

MỤC LỤC

LỜI CẢM ƠN	3
PHẦN MỞ ĐẦU	4
CHƯƠNG 1 : KHÁI QUÁT VỀ XỬ LÝ ẢNH VÀ BÀI TOÁN VỀ ĐỘ ĐO KHOẢNG CÁCH	5
1.1. Khái quát về xử lý ảnh.....	5
1.1.1. Xử lý ảnh là gì?	5
1.1.2. Một số vấn đề cơ bản trong xử lý ảnh:.....	5
1.2. Bài toán về độ đo khoảng cách.....	6
1.2.1. Bài toán.....	6
1.2.2. Một số ứng dụng của độ đo khoảng cách.....	6
CHƯƠNG 2: MỘT SỐ ĐỘ ĐO KHOẢNG CÁCH	7
2.1. Các độ đo khoảng cách giữa các đối tượng	7
2.1.1. Phân phối chuẩn	8
2.1.2. Độ đo Divergence (độ phân kỳ)	9
2.1.3. Phân phối xác suất rời rạc	14
2.1.4. Khoảng cách Euclid.....	15
2.2. Độ đo khoảng cách giữa các dãy	15
2.2.1. Khoảng cách Hamming	15
2.2.2. Khoảng cách Hamming mờ.....	16
2.2.3. Khoảng cách Levenshtein(chỉnh sửa)	16
2.2.4. Khoảng cách liên quan khác.....	17
2.2.5. Khoảng cách thông tin và xấp xỉ thông tin.....	17
2.3. Độ đo theo lý thuyết thông tin	18
2.4. Độ đo khoảng cách giữa các tập hợp.....	19
2.4.1. Khoảng cách Hausdorff.....	19
2.4.2. Các biến thể của khoảng cách Hausdorff.....	22
2.4.3. Các độ đo trên tập mờ	23
2.5. Độ đo khoảng cách trong các ứng dụng	24
2.5.1. Bất biến.....	24

2.5.2. Ví dụ về độ đo	24
CHƯƠNG III. CHƯƠNG TRÌNH THỬ NGHIỆM	34
3.1. Bài toán:	34
3.2. Phương pháp thực hiện	34
3.3. Kết quả	34
KẾT LUẬN	37
TÀI LIỆU THAM KHẢO:	38

LỜI CẢM ƠN

Em xin chân thành cảm ơn PGS TS Ngô Quốc Tạo, Trưởng phòng Nhân dạng và Công nghệ tri thức, Viện Công nghệ thông tin, Viện Hàn Lâm Khoa học và Công nghệ Việt Nam đã định hướng và giúp đỡ em tận tình trong suốt quá trình làm đồ án.

Em xin chân thành cảm ơn các thầy, cô giáo bộ môn khoa Công Nghệ Thông Tin đã truyền dạy những kiến thức thiết thực trong suốt quá trình học, đồng thời em xin cảm ơn nhà trường đã tạo điều kiện tốt nhất cho em hoàn thành đồ án này.

Trong phạm vi hạn chế của một đồ án tốt nghiệp, những kết quả thu được còn là rất ít và quá trình làm việc khó tránh khỏi những thiếu sót, em rất mong nhận được sự góp ý của các thầy cô giáo và các bạn.

PHẦN MỞ ĐẦU

Trong khoa học vật lý giai đoạn cần thiết đầu tiên theo hướng chủ đề học tập nào là tìm nguyên tắc số cách tính toán và phương pháp thực hành để đánh giá chất lượng một số kết nối với nó. Tôi thường nói rằng, khi bạn có thể đánh giá những gì bạn đang nói đến, và biểu diễn nó bằng các con số, bạn biết gì về nó, trong khi bạn không thể đo lường nó, khi bạn không thể biểu diễn nó với số, tri thức của bạn là sơ sài và không thỏa đáng, nó có thể là sự khởi đầu của tri thức, nhưng hầu như trong suy nghĩ của bạn luôn tiến đến trạng thái của khoa học, bất cứ vấn đề gì có thể được.

“POPULAR LECTURES AND ADDRESSES”, LORD KELVIN

Sự giống nhau tương đối có thể được định nghĩa là mối quan hệ giữa hai thực thể có cùng tính chất hoặc có các đặc điểm giống nhau, nhưng khác nhau về độ đo hoặc mức độ. Lớn hơn giá trị tương đồng, lớn hơn sự tương đồng giữa các đối tượng. Mặt khác, sự không giống nhau tương đối tập trung vào sự khác biệt; nhỏ hơn sự khác nhau, giống nhau hơn các đối tượng. Cả giá trị giống nhau và giá trị không giống nhau thể hiện khái niệm về chân dung giữa các đối tượng, nhưng sự nhấn mạnh là khác nhau. Đó là phù hợp hơn để xác định phụ thuộc vào loại dữ liệu và các vấn đề ở bàn tay. Nói chung, sự lân cận là một chức năng của các biến quan sát hoặc các thông số thu thập. Chúng ta sẽ đề cập đến nó như là một độ đo, mặc dù nó có thể không được như vậy theo nghĩa toán học nghiêm ngặt. Nội dung đồ án sẽ trình bày tổng quan về các độ đo không giống nhau đối với các loại dữ liệu khác nhau, cùng với đặc điểm của nó. Một số trong số đó đã được biết đến, trong khi những độ đo khác còn tương đối mới.

Nội dung đồ án bao gồm 3 chương:

Chương 1: Trình bày các độ đo khoảng cách không giống nhau.

Chương 2: Các độ đo đặc biệt được sử dụng trong lĩnh vực học tập mô hình và ứng dụng của các độ đo.

Chương 3: Chương trình thực nghiệm và kết quả.

Cuối cùng là phần kết luận.

CHƯƠNG I : KHÁI QUÁT VỀ XỬ LÝ ẢNH VÀ BÀI TOÁN VỀ ĐỘ ĐO KHOẢNG CÁCH

1.1. Khái quát về xử lý ảnh.

1.1.1. Xử lý ảnh là gì?

Xử lý ảnh là một khoa học còn tương đối mới mẻ so với nhiều ngành khoa học khác, nhất là trên qui mô công nghiệp.

Xử lý ảnh là quá trình thực hiện các thao tác trên ảnh đầu vào cho ra kết quả như mong muốn. Ảnh kết quả có thể khác so với ảnh ban đầu tốt hơn hoặc xấu hơn so với ảnh đầu vào.

1.1.2. Một số vấn đề cơ bản trong xử lý ảnh:

a) Một số khái niệm cơ bản:

- Ảnh: là một tập hợp hữu hạn các điểm ảnh kề nhau. Ảnh thường được biểu diễn bằng một ma trận 2 chiều, mỗi phần tử của ma trận tương ứng với một điểm ảnh.
- Điểm ảnh: được xem như là đặc trưng cường độ sáng hay một dấu hiệu nào đó tại một vị trí nào đó của đối tượng trong không gian.
- Mức xám: là kết quả sự mã hóa tương ứng một cường độ sáng của mỗi điểm ảnh với 1 giá trị số - kết quả của quá trình lượng hóa.
- Biểu diễn ảnh: Trong biểu diễn ảnh người ta thường dùng các phần tử đặc trưng của ảnh là pixel. Việc xử lý ảnh số yêu cầu ảnh phải được mẫu hóa và lượng tử hóa. Một số mô hình được dùng trong biểu diễn ảnh: mô hình toán, mô hình thống kê.

b) Tăng cường ảnh – khôi phục ảnh:

- Tăng cường ảnh là bước quan trọng tạo tiền đề cho xử lý ảnh. Nó gồm các kỹ thuật: lọc độ tương phản, khử nhiễu, nổi màu....
- Khôi phục ảnh là nhằm loại bỏ các suy giảm trong ảnh.

c) Biến đổi ảnh: Thuật ngữ biến đổi ảnh thường được dùng để nói tới một lớp các ma trận đơn vị và các kỹ thuật dùng để biến đổi ảnh. Có nhiều loại biến dạng được dùng như: biến đổi Fourier, sin, cosin

d) Nhận dạng ảnh: Nhận dạng ảnh là quá trình liên quan đến các mô tả đối tượng mà người ta muốn đặc tả nó. Người ta đã áp dụng kỹ thuật nhận dạng khá thành công với nhiều đối tượng khác nhau như: nhận dạng vân tay, nhận dạng chữ viết... Có bốn cách tiếp cận khác nhau:

- +/ Đối sánh mẫu dựa trên các đặc trưng được trích chọn.
- +/ Phân loại thống kê.
- +/ Đối sánh cấu trúc.
- +/ Phân loại dựa trên mạng nơron nhân tạo.

e) Nén ảnh:

Dữ liệu ảnh cũng như các dữ liệu khác cần phải lưu trữ hay truyền đi trên mạng mà lượng thông tin để biểu diễn cho một ảnh là rất lớn. Do đó cần phải giảm lượng thông tin hay nén dữ liệu là một nhu cầu cần thiết. Nén ảnh thường được tiến hành theo cả hai khuynh hướng là nén có bảo toàn và không bảo toàn thông tin.

1.2. Bài toán về độ đo khoảng cách

1.2.1. Bài toán

Độ đo tương tự là một trong những phương pháp tốt để máy tính phân biệt được các hình ảnh qua nội dung của chúng. Thông thường hệ thống tra cứu ảnh sẽ truy vấn hình ảnh bằng phương pháp đo tương tự dựa trên các chức năng, việc xác định nó có thể dưới nhiều hình thức như phát hiện biên, màu sắc, vị trí điểm ảnh... các phương pháp như histogram, màu sắc và phân tích histogram dòng cột sử dụng biểu đồ để xác định độ tương tự. Áp dụng cho bài toán độ đo khoảng cách: cho một ảnh đầu vào và một danh sách ảnh, sau đó sử dụng một trong số các độ đo khoảng cách để xác định độ tương tự của ảnh trong danh sách ảnh với ảnh đầu vào. Ảnh nào trong danh sách ảnh có độ đo khoảng cách gần với ảnh đầu vào nhất thì sẽ được sắp xếp theo thứ tự.

1.2.2. Một số ứng dụng của độ đo khoảng cách

Độ đo khoảng cách được ứng trong rất nhiều lĩnh vực như xử lý ảnh và nhận dạng mẫu, nhận dạng chữ viết tay, trong y học giúp bác sĩ phát hiện các mô bệnh để tìm ra các tế bào ung thư (sử dụng công cụ tự phát huỳnh quang),... Như vậy, ta có thể thấy tầm quan trọng của độ đo khoảng cách trong thực tiễn là rất lớn.

CHƯƠNG 2: MỘT SỐ ĐỘ ĐO KHOẢNG CÁCH

2.1. Các độ đo khoảng cách giữa các đối tượng

Để phân tích sự khác biệt giữa các đối tượng được mô tả bởi các vector trong một không gian đặc trưng, một số độ đo khác nhau có thể được xem xét. Nếu các vector trung bình được sử dụng để làm đại diện cho toàn bộ các đối tượng, chúng có thể được sử dụng để tính toán khoảng cách giữa các nhóm theo các công thức từ bảng 2.1.

