

Bài báo tổng quan**PHÂN ĐOẠN ẢNH VÀ NCUTS****Trần Như Ý¹, Nguyễn Việt Hưng^{2*}, Nguyễn Quốc Huy³, Phạm Thế Bảo³**¹Khoa Công nghệ thông tin, Trường Đại học Công nghiệp Thực phẩm Thành phố Hồ Chí Minh, Việt Nam²Khoa Công nghệ thông tin, Trường Đại học Sư phạm Thành phố Hồ Chí Minh, Việt Nam³Khoa Công nghệ thông tin, Trường Đại học Sài Gòn, Việt Nam*Tác giả liên hệ: Nguyễn Việt Hưng – Email: hungnv@hcmue.edu.vn

Ngày nhận bài: 17-4-2021; ngày nhận bài sửa: 13-5-2021; ngày duyệt đăng: 10-6-2021

TÓM TẮT

Trong nhiều thập kỷ qua, nhiều công trình nghiên cứu khoa học đóng góp không ngừng trong lĩnh vực thị giác máy tính nói chung cũng như nghiên cứu phân đoạn ảnh nói riêng. Trong đó, phân đoạn ảnh là quá trình tiền xử lý quan trọng trong hầu hết các ứng dụng xử lý ảnh. Chúng tôi tóm tắt và đánh giá các kỹ thuật phân đoạn ảnh và phân chia các kỹ thuật này thành các nhóm gồm: kỹ thuật dựa trên phát hiện cạnh/biên, kỹ thuật phân ngưỡng, phương trình vi phân, phương pháp gom nhóm, kỹ thuật dựa trên phân hoạch đồ thị. Tiếp theo, chúng tôi trình bày ưu điểm và khuyết điểm của thuật toán Ncuts, là thuật toán kinh điển khá phổ biến trong phân đoạn ảnh dựa trên đồ thị. Thuật toán Ncuts (Shi, & Malik, 2000) được đưa ra năm 2000 nhưng đã được áp dụng thành công và cho kết quả tối ưu cho nhiều ứng dụng xử lý ảnh cũng như các ứng dụng khoa học kỹ thuật.

Từ khóa: eigenvalue; graph-cut; Ncuts**1. Bài toán phân đoạn ảnh**

Ảnh số ngày càng trở nên phổ biến và phong phú hơn, đặc biệt là có liên quan đến nhiều ứng dụng khoa học kỹ thuật. Ảnh số được xem là một trong những phương tiện quan trọng nhất trong việc truyền tải thông tin trong lĩnh vực thị giác máy tính. Việc hiểu các thông tin từ ảnh giúp thực hiện được nhiều nhiệm vụ trong các ứng dụng khoa học kỹ thuật như xác định các tế bào ung thư trong y khoa, xác định vị trí sân bay từ dữ liệu điều khiển cảm biến. Và mục tiêu cần thiết phải xác định các phân vùng hay các đối tượng trong ảnh để xử lý hay rút trích thông tin ở mức cao là rất quan trọng trong các quá trình xử lý ảnh và thị giác máy tính (Wang, Jensen, & Im, 2010; Wang, Kong, Lu, Qi, & Zhang, 2008; Guo, Ng, Goubran, Petersen, Piechnik, Neubauer, & Wright, 2020).

Phân đoạn ảnh là quá trình phân chia ảnh thành các vùng không giao nhau hay các đối tượng như tập các điểm ảnh, hoặc vùng gồm các điểm ảnh có tính chất tương đồng nhau theo một tiêu chuẩn đồng nhất nào đó, ví dụ như màu sắc, cường độ, kết cấu, mức xám... để từ

Cite this article as: Tran Nhu Y, Nguyen Viet Hung, Nguyen Quoc Huy, & Pham The Bao (2021). Image Segmentation and Ncuts. *Ho Chi Minh City University of Education Journal of Science*, 18(6), 1100-1112.

đó xác định vị trí đối tượng và biên trong ảnh (Dass, Priyanka, & Devi, 2012). Đây là một trong những nhiệm vụ quan trọng nhất trong phân tích ảnh số tự động vì các kết quả của phân đoạn ảnh sẽ có ảnh hưởng cốt yếu đến tất cả các quá trình tiếp theo trong nhiều ứng dụng của xử lý ảnh và thị giác máy tính, như biểu diễn và mô tả đối tượng trong ảnh, xử lý nhận dạng đối tượng, phân loại ảnh, nén ảnh dựa trên đối tượng hay truy vấn ảnh dựa trên nội dung...

Trong lĩnh vực thị giác máy tính và xử lý ảnh, phân đoạn ảnh là một trong những công đoạn quan trọng và được quan tâm nghiên cứu rất nhiều do tính cần thiết trong nhiều ứng dụng. Quá trình phân đoạn ảnh số còn có nhiều đóng góp quan trọng trong nhiều lĩnh vực (Hoang, & Nguyen, 2018; Almotiri, Elleithy, & Elleithy, 2018; Jain, Prabhakar, Member, & Hong, 1999; Chandra, & Bajpai, 2019; Wang, Jensen, & Im, 2010; Wang, Kong, Lu, Qi, & Zhang, 2008; Dimauro, & Simone, 2020; Guo, Ng, Goubran, Petersen, Piechnik, Neubauer, & Wright, 2020) như y tế, viễn thám, dự báo thời tiết, các bài toán nhận dạng, các bài toán giao thông...

