấn đề và sự việc liên quan. Các loại máy móc hiện nay cũng đã có thể lưu trữ hai dạng thông tin này, tuy nhiên, việc đưa ra các đánh giá và kết luận về nội dung của chúng vẫn còn là một mục tiêu to lớn của xử lý tín hiệu nói chung và xử lý tín hiệu số nói riêng.

Trước khi làm được tất cả những điều đó, ta cần xác định được đường biên ảnh. Xác định biên là việc phát hiện những thay đổi rõ nét về độ sáng của hình ảnh. Mục đích của việc làm này là để nắm bắt những sự kiện quan trọng và những thay đổi về đặc tính của hình ảnh. Đây là bước cơ sở cho quá trình xử lý ảnh nói chung và quá trình phân đoạn ảnh nói riêng. Nhờ xác định đường biên, ta phân biệt được các vùng khác nhau, đối tượng với nền, các đối tượng với nhau, định vị được đối tượng và cuối cùng là nhận dạng được đối tượng.

Trong trường hợp lý tưởng, kết quả của việc áp dụng bộ dò cạnh vào hình ảnh có thể dẫn đến một tập hợp các đường cong được kết nối cho biết ranh giới của các đối tượng, ranh giới của các dấu bề mặt cũng như các đường cong tương ứng với sự không liên tục trong định hướng bề mặt. Do đó, việc áp dụng thuật toán phát hiện cạnh cho hình ảnh có thể làm giảm đáng kể lượng dữ liệu được xử lý và do đó có thể lọc ra thông tin có thể được

coi là ít liên quan hơn, trong khi vẫn bảo toàn các đặc tính cấu trúc quan trọng của hình ảnh. Nếu bước phát hiện cạnh thành công, nhiệm vụ giải thích nội dung thông tin trong ảnh gốc do đó có thể được đơn giản hóa đáng kể. Tuy nhiên, không phải lúc nào cũng có thể có được những cạnh lý tưởng như vậy từ những hình ảnh đời thực có chất lượng và độ phân giải thấp.

1.3 Giới thiệu chương

Phần còn lại của báo cáo này được chia thành các chương như sau

- Trong Chương 2, chúng tôi sẽ giới thiệu các kiến thức tổng quan về xử lý tín hiệu và quy trình xác định biên.
- Trong Chương 3, chúng tôi sẽ trình bày về chương trình dò biên trên dữ liệu hình ảnh, video và webcam.
- Trong Chương 4, chúng tôi sẽ giới thiệu dữ liệu và kết quả thực nghiệm trên bộ dữ liệu đã đề cập.
- Trong Chương 5, chúng tôi sẽ thảo luận về ưu - khuyết điểm của chương trình và hướng phát triển trong tương lai.

CHƯƠNG 2: KIẾN THỨC TỔNG QUAN

2.1 Xử lý tín hiệu và Xử lý tín hiệu số

Xử lý tín hiệu là một trường con kỹ thuật điện tập trung vào việc phân tích, sửa đổi và tổng hợp các tín hiệu như âm thanh, hình ảnh và các phép đo khoa học. Các kỹ thuật xử lý tín hiệu có thể được sử dụng để cải thiện khả năng truyền, hiệu quả lưu trữ và chất lượng chủ quan và cũng để nhấn mạnh hoặc phát hiện các thành phần quan tâm trong một tín hiệu được đo.

Xử lý tín hiệu kỹ thuật số là xử lý các tín hiệu được lấy mẫu theo thời gian rời rạc đã được số hóa. Quá trình xử lý được thực hiện bởi các máy tính đa năng hoặc bằng các mạch kỹ thuật số như ASIC, mảng cổng lập trình trường hoặc bộ xử lý tín hiệu kỹ thuật số chuyên dụng (chip DSP). Ví dụ về các thuật toán bao gồm biến đổi Fourier nhanh (FFT), bộ lọc đáp ứng xung hữu hạn (FIR), bộ lọc đáp ứng xung vô hạn (IIR) và các bộ lọc thích ứng như bộ lọc Wiener và Kalman. Các ứng dụng DSP bao gồm xử lý âm thanh và giọng nói, xử lý sonar, radar và các mảng cảm biến khác, ước tính mật độ phổ, xử lý tín hiệu thống kê, xử lý hình ảnh kỹ thuật số, nén dữ liệu, mã hóa video, mã hóa âm thanh, nén hình ảnh, xử lý tín hiệu cho viễn thông, hệ thống điều khiển, y sinh kỹ thuật và địa chấn học, trong số những người khác. Xử lý tín hiệu kỹ thuật số cũng là nền tảng của công nghệ kỹ thuật số, chẳng hạn như viễn thông kỹ thuật số và truyền thông không dây.

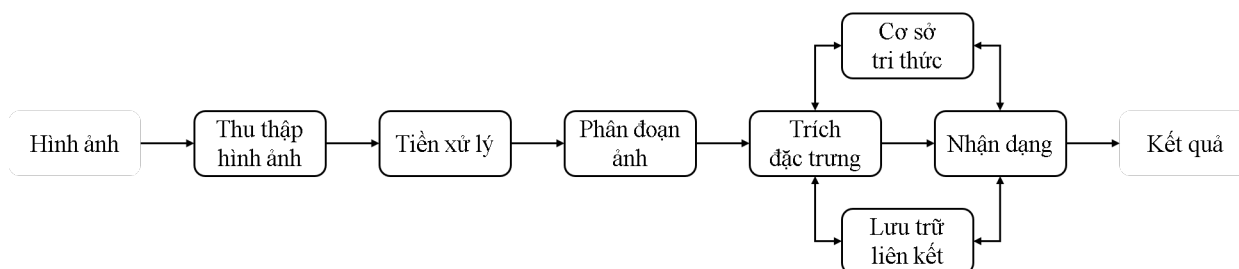
2.2 Xử lý hình ảnh kỹ thuật số

Một hình ảnh được xem như một tín hiệu hai chiều. Nó được định nghĩa bởi hàm toán học $f(x, y)$ trong đó x và y là hai tọa độ theo chiều ngang và chiều dọc. Giá trị của $f(x, y)$ tại bất kỳ điểm nào là giá trị pixel tại điểm đó của hình ảnh. Như vậy, thực ra hình ảnh chẳng qua là một mảng số hai chiều nằm trong khoảng từ 0 đến 255. Ngoài ra, kích thước của hình ảnh chính là kích thước của mảng hai chiều này.

Xử lý hình ảnh kỹ thuật số là một nhánh con của xử lý tín hiệu kỹ thuật số, dùng để chỉ việc sử dụng máy tính kỹ thuật số để xử lý hình ảnh kỹ thuật số thông qua các thuật toán. Nó cho phép một loạt các thuật toán được áp dụng cho dữ liệu đầu vào và có thể tránh các vấn đề như sự tích tụ của nhiễu và biến dạng trong quá trình xử lý. Vì hình ảnh được xác định trên hai chiều (hoặc nhiều hơn) nên việc xử lý hình ảnh kỹ thuật số có thể được mô hình hóa dưới dạng hệ thống đa chiều. Sự ra đời và phát triển của xử lý ảnh kỹ thuật số dựa trên ba yếu tố chính. Một là sự phát triển của máy tính. Hai là sự phát triển của toán học.

