

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

**KHOA CÔNG NGHỆ THÔNG TIN**

-----🙞🙜🕮🙞🙜-----

****

**BÁO CÁO BÀI TẬP LỚN**

**MÔN HỌC: HỆ CSDL ĐA PHƯƠNG TIỆN**

**TÊN LỚP VÀ TÊN NHÓM : D17-070 NHÓM 20**

**ĐỀ TÀI: XÂY DỰNG HỆ CSDL**

**NHẬN DẠNG ẢNH BẦU TRỜI**

|  |  |
| --- | --- |
| *Giảng viên:* | **NGUYỄN ĐÌNH HÓA**  **ĐOÀN XUÂN PHI - B17DCCN484**  **LÊ TRỌNG TUÂN - B17DCCN646**  **LÊ THUẦN TUẤN - B17DCCN653**  HÀ NỘI- 2021 |
| *Thành viên:* |
|  |
|  |

Mục Lục

[MỞ ĐẦU : 4](#_Toc75195759)

[CHƯƠNG I. TỔNG QUAN VỀ HỆ CƠ SỞ DỮ LIỆU ĐA PHƯƠNG TIỆN 4](#_Toc75195760)

[1.1 Hệ CSDL đa phương tiện 4](#_Toc75195761)

[1.2 Công dụng của hệ CSDL đa phương tiện 5](#_Toc75195762)

[1.3 Tra cứu thông tin dữ liệu đa phương tiện 6](#_Toc75195763)

[1.4 Kiến trúc hệ cơ sở dữ liệu ĐPT 10](#_Toc75195764)

[1.5 Mô hình hóa dữ liệu đa phương tiện 11](#_Toc75195765)

[1.5.1 Tầng đối tượng: 11](#_Toc75195766)

[1.5.2 Tầng kiểu phương tiện : 12](#_Toc75195767)

[1.5.3 Tầng dạng thức phương tiện : 12](#_Toc75195768)

[1.6 Trích chọn đặc trưng, chỉ số hóa và đo tính tương tự 12](#_Toc75195769)

[1.6.1 Trích chọn đặc trưng 12](#_Toc75195770)

[1.6.2 : Chỉ số hóa 14](#_Toc75195771)

[1.6.3 Đo tính tương tự 14](#_Toc75195772)

[CHƯƠNG 2 : MỘT SỐ KĨ THUẬT XỬ LÍ VÀ NHẬN DẠNG ẢNH NỀN BẦU TRỜI HIỆN NAY : 14](#_Toc75195773)

[2.1 : Nhận dạng ảnh dựa trên màu sắc : 14](#_Toc75195774)

[2.1.1 Biểu đồ màu toàn cục (Global Color Histogram) 16](#_Toc75195775)

[2.1.2 Biểu đồ màu cục bộ (Local Color Histogram) 17](#_Toc75195776)

[2.1.3 Biểu đồ màu tương quan (Color Correlogram Histogram) 18](#_Toc75195777)

[2.2 Nhận dạng ảnh dựa trên hình dạng 18](#_Toc75195778)

[Tổng quan : 18](#_Toc75195779)

[2.2.1 Thuật toán canny : 19](#_Toc75195780)

[2.2.2 Kĩ thuật La bàn : 20](#_Toc75195781)

[2.2.3 Kĩ thuật prewitt : 20](#_Toc75195782)

[2.2.4 Kĩ thuật Sobel 20](#_Toc75195783)

[2.2.5 Kĩ thuật Phát hiện biên Laplace 21](#_Toc75195784)

[2.3 Truy vấn ảnh dựa trên bố cục, kết cấu hoa văn 21](#_Toc75195785)

[2.3.1 Các thuộc tính chủ yếu liên quan đến bố cục ảnh 21](#_Toc75195786)

[2.3.2 Một số phương pháp được sử dụng phổ biến: 22](#_Toc75195787)

[2.4 Tra cứu ảnh dựa trên đặc trưng bất biến 24](#_Toc75195788)

[CHƯƠNG 3 : SƠ ĐỒ KHỐI QUY TRÌNH THỰC HIỆN BÀI TOÁN NHẬN DẠNG VÀ PHÂN LOẠI ẢNH BẦU TRỜI 25](#_Toc75195789)

[3.1 Sơ đồ khối bài toán 25](#_Toc75195790)

[3.2 Quy trình thực hiện bài toán : 25](#_Toc75195791)

[3.3 Thuộc tính nhận dạng và phân loại 25](#_Toc75195792)

[3.4 Kĩ thuật trích rút thuộc tính màu sắc 27](#_Toc75195793)

[Tổng quan về kĩ thuật : Thống kê kênh màu(Color channel statistics) 27](#_Toc75195794)

[- Tách hình ảnh đầu vào thành các kênh tương ứng. Đối với hình ảnh RGB, tách từng kênh Đỏ, Xanh lục và Xanh lam một cách độc lập. 27](#_Toc75195795)

[- Tính giá trị trung bình màu và độ lệch chuẩn của từng kênh màu 27](#_Toc75195796)

[- Ghép các giá trị trung bình và độ lệch chuẩn của từng kênh thành 1 list để tạo thành Feature vector 6-d 27](#_Toc75195797)

[3.5 Kĩ thuật lưu trữ thuộc tính màu sắc 29](#_Toc75195798)

[3.7 Kĩ thuật nhận dạng ảnh bầu trời từ thuộc tính màu sắc 30](#_Toc75195799)

[3.7.1 Giới thiệu về khoảng cách Euclidean 30](#_Toc75195800)

[3.7.2 Giới thiệu thuật toán KNN(K-Nearest Neighbors) 32](#_Toc75195801)

[- Ưu điểm của KNN 33](#_Toc75195802)

[- Nhược điểm của KNN 33](#_Toc75195803)

[Chương 4 : Cài đặt chương trình và đánh giá kết quả 33](#_Toc75195804)

[4.1 Cài đặt 33](#_Toc75195805)

[Link DataSet 33](#_Toc75195806)

[Đọc và trích xuất dữ liệu 35](#_Toc75195807)

[Lưu trữ đặc trưng 38](#_Toc75195808)

[Cài đặt Thuật toán KNN 39](#_Toc75195809)

[Cài đặt giao diện 41](#_Toc75195810)

[4.2 Đánh giá kết quả : 49](#_Toc75195811)

[4.2.1 Hạn chế 49](#_Toc75195812)

[4.2.2 Hướng phát triển 50](#_Toc75195813)

[TÀI LIỆU THAM KHẢO 50](#_Toc75195814)

# MỞ ĐẦU :

Trong nhiều năm, nghiên cứu và phát triển đa phương tiện là cần thiết trong ứng dụng truyền thông và để thể hiện thông tin đa phương tiện. Ngày càng nhiều dữ liệu số đa phương tiện được thể hiện dưới dạng hình ảnh, video, âm thanh… đòi hỏi các kĩ thuật lưu trữ, tìm kiếm, nhận dạng hiệu quả và mạnh.

**Đặt vấn đề** : Hình ảnh là một trong số các dạng của dữ liệu đa phương tiện.Hình ảnh là những gì mà chúng ta nhìn thấy thông qua thị giác chuyển về não bộ giúp ta cảm nhận hình ảnh một cách chân thực nhất .Hình ảnh xuất hiện nhiều trong các tài liệu sách báo, các thông tin cần mô tả trực quan sinh động. Bài toán nhận dạng là lĩnh vực được quan tâm để giải quyết các yêu cầu trong cuộc sống hiện nay, đặc biệt là nhận dạng ảnh. Cùng với sự ra đời và phát triển của máy tính, các công cụ xử lý cũng ngày càng hoàn thiện dựa trên những kỹ thuật hiện đại để phục vụ cho nhu cầu đó .

Các mô hình nhận dạng hình ảnh thường được sử dụng đó là : Nhận dạng dựa trên văn bản, nhận dạng dựa trên màu sắc, nhận dạng dựa trên hình dạng , nhận dạng dựa trên bố cục. Mỗi mô hình sẽ cho những ưu và nhược điểm khác nhau tùy vào từng bộ dữ liệu : Ví dụ như ảnh có sự tương đồng về màu sắc nhưng chưa chắc đã cùng nội dung và ngược lại. Trong bài báo cáo này nhóm em xin trình bày bài toán về : Nhận dạng ảnh bầu trời. Bố cục bài báo cáo gồm 3 phần :

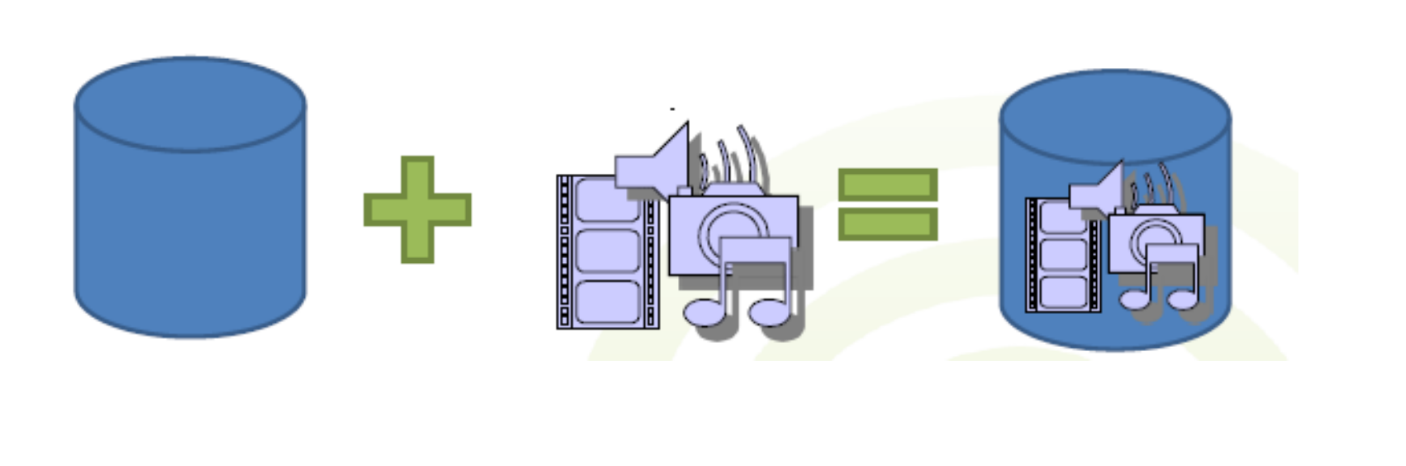
Chương 1 : Tổng quan về hệ cơ sở dữ liệu đa phương tiện

Chương này sẽ giới thiệu tổng quát về hệ csdl đa phương tiện : Các khái niệm, công cụ

# CHƯƠNG I. TỔNG QUAN VỀ HỆ CƠ SỞ DỮ LIỆU ĐA PHƯƠNG TIỆN

## 1.1 Hệ CSDL đa phương tiện

- Hệ CSDL đa phương tiện = CSDL + Đa phương tiện



Khái niệm về đa phương tiện :

* Là sự tích hợp của nhiều loại dữ liệu truyền thông khác nhau,
* Sự tích hợp thường được biểu diễn dưới dạng tài liệu
* Những loại hình truyền thông cơ bản: văn bản, hình ảnh, âm thanh, phim,…

Các dạng tài liệu:

* Đơn phương tiện: bao gồm một loại dữ liệu truyền thông,
* Đa phương tiện: tích hợp nhiều loại dữ liệu truyền thông.

-Các dữ liệu đa phương tiện được truyền đi bởi các phương tiện truyền dẫn

Phương tiện truyền dẫn là gì?

* Là một thực thể / môi trường để truyền thông tin khi có một liên kết được thiết lập,
* Nó độc lập với nội dung thông tin cần truyền tải,
* Trong quá trình truyển tải thông tin, phương tiện truyền dẫn có thể thay đổi từ dạng này sang dạng khác.

Ví dụ: Sách: truyền thông tin giữa người viết và người đọc, độc lập với nội dung được viết bên trong, nội dung bao gồm văn bản và hình ảnh.

## 1.2 Công dụng của hệ CSDL đa phương tiện

* Lưu trữ thông tin đa phương tiện: văn bản, hình ảnh, âm thanh, video…
* Nhận dạng ,phân loại dữ liệu đa phương tiện: tìm kiếm thông tin hiệu quả, chuẩn hoá dữ liệu , phân loại dữ liệu
* Các yêu cầu đối với hệ CSDL đa phương tiện:
  + Đảm bảo chức năng cơ bản của một hệ cơ sở dữ liệu
  + Bảo trì được các dữ liệu chưa được định dạng
  + Cung cấp được các kho dữ liệu đặc biệt, và lưu trữ các thiết bị trình chiếu.