2. Các hướng tiếp cận

Việc nghiên cứu phân đoạn ảnh đã có nhiều thành tựu và nhiều kỹ thuật phân đoạn ảnh được đưa ra. Tuy nhiên, không có một phương pháp nào là tốt nhất cho các loại ảnh khác nhau, cũng như không phải tất cả các phương pháp là tốt nhất cho một loại ảnh đặc trưng nào. Vì vậy, việc lựa chọn kỹ thuật phân đoạn ảnh nào đó còn dựa trên tính chất của ảnh và vấn đề cần giải quyết. Có nhiều cách để phân loại các kỹ thuật này, theo (Dass, Priyanka, & Devi, 2012; Misal, & Singh, 2013; Cheng, Jiang, Sun, & Wang, 2001) các kỹ thuật có thể phân loại.

a. Hướng tiếp cận phát hiện cạnh/biên

Kỹ thuật này dựa trên việc phát hiện các cạnh hay các điểm ảnh giữa các vùng khác nhau về mức độ biến đổi cường độ nhằm thành các biên giữa các vùng hay đối tượng (Zaitoun, & Aqel, 2015). Hai phương pháp phân đoạn dựa trên cạnh cơ bản là: lược đồ xám của ảnh và kỹ thuật dựa trên biến đổi gradient. Để phát hiện các cạnh, một trong những kỹ thuật phát hiện cạnh cơ bản (Senthilkumaran, & Rajesh, 2009; Kirti, & Bhatnagar, 2017; Dhankhar, & Sahu, 2013; Ganesan, & Sajiv, 2017) như toán tử sobel, toán tử prewitt, toán tử Robert, toán tử Canny, Wavelets...

Theo bài báo (Hoang, & Nguyen, 2018) đã xây dựng một phương pháp tiếp cận tự động để khảo sát định kỳ kết cấu tường bê tông. Cách tiếp cận mới nhằm xác định nhanh chóng và chính xác các vết nứt trên bề mặt tường bê tông bằng cách phân tích hình ảnh thu được bằng máy ảnh kỹ thuật số. Nghiên cứu so sánh về hiệu suất của các thuật toán Roberts, Prewitt, Canny và Sobel cho phép phát hiện cạnh được tối ưu hóa mô hình để nhận biết vết nứt tường bê tông. Tác giả (Ganesan, & Sajiv, 2017) trình bày đánh giá hiệu suất của các phương pháp phát hiện cạnh khác nhau được thực hiện trên các hình ảnh khác nhau bằng Matlab. Kết quả đánh giá điểm tích cực và điểm yếu của các phương pháp phát hiện cạnh

được lập trong bảng. Trong số các phương pháp phát hiện cạnh tiêu chuẩn, toán tử canny cho kết quả rất tốt, đặc biệt là trong điều kiện nhiễu. Về độ chính xác, phương pháp dựa trên Wavelets chính xác hơn và tốt hơn so với các phương pháp khác.

Hướng tiếp cận phương pháp phát hiện cạnh/biên là phương pháp hoạt động dựa trên tích chập mặt nạ nhằm tìm ra biên đối tượng trong phân đoạn ảnh. Ưu điểm của phương pháp trên đơn giản, cho hiệu suất tốt khi phát hiện cạnh dọc và ngang. Nhược điểm của hướng tiếp cận phụ thuộc mức độ nhiễu của ảnh nên độ chính xác không cao.

b. Hướng tiếp cận phân ngưỡng

Kỹ thuật này (Sezgin, & Sankur, 2004) tách biệt giữa nền ảnh và đối tượng bằng việc đặt một giá trị ngưỡng. Những điểm ảnh thuộc về phần tiền cảnh có cường độ cao hơn hoặc bằng giá trị ngưỡng cho trước, và ngược lại là những điểm ảnh thuộc về nền ảnh, ví dụ các kỹ thuật k-mean, kỹ thuật P-Tile, Histogram-shaped-based... Tuy nhiên, nếu hình ảnh phức tạp và chứa nhiều đối tượng thì ngưỡng hai cấp không hiệu quả (Khairuzzaman, & Chaudhury, 2017). Vì vậy, ngưỡng nhiều cấp được sử dụng để phân đoạn một hình ảnh và việc lựa chọn các giá trị thích hợp của các ngưỡng là rất quan trọng để có được một phân đoạn tốt. Các kỹ thuật lựa chọn ngưỡng tối ưu tìm kiếm các ngưỡng bằng cách tối ưu hóa một hàm mục tiêu. Phương pháp Kapur (Manic, Priya, & Rajinikanth, 2016) và phương pháp Otsu (Yang, Shen, Long, & Chen, 2012) dựa trên phương sai tạo ngưỡng tối ưu được sử dụng rộng rãi.

Ứng dụng (Almotiri, Elleithy, & Elleithy, 2018) kỹ thuật phân đoạn ngưỡng được sử dụng rộng rãi trong nghiên cứu liên quan đến phân đoạn hình ảnh y tế, nơi các mô và cơ quan khác nhau trong cơ thể con người được thể hiện ở các mức xám khác nhau. Hoặc ứng dụng (Jain, Prabhakar, Member, & Hong, 1999) trong nhận dạng vân tay sử dụng phương pháp xác thực và xác thực sinh trắc học phổ biến nhất vì khả năng chấp nhận cao, tính bất biến và tính duy nhất. Giai đoạn quan trọng nhất trong hệ thống nhận dạng vân tay tự động đó là phân loại vân tay. Phân loại cho phép một dấu vân tay đầu vào chỉ được so khớp với một tập hợp con của cơ sở dữ liệu. Nhằm giảm bớt sự phức tạp trong tìm kiếm và không gian, việc phân chia một cách có hệ thống cơ sở dữ liệu thành các lớp khác nhau là rất cần thiết.

Hướng tiếp cận phân ngưỡng dựa trên các đỉnh trong biểu đồ histogram để tìm ra giá trị phân ngưỡng cho phân đoạn đối tượng trong ảnh. Vì vậy, ưu điểm hướng tiếp cận này không cần biết thông tin về ảnh cần phân đoạn trước khi phân đoạn ảnh. Nhược điểm của hướng tiếp cận phụ thuộc vào các đỉnh trong biểu đồ histogram và đặc điểm không gian không được cân nhắc.

c. Hướng tiếp cận phương trình vi phân

Phương trình vi phân (Partial Differential Equation – PDE) là phương pháp (Ruthotto, & Haber, 2020) không thể thiếu để mô hình hóa nhiều hiện tượng vật lý và giải quyết phân đoạn ảnh trong xử lý ảnh bằng phương pháp chủ yếu dựa trên những thay đổi về cường độ hoặc màu sắc của ảnh dựa trên mô hình PDE và giải phương trình PDE. Các ví dụ chứng

minh rằng việc tiền xử lí bằng các kĩ thuật PDE cải thiện đáng kể việc phân đoạn ban đầu và phương pháp phân đoạn thu được cho kết quả tốt hơn so với một số kĩ thuật truyền thống (Weickert, 2001).

