

NÂNG CAO CHẤT LƯỢNG ẢNH MÀU MẶT NGƯỜI BỞI SVD CỦA DCT TRONG MIỀN LOGARIT ỨNG DỤNG TRONG HỆ THỐNG NHẬN DẠNG MẶT NGƯỜI

COLOR FACE IMAGE ENHANCEMENT BASED ON SVD OF DCT COEFFICIENTS IN LOGARITHM DOMAIN

Nguyễn Nam Phúc^{1,*},
Nguyễn Quốc Trung², Hà Hữu Huy³

TÓM TẮT

Trong bài báo này, chúng tôi đưa ra một phương pháp cân bằng ánh sáng hữu hiệu để nâng cao chất lượng ảnh mặt người trong không gian màu RGB ứng dụng trong nhận dạng mặt người. Trước tiên, ảnh mặt người trong không gian màu RGB được phân tách thành ba kênh màu và tham chiếu Gaussian được biến đổi sang miền logarit. Sau đó sử dụng biến đổi Cosin rời rạc để xác định các thành phần tần số thấp mang thông tin độ sáng của ảnh mặt người. Các hệ số bù sáng tương ứng với ba kênh màu RGB sẽ được tính toán một cách tự động thông qua tỉ lệ các giá trị riêng lớn nhất của các ma trận hệ số biến đổi Cosin rời rạc, tần số thấp của ba kênh màu và tham chiếu Gaussian. Kết quả thử nghiệm trên hai bộ cơ sở dữ liệu màu nổi tiếng CMU-PIE và FERET cho thấy, ảnh khuôn mặt người thu được không chỉ rõ hơn, lấy lại được màu da người tự nhiên, ứng dụng rất nhiều trong lĩnh vực thị giác máy tính, mà còn nâng cao được hiệu suất của hệ thống nhận dạng mặt người, tốt hơn các phương pháp hiện nay như ASVD và TSVD.

Từ khóa: Biến đổi cosin rời rạc (DCT), phân rã giá trị riêng (SVD), nhận dạng ảnh mặt người.

ABSTRACT

In this paper, we provided an effective illumination compensation method to improve the quality of human face images in RGB color space for face recognition. First, three color channels of color face image and Gaussian reference are converted to the logarithm domain by logarithm transform (LT), then using the discrete cosine transform (DCT) to determine their low-frequency components. The compensated factors corresponding to the three RGB color channels will be calculated automatically through the ratio of the largest singular values of the low-frequency DCT coefficient matrices of the three color channels and the Gaussian reference. Experimental results on two databases, namely CMU-PIE and Color FERET show that our method can improve the performance of the face recognition system, better than relevant studies such as ASVD and TSVD.

Từ khóa: Discrete cosine transform (DCT), Singular Value Decomposition (SVD), face recognition.

¹Cục Công nghệ thông tin, Bộ Công an

²Trường Đại học Bách khoa Hà Nội

³Viện Khoa học và Công nghệ Quân sự

*Email: phucnguyenh46@gmail.com

Ngày nhận bài: 12/7/2019

Ngày nhận bài sửa sau phản biện: 15/8/2019

Ngày chấp nhận đăng: 15/10/2019

1. GIỚI THIỆU

Trong những năm gần đây, nhận dạng khuôn mặt người đã trở thành một lĩnh vực nghiên cứu quan trọng trong các lĩnh vực như nhận dạng mẫu, thị giác máy tính, học máy, hình sự, camera giám sát. Tuy nhiên, vẫn còn rất nhiều yếu tố tác động không nhỏ tới độ chính xác của các hệ thống nhận dạng mặt người như độ sáng, biểu hiện của khuôn mặt, tư thế của khuôn mặt, trong đó tác động của độ sáng được coi là một trong những nhân tố quan trọng nhất. Dưới điều kiện độ sáng khác nhau, ảnh khuôn mặt người có thể bị quá sáng, quá tối, thậm chí có thể bị che khuất toàn bộ hoặc một phần khuôn mặt. Do vậy, việc tăng cường chất lượng ảnh mặt người có ý nghĩa vô cùng quan trọng tại bước tiền xử lý, qua đó nâng cao độ chính xác của các hệ thống nhận dạng mặt người. Nhiều nhà nghiên cứu đã đưa ra nhiều phương pháp nâng cao chất lượng ảnh khác nhau, chủ yếu tập trung vào hai loại chính: bù độ sáng của ảnh mặt người, qua đó làm cho ảnh rõ nét hơn hoặc loại bỏ sự tác động của độ sáng, qua đó thu được ảnh chỉ chứa đặc trưng cấu trúc của khuôn mặt người.

Nhóm thứ nhất bao gồm các phương pháp cân bằng ánh sáng ảnh mặt người. Cân bằng lược đồ xám (HE) [1] là phương pháp cơ bản nhất nhằm cân bằng độ tương phản của ảnh. Tuy nhiên, HE là công nghệ xử lý toàn cục nên với những ảnh có độ sáng biến thiên cao, kết quả ảnh sau khi xử lý không được như mong đợi. Pizer và cộng sự [2] giới thiệu phương pháp mang tên cân bằng lược đồ xám tương thích (AHE) bằng việc tính toán một số lược đồ xám, mỗi lược đồ xám tương ứng với một phần riêng biệt của ảnh và sử dụng chúng để phân phối lại các giá trị độ sáng của ảnh. Tuy nhiên, AHE có xu hướng khuếch đại quá mức. Shan và cộng sự [3] đưa ra phương pháp HE từng vùng và hiệu chỉnh cường độ gamma để điều chỉnh độ sáng tổng thể của hình ảnh khuôn mặt thành hình ảnh khuôn mặt "chính tắc" được xác định trước, đồng thời cho rằng, đối với những ảnh mặt người bị che khuất cần thì cần phải chia thành nhiều ảnh con để xử lý. Choi và Jeong [4] đưa ra phương pháp để bù vào phần khuôn mặt bị che khuất trong ảnh mặt người bằng cách sử dụng biến đổi Fourier để biến đổi

ảnh mặt người từ miền không gian sang miền Fourier, sau đó cố định phổ pha và thay thế phổ biên độ của ảnh mặt người bằng phổ biên độ trung bình được tính toán bằng cách lấy trung bình của phổ biên độ của ảnh mặt người và phổ biên độ trung bình của tất cả các ảnh mặt người khác trong cơ sở dữ liệu. Tuy nhiên, khuôn mặt nhận được không tự nhiên do việc thay đổi các giá trị của phổ biên độ là không tuyến tính, dẫn đến cấu trúc ảnh nhận được không đúng với ảnh thực tế.