Ref.	D	Dissimilarity $d(\mathbf{x}, \mathbf{y})$	M	E
Euclidean	D_E, D_2	$\sqrt{(\mathbf{x} - \mathbf{y})^T(\mathbf{x} - \mathbf{y})}$	Yes	Yes
Weighted Euclidean	D_{wE}	$\sqrt{(\mathbf{x} - \mathbf{y})^T \text{diag}(w_i^2)(\mathbf{x} - \mathbf{y})}$	Yes	Yes
City block	D_1	$\sum_{i=1}^m x_i - y_i $	Yes	No
Max norm	D_{\max}	$\max_i x_i - y_i $	Yes	No
ℓ_p or Minkowski	D_p	$(\sum_{i=1}^m x_i - y_i ^p)^{1/p}, p \geq 1, p \neq 2$	Yes	No
Mahalanobis	D_M	$\sqrt{(\mathbf{x} - \mathbf{y})^T C^{-1}(\mathbf{x} - \mathbf{y})}$; C is psd	Yes	Yes
Median distance	D_{med}	$D_{\lfloor n/2 \rfloor - \text{rank}}$	No	No
Correlation-based	D_{corr}	$\frac{1}{2} \left(1 - \frac{\mathbf{x}^T \mathbf{y}}{\ \mathbf{x}\ ^2 + \ \mathbf{y}\ ^2} \right)$	No	No
Correlation-based	D_{corr2}	$\frac{1}{2} \left(1 - \frac{\mathbf{x}^T \mathbf{y}}{\ \mathbf{x}\ ^2 + \ \mathbf{y}\ ^2 - 2\mathbf{x}^T \mathbf{y}} \right)$	No	No
Cosine	D_{cos}	$\frac{1}{2} \left(1 - \frac{\mathbf{x}^T \mathbf{y}}{\ \mathbf{x}\ \ \mathbf{y}\ } \right)$	No	No
Divergence	D_{div}	$\sqrt{\sum_{i=1}^n \frac{(x_i - y_i)^2}{(x_i + y_i)^2}}$	No	No
Bray and Curtis	D_{BC}	$\frac{\sum_{i=1}^n x_i - y_i }{\sum_{i=1}^n x_i + y_i}$	No	No
Soergel	D_S	$\frac{\sum_{i=1}^n x_i - y_i }{\sum_{i=1}^n \max\{x_i, y_i\}}$	No	No
Ware and Hedges	D_{WH}	$\sum_{i=1}^n \left(1 - \frac{\min\{x_i, y_i\}}{\max\{x_i, y_i\}} \right)$	No	No

Bảng 2.1

Một khả năng khác là đặc trưng cho một đối tượng bằng một hàm phân bố xác suất nhiều biến (pdf) $F(\mathbf{x})$. Sau đó, sự khác biệt giữa hai quần thể được đo bằng sự khác nhau giữa hai hàm phân bố xác suất pdf F_1 và F_2 . Độ đo Kolmogorov thường được sử dụng [Gibbs và Su, 2002]. Cho hai hàm phân phối F_1 và F_2 nó được định nghĩa như sau:

$$D_K(F_1, F_2) = \sup_{\mathbf{x}} |F_1(\mathbf{x}) - F_2(\mathbf{x})|. \quad (2.1)$$

Như một phần mở rộng, việc đánh giá sự khác nhau giữa các đối tượng cũng có thể dựa vào mô tả từng phân phối như là một điểm trong một không gian Riemann với các tọa độ xác định bởi các thông số đối tượng. Ví dụ: một đối tượng đặc trưng bởi một hàm mật độ bình thường được xác định bởi các tọa độ (μ, Σ) trong $m + m(m+1)/2$ không gian chiều. Đối tượng được mô tả bởi các thông số tương tự sẽ được ánh xạ thành các điểm lân cận trong không gian này. Với điều kiện một độ đo metric phù hợp có thể được xác định, sự khác nhau giữa các nhóm là chiều dài trắc địa (kết nối con đường ngắn nhất hai điểm trên một đa tạp) giữa các điểm biểu diễn cho đối tượng.

2.1.1. Phân phối chuẩn

Giả định của dữ liệu được rút ra từ một phân bố chuẩn thường được thực hiện trong thực tế. Do đó, cần có các độ đo không giống nhau thích hợp. Một độ đo cổ điển giữa hai phân phối chuẩn $N(\mu_1, \Sigma)$ và $N(\mu_2, \Sigma)$ với ma trận hiệp phương sai bằng Σ là khoảng cách Mahalanobis vuông D_M giữa các phương thức:

$$D_M^2(\mu_1, \mu_2; \Sigma) = (\mu_1 - \mu_2)^\top \Sigma^{-1} (\mu_1 - \mu_2). \quad (2.2)$$

Vì các tham số phân phối hầu như không được biết đến nên trong thực tế chúng được thay thế bằng cách ước lượng mẫu: $\bar{\mathbf{x}}_{(i)} = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{x}_j$, $i = 1, 2$ và $C = \frac{1}{n_1 + n_2 - 2} ((n_1 - 1)C_1 + (n_2 - 1)C_2)$, trong đó n_i biểu thị kích cỡ $\bar{\mathbf{x}}_i$ và C_i , $i = 1, 2$, biểu diễn cho mẫu có nghĩa là vectơ và ma trận hiệp phương sai mẫu tương ứng. Khoảng cách Mahalanobis ước tính sau đó trở thành:

$D_M^2(\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2; C) = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)^\top C^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$. Nếu $C = I$ hoặc $C = \text{diag}(\sigma_i)$, thì các D_M^2 trở thành Euclide hay khoảng cách Euclide trọng lượng giữa các vectơ tâm tương ứng. Lưu ý, nếu khoảng cách Mahalanobis được xét đối với một không gian $X = N(\mu,$

Σ), sau đó không gian (X, d_M) là tiền metric.

Khoảng cách Mahalanobis là dựa trên giả định của ma trận hiệp phương sai bằng nhau. Cho ma trận hiệp phương sai không đồng nhất, tổng quát của nó dẫn đến bán kính thông tin chuẩn [Jardine và Sibson, 1971]. Cho hai phân phối chuẩn $N_1 \equiv N(\mu_1, \Sigma_1)$ và $N_2 \equiv N(\mu_2, \Sigma_2)$, ta có:

$$d_{NIR}(\mathcal{N}_1, \mathcal{N}_2) = \begin{cases} \frac{1}{2} \log \frac{\frac{1}{2} \det(\Sigma_1 + \Sigma_2) + \frac{1}{4} \|\mu_1 - \mu_2\|^2}{(\det(\Sigma_1))^{\frac{1}{2}} (\det(\Sigma_2))^{\frac{1}{2}}}, & \text{if } \Sigma_1 \neq \Sigma_2 \\ \frac{1}{2} \log \left(1 + \frac{1}{4} D_M^2(\mu_1, \mu_2; \Sigma) \right), & \text{if } \Sigma \equiv \Sigma_1 = \Sigma_2 \end{cases} \quad (2.3)$$

Một thước đo khoảng cách giữa các phân phối chuẩn, phù hợp với ma trận hiệp biến không đồng nhất, được đề xuất trong [Anderson và Bahadur, 1962]. Cho $b_\alpha = (\alpha \Sigma_1 + (1 - \alpha) \Sigma_2)^{-1} (\mu_1 - \mu_2)$ với $\alpha \in (0, 1)$. Sau đó,

$$d_{AB}(\mathcal{N}_1, \mathcal{N}_2) = \max_{\alpha} \frac{2 \mathbf{b}_\alpha^T (\mu_1 - \mu_2)}{(\mathbf{b}_\alpha^T \Sigma_1 \mathbf{b}_\alpha)^{\frac{1}{2}} + (\mathbf{b}_\alpha^T \Sigma_2 \mathbf{b}_\alpha)^{\frac{1}{2}}}. \quad (2.4)$$

Như trước đây, các thông số phân phối được thay thế bằng ước lượng mẫu. Các độ đo khác đối với phân phối chuẩn được thể hiện trong phần tiếp theo.

2.1.2. Độ đo Divergence (độ phân kỳ)

Nhiều độ đo cổ điển thể hiện sự khác biệt giữa hai phân phối xác suất F_1 và F_2 với các hàm mật độ f_1 và f_2 là trường hợp đặc biệt của \emptyset -phân kỳ được đề xuất bởi

Csiszar [Csiszar, 1967], dựa trên tỷ lệ khả năng $\lambda(\mathbf{x}) = \frac{f_2(\mathbf{x})}{f_1(\mathbf{x})}$:

$$\begin{aligned} d_\phi(F_1, F_2) &= E[\phi(\lambda(X))] = \int_{\mathcal{D}} \phi(\lambda(\mathbf{x})) dF_1(\mathbf{x}) \\ &= \int_{\mathcal{D}} \phi\left(\frac{f_2(\mathbf{x})}{f_1(\mathbf{x})}\right) f_1(\mathbf{x}) d\mu(\mathbf{x}), \end{aligned} \quad (2.5)$$

trong đó $\emptyset(\lambda)$ là một số thực, hàm lồi được xác định trên \mathbb{R}_+ sao cho $\emptyset(1) = 0$, và μ là một thước đo trên miền \mathcal{D} . Lưu ý rằng bằng cách đảo ngược các đối số F_1 và F_2 của $d_\phi(F_1, F_2)$, thu được \emptyset -phân kỳ khác, tức là $d_\phi(F_2, F_1)$ trở thành $d_{\lambda\phi}(1/\lambda)(F_1, F_2)$. Hơn

nữa, sự phân kỳ đối xứng, $d_{\phi}(F_1, F_2) + d_{\phi}(F_2, F_1)$, có thể được xem xét như $d_{\phi(\lambda) + \lambda \phi(1/\lambda)}(F_1, F_2)$ [Esposito và cộng sự, 2000.].

Một số độ đo phân kỳ nổi tiếng với biểu đồ phân bố liên tục một biến được đưa ra dưới đây, cùng với các công thức tương đương cho hai phân phối chuẩn. Công thức cho phân phối rời rạc bị bỏ qua vì chúng là những khái quát đơn giản của những giá trị liên tục, bằng cách sử dụng tổng thay vì tích phân. Các nghiên cứu về mối quan hệ giữa các độ đo phân kỳ được trình bày cũng như tổng quát hóa của chúng có thể được tìm thấy ví dụ trong [Taneja, 1989, 1995] hoặc trong các cuốn sách trực tuyến [Taneja].

Cho ngắn gọn, chúng ta hãy biểu thị $\mathcal{N}_i \equiv \mathcal{N}(\boldsymbol{\mu}_i, \Sigma_i)$, cho $i = 1, 2$, và $\Sigma = \Sigma_1 = \Sigma_2$, cho ma trận hiệp phương sai bằng nhau và bình phương khoảng cách Mahalanobis D^2_M . Các biểu đồ phân bố giống như f_1 và f_2 là liên tục trên khoảng thời gian phân chia $I_1^{(1)}, I_2^{(1)}, \dots, I_{N_1}^{(1)}$ và $I_1^{(2)}, I_2^{(2)}, \dots, I_{N_2}^{(2)}$, tương ứng như

vậy $f_i(x) = \sum_{\tau=1}^{N_i} h_{\tau}^{(i)} \mathcal{I}(x \in I_{\tau}^{(i)})$, $i = 1, 2$, trong đó $h_{\tau}^{(i)}$ là trọng lượng dương. $J_{st} = I_s^{(1)} \cap I_t^{(2)}$ là viết tắt của giao giữa hai khoảng thời gian $I_s^{(1)}$ và $I_t^{(2)}$ và $\mu(J_{st})$ là độ dài (độ đo Lebesgue) của J_{st} .

+/ Độ khác nhau Kullback-Leibler:

Độ đo này, còn được gọi là khoảng cách thông tin hoặc dữ liệu ngẫu nhiên tương đối [Esposito và cộng sự., 2000], ta thu được $\phi(\lambda) = \lambda \log(\lambda)$, $\lambda > 0$ and $\phi(0) = 0$:

$$d_{KL}(F_1, F_2) = \int_{\mathcal{D}} \log \left(\frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} \right) f_1(\mathbf{x}) d\mathbf{x}. \tag{2.6}$$

Quy ước thông thường là $\log(0/b) = 0$ cho tất cả các b và $\log(a/0) = \infty$ cho tất cả a khác không. Do đó, d_{KL} là giá trị lợi tức trong $[0, \infty]$.

