

## MỘT KỸ THUẬT PHÂN LỚP ẢNH VIỄN THÁM SỬ DỤNG TIẾP CẬN LAI

Nguyễn Tu Trung<sup>1</sup>, Ngô Hoàng Huy<sup>1</sup>, Đặng Văn Đức<sup>1</sup>, Vũ Văn Thỏa<sup>2</sup>, Lại Anh Khôi<sup>3</sup>

<sup>1</sup>Viện Công nghệ thông tin, Viện Khoa học và Công nghệ Việt Nam

<sup>2</sup>Học viện Công nghệ Bưu chính Viễn thông

<sup>3</sup>Viện Công nghệ Vũ trụ, Viện Khoa học và Công nghệ Việt Nam

**Tóm tắt:** Phân loại ảnh viễn thám là vấn đề được các nhà nghiên cứu viễn thám quan tâm. Có 2 hướng tiếp cận chủ yếu là dựa trên điểm ảnh và hướng đối tượng cho vấn đề này. Ảnh viễn thám có thể có nhiều kênh và độ phân giải rất cao. Hướng tiếp cận dựa trên điểm ảnh thường cho độ chính xác cao nhưng gặp vấn đề khi phân loại ảnh có kích thước lớn như ảnh viễn thám. Trong khi đó, tiếp cận dựa trên đối tượng khắc phục được vấn đề kích thước ảnh nhưng thường có độ chính xác thấp hơn so với hướng tiếp cận điểm ảnh. Bài báo này đề xuất một tiếp cận lai với sự kết hợp cả hai hướng tiếp cận để đưa ra một thuật toán phân loại ảnh viễn thám hiệu quả hơn. Thuật toán được thử nghiệm trên tập dữ liệu là các ảnh viễn thám thuộc tỉnh Hoà Bình.

**Từ khóa:** Viễn thám, phân loại ảnh, tiếp cận hướng điểm ảnh, tiếp cận hướng đối tượng, tiếp cận lai.

### I. MỞ ĐẦU

Độ phân giải của ảnh đa phổ ngày càng tăng. Gần đây, các ảnh viễn thám có độ phân giải nhiều mét. Hiện tại, khi mà các vệ tinh mới đã đạt được 60 centimet của độ phân giải thì mức độ chi tiết tăng lên 10 lần. Với các ảnh như vậy, chúng ta có thể cho rằng mỗi điểm ảnh là một phần của một đối tượng đơn giản. Do đó, tính hỗn tạp của ảnh tăng lên rõ rệt.

Các ảnh vệ tinh được sử dụng chủ yếu trong các hệ thống tin địa lý (GIS). Việc phân loại chúng rất

Tác giả liên hệ: Nguyễn Tu Trung

Email: trungnt.sremis@gmail.com

Đến tòa soạn: 23/7/2016, chỉnh sửa: 30/8/2016, chấp nhận đăng: 03/9/2016.

có ích cho khoa nghiên cứu bản đồ. Với các ảnh đa phổ có độ phân giải thấp, cường độ của các điểm ảnh đủ để phân loại riêng rẽ từng điểm của chúng. Ngược lại, việc phân loại ảnh có độ phân giải cao khó hơn rất nhiều. Việc tăng độ phức tạp của cảnh tạo ra các mức độ chi tiết khác nhau. Ví dụ một cây trong một cánh đồng hay các bóng râm của các đối tượng có thể nhìn thấy và thông tin ngữ cảnh của các điểm ảnh trở nên cần thiết cho một phân loại tốt. Các phần mềm phân loại GIS và phần mềm dùng trong lĩnh vực y học đang tồn tại nói chung đều sử dụng các phương pháp giống nhau cho các ảnh có độ phân giải thấp và cao. Nếu như các kết quả vừa ý có thể đạt được với các ảnh độ phân giải thấp, thì hiệu quả của các phần mềm này với các ảnh độ phân giải cao vẫn cần phải xem xét và nghiên cứu thêm. Chính vì vậy, để đảm bảo một sự chính xác tốt, việc phân loại bằng tay đôi khi được ưu tiên hơn các phương pháp tự động.

Trong phân loại ảnh viễn thám, có 2 hướng tiếp cận chủ yếu dựa trên điểm ảnh và hướng đối tượng cho vấn đề này. Hướng tiếp cận dựa trên điểm ảnh [9] thường cho độ chính xác cao nhưng gặp vấn đề khi phân loại ảnh có kích thước lớn như ảnh viễn thám. Hướng tiếp cận dựa trên đối tượng bao gồm hai giai đoạn chính [2]. Một là, xác định đối tượng là các vùng (cụm) sử dụng các thuật toán phân loại không giám sát (phân vùng, phân cụm). Phân cụm là một quy trình dùng để trích chọn những nét chính của các đối tượng nên bởi việc định nghĩa các vùng tương ứng. Có nhiều phương pháp phân vùng khác nhau như: Các phương pháp hình thái, Các phương pháp họ K-means, Mô hình pha trộn Gaussian có giới hạn (FGMM), Tách và hợp, Các mô hình Markov,... Hiện nay, một số thuật toán bao