Nhóm thứ hai tìm cách loại bỏ cường độ sáng mà chỉ giữ lại các đặc trưng mang tính cấu trúc của khuôn mặt. Zhang và cộng sự [5] định nghĩa độ đo mức nhạy của độ sáng bằng tỷ lệ gradient theo hướng ngang và hướng dọc và thu được ảnh cấu trúc khuôn mặt người gọi là gradient face không còn tác động của độ sáng bằng cách tính arctangent của tỷ lệ này. Wang và cộng sự [6], Wu cùng cộng sự [7] thu được ảnh mặt người không còn tác động của độ sáng dựa trên luật Weber, ảnh thu được gọi là weberface. Tuy nhiên, các phương pháp này sử dụng rất nhiều tham số để xác định các ngưỡng trong mô hình và việc thiết lập các giá trị ngưỡng này là rất mơ hồ. Savvides và Kumar [8] sử dụng biến đổi logarit (LT) biến đổi ảnh sang miền logarit, sau đó sử dụng bộ lọc chuẩn dựa trên biến đổi contourlet để thu được ảnh bất biến với độ sáng. Chen cùng cộng sự [9] cũng sử dụng LT để biến đổi ảnh sang miền logarit, sau đó sử dụng biến đổi cosin rời rạc (DCT) để xác định tần số thấp và tần số cao của ảnh mặt người. Dựa trên tính chất các thông tin về độ sáng thường nằm tại băng tần tần số thấp và thông tin độ phản xạ nằm tại băng tần tần số cao, gán những giá trị băng tần tần số thấp bằng 0, chỉ sử dụng các giá trị băng tần tần số cao để tái tạo lại ảnh mặt người bất biến với độ sáng.

Các phương pháp trên tập trung giải quyết vấn đề biến đổi độ sáng trên ảnh khuôn mặt đa mức xám. Gần đây, nhiều phương pháp được xử lý trên ảnh khuôn mặt màu. Torres và cộng sự [10] chỉ ra rằng, thông tin màu thể hiện trong một số không gian màu nhất định có tác dụng lớn trong nhận dạng khuôn mặt người. Để giảm hiệu ứng chiếu sáng lên ảnh màu, Demirel và Anbarjafari [11] đã sử dụng phân rã giá trị riêng (SVD) để bù sáng ảnh khuôn mặt màu dựa trên một hệ số bù sáng được tính toán dựa trên tỷ lệ của giá trị riêng lớn nhất của tham chiếu Gaussian và giá trị riêng lớn nhất của kênh màu. Tuy nhiên, phương pháp này chỉ sử dụng một hệ số bù sáng cho toàn bộ ba kênh màu RGB, dẫn đến mất thông tin màu trong hình ảnh khuôn mặt. Để khắc phục những thiếu sót này, Wang và cộng sự [12] đưa ra phương pháp phân rã giá trị riêng tương thích (ASVD) bằng cách sử dụng thêm các giá trị trung bình của 03 kênh màu RGB, kết hợp với tỷ lệ giữa giá trị riêng lớn nhất của tham chiếu Gaussian với giá trị riêng lớn nhất của các kênh màu để tính toán các hệ số bù sáng tương ứng. Tương tự, Wang và cộng sự [13] đề xuất phương pháp gọi là TSVD, trước tiên phân tách ảnh màu thành hai phần, một phần chứa thông tin các điểm ảnh tối và một phần mang thông tin các điểm ảnh sáng, sau đó áp dụng ASVD thích hợp trên hai phần này. Tuy nhiên, hình

ảnh ASVD và TSVD không cải tiến được nhiều ảnh mặt người có độ biến thiên cao.

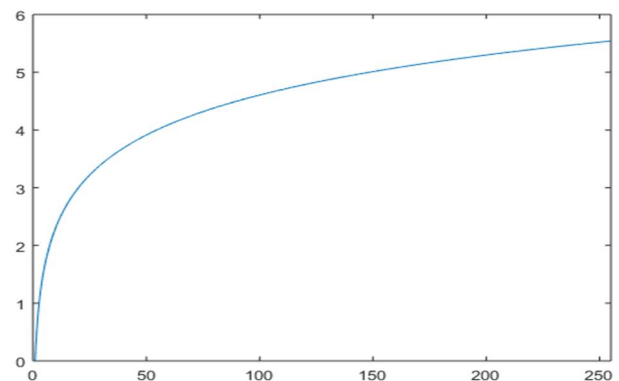
Trong bài báo này, chúng tôi đưa ra một phương pháp hữu hiệu để bù sáng cho ảnh mặt người trong không gian màu RGB, đặc biệt là các ảnh bị che khuất bằng cách tính toán các hệ số bù qua việc tham chiếu đến một ảnh màu có phân bố chuẩn Gaussian (tham chiếu Gaussian) có cùng kích thước. Trước tiên, ảnh mặt người trong không gian màu RGB và tham chiếu Gaussian được biến đổi sang miền logarit bởi biến đổi Logarit (LT), sau đó sử dụng biến đổi DCT để xác định các thành phần tần số thấp mang thông tin độ sáng của ảnh mặt người. Các hệ số bù sáng tương ứng với ba kênh màu RGB sẽ được tính toán một cách tự động thông qua tỉ lệ các giá trị riêng lớn nhất của các ma trận hệ số DCT của 3 kênh màu và tham chiếu Gaussian. Kết quả thử nghiệm trên hai bộ cơ sở dữ liệu màu nổi tiếng CMU-PIE và FERET cho thấy, ảnh khuôn mặt người thu được không chỉ rõ hơn, lấy lại được màu da người tự nhiên, ứng dụng rất nhiều trong lĩnh vực thị giác máy tính, mà còn nâng cao được hiệu suất của hệ thống nhận dạng mặt người.

Bố cục của bài báo được bố trí như sau: Phần 1 giới thiệu tổng quan về các phương pháp nâng cao chất lượng ảnh mặt người. Phần 2 giới thiệu về ứng dụng của LT, DCT và SVD trên ảnh mặt người. Phần 3 trình bày về phương pháp nâng cao chất lượng ảnh mặt người trong không gian màu RGB và kết luận được trình bày trong phần 4.

