

ỨNG DỤNG CẮT ẢNH TỰ ĐỘNG TRONG THIẾT KẾ THẺ SINH VIÊN

Nguyễn Văn Lễ*, Phạm Nguyễn Huy Phương, Vũ Văn Vinh

Trường Đại học Công nghiệp Thực phẩm TP.HCM

*Email: *lecntp@gmail.com*

Ngày nhận bài: 25/9/2018; Ngày chấp nhận đăng: 15/11/2018

TÓM TẮT

Hình ảnh người trên các dạng thẻ như thẻ sinh viên, thẻ học sinh, thẻ nhân viên... là một trong những thành phần quan trọng để nhận diện người. Hiện nay, cách phổ biến để xử lý các hình thẻ này là cắt thủ công trên các phần mềm xử lý đồ họa (Adobe Photoshop, Corel, MS Paint,...) dẫn đến mất nhiều thời gian khi xử lý tập ảnh lớn và chưa có căn cứ để xác định vị trí của đối tượng bên trong ảnh. Trong bài báo này, nhóm tác giả đề xuất giải pháp kết hợp thuật toán phát hiện khuôn mặt và thuật toán phát hiện biên để cắt ảnh thẻ tự động theo tỷ lệ kích thước 3:4. Trong đó, vị trí đối tượng bên trong ảnh được tính sao cho cách đều 2 biên dọc và cách biên ngang phía trên của vùng cắt một khoảng cách phù hợp. Phương pháp đề xuất thử nghiệm trên tập ảnh có nền đồng nhất đạt được tỷ lệ chính xác cao.

Từ khóa: Phát hiện khuôn mặt, biên Canny, cắt ảnh, cắt ảnh tự động, phát hiện biên.

1. GIỚI THIỆU

Hiện nay, hình ảnh người được sử dụng trong các dạng thẻ như thẻ học sinh, sinh viên, nhân viên... chỉ quy định kích thước khung ảnh bằng 2 x 3 cm, 3 x 4 cm, 4 x 6 cm mà chưa quan tâm đến vị trí, kích thước của người bên trong ảnh, điều này dẫn đến sự không đồng đều giữa các đối tượng trong các ảnh thẻ và một số ảnh bị lệch đối tượng (Hình 1). Tuy nhiên, Tổ chức Hàng không Dân dụng Quốc tế có các tiêu chuẩn quy định rất chặt chẽ về ảnh chụp sử dụng trong hộ chiếu và đã được rất nhiều nước trên thế giới áp dụng [1]. Theo tiêu chuẩn này, chiều rộng ảnh 35-40 mm, khuôn mặt chiếm từ 70-80% ảnh, mặt nhìn thẳng về trước... Trong bài báo này, nhóm tác giả đề xuất phương pháp cắt ảnh chứa đối tượng người một cách tự động theo tỷ lệ kích thước ảnh 3:4 từ một ảnh kỹ thuật số bất kỳ. Trong đó, vị trí của đối tượng người được tính cân đối với các đường biên của vùng cắt hình chữ nhật. Thuật toán Viola – Jones được sử dụng để phát hiện và xác định tọa độ của khuôn mặt người có trong ảnh [2]. Thuật toán này phát hiện được nhiều khuôn mặt trong một ảnh với nền ảnh bất kỳ. Tuy nhiên, vùng cắt ảnh thẻ được xác định là một hình chữ nhật duy nhất nên tập ảnh đầu vào được chọn để phát hiện khuôn mặt là các ảnh đơn, nghĩa là mỗi ảnh gồm một đối tượng người. Để xác định vị trí đối tượng, nhóm tác giả sử dụng thuật toán Canny Edge Detection [3], thuật toán này tìm ra các đường biên của đối tượng (Hình 7b). Dựa vào tập đường biên này để xác định vị trí đỉnh đầu và tính khoảng cách từ đỉnh đầu đến biên ngang phía trên của khung ảnh.



