



(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) 
1-0020878

(51)⁷ **G08B 21/06**

(13) **B**

(21) 1-2015-04715

(22) 09.12.2015

(45) 27.05.2019 374

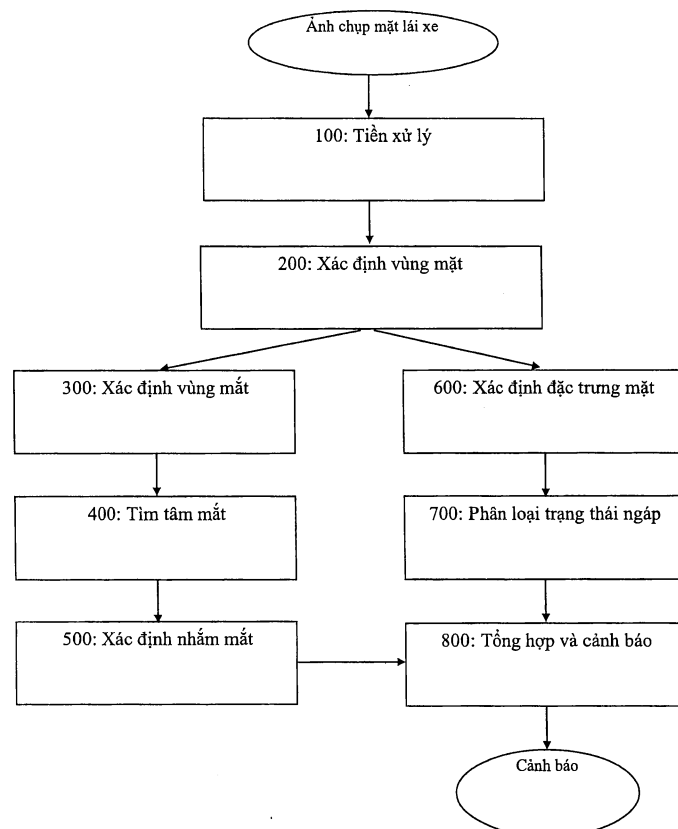
(43) 27.06.2016 339

(73) **VIỆN NGHIÊN CỨU CÔNG NGHỆ FPT - TRƯỜNG ĐẠI HỌC FPT (VN)**
Số 8, đường Tôn Thất Thuyết, phường Mỹ Đình 2, quận Nam Từ Liêm, thành phố Hà Nội

(72) **Từ Minh Phương (VN), Hoàng Anh Tuấn (VN), Trần Nguyên Ngọc (VN), Đặng Minh Đức (VN), Phạm Bảo Thạch (VN), Nguyễn Tuấn Anh (VN)**

(54) **QUY TRÌNH CẢNH BÁO TRẠNG THÁI BUỒN NGỦ CỦA LÁI XE**

(57) Sáng chế đề xuất quy trình nhận ảnh chụp từ camera thông thường về khuôn mặt người lái xe khi đang lái, và đưa ra cảnh báo khi lái xe có dấu hiệu buồn ngủ. Quy trình sau khi chuẩn hoá ảnh chụp đầu vào kích thước chuẩn, có độ xám phù hợp, thì thực hiện xác định vùng mắt và so khớp với mẫu để xác định mắt nhắm hay mở, đồng thời phân loại đặc trưng ảnh mặt lái xe để xác định ở trạng thái ngáp hay không ngáp. Cảnh báo về trạng thái buồn ngủ của lái xe được đưa ra nếu có phát hiện ngáp hoặc nhắm mắt.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến quy trình nhận video hoặc ảnh chụp khuôn mặt người lái xe khi đang lái, và đưa ra cảnh báo khi lái xe có dấu hiệu buồn ngủ.

Tình trạng kỹ thuật của sáng chế

Việc phát hiện và cảnh báo trạng thái buồn ngủ của lái xe, trong lúc đang điều khiển phương tiện, là một hoạt động có ý nghĩa quan trọng trong phòng ngừa tai nạn giao thông xảy ra do sự mất tập trung của lái xe.

Đã có nhiều sáng chế trên thế giới đề xuất các giải pháp cho việc phát hiện lái xe có buồn ngủ khi đang trên đường hay không.

