



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0052796

(51)^{2022.01} G06F 17/00

(13) B

(21) 1-2023-06888

(22) 03/10/2023

(45) 27/10/2025 451

(43) 25/01/2024 430A

(73) TẬP ĐOÀN CÔNG NGHIỆP - VIỆN THÔNG QUÂN ĐỘI (VN)

Lô D26 Khu đô thị mới Cầu Giấy, phường Yên Hoà, quận Cầu Giấy, thành phố Hà Nội.

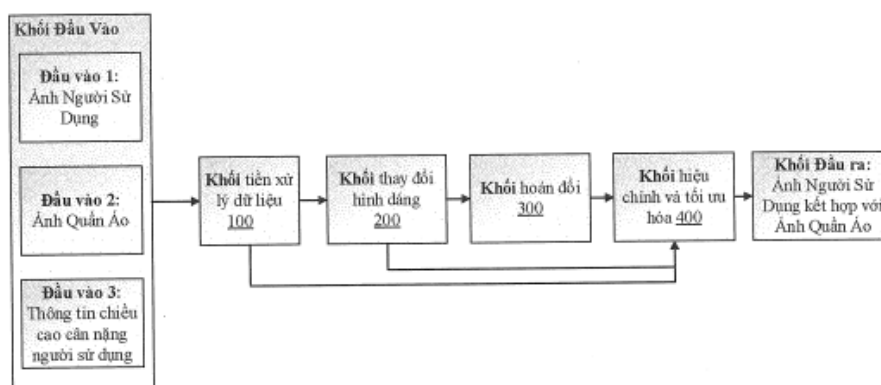
(72) PHẠM QUANG TIẾN (VN); TRẦN VĂN ĐỨC (VN); NGUYỄN HOÀNG VŨ (VN); NGUYỄN TIẾN ĐẠT (VN); DƯƠNG ANH TRÀ (VN).

(74) Công ty TNHH NACILAW (NACILAW)

(54) HỆ THỐNG VÀ PHƯƠNG PHÁP THỬ ẢO TRANG PHỤC HAI CHIỀU (2D) SỬ DỤNG CÔNG NGHỆ MÁY HỌC KẾT HỢP GIỮA THUẬT TOÁN TỐI ƯU VÀ THUẬT TOÁN PHÂN LOẠI XÁC ĐỊNH

(21) 1-2023-06888

(57) Sáng chế đề cập hệ thống và phương pháp thử ảo trang phục hai chiều (2D) sử dụng công nghệ máy học và thuật toán phân loại xác định. Hệ thống và phương pháp đề xuất cho phép xây dựng hình ảnh hai chiều (2D) tương tác số giữa quần áo mẫu và cơ thể người sử dụng thay vì thay vì sử dụng kỹ sư đồ họa mô phỏng. Phương pháp sử dụng các mô hình máy học về xử lý ảnh kết hợp với thuật toán phân loại xác định để tái tạo ảnh người sử dụng kết hợp với ảnh quần áo và có biến đổi ảnh quần áo mẫu theo kích thước của người sử dụng.



Hình 2

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến hệ thống và phương pháp thử ảo trang phục hai chiều (2D) sử dụng công nghệ máy học kết hợp giữa thuật toán tối ưu và thuật toán phân loại xác định. Cụ thể, hệ thống và phương pháp thử ảo trang phục hai chiều (2D) với việc xây dựng các hình ảnh người sử dụng trên các trang phục thời trang sử dụng mô hình máy học thông minh trong xử lý ảnh kết hợp giữa thuật toán tối ưu và thuật toán phân loại xác định để đồng bộ được ứng dụng trong lĩnh vực mô phỏng ứng dụng như trình diễn mô hình trang phục, thử đồ hai chiều (2D) trực tuyến.

Tình trạng kỹ thuật của sáng chế

Hệ thống và phương pháp thử ảo trang phục trình bày chi tiết cách thức mô phỏng tương tác quần áo trên người sử dụng với đầu ra là các ảnh hai chiều (2D) của người sử dụng kết hợp với ảnh quần áo ứng dụng trong lĩnh vực mô phỏng ứng dụng. Theo phương pháp truyền thống, để có thể xây dựng các hình ảnh hai chiều (2D) của người sử dụng trên các ảnh mẫu có sẵn, có hai cách chính được sử dụng: một là thực hiện thủ công bởi các kỹ sư đồ họa mô phỏng; hai là thực hiện bởi các mô hình máy học cơ bản. Cụ thể, kỹ sư đồ họa mô phỏng thực hiện việc chỉnh sửa, cắt ghép, hiệu chỉnh đầu của người sử dụng lên các ảnh trang phục mẫu có sẵn, hoặc ngược lại, ảnh trang phục có sẵn được ghép trực tiếp trên ảnh của người sử dụng và sử dụng các phần mềm thương mại để hiệu chỉnh ánh sáng, phong nền... Ngoài ra, một số mô hình máy học cơ bản ghép trực tiếp mặt người sử dụng lên ảnh mẫu hoặc ngược lại, với các bước làm tương tự như làm thủ công, tuy nhiên ở mức độ chính xác và chân thực thấp hơn. Tổng quan của phương pháp truyền thống được trình bày trong Hình 1.

Nhược điểm của phương pháp truyền thống trong quá trình thực hiện thử ảo quần áo hai chiều (2D) gồm hai phần chính: một là độ chân thực: được tác động bởi vị trí và kích thước của phần đầu người sử dụng khi thực hiện ghép trực tiếp trên ảnh quần áo mẫu, hai là độ chính xác: không có phương pháp nào hiện tại có thể mô tả được chính xác kích thước của người sử dụng trên ảnh mẫu. Các phương pháp truyền thống gặp những giới hạn rất lớn khi gặp phải những vấn đề phía trên. Kỹ sư đồ họa mô phỏng có thể xử lý từng

trường hợp riêng lẻ, tuy nhiên chi phí nhân công và thời gian thực hiện là cực kỳ lớn và không thể mở rộng và ứng dụng ở phạm vi công nghiệp.

Bản chất kỹ thuật của sáng chế

Mục đích thứ nhất của sáng chế là đề xuất hệ thống thử ảo trang phục sử dụng công nghệ máy học và thuật toán tối ưu. Trong đó, các mô hình máy học về xử lý ảnh kết hợp với thuật toán tối ưu và thuật toán phân loại xác định để tái tạo ảnh người sử dụng kết hợp với ảnh quần áo và có biến đổi ảnh quần áo mẫu theo kích thước của người sử dụng. Việc xử lý và tạo đầu ra này cơ bản gồm hai phần chính: thứ nhất là sử dụng thuật toán tối ưu và mô hình tham số cơ thể người để xây dựng kích thước của người sử dụng, từ đó biến đổi hình dạng kích thước ảnh hai chiều (2D) trang phục, thứ hai là sử dụng các mô hình học máy để thay đổi và hiệu chỉnh đầu người sử dụng với ảnh mẫu hai chiều (2D) trang phục.

Để đạt được mục đích nêu trên, hệ thống thử ảo trang phục sử dụng công nghệ máy học kết hợp giữa thuật toán tối ưu và thuật toán phân loại xác định bao gồm bốn khối chính: khối tiền xử lý dữ liệu; khối cập nhật dáng đầu; khối thay đổi hình dáng; khối hoán đổi. Cụ thể:

Khối tiền xử lý dữ liệu gồm năm mô-đun: mô-đun phân đoạn, mô-đun xác định trục chính giữa cổ, mô-đun xác định điểm dấu mặt và cơ thể, mô-đun xác định hệ số mẫu, mô-đun xác định và phân loại mặt người sử dụng.

Khối thay đổi hình dáng gồm ba mô-đun: mô-đun ước lượng dữ liệu ba chiều (3D) người, mô-đun sinh dữ liệu bề mặt lưới hai chiều (2D), mô-đun trích xuất hình ảnh hai chiều (2D) được sử dụng để xác định kích thước người sử dụng, xác định hệ số trục chính giữa cổ mẫu cập nhật, thông tin phân đoạn ảnh mẫu cập nhật và thay đổi hình dạng ảnh mẫu quần áo hai chiều (2D) tương ứng với kích thước của người sử dụng.

Khối hoán đổi gồm bảy mô-đun: mô-đun xác định điểm dấu mặt người sử dụng, mô-đun phân vùng mặt cổ người sử dụng, mô-đun tái tạo vùng cổ người sử dụng bị che khuất, mô-đun đổi màu da, mô-đun cập nhật đầu người sử dụng, mô-đun cập nhật quần áo không đầu, mô-đun hoán đổi ảnh sử dụng các mô hình máy học và thuật toán tối ưu để thực hiện việc hoán đổi ảnh mẫu quần áo hai chiều (2D) và ảnh người sử dụng được cập nhật.

Khối hiệu chỉnh và tối ưu hoá gồm ba mô-đun: mô-đun điều chỉnh kích thước đầu người sử dụng, mô-đun điều chỉnh vị trí đầu người sử dụng, mô-đun xử lý màu da vùng tiếp giáp đồng đều.

Mục đích thứ hai của sáng chế là đề xuất phương pháp thử ảo trang phục hai chiều sử dụng công nghệ máy học kết hợp giữa thuật toán tối ưu và thuật toán phân loại xác định. Để đạt được mục đích này, phương pháp được đề xuất bao gồm sáu bước chính:

Bước 1: xử lý dữ liệu hình ảnh đầu vào và dữ liệu mẫu; mục đích của bước này là chuẩn hóa dữ liệu đầu vào cho bước 2 và bước 3 để tối ưu cập nhật dữ liệu người dùng. Các dữ liệu lưới dạng thô được hiệu chuẩn và biến đổi theo các đặc tính yêu cầu để đảm bảo việc kế thừa và sử dụng khi có sự cập nhật dữ liệu.

Bước 2: tối ưu cập nhật dữ liệu đầu người dùng; mục đích của bước này là hiệu chỉnh dữ liệu ảnh đầu người sử dụng để phù hợp, đầu ra được sử dụng tại bước 4.

Bước 3: tối ưu và thay đổi hình dạng cơ thể tương ứng; mục đích của bước này là tính toán mô hình người, sau đó biến đổi hình dạng ảnh hai chiều (2D) quần áo theo kích thước tương ứng, đầu ra của bước này được sử dụng tại bước 4.

Bước 4: hoán đổi đầu sử dụng và hình dạng cơ thể tương ứng; mục đích của bước này hoán đổi hình ảnh đầu người sử dụng với hình ảnh mẫu quần áo hai chiều (2D) được cập nhật, đầu ra được sử dụng tại bước 5.

Bước 5: hiệu chỉnh và tối ưu hoá giữa đầu và hình dạng cơ thể tương ứng; mục đích của bước này là thực hiện và chuẩn hóa tỷ lệ và vị trí của ảnh đầu người sử dụng tương đương và phù hợp với tỷ lệ và vị trí của thân ảnh mẫu, đầu ra được sử dụng tại bước 6

Bước 6: hiển thị hình ảnh tương tác hai chiều (2D); mục đích của bước này là việc hiển thị dưới dạng ba chiều (3D) của hình ảnh người sử dụng và hình ảnh quần áo mẫu đã được tính toán theo bước 5.

