JOURNAL OF SCIENCE OF HNUE DOI: 10.18173/2354-1075.2015-0061
Educational Sci., 2015, Vol. 60, No. 7A, pp. 137-144
This paper is available online at http://stdb.hnue.edu.vn

MỘT CHỨNG MINH HÌNH THỨC CHO PHÉP BÙ TRỪ PHỔ CMN CỦA TÍN HIỆU SỐ VÀ ỨNG DỤNG TRONG PHÂN VÙNG ẢNH VIỄN THÁM

Nguyễn Tu Trung, Ngô Hoàng Huy
Viện Công nghệ Thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

Tóm tắt. Phép chuẩn hoá CMN (Cepstral Mean Normalisation) từ lâu đã được sử dụng rộng rãi và hiệu quả trong xử lí tín hiệu số và nhận dạng tiếng nói. Tuy nhiên, khi áp dụng trong xử lí tín hiệu số thời gian thực, các tham số và tính đúng đắn của CMN được chọn và kiểm chứng thông qua thực nghiệm trên tín hiệu thực cụ thể mà thiếu các phép chứng minh hình thức chặt chẽ bằng toán học. Bài báo này đưa ra tính chất toán học của phép chuẩn hoá CMN và đưa ra chứng minh chặt chẽ của các tính chát trên bằng toán học. Dựa theo các tính chất trên, cho phép đưa ra cách chọn tham số CMN phù hợp thực tế. Đồng thời chỉ ra một ứng dụng của phép chuẩn hóa CMN trong phân cụm ảnh.
Từ khóa: Phân cụm, phân vùng ảnh, xử lí tín hiệu số, nhiễu, ảnh viễn thám, CMN, Kmeans, KMeansCMN, Quickbird, Landsat.

1. Mở đầu

Một vấn đề chung với các hệ thống xử lí tiếng nói là các đặc trưng của các kênh có thể biến đổi từ một phiên sang phiên tiếp theo. Một phương pháp được sử dụng để cự tiểu hóa ảnh hưởng của những khác biệt này trên hiệu năng nhận dạng là phép chuẩn hóa trung bình phổ (Cepstral Mean Normalisation - CMN).

Phương pháp này được áp dụng rộng rãi và hiệu quả trong xử lí tín hiệu số và nhận dạng tiếng nói. Tuy nhiên, khi áp dụng trong xử lí tín hiệu số thời gian thực, các tham số và tính đúng đắn của CMN được chọn và kiểm chứng thông qua thực nghiệm trên tín hiệu thực cụ thể mà thiếu các phép chứng minh hình thức chặt chẽ bằng toán học. Bài báo này đưa ra tính chất toán học của phép chuẩn hoá CMN và đưa ra chứng minh chặt chẽ của các tính chát trên bằng toán học. Ngoài ra, bài báo còn chỉ ra một ứng dụng của phép chuẩn hóa CMN trong phân cụm ảnh viễn thám.

2. Nội dung nghiên cứu

2.1. Chứng minh phép chuẩn hóa trung bình phổ

Cho trước $\{x_n\}_{n=1}^{\infty}$ là dãy vector số có số chiều hữu hạn, xác định dãy vector $\{y_n\}_{n=1}^{\infty}$ như sau: $y_1 = \alpha y_0 + \beta x_1, y_n = \alpha y_{n-1} + \beta x_n, \forall n = 2, 3 \ldots, \alpha, \beta \in (0, 1), \alpha + \beta = 1, y_0 = 0$ hoặc

Ngày nhận bài: 15/7/2015. Ngày nhận đăng: 20/11/2015.
Liên hệ: Nguyễn Tu Trung, e-mail: nttrung@ioit.ac.vn.

được xác định trước.

Trong các ứng dụng xử lí tín hiệu số, tiếng nói hoặc dữ liệu ảnh thường các vectorx_n biến đổi xung quanh một giá trị trung bình (tổng quát là kiểu các biến ngẫu nhiên có cùng phân bố) sau khi phép tiền xử lí tín hiệu đã đi qua một phép phân cụm, phân loại tín hiệu (chẳng hạn phép phân loại tín hiệu nền/nhiễu/tiếng nói trong xử lí tiếng nói).