Ứng dụng phân đoạn ảnh với nhiều đối tượng có dạng lồi trong thế giới thực (Luo, Tai, Huo, Wang, & Glowinski, 2019). Bài báo đề xuất một phương pháp kết hợp hình dạng lồi trước để phân đoạn nhiều đối tượng được phân tích về mặt lí thuyết. Các điều kiện biên đặc biệt cũng được đặt ra để có được thuật toán hiệu quả cho phương trình đạo hàm riêng. Ngoài ra, phương pháp chỉ cần một hàm thiết lập mức không phụ thuộc vào số lượng đối tượng. Vì vậy, sự gia tăng số lượng đối tượng không dẫn đến sự gia tăng của mô hình và độ phức tạp của thuật toán. Ngoài ra, một ứng dụng khác (Chandra, & Bajpai, 2019) cho phép phân đoạn các khối u não lành tính, cho phép phân biệt với khối u ác tính là một trong những căn bệnh nguy hiểm đến tính mạng. Các tế bào khối u lành tính có các đặc điểm rất giống với các tế bào khỏe mạnh xung quanh của nó. Độ chính xác cao đạt được bằng cách sử dụng phương trình đạo hàm riêng trong việc bóc tách vùng u não lành tính.

Hướng tiếp cận phương trình vi phân dựa trên mô hình hoạt động của phương trình vi phân. Vì vậy, ưu điểm hướng tiếp cận này là phương pháp tốt cho các ứng dụng cần ít thời gian xử lí phân đoạn ảnh và phương pháp trên không phụ thuộc vào số lượng đối tượng cần phân đoạn.

d. Hướng tiếp cận gom nhóm

Các điểm ảnh có tính tương đồng sẽ được gom nhóm thành các cụm theo một tiêu chuẩn tương đồng nào đó dựa trên đặc trưng của ảnh như cường độ, kết cấu, màu sắc... và xét thêm quan hệ giữa các điểm ảnh khi phân nhóm và cập nhật giá trị đại diện cho mỗi nhóm. Các kĩ thuật gom nhóm thường dùng như k-Means, FCM (Kamil, & Salih, 2019; Bezdek, 2013)...

Trong ứng dụng viễn thám khi sử dụng dữ liệu có độ phân giải không gian cao. Một trong những bước đầu tiên của phân tích ảnh dựa trên đối tượng là tạo ra các vùng đồng nhất từ ảnh. Bài báo sử dụng thuật toán phân cụm k-means cho phép phân đoạn ảnh dựa trên vùng tự động, được thiết kế đặc biệt cho các ứng dụng viễn thám (Wang, Jensen, & Im, 2010). Một ứng dụng (Wang, Kong, Lu, Qi, & Zhang, 2008) phân đoạn ảnh trong giai đoạn đầu của quá trình phân tích lâm sàng hình ảnh não MRI. Để giảm hiệu ứng nhiễu trong quá trình phân đoạn, phương pháp được đề xuất kết hợp cả bối cảnh không gian cục bộ và thông tin không cục bộ vào thuật toán phân cụm FCM. Hiệu quả của thuật toán đề xuất được chứng minh bằng các thực nghiệm so sánh với các thuật toán hiện đại khác. Ngoài ra, có giải pháp cải tiến kĩ thuật gom cụm FCM để phân đoạn hình ảnh não MRI (Chen, Zhang, Zheng, Jeon, & Wu, 2016; Ding, & Fu, 2016; Huang, Meng, Zhou, Jiang, & Manogaran, 2019; Bai, Zhang, Liu, & Chen, 2019; Kouhi, Seyedarabi, & Aghagolzadeh, 2020).

Hướng tiếp cận gom cụm dựa trên sự phân chia thành các cụm đồng nhất. Vì vậy, hướng tiếp cận này cho thấy kết quả vượt trội so với các phương pháp hiện có khác nhưng nhạy với nhiễu nên chưa cải thiện độ chính xác.

e. Hướng tiếp cận phân hoạch đồ thị

Các hướng tiếp cận trên giải quyết tốt bài toán phân đoạn nếu các đối tượng bên trong ảnh phân biệt rõ, ngược lại các đối tượng có độ tương phản gần nhau (nghĩa là biên giữa 2 đối tượng có độ tương phản giống nhau), các phương pháp trên rất khó phân biệt trong quá trình phân đoạn ảnh. Tuy nhiên, hướng tiếp cận phân hoạch đồ thị sẽ giải quyết tốt bài toán này. Cụ thể, xây dựng đồ thị mà mỗi node là một điểm ảnh, cạnh nối giữa hai điểm ảnh này chính là độ chênh lệch. Bài toán phân đoạn ảnh giờ trở thành bài toán tìm các đồ thị con, với mỗi đối tượng được phân đoạn là 1 cụm đồ thị con.

Việc nghiên cứu phân đoạn ảnh dựa vào lý thuyết đồ thị đã có nhiều kỹ thuật tiên bộ. Cụ thể, nhiều thuật toán mới ra đời đã chứng minh rằng phương pháp này là một hướng nghiên cứu đầy hứa hẹn trong cộng đồng nghiên cứu phân đoạn ảnh. Các phương pháp dựa trên lý thuyết đồ thị đã có nhiều ứng dụng quan trọng trong lĩnh vực thị giác máy tính, đặc biệt là các ứng dụng trong phân đoạn ảnh (Zhu, Zhang, Xu, & Deng, 2021; Hawas, Guo, Du, Polat, & Ashour, 2020; Bejar, Guimaraes, & Miranda, 2020; Wang, Oda, Hayashi, Yoshino, Yamamoto, Frangi, & Mori, 2020). Dựa trên các nguyên lý cấu trúc hình thức về mức độ tương đồng hoặc lân cận trong việc gom nhóm thị giác, đồ thị được chia thành các vùng tương ứng với một vùng hay đối tượng nào đó trong ảnh. Kết quả phân đoạn của các phương pháp này chỉ sử dụng ngưỡng cố định và tính toán cục bộ, chưa thể hiện được tính chất toàn cục, nên các phân vùng thường chia nhỏ và không như mong muốn. Những năm gần đây, phương pháp phân đoạn ảnh bằng đồ thị sử dụng hàm tính toán cục bộ được đề xuất với kết quả phân đoạn cuối cùng tốt hơn và đáp ứng được yêu cầu của nhiều ứng dụng thị giác (Zhang, Zhang, & Peng, 2013; Huttenlocher, & Felzenszwalb, 2004):

