

*Hội nghị Khoa học kỷ niệm 40 năm thành lập Viện Hàn lâm KH&CN Việt Nam
Tiểu ban Công nghệ thông tin, Điện tử, Tự động hóa và Công nghệ vũ trụ
Hà Nội, 7/10/2015*

Phân loại ảnh viễn thám dựa trên phương pháp tựa đối tượng

Nguyễn Tu Trung¹, Ngô Hoàng Huy¹, Lại Anh Khôi², Đặng Văn Đức¹, Vũ Văn Thoa³

¹Viện Công nghệ Thông tin, Viện Khoa Học và Công nghệ Việt Nam

²Viện Công nghệ Vũ Trụ, Viện Khoa Học và Công nghệ Việt Nam

³Học viện Công nghệ Bưu chính Viễn thông

Email liên lạc: nguyentutrong. nhuy1@ioit.ac.vn, lahoi@sti.vast.vn, dvdac@ioit.ac.vn,
thoa236@gmail.com

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 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 điểm ảnh. Bài báo này trình bày một kỹ thuật lai với sự kết hợp cả hai hướng tiếp cận để đưa ra một thuật toán phân loại ảnh viễn thám hiệu quả hơn mà chúng tôi tạm gọi là phương pháp tựa đối tượng. Kết quả đạt được cũng được thử nghiệm trên ảnh viễn thám tỉnh Hòa Bình.

Từ khóa: Viễn thám, phân loại ảnh.

1. Giới thiệu

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 có ích cho khoa học 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 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. Hai là, phân loại đối tượng sử dụng các thuật toán phân loại có giám sát. 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.

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 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 pixel. Trong nghiên cứu này, chúng tôi đề xuất phương pháp phân loại ảnh viễn thám mới mà chúng tôi tạm gọi là phân loại tọa độ tương, được cải tiến từ phương pháp phân loại hợp lý tối đa với sự kết hợp cả hai hướng tiếp cận trên.

2. Phân loại hợp lý tối đa

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

2.1. 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, như được thể hiện trong ví dụ hai chiều trong hình 3.5. Xác suất $p(\omega_i/x)$ cho sự hợp lý mà lớp chính xác là ω_i cho một điểm ảnh tại vị trí x . Phân loại được thực hiện theo

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

Nghĩa là, điểm ảnh tại x thuộc lớp ω_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. Tiếp cận chung được gọi là phân loại Bayes và là chủ đề được trình bày trong phụ lục F.

2.2. 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 ω_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 ω_i xảy ra trong ảnh. Nếu, chẳng hạn, 15% điểm ảnh của ảnh thuộc về lớp ω_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:

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

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 (3) 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 (3) được trình bày lại như sau:

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

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).

2.3. 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 (xem phụ lục E)

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

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

Trong đó, m_i và Σ_i là vector trung bình và ma trận hiệp phương sai của dữ liệu trong lớp ω_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):

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

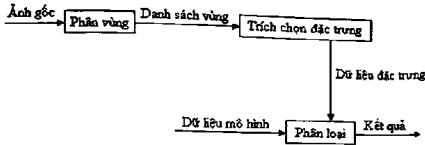
Thông thường, người phân tích không có thông tin hữu ích về $p(\omega_i)$, trong trường hợp một tình huống của các khả năng ưu tiên cân bằng được giả định; như một hệ quả $\ln p(\omega_i)$ có thể được xóa bỏ từ (7) khi nó giống nhau với mọi i . Trong trường hợp này thừa số $1/2$ cũng có thể được loại 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. Điều này liên quan đến việc sử dụng ngưỡng như được thảo luận trong phần 2.5 sau đây.

3. Đề xuất phương pháp phân loại dựa trên đối tượng

Trong nghiên cứu này, chúng tôi đề xuất một quy trình phân loại ảnh với các giai đoạn khác nhau như hình 1. Trong bước đầu tiên, mục tiêu là để phân chia ảnh thành các vùng tương ứng, các đặc trưng như là cường độ, kết cấu hay hình dạng được trích chọn từ các vùng này để thực hiện một phân loại dựa trên vùng. Cuối cùng của bước này một nhãn được phân gán cho mỗi vùng, tương ứng với phân loại cuối cùng.



Hình 1: Quy trình phân loại ảnh đa phổ.

3.1. 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.

Trong nghiên cứu này, chúng tôi đã thử nghiệm cài đặt thuật toán K-means. 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_j ($j=1..k$) sao cho hàm mục tiêu E sau đây đạt cực tiểu:	
$E = \sum_{i=1}^k \sum_{x \in C_i} d^2(x, m_i)$	(9)
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.	
$d(x, C_s) = \min d(x, C_j), 1 \leq j \leq k$	(10)
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.	
$C_j = \frac{\sum_{x \in \text{cluster}(j)} x}{\text{count}(\text{cluster}(j))}$	(11)
Bước 4: Lập 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.	