Ba là nhu cầu về hàng loạt các ứng dụng trong các lĩnh vực nông nghiệp, công nghiệp, y học...

Hệ thống xử lý ảnh kỹ thuật số bao gồm bảy bước: thu thập hình ảnh, tiền xử lý, phân đoạn ảnh, trích đặc trưng, lưu trữ liên kết, áp dụng tri thức và nhận dạng được mô tả trong Hình 1.



Hình 1. Hệ thống xử lý ảnh kỹ thuật số

Bước 1: Thu thập hình ảnh

Đầu tiên, ta cần sử dụng các thiết bị thu nhận ảnh để chuyển các thông tin dưới dạng hình ảnh thành các cấu trúc lưu trữ được trong máy tính và được hiển thị ra màn hình, máy in,... Ảnh có thể thu nhận từ vệ tinh, các bộ cảm biến, camera, máy chụp ảnh đơn sắc (màu), hay các tranh, ảnh được quét trên máy quét ảnh. Nếu ảnh thu nhận được chưa phải là dạng số hóa ta phải chuyển đổi hay số hóa ảnh trước khi chuyển sang giai đoạn tiếp theo.

Bước 2: Tiền xử lý

Ở bước này, ảnh sẽ được cải thiện bởi các kỹ thuật xử lý ảnh để làm ảnh tốt hơn tùy theo mục đích sử dụng nhằm phục vụ cho quá trình xử lý tiếp theo. Một số kỹ thuật cơ bản trong quá trình tiền xử lý là:

- *Cân bằng độ sáng*: khắc phục hậu quả của sự chiếu sáng không đồng đều.
- *Khử nhiễu*: Nhiễu được chia làm 2 loại cơ bản là nhiễu hệ thống và nhiễu ngẫu nhiên. Trong đó, nhiễu hệ thống là nhiễu có quy luật có thể khử bằng các phép biến đổi Fourier và loại bỏ các đỉnh điểm. Đối với nhiễu ngẫu nhiên – vết bẩn không rõ nguyên nhân thì được khắc phục bằng các phép lọc (lọc trung bình, lọc trung vị,...).
- *Hiệu chỉnh độ xám*: có thể tăng hay giảm số độ xám nhằm khắc phục tính không đồng bộ gây nên từ hiệu ứng của thiết bị thu nhận hình ảnh hoặc độ tương phản giữa các vùng ảnh.
- *Chuẩn hóa*: độ lớn, hình dạng và màu sắc.

- *Nấn chỉnh hình học*: ảnh thu nhận thường bị biến dạng do các thiết bị quang học và điện tử, để khắc phục điều này người ta sử dụng các phép chiếu được xây dựng trên tập các điểm điều khiển.

Bước 3: Phân đoạn ảnh

Phân đoạn ảnh là bước cơ bản trong toàn bộ hệ thống xử lý ảnh nhằm chia ảnh thành các vùng không trùng lặp, mỗi vùng gồm 1 nhóm pixel liên thông và đồng nhất theo 1 tiêu chí nào đó. Ví dụ: đồng nhất về màu sắc, độ xám. Nói một cách khác, phân đoạn ảnh có thể được hiểu là xác định các biên của các vùng ảnh đó. Sau khi phân đoạn mỗi pixel chỉ thuộc về một vùng duy nhất.

Kết quả của quá trình phân đoạn ảnh thường được chia thành hai dạng. Một là hàm chứa biên của vùng ảnh. Hai là tập hợp tất cả các điểm ảnh thuộc về vùng ảnh. Trong cả hai trường hợp, sự chuyển đổi dữ liệu thô này thành một dạng thích hợp cho việc xử lý trong máy tính là vô cùng cần thiết, cụ thể:

- *Biểu diễn dạng biên* cho một vùng phù hợp với những ứng dụng chỉ quan tâm đến các đặc trưng hình dạng bên ngoài của đối tượng, ví dụ như các góc cạnh và điểm uốn trên biên.
- *Biểu diễn dạng vùng* 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 nó.
- Trong một số ứng dụng thì cả hai cách biểu diễn trên đều cần thiết.

Bước 4: Trích đặc trưng

Dựa vào kết quả phân đoạn ảnh, kết hợp với các kỹ thuật xử lý và phân tích dữ liệu để chọn ra các vector đặc trưng biểu diễn các thông tin cần thiết trong quá trình xử lý. Nhờ đó việc nhận dạng sẽ chính xác hơn, tốc độ tính toán cao và dung lượng nhớ lưu trữ giảm xuống.

Các đặc trưng được trích chọn tùy theo mục đích nhận dạng trong hệ thống xử lý ảnh. Sau đây là một vài đặc trưng:

- *Đặc trưng không gian*: Phân bố độ xám, phân bố xác suất, biên độ, điểm uốn...
- *Đặc trưng biên và đường biên*: Biểu diễn đường biên của đối tượng vì vậy rất hữu ích trong việc trích chọn các thuộc tính bất biến được dùng khi nhận dạng đối tượng.

Bước 5: Lưu trữ liên kết

Bộ nhớ liên kết là những bộ nhớ có thể giải quyết bằng nội dung. Đó là khả năng chuyển

từ biểu diễn nội bộ sang biểu diễn khác. Điều này cũng dễ suy ra một biểu diễn phức tạp từ một phần hình thành cơ sở của bộ nhớ liên kết. Chức năng cơ bản của nó là lưu trữ các cặp mẫu liên kết về việc trình bày các kích thích tương ứng mẫu.

Bước 6: Cơ sở tri thức

Quy tắc phân loại được lưu trữ trong dữ liệu cơ sở tri thức. Dữ liệu cơ sở tri thức không chỉ tương tác với các giai đoạn trích xuất và nhận dạng tính năng mà còn tương tác với bộ nhớ liên kết. Các cơ sở tri thức đơn giản như mô tả chi tiết các vùng của hình ảnh nơi chứa thông tin quan tâm.

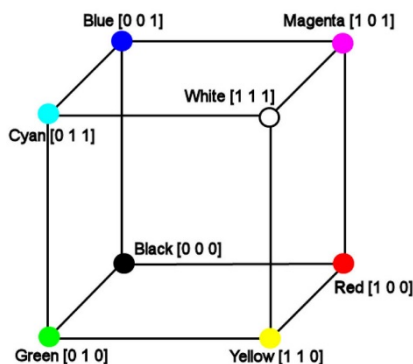
Bước 7: Nhận dạng

Nhận dạng liên quan đến phân loại. Nó gắn nhãn đối tượng dựa trên thông tin được cung cấp bởi các bộ mô tả. Các kỹ thuật phân loại thường được chia thành hai nhóm: có giám sát và không giám sát. Trong chế độ giám sát, bộ phân loại học với các tập huấn luyện. Ở chế độ không được giám sát, bộ phân loại học mà không cần tập hợp đào tạo.