(a) (b) (c) (d)
Hình 1. Ảnh thẻ với đối tượng người bị lệch
(a): lệch trái, (b): lệch phải, (c) và (d): lệch trên



Hình 2. Ảnh gốc ban đầu



Hình 3. Kết quả ảnh sau khi cắt tự động theo tỷ lệ 3:4

2. CÁC CÔNG TRÌNH LIÊN QUAN

Năm 2003, Bongwon Suh và B.Bederson đề xuất phương pháp cắt ảnh thu nhỏ tự động (Thumbnail Cropping) dựa trên những điểm nổi bật của đối tượng (Saliency Map) [4]. Phương pháp này tìm ra hình chữ nhật tối ưu bao quanh đối tượng với các ngưỡng cho trước và cắt thành ảnh thu nhỏ tương ứng. Ngoài ra, kết hợp phát hiện khuôn mặt (face detection) để cắt ảnh chứa đối tượng người hỗ trợ cho các hệ thống nhận dạng. Năm 2005, Mingju Zhang và Lei Zhang đề xuất phương pháp cắt ảnh tự động dựa trên 14 loại mẫu hình ảnh được xác định trước, kết hợp với kết quả phát hiện khuôn mặt và phát hiện các vùng nổi bật để cắt ảnh người theo tỷ lệ cho trước như 2:1, 4:3, 1:1, 3:4, 2:3 [5]. Tuy nhiên, việc dựa vào các điểm nổi bật để xác định vùng cắt đôi khi dẫn đến kết quả không chính xác vì một số đối tượng trong ảnh rất nổi bật nhưng không phải là đối tượng trung tâm. Năm 2006, A.Santella và D.DeCarlo đề xuất phương pháp cắt ảnh bán tự động dựa trên tương tác góc nhìn. Xác định vùng cắt bằng cách phân đoạn ảnh kết hợp với kỹ thuật Eye Tracking [6]. Năm 2007, Fred Stentiford đề xuất phương pháp cắt ảnh tự động dựa trên điểm chú ý [7]. Phương pháp này dựa trên nền tảng Saliency Map và tìm vùng chứa các điểm chú ý là các điểm có sự chênh lệch màu sắc lớn hơn một ngưỡng xác định. Ngoài ra, kết hợp với hệ số phóng to để cắt ảnh chứa các điểm chú ý theo nhiều kích thước khác nhau. Năm 2009, Nishiyama *et al.* trình bày phương pháp cắt hình ảnh dựa trên cảm giác [8]. Phương pháp này sử dụng Saliency Map để trích xuất các đối tượng có trong ảnh, mỗi đối tượng được chỉ định một

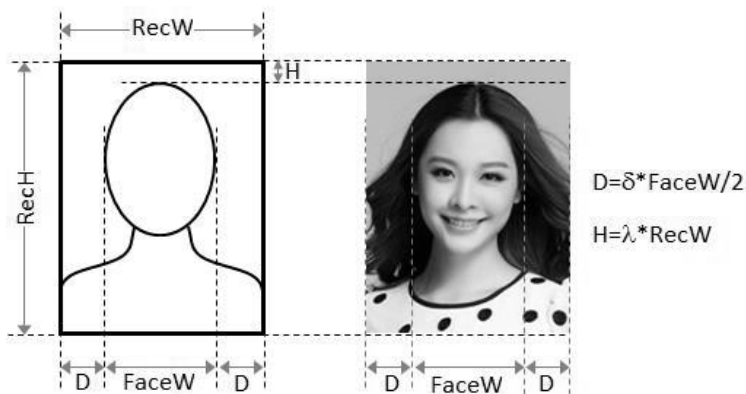
vùng cắt tương ứng. Sau đó, chọn vùng cắt phù hợp nhất dựa vào kết quả phân lớp hình ảnh. Nguồn dữ liệu huấn luyện để phân lớp được thu thập từ các trang web lưu trữ ảnh với số lượng lớn. Năm 2010, nhóm tác giả G. Liu và J. Chen đề xuất tối ưu hóa thành phần ảnh bằng cách xác định vùng cắt tối ưu là hình chữ nhật có cùng tỷ lệ với ảnh gốc ban đầu [9]. Phương pháp này thực hiện phân đoạn ảnh để tìm ra các vùng nổi bật (salient regions) và đường nổi bật (prominent lines), sau đó đo lường mức độ thâm mỹ và xác định vùng cắt tối ưu. Năm 2011, J.She và D.Wang đề xuất cắt ảnh tự động dựa trên mã hóa mật độ thấp (Sparse coding) [10]. Phương pháp này thực hiện phân loại ảnh thành 13 danh mục, trích xuất các điểm nổi bật của ảnh dựa trên đồ thị để tạo từ điển cho từng danh mục. Vùng cắt được xác định dựa trên mã hóa mật độ thấp và bộ từ điển. Năm 2013, J.Yan và S.Lin xây dựng 3 bộ huấn luyện liên quan đến các chủ đề ảnh khác nhau dựa trên cặp ảnh gốc và ảnh được cắt bởi các chuyên gia về nhiếp ảnh [11]. Phương pháp này kết hợp phát hiện khuôn mặt, phát hiện điểm nổi bật để trích xuất các đối tượng tiền cảnh, sau đó loại trừ các đối tượng phụ và xác định vùng cắt phù hợp. Gần đây, Chen *et al.* nghiên cứu độ phức tạp tính toán trong việc tìm kiếm vùng cắt tối ưu dựa trên điểm nổi bật của ảnh và chỉ ra 3 vấn đề trong các thuật toán trước đó có độ phức tạp tính toán cao [12]. Các vấn đề là: tìm kiếm vùng cắt tối thiểu, cố định tỷ lệ vùng cắt tìm kiếm và xác định nhiều vùng cắt trên một ảnh cho các đối tượng rời rạc, đồng thời đưa ra thuật toán để giải quyết các vấn đề này.