Mô tả vắn tắt các hình vẽ

Hình 1 là hình vẽ mô tả sơ đồ hệ thống truyền thống;

Hình 2 là hình vẽ mô tả sơ đồ hệ thống thử ảo trang phục được đề cập trong sáng chế;

Hình 3 là hình vẽ mô tả bối cảnh thu thập dữ liệu;

Hình 4 là hình vẽ mô tả chi tiết khối tiền xử lý dữ liệu;

Hình 5 là hình vẽ mô tả chín điểm đầu mặt có vị trí tương quan với xương hàm

Hình 6 là hình vẽ mô tả chi tiết khối thay đổi hình dáng;

Hình 7 là hình vẽ mô tả chi tiết khối hoán đổi;

Hình 8 là hình vẽ mô tả chi tiết khối hiệu chỉnh và tối ưu;

Hình 9 là hình vẽ mô tả sơ đồ hệ thống thử ảo trang phục được đề cập trong sáng chế.

Mô tả chi tiết sáng chế

Tham chiếu mô tả trong Hình 1 và Hình 2, sáng chế đề cập đến hệ thống và phương pháp mới nhằm tái tạo hình ảnh thử đồ của người sử dụng trên các mẫu quần áo mẫu tương ứng sử dụng công nghệ máy học kết hợp giữa thuật toán tối ưu và thuật toán phân loại xác định các mô hình máy học và thuật toán tối ưu, thay vì sử dụng các kỹ sư đồ họa mô phỏng và mô hình máy học cơ bản. Việc ứng dụng hệ thống và phương pháp sẽ giải quyết hai vấn đề chính đang gặp phải là độ chân thực và độ chính xác của đầu ra mô hình. Trong phần mô tả chi tiết giải pháp kỹ thuật được đề xuất, một số thuật ngữ được hiểu theo cách diễn giải như sau:

Khớp trong cơ thể người là điểm, hay đúng hơn là bề mặt, bộ phận kết nối vật lý giữa các xương với nhau.

Khớp trong trang phục, tương tự với khớp trong cơ thể người, là các đoạn nối giữa các khớp được định sẵn của trang phục.

Điểm đầu người/trang phục (điểm mốc) là các điểm đặc trưng trên ảnh chụp của người/trang phục. Điểm đầu thông thường là các điểm nằm trên đường biên và có ý nghĩa trong quá trình xác định hoặc phân vùng hoặc tham chiếu cho một bài toán cụ thể.

Đường biên người/trang phục là khái niệm liên quan đến đường giới hạn trang phục trên ảnh chụp. Đường biên trang phục quan trọng trong quá trình trích xuất và phân đoạn xác định các vùng dữ liệu khác nhau từ ảnh.

Ảnh mẫu trang phục là một ảnh hoặc một tập hợp các ảnh thường chụp trang phục theo quy chuẩn định trước như: ảnh mặt trước, là ảnh chụp mặt trước của trang phục được người mẫu mặc, trong đó người mẫu nhìn thẳng và tạo dáng với trang phục.

Bản đồ UV là bản đồ lưu trữ thông số để chiếu ảnh hai chiều (2D) lên bề mặt mô hình ba chiều (3D).

Tham chiếu Hình 2, hệ thống thử ảo trang phục sử dụng các mô hình học máy và thuật toán tối ưu có sự khác biệt rõ rệt so với hệ thống và phương pháp truyền thống. Cụ thể, hệ thống bao gồm bốn khối chính: khối tiền xử lý dữ liệu 100, khối thay đổi hình dáng 200, khối hoán đổi 300, khối hiệu chỉnh và tối ưu hóa 400; ngoài ra còn có khối đầu vào và khối đầu ra; bên cạnh đó phương pháp cũng được đề cập trong sáng chế. Trong đó:

Khối đầu vào gồm ba phần chính: ảnh người sử dụng, ảnh quần áo, thông tin chiều cao cân nặng người sử dụng. Dữ liệu đi qua khối tiền xử lý 100 để hiệu chỉnh, tiền xử lý để lấy các thông tin cần thiết từ người dùng và mẫu để sử dụng cho khối thay đổi hình dáng 200 và khối hiệu chỉnh và tối ưu hóa 400. Trong đó, khối thay đổi hình dáng 200 được sử dụng để biến đổi ảnh mẫu quần áo theo kích thước thực của người sử dụng. Đầu ra của khối này được sử dụng làm đầu vào cho khối hoán đổi 300 và khối hiệu chỉnh và tối ưu hóa 400, trong đó khối hoán đổi 300 có vai trò ghép giữa ảnh người sử dụng cập nhật và ảnh mẫu quần áo cập nhật. Đầu ra của khối 300 tiếp tục được đưa qua khối hiệu chỉnh và tối ưu 400 để hậu xử lý kết quả ghép đầu người dùng và mẫu quần áo, giúp đạt sự tự nhiên và chân thực trong kết quả. Đầu ra của khối là ảnh người sử dụng kết hợp với ảnh quần áo và là đầu ra của hệ thống.

Tham chiếu Hình 3, khối đầu vào thu thập dữ liệu bằng một thiết bị máy tính có màn hình và máy ảnh. Người dùng tự nhập chiều cao, cân nặng của bản thân vào thiết bị, sau đó chụp ảnh người sử dụng bằng máy ảnh gắn bên trên thiết bị với khoảng cách từ người dùng đến thiết bị từ 1,7 – 2m. Dữ liệu ảnh quần áo được chọn từ bộ nhớ trong của thiết bị.

Tham chiếu Hình 4, dữ liệu đầu vào trong khối tiền xử lý dữ liệu 100 gồm ba phần như đã đề cập trong khối đầu vào: ảnh người sử dụng, ảnh quần áo, thông tin chiều cao cân nặng người sử dụng. Năm mô-đun trong khối tiền xử lý dữ liệu 100 gồm mô-đun phân đoạn 101, mô-đun xác định trục chính giữa cổ 102, mô-đun xác định điểm dấu mặt và cơ thể 103, mô-đun xác định hệ số mẫu 104, mô-đun xác định và phân loại kiểu mặt 105.

Trong đó mô-đun phân đoạn 101 sử dụng mô hình học máy với mục đích xác định đường biên trang phục và người mẫu với nền xung quanh, sử dụng đầu vào là ảnh quần áo trang phục. Mô-đun xác định điểm dấu mặt và cơ thể 103 sử dụng mô hình học máy xác định và phân loại các điểm dấu trên mặt và cơ thể dựa trên ảnh quần áo mẫu. Đầu ra của mô-đun 101 và mô-đun 103 được sử dụng kết hợp với ảnh quần áo làm đầu vào cho mô-đun xác định trục chính giữa cổ 102. Từ các điểm dấu thu được từ mô-đun xác định điểm dấu mặt và cơ thể 103, chín điểm có vị trí tương quan với xương hàm được lựa chọn (tham chiếu Hình 5) và nối một đường giữa các điểm này, đường này được định nghĩa là đường xương quai hàm (*jawline*). Đường xương quai hàm được mở rộng bởi công thức:

$$jawline_{ext}(x, y) = \max(jawline(x - i, y - j) | S(i, j))$$

Trong đó:

(x, y) là tọa độ điểm ảnh trên ảnh biểu diễn đường xương quai hàm (*jawline*);

(i, j) là tọa độ điểm ảnh trên yếu tố cấu trúc S , kích thước của yếu tố cấu trúc S tỷ lệ thuận với khoảng cách đồng tử được ước lượng bằng các điểm dấu mặt thu được từ mô-đun 103;

$jawline_{ext}(x, y)$ là giá trị điểm ảnh sau khi mở rộng tại tọa độ (x, y) ;

$S(i, j)$ là giá trị điểm ảnh của yếu tố cấu trúc tại tọa độ (i, j) .

$\max()$ là toán tử tìm giá trị lớn nhất trong các điểm ảnh tương quan giữa ảnh biểu diễn đường xương quai hàm và yếu tố cấu trúc.

Vùng cằm (*chin*) được định nghĩa là vùng giao giữa đường xương quai hàm mở rộng và vùng cổ phân đoạn được thu lại từ mô-đun phân đoạn 101 và được biểu diễn bởi công thức:

$$chin_{mask} = jawline_{ext}(x, y) \cap neck_{mask}$$

Với $chin_{mask}$ và $neck_{mask}$ lần lượt là vùng cằm và vùng phân đoạn cổ. Giả thiết rằng, trên $chin_{mask}$, các điểm ảnh mang giá trị 1 thể hiện nằm trong vùng cằm và các điểm ảnh mang giá trị 0 thể hiện nằm ngoài vùng cằm. Đường trục chính giữa cổ được thể hiện bằng phương trình:

$$x = \frac{(x_{min} + x_{max})}{2}$$

$$x_{min} = \min(x) \forall (x, y) | chin_{mask}(x, y) = 1$$

$$x_{max} = \max(x) \forall (x, y) | chin_{mask}(x, y) = 1$$

Trong đó:

x_{min} và x_{max} lần lượt là giá trị nhỏ nhất và lớn nhất theo trục x của vùng cầm $chin_{mask}$.

Mô-đun xác định hệ số mẫu 104 sử dụng thuật toán tối ưu phi tuyến để xác định các thông số của người mẫu trong ảnh, các thông số được sử dụng làm gốc chuẩn để có thể sử dụng làm đầu vào cho khối thay đổi hình dáng 200. Bộ thông số của người mẫu được đặc trưng bởi các tham số tư thế, hình dáng của mô hình tham số người và các tham số máy ảnh ảo để chiếu mô hình tham số người lên ảnh 2 chiều, Các tham số được khởi tạo thô và được thay đổi liên tục thông qua thuật toán tối ưu Adam để tìm nghiệm làm tối thiểu hàm mục tiêu sau:

$$E = \omega_J \cdot E_J(\theta, \beta, K, J_e) + \omega_{weight} \cdot E_{weight} + \omega_{height} \cdot E_{height} + \omega_{\beta} \cdot E_{\beta}$$

Trong đó:

$$E_J = \sum_i \omega_i \rho(J_M - J_e): \text{sai số vị trí khớp trên ảnh hai chiều (2D);}$$

$$E_{weight} = \|pd_weight - gt_weight\|: \text{sai số cân nặng;}$$

$$E_{height} = \|pd_height - gt_height\|: \text{sai số chiều cao;}$$

$$E_{\beta}^m = \beta^T \Sigma_{\beta}^{-1} \beta: \text{sai số hình dạng;}$$

M kí hiệu mô hình tham số người

$\omega_J, \omega_{weight}, \omega_{height}, \omega_{\beta}$: là các trọng số thực;

θ, β : các tham số về tư thế hình dạng của mô hình tham số M;

K : các tham số của camera;

J_M : kí hiệu hình chiếu vị trí các khớp ba chiều (3D) của mô hình tham số M lên ảnh;

J_e : kí hiệu vị trí các khớp hai chiều (2D) của người thực trên ảnh ρ kí hiệu hàm Geman-McClure.