-Các công cụ cần có:

* Kiến trúc phần mềm
* Đánh chỉ số nội dung lưu trữ
* Giao diện với người dùng
* Trích xuất thông tin
* Thiết bị lưu trữ
* Tra cứu thông tin

## 1.3 Tra cứu thông tin dữ liệu đa phương tiện

-Tra cứu thông tin:

* Tìm kiếm, lựa chọn thông tin trong hệ CSDL khớp với yêu cầu người sử dụng đặt ra
* Chuyển kết quả tìm kiếm đến người dùng yêu cầu
* Chức năng tìm kiếm:
  + Tìm kiếm theo ngữ nghĩa (‘semantic”)
  + Tích hợp giữa tìm kiếm cơ bản và nâng cao
  + Trích xuất, chuẩn hoá, đánh chỉ số các thuộc tính thông tin
  + Dựa trên các chuẩn đo khoảng cách

-Các loại tra cứu:

* Tra cứu dựa trên nội dung
* Tra cứu dựa trên ngữ nghĩa

-Tra cứu dữ liệu bao gồm 2 bước chính:

VD: Tra cứu ảnh

+image collection:

+Digitization:

+Image analysis and feature extraction:

+image database:

+Creating the database:

+Querying the database:

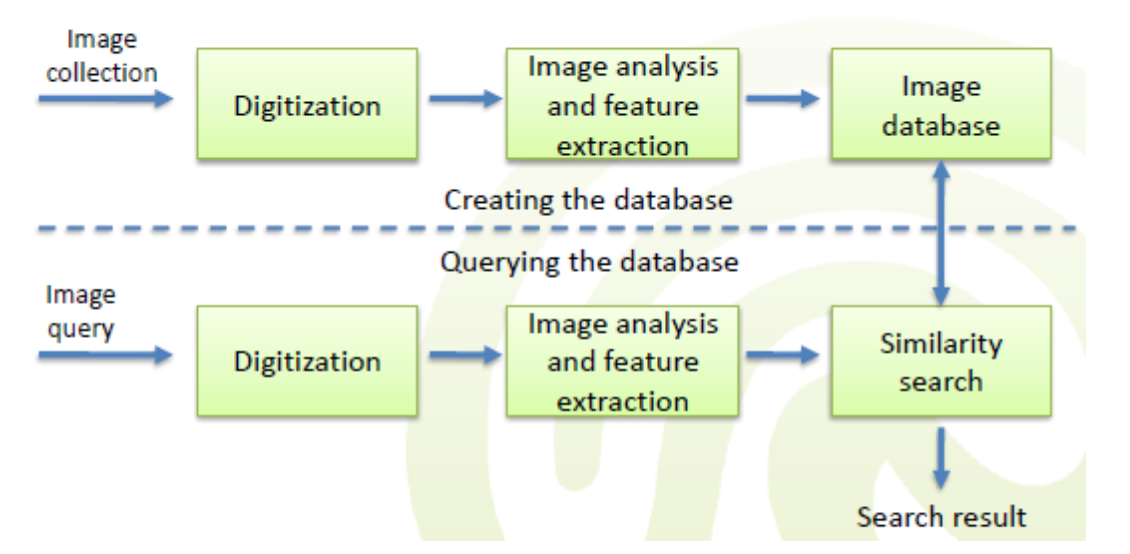
+Image query:

+Digitization:

+Image analysis and feature extraction:

+Similarity search:

+Search result:



Tổng quát thì quy trình tra cứu dữ liệu ĐPT gồm :

+Query:

+Result:

+Query preparation:

+Result preparation:

+Query plan & feature values:

+Result data:

+Similarity computation & query processing:

+Feature values:

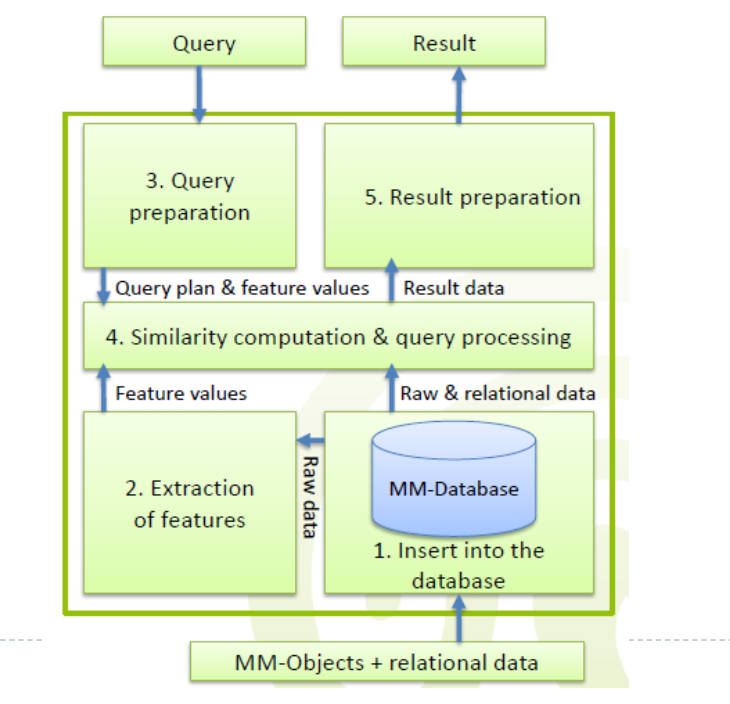
+Raw & relational data:

+Extraction of features:

+Raw data:

+Insert into the database:

+MM-Objects + relational data:



Ứng dụng của tra cứu CSDL đa phương tiện:

* Tra cứu thông tin trên web:
  + Mạng xã hội: Facebook, ig, weibo, twitter…
  + Chia sẻ hình ảnh: Flickr, Imeem, Picasa, pinterest…
  + Chia sẻ video: Youtube, Dailymotion…
* Camera giám sát: giao thông, an ninh…
* Nhận dạng khuôn mặt: Face id…

Thước đo sự hiêu quả của hệ csdl ĐPT

* Các thông số thực của hệ thống:
  + Bộ nhớ cần dùng
  + Tốc độ xử lý
  + Số thao tác input,output
  + Thời gian phản hồi
* Độ phụ thuộc vào phần cứng
* Mục tiêu: tính hiệu quả tương đối của hệ thống

-Thước đo sự hiệu quả của tra cứu dữ liệu

* Tuỳ thuộc vào yêu cầu tra cứu
* Đơn vị đánh giá chất lượng tìm kiếm
* Dựa trên câu truy vấn cụ thể
* Dựa trên thông tin cụ thể cần tìm kiếm

-Thước đo độ chính xác:

* Precision: là phần trăm của dữ liệu tìm đúng trong tổng số kết quả tìm được => dễ đánh giá hơn
* Recall: là phần trăm của dữ liệu tìm đúng trong tổng số dữ liệu cần tìm => khó đánh giá hơn
* Cả Precision và Recall chỉ có ý nghĩa khi chúng đồng thời được sử dụng

## 1.4 Kiến trúc hệ cơ sở dữ liệu ĐPT

Hai loại kiến trúc CSDL cơ bản :

* Hệ CSDL tập trung
* Hệ CSDL phân tán

Các nguyên lý hệ CSDL ĐPT

* Nguyên lý riêng biệt: mỗi loại dữ liệu đơn phương tiện được lưu trữ tại các máy chủ riêng biệt theo các cấu trúc phù hợp với đặc tính của từng loại
* Nguyên lý đồng nhất: tất cả các loại dữ liệu đa phương tiện được lưu trữ trong cùng một cấu trúc dữ liệu
* Nguyên lý kết hợp: kế hợp của cả hai loại trên

Các yêu cầu đối với máy chủ ĐPT

* Đáp ứng yêu cầu của nhiều loại ứng dụng
* Thỏa mãn Nhu cầu của nhiều người sử dụng khác nhau
* Băng thông truy cập tốt
* Dung lượng lưu trữ lớn (ứng với nhiều loại định dạng dữ liệu đa phương tiện)

Hệ CSDL đa phương tiện phân tán :

CSDL có thể được chia nhỏ với các mối quan hệ phụ

* Chia nhỏ theo chiều ngang
* Chia nhỏ theo chiều dọc

CSDL được phân mảnh

Các phân đoạn dữ liệu được nhân thành nhiều bản sao, có thể được lưu trữ ở các điểm khác nhau

LOB: Các dữ liệu (video, audio) được lưu trữ tại các điểm gần nơi yêu cầu dịch vụ.

## 1.5 Mô hình hóa dữ liệu đa phương tiện

**-**Mô hình dữ liệu MIRS (*Multimedia Information Retrieval System* ) là mô hình hướng đối tượng và phân cấp nhiều tầng

### 1.5.1 **Tầng đối tượng**:

Định nghĩa: đối tượng đa phương tiện: Một đối tượng gồm một hay nhiều mục phương tiện với các mối quan hệ không gian, thời gian được mô tả. Khía cạnh mấu chốt là cách mô tả các mỗi quan hệ không gian và thời gian. Các quan hệ không gian được mô tả nhờ hiển thị kích thước cửa sổ và vị trí của mỗi mục. Phương pháp chung đối với đặc tả thời gian là đặc tả dựa trên trục thời gian, có thời gian bắt đầu, thời lượng của mỗi mục theo đồng hồ chung. Các phương pháp khác gồm các mô hình theo sự kiện, theo kịch bản. Khi thể hiện, ngữ nghĩa của một ñối tượng hiện lên khi hiện các mối quan hệ không gian và thời gian.

### 1.5.2 Tầng kiểu phương tiện :

Tầng này mang các kiểu dữ liệu chung, như văn bản, đồ họa, hình ảnh, âm thanh và video. Mức này mô tả các đặc trưng hay thuộc tính. Chẳng hạn kiểu dữ liệu hình ảnh: có kích thước ảnh, lược ñồ màu, các đối tượng chính bao gồm trong hình ảnh. Các ñặc trưng này ñược dùng trực tiếp cho tính toán khoảng cách và tìm kiếm.

### 1.5.3 Tầng dạng thức phương tiện :

Tầng dạng thức đa phương tiện mô tả các dạng phương tiện mà dữ liệu đưcọ lưu trữ. Một phương tiện thường có nhiều dạng thức có thể. Chẳng hạn một hình ảnh có thể ở dạng bitmap thô hay dạng nén. Cũng có nhiều kĩ thuật và chuẩn nén. Thông tin trên tầng này đưcọ dùng cho mã hóa, phân tích và thể hiện

## 1.6 Trích chọn đặc trưng, chỉ số hóa và đo tính tương tự

### 1.6.1 Trích chọn đặc trưng

Đặc trưng và thuộc tính trong MIRS được trích, tham số hóa, lưu trữ với bản thân mục tin. Các đặc trưng và thuộc tính biễu diễn nội dung của từng khoản mục multimedia.

Các mục thông tin đa phương tiện trong cơ sở dữ liệu thường được tiền xử lí để trích đặc trưng và thuộc tính. Trong quá trình tìm kiếm, người ta tìm và so sánh theo đặc trưng và thuộc tính này. Do vậy, chất lượng của việc trích đặc trưng sẽ quyết định hiệu quả tìm kiếm. Nếu không trích đặc trưng thì không đáp ứng được yêu cầu cần thiết hoặc kết quả có thể bị sai khác. Đây là khác nhau chính giữa DBMS và MIRS trong DBMS có tất cả các thuộc tính và đầy đủ , trong MIRS cần trích đặc trưng và thuộc tính theo các kiểu câu truy vấn, và không thường không đầy đủ

-Trích chọn đặc trưng phải thỏa mãn các yêu cầu sau:

* Các đặc trưng và thuộc tính trích ra cần đầy đủ ñể thể hiện nội dung của mục thông tin
* Các đặc trưng cần được thể hiện và lưu trữ. Các đặc trưng phức tạp và lớn không phù hợp với việc trích đặc trưng
* Việc tính khoảng cách giữa các ñặc trưng cần hiệu quả, nếu không sẽ làm trễ thời gian phản hồi kết quả

Nhìn chung có 4 mức đặc trưng như sau :

* **Metadata*:*** bao gồm các thuộc tính của các đối tượng đa phương tiện như tên tác giả, ngày tạo lập, tiêu đề đối tượng. Không mô tả hay diễn giải nội dung của đối tượng. Các thuộc tính này được quản lý bằng kỹ thuật DBMS.
* **Mô tả bằng văn bản :** mô tả về nội dung đối tượng **,** có thể ở dạng một số các từ khóa hay mô tả bằng văn bản thông thường . Việc chỉ số hóa và tìm kiếm dựa trên ghi chú có thể được kỹ thuật IR thực hiện
* **Đặc trưng nội dung mức thấp :** Nắm các mẫu dữ liệu và các thống kê về đối tượng đa phương tiện, và quan hệ không gian, thời gian giữa chúng. Phương tiện khác nhau có đặc trưng nội dung mức thấp khác nhau.
  + Với âm thanh : Đặc trưng mức thấp là ồn trung bình, phân bố tần số, tỉ lệ im lặng.
  + Với dữ liệu ảnh : Đặc trưng mức thấp gồm phân bố màu, bề mặt, dạng ñối tượng và cấu trúc không gian.
  + Với dữ liệu video, ñặc trưng mức thấp gồm cấu trúc thời gian và các đặc trưng như đối với ảnh.

