# Description of Content : Image Processing (Mô tả nội dung : Xử lý ảnh)

Điều quan trọng là thiết lập content-based không dựa vào việc mô tả toàn bộ nội dung của hình ảnh. Nó có thể chỉ đủ cho một hệ thống truy xuất trình bày hình ảnh tương tự, theo một ý nghĩa của user-defined. Nội dung mô tả nên phục vụ vào mục tiêu chính.

Chúng ta xem xét nội dung mô tả trên 2 bước. Đầu tiên, chúng ta cần chuyển image-processing sang dữ liệu mảng ở không gian khác, xem Lig.6. Chúng ta chia các phương thức theo màu cục bộ, kết cấu cục bộ hoặc hình học cục bộ.

Công thức đặc trưng chung bởi:



ở đây, i(x) là hình ảnh, không gian hình ảnh I, g là một toán tử trên hình ảnh, f(x) là phương trình kết quả cần tìm. Các thông số tính toán của g có thể bao gồm kích thước của vùng lân cận xung quanh x để tính f(x) hoặc là một tiêu chí đồng nhất khi kích thước của từng phần để tính f(x) phụ thuộc vào dữ liệu thực tế, như trong [16S], [126], chẳng hạn.

Vì vậy, mục đích của xử lý hình ảnh trong truy xuất hình ảnh là để tăng cường các khía cạnh trong dữ liệu hình ảnh liên quan đến truy vấn và để giảm các khía cạnh còn lại.

Một mục tiêu như vậy có thể được đáp ứng bằng cách sử dụng bất biến như một công cụ để đối phó với các biến dạng ngẫu nhiên trong thông tin được intro-duced bởi khoảng cách cảm giác. Từ các cuộc thảo luận ở trên về khoảng cách cảm giác, rõ ràng là các tính năng bất biến có thể mang nhiều thông tin cụ thể hơn đối tượng so với các tính năng khác vì chúng không nhạy cảm với các điều kiện tình cờ của cảm biến. Mục đích của các mô tả bất biến là xác định các đối tượng, bất kể chúng được quan sát như thế nào và ở đâu, khi mất một số nội dung thông tin. Nếu hai đối tượng ti (hoặc hai lần xuất hiện của cùng một đối tượng) là tương đương dưới một nhóm các phép biến đổi W, chúng nằm trong một lớp tương đương [18]:



Một thuộc tính ƒ của t là bất biến theo W khi và chỉ khi ƒt giữ nguyên như bất kể điều kiện không mong muốn được biểu thị bởi W,



Mức độ bất biến, nghĩa là tính chiều của nhóm W, nên được điều chỉnh theo bản ghi circum-stances. Nói chung, một tính năng với một lớp bất biến rộng sẽ mất khả năng phân biệt giữa các khác biệt cơ bản. Kích thước của lớp hình ảnh được coi là tương đương tăng theo chiều của W. Cuối cùng, tính bất biến có thể rộng đến mức không có sự phân biệt đối xử trong số các đối tượng được giữ lại. Mục đích là để chọn chặt chẽ nhất tập hợp các bất biến phù hợp với tập hợp các điều kiện không quan trọng dự kiến. Điều cần thiết trong tìm kiếm hình ảnh là một đặc điểm của các điều kiện bất biến tối thiểu trong đặc tả của truy vấn được thảo luận trong [159]. Tập hợp tối thiểu các điều kiện bất biến chỉ có thể được chỉ định bởi người dùng vì nó là một phần trong ý định của người đó. Công trình lâu đời nhất về tính bất biến trong tầm nhìn máy tính đã được thực hiện trong nhận dạng đối tượng, như đã báo cáo, trong số những người khác, trong [117]. Mô tả bất biến trong truy xuất hình ảnh là tương đối mới, nhưng nhanh chóng đạt được mục đích giới thiệu tốt, xem [18], [32]. Một thay thế cho các tính năng bất biến là thể hiện các điều kiện xem riêng biệt với các đối tượng trong cảnh. Bằng cách này, không có thông tin nào bị mất trong việc giảm các tính năng bất biến, trong khi thông tin chỉ được sắp xếp lại. Nó nên được để lại cho các giai đoạn sau để quyết định điều gì là quan trọng.