Mô-đun xác định và phân loại kiểu mặt 105 sử dụng đầu vào từ mô-đun phân đoạn 101 và mô-đun xác định điểm dấu mặt và cơ thể 103 để tính toán các thông số về độ rộng trán ($d_{trán}$), độ rộng gò má ($d_{má}$), độ rộng cằm ($d_{cằm}$), và độ dài gương mặt ($d_{mặt}$). So sánh giá trị các thông số này với nhau, ta kết luận được kiểu mặt của người đang xét. Các loại kiểu mặt được xét là mặt trái xoan, mặt chữ nhật dài và mặt tròn và được định nghĩa cụ thể như sau:

Mặt trái xoan: $(d_{mặt} > d_{má}) \& (d_{trán} > d_{cằm})$

Mặt chữ nhật dài: $d_{mặt} > d_{má} \approx d_{trán} \approx d_{cằm}$

Mặt tròn: $d_{má} \approx d_{mặt} > d_{trán} \approx d_{cằm}$

Đầu ra của khối tiền xử lý dữ liệu 100 được sử dụng tại các khối cập nhật dáng đầu 200 và khối hiệu chỉnh và tối ưu hóa 400.

Tham chiếu Hình 6, dữ liệu đầu vào trong khối thay đổi hình dáng 200 gồm năm phần: thông tin phân đoạn ảnh mẫu, hệ số trục chính giữa cổ mẫu, thông tin điểm dấu trên mặt và cơ thể mẫu, thông tin hệ số mẫu là đầu ra của khối tiền xử lý dữ liệu 100 và thông tin chiều cao cân nặng người sử dụng từ khối đầu vào. Khối thay đổi hình dáng 200 gồm ba mô-đun chính: mô-đun ước lượng dữ liệu ba chiều (3D) người 201, mô-đun sinh dữ liệu bề mặt lưới hai chiều (2D) 202, và mô-đun trích xuất hình ảnh hai chiều (2D) 203. Trong đó, mô-đun ước lượng dữ liệu ba chiều (3D) người 201 sử dụng thuật toán tối ưu để xác định mô hình lưới ba chiều (3D) người dựa trên thông tin chiều cao cân nặng với hàm mục tiêu tương tự như mô-đun xác định hệ số mẫu 103, tuy nhiên chỉ sử dụng ba hàm mục tiêu E_{weight} , E_{height} , E_{β} . Mô-đun sinh dữ liệu bề mặt lưới hai chiều (2D) 202 sử dụng thông tin hệ số mẫu để sinh mô hình lưới ba chiều (3D) và sử dụng phép chiếu phối cảnh chiếu lưới ba chiều (3D) thành các điểm hai chiều (2D) khớp với hình dạng người mẫu trong ảnh. Sau đó sử dụng thuật toán tam giác Delaunay sinh lưới tam giác cho các điểm hai chiều (2D) tạo được, thực hiện chuẩn hóa về dải $[0, 1]$ để tạo bản đồ UV và gắn ảnh người mẫu thành ảnh kết cấu (texture) cho mô hình lưới hai chiều (2D). Mô-đun trích xuất hình ảnh hai chiều (2D) 203 sử dụng phép chiếu phối cảnh chiếu mô hình lưới ba chiều (3D) người thu được trong mô-đun ước lượng dữ liệu ba chiều (3D) 301 người thành các điểm hai chiều (2D).

Để biến đổi hình dạng của người trên ảnh hai chiều (2D) cần tính toán ma trận biến đổi của từng điểm ảnh trên ảnh. Quá trình được thực hiện thông qua hai bước: thứ nhất là tính toán ma trận biến đổi của các điểm chiếu từ mô hình người ba chiều (3D) và thứ hai là thực hiện nội suy ma trận của toàn bộ điểm ảnh. Ma trận biến đổi giữa các điểm chiếu của mô hình người mẫu và người sử dụng được tính toán thông qua thuật toán tối ưu sử dụng hàm mục tiêu sau:

$$E_A = \omega_p E_p(J_m, J_u) + \omega_e E_e$$

Trong đó:

ω_p, ω_e : là các trọng số thực;

$E_p(J_m, J_u) = \|J_m - J_u\|_2$: sai số vị trí điểm ảnh;

J_m, J_u : vị trí 2 chiều của các điểm chiếu của mô hình người mẫu và người dùng;

$E_e = \sum_i \sum_j \|A_i - A_j\|$: sai số đảm bảo giữ cấu trúc của hình ảnh. A_i, A_j là các ma trận biến đổi điểm ảnh i, j thuộc cùng một cạnh của tam giác trên mô hình lưới hai chiều (2D) sinh từ mô-đun sinh dữ liệu bề mặt lưới hai chiều (2D) 202.

Ma trận của toàn bộ điểm ảnh được nội suy từ các ma trận biến đổi của các điểm chiếu mô hình người theo:

$$A_k = \frac{1}{n} \sum_l A_l$$

A_k : ma trận biến đổi của điểm ảnh thứ k ;

A_l : ma trận biến đổi của điểm lân cận với điểm ảnh thứ k của điểm ảnh thứ l trên các điểm chiếu của mô hình người.

Đầu ra của khối thay đổi hình dáng 200 bao gồm ảnh mẫu cập nhật, hệ số trục chính giữa cổ mẫu cập nhật, thông tin điểm dấu trên mặt và cơ thể mẫu cập nhật, và thông tin phân đoạn ảnh mẫu cập nhật được sử dụng làm đầu vào cho khối hoán đổi 300 và khối hiệu chỉnh và tối ưu hóa 400.

Tham chiếu Hình 7, dữ liệu đầu vào của khối hoán đổi 300 gồm bốn phần: thông tin phân đoạn ảnh mẫu cập nhật và ảnh mẫu cập nhật từ khối thay đổi hình dáng 200, ảnh

người sử dụng từ khối đầu vào, và thông tin điểm dấu trên mặt và cơ thể mẫu cập nhật từ khối thay đổi hình dáng 200. Khối hoán đổi 300 bao gồm sáu mô-đun chính: mô-đun phân vùng mặt cổ người sử dụng 301, mô-đun xác định điểm dấu mặt người sử dụng 302, mô-đun tái tạo vùng cổ người sử dụng bị che khuất 303, mô-đun đổi màu da 304, mô-đun xác định kiểu mặt 305, và mô-đun hoán đổi ảnh 306. Trong đó, mô-đun phân vùng mặt cổ người sử dụng 301 dùng mạng học sâu để chia hình ảnh chân dung thành mười tám phân vùng (vùng da, mũi, kính mắt, mắt trái, mắt phải, lông mày trái, lông mày phải, tai trái, tai phải, miệng, môi trên, môi dưới, tóc, mũ, khuyên tai, vòng cổ, cổ, và quần áo) qua đó loại bỏ các phần không cần thiết (bao gồm vùng nền (vùng không thuộc mười tám vùng trên), quần áo, vòng cổ). Mô-đun xác định điểm dấu mặt người sử dụng 302 sử dụng mạng học sâu trích xuất sáu tám tọa độ hai chiều các điểm dấu trên hình ảnh khuôn mặt người sử dụng. Với đầu ra từ mô-đun phân vùng mặt cổ người sử dụng 301, mô-đun tái tạo vùng cổ người sử dụng bị che khuất 303 dùng mô hình mạng đối nghịch tạo sinh (GAN) để tự động nhận dạng và phục hồi vùng cổ người sử dụng bị che khuất bởi quần áo. Mô hình mạng đối nghịch tạo sinh (GAN) được định nghĩa bởi mỗi không gian xác suất (Ω, μ_{ref}) và bao gồm hai bộ chính: bộ tạo sinh và bộ phân biệt. Bộ tạo sinh sử dụng $P(\Omega)$ là bộ của tất cả xác suất đo μ_G trên Ω ; bộ phân biệt sử dụng nhân Markov $\mu_D : \Omega \rightarrow P[0,1]$, với $P[0,1]$ là bộ xác suất đo trên $[0,1]$. Hàm mục tiêu của GAN được biểu diễn bằng công thức (được trích xuất trong bài báo Generative Adversarial Nets của Ian J. Goodfellow):

$$L(\mu_G, \mu_D) := E_{x \sim \mu_{ref}, y \sim \mu_D(x)} [\ln y] + E_{x \sim \mu_G, y \sim \mu_D(x)} [\ln(1 - y)]$$

Trong hàm mục tiêu, nhiệm vụ của bộ tạo sinh là tối thiểu hóa giá trị còn bộ phân biệt có tác dụng tối đa hóa giá trị. Sự tương quan giữa hai bộ này khiến mô hình GAN có thể được coi như một trò chơi có tổng bằng 0.

Mô-đun đổi màu da 304 sử dụng thuật toán phân cụm K-means để phân vùng xác định màu da của người dùng và người mẫu qua đó thay đổi màu da của người mẫu thành màu da của người dùng bằng cách điều chỉnh đường đặc tính ánh sáng cho từng kênh màu trong ảnh. Mô-đun xác định kiểu mặt 305 phân loại mặt người dùng tương tự như mô-đun xác định và phân loại kiểu mặt 105 được sử dụng trong khối tiền xử lý dữ liệu 100. Mô-đun hoán đổi ảnh 306 tính toán ma trận biến đổi M từ các điểm dấu thu được trong mô-đun xác định điểm dấu mặt 301 và sử dụng thuật toán biến dạng hình ảnh để biến đổi

ảnh khuôn mặt người dùng sao cho khớp với khuôn mặt người mẫu trên ảnh. Ma trận M được tính toán dựa trên các phép biến đổi ma trận không gian như sau:

$$M = \begin{bmatrix} scale.R & c_u^T - scale.R.c_m^T \\ 0 & 1 \end{bmatrix}$$

$$R = (U.(V)^T)^T$$

$$U, \Sigma, (V)^T = SVD((lmk_m)^T.lmk_u)$$

$$c_u = mean(lmk_u), c_m = mean(lmk_m)$$

$$\|lmk_m\| = \frac{lmk_m - c_m}{std(lmk_m - c_m)}; \|lmk_u\| = \frac{lmk_u - c_u}{std(lmk_u - c_u)}$$

$$scale = \frac{std(lmk_u - c_u)}{std(lmk_m - c_m)}$$

Trong đó: lmk_m , lmk_u là các điểm dấu xác định được trên khuôn mặt người mẫu và người dùng; $mean$: tính giá trị trung bình của tập dữ liệu; std : đo mức độ phân tán của tập dữ liệu quanh giá trị trung bình; SVD : phép phân tích một ma trận thành tích của các ma trận trực giao và ma trận đường chéo không vuông.

Đầu ra của khối hoán đổi 300 là thông tin kiểu mặt người dùng, ảnh người sử dụng kết hợp với ảnh quần áo, và điểm dấu mặt người sử dụng tương ứng với ảnh quần áo.