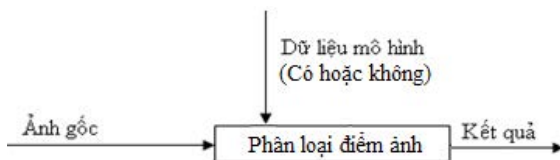
gồm thông tin ngữ cảnh trong quy trình để giảm bớt tính hỗn tạp của các phân đoạn [2]. Trong [4], Chen và cộng sự đã trình bày thuật toán phân cụm KMeans sử dụng thay thế tâm cụm. Trong [5], Balaji và cộng sự đã trình bày thuật toán phân cụm mới dựa trên việc chuyển đổi ảnh từ không gian màu RGB sang không gian  $L^*a^*b$  và phân cụm trên không gian này. Hai là, phân loại đối tượng sử dụng các thuật toán phân loại có giám sát. Trong [6], các tác giả đã sử dụng tiếp cận mạng Neural để phân lớp ảnh Landsat. Một trong những phương pháp phân loại được sử dụng phổ biến nhất trong ảnh viễn thám có thể kể đến là phương pháp phân loại hợp lý tối đa [1]. Đây là phương pháp phân loại dựa trên hướng tiếp cận điểm ảnh. Trong [7], Nedeljkovic đề xuất thuật toán phân lớp ảnh dựa trên logic mờ và thuật toán phân loại hợp lý tối đa. Hướng tiếp cận dựa trên đối tượng khắc phục được vấn đề kích thước ảnh nhưng thường có độ chính xác thấp hơn so với hướng điểm ảnh.

Trong nghiên cứu này, chúng tôi đề xuất tiếp cận phân loại ảnh viễn thám mới mà thuật toán cài đặt từ tiếp cận này được cải tiến từ phương pháp phân loại hợp lý tối đa với sự kết hợp cả hai hướng tiếp cận trên.

## II. TIẾP CẬN HƯỚNG ĐIỂM ẢNH VÀ HƯỚNG ĐỐI TƯỢNG

### A. Tiếp cận hướng điểm ảnh

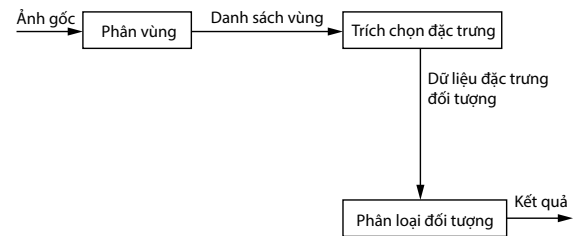
Tiếp cận kinh điển của phân loại ảnh viễn thám dựa trên điểm ảnh [9]. Tiếp cận này chỉ thông tin phổ được sử dụng trong việc phân loại [8]. Tiếp cận này bao gồm phương pháp phân loại có giám sát và không giám sát truyền thống [8][9]. Phương pháp phân loại hợp lý tối đa (maximum likelihood) thuộc tiếp cận này. Hình 1 minh họa quy trình phân loại ảnh hướng điểm ảnh.



Hình 1. Quy trình phân loại ảnh đa phổ hướng điểm ảnh

### B. Tiếp cận hướng đối tượng

Trong tiếp cận hướng đối tượng, đơn vị xử lý không còn là các điểm ảnh đơn giản mà là các đối tượng ảnh [8]. Đầu tiên, ảnh được phân đoạn thành các nhóm điểm ảnh có nghĩa. Thứ hai, một tập các luật phân đoạn dựa trên tri thức để mô tả mỗi lớp được định nghĩa. Luật bao gồm thông tin phổ, không gian, ngữ cảnh và kết cấu [8]. Và sau đó, bộ phân loại sẽ được chọn để gán mỗi đoạn về lớp phù hợp theo luật [10]. Hình 2 minh họa quy trình phân loại dựa trên tiếp cận hướng đối tượng.



Hình 2. Quy trình phân loại ảnh đa phổ hướng đối tượng

## III. PHÂN LOẠI HỢP LÝ TỐI ĐA

Trong [1], phương pháp phân loại hợp lý tối đa được trình bày một cách chi tiết.

### A. Phân loại Bayes

Ký hiệu các lớp phổ cho một ảnh qua việc biểu diễn như sau:  $\omega_i, i = 1, \dots, M$ . Trong đó,  $M$  là tổng số lớp. Trong việc cố gắng xác định lớp hoặc loại mà vector điểm ảnh  $x$  thuộc lớp đó là xác suất có điều kiện  $p(\omega_i / x), i = 1, \dots, M$ . Vector độ đo  $x$  là một cột giá trị độ sáng cho điểm ảnh. Nó mô tả điểm ảnh như một điểm trong không gian đa phổ với hệ tọa độ được bởi độ sáng. Xác suất  $p(\omega_i / x)$  cho sự hợp lý mà lớp chính xác là  $\omega_i$  cho một điểm ảnh tại vị trí  $x$ . Phân loại được thực hiện theo

$$x \in \omega_i, \text{ nếu } p(\omega_i / x) > p(\omega_j / x) \quad \forall j \neq i \quad (1)$$

Nghĩa là, điểm ảnh tại  $x$  thuộc lớp  $\omega_i$  nếu  $p(\omega_i / x)$  là lớn nhất. Luật quyết định trực giác này là một trường hợp đặc biệt của một luật tổng quát hơn trong đó các quyết định có thể được chịu ảnh hưởng theo mức độ khác nhau của ý nghĩa được gán đến sự phân loại không chính xác khác.