Các phương pháp mô tả dựa trên các quy tắc phân loại ánh xạ các vector đặc trưng đầu vào thành các danh mục đầu ra tạm gọi là kết quả.

2.3 Mô hình màu RGB

Để tạo thành một màu với RGB, ba chùm ánh sáng (đỏ, lục và lam) phải được chồng lên. Mỗi trong số ba chùm tia được gọi là một thành phần của màu đó, và mỗi chùm có thể có cường độ tùy ý.



Hình 2. Biểu diễn mô hình màu RGB trên hệ trục tọa độ

Mô hình màu RGB có tính chất cộng theo nghĩa là ba chùm ánh sáng được cộng lại với nhau và quang phổ ánh sáng của chúng cộng với bước sóng của bước sóng để tạo nên phổ của màu cuối cùng.

CHƯƠNG 3: CHƯƠNG TRÌNH XÁC ĐỊNH BIÊN

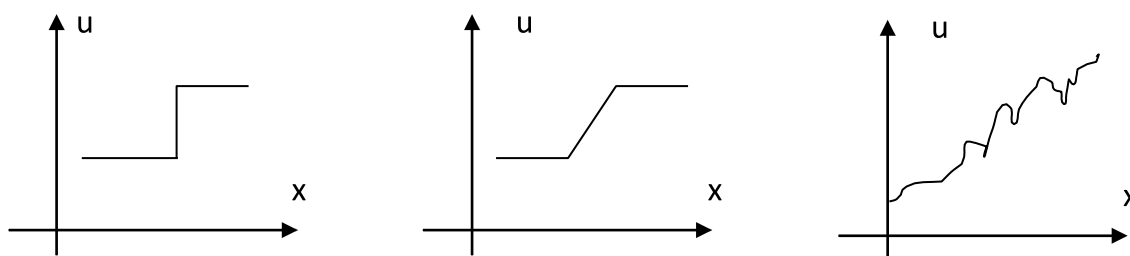
3.1 Khái niệm biên

Biên có thể được tạo ra bởi bóng tối, kết cấu hình học... Biên cũng có thể được định nghĩa là không liên tục ở cường độ hình ảnh do sự thay đổi trong cấu trúc hình ảnh. Biên trong một hình ảnh thường xảy ra với độ phân giải hoặc quy mô khác nhau và đại diện cho quá trình chuyển đổi của độ xám khác nhau, hay mức độ gradient. Tuy nhiên, nhìn chung biên có thể được định nghĩa như sau:

- *Điểm biên*: điểm có sự thay đổi đột ngột về độ xám.

Ví dụ: đối với ảnh đen trắng, một điểm được gọi là điểm biên nếu nó là điểm đen có ít nhất một điểm trắng bên cạnh.

- *Đường biên của đối tượng*: được tạo thành bởi một tập các điểm biên liên tiếp.



a. Đường biên lý tưởng

b. Đường biên dốc

c. Đường biên thực

Hình 3. Một số dạng đường biên

Việc xác định các đường biên là bước đầu tiên trong rất nhiều thuật toán trong computer vision như nhận dạng khuôn mặt, xác định chương ngại vật... Ngoài ra, dò biên còn được ứng dụng trong việc xác định dấu vân tay, xử lý ảnh vệ tinh, robotics vision và xử lý ảnh y khoa.

3.2 Quy trình và Phương pháp xác định biên ảnh

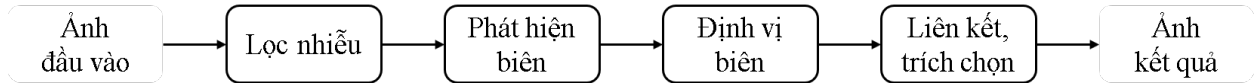
Quy trình xác định biên gồm bốn bước cụ thể từng bước như sau:

Bước 1: Lọc nhiễu ảnh đầu vào

Bước 2: Phát hiện biên bằng cách sử dụng các toán tử

Bước 3: Định vị biên (loại bỏ biên giả)

Bước 4: Liên kết và trích chọn biên



Hình 4. Quy trình xác định biên

Có nhiều phương pháp để phát hiện biên cơ bản, nhưng hầu hết trong số đó có thể chia thành hai loại:

- *Phương pháp phát hiện biên trực tiếp*: tìm biên dựa vào sự biến thiên về giá trị độ sáng (cấp xám) của ảnh. Chủ yếu dựa vào kỹ thuật lấy đạo hàm. Ở phương pháp này ta có thể nhóm thành 2 loại:
 - Phương pháp Gradient: chúng ta tính toán ước lượng độ lớn gradient bằng cách sử dụng bộ lọc làm mịn và sử dụng dự đoán tính toán để xác định vị trí của biên. Nói cách khác phương pháp Gradient là phương pháp dò biên cục bộ bằng cách tìm kiếm cực đại và cực tiểu khi lấy đạo hàm bậc nhất của ảnh.
 - Phương pháp Laplace: chúng ta lấy đạo hàm bậc hai của các tín hiệu và biên độ đạo hàm là cực đại khi đạo hàm bậc hai bằng 0. Trong ngắn hạn, phương pháp Laplace tìm kiếm toán tử chéo không tại đạo hàm bậc hai của ảnh để tìm biên.
- *Phương pháp phát hiện biên gián tiếp*: Nếu bằng cách nào đó ta phân được ảnh thành các vùng thì ranh giới giữa các vùng đó gọi là biên. Kỹ thuật dò biên và phân vùng ảnh là hai bài toán đối ngẫu nhau vì dò biên để thực hiện phân lớp đối tượng mà khi đã phân lớp xong nghĩa là đã phân vùng được ảnh và ngược lại, khi đã phân vùng ảnh đã được phân lớp thành các đối tượng, do đó có thể phát hiện được biên.

Ngoài ra còn các phương pháp nâng cao khác như phương pháp Canny, Wavelet..

3.3 Phương pháp xác định biên Sobel

Toán tử Sobel, đôi khi được gọi là toán tử Sobel-Feldman hoặc bộ lọc Sobel, được sử dụng trong xử lý hình ảnh và thị giác máy tính, đặc biệt là trong các thuật toán phát hiện

cạnh, nơi nó tạo ra một hình ảnh nhấn mạnh các cạnh. Sobel và Feldman đã trình bày ý tưởng về "Toán tử Gradient Hình ảnh 3x3 đẳng hướng" tại một cuộc nói chuyện tại SAIL năm 1968. Về mặt kỹ thuật, nó là một toán tử phân biệt rời rạc, tính toán xấp xỉ độ dốc của hàm cường độ hình ảnh. Tại mỗi điểm trong hình ảnh, kết quả của toán tử Sobel-Feldman là vector gradient tương ứng hoặc chuẩn của vector này. Toán tử Sobel-Feldman dựa trên việc biến đổi hình ảnh với một bộ lọc nhỏ, có thể phân tách và có giá trị số nguyên theo hướng ngang và dọc và do đó tương đối rẻ về mặt tính toán. Mặt khác, xấp xỉ gradient mà nó tạo ra tương đối thô, đặc biệt là đối với các biến thể tần số cao trong hình ảnh.