Các công trình nghiên cứu vừa trình bày cho thấy việc cắt ảnh chủ yếu dựa vào phát hiện các điểm nổi bật (Saliency Map) để khoanh vùng các đối tượng chính có trong ảnh hoặc kết hợp với phát hiện khuôn mặt (đối với ảnh người), từ đó xác định vùng cắt bằng cách tìm kiếm hình chữ nhật tối ưu theo nhiều phương pháp khác nhau. Trong bài báo này, nhóm tác giả đề xuất giải pháp mới để cắt ảnh thẻ, vùng cắt là hình chữ nhật được xác định theo tỷ lệ kích thước 3:4. Phương pháp này kết hợp phát hiện khuôn mặt (Face Detection) và phát hiện biên (Edge Detection) sau đó tính toán đưa ra vùng cắt phù hợp nhất.

3. PHƯƠNG PHÁP CẮT ẢNH THẺ

3.1. Cấu trúc ảnh thẻ

Trong bài báo này, nhóm tác giả đề xuất cấu trúc ảnh thẻ bao gồm kích thước biên ảnh với chiều rộng RecW và chiều cao RecH theo tỷ lệ 3:4; khoảng cách từ đỉnh đầu của người đến biên trên là $H = \lambda * \text{RecW}$; khoảng cách từ hai biên trái và phải của khuôn mặt đến 2 biên trái và phải của khung ảnh là $D = \delta * \text{FaceW} / 2$, với FaceW là độ rộng của khuôn mặt.



Hình 4. Cấu trúc ảnh thẻ

Tìm hệ số λ : Với ảnh thẻ giấy chuẩn kích thước 30 x 40 mm thì khoảng cách K tính từ đỉnh đầu người trong ảnh đến biên trên của khung ảnh phù hợp nhất là trong khoảng từ 2 đến 8 mm. Tuy nhiên, ảnh kỹ thuật số sử dụng đơn vị Pixel nên kích thước các ảnh có thể chênh lệch với nhau tùy thuộc vào độ phân giải của ảnh. Do đó, với một ảnh kỹ thuật số có chiều rộng là $RecW$ pixel thì khoảng cách từ đỉnh đầu người đến biên trên của khung ảnh được tính lại theo công thức $H=(K/30)*RecW$. Đặt $\lambda=K/30$, vậy $H=\lambda*RecW$. Hệ số λ có giá trị phụ thuộc vào K được liệt kê như trong bảng 1. Giá trị λ càng lớn thì khoảng cách H tính từ đỉnh đầu người đến biên trên càng lớn.