## Color Image Processing (Xử lý màu sắc ảnh)

Màu sắc là một lĩnh vực nghiên cứu tích cực trong phục hồi hình ảnh, nhiều hơn bất kỳ nhánh nào khác của thị giác máy tính. Màu làm cho hình ảnh i (x) nhận các giá trị trong không gian vectơ màu. Sự quan tâm đến màu sắc có thể được gán cho cấp trên khả năng phân biệt đối xử của miền ba chiều so với miền một chiều của hình ảnh mức xám.

Hai khía cạnh của trả về màu sắc trong nhiều đóng góp. Một là màu sắc được ghi lại thay đổi đáng kể theo hướng của bề mặt, góc nhìn của máy ảnh, vị trí chiếu sáng, quang phổ của đèn chiếu sáng và cách ánh sáng tương tác với vật thể. Sự thay đổi này nên được xử lý theo cách này hay cách khác. Thứ hai, nhận thức của con người về màu sắc là một chủ đề phức tạp, trong đó nhiều nỗ lực đã được thực hiện để nắm bắt sự tương đồng thường xuyên.

Chỉ khi không có biến thể trong bản ghi hoặc trong nhận thức thì đại diện màu RGB là một lựa chọn tốt vì đại diện đó được thiết kế để phù hợp với kênh đầu vào của mắt. RGB-representations đang được sử dụng rộng rãi. Ta mô tả hình ảnh trong các thuộc tính màu sắc của nó. Một hình ảnh được thể hiện dưới dạng (R (x), G (x), B (x)) (các chỉ số sẽ được bỏ qua từ bây giờ) có ý nghĩa nhất khi ghi trong trường hợp không có phương sai, ví dụ như đối với các bức tranh nghệ thuật [64], thành phần màu của ảnh [48] và nhãn hiệu [79], [S9], trong đó hình ảnh hai chiều được ghi ở chế độ xem trước trong điều kiện tiêu chuẩn.

Một cải tiến đáng kể so với RGB-color space (ít nhất là cho các ứng dụng truy xuất) xuất phát từ việc sử dụng các biểu diễn màu của đối thủ [169], sử dụng các trục màu của đối thủ (R - G, 2B - R - G, R + G + B). Đại diện này có lợi thế là cách ly thông tin độ sáng trên trục thứ ba. Với giải pháp này, hai trục màu đầu tiên có thể được lấy mẫu xuống vì con người nhạy cảm với độ sáng hơn so với sắc độ. Chúng là bất biến đối với những thay đổi về cường độ chiếu sáng và bóng tối.

Các phương pháp khác sử dụng Munsell hoặc Lab-space vì tính đồng nhất về nhận thức tương đối của chúng. Lab- representation được thiết kế sao cho khoảng cách Euclide giữa hai biểu diễn màu mô hình nhận thức của con người về sự khác biệt màu sắc. Cần thận trọng khi số hóa chuyển đổi phi tuyến sang Lab-space [115].

The HSV-representation thường được chọn cho các thuộc tính bất biến của nó. Màu sắc là bất biến dưới sự định hướng của đối tượng đối với hướng chiếu sáng và hướng camera và do đó phù hợp hơn cho việc thu hồi đối tượng.

Một loạt các bất biến màu trắc quang chặt chẽ để truy xuất đối tượng được lấy từ [57] từ một phân tích mô hình Schafer về sự phản xạ của đối tượng. Chúng lấy được các mảng mờ dưới ánh sáng trắng không gian màu bất biến



chỉ phụ thuộc vào cảm biến và suất phản chiếu bề mặt. Cho bóng chiếu sáng bề mặt và trắng, chúng lấy được đại diện bất biến như



và hai hoán vị nữa. Các mô hình màu sắc mạnh mẽ chống lại sự biến dạng quan điểm chính.