Toán tử sử dụng hai hạt nhân 3×3 được đối chiếu với hình ảnh ban đầu để tính toán các giá trị gần đúng của các đạo hàm - một cho các thay đổi theo chiều ngang và một cho các thay đổi theo chiều dọc. Nếu chúng ta xác định \mathbf{A} là hình ảnh nguồn, \mathbf{G}_x và \mathbf{G}_y là hai hình ảnh mà tại mỗi điểm chứa các xấp xỉ đạo hàm ngang và dọc tương ứng, các phép tính như sau:

$$\mathbf{G}_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A} \quad \text{và} \quad \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$

trong đó $*$ ở đây biểu thị phép toán tích chập xử lý tín hiệu hai chiều.

Vì các hạt nhân Sobel có thể được phân rã thành một hạt nhân trung bình và một hạt nhân phân biệt, chúng sẽ tính toán độ dốc bằng làm mịn. Ví dụ: \mathbf{G}_x có thể được viết là

$$\mathbf{G}_x = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} * ([+1 \quad 0 \quad -1] * \mathbf{A}) \quad \text{và} \quad \mathbf{G}_y = \begin{bmatrix} +1 \\ 0 \\ -1 \end{bmatrix} * ([1 \quad 2 \quad 1] * \mathbf{A})$$

Ở đây, tọa độ x được định nghĩa là tăng theo hướng qua phải và tọa độ y được định nghĩa là tăng theo hướng đi xuống. Tại mỗi điểm trong hình ảnh, các giá trị xấp xỉ gradient thu được có thể được kết hợp để cung cấp độ lớn gradient, bằng cách sử dụng:

$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2}$$

Sử dụng thông tin này, ta cũng có thể tính toán hướng của gradient:

$$\Theta = \arctan \frac{\mathbf{G}_y}{\mathbf{G}_x}$$

Ngoài ra, để đạt được kết quả tốt hơn, hạt nhân 3x3 được đổi thành hạt nhân 5x5 hiển thị các vùng tuyến tính rõ ràng hơn so với trường hợp trước. Các hạt nhân 5 X 5 này được cài đặt như sau:

+2	+2	+4	+2	+2
+1	+1	+2	+1	+1
0	0	0	0	0
-1	-1	-2	-1	-1
-2	-2	-4	-2	-2

G_x

và

+2	+1	0	-1	-2
+2	+1	0	-1	-2
+4	+2	0	-2	-4
+2	+1	0	-1	-2
+2	+1	0	-1	-2

G_y

3.4 Phương pháp xác định biên Prewitt

Toán tử Prewitt được sử dụng trong xử lý hình ảnh, đặc biệt là trong các thuật toán phát hiện cạnh. Về mặt kỹ thuật, nó là một toán tử phân biệt rời rạc, tính toán xấp xỉ độ dốc của hàm cường độ hình ảnh. Tại mỗi điểm trong hình ảnh, kết quả của toán tử Prewitt là vector gradient tương ứng hoặc chuẩn của vector này. Toán tử Prewitt dựa trên việc biến đổi hình ảnh với một bộ lọc nhỏ, có thể phân tách và có giá trị nguyên theo hướng ngang và dọc và do đó tương đối rẻ về mặt tính toán như toán tử Sobel. Mặt khác, xấp xỉ gradient mà nó tạo ra tương đối thô, đặc biệt đối với các biến thể tần số cao trong hình ảnh. Toán tử Prewitt được phát triển bởi Judith M. S. Prewitt.

Về mặt toán học, toán tử sử dụng hai hạt nhân 3×3 được đối chiếu với hình ảnh ban đầu để tính toán các giá trị gần đúng của các đạo hàm - một cho các thay đổi theo chiều ngang và một cho các thay đổi theo chiều dọc. Nếu chúng ta xác định **A** là hình ảnh nguồn, **G_x** và **G_y** là hai hình ảnh mà tại mỗi điểm chứa các giá trị xấp xỉ đạo hàm ngang và dọc, giá trị được tính là:

$$\mathbf{G}_x = \begin{bmatrix} +1 & 0 & -1 \\ +1 & 0 & -1 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A}$$

và

$$\mathbf{G}_y = \begin{bmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} * \mathbf{A}$$

trong đó * ở đây biểu thị phép toán tích chập hai chiều.

Vì các hạt nhân Prewitt có thể được phân rã thành một hạt nhân trung bình và một hạt nhân phân biệt, chúng sẽ tính toán độ dốc bằng làm mịn. Do đó, nó là một bộ lọc có thể phân tách. Ví dụ: G_x có thể được viết là

$$\begin{bmatrix} +1 & 0 & -1 \\ +1 & 0 & -1 \\ +1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} +1 & 0 & -1 \end{bmatrix}$$

Ở đây, tọa độ x được định nghĩa là tăng theo hướng qua trái và tọa độ y được định nghĩa là tăng theo hướng lên trên. Tại mỗi điểm trong hình ảnh, các giá trị xấp xỉ gradient thu được có thể được kết hợp để cung cấp độ lớn gradient, bằng cách sử dụng:

$$G = \sqrt{G_x^2 + G_y^2}$$

Sử dụng thông tin này, chúng ta cũng có thể tính toán hướng của gradient:

$$\Theta = \text{atan2}(G_y, G_x)$$

Ngoài ra, để đạt được kết quả tốt hơn, hạt nhân 3x3 được đổi thành hạt nhân 5x5 hiển thị các vùng tuyến tính rõ ràng hơn so với trường hợp trước. Các hạt nhân 5 X 5 này được cài đặt như sau:

+2	+2	+2	+2	+2
+1	+1	+1	+1	+1
0	0	0	0	0
-1	-1	-1	-1	-1
-2	-2	-2	-2	-2

G_x

và

+2	+1	0	-1	-2
+2	+1	0	-1	-2
+2	+1	0	-1	-2
+2	+1	0	-1	-2
+2	+1	0	-1	-2

G_y

3.5 Xác định biên ảnh màu

Trong ảnh xám, biên là những điểm có độ xám thay đổi đột ngột so với các điểm lân cận. Định nghĩa của một biên trong ảnh màu có nhiều khó khăn hơn vì khoảng màu sắc khác nhau và có sự biến thiên về màu sắc.

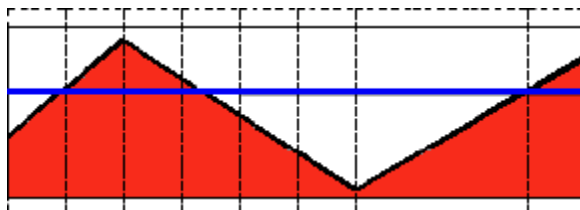
Các phương pháp đơn giản để phát hiện biên màu được dựa trên phần mở rộng của các toán tử đơn sắc. Những kỹ thuật này được áp dụng cho ba kênh màu độc lập, thực hiện phát hiện biên xám trên mỗi kênh màu. Các kết quả này sau đó được kết hợp để cung cấp một sơ đồ biên duy nhất. Nói chung, để phát hiện được biên ảnh màu, các kỹ thuật dựa trên biên độ được mở rộng bằng cách lấy biên độ tối đa của gradient, tổng trọng số của chúng, các góc nghĩa là ô của tổng bình phương biên độ của các sơ đồ biên, hoặc bằng cách thực hiện các phép biến đổi riêng lẻ kết quả đầu ra.