*B. Luật quyết định hợp lý tối đa*

Mặc dù rất đơn giản,  $p(\omega_i / x)$  không được biết đến. Tuy nhiên, nếu giả sử rằng dữ liệu huấn luyện đầy đủ sẵn sàng cho mỗi loại che phủ đất. Điều này có thể được dùng để ước lượng một phân bố xác suất cho một loại che phủ mà mô tả cơ hội tìm một điểm ảnh từ lớp  $\omega_i$  tại vị trí  $x$ . Sau đó, việc hình thức hóa hàm phân bố này sẽ được tạo cụ thể hơn. Tuy nhiên, hiện tại, hàm này được giữ lại trong dạng chung và được biểu diễn bởi ký hiệu  $p(\omega_i / x)$ . Sẽ có nhiều  $p(x / \omega_i)$  như có các lớp che phủ đất. Nói cách khác, với một điểm ảnh tại vị trí  $x$  trong không gian đa phổ một tập các khả năng có thể được tính mà đưa ra sự hợp lý liên quan mà điểm ảnh thuộc về mỗi lớp sẵn có.

$p(\omega_i / x)$  mong muốn trong phần 1 và  $p(x / \omega_i)$  sẵn có – được ước lượng từ dữ liệu huấn luyện – liên quan bởi định lý Bayes (Freund, 1992):

$$p(\omega_i / x) = p(x / \omega_i) p(\omega_i) / p(x) \quad (2)$$

Trong đó,  $p(\omega_i)$  là xác suất mà lớp xảy ra trong ảnh. Nếu, chẳng hạn, 15% điểm ảnh của ảnh thuộc về lớp  $\omega_i$  thì  $p(\omega_i) = 0.15$ ;  $p(x)$  trong (2) là xác suất để tìm được điểm ảnh trong bất kỳ lớp nào tại vị trí  $x$ . Nó được quan tâm vì:

$$p(x) = \sum_{i=1}^M p(x / \omega_i) p(\omega_i) \quad (3)$$

mặc dù chính  $p(x)$  không quan trọng trong những điều sau đây.  $p(\omega_i)$  được gọi là xác suất ưu tiên, khi chúng là xác suất với thành viên lớp của một điểm ảnh có thể được dự đoán trước khi phân loại. Bằng việc so sánh  $p(\omega_i / x)$  là các khả năng sau đó. Sử dụng (2) có thể thấy rằng luật phân loại của phần 1 là:

$$x \in \omega_i, \text{ nếu } p(x / \omega_i) p(\omega_i) > p(x / \omega_j) p(\omega_j) \quad \forall j \neq i \quad (4)$$

Trong đó,  $p(x)$  đã được xóa như một thừa số chung. Luật (4) dễ chấp nhận hơn so với luật (1) khi  $p(x / \omega_i)$  được biết từ dữ liệu huấn luyện và nó có thể được hiểu rằng  $p(\omega_i)$  cũng đã biết hoặc

có thể ước lượng từ tri thức phân tích ảnh. Kết quả phù hợp về mặt toán học nếu trong (4) định nghĩa:

$$g_i(x) = \ln \{ p(x / \omega_i) p(\omega_i) \} = \ln p(x / \omega_i) + \ln p(\omega_i) \quad (5)$$

được sử dụng, trong đó  $\ln$  là logarit tự nhiên và (4) được trình bày lại như sau:

$$x \in \omega_i, \text{ nếu } g_i(x) > g_j(x) \quad \forall j \neq i \quad (6)$$

Nghĩa là, với thay đổi như trên, luật quyết định được sử dụng trong phân loại hợp lý tối đa;  $g_i(x)$  tham chiếu đến hàm phân biệt (discriminant).

*C. Mô hình lớp chuẩn đa biến*

Ở giai đoạn này, giả sử rằng phân bố xác suất của các lớp thuộc dạng mô hình chuẩn đa biến. Đây là một giả định, hơn là một thuộc tính có thể chứng minh của các lớp phổ tự nhiên hoặc thông tin. Tuy nhiên, nó dẫn đến sự đơn giản hóa về mặt toán học như sau đây. Hơn nữa nó là một phân bố cho các thuộc tính của dạng đa biến đã biết.

Trong (4) vì thế, bây giờ được giả sử cho  $N$  kênh rằng

$$p(x / \omega_i) = (2\pi)^{-N/2} |\Sigma_i|^{-1/2} \exp \left\{ -\frac{1}{2} (x - m_i)^t \Sigma_i^{-1} (x - m_i) \right\} \quad (7)$$

$$\ln p(x / \omega_i) = -\frac{N}{2 \ln(2\pi)} - \frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (x - m_i)^t \Sigma_i^{-1} (x - m_i) \quad (7.1)$$

Trong đó,  $m_i$  và  $\Sigma_i$  là vector trung bình và ma trận hiệp phương sai của dữ liệu trong lớp  $\omega_i$ . Ta có