Mệnh đề 1: $\forall N > 1, n > N$

$$\left\| y_n - \frac{\sum\limits_{k=1}^{n} x_k}{n} \right\| \leq \alpha \left\| y_{n-1} - \frac{\sum\limits_{k=1}^{n-1} x_k}{n-1} \right\| + \frac{\alpha}{n} \frac{2\,N \max\limits_{1\leq k} \|x_k\| + (n-1-N) \max\limits_{N\leq k\leq n} \|x_n - x_k\|}{n-1}$$
$$+ \beta \frac{2\,N \max\limits_{1\leq k} \|x_k\| + (n-N) \max\limits_{N\leq k\leq n} \|x_n - x_k\|}{n} \tag{2.1}$$

Chứng minh: Do $\alpha + \beta = 1$ ta có,

$$y_n - \frac{\sum\limits_{k=1}^{n} x_k}{n} = \alpha \left(y_{n-1} - \frac{\sum\limits_{k=1}^{n-1} x_k}{n-1} \right) + \frac{\alpha}{n} \left(\frac{\sum\limits_{k=1}^{n-1} x_k}{n-1} - x_n \right) + \beta \left(x_n - \frac{\sum\limits_{k=1}^{n} x_k}{n} \right) \tag{2.2}$$

Từ đó suy ra ước lượng trên.

Từ ước lượng này nên trong thực hành thường chọn β rất gần 0.

Mệnh đề 2: $\forall N > 1, n > N$.

$$\|y_{n+N} - y_{2N}\| \leq \alpha^N \|y_n - y_N\| + \max_{N+1\leq l\leq n+N} \|x_{n+l-N} - x_l\| \tag{2.3}$$

Chứng minh: $y_m - y_n = \alpha\,(y_{m-1} - y_{n-1}) + \beta\,(x_m - x_n)$

Suy ra

$$\|y_m - y_n\| \leq \alpha \|y_{m-1} - y_{n-1}\| + \beta \|x_m - x_n\| \tag{2.4}$$

Tương tự

$$\|y_{m-1} - y_{n-1}\| \leq \alpha \|y_{m-2} - y_{n-2}\| + \beta \|x_{m-1} - x_{n-1}\| \tag{2.5}$$

$$\|y_m - y_n\| \leq \alpha^2 \|y_{m-2} - y_{n-2}\| + \beta\,(\alpha \|x_{m-1} - x_{n-1}\| + \|x_m - x_n\|) \tag{2.6}$$

Bằng quy nạp ta có

$$\|y_m - y_n\| \leq \alpha^N \|y_{m-N} - y_{n-N}\| + \beta \sum_{k=0}^{N-1} \alpha^k \|x_{m-k} - x_{n-k}\| \tag{2.7}$$

Suy ra

$$\begin{aligned}\|y_{n+N} - y_{2N}\| &\leq \alpha^N \|y_n - y_N\| + \beta \sum_{k=0}^{N-1} \alpha^k \|x_{n+N-k} - x_{2N-k}\| \\ &\leq \alpha^N \|y_n - y_N\| + \beta \max_{N+1 \leq l \leq n+N} \|x_{n+l-N} - x_l\| \sum_{k=0}^{N-1} \alpha^k \end{aligned} \tag{2.8}$$

Do

$$\beta \sum_{k=0}^{N-1} \alpha^k \leq \beta \frac{1}{1-\alpha} = 1 \tag{2.9}$$

Nên

$$\|y_{n+N} - y_{2N}\| \leq \alpha^N \|y_n - y_N\| + \max_{N+1 \leq l \leq n+N} \|x_{n+l-N} - x_l\| \tag{2.10}$$

Mệnh đề 3: $\{x_n - y_n\}_{n=1}^{\infty}$ là dãy có tổng trung bình các phần tử xấp xỉ 0 tại mọi thời điểm n.

Chứng minh:

$$x_n - y_n = \alpha (x_n - x_{n-1}) + \alpha (x_{n-1} - y_{n-1}) \tag{2.11}$$

$$\sum_{k=2}^{n} x_k - y_k = \alpha (x_n - x_1) + \alpha \sum_{k=1}^{n-1} x_k - y_k \tag{2.12}$$

$$\sum_{k=1}^{n} x_k - y_k = \alpha (x_n - x_1) + \alpha \sum_{k=1}^{n} x_k - y_k + (x_1 - y_1) - \alpha (x_n - y_n) \tag{2.13}$$

$$\sum_{k=1}^{n} x_k - y_k = x_1 + \frac{\alpha y_n - y_1}{1-\alpha} \tag{2.14}$$