Tham chiếu Hình 8, dữ liệu đầu vào của khối hiệu chỉnh và tối ưu hóa 400 bao gồm: thông tin điểm dấu trên mặt và cơ thể mẫu cập nhật từ khối 200, điểm dấu mặt người sử dụng tương ứng với ảnh quần áo từ khối 300, thông tin kiểu mặt mẫu từ khối 100, thông tin kiểu mặt người dùng từ khối 300, hệ số trục chính giữa cơ mẫu cập nhật từ khối 200, và ảnh người sử dụng kết hợp với ảnh quần áo từ khối 300. Trong khối hiệu chỉnh và tối ưu hóa 400, mô-đun tính toán kích thước đầu người sử dụng 401 dùng tập điểm dấu của người sử dụng và mẫu để điều chỉnh tỷ lệ đầu phù hợp. Tỷ lệ đầu được tính với các trường hợp khác nhau như sau:

Đối với nam:

Hình chụp cả người:

$$scale_{head} = scale_{chin}$$

Hình chụp nửa người

$$scale_{head} = (scale_{eye} + scale_{chin} * 3) / 4$$

Đối với nữ:

$$scale_{head} = (scale_{eye} + scale_{chin}) / 2$$

Trong đó, $scale_{head}$ là tỷ lệ đầu cần điều chỉnh, $scale_{eye}$ và $scale_{chin}$ lần lượt là tỷ lệ mắt và cằm giữa người sử dụng và ảnh mẫu được tính toán dựa vào hai tập điểm dấu tương ứng của người sử dụng và ảnh mẫu. Mô-đun tính toán vị trí đầu người sử dụng 403 điều chỉnh vị trí đầu người sử dụng vào chính giữa cổ ảnh mẫu theo trục hoành của ảnh. Mô-đun so sánh kiểu mặt người dùng và mẫu 402 so sánh thông tin kiểu mặt giữa người dùng và mẫu để bổ sung thêm thông tin về tỷ lệ đầu và vị trí đầu phù hợp. Mô-đun hiệu chỉnh tính toán vị trí và kích thước đầu người dùng 404 nhận thông tin từ các mô-đun cập nhật kích thước và vị trí đầu người dùng 401, 402, và 403 để tính toán ma trận biến đổi M mới tương tự như mô-đun hoán đổi ảnh 306 trong khối hoán đổi ảnh 300. Mô-đun xử lý màu da vùng tiếp giáp đồng đều 405 sử dụng phương trình Poisson kết hợp điều kiện biên Dirichlet. Giá trị trường độ dốc trong vùng ảnh ghép được tính toán và điều chỉnh cho phù hợp với ảnh người sử dụng, hạn chế tối đa sự chênh lệch màu sắc trong vùng da tiếp giáp giữa người sử dụng và ảnh mẫu. Đầu ra của khối hiệu chỉnh và tối ưu hóa 400 là ảnh người sử dụng kết hợp với ảnh quần áo đã được tối ưu và hiệu chỉnh.

Tham chiếu Hình 9, phương pháp thử ảo trang phục hai chiều (2D) sử dụng công nghệ máy học kết hợp giữa thuật toán tối ưu và thuật toán phân loại xác định bao gồm năm bước chính:

Bước 1: xử lý dữ liệu hình ảnh dữ liệu mẫu; mục đích của bước này là chuẩn hóa dữ liệu đầu vào cho bước 2 và bước 3. Các dữ liệu lưới dạng thô được hiệu chuẩn và biến đổi theo các đặc tính yêu cầu để đảm bảo việc kế thừa và sử dụng khi có sự cập nhật dữ liệu. Quá trình được thực hiện trên khối tiền xử lý dữ liệu 100. Dữ liệu đầu vào trong khối tiền xử lý dữ liệu 100 gồm ba phần như đã đề cập trong khối đầu vào: ảnh người sử dụng, ảnh quần áo, thông tin chiều cao cân nặng người sử dụng. Năm mô-đun trong khối tiền xử lý dữ liệu 100 gồm mô-đun phân đoạn 101, mô-đun xác định trục chính giữa cổ 102, mô-đun xác định điểm dấu mặt và cơ thể, mô-đun xác định hệ số mẫu 104, và mô-đun xác định và phân loại kiểu mặt 105. Trong đó, mô-đun phân đoạn 101 sử dụng các mô hình máy học

phân đoạn để xác định đường biên trang phục và người mẫu với nền xung quanh, sử dụng đầu vào là ảnh quần áo trang phục. Mô-đun trích xuất ảnh và gắn điểm dấu mặt 103 sử dụng mô hình máy học xác định và phân loại để xác định các điểm dấu trên ảnh mẫu quần áo, gồm hai phần: một là điểm dấu trên mặt mẫu mặc quần áo, hai là điểm dấu trên quần áo mẫu. Từ các điểm dấu thu được từ mô-đun xác định điểm dấu mặt và cơ thể 103, tám điểm có vị trí tương quan với xương hàm được lựa chọn và nối một đường giữa các điểm này, đường này được định nghĩa là đường xương quai hàm (*jawline*). Đường xương quai hàm được mở rộng bởi công thức:

$$jawline_{ext}(x, y) = \max(jawline(x - i, y - j) \mid S(i, j))$$

Trong đó:

(x, y) là tọa độ điểm ảnh trên ảnh biểu diễn đường xương quai hàm *jawline*.

(i, j) là tọa độ điểm ảnh trên yếu tố cấu trúc S , kích thước của yếu tố cấu trúc S tỷ lệ thuận với khoảng cách đồng tử được ước lượng bằng các điểm dấu mặt thu được từ mô-đun 103.

$jawline_{ext}(x, y)$ là giá trị điểm ảnh sau khi mở rộng tại tọa độ (x, y) .

$S(i, j)$ là giá trị điểm ảnh của yếu tố cấu trúc tại tọa độ (i, j) .

$\max()$ là toán tử tìm giá trị lớn nhất trong các điểm ảnh tương quan giữa ảnh biểu diễn đường xương quai hàm và yếu tố cấu trúc.

Vùng cằm được định nghĩa là vùng giao giữa đường xương quai hàm mở rộng và vùng cổ phân đoạn được thu lại từ mô-đun phân đoạn 101 và được biểu diễn bởi công thức:

$$chin_{mask} = jawline_{ext}(x, y) \cap neck_{mask}$$

Với $chin_{mask}$ và $neck_{mask}$ lần lượt là vùng cằm và vùng phân đoạn cổ. Giả thiết rằng, trên $chin_{mask}$, các điểm ảnh mang giá trị 1 thể hiện nằm trong vùng cằm và các điểm ảnh mang giá trị 0 thể hiện nằm ngoài vùng cằm. Đường trục chính giữa cổ được thể hiện bằng phương trình:

$$x = \frac{(x_{min} + x_{max})}{2}$$

$$x_{min} = \min(x) \forall (x, y) | chin_{mask}(x, y) = 1$$

$$x_{max} = \max(x) \forall (x, y) | chin_{mask}(x, y) = 1$$

Mô-đun xác định hệ số mẫu 104 sử dụng thuật toán tối ưu phi tuyến để xác định các thông số của người mẫu trong ảnh. Mô-đun xác định và phân loại kiểu mặt 105 sử dụng đầu vào từ mô-đun phân đoạn 101 và mô-đun xác định điểm dấu mặt và cơ thể 103 để tính toán các thông số về độ rộng trán, độ rộng gò má, độ rộng cằm. Mô-đun xác định và phân loại kiểu mặt 105 sử dụng đầu vào từ mô-đun phân đoạn 101 và mô-đun xác định điểm dấu mặt và cơ thể 103 để tính toán các thông số về độ rộng trán ($d_{trán}$), độ rộng gò má ($d_{má}$), độ rộng cằm ($d_{cằm}$), và độ dài gương mặt ($d_{mặt}$). So sánh giá trị các thông số này với nhau, ta kết luận được kiểu mặt của người đang xét. Các loại kiểu mặt được xét là mặt trái xoan, mặt chữ nhật dài và mặt tròn và được định nghĩa cụ thể như sau:

Mặt trái xoan: $(d_{mặt} > d_{má}) \& (d_{trán} > d_{cằm})$

Mặt chữ nhật dài: $d_{mặt} > d_{má} \approx d_{trán} \approx d_{cằm}$

Mặt tròn: $d_{má} \approx d_{mặt} > d_{trán} \approx d_{cằm}$

Đầu ra của khối tiền xử lý dữ liệu được sử dụng bước 2 và bước 4.

Bước 2: tối ưu và thay đổi hình dạng cơ thể tương ứng; mục đích của bước này là tính toán mô hình người, sau đó biến đổi hình dạng ảnh hai chiều (2D) quần áo theo kích thước tương ứng, đầu ra của bước này được sử dụng tại bước 3. Bước 2 được thực hiện trên khối thay đổi hình dáng. Dữ liệu đầu vào của khối thay đổi hình dáng gồm năm phần: thông tin phân đoạn ảnh mẫu, hệ số trục chính giữa cổ mẫu, thông tin điểm dấu trên mặt và cơ thể mẫu, thông tin chiều cao cân nặng người sử dụng, và thông tin hệ số mẫu. Khối thay đổi hình dáng 300 gồm ba mô-đun chính: mô-đun ước lượng dữ liệu ba chiều (3D) người 301, mô-đun sinh dữ liệu bề mặt lưới hai chiều (2D) 302, mô-đun trích xuất hình ảnh hai chiều (2D) 303. Trong đó, mô-đun ước lượng dữ liệu ba chiều (3D) người 301 sử dụng thuật toán tối ưu để xác định mô hình lưới ba chiều (3D) người dựa trên thông tin chiều cao cân nặng và các đặc tính trích xuất được từ ảnh người dùng sử dụng mạng học sâu. Mô-đun sinh dữ liệu bề mặt lưới hai chiều (2D) 302 sử dụng thông tin hệ số mẫu để sinh mô hình lưới ba chiều (3D) và sử dụng phép chiếu phối cảnh chiếu lưới ba chiều (3D) thành các điểm hai chiều (2D) khớp với hình dạng người mẫu trong ảnh. Sau đó sử dụng

thuật toán tam giác Delaunay sinh lưới tam giác cho các điểm hai chiều (2D) tạo được, thực hiện chuẩn hóa về dải $[0, 1]$ để tạo bản đồ UV và gắn ảnh người mẫu thành ảnh kết cấu (texture) cho mô hình lưới hai chiều (2D). Mô-đun trích xuất hình ảnh hai chiều (2D) 303 sử dụng phép chiếu phối cảnh chiếu mô hình lưới ba chiều (3D) người thu được trong mô-đun ước lượng dữ liệu ba chiều (3D) người thành các điểm hai chiều (2D) và thực hiện gán các điểm thu được lên mô hình lưới hai chiều (2D) trong mô-đun sinh dữ liệu bề mặt lưới hai chiều (2D) để biến đổi hình dạng của lưới. Sau đó thực hiện cài đặt môi trường ảo (máy ảnh ảo, ánh sáng) để chiếu mô hình lưới thành hình ảnh hai chiều. Đầu ra của khối cập nhật hình dáng được sử dụng làm đầu vào cho khối hoán đổi.