3.6 Ngưỡng

Để phân biệt các pixel mà chúng ta quan tâm với phần còn lại, ta thực hiện so sánh từng giá trị cường độ pixel đối với một ngưỡng (được xác định tùy theo vấn đề cần giải quyết).

Khi đã tách đúng các pixel quan trọng, ta có thể đặt chúng với một giá trị mới để xác định chúng (tức là có thể gán cho chúng giá trị 0 (đen), 255 (trắng) hoặc bất kỳ giá trị nào phù hợp với nhu cầu).

Để minh họa cách thức hoạt động của các quá trình tạo ngưỡng này, chúng ta hãy xem xét rằng ta có một hình ảnh với các pixel có giá trị cường độ src (x, y) như trong Hình 5. Đường ngang màu xanh lam đại diện cho ngưỡng thresh (cố định).



Hình 5. Giá trị cường độ của các pixel trong ảnh

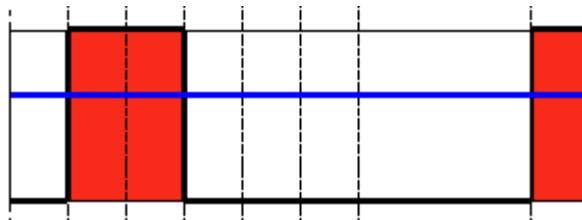
Ta xem xét 5 loại ngưỡng bao gồm:

- *Ngưỡng nhị phân* được định nghĩa như sau:

$$\text{dst}(x, y) = \begin{cases} \text{maxVal} & \text{if } \text{src}(x, y) > \text{thresh} \\ 0 & \text{otherwise} \end{cases}$$

Nếu cường độ của pixel src (x, y) cao hơn thresh, thì cường độ pixel mới được đặt thành maxVal. Ngược lại, các pixel được đặt thành 0.

Khi đó, Hình 5 trở thành:

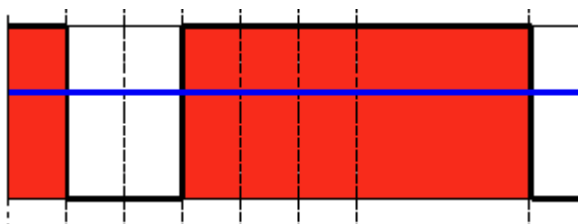


- *Ngưỡng nghịch đảo nhị phân:*

$$\text{dst}(x, y) = \begin{cases} 0 & \text{if } \text{src}(x, y) > \text{thresh} \\ \text{maxVal} & \text{otherwise} \end{cases}$$

Nếu cường độ của pixel src (x, y) cao hơn thresh, thì cường độ pixel mới được đặt thành 0. Ngược lại, cường độ đó được đặt thành maxVal.

Khi đó, Hình 5 trở thành:

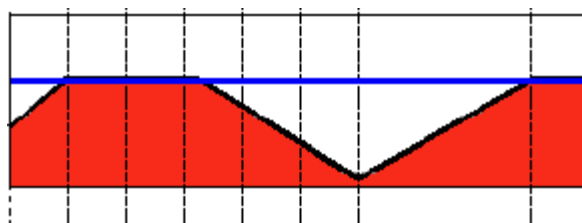


- *Ngưỡng cắt:*

$$\text{dst}(x, y) = \begin{cases} \text{threshold} & \text{if } \text{src}(x, y) > \text{thresh} \\ \text{src}(x, y) & \text{otherwise} \end{cases}$$

Giá trị cường độ tối đa cho các pixel là thresh, nếu src (x, y) lớn hơn, thì giá trị của nó sẽ bị cắt bớt.

Khi đó, Hình 5 trở thành:

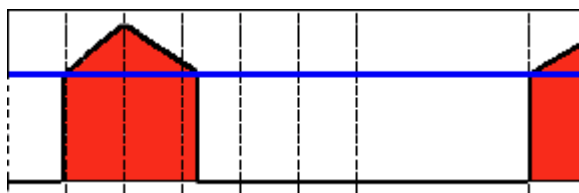


- Ngưỡng về không:

$$\text{dst}(x, y) = \begin{cases} \text{src}(x, y) & \text{if } \text{src}(x, y) > \text{thresh} \\ 0 & \text{otherwise} \end{cases}$$

Nếu $\text{src}(x, y)$ thấp hơn thresh , giá trị pixel mới sẽ được đặt thành 0.

Khi đó, Hình 5 trở thành:

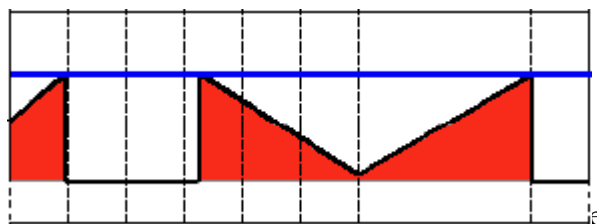


- Ngưỡng nghịch đảo về không:

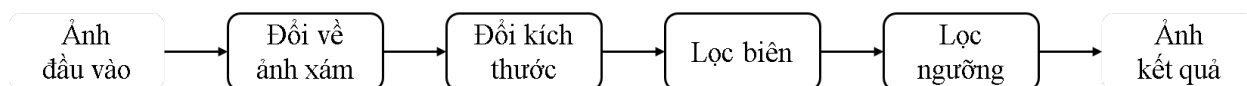
$$\text{dst}(x, y) = \begin{cases} 0 & \text{if } \text{src}(x, y) > \text{thresh} \\ \text{src}(x, y) & \text{otherwise} \end{cases}$$

Nếu $\text{src}(x, y)$ lớn hơn thresh , giá trị pixel mới sẽ được đặt thành 0.