Bảng 1. Giá trị hệ số λ

K	2	3	4	5	6	7	8
Hệ số λ	0,067	0,100	0,133	0,167	0,200	0,233	0,267

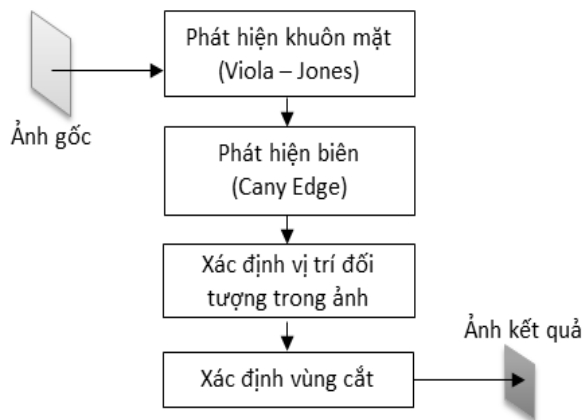
Tìm hệ số δ : Để tìm hệ số này, nhóm tác giả thu thập 100 ảnh thẻ gồm 50 ảnh của sinh viên chụp tại Trường Đại học Công nghiệp Thực phẩm TP.HCM, 50 ảnh thẻ trên mạng internet và tiến hành đo thực nghiệm 2 giá trị gồm: tổng khoảng cách W tính từ 2 biên trái phải của khuôn mặt đến 2 biên trái phải của khung ảnh ($W = 2D$) và độ rộng F của khuôn mặt. Sau đó tính trung bình các giá trị này trên 100 ảnh ($N = 100$).

$$W_{tb} = \frac{\sum_{i=1}^N W_i}{N} \quad F_{tb} = \frac{\sum_{i=1}^N F_i}{N}$$

Ta có công thức tính hệ số δ như sau:

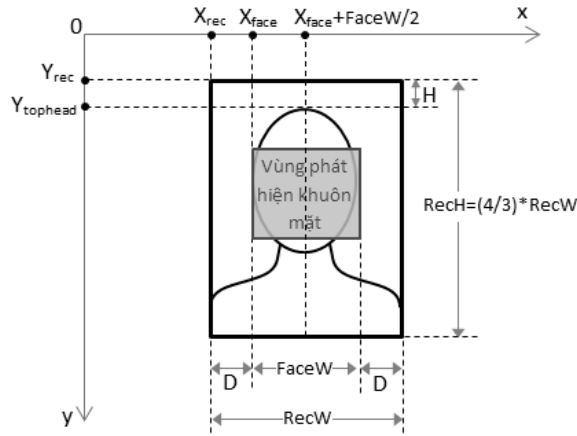
$$\delta = \frac{W_{tb}}{F_{tb}} \approx \frac{1.2}{1.7} \quad (1)$$

3.2. Quy trình xử lý cắt ảnh thẻ



Hình 5. Sơ đồ xử lý cắt ảnh thẻ

Mục tiêu hoạt động của hệ thống là xác định vùng cắt hình chữ nhật trên mỗi ảnh đầu vào, vùng cắt này là duy nhất nên ảnh đầu vào được chọn là ảnh đơn, nghĩa là mỗi ảnh chỉ chứa một đối tượng người. Các thông số của vùng cắt cần xác định gồm: chiều rộng $RecW$, chiều cao $RecH$ và tọa độ của hình chữ nhật này trong không gian 2 chiều gồm: hoành độ X_{rec} và tung độ Y_{rec}



Hình 6. Vùng cắt hình chữ nhật trong hệ trục tọa độ không gian hai chiều

Thuật toán Viola – Jones xác định tọa độ của khuôn mặt với hoành độ X_{face} và hình chữ nhật màu đỏ (Hình 6, Hình 7a) chứa khuôn mặt với độ rộng $FaceW$. Dựa vào độ rộng $FaceW$ của hình chữ nhật này và hệ số δ để tính độ rộng $RecW$ vùng cắt ảnh thẻ:

$$RecW = FaceW + 2D = FaceW + 2 * \frac{\delta * FaceW}{2} = (1 + \delta) * FaceW \quad (2)$$