Nhóm sáng chế thứ nhất tập trung vào sử dụng sự tiếp xúc giữa tay người lái với vô lăng để phát hiện sự buồn ngủ của lái xe. Nếu tay của người lái xe không còn nắm chắc vô lăng lâu hơn một khoảng thời gian nhất định, cảnh báo sẽ được đưa ra. Ví dụ Bằng độc quyền sáng chế Mỹ số US 6016103 A đề người lái xe sử dụng găng tay có các cảm biến ghi nhận lực nắm của tay lên vô lăng, nếu lực nhỏ trong thời gian lâu thì đưa ra cảnh báo. Hoặc Bằng độc quyền sáng chế Mỹ số US 6590499 B1 đề xuất sử dụng cảm biến ngay trên vô lăng, cũng ghi nhận lực nắm của tay lên vô lăng, nếu lực nhỏ trong thời gian lâu thì đưa ra cảnh báo. Nhóm sáng chế này đòi hỏi sử dụng găng tay hay vô lăng chuyên dụng, không phải lúc nào cũng tiện lợi trong thời tiết nóng, hoặc dễ dàng thay thế cho vô lăng nguyên bản của xe.

Nhóm sáng chế thứ hai sử dụng camera để theo dõi biểu hiện trên khuôn mặt tài xế. Bằng độc quyền sáng chế Mỹ số US 5689241 A đề xuất sử dụng camera hồng ngoại để ghi nhận biến đổi nhiệt độ trên các vùng trên mặt người lái xe, và đo nhiệt độ và mức độ to nhỏ của thể tích khí nóng thở ra từ mũi của người lái, đồng thời kiểm tra mức độ vận động của người lái. Nếu mức độ vận động giảm cùng với nhiệt độ giảm, thể tích khí hít thở giảm, thì đây là các dấu hiệu của sự buồn ngủ và cảnh báo trạng thái buồn ngủ được đưa ra. Sáng chế

này sử dụng camera hồng ngoại, không phải là camera phổ thông, nên giá thành cao.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất quy trình xác định và cảnh báo trạng thái buồn ngủ của lái xe, sử dụng được video thu từ camera phổ thông, do đó giá thành được hạ thấp.

Cụ thể, quy trình nhận ảnh chụp khuôn mặt người lái xe khi đang lái, và đưa ra cảnh báo khi lái xe có dấu hiệu buồn ngủ, bao gồm các bước:

đưa ảnh chụp đầu vào về kích thước chuẩn, bằng cách với mỗi điểm ảnh trên ảnh kích thước chuẩn:

ánh xạ vị trí của điểm ảnh trên ảnh kích thước chuẩn đang xét, gọi là p , sang vị trí tương ứng trên ảnh chụp đầu vào, gọi là P , qua ánh xạ tuyến tính;

giá trị màu sắc RGB của điểm ảnh trên ảnh kích thước chuẩn đang xét được nội suy từ các giá trị màu sắc RGB ở các điểm ảnh trên ảnh chụp đầu vào, lân cận với vị trí P ;

xám hoá ảnh kích thước chuẩn, bằng cách ánh xạ giá trị RGB ở mỗi điểm ảnh đến một số thực, gọi là độ sáng điểm ảnh, theo một hàm ánh xạ nhất định;

xác định vùng chứa mặt người trên ảnh đã xám hoá ở trên, là vùng có độ cao H và chiều rộng D , bằng cách thực hiện:

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở nửa trái của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trái của cửa sổ;

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường

độ sáng ở nửa dưới của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trên của cửa sổ;

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở hai bên của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở vùng giữa của cửa sổ;

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở nửa trên bên trái và nửa dưới bên phải của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trên bên phải và nửa dưới bên trái của cửa sổ;

đưa các con số tính được ở trên, gọi là các giá trị đặc trưng, vào một mô hình phân loại học máy thống kê, đã được xây dựng sẵn, hay còn gọi là được huấn luyện sẵn, từ các giá trị đặc trưng tính từ vùng hình ảnh đã biết chính xác nằm trong vùng chứa mặt người hay không, gọi là tập hợp mẫu, để thu được kết quả phân loại vùng nào chứa mặt người;

xác định vùng chứa mắt, từ vùng chứa mặt người, nằm trong độ cao $y_1 * H$ đến $y_2 * H$, và mắt trái nằm trong khoảng $x_1 * D$ đến $x_2 * D$, mắt phải trong khoảng $x_3 * D$ đến $x_4 * D$ với x_1, x_2, x_3, x_4 và y_1, y_2 là các con số nằm giữa 0 và 1, chọn trước;