Bước 3: hoán đổi đầu sử dụng và hình dạng cơ thể tương ứng; mục đích của bước này hoán đổi hình ảnh đầu người sử dụng với hình ảnh mẫu quần áo hai chiều (2D) được cập nhật, đầu ra được sử dụng tại bước 4. Bước 3 được thực hiện trên khối hoán đổi 300. Dữ liệu đầu vào của khối hoán đổi gồm bốn phần: thông tin phân đoạn ảnh mẫu cập nhật, ảnh mẫu cập nhật, ảnh người sử dụng, và thông tin điểm dấu trên mặt và cơ thể mẫu cập nhật. Khối hoán đổi 300 gồm sáu mô-đun chính: mô-đun phân vùng mặt cổ người sử dụng 301, mô-đun xác định điểm dấu mặt người sử dụng 302, mô-đun tái tạo vùng cổ người sử dụng bị che khuất 303, mô-đun đổi màu da 304, mô-đun xác định kiểu mặt 305, và mô-đun hoán đổi ảnh. Trong đó, mô-đun phân vùng mặt cổ người sử dụng 301 sử dụng mạng học sâu để chia hình ảnh chân dung thành mười tám phân vùng (hình nền, cổ, tóc, ...) qua đó loại bỏ các phần không cần thiết (hình nền, trang phục của người dùng). Mô-đun xác định điểm dấu mặt người sử dụng 302 sử dụng mạng học sâu trích xuất sáu tám tọa độ hai chiều các điểm dấu trên hình ảnh khuôn mặt người dùng cập nhật và ảnh mẫu cập nhật. Mô-đun tái tạo vùng cổ người sử dụng bị che khuất 303 dùng mô hình mạng đối nghịch tạo sinh (GAN) để tự động nhận dạng và phục hồi vùng cổ người sử dụng bị che khuất bởi quần áo. Mô hình mạng đối nghịch tạo sinh (GAN) được định nghĩa bởi mỗi không gian xác suất (Ω, μ_{ref}) và bao gồm hai bộ chính: bộ tạo sinh và bộ phân biệt. Bộ tạo sinh sử dụng $P(\Omega)$ là bộ của tất cả xác suất đo μ_G trên Ω ; bộ phân biệt sử dụng nhân Markov $\mu_D : \Omega \rightarrow P[0,1]$, với $P[0,1]$ là bộ xác suất đo trên $[0,1]$. Hàm mục tiêu của GAN được biểu diễn bằng công thức thức (được trích xuất trong bài báo Generative Adversarial Nets của Ian J. Goodfellow):

$$L(\mu_G, \mu_G) := E_{x \sim \mu_{ref}, y \sim \mu_D(x)} [\ln y] + E_{x \sim \mu_G, y \sim \mu_D(x)} [\ln(1 - y)]$$

Trong hàm mục tiêu, nhiệm vụ của bộ tạo sinh là tối thiểu hóa giá trị còn bộ phân biệt có tác dụng tối đa hóa giá trị. Sự tương quan giữa hai bộ này khiến mô hình GAN có thể được coi như một trò chơi có tổng bằng 0.

Mô-đun đổi màu da 304 sử dụng thuật toán phân cụm *K-means* để phân vùng xác định màu da của người dùng và người mẫu qua đó thay đổi màu da của người mẫu thành màu da của người dùng bằng cách điều chỉnh đường đặc tính ánh sáng cho từng kênh màu trong ảnh. Mô-đun xác định kiểu mặt 305 phân loại mặt người dùng tương tự như Mô-đun xác định và phân loại kiểu mặt 105 được sử dụng trong khối tiền xử lí dữ liệu 100. Mô-đun hoán đổi ảnh 306 tính toán ma trận biến đổi M từ các điểm dấu thu được trong mô-đun xác định điểm dấu mặt người sử dụng 302 và sử dụng thuật toán biến dạng hình ảnh để biến đổi ảnh khuôn mặt người dùng sao cho khớp với khuôn mặt người mẫu trên ảnh. Ma trận M được tính toán dựa trên các phép biến đổi ma trận không gian như sau:

$$M = \begin{bmatrix} scale.R & c_u^T - scale.R.c_m^T \\ 0 & 1 \end{bmatrix}$$

$$R = (U.(V)^T)^T$$

$$U, \Sigma, (V)^T = SVD((lmk_m)^T.lmk_u)$$

$$c_u = mean(lmk_u), c_m = mean(lmk_m)$$

$$\|lmk_m\| = \frac{lmk_m - c_m}{std(lmk_m - c_m)}; \|lmk_u\| = \frac{lmk_u - c_u}{std(lmk_u - c_u)}$$

$$scale = \frac{std(lmk_u - c_u)}{std(lmk_m - c_m)}$$

Trong đó: lmk_m, lmk_u là các điểm dấu xác định được trên khuôn mặt người mẫu và người dùng; *mean*: tính giá trị trung bình của tập dữ liệu; *std*: đo mức độ phân tán của tập dữ liệu quanh giá trị trung bình; *SVD*: phép phân tích một ma trận thành tích của các ma trận trực giao và ma trận đường chéo không vuông.

Bước 4: hiệu chỉnh và tối ưu hóa giữa đầu và hình dạng cơ thể tương ứng; mục đích của bước này là tinh chỉnh về vị trí, kích thước và màu da của đầu người sử dụng phù hợp với cơ thể mẫu tương ứng. Bước 4 được thực hiện trên khối hiệu chỉnh và tối ưu hóa 400. Dữ liệu đầu vào của khối hiệu chỉnh và tối ưu hóa gồm sáu phần: thông tin điểm dấu trên

mặt và cơ thể mẫu cập nhật, điểm dấu mặt người sử dụng tương ứng với ảnh quần áo, thông tin kiểu mặt mẫu, thông tin kiểu mặt người dùng, hệ số trục chính giữa cổ mẫu cập nhật, và ảnh người sử dụng kết hợp với ảnh quần áo. Khối hiệu chỉnh và tối ưu hóa 400 gồm năm mô-đun chính: mô-đun tính toán kích thước đầu người sử dụng 401, mô-đun so sánh kiểu mặt người dùng và mẫu 402, mô-đun tính toán vị trí đầu người sử dụng, mô-đun hiệu chỉnh tính toán vị trí và kích thước đầu người dùng 404, và mô-đun xử lý màu da vùng tiếp giáp đồng đều 405. Mô-đun tính toán kích thước đầu người sử dụng 401 tính toán tỷ lệ kích thước đầu ảnh mẫu so với đầu người sử dụng từ các điểm dấu trên mặt của ảnh mẫu và người sử dụng. Mô-đun so sánh kiểu mặt người dùng và mẫu 402 sử dụng thông tin kiểu mặt mẫu và thông tin kiểu mặt người dùng để so sánh về độ rộng trán, độ rộng gò má và độ rộng cằm của hai đầu. Mô-đun tính toán vị trí đầu người sử dụng 403 so sánh vị trí đầu người với hệ số trục chính giữa cổ mẫu cập nhật để dịch hình ảnh đầu người dùng về khoảng cách phù hợp. Mô-đun hiệu chỉnh tính toán vị trí và kích thước đầu người dùng 404 sử dụng thông tin về kích thước, vị trí đầu người dùng so với ảnh mẫu để điều chỉnh đầu người dùng về vị trí mong muốn với ma trận M tương tự mô-đun hoán đổi ảnh 306 trong khối hoán đổi 300. Mô-đun xử lý màu da vùng tiếp giáp đồng đều 405 sử dụng phương trình Poisson kết hợp điều kiện biên Dirichlet để cân bằng trường độ dốc trong vùng da tiếp giáp giữa người sử dụng và ảnh mẫu. Đầu ra của bước 4 là ảnh người sử dụng kết hợp với ảnh quần áo đã được tối ưu, hiệu chỉnh.

Bước 5: hiển thị hình ảnh tương tác hai chiều (2D); mục đích của bước này là việc hiển thị dưới dạng ba chiều (2D) của hình ảnh người sử dụng và hình ảnh quần áo mẫu đã được tính toán theo bước 4.

Yêu cầu bảo hộ

1. Hệ thống thử ảo trang phục hai chiều (2D) sử dụng công nghệ máy học kết hợp giữa thuật toán tối ưu và thuật toán phân loại xác định bao gồm:

khởi tiên xử lý dữ liệu (100) sử dụng kỹ thuật xử lý ảnh và mô hình học máy để xử lý, chuẩn hóa và quy chuẩn các dữ liệu ảnh đầu vào, xác định và phân đoạn các vùng trên ảnh người sử dụng và ảnh mẫu quần áo hai chiều (2D); gồm có:

khởi tiên xử lý dữ liệu (100) bao gồm năm mô-đun, gồm có,

mô-đun phân đoạn (101),

mô-đun xác định trục chính giữa cổ (102),

mô-đun xác định điểm dấu mặt và cơ thể (103),

mô-đun xác định hệ số mẫu (104),

mô-đun xác định và phân loại kiểu mặt (105),

trong đó mô-đun phân đoạn (101) sử dụng mô hình học máy với mục đích xác định đường biên trang phục và người mẫu với nền xung quanh, sử dụng đầu vào là ảnh quần áo trang phục; mô-đun xác định điểm dấu mặt và cơ thể (103) sử dụng mô hình học máy xác định và phân loại các điểm dấu trên mặt và cơ thể dựa trên ảnh quần áo mẫu; đầu ra của mô-đun phân đoạn (101) và mô-đun xác định điểm dấu mặt và cơ thể (103) được sử dụng kết hợp với ảnh quần áo làm đầu vào cho mô-đun xác định trục chính giữa cổ (102); từ các điểm dấu thu được từ mô-đun xác định điểm dấu mặt và cơ thể (103), chín điểm có vị trí tương quan với xương hàm được lựa chọn và nối một đường giữa các điểm này, đường này được định nghĩa là đường xương quai hàm (*jawline*); đường xương quai hàm được mở rộng bởi công thức:

$$jawline_{ext}(x, y) = \max(jawline(x - i, y - j) | S(i, j))$$

trong đó:

(x, y) là tọa độ điểm ảnh trên ảnh biểu diễn đường xương quai hàm (*jawline*)

(i, j) là tọa độ điểm ảnh trên yếu tố cấu trúc S , kích thước của yếu tố cấu trúc S tỷ lệ thuận với khoảng cách đồng tử được ước lượng bằng các điểm dấu mặt thu được từ mô-đun xác định điểm dấu mặt và cơ thể (103);

$jawline_{ext}(x, y)$ là giá trị điểm ảnh sau khi mở rộng tại tọa độ (x, y) ;

$S(i, j)$ là giá trị điểm ảnh của yếu tố cấu trúc tại tọa độ (i, j) ;