Do các giá trị y_n bị chặn,

$$\lim_{n \to \infty} \frac{\sum_{k=1}^{n} x_k - y_k}{n} = 0 \tag{2.15}$$

Nhận xét: Với tín hiệu tiếng nói, thường x_1, y_1 xấp xỉ vector 0, nên

$$\sum_{k=1}^{n} x_k - y_k \approx \frac{\alpha y_n}{1-\alpha} \tag{2.16}$$

2.2. Thuật toán phân cụm Kmeans

Thuật toán KMeans [3] bao gồm 4 bước, được trình bày như sau:

Bảng 1. Thuật toán KMeans cơ bản

Đầu vào: n đối tượng và số cụm k
Đầu ra: Các cụm C_i $(i = 1..k)$ sao cho hàm mục tiêu E sau đây đạt cực tiểu:

$$E = \sum_{i=1}^{k} \sum_{x \in C_i} d^2 \tag{2.17}$$

Bước 1: Khởi tạo
Chọn k đối tượng C_j $(j = 1..k)$ là tâm ban đầu của k cụm dữ liệu đầu vào (lựa chọn ngẫu nhiên hoặc theo kinh nghiệm).
Bước 2: Gán tâm cụm theo khoảng cách
Với mỗi đối tượng x_i $(1 \leq i \leq n)$, tính khoảng cách của nó tới mỗi tâm C_j với $j = 1..k$. Đối tượng thuộc về cụm C_S mà khoảng cách từ tâm C_S tương ứng đến đối tượng đó là nhỏ nhất.

$$d(x, C_s) = \min d(c, C_j), 1 \leq \tag{2.18}$$

Bước 3: Cập nhật tâm cụm
Đối với mỗi $j = 1..k$, cập nhật lại tâm cụm C_j bằng cách xác định trung bình cộng của các vector đối tượng dữ liệu đã được gán về cụm.

$$C_j = \frac{\sum_{x \in clust}}{count}(clust \tag{2.19}$$

Bước 4: Lặp và kiểm tra điều kiện dừng
Lặp lại các bước 2 và 3 cho đến khi các tâm cụm không thay đổi giữa hai lần lặp liên tiếp.

2.3. Thuật toán phân cụm Kmeans cải tiến

Với công thức tính tâm như (2.13), tâm thu được dễ bị ảnh hưởng bởi nhiễu. Chúng ta có thể áp dụng phép chuẩn hóa trung bình phổ để cho một công thức tính tâm cụm có thể giảm nhiễu. Tuy nhiên, phép chuẩn hóa này chỉ tốt khi số phần tử là rất lớn. Trong phần này, chúng tôi đề xuất thuật toán phân cụm KMeansCMN cải tiến cho ảnh viễn thám kích thước lớn với công thức tính tâm cụm áp dụng kĩ thuật chuẩn hóa trung bình phổ như sau:

Bảng 2. Thuật toán KMeanCMN

Đầu vào: n đối tượng và số cụm k
Đầu ra: Các cụm C_i $(i = 1..k)$ sao cho hàm mục tiêu E sau đây đạt cực tiểu:

$$E = \sum_{i=1}^{k} \sum_{x \in C_i} d^2 \tag{2.20}$$

Bước 1: Khởi tạo
Chọn k đối tượng C_j $(j = 1..k)$ là tâm ban đầu của k cụm dữ liệu đầu vào (lựa chọn ngẫu nhiên hoặc theo kinh nghiệm).
Bước 2: Gán tâm cụm theo khoảng cách
Với mỗi đối tượng x_i $(1 \leq i \leq n)$, tính khoảng cách của nó tới mỗi tâm C_j với $j = 1..k$. Đối tượng thuộc về cụm C_S mà khoảng cách từ tâm C_S tương ứng đến đối tượng đó là nhỏ nhất.

$$d(x, C_s) = \min d(x, C_j), 1 \leq \tag{2.21}$$

Bước 3: Cập nhật tâm cụm
Đối với mỗi $j = 1..k$, cập nhật lại tâm cụm C_j bằng cách xác định trung bình cộng của các vector đối tượng dữ liệu đã được gán về cụm.
- Nếu số lượng điểm ảnh trong cụm nhỏ hơn hằng số rất lớn Max thì tâm vẫn tính theo công thức (2.13) như sau:

$$C_i = \frac{\sum_{x \in cluste}}{count(clus} \tag{2.22}$$

- Nếu số lượng điểm ảnh trong cụm lớn hơn hằng số rất lớn Max thì tâm tính theo công thức (2.17) như sau:

$$C_j = CMN(Clu \tag{2.23}$$

Bước 4: Lặp và kiểm tra điều kiện dừng
Lặp lại các bước 2 và 3 cho đến khi các tâm cụm không thay đổi giữa hai lần lặp liên tiếp.