xác định tọa độ tâm của mắt bằng cách thực hiện:

với mỗi một điểm ảnh, gọi là điểm c , ở trong vùng chứa mắt tìm được sau bước trên, tính trung bình cộng của tất cả các giá trị tích vô hướng của véc tơ thứ nhất gọi là d_i và véc tơ thứ hai gọi là g_i , trong đó:

d_i là véc tơ đơn vị có hướng trùng với hướng từ điểm ảnh i đến điểm c ;

g_i là véc tơ gradien của cường độ sáng của tại điểm ảnh i ;

xác định tâm mắt là điểm ảnh c mà ở đó giá trị trung bình cộng tính được như trên là lớn nhất;

lấy tích chập thứ nhất giữa vùng ảnh nằm xung quanh tâm mắt với mẫu mắt nhắm, và lấy tích chập thứ hai giữa vùng ảnh nằm xung quanh tâm mắt với mẫu mắt mở, và so sánh giá trị cực đại của hai kết quả; nếu tích chập thứ nhất có giá trị cực đại lớn hơn thì kết luận mắt nhắm, nếu không thì kết luận mắt mở;

phân tích ảnh đã xám hoá để thu được véc tơ đặc trưng của ảnh, bằng cách thực hiện như sau:

thu thập các ảnh chụp người ngáp; dữ liệu trong mỗi ảnh này được biểu diễn là chuỗi các con số tạo thành một véc tơ; rồi tính véc tơ trung bình của các véc tơ này, gọi là v_1 ;

thu thập các ảnh chụp người không ngáp; dữ liệu trong mỗi ảnh này được biểu diễn là chuỗi các con số tạo thành một véc tơ; rồi tính véc tơ trung bình của các véc tơ này, gọi là v_2 ;

tìm véc tơ trung bình cộng của v_1 và v_2 , gọi là v ;

tính ma trận M_1 , là kết quả nhân ma trận của véc tơ cột $(v_1 - v)$ với véc tơ hàng $(v_1 - v)$; và ma trận M_2 , là kết quả nhân ma trận của véc tơ cột $(v_2 - v)$ với véc tơ hàng $(v_2 - v)$; rồi tính ma trận $M = (M_1 + M_2)/2$;

tính ma trận S bằng nghịch đảo của ma trận hiệp phương sai, của tất cả các véc tơ đã thu thập được, nhân với M ;

tính k véc tơ riêng ứng với k trị riêng lớn nhất của ma trận S , với k là một số chọn trước;

với ảnh đầu vào cần phân tích, dữ liệu trong ảnh này được biểu diễn là chuỗi các con số tạo thành một véc tơ gọi là x , tính tích vô hướng của x với từng véc tơ riêng trong bộ k véc tơ riêng nêu trên; kết quả thu được là một véc tơ đặc trưng k chiều của ảnh đầu vào.

phân loại véc tơ đặc trưng của ảnh, thuộc về loại ngáp hay không ngáp, bằng kỹ thuật học máy thống kê có giám sát, là phương pháp Máy Véc tơ Hỗ trợ, sử dụng một mô hình thống kê Máy Véc tơ Hỗ trợ được xây dựng sẵn, từ tập hợp các vectơ đặc trưng tính từ những ảnh chụp người lái xe đã biết chính xác đang ngáp hay không đang ngáp;

đưa ra cảnh báo về trạng thái buồn ngủ của lái xe nếu có kết luận ngáp hoặc kết luận nhầm mắt trong các bước trên.

Mô tả vắn tắt các hình vẽ

Hình 1 là sơ đồ khối của quy trình được đề xuất để cảnh báo trạng thái buồn ngủ của lái xe;

Hình 2 minh họa bốn loại cửa sổ để tính ra các giá trị đặc trưng của vùng ảnh chứa mặt người, trích từ bài báo Rapid Object Detection using a Boosted Cascade of Simple Features của Paul Viola và Michael Jones in trong kỷ yếu hội nghị Computer Vision and Pattern Recognition năm 2001.

