



(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)



1-0047661

(51)^{2020.01} G06K 9/00

(13) B

(21) 1-2020-05043

(22) 01/09/2020

(45) 25/06/2025 447

(43) 25/03/2022 408A

(73) CÔNG TY CỔ PHẦN NGHIÊN CỨU VÀ ỨNG DỤNG TRÍ TUỆ NHÂN TẠO
VINAI (VN)

Tòa nhà văn phòng Symphony, đường Chu Huy Mân, KĐT sinh thái Vinhomes
Riverside, phường Phúc Lợi, quận Long Biên, thành phố Hà Nội, Việt Nam

(72) Hung Hai Bui (US); Nguyễn Minh Hoài (VN); Trần Thế Phong (VN); Trần Tuấn
Anh (VN); Nguyễn Thị Phương Thảo (VN).

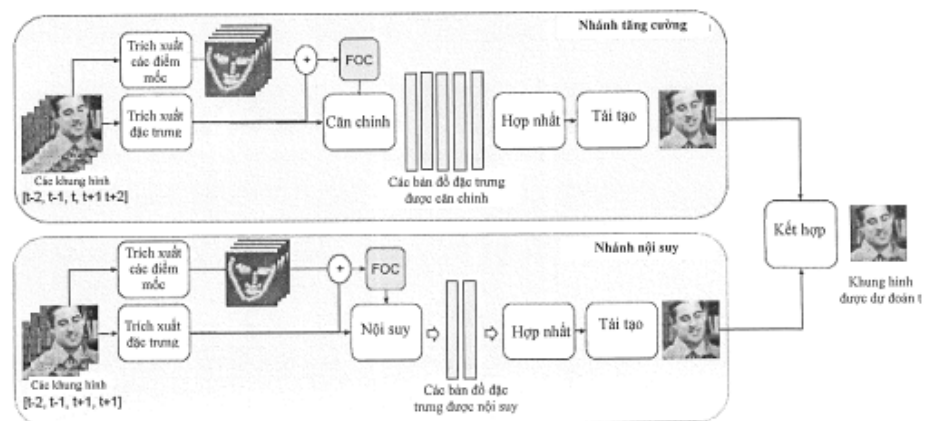
(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)

(54) MÔĐUN VÀ PHƯƠNG PHÁP TÍNH ĐỘ DỊCH NHẬN DIỆN KHUÔN MẶT
DÙNG CHO NỘI SUY VÀ TĂNG CƯỜNG KHUNG HÌNH KHUÔN MẶT VÀ HỆ
THỐNG VÀ PHƯƠNG PHÁP KHỬ NHÒE VIDEO KHUÔN MẶT SỬ DỤNG
MÔĐUN VÀ PHƯƠNG PHÁP TÍNH ĐỘ DỊCH NHẬN DIỆN KHUÔN MẶT NÀY

(21) 1-2020-05043

(57) Sáng chế đề cập đến mô đun và phương pháp tính độ dịch nhận diện khuôn mặt (FOC) dùng cho nội suy và tăng cường khung hình khuôn mặt và hệ thống và phương pháp khử nhiễu video khuôn mặt sử dụng mô đun và phương pháp tính độ dịch nhận diện khuôn mặt này. Hệ thống bao gồm: thiết bị tăng cường khung hình khuôn mặt, bao gồm mô đun FOC, để tăng cường khung hình đích; thiết bị nội suy khung hình khuôn mặt, bao gồm mô đun FOC, để nội suy khung hình đích; và thiết bị kết hợp để kết hợp khung hình đích được tăng cường với khung hình đích được nội suy.

【Fig.3】



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến môđun và phương pháp tính độ dịch nhận diện khuôn mặt (Face-aware Offset Calculation, FOC) dùng cho nội suy và tăng cường khung hình khuôn mặt và hệ thống và phương pháp khử nhòe video khuôn mặt sử dụng môđun và phương pháp tính độ dịch nhận diện khuôn mặt này.

Tình trạng kỹ thuật của sáng chế

Nói chung, nhiều video khuôn mặt gồm các khung hình nhòe do sử dụng các máy ảnh cầm tay nghiệp dư hoặc do chuyển động nhanh của các khuôn mặt có độ động cao.

Các khuôn mặt nhòe gây khó chịu khi nhìn và dẫn đến lỗi trong nhiều thuật toán phân tích khuôn mặt. Do đó, dẫn đến nhu cầu cao trong việc khử nhòe khuôn mặt trong nhiều lĩnh vực ứng dụng.

Cụ thể, khử nhòe là một chủ đề nghiên cứu quan trọng và đang được nghiên cứu rộng rãi trong lĩnh vực xử lý tín hiệu và thị giác máy tính. Theo một phương pháp công nghệ cao gần đây dựa trên công nghệ học sâu (deep learning), mạng nơ-ron tích chập (convolutional neural network, CNN) được huấn luyện sử dụng độ dư thừa theo thời gian giữa các khung hình lân cận để tăng cường khung hình nhòe. Trong quy trình xử lý của phương pháp này, bước quan trọng là căn chỉnh, trong đó các khung hình lân cận có thể được làm biến dạng để căn chỉnh với khung hình đích. Tuy nhiên, khi khung hình đích rất nhòe, việc căn chỉnh có thể gặp khó khăn, và việc căn chỉnh không chính xác có thể dẫn đến các kết quả khử nhòe kém.

Bản chất kỹ thuật của sáng chế

Theo khía cạnh thứ nhất, sáng chế đề xuất phương pháp tính độ dịch nhận diện khuôn mặt (FOC), phương pháp này bao gồm bước nhận các đặc trưng và các bản đồ nhiệt điểm mốc của tập hợp các khung hình; và tạo ra bản đồ độ dịch dựa vào các đặc trưng và các bản đồ nhiệt điểm mốc nhận được.

Theo khía cạnh thứ hai, sáng chế đề xuất phương pháp tăng cường khung hình khuôn mặt, phương pháp này bao gồm bước trích xuất các bản đồ đặc trưng từ các khung hình đầu vào chứa khung hình đích; căn chỉnh mỗi trong số các bản đồ đặc trưng được trích xuất với bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo ra bằng cách sử dụng phương pháp FOC theo khía cạnh thứ nhất; hợp nhất các bản đồ đặc trưng được căn chỉnh, bao gồm bản đồ đặc trưng của khung hình đích, để ước lượng bản đồ đặc trưng của khung hình đích được tăng cường; và tái tạo khung hình đích được tăng cường bằng cách sử dụng bản đồ đặc trưng của nó.

Theo một phương án, các khung hình đầu vào bao gồm năm khung hình liên tiếp.

Theo một phương án, bước căn chỉnh bao gồm thu được các độ dịch căn chỉnh từ mỗi trong số các khung hình lân cận đến khung hình đích bằng phương pháp FOC; và căn chỉnh, bởi lớp tích chập biến dạng được, mỗi trong số các bản đồ đặc trưng được trích xuất của các khung hình lân cận với bản đồ đặc trưng của khung hình đích sử dụng bản đồ độ dịch được tạo.

Theo một phương án, bước căn chỉnh là căn chỉnh hình tháp, phân tầng và biến dạng được (pyramid, cascading, and deformable - PCD) sử dụng FOC để tính độ dịch.

Theo một phương án, bước căn chỉnh đơn cấp trong căn chỉnh PCD bao gồm sử dụng, ở mỗi cấp không phải thấp nhất, các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của các ảnh được lấy mẫu xuống để tính các nhân tích chập biến dạng được; tinh chỉnh, ở cấp thấp nhất, các nhân tích chập biến dạng được bằng cách hợp nhất chúng với giá trị ước lượng độ dịch được lấy mẫu lên từ cấp trước đó; và tính tích chập các nhân biến dạng được với bản đồ đặc trưng của khung hình lân cận.

Theo một phương án, số lượng cấp là ba.

Theo một phương án, đối với mỗi cấp không phải cao nhất, phương pháp FOC đơn giản hóa sử dụng các đặc trưng của các khung hình làm đầu vào được sử dụng.

Theo khía cạnh thứ ba, sáng chế đề xuất phương pháp nội suy khung hình khuôn mặt, phương pháp này bao gồm bước trích xuất các bản đồ đặc trưng từ các khung hình đầu vào trong đó khung hình đích bị thiếu; nội suy, từ mỗi tập hợp các bản đồ đặc trưng được trích xuất, bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo

ra bằng cách sử dụng phương pháp FOC theo khía cạnh thứ nhất; hợp nhất các bản đồ đặc trưng được nội suy để ước lượng bản đồ đặc trưng được kết hợp của khung hình đích được nội suy; và tái tạo khung hình đích được nội suy bằng cách sử dụng bản đồ đặc trưng được kết hợp của khung hình đích này.

Theo một phương án, các khung hình đầu vào bao gồm bốn khung hình.

Theo một phương án, bước nội suy bao gồm sử dụng mỗi tập hợp gồm ba khung hình lân cận từ các khung hình đầu vào cho phương pháp FOC; và biến đổi, bởi lớp tích chập biến dạng được, bản đồ đặc trưng của khung hình giữa trong mỗi tập hợp thành bản đồ đặc trưng của khung hình đích sử dụng bản đồ độ dịch được tạo.

Theo một phương án, bước nội suy bao gồm nội suy xuôi và nội suy ngược.

Theo một phương án, bước nội suy là nội suy hình tháp, phân tầng và biến dạng được (PCD) sử dụng FOC để tính độ dịch.

Theo một phương án, bước nội suy đơn cấp trong nội suy PCD bao gồm sử dụng, ở mỗi cấp không phải thấp nhất, các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của các hình ảnh được lấy mẫu xuống để tính các nhân tích chập biến dạng được; tinh chỉnh, ở cấp thấp nhất, các nhân tích chập biến dạng được bằng cách hợp nhất chúng với giá trị ước lượng độ dịch được lấy mẫu lên từ cấp trước đó; và tính tích chập các nhân biến dạng được với bản đồ đặc trưng của các khung hình lân cận.

Theo một phương án, số lượng cấp là ba.

Theo một phương án, đối với cấp không phải cao nhất, phương pháp FOC đơn giản hóa sử dụng các đặc trưng của các khung hình làm đầu vào được sử dụng.

Theo khía cạnh thứ tư, sáng chế đề xuất phương pháp khử nhòe video khuôn mặt, phương pháp này bao gồm bước tăng cường khung hình đích sử dụng phương pháp theo khía cạnh thứ hai; nội suy khung hình đích sử dụng phương pháp theo khía cạnh thứ ba; và kết hợp khung hình đích được tăng cường với khung hình đích được nội suy.

Theo một phương án, phương pháp khử nhòe video khuôn mặt còn bao gồm bước tiền xử lý để khử nhiễu và khử nhòe khung hình đích không quá nhòe trước.

Theo khía cạnh thứ năm, sáng chế đề xuất mô đun tính độ dịch nhận diện khuôn mặt (FOC), mô đun FOC được tạo cấu hình để nhận các đặc trưng và các bản đồ nhiệt điểm mốc của tập hợp các khung hình; và tạo ra bản đồ độ dịch dựa vào các đặc trưng

và các bản đồ nhiệt điểm mốc nhận được.

Theo khía cạnh thứ sáu, sáng chế đề xuất thiết bị tăng cường khung hình khuôn mặt, thiết bị này bao gồm môđun trích xuất để trích xuất các bản đồ đặc trưng từ các khung hình đầu vào chứa khung hình đích; môđun căn chỉnh, bao gồm môđun FOC theo khía cạnh thứ năm, để căn chỉnh mỗi trong số các bản đồ đặc trưng được trích xuất với bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo ra bởi môđun FOC; môđun hợp nhất để hợp nhất các bản đồ đặc trưng được căn chỉnh, bao gồm bản đồ đặc trưng của khung hình đích, để ước lượng bản đồ đặc trưng của khung hình đích được tăng cường; và môđun tái tạo để tái tạo khung hình đích được tăng cường bằng cách sử dụng bản đồ đặc trưng của nó.