- Phương pháp tìm cây bao trùm tối thiểu của đồ thị (Minimal Spanning Tree – MST): các thuật toán MST cho phân đoạn ảnh định nghĩa ảnh bằng đồ thị có trọng số theo không gian đặc trưng nào đó, xem việc gom nhóm các điểm ảnh được thực hiện như việc tìm cây bao trùm tối thiểu của đồ thị tương ứng với ảnh đầu vào. Đồ thị được chia thành các đồ thị con bằng cách xóa đi các cạnh sao cho thỏa tổng trọng số nhỏ nhất.

- Phương pháp graph-cut với các hàm giá trị: tương tự với MST cũng định nghĩa dựa trên đồ thị có trọng số và phân vùng đồ thị mang tính toán cục bộ. Phương pháp graph-cut có thể xem là mẫu chung cho các ứng dụng phân đoạn ảnh bằng kỹ thuật phân vùng đồ thị. Đây là một thuận lợi cho các ứng dụng khác nhau vì có thể định nghĩa các tiêu chuẩn ‘cut’ khác nhau, hay tối ưu hóa các hàm tính toán giá trị toàn cục cho việc phân vùng đồ thị cho ứng dụng riêng biệt nào đó, ví dụ tiêu biểu là thuật toán max-flow/min-cut, mô hình Markov random field. Các phương pháp tiêu biểu cho phương pháp graph-cut như minimal cut, Ncuts, mean cut, ratio cut (Wang, & Siskind, 2001; Wang, & Siskind, 2003).

- Phương pháp dựa trên tìm đường đi ngắn nhất của đồ thị: biên của đối tượng hay vùng của ảnh được định nghĩa như tập cạnh là đường đi ngắn nhất giữa hai đỉnh của đồ thị bằng các thuật toán tiêu biểu trong lý thuyết đồ thị như Dijkstra. Việc tìm biên của đối tượng hay

phân vùng trong ảnh được chuyển thành việc tìm đường đi ngắn nhất giữa cặp đỉnh trong đồ thị. Và trong nhiều ứng dụng đòi hỏi có sự tương tác của người dùng như chọn điểm khởi tạo của biên.

Rõ ràng các phương pháp phân đoạn ảnh sử dụng lý thuyết đồ thị đã mang lại hiệu quả đáng kể trong xử lý ảnh số, hầu hết các kết quả đều thể hiện được các phân vùng theo thị giác con người (Zhang, Zhang, & Peng, 2013). Đồ thị được dùng như mô hình toán học mô tả những đặc tính rời rạc của ảnh trong thế giới thực. Hơn nữa, trong lý thuyết đồ thị đã có sẵn nhiều định lý và các phương pháp gom nhóm cũng là vấn đề quan trọng trong phân đoạn ảnh bằng đồ thị. Việc dùng đồ thị để biểu diễn ảnh số cho ta một hướng nghiên cứu hữu hiệu trong phân đoạn ảnh. Tuy nhiên, đa số các thuật toán phân đoạn ảnh dựa trên lý thuyết đồ thị chiếm bộ nhớ và thời gian lớn đối với dữ liệu ảnh đầu vào lớn, điển hình là ảnh y khoa, ảnh thiên văn đòi hỏi phải cho kết quả nhanh và tốt nhất là kết quả tương đối với thời gian thực.

3. Ứng dụng Ncuts cho bài toán phân đoạn ảnh

a. Tổng quan

Thuật toán Ncuts mà nhóm tác giả Shi (Shi, & Malik, 2000) đưa ra là thuật toán phân đoạn ảnh dựa vào lý thuyết đồ thị khá phổ biến trong cộng đồng xử lý ảnh. Nhóm tác giả Shi đã đưa ra một tiêu chuẩn mới cho việc tìm phân vùng tối ưu – tiêu chuẩn Ncuts, không chỉ tập trung trên đặc tính cục bộ trên ảnh mà còn tập trung vào việc phân vùng và gom nhóm dựa trên đặc trưng toàn cục ảnh. Thuật toán Ncuts là thuật toán sử dụng phương pháp phân vùng dựa trên lý thuyết đồ thị. Tập dữ liệu được thể hiện một mối quan hệ tương ứng bằng việc chuyển tập điểm dữ liệu ban đầu thành một đồ thị trọng số vô hướng $G = (E, V)$. Mỗi đỉnh của đồ thị G thể hiện một điểm trong không gian đặc trưng. Mỗi cạnh được hình thành giữa hai đỉnh với trọng số $w(i, j)$ là hàm tính khoảng cách giữa hai đỉnh i và j theo nhiều cách dựa trên đặc trưng của ảnh và ứng dụng.

Nhóm tác giả Shi đề xuất tiêu chuẩn mới – tiêu chuẩn Ncuts để khắc phục trường hợp chỉ quan tâm đến giá trị tổng trọng lượng các cạnh giữa hai vùng, tiêu chuẩn Ncuts tính giá trị ‘cut’ theo tỉ số trên tổng trọng lượng kết nối đến tất cả các đỉnh trong đồ thị, công thức (1).

$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)} \tag{1}$$

Với $assoc(X, V) = \sum_{u \in X, v \in V} w(u, v)$ là tổng trọng lượng của các kết nối từ đỉnh trong A đến tất cả các đỉnh trong đồ thị.

Đặt x là vector có $|V|$ chiều, đại diện cho các phân vùng của đồ thị trong hai tập A và B . Ta có $x_i = 1$ nếu đỉnh i thuộc A và $x_i = -1$ nếu đỉnh i thuộc B . $Ncut(x)$ được định nghĩa như $Ncut(A, B)$.