2. BIẾN ĐỔI LOGARIT, BIẾN ĐỔI COSIN RỜI RẠC, PHÂN RÃ GIÁ TRỊ RIÊNG CỦA ẢNH

2.1. Biến đổi Logarit

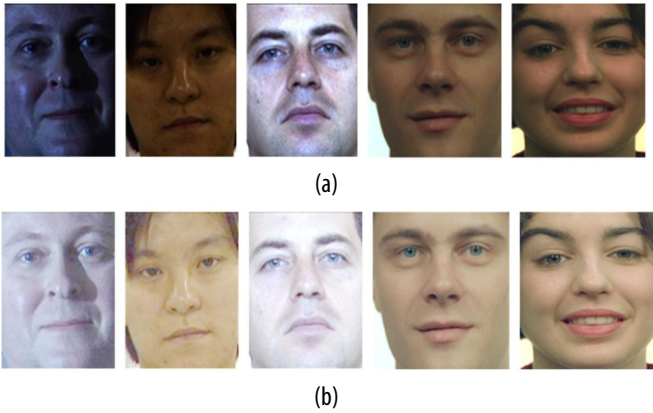
Biến đổi Logarit được sử dụng để mở rộng dải trong lĩnh vực xử lý ảnh, đặc biệt trong các bài toán nâng cao chất lượng ảnh [14]. Do tính chất của hàm logarit với đồ thị của được biểu diễn trong hình 1, LT ánh xạ một khoảng hẹp của các giá trị cấp xám thấp trong ảnh đầu vào thành một khoảng rộng hơn của ảnh đầu ra và ngược lại biến một khoảng rộng các giá trị cấp xám cao trong ảnh đầu vào thành một khoảng hẹp các giá trị cấp xám của ảnh đầu ra.



Hình 1. Đồ thị của hàm $y = \log(x)$ với x chạy từ 0 đến 255

Trong bài báo này, chúng tôi sử dụng LT để mở rộng các giá trị của các điểm ảnh tối trong một ảnh màu trong khi nén các giá trị màu cao lại, qua đó nâng cao chất lượng ảnh mặt người, đặc biệt với các ảnh màu mặt người trong

trường hợp bị che khuất, do đó nâng cao chất lượng ảnh màu khuôn mặt người. Hình 2 biểu diễn một số ảnh màu mặt người trong cơ sở dữ liệu CMU-PIE và FERET (a) và các ảnh thu sau khi được xử dụng LT (b). Kết quả cho thấy, phần bị che khuất trong ảnh đã được khuếch đại làm cho rõ hơn. Tuy nhiên, các phần không bị che khuất cũng bị khuếch đại làm cho ảnh bị quá sáng, ảnh hưởng lớn độ chính xác của hệ thống nhận dạng khuôn mặt.



Hình 2. Ảnh màu mặt người trong cơ sở dữ liệu CMU-PIE và FERET (a) và ảnh thu được sau khi sử dụng LT (b)

2.2. Biến đổi cosin rời rạc

Biến đổi cosin rời rạc biến đổi ảnh từ miền không gian sang miền tần số, tập trung năng lượng vào một hệ số, được áp dụng rộng rãi trong nén ảnh số chuẩn JPEG và MPEG [15]. Các dạng DCT được chia thành 4 loại, ký hiệu là DCT-I, DCT-II, DCT-III và DCT-IV. DCT-II được sử dụng rộng rãi trong mã hóa tín hiệu số bởi gần tương đương biến đổi Karhunen-Loeve trong mô hình tín hiệu Markov-I với hệ số tương quan gần bằng 1 [16]. Trong xử lý ảnh, DCT-II được sử dụng rộng rãi nhất và thường được gọi vắn tắt là DCT.

Cho ảnh xám có kích thước $m \times n$, biến đổi hai chiều DCT (2D-DCT) được định nghĩa như sau:

$$C(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} f(x,y) \cos \left[\frac{\pi(2x+1)u}{2m} \right] \cos \left[\frac{\pi(2y+1)v}{2n} \right] \quad (1)$$

trong đó, $C(u, v)$ được gọi là hệ số cosin rời rạc.

Và biến đổi DCT ngược (iDCT) được định nghĩa bởi

$$f(x,y) = \sum_{u=0}^{m-1} \sum_{v=0}^{n-1} \alpha(u)\alpha(v)C(u,v) \cos \left[\frac{\pi(2x+1)u}{2m} \right] \cos \left[\frac{\pi(2y+1)v}{2n} \right] \quad (2)$$

với

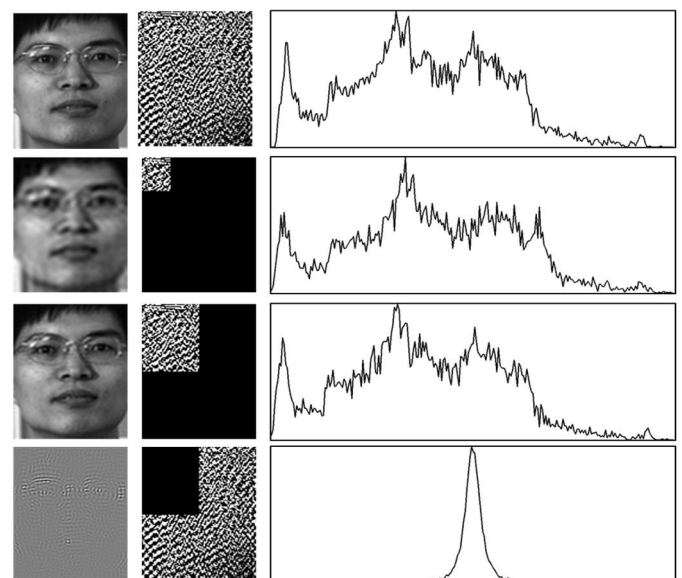
$$\alpha(u) = \begin{cases} \frac{1}{\sqrt{m}}, & u = 0 \\ \frac{2}{\sqrt{m}}, & u = 1,2, \dots, m - 1 \end{cases} \quad (3)$$

và

$$\alpha(v) = \begin{cases} \frac{1}{\sqrt{n}}, & v = 0 \\ \frac{2}{\sqrt{n}}, & v = 1,2, \dots, n - 1 \end{cases} \quad (4)$$

Trong chuẩn nén ảnh JPEG, ảnh gốc được chia thành những khối nhỏ không chồng khít lên nhau có kích thước 8

