

## NHẬN DẠNG MẶT NGƯỜI DÙNG POLAR COSINE TRANSFORM VÀ MẠNG RADIAL BASIS FUNCTION

Võ Hoàng Minh<sup>1</sup>, Trần Bình Long<sup>1</sup>, Lê Hoàng Thái<sup>2</sup>, Trần Hành<sup>1</sup>

<sup>1</sup> Trường Đại học Lạc Hồng, Đồng Nai

<sup>2</sup> Trường Đại học Tự nhiên, Đại học Quốc gia Thành phố Hồ Chí Minh

**Tóm tắt.** Nhận dạng mặt người tự động đã được nghiên cứu và ứng dụng rộng rãi trong nhiều lĩnh vực của cuộc sống: bảo mật, chứng thực dựa trên sinh trắc học, quản lý truyền thông đa phương tiện, giải trí điện toán. Mặc dầu nghiên cứu về nhận dạng mặt người tự động đã được thực hiện từ năm 1960, tuy nhiên đến nay bài toán này vẫn là một bài toán khó. Thách thức của việc nhận dạng là khả năng rút trích thông tin hiệu quả dùng để phân biệt khuôn mặt của những người khác nhau theo sự khác biệt dựa trên đặc tính hình học và quang trắc. Trong bài báo này chúng tôi đề xuất mô hình dùng Polar Cosine Transform (PCT) và mạng Radial Basis Function (RBF) để nhận dạng mặt người. Ảnh mặt người sau khi được xử lý, tiếp theo chúng tôi dùng Polar Cosine Transform (PCT), với ưu điểm trực giao, bất biến, để rút trích vector đặc trưng cho mạng Radial Basis Function (RBF). Kết quả thử nghiệm được tiến hành trên cơ sở dữ liệu ảnh YALE, gồm tập ảnh mặt của 15 người khác nhau, mỗi người 11 ảnh, tổng số ảnh là 165, với tỉ lệ nhận dạng đạt 96,92% cho thấy tính khả thi của phương pháp đề xuất.

**Từ khóa:** Nhận dạng mặt người, Polar Cosine Transform, Radial Basis Function.

### 1. Giới thiệu

Ảnh mặt người là đặc trưng sinh trắc học được sử dụng rộng rãi để nhận dạng một người vì tính chất dễ thu thập dữ liệu của ảnh. Ngày nay, người ta dễ dàng trang bị các thiết bị thu nhận ảnh ở bất kì đâu: trong cơ quan, camera ở các điểm công cộng, thiết bị chuyên dụng trong điều tra theo dõi tội phạm, các thiết bị chụp ảnh và quay phim cá nhân cũng ngày càng trở nên phổ biến. Cho đến nay lĩnh vực nhận dạng mặt người đã nhận được rất nhiều sự quan tâm. Nhiều phương pháp đã được đề xuất cho mục đích này như Gabor wavelet [1], PCA [2], LDA [3], ICA [4], Pseudo-Zernike-RBF [5]. . . , và bài toán nhận dạng mặt người nhìn thẳng, trong tư thế nghiêm có thể nói là đã được giải quyết triệt để. Tuy nhiên không phải lúc nào ta cũng thu được những tấm ảnh như ý muốn do những vấn đề liên quan đến nhiễu, che khuất, các trạng thái và góc độ khác nhau của khuôn mặt, .... Từ đó phát sinh vô số vấn đề thách thức, thu hút nhiều sự quan tâm của các nhà nghiên cứu. Trong số các vấn đề liên quan đến nhận dạng ảnh mặt người nhìn thẳng hiện nay có 2 vấn đề lớn hiện vẫn chưa được giải quyết tốt: Một là trường hợp mắt mắt thông tin, ảnh bị

mắt mắt thông tin có thể do ảnh bị nhiễu trong quá trình thu nhận ảnh, khuôn mặt đối tượng bị che khuất do đeo kính, đeo khẩu trang hoặc bị che bởi vật thể khác hoặc do điều kiện sáng tối khi chụp làm ảnh hưởng một vùng nào đó trong ảnh; Hai là trạng thái cảm xúc trên khuôn mặt, ảnh mặt người có thể được chụp ở các trạng thái khác nhau như cười, khóc, buồn, vui quá mức hoặc giận dữ...