Màu sắc không đổi là khả năng con người nhận biết cùng một màu rõ ràng với sự có mặt của các biến thể chiếu sáng làm thay đổi quang phổ vật lý của ánh sáng cảm nhận. Trong thị giác máy tính, màu sắc không đổi được xem xét lần đầu trong [49]. Đối với các vật thể phẳng, mờ và được chiếu sáng đồng đều, mặt phẳng tạo thành gam màu chính tắc được định nghĩa là tập hợp của các phản ứng RGB bình thường hóa khả thi về mặt vật lý, tức là, rgb-responses. Tham chiếu sau đó ánh xạ tất cả các rgb-responses quan sát trong hình ảnh thành chính tắc. Bản đồ giải thích tất cả các quan sát xác định hằng số màu. Trong [47], điều này được cải thiện để bao gồm phản xạ gương, hình dạng và độ chiếu sáng khác nhau. Bằng cách tính toán blue-ratio vector (v , g , 1), chỉ sử dụng hướng màu và cường độ được loại trừ. Trong 2D-space(không gian 2 chiều) này, màu sắc constancy map một lần nữa có thể được chọn từ một gam màu chính và màu phụ. Trong [56], các tỷ lệ



cung cấp sự ổn định hơn cho các biến thể hình học bề mặt. Màu sắc không đổi được áp dụng để truy xuất trong [54] bằng cách sử dụng biểu diễn màu bất biến chiếu sáng. Các chỉ số tỷ lệ của các màu lân cận. Lập chỉ mục màu liên tục dẫn đến một số mất mát trong khả năng phân biệt giữa các đối tượng, nhưng thay vào đó mang lại sự phục hồi độc lập chiếu sáng. Đề án đã được cải thiện trong [158] bằng cách sử dụng các bất biến đại số.

Thay vì mô tả bất biến, một cách tiếp cận khác để đối phó với sự bất bình đẳng trong quan sát do sự phản xạ bề mặt là tìm kiếm các cụm trong biểu đồ màu của hình ảnh. Trong RGB-histogram (biểu đồ RGB), các cụm pixel phản chiếu ra một vệt dài hình đối tượng. Do đó, trong [126], thuật toán cụm không theo tỷ lệ trong RGB-space được sử dụng để xác định các pixel nào trong ảnh bắt nguồn từ một đối tượng có màu đồng nhất.

## Image Processing for Local Shape (Xử lý ảnh cho hình dạng )

Dưới tên “local shape,” chúng tôi thu thập tất cả các thuộc tính nắm bắt các chi tiết hình học dễ thấy trong hình ảnh. Chúng tôi thích tên hình dạng cục bộ hơn các thuộc tính hình học vi phân để biểu thị kết quả hơn là phương thức. Kết quả của việc đánh giá hình dạng cục bộ là trường dữ liệu hình ảnh dày đặc khác với hình dạng đối tượng (được thảo luận trong Phần 4).

Đặc điểm hình dạng cục bộ bắt nguồn từ màu định hướng derivatives—in được kết cáu quan hệ phù hợp được sử dụng trong [115] để rút ra các chi tiết dễ nhận biết trong các bản vá có kết cấu cao của các vật liệu đa dạng. Một loạt các máy dò hình ảnh khá rộng, không có cấu trúc có thể được tìm thấy trong [165].

Lý thuyết không gian quy mô đã được nghĩ ra là bước cơ bản hoàn chỉnh và duy nhất trong tầm nhìn tiên phong, nắm bắt mọi thông tin [187]. Nó cung cấp cơ sở lý thuyết cho việc phát hiện các chi tiết dễ thấy trên bất kỳ quy mô nào. Trong [105], một loạt các bộ lọc Gabor có các hướng khác nhau và tỷ lệ đã được sử dụng để tăng cường các thuộc tính hình ảnh [137]. Các bất biến hình học hình dạng dễ thấy được trình bày trong [136]. Một phương pháp sử dụng thông tin hình dạng và cường độ cục bộ cho truy xuất đối tượng có quan điểm bất biến và phù hợp được đưa ra trong [148]. Phương pháp này dựa vào việc bỏ phiếu giữa một họ hoàn chỉnh các biến thể hình học vi phân. Ngoài ra, [178] tìm kiếm các mô tả bất biến affine-invariant. Từ sự phản xạ bề mặt, trong [5], dấu hiệu cục bộ của độ cong Gaussian được tính toán trong khi không đưa ra giả định nào về suất phản chiếu hoặc mô hình phản xạ khuếch tán.