$-N / 2 \ln(2\pi)$  là chung cho tất cả  $g_i(x)$  và không giúp cho sự phân biệt. Do đó, thừa số này được bỏ qua và dạng cuối của hàm phân biệt cho phân loại hợp lý tối đa, dựa trên giả định của thống kê chuẩn, là (thay 7.1 vào 5):

$$g_i(x) = \ln p(\omega_i) - \frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (x - m_i)^t \Sigma_i^{-1} (x - m_i) \quad (8)$$

Thông thường, người phân tích không có thông tin hữu ích về  $p(\omega_i)$ , trong trường hợp một tình huống của các khả năng ưu tiên cân bằng được giả định; như một hệ quả  $lnp(\omega_i)$  có thể được xóa bỏ từ (7) khi nó giống nhau với mọi  $i$ . Trong trường hợp này thừa số  $\frac{1}{2}$  cũng có thể được loại bỏ, như hàm phân biệt:

$$g_i(x) = -ln |\Sigma_i| -(x - m_i)^t \Sigma_i^{-1} (x - m_i) \quad (9)$$

Việc cài đặt luật quyết định hợp lý tối đa liên quan đến việc sử dụng hoặc (8) hoặc (9) trong (6). Tuy nhiên có một suy xét xa hơn liên quan đến liệu rằng bất kỳ nhãn hay lớp nào sẵn sàng cũng hợp lý.

#### IV. ĐỀ XUẤT PHƯƠNG PHÁP PHÂN LOẠI DỰA TRÊN TIẾP CẬN LAI

Trong nghiên cứu này, chúng tôi đề xuất một tiếp cận phân loại ảnh với sự kết hợp cả hai tiếp cận trên mà chúng tôi tạm gọi là tiếp cận lai. Đầu tiên, ảnh gốc được tiến hành phân vùng. Thay vì trích chọn đặc trưng đối tượng như trong tiếp cận hướng đối tượng, các vùng sẽ được đưa vào bộ phân loại lai để cho quyết định phân lớp cuối cùng. Hình 3 mô tả quy trình phân loại theo tiếp cận lai.



Hình 3. Quy trình phân loại ảnh đa phổ theo tiếp cận lai.

##### A. Phân vùng ảnh (Unsupervised Segmentation)

Phân đoạn là một quy trình dùng để trích chọn những nét chính của các đối tượng nền bởi việc định nghĩa các vùng tương ứng. Nhiệm vụ của chức năng phân vùng ảnh là từ ảnh đa phổ ban đầu, tiến hành xử lý và phân chia thành các vùng, các cụm khác nhau. Hiện nay, có nhiều phương pháp phân vùng khác nhau như: Các phương pháp hình thái, Các phương pháp họ K-means, Mô hình pha trộn Gaussian có giới hạn (FGMM), Tách và hợp, Các mô hình Markov,... Hầu hết các phương pháp chỉ sử dụng cường độ của mỗi điểm ảnh để định nghĩa các vùng, nhưng đưa ra các phân đoạn rất

hỗn tạp, cụ thể với các ảnh đa phổ có độ phân giải cao. Hiện nay một số thuật toán bao gồm thông tin ngữ cảnh trong quy trình để giảm bớt tính hỗn tạp của các phân đoạn. Trong đó một số thông tin ngữ cảnh của các phân đoạn này được trích chọn từ ảnh cũng được sử dụng.

Bảng 1. Thuật toán KMeansCMN

<p>Đầu vào: <math>n</math> đối tượng và số cụm <math>k</math></p> <p>Đầu ra: Các cụm <math>C_j (j=1...k)</math> sao cho hàm mục tiêu <math>E</math> sau đây đạt cực tiểu:</p> $E = \sum_{(i=1)}^k \sum_{(x \in C_i)} d^2(x, m_i)$
<p><b>Bước 1: Khởi tạo</b></p> <p>Chọn <math>k</math> đối tượng <math>C_j (j=1...k)</math> là tâm ban đầu của <math>k</math> cụm dữ liệu đầu vào (lựa chọn ngẫu nhiên hoặc theo kinh nghiệm).</p> <p><b>Bước 2: Gán tâm cụm theo khoảng cách</b></p> <p>Với mỗi đối tượng <math>x_i (1 \leq i \leq n)</math>, tính khoảng cách của nó tới mỗi tâm <math>C_j</math> với <math>j=1...k</math>. Đối tượng thuộc về cụm <math>C_s</math> mà khoảng cách từ tâm <math>C_s</math> tương ứng đến đối tượng đó là nhỏ nhất.</p> $d(x, C_s) = \min d(x, C_j), 1 \leq j \leq k \quad (10)$ <p><b>Bước 3: Cập nhật tâm cụm</b></p> <p>Đối với mỗi <math>j=1...k</math>, cập nhật lại tâm cụm <math>C_j</math> bằng cách xác định trung bình cộng của các vector đối tượng dữ liệu đã được gán về cụm.</p> <p>Nếu số lượng điểm ảnh trong cụm nhỏ hơn hằng số rất lớn <math>Max</math> thì tâm vẫn tính theo công thức như sau:</p> $C_j = \frac{\sum_{x \in cluster(j)} x}{count(cluster(j))} \quad (11)$ <p>Nếu số lượng điểm ảnh trong cụm lớn hơn hằng số rất lớn <math>Max</math> thì tâm tính theo công thức như sau:</p> $C_j = CMN(Cluster_j) \quad (11a)$ <p><b>Bước 4: Lặp và kiểm tra điều kiện dừng</b></p> <p>Lặp lại các bước 2 và 3 cho đến khi các tâm cụm không thay đổi giữa hai lần lặp liên tiếp.</p>