Khi đó, Hình 5 trở thành:



3.7 Quy trình xác định biên ảnh xám



Hình 6. Quy trình xác định biên ảnh xám

Bước 1: Đổi về ảnh xám

Đối với ảnh đầu vào không phải là ảnh xám, ta cần chuyển ảnh này về dạng ảnh độ xám

Bước 2: Đổi kích thước

Ta có thể tùy chỉnh kích thước ảnh như một cách cải thiện chất lượng ảnh

Bước 3: Lọc biên

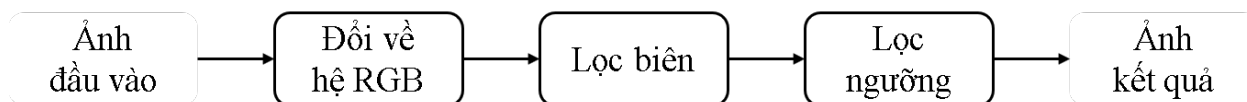
Ta dùng các cửa sổ trượt qua toàn bộ tấm ảnh để xác định biên. Ở đây, các cửa sổ được định nghĩa là các toán tử Sobel và Prewitt với hai kích thước 3x3 và 5x5

Bước 4: Lọc ngưỡng

Ảnh sau khi lọc biên được đưa qua bước lọc ngưỡng nhằm nâng cao chất lượng biên. Ở đây chúng tôi dùng năm loại ngưỡng bao gồm:

- Ngưỡng nhị phân
- Ngưỡng nghịch đảo nhị phân
- Ngưỡng cắt
- Ngưỡng về không
- Ngưỡng nghịch đảo về không

3.8 Quy trình xác định biên ảnh màu



Hình 7. Quy trình xác định biên ảnh màu

Bước 1: Đổi về hệ màu RGB

Đối với các ảnh có hệ màu khác RGB, ta cần chuyển ảnh về hệ màu RGB để thuận tiện hơn cho các bước tiếp theo

Bước 2: Lọc biên

Ta dùng các cửa sổ trượt qua toàn bộ tấm ảnh để xác định biên. Ở đây, các cửa sổ được định nghĩa là các toán tử Sobel và Prewitt với hai kích thước 3×3 và 5×5

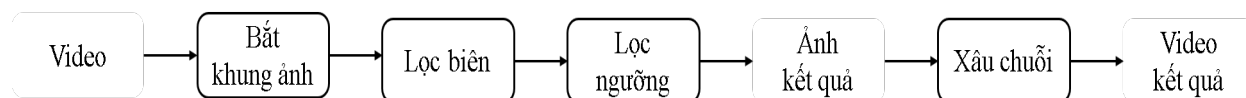
Bước 3: Lọc ngưỡng

Ảnh sau khi lọc biên được đưa qua bước lọc ngưỡng nhằm nâng cao chất lượng biên. Ở đây chúng tôi dùng năm loại ngưỡng bao gồm:

- Ngưỡng nhị phân
- Ngưỡng nghịch đảo nhị phân
- Ngưỡng cụt
- Ngưỡng về không
- Ngưỡng nghịch đảo về không

3.9 Quy trình xác định biên video và webcam

Vì video có thể được xem như chuỗi các hình ảnh liên tiếp nhau, ta có thể xác định biên của một video bằng cách xác định biên từng khung ảnh và xâu chuỗi chúng lại với nhau. Đối với webcam, ta cũng có thể áp dụng tương tự như video. Quy trình cụ thể như sau:



Hình 8. Quy trình xác định biên dành cho video hoặc webcam

Bước 1: Bắt khung ảnh

Ở mỗi khung ảnh của video/webcam, ta lưu như một ảnh cần được xác định biên

Bước 2: Lọc biên

Ta dùng các cửa sổ trượt qua toàn bộ tấm ảnh để xác định biên. Ở đây, các cửa sổ được định nghĩa là các toán tử Sobel và Prewitt với hai kích thước 3×3 và 5×5

Bước 3: Lọc ngưỡng

Ảnh sau khi lọc biên được đưa qua bước lọc ngưỡng nhằm nâng cao chất lượng biên. Ở đây chúng tôi dùng năm loại ngưỡng bao gồm:

- Ngưỡng nhị phân
- Ngưỡng nghịch đảo nhị phân
- Ngưỡng cắt
- Ngưỡng về không
- Ngưỡng nghịch đảo về không

Bước 4: Xâu chuỗi

Các ảnh kết quả được xâu chuỗi lại thành video kết quả.

CHƯƠNG 4: KẾT QUẢ THỰC NGHIỆM

4.1 Giới thiệu dữ liệu

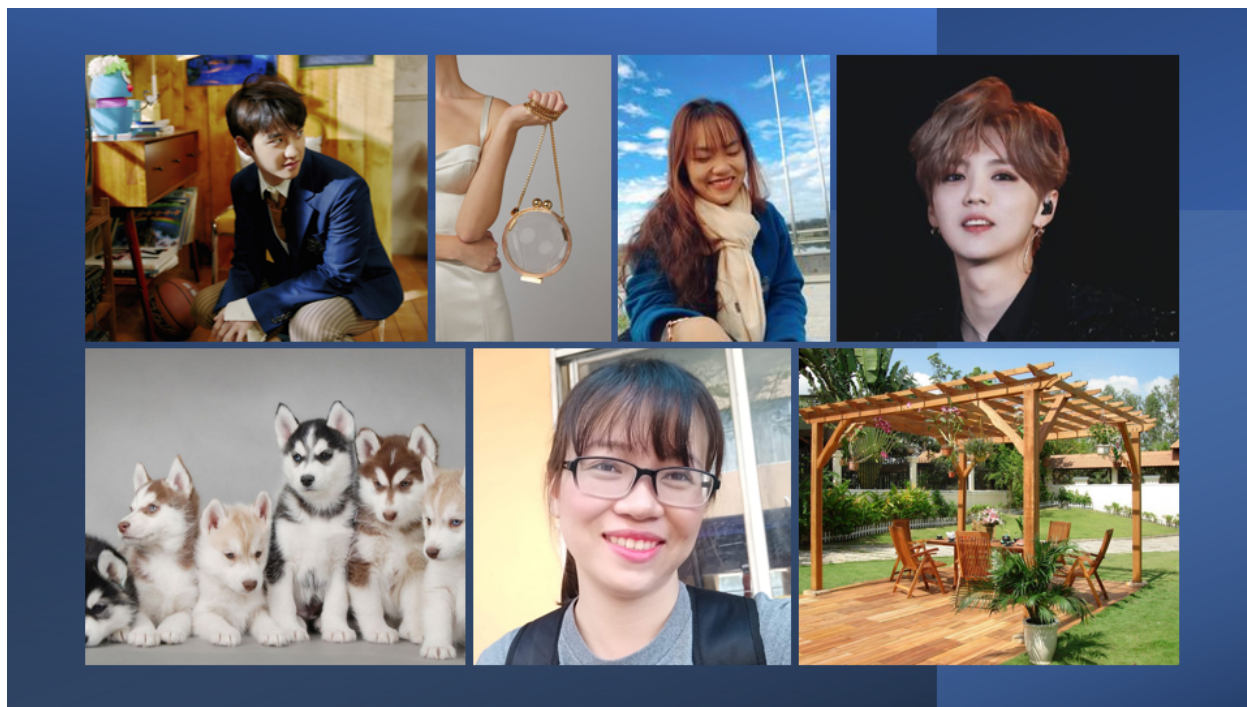
Để kiểm tra chương trình xác định biên, chúng tôi đã thực nghiệm trên ba loại dữ liệu gồm: hình ảnh, video và webcam.

- *Dữ liệu hình ảnh:*

Gồm 15 ảnh màu (24 bit) với thông số được mô tả trong Bảng 1.