Ưu điểm chính của đặc trưng mức thấp là cho phép trích tự động

* **Trích đặc trưng mức cao** : nhằm nhận dạng và hiểu đối tượng. Trừ việc nhận dạng văn bản và tiếng nói, nhìn chung khó nhận dạng và hiểu đoạn âm thanh hay các đối tượng video. Tuy nhiên người ta cũng có ý tưởng giải quyết. Trong ứng dụng đặc biệt với số đối tượng hạn chế, việc mô tả và nhận dạng đối tượng chung là có ích và có thể ñược. Chẳng hạn trên 95% video có chủ đề con người, nhóm người. Vậy hệ thống nhận ra và diễn giải con người là có ý nghĩa. Ban đầu công việc này là bán tự động

### 1.6.2 : Chỉ số hóa

Sau khi trích đặc trưng, cần sử dụng các cấu trúc chỉ số để tổ chức các đặc trưng này sao xho việc tìm kiếm hiệu quả. Nhiều đặc trưng cần thiết cho việc thể hiện một đối tượng và mỗi ñặc trưng cần thể hiện quan nhiều tham số. Chẳng hạn phân bố màu thường được thể hiện qua lược đồ với nhiều màu. Các cấu trúc chỉ số dùng trong DBMS là không phù hợp cho các đặc trưng theo nội dung. Việc chỉ số hóa trong MIRS sẽ là phân cấp và có nhiều mức: mức cao nhất là phân loại ứng dụng; mức chỉ số thứ hai khác với mức đặc trưng; chỉ số hóa khác nhau cần cho nhiều ñặc trưng; mức chỉ số hóa thứ ba theo mối quan hệ thời gian và không gian giữa các đối tượng.

### 1.6.3 Đo tính tương tự

**-**Truy vấn đa phương tiện trên cơ sở tính tương tự thay cho đối sánh chính xác giữa các item truy vấn và các item trong CSDL. Tính tương tự được tính toán trên cơ sở các đặc trưng, thuộc tính trích chọn và dưới dạng một hay nhiều giá trị. Tuy nhiên, tương quan của kết quả truy vấn do con người quyết định. Các kiểu đặc trưng được sử dụng để mô tả các đối tượng đóng vai trò quan trọng để phù hợp với yêu cầu này. Thước đo tính tương tự rất phức tạp vì quyết định của người sử dụng là chủ quan và phụ thuộc ngữ cảnh.

# CHƯƠNG 2 : MỘT SỐ KĨ THUẬT XỬ LÍ VÀ NHẬN DẠNG ẢNH NỀN BẦU TRỜI HIỆN NAY :

Một số kĩ thuật phổ biến như sau :

* Nhận dạng ảnh dựa trên màu sắc
* Nhận dạng ảnh dựa trên hình dạng vật thể trong ảnh
* Nhận dạng ảnh dựa trên bố cục,kết cấu của ảnh
* Nhận dạng ảnh dựa trên các đặc trưng bất biến

## 2.1 : Nhận dạng ảnh dựa trên màu sắc :

Hình ảnh bao gồm một mảng các điểm ảnh(pixel) , mỗi pixel thể hiện một màu sắc Có nhiều không gian được sử dụng để tính toán các giá trị màu của pixel như : RGB,HSV,…

Mỗi dải mầu cơ bản được chia nhỏ thành 𝑚 mức. Tổng số sẽ có 𝑛 = 𝑚\*m\*m : mức mầu tương ứng với số ô mầu

Xây dựng biểu đồ tần suất mầu 𝐻(𝑀) = [ℎ1 , ℎ2 , … , ℎ𝑗 , … , ℎ𝑛 ] với mỗi ℎ𝑗 là số điểm ảnh mức mầu j.

𝐻(𝑀) được lưu trữ nhận dạng sử dụng các kỹ thuật nhận dạng dựa trên không gian vector, hoặc dựa trên phân cụm.

Nhận dạng ảnh theo lược đồ màu là phương pháp phổ biến và được sử dụng nhiều

nhất trong các hệ thống nhận dạng ảnh. Đây là phương pháp đơn giản, tốc độ cho ra kết quả tương đối nhanh tuy nhiên kết quả có độ chính xác không cao.

Một số lược đồ màu được sử dụng như: lược đồ màu RGB, lược đồ màu HSV

Trong đó, lược đồ màu RGB được sử dụng phổ biến nhất.

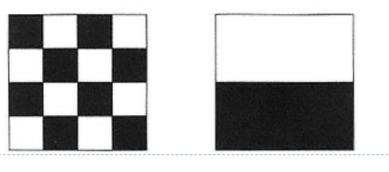
Sử dụng biểu đồ tần suất theo trọng số (PWH)

* Ứng với từng ô mầu trong cách tính biểu đồ tần suất, người ta chọn ra một mầu đại diện.
* 10 mầu đại diện gần nhất được chọn làm tâm.
* Tính khoảng cách giữa mầu của từng pixel đến các mầu đại diện đó.
* Trọng số của 10 mầu đại diện cho mỗi điểm ảnh sẽ được tính bằng nghịch đảo của các khoảng cách đó.

**Nhược điểm** : của phương pháp nhận dạng ảnh dựa trên mầu sắc là không tính đến mối liên hệ về không gian giữa các điểm ảnh.

**Khắc phục** : Sử dụng mối liên hệ về không gian giữa các điểm ảnh

* Ảnh được chia thành nhiều vùng nhỏ cố định
* Tại mỗi vùng ảnh, biểu đồ tần suất mầu được thiết lập
* Trong quá trình nhận dạng, biểu đồ tần suất mầu tại các vùng tương ứng sẽ được so sánh với nhau



**Nhược điểm** : của phương pháp này là sự ảnh hưởng của hiệu ứng che mặt

* Các ô mầu lớn sẽ lấn át nội dung của ô mầu nhỏ.
* Các ảnh có mầu nền giống nhau sẽ mang nội dung giống nhau, mặc dù các vật thể trong đó là khác hẳn nhau.

Có thể sử dụng biểu đồ tần suất mầu sắc riêng biệt cho hình nền và các vật thể: phân chia ảnh thành nhiều khung chữ nhật, mỗi khung chứa riêng vật thể hoặc hình nền.

Việc xác định vật thể hay hình nền được thực hiện dựa trên sự biến thiên giá trị mầu theo trục ngang hay trục dọc

Mỗi ảnh được đại diện bởi hai biểu đồ tần suất.

Quá trình nhận dạng có thể thực hiện theo một trong các cách:

* + Dựa vào biểu đồ tần suất mầu tổng thể của ảnh
  + Dựa vào biểu đồ tần suất của các vật thể trong ảnh
  + Dựa vào biểu đồ tần suất mầu của hình nền
  + Dựa vào biểu đồ tần suất mầu của cả hình nền lẫn vật thể (có trọng số)

Sử dụng thông tin về thống kê của sự phân bố mầu sắc (giống lượng tử hoá phi tuyến): dải mầu có phân bố nhiều điểm ảnh hơn sẽ được phân chia nhỏ hơn.

Nếu chúng ta coi thông tin màu của ảnh là tín hiệu một, hai, hoặc ba chiều đơn giản thì việc phân tích các tín hiệu sử dụng ước lượng mật độ xác xuất là một cách dễ nhất để mô tả thông tin màu của ảnh.

Có ba kỹ thuật truyền thống được sử dụng trong nhận dạng ảnh dựa trên màu sắc đó là biểu đồ màu toàn cục (Global Color Histogram), biểu đồ màu cục bộ (Local Color Histogram) và biểu đồ màu tương quan (Color Correlogram Histogram). Những kỹ thuật này thích hợp với các kiểu ảnh khác nhau

### 2.1.1 Biểu đồ màu toàn cục (Global Color Histogram)

Biểu đồ màu loại này mô tả phân bố màu sử dụng tập các màu. Việc sử dụng biểu đồ màu toàn cục thì một ảnh sẽ được mã hóa với biểu đồ màu của nó và khoảng cách giữa hai ảnh sẽ được xác định bởi khoảng cách giữa những biểu đồ màu của chúng. Với kỹ thuật này chúng ta có thể sử dụng các thước đo khác nhau để tính toán khoảng cách giữa hai biểu đồ màu. Ví dụ dưới đây sẽ mô tả hoạt động của kỹ thuật này:

Trong biểu đồ màu mẫu có 3 màu : black, white and grey.

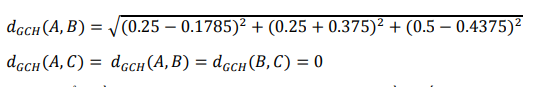
Ta kí hiệu biểu đồ màu của ảnh

A:{25%, 25%, 50%};

B: {18.75%, 37.5%, 43.75%}

Ảnh C có biểu đồ màu như ảnh B.

Nếu sử dụng thước đo khoảng cách Euclidean để tính toán khoảng cách biểu đồ thì khoảng cách giữa hai ảnh A và B cho biểu đồ màu toàn bộ là

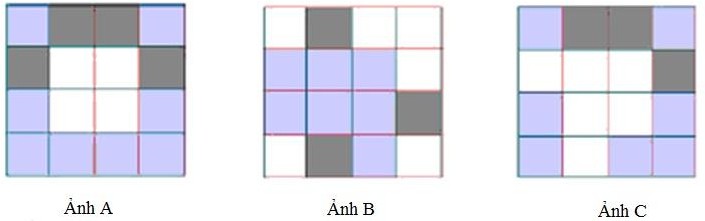


**Ưu điểm:**

* Tính toán nhanh, đơn giản.
* Thích hợp trong các ứng dụng thời gian thực.

**Nhược điểm :**

* Không chứa các thông tin liên quan đến sự phân bố màu của các vùng. Vì vậy khoảng cách giữa các ảnh đôi khi không thể chỉ ra được sự khác nhau thực sự giữa các ảnh.
* Ví dụ khoảng cách giữa ảnh Avà C khác so với khoảng cách giữa ảnh A và B nhưng bằng việc xây dựng biểu đồ màu toàn cục thì lại thu được khoảng cách tương tự. Ngoài ra còn có trường hợp hai ảnh khác nhau có biểu đồ màu toàn cục giống nhau như ví dụ trên ảnh B và C. và đây chính là hạn chế của biểu đồ màu toàn bộ.



### 2.1.2 Biểu đồ màu cục bộ (Local Color Histogram)

Biểu đồ màu cục bộ bao gồm thông tin liên quan đến sự phân bố màu của các vùng. Trước tiên là nó phân đoạn ảnh thành nhiều khối và sau đó biểu diễn biểu đồ màu cho mỗi khối, một ảnh sẽ được biểu diễn bởi những biểu đồ màu này. Khi so sánh hai hình ảnh, khoảng cách được tính toán bằng cách sử dụng những biểu đồ của chúng giữa một vùng trong một ảnh và một vùng tương ứng trong ảnh khác. Khoảng cách giữa hai ảnh được xác định bằng tổng tất cả các khoảng cách này. Nếu sử dụng căn bậc hai của khoảng cách Euclidean để tính toán khoảng cách biểu đồ thì khoảng cách giữa hai ảnh Q và I cho biểu đồ màu cục bộ là:



Ở đây M là số vùng được phân đoạn trong ảnh, N là số màu trong biểu đồ màu và H[i] là giá trị của màu i trong biểu đồ màu đại diện cho vùng k của ảnh.