Các độ đo Kullback-Leibler dựa trên khái niệm trọng lượng thông tin. Nếu hai đối tượng được mô tả bởi các phân bố xác suất, d_{KL} thể hiện các thông tin trung bình cho việc loại bỏ đối tượng đầu tiên để nghiêng về đối tượng thứ hai, khi x thuộc về đối

tượng thứ hai độ đo này là không đối xứng, do đó nó không phải là độ đo metric. Cho hai phân phối chuẩn m chiều, d_{KL} trở thành:

$$d_{KL}(\mathcal{N}_1, \mathcal{N}_2) = \frac{1}{2} \left(D_M^2(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; \Sigma_2) + \text{tr}(\Sigma_2^{-1}\Sigma_1 - I) + \log \frac{\det(\Sigma_2)}{\det(\Sigma_1)} \right) \quad (2.7)$$

hoặc $d_{KL}(\mathcal{N}_1, \mathcal{N}_2) = \frac{1}{2} D_M^2(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; \Sigma)$ khi ma trận hiệp phương sai bằng nhau. Đối với hai phân phối biểu đồ giống nhau, d_{KL} được cho là:

$$d_{KL}(F_1, F_2) = \sum_{s=1}^{N_1} \sum_{t=1}^{N_2} \log \left(\frac{h_t^{(2)}}{h_s^{(1)}} \right) h_t^{(2)} \mu(J_{st}).$$

+/ Hệ số J:

Cho $\varnothing(\lambda) = (\lambda - 1) \log(\lambda)$, chúng ta có được một đối xứng Kullback-Leibler phân kỳ:

$$d_J(F_1, F_2) = d_{KL}(F_1, F_2) + d_{KL}(F_2, F_1). \quad (2.8)$$

Đối với hai phân phối chuẩn m chiều, d_J trở thành:

$$d_J(\mathcal{N}_1, \mathcal{N}_2) = \frac{1}{2} \left(D_M^2(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; \Sigma_1) + D_M^2(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; \Sigma_2) + \text{tr}(\Sigma_1^{-1}\Sigma_2 - I) + \text{tr}(\Sigma_2^{-1}\Sigma_1 - I) \right) \quad (2.9)$$

hoặc $d_J(\mathcal{N}_1, \mathcal{N}_2) = D_M^2(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; \Sigma)$, khi ma trận hiệp phương sai bằng nhau. Đối với hai phân phối biểu đồ giống nhau, ta có:

$$d_J(F_1, F_2) = \sum_{s=1}^{N_1} \sum_{t=1}^{N_2} \log \left(\frac{h_t^{(2)}}{h_s^{(1)}} \right) (h_t^{(2)} - h_s^{(1)}) \mu(J_{st}).$$

+/ Bán kính thông tin.

Đây là một độ đo đối xứng thu được cho $\varnothing(\lambda) = -\frac{1}{2}(1 + \lambda) \log(1 + \frac{\lambda}{2})$:

$$d_{IR}(F_1, F_2) \equiv d_\varnothing(F_1, F_2) + d_\varnothing(F_2, F_1). \quad (2.10)$$

Đối với hai phân phối chuẩn, d_{IR} trở thành bán kính thông tin chuẩn.

+/ χ^2 -phân kỳ.

Độ đo này không đối xứng (như vậy không phải là độ đo metric) thu được cho $\phi(\lambda) = (\lambda - 1)^2$:

$$d_{\chi^2}(F_1, F_2) = \int_{\mathcal{D}} \left(\frac{f_2(\mathbf{x})}{f_1(\mathbf{x})} - 1 \right)^2 f_1(\mathbf{x}) d\mathbf{x} = \int_{\mathcal{D}} \frac{(f_1(\mathbf{x}) - f_2(\mathbf{x}))^2}{f_1(\mathbf{x})} d\mathbf{x}. \quad (2.11)$$

Đối với hai phân phối chuẩn, với $(2\Sigma_2^{-1} - \Sigma_1^{-1})$ việc xác định đại lượng dương, d_{χ^2} trở thành:

$$\begin{aligned} d_{\chi^2}(\mathcal{N}_1, \mathcal{N}_2) &= \frac{\det(\Sigma_1 \Sigma_2^{-1})}{2 (\det(\Sigma_1 \Sigma_2^{-1} - I))^{\frac{1}{2}}} \exp \left\{ \frac{1}{2} \left(D_M^2(2\Sigma_2^{-1}\boldsymbol{\mu}_2, \Sigma_1^{-1}\boldsymbol{\mu}_1; 2\Sigma_2^{-1} - \Sigma_1^{-1}) \right. \right. \\ &\quad \left. \left. + D_M^2(\boldsymbol{\mu}_1, \mathbf{0}; \Sigma_1) - 2 D_M^2(\boldsymbol{\mu}_2, \mathbf{0}; \Sigma_2) \right) \right\} - 1 \end{aligned} \quad (2.12)$$

hoặc $d_{\chi^2}(\mathcal{N}_1, \mathcal{N}_2) = \exp \{ D_M^2(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; \Sigma) \} - 1$ khi ma trận hiệp phương sai giống nhau. Đối với hai phân phối biểu đồ giống nhau, d_{χ^2} tương đương với:

$$d_{\chi^2}(F_1, F_2) = \sum_{s=1}^{N_1} \sum_{t=1}^{N_2} \frac{|h_t^{(2)} - h_s^{(1)}|}{h_s^{(1)}} \mu(J_{st}). \quad (2.13)$$

+/ Hệ số Hellinger.

Độ đo tương tự này là thu được đối với $\phi(\lambda) = \lambda^t$ trong đó $t \in (0,1)$:

$$s_H^{(t)}(F_1, F_2) = \int_{\mathcal{D}} f_2(\mathbf{x})^t f_1(\mathbf{x})^{1-t} d\mathbf{x}. \quad (2.14)$$

Đối với hai phân phối chuẩn m chiều. $s_H^{(t)}$ trở thành:

$$s_H^{(t)}(\mathcal{N}_1, \mathcal{N}_2) = \exp \left\{ -\frac{t(1-t)}{2} D_M^2(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; t\Sigma_2 + (1-t)\Sigma_1) + \log \left(\frac{\det(t\Sigma_2 + (1-t)\Sigma_1)}{\det(\Sigma_2)^t \det(\Sigma_1)^{1-t}} \right)^{\frac{1}{2}} \right\} \quad (2.15)$$

hoặc $s_H^{(t)}(\mathcal{N}_1, \mathcal{N}_2) = \exp \left\{ -\frac{t(1-t)}{2} D_M^2(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; \Sigma) \right\}$, khi ma trận hiệp phương sai đều giống nhau.

+/ Hệ số Chernoff và Bhattacharyya:

Cho $t = 1/2$, hệ số Hellinger tương tự trở thành hệ số Bhattacharyya đối xứng [Fukunaga, 1990]. Khoảng cách Bhattacharyya khi đó là:

$$d_{BH}(F_1, F_2) = -\log(s_H^{(1/2)}(F_1, F_2)). \quad (2.16)$$

Đối với hai phân phối chuẩn, nó sẽ trở thành:

$$d_{BH}(\mathcal{N}_1, \mathcal{N}_2) = \frac{1}{8} D_M^2 \left(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; \frac{1}{2}(\Sigma_1 + \Sigma_2) \right) + \frac{1}{2} \log \frac{\frac{1}{2} \det(\Sigma_1 + \Sigma_2)}{(\det(\Sigma_1) \det(\Sigma_2))^{\frac{1}{2}}}. \quad (2.17)$$

Khoảng cách Bhattacharyya là một trường hợp đặc biệt của khoảng cách Chernoff [Fukunaga, 1990]:

$$d_{CH}^{(t)}(F_1, F_2) = -\log(s_H^{(t)}(F_1, F_2)). \quad (2.18)$$

Các khoảng cách Chernoff và Bhattacharyya là rất quan trọng trong khu vực phân loại vì chúng cung cấp giới hạn trên về lỗi Bayes của hai lớp được mô tả bởi các phân phối chuẩn [Fukunaga, 1990; Duda et al, 2001.].

+/ Khoảng cách thay đổi và khoảng cách l_2 :

Đối với lựa chọn $\emptyset(\lambda) = |1 - \lambda|$ hoặc $\emptyset(\lambda) = |1 - \lambda|^2$, tương đương đối xứng của l_1 và khoảng cách l_2 thu được:

$$d_p(F_1, F_2) = \int_{\mathcal{D}} |f_2(\mathbf{x}) - f_1(\mathbf{x})|^p d\mathbf{x}, \quad p = 1, 2. \quad (2.19)$$

Đối với hai phân phối chuẩn m chiều, d_2 trở thành:

$$d_2(\mathcal{N}_1, \mathcal{N}_2) = \frac{1}{2^m \pi^{\frac{m}{2}}} \left(\frac{1}{(\det(\Sigma_1))^{\frac{1}{2}}} + \frac{1}{(\det(\Sigma_2))^{\frac{1}{2}}} \right) - \frac{2}{(2\pi)^{\frac{m}{2}} (\det(\Sigma_1 + \Sigma_2))^{\frac{1}{2}}} \exp \left\{ \frac{1}{2} D_M^2(\Sigma_2^{-1} \boldsymbol{\mu}_2, \Sigma_1^{-1} \boldsymbol{\mu}_1; \Sigma_2^{-1} + \Sigma_1^{-1}) - \sum_{i=1}^2 D_M^2(\boldsymbol{\mu}_i, \mathbf{0}; \Sigma_i) \right\}, \quad (2.20)$$

hoặc khi ma trận hiệp phương sai bằng nhau, ta có:

$$d_2(\mathcal{N}_1, \mathcal{N}_2) = \frac{2}{2^m \pi^{\frac{m}{2}} (\det(\Sigma))^{\frac{1}{2}}} (1 - \exp \left\{ \frac{1}{4} D_M^2(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2; \Sigma) \right\}). \quad (2.21)$$

2.1.3. Phân phối xác suất rời rạc

Chúng ta hãy xét đối tượng n , mô tả bởi m biến phân loại và thuộc hai nhóm.

Các nhóm này sau đó được xử lý như phân phối riêng biệt. Cho $p_i^{kj} = n_i^{kj} / n$ là

tần số tương đối, trong đó n_i^{kj} là số lượng các trường thuộc lớp thứ j hiện có của biến thứ k trong nhóm thứ i , trong đó $i = 1, 2$. Cho

$\mathbf{p}_i = [p_i^{11} \dots p_i^{1c_1} p_i^{21} \dots p_i^{2c_2} \dots p_i^{mc_m}]$ và c_k là số lượng các loại khác nhau cho các

biến thứ k và $c = \sum_{k=1}^m c_k$. Khoảng cách giữa các nhóm có thể được tính như sau:

$$d^2(\mathbf{p}_1, \mathbf{p}_2) = \sum_{k=1}^m \sum_{j=1}^{c_k} \frac{(p_1^{kj} - p_2^{kj})^2}{\frac{1}{2}(p_1^{kj} + p_2^{kj})}. \quad (2.22)$$

Một khả năng khác để mở rộng khoảng cách Mahalanobis bằng cách thay thế các biến liên tục của những biến phân loại. Nếu C là một ma trận hiệp phương sai mẫu $c \times c$, độ đo này được đưa ra như sau:

$$D_{M-cat}^2(\mathbf{p}_1, \mathbf{p}_2) = (\mathbf{p}_1 - \mathbf{p}_2)^T C^{-1} (\mathbf{p}_1 - \mathbf{p}_2). \quad (2.23)$$

Hệ số phép biến đổi có thể được sử dụng tốt. Nó liên quan đến sự giống nhau Hellinger và các thước đo sự giống nhau giữa hai phạm trù hoặc các đặc trưng phương thức, hoặc

hai biểu đồ. Cho $p_i^{kj} = n_i^{kj} / n$, như ở trên. Do đó, những tần số tạo ra một phân bố xác suất rời rạc. Các mối quan hệ giữa hai phân phối tần số cho các biến f_k được thể hiện là $a_{f_k} = \sum_{j=1}^{c_k} (p_1^{kj} p_2^{kj})^{\frac{1}{2}}$. Điều này dẫn đến sự khác nhau về mối quan hệ giữa các nhóm:

$$d_{\text{aff}}(\mathbf{p}_1, \mathbf{p}_2) = 1 - \sum_{k=1}^m w_k a_{f_k} \tag{2.24}$$

trong đó w_k là trọng lượng thích hợp.