Theo một phương án, môđun căn chỉnh được tạo cấu hình để thu được các độ dịch căn chỉnh từ mỗi trong số các khung hình lân cận đến khung hình đích với môđun FOC; và căn chỉnh, bởi lớp tích chập biến dạng được, mỗi trong số các bản đồ đặc trưng được trích xuất của các khung hình lân cận với bản đồ đặc trưng của khung hình đích sử dụng bản đồ độ dịch được tạo.

Theo một phương án, môđun căn chỉnh có cấu trúc hình tháp, phân tầng và biến dạng được (PCD) sử dụng môđun FOC để tính độ dịch.

Theo một phương án, môđun căn chỉnh đơn cấp theo cấu trúc PCD được tạo cấu hình để sử dụng, ở mỗi cấp không phải thấp nhất, các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của các hình ảnh được lấy mẫu xuống để tính các nhân tích chập biến dạng được; tính chỉnh, ở cấp thấp nhất, các nhân tích chập biến dạng được bằng cách hợp nhất chúng với giá trị ước lượng độ dịch được lấy mẫu lên từ cấp trước đó; và tính tích chập các nhân biến dạng được với bản đồ đặc trưng của khung hình lân cận.

Theo một phương án, trong đó đối với mỗi cấp không phải cao nhất, môđun FOC đơn giản hóa sử dụng các đặc trưng của các khung hình làm đầu vào được sử dụng.

Theo khía cạnh thứ bảy, sáng chế đề xuất thiết bị nội suy khung hình khuôn mặt, thiết bị này bao gồm môđun trích xuất để trích xuất các bản đồ đặc trưng từ các khung hình đầu vào trong đó khung hình đích bị thiếu; môđun nội suy, bao gồm môđun FOC theo khía cạnh thứ năm, để nội suy, từ mỗi tập hợp các bản đồ đặc trưng được trích xuất, bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo ra bởi môđun FOC; môđun hợp nhất để hợp nhất các bản đồ đặc trưng được nội suy để ước lượng bản

đồ đặc trưng được kết hợp của khung hình đích được nội suy; và môđun tái tạo để tái tạo khung hình đích được nội suy bằng cách sử dụng bản đồ đặc trưng được kết hợp của khung hình đích này.

Theo một phương án, môđun nội suy được tạo cấu hình để sử dụng mỗi tập hợp gồm ba khung hình lân cận từ các khung hình đầu vào cho môđun FOC; và biến đổi, bởi lớp tích chập biến dạng được, bản đồ đặc trưng của khung hình giữa trong mỗi tập hợp thành bản đồ đặc trưng của khung hình đích sử dụng bản đồ độ dịch được tạo.

Theo một phương án, môđun nội suy được tạo cấu hình để thực hiện nội suy xuôi và nội suy ngược.

Theo một phương án, môđun nội suy có cấu trúc hình tháp, phân tầng và biến dạng được (PCD) sử dụng môđun FOC để tính độ dịch.

Theo một phương án, môđun nội suy đơn cấp theo cấu trúc PCD được tạo cấu hình để sử dụng, ở mỗi cấp không phải thấp nhất, các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của các hình ảnh được lấy mẫu xuống để tính các nhân tích chập biến dạng được; tinh chỉnh, ở cấp thấp nhất, các nhân tích chập biến dạng được bằng cách hợp nhất chúng với giá trị ước lượng độ dịch được lấy mẫu lên từ cấp trước đó; và tính tích chập các nhân biến dạng được với bản đồ đặc trưng của các khung hình lân cận.

Theo một phương án, đối với mỗi cấp không phải cao nhất, môđun FOC đơn giản hóa sử dụng các đặc trưng của các khung hình làm đầu vào được sử dụng.

Theo khía cạnh thứ tám, sáng chế đề xuất hệ thống khử nhòe video khuôn mặt, hệ thống này bao gồm thiết bị tăng cường khung hình khuôn mặt theo khía cạnh thứ sáu để tăng cường khung hình đích; thiết bị nội suy khung hình khuôn mặt theo khía cạnh thứ bảy để nội suy khung hình đích; và thiết bị kết hợp để kết hợp khung hình đích được tăng cường với khung hình đích được nội suy.

Theo một phương án, hệ thống khử nhòe video khuôn mặt còn bao gồm thiết bị tiền xử lý để khử nhiễu và khử nhòe khung hình đích không quá nhòe trước.

Mô tả vắn tắt các hình vẽ

Các đối tượng, dấu hiệu và ưu điểm nêu trên và khác nữa của sáng chế sẽ trở nên rõ ràng hơn đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này bằng cách mô tả chi tiết các phương án làm ví dụ dựa vào các hình vẽ kèm theo, trong đó:

Fig.1 là sơ đồ minh họa phương pháp tăng cường và nội suy để xử trí khử nhòe theo tình trạng kỹ thuật;

Fig.2 là tập hợp các ảnh minh họa ưu điểm và nhược điểm của phương pháp tăng cường và nội suy;

Fig.3 là sơ đồ thể hiện cấu trúc của phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo một phương án làm ví dụ của sáng chế;

Fig.4 là sơ đồ thể hiện cấu trúc môđun tính độ dịch nhận diện khuôn mặt (FOC) trong phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế;

Fig.5 là tập hợp các sơ đồ thể hiện cấu trúc của các môđun căn chỉnh và nội suy đơn cấp;

Fig.6 là sơ đồ thể hiện cấu trúc hình tháp và phân tầng của nhánh tăng cường;

Fig.7 thể hiện các kết quả khử nhòe trên hai ví dụ từ tập dữ liệu VoxCeleb (trên cùng) và 300VW (dưới cùng) (trong đó chỉ các tỷ số tín hiệu cực đại trên nhiễu (PSNR) của các hình ảnh đầu ra so với các hình ảnh thực được thể hiện);

Fig.8 là tập hợp các sơ đồ thể hiện các so sánh theo cặp giữa nhánh tăng cường, nhánh nội suy và kết hợp của chúng (trong đó đối với mỗi cặp phương pháp, biểu đồ hình tròn báo cáo tỷ lệ thành công của một phương pháp so với các phương pháp khác, và số liệu thống kê được dựa vào các điểm số PSNR của dữ liệu kiểm tra 300VW được khử nhòe);

Fig.9 thể hiện các kết quả của mỗi môđun trên hai ví dụ trong tập kiểm tra 300VW (các PSNR có thể được sử dụng để đánh giá chất lượng của các hình ảnh); và

Fig.10 thể hiện các kết quả định tính trên các video nhòe thực.

Mô tả chi tiết sáng chế

Mặc dù có thể thực hiện nhiều thay đổi và nhiều phương án theo sáng chế, các phương án làm ví dụ sẽ được thể hiện trên các hình vẽ và được mô tả. Tuy nhiên, cần hiểu rằng sáng chế không bị giới hạn ở các phương án làm ví dụ và bao gồm tất cả các

thay đổi, các dạng tương đương hoặc các dạng thay thế nằm trong phạm vi kỹ thuật của sáng chế.

Các thuật ngữ bao gồm các số thứ tự, chẳng hạn như thứ hai và thứ nhất, có thể được sử dụng để mô tả nhiều thành phần, nhưng các thành phần không bị giới hạn ở các thuật ngữ này. Các thuật ngữ được sử dụng chỉ để phân biệt một thành phần với một thành phần khác. Ví dụ, không vượt ra khỏi phạm vi của sáng chế, thành phần thứ hai có thể được gọi là thành phần thứ nhất, và tương tự, thành phần thứ nhất có thể được gọi là thành phần thứ hai. Thuật ngữ “và/hoặc” bao gồm kết hợp bất kỳ của nhiều mục được liệt kê liên quan hoặc một mục bất kỳ trong số nhiều mục được liệt kê liên quan.

Khi nói rằng một thành phần “được nối” hoặc “được liên kết” với một thành phần khác, cần hiểu rằng thành phần này có thể được nối hoặc được liên kết trực tiếp với thành phần kia nhưng có thể có một thành phần khác ở giữa chúng. Mặt khác, khi nói rằng một thành phần “được nối trực tiếp” hoặc “được liên kết trực tiếp” với một thành phần khác, cần hiểu rằng không có thành phần khác ở giữa chúng.

Các thuật ngữ được sử dụng ở đây chỉ để mô tả các phương án làm ví dụ và không nhằm mục đích giới hạn sáng chế. Các cách diễn đạt ở dạng số ít bao gồm cả các cách diễn đạt ở dạng số nhiều trừ khi ngữ cảnh chỉ rõ ràng khác đi. Xuyên suốt bản mô tả này, cần hiểu rằng các thuật ngữ “bao gồm”, “có” v.v được sử dụng ở đây để chỉ sự có mặt của các dấu hiệu, số, bước, hoạt động, thành phần, phần, hoặc kết hợp của chúng nhưng không loại trừ sự có mặt hoặc bổ sung một hoặc nhiều dấu hiệu, số, bước, hoạt động, thành phần, phần khác, hoặc kết hợp của chúng.

Trừ khi được định nghĩa khác, các thuật ngữ được sử dụng ở đây bao gồm các thuật ngữ khoa học hoặc kỹ thuật có thể có nghĩa giống các thuật ngữ được hiểu chung bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này. Các thuật ngữ chẳng hạn như các thuật ngữ được định nghĩa trong các từ điển thông dụng cần được hiểu là có nghĩa giống với các nghĩa theo ngữ cảnh của lĩnh vực kỹ thuật liên quan và không nên được giải thích theo nghĩa lý tưởng hóa hoặc quá trang trọng trừ khi được định nghĩa rõ ràng ở đây.

Sau đây, các phương án của sáng chế sẽ được mô tả chi tiết dựa vào các hình vẽ kèm theo. Xuyên suốt các hình vẽ, các số chỉ dẫn giống nhau sẽ chỉ các thành phần giống hoặc tương tự, và phần mô tả lặp lại của chúng sẽ được lược bỏ.

Fig.1 là sơ đồ minh họa phương pháp tăng cường và nội suy để xử trí khử nhòe. Fig.2 là tập hợp các ảnh minh họa ưu điểm và nhược điểm của phương pháp tăng cường và nội suy. Fig.3 là sơ đồ thể hiện cấu trúc của phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo một phương án làm ví dụ của sáng chế. Fig.4 là sơ đồ thể hiện cấu trúc môđun tính độ dịch nhận diện khuôn mặt (FOC) trong phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế. Fig.5 là tập hợp các sơ đồ thể hiện cấu trúc của các môđun căn chỉnh và nội suy đơn cấp. Fig.6 là sơ đồ thể hiện cấu trúc hình tháp và phân tầng của nhánh tăng cường. Fig.7 thể hiện các kết quả khử nhòe trên hai ví dụ từ tập dữ liệu VoxCeleb (trên cùng) và 300VW (dưới cùng) (trong đó chỉ các tỷ số tín hiệu cực đại trên nhiễu (PSNR) của các hình ảnh đầu ra so với các hình ảnh thực được thể hiện). Fig.8 là tập hợp các sơ đồ thể hiện các so sánh theo cặp giữa nhánh tăng cường, nhánh nội suy, và kết hợp của chúng (trong đó đối với mỗi cặp phương pháp, biểu đồ hình tròn báo cáo tỷ lệ thành công của một phương pháp so với các phương pháp khác, và số liệu thống kê được dựa vào các điểm số PSNR của dữ liệu kiểm tra 300VW được khử nhòe). Fig.9 thể hiện các kết quả của mỗi môđun trên hai ví dụ trong tập kiểm tra 300VW (các PSNR có thể được sử dụng để đánh giá chất lượng của các hình ảnh). Fig.10 thể hiện các kết quả định tính trên các video nhòe thực.

Phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế bao gồm công đoạn nội suy và tăng cường khung hình và công đoạn kết hợp khung hình được nội suy và khung hình được tăng cường với nhau.

Theo Fig.1, trong phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế, phương pháp tăng cường sử dụng khung hình nhòe làm đích căn chỉnh. Trong phương pháp nội suy, khung hình nhòe được xử lý là khung hình bị thiếu và được ước lượng bằng cách nội suy từ các khung hình lân cận. Phương pháp tăng cường có lợi thế hơn so với phương pháp nội suy khi khung hình đích có chất lượng cao, và do đó việc tăng cường khung hình đích dễ dàng hơn so với việc tạo ra khung hình đích từ đầu.

Phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế có thể cải thiện chất lượng của video đầu ra được khử nhòe bằng cách kết hợp phương pháp tăng cường và phương pháp nội

suy.

Theo Fig.2, các dấu “√” chỉ báo các hình ảnh chất lượng cao, và các dấu “x” chỉ báo các hình ảnh chất lượng thấp. Trên khung hình đầu vào thứ hai nhòe hơn, phương pháp nội suy có thể được thực hiện tốt hơn so với phương pháp tăng cường. Mặt khác, phương pháp tăng cường tốt hơn đối với các khung hình ít nhòe hơn. Do đó, phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế có thể cung cấp các ưu điểm của cả phương pháp nội suy và phương pháp tăng cường để tạo ra video có chất lượng được cải thiện.

Ngoài ra, phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế có thể còn bao gồm FOC để căn chỉnh và nội suy khung hình.

FOC có thể được mô tả là phương pháp hoặc môđun FOC dưới đây.

FOC có thể được thực hiện bằng cách sử dụng các điểm mốc của khuôn mặt làm tiên nghiệm để tìm các điểm tương ứng và tính các độ dịch vị trí giữa các khung hình. Các điểm mốc của khuôn mặt có thể thu được bằng cách chạy bộ phát hiện điểm mốc của khuôn mặt.

Theo phương án làm ví dụ, có thể ước lượng các độ dịch vị trí đối với các nhân tích chập biến dạng được bằng cách sử dụng các bản đồ nhiệt điểm mốc thay vì định ra các ràng buộc rõ ràng trên cơ sở các điểm mốc được ước lượng. Phương pháp này có thể khắc phục sự không chính xác của các điểm mốc của khuôn mặt được phát hiện và cải thiện chất lượng của các đầu ra khử nhòe cuối cùng.

Ngoài ra, phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế có thể tạo ra hình ảnh rõ ràng và sắc nét ngay cả từ video trên Internet mà thực sự bị nhòe do rung camera hoặc do chuyển động nhanh.

Phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế trích xuất và sử dụng thông tin theo thời gian để khử nhòe video qua phương pháp nội suy và tăng cường khung hình khuôn mặt.

Theo phương án làm ví dụ của sáng chế, thông tin theo thời gian được sử dụng trong phương pháp nội suy khung hình khuôn mặt.

Tăng cường hình ảnh khuôn mặt

Trước tiên, khi việc phân phối dữ liệu kiểm tra khác với việc phân phối dữ liệu huấn luyện, thì hiệu suất của phương pháp khử nhòe có thể bị giảm. Để ngăn chặn vấn đề này, nhiều phương pháp theo lĩnh vực đã được phát triển, và khuôn mặt là một lĩnh vực quan trọng được quan tâm.

Ngoài ra, phương pháp phân tách khuôn mặt bị nhòe thành bốn phần và khử nhòe riêng rẽ mỗi phần trước khi kết hợp các kết quả hoặc phương pháp ghép nối các nhãn ngữ nghĩa khuôn mặt với các hình ảnh khuôn mặt có thể khó để áp dụng cho khuôn mặt bị nhòe.

Không giống phương pháp này, phương án làm ví dụ theo sáng chế có thể dễ dàng có hiệu quả với các khuôn mặt bị nhòe. Nói cách khác, việc khử nhòe sử dụng tiên nghiệm khuôn mặt không thể xử lý nhòe nghiêm trọng, trong khi phương án làm ví dụ của sáng chế sử dụng các điểm mốc của khuôn mặt để học các nhân động của các lớp tích chập biến dạng được. Theo cách này, có thể giải quyết vấn đề về việc phát hiện điểm mốc của khuôn mặt.

Nội suy khung hình khuôn mặt

Nội suy video theo thời gian được nghiên cứu trong lĩnh vực thị giác máy tính. Trước tiên, các dòng quang hai chiều được ước lượng từ hai khung hình lân cận. Tuy nhiên, việc ước lượng dòng quang từ khung hình nhòe có thể rất không chính xác. Trong trường hợp này, một phương pháp thay thế để nội suy khung hình khuôn mặt, mà sử dụng mạng tích chập biến dạng được thay vì dòng quang, có thể được sử dụng. Nói cách khác, thay vì sử dụng mạng được huấn luyện trước để nội suy, có thể học và tối ưu hóa mạng nội suy để khử nhòe video khuôn mặt.

Mạng tích chập biến dạng được

Mạng tích chập biến dạng được (deformable convolution network, DCN) thực hiện toán tử tích chập ở vị trí biến dạng được với độ dịch vị trí được chỉ định bởi nhánh mạng bổ sung thay vì sử dụng nhân tích chập.

Ngoài ra, DCN được mở rộng bằng cách đưa một đại lượng vô hướng biến đổi có thể học được bổ sung Δm để cải thiện hiệu suất của phép toán tích chập biến dạng được.

Trong phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt, tích chập biến dạng được có thể được sử dụng để căn chỉnh. Nói cách khác, tích chập biến dạng được có thể được sử dụng trong cả phương pháp tăng cường và nội suy.

Mạng tích chập biến dạng được sẽ được mô tả ngắn gọn dưới đây.

Khi bản đồ đặc trưng F được cho trước, đặc trưng ở vị trí p trong lớp tiếp theo $F'(p)$ có thể thu được bởi lớp tích chập với nhân có kích thước 3×3 như sau: (xem phương trình 1)

[Phương trình 1]

$$F'(p) = \sum_{k=1}^9 w_k * F(p + p_k).$$

Ở đây, $p_k \in \{(-1, -1), (-1, 0), \dots\}$ và w_k là các trọng số của vị trí p_k trong nhân. Không giống mạng tích chập bình thường, DCN được biến đổi có hai độ dịch có thể học được Δp_k và các đại lượng vô hướng biến đổi Δm_k . Bây giờ, đặc trưng ở vị trí p trong lớp tiếp theo $F'(p)$ được đưa ra bởi phương trình 2 dưới đây.

[Phương trình 2]

$$F'(p) = \sum_{k=1}^9 w_k * F(p + p_k + \Delta p_k) * \Delta m_k.$$

Phát hiện điểm mốc của khuôn mặt

Phương pháp phát hiện điểm mốc có thể tạo ra các điểm đặc trưng của khuôn mặt chính xác ngay cả từ các hình ảnh nhòe. Phương pháp như vậy là mô hình hồi quy bản đồ nhiệt trên cơ sở mạng nơron tích chập (CNN). Theo phương pháp này, các vị trí điểm mốc không được ước lượng trực tiếp, thay vào đó, bản đồ nhiệt của mỗi điểm đặc trưng được hồi quy và sau đó được sử dụng để tạo ra vị trí của các điểm mốc. Theo phương án làm ví dụ của sáng chế, bản đồ nhiệt trung gian có thể được sử dụng.

Phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế đề xuất mô hình tuần hoàn mà có thể cải thiện cả tác vụ khử nhòe và tác vụ phát hiện các điểm mốc. Mô hình được đề xuất bao gồm hai nhánh tương ứng với hai tác vụ được mô tả ở trên. Đầu ra của nhánh thứ nhất được sử dụng làm đầu vào của nhánh thứ hai.

Các đầu ra của nhánh thứ hai được sử dụng liên tục làm các đầu vào trước đó của nhánh thứ nhất. Do đó, hai tác vụ được cải thiện lẫn nhau. Nhánh phát hiện điểm mốc làm việc tốt hơn khi đầu vào rõ ràng, trong khi nhánh khử nhòe làm việc tốt hơn khi điểm mốc trước đó chính xác. Tuy nhiên, quy trình của mô hình này tốn quá nhiều thời gian.

Theo một phương án làm ví dụ khác của sáng chế, phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt bao gồm hai giai đoạn.

Ở giai đoạn thứ nhất, nhiễu được loại bỏ, và các khung hình không quá nhòe được khử nhòe. Sau đó, ở giai đoạn thứ hai, mô hình đầy đủ được áp dụng.

Mạng nội suy và tăng cường khung hình khuôn mặt

Trong phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt theo phương án làm ví dụ của sáng chế, có hai nhánh: tăng cường và nội suy, như được thể hiện trên Fig.3. Nhánh tăng cường và nhánh nội suy sẽ được mô tả dưới đây dưới dạng phương pháp tăng cường và phương pháp nội suy.

Theo phương án làm ví dụ của sáng chế, công đoạn nội suy và tăng cường khung hình bao gồm công đoạn trích xuất điểm mốc và đặc trưng, công đoạn tính độ dịch nhận diện khuôn mặt bằng cách sử dụng điểm mốc và đặc trưng, và công đoạn tạo ra khung hình đích bằng cách ước lượng bản đồ đặc trưng từ độ dịch nhận diện khuôn mặt tính được. Điểm mốc và đặc trưng có thể được trích xuất từ các khung hình lân cận.

Cụ thể, ở nhánh tăng cường, với sự hỗ trợ của khung hình bị biến dạng bằng cách căn chỉnh các khung hình lân cận với khung hình đích, mỗi khung hình nhòe được tăng cường trực tiếp.

Trong nhánh nội suy, khung hình nhòe được xử lý là khung hình thiếu và được ước lượng từ các khung hình lân cận. Các hình ảnh đầu ra từ nhánh tăng cường và nhánh nội suy được hợp nhất với nhau để tạo ra đầu ra cuối cùng.

Để cải thiện cả nhánh tăng cường và nhánh nội suy, các điểm mốc của khuôn mặt được sử dụng để tính độ dịch vị trí cho nhân tích chập biến dạng được.

FOC sẽ được mô tả dưới đây là phương pháp hoặc môđun mới để lấy khung hình đích, các khung hình lân cận và các bản đồ nhiệt điểm mốc của khuôn mặt tương ứng để ước lượng các độ dịch căn chỉnh. FOC được sử dụng trong cả nhánh tăng cường và nhánh nội suy. FOC sẽ được mô tả dưới đây, tiếp theo là phần mô tả nhánh tăng cường và nhánh nội suy.

Tính độ dịch nhận diện khuôn mặt (FOC)

Ước lượng độ dịch nhận diện khuôn mặt là công đoạn quan trọng trong căn chỉnh đặc trưng trên cơ sở DCN nhưng lại bị bỏ qua trong các tác vụ khử nhòe hiện tại.

Theo phương án làm ví dụ của sáng chế, cả hai phương pháp có thể chỉ sử dụng hai lớp tích chập để ước lượng độ dịch. Do đó, FOC có thể được thực hiện.