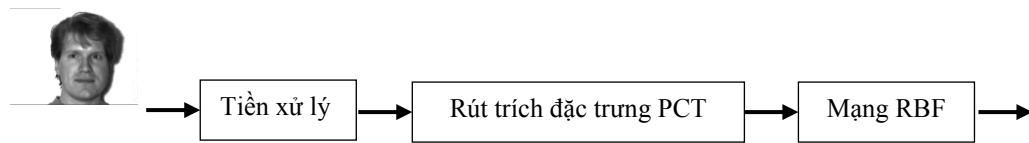
Trên thực tế quá trình ghi hình đối tượng, thường ta không có điều kiện thu được một tấm ảnh rõ ràng, đầy đủ do bị che khuất hoặc bắt buộc khuôn mặt đối tượng ở trạng thái nghiêm nghị được. Do đó, một hệ thống nhận dạng ảnh mặt người nhìn thẳng trong trạng thái bị mất mát thông tin và xác định được trạng thái cảm xúc của khuôn mặt (ở đây, chúng tôi chỉ giới hạn trong các trạng thái khuôn mặt sau: giận dữ, hạnh phúc, ngạc nhiên, sợ hãi, buồn bã, kinh tởm và bình thường) thực sự hữu dụng và có giá trị ứng dụng thực tiễn cao.

Trong phạm vi nghiên cứu của đề tài này, chúng tôi tìm hiểu, thử nghiệm và đề xuất mô hình dùng Polar Cosine Transform (PCT) rút trích đặc trưng khuôn mặt và mạng Radial Basis Function (RBF) để nhận dạng mặt người trong tư thế nhìn thẳng có khả năng giải quyết được 2 vấn đề nêu trên. PCT có ưu điểm hơn các moment trực giao khác: [6] Legendre moments, Orthogonal Fourier Mellin moment [7], Legendre moments, Zernike moments [8] and Pseudo-Zernike moments [9] nhờ vào sự đơn giản nên tốc độ tính nhanh, không tốn nhiều thời gian, đa thức PCT ổn định ít sai số. Chúng tôi hy vọng mô hình này có thể được ứng dụng rộng rãi trong thực tế và mang lại nhiều lợi ích, góp phần hoàn thiện hệ thống nhận dạng mặt người trong tương lai.

Phần còn lại của bài báo được tổ chức như sau: Hệ thống nhận dạng mặt người và tiền xử lý ảnh được giới thiệu trong phần 2. Phần 3 trình bày kỹ thuật rút trích đặc trưng và thiết kế bộ phân lớp. Phần 4 trình bày kết quả nghiên cứu trên cơ sở dữ liệu YALE [10] và phần 5 các kết luận.

## **2. Hệ thống nhận dạng khuôn mặt**

Hệ thống nhận dạng mặt người gồm ba phần. Phần đầu là bước tiền xử lý, ảnh tĩnh mặt người được xử lý để chuẩn hóa, loại bỏ nhiễu, . . . phần hai rút trích các đặc trưng từ ảnh thu được ở phần một. Phần ba phân lớp ảnh mặt người dựa vào các đặc trưng thu được ở phần 2. Hình.1



**Hình 1-**Hệ thống nhận dạng mặt người

Trong hệ thống này chúng tôi dùng phương pháp biến đổi Wavelet cho tiền xử lý làm giảm ảnh hưởng của độ sáng và chuẩn hóa mặt cho việc rút trích đặc trưng bằng cách thay đổi kỹ thuật cân bằng Histogram toàn cục với cân bằng histogram thích nghi giới hạn tương phản.

### 3. Rút trích đặc trưng và phân lớp

Bằng phân tích moment [11], một hình ảnh bất kỳ có thể được mô tả lại đầy đủ, phân tích moment là phương pháp nổi tiếng và được sử dụng nhiều trong xử lý ảnh. Trong bài báo này chúng tôi sử dụng moment biến đổi Polar Cosine[12] cho việc rút trích đặc trưng.