2.1.4. Khoảng cách Euclid

Đây là cách tính khoảng cách Euclid thông thường giữa các K bin:

$$\text{Intersection } \langle \mathbf{Q}, h \rangle = \sum_{j=1}^K \sqrt{\langle \mathbf{Q} \rangle - h}^2 \tag{2.25}$$

2.2. Độ đo khoảng cách giữa các dãy

Gọi A là một bảng chữ cái, ví dụ như một bộ sưu tập hữu hạn các ký hiệu, còn được gọi là chữ cái, từ đó trình tự hoặc chuỗi được tạo thành. Cho $s = s_1 s_2 \dots s_n$ là một chuỗi các ký tự từ A. Một từ rỗng được ký hiệu là ϵ và nó có một chiều dài vô giá trị. Chuỗi này được sử dụng trong các mô hình nhận biết và khu vực máy học để mã hóa các đối tượng của cấu trúc tương đối đồng nhất. Ở đây, tôi sẽ giới thiệu ngắn gọn các độ đo khoảng cách phổ biến nhất.

2.2.1. Khoảng cách Hamming

Đây là một trong những độ đo đơn giản nhất: cho hai chuỗi chiều dài bằng nhau, nó đếm vị trí ký hiệu mà chúng khác nhau. Để không mất tính tổng quát, cho $s = s_1 s_2 \dots s_n$ và $t = t_1 t_2 \dots t_n$, là chuỗi nhị phân. Khoảng cách Hamming sau đó được xác định như sau $d_{\text{Ham}}(s, t) = \sum_{k=1}^n \mathbf{I}(s_k \neq t_k)$. Nó không phải là một độ đo linh hoạt vì nó giả định các chuỗi có chiều dài cố định. Trong nhiều bài toán, tuy nhiên, các chuỗi có chiều dài thay đổi, và hơn nữa, có thể giữa các vị trí biểu tượng sự tương ứng không ổn định. Sự thay đổi nhỏ của các vị trí của một trong hai chuỗi gần như giống hệt nhau có thể dẫn đến các giá trị phóng đại trong khoảng cách Hamming.

2.2.2. Khoảng cách Hamming mờ

Một khoảng cách Hamming mờ đã được đề xuất để làm cho khoảng cách Hamming nhạy cảm với khu vực lân cận [Bookstein et al. 2001]. Đây là kiểu của khoảng cách chỉnh sửa cho trình tự của chiều dài bằng nhau. Sửa khoảng cách dựa trên chuyển một chuỗi thành chuỗi khác bằng cách sử dụng cái gọi là phép toán sửa. Các hoạt động của phép toán chỉnh sửa sau đây được giới thiệu: chèn, xóa và thay đổi, với giá trị c_{ins} , c_{del} và c_{sub} được phân công tương ứng. Các phép toán dịch chuyển cho phép chuyển đổi một 1-bit trong một chuỗi đến gần 1-bit trong chuỗi khác với chi phí nhỏ hơn bởi có cả hai phép xóa và chèn. Các phép toán được sử dụng để chuyển đổi một chuỗi ký tự thành chuỗi khác và sự khác nhau của kết quả d_{fHam} được tính bằng cách cộng các chi phí của các phép toán, như vậy nó có tổng chi phí tối thiểu. Khoảng cách Hamming mờ là độ đo metric nếu $c_{del} = c_{ins}$ và cho kích thước tuyệt đối của sự chuyển dịch $h \geq 0$, $c_{sub}(h) \geq 0$ và $c_{sub}(h) = 0$ khi và chỉ khi $h = 0$, $c_{sub}(h)$ tăng đơn điệu và nó là lõm trên các số nguyên [Bookstein et al. 2001].

2.2.3. Khoảng cách Levenshtein(chỉnh sửa)

Khoảng cách chỉnh sửa phổ biến nhất là khoảng cách Levenshtein [Levenshtein, 1966; Wagner và Fisher, 1974], thể hiện một sự tương đồng nội vùng giữa các chuỗi có độ dài tùy ý. Nó được dựa trên ba phép toán sửa: chèn, xóa và thay thế. Các giá trị c_{ins} , c_{del} và c_{sub} tương ứng với mỗi phép toán trong số đó, dẫn đến một phiên bản trọng lượng của khoảng cách này. Trong khoảng cách chỉnh sửa, $c_{sub} > c_{del} + c_{ins}$, có nghĩa là xóa của a và chèn của b là sự thay thế của a cho b. Nếu tất cả các giá trị như vậy là một biến đơn không lớn hơn tổng của hai giá trị khác, sau đó d_L là độ đo metric [Bunke et al., 2002]. Tương tự như d_{fHam} , trọng lượng khoảng cách Levenshtein d_L được xác định bằng tổng chi phí tối thiểu liên quan đến hoạt động chuyển đổi một chuỗi s thành t. (Lưu ý rằng các giải pháp có thể không được duy nhất). Giả định rằng một chuyển đổi như vậy đòi hỏi phải thay thế n_{sub} , chèn n_{ins} và phép xóa n_{del} , d_L được thể hiện như sau:

$$d_L(s, t) = \min_{n_{sub}, n_{ins}, n_{del}} (n_{sub} c_{sub} + n_{ins} c_{ins} + n_{del} c_{del}). \quad (2.26)$$

Khoảng cách soạn thảo truyền thống với tất cả các giá trị tương đương thường

được xét đến. Tuy nhiên vấn đề chính là d_L phụ thuộc vào độ dài của chuỗi so sánh và có thể bị ảnh hưởng nhiều bằng cách so sánh hai chuỗi, trong đó một là ngắn và chuỗi khác là rất dài. Để làm cho nó độc lập với độ dài, một chuẩn có thể được sử dụng, năng suất các khoảng cách Levenshtein chuẩn [Marzal và Vidal, 1993; Vidal và cộng sự; năm 1995.]:

$$d_{nL}(s, t) = \frac{d_L(s, t)}{\max\{n, m\}}. \quad (2.27)$$

Tuy nhiên, kể từ khi bất đẳng thức tam giác không đúng, d_{nL} là nửa metric.

2.2.4. Khoảng cách liên quan khác

Hai dãy có thể được so sánh dựa trên tiền tố chung dài nhất, hậu tố hay chỉ là một dãy. Giả sử chúng ta được cho hai chuỗi s và t của chiều dài n và $m \leq n$, tương ứng. Sau đó, khoảng cách giữa chúng có thể được định nghĩa là $d(s, t) = m + n - 2 | \text{common}(s, t) |$. Vấn đề của việc tìm kiếm dãy chung dài nhất là bổ sung cho việc xác định khoảng cách soạn thảo. Điều đó cũng có thể được giải quyết bằng việc sử dụng các quy hoạch động, xem thêm tài liệu [Stephen, 1998]. Bài tổng quan về đối sánh chuỗi có thể được tìm thấy trong [Navarro, 2001].

2.2.5. Khoảng cách thông tin và xấp xỉ thông tin

Giả sử một tập hợp các xâu nhị phân. Độ tính phức tạp Kolmogorov $K(s)$ của một chuỗi nhị phân s là chiều dài (theo bit) của các chương trình máy tính nhanh nhất của một tham chiếu cố định hệ thống máy tính sản xuất s như là một kết quả. Sự thay đổi của một hệ thống máy tính thay đổi giá trị này bằng một chất phụ gia cố định liên tục xem trong [Li và Vitányi, 1997]. Một cách giải thích có thể có của $K(s)$ là chiều dài của phiên bản nén cuối cùng của s mà s có thể được phục hồi bằng một chương trình giải nén. Để đo sự khác biệt giữa hai xâu, s và t , khoảng cách thông tin chuẩn đã được đề xuất trong [Li và cộng sự, 2003.]:

$$\text{NID}(s, t) = \frac{K(s, t) - \min\{K(s), K(t)\}}{\max\{K(s), K(t)\}}. \quad (2.28)$$

Lưu ý rằng $K(s, t)$ là độ dài của chương trình ngắn nhất mà bản in S và t mô tả làm thế nào để phân biệt chúng. Từ khoảng cách NID là không thể tính được, một xấp xỉ được

đề nghị sử dụng chương trình nén dữ liệu để ước tính K. Điều này dẫn đến khoảng cách nén chuẩn được định nghĩa là [Cilibrasi và VitBnyi, 2004]:

$$\text{NCD}(s, t) = \frac{C(s, t) - \min\{C(s), C(t)\}}{\max\{C(s), C(t)\}}, \quad (2.29)$$

trong đó C là chương trình nén lựa chọn và C(s) là chiều dài của chuỗi nén. Bất kỳ chuỗi (sau khi mã hóa lại thích hợp để các chuỗi nhị phân) có thể được so sánh bằng khoảng cách này, chẳng hạn như các tập tin nhị phân như phân âm nhạc ở định dạng MIDI (Musical Instrument Digital Interface: Giao Diện Số Hoá Nhạc Cụ)

2.3. Độ đo theo lý thuyết thông tin

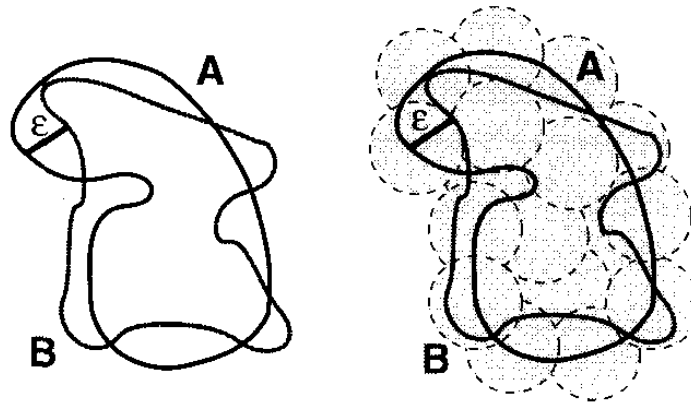
Theo ý nghĩa lý thuyết thông tin, một định nghĩa phổ biến của sự tương đồng, áp dụng đối với các lĩnh vực trong đó có một mô hình xác suất, được đề xuất bởi [Lin, 1998]. Nó được dựa trên quan sát chung rằng sự tương đồng giữa hai đối tượng là kết nối phổ biến và chúng khác nhau và hai đối tượng đồng nhất dẫn đến sự giống nhau tối đa. Điều này dẫn đến các giả định sau [Lin, 1998]:

- (1) Sự tương đồng giữa A và B được đo bằng $I(\text{com}(A, B))$, trong đó I là số lượng thông tin, thường là logarit tiêu cực xác suất của các sự kiện nó đề cập đến.
- (2) Sự khác biệt giữa A và B được đo bằng $I(\text{desc}(A, B)) - I(\text{com}(A, B)) \geq 0$, trong đó $\text{desc}(A, B)$ là một đề xuất mô tả A và B.
- (3) Sự tương đồng là một hàm $f: \mathbb{R}_+^0 \times \mathbb{R}_+ \rightarrow [0, 1]$ của tương đồng và sự khác biệt cho là $\text{sim}(A, B) = f(I(\text{com}(A, B)), I(\text{desc}(A, B)))$, chẳng hạn $f(x, x) = 1$ và $f(0, y) = 0$.
- (4) Sự tương tự chung của hai đối tượng là trung bình có trọng số được tính tương tự từ những quan điểm khác nhau.

Sự giống nhau bắt nguồn từ những giả định được tính như tỷ lệ giữa số lượng thông tin cần thiết tới tính phổ biến của hai đối tượng và số lượng thông tin cần thiết để mô tả chúng. Nó được cho là $\text{sim}(A, B) = \log P(\text{com}(A, B)) / \log P(\text{desc}(A, B))$. [Lin, 1998] trình bày định nghĩa chung này được áp dụng cho một số lĩnh vực, kết quả là một sự tương đồng giữa các chuỗi, từ hoặc khái niệm trong phân loại.

Khái quát và khoảng cách phổ biến độ đo metric đã được đề xuất trong [Bennett

và các cộng sự năm 1998; Li et al, 2003] và tiếp tục khám phá trong [Cilibrasi và Vitiinyi, 2005. Cilibrasi et al., 2004]. Như các tác giả khẳng định, độ đo metric của họ là tổng thể, nó có thể được áp dụng trong nhiều lĩnh vực như: âm nhạc, văn bản, hệ gen, thực thi chương trình hoặc mô tả ngôn ngữ tự nhiên và nó không tập trung vào đặc trưng đặc biệt hoặc tương đồng giữa các trường, nhưng phải mất tất cả chúng cùng một lúc vào tài khoản. Ý tưởng cơ bản là để thể hiện sự gần gũi của hai đối tượng nếu chúng có ý nghĩa 'nén' cho các thông tin khác. Điều này được chính thức hóa bởi quan điểm của Kolmogorov phức tạp. Trong thực tế, nó được xấp xỉ bằng khoảng cách nén chuẩn. Cùng một nguyên tắc là tiếp tục sử dụng để xác định một độ đo khoảng cách, Googlebased so sánh hai thuật ngữ tìm kiếm x và y như lập chỉ mục của Google [Cilibrasi và VitAnyi năm 2005; VitBnyi, 2005].