Theo phương án làm ví dụ của sáng chế, công đoạn tính độ dịch nhận diện khuôn mặt bằng cách sử dụng bản đồ nhiệt điểm mốc và đặc trưng có thể bao gồm công đoạn nhận đặc trưng và bản đồ nhiệt điểm mốc của khung hình lân cận và công đoạn ghép nối đặc trưng và bản đồ nhiệt điểm mốc nhận được với nhau và đưa chúng vào FOC để thu được độ dịch.

Công đoạn tạo ra khung hình đích bằng cách ước lượng bản đồ đặc trưng thông qua độ dịch nhận diện khuôn mặt tính được có thể bao gồm công đoạn biến đổi bản đồ đặc trưng đích thành khung hình đích.

Trước tiên, trong FOC, các đặc trưng và các bản đồ nhiệt điểm mốc của một số khung hình lân cận có thể được sử dụng làm các đầu vào. Các đặc trưng và các bản đồ nhiệt điểm mốc có thể được trích xuất từ các khung hình lân cận. Ngoài ra, trong nhánh tăng cường (hoặc trong công đoạn thực hiện tăng cường), bước trích xuất được thực hiện từ các khung hình nhòe như được thể hiện trên Fig.3. Mặt khác, trong nhánh nội suy (hoặc trong công đoạn thực hiện nội suy), các khung hình nhòe được loại trừ, và do đó bước trích xuất được thực hiện, ngoại trừ các khung hình nhòe.

Đầu ra của FOC là bản đồ độ dịch để căn chỉnh hoặc nội suy.

Cụ thể hơn, giả sử rằng có một tập hợp các khung hình $F=\{f_0, f_1, \dots, f_k\}$, các bản đồ nhiệt điểm mốc $H=\{h_0, h_1, \dots, h_k\}$ có thể được tạo ra. Mỗi bản đồ nhiệt điểm

mốc bao gồm $L (=68)$ kênh, mỗi trong số đó tương ứng với một điểm đặc trưng.

Sau đó, phương pháp hoặc mô đun trích xuất đặc trưng có thể được sử dụng để biến đổi các đặc trưng thành các bản đồ đặc trưng $H^l = \{h^l_0, h^l_1, \dots, h^l_k\}$.

Cuối cùng, $\{f_0, f_1, \dots, f_k\}$ và $\{h^l_0, h^l_1, \dots, h^l_k\}$ có thể được ghép nối với nhau và được đưa vào FOC để xuất ra độ dịch.

Trong nhánh tăng cường, FOC lấy một khung hình lân cận và khung hình đích làm các đầu vào. Lớp tích chập biến dạng được có thể sử dụng giá trị độ dịch ước lượng được để căn chỉnh bản đồ đặc trưng lân cận với bản đồ đặc trưng đích.

Trong nhánh nội suy, FOC lấy ba khung hình đầu vào $[t-2, t-1, t+1]$ hoặc $[t+2, t+1, t-1]$. Lớp tích chập biến dạng được có thể sử dụng giá trị độ dịch tính được để biến đổi khung hình giữa của tập đầu vào thành khung hình đích t (FOC được minh họa trên Fig.4).

Nhánh tăng cường

Để tinh chỉnh khung hình nhòe t , nhánh tăng cường có thể sử dụng thông tin của năm khung hình liên tiếp $t-2, t-1, t, t+1$ và $t+2$. Nhánh này có bốn công đoạn xử lý: trích xuất đặc trưng, căn chỉnh, hợp nhất và tái tạo. Mỗi khung hình sau khi lấy mẫu xuống đi qua mô đun trích xuất đặc trưng để bản đồ đặc trưng có thể được tạo ra. Các bản đồ đặc trưng của tất cả các khung hình được căn chỉnh với bản đồ đặc trưng của khung hình t sử dụng mô đun căn chỉnh.

Mô đun căn chỉnh có thể là phần mở rộng của mô đun căn chỉnh hình tháp, phân tầng và biến dạng được (PCD) với mô đun FOC.

Theo phương án làm ví dụ của sáng chế, tất cả các bản đồ đặc trưng được căn chỉnh bao gồm bản đồ đặc trưng của khung hình đích t được kết hợp bởi mô đun hợp nhất để ước lượng bản đồ đặc trưng của khung hình đích được tăng cường. Ở công đoạn cuối cùng, khung hình đích đã ước lượng có thể được tạo ra từ bản đồ đặc trưng bằng cách tái tạo.

Mặc dù mô đun căn chỉnh được phát triển trên cơ sở căn chỉnh PCD theo phương án làm ví dụ, mô đun căn chỉnh có thể được cải thiện bằng cách sử dụng mô đun FOC để tính các độ dịch vị trí cho các nhân tích chập biến dạng được. Phần mở rộng này là phương pháp tốt hơn các phương pháp khử nhòe công nghệ cao hiện có.

Theo phương án làm ví dụ của sáng chế, để căn chỉnh khung hình $t+i$ với khung hình t , các độ dịch có thể được tính và các đặc trưng có thể được căn chỉnh ở một số tỷ lệ từ thô đến tinh. Đối với mỗi cấp tỷ lệ, các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của các hình ảnh được lấy mẫu xuống được sử dụng để tính các nhân tích chập biến dạng được. Ngoại trừ cấp thấp nhất của FOC, các nhân sẽ được tinh chỉnh bằng cách hợp nhất với các độ dịch được lấy mẫu lên được ước lượng từ cấp trước đó.

Cuối cùng, các nhân biến dạng được tính tích chập với bản đồ đặc trưng của khung hình $t + i$ (Fig.5A thể hiện quy trình căn chỉnh này).

Như được thể hiện trên Fig.6, cấu trúc hình tháp và phân tầng được sử dụng để căn chỉnh. Cấu trúc hình tháp của các bản đồ đặc trưng được tính cho mỗi khung hình.

Cấp thứ nhất của hình tháp là bản đồ đặc trưng được trích xuất bởi mô đun trích xuất đặc trưng. Từ cấp thứ hai, bản đồ đặc trưng của hình tháp được tính trực tiếp từ cấp trước đó bởi phép toán tích chập bước trượt. Số lượng các cấp được sử dụng trong các phương án làm ví dụ là ba. Cấp cao nhất được đánh số là cấp 1 và cấp thấp nhất là cấp 3. Để căn chỉnh khung hình $t'=t+i$ với khung hình t , các cấp trong cấu trúc hình tháp của khung hình t' được căn chỉnh với các cấp trong hình tháp của khung hình t , một cách tương ứng.

Ở mỗi cấp không phải là 1-1, các đặc trưng được căn chỉnh được ghép nối với các đặc trưng được lấy mẫu lên tương ứng của cấp 1, và tiếp theo là lớp tích chập. Đối với việc tính độ dịch, lớp tích chập có thể được thay thế bởi FOC.

Ở đây, F_t^1 và $F_{t'}^1$ là các bản đồ đặc trưng của khung hình t và khung hình t' ở cấp 1, đặc trưng được căn chỉnh mong muốn là A_t^1 , và bản đồ của các độ dịch vị trí là ΔP_t^1 .

Ngoài ra, các bản đồ nhiệt điểm mốc của khung hình t và khung hình t' là h_t và $h_{t'}$. $A_{t'}^1$ và $\Delta P_{t'}^1$ được tính như sau:

$$\begin{aligned}\Delta P_{t'}^3 &= FOC(F_{t'}^3, F_t^3), \\ A_{t'}^3 &= DConv(F_{t'}^3, \Delta P_{t'}^3), \\ \Delta P_{t'}^2 &= FOC(Conv(F_{t'}^2 \oplus F_t^2), (F_{t'}^3)^{\uparrow 2}), \\ A_{t'}^2 &= Conv(DConv(F_{t'}^2, \Delta P_{t'}^2) \oplus (A_{t'}^3)^{\uparrow 2}), \\ \Delta P_{t'}^1 &= FOC(Conv(F_{t'}^1 \oplus F_t^1), (F_{t'}^2)^{\uparrow 2}, h_{t'}, h_t), \\ A_{t'}^1 &= Conv(DConv(F_{t'}^1, \Delta P_{t'}^1) \oplus (A_{t'}^2)^{\uparrow 2}),\end{aligned}$$

trong đó $\oplus$ và $\uparrow$ là phép toán ghép nối và phép toán lấy mẫu lên.

Chỉ bản đồ nhiệt điểm mốc của khuôn mặt có thể được sử dụng trong FOC của cấp 0. Để căn chỉnh bản đồ đặc trưng của mỗi cấp không phải cao nhất, FOC đơn giản hóa với mạng bộ mã hóa-bộ giải mã nhỏ không sử dụng tiên nghiệm điểm mốc có thể được sử dụng để tiết kiệm chi phí tính toán.

Nhánh nội suy

Nhánh nội suy có thể sử dụng tích chập biến dạng được với cấu trúc hình tháp và phân tầng đặc trưng. Khung hình “thiếu” được nội suy từ I_{t-2} , I_{t-1} , I_{t+1} , I_{t+2} và các bản đồ nhiệt điểm mốc tương ứng h_{t-2} , h_{t-1} , h_{t+1} và h_{t+2} của chúng.

Phương án làm ví dụ của sáng chế sử dụng cả nội suy xuôi và nội suy ngược, mà được ký hiệu lần lượt là $I^{\rightarrow}_t$ và $I^{\leftarrow}_t$. Phép nội suy xuôi và nội suy ngược có thể được hợp nhất như sau.

Thứ nhất, $I^{\rightarrow}_t$ có thể được tính bằng cách áp dụng môđun nội suy cho ba khung hình $[I_{t-2}, I_{t-1}, I_{t+1}]$ và các bản đồ nhiệt điểm mốc của chúng.

Thứ hai, $I^{\leftarrow}_t$ có thể được tính bằng cách áp dụng môđun nội suy cho ba khung hình $[I_{t+2}, I_{t+1}, I_{t-1}]$ và các bản đồ nhiệt điểm mốc của chúng (theo thứ tự giảm dần của các khung hình).

Cuối cùng, $I^{\rightarrow}_t$ và $I^{\leftarrow}_t$ được kết hợp bởi môđun hợp nhất, và khung hình thiếu có thể được ước lượng sử dụng bản đồ đặc trưng được kết hợp (Fig.5B thể hiện quy trình nội suy xuôi).

Đối với phép nội suy xuôi và nội suy ngược, ba khung hình lân cận có thể được sử dụng thay vì bốn. Theo các nghiên cứu được mô tả trong các ví dụ thực nghiệm sau đây, không có lợi ích gì khi sử dụng bốn khung hình thay vì ba. Do đó, ba khung hình có thể được sử dụng để giảm chi phí tính toán.

Phương pháp nội suy khung hình khuôn mặt theo phương án làm ví dụ của sáng chế được thiết kế dành riêng để khử nhòe. Phương pháp này có thể dễ dàng được chuyển thành ứng dụng tương tự của nội suy video chuyển động chậm.

Kết hợp khung hình được tăng cường và khung hình được nội suy

Các hình ảnh đầu ra của nhánh tăng cường và nhánh nội suy có thể được kết hợp sử dụng môđun mạng đơn giản. Môđun kết hợp có cấu trúc đơn giản bao gồm hai lớp lấy mẫu xuống, chín khối ResNet và hai lớp lấy mẫu lên.

Hợp nhất sớm là một phương pháp khác để hợp nhất các kết quả của nhánh tăng cường và nhánh nội suy. Có một phương pháp để hợp nhất các bản đồ đặc trưng trung gian, năm bản đồ từ nhánh tăng cường và hai bản đồ từ nhánh nội suy, thay vì hợp nhất các hình ảnh đầu ra cuối cùng.