Tên ảnh	Kích thước	Độ phân giải
0001.jpg	4096 x 2732	72 dpi
0002.jpg	400 x 600	72 dpi
0003.jpg	712 x 1067	72 dpi
0004.jpg	1024 x 768	72 dpi
0005.jpg	1024 x 1024	72 dpi
0006.jpg	1024 x 1024	72 dpi
0007.jpg	1440 x 1430	96 dpi
0008.jpg	600 x 900	72 dpi
0009.jpg	564 x 903	72 dpi
0014.jpg	155 x 206	96 dpi
0019.jpg	1024 x 1024	96 dpi
1000(1).jpg	150 x 150	96 dpi
1003(1).jpg	150 x 150	96 dpi
1007(1).jpg	150 x 150	96 dpi
1014(1).jpg	150 x 150	96 dpi

Bảng 1. Thông số dữ liệu hình ảnh



Hình 9. Một số ảnh từ bộ dữ liệu hình ảnh

- *Dữ liệu video:*

Video có thể được tạo ra từ các nguồn khác nhau bao gồm máy quay kỹ thuật số, các phần mềm máy tính... Trong bộ dữ liệu thử nghiệm, chúng tôi đã chọn đại diện mỗi loại một video với thông số khác nhau được mô tả như trong bảng sau:

Tên video	Kích thước khung hình	Tốc độ khung hình
0001.mp4	1920 x 1080	23.98 fps
0002.mp4	1280 x 720	30.00 fps
0003.mp4	810 x 1080	30.00 fps

Bảng 2. Thông số dữ liệu video

- *Dữ liệu webcam:*

Các video được ghi lại từ webcam với các thông số như sau:

Kích thước khung hình	Tốc độ khung hình
640 x 480	30.00 fps

Bảng 3. Thông số video được ghi lại từ webcam

4.2 Kết quả trên tập dữ liệu hình ảnh xám

Bước 1: Đổi về ảnh xám – Bước 2: Đổi kích thước

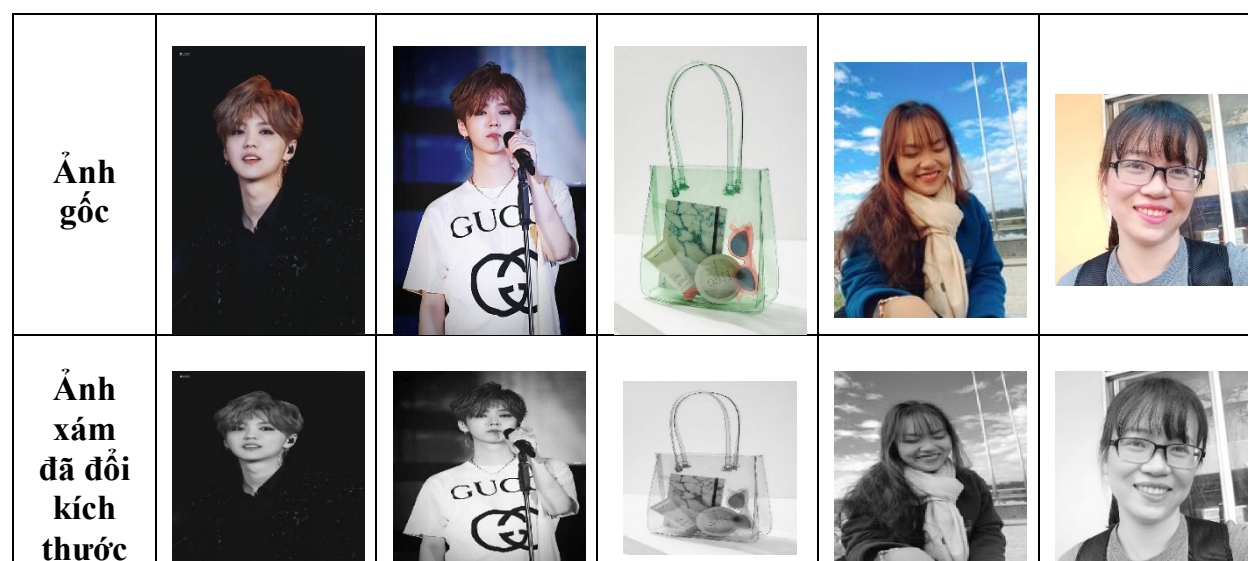
Vì bộ dữ liệu là ảnh màu có kích thước khác nhau, chúng tôi dùng hàm `grayscale_resize` chuyển đổi chúng thành ảnh xám và đổi kích thước ảnh. Ở đây chúng tôi đổi tất cả ảnh thành ảnh có kích thước 256x256.




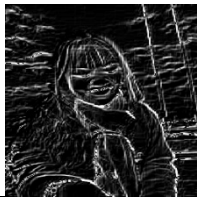



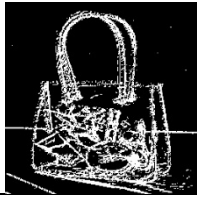




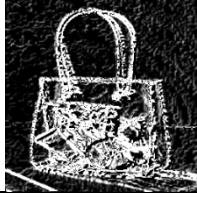


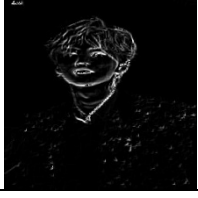


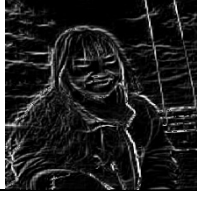



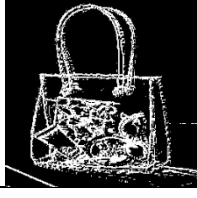







Bước 3: Lọc biên – Bước 4: Lọc ngưỡng

Chúng tôi dùng hàm `detect_edge_using_filter` để lọc biên sau đó lọc ngưỡng cho ảnh xám 256x256 vừa nhận được. Trong đó:

- `filter_type` là loại toán tử cần sử dụng bao gồm: `sobel_3` (toán tử Sobel kích thước 3x3), `sobel_5` (toán tử Sobel kích thước 5x5), `prewitt_3` (toán tử Prewitt kích thước 3x3) và `prewitt_5` (toán tử Prewitt kích thước 5x5)
- Giá trị ngưỡng được nhập trong lúc chương trình được thực thi. Ở đây chúng tôi chọn giá trị ngưỡng nhị phân là 20.

Dưới đây là một số kết quả điển hình:



Sobel 3x3					
Sobel 3x3 ngưỡng 20					
Sobel 5x5					
Prewitt 3x3					
Prewitt 3x3 ngưỡng 20					
Prewitt 5x5					

Bảng 4. Một số kết quả của quy trình xác định biên ảnh xám

Ghi chú: Hình ảnh đã được thay đổi để phù hợp với kích thước bảng.

Nhận xét:

- Đối với các toán tử Sobel và Prewitt có kích thước 3x3, các biên cơ bản được xác định. Tuy nhiên với trường hợp vật thể có độ xám trùng với độ xám của nền, kết quả không thực sự khả quan.
- Đối với các toán tử Sobel và Prewitt kích thước 3x3 có áp dụng giá trị ngưỡng 20, các biên được xác định rõ ràng, kể cả trường hợp vật thể có độ xám tương đồng với nền.
- Đối với các toán tử Sobel và Prewitt có kích thước 5x5, kết quả khá tương đồng với phương pháp toán tử kích thước 3x3 có áp dụng giá trị ngưỡng 20.