Hình 2.1: Minh họa của khoảng cách Hausdorff giữa 2 tập hợp A và B: $d_H(A, B) = \varepsilon$

2.4. Độ đo khoảng cách giữa các tập hợp

Sự khác nhau cũng có thể được xét giữa hai tiêu vùng đóng kín và bị chặn bởi một không gian (Euclide), tập hợp của các điểm hoặc các yếu tố. Đầu tiên xin giới thiệu khoảng cách Hausdorff [ghi chú Robinson, trang web; Klein và Thompson, 1984].

2.4.1. Khoảng cách Hausdorff

Cho (X, p) là một không gian độ đo metric và $C(X) \subseteq X$ là một không gian rỗng, tập con đóng kín và bị chặn của X . Cho $N_\varepsilon(A) = \bigcup_{x \in A} B_\varepsilon(x)$ với lớp bảo vệ của $A \in X$ bằng cách mở ε -hình cầu $B_\varepsilon(x) = \{y \in X: p(x, y) < \varepsilon\}$. Từ $B_\varepsilon(x)$ là lân cận của x , $N_\varepsilon(A)$ là lân cận của A . Khoảng cách Hausdorff giữa A và B được định nghĩa

là nhỏ nhất ε -lân cận của A bao gồm B và ngược lại, xem thêm hình 2.1. Mặt khác, hướng khoảng cách Hausdorff giữa A và B, $d_H^{\leftarrow}(A, B)$ có thể được thể hiện như tối đa thực hiện trên bộ sưu tập của khoảng cách tối thiểu giữa các phần tử của A và tập B. Sau đó, khoảng cách Hausdorff $d_H(A, B)$ là lớn nhất trong hai khoảng cách chỉ dẫn. Chính thức, chúng ta có:

+/ **Định nghĩa 2.1 (khoảng cách Hausdorff):** Trong không gian (nửa) metric (X, ρ) , khoảng cách Hausdorff với các cơ sở ρ được xác định cho tất cả $A, B \in C(X)$ theo những cách sau đây:

$$(1) \quad d_H(A, B) = \inf_{\varepsilon > 0} \{A \subset N_\varepsilon(B) \ \& \ B \subset N_\varepsilon(A)\}.$$

$$(2) \quad d_H(A, B) = \max \{d_H^{\leftarrow}(A, B), d_H^{\rightarrow}(B, A)\}, \text{ trong đó } d_H^{\leftarrow} \text{ là một khoảng cách Hausdorff định hướng}$$

$$d_H^{\leftarrow}(A, B) = \sup_{a \in A} \inf_{b \in B} \rho(a, b).$$

Nếu miền d_H^{\leftarrow} bị hạn chế, sau đó cận trên đúng trở thành tối đa và vô cùng trở thành tối thiểu, cụ thể là:

$$d_H^{\leftarrow}(A, B) = \max_{a \in A} \min_{b \in B} \rho(a, b).$$

+/ **Hệ quả 2.1:** Hai công thức khoảng cách Hausdorff được đưa ra trong định nghĩa 2.1 là tương đương.

+/**Chứng minh.** Chúng tôi bắt đầu từ định nghĩa (1) và biến đổi tương đương, việc xây dựng các định nghĩa (2) là đạt. $\inf_{\varepsilon > 0} \{A \subset N_\varepsilon(B)\} =$

$$\inf_{\varepsilon > 0} \{\forall a \in A \ a \in N_\varepsilon(B)\} = \inf_{\varepsilon > 0} \{\forall a \in A \ a \in \bigcup_{b \in B} (x : \rho(x, b) < \varepsilon)\} =$$

$$\inf_{\varepsilon > 0} \{\forall a \in A \ \inf_{b \in B} \rho(a, b) < \varepsilon\} = \{\sup_{a \in A} \inf_{b \in B} \rho(a, b)\} = d_H^{\leftarrow}(A, B).$$

Trên cơ sở này, chúng ta có:

$$d_H(A, B) = \inf_{\varepsilon > 0} \{B \subset N_\varepsilon(A) \ \& \ A \subset N_\varepsilon(B)\} =$$

$$\max \{\inf_{\varepsilon > 0} \{B \subset N_\varepsilon(A)\}, \inf_{\varepsilon > 0} \{A \subset N_\varepsilon(B)\}\} =$$

$$\max \{d_H^{\leftarrow}(A, B), d_H^{\leftarrow}(B, A)\}, \text{ được kết thúc chứng minh.}$$

+/**Định lý 2.1:** Nếu (X, ρ) là một độ đo metric (nửa metric) không gian, sau đó d_H là độ

đo metric (nửa metric).

+/**Chứng minh:** Đầu tiên, ta sẽ chứng minh rằng nếu p là nửa metric, sau đó d_H là nửa metric. Chúng ta sẽ sử dụng các công thức thứ hai trong định nghĩa 2.1. Từ đó cho tất cả $a \in A$, $\inf_{a \in A} \rho(a, a) = 0$, sau đó $d_H(A, A) = 0$. Hoạt động tối đa là đối xứng, vì vậy d_H là đối xứng. Cho $A, B, C \in \mathcal{C}(X)$. Cho $p(a, B) = \inf_{b \in B} \rho(a, b)$. Nếu $a \in A$, sau đó có tồn tại b như vậy

$\inf_{b \in B} \rho(a, b) \leq \sup_{a \in A} \rho(a, B) = d_H^p(A, B) \leq d_H(A, B)$. Cho b như vậy, chúng ta cũng có thể viết $\rho(b, C) = \inf_{c \in C} \rho(b, c) \leq d_H(B, C)$. Bằng cách áp dụng bất đẳng thức tam giác tới p , cho mỗi $a \in A$ ta có:

$\rho(a, C) \leq \rho(a, B) + \rho(B, C) \leq d_H(A, B) + d_H(B, C)$. Khi bất đẳng thức ở trên vẫn còn đúng cho tất cả $a \in A$, sau đó

$d_H^p(A, C) = \sup_{a \in A} \rho(a, C) \leq d_H(A, B) + d_H(B, C)$. Bởi vì sự sắp đặt của A và C là tùy ý, chúng ta cũng biết rằng $d_H^p(C, A) \leq d_H(A, B) + d_H(B, C)$. Do đó, $d_H(A, C) \leq d_H(A, B) + d_H(B, C)$.

Để chứng minh rằng d_H là độ đo metric nếu p là cơ sở cho $d_H(A, B) = 0$, sau đó $d_H^p(A, B) = d_H^p(B, A) = 0$. Do đó, đối với mỗi $a \in A$, $\inf_{b \in B} \rho(a, b) = 0$. Điều này có nghĩa rằng tất cả các lân cận của a chứa một phần tử từ B . Chúng ta biết rằng $a \in \bar{(B)} = B$, vì B là một tập đóng. Vì điều này giữ cho tất cả $a \in A$, thì $A \subset B$. Từ tính đối xứng hóa của định nghĩa, ta cũng nhận được $B \subset A$. Như vậy $A = B$.

Khoảng cách Hausdorff là bất biến đối với một chuyển đổi chỉ khi các số liệu cơ bản là bất biến. Do đó, tất cả phép đẳng cự trong các cơ sở độ đo metric là một đẳng cự trong các độ đo metric Hausdorff. Hơn nữa, hai tập hợp nằm trong khoảng cách Hausdorff d nếu bất kỳ điểm nào của một tập hợp nằm trong d khoảng cách từ một số điểm của các thiết lập khác. Một khoảng cách như vậy là nhạy cảm tới bên ngoài cá biệt. Ví dụ, suy nghĩ của một trường hợp trong đó điểm a là một khoảng cách lớn d_a , đến tất cả các điểm trong tập A . Sau đó, $d_H(A, B) = d_a$ được xác định bởi thời điểm này. Do đó, khái quát khoảng cách Hausdorff đã được xem xét, đó là có thể tốt hơn so với giá trị ngoại lai hoặc tiếng ồn.

2.4.2. Các biến thể của khoảng cách Hausdorff

Cho (X, ρ) là một độ đo không gian metric (thường là Euclide) và $C(X) \subseteq X$ là một không gian rỗng, các tập con đóng kín và bị chặn của X . Cho $A, B \in C(X)$ có các tập hợp n_A và các yếu tố n_B tương ứng. Khoảng cách giữa một yếu tố $a \in A$ và tập hợp B có thể được định nghĩa là:

$$d(a, B) = d(\{a\}, B) = \min_{b \in B} \rho(a, b). \quad (2.30)$$

Sự khác nhau có hướng giữa hai tập hợp sau đó có thể được tìm thấy như sau [Dubuisson and Jain, 1994]:

$$\begin{aligned} d_{\min}^{\triangleright}(A, B) &= \min_{a \in A} d(a, B), \\ d_{\max}^{\triangleright}(A, B) &= \max_{a \in A} d(a, B), \\ d_{avr}^{\triangleright}(A, B) &= \frac{1}{n_A} \sum_{a \in A} d(a, B), \\ d_{0.5}^{\triangleright}(A, B) &= M_{a \in A}^{0.5} d(a, B), \\ d_{0.75}^{\triangleright}(A, B) &= M_{a \in A}^{0.75} d(a, B), \\ d_{0.9}^{\triangleright}(A, B) &= M_{a \in A}^{0.9} d(a, B), \end{aligned} \quad (2.31)$$

trong đó $M_{a \in A}^s$ là khoảng cách được xếp hạng thứ k như vậy $k = s n_A$. Ví dụ, cho $s = 0,5$, $M_{a \in A}^{0.5}$ trở thành trung bình của dãy khoảng cách $d(x, Y)$ và cho $s = 0,75$, đây là tứ phân vị trên.

Vì giá trị $d^{\triangleright}(A, B)$ và $d^{\triangleright}(B, A)$ thường không giống nhau, đối xứng được áp dụng bằng cách áp dụng một trong các phép toán như sau:

$$f_{\min}(x, y) = \min\{x, y\}, \quad f_{\max}(x, y) = \max\{x, y\}, \quad f_{avr}(x, y) = \frac{1}{2}(x + y)$$

hay $f_{wavr}(x, y) = \frac{1}{n_A + n_B}(n_A x + n_B y)$. Khoảng cách Hausdorff (chỉ các độ đo metric), đã được giới thiệu trong định nghĩa 2.1, và khoảng cách Hausdorff sửa đổi là quan trọng với mục đích của đối tượng đối sánh trong hình ảnh nhị phân. Sau này, mặc dù không có độ đo metric, đã được tìm thấy hữu ích [Dubuisson và Jain, 1994] và có thể tốt hơn so với giá trị bên ngoài. Ngoài ra các biến thể khác thu được bằng cách thay thế các hoạt động tối đa trong các độ đo Hausdorff bởi một thứ hạng thứ k thường ít tiếng ồn nhạy cảm [Huttenlocher et al., 1993].