Một vấn đề chung với các hệ thống xử lý 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 cụ thể hóa ảnh hưởng của những khác biệt này trên hiệu năng nhận dạng là phép chuẩn

hóa trung bình phổ (Cepstral Mean Normalisation - CMN) [12]. Phương pháp này được áp dụng rộng rãi và hiệu quả trong xử lý tín hiệu số và nhận dạng tiếng nói. Tuy nhiên, khi áp dụng trong xử lý 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 chỉ ra một ứng dụng của phép chuẩn hóa CMN trong phân cụm ảnh viễn thám.

Trong nghiên cứu này, chúng tôi đã thử nghiệm cải tiến thuật toán KMeans [11] và cải tiến thành KMeansCMN. Thuật toán KMeansCMN được trình bày như bảng I:

Thu tục tính tâm cụm CMN(Cluster<sub>j</sub>) tại vòng lặp thứ n như sau:

**Bước 1:** Khởi tạo tâm theo công thức

$$C_j^n = \beta C_j^{n-1}$$

**Bước 2:** Với mỗi  $x \in \text{Cluster}_j$  tính theo công thức

$$C_j^n = \alpha C_j^{n-1} + \beta x$$

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

### B. Phân loại lai

Hàm phân biệt  $g_i(x)$  như được trình bày trong mục III.C của phương pháp phân loại hợp lý tối đa chỉ áp dụng cho từng điểm ảnh. Trong phần này, chúng tôi đề xuất một cải tiến cho hàm phân biệt này để áp dụng cho việc phân lớp các cụm-đối tượng đã thu được từ giai đoạn phân loại không giám sát trong mục IV.A.

Sau khi phân cụm ảnh gốc I ta được tập O cụm-đối tượng như sau:

$$O = \{o_i : 0 \leq i \leq K\} \quad (12)$$

Trong đó,  $o_i \cap o_j = \emptyset; \forall i \neq j, i, j = 1, 2, \dots, K; o_1 \cup o_2 \cup \dots \cup o_K = I$ .

Chúng tôi xây dựng hàm phân biệt cho mỗi cụm như sau:

$$f_i(o) = \sum_{x \in o} g_i(x) / \text{count}(o) \quad (13)$$

Hoặc  $f_i(o) = \max_{x \in o} g_i(x) \quad (13.1)$

Từ (6) chúng tôi đề xuất luật quyết định lớp cho mỗi cụm như sau:

$$o \in \omega_i, \text{ nếu } f_i(o) > f_j(o) \quad \forall j \neq i \quad (14)$$

## V. THỬ NGHIỆM

Tập dữ liệu phục vụ cho thử nghiệm gồm 2 loại. Một là, loại ảnh LANDSAT ETM+ chụp khu vực Hòa Bình, bao gồm các ảnh ranh giới từng huyện khu vực tỉnh Hòa Bình. Hai là ảnh SPOT, loại ảnh có độ phân giải cao, gồm 4 kênh: Lục, Đỏ, Cận hồng ngoại, Hồng ngoại, chụp khu vực Hòa Bình và Sơn La với 21 ảnh chụp năm 2003 và 14 ảnh chụp năm 2008. Do khuôn khổ bài báo có hạn, nhóm tác giả trình việc thử nghiệm với hai mẫu ảnh đầu vào khác nhau.

Để đánh giá chất lượng phân cụm, trong [13], các tác giả đã sử dụng chỉ số  $F(I)$  [14], tuân theo các tiêu chí về sự đồng nhất cụm [13] [15], để so sánh kết quả phân cụm của các thuật toán.  $F(I)$  càng nhỏ thì độ đồng nhất càng cao. Chỉ số này được tính như sau:

$$F(I) = \frac{1}{1000(N \times M)} \sqrt{R} \sum_{i=1}^R \frac{e_i}{\sqrt{A_i}}$$

### A. Thử nghiệm thuật toán phân cụm

#### 1) Thử nghiệm 1

Hình 4 là ảnh kết quả phân cụm của KMeans và KMeansCMN trong trường hợp 5 cụm với ảnh SPOT. 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.