$\max()$ là toán tử tìm giá trị lớn nhất trong các điểm ảnh tương quan giữa ảnh biểu diễn đường xương quai hàm và yếu tố cấu trúc;

vùng cằm được định nghĩa là vùng giao giữa đường xương quai hàm mở rộng và vùng cổ phân đoạn được thu lại từ mô-đun phân đoạn(101) và được biểu diễn bởi công thức:

$$chin_{mask} = jawline_{ext}(x, y) \cap neck_{mask}$$

với $chin_{mask}$ và $neck_{mask}$ lần lượt là vùng cằm và vùng phân đoạn cổ; giả thiết rằng, trên $chin_{mask}$, các điểm ảnh mang giá trị 1 thể hiện nằm trong vùng cằm và các điểm ảnh mang giá trị 0 thể hiện nằm ngoài vùng cằm; đường trục chính giữa cổ được thể hiện bằng phương trình:

$$x = \frac{(x_{min} + x_{max})}{2}$$

$$x_{min} = \min(x) \forall (x, y) | chin_{mask}(x, y) = 1$$

$$x_{max} = \max(x) \forall (x, y) | chin_{mask}(x, y) = 1$$

trong đó, mô-đun xác định và phân loại kiểu mặt (105) sử dụng đầu vào từ mô-đun phân đoạn (101) và mô-đun xác định điểm dấu mặt và cơ thể (103) để tính toán các thông số về độ rộng trán ($d_{trán}$), độ rộng gò má ($d_{má}$), độ rộng cằm ($d_{cằm}$), và độ dài gương mặt ($d_{mặt}$); so sánh giá trị các thông số này với nhau, ta kết luận được kiểu mặt của người đang xét; các loại kiểu mặt được xét là mặt trái xoan, mặt chữ nhật dài và mặt tròn và được định nghĩa cụ thể như sau:

$$\text{mặt trái xoan: } (d_{mặt} > d_{má}) \& (d_{trán} > d_{cằm})$$

$$\text{mặt chữ nhật dài: } d_{mặt} > d_{má} \approx d_{trán} \approx d_{cằm}$$

mặt tròn: $d_{má} \approx d_{mặt} > d_{trán} \approx d_{cằm}$

khối thay đổi hình dáng(200)gồm ba mô-đun chính:

mô-đun ước lượng dữ liệu ba chiều (3D) người (201) sử dụng thuật toán tối ưu để xác định mô hình lưới ba chiều (3D) người dựa trên thông tin chiều cao cân nặng với hàm mục tiêu tương tự như mô-đun xác định hệ số mẫu(103), tuy nhiên chỉ sử dụng ba hàm mục tiêu $E_{weight}, E_{height}, E_{\beta}$;

mô-đun sinh dữ liệu bề mặt lưới hai chiều (2D) (202) sử dụng thông tin hệ số mẫu để sinh mô hình lưới ba chiều (3D) và sử dụng phép chiếu phối cảnh chiếu lưới ba chiều (3D) thành các điểm hai chiều (2D) khớp với hình dạng người mẫu trong ảnh; sau đó sử dụng thuật toán tam giác delaunay sinh lưới tam giác cho các điểm hai chiều (2D) tạo được, thực hiện chuẩn hóa về dải $[0, 1]$ để tạo bản đồ UV và gắn ảnh người mẫu thành ảnh kết cấu (texture) cho mô hình lưới hai chiều (2D);

mô-đun trích xuất hình ảnh hai chiều (2D)(203) sử dụng phép chiếu phối cảnh chiếu mô hình lưới ba chiều (3D) người thu được trong mô-đun ước lượng dữ liệu ba chiều (3D) (301) người thành các điểm hai chiều (2D);

khối hoán đổi (300) bao gồm sáu mô-đun chính:

mô-đun phân vùng mặt cổ người sử dụng (301),

mô-đun xác định điểm dấu mặt người sử dụng (302),

mô-đun tái tạo vùng cổ người sử dụng bị che khuất (303),

mô-đun đổi màu da (304),

mô-đun xác định kiểu mặt (305),

mô-đun hoán đổi ảnh (306),

trong đó, mô-đun phân vùng mặt cổ người sử dụng (301) dùng mạng học sâu để chia hình ảnh chân dung thành 18 phân vùng (hình nền, cổ, tóc,...) qua đó loại bỏ các phần không cần thiết; mô-đun xác định điểm dấu mặt người sử dụng (302) sử dụng mạng học sâu trích xuất sáu tám tọa độ hai chiều các điểm dấu trên hình ảnh khuôn mặt người sử dụng; với đầu ra từ mô-đun phân vùng mặt cổ người sử dụng (301), mô-đun tái tạo vùng

cỗ người sử dụng bị che khuất (303) dùng mô hình mạng đối nghịch tạo sinh (GAN) để tự động nhận dạng và phục hồi vùng cỗ người sử dụng bị che khuất bởi quần áo; mô hình mạng đối nghịch tạo sinh (GAN) được định nghĩa bởi mỗi không gian xác suất (Ω, μ_{ref}) và bao gồm hai bộ chính: bộ tạo sinh và bộ phân biệt; bộ tạo sinh sử dụng $P(\Omega)$ là bộ của tất cả xác suất đo μ_G trên Ω ; bộ phân biệt sử dụng nhân Markov $\mu_D : \Omega \rightarrow P[0,1]$, với $P[0,1]$ là bộ xác suất đo trên $[0,1]$; hàm mục tiêu của GAN được biểu diễn bằng công thức:

$$L(\mu_G, \mu_D) := E_{x \sim \mu_{ref}, y \sim \mu_D(x)} [\ln y] + E_{x \sim \mu_G, y \sim \mu_D(x)} [\ln(1 - y)]$$

khối hiệu chỉnh và tối ưu hóa (400) bao gồm các mô-đun chính:

mô-đun tính toán kích thước đầu người sử dụng (401) dùng tập điểm đầu của người sử dụng và mẫu để điều chỉnh tỷ lệ đầu phù hợp; tỷ lệ đầu được tính với các trường hợp khác nhau như sau:

đối với nam:

hình chụp cả người:

$$scale_{head} = scale_{chin}$$

hình chụp nửa người:

$$scale_{head} = (scale_{eye} + scale_{chin} * 3) / 4$$

đối với nữ:

$$scale_{head} = (scale_{eye} + scale_{chin}) / 2$$

trong đó, $scale_{head}$ là tỷ lệ đầu cần điều chỉnh, $scale_{eye}$ và $scale_{chin}$ lần lượt là tỷ lệ mắt và cằm giữa người sử dụng và ảnh mẫu được tính toán dựa vào hai tập điểm đầu tương ứng của người sử dụng và ảnh mẫu;

mô-đun tính toán vị trí đầu người sử dụng (403) điều chỉnh vị trí đầu người sử dụng vào chính giữa cỗ ảnh mẫu theo trục hoành của ảnh;

mô-đun so sánh kiểu mặt người dùng và mẫu (402) so sánh thông tin kiểu mặt giữa người dùng và mẫu để bổ sung thêm thông tin về tỷ lệ đầu và vị trí đầu phù hợp;

mô-đun hiệu chỉnh tính toán vị trí và kích thước đầu người dùng (404) nhận thông tin từ các mô-đun cập nhật kích thước và vị trí đầu người dùng (401, 402, và 403) để tính toán ma trận biến đổi M mới tương tự như mô-đun hoán đổi ảnh(306) trong khối hoán đổi ảnh(300);

mô-đun xử lý màu da vùng tiếp giáp đồng đều (405) sử dụng phương trình poisson kết hợp điều kiện biên dirichlet; giá trị trường độ dốc trong vùng ảnh ghép được tính toán và điều chỉnh cho phù hợp với ảnh người sử dụng, hạn chế tối đa sự chênh lệch màu sắc trong vùng da tiếp giáp giữa người sử dụng và ảnh mẫu; đầu ra của khối hiệu chỉnh và tối ưu hóa (400) là ảnh người sử dụng kết hợp với ảnh quần áo đã được tối ưu và hiệu chỉnh.

2. Phương pháp thử ảo trang phục hai chiều (2D) sử dụng công nghệ máy học kết hợp giữa thuật toán tối ưu và thuật toán phân loại xác định bao gồm năm bước chính:

bước 1: xử lý dữ liệu hình ảnh dữ liệu mẫu; mục đích của bước này là chuẩn hóa dữ liệu đầu vào cho bước 2 và bước 3; các dữ liệu lưới dạng thô được hiệu chuẩn và biến đổi theo các đặc tính yêu cầu để đảm bảo việc kế thừa và sử dụng khi có sự cập nhật dữ liệu; quá trình được thực hiện trên khối tiền xử lý dữ liệu (100); dữ liệu đầu vào trong khối tiền xử lý dữ liệu (100) gồm ba phần như đã đề cập trong khối đầu vào: ảnh người sử dụng, ảnh quần áo, thông tin chiều cao cân nặng người sử dụng; năm mô-đun trong khối tiền xử lý dữ liệu (100) gồm:

mô-đun phân đoạn (101),

mô-đun xác định trục chính giữa cổ (102),

mô-đun xác định điểm dấu mặt và cơ thể (103),

mô-đun xác định hệ số mẫu (104),

mô-đun xác định (105),

đầu ra của khối tiền xử lý dữ liệu được sử dụng bước 2 và bước 4;

bước 2: tối ưu và thay đổi hình dạng cơ thể tương ứng; mục đích của bước này là tính toán mô hình người, sau đó biến đổi hình dạng ảnh hai chiều (2D) quần áo theo kích thước tương ứng, đầu ra của bước này được sử dụng tại bước 3; bước 2 được thực hiện trên khối thay đổi hình dáng; dữ liệu đầu vào của khối thay đổi hình dáng gồm năm phần: thông

tin phân đoạn ảnh mẫu, hệ số trục chính giữa cổ mẫu, thông tin điểm dấu trên mặt và cơ thể mẫu, thông tin chiều cao cân nặng người sử dụng, và thông tin hệ số mẫu; khối thay đổi hình dáng (200) gồm ba mô-đun chính:

mô-đun ước lượng dữ liệu ba chiều (3D) người (301),

mô-đun sinh dữ liệu bề mặt lưới hai chiều (2D) (302),

mô-đun trích xuất hình ảnh hai chiều (2D) (303),

đầu ra của khối cập nhật hình dáng được sử dụng làm đầu vào cho khối hoán đổi;

bước 3: hoán đổi đầu sử dụng và hình dạng cơ thể tương ứng; mục đích của bước này hoán đổi hình ảnh đầu người sử dụng với hình ảnh mẫu quần áo hai chiều (2D) được cập nhật, đầu ra được sử dụng tại bước 4; bước 3 được thực hiện trên khối hoán đổi (300); dữ liệu đầu vào của khối hoán đổi gồm bốn phần: thông tin phân đoạn ảnh mẫu cập nhật, ảnh mẫu cập nhật, ảnh người sử dụng, và thông tin điểm dấu trên mặt và cơ thể mẫu cập nhật; khối hoán đổi (300) gồm sáu mô-đun chính:

mô-đun phân vùng mặt cổ người sử dụng (301),

mô-đun xác định điểm dấu mặt người sử dụng (302),

mô-đun tái tạo vùng cổ người sử dụng bị che khuất (303),

mô-đun đổi màu da (304),

mô-đun xác định kiểu mặt (305),

và mô-đun hoán đổi ảnh (306),

bước 4: hiệu chỉnh và tối ưu hóa giữa đầu và hình dạng cơ thể tương ứng; mục đích của bước này là tinh chỉnh về vị trí, kích thước và màu da của đầu người sử dụng phù hợp với cơ thể mẫu tương ứng; bước 4 được thực hiện trên khối hiệu chỉnh và tối ưu hóa (400); dữ liệu đầu vào của khối hiệu chỉnh và tối ưu hóa gồm sáu phần: thông tin điểm dấu trên mặt và cơ thể mẫu cập nhật, điểm dấu mặt người sử dụng tương ứng với ảnh quần áo, thông tin kiểu mặt mẫu, thông tin kiểu mặt người dùng, hệ số trục chính giữa cổ mẫu cập nhật, và ảnh người sử dụng kết hợp với ảnh quần áo; khối hiệu chỉnh và tối ưu hóa (400) gồm năm mô-đun chính:

mô-đun tính toán kích thước đầu người sử dụng (401),

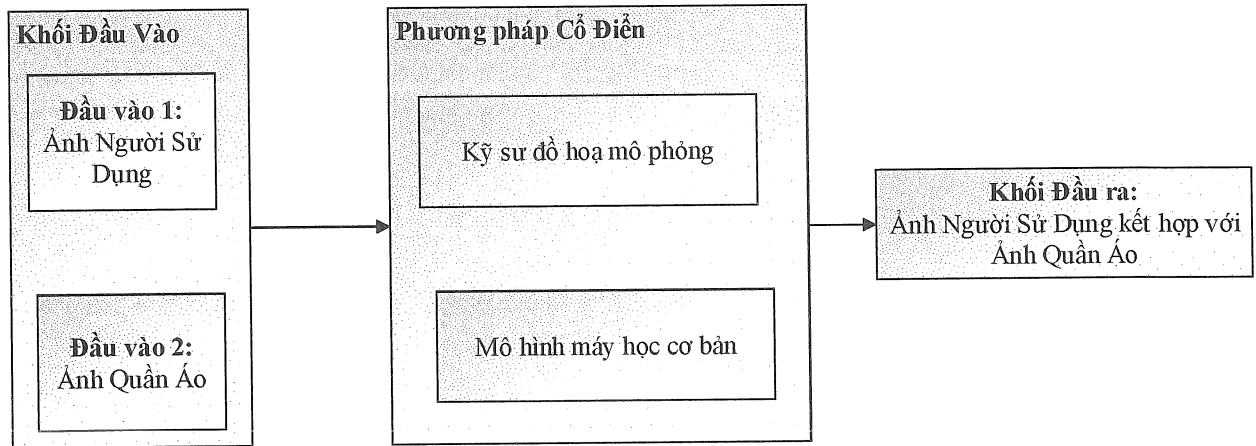
mô-đun so sánh kiểu mặt người dùng và mũ (402),

mô-đun tính toán vị trí đầu người sử dụng, mô-đun hiệu chỉnh tính toán vị trí và kích thước đầu người dùng (404),

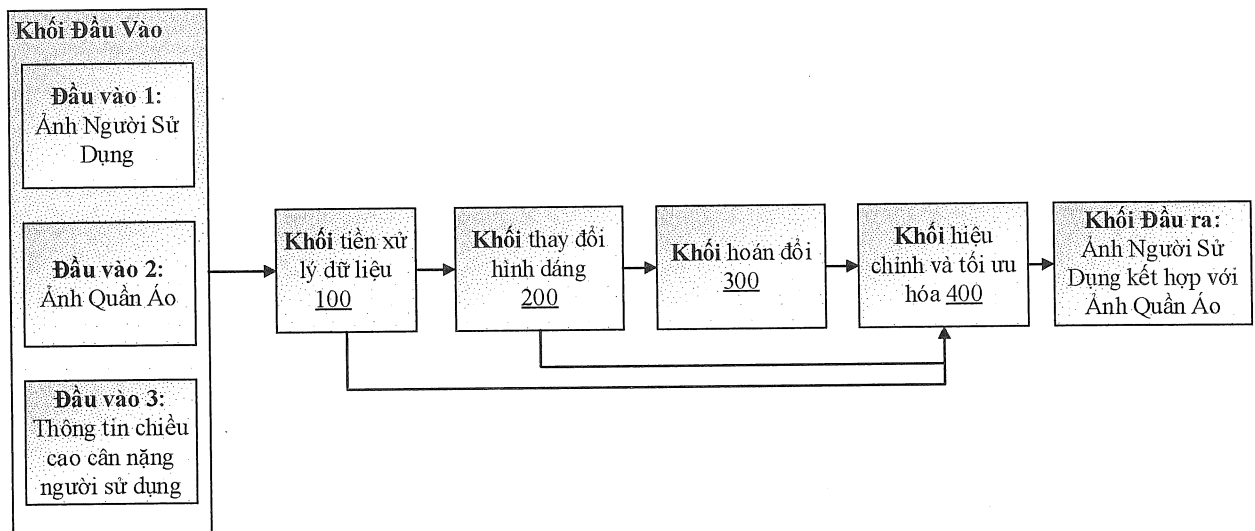
mô-đun xử lý màu da vùng tiếp giáp đồng đều (405),

đầu ra của bước 4 là ảnh người sử dụng kết hợp với ảnh quần áo đã được tối ưu, hiệu chỉnh;

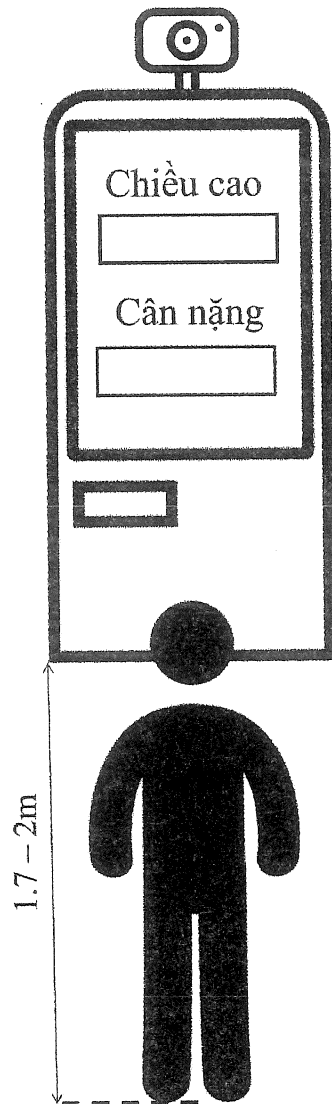
bước 5: hiển thị hình ảnh tương tác hai chiều (2D); mục đích của bước này là việc hiển thị dưới dạng ba chiều (2D) của hình ảnh người sử dụng và hình ảnh quần áo mẫu đã được tính toán theo bước 4.