+/ **Định nghĩa 2.2 (sửa đổi Hausdorff)** Trong một (nửa) không gian độ đo metric (X, ρ) , khoảng cách Hausdorff được thay đổi với cơ sở ρ được xác định cho tất cả $A, B \in \mathcal{C}(X)$ như sau:

$$d_{MH}(A, B) = \max \{d_{avr}^\rho(A, B), d_{avr}^\rho(B, A)\},$$

$$d_{avr}^\rho(A, B) = \frac{1}{n_A} \sum_{a \in A} \min_{b \in B} \rho(a, b). \tag{2.32}$$

2.4.3. Các độ đo trên tập mờ

Một khoảng cách Hausdorff-cùng tên cũng có thể được xác định cho tập mờ, xem [Chaudhuri và Rosenfeld, 1996, 1999] để biết chi tiết. Xét hai tập hợp không mờ A_f và B_f trên một tập hợp hỗ trợ S trong một thước đo không gian. Cho $x^* = \max\{A_f(t) : A_f \in S\}$ là thành viên lớn nhất của x . Cho $A_{max} = \{t : A_f(t) = x^*\}$ là tập không mờ và cho A_a là một tập hợp không trống, không mờ của S , như vậy $A_{max} \subset A_a$, và cho hai tập mờ A_f và B_f , $A_a = B_a$ khi và chỉ khi $A_{max} = B_{max}$. Xác định các nhóm tập hợp không mờ $A_\mu, \mu \in [0,1]$ bởi:

$$A_\mu = \begin{cases} \{t : A_f(t) \in [\mu, x^*]\}, & \text{if } \mu \leq x^* \\ A_a, & \text{if } \mu > x^*. \end{cases} \tag{2.33}$$

Lưu ý rằng $A_\mu = A_{max}$ nếu $\mu = x^*$, cho $x^* \neq 1$.

Giả định rằng tập mờ có thể chỉ có giá trị từ một tập hợp rời rạc của các giá trị thành viên $\mu_1, \mu_2, \dots, \mu_c$. Cho $d_H(A_{\mu_i}, B_{\mu_i})$ là khoảng cách Hausdorff sắc nét giữa các bộ A_{μ_i} và B_{μ_i} . Sau đó, khoảng cách Hausdorff-cùng tên mờ giữa A_f và B_f được định nghĩa là:

$$d_{H_f}(A_f, B_f) = \frac{\sum_{i=1}^c \mu_i d_H(A_{\mu_i}, B_{\mu_i})}{\sum_{i=1}^c \mu_i}, \tag{2.34}$$

đó là độ đo metric [Chaudhuri và Rosenfeld, 1999]. d_{H_f} có thể được xem như một

thành viên trọng lượng trung bình của khoảng cách Hausdorff giữa mức độ biến đổi tập mờ đặt ra xem xét. Lưu ý khoảng cách được thay đổi Hausdorff-cùng tên mờ có thể được định nghĩa bằng cách sử dụng các d_{MH} thay vì d_H trong công thức trên.

2.5. Độ đo khoảng cách trong các ứng dụng

Có tồn tại một kho tàng lớn của các độ đo gần khác nhau được phát triển cho mục đích của tổ chức dữ liệu, hình ảnh và văn bản thu hồi, phân nhóm và phân loại. Trước khi trình bày một bản tóm tắt những độ đo, tôi muốn nhấn mạnh tầm quan trọng của bất biến trong quá trình thiết kế của họ.

2.5.1. Bất biến

Bất biến là một vấn đề quan trọng cho việc thiết kế các độ đo thông tin khác nhau. Để so sánh hai đối tượng, ta nên tập trung vào đặc điểm cơ bản của chúng. Điều này thì khó khăn để xử lý bằng các phương tiện tự động như các đối tượng thường khác nhau về hình thức và kích thước, phản ánh tất cả thay đổi. Nếu bắt đầu từ các phép đo cảm giác của các đối tượng, tiếng ồn có thể có mặt, cũng là góp phần vào sự thay đổi tổng thể. Nói chung, việc so sánh các đối tượng không nên bị ảnh hưởng bởi vị trí của chúng, một chút tỷ lệ khác nhau, hoặc phép quay. Hơn nữa, các độ đo nên mạnh hơn đối với các biến dạng nhỏ và quang sai của các phép đo các đối tượng. Điều này dẫn đến việc nghiên cứu các độ đo bất biến và mạnh hơn, đặc biệt là đối với các phép đo cảm giác. Mặc dù ở đây không triệt để thảo luận về vấn đề này, như tập trung vào phương pháp học tập, tôi muốn nhấn mạnh tầm quan trọng của nó. Các độ đo bất biến sẽ dẫn đến giới thiệu lớp học nhỏ, sẽ đảm bảo khả năng tổng quát tốt các chức năng học thống kê. Cho một đánh giá chung về nhận dạng mẫu bất biến, xem [Wood, 1996]. Nhiều phương pháp áp dụng có thể được tìm thấy trong [Rodrigues, 2001]. Ví dụ về các độ đo khoảng cách bất biến xem thêm [Hagedoorn và Veltkamp, 1999a, b; Simard và cộng sự, 1993, 1998.].

2.5.2. Ví dụ về độ đo

Ngoại trừ các độ đo đã được trình bày trong các phần trên. Một số ví dụ về độ đo khác nhau sẽ được trình bày trong một số lĩnh vực ứng dụng.

+/Loại dữ liệu đặc trưng

Biểu diễn cho dữ liệu trong không gian đặc trưng, một số các độ đo khoảng cách đã được thiết kế vào tài khoản phân phối cho các điểm trong khu vực lân cận. Khoảng

cách này sau đó được sử dụng bởi các k-NN quy định hoặc một số biến thể của học tập có trọng số cục bộ, xem [Atkeson et al. 1997] cho một cuộc khảo sát của các phương pháp.

Một số kỹ thuật để xây dựng độ đo metric linh hoạt được đề xuất trong [Friedman, 1994]. Những phương pháp này dựa trên một chiến lược phân vùng đệ quy để thích nghi và thu nhỏ hình chữ nhật xung quanh vùng lân cận các điểm kiểm tra. Như vậy Hastie và Tibshirani phát triển một quy tắc NN(Nearest Neighbor: láng giềng gần nhất) có khả năng thích nghi sử dụng thông tin phân biệt lân cận để chỉnh sửa các vùng lân cận thích hợp [Hastie và Tibshirani, 1996]. Khoảng cách độ đo metric là khoảng cách Euclide vuông có trọng số bằng một sản phẩm của trọng số thích hợp giữa và trong-tổng các hình vuông ma trận. Chúng cho thấy rằng độ đo metric xấp xỉ một khoảng cách chi bình phương giữa sự thật và ước tính xác suất sau cho các lớp Gaussian hình cầu. Khái quát cả hai phương pháp trước đây, một độ đo metric linh hoạt cho các lân cận tính toán dựa trực tiếp vào khoảng cách chi bình phương được ước tính trong [Dorneniconi et al. 2002]. Đặc tính của các vùng lân cận là chúng đang kéo dài cùng các đặc trưng ít thông tin nhỏ gọn và cùng những thuộc tính có ảnh hưởng nhất. Thật vậy Avesani và các đồng nghiệp đã đề xuất hai độ đo metric cho quy tắc NN: một khu vực không trọng số đối xứng độ đo metric tương tự và một độ đo rủi ro tối thiểu dựa trên một dự toán xác suất mà ước lượng khả năng giảm thiểu nguy cơ phân loại sai [Avesani et al 1999.]. Họ phát hiện bằng thực nghiệm rằng các quy tắc 1-NN dựa trên các độ đo của họ hoạt động tốt. Lowe giới thiệu một thiết bị biến hạt nhân dựa trên sự tương tự độ đo, bằng cách kết hợp các nguyên tắc k-NN với trọng số mịn xác định bởi các hạt nhân Gaussian [Lowe, 1995]. Những hạt nhân Gaussian được dựa trên một trọng số khoảng cách Euclide, trong đó các trọng số được học trong các thủ tục xác thực chéo. Trong các vấn đề phân loại, nguyên tắc láng giềng gần nhất (NN) thường được dựa trên (trọng số) khoảng cách Euclide. Tuy nhiên, các độ đo không giống nhau khác có thể được tính cho dữ liệu của các loại hỗn hợp.

Tất cả những phương pháp tiếp cận có thể được bao phủ bởi một khuôn khổ chung dựa trên sự tương đồng được tính giữa các đặc trưng, như đề xuất trong [Duch và cộng sự, 1998; Duch, năm 2000, Duch và cộng sự, 2000.]. Mô hình này bao gồm các bước lựa chọn đặc trưng đặc biệt, trọng số và mở rộng quy mô hợp lý, và tính toán một khoảng cách thích hợp cho các loại đặc trưng và các vấn đề ở bàn tay.

+/ Văn bản

Nhiều người trong số các mô hình tìm kiếm thông tin sử dụng các tính chất thống kê của văn bản [Manning và Schutze, 1999]. Đối với một bộ sưu tập các tài liệu văn bản, một bộ từ vựng thường được lựa chọn cho mục đích lập chỉ mục. Tài liệu văn bản sau đó được biểu diễn như là vector trọng lượng giới hạn cho tất cả các kỳ hạn từ bộ từ vựng. Trọng lượng giới hạn thường là tỷ lệ thuận với tần số xuất hiện trong tài liệu, và tỉ lệ nghịch thì số tài liệu giới hạn xảy ra nhập. Các độ đo giống nhau giữa các văn bản thường là một phiên bản thích hợp trọng số của một sự tương đồng cosin mà các độ đo cosin của góc giữa các vectơ tài liệu hoặc một lp-khoảng cách [Strehl et al., 2000]. Nhiều chương trình trọng số có thể được sử dụng, cũng như các độ đo nhị phân tập trung vào các từ xuất hiện được nhìn thấy trong biên bản của hội nghị SIGIR [SIGIR, trang web]. Ví dụ về các độ đo tương tự từ thống kê có thể được tìm thấy ví dụ trong Terra và Clarke [2003]. Một khả năng khác cũng là độ đo theo lý thuyết thông tin như mô tả trong [Lin, 1998; Bennett và cộng sự, 1998; Li và cộng sự, 2003.]. Khi bộ sưu tập tài liệu được mô tả bằng đồ thị, biểu đồ khác nhau không giống nhau liên quan đến việc tối đa đồ thị con chung [Bunke và Shearer, 1997, 1998] cũng như phép hợp đồ thị hoặc tối thiểu chung siêu đồ thị [Schenker et al., 2003], có thể được sử dụng.

+/ Hình dạng.

Trong tầm nhìn máy tính, xử lý ảnh và nhận dạng mẫu khu vực, nhiều kỹ thuật mô tả hình dạng đã được phát triển cho cả hai phép đo định lượng và định tính. Mô tả như vậy chủ yếu dựa hoặc phân khúc tiếp theo đặc điểm bên ngoài của kết quả hình dạng nhị phân được xác định bằng cách sắp đặt không gian của các yếu tố như cạnh các nút giao thông, hoặc trên các đặc điểm hình dạng nội bộ, như kết cấu hoặc cường độ đặc trưng, trong các hình ảnh màu xám cấp nhất định. Cho một giới thiệu chung các phương pháp mô tả hình dạng, xem [Costa và Cesar, 2001].

Ở đây, chúng ta quan tâm trong việc so sánh các đối tượng, do đó trong các độ đo tương tự. Nhiều độ đo như vậy tồn tại, cả hai nói chung và ứng dụng cụ thể, chủ yếu là phát triển để giải quyết các vấn đề mô hình phù hợp. Một ví dụ điển hình của một sự khác biệt theo định hướng mô hình kết hợp dựa trên việc tìm kiếm biến đổi hình học (từ một lớp nào đó) của một mô hình (hình dạng, đường viền, hình ảnh) vào nhau như vậy mà chi phí được xác định trước được giảm thiểu. Một cuộc khảo sát các cách tiếp cận phù hợp với hình dạng có thể được tìm thấy trong [Veltkamp và Hagedoorn, 1999],

trong khi một số độ đo tương tự và các thuật toán được mô tả trong [Veltkamp, 2001].

Với mục đích kết hợp các hình ảnh nhị phân (vì thế cũng đường nét), các biến thể của khoảng cách Hausdorff có thể được sử dụng. Cho một số cân nhắc thực tế, xem [Dubuisson và Jain, 1994; Huttenlocher et al, 1993.]. Kể từ khi các độ đo này trong các độ đo thực tế giữa các tập hợp của các điểm, một số phần mở rộng hơn nữa có thể được tìm thấy trong [Eiter và Mannila, 1997]. Biểu thức toán học cũng cho khoảng cách giữa các tập hợp điểm 2D với thư tín được biết đến đã được đề xuất trong [Werman và Weinshall, 1995]. Chúng là bất biến hoặc là phép biến đổi afin, chuyển thể giống nhau của các tập hợp. Đầu tiên, hình ảnh được chuẩn hóa và liên kết bằng cách sử dụng biến đổi afin, chẳng hạn như xoay, bản dịch và mở rộng quy mô. Tiếp theo, khoảng cách Euclide vuông giữa các điểm trong hình ảnh được tính toán. Kể từ khi hình ảnh được biểu diễn như là ma trận phối hợp, tất cả các biến đổi và khoảng cách có thể được thể hiện bằng ký hiệu ma trận.