4.3 Kết quả trên tập dữ liệu ảnh màu



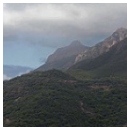


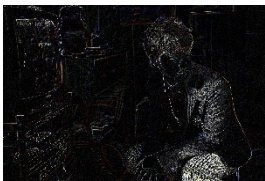

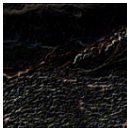

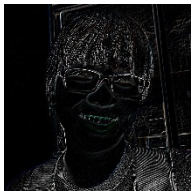
Bước 1: Đổi về hệ màu RGB


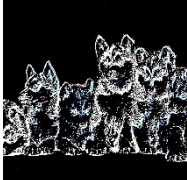
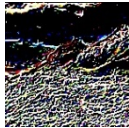


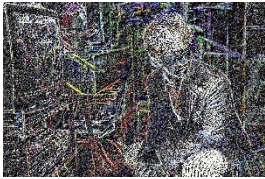
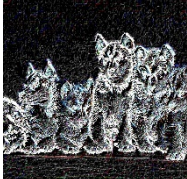
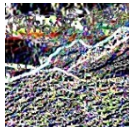


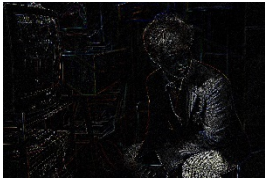

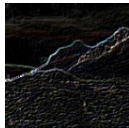



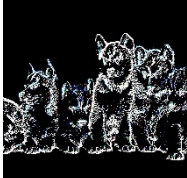
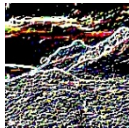


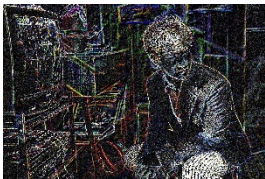

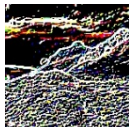


Bước 2: Lọc biên – Bước 3: Lọc ngưỡng

Chúng tôi dùng hàm `detect_edge_using_filter` để lọc biên sau đó lọc ngưỡng cho ảnh màu. Trong đó:

- `filter_type` là loại toán tử cần sử dụng bao gồm: `sobel_3` (toán tử Sobel kích thước 3x3), `sobel_5` (toán tử Sobel kích thước 5x5), `prewitt_3` (toán tử Prewitt kích thước 3x3) và `prewitt_5` (toán tử Prewitt kích thước 5x5)
- Giá trị ngưỡng được nhập trong lúc chương trình được thực thi. Ở đây chúng tôi chọn giá trị ngưỡng nhị phân là 20.

Dưới đây là một số kết quả điển hình:

Ảnh gốc					
Sobel 3x3					

Sobel 3x3 ngưỡng 20					
Sobel 5x5					
Prewitt 3x3					
Prewitt 3x3 ngưỡng 20					
Prewitt 5x5					

Bảng 5. Một số kết quả của quy trình xác định biên ảnh màu

Ghi chú: Hình ảnh đã được thay đổi để phù hợp với kích thước bảng.

Nhận xét:

- Kết quả xác định biên ảnh màu của chương trình nhìn chung khá tương đồng với kết quả trên ảnh xám. Tuy nhiên, vẫn có một số điểm đặc biệt như sau
 - Đối với ảnh có kích thước ảnh có cạnh mờ (ví dụ ảnh chụp trong sương mù), kết quả khi dùng toán tử Prewitt kích thước 3x3 tốt hơn so với toán tử Sobel kích thước tương tự.
 - Trong một số trường hợp, việc tăng kích thước toán tử không thực sự cần thiết.

4.4 Kết quả trên dữ liệu video và webcam

Dữ liệu và kết quả chương trình xác định biên trên tập dữ liệu video và webcam được đính kèm trong thư mục Video_Dataset.

CHƯƠNG 5: THẢO LUẬN

5.1 Ưu điểm của chương trình

- Tốc độ tính toán khá nhanh
- Xác định được các đường biên cơ bản trên ảnh xám và ảnh màu
- Xác định được các đường biên của video (kể cả video trực tuyến)
- Có thể chọn giá trị ngưỡng đối với hình ảnh.

5.2 Hạn chế của chương trình

- Thực thi tất cả các loại ngưỡng đã định nghĩa
- Chưa tùy chọn được giá trị ngưỡng đối với video

5.3 Hướng phát triển

- Cài đặt cho phép tùy chọn loại ngưỡng
- Cài đặt cho phép tùy chọn giá trị ngưỡng đối với video
- Cài đặt thêm các phương pháp khác như Canny, Wavelet, Holistically-Nested ...

KẾT LUẬN

Trong báo cáo này, chúng tôi thảo luận và đánh giá hai phương pháp Sobel và Prewitt nhằm xác định biên phục vụ cho các bước tiếp theo của quy trình xử lý ảnh kỹ thuật số. Ta thấy rằng, tùy vào mục đích sử dụng cũng như thông số dữ liệu mà các phương pháp sẽ cho kết quả khác nhau. Vì vậy, không thể áp dụng một phương pháp duy nhất cho tất cả các dữ liệu. Tuy nhiên, với tính chất và đặc điểm riêng, ta vẫn có thể chọn phương pháp thích hợp cho một ảnh cụ thể.

NGUỒN THAM KHẢO

1. Lọc filter: <https://www.sciencedirect.com/topics/computer-science/sobel-filter>
2. Mở rộng kích thước Sobel và Prewitt 3x3 thành 5x5:
<https://www.iasj.net/iasj?func=fulltext&aId=52927>
3. Phát hiện biên: <https://paperswithcode.com/task/edge-detection>
4. Thuật toán phát hiện biên trong xử lý ảnh kỹ thuật số:
<https://www.slideshare.net/rahulmonikasharma/a-review-on-edge-detection-algorithms-in-digital-image-processing-applications>
5. Phát hiện biên trong xử lý ảnh: <https://ieeexplore.ieee.org/document/8275968>
6. Sobel và Laplacian:
https://www.bogotobogo.com/python/OpenCV_Python/python_opencv3_Image_Gradient_Sobel_Laplacian_Derivatives_Edge_Detection.php
7. Phát hiện biên dùng opencv: <https://learnopencv.com/edge-detection-using-opencv/>
8. Phát hiện biên bằng Sobel: <https://www.geeksforgeeks.org/python-program-to-detect-the-edges-of-an-image-using-opencv-sobel-edge-detection/>
9. Ngưỡng: https://docs.opencv.org/3.4/db/d8e/tutorial_threshold.html
10. Các toán tử Sobel và Prewitt:
https://www.researchgate.net/publication/49619233_Image_Segmentation_using_Extended_Edge_Operator_for_Mammographic_Images
11. Phát hiện biên dùng Deep Learning:
https://www.linkedin.com/pulse/deep-learning-cnn-edge-detection-chen-yang?trk=portfolio_article-card_title