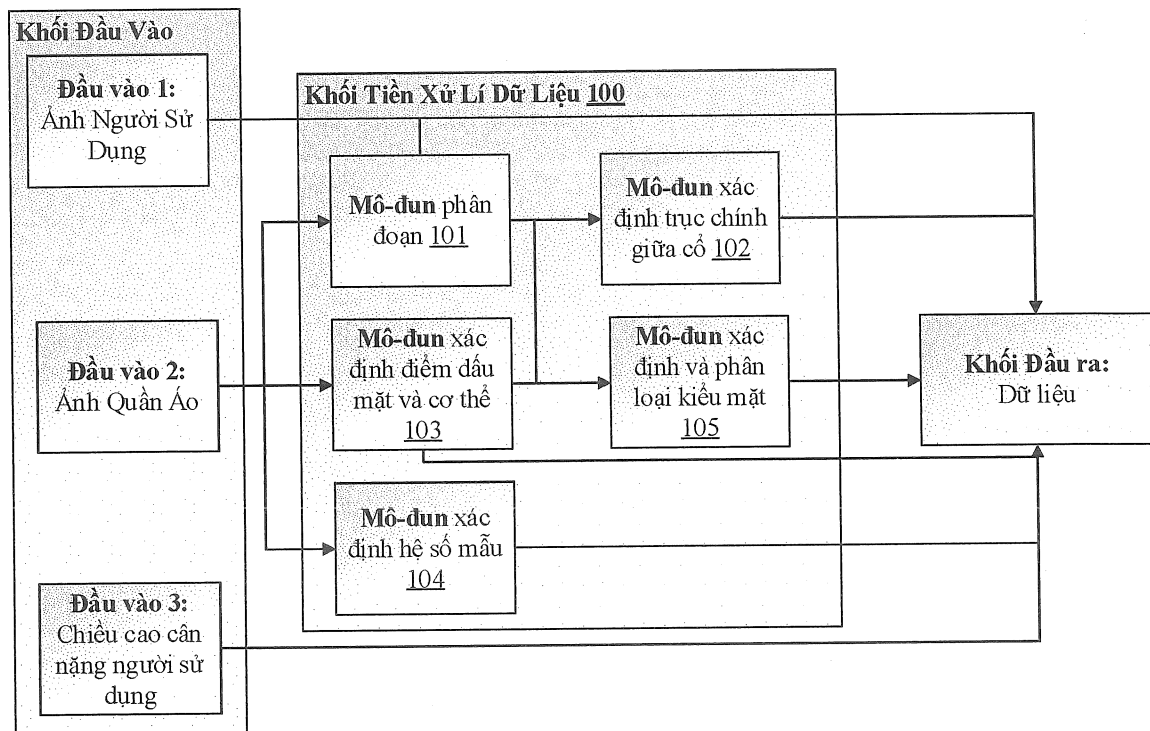
Hình 1



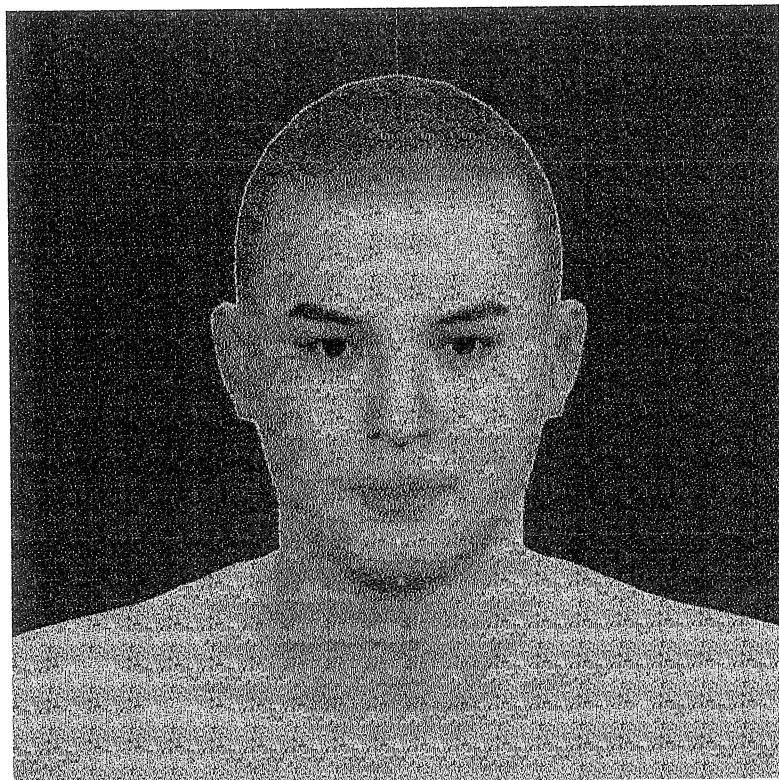
Hình 2



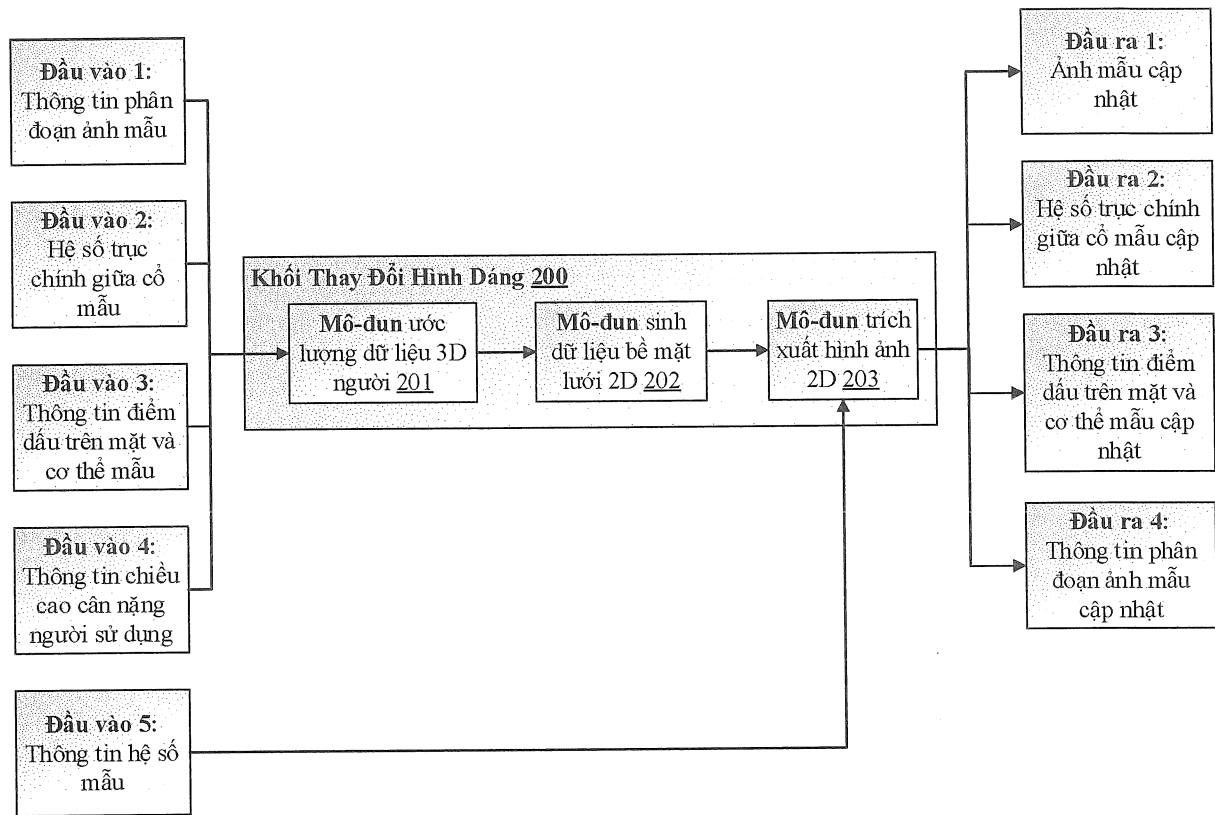
Hình 3



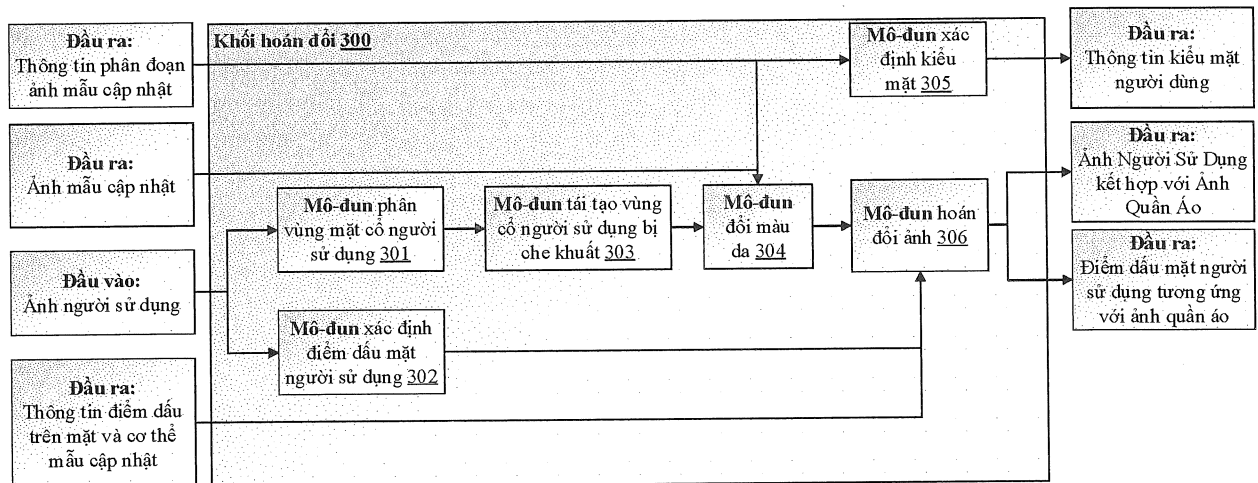
Hình 4



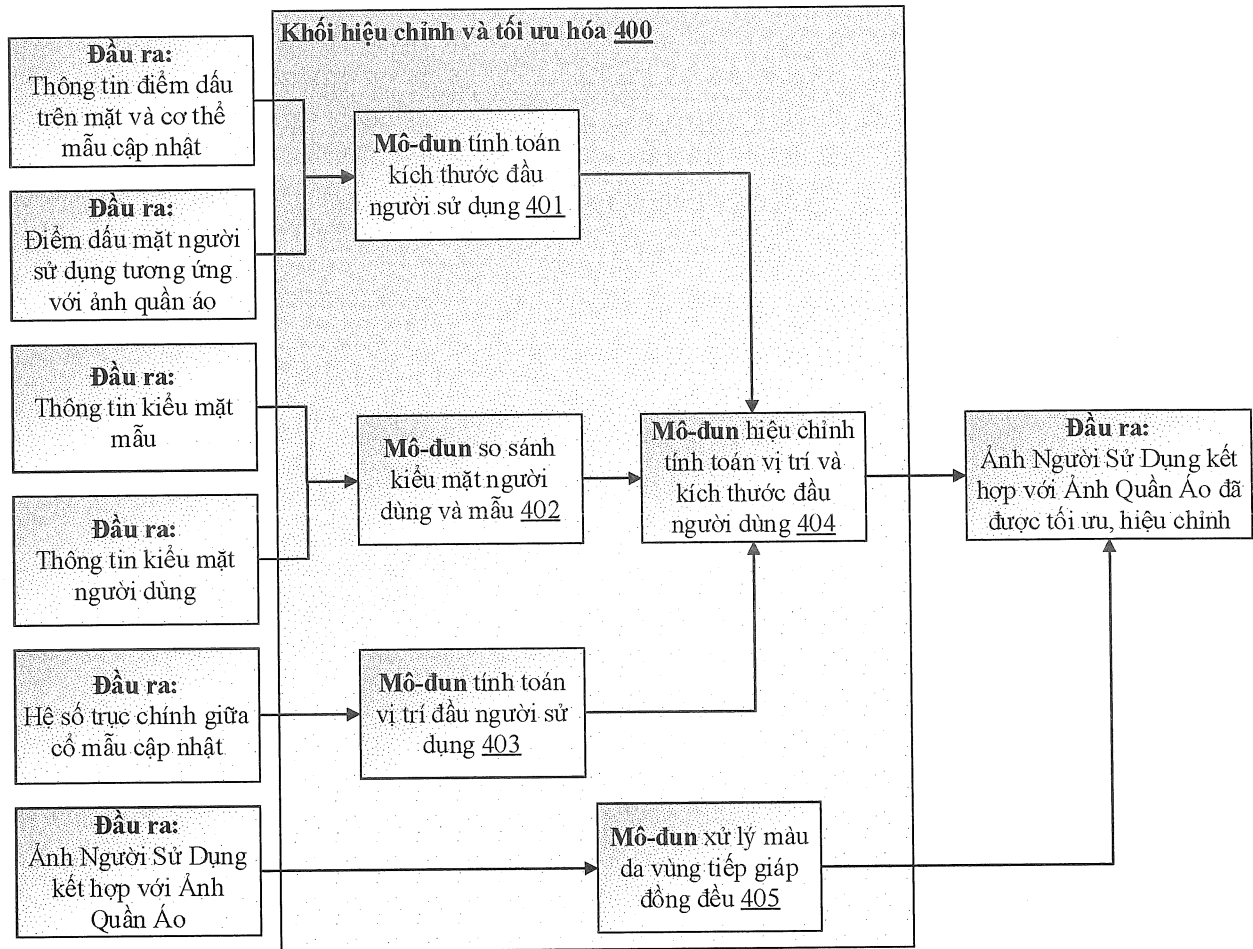
Hình 5



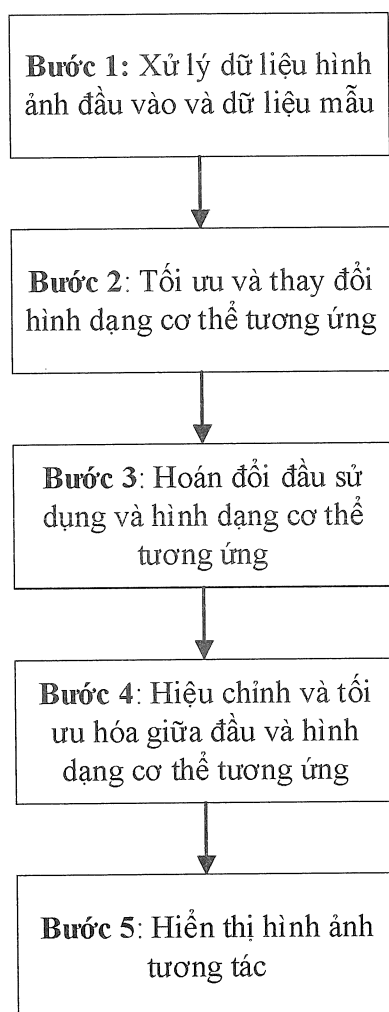
Hình 6



Hình 7



Hình 8



Hình 9