Một độ đo tổng quát hơn khoảng cách metric, sự khác biệt tuyệt đối cái gọi là đã được giới thiệu trong [Hagedoorn và Veltkamp, 1999a]. Độ đo này là bất biến với phép biến đổi afin và những giao dịch tốt với các đối tượng có nhiều thành phần kết nối. Nó là mạnh mẽ chống lại nhiễu loạn và tắc.

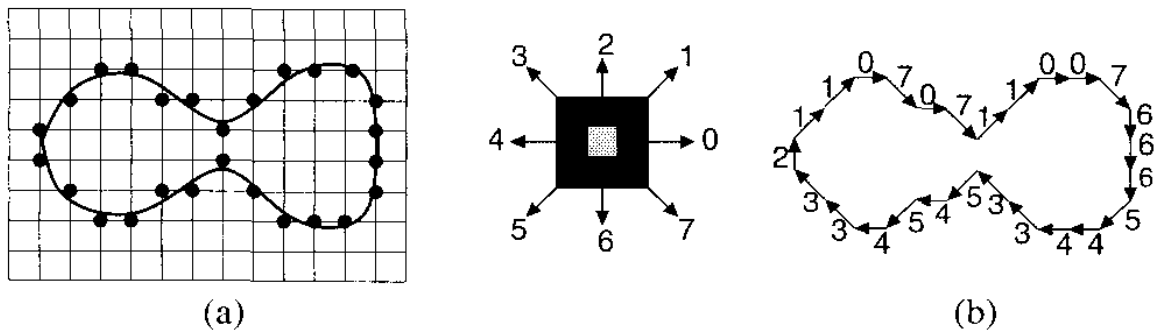
Phán đoán tương tự của con người đang cố gắng để được chụp trong [Basri và Jacobs, năm 1997; Basri và cộng sự, 1996, 1998.]. Ví dụ, trong [Basri et al., 1998], sự khác nhau giữa đường nét hình ảnh được nghiên cứu như là một chi phí phù hợp bằng cách tổng hợp các chi phí của biến dạng khu vực phản ánh sự khác nhau giữa hai đường nét. Một hàm chi phí được đề xuất mà phụ thuộc vào độ cong của khu vực và tuân theo những hạn chế về tính liên tục, tính số liệu và bất biến theo một số lớp học của biến đổi. Hàm chi phí cũng nên phát triển với sự gia tăng của uốn hoặc duỗi, nhưng uốn nên được ít tốn kém tại một điểm cong cao.

Một số ý kiến khác phù hợp với đường cong có thể được tìm thấy trong [Gdalyahu và Weinshall, 1999], định nghĩa của khoảng cách đàn hồi được xem xét trong [Younes, 1998, 1999] và sử dụng các mẫu biến dạng cho chữ số viết tay trong [Jain và Zongker, 1997].

Một mô tả hình dạng, bối cảnh hình dạng, cùng với một khuôn khổ cho phù hợp với biến dạng được phát triển trong [Belongie và Malik, 2000; Belongie và cộng sự, 2002;., Mori và cộng sự, 2001]. Bối cảnh hình thành tại một vị trí điểm cụ thể trên hình

dạng được xác định bởi các biểu đồ của tọa độ cực của tất cả các điểm khác. Từ điểm tương ứng của hai hình dạng khác nhau này có đặc điểm tương tự, sự liên kết của hình dạng được đơn giản hóa. Khoảng cách tổng thể được đưa ra như là trung bình trọng lượng của ba đóng góp: tổng của hình dạng tốt nhất phù hợp với chi phí, khoảng cách xuất hiện do sự khác biệt độ sáng và năng lượng uốn.

Một cách tiếp cận mới để sự liên kết giữa hai đường cong, dẫn đến nguồn gốc của đồng dạng của chúng được đề xuất trong [Sebastian và cộng ul., 2003].



Hình 2.2 Mã chuỗi biểu diễn: a: kết quả của việc lấy mẫu lại; b: mã chuỗi cơ sở

Theo báo cáo đó, phương pháp này là mạnh mẽ dưới nhiều biến đổi affine, cũng như các biến quan điểm và biến dạng nhỏ và nó có thể được áp dụng cho các vấn đề công nhận đối tượng. Thuật toán, các liên kết được giải quyết bằng cách lập trình động [Bellman, 1957].

Một khả năng so sánh hai hình dạng nhị phân là bằng cách sử dụng một chuyển đổi khoảng cách. Đây là hoạt động trên một hình ảnh nhị phân mà biến nó thành một hình ảnh màu xám cấp, một bản đồ khoảng cách, mà không đối tượng pixel có một giá trị tương ứng với khoảng cách đến đối tượng pixel gần nhất. Đối tượng không những có hình dạng, mà còn đường cong, cạnh hoặc điểm phù hợp dựa trên định vị hình mẫu tại các địa điểm khác nhau của khoảng cách bản đồ. Chi phí phù hợp, do đó sự khác nhau giữa các hình dạng đối tượng và các mẫu, được xác định bởi các giá trị điểm ảnh của bản đồ khoảng cách mà nằm dưới các điểm ảnh dữ liệu của các mẫu. Các mục tiêu được coi là phát hiện khi ví dụ giá trị khoảng cách trung bình là dưới một ngưỡng chọn. Khoảng cách phổ biến nhất là Euclidean, nhưng do chi phí tính toán của nó, khoảng cách thường vát cạnh, như xấp xỉ tốt nhất của nó, được sử dụng, xem [Borgefors, 1986]. Một ví dụ về hình dạng phù hợp với sử dụng vát cạnh khoảng cách biến đổi có thể được tìm thấy trong [Gavrila, 2000; Gavrila và Philomin, 1999]. Nó bao gồm việc

phát hiện các đối tượng tùy tiện hình, hoặc tham số hay không, như đường nét cho người đi bộ. Việc so sánh các bối cảnh và hình dạng vát cạnh các phương pháp phù hợp áp dụng cho các đối tượng phát hiện, nơi các đối tượng được mô tả bởi đường nét là thực hiện trong [Thayananthan et al., 2003]. Đây là báo cáo, mà trong trường hợp cảnh lộn xộn, kết hợp Chamfer dựa trên một số mẫu là trở nên tốt hơn so với phương pháp tiếp cận bối cảnh hình thành.

Hình dạng cũng có thể được mô tả một cách cơ cấu. Một mã chuỗi biểu diễn cho một ranh giới kỹ thuật số như là một chuỗi các vector hướng dựa trên 4 - hoặc 8-kết nối nguyên tắc [Freeman và Thủy tinh, 1961], xem thêm hình 2.2. Nói chung, nó không phải là duy nhất, vì nó phụ thuộc vào điểm khởi đầu. Tuy nhiên, với một điểm khởi đầu, nó xây dựng lại một hình dạng hoàn hảo. Không hay, mã chuỗi trở nên rất dài cho các đối tượng phức tạp, nhưng quan trọng hơn, chúng phản ánh tất cả hiện tại tiếng ồn trên ví dụ như rối loạn nhỏ. Tuy nhiên, để so sánh hai hình dạng, mã được xâu chuỗi có thể được so sánh. Từ điểm khởi đầu của chúng có thể tùy ý, phù hợp nên được thực hiện giữa tất cả các hoán vị theo chu kỳ của chúng. Cho $s = s_1 s_2 \dots s_n$ và $t = t_1 t_2 \dots t_n$ là chuỗi mã của hai đường viền. Cho S và T biểu diễn cho tập hợp của tất cả các hoán vị theo chu kỳ của s và t , tương ứng. Sau đó, so sánh hai chuỗi mã dựa trên khoảng cách Levenshtein trọng số như sau:

$$d_{\text{chain}}(s, t) = \min \{d^{\mathbb{D}}(s, t), d^{\mathbb{D}}(t, s)\}, \text{ trong đó}$$

$$d^{\mathbb{D}}(s, t) = \min_{s^* \in S, t^* \in T} d_{wL}(s^*, t^*) \text{ là một khoảng cách chỉ dẫn.}$$

Bằng cách này, d_{chain} là mạnh mẽ dựa vào phép quay của hình dạng, tuy nhiên, không dựa vào nhân rộng.

Ngoài ra, một đường viền có thể được biểu diễn như là một chuỗi các điểm $s = (x_1, y_1) \dots (x_m, y_m)$ trong một không gian hai chiều, lấy mẫu lại nếu cần thiết như vậy mà khoảng cách giữa bất kỳ cặp liên tiếp điểm giống hệt nhau. Sau đó, một chuỗi $z = z_1 \dots z_m$, mô tả một đường viền, có nguồn gốc như vậy mà z_i là vector chỉ hướng từ (x_i, y_i) đến (x_{i+1}, y_{i+1}) . Khoảng cách giữa các dãy là một khoảng cách chỉnh sửa với chèn cố định và chi phí xóa và một số chi phí thay thế. Chi phí thay thế khác nhau, ví dụ dựa trên một góc hoặc khoảng cách Euclide giữa vector tới các độ đo khoảng cách khác nhau, xem [Bunke và cộng sự, 2001, 2002.]. Đó là tuyên bố trong đó một cách tiếp cận như vậy có một số lợi thế như độ phân giải góc cao hơn [Buiike et al, 2002.]; Mạnh mẽ

để định hình biến dạng dưới luân chuyển và bất biến dưới chia tỷ lệ.

Fourier cũng mô tả cho đường nét kín được tìm thấy mà các độ đo khoảng cách có thể được xác định, chẳng hạn như khoảng cách từ lp-family [Zahn và Roskies, 1972; Persoon và Fu, 1974].

Mô tả cấu trúc của hình dạng hoàn thành dựa trên một mô tả thô của các mối quan hệ hình học giữa các bộ phận biên soạn chúng. Sự tương đồng giữa các hình khối có thể được đánh giá như một độ đo sửa khoảng cách giữa các đồ thị cú sôc biểu diễn cho các hình dạng như chủ trương của Kimia và các đồng nghiệp [Sharvit và cộng sự, 1998; Klein và Thompson, 1984; Sebastian và cộng ol, năm 2001, 2002]. Độ đo này được tính là chi phí tối ưu của con đường bị biến dạng giữa hai đường cong và nó là có sức mạnh chống lại sự biến dạng nhỏ, bị tắt và rối loạn ranh giới. Ngoài ra, sự so sánh giữa sự truy tìm dựa trên đồ thị sôc (cách tiếp cận cấu trúc) phù hợp với đường cong và (cách tiếp cận số liệu) được trình bày trong [Sebastian và Kiniia: 2001, 2003]. Một số phương pháp khác dựa trên các biểu diễn của các hình dạng của trục trung gian cũng có thể được tìm thấy trong [Liu và Geiger, 1999; Torsello và Hancock, 2003; Zhu và Yuille, 1996].

Cuối cùng, tính thống kê của hình dạng của đối tượng cũng có thể được sử dụng để so sánh. Điều này có nghĩa là thông tin hình dạng có thể được mã hóa bằng cách mô tả thời điểm, trong đó mô tả trung tâm của khối lượng lớn, các khía cạnh kéo dài và định hướng tổng thể. Ngoài ra, các đặc trưng cụ thể hơn đối với hình dạng tổng thể có thể được tìm thấy, chẳng hạn như: chu vi, diện tích, ranh giới thẳng, cong về số không-quá của độ cong xung quanh các đường viền hình dạng hoặc năng lượng uốn. Tất cả những đặc trưng chất lượng có thể được sử dụng để xây dựng một độ đo không đồng dạng như ví dụ được đưa ra trong Bảng 2.1.

+/ Biểu đồ và quang phổ.