Chiều cao $RecH$ của vùng cắt được xác định dựa vào chiều rộng $RecW$ theo tỷ lệ 3:4:

$$RecH = \frac{4}{3} * RecW \quad (3)$$

Hoành độ vùng cắt (Hình 6) được xác định theo công thức:

$$X_{rec} = X_{face} + \frac{FaceW}{2} - \frac{RecW}{2} \quad (4)$$

Thuật toán Canny edge detection xác định tất cả các đường biên của đối tượng (Hình 7b). Mục tiêu bước này là tìm đường biên qua đỉnh (Hình 7c). Đường biên qua đỉnh được xác định là đường biên đi qua điểm có tung độ nhỏ nhất $Y_{tophead}$, sau đó ta tính được Y_{rec} theo công thức:

$$Y_{rec} = Y_{tophead} - H = Y_{tophead} - \lambda * RecW \quad (5)$$

Sau cùng là cắt ảnh dựa vào vùng cắt đã xác định: X_{rec} , Y_{rec} , $RecW$, $RecH$



(a) Phát hiện khuôn mặt (b) Phát hiện tất cả biên của đối tượng (c) Tìm đường biên qua đỉnh (d) Vùng cắt với đường nét đứt

Hình 7. Xác định vùng cắt

Đối với ảnh có nền không đồng nhất, một số ảnh sẽ xuất hiện các đường biên phía trên tập đường biên của đối tượng người do nền ảnh sinh ra dẫn đến sai số trong việc xác định $Y_{tophead}$. Để hạn chế sai số này, việc tìm kiếm đường biên qua đỉnh chỉ giới hạn trong một khoảng cho trước ở vị trí phía trên khuôn mặt.

3.3. Đánh giá phương pháp

Để đánh giá độ chính xác của phương pháp cắt ảnh thẻ, nhóm tác giả thu thập 200 ảnh người chụp với tư thế mặt hướng ra phía trước. Trong đó, 100 ảnh được chụp với nền không đồng nhất [13], 100 ảnh còn lại có nền đồng nhất [14] và tiến hành thử nghiệm cắt ảnh thẻ tự động theo tỷ lệ kích thước 3:4 trên 2 tập ảnh này. Các giá trị được ghi nhận gồm: số lượng ảnh cắt đúng (sau khi cắt, đối tượng người được canh cân đối so với các biên của khung ảnh); số lượng ảnh cắt lệch (đối tượng người bị lệch so với các biên của khung ảnh sau khi cắt); số lượng ảnh không xác định (không phát hiện được mặt người trong ảnh); tỷ lệ đúng (tính bằng tỷ lệ phần trăm của số lượng ảnh cắt đúng trên tổng số lượng ảnh). Kết quả thử nghiệm được trình bày trong bảng dưới đây.

Bảng 2. Đánh giá cắt ảnh thẻ

Loại ảnh	Số lượng	Cắt đúng	Cắt lệch	Không xác định	Tỷ lệ đúng
Ảnh có nền không đồng nhất	100	85	10	5	85%
Ảnh có nền đồng nhất	100	93	4	3	93%

Bảng kết quả thử nghiệm trên cho thấy, ảnh có nền đồng nhất có tỷ lệ cắt đúng (93%) cao hơn so với ảnh có nền không đồng nhất (85%). Thông thường các ảnh dùng làm thẻ được chụp với tư thế nhìn thẳng về trước, màu nền đồng nhất nên việc áp dụng nghiên cứu này trong các hệ thống cắt ảnh thẻ sẽ rất phù hợp và cho hiệu suất cao.



Hình 8. Ứng dụng cắt ảnh thẻ trong thiết kế thẻ sinh viên

4. KẾT LUẬN

Nhóm tác giả vừa trình bày phương pháp cắt ảnh thẻ theo tỷ lệ kích thước 3:4 dựa trên sự kết hợp thuật toán Viola – Jones để phát hiện khuôn mặt và thuật toán Canny edge detection nhằm xác định biên đối tượng. Kết quả thử nghiệm trên 2 tập dữ liệu ảnh có nền đồng nhất và có nền không đồng nhất là 93% và 85%. Trong đó tập ảnh có nền đồng nhất cho tỷ lệ chính xác cao hơn, phù hợp để ứng dụng trong thiết kế thẻ sinh viên.