Lượng mất mát huấn luyện

Lượng mất mát L1 giữa khung hình đã ước lượng và khung hình thực sự có thể được sử dụng ở cả nhánh nội suy và nhánh tăng cường. Ngoài ra, độ sắc nét của khuôn mặt được tái tạo được tăng cường sử dụng lượng mất mát gradient theo không gian. Các bản đồ gradient đối với các tọa độ x và y được ký hiệu lần lượt là G_x và G_y . Lượng mất mát huấn luyện của nhánh tăng cường và nhánh nội suy như sau.

$$\begin{aligned}\mathcal{L}_{in} &= \|I^{gt} - I^{in}\|_1 + \|G_x(I^{gt}) - G_x(I^{in})\|_1 + \|G_y(I^{gt}) - G_y(I^{in})\|_1, \\ \mathcal{L}_{en} &= \|I^{gt} - I^{en}\|_1 + \|G_x(I^{gt}) - G_x(I^{en})\|_1 + \|G_y(I^{gt}) - G_y(I^{en})\|_1,\end{aligned}$$

trong đó I^{en} và I^{in} là các đầu ra hình ảnh của nhánh tăng cường và nhánh nội suy, và I^{gt} là khung hình thực sự.

Sau khi nhánh tăng cường và nhánh nội suy được huấn luyện, môđun kết hợp được huấn luyện để kết hợp I^{en} và I^{in} . Đối với môđun kết hợp, chỉ lượng mất mát L1 có thể được sử dụng để huấn luyện (ví dụ, $\mathcal{L}_{combine} = \|I^{gt} - \hat{I}\|_1$, trong đó $\hat{I}$ là đầu ra cuối cùng của mạng khử nhiễu).

Các thực nghiệm

Chi tiết phương án thực hiện

Mạng đã được huấn luyện với kích cỡ lô nhỏ gồm 12 hình ảnh. Thuật toán tối ưu Adam đã được sử dụng với $\beta_0=0,9$ và $\beta_1=0,999$. Tốc độ học (learning rate) ban đầu là 10^{-4} và việc huấn luyện được hội tụ sau khoảng 200.000 lần lặp.

Chiến lược hai giai đoạn đã được sử dụng. Ở giai đoạn thứ nhất, môđun khử nhiễu nhỏ đã được sử dụng để loại bỏ nhiễu và khử nhiễu một phần video. Ở giai đoạn

thứ hai, các khung hình rất nhòe đã được xử lý. Chiến lược này đã cải thiện hiệu suất của mô đun căn chỉnh và nội suy.

Các tập dữ liệu để huấn luyện và đánh giá

Để huấn luyện chương trình khung và so sánh nó với các phương pháp khác, các tập dữ liệu thường được sử dụng VoxCeleb và 300VW có thể được sử dụng. Để xác minh thêm hệ thống trong một kịch bản thực tế, hệ thống đã được kiểm tra với các tập dữ liệu YouTube nhòe được thu thập.

Tập dữ liệu VoxCeleb là tập dữ liệu lớn của các trích ảnh video khuôn mặt được thu thập từ YouTube với khoảng 146K video huấn luyện và 6K video kiểm tra.

Do kích cỡ lớn, 1300 video đã được chọn ngẫu nhiên từ các video huấn luyện gồm khoảng 700K khung hình để tạo ra tập huấn luyện. Nhân nhòe ngẫu nhiên đã được sử dụng để tạo ra dữ liệu huấn luyện trong các thực nghiệm. Ngoài ra, 130 video đã được chọn từ tập kiểm tra của VoxCeleb để đánh giá.

Để đánh giá khả năng tổng quát hóa của mô hình được huấn luyện đối với miền mới, 30 video của tập dữ liệu 300VW đã được thêm vào dữ liệu kiểm tra.

Phương pháp này không chỉ được đánh giá với các hình ảnh bị nhòe nhân tạo mà còn được kiểm tra với 100 video khuôn mặt chất lượng thấp từ YouTube bao gồm các khung hình nhòe. Các kết quả thực nghiệm định tính thể hiện rằng phương án làm ví dụ của sáng chế cung cấp các kết quả khử nhòe rõ ràng hơn so với các phương pháp khác.

So sánh với các phương pháp công nghệ cao

Theo phương án làm ví dụ, PSNR và sự giống nhau về cấu trúc (SSIM) được sử dụng làm các tiêu chí đánh giá.

[Bảng 1]

Phương pháp	VoxCeleb		300VW	
	PSNR	SSIM	PSNR	SSIM
Kupyn và cộng sự [11]	26,45	0,4418	26,83	0,6228
Yasarla và cộng sự [26]	31,08	0,9035	31,35	0,9538
Ren và cộng sự [13]	32,26	0,9089	33,05	0,9278

Liu và cộng sự [12]	34,71	0,9342	34,15	0,9501
Wang và cộng sự [23]	34,77	0,8409	29,12	0,8775
Phương án làm ví dụ	37,04	0,9612	38,22	0,9700

Các giá trị PSNR và SSIM trung bình của tất cả các phương pháp cho dữ liệu kiểm tra của các tập dữ liệu VoxCeleb và 300VW được thể hiện trong bảng 1.

Cụ thể hơn, phương án làm ví dụ của sáng chế đạt được PSNR là 37,04 trong tập dữ liệu VoxCeleb, trong khi PSNR của phương pháp tốt thứ hai chỉ là 34,77. Khoảng cách hiệu suất giữa phương án làm ví dụ và phương pháp tốt thứ hai rộng hơn nhiều đối với tập kiểm tra 300VW.

Ngoài ra, phương án làm ví dụ tạo ra đầu ra chất lượng cao với các đường biên sắc nét giữa các bộ phận trên khuôn mặt ngay cả trong điều kiện bất lợi như được thể hiện trên Fig.7.

Bảng 2 dưới đây thể hiện các lợi ích của FOC và các đầu vào điểm mốc trên cả nhánh tăng cường và nhánh nội suy. Các số được báo cáo trong tập dữ liệu 300VW.

[Bảng 2]

Được sử dụng trong nhánh tăng cường?				Được sử dụng trong nhánh nội suy?			
FOC	Các điểm mốc	PSNR	SSIM	FOC	Các điểm mốc	PSNR	SSIM
x	x	29,12	0,8775	x	x	35,49	0,9490
✓	x	37,37	0,9590	✓	x	36,36	0,9598
✓	✓	37,81	0,9649	✓	✓	36,87	0,9633

Bảng 2 cho thấy kết quả của các thực nghiệm về tầm quan trọng của việc có các mô đun FOC để tính các độ dịch vị trí cho các nhân tích chập biến dạng được. Lưu ý rằng các mô đun FOC được đề xuất để sử dụng ở cả nhánh tăng cường và nhánh nội suy. Có thể thấy, có một lợi ích rõ ràng của việc sử dụng các mô đun FOC; PSNR giảm mạnh nếu FOC không được sử dụng trong nhánh tăng cường. Các bản đồ nhiệt điểm mốc cũng hữu ích cho cả nhánh tăng cường và nhánh nội suy. Bằng cách sử dụng các bản đồ nhiệt điểm mốc của hai khung hình, FOC có thể dự đoán chuyển động của khuôn mặt và tạo ra các độ dịch chính xác cho tích chập biến dạng được. Lưu ý rằng các bản đồ nhiệt điểm mốc chỉ được sử dụng làm tiên nghiệm ẩn để tính các độ dịch vị trí cho các nhân biến

dạng được. Phương pháp này rất bền vững đối với sự hư hỏng của bộ phát hiện điểm mốc, không giống với các phương pháp sử dụng các tiên nghiệm cấu trúc khuôn mặt một cách rõ ràng.

Phương án làm ví dụ có thể tránh được các hạn chế của hai nhánh được mô tả ở trên. Trên Fig.9, hàng thứ nhất là trường hợp trong đó các đầu ra của nhánh nội suy tốt hơn so với các đầu ra của nhánh tăng cường, trong khi ở hàng thứ hai, nhánh nội suy không tạo ra các bộ phận trên khuôn mặt sắc nét. Sự kết hợp của hai nhánh có thể mang lại các kết quả tốt nhất trong cả hai trường hợp. Như được thể hiện ở biểu đồ thứ hai và thứ ba trên Fig.8, phương pháp kết hợp vượt trội hơn so với phương pháp tăng cường và nội suy lần lượt là 95,3% và 85,0% thời gian.

Bảng 3 dưới đây thể hiện các lợi ích của phương pháp hợp nhất tăng cường-nội suy. Các số được báo cáo trong tập dữ liệu 300VW.

[Bảng 3]

Tăng cường	Nội suy	Hợp nhất muộn	PSNR	SSIM
✓	x		37,81	0,9649
x	✓		36,87	0,9633
✓	✓	x	37,88	0,9655
✓	✓	✓	38,22	0,9700

Để kiểm nghiệm sơ đồ hợp nhất, hiệu suất của phương pháp hợp nhất sớm được thể hiện trong hàng thứ ba của bảng 3. Không giống phương pháp hợp nhất muộn, phương pháp này chỉ tạo ra độ tăng ích nhỏ. Điều này có thể được giải thích là do số lượng các bản đồ đặc trưng từ nhánh tăng cường trội hơn số lượng các bản đồ được nội suy, làm cho phương pháp hợp nhất sớm không hiệu quả.

Theo phương án làm ví dụ, trong trường hợp mô đun nội suy, thực nghiệm đã được thực hiện trên cả nội suy xuôi và nội suy ngược với bốn khung hình. Tuy nhiên, không có sự khác biệt đáng kể về hiệu suất khi ba hay bốn khung hình được sử dụng. PSNR/SSIM của hai phương pháp lần lượt là 36,83/0,9632 và 36,87/0,9633.

Đánh giá trên các video nhòe thực

Cuối cùng, để hợp thức hóa khả năng tổng quát và sự hữu ích thực tế của mô hình khử nhòe, mô hình khử nhòe đã được chạy cùng với các phương pháp công nghệ

cao khác trên các video Youtube chất lượng thấp thực. Các kết quả được thể hiện trên Fig.10.

Các video chứa vết nhòe thực do rung camera, chuyển động khuôn mặt nhanh hoặc cả hai. Một khảo sát đối với người dùng đã được thực hiện trong đó 46 người dùng đã được hỏi để chọn hình ảnh được khử nhòe tốt nhất trong số các kết quả của phương án làm ví dụ và các phương pháp công nghệ cao khác đối với 23 trường hợp kiểm tra. Phương pháp theo phương án làm ví dụ đã được chọn là phương pháp tốt nhất bởi 91,3% trong số 23 trường hợp kiểm tra.

Có thể dễ dàng thấy rằng các bộ phận trên khuôn mặt, chẳng hạn như mắt hoặc mũi, được tạo ra bởi phương án làm ví dụ rõ ràng hơn so với các bộ phận được tạo ra bởi các mô hình khác.

Phương án làm ví dụ có hai nhánh mạng song song để tăng cường khung hình khuôn mặt và nội suy khung hình khuôn mặt. Ngoài ra, phương pháp (hoặc mô đun) mới để tính các độ dịch theo không gian giữa hai bản đồ đặc trưng của khuôn mặt bằng cách sử dụng các bản đồ nhiệt điểm mốc của khuôn mặt đã được giới thiệu. Các thực nghiệm trên một số khung hình nhòe được tạo ra thực và nhân tạo cho thấy hiệu suất tốt của phương pháp.

Theo một phương án làm ví dụ của sáng chế, có thể thực hiện phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt nhờ đó chất lượng của các kết quả cuối cùng được cải thiện.