Đặt $d(i) = \sum_j(i, j)$ là tổng các liên kết từ đỉnh i tới tất cả các đỉnh còn lại. Nếu $D = diag(d_i)$ và $b = \frac{\sum_{x_i > 0} d_i}{\sum_{x_i < 0} d_i}$, công thức (2) (Shi, & Malik, 2000),

$$\min_x Ncut(x) = \min_y \frac{y^T(D - W)y}{y^T D y} \tag{2}$$

Với $y_i \in \{1, -b\}$ và $y^T D 1 = 0$. Gọi y là biến được tính theo công thức (3).

$$y = \frac{1 + x}{2} - b \frac{1 - x}{2} \tag{3}$$

Từ công thức (3), có thể thấy rằng $y_i = 1$ nếu đỉnh i thuộc A và $y_i = -b$ nếu đỉnh i thuộc B. Công thức (2) có thể được tối thiểu hóa bằng cách giải bài toán tìm trị riêng bằng phương trình (4).

$$(D - W)y = \lambda D y \tag{4}$$

Phương trình (4) có thể được viết lại thành (5).

$$D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}z = \lambda z, \text{ với } z = D^{\frac{1}{2}}y. \tag{5}$$

Một số tính chất:

- $D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}$ là ma trận các trục đường chéo đối xứng với các giá trị không âm, nghĩa là ma trận này có dạng bán xác định dương, vì vậy, ma trận chỉ có các trị riêng dương.
- $D^{\frac{1}{2}}1$ một vector riêng của phương trình (5) với trị riêng là 0.
- Vì ma trận $D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}$ đối xứng, nên các vector riêng của nó trực giao với nhau.

Cụ thể vector riêng của phương trình (5) ở trị riêng nhỏ thứ hai phải trực giao với $D^{-\frac{1}{2}}1$ và thỏa điều kiện $y_1^T D 1 = 0$ với $z_1 = D^{\frac{1}{2}}y_1$.

Theo định lý về Rayleigh quotient (Golub, & Loan, 2013), nếu A là một ma trận đối xứng thực, với điều kiện x trực giao với $j - 1$ vector riêng nhỏ nhất x_1, \dots, x_{k-1} . Thương số $\frac{x^T A x}{x^T x}$ được tối thiểu hóa bằng vector x_j tương ứng với trị riêng nhỏ nhất λ_j .

Kết quả thu được công thức (6) và (7).

$$z_1 = \arg. \min_{z^T z_0 = 0} \frac{z^T D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}z}{z^T z} \tag{6}$$

$$\text{Do đó } y_1 = \arg. \min_{y^T D 1 = 0} \frac{y^T(D-W)y}{y^T D y} \text{ và } z_0 = D^{\frac{1}{2}}1 \tag{7}$$

Tập các phần tử có thể được chia thành hai tập con, một tập chứa các phần tử mang giá trị dương trong vector riêng và tập còn lại chứa các phần tử có giá trị âm trong vector riêng. Cách chia này được lặp lại cho đến khi đạt được số nhóm cần phân đoạn. Nhóm tác giả J. Shi (Shi, & Malik, 2000). cho rằng, vector riêng của trị riêng nhỏ nhất thứ hai y trong phương trình (5) xấp xỉ một lời giải tối ưu của bài toán tối thiểu hóa Ncuts, được mô tả thông qua công thức (8).

$$\inf_{y^T D 1 = 0} \frac{\sum_i \sum_j (y(i) - y(j))^2 w_{ij}}{\sum_i y(i)^2 d(i)} \tag{8}$$

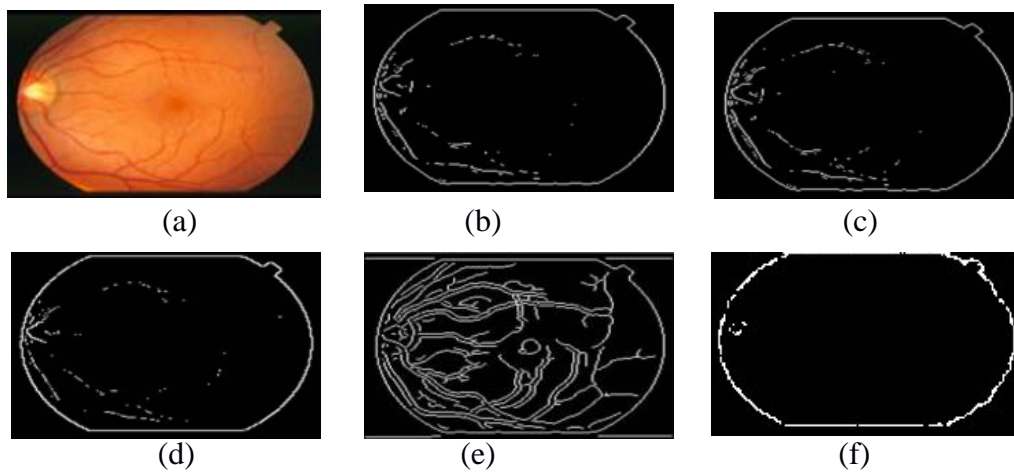
Cuối cùng, để phân vùng đồ thị, ta có thể đặt một ngưỡng trên vector riêng này để chia

đôi đồ thị và đệ qui theo kiểu two-way-cut (Shi, & Malik, 2000) để đạt kết quả.