Dựa trên kết quả này, hướng phát triển tiếp theo là xử lý cắt và xoay đối với các ảnh nghiêng, tính tỷ lệ phần trăm đối tượng trong ảnh và một số ràng buộc khác đáp ứng tiêu chuẩn ảnh thẻ được quy định cho thẻ visa hoặc hộ chiếu.

TÀI LIỆU THAM KHẢO

1. Poon J. - A photograph guidelines, International Civil Aviation Organization (ICAO), 2011 ([https://www.icao.int/Security/mrtd/Downloads/Technical Reports/Annex_A-Photograph_Guidelines.pdf](https://www.icao.int/Security/mrtd/Downloads/Technical%20Reports/Annex_A-Photograph_Guidelines.pdf)).
2. Viola P. and Jones M. - Rapid object detection using a boosted cascade of simple features, Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) **1** (2001) 511-518.
3. Canny J. - A computational approach to edge detection, IEEE Transactions on Pattern Analysis And Machine Intelligence **6** (1986) 679–698.
4. Suh B., Ling H., Bederson B.B. and Jacobs D.W. - Automatic thumbnail cropping and its effectiveness, Proceedings of the 16th annual ACM symposium on User interface software and technology (2003) 95-104.
5. Zhang M., Zhang L., Sun Y., Feng L. and Ma W. - Auto cropping for digital photographs, IEEE International Conference on Multimedia and Expo (ICME) (2005).
6. Santella A., Agrawala M., DeCarlo D., Salesin D. and Cohen M. - Gaze-based interaction for semi-automatic photo cropping, Proceedings of the SIGCHI conference on Human Factors in Computing Systems (2006) 771-780.
7. Stentiford F. - Attention based auto image cropping, Workshop on Computational Attention and Applications on International Conference on Computer Vision Systems (ICVS) **1** (2007) 253-261.
8. Nishiyama M., Okabe T., Sato Y. and Sato I. - Sensation-based photo cropping, Proceedings of the 17th ACM International Conference on Multimedia (2009) 669-672.
9. Liu L., Chen R., Wolf L. and Cohen-Or D. - Optimizing photo composition, Computer Graphics Forum **29** (2010) 469-478.
10. She J., Wang D. and Song M. - Automatic image cropping using sparse coding, First Asian Conference on Pattern Recognition (ACPR) (2011) 490-494.
11. Yan J., Lin S., Bing Kang S. and Tang X. - Learning the change for automatic image cropping, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2013) 971-978.
12. Chen J., Bai G., Liang S. and Li Z. - Automatic Image Cropping: A Computational Complexity Study, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016) 507-515.

13. Nguyễn Văn Lễ, 100 ảnh có nền không đồng nhất, 2018
(<https://www.flickr.com/gp/108852439@N03/30W18x>).
14. Nguyễn Văn Lễ, 100 ảnh có nền đồng nhất, 2018
(<https://www.flickr.com/gp/108852439@N03/p8N2j6>).

ABSTRACT

APPLYING AUTO IMAGE CROPPING TO DESIGN STUDENT CARD

Nguyen Van Le*, Pham Nguyen Huy Phuong, Vu Van Vinh
Ho Chi Minh City University of Food Industry
*Email: lecntp@gmail.com

The people image on cards such as student cards, staff cards is one of the important components to identify people. Currently, the common way to create these images is manual image cropping on the graphics processing software (Adobe Photoshop, Corel, MS Paint, etc.), this task takes a long time to process the large numbers of image files and there is baseless to determine the location of the object inside the image. In this paper, the face detection and edge detection algorithm were combined to automatically crop the image at the aspect ratio of 3:4. In which the position of the object inside the image was calculated to equal the distance of both vertical edges side and far away to the upper edge on a suitable distance. The proposed method was tested on a set of images with uniform backgrounds achieved a high accuracy rate.

Keywords: Face detection, Canny edge, image cropping, auto cropping, edge detection.