Kết hợp hình dạng và màu sắc theo kiểu bất biến là một sự kết hợp mạnh mẽ, như được mô tả bởi [56], trong đó màu sắc bên trong và bên ngoài tối đa độ cong affine ở các cạnh màu được lưu trữ để xác định các đối tượng.

## Image Tecture Processing (Xử lý văn bản hình ảnh)

Trong thị giác máy tính, kết cấu được định nghĩa là tất cả những gì còn lại sau khi màu sắc và hình dạng cục bộ đã được xem xét hoặc nó được xác định bởi các thuật ngữ như cấu trúc và tính ngẫu nhiên. Nhiều kết cấu phổ biến bao gồm các texton nhỏ thường có số lượng quá lớn để được coi là các đối tượng bị cô lập. Các yếu tố có thể được đặt nhiều hơn hoặc ít hơn thường xuyên hoặc ngẫu nhiên. Chúng có thể gần như giống hệt nhau hoặc chịu sự thay đổi lớn về ngoại hình và tư thế. Trong bối cảnh truy xuất hình ảnh, nghiên cứu chủ yếu hướng đến các phương pháp thống kê hoặc tổng quát để mô tả đặc điểm của các bản vá.

Các thuộc tính kết cấu cơ bản bao gồm phân tích Markovian, có từ thời Haralick vào năm 1973 và các phiên bản tổng quát của chúng [91], [58]. Khi truy xuất, thuộc tính được tính toán trong mặt nạ trượt để bản địa hóa [99], [59].

Một kỹ thuật phân tích kết cấu quan trọng khác sử dụng MRSAR-models tự phát đa biến, coi kết cấu là kết quả của một hệ thống động xác định chịu tiếng ồn quan sát và trạng thái [174], [106]. Các mô hình khác khai thác các quy tắc thống kê trong trường kết cấu [9].

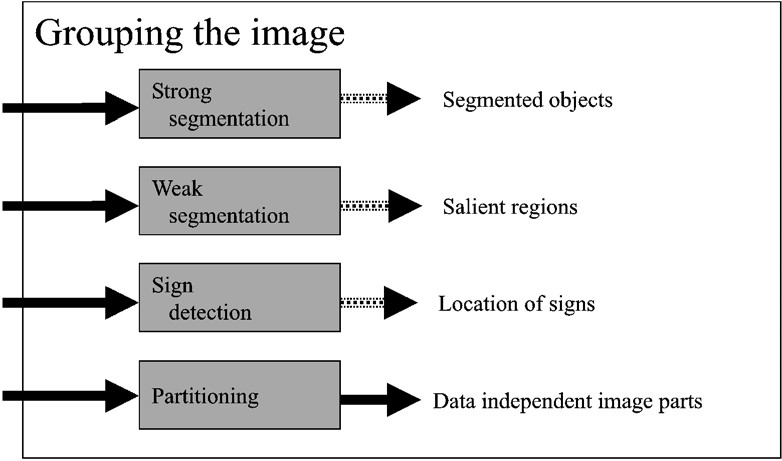
Wavelets [34] đã nhận được sự chú ý rộng rãi. Họ thường được xem xét cho địa phương và hiệu quả của họ. Nhiều biến đổi wavelet được tạo ra bởi các nhóm giãn hoặc giãn và xoay được cho là có một số tương ứng ngữ nghĩa. Các mức thấp nhất của biến đổi wavelet [34], [26] đã được áp dụng cho biểu diễn kết cấu [92], [162] đôi khi kết hợp với phân tích Markovian [25]. Các biến đổi khác cũng đã được khám phá, đáng chú ý nhất là fractals [44]. Một nghiên cứu so sánh vững chắc về phân loại kết cấu từ các thuộc tính chủ yếu dựa trên biến đổi có thể được tìm thấy trong [133].