Thủ tục tính tâm cụm CMN(Cluster$_j$) tại vòng lặp thứ n như sau:

Bước 1: Khởi tạo tâm theo công thức

$$C_j^n = \beta \tag{2.24}$$

Bước 2: Với mỗi $x \in$ ..... tính theo công thức

$$C_j^n = \alpha C_j^{n-1} \tag{2.25}$$

Trong nghiên cứu này, chúng tôi chọn $\max = 50000$ và $= 0.95$.

2.4. Thực nghiệm

Chúng tôi tiến hành thử nghiệm thuật toán đề xuất KMeansCMN và so sánh kết quả với thuật toán Kmeans đã được sử dụng phổ biến cho phân vùng ảnh viễn thám. Giả sử ảnh đầu vào có kích thước M x N điểm ảnh.Chúng tôi thực hiện phân rã wavelet 3 mức. Như vậy, ảnh xấp xỉ cực tiểu chúng tôi chọn có kích thước M/8 x N/8 điểm ảnh.

Tập dữ liệu phục vụ cho thử nghiệm gồm hai loại. Một là, loại ảnh LANDSAT ETM+ chụp khu vực Hòa Bình ngày 15/02/2001, bao gồm 11 ảnh ranh giới từng huyện và một ảnh theo ranh giới tỉnh của tỉnh Hòa Bình. Hai là, loại ảnh Quickbird, gồm 4 kênh: Lam, Lục, Đỏ, và cận hồng ngoại, được tải từ dữ liệu mẫu trên trang http://opticks.org.

Trong thử nghiệm 1, ảnh gốc là ảnh vệ tinh Quickbird. Trong thử nghiệm 2, ảnh gốc là ảnh vệ tinh LANSAT về huyện Đà Bắc thuộc tỉnh Hoà Bình. Bảng 3 minh họa ảnh đầu vào trong các thử nghiệm 1 và 2.

Bảng 3. Các ảnh đầu vào trong tử nghiệm 1 và 2

Thử nghiệm 1	Thử nghiệm 2

2.4.1. Thử nghiệm 1

Dưới đây là ảnh kết quả phân cụm của KMeans và KMeansCMN trong trường hợp 5 cụm với ảnh Quickbird. Các ảnh từ 1 đến 5 là ảnh từng cụm. Ảnh thứ 6 là ảnh đã thay các điểm ảnh gốc bằng tâm các cụm.

Hình 1. Kết quả phân cụm bởi KMeans

Hình 2. Kết quả phân cụm bởi KMeansCMN

Bảng 4 thống kê tâm cụm thu được từ KMeans và KMeansCMN trong trường hợp 5 cụm. Bảng 5 thống kê số bước lặp cũng như thời gian thực thi của KMeans và KMeansCMN với 5 cụm và 8 cụm.

Bảng 4. Tâm cụm sinh từ KMeans và KMeansCMN

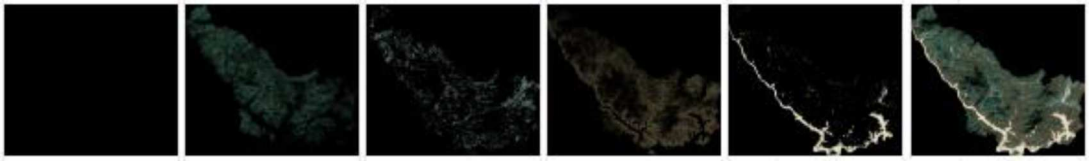
Cụm	KMeans	KMeansCMN
1	178, 170, 132	160, 147, 108
2	143, 134, 99	137, 123, 83
3	107, 103, 78	101, 91, 59
4	61, 62, 48	59, 55, 37
5	26, 31, 22	24, 27, 17

Bảng 5. Thời gian phân cụm

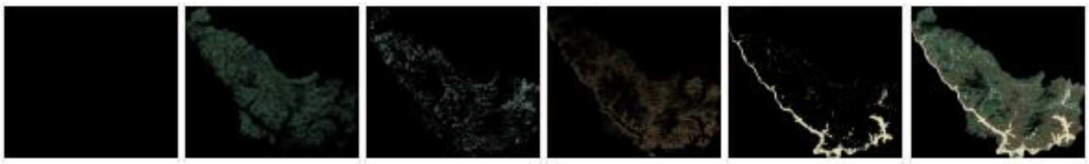
		KMeans	KMeansCMN
5 cụm	Thời gian (ms)	607,818	498,719
	Số vòng lặp	10	8
8 cụm	Thời gian (ms)	2,114,812	2,004,578
	Số vòng lặp	11	10

2.4.2. Thử nghiệm 2

Dưới đây là ảnh kết quả phân cụm của KMeans và KMeansCMN trong trường hợp 5 cụm với ảnh LANSAT. Các ảnh từ 1 đến 5 là ảnh từng cụm. Ảnh thứ 6 là ảnh đã thay các điểm ảnh gốc bằng tâm các cụm.