$\times 8$ và áp dụng DCT cho từng khối này một cách riêng rẽ. Trong khuôn khổ nghiên cứu này, chúng tôi áp dụng DCT trên toàn bộ ảnh thể thu được tất cả các thành phần tần số của ảnh mặt người. Hơn thế nữa, trong ảnh khuôn mặt người, độ sáng thường thay đổi chậm hơn khi so sánh với độ phản xạ, ngoại trừ phần khuôn mặt bị che khuất. Khi biến đổi sang miền tần số, thông tin chứa độ sáng thường nằm ở băng tần tần số thấp và thông tin chứa độ phản xạ thường nằm ở tần số cao. Nếu muốn tác động vào độ sáng, chúng ta tác động vào các giá trị băng tần thấp và loại bỏ các giá trị băng tần cao, ngược lại, nếu ta muốn sử dụng độ phản xạ để biểu diễn ảnh khuôn mặt người, chúng ta tác động vào băng tần cao và loại bỏ các giá trị băng tần thấp. Kết quả của DCT là sự chuyển đổi ảnh gốc sang miền tần số. Giá trị trên cùng bên trái lưu trữ "biên độ" tần số và tần số "cơ sở" tăng cả dọc theo trục ngang và dọc. Kết quả của DCT thường là một tập hợp các biên độ ở tần số thấp hơn bình thường (góc phần tư trên cùng bên trái) và các mục nhỏ hơn ở tần số cao hơn. Do thông tin độ sáng nằm chủ yếu tại băng tần thấp, chúng ta có thể tái tạo lại ảnh khuôn mặt bằng việc chỉ sử dụng các hệ số băng tần thấp, còn gán những hệ số băng tần cao bằng 0. Hình 3 cho thấy sự đúng đắn của khẳng định này. Hàng thứ nhất biểu diễn ảnh mặt người và ma trận hệ số DCT tương ứng; hàng 2 là ảnh mặt người được xây dựng lại khi chỉ giữ lại một số thành phần tần số thấp (góc phần tư trên cùng bên trái) còn các hệ số khác được gán bằng 0; hàng 3 là ảnh mặt người được xây dựng lại khi chỉ giữ lại một số thành phần tần số thấp (góc phần tư trên cùng bên trái) còn các hệ số khác được gán bằng 0; hàng 4 là ảnh mặt người được xây dựng lại khi gán các thành phần tần số thấp bằng 0 (góc phần tám trên cùng bên trái) còn các hệ số khác được giữ nguyên. Kết quả cho thấy, ảnh mặt người vẫn được duy trì khi ta giữ lại các thành phần tần số thấp và ngược lại ta sẽ thu được cấu trúc của ảnh nếu loại bỏ tần số thấp và giữ lại tần số cao.



Hình 3. Từ trái qua phải là ảnh mặt người, ma trận hệ số DCT và lược đồ histogram của ảnh mặt người

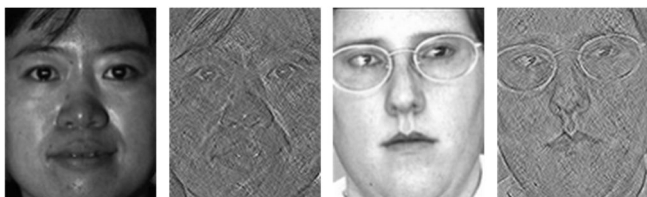
Trong nghiên cứu này, chúng tôi thực hiện việc bù sáng trên ảnh màu mặt người, do đó loại bỏ các tần số cao, chỉ giữ lại các giá trị tần số thấp (góc phần tư trên cùng bên trái) và đưa ra một thuật toán rất hữu hiệu để bù sáng cho ảnh màu mặt người, làm cho ảnh khuôn mặt người không chỉ rõ hơn, mà còn tái tạo lại làn da tự nhiên của khuôn mặt người, qua đó nâng cao hiệu suất của hệ thống nhận dạng khuôn mặt người. Việc làm này cũng giống như việc sử dụng bộ lọc băng tần thấp trong miền DCT.

2.3. Phân rã giá trị riêng

Phân rã giá trị riêng được sử dụng rộng rãi trong xử lý ảnh do cấu trúc ảnh số giống như một ma trận. Không làm mất tính tổng quát, giả sử f là ảnh mặt người có kích thước $M \times N$, ($M \geq N$). SVD của ảnh f được tính như sau:

$$f = U * \Sigma * V^T; \tag{5}$$

trong đó, $U = [u_1, u_2, \dots, u_N]$ và $V = [v_1, v_2, \dots, v_N]$ là các ma trận trực giao chứa các vectơ riêng; $\Sigma = [D, 0]$ gồm các giá trị riêng được sắp xếp theo chiều giảm dần, với $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$ là các giá trị riêng và k là hạng của f . Khi ảnh được phân rã bởi SVD, các giá trị riêng sẽ chứa thông tin độ sáng của ảnh số, còn vectơ riêng chứa thông tin về độ phản xạ. Hình 4 biểu diễn 02 ảnh trong cơ sở dữ liệu CMU-PIE và ảnh của chúng thu được sau khi gán các giá trị riêng của ảnh bằng 1. Từ công thức (5) cho thấy, khi gán ma trận Σ bởi ma trận đơn vị, ảnh thu được sẽ không còn tác động của ma trận Σ . Sau khi loại bỏ sự tác động của các giá trị riêng, ta thu được ảnh cấu trúc của khuôn mặt không còn tác động của độ sáng. Điều này khẳng định, các giá trị riêng mang thông tin độ sáng trong ảnh. Không những vậy, giá trị riêng lớn nhất của ảnh còn chứa 99,72% năng lượng của ảnh [18].



Hình 4. Ảnh mặt người và ảnh tái tạo của nó khi gán tất cả các giá trị riêng bằng 1

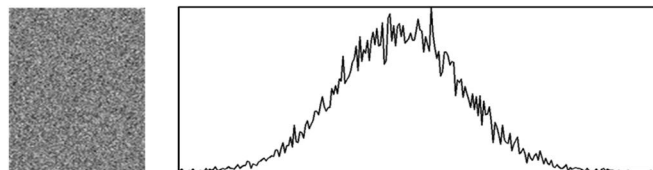
3. NÂNG CAO CHẤT LƯỢNG ẢNH MẶT NGƯỜI TRONG KHÔNG GIAN MÀU RGB

Trong phần này, chúng tôi giới thiệu phương pháp hữu hiệu để nâng cao chất lượng ảnh mặt người trong không gian màu RGB, qua đó nâng cao độ chính xác của hệ thống nhận dạng khuôn mặt người. Một ảnh có chất lượng tốt thường có độ phân phối gần phân phối chuẩn tắc, còn được gọi là phân bố Gaussian, là phân phối chuẩn với giá trị trung bình μ bằng 0 và độ lệch chuẩn σ bằng 1. Do vậy, trong phương pháp của chúng tôi sẽ tham chiếu đến một ảnh có phân phối Gaussian để tính toán hệ số bù sáng cho các kênh màu của ảnh màu mặt người, qua đó tự động điều chỉnh các kênh màu có độ phân bố gần với phân bố chuẩn tắc, và do vậy nâng cao chất lượng ảnh. Do giá trị các điểm ảnh nằm trong khoảng giá trị từ 0 đến 255 nên ảnh có phân phối chuẩn tắc có giá trị trung bình μ bằng 32 và độ lệch

chuẩn σ bằng 128. Ta ký hiệu ảnh có phân phối chuẩn tắc là G_a , như biểu diễn trong hình 5.

Trước tiên, phân tách ảnh mặt người trong không gian màu RGB thành các kênh màu R, G, B riêng biệt. Ký hiệu ảnh mặt là f , ta có:

$$f = \{R, G, B\}; \tag{6}$$



Hình 5. Ảnh phân phối chuẩn tắc Gaussian và lược đồ xám của nó

Tiếp theo, sử dụng LT để biến đổi các kênh màu sang miền logarit. Như đã thảo luận trong phần 2.1, LT có vai trò rất lớn trong việc nâng cao chất lượng các ảnh bị che khuất.