Tìm kiếm kết cấu tỏ ra hữu ích trong ảnh vệ tinh [98] và hình ảnh của tài liệu [33]. Hoạ tiết cũng đóng vai trò là một tính năng hỗ trợ cho nhận dạng dựa trên phân đoạn [102], nhưng các thuộc tính kết cấu được thảo luận cho đến nay cung cấp ít tham chiếu ngữ nghĩa. Do đó, chúng không phù hợp để sử dụng các ứng dụng trong đó người dung muốn sử dụng các mô tả bằng lời nói của hình ảnh. Do đó, trong nghiên cứu truy xuất, trong [101], các tính năng Wold về tính tuần hoàn, tính định hướng và tính ngẫu nhiên được sử dụng, phù hợp một cách hợp lý với các mô tả ngôn ngữ về kết cấu như được thực hiện trong [128].

## Discussion on Image Processing (Thảo luận về xử lý ảnh)

Trước hết, xử lý hình ảnh trong truy xuất dựa trên nội dung chủ yếu nên tham gia vào việc tăng cường thông tin hình ảnh của truy vấn, chứ không phải mô tả toàn bộ nội dung của hình ảnh.

Để tăng cường thông tin hình ảnh, truy xuất đã thiết lập đèn chiếu vào màu sắc, vì màu sắc có sức mạnh phân biệt cao giữa các vật thể trong một cảnh, cao hơn nhiều so với mức xám. Mục đích của hầu hết xử lý màu sắc hình ảnh là để giảm ảnh hưởng của các điều kiện tình cờ của cảnh và cảm giác (nghĩa là khoảng cách cảm giác). Tiến trình đã được thực hiện trong biểu diễn không gian màu phù hợp cho các lớp mô tả các điều kiện biến thể. Ngoài ra, việc áp dụng mô tả hình học xuất phát từ lý thuyết không gian tỷ lệ sẽ tiết lộ quan điểm và cảnh quan điểm độc lập



do đó, mở đường cho sự tương tự của hình ảnh trên một vài vùng hoặc điểm thông tin nhất. Trong phần này, chúng tôi đã thực hiện một sự tách biệt giữa màu sắc, hình học địa phương và kết cấu.

Tại thời điểm này, có thể an toàn để kết luận rằng bộ phận là một nhãn nhân tạo. Wavelets nói điều gì đó về hình dạng cục bộ cũng như kết cấu và, do đó, có thể mở rộng quy mô không gian và chiến lược bộ lọc cục bộ. Cho các mục đích của content-based phục hồi, một cái nhìn tích hợp về màu sắc, kết cấu và hình học cục bộ là rất cần thiết vì chỉ một chế độ xem tích hợp trên các thuộc tính cục bộ có thể cung cấp các phương tiện để phân biệt giữa hàng trăm ngàn hình ảnh khác nhau. Một tiến bộ gần đây theo hướng đó là sự hợp nhất của sự chiếu sáng và quy mô màu sắc bất biến và thông tin kết cấu thành một tập hợp các thuộc tính cục bộ nhất quán [66]. Ngoài ra, trong [20], các vùng đồng nhất được biểu diễn dưới dạng tập hợp các elip có màu hoặc kết cấu đồng nhất, nhưng các thuộc tính kết cấu bất biến đáng được chú ý hơn, [173] và [185]. Cần nghiên cứu them trong việc thiết kế các bộ thuộc tính hình ảnh hoàn chỉnh với các điều kiện biến thể mà họ có khả năng xử lý. Bất biến chỉ là một mặt của đồng tiền, trong đó quyền lực phân biệt đối xử là mặt khác. Trong content-based truy xuất hình ảnh, các bước đầu tiên được thực hiện để thiết lập sức mạnh phân biệt của các thuộc tính bất biến [55]. Điều này rất cần thiết vì sự cân bằng giữa sự ổn định chống lại các biến thể và sức mạnh phân biệt đối xử được giữ lại quyết định tính hiệu quả của một tùy chọn.