Mô tả chi tiết sáng chế

Hình 1 là sơ đồ khối của quy trình được đề xuất để cảnh báo trạng thái buồn ngủ của lái xe, từ đầu vào là ảnh chụp khuôn mặt người lái xe, hoặc khung hình từ video quay lái xe khi đang hoạt động trên đường. Quy trình này gồm các bước:

bước 100, ảnh chụp đầu vào được đưa qua tiền xử lý, để đưa về kích thước chuẩn, và xám hoá với cường độ sáng chuẩn;

bước 200, xác định vùng chứa mặt người trên ảnh đã chuẩn hoá theo phương pháp như mô tả trong bài báo Rapid Object Detection using

a Boosted Cascade of Simple Features của Paul Viola và Michael Jones in trong kỷ yếu hội nghị Computer Vision and Pattern Recognition năm 2001; đầu ra của bước 200, là phần ảnh chụp đã chuẩn hoá chứa mặt người, được cho thành đầu vào của bước 300 và bước 600 dưới đây; chi tiết của bước 200 được giải thích rõ hơn trong phần mô tả bên dưới đây;

bước 300, vùng chứa mắt được xác định, từ đầu ra của bước 200, nhờ vào các tỷ lệ định trước của tọa độ vùng chứa mắt so với tọa độ vùng chứa mặt;

bước 400, tâm của mắt được xác định, theo phương pháp như mô tả trong bài báo Accurate eye centre localisation by means of gradients bởi tác giả Fabian Timm và Erhardt Barth in trong kỷ yếu hội nghị International Conference on Computer Vision Theory and Applications năm 2011; chi tiết của bước 400 được giải thích rõ hơn trong phần mô tả bên dưới đây;

bước 500, vùng xung quanh tâm mắt được so khớp với một vài mẫu ảnh mắt nhắm và mắt mở, xem vùng này giống với mẫu ảnh nào nhất thì đưa ra thông tin mắt đang nhắm hay đang mở;

bước 600, ảnh đã chuẩn hoá được phân tích theo phương pháp Linear Discriminant Analysis, viết tắt là LDA, như được mô tả trong bài báo The Use of Multiple Measurements in Taxonomic Problems của tác giả Ronald Fisher in trong tạp chí Annals of Eugenics năm 1936, để thu được véc tơ đặc trưng của ảnh; chi tiết của bước 600 được giải thích rõ hơn trong phần mô tả bên dưới đây;

bước 700, là bước phân loại bằng kỹ thuật học máy thống kê có giám sát, nhận đầu vào là véc tơ đặc trưng thu được ở bước 600 nêu trên, phân loại nó thuộc về loại ngáp hay không ngáp. Kỹ thuật phân loại bằng học máy thống kê có giám sát có thể sử dụng là phương pháp Máy Véc tơ Hỗ trợ - Support Vector Machine như trình bày ở bài báo năm 1992 "A

training algorithm for optimal margin classifiers” của Boser, Guyon và Vapnik trong kỷ yếu hội nghị COLT '92 trang 144. Phương pháp này sử dụng một mô hình thống kê Máy Véc tơ Hỗ trợ được xây dựng sẵn, hay còn gọi là được huấn luyện sẵn, từ tập hợp các véc tơ đặc trưng tính từ những ảnh chụp người lái xe đã biết chính xác đang ngập hay không đang ngập, gọi là tập hợp mẫu;

bước 800, từ kết quả xác định mất nhắm hay mở ở bước 500, và ngập hay không ngập ở bước 700, đưa ra cảnh báo về trạng thái buồn ngủ của lái xe; nếu không ngập và không nhắm mắt thì không cảnh báo, nếu có ngập hoặc nhắm mắt thì cảnh báo.

Ở bước 100, có hai tác vụ chính là chuẩn hoá kích thước ảnh và xám hoá.

Kích thước chuẩn có thể chọn là 256 nhân 256 điểm ảnh. Những ảnh chụp đầu vào có kích thước khác với kích thước chuẩn sẽ được đưa về kích thước chuẩn. Với mỗi điểm ảnh trên ảnh kích thước chuẩn:

ánh xạ vị trí của điểm ảnh trên ảnh kích thước chuẩn đang xét, gọi là p , sang vị trí tương ứng trên ảnh gốc, gọi là P , qua ánh xạ tuyến tính;

giá trị màu sắc RGB của điểm ảnh trên ảnh kích thước chuẩn đang xét được nội suy từ các giá trị màu sắc RGB ở các điểm ảnh trên ảnh gốc, lân cận với vị trí P đã được ánh xạ trên ảnh gốc.