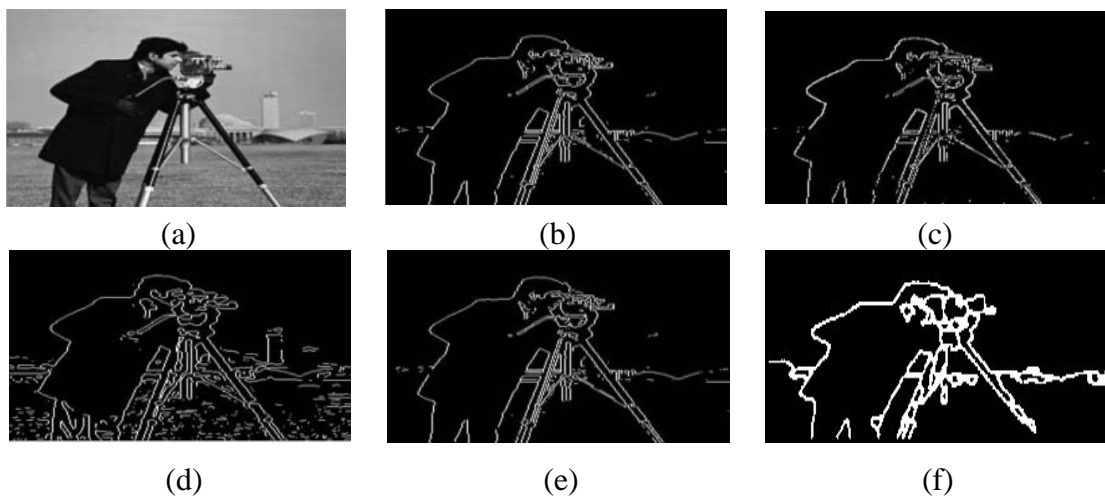
b. Thực nghiệm

Chúng tôi thực nghiệm trên máy tính sử dụng hệ điều hành Windows 10 Pro bản 64bit, RAM 8 GB, Chip Intel Core (TM) 5i-3210M CPU @ 2.5GHz. Ngôn ngữ lập trình Matlab phiên bản R2016a.

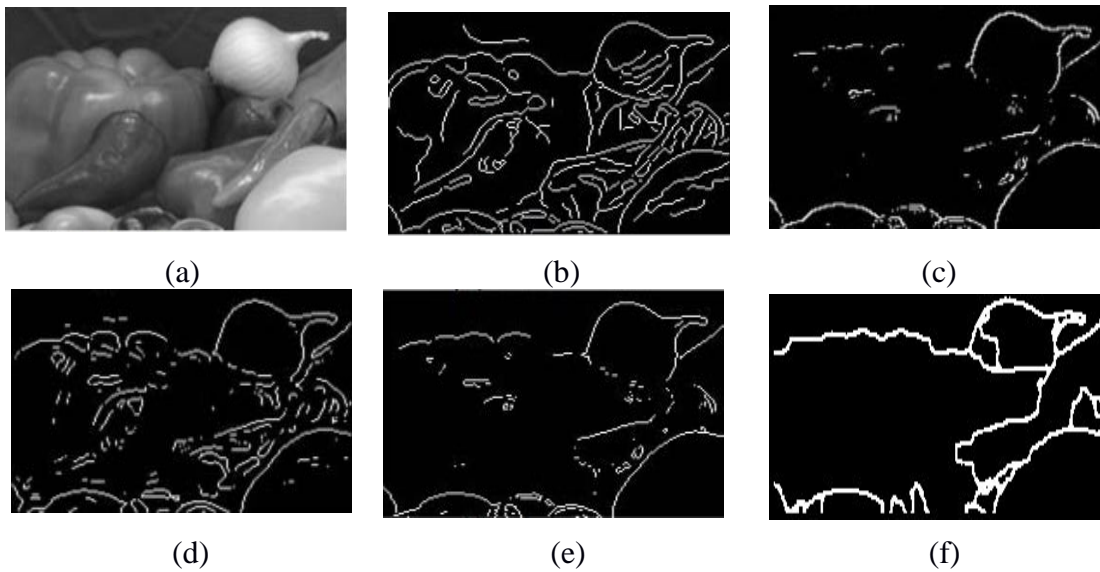
Kết quả thực nghiệm các phương pháp Prewitt, Roberts, Sobel, Laplacian of Gaussian, Ncuts nhằm phát hiện cạnh/biên đối tượng qua dữ liệu Hình 1, 2, 3, 4 được sử dụng nguồn tại (Zaitoun, & Aqel, 2015; Ganesan, & Sajiv, 2017).



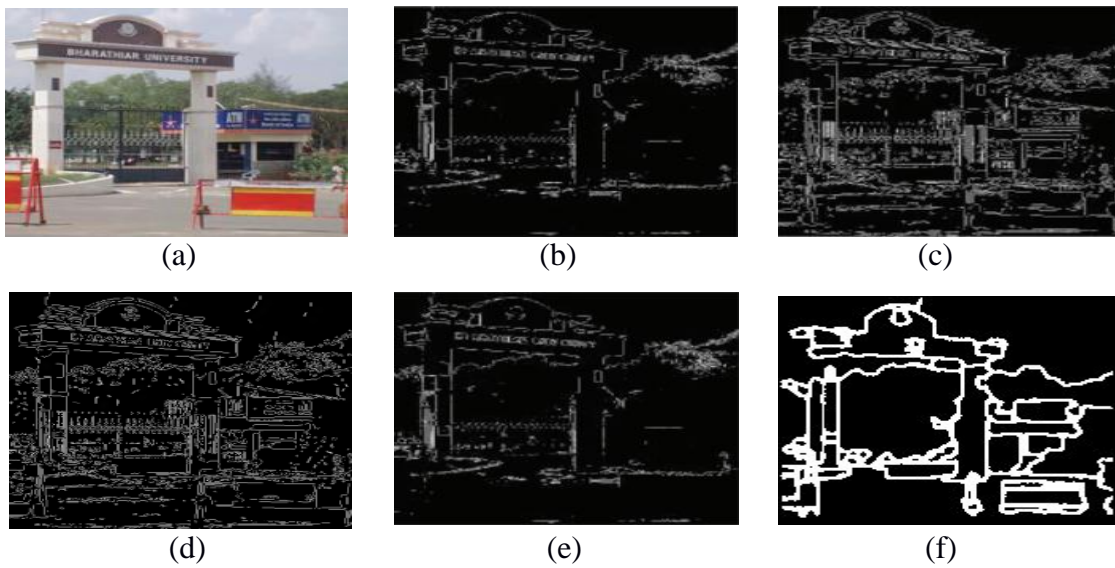
Hình 1. Kết quả phân đoạn cạnh ảnh võng mạc: (a) ảnh gốc, (b) Prewitt, (c) Roberts, (d) LoG, (e) Sobel, (f) NCuts



Hình 2. Kết quả phân đoạn cạnh ảnh cameraman: (a) ảnh gốc, (b) Prewitt, (c) Roberts, (d) LoG, (e) Sobel, (f) NCuts



Hình 3. Kết quả phân đoạn cạnh ảnh củ quả: (a) ảnh gốc, (b) Prewitt, (c) Roberts, (d) LoG, (e) Sobel, (f) Ncuts