Phổ phát xạ và phản xạ trở thành phổ biến hơn cho việc xác định các vật liệu nhất định, ví dụ như loại nhựa hoặc khoáng chất và đá. Tụ phát huỳnh quang cũng đang nổi lên như một công cụ hữu ích để phát hiện ung thư ví dụ trong khoang miệng hoặc trong phế quản. Nó dựa trên các quang phổ của các mô của lợi ích. Các phép đo thường thực hiện trên các mô khỏe và mô bệnh (trong giai đoạn khác nhau của ung thư) tại nhiều bước sóng kích thích. Phổ phát xạ này sau đó được phân tích để hỗ trợ chẩn đoán của bác sĩ.

Biểu đồ và quang phổ có thể được hiểu trong khuôn khổ xác suất, nơi mà phiên bản chuẩn của chúng được coi là phân bố xác suất. Điều này cho phép một người sử dụng các độ đo phân kỳ hoặc các độ đo chung giữa các phân. Vì cấu trúc của dữ liệu đó được tổ chức bởi các yếu tố cơ bản, chẳng hạn như thứ tự của các ngăn, bước sóng hoặc thời gian, nó có thể có lợi để kết hợp kiến thức thành các độ đo. Điều này phần nào có thể ví dụ bởi sự khác biệt máy tính, chẳng hạn như các lp-khoảng cách, giữa các dẫn xuất xấp xỉ của biểu đồ hoặc phổ [Paclik và Duin, 2003b, một; Pekalska et al., 2004a.]. Ví dụ, khoảng cách giữa các dẫn xuất đầu tiên để nhấn mạnh sự khác biệt ở các vị trí giữa các cực tiểu lân cận và cực đại của biểu đồ. Cũng là khoảng cách giữa các biểu đồ tích lũy có thể được sử dụng cũng như chúng ta sử dụng để so sánh các nhiệm vụ sắc thể trong hồ sơ băng thông [Pekalska và Duin, 2002a].

+/ Hình ảnh

Cho rằng hình ảnh màu xám cấp được biểu diễn như là vectơ trong một không gian. Một khoảng cách tiếp xúc, đó là bất biến tại lân cận với bất kỳ tập hợp các lựa chọn chuyển đổi (như xoay và tỉa thưa) và tương đối rõ để tính toán được đề xuất trong [Simard et al., 1993, 1998]. Nó đã được tìm thấy đặc biệt hiệu quả trong lĩnh vực nhận dạng chữ số viết tay (Simard et al., 1993]. Khi một hình ảnh được chuyển đổi (ví dụ như thu nhỏ và xoay) với phép biến đổi phụ thuộc vào một số thông số (như các yếu tố rộng và góc quay), các thiết lập của tất cả các mô hình chuyển đổi tạo ra một đa dạng của một kích thước tối đa bằng số lượng các thông số miễn phí trong vector không gian. Khoảng cách giữa hai mẫu hình ảnh có thể được định nghĩa là khoảng cách tối thiểu giữa các đa tạp và bất biến này liên quan đến việc biến đổi cân nhắc. Một khoảng cách như vậy là khó để tính toán. Sự thỏa hiệp được cung cấp bởi các khoảng cách tiếp xúc được định nghĩa là khoảng cách tối thiểu giữa các không gian con tiếp xúc mà tốt nhất xấp xỉ với đa tạp phi tuyến tính. Xem [Simard et al., 1993.] để biết chi tiết.

Vì hai hình ảnh có giá trị màu xám có thể được coi như tập mờ (do rescaling khoảng $[0, 1]$), để so sánh khoảng cách Hausdorff mờ (hoặc sửa đổi-Hausdorff) có thể được sử dụng. Còn hình ảnh nhị phân có thể được coi là tập mờ theo cách sau đây: điểm ảnh màu trắng có giá trị không thành viên và một điểm ảnh màu đen có giá trị là $k / (K - 1)$ nếu nó có k lân cận đen trong khu $K \times K$ của nó. Bằng cách này, điểm ảnh đen có nhiều sẽ có một trong hai giá trị thành viên không hoặc rất nhỏ. Nếu những hình ảnh nhị phân được chuyển đổi thành các tập mờ như mô tả, [Chaudhuri và Rosenfeld,

1999] báo cáo rằng những tiếng ồn có ít nhiều ảnh hưởng đến khoảng cách Hausdorff mờ hơn trên khoảng cách Hausdorff ban đầu giữa các hình ảnh nhị phân. Do đó, khoảng cách Hausdorff mờ là tương đối mạnh với tiếng ồn.

Mặt khác, hình ảnh có giá trị màu xám có thể được giải thích theo quan điểm xác suất ví dụ như như biểu đồ hai chiều. Điều này cho phép một người sử dụng các độ đo phân kỳ khác nhau hoặc các độ đo chung giữa các phân phối. Từ các giá trị cường độ của hình ảnh có thể khác nhau, một số chuẩn có thể là rất quan trọng.

Các mô tả về hình ảnh cũng có thể được đơn giản hóa để đơn biến biểu đồ, Ví dụ biểu đồ cường độ. Sau đó, khoảng cách giữa hai hình ảnh A và B có thể được tính như dựa trên giao điểm giữa hai biểu đồ cường độ với b bin là :

$$d_I(A, B) = 1 - \frac{\sum_{i=0}^{b-1} \min(h_i(A), h_i(B))}{\# \text{ pixels}}$$
, trong đó $h_i(A)$ mô tả số lượng điểm ảnh có cường độ tương đương với giá trị giao trong bin thứ i. Lưu ý rằng các giao lộ là ước tính của lỗi Bayes, tức là sự chồng chéo giữa hai khả năng mật độ chức năng $P(A)$ và $P(B)$ xấp xỉ bằng biểu đồ. Một phần mở rộng của độ đo này được đề xuất trong [Cha và Srihari, 2000], có tính đến sự giống nhau của cả hai chồng chéo, không chồng chéo phần.

Có tồn tại một số độ đo không đồng dạng để hỗ trợ việc thu hồi hình ảnh dựa trên nội dung. Đối với một bản tóm tắt ngắn gọn, xem ví dụ [Vasconcelos và Kunt, 2001]. Trong khuôn khổ xác suất, thường là độ đo được xác định giữa các bản phân phối, chẳng hạn như sự phân kỳ Kullback-Leibler, khoảng cách Bhattacharyya, hoặc khoảng cách Mahalanobis được sử dụng. Một phân tích ngắn gọn về mối tương quan của chúng được báo cáo trong [Vasconcelos và Lippman, 2000]. Ngoài ra, khoảng cách các động lực của trái đất [Rubner et al., 1998b] được thiết kế để đánh giá một sự khác biệt giữa hai bản phân phối dựa trên cái gọi là mặt đất đo khoảng cách giữa các đặc trưng duy nhất. Nói chưa chính xác, một phân phối có thể được hiểu như là một khối lượng của trái đất trải rộng trong không gian, trong khi việc phân phối khác như một tập hợp các lỗ hổng trong cùng một không gian. Sau đó, khoảng cách các động lực của trái đất xác định số tiền ít nhất công việc cần thiết để điền vào các lỗ với trái đất. Tính toán khoảng cách này được dựa trên một giải pháp cho vấn đề giao thông vận tải [Rubner et al., 1998b]. Độ đo này được áp dụng thành công cho một đánh giá về kết cấu tương đồng màu và trong hình ảnh [Rubner và cộng sự, 1998b, một; Rnbrier,

1999]. Nó có, tuy nhiên, một giải thích xác suất nghiêm ngặt, như trong [cô Levina và Bickel, 2001].

Trong khuôn khổ xác suất, cũng Puzicha và các đồng nghiệp thực nghiệm nghiên cứu một số độ đo không đồng dạng với mục đích phân đoạn kết cấu và hình ảnh thu hồi [Puzicha et al, 1997.] và cho màu sắc và kết cấu [Puzicha et một l., 1999a]. Trong cả hai giấy tờ, hình ảnh được so sánh bằng phân phối dựa trên các độ đo không đồng dạng, hệ số Gabor trong những hình ảnh được lọc trong bài báo đầu tiên, và giữa biểu đồ ở sau.

Một cách tiếp cận kết hợp đánh giá tương tự con người trong các độ đo không đồng dạng dựa trên các phần mở rộng của mô hình của Tversky bởi logic mờ như trình bày trong [Santini một, thứ Jain, 1997, 1996, 1999].

CHƯƠNG III. CHƯƠNG TRÌNH THỬ NGHIỆM

3.1. Bài toán:

Đầu vào là một ảnh bất kỳ

Đầu ra là danh sách các ảnh đã được sắp xếp theo độ đo khoảng cách

3.2. Phương pháp thực hiện

Để minh họa các phép toán : Độ đo Euclide và Độ đo Divergence, em sử dụng phần mềm Matlab 7.7 để thiết kế chương trình.

a) Thiết kế giao diện:

+/ Cửa sổ axes: nơi hiển thị ảnh truy vấn

+/ Nút Mở ảnh (button): Mở ảnh cần truy vấn

+/ Listbox: nơi hiển thị danh sách ảnh

+/ Nút Mở tập ảnh (button): Mở danh sách ảnh

+/ Nút Do đo Euclide (button): Tính khoảng cách giữa ảnh truy vấn và ảnh trong danh sách theo độ đo Euclide.

+/ Nút Do đo Divergence (button): Tính khoảng cách giữa ảnh truy vấn và các ảnh trong danh sách theo độ đo Divergence

b) Các bước thực hiện :

- Đầu tiên ta hiển thị một ảnh bất kỳ lên cửa sổ axes (ảnh truy vấn)

- Tiếp theo ta hiển thị một danh sách ảnh lên Listbox

- Khi kích vào button : Do đo Euclid hoặc Do đo Divergence thì các kết quả tính toán sẽ được hiển thị lên các axes kế bên.

3.3. Kết quả

- Giao diện chương trình:



Hình 3.1. Giao diện chương trình



Hình 3.2. Hình ảnh kết quả của độ đo Euclide



Hình 3.3. Hình ảnh kết quả của độ đo Divergence

KẾT LUẬN

❖ **Đánh giá**

Bài đồ án đã trình bày được một số độ đo khoảng cách: độ đo khoảng cách giữa các đối tượng, độ đo khoảng cách giữa các dãy, độ đo theo lý thuyết thông tin, độ đo khoảng cách giữa các tập hợp và độ đo khoảng cách trong các ứng dụng.

❖ **Những mục tiêu đã đạt được**

- Giới thiệu tổng quan về độ đo khoảng cách giữa các đối tượng dựa trên đặc trưng
- Áp dụng một số độ đo thông dụng vào bài toán độ đo khoảng cách

❖ Hướng nghiên cứu tiếp theo, em sẽ tiếp tục nghiên cứu và thử nghiệm với các độ đo khác để có thể đánh giá chúng với các độ đo đã trình bày và khả năng ứng dụng của chúng trong thực tiễn.

❖ Trong một khoảng thời gian có hạn, nên khi trình bày các vấn đề em đã nghiên cứu được không tránh khỏi những thiếu sót. Em rất mong nhận được những ý kiến đóng góp quý báu của thầy cô và các bạn.

Em xin chân thành cảm ơn!

TÀI LIỆU THAM KHẢO:

Tiếng Việt

- [1]. TS. Đỗ Năng Toàn, TS. Phạm Việt Bình (2007) - Giáo Trình Môn Học Xử Lý Ảnh, trường ĐH thái nguyên, khoa CNTT
- [2]. PGS. Nguyễn Quang Hoan (2006)- Giáo Trình Xử Lý Ảnh, học viện công nghệ bưu chính viễn thông
- [3]. Lương Mạnh Bá, Nguyễn Thanh Thủy (2003)- Nhập Môn Xử Lý Ảnh Số, Nhà xuất bản Khoa học và Kỹ thuật

Tiếng Anh:

- [1]. Jones, K. S. and Willett, P. (1977). Reading in Information Retrieval. Morgan Kaufmann Pub. Inc.
- [2] The Dissimilarity Representation for Pattern Recognition Foundations And Applications – Elzbieta Pekalska – Robert P.W.Duin

Trang Web:

- <http://luanvan.net.vn>
- <http://www.toantin.org/forums/index.php/topic/8843-matlab>
- <http://wang.ist.psu.edu/docs/related>
- <http://www.ictu.edu.vn>