#### 3.1. Polar Cosine Transform (PCT)

Cho một ảnh 2D  $f(x,y)$ , ta có thể chuyển đổi từ hệ tọa độ đề các đến hệ tọa độ cực  $f(r,\theta)$ , trong đó  $r$  và  $\theta$  là bán kính và góc tương ứng. Công thức chuyển đổi từ hệ tọa độ đề các sang hệ tọa độ cực:

$$r = \sqrt{x^2 + y^2} \quad (1)$$

Và 
$$\theta = \arctan\left(\frac{y}{x}\right) \quad (2)$$

Ảnh được định nghĩa trên đường tròn đơn vị với  $r \leq 1$ , và có thể tái tạo với hàm cơ bản  $H_{nl}(r,\theta)$

$$f(r,\theta) = \sum_{n=0}^{\infty} \sum_{l=-\infty}^{\infty} M_{nl} H_{nl}(r,\theta) \quad (3)$$

trong đó hệ số

$$M_{nl} = \frac{\alpha}{\pi} \int_0^{2\pi} \int_0^1 f(r,\theta) H_{nl}^*(r,\theta) r dr d\theta \quad (4)$$

Hàm cơ bản được tính bởi: 
$$H_{nl}(r,\theta) = R_n(r) e^{il\theta} \quad (5)$$

Với 
$$R_n(r) = \cos(\pi n r^2) \quad (6)$$

Và thỏa mãn điều kiện trực giao:

$$\int_0^1 R_n(r) R_{n'}^*(r) r dr = \frac{1}{2} \delta_{nn'} \quad (7)$$

Và 
$$\int_0^{2\pi} \int_0^1 H_{nl}(r,\theta) H_{n'l'}^*(r,\theta) r dr d\theta = \pi \delta_{nn'} \delta_{ll'} \quad (8)$$

Trong đó  $\delta_{ij}$  là Kronecker delta. Từ (4) và (5)(6) viết lại như sau

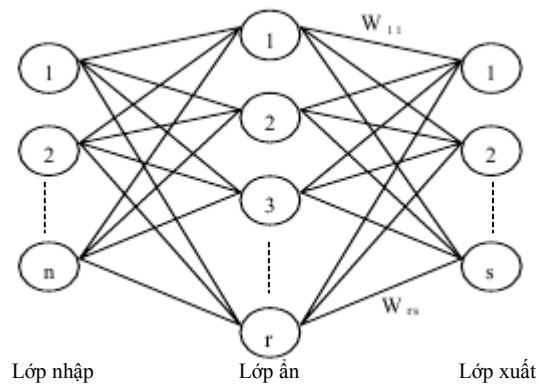
$$M_{nl} = \frac{\alpha}{\pi} \int_0^{2\pi} \int_0^1 f(r, \theta) \cos(\pi nr^2) (\cos(l\theta) - i\sin(l\theta)) r dr d\theta \quad (9)$$

$$\text{Với } \alpha = \begin{cases} 1 & \text{if } n = 0 \\ 2 & \text{if } n \neq 0 \end{cases} \quad (10)$$

Nhờ sự đơn giản và tính ổn định của PCT, qua thực nghiệm với  $n \geq 16$  sẽ cho vec tơ đặc trưng tốt hơn.

### 3.2. Mạng Radial Basis Function

Mạng Radial Basis Function (RBF) [13] có cấu trúc như hình 2, gồm 3 lớp. Lớp nhập là một tập hợp gồm  $n$  nút tương ứng với  $n$  yếu tố của vec tơ đặc trưng được rút trích từ ảnh bằng phương pháp PCT. Các nút nhập này được kết nối đầy đủ với  $r$  nút ẩn của lớp ẩn, số nút ẩn bằng với số nút nhập. Trong lớp ẩn các nút ẩn này được gọi là đơn vị RBF, các nút ẩn này cũng được kết nối đầy đủ với  $s$  nút xuất của lớp xuất. Mỗi nút xuất là một khuôn mẫu đặc trưng đại diện cho các trạng thái của khuôn mặt.