Hình 4. Kết quả phân đoạn cạnh ảnh cổng đại học Bharathiar: (a) ảnh gốc, (b) Prewitt, (c) Roberts, (d) LoG, (e) Sobel, (f) Ncuts

Nhận xét: nhóm Hình 1, 2 dấu hiệu sai khác các đối tượng nhiều trong ảnh phương pháp Prewitt, Roberts, Sobel, Laplacian of Gaussian, Ncuts cho kết quả độ chính xác tốt. Nhóm Hình 3, 4 độ tương phản giữa các đối tượng sai khác không nhiều trong ảnh thì phương pháp Ncuts cho kết quả tốt nhất.

c. Ứng dụng

Phương pháp Ncuts (Dimauro, & Simone, 2020; Guo, Ng, Goubran, Petersen, Piechnik, Neubauer, & Wright, 2020) được sử dụng để tự động phân đoạn kết mạc mí mắt

cho phép chẩn đoán bệnh thiếu máu. Hay trong việc chăm sóc bệnh nhân lâm sàng liên quan đến tim mạch bằng ảnh MRI tim. Để tăng cường phân đoạn đa lớp MRI tim bằng cách kết hợp các điểm mạnh của CNN và phương pháp Neuts phân đoạn cấu trúc tim. Ứng dụng (Wang, Lin, & Chen, 2019) phương pháp Neuts để tự động phân đoạn hình ảnh soi. Với độ chính xác của phân đoạn cạnh so với phương pháp đo thủ công và phương pháp xử lý ảnh khác. Ứng dụng này có vai trò quan trọng trong nghiên cứu định lượng về sự phân bố cỡ hạt, có ý nghĩa lớn đối với địa chất, thủy văn và sinh thái học.

4. Kết luận

Với ảnh có độ tương phản giữa các đối tượng sai khác không nhiều, thuật toán Neuts cho kết quả tốt hơn so với các phương pháp gồm Prewitt, Roberts, LoG và Sobel. Tuy nhiên, thuật toán Neuts cũng tồn tại những khuyết điểm:

- Tập ảnh lớn thuật toán thực thi với tốc độ chậm, mất nhiều thời gian và chiếm nhiều bộ nhớ trong tính toán.
- Vấn đề tìm số phân vùng cần phân vùng có thể chủ yếu còn dựa vào cảm tính trực quan, chưa có cơ sở lý thuyết để áp dụng chung.

Chúng tôi nhận thấy quá trình phân đoạn ảnh bằng thuật toán Neuts có hiệu suất thời gian còn thấp; Vì vậy, việc song song hóa thuật toán Neuts trong quá trình tính toán ma trận sẽ cho phép tăng hiệu suất bài toán. Ngoài ra, chúng ta có thể sử dụng 1 số phương pháp gom cụm tìm số phân vùng cần phân vùng, thuật toán Neuts có cơ sở để phân vùng ảnh tốt hơn.

❖ **Tuyên bố về quyền lợi:** Các tác giả xác nhận hoàn toàn không có xung đột về quyền lợi.

TÀI LIỆU THAM KHẢO

- Almotiri, J., Elleithy, K., & Elleithy, A. (2018). Retinal Vessels Segmentation Techniques and Algorithms: A Survey. *Applied Science, MDPI*, 8(2).
- Bai, X., Zhang, Y., Liu, H., & Chen, Z. (2019). Similarity Measure-Based Possibilistic FCM with Label Information for Brain MRI Segmentation. *IEEE Transactions on Cybernetics*, 49, 2618-2630.
- Bejar, H. H. C., Guimaraes, S. J., & Miranda, P. A. V. (2020). *Efficient hierarchical graph partitioning for image segmentation by optimum oriented cuts*. Pattern Recognition Letters, 131, 185-192.
- Bezdek, J. C. (2013). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Springer Science & Business Media.
- Chandra, S. K., & Bajpai, M. K. (2019). Mesh free alternate directional implicit method based three dimensional super-diffusive model for benign brain tumor segmentation. *Computers & Mathematics with Applications*, 77(12), 3212-3223.
- Cheng, H. D., Jiang, X. H., Sun, Y., & Wang, J. L. (2001). Color image segmentation: advances and prospects, *Pattern Recognition*, 34, 2259-2281.

