

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

**KHOA CÔNG NGHỆ THÔNG TIN**

-----------------------

**BÁO CÁO**

**BTL môn hệ cơ sở dữ liệu đa phương tiện**

**Đề tài: Xây dựng hệ CSDL lưu trữ và tìm kiếm ảnh thiên nhiên**

|  |  |  |
| --- | --- | --- |
| **Giảng viên:** | **Thầy Nguyễn Đình Hóa** |  |
| **Họ và tên:** | **Nguyễn Văn Hùng** | **B21DCCN417** |
|  |  |  |
|  |  |  |
| **Số điện thoại:** | **0363524788** |  |



**Yêu Cầu**

Xây dựng hệ CSDL lưu trữ và tìm kiếm ảnh thiên nhiên.

1.Hãy xây dựng/sưu tầm một bộ dữ liệu ảnh gồm ít nhất 100 files ảnh về cảnh thiên nhiên khác nhau về cùng một chủ đề, các ảnh có cùng kích thước, mỗi ảnh có một khung hình đơn giản về một cảnh thiên nhiên với cùng độ phân giải, cùng tỉ lệ khung hình của các vật trong ảnh (SV tùy chọn định dạng ảnh).

2.Hãy xây dựng một bộ thuộc tính để nhận diện ảnh thiên nhiên từ bộ dữ liệu đã thu thập. Hãy trình bày lý do lựa chọn và giá trị thông tin của các thuộc tính được sử dụng trong bài.

3.Xây dựng hệ thống tìm kiếm ảnh thiên nhiên với đầu vào là một ảnh mới về một cảnh thiên nhiên nào đó đã có và không có trong dữ liệu, đầu ra là 3 ảnh giống nhất, xếp thứ tự giảm dần về độ tương đồng nội dung với ảnh đầu vào.

a.Trình bày sơ đồ khối của hệ thống và quy trình thực hiện yêu cầu của đề bài.

b.Trình bày quá trình trích rút, lưu trữ và sử dụng các thuộc tính để tìm kiếm ảnh trong hệ thống.

4.Demo hệ thống và đánh giá kết quả đã đạt được.

**1. Đặc điểm của kho ảnh**

- Có khoảng 130 ảnh. Tro ng đó là 100 ảnh đã lọc theo đặc điểm và 30 ảnh là nhiều chi tiết hơn.

- Gồm các hình ảnh về biển ở một số địa điểm khác nhau trên thế giới.

- Mỗi ảnh có thể có bờ biển, biển, núi, cây, bầu trời và đa phần các ảnh sẽ theo góc nhìn bên phải của người chụp.

- Kích thước 1:1.

- Có ảnh chụp gần, có ảnh chụp xa.

1.1. Đặc điểm về màu sắc (Color Features)

- Tông xanh dương (Hue ~ 100–140): Biển trong, nước biển.

- Tông vàng – nâu (Hue ~ 20–30): Bãi cát, nắng.

- Đỏ, cam (hoàng hôn): Một số ảnh chụp lúc bình minh hoặc chiều tà.

1.2. Đặc điểm về kết cấu (Texture Features)

Ảnh có nhiều biên – nhiều hướng gradient:

→ Vector HOG có nhiều vùng biên nổi bật

→ Ví dụ: sóng biển, cây rừng, đá, tàu thuyền

Ảnh phẳng, ít thay đổi hướng sáng:

→ Vector HOG có giá trị thấp, gần đồng đều

→ Ví dụ: mặt biển phẳng, bầu trời xanh không mây

**2. Xây dựng bộ thuộc tính**

Để xây dựng một bộ thuộc tính (feature set) nhận diện ảnh thiên nhiên, đặc biệt là từ bộ dữ liệu ảnh biển và cảnh quan thiên nhiên, bạn nên lựa chọn các đặc trưng có khả năng mô tả tốt màu sắc, kết cấu, đường nét và bố cục, vì ảnh thiên nhiên thường chứa các yếu tố phong phú như:

+ Bầu trời, nước, cây, cát, đá, mặt trời...

+ Kết cấu sóng, sương mù, đồi núi, lá cây...

**Bộ thuộc tính đề xuất để nhận diện ảnh thiên nhiên và lý do lựa chọn**

1.1. Đặc trưng màu sắc – HSV Histogram (192 chiều)

Kỹ thuật: Histogram theo từng kênh Hue, Saturation, Value với 64 bins/kênh.

Lý do lựa chọn:

- Màu sắc là đặc trưng nổi bật để nhận diện cảnh thiên nhiên: bầu trời, nước, cát, rừng.

- HSV ổn định hơn RGB khi thay đổi ánh sáng hoặc độ tương phản.

- Phân tách màu sắc hiệu quả:

+ Biển: màu xanh lam (Hue ~ 100–140)

+ Cát: màu vàng nhạt

+ Rừng, núi: màu xanh lá – nâu

- Giá trị thông tin:

+ Cho phép phân biệt rõ các loại cảnh thiên nhiên: biển, rừng, núi, hoàng hôn...

+ Là cơ sở chính để nhóm ảnh theo tone màu chủ đạo.

1.2. Đặc trưng kết cấu – HOG (Histogram of Oriented Gradients) (~8100 chiều)

Kỹ thuật: HOG với orientation = 9, pixels\_per\_cell = (8,8), cells\_per\_block=(2,2), L2-Hys.

Lý do lựa chọn:

- HOG mạnh trong mô tả hướng cạnh, biên, cấu trúc hình học trong ảnh.

- Đặc biệt phù hợp với ảnh thiên nhiên có cấu trúc như:

+ Sóng biển

+ Đồi núi, rừng

+ Tàu thuyền, mặt trời, bầu trời có mây

- Giá trị thông tin:

+ Giúp phân biệt ảnh có nhiều chi tiết, cạnh mạnh (ví dụ: sóng, rừng rậm) và ảnh phẳng, mịn (ví dụ: biển phẳng, trời quang).

+ Phân biệt tốt các vùng có biên định hướng rõ so với nền phẳng hoặc ít biến động.

1.3. Đặc trưng cạnh – Canny edge histogram (64 chiều)

Kỹ thuật: Dùng bộ lọc biên Canny → tính histogram độ sáng của biên.

Lý do lựa chọn:

- Ảnh thiên nhiên có thể có các biên mạnh (núi, sóng) hoặc vùng ít biên (bầu trời, mặt biển).

- Canny giúp nhận diện vùng chuyển sáng rõ nét.

- Giá trị thông tin:

+ Bổ sung cho HOG, giúp xác định các ảnh có biên sắc nét/tương phản.

+ Giúp đánh giá mức độ "chi tiết" trong ảnh.