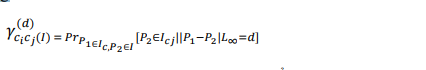
### 2.1.3 Biểu đồ màu tương quan (Color Correlogram Histogram)

Quan sát thấy rằng lược đồ màu thiếu thông tin về cách mà màu sắc được phân bố theo không gian, một đặc trưng mới được giới thiệu gọi là lược đồ tương quan màu. Lược đồ tương quan màu hứa hẹn mô tả không chỉ là phân phối màu của các điểm ảnh mà còn là sự tương quan về không quan giữa các cặp màu. Lược đồ này chỉ quan tâm đến sự tương quan về không gian giữa những màu giống nhau và do đó giảm được số chiều và chi phí tính toán.

Cách tính lược đồ tương quan màu

Gọi [D] là tập gồm D khoảng cách d1 , d 2 ,..., d D được đo bằng độ đo L∞ .

Lược đồ tương quan màu của ảnh I được xác định với cặp màu ci , c j và khoảng cách d như sau:



Trong đó I là ảnh, kích thước MxN (Điểm ảnh), I c = p ∈ I | Ip= c, lược đồ tương quan màu thể hiện xác suất cặp điểm ảnh bất kỳ p1 và p2 chịu sự ràng buộc về màu (p1 có màu ci, p2 có màu c j ) và vị trí (p1− p2|L= d).

## 2.2 Nhận dạng ảnh dựa trên hình dạng

### Tổng quan :

Ảnh cần được chia thành các mảnh chứa vật thể :

Mỗi hình ảnh vật thể cần phải được diễn tả thống thất, không phụ thuộc vào sự biến đổi của sự vật, vị trí, kích thước,…

Các hình giống nhau cần phải được diễn đạt giống nhau

Trong quá trình nhận dạng, người dùng sử dụng một hình ảnh mẫu hoặc tự vẽ ra một hình mẫu để hệ thống tìm kiếm các ảnh có hình giống như vậy.

Kết quả trả về thường là nhiều ảnh, sắp xếp theo thứ tự giảm dần về sự giống nhau với hình mẫu

Các khái niệm chung và một số thước đo về hình dạng đơn giản:

* Trục chính: đường thẳng nối hai điểm xa nhất trong hình
* Trục phụ: đường thẳng vuông góc với trục chính, có độ dài sao cho một hình chữ nhật với các cạnh song song và dài bằng trục chính và trục phụ chứa được toàn bộ đường bao của hình.
* Hình chữ nhật cơ bản: được hình thành như mô tả ở trên
* Tâm sai / độ lệch tâm: tỉ lệ về độ dài giữa trục chính và trục phụ
* Những thông tin cơ bản nói trên có thể được sử dụng để nhận dạng hình ảnh

Diễn tả hình dạng theo khu vực và các tiêu chí so sánh tương tự

Khái niệm cơ bản của diễn tả hình dạng theo khu vực:

* Ứng với mỗi hình thể nào đó, ta đặt một lưới (gồm các hình vuông có kích thước cố định) lên đó
* Ứng với mỗi hình vuông có trên 15% là diện tích vật thể, ta gán giá trị 1
* Các hình vuống còn lại có giá trị là 0.
* Đọc giá trị của các hình vuông từ trái sang phải, từ trên xuống dưới để có một chuỗi nhị phân đại diện cho hình thể đó.
* Chuẩn hóa xoay vòng
* Chuẩn hóa thước đo
* Diễn tả hình dạng thống nhất – đánh chỉ số hình dạng
* Các thước đo về sự tương tự.

Phân đoạn ảnh là quá trình phân nhóm các pixel trong ảnh dựa trên các tiêu chuẩn tương đồng về màu, về texture, hoặc dựa trên các đường biên kết nối, Khi đóhình dạng là thuộc tính chính của các vùng ảnh phân đoạn , hình dạng không phải là một thuộc tính của ảnh. Do đó, hình dạng thường được mô tả sau khi các ảnh được phân đoạn thành các vùng hoặc các đối tượng

Hình dạng chỉ là biên của đối tượng nào đó trong ảnh. Một biểu diễn đặc trưng hình dạng tốt cho một đối tượng phải bất biến với dịch chuyển, quay và tỷ lệ. Các bài toán trích trọn đặc trưng dựa trên hình dạng thường được bắt đầu với việc tìm và phát hiện biên của đối tượng, qua đó định hình cấu trúc và các thông tin bất biến của đối tượng ảnh

Biên cạnh là đối tượng phân cách giữa 2 vùng ảnh thuần nhất có độ sáng khác nhau (Biên là nơi có biến thiên về độ sáng). Tập hợp các điểm biên tạo thành biên hay đường bao của ảnh (boundary). Như vậy phát hiện biên một cách lý tưởng là xác định được tất cả các đường bao trong các đối tượng

Việc phân loại các phương pháp biểu diễn hình dạng phổ biến nhất là dựa trên việc sử dụng các điểm biên hình dạng và điểm vùng

Một số kĩ thuật để phát hiện biên đối tượng phổ biến:

### 2.2.1 Thuật toán canny :

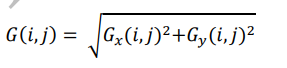
Phát hiện các đường biên mảnh với các ảnh có nhiễu

Thuật toán Canny gồm  5 bước :

Bước 1: Làm trơn ảnh

Bước 2***:*** Tính gradient của ảnh bằng mặt nạ PreWitt (hoặc các ma trận phát hiện biên khác như Roberts, Sobel..), kết quả đặt vào Gx,Gy.

+ Bước3:Tính độ lớn và hướng của gradient tại mỗi điểm (i,j) của ảnh. Độ lớn G(i,j) tại điểm (i,j) được tính dựa vào định lý Pitago



Ngoài ra hướng của véc tơ gradient được tính với công thức:

https://lh3.googleusercontent.com/tJPuMFKgjbmF1hdk1s74K0uucbW0l6ftCcOstqjI8Hhem2FH5H9cRSHs9jdMkvOfr02ls_U-n3moK7nHMNhIDDrZoYPFjUuNMptn4eeG-dCtqQBNBk0u67vYdfP0_5eTsan9PWmW

Bước 4: Bước này loại bỏ những điểm không phải là cực đại địa phương để xóa bỏ những điểm không thực sự là biên

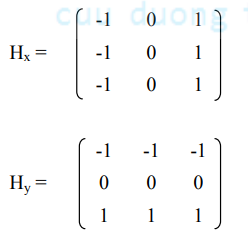
Bước 5: Tại bước này, Canny sử dụng hai ngưỡng cao và thấp Đầu tiên, Canny lọc các điểm được giữ lại sử dụng ngưỡng cao .  Khi thực hiện việc dò theo biên, Canny sử dụng ngưỡng thấp để xác định điểm dừng của biên

### 2.2.2 Kĩ thuật La bàn :

Sử dụng 8 mặt nạ nhân chập theo 8 hướng 0dèg , 450deg , 90deg , 135deg, 180deg , 225deg , 270deg , 315deg

### 2.2.3 Kĩ thuật prewitt :

Sử dụng 2 mặt nạ nhân chập theo 2 hướng x và y



Các bước tính toán của kỹ thuật Prewitt

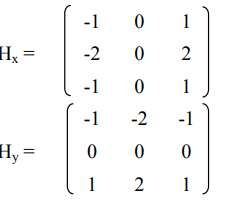
 + Bước 1: Tính I https://lh3.googleusercontent.com/BZp8sj2XsFp8cF1dZqEfXE5B1LdMBa_mRwo0px2vofUHbzK0ujtGKA-STsJTcMi_qnGQo8QSVnpfcwpM04VtayPsKEbrn9rQYFtmvzeyPVX8dCeJxh3RvJxH92FGhnK3lKLEe8_c Hx và I https://lh5.googleusercontent.com/TR5CGlGZmuS3D9PGbJmIoVZy3ygMiexo5XJHwdAc5uz9z7jOuLwwhYBdWT5lC6oDSZiol5BlXj13BbMWBMKKHQwBJh-YVrHwNQqEqt_dg6x34VxWTH2mASAYEmw2BFEXlB9cJvsg Hy

+ Bước 2: Tính I https://lh6.googleusercontent.com/z7wA7c_bBAbx7THnrfNs-P0mnRvP5xRNtpQdkflO2Dy_ARXEZPP1OUP6OBXF7WBTUFv6HhA87fAzrPBMi4lUEdfUnCHXhd9Nn7DzN1p6aETYevQxH65ddSbBUTNGaR9yc7sa8NI3  Hx + I https://lh6.googleusercontent.com/Xemqtd1EWGQV5ID0qHP-pkXFBXIzhic6Dcg9w_seiAkSTFcgPD_0fskRHJbSfAZ3PL0xKHv4hWzYQ_0-Tq2ws5MhqyA6jOmyEhqot7J5jrbePZYnrqCqpAT_sdT0ozgW0PU4G_xr Hy

+ Bước 3: Phân ngưỡng theo θ để có ảnh biên

### 2.2.4 Kĩ thuật Sobel

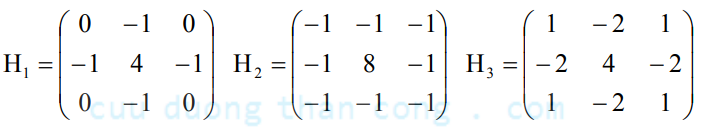
Sử dụng 2 mặt nạ nhân chập  sobelx,sobely



Các bước tính toán tương tự như prewitt

### 2.2.5 Kĩ thuật Phát hiện biên Laplace

Sử dụng đạo hàm bậc 2 laplace (hiệu quả khi mức xám ảnh thay đổi chậm) .Trong thực tế, người ta thường dùng nhiều kiểu mặt nạ khác nhau để xấp xỉ rời rạc đạo hàm bậc hai Laplace. Dưới đây là ba kiểu mặt nạ thường dùng:



## 2.3 Truy vấn ảnh dựa trên bố cục, kết cấu hoa văn

Kết cấu ảnh là một đặc trưng ảnh quan trọng để mô tả các thuộc tính bề mặt của một đối tượng và mối quan hệ của nó với các vùng xung quanh . Do các đặc trưng kết cấu được xuất hiện trong nhiều ảnh thực, chúng rất quan trọng và có lợi ích trong các nhiệm vụ tra cứu ảnh và nhận dạng mẫu. Tuy nhiên, độ phức tạp tính toán và độ chính xác tra cứu là những nhược điểm chính của các hệ thống tra cứu ảnh dựa vào kết cấu .

### 2.3.1 Các thuộc tính chủ yếu liên quan đến bố cục ảnh

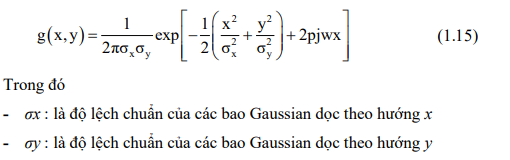
* Độ thô của ảnh (ngược với độ mịn): sự khác biệt giữa các phần tử của ảnh
* Độ tương phản: khoảng động của dải mầu xám, sự phân cực giữa hai mầu sáng và tối trong biểu đồ tần suất, độ sắc của các nét, tần suất lặp lại của các khối mầu
* Hình dạng và chiều sắp xếp của các hình trong ảnh
* Sự tương đồng của các đường viền
* Tính trật tự của ảnh
* Bố trí mầu sắc trong ảnh (liên tục hay gián đoạn)

### 2.3.2 Một số phương pháp được sử dụng phổ biến:

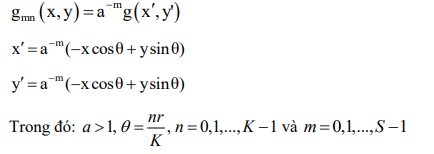
#### 2.3.2.1 : Bộ Lọc Gabor :

Lọc Gabor được sử dụng rộng rãi để trích rút các đặc trưng ảnh, đặc biệt là các đặc trưng kết cấu. Nó tối ưu về mặt cực tiểu hoá sự không chắc chắn chung trong miền không gian và miền tần số, và thường được sử dụng như một hướng và tỷ lệ biên điều hướng và phát hiện đường. Có nhiều cách tiếp cận đã được đề xuất để mô tả các kết cấu của các ảnh dựa trên các lọc Gabor. Ý tưởng cơ bản của sử dụng các lọc Gabor để trích rút các đặc trưng kết cấu

Hàm Gabor hai chiều g(x, y) được định nghĩa:

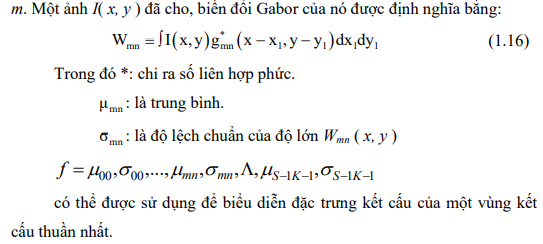


Sau đó một tập các lọc Gabor có thể thu được bởi sự co giãn và quay thích hợp của g( x, y ) :



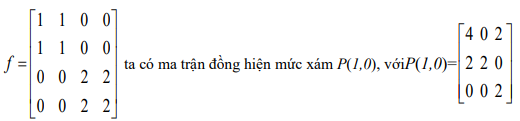
K và S: là số các hướng và các tỷ lệ



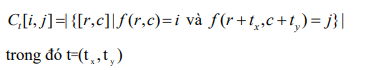


#### 2.3.2.2 Ma trận đồng nhất mức xám :

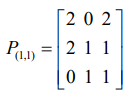
Ma trận đồng hiện mức xám là ma trận lưu trữ số lần xuất hiện của những cặp điểm ảnh trên một vùng đang xét. Các cặp điểm này được tính theo những quy luật cho trước. Ví dụ với ảnh f như sau:



(lưu ý là có rất nhiều ma trận đồng hiện mức xám khác nhau cho một ma trận ban đầu) Ma trận đồng hiện mức xám trên tạo ra bởi những cặp điểm lệch nhau (1,0) nghĩa là 2 điểm kế nhau trên cùng hàng. Giá trị tại dòng 0, cột 0 của ma trận đồng hiện trên là 4 vì ảnh f có 4 cặp điểm 0 0 kế nhau trên cùng một hàng. Tương tự như vậy, giá trị ở dòng 1, cột 2 của ma trận là 0 vì không có cặp 1 2 nào xuất hiện nhau trên cùng một hàng. Công thức tổng quát của ma trận đồng hiện mức xám là :



Ví dụ với ma trận f đã cho như trên thì khi t=(1,0) ta sẽ có ma trận đồng hiện như ví dụ trên, và khi t=(1,1), nghĩa là tìm những cặp điểm kế nhau trên cùng một đường chéo, ta có ma trận đồng hiện là:



Từ ma trận đồng hiện mức xám người ta định nghĩa ra các đặc trưng về : Energy (năng lượng), Entropy, Maximan Probability:

* Ngoài ra còn 1 số kĩ thuật như **:**biến đổi Wavelet, trường ngẫu nhiên Markov ,mô tả lược đồ cạnh ,phân rã tháp , các đặc trưng Tamura .

## 2.4 Tra cứu ảnh dựa trên đặc trưng bất biến

Phương pháp tra cứu này có tên là Scale-Invariant Feature Transform (SIFT) và đặc trưng trích rút đựợc gọi là đặc trưng SIFT. Phương pháp này trích rút các đặc trưng cục bộ bất biến của ảnh. Các đặc trưng này bất biến với việc thay đổi tỉ lệ ảnh, quay ảnh, đôi khi là thay đổi điểm nhìn và thêm nhiễu ảnh hay thay đổi cường độ chiếu sáng của ảnh. Các đặc trưng này được trích rút ra từ các điểm đặc trưng cục bộ. Điểm đặc trưng: Là vị trí (điểm ảnh) "đặc trưng" trên ảnh. "Đặc trưng" ở đây có nghĩa là điểm đó có thể có các đặc trưng bất biến với việc quay ảnh, co giãn ảnh hay thay đổi cường độ chiếu sáng của ảnh

Giống như nhiều thuật toán về xử lý ảnh, SIFT là thuật toán khá phức tạp, phải trải qua nhiều bước xử lý và sử dụng nhiều kiến thức về toán học. Sau đây sẽ là các bước chính trong thuật toán:

- Dò tìm cực trị trong không gian đo (Scale space Extrema Detection)

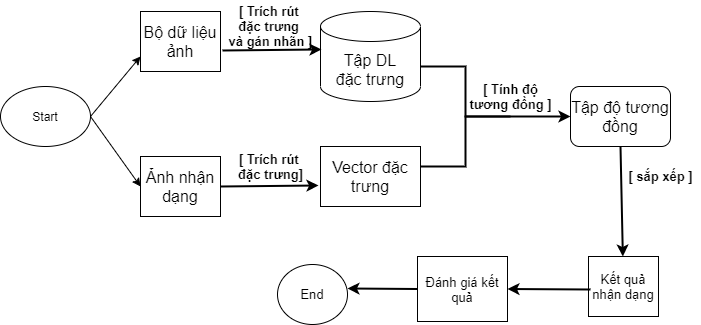
- Lọc và trích xuất các điểm đặc biệt (Keypoint Localization)

- Gán hướng cho các điểm đặc trưng (Oriented Assignment)

- Bộ mô tả điểm đặc trưng (Keypoint Descriptor)

# CHƯƠNG 3 : SƠ ĐỒ KHỐI QUY TRÌNH THỰC HIỆN BÀI TOÁN NHẬN DẠNG VÀ PHÂN LOẠI ẢNH BẦU TRỜI

## 3.1 Sơ đồ khối bài toán

****

## 3.2 Quy trình thực hiện bài toán :

* + Đọc tập ảnh đầu vào
  + Trích rút vector đặc trưng và gán nhãn từng ảnh tương ứng
  + Lưu tập dữ liệu đặc trưng và nhãn dưới dạng file csv
  + Đọc ảnh nhận dạng
  + Trích rút vector đặc trưng
  + Đo độ tương đồng ảnh đầu vào với tập dl đặc trưng đã lưu
  + Sắp xếp tập độ tương đồng tăng dần
  + Trả kết quả dự đoán và đánh giá thuật toán

## 3.3 Thuộc tính nhận dạng và phân loại

Đối với bài toán nhân dạng ảnh bầu trời thuộc tính chủ yếu dùng để nhận dạng là dựa vào màu sắc của ảnh trong đó :

-Trời ban ngày âm u : Màu xám + màu trắng

- Hoàng hôn : Màu đỏ +màu vàng+màu cam +màu tím

-Bình minh : màu xanh nhạt+ màu vàng nhạt+ màu trắng

- Buổi trưa : màu xanh + màu trắng

## 3.4 Kĩ thuật trích rút thuộc tính màu sắc

### Tổng quan về kĩ thuật : **Thống kê kênh màu(Color channel statistics)**

Bằng cách tính toán các số liệu thống kê cơ bản :  **trung bình màu**  và **độ lệch chuẩn (**Standard Deviation**)** cho mỗi kênh màu của hình ảnh từ đó có thể định lượng và biểu thị sự  **phân bố màu sắc**  của hình ảnh. Do đó, nếu hai hình ảnh có trung bình và độ lệch chuẩn gần tương tự nhau, chúng ta có thể giả định rằng những hình ảnh này có sự phân bố màu sắc tương tự

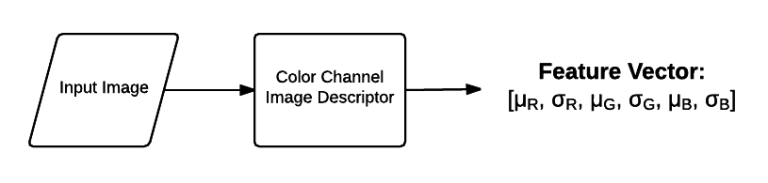
Note :Hai bức ảnh có sự phân bố màu sắc tương tự nhau không có nghĩa là chúng có nội dung tương đương . Tuy nhiên đối với ảnh bầu trời tại các thời điểm khác nhau trong ngày ảnh sẽ có màu sắc khác nhau nên ta coi hai bức ảnh có sự phân bố màu sắc tương tự nhau thì sẽ nội dung tương tự nhau.

Các bước để thực hiện thống kê kênh màu gồm:

### Tách hình ảnh đầu vào thành các kênh tương ứng. Đối với hình ảnh RGB, tách từng kênh Đỏ, Xanh lục và Xanh lam một cách độc lập.

### Tính giá trị trung bình màu và độ lệch chuẩn của từng kênh màu

### Ghép các giá trị trung bình và độ lệch chuẩn của từng kênh thành 1 list để tạo thành Feature vector 6-d

****

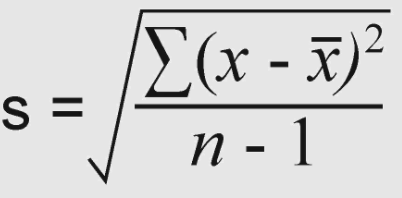
Công thức tính trung bình màu của từng kênh:



Với a1,a2,… an là giá trị màu từng điểm ảnh của kênh màu đang xét

N : Tổng số điểm ảnh của kênh màu

Công thức tính độ lệch chuẩn :



Với n : tổng số điểm ảnh của kênh

X : giá trị màu từng điểm ảnh của kênh màu đang xét

¯¯X Giá trị trung bình màu của kênh màu đang xét

Với việc tính độ lệch chuẩn và giá trị trung bình màu từng kênh sẽ giúp chúng ta đo lường sự phân tán màu sắc của ảnh

Ví dụ :



Với ảnh bình minh như trên ta có vector đặc trưng :

[85.61716757 81.8975008 81.48909837 46.90843985 52.24979095 76.02248777]

Gồm :

Trung bình kênh màu Blue : 85.61716757

Trung bình kênh màu Green : 81.8975008

Trung bình kênh màu Red : 81.48909837

Độ lệch chuẩn kênh màu Blue: 46.90843985

Độ lệch chuẩn kênh màu Green: 52.24979095

Độ lệch chuẩn kênh màu Red: 76.02248777

Ta thấy : Ảnh trên kênh màu Blue có độ đồng đều hơn khi có độ lệch chuẩn thấp nhất , trong khi kênh màu Red có sự phân bố giá trị màu sắc rời rạc nhất.

Code minh họa chi tiết việc trích rút thuộc tính màu săc sẽ được trình bày trong chương 4

## 3.5 Kĩ thuật lưu trữ thuộc tính màu sắc

Sau khi có tập vector đặc trưng và nhãn tương ứng với đặc trưng đó ta sẽ lưu lại tập vector dưới dạng file CSV(Comma Seperated Value )

Tệp CSV là một loại tệp văn bản thuần túy sử dụng cấu trúc cụ thể để sắp xếp dữ liệu dạng bảng **.**

**Đặc điểm của file csv là:**

1. Phổ biến ,dễ dàng chia sẻ
2. Truy vấn tốc độ cao và tiết kiệm dung lượng lưu trữ
3. Đa nền tảng, thư viện thao tác đọc ghi nhanh chóng , có nhiều ngôn ngữ sử dụng như C/C++, Java, Python, Golang....
4. Không giới hạn kích thước

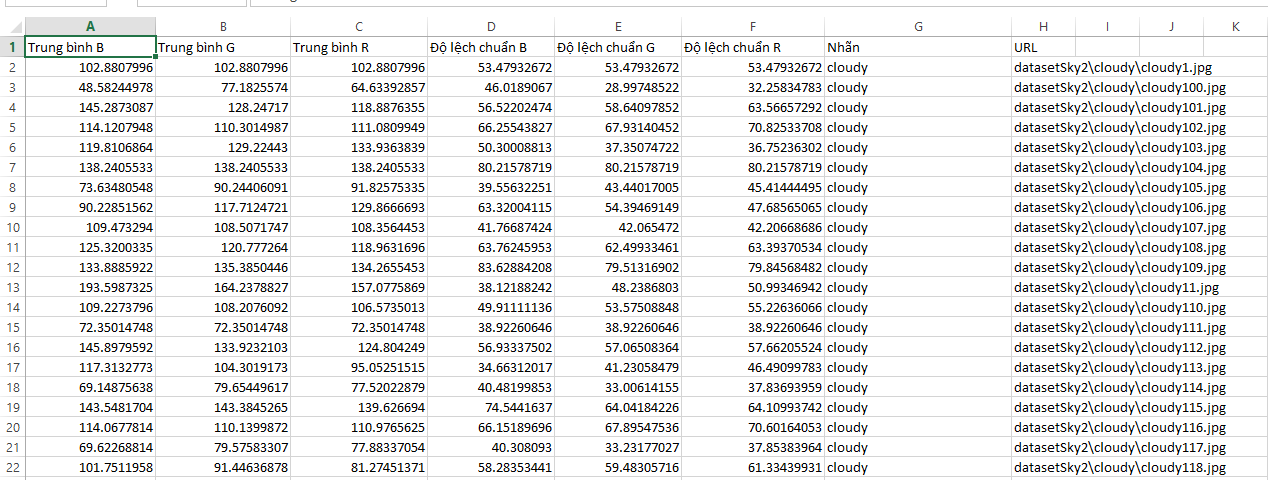
Việc lưu trữ đặc trưng giúp cho quá trình nhận dạng nhanh hơn khi chỉ cần đọc dữ liệu từ file đã lưu đỡ mất thời gian tính toán lại các đặc trưng