Theo một phương án làm ví dụ của sáng chế, có thể thực hiện phương pháp nội suy và tăng cường khung hình khuôn mặt để khử nhòe video khuôn mặt mà cung cấp các kết quả định tính và định lượng được cải thiện của dữ liệu video nhòe.

Các ưu điểm và lợi ích khác nhau của sáng chế không bị giới hạn ở các ưu điểm và lợi ích được mô tả ở trên và có thể được hiểu dễ dàng trong phần mô tả chi tiết các phương án của sáng chế.

Thuật ngữ “bộ phận” được sử dụng trong phương án làm ví dụ theo sáng chế có nghĩa là bộ phận phần mềm hoặc phần cứng, chẳng hạn như mảng cổng lập trình được định cấu hình (field-programmable gate array, FPGA) hoặc mạch tích hợp chuyên dụng (application-specific integrated circuit, ASIC), và “bộ phận” thực hiện một vai trò cụ

thể. Tuy nhiên, “bộ phận” không bị giới hạn ở phần mềm hoặc phần cứng. “Bộ phận” có thể được tạo cấu hình để nằm trong vật ghi khả lập địa chỉ và cũng có thể được tạo cấu hình để chạy một hoặc nhiều bộ xử lý. Do đó, như một ví dụ, “bộ phận” bao gồm các thành phần, chẳng hạn như các thành phần phần mềm, các thành phần phần mềm định hướng đối tượng, các thành phần lớp và các thành phần tác vụ, các quy trình, các chức năng, các thuộc tính, các thủ tục, các chương trình con, các đoạn mã chương trình, các bộ điều khiển, phần sụn, vi mã, các mạch, dữ liệu, cơ sở dữ liệu, các cấu trúc dữ liệu, các bảng, các mảng, và các biến thể. Các thành phần và các chức năng được cung cấp trong “các bộ phận” có thể được ghép với số lượng các thành phần và “các bộ phận” nhỏ hơn hoặc có thể được chia nhỏ thành số lượng các thành phần và “các bộ phận” lớn hơn. Ngoài ra, các thành phần và “các bộ phận” có thể được triển khai để chạy một hoặc nhiều bộ xử lý trung tâm (CPU) trong thiết bị hoặc thẻ đa phương tiện bảo mật.

Tất cả hoặc ít nhất một số trong số các thành phần của thiết bị được mô tả ở trên có thể được triển khai dưới dạng môđun phần cứng, môđun phần mềm, hoặc kết hợp của môđun phần cứng và môđun phần mềm. Môđun phần mềm có thể được hiểu là, ví dụ, lệnh được thực thi bởi bộ xử lý để điều khiển việc tính toán trong thiết bị, và lệnh có thể có dạng được cài đặt trong bộ nhớ của thiết bị.

Thiết bị và phương pháp theo các phương án làm ví dụ của sáng chế có thể được triển khai dưới dạng lệnh chương trình mà có thể được thực hiện qua các phương tiện tính toán khác nhau và có thể được ghi trong vật ghi đọc được bằng máy tính. Vật ghi đọc được bằng máy tính có thể bao gồm lệnh chương trình, tệp dữ liệu, cấu trúc dữ liệu, v.v ở dạng riêng rẽ hoặc kết hợp. Lệnh chương trình được ghi trong vật ghi có thể được thiết kế và xây dựng đặc biệt cho sáng chế hoặc có thể đã được biết đến và sử dụng bởi người có hiểu biết trung bình trong lĩnh vực phần mềm máy tính. Các ví dụ về vật ghi đọc được bằng máy tính bao gồm các vật ghi từ chẳng hạn như đĩa cứng, đĩa mềm và băng từ, phương tiện quang chẳng hạn như bộ nhớ chỉ đọc-đĩa compac (CD-ROM) và đĩa đa dụng số (DVD), phương tiện từ-quang chẳng hạn như đĩa mềm quang học, và các thiết bị phần cứng, chẳng hạn như ROM, bộ nhớ truy cập ngẫu nhiên (RAM), và bộ nhớ nhanh, được tạo cấu hình đặc biệt để lưu và thực thi lệnh chương trình. Các ví dụ về lệnh chương trình bao gồm không chỉ mã ngôn ngữ máy được tạo ra bởi trình biên dịch mà còn bao gồm mã ngôn ngữ cấp cao có thể thực thi được bởi máy tính qua trình thông dịch và dạng tương tự. Thiết bị phần cứng nêu trên có thể được tạo cấu hình để vận hành

như một hoặc nhiều môđun phần mềm để thực hiện hoạt động của sáng chế, và ngược lại.

Mặc dù sáng chế đã được mô tả ở trên với các phương án làm ví dụ, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này cần hiểu rằng các biến thể và các thay đổi khác nhau có thể được thực hiện từ phần mô tả ở trên mà không nằm ngoài phạm vi của sáng chế được xác định trong yêu cầu bảo hộ kèm theo.

YÊU CẦU BẢO HỘ

1. Phương pháp tính độ dịch nhận diện khuôn mặt (FOC), phương pháp này bao gồm các bước:

nhận các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của tập hợp các khung hình;

thực hiện phép ghép vector của các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc với nhau để tạo ra các đầu vào, trong đó phép ghép vector nối các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc từ đầu đến cuối bằng cách ghép vector của các bản đồ đặc trưng và vector của các bản đồ nhiệt điểm mốc với nhau từ đầu đến cuối để tạo thành vector mới được ghép bao gồm các phần tử của vector của các bản đồ đặc trưng và vector của các bản đồ nhiệt điểm mốc; và

tạo ra bản đồ độ dịch bằng cách thực hiện việc phân tích hồi quy trên các đầu vào,

trong đó bản đồ độ dịch biểu diễn bản đồ của các vector độ dịch 2D chỉ báo các độ lớn mà theo đó các vị trí lấy mẫu của nhân tích chập sẽ bị biến dạng tại mỗi điểm ảnh để biến đổi đặc trưng từ một khung hình sang một khung hình khác trong mạng tích chập biến dạng được (DCN), và

trong đó bản đồ độ dịch được tạo cấu hình để được sử dụng bởi DCN để căn chỉnh và/hoặc nội suy bản đồ đặc trưng khuôn mặt.

2. Phương pháp tăng cường khung hình khuôn mặt, phương pháp này bao gồm các bước:

trích xuất các bản đồ đặc trưng từ các khung hình đầu vào chứa khung hình đích;

căn chỉnh mỗi trong số các bản đồ đặc trưng được trích xuất với bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo ra bằng cách sử dụng phương pháp FOC theo điểm 1;

hợp nhất các bản đồ đặc trưng được căn chỉnh, bao gồm bản đồ đặc trưng của khung hình đích, để ước lượng bản đồ đặc trưng của khung hình đích được tăng cường; và

tái tạo khung hình đích được tăng cường bằng cách sử dụng bản đồ đặc trưng của khung hình đích này.

3. Phương pháp tăng cường khung hình khuôn mặt theo điểm 2, trong đó các khung

hình đầu vào bao gồm năm khung hình liên tiếp.

4. Phương pháp tăng cường khung hình khuôn mặt theo điểm 2, trong đó bước căn chỉnh bao gồm các bước:

thu được các độ dịch căn chỉnh từ mỗi trong số các khung hình lân cận đến khung hình đích với phương pháp FOC; và

căn chỉnh, bởi lớp tích chập biến dạng được, mỗi trong số các bản đồ đặc trưng được trích xuất của các khung hình lân cận với bản đồ đặc trưng của khung hình đích sử dụng bản đồ độ dịch được tạo.

5. Phương pháp tăng cường khung hình khuôn mặt theo điểm 2, trong đó bước căn chỉnh là căn chỉnh hình tháp, phân tầng và biến dạng được (PCD) sử dụng phương pháp FOC cho việc tính độ dịch.

6. Phương pháp tăng cường khung hình khuôn mặt theo điểm 5, trong đó việc căn chỉnh đơn cấp theo căn chỉnh PCD bao gồm các bước:

sử dụng, ở mỗi cấp không phải thấp nhất, các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của các hình ảnh được lấy mẫu giảm để tính các nhân tích chập biến dạng được;

tinh chỉnh, ở cấp thấp nhất, các nhân tích chập biến dạng được bằng cách hợp nhất các nhân này với sự ước lượng độ dịch được lấy mẫu tăng từ cấp trước đó; và

tính tích chập các nhân biến dạng được với bản đồ đặc trưng của khung hình lân cận.

7. Phương pháp tăng cường khung hình khuôn mặt theo điểm 5, trong đó số lượng cấp là ba.

8. Phương pháp tăng cường khung hình khuôn mặt theo điểm 5, trong đó đối với mỗi cấp không phải cao nhất, phương pháp FOC đơn giản hóa sử dụng các đặc trưng của các khung hình làm đầu vào được sử dụng.

9. Phương pháp nội suy khung hình khuôn mặt, phương pháp này bao gồm các bước:

trích xuất các bản đồ đặc trưng từ các khung hình đầu vào trong đó khung hình đích bị thiếu;

nội suy, từ mỗi tập hợp các bản đồ đặc trưng được trích xuất, bản đồ đặc trưng của

khung hình đích dựa vào bản đồ độ dịch được tạo ra bằng cách sử dụng phương pháp FOC theo điểm 1;

hợp nhất các bản đồ đặc trưng được nội suy để ước lượng bản đồ đặc trưng được kết hợp của khung hình đích được nội suy; và

tái tạo khung hình đích được nội suy bằng cách sử dụng bản đồ đặc trưng được kết hợp của khung hình đích này.

10. Phương pháp nội suy khung hình khuôn mặt theo điểm 9, trong đó các khung hình đầu vào bao gồm bốn khung hình.

11. Phương pháp nội suy khung hình khuôn mặt theo điểm 9, trong đó bước nội suy bao gồm các bước:

sử dụng mỗi tập hợp gồm ba khung hình lân cận từ các khung hình đầu vào cho phương pháp FOC; và

biến đổi, bởi lớp tích chập biến dạng được, bản đồ đặc trưng của khung hình giữa trong mỗi tập hợp thành bản đồ đặc trưng của khung hình đích sử dụng bản đồ độ dịch được tạo.

12. Phương pháp nội suy khung hình khuôn mặt theo điểm 9, trong đó bước nội suy bao gồm nội suy xuôi và nội suy ngược.

13. Phương pháp nội suy khung hình khuôn mặt theo điểm 9, trong đó bước nội suy là nội suy hình tháp, phân tầng và biến dạng được (PCD) sử dụng phương pháp FOC cho việc tính độ dịch.

14. Phương pháp nội suy khung hình khuôn mặt theo điểm 13, trong đó việc nội suy đơn cấp theo nội suy PCD bao gồm các bước:

sử dụng, ở mỗi cấp không phải thấp nhất, các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của các hình ảnh được lấy mẫu giảm để tính các nhân tích chập biến dạng được;

tinh chỉnh, ở cấp thấp nhất, các nhân tích chập biến dạng được bằng cách hợp nhất các nhân này với sự ước lượng độ dịch được lấy mẫu tăng từ cấp trước đó; và

tính tích chập các nhân biến dạng được với bản đồ đặc trưng của các khung hình lân cận.

15. Phương pháp nội suy khung hình khuôn mặt theo điểm 13, trong đó số lượng cấp là ba.

16. Phương pháp nội suy khung hình khuôn mặt theo điểm 13, trong đó đối với mỗi cấp không phải cao nhất, phương pháp FOC đơn giản hóa sử dụng các đặc trưng của các khung hình làm đầu vào được sử dụng.