3.2. Phân loại ảnh (Supervised segmentation)

Trong phần này, chúng tôi đề xuất một cải tiến của phương pháp phân loại hợp lý tối đa áp dụng cho việc phân lớp các cụm-đối tượng đã thu được từ giai đoạn phân loại không giám sát trong mục 3.1

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:

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

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

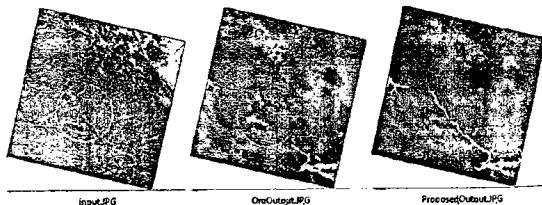
$$o \in \omega_i, \text{ nếu } f_i(o) > f_j(o) \quad \forall j \neq i \quad (14)$$

4. Thử nghiệm và so sánh

Kết quả thử nghiệm có sự so sánh giữa phương pháp phân loại Maximun Likelihood của phần mềm Grass và phương pháp phân loại cải tiến.

Tập dữ liệu phục vụ cho thử nghiệm gồm ba loại. Một là, loại ảnh Landsat ETM+ chụp khu vực Hòa Bình năm 2001, bao gồm 11 ảnh ranh giới từng huyện và một ảnh theo ranh giới tỉnh của tỉnh Hòa Bình. Ảnh Landsat ETM+ gồm 7 kênh [10]: Châm, Lục đỏ, Đỏ, Cận hồng ngoại, Hồng ngoại trung, Hồng ngoại nhiệt, Hồng ngoại trung. Hai là, loại ảnh SPOT, 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. Hai loại ảnh này, nhóm tác giả có được khi tham gia thực hiện đề tài "Phát triển phần mềm xử lý ảnh viễn thám trên nền phần mềm GRASS". Đây là đề tài cấp nhà nước, thuộc Chương trình KHCN Vũ Trụ. Trong đó, ảnh SPOT với độ phân giải cao, được bán với giá khá cao bởi Cục Viễn thám – Bộ Tài nguyên và Môi trường, khoảng vài chục triệu một ảnh. Ba 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>. Do khuôn khổ bài báo có hạn, nhóm tác giả trình bày thử nghiệm với hai mẫu ảnh đầu vào khác nhau.

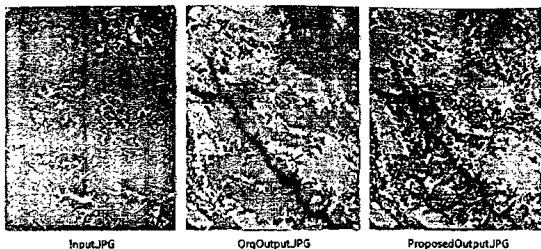
Mẫu thử nhất là ảnh SPOT với kích thước 7173 x 7202. Phân loại ảnh đầu vào với 3 lớp: đất (màu đỏ trong ảnh kết quả), nước (màu vàng trong ảnh kết quả), rừng (màu xanh lam trong ảnh kết quả). Kết quả được thể hiện trong hình 1. Từ kết quả trong hình 1, chúng ta thấy, tại vùng đã được khoanh tròn trên ảnh, độ chính xác của phương pháp gốc là thấp hơn so với phương pháp đề xuất.



Hình 1: Mẫu 1: Ảnh đầu vào, kết phân loại thuật toán gốc và cải tiến.

Mẫu thử nhất là ảnh SPOT với kích thước 2592 x 3348. Phân loại ảnh đầu vào với 3 lớp: đất (màu vàng trong ảnh kết quả), nước (màu đỏ trong ảnh kết quả), rừng (màu xanh lam trong ảnh kết quả). Kết quả được thể hiện trong hình 2. Từ kết quả trong hình 1, chúng ta thấy, vùng đã được khoanh

tròn trên ảnh là vùng không đồng nhất. Tuy nhiên, phương pháp gốc cho kết quả đồng nhất. Trong khi phương pháp đề xuất thể hiện rõ sự không đồng nhất của vùng được quan tâm.



Hình 2: Mẫu 2: Ảnh đầu vào, kết phân loại thuật toán gốc và cải tiến.

5. Kết luận

Trong nghiên cứu này, chúng tôi đã đề xuất một quy trình phân loại ảnh đa phổ theo tiếp cận tọa đối tượng. Trong giai đoạn phân vùng ảnh, chúng tôi đã thuật toán KMeans. Sau đó, các vùng ảnh này sẽ được phân lớp sử dụng thuật toán phân loại hợp lý tối đa cải tiến mà chúng tôi đề xuất có việc phân lớp các cụm-đối tượng. Các kết quả thử nghiệm cho thấy kết quả phân loại của phương pháp đề xuất cho độ chính xác cao hơn so với phương pháp gốc.

Các hướng nghiên cứu tiếp theo có thể tập trung vào giai đoạn phân vùng ảnh để tăng cường các hiệu quả tiềm năng cho giai đoạn phân loại. Một trong số đó có thể là việc nghiên cứu sâu hơn về các phép tiền xử lý, nâng cao chất lượng ảnh để hỗ trợ cho các phương pháp phân vùng.

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] <http://www.onmyphd.com/?p=k-means.clustering>.