Hình 3. Kết quả phân cụm bởi KMeans



Hình 4. Kết quả phân cụm bởi KMeansCMN

Bảng 6. Tâm cụm sinh từ KMeans và KMeansCMN

Cụm	**KMeans**	**KMeansCMN**
1	0, 0, 0	0, 0, 0
2	43, 76, 62	50, 77, 59
3	100, 125, 118	101, 123, 113
4	79, 70, 50	85, 69, 47
5	225, 220, 192	225, 217, 183

Bảng 7. Thời gian phân cụm

		KMeans	KMeansCMN
5 cụm	Thời gian (ms)	263,672	213,109
	Số vòng lặp	10	8
8 cụm	Thời gian (ms)	1,658,609	1,568,062
	Số vòng lặp	25	23

Bảng 6 thống kê tâm cụm thu được từ KMeans và KMeansCMN trong trường hợp 5 cụm.

Bảng 7 thống kê số bước lặp cũng như thời gian thực thi của KMeans và KMeansCMN với 5 cụm và 8 cụm.

Nhận xét: Việc tính tâm theo công thức (2.17) chắc chắn lâu hơn công thức (2.13). Do đó, về mặt lí thuyết, nhiều khả năng thời gian phân cụm của KMeansCMN sẽ lâu hơn KMeans. Tuy nhiên, theo thống kê trong bảng 5 và 7, số vòng lặp và thời gian thực thi của KMeansCMN ít hơn. Nói cách khác, để thuật toán hội tụ, KMeansCMN cần số vòng lặp thực thi ít hơn dẫn tới thời gian thực thi được cải thiện.

3. Kết luận

Trong nghiên cứu này, chúng tôi đã đề xuất thuật toán KMeansCMN với mục tiêu áp dụng phương thức chuẩn hóa trung bình phổ để tính tâm cụm cho việc phân vùng ảnh viễn thám kích thước lớn. Các kết quả thử nghiệm cho thấy KMeansCMN phân cụm tốt với ảnh viễn thám kích thước lớn. Ngoài ra, tốc độ phân cụm của KMeansCMN là tốt hơn so với KMeans thông thường. Hiện tại, thủ tục tính tâm theo CMN vẫn sử dụng nhiều tính toán với số thưc nên tốc độ chậm. Trong nghiên cứu tiếp theo, nhóm tác giả dự kiến sử dụng phương pháp tính toán chấm tĩnh để tăng cường tốc độ thủ tục này để tăng tốc độ phân cụm.

TÀI LIỆU THAM KHẢO

[1] A.E. Hasanien, A. Badr, 2003. *A Comparative Study on Digital Mamography Enhancement Algorithms Based on Fuzzy Theory*. Studies in Informatics and Control, Vol.12, No.1, pp. 21-31.

[2] Chih-Tang Chang và cộng sự, 2011. *A Fuzzy K-means Clustering Algorithm Using Cluster Center Displacement*. Journal of Information Science and Engineering 27, pp. 995-1009.

[3] http://www.onmyphd.com/?p=k-means.clustering.

[4] S.G.Mallat, 1989. *A theory for multi resolution signal decomposition, the wavelet representation*. IEEE transactions on Pattern Analysis and machine Intelligence, 11(7): 674-693.

[5] T. Balaji, M. Sumathi, 2013. *Relational Features of Remote Sensing Image lassification using Effective K-Means Clustering*. International Journal of Advancements in Research & Technology, Vol. 2, Issue 8, pp. 103-107.

ABSTRACT

A formal proof of a cepstal mean normalisation of a digital signal and the application of the signal to segment remote sensing images

Cepstal Mean Normalisation has long been used extensively and effectively in digital signal processing and speech recognition. However, when applied in digital signal processing in real time, parameters and the rightness of CMN can be selected and verified through experiments on real signals without a strict formal mathematical proof. This paper presents mathematical properties of CMN and given strict proof on mathematically. Based on the above properties, give selecting parameter of CMN. It also indicates that CMN can be applied in a clustering image.

Keywords: Cepstral Mean Normalisation, CMN, KMeans, KMeansCMN, Quickbird, Landsat.