Việc xám hoá được thực hiện bằng cách ánh xạ giá trị RGB ở mỗi điểm ảnh, trên ảnh đã có kích thước chuẩn, đến một số thực, gọi là độ sáng điểm ảnh, theo một hàm ánh xạ nhất định.

Ở bước 200, như đã mô tả trong bài báo Rapid Object Detection using a Boosted Cascade of Simple Features của Paul Viola và Michael Jones in trong kỷ yếu hội nghị Computer Vision and Pattern Recognition năm 2001, thực hiện:

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở nửa trái của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trái của cửa sổ;

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở nửa dưới của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trên của cửa sổ;

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở hai bên của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở vùng giữa của cửa sổ;

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở nửa trên bên trái và nửa dưới bên phải của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trên bên phải và nửa dưới bên trái của cửa sổ;

các con số tính được ở trên, gọi là các giá trị đặc trưng, được đưa vào mô hình phân loại học máy thống kê, đã được xây dựng sẵn, hay còn gọi là được huấn luyện sẵn, từ các giá trị đặc trưng tính từ vùng hình ảnh đã biết chính xác nằm trong vùng chứa mặt người hay không, gọi là tập hợp mẫu, để thu được kết quả phân loại vùng nào chứa mặt người; mô hình phân loại học máy thống kê được lựa chọn là Adaboost được mô tả trong nhiều sách giáo khoa về học máy thống kê và quen thuộc đối với người có trình độ trung bình trong lĩnh vực học máy thống kê.

Hình 2 minh họa bốn loại cửa sổ để tính ra các giá trị đặc trưng ở bước 200. Hình này được trích từ bài báo Rapid Object Detection using a Boosted Cascade of Simple Features của Paul Viola và Michael Jones in trong kỷ yếu hội nghị Computer Vision and Pattern Recognition năm 2001. Loại A trên Hình 2 để tính các giá trị đặc trưng bằng tổng giá trị cường độ sáng ở nửa trái của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trái của cửa sổ. Loại B trên

Hình 2 để tính các giá trị đặc trưng bằng tổng giá trị cường độ sáng ở nửa dưới của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trên của cửa sổ. Loại C trên Hình 2 để tính các giá trị đặc trưng bằng tổng giá trị cường độ sáng ở hai bên của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở vùng giữa của cửa sổ. Loại D trên Hình 2 để tính các giá trị đặc trưng bằng tổng giá trị cường độ sáng ở nửa trên bên trái và nửa dưới bên phải của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trên bên phải và nửa dưới bên trái của cửa sổ.

Ở bước 300, nếu coi vùng chứa mặt người có chiều cao H , chiều rộng D , vùng chứa mắt có thể đặt nằm trong độ cao $y_1 \cdot H$ đến $y_2 \cdot H$, và mắt trái nằm trong khoảng $x_1 \cdot D$ đến $x_2 \cdot D$, mắt phải trong khoảng $x_3 \cdot D$ đến $x_4 \cdot D$ với x_1, x_2, x_3, x_4 và y_1, y_2 là các tỷ lệ, tức là các con số nằm giữa 0 và 1, chọn trước, được rút ra từ quan sát nhiều mẫu khuôn mặt người.

Ở bước 400, với mỗi một điểm ảnh, gọi là điểm c , ở trong vùng chứa mắt, là vùng tìm được sau bước 300, tính trung bình cộng của tất cả các giá trị tích vô hướng của véc tơ thứ nhất gọi là d_i và véc tơ thứ hai gọi là g_i , trong đó:

d_i là véc tơ đơn vị có hướng trùng với hướng từ điểm ảnh i bất kỳ đến điểm c ;

g_i là véc tơ gradient của cường độ sáng của tại điểm ảnh i .

Tâm mắt chính là điểm ảnh c mà ở đó giá trị trung bình cộng tính được như trên là lớn nhất.

Ở bước 500, việc so khớp vùng mắt với mẫu cho trước được thực hiện bằng cách lấy tích chập thứ nhất giữa vùng mắt với mẫu mắt nhắm, và lấy tích chập thứ hai giữa vùng mắt với mẫu mắt mở, và so sánh giá trị cực đại của hai kết quả. Nếu tích chập thứ nhất có giá trị cực đại lớn hơn thì kết luận mắt nhắm, ngược lại thì kết luận mắt mở.