17. Phương pháp khử nhòe video khuôn mặt, phương pháp này bao gồm các bước:

trích xuất các bản đồ đặc trưng từ các khung hình đầu vào chứa khung hình đích;

căn chỉnh mỗi trong số các bản đồ đặc trưng được trích xuất với bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo ra bằng cách sử dụng phương pháp FOC theo điểm 1;

hợp nhất các bản đồ đặc trưng được căn chỉnh, bao gồm bản đồ đặc trưng của khung hình đích, để ước lượng bản đồ đặc trưng của khung hình đích được tăng cường;

tái tạo khung hình đích được tăng cường bằng cách sử dụng bản đồ đặc trưng của khung hình đích này;

trích xuất các bản đồ đặc trưng từ các khung hình đầu vào trong đó khung hình đích bị thiếu;

nội suy, từ mỗi tập hợp các bản đồ đặc trưng được trích xuất, bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo ra bằng cách sử dụng phương pháp FOC;

hợp nhất các bản đồ đặc trưng được nội suy để ước lượng bản đồ đặc trưng được kết hợp của khung hình đích được nội suy;

tái tạo khung hình đích được nội suy bằng cách sử dụng bản đồ đặc trưng được kết hợp của khung hình đích này; và

kết hợp khung hình đích được tăng cường với khung hình đích được nội suy.

18. Phương pháp khử nhòe video khuôn mặt theo điểm 17, còn bao gồm bước tiền xử lý để khử nhiễu và khử nhòe một số khung hình đích trước.

19. Vật ghi đọc được bằng máy tính lâu dài chứa các lệnh chương trình, mà bao gồm mô đun tính độ dịch nhận diện khuôn mặt (FOC), được thực thi bởi bộ xử lý hoặc bộ

điều khiển, môđun FOC của các lệnh chương trình này khi được thực thi bởi bộ xử lý hoặc bộ điều khiển được tạo cấu hình để:

nhận các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của tập hợp các khung hình;

thực hiện phép ghép vector của các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc với nhau để tạo ra các đầu vào, trong đó phép ghép vector nối các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc từ đầu đến cuối bằng cách ghép vector của các bản đồ đặc trưng và vector của các bản đồ nhiệt điểm mốc với nhau từ đầu đến cuối để tạo thành vector mới được ghép bao gồm các phần tử của vector của các bản đồ đặc trưng và vector của các bản đồ nhiệt điểm mốc; và

tạo ra bản đồ độ dịch bằng cách thực hiện việc phân tích hồi quy trên các đầu vào,

trong đó bản đồ độ dịch biểu diễn bản đồ của các vector độ dịch 2D chỉ báo các độ lớn mà theo đó các vị trí lấy mẫu của nhân tích chập sẽ bị biến dạng tại mỗi điểm ảnh để biến đổi đặc trưng từ một khung hình sang một khung hình khác trong mạng tích chập biến dạng được (DCN), và

trong đó bản đồ độ dịch được tạo cấu hình để được sử dụng bởi DCN để căn chỉnh và/hoặc nội suy bản đồ đặc trưng khuôn mặt.

20. Thiết bị tăng cường khung hình khuôn mặt, thiết bị này bao gồm:

môđun trích xuất để trích xuất các bản đồ đặc trưng từ các khung hình đầu vào chứa khung hình đích;

môđun căn chỉnh, bao gồm vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC theo điểm 19, để căn chỉnh mỗi trong số các bản đồ đặc trưng được trích xuất với bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo ra bởi vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC;

môđun hợp nhất để hợp nhất các bản đồ đặc trưng được căn chỉnh, bao gồm bản đồ đặc trưng của khung hình đích, để ước lượng bản đồ đặc trưng của khung hình đích được tăng cường; và

môđun tái tạo để tái tạo khung hình đích được tăng cường bằng cách sử dụng bản đồ đặc trưng của khung hình đích này.

21. Thiết bị tăng cường khung hình khuôn mặt theo điểm 20, trong đó các khung hình đầu vào bao gồm năm khung hình liên tiếp.

22. Thiết bị tăng cường khung hình khuôn mặt theo điểm 20, trong đó môđun căn chỉnh được tạo cấu hình để:

thu được các độ dịch căn chỉnh từ mỗi trong số các khung hình lân cận đến khung hình đích với vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC; và

căn chỉnh, bởi lớp tích chập biến dạng được, mỗi trong số các bản đồ đặc trưng được trích xuất của các khung hình lân cận với bản đồ đặc trưng của khung hình đích sử dụng bản đồ độ dịch được tạo.

23. Thiết bị tăng cường khung hình khuôn mặt theo điểm 20, trong đó môđun căn chỉnh có kiến trúc hình tháp, phân tầng và biến dạng được (PCD) sử dụng vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC cho việc tính độ dịch.

24. Thiết bị tăng cường khung hình khuôn mặt theo điểm 23, trong đó môđun căn chỉnh đơn cấp theo kiến trúc PCD được tạo cấu hình để:

sử dụng, ở mỗi cấp không phải thấp nhất, các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của các hình ảnh được lấy mẫu giảm để tính các nhân tích chập biến dạng được;

tinh chỉnh, ở cấp thấp nhất, các nhân tích chập biến dạng được bằng cách hợp nhất các nhân này với sự ước lượng độ dịch được lấy mẫu tăng từ cấp trước đó; và

tính tích chập các nhân biến dạng được với bản đồ đặc trưng của khung hình lân cận.

25. Thiết bị tăng cường khung hình khuôn mặt theo điểm 23, trong đó số lượng cấp là ba.

26. Thiết bị tăng cường khung hình khuôn mặt theo điểm 23, trong đó đối với mỗi cấp không phải cao nhất, môđun FOC đơn giản hóa sử dụng các đặc trưng của các khung hình làm đầu vào được sử dụng.

27. Thiết bị nội suy khung hình khuôn mặt, thiết bị này bao gồm:

môđun trích xuất để trích xuất các bản đồ đặc trưng từ các khung hình đầu vào trong đó khung hình đích bị thiếu;

môđun nội suy, bao gồm vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC theo điểm 19, để nội suy, từ mỗi tập hợp các bản đồ đặc trưng được trích xuất, bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo ra bởi vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC;

môđun hợp nhất để hợp nhất các bản đồ đặc trưng được nội suy để ước lượng bản đồ đặc trưng được kết hợp của khung hình đích được nội suy; và

môđun tái tạo để tái tạo khung hình đích được nội suy bằng cách sử dụng bản đồ đặc trưng được kết hợp của khung hình đích này.

28. Thiết bị nội suy khung hình khuôn mặt theo điểm 27, trong đó các khung hình đầu vào bao gồm bốn khung hình.

29. Thiết bị nội suy khung hình khuôn mặt theo điểm 27, trong đó môđun nội suy được tạo cấu hình để:

sử dụng mỗi tập hợp gồm ba khung hình lân cận từ các khung hình đầu vào cho vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC; và

biến đổi, bởi lớp tích chập biến dạng được, bản đồ đặc trưng của khung hình giữa trong mỗi tập hợp thành bản đồ đặc trưng của khung hình đích sử dụng bản đồ độ dịch được tạo.

30. Thiết bị nội suy khung hình khuôn mặt theo điểm 27, trong đó môđun nội suy được tạo cấu hình để thực hiện nội suy xuôi và nội suy ngược.

31. Thiết bị nội suy khung hình khuôn mặt theo điểm 27, trong đó môđun nội suy có kiến trúc hình tháp, phân tầng và biến dạng được (PCD) sử dụng môđun FOC cho việc tính độ dịch.

32. Thiết bị nội suy khung hình khuôn mặt theo điểm 31, trong đó môđun nội suy đơn cấp theo kiến trúc PCD được tạo cấu hình để:

sử dụng, ở mỗi cấp không phải thấp nhất, các bản đồ đặc trưng và các bản đồ nhiệt điểm mốc của các hình ảnh được lấy mẫu giảm để tính các nhân tích chập biến dạng được;

tinh chỉnh, ở cấp thấp nhất, các nhân tích chập biến dạng được bằng cách hợp nhất các nhân này với sự ước lượng độ dịch được lấy mẫu tăng từ cấp trước đó; và

tính tích chập các nhân biến dạng được với bản đồ đặc trưng của các khung hình lân cận.

33. Thiết bị nội suy khung hình khuôn mặt theo điểm 31, trong đó số lượng cấp là ba.

34. Thiết bị nội suy khung hình khuôn mặt theo điểm 31, trong đó đối với mỗi cấp không phải cao nhất, môđun FOC đơn giản hóa sử dụng các đặc trưng của các khung hình làm đầu vào được sử dụng.

35. Hệ thống khử nhòe video khuôn mặt, hệ thống này bao gồm:

môđun trích xuất để trích xuất các bản đồ đặc trưng từ các khung hình đầu vào chứa khung hình đích;

môđun căn chỉnh, bao gồm vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC theo điểm 19, để căn chỉnh mỗi trong số các bản đồ đặc trưng được trích xuất với bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo ra bởi vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC;

môđun hợp nhất để hợp nhất các bản đồ đặc trưng được căn chỉnh, bao gồm bản đồ đặc trưng của khung hình đích, để ước lượng bản đồ đặc trưng của khung hình đích được tăng cường;

môđun tái tạo để tái tạo khung hình đích được tăng cường bằng cách sử dụng bản đồ đặc trưng của khung hình đích này;

môđun trích xuất để trích xuất các bản đồ đặc trưng từ các khung hình đầu vào trong đó khung hình đích bị thiếu;

môđun nội suy, bao gồm vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC, để nội suy, từ mỗi tập hợp các bản đồ đặc trưng được trích xuất, bản đồ đặc trưng của khung hình đích dựa vào bản đồ độ dịch được tạo ra bởi vật ghi đọc được bằng máy tính lâu dài gồm môđun FOC;

môđun hợp nhất để hợp nhất các bản đồ đặc trưng được nội suy để ước lượng bản đồ đặc trưng được kết hợp của khung hình đích được nội suy;