Hình 2 -Mạng Radial Basis Function

Hàm hoạt động của các đơn vị ẩn được thể hiện như sau:

$$R_i(\mathbf{x}) = R_i\left(\frac{\|\mathbf{x}-\mathbf{c}_i\|}{\sigma_i}\right), \quad i = 1, 2, \dots, r \quad (11)$$

Với  $\mathbf{x}$  là vector nhập  $n$  chiều,  $\mathbf{c}_i$  là vector  $n$  chiều gọi là tâm của các đơn vị RBF,  $\sigma_i$  là độ rộng của của các đơn vị RBF và  $r$  là số lượng các đơn vị RBF. Hàm hoạt động đặc trưng của các đơn vị RBF được chọn theo hàm Gaussian với vec tơ trung bình  $\mathbf{c}_i$  và vec tơ biến  $\sigma_i$  theo biểu thức:

$$R_i(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{c}_i\|^2}{\sigma_i^2}\right) \quad (12)$$

Trong biểu thức trên  $\sigma_i^2$  đại diện cho các đường chéo của ma trận hiệp phương sai của hàm Gaussian. Nút xuất thứ  $j$  của nút nhập  $\mathbf{x}$  được thể hiện:

$$\mathbf{y}_j(\mathbf{x}) = \mathbf{b}(j) + \sum_{i=1}^r R_i(\mathbf{x})w_2(i, j) \quad (13)$$

Với  $w_2(i, j)$  là trọng số kết nối của đơn vị RBF thứ  $i$  với nút xuất thứ  $j$  và  $b(j)$  là độ lệch của nút xuất thứ  $j$ . Độ lệch này được bỏ đi để làm giảm độ phức tạp của mạng này, do đó:

$$y_i(\mathbf{x}) = \sum_{j=1}^r R_j(\mathbf{x})w_2(i, j) \quad (14)$$

#### 4. Kết quả nghiên cứu

Để kiểm tra hiệu quả của hệ thống, chúng tôi thực hiện nghiên cứu trên cơ sở dữ liệu Yale.

##### 4.1 Cơ sở dữ liệu Yale

Cơ sở dữ liệu Yale được lấy từ trung tâm thị giác máy tính và điều khiển của Đại học Yale, gồm các hình ảnh của 15 người khác nhau, mỗi người 11 ảnh, tổng số 165 ảnh. Tập ảnh thể hiện các hình dạng hoặc cảm xúc khác nhau của một người: chiếu sáng chính giữa, chiếu sáng bên phải, chiếu sáng bên trái, có đeo kính, không đeo kính, vui, buồn, buồn ngủ, ngạc nhiên, nháy mắt và bình thường. Hình 3



Hình 3 - Những ảnh mặt trong cơ sở dữ liệu Yale

Để đánh giá sự hiệu quả của hệ thống, chúng tôi kiểm tra trên cơ sở dữ liệu Yale. Ảnh được chuẩn hóa và giảm kích thước còn 80x80 pixels, với những điều kiện sáng khác nhau và biểu lộ khác nhau.



Hình 4 - Chuẩn hóa ảnh khuôn mặt

#### 4.2 Kết quả nghiên cứu

Để chuẩn hóa vị trí và độ sáng của ảnh chúng tôi dùng wavelet dựa trên kỹ thuật chuẩn hóa độ sáng. (Hình 4) Tiếp theo, dùng PCT để rút trích tạo véc tơ đặc trưng cho mạng RBF. Chúng tôi dùng 100 ảnh huấn luyện và 65 ảnh test cho mỗi lần thử, vì cơ sở dữ liệu Yale có giới hạn, chúng tôi tiến hành thử nghiệm nhiều lần để tính tỉ lệ nhận dạng trung bình. Kết quả cho thấy tỉ lệ nhận dạng trung bình của phương pháp đề nghị đạt 96.92%. Bảng 1.