Ký hiệu $f_{LT} = LT(f)$ và $G_{aLT} = LT(G_a)$, ta có:

$$f_{LT} = \log(f) = \log(\{R, G, B\}) = \{R_{LT}, G_{LT}, B_{LT}\} \tag{7}$$

$$G_{aLT} = \log(G_a); \tag{8}$$

Để xác định các thành phần tần số thấp mang thông tin của độ sáng, chúng tôi sử dụng DCT để biến đổi các kênh màu trong miền logarit sang miền DCT. Như thảo luận trong phần 2.2 các tần số thấp trong miền DCT chứa thông tin độ sáng, cho nên chúng tôi chỉ giữ lại các thành phần tần số thấp (góc phần tư bên trái) của các kênh màu, ký hiệu là $DCT_{R_{LT}}, DCT_{G_{LT}}, DCT_{B_{LT}}$, có kích thước bằng 1/4 ảnh gốc (dòng 3 của hình 3) tương ứng với 3 kênh màu R, G, B và tác động vào các hệ số này để thu được ảnh có chất lượng tốt hơn. Ta có:

$$\{DCT_{R_{LT}}, DCT_{G_{LT}}, DCT_{B_{LT}}\} = DCT(R_{LT}, G_{LT}, B_{LT}) \tag{9}$$

$$DCT_{G_{aLT}} = DCT(G_{aLT}) \tag{10}$$

Như thảo luận trong phần 2.3, giá trị riêng của ảnh chứa thông tin độ sáng và giá trị riêng lớn nhất chứa 99,72% năng lượng của ảnh, do đó tính toán SVD của $DCT_{R_{LT}}, DCT_{G_{LT}}, DCT_{B_{LT}}$ và $DCT_{G_{aLT}}$:

$$DCT_{R_{LT}} = U_R * \Sigma_R * V_R^T; \tag{11}$$

$$DCT_{G_{LT}} = U_G * \Sigma_G * V_G^T; \tag{12}$$

$$DCT_{B_{LT}} = U_B * \Sigma_B * V_B^T; \tag{13}$$

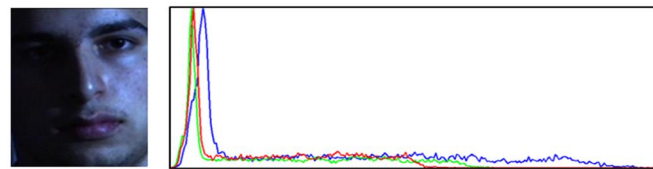
$$DCT_{G_{aLT}} = U_{G_a} * \Sigma_{G_a} * \tag{14}$$

Ký hiệu các giá trị riêng lớn nhất của $\Sigma_R, \Sigma_G, \Sigma_B$ và Σ_{G_a} tương ứng là $\lambda_R, \lambda_G, \lambda_B$ và λ_{G_a} . Khi đó, hệ số bù sáng cho các kênh màu được tính như sau:

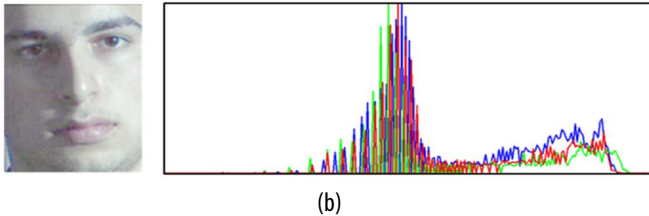
$$\mu_R = \frac{\lambda_{G_a}}{\lambda_R} \tag{15}$$

$$\mu_G = \frac{\lambda_{G_a}}{\lambda_G} \tag{16}$$

$$\mu_B = \frac{\lambda_{G_a}}{\lambda_B} \tag{17}$$



(a)



Hình 6. Ảnh màu mặt người và lược đồ histogram của nó (a); ảnh ALDS của ảnh gốc ở hàng trên và lược đồ histogram của ALDS (b)

Thông qua các hệ số bù sáng này, cách hệ số DCT được tính toán lại bằng cách nhân với hệ số bù sáng như sau:

$$DCT_{R_{bù}} = U_R * (\mu_R \Sigma_R) * V_R^T; \tag{18}$$

$$DCT_{G_{bù}} = U_G * (\mu_G \Sigma_G) * V_G^T; \tag{19}$$

$$DCT_{B_{bù}} = U_B * (\mu_B \Sigma_B) * V_B^T; \tag{20}$$

Sử dụng biến đổi DCT ngược của các hệ số DCT bù sáng, ta được các kênh màu sau khi bù sáng:

$$R_{bù} = iDCT(DCT_{R_{bù}}); \tag{21}$$

$$G_{bù} = iDCT(DCT_{G_{bù}}); \tag{22}$$

$$B_{bù} = iDCT(DCT_{B_{bù}}); \tag{23}$$

Cuối cùng, trộn lẫn ba kênh màu và chuẩn hóa dữ liệu, ta thu được ảnh màu mặt người, ký hiệu là ALDS, không chỉ rõ nét hơn rất nhiều so với ảnh gốc, mà còn tái tạo lại được mẫu da người, có thể rất hữu hiệu trong lĩnh vực thị giác máy tính và phát hiện màu da, như chỉ trong hình 6. Do bề mặt người không phải là bề mặt khuếch tán hoàn hảo (bề mặt lambertian) nên trong một số trường hợp, có những đặc trưng của khuôn mặt không nằm ở băng tần thấp. Hơn nữa, những phần bị che khuất cũng nằm trên cùng băng tần với các đặc trưng quan trọng của khuôn mặt, do đó có lúc độ sáng sẽ không được bù đúng bởi bỏ đi các tần số cao. Do vậy, trong nghiên cứu này, chúng tôi không dùng biến đổi logarit ngược. Hình 6 cho ta thấy, ảnh màu mặt người ALDS rõ nét hơn, màu da đúng với tự nhiên hơn và có phân bố giá trị độ sáng gần với phân bố chuẩn chính tắc. Do vậy, nâng cao được hiệu suất của hệ thống nhận dạng khuôn mặt người.

