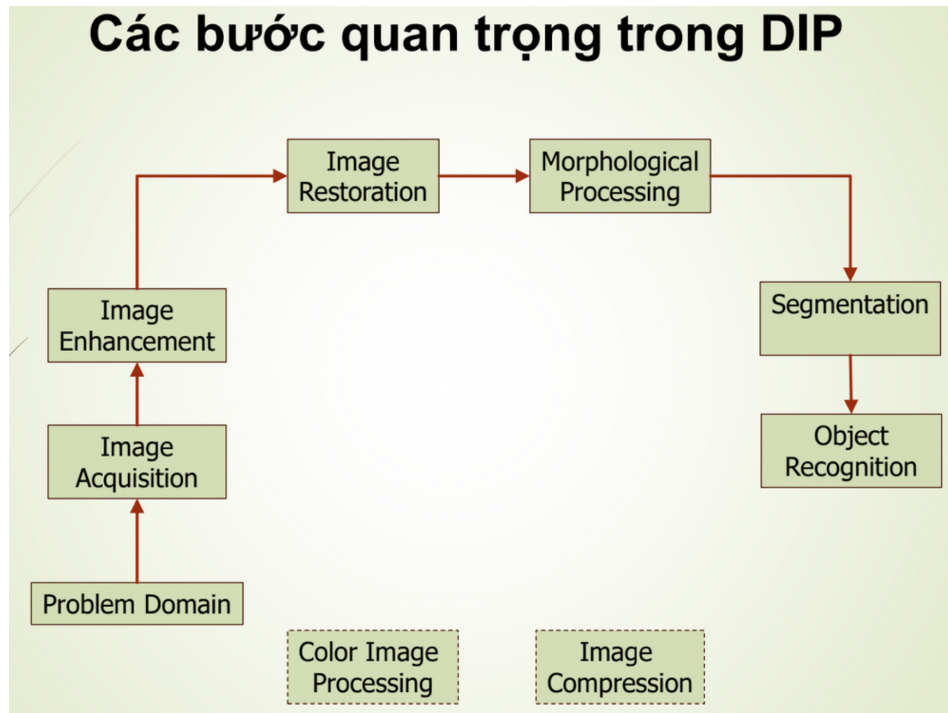


Yêu cầu dự án



Hình 1:

Yêu cầu:

Xây dựng phần mềm ứng dụng Web/Mobile/GUI triển khai các thuật toán như một bộ công cụ giúp xử lý ảnh, bao gồm các chức năng: input ảnh, xử lý ảnh, và output ảnh sau khi đã xử lý.

Input: cho phép nhập vào 1 ảnh, hiển thị thông số tên ảnh, độ phân giải, biểu diễn ảnh.

Xử lý ảnh là bộ công cụ bao gồm các bước xử lý ảnh như trên hình 1:

- Thu thập hình ảnh
- Cải thiện ảnh
- Phục hồi ảnh
- Xử lý hình thái học
- Phân vùng ảnh
- Nhận dạng vật thể
- Xử lý màu
- Nén ảnh...

Trong mỗi một bước xử lý ảnh có thể gồm nhiều phương pháp. Ví dụ với “Cải thiện ảnh” có cải thiện ảnh trên miền không gian, cải thiện ảnh trên miền tần số.

Trong cải thiện ảnh trên miền không gian có thể gồm các bộ lọc: Smoothing linear filter, Median filter, The Laplacian...

Cách thức tính điểm:

- Trình bày báo cáo 5-10 trang, mạch lạc rõ ràng theo đúng các quy tắc trình bày văn bản. **(2 điểm)**
- Đẩy mã nguồn của ứng dụng lên Git, viết Readme rõ ràng, dễ dàng sử dụng **(2 điểm)**
- Nội dung **(6 điểm)**: Lựa chọn bất kỳ kỹ thuật, phương pháp xử lý ảnh mà bạn thấy cần thiết, hữu ích. Mỗi một phương pháp, kỹ thuật xử lý ảnh được tính **1 điểm**. Một phương pháp, kỹ thuật xử lý ảnh phải được viết dưới hai dạng:
 - Sử dụng thư viện OpenCV
 - Không sử dụng thư viện OpenCV, dựa trên phân tích toán học

Yêu cầu bài tập

Nếu sinh viên không lựa chọn yêu cầu dự án có thể lựa chọn giải các bài tập sau đây. Viết tay hoặc đánh máy, ghi rõ họ tên, mã số sinh viên. Scan nộp lại qua google driver.

Câu 1: (2.5 điểm)

- Liệt kê hai phương pháp biểu diễn ảnh.
- Nêu các loại kết nối (connectivity) và giải thích.
- Nêu ba loại bộ lọc làm sắc nét (sharpening filters).
- Đưa ra công thức tính khoảng cách D_4 (khoảng cách thành phố - city block distance) và khoảng cách D_8 (khoảng cách bàn cờ - chess board distance).
- Giải thích các bước khái niệm và các phương trình liên quan trong kỹ thuật **Unsharp Masking** trong việc nâng cao chất lượng ảnh.

Câu 2: (2.5 điểm)

Chọn đáp án đúng sau đây:

1. Computer vision is defined as a discipline in which ...
 - a) Both the input and output of a process are images
 - b) The input of a process is an image description and the output is an image
 - c) Both the input and output of a process are descriptions
 - d) The input of a process is an image and the output is an image description
2. MRIT technology uses the ... band
 - a) Gamma Rays
 - b) CT Scan
 - c) Visible
 - d) Radio waves
3. ... filter cannot be implemented using the convolution mechanism
 - a) Average
 - b) Gaussian
 - c) Median
 - d) Disk

4. To remove “salt and pepper” noise without blurring, we use ...

- | | |
|---------------|---------------------|
| a) Max Filter | b) Median Filter |
| c) Min Filter | d) Smoothing Filter |

5. Edge detection in images is commonly accomplished by performing a spatial ... of the image field.

- | | |
|---------------------|------------------------|
| a) Smoothing Filter | b) Integration |
| c) Differentiation | d) Sobel and Laplacian |

Câu 3: (2.5 điểm)

Tham khảo sách

Computer Vision - Richard Szeliski

Algorithms and Applications (Second Edition)

Bài tập 3.7

Image Processing - Histogram equalization.

Tính toán biểu đồ mức xám (luminance) cho một hình ảnh và cân bằng nó để các tông màu nhìn đẹp hơn (và hình ảnh ít nhạy cảm hơn với các thiết lập phơi sáng - the image is less sensitive to exposure settings). Thực hiện các bước sau (nên thực hiện hết 6 yêu cầu trong bài).

- Chuyển đổi ảnh màu sang độ chói sáng (luminance) (mục 3.1.2, sách tham khảo).
- Tính toán biểu đồ, phân phối tích lũy, và hàm bù chuyển (mục 3.1.4, sách tham khảo).
- Thử tăng “độ nổi bật” trong hình ảnh bằng cách đảm bảo rằng một tỷ lệ nhất định của các pixel (ví dụ, 5%) được ánh xạ thành màu đen và trắng thuần túy.

Câu 4: (2.5 điểm)

Bài tập 6.7: Phát hiện đối tượng. Đối với phát hiện đối tượng, số lượng tham số của DETR, Faster-RCNN và YOLOv4 so sánh như thế nào? Hãy thử huấn luyện từng mô hình trên bộ dữ liệu MS COCO. Mô hình nào có xu hướng huấn luyện chậm nhất? Mất bao lâu để mỗi mô hình đánh giá một hình ảnh đơn lẻ trong thời gian suy luận?