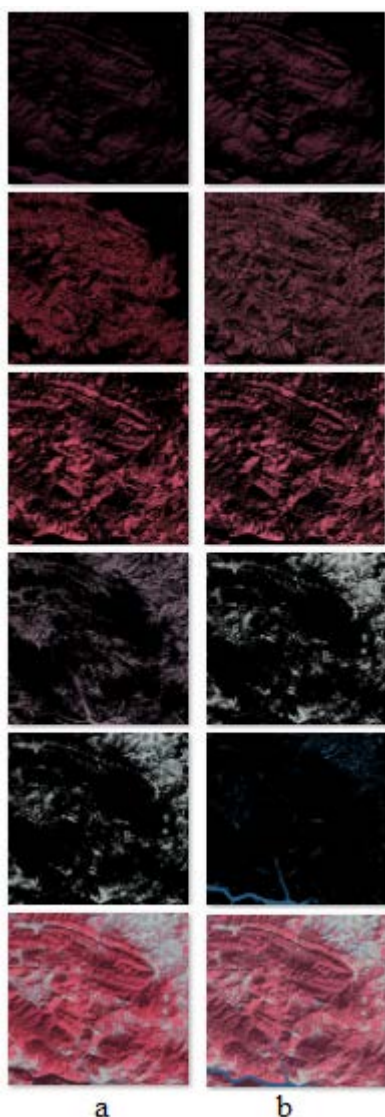
Bảng II. So sánh độ đồng nhất cụm (\*1.0E + 3)

Số cụm	KMeans	KMeansCMN
5	2.24	2.04
10	1.58	1.81

Bảng II so sánh độ đồng nhất tâm cụm giữa KMeans và KMeansCMN. Bảng III thống kê số thời gian thực thi của KMeans và KMeansCMN với 5 cụm và 10 cụm. Chúng ta thấy độ đồng nhất của KMeansCMN là tốt hơn so với KMeans. Ngoài ra, thời gian phân cụm của KMeansCMN cũng nhỏ hơn KMeans.

Bảng III. Thời gian phân cụm (MS)

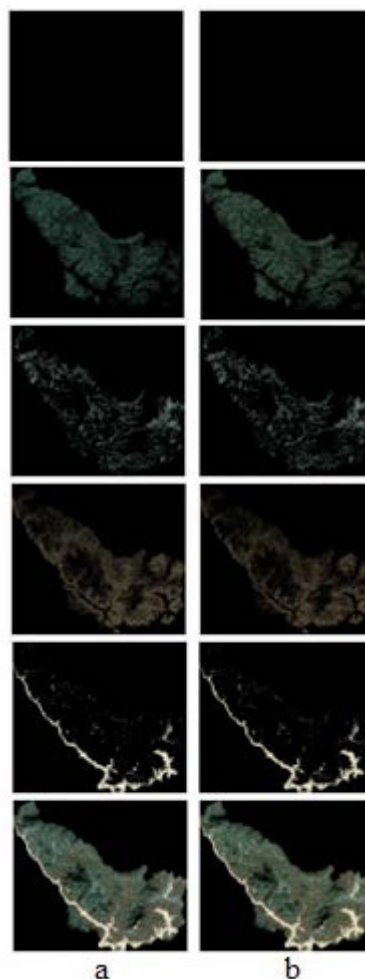
Số cụm	KMeans	KMeansCMN
5	2,616,938	2,413,791
10	11,275,333	8,618,345



Hình 4. Kết quả phân cụm bởi KMeans (a) và KMeansCMN (b)

## 2) Thử nghiệm 2

Hình 5 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 5. Kết quả phân cụm bởi KMeans (a) và KMeansCMN (b)

Bảng IV. So sánh độ đồng nhất cụm (\*1.0E-05)

Số cụm	KMeans	KMeansCMN
5	1.66	1.5
8	1.61	1.53

Bảng IV so sánh độ đồng nhất tâm cụm giữa KMeans và KMeansCMN. Bảng V thống kê số bước lặp cũng như thời gian thực thi của KMeans

và KMeansCMN với 5 cụm và 8 cụm. Chúng ta thấy độ đồng nhất của KMeansCMN là tốt hơn so với KMeans. Ngoài ra, thời gian phân cụm của KMeansCMN cũng nhỏ hơn KMeans.

Bảng V. Thời gian phân cụm (MS)

Số cụm	KMeans	KMeansCMN
5	263,672	213,109
8	1,658,609	1,568,062

**Nhận xét:** Tốc độ hội tụ của KMeansCMN phân cụm tốt hơn và nhanh hơn so với KMeans.

*B. Thử nghiệm thuật toán phân lớp*

Kết quả thử nghiệm có sự so sánh giữa thuật toán phân loại theo phương pháp phân loại hợp lý tối đa (thuật toán góc), được cài đặt trong phần mềm Grass và thuật toán dựa trên tiếp cận lai (thuật toán cải tiến).

Mẫu thứ nhất là ảnh SPOT với kích thước 2201 × 2101. Thực hiện việc phân loại ảnh đầu vào với 3 lớp: đất (màu nâu trong ảnh kết quả), nước (màu vàng trong ảnh kết quả), rừng (màu xanh lam trong ảnh kết quả). Kết quả phân loại được thể hiện trong hình 7. Từ kết quả trong hình 7, quan sát vùng được khoanh tròn trên ảnh, vùng rừng có xen lẫn đất với tính hỗn tạp rất cao. Tuy nhiên, thuật toán góc không phản ánh rõ sự đan xen này mà quy về một lớp. Trong khi thuật toán cải tiến phân biệt rất rõ.