**Bảng 1** - Tỉ lệ nhận dạng của phương pháp PCT-RBF

Test	Sáng giữa	Sáng phải	Sáng trái	Đeo kính	Không kính	Vui	Buồn	Buồn ngủ	Ngạc nhiên	Nháy mắt	Bình thường
1	96.98	96.58	96.95	97.01	97.68	96.88	96.45	96.93	96.90	97.27	96.77
2	96.8	97.88	97.76	96.87	96.64	96.46	96.95	96.78	97.20	96.98	97.02
3	97.7	96.85	95.92	97.42	96.84	97.45	97.86	97.25	96.98	96.85	97.15

Kiểm tra kết quả nhận dạng với moment n=20 với tỉ lệ chấp nhận sai (FAR) và từ chối sai (FRR) với ngưỡng (thres) 0.2954, bảng 2

Tỉ lệ thành công tổng công (TSR) được tính theo công thức:

$$TSR = \left( 1 - \frac{FAR+FRR}{\text{total number of accesses}} \right) \times 100\% \quad (15)$$

**Bảng 2** -Kết quả kiểm tra

Moment	thres	FAR(%)	FRR(%)	TSR(%)
PCT	0.2954	2.7998	3.1674	96.92

#### 5. Kết luận

Trong bài báo này chúng tôi đề xuất phương pháp rút trích đặc trưng dùng biến đổi Polar Cosine do khả năng bất biến, ít bị ảnh hưởng bởi nhiễu của nó và mạng nơ ron RBF để phân lớp thu được kết quả tốt. Ở bước tiền xử lý để làm rõ khuôn mặt và chuẩn hóa độ sáng (mức xám) chúng tôi dùng wavelet dựa kỹ thuật chuẩn hóa độ sáng. Phương pháp của chúng tôi đã được test thành công trong cơ sở dữ liệu Yale, với các loại ảnh thường, nhiễu, thay đổi. Kết quả tính toán cho thấy hệ thống của chúng tôi cho tỉ lệ nhận dạng đạt 96,92%. Hy vọng trong tương lai có thể tiếp tục cải tiến thêm và ứng dụng rộng rãi trong các lĩnh vực cần có sự nhận dạng với độ chính xác cao.

### **Tài liệu tham khảo**

- [1] Wei Jiang, Jian Zhang, Ting-zhi Shen, Xiao-hua Wang, “A Novel Facial Features Extraction Algorithm Using Gabor Wavelets”, IEEE Congress on Image and Signal Processing, vol. 2, pp. 649 – 653, 2008.
- [2] M. Turk and A. Pentland, “Face recognition using eigenfaces”, IEEE Conference on Computer Vision and Pattern Recognition, pp. 586 – 591, 1991.
- [3] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection”, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 19, pp. 711 – 720, 1997.
- [4] Bruce A. Draper, Kyungim Baek, Marian Stewart Bartlett, J. Ross BeveRidge, “Recognizing Face with PCA and ICA”, Computer Vision and Image Understanding, pp. 115 – 137, 2003.
- [5] Javad Haddadnia, Majid Ahmadi, Karim Faez, “An efficient feature extraction method with Pseudo-Zernike moment in RBF neural network - based human face recognition system”, Eurasip journal on applied signal processing 2003:9, pp. 890 – 901.
- [6] Liao, S.X., Pawlak, M., (1996) “On image analysis by moments”, IEEE Trans. Pattern Anal. mach. Intell. 18, 254-266.
- [7] Sheng, Y.L. and Shen, L.X., (1994) “Orthogonal Fourier- Mellin moments for invariant pattern recognition,” J. Opt. Soc. Am. A. 11.1748-1757.
- [8] Hu, M.K., (1962) “Visual pattern recognition by moment invariants”, IRE Trans. Inf. Theory IT-8, 179-187.
- [9] Chong, C.W., Raveendran, P. and Mukundan, R. (2003), “The scale invariants of pseudo-Zernike moments,” Pattern Anal. Appl. 6, 176-184.
- [10] Yale University [Online] available, <http://cvc.yale.edu/projects/yalefaces/yaleface.html>
- [11] C. Teh and R. Chin, “On image analysis by the method of moments,” IEEE Trans. PAMI, 10(4):496-513, 1988.
- [12] Zhuo Yang, Sei-ichiro KAMATA, “Fast Polar Harmonic Transforms,” 2010 11th Int. Conf. Control, Automation, Robotics and Vision, pp.673-677, Singapore, 7-10th December 2010
- [13] Yegnanarayana, B. (1999), “Artificial Neural Networks” (First Edition) –by Prentice Hall of India Private Limited, ISBN-81-203-253-8.