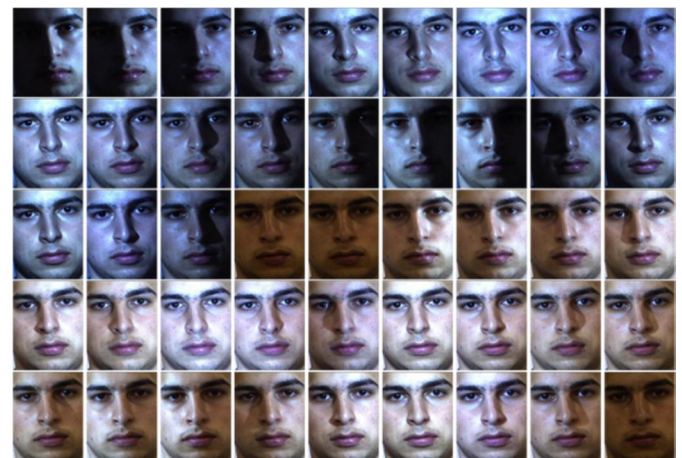
4. KẾT QUẢ THỰC NGHIỆM

Để làm rõ sự hiệu quả của phương pháp ALDS, chúng tôi tiến hành thực nghiệm trên hai cơ sở dữ liệu ảnh màu nổi tiếng là CMU-PIE [19] và FERET [20]. Chúng tôi cũng so sánh kết quả của ALDS với các phương pháp trước như ASVD, TSVD và ảnh màu mặt người thu được bằng việc chỉ sử dụng LT, ký hiệu là CFLT, như thảo luận trong phần 2.1. Để tiến hành so sánh, chúng tôi sử dụng các phương pháp trích xuất đặc trưng khuôn mặt để tiến hành nhận dạng khác nhau như Eigenface [21] và LBP [22].

4.1. Thực nghiệm trên cơ sở dữ liệu CMU-PIE

Cơ sở dữ liệu mặt người CMU-PIE bao gồm 41.368 bức ảnh màu của 68 người. Mỗi người được chụp dưới 13 tư thế, 43 cường độ sáng, 4 sắc thái khuôn mặt khác nhau. Để thực nghiệm kết quả, mỗi người lấy 45 ảnh chụp trực diện, mỗi ảnh có độ sáng khác nhau được chụp bởi máy ảnh trung tâm (c27), bao gồm hai điều kiện "tắt đèn" và "bật

đèn". Loại "tắt đèn" bao gồm 21 độ sáng từ f02 đến f22, trong khi loại "bật đèn" bao gồm 24 độ sáng từ f00 đến f23 như trong hình 7(a). Khuôn mặt người được trích xuất và chuẩn hóa dưới cùng một kích thước 96x132. Tổng cộng chúng tôi có 3600 ảnh. Chúng tôi sử dụng lần lượt từng phương pháp ASVD, TSVD, CFLT và ALDS để thu được các ảnh từ ảnh màu mặt người gốc để tiến hành nhận dạng. Hình 7 biểu diễn các ảnh gốc và kết quả của của các phương pháp. Hình 7(a) biểu diễn 45 ảnh dưới 45 độ sáng khác nhau, trong cả điều kiện "tắt đèn" và "bật đèn" của một người trong cơ sở dữ liệu CMU-PIE. Hình 7(b) biểu diễn kết quả thu được sau khi áp dụng phương pháp ASVD và kết quả của phương pháp TSVD được biểu diễn trong hình 7(b). Kết quả cho thấy, ảnh sau khi được nâng cao không quá khác biệt so với ảnh gốc vì các phương pháp này xử lý ảnh trong miền không gian. Hình 7(c) biểu diễn ảnh thu được khi áp dụng phương pháp CFLT như thảo luận trong mục 2.1. Kết quả cho thấy, ảnh đã rõ nét hơn nhưng chúng ta vẫn thấy được tác động của điều kiện "tắt đèn" và "bật đèn". Kết của phương pháp ALDS của chúng tôi được biểu diễn trong hình 7(d). Kết quả cho thấy, tất cả các ảnh đã rõ nét hơn, đặc biệt là các ảnh bị che khuất và tái tạo được làn da gốc của người, trông rất tự nhiên. Tất cả các ảnh có màu sắc giống nhau như được chụp cùng một thời điểm trong điều kiện "bật đèn".



(a)



(b)



(c)



(d)



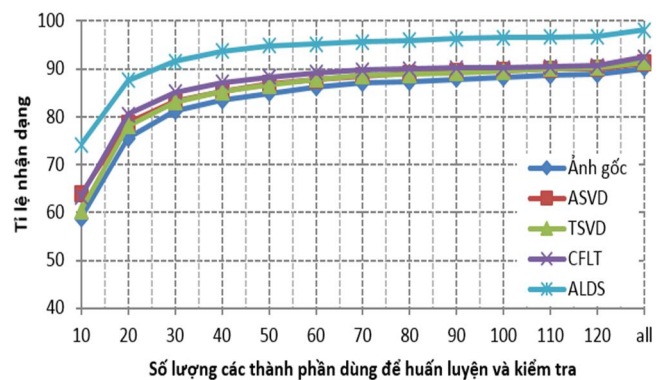
(e)

Hình 7. (a) 45 ảnh gốc của một người trong cơ sở dữ liệu CMU-PIE; (b) ASVD của (a); (c) TSVD của (a); (d) CFLT của (a); (e) ALDS của (a)

Bảng 1. Kết quả nhận dạng mặt người trên cơ sở dữ liệu CMU-PIE bằng phương pháp eigenface

Số thành phần	Tỉ lệ nhận dạng				
	Ảnh gốc	ASVD	TSVD	CFLT	ALDS
10	58,77	63,82	60,39	63,33	74,12
20	75,54	78,73	78,14	80,54	87,60

30	81,13	83,24	83,14	85,15	91,57
40	83,38	85,34	85,25	87,11	93,73
50	84,85	86,91	86,57	88,33	94,80
60	86,13	87,75	87,84	89,22	95,20
70	87,06	88,48	88,58	89,80	95,64
80	87,30	89,07	88,82	90,00	95,98
90	87,79	89,41	89,22	90,20	96,37
100	88,19	89,66	89,51	90,25	96,52
110	88,63	90,00	90,05	90,44	96,72
120	88,82	90,15	90,34	90,69	96,81
Tất cả	90,05	91,22	91,43	92,42	98,14



Hình 8. Tỉ lệ nhận dạng bằng phương pháp eigenface trên cơ sở dữ liệu CMU-PIE

Đầu tiên, chúng tôi sử dụng phương pháp eigenface để đánh giá hiệu suất của nhận dạng khuôn mặt. Phương pháp eigenface dựa trên việc ánh xạ tuyến tính ảnh mặt người vào không gian đặc trưng có số chiều thấp hơn bằng cách sử dụng phương pháp phân tích thành phần chính (PCA). Nó sử dụng các thành phần chính là các véc-tơ riêng tương ứng với các giá trị riêng lớn nhất làm đặc trưng và sau đó dùng giải thuật hàng xóm gần nhất giữa ảnh huấn luyện và ảnh kiểm tra. Để tiến hành nhận dạng, với mỗi người trong cơ sở dữ liệu, chúng tôi chọn 15 ảnh để huấn luyện và 30 ảnh còn lại để kiểm tra. Kết quả nhận dạng được biểu diễn trong bảng 1 và hình 8. Kết quả cho thấy, phương pháp ALDS đã nâng cao đáng kể tỉ lệ nhận dạng mặt người, cao hơn ảnh gốc, ASVD, TSVD, CFLT lần lượt là 8,09%, 6,42%, 6,71% và 5,72%.