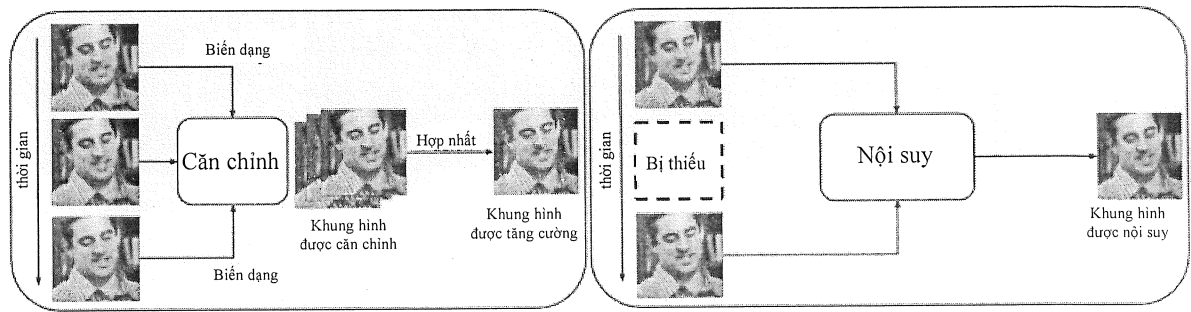
môđun tái tạo thứ hai để tái tạo khung hình đích được nội suy bằng cách sử dụng bản đồ đặc trưng được kết hợp của khung hình đích này; và

thiết bị kết hợp để kết hợp khung hình đích được tăng cường với khung hình đích

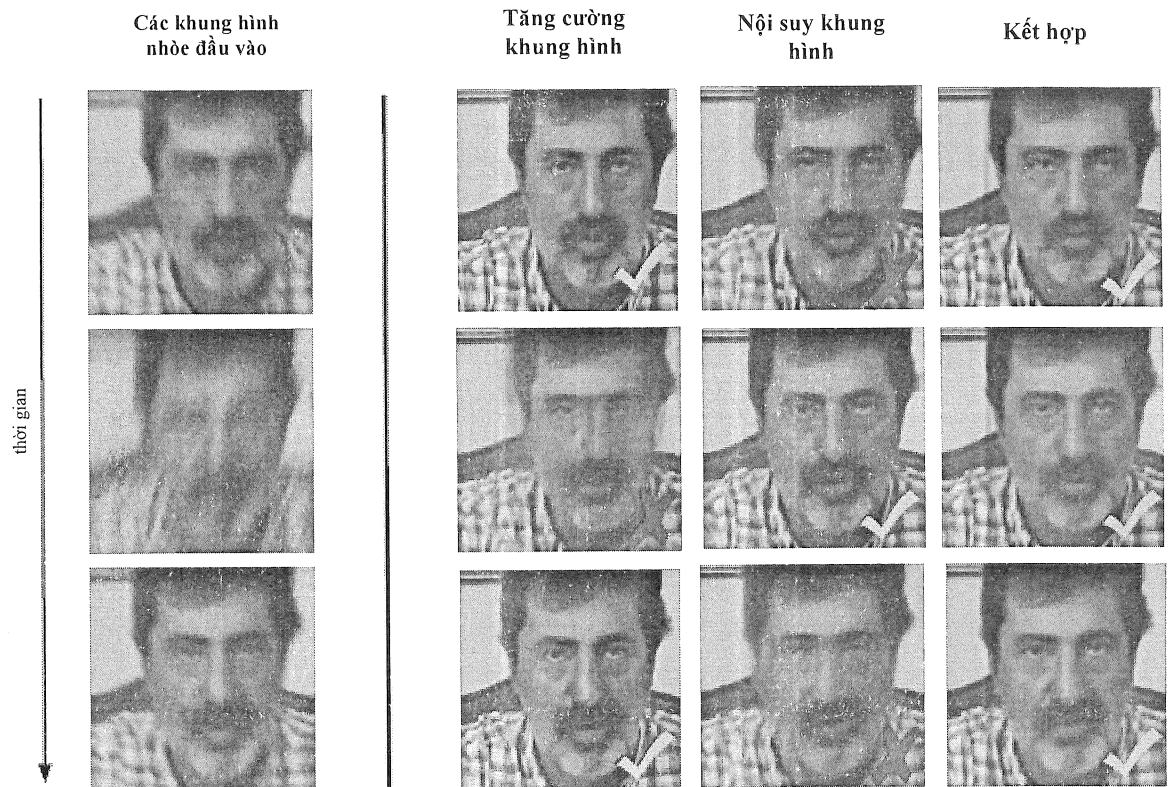
được nội suy.

36. Hệ thống khử nhòe video khuôn mặt theo điểm 35, còn bao gồm thiết bị tiền xử lý để khử nhiễu và khử nhòe một số khung hình đích trước.

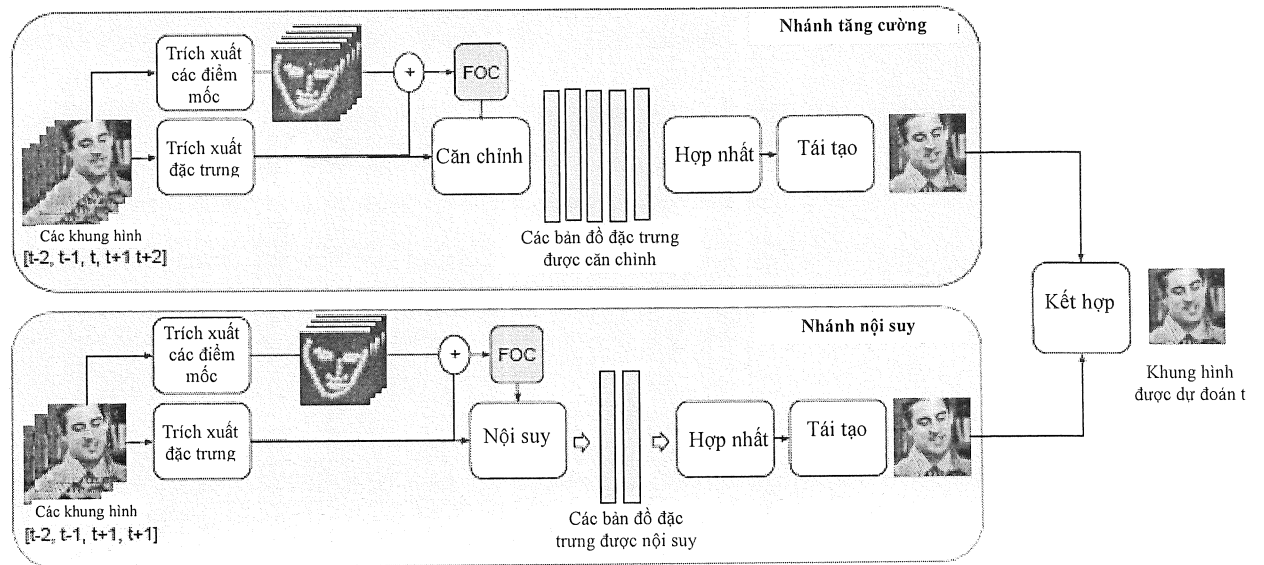
【Fig.1】



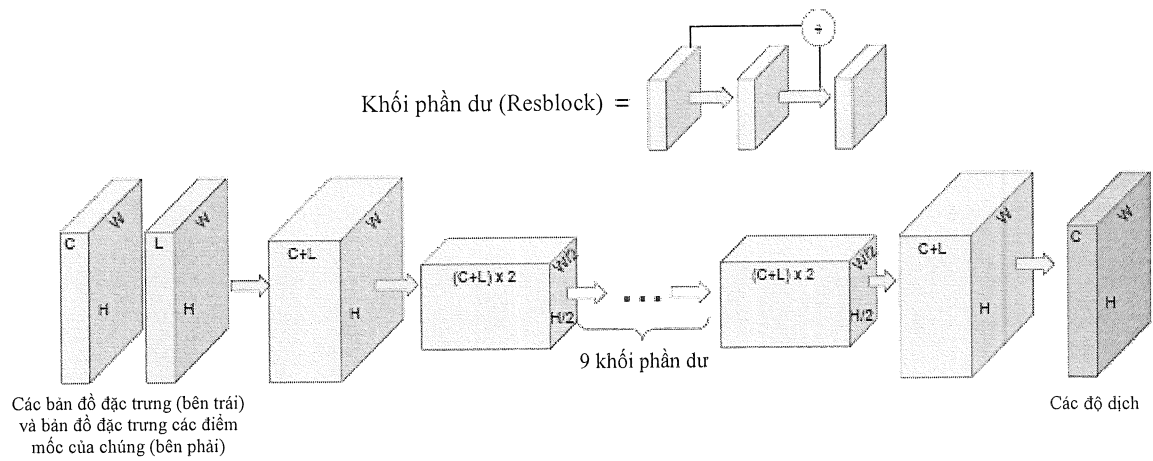
【Fig.2】



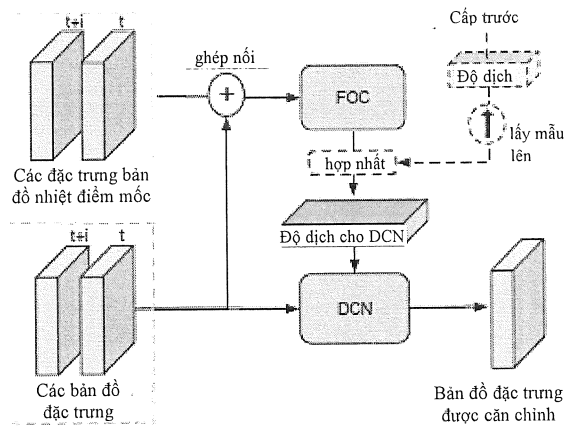
【Fig.3】



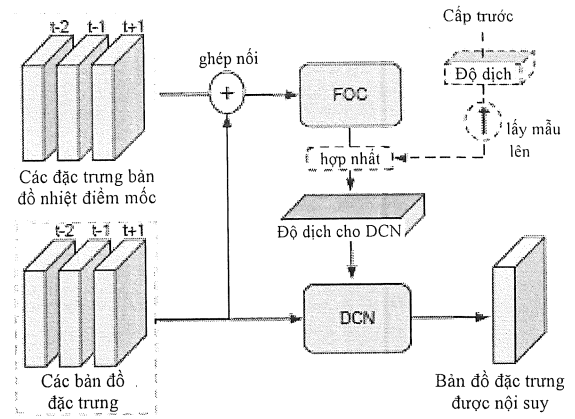
【Fig.4】



【Fig.5】

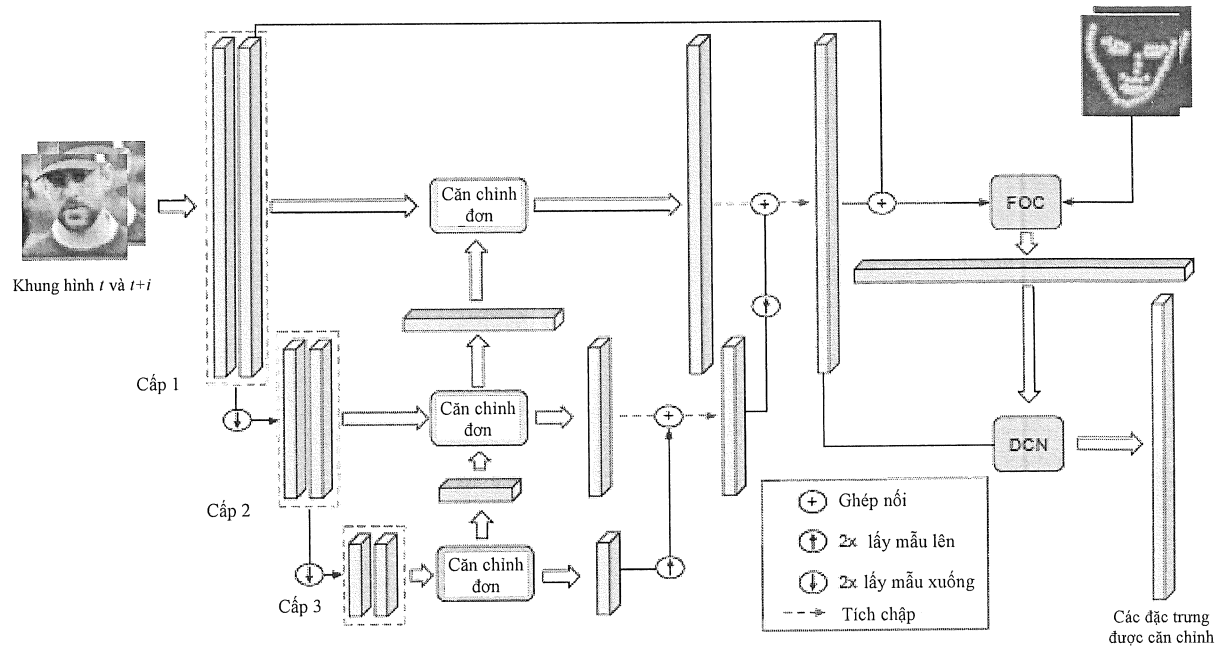


a. Môđun căn chỉnh đơn cấp



b. Môđun nội suy đơn cấp