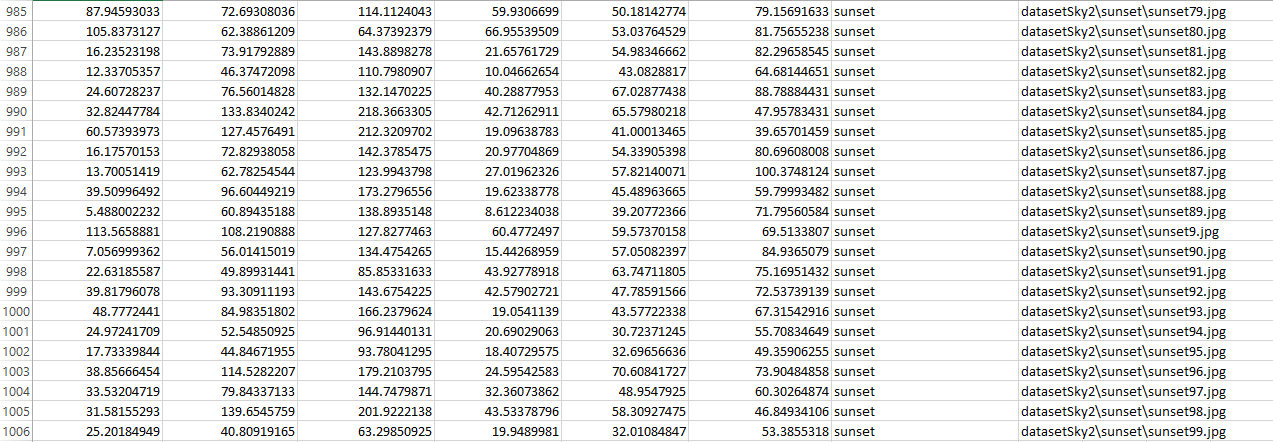
Thư viện pandas trong python giúp chúng ta có thể lưu trữ và đọc dữ liệu 1 cách dễ dàng

Ta sẽ lưu từng đặc trưng vào mỗi dòng trong file csv , kèm theo nhãn tương ứng, ngoài ra ta lưu thêm 1 cột đường đẫn của ảnh có đặc trưng tương ứng để phục vụ cho việc in ảnh ra màn hình

Code minh họa chi tiết việc lưu trữ sẽ được trình bày ở chương 4

Một đoạn hình ảnh file đặc trưng sau khi lưu



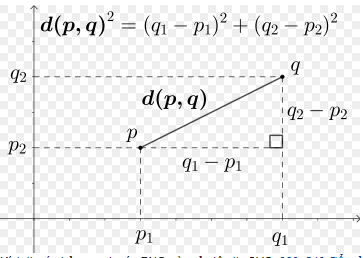


## 3.7 Kĩ thuật nhận dạng ảnh bầu trời từ thuộc tính màu sắc

Để nhận dạng ảnh qua vector đặc trưng ta có thể sử dụng cách đo độ tương đồng giữa 2 vector công thức thường được sử dụng là công thức tính khoảng cách Euclid

### 3.7.1 Giới thiệu về khoảng cách Euclidean

Định nghĩa : Trong [toán học](https://vi.wikipedia.org/wiki/To%C3%A1n_h%E1%BB%8Dc), khoảng cách Euclidean là [khoảng cách](https://vi.wikipedia.org/wiki/Kho%E1%BA%A3ng_c%C3%A1ch) "thường" giữa hai điểm mà người ta có thể đo được bằng cây thước, và được tính bằng [công thức Pytago](https://vi.wikipedia.org/wiki/%C4%90%E1%BB%8Bnh_l%C3%BD_Pytago).



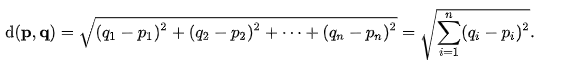
Giả sử : Khoảng cách Euclid giữa hai điểm p và q là [chiều dài](https://vi.wikipedia.org/wiki/Chi%E1%BB%81u_d%C3%A0i) [đoạn thẳng](https://vi.wikipedia.org/wiki/%C4%90o%E1%BA%A1n_th%E1%BA%B3ng) {\displaystyle {\overline {\mathbf {p} \mathbf {q} }}}  .

Trong [hệ tọa độ Descartes](https://vi.wikipedia.org/wiki/H%E1%BB%87_t%E1%BB%8Da_%C4%91%E1%BB%99_Descartes), nếu :

p = (*p*1, *p*2,..., *pn*)

 q = (*q*1, *q*2,..., *qn*)

Là hai điểm trong [không gian Euclid *n* chiều](https://vi.wikipedia.org/wiki/Kh%C3%B4ng_gian_Euclide), thì khoảng cách từ p đến q được tính theo công thức:



Ví dụ với 2 ảnh

Ta có lần lượt 2 vector đặc trưng

Vector 1 : [108.88416773 ,106.71035953, 105.79655612, 45.65301785, 46.99392051,

47.33766131]

Vector 2 : [104.96183434 , 108.73740434 ,107.17293128 , 49.53197163 , 45.98090932,

44.47265154]

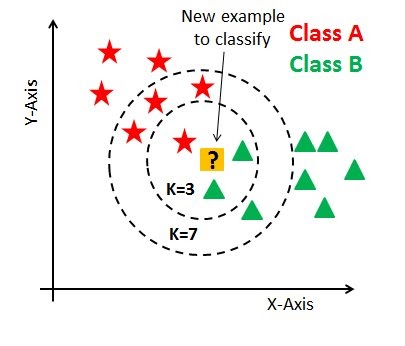
Khi đó khoảng cách Euclidean là

d(vector1,vector2) = = 6.120

Dựa vào cách tính khoảng cách Euclidean ta sử dụng thuật toán KNN(K-Nearest Neighbors) để nhận dạng phân loại ảnh đầu vào :

### 3.7.2 Giới thiệu thuật toán KNN(K-Nearest Neighbors)

KNN là thuật toán học giám sát. KNN cho rằng những dữ liệu tương tự nhau sẽ tồn tại gần nhau trong một không gian, từ đó công việc của chúng ta là sẽ tìm k điểm gần với dữ liệu cần kiểm tra nhất. Cách chọn k sẽ ảnh hưởng trực tiếp tới kết quả bài toán.Việc tìm khoảng cách giữa 2 điểm cũng có nhiều công thức có thể sử dụng, tùy trường hợp mà chúng ta lựa chọn cho phù hợp trong bài toán này ta chọn cách tính theo Công thức euclidean :



* + Ưu điểm của KNN

1. Độ phức tạp tính toán của quá trình training là bằng 0.
2. Việc dự đoán kết quả của dữ liệu mới rất đơn giản.
3. Không cần giả sử gì về phân phối của các class.
   * Nhược điểm của KNN

* KNN rất nhạy cảm với nhiễu khi K nhỏ.
* Như đã nói, KNN là một thuật toán mà mọi tính toán đều nằm ở khâu test. Trong đó việc tính khoảng cách tới *từng* điểm dữ liệu trong training set sẽ tốn rất nhiều thời gian, đặc biệt là với các cơ sở dữ liệu có số chiều lớn và có nhiều điểm dữ liệu. Với K càng lớn thì độ phức tạp cũng sẽ tăng lên. Ngoài ra, việc lưu toàn bộ dữ liệu trong bộ nhớ cũng ảnh hưởng tới hiệu năng của KNN.
  + Các bước cài đặt thuật toán :
  + Chia tập vector đặc trưng thành 2 phần training và phần Test

Phần training : để làm tập dữ liệu cho vector ảnh nhận dạng so khớp

Phần test : để thử dự đoán kết quả các ảnh đã biết nhằm đánh giá độ chính xác của thuật toán để điều chỉnh các tham số phù hợp

* + Chọn tham số K phù hợp để tìm k điểm sao cho khoảng cách gần với vector ảnh đầu vào nhất
  + Tiến hành dự đoán và trả ra kết quả

Code cài đặt chi tiết sẽ được trình bày vào chương 4

# Chương 4 : Cài đặt chương trình và đánh giá kết quả

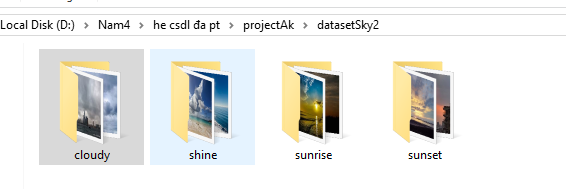
# 4.1 Cài đặt

## Link DataSet

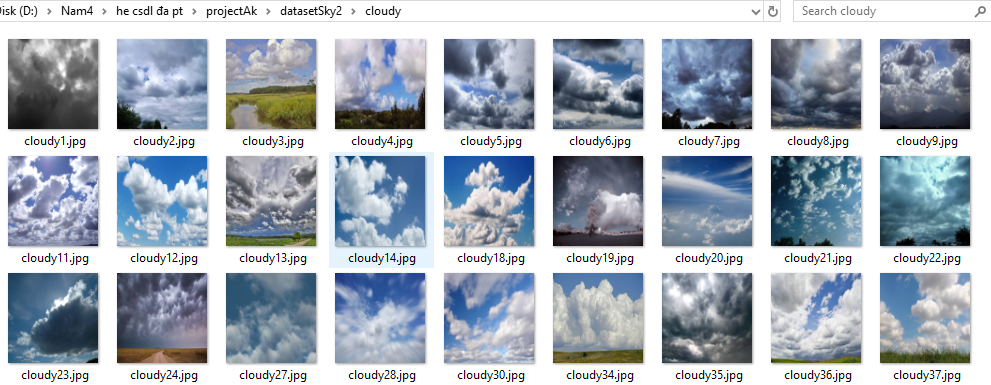
Bộ dữ liệu được tải tại : <https://www.kaggle.com/vijaygiitk/multiclass-weather-dataset>

Trong đó được chia thành 4 loại : Ảnh âm u(285 ảnh), ảnh buổi trưa(252 ảnh), Ảnh bình minh(255 ảnh),Ảnh hoàng hôn(214 ảnh)

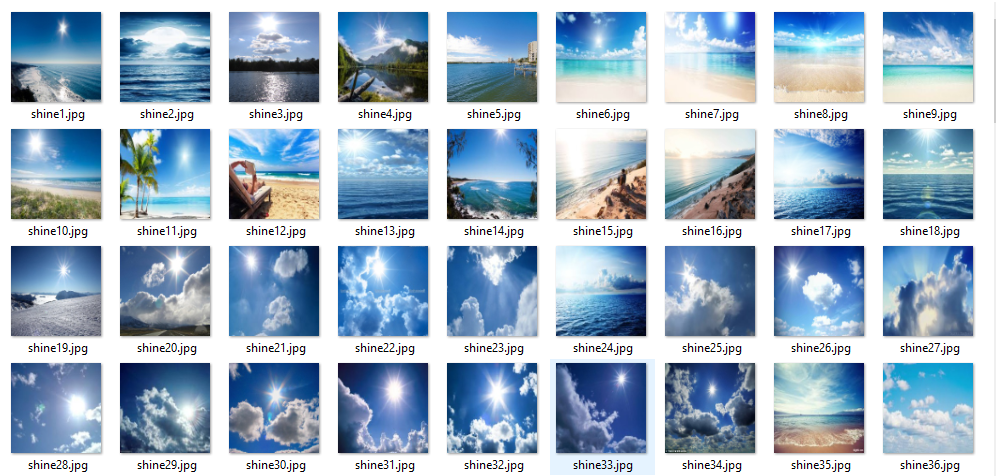
Tất cả các ảnh được chuyển về định dạng đuôi JPG , hệ màu RGB,kích thước 224x224 pixel được sắp xếp vào 4 thư mục tương đương với 4 nhãn :