- Chen, Y., Zhang, H., Zheng, Y., Jeon, B., & Wu, Q. M. J. (2016). An improved anisotropic hierarchical fuzzy c-means method based on multivariate student t-distribution for brain MRI segmentation. *Pattern Recognition*, 60, 778-792.
- Dass, R., Priyanka, & Devi, S. (2012). Image Segmentation Techniques. *International Journal on Electronics & Communication Technology*, 3(1).
- Dhankhar, P., & Sahu, N. (2013). A Review and Research of Edge Detection Techniques for Image Segmentation. *International Journal of Computer Science and Mobile Computing*, 2(7), 86-92.
- Ding, Y., & Fu, X. (2016). Kernel-based fuzzy c-means clustering algorithm based on genetic algorithm. *Neurocomputing*, 188, 233-238.
- Dimauro, G., & Simone, L. (2020). Novel Biased Normalized Cuts Approach for the Automatic Segmentation of the Conjunctiva. *MDPI Multidisciplinary Digital Publishing Institute*, 9(6).
- Ganesan, P., & Sajiv, G. (2017). *A comprehensive study of edge detection for image processing applications*. International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS 2017), 2018, 1-6.
- Golub, G. H. & Loan, C. F. V., (2013). *Matrix Computations*. John Hopkins University Press.
- Guo, F., Ng, M., Goubran, M., Petersen, S. E., Piechnik, S. K., Neubauer, S., & Wright, G. (2020). Improving cardiac MRI convolutional neural network segmentation on small training datasets and dataset shift: A continuous kernel cut approach. *Medical Image Analysis*, 61.
- Hawas, A. R., Guo, Y., Du, C., Polat, K., & Ashour, A. S. (2020). OCE-NGC: A neutrosophic graph cut algorithm using optimized clustering estimation algorithm for dermoscopic skin lesion segmentation. *Applied Soft Computing*, 86.
- Hoang, N. D., & Nguyen, Q. L. (2018). *Metaheuristic Optimized Edge Detection for Recognition of Concrete Wall Cracks: A Comparative Study on the Performances of Roberts, Prewitt, Canny, and Sobel Algorithms*. Advances in Civil Engineering, 2018.
- Huang, H., Meng, F., Zhou, S., Jiang, F., & Manogaran, G. (2019). Brain Image Segmentation Based on FCM Clustering Algorithm and Rough Set. *IEEE Access*, 7, 12386-12396.
- Huttenlocher, D. P., & Felzenszwalb, P. F. (2004). Efficient graph based image segmetnation. *International Journal of Computer Vision*, 59(2), 167-181.
- Jain, A. K., Prabhakar, S., Member, S., & Hong, L. (1999). A Multichannel Approach to Fingerprint Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(4), 348-359.
- Kamil, M. Y., & Salih, A. M. (2019). Mammography Images Segmentation via Fuzzy C-mean and K-mean. *International Journal of Intelligent Engineering and Systems*, 12(1).
- Khairuzzaman, A. K. M., & Chaudhury, S. (2017). Multilevel thresholding using grey wolf optimizer for image segmentation. *Expert Systems with Applications*, 86, 64-76.
- Kirti, & Bhatnagar, A. (2017). Image Segmentation Using Canny Edge Detection Technique. *International Journal of Techno-Management Research*, 4(4).
- Kouhi, A., Seyedarabi, H., & Aghagolzadeh, A. (2020). Robust FCM clustering algorithm with combined spatial constraint and membership matrix local information for brain MRI segmentation. *Expert Systems with Applications*, 146.
- Luo, S., Tai, X. Ch., Huo, L., Wang, Y., & Glowinski, R. (2019). *Convex Shape Prior for Multi-object Segmentation Using a Single Level Set Function*. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 613-621.

- Manic, K. S., Priya, R. K., & Rajinikanth, V. (2016). Image Multithresholding based on Kapur/Tsallis Entropy and Firefly Algorithm. *Indian Journal of Science and Technology*, 9(12).
- Misal, A., & Singh, M. (2013). A survey paper on various visual image segmentation techniques. *International Journal of Computer Science and Management Research*, 2.
- Ruthotto, L., & Haber, E. (2020). Deep Neural Networks Motivated by Partial Differential Equations. *Journal of Mathematical Imaging and Vision*, 62, 352-364.
- Senthilkumaran, N. & Rajesh, R. (2009). Edge Detection Techniques for Image Segmentation A Survey of Soft Computing Approaches. *International Journal of Recent Trends in Engineering*, 1(2).
- Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), 146-165.
- Shi, J., & Malik, J. (2000). Normalized cuts and Image Segmentation. Pattern Analysis and Machine Intelligence. *IEEE Trans.*, 22, 888-905.
- Wang, Z., Jensen, J. R., & Im, J. (2010). An automatic region-based image segmentation algorithm for remote sensing applications. *Environmental Modelling & Software*, 25, 1149-1165.
- Wang, J., Kong, J., Lu, Y., Qi, M., & Zhang, B. (2008). A modified FCM algorithm for MRI brain image segmentation using both local and non-local spatial constraints. *Computerized Medical Imaging and Graphics*, 8, 685-698.
- Wang, S., & Siskind, J. M. (2001). *Image Segmentation with Minimum Mean Cut*. IEEE International Conference on Computer Vision (ICCV 2001), 1, 517-524.
- Wang, C., Oda, M., Hayashi, Y., Yoshino, Y., Yamamoto, T., Frangi, A. F., & Mori, K. (2020). Tensor-cut: A tensor-based graph-cut blood vessel segmentation method and its application to renal artery segmentation. *Medical Image Analysis*, 60.
- Wang, S., & Siskind, J. M. (2003). Image Segmentation with Ratio Cut. Pattern Analysis and Machine Intelligence, *IEEE Trans.*, 25(6), 675-690.
- Wang, C., Lin, X., & Chen, C. (2019). Gravel Image Auto-Segmentation Based on an Improved Normalized Cuts Algorithm. *Journal of Applied Mathematics and Physics*, 7(3).
- Weickert, J. (2001). Efficient image segmentation using partial differential equations and morphology. *Pattern Recognition*, 34(9), 1813-1824.
- Yang, X., Shen, X., Long, J., & Chen, H. (2012). *An Improved Median-based Otsu Image Thresholding Algorithm*. AASRI Procedia, 3, 468-473.
- Zaitoun, N. M., & Aqel, M. J. (2015). Survey on Image Segmentation Techniques. International Conference on Communication. *Management and Information Technology (ICCMIT 2015)*, 797-806.
- Zhang, L., Zhang, D., & Peng, B. (2013). A survey of graph theoretical approaches to image segmentation. *Pattern Recognition*, 46(3), 1020-1038.
- Zhu, H., Zhang, J., Xu, G., & Deng, L. (2021). Tensor Field Graph-Cut for Image Segmentation: A Non-Convex Perspective. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(3), 1103-1113.

IMAGE SEGMENTATION AND NCUTS

Tran Nhu Y¹, Nguyen Viet Hung^{2*}, Nguyen Quoc Huy³, Pham The Bao³

¹Faculty of Information Technology, Ho Chi Minh city, University of Food Industry, Vietnam

²Faculty of Information Technology, Ho Chi Minh city, University of Education, Vietnam

³Faculty of Information Technology, Saigon University, Vietnam

**Corresponding author: Nguyen Viet Hung – Email: hungnv@hcmue.edu.vn*

Received: April 17, 2021; Revised: May 13, 2021; Accepted: June 10, 2021

ABSTRACT

In the past decades, many studies have been conducted in computer vision and image segmentation. Image segmentation is the process of image preprocessing in most image processing applications. We summarize and evaluate image segmentation techniques and categorize them, including edge detection, thresholding, partial differential equation, clustering, and graph-based. Next, we present the pros and cons of the Ncuts algorithm, which is typical of graph-based image segmentation. The Ncuts algorithm was introduced in 2000 and showed optimal results for image processing and other applications.

Keywords: eigenvalue; graph-cut; Ncuts