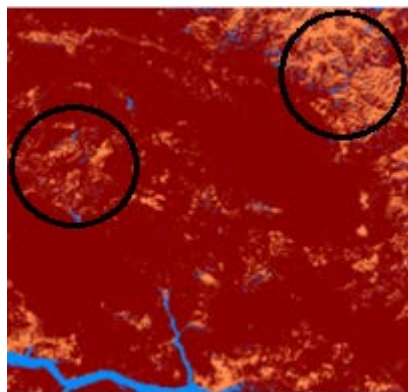
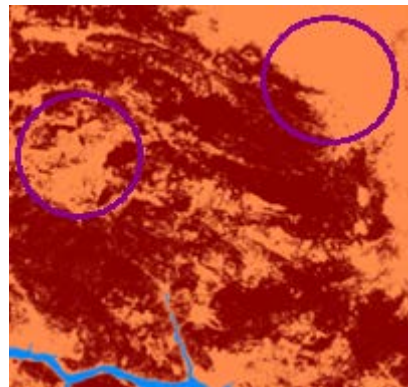
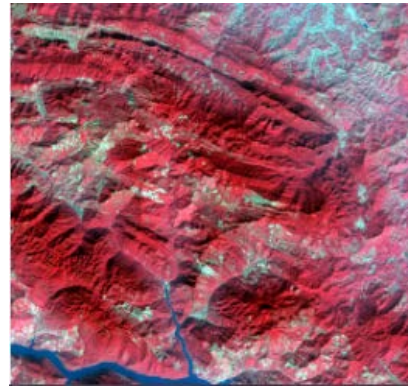
Bảng VI. So sánh độ chính xác

Lớp	MLK	Lai
Đất	95%	98%
Rừng	100%	100%
Nước	98%	98%

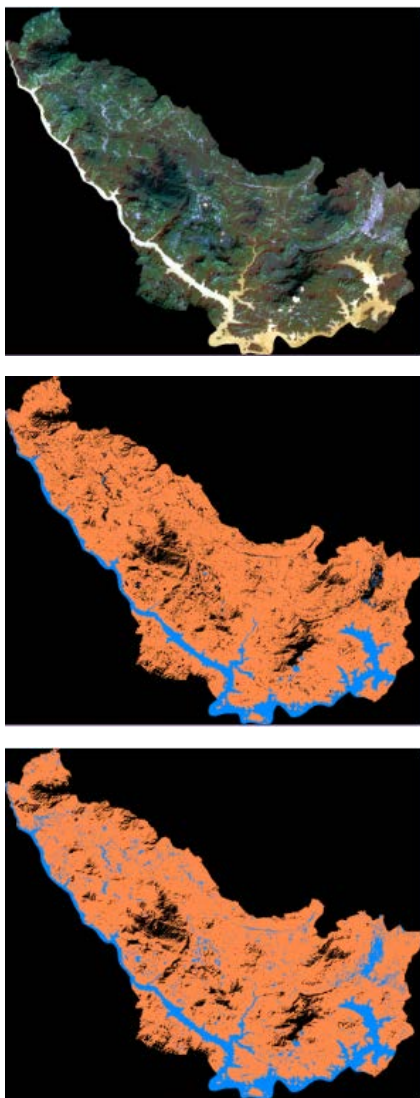
Bảng VII. So sánh độ chính xác

Lớp	MLK	Lai
Núi đá	85%	84%
Rừng	90%	93%
Nước	100%	100%

Mẫu thứ hai là ảnh LANDSAT với kích thước 1596 × 1333. Thực hiện việc phân loại ảnh đầu vào với 3 lớp: Núi đá, nước, rừng. Kết quả được thể hiện trong hình 7. Ảnh đầu vào cũng có tính hỗn tạp cao và được thể hiện rõ hơn qua ảnh kết quả của thuật toán phân loại đề xuất.



Hình 6. Ảnh đầu vào, kết quả phân loại thuật toán góc và cải tiến



Hình 7. Ảnh đầu vào, kết quả phân loại thuật toán gốc và cải tiến

## VI. KẾT LUẬN

Trong nghiên cứu này, chúng tôi đã đề xuất một tiếp cận phân loại lai với sự kết hợp cả hai tiếp cận hướng điểm ảnh và hướng đối tượng. Quy trình phân loại theo tiếp cận này gồm hai giai đoạn. Một là, phân vùng ảnh, trong thuật toán cải tiến theo tiếp cận này, chúng tôi sử dụng thuật toán KMeans. Hai là, phân loại lai, các vùng sau đó sẽ được phân lớp theo bộ phân loại lai sử dụng hàm quyết định phân lớp mới mà chúng tôi đề xuất để phân lớp các cụm-đối tượng. Hàm quyết định phân lớp này được

cải tiến dựa trên hàm phân biệt của phương pháp phân loại hợp lý tối đa. Các kết quả thử nghiệm cho thấy kết quả phân loại của thuật toán dựa trên tiếp cận đề xuất cho độ chính xác cao hơn so với thuật toán của phương pháp hợp lý tối đa.

Ngoài ra, trong giai đoạn phân vùng, 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. 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 thực nên tốc độ chậm. Trong nghiên cứu tiếp theo, nhóm tác giả dự kiến sử dụng phương pháp tính toán chấm tính để tăng cường tốc độ thủ tục này nhằm tăng tốc độ phân cụm.