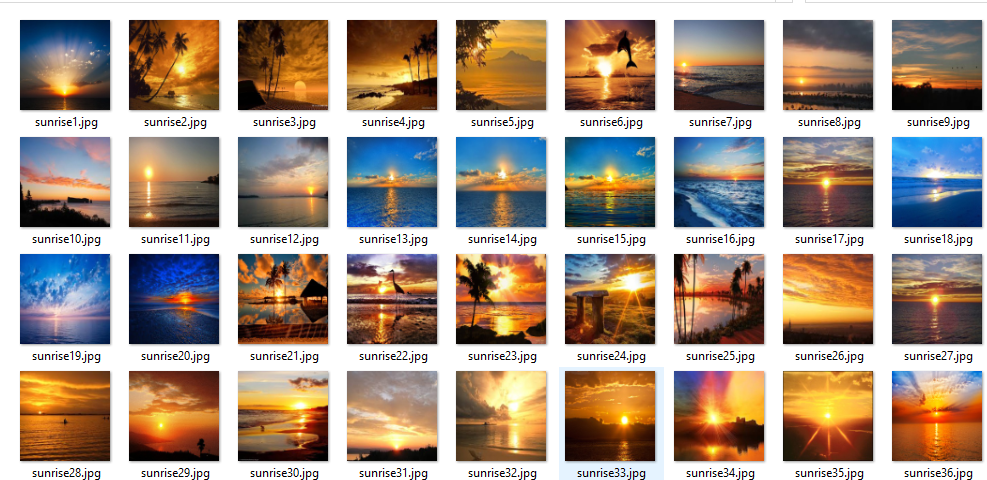
Một vài hình ảnh trong tập cloudy( âm u,nhiều mây)



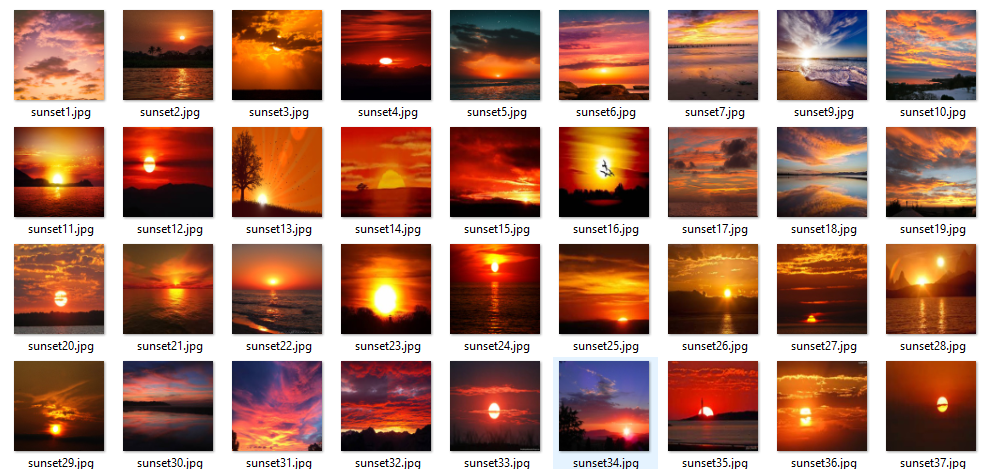
Một vài hình ảnh trong tập Shine( trời buổi trưa-nắng)



Một vài hình ảnh trong tập Sunrise(bình minh)



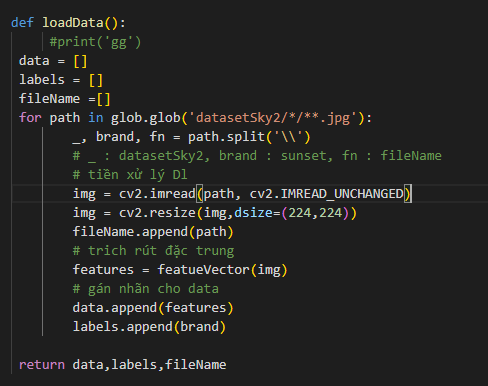
Một vài hình ảnh trong tập Sunset(Hoàng hôn)



## Đọc và trích xuất dữ liệu

Đầu tiên duyệt tất cả ảnh trong thư mục:

Code minh họa

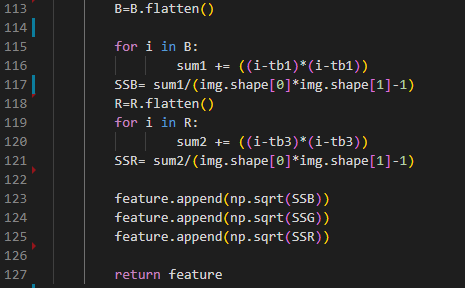


Trong khi duyệt ảnh ta đọc và resize ảnh về cùng kích thước 224x224 pixel

Hàm featureVector dùng trích rút đặc trưng của ảnh

Code minh họa





Dòng 92 : Tách ảnh thành 3 kênh màu độc lập B,G,R

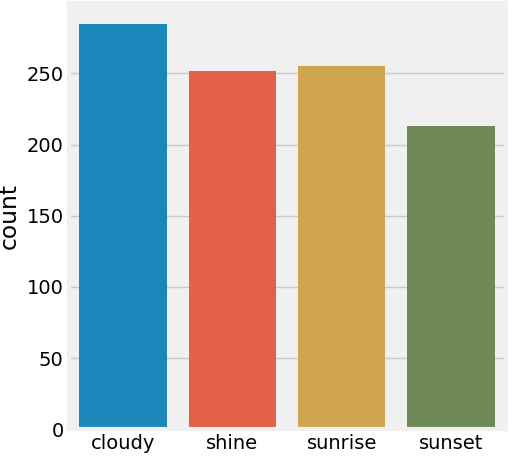
Dòng 94 – 96 : Tính toán giá trị trung bình quân màu của 3 kênh vừa tách

Dòng 98 – 101 : đưa 3 giá trị trung bình vừa tính vào cùng 1 mảng

Dòng 106 – 121 : Tính độ lệch chuẩn từng kênh màu

Dòng 123 – 127 : Đưa 3 giá trị độ lệch chuẩn vào mảng chứa 3 giá trị trung bình quân màu từng kênh và trả ra kết quả

Kết quả thu được sau khi đọc và trích xuất đặc trưng ảnh



Trong đó gồm : 285 đặc trưng ảnh cloudy ( âm u)

: 252 đặc trưng ảnh shine (trời nắng buổi trưa)

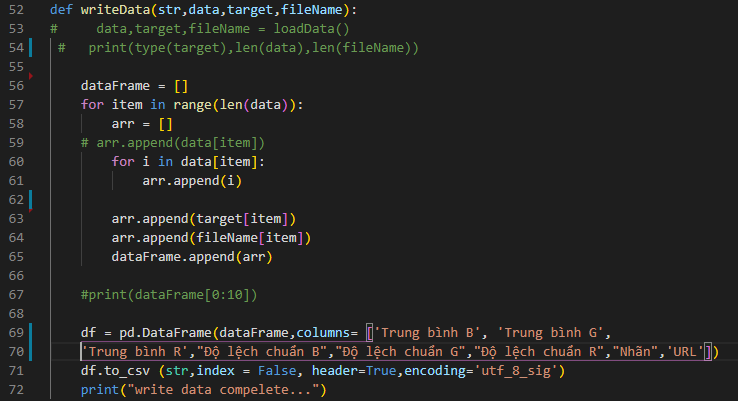
: 255 đặc trưng ảnh sunrise(bình minh)

: 214 đặc trưng ảnh sunset(hoàng hôn)

## Lưu trữ đặc trưng

Sau khi có tập dữ liệu đặc trưng ta tiến hành lưu đặc trưng dưới dạng file csv

Code minh họa



Với tham số đầu vào là 1 tập đặc trưng ,nhãn và đường dẫn ảnh tương ứng ,sử dụng thư viện pandas để lưu file

Kết quả file csv sau khi lưu



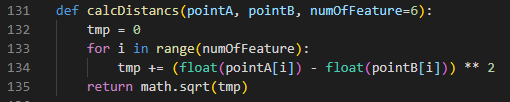
File gồm 1005 dòng không tính dòng header tương ứng với 1005 ảnh với đặc trưng được trích rút tương ứng

Cột thứ 7 là nhãn , cột thứ 8 là đường dẫn ảnh tương ứng khi muốn hiển thị ảnh

## Cài đặt Thuật toán KNN

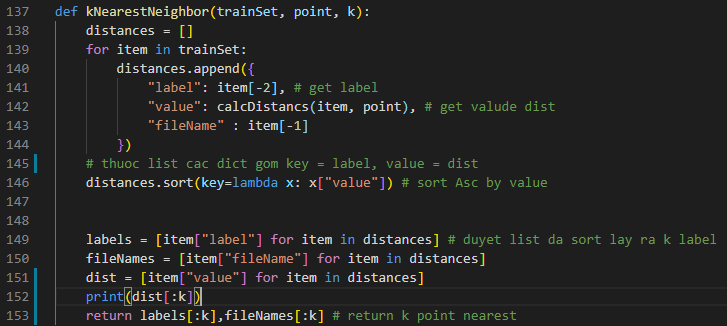
Sau khi lưu đặc trưng ta tiến thành cài đặt thuật toán KNN

* Hàm tính khoảng cách Euclidean của 2 đặc trưng



Trong đó numOfFeature = 6 thể hiện 1 vector đặc trung có 6 đặc trưng : trung bình màu Blue, Green,Red , Độ lệch chuẩn Blue,Green,Red

* Hàm tính toán khoảng cách của 1 vector đặc trưng tới tập dữ liệu Vector đã lưu



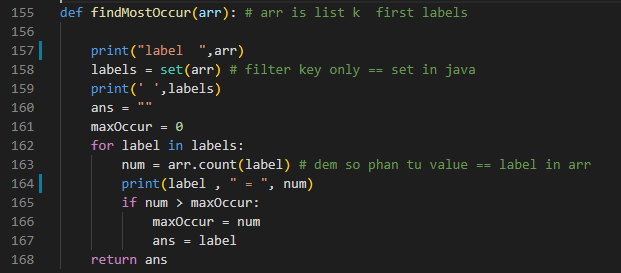
Đầu vào là tập trainSet được lấy từ file Csv

Dòng 138- 144 : Duyệt tập trainSet để lấy nhãn,đường dẫn ảnh ,tính khoảng cách tới point ( là vector đặc trưng cần xét) bằng hàm khoảng cách đã định nghĩa trước đó rồi lưu vào thành 1 list

Dòng 146 :Sau đó sắp xếp list theo khoảng cách tăng dần

Dòng 149 – 153 :Lọc k phần tử đầu tiên của danh sách tưng ứng là k hàng xóm có khoảng cách ngắn nhất với vector đang xét và trả ra kết quả

* Tiếp theo tạo hàm tính số lần xuất hiện nhiều nhất của nhãn trong k nhãn



Đầu vào là list k nhãn gần nhất đã lọc

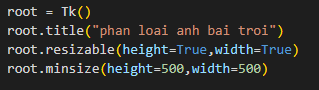
Dòng 158 : Lọc các phần tử xuất hiện trong list

Dòng 160 – 168 : duyệt list để tính số lần xuất hiện nhiều nhất của nhãn và trả ra kết quả

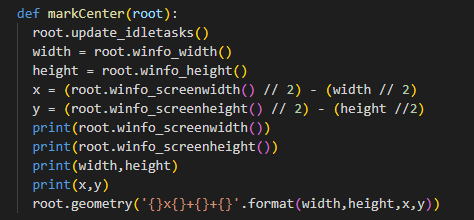
## Cài đặt giao diện

Chương trình sử dụng gói giao diện Tkiner

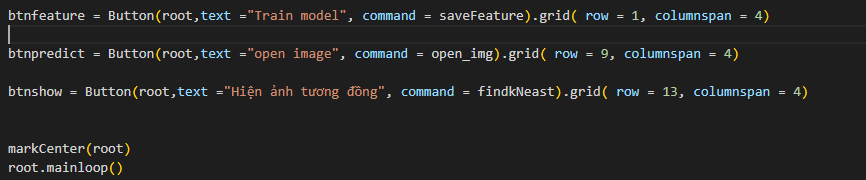
Khởi tạo giao diện



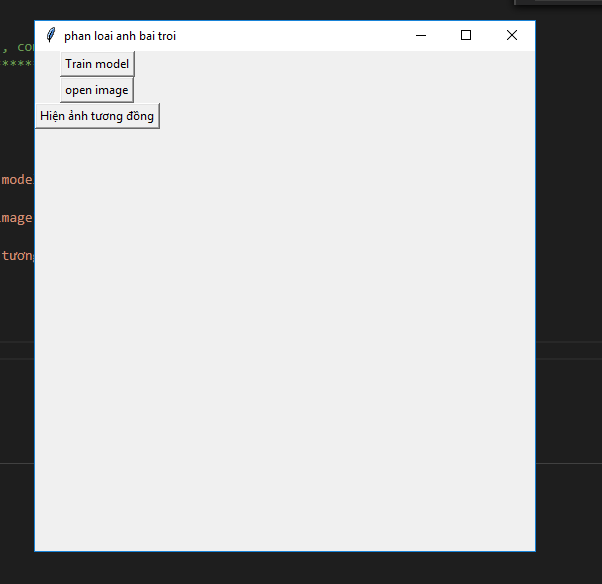
Đưa frame ra giữa màn hình máy tính



Thêm 3 button chính cho giao diện và chạy Gd



Giao diện khi chạy lên



Ở đây chúng em chủ yếu tập trung vào phần logic xử lí trích xuất lưu trữ đặc trưng và thuật toán nhận dạng ,chưa chú trong về UX,UI.