Tiếp theo, chúng tôi tiếp tục so sánh tỉ lệ nhận dạng của phương pháp ALDS với các phương pháp ASVD, TSVD, CFLT bằng phương pháp trích xuất đặc trưng khác như phương pháp mẫu nhị phân cục bộ (LBP). Không giống như phương pháp eigenface bị ảnh hưởng rất lớn bởi sự thay đổi của độ sáng. Phương pháp LBP loại bỏ được thông tin chứa độ sáng và thu được cấu trúc của ảnh khuôn mặt người. Hình 9(a) biểu diễn các ảnh mặt người gốc và ảnh LBP của chúng và hình 9(b) biểu diễn các ảnh ALDS và ảnh LBP của chúng cho thấy, hình ảnh LBP của ALDS rõ ràng hơn. Kết quả nhận dạng bằng LBP trong bảng 2 cho thấy, sau khi sử dụng ALDS, tỉ lệ nhận dạng tăng 2,92%, 2,18%,

1,98% và 1,9% khi so sánh với ảnh gốc, ASVD, TSVD và CFLT. Điều này chứng tỏ, ALDS không những tăng cường màu sắc của ảnh mặt người, mà còn tăng cường được cả cấu trúc của ảnh mặt người.

Bảng 2. Kết quả nhận dạng mặt người trên cơ sở dữ liệu FERET bằng phương pháp LBP

Cơ sở dữ liệu	Tỉ lệ nhận dạng				
	Ảnh gốc	ASVD	TSVD	CFLT	ALDS
CMU-PIE	95,08	95,82	96,02	96,10	98
FERET	90,55	90,85	90,85	91,20	94,5



(a)

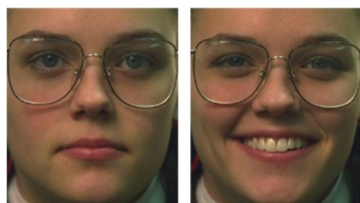


(b)

Hình 9. Ảnh thu được khi áp dụng phương pháp LBP. (a) Ảnh gốc và LBP của ảnh gốc; (b) Ảnh ALDS và LBP của chúng

4.2. Thực nghiệm trên cơ sở dữ liệu FERET

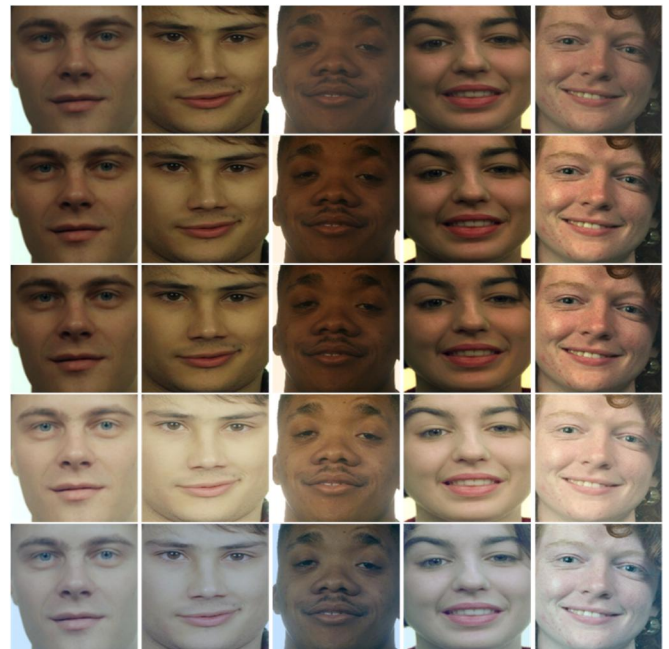
Để kiểm tra thêm khả năng nhận dạng khuôn mặt của ALDS, chúng tôi cũng đã sử dụng các phương pháp eigenface và LBP để đánh giá hiệu suất nhận dạng khuôn mặt trên bộ cơ sở dữ liệu FERET. Cơ sở dữ liệu FERET được đưa ra bởi Viện Tiêu chuẩn và Công nghệ Quốc gia Mỹ (NIST), gồm 11.338 ảnh được thu thập từ 994 người với nhiều sắc thái khuôn mặt và nhiều độ sáng khác nhau. Để tiến hành thực nghiệm, chúng tôi chọn 810 người trong cơ sở dữ liệu, mỗi người gồm 02 ảnh trong tập fa và fb như biểu diễn trong hình 10, trong đó fa được dùng để huấn luyện còn fb dùng để kiểm tra.



(a) (b)

Hình 10. Ví dụ về ảnh fa và fb của một người trong cơ sở dữ liệu FERET

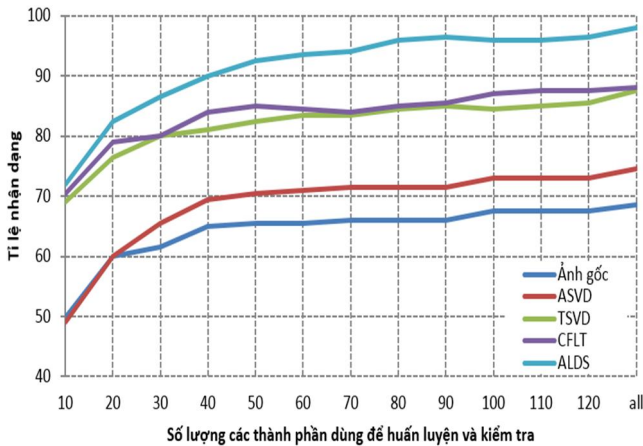
Hình 10 biểu diễn các hình ảnh gốc trong cơ sở dữ liệu FERET và hình ảnh ASVD, TSVD, CFLT và ALDS. Cũng giống như trong cơ sở dữ liệu CMU-PIE, hình ảnh ASVD ở hàng 1 và TSVD ở hàng 2 không có quá nhiều sự khác biệt so với ảnh gốc. Tuy nhiên, ảnh CFLT ở hàng 3 và đặc biệt ảnh ALDS ở hàng cuối cùng cho thấy, tất cả các ảnh đều cùng một độ sáng, qua đó làm khuôn mặt người rõ nét hơn. Kết quả nhận dạng bằng eigenface trên cơ sở dữ liệu FERET được biểu diễn trong bảng 3 và hình 12. Kết quả cho thấy, sau khi sử dụng ALDS, tỉ lệ nhận dạng tăng 29,5%, 23,5%, 10,5% và 1% khi so sánh với ảnh gốc, ASVD, TSVD và CFLT.