Ở bước 600, việc phân tích ảnh theo phương pháp Linear Discriminant Analysis, được thực hiện như sau:

thu thập các ảnh chụp người ngáp; dữ liệu trong mỗi ảnh này được biểu diễn là chuỗi các con số tạo thành một véc tơ; rồi tính véc tơ trung bình của các véc tơ này, gọi là v_1 ;

thu thập các ảnh chụp người không ngáp; dữ liệu trong mỗi ảnh này được biểu diễn là chuỗi các con số tạo thành một véc tơ; rồi tính véc tơ trung bình của các véc tơ này, gọi là v_2 ;

tìm véc tơ trung bình cộng của v_1 và v_2 , gọi là v ;

tính ma trận M_1 , là kết quả nhân ma trận của véc tơ cột (v_1-v) với véc tơ hàng (v_1-v) ; và ma trận M_2 , là kết quả nhân ma trận của véc tơ cột (v_2-v) với véc tơ hàng (v_2-v) ; rồi tính ma trận $M = (M_1 + M_2)/2$;

tính ma trận S bằng nghịch đảo của ma trận hiệp phương sai, của tất cả các véc tơ đã thu thập được, nhân với M ;

tính k véc tơ riêng ứng với k trị riêng lớn nhất của ma trận S ;

với ảnh đầu vào cần phân tích, dữ liệu trong ảnh này được biểu diễn là chuỗi các con số tạo thành một véc tơ gọi là x , tính tích vô hướng của x với từng véc tơ riêng trong bộ k véc tơ riêng nêu trên; kết quả thu được là một véc tơ đặc trưng k chiều của ảnh đầu vào.

Số chiều của véc tơ đặc trưng, k , thu được sau khi phân tích ảnh theo phương pháp Linear Discriminant Analysis có thể chọn cố định trước, ví dụ bằng 320 chiều.

Số chiều của véc tơ đặc trưng đã chọn như vậy, cùng toàn bộ việc phân tích ảnh theo phương pháp Linear Discriminant Analysis, cũng được sử dụng trong việc tính các véc tơ đặc trưng tính từ những ảnh chụp người lái xe đã biết chính xác đang ngáp hay không đang ngáp, để huấn luyện ra mô hình thống kê Máy Véc tơ Hỗ trợ sử dụng trong bước 700.

Trong một phương án thực thi mở rộng của quy trình, có thể nhận đầu vào là nhiều khung hình của video quay mặt người lái xe. Khi đó bước 800 có

thể nhận kết quả của hai bước 500 và bước 700 cho nhiều khung hình liên tiếp, và có thể đưa ra cảnh báo khi số lượng khung hình liên tiếp có kết luận đang nhầm mắt hoặc gấp lớn hơn ngưỡng nhất định.

Yêu cầu bảo hộ

Quy trình cảnh báo trạng thái buồn ngủ của lái xe bao gồm các bước:

đưa ảnh chụp đầu vào về kích thước chuẩn, bằng cách với mỗi điểm ảnh trên ảnh kích thước chuẩn:

ánh xạ vị trí của điểm ảnh trên ảnh kích thước chuẩn đang xét, gọi là p , sang vị trí tương ứng trên ảnh chụp đầu vào, gọi là P , qua ánh xạ tuyến tính;

giá trị màu sắc RGB của điểm ảnh trên ảnh kích thước chuẩn đang xét được nội suy từ các giá trị màu sắc RGB ở các điểm ảnh trên ảnh chụp đầu vào, lân cận với vị trí P ;

xám hoá ảnh kích thước chuẩn, bằng cách ánh xạ giá trị RGB ở mỗi điểm ảnh đến một số thực, gọi là độ sáng điểm ảnh, theo một hàm ánh xạ nhất định;

xác định vùng chứa mặt người trên ảnh đã xám hoá ở trên, là vùng có độ cao H và chiều rộng D , bằng cách thực hiện:

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở nửa trái của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trái của cửa sổ;

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở nửa dưới của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trên của cửa sổ;