## TÀI LIỆU THAM KHẢO

- [1] John A. Richards, Xiuping Jia, Remote Sensing Digital Image Analysis, Springer-Verlag Berlin Heidelberg 2006.
- [2] Meritxell Bach Cuadra, Jean-Philippe Thiran, Satellite Image Segmentation and Classification, Fall 2004.
- [3] MacQueen, J.: Some methods for classification and analysis of multivariate observations. In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Vol.1. University of California Press (1967) 281-297.
- [4] Chih-Tang Chang, Jimz. C. Lai, Muderjeng, A Fuzzy K-means Clustering Algorithm Using Cluster Center Displacement, Journal of Information Science and Engineering 27, 2011, pp. 995-1009.
- [5] Balaji T., Sumathi M., “Relational Features of Remote Sensing Image classification using Effective K-Means Clustering”, International Journal of Advancements in Research & Technology, Volume 2, Issue 8, August-2013, pp. 103-107.
- [6] Smriti Sehgal, “Remotely sensed Landsat Image Classification using Neural network approaches”, International Journal of Engineering Research and Applications, Vol. 2, Issue 5, 2012, pp.043-046.



- [7] Nedeljkovic, “Image Classification based on fuzzy logic”, The International Archives of the Photogrammetry, Remote sensing and Spatial Information Sciences, Vol. 34, Part XXX, 2003.
- [8] Sun Xiaoxia, Zhang Jixian, Liu Zhengjun, “A comparison of object-oriented and pixel-based classification approaches using quickbird imagery”, Chinese Academy of Surveying and Mapping, Beijing, China.
- [9] H.R. Matinfar, F. Sarmadian, S.K. Alavi Panah, R.J. Heck, “Comparison of object-oriented and pixel-based classification on Lansatsat7, Etm+ Spectral Bands (Case Study: Arid Region of Iran)”, American-Eurasian J. Agric. & Environ, 2007.
- [10] Leukert k., “Transferability of knowledge-based classification rules”, ISPRS2004, Istanbul, 2004.
- [11] <http://www.onmyphd.com/?p=k-means.clustering>.
- [12] <http://recognize-speech.com/preprocessing/cepstral-mean-normalization>
- [13] Intan aidha yusoff, Nor ashidi mat isa, Two-Dimensional Clustering Algorithms for Image Segmentation, WSEAS Transactions on Computers, Issue 10, Volume 10, October 2011.
- [14] J. Liu, and Y. H. Yang, Multiresolution color image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.16, no.7, pp.689-700, Jul 1994.
- [15] R. H. Haralick, and L. G. Shapiro, Image segmentations techniques, Computer Vision Graphics Image Processing 29, pp. 100-132, 1985.

## A TECHNIQUE OF CLASSIFYING REMOTE SENSING IMAGES BASED ON HYDRID APPROACH

**Abstract:** Remote sensing image classification is interested by reseachers. Having two main approaches include the pixel based approach and the object oriented based approach. Remote sensing images can have multichannel and high resolution. The pixel based approach usually has high accuracy but having problem with large size images as remote sensing images. While the object oriented based approach the problem but usually having

accuracy lower than the pixel based approach. This paper presents a hybrid technique which combines both of the approaches to propose a more effective classifying algorithm of the remote sensing image classification. The algorithm is experimented on data set which is remote sensing images of Hoa Binh province.

**Keyword:** remote sensing, image classification, Object Oriented approach, Hybrid approach.



**Nguyễn Tu Trung**, tốt nghiệp đại học Trường ĐH Sư phạm 2 Hà Nội năm 2007 và Thạc sỹ tại trường ĐH Công Nghệ, ĐHQGHN năm 2011, nghiên cứu sinh khóa 2013, Học viện Công nghệ Bưu chính Viễn thông. Lĩnh vực nghiên cứu: Xử lý ảnh, xử lý tiếng nói, hệ thống thông tin, hệ thống nhúng.



**Ngô Hoàng Huy**, tốt nghiệp đại học trường ĐH Sư phạm Hà Nội năm 1990. Nơi công tác: Viện CNTT, Viện Hàn lâm KHCNVN. Hiện đang làm nghiên cứu sinh tại VCNTT, Viện Hàn lâm KHCNVN. Lĩnh vực nghiên cứu: Xử lý ảnh, xử lý tiếng nói, hệ thống thông tin, hệ thống nhúng.



**Đặng Văn Đức**, nhận học vị Tiến sĩ năm 1996, Việt Nam. Nhận chức danh PGS năm 2002. Nơi công tác: Viện CNTT, Viện Hàn lâm KHCNVN. Lĩnh vực nghiên cứu: GIS và Viễn thám, Đa phương tiện, Công nghệ phần mềm.



**Vũ Văn Thoả**, tốt nghiệp Đại học Sư phạm Vinh năm 1975, Tiến sĩ năm 1990 Viện Điều khiển tại Liên Xô cũ. Hiện công tác tại Khoa Quốc tế và Đào tạo Sau Đại học, Học viện Công nghệ Bưu chính Viễn thông. Lĩnh vực nghiên cứu: Lý thuyết thuật toán, tối ưu hóa, hệ thống tin địa lý, mạng viễn thông.



**Lại Anh Khôi**, hiện công tác tại Viện Công nghệ Vũ Trụ, Viện Hàn lâm KHCNVN. Lĩnh vực nghiên cứu: Viễn thám.