Hình 11. Hàng 1 gồm 05 ảnh trong FERET; hàng 2 gồm ASVD của hàng 1; hàng 3 gồm TSVD của hàng 1; hàng 4 gồm CFLT của hàng 1; hàng 5 gồm ALDS của hàng 1

Bảng 3. Kết quả nhận dạng mặt người trên cơ sở dữ liệu FERET bằng phương pháp eigenface

Số thành phần	Tỉ lệ nhận dạng				
	Ảnh gốc	ASVD	TSVD	CFLT	ALDS
10	50,0	49,0	69,0	70,5	72,0
20	60,0	60,0	76,5	79,0	82,5
30	61,5	65,5	80,0	80,0	86,5
40	65,0	69,5	81,0	84,0	90,0
50	65,5	70,5	82,5	85,0	92,5
60	65,5	71,0	83,5	84,5	93,5
70	66,0	71,5	83,5	84,0	94,0
80	66,0	71,5	84,5	85,0	96
90	66,0	71,5	85,0	85,5	96,5
100	67,5	73,0	84,5	87,0	96,0
110	67,5	73,0	85,0	87,5	96,0
120	67,5	73,0	85,5	87,5	96,5
Tất cả	68,5	74,5	87,5	88,0	98,0



Hình 12. Tỷ lệ nhận dạng bằng phương pháp eigenface trên cơ sở dữ liệu FERET

Kết quả nhận dạng bằng LBP trên tập dữ liệu FERET trong bảng 2 cho thấy, sau khi sử dụng ALDS, tỷ lệ nhận dạng tăng 3,95%, 3,65%, 3,65% và 3,3% khi so sánh với ảnh gốc, ASVD, TSVD và CFLT.

5. KẾT LUẬN

Trong bài báo này, một phương pháp tiền xử lý hình ảnh mới có tên ALDS được đề xuất để nhận dạng khuôn mặt màu dưới nhiều độ sáng khác nhau. Phương pháp này có thể làm cho hình ảnh khuôn mặt màu rõ hơn, tự nhiên hơn và mịn hơn, ngay cả khi hình ảnh khuôn mặt bị che khuất. Các kết quả thử nghiệm dựa trên hai cơ sở dữ liệu màu mặt phổ biến hiện nay là CMU-PIE và FERET cho thấy, phương pháp được đề xuất là cực kỳ hiệu quả trong các ứng dụng thực tế. Phương pháp này cho thấy hiệu suất cao cho tác vụ khớp khuôn mặt và chắc chắn cũng hữu ích trong lĩnh vực thị giác máy tính, phát hiện khuôn mặt và phát hiện màu da.

TÀI LIỆU THAM KHẢO

[1]. R. C. Gonzalez and R. E. Wood, 2007. *Digital image processing*. third ed. Prentice Hall.

[2]. S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. T. H. Romeny, J. B. Zimmerman, and K. Zuiderveld, 1987. *Adaptive histogram equalization and its variations*. *Comput. Vision. Graph. Image Process.*, vol. 39, pp. 355-368.

[3]. S. Shan, W. Gao, B. Cao, and D. Zhao, 2003. *Illumination normalization for robust face recognition against varying lighting conditions*. in *IEEE. Workshop on AMFG*, pp.157-164.

[4]. S. -I. Choi and G. -M. Jeong, 2011. *Shadow compensation using Fourier analysis with application to face recognition*. *IEEE Signal Process. Lett.*, vol. 18, pp. 23-26.

[5]. T. Zhang, Y. Y. Tang, B. Fang, Z. Shang and X. Liu, 2009. *Face recognition under varying illumination using Gradientfaces*. *IEEE Trans. Image Process.*, vol. 18, pp. 2599-2606.

[6]. B. Wang, W. Li, W. Yang and Q. Liao, 2011. *Illumination normalization based on Weber's law with application to face recognition*. *IEEE Signal Process. Lett*, vol. 18, pp. 462-465.

[7]. Y. Wu, Y. Jiang, Y. Zhou, W. Li, Z. Lu, and Q. Liao, 2014. *Generalized Weber-face for illumination-robust face recognition*. *Neurocomputing*, vol. 136, pp. 262-267.

[8]. M. Savvides and B. V. K. V. Kumar, 2003. *Illumination normalization using logarithm transforms for face authentication*. in *Proc. IAPR AVBPA*, pp. 549-556.

[9]. W. Chen, M. J. Er, and S. Wu, 2006. *Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain*. *IEEE Trans. Syst., Man, Cybern., Syst*, vol. 36, pp. 458-466.

[10]. L. Torres, J. Y. Reutter, and L. Lorente, 1999. *The importance of the color information in face recognition*. *Int. Conf. ICIP*, vol. 3, pp. 627-631.

[11]. H. Demirel and G. Anbarjafari, 2008. *Pose invariant face recognition using probability distribution functions in different color channels*. *IEEE Signal Process. Lett*, vol. 15, pp. 537-540.

[12]. J. -W. Wang, J. -S. Lee, and W. -Y. Chen, 2011. *Face recognition based on projected color space with lighting compensation*. *IEEE Signal Process. Lett*, vol. 18, pp. 567-570.

[13]. J. -W. Wang, J. -S. Lee, and W. -Y. Chen, 2014. *Recognition based on two separated singular value decomposition-enriched faces*. *Journal of Electronic Imaging*, vol. 23, no. 6, pp. 063010-1~063010-15.

[14]. Y. Adini, Y. Moses, and S. Ullman, 1997. *Face recognition: the problem of compensating for changes in illumination direction*. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 721-732.

[15]. W. Pennebaker and J. Mitchell, 1993. *JPEG Still Image Data Compression Standard*. New York: Van Nostrand Reinhold.

[16]. K. R. Rao and P. Yip, 1990. *Discrete Cosine Transform: Algorithms, Advantages, Applications*. Boston, MA: Academic.

[17]. H. Demirel and G. Anbarjafari, 2008. *Pose invariant face recognition using probability distribution functions in different color channels*. *IEEE Signal Process. Lett*, vol. 15, pp. 537-540.

[18]. T. Sim, S. Baker, and M. Bsat, 2003. *The CMU pose, illumination, and expression database*. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, pp. 1615-1618.

[19]. P. J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, 2000. *The FERET evaluation methodology for face recognition algorithms*. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, pp. 1090-1104.

[20]. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, 1997. *Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection*. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19.

[21]. T. Ahonen, A. Hadid, and M. Pietikainen, 2006. *Face description with local binary patterns: application to face recognition*. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, pp. 2037-2041, 2006.

AUTHORS INFORMATION

Nguyen Nam Phuc¹, Nguyen Quoc Trung², Ha Huu Huy³

¹Department of Information Technology, Ministry of Public Security of Socialist Republic of Vietnam

²Hanoi University of Science and Technology

³Military Institute of Technology and Science, Vietnam