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở hai bên của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở vùng giữa của cửa sổ;

trên ảnh đầu vào, cho các cửa sổ hình chữ nhật di chuyển lần lượt trên từng điểm ảnh, và tính tổng giá trị cường độ sáng ở nửa trên bên trái và nửa dưới bên phải của cửa sổ rồi trừ đi tổng giá trị cường độ sáng ở nửa trên bên phải và nửa dưới bên trái của cửa sổ;

đưa các con số tính được ở trên, gọi là các giá trị đặc trưng, vào một mô hình phân loại học máy thống kê, đã được xây dựng sẵn, hay còn gọi là được huấn luyện sẵn, từ các giá trị đặc trưng tính từ vùng hình ảnh đã biết chính xác nằm trong vùng chứa mặt người hay không, gọi là tập hợp mẫu, để thu được kết quả phân loại vùng nào chứa mặt người;

xác định vùng chứa mắt, từ vùng chứa mặt người, nằm trong độ cao $y_1 * H$ đến $y_2 * H$, và mắt trái nằm trong khoảng $x_1 * D$ đến $x_2 * D$, mắt phải trong khoảng $x_3 * D$ đến $x_4 * D$ với x_1, x_2, x_3, x_4 và y_1, y_2 là các con số nằm giữa 0 và 1, chọn trước;

xác định tọa độ tâm của mắt bằng cách thực hiện:

với mỗi một điểm ảnh, gọi là điểm c , ở trong vùng chứa mắt tìm được sau bước trên, tính trung bình cộng của tất cả các giá trị tích vô hướng của véc tơ thứ nhất gọi là d_i và véc tơ thứ hai gọi là g_i , trong đó:

d_i là véc tơ đơn vị có hướng trùng với hướng từ điểm ảnh i đến điểm c ;

g_i là véc tơ gradient của cường độ sáng của tại điểm ảnh i ;

xác định tâm mắt là điểm ảnh c mà ở đó giá trị trung bình cộng tính được như trên là lớn nhất;

lấy tích chập thứ nhất giữa vùng ảnh nằm xung quanh tâm mắt với mẫu mắt nhắm, và lấy tích chập thứ hai giữa vùng ảnh nằm xung quanh tâm mắt với mẫu mắt mở, và so sánh giá trị cực đại của hai kết quả; nếu tích chập thứ nhất có giá trị cực đại lớn hơn thì kết luận mắt nhắm, nếu không thì kết luận mắt mở;

phân tích ảnh đã xám hoá để thu được véc tơ đặc trưng của ảnh, bằng cách thực hiện như sau:

thu thập các ảnh chụp người ngáp; dữ liệu trong mỗi ảnh này được biểu diễn là chuỗi các con số tạo thành một véc tơ; rồi tính véc tơ trung bình của các véc tơ này, gọi là v_1 ;

thu thập các ảnh chụp người không ngáp; dữ liệu trong mỗi ảnh này được biểu diễn là chuỗi các con số tạo thành một véc tơ; rồi tính véc tơ trung bình của các véc tơ này, gọi là v_2 ;

tìm véc tơ trung bình cộng của v_1 và v_2 , gọi là v ;

tính ma trận M_1 , là kết quả nhân ma trận của véc tơ cột $(v_1 - v)$ với véc tơ hàng $(v_1 - v)$; và ma trận M_2 , là kết quả nhân ma trận của véc tơ cột $(v_2 - v)$ với véc tơ hàng $(v_2 - v)$; rồi tính ma trận $M = (M_1 + M_2)/2$;