**Button 1 : TrainModel**

btnfeature = Button(root,text ="Train model", command = saveFeature).grid( row = 1, columnspan = 4)

Khi click button1 hàm SaveFeatue được chạy

def saveFeature():

        data,target,fileName = loadData()

        choose =  mbox.askquestion("Question" , "Lưu đặc trưng ??")

        if choose == 'yes':

                nameFile =  simpledialog.askstring(title="Title",prompt="Nhap ten file ")

                if nameFile is not None and nameFile != ''  :

                        nameFile = nameFile+".csv"

                        writeData(nameFile,data,target,fileName)

                        fit(nameFile)

                       # mbox.showinfo( "Message", "Thành công!!!")

                else:

                        mbox.showerror( "Message", "Nhap lai ten file")

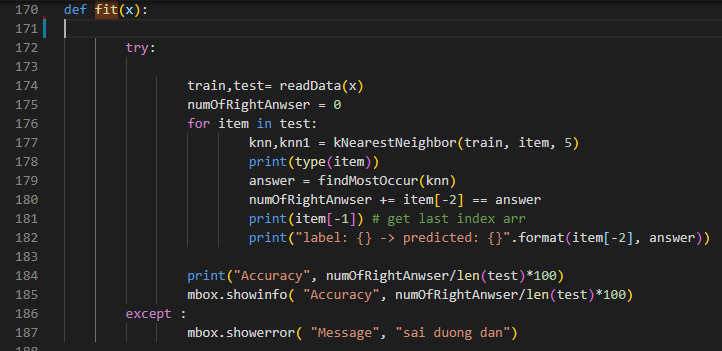
Đầu tiên ta gọt hàm loadData đã định nghĩa từ trước để lấy đặc trưng ,nhãn cùng đường dẫn ảnh tương ứng .

Sau đó yêu cầu người dùng nhập tên file csv muốn lưu chương trình gọi hàm writeData đã định nghĩa từ trước

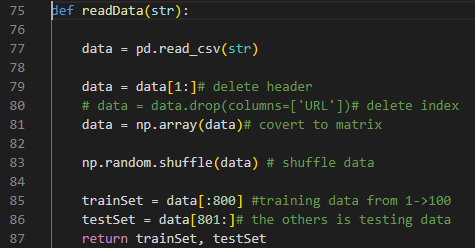
Tiếp theo sau khi lưu chúng ta sẽ đọc dữ liệu đã lưu để tiến hành test thuật toán KNN thông qua hàm **fit()**

Định nghĩa hàm fit như sau :

Đầu vào là đường dẫn tới file csv



Hàm **readData**() có trách nghiệm lấy dữ liệu từ file csv đã lưu chia thành 80% cho train và 20 % cho phần Test thử nghiệm



Dòng 77 đọc dữ liệu từ file csv :

Dòng 79 : Loại bỏ header

Dòng 81 Đưa kiểu dữ liệu từ DataFrame về dạng array

Dòng 83 : Trộn ngẫu nhiên dữ liệu

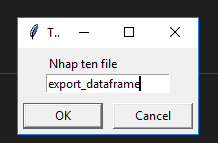
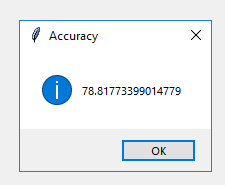
Dòng 85-87 : chia dữ liệu ra tập train và test sau đó trả ra kết quả

Sau khi có dữ liệu tại hàm fit ta duyệt tập test để bắt đầu qua s trình test thuật toán

* + Gọi hàm **kNearestNeighbor**() đã định nghĩa để lấy k=5 phần tử gần nhất
  + Hàm **findMostOccur**() đã định nghĩa lấy ra nhãn có số lần xuất hiện lớn nhất
  + Sau đó so với nhãn thực của tập Test
  + numOfRightAnwser là biến tính số nhãn đoán đúng
  + Kết thúc dự đoán ta lấy numOfRightAnwser / tổng số lượng tập test để ra độ chính xác

Kết quả chạy chương trình khi ấn button 1 :



Độ chính xác là 78.81 , độ chính xác này có thể theo đổi theo lần chạy tùy thuộc vào tập dữ liệu train,test .Sở dĩ có sự thay đổi Accuracy là do khi đọc dữ liệu có thêm hàm **np.random.shuffle(data)** nhằm tránh tình trạng luôn luôn lấy 20% dữ liệu cuối để test khi đó dl cuối nếu ko được xáo trộn sẽ không có nhãn cloudy ,shine làm mất tính tổng quát.

**Button 2: Open Image**

btnpredict = Button(root,text ="open image", command = open\_img).grid( row = 9, columnspan = 4)

Khi click button này sẽ gọi hàm open\_img để nhận dạng ảnh mới

Định nghĩa hàm open\_img

def open\_img():

    # Select path  img

    global path\_img\_predict

    x = openfilename()

    path\_img\_predict = x

    print('x= ',x)

    # opens the image

    img = Image.open(x)

    print('img',img)

    # resize the image and apply a high-quality down sampling filter

    img = img.resize((250, 250), Image.ANTIALIAS)

    # PhotoImage class is used to add image to widgets, icons etc

    img = ImageTk.PhotoImage(img)

    # create a label

    panel = Label(root, image = img)

    # set the image as img

    panel.image = img

    panel.grid(row = 2)

    # doc anh

    image = mpimg.imread(x)

    print('Moi nhan dạng')

    train,test= readData("export\_dataframe.csv")

    arrData = []

    arrData = np.array(train)

    arrData = np.append(arrData,test,axis = 0)

    print("\*"\*30)

    print(arrData.shape)

    str,img\_Like = predict(image,arrData)

    res = Label(root,text=" Ảnh  -- "+ str  ).grid( row = 4, columnspan = 4)

Đầu tiên gọi hàm openfilename() để lấy đường dẫn ảnh

Sau đó gọi hàm readData đã định nghĩa để lấy tập đặc trưng

Sau đó tiến hành gọi hàm predict để dự đoán

Định nghĩa hàm predict

def predict(img,train):

#   img = cv2.imread(anh,cv2.IMREAD\_UNCHANGED)

    global anh\_like

    img = preImg(img)

    featue = featueVector(img)

    print(" Dac trung dau vao ",featue)

    knn,knn1 = kNearestNeighbor(train,featue,5)

    answer = findMostOccur(knn)

    print(answer,knn1)

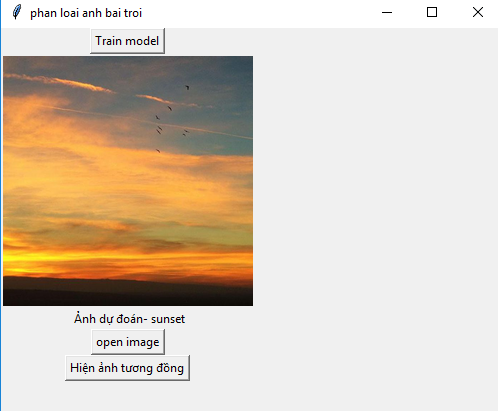
    anh\_like = knn1[0]

    return answer,knn1[0]

Hàm sẽ trích rút đặc trưng theo hàm featureVector đã định nghĩa và gọi hàm kNearestNeighbor() lọc ra 5 vector có độ tương đồng nhất

Sau đó gọi hàm findMostOccur() để lấy nhãn xuất hiện nhiều nhất

Kết quả khi chạy giao diện



**Button 3 : Lấy ra ảnh có độ tương đồng nhất**

btnshow = Button(root,text ="Hiện ảnh tương đồng", command = findkNeast).grid( row = 13, columnspan = 4)

Đây là phần chúng em tự thêm so với yêu cầu bài toán của thầy. Để hiển thị ảnh tương đồng 1 cách trực quan hơn

Khi click button 3 hàm findNeast() được gọi

def findkNeast():

        gray = plt.imread(anh\_like, cv2.IMREAD\_UNCHANGED)

        import matplotlib.image as mpimg

        plt.figure(figsize = (12, 4)) # chia khoảng cho 2 ảnh

        plt.subplot(1, 2, 1) # tạo 2 ô gồm 1 dòng 2 cột ảnh 1 ở cột 1

        img = mpimg.imread(path\_img\_predict)

        imgplot = plt.imshow(img)

        plt.title(" anh du doan : ")

        plt.subplot(1, 2, 2)

        gray = mpimg.imread(anh\_like)

        plt.imshow(gray)

        plt.title("anh gan giong")

        plt.show()

Khi click open\_image để nhận dạng ảnh mới ta gọi thêm hàm findNeast() để truyền đường dẫn ảnh tương đồng nhất so với ảnh đang xét

Giao diện khi chạy:



# 4.2 Đánh giá kết quả :

- Chương trình có độ chính xác khá tốt trong khoảng 80% , nhận diện tốt hầu hết các ảnh Trời âm u và ảnh buổi trưa , tuy vậy vẫn còn sự nhầm lẫn giữa các ảnh

- Việc trích xuất đặc trưng mất khá nhiều thời gian

## 4.2.1 Hạn chế

- Do chỉ dựa vào màu săc để phân biệt nên chương trình đôi khi bị nhầm lẫn giữa các ảnh nếu như ảnh chứa khung cảnh lạ

- Chương trình vẫn đang trong quá trình kiểm nghiệm , chưa thể áp dụng vào thực tế do thời gian chạy còn chậm , thuật toán chưa thực sự tối ưu , đặc biệt trong quá trình trích rút đặc trưng .

-Quá trình trích rút đặc trưng chạy khá chậm làm giảm hiệu năng hệ thống.

## 4.2.2 Hướng phát triển

- Trong tương lai ,cần kết hợp nhiều thuộc tính để nhận dạng, có thể sử dụng 1 số kĩ thuật mới trong việc trích rút đặc trưng.

- Kết hợp các kĩ thuật phân đoạn, phân vùng ảnh ,để nhận biết đâu là ảnh bầu trời giúp lọc bỏ nhiễu không cần thiết.

- Sử dụng các thuật toán cao cấp hơn: SVM, Random forest,… các kĩ thuật học sâu để nhận dạng.

- Cải thiện thuật toán đặc biệt là quá trình trích rút đặc trưng để giảm thời gian chạy chương trình xuống.

# TÀI LIỆU THAM KHẢO

[1] Giáo Trình CSDL Đa Phương Tiện : Học Viện Công Nghệ Bưu Chính Viễn Thông – Tác giả : Đỗ Trung Tuấn, Nguyễn Đình Hóa

[2] Giáo Trình Xử lý ảnh : Học Viện Công Nghệ Bưu Chính Viễn Thông – Tác giả : PGS. TS. ĐỖ NĂNG TOÀN

[3] COLOR CHANNEL STATISTICS : <https://codelungtung.wordpress.com/2018/06/10/color-channel-statistics-and-color-histograms/>

[4] Thuật toán K-nearest neighbors : <https://machinelearningcoban.com/2017/01/08/knn/>

[5] Trích chọn đặc trưng theo màu sắc : <https://viblo.asia/p/trich-chon-dac-trung-va-nhan-dang-bang-mau-sac-gDVK2LbXlLj>

[6] Image Search Engine : <https://codelungtung.wordpress.com/2018/08/17/image-search-engine-4-buoc-de-xay-dung-mot-image-search-engine/>

[7] Euclidean distance : <https://nkvtruong.wordpress.com/2011/11/03/khai-niem-khoang-cach-trong-thong-ke/>

{\displaystyle \mathrm {d} (\mathbf {p} ,\mathbf {q} )={\sqrt {(q\_{1}-p\_{1})^{2}+(q\_{2}-p\_{2})^{2}+\cdots +(q\_{n}-p\_{n})^{2}}}={\sqrt {\sum \_{i=1}^{n}(q\_{i}-p\_{i})^{2}}}.}