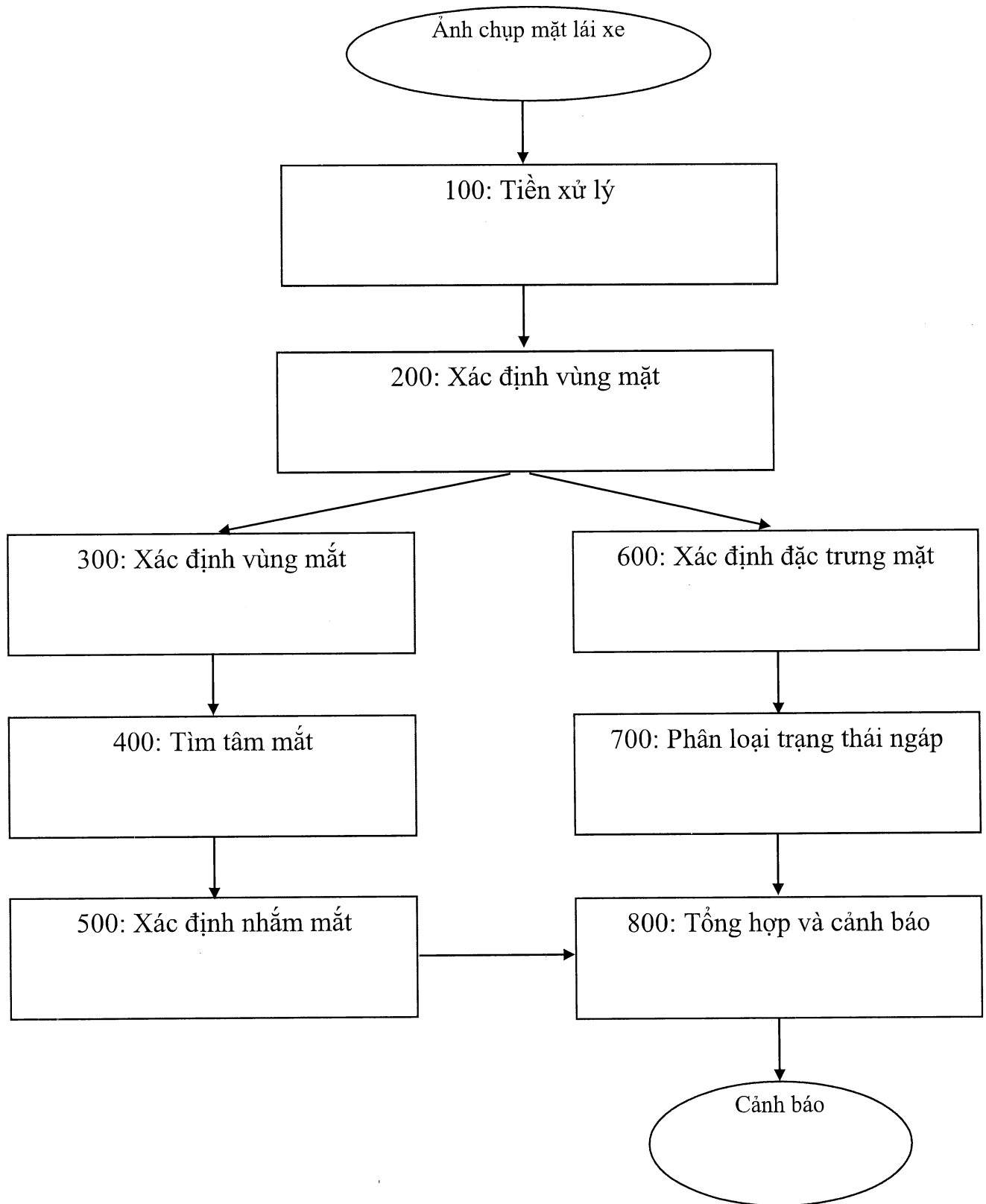
tính ma trận S bằng nghịch đảo của ma trận hiệp phương sai, của tất cả các véc tơ đã thu thập được, nhân với M ;

tính k véc tơ riêng ứng với k trị riêng lớn nhất của ma trận S , với k là một số chọn trước;

với ảnh đầu vào cần phân tích, dữ liệu trong ảnh này được biểu diễn là chuỗi các con số tạo thành một véc tơ gọi là x , tính tích vô hướng của x với từng véc tơ riêng trong bộ k véc tơ riêng nêu trên; kết quả thu được là một véc tơ đặc trưng k chiều của ảnh đầu vào.

phân loại véc tơ đặc trưng của ảnh, thuộc về loại ngáp hay không ngáp, bằng kỹ thuật học máy thống kê có giám sát, là phương pháp Máy Véc tơ Hỗ trợ, sử dụng một mô hình thống kê Máy Véc tơ Hỗ trợ được xây dựng sẵn, từ tập hợp các véc tơ đặc trưng tính từ những ảnh chụp người lái xe đã biết chính xác đang ngáp hay không đang ngáp;

đưa ra cảnh báo về trạng thái buồn ngủ của lái xe nếu có kết luận ngáp hoặc kết luận nhắm mắt trong các bước trên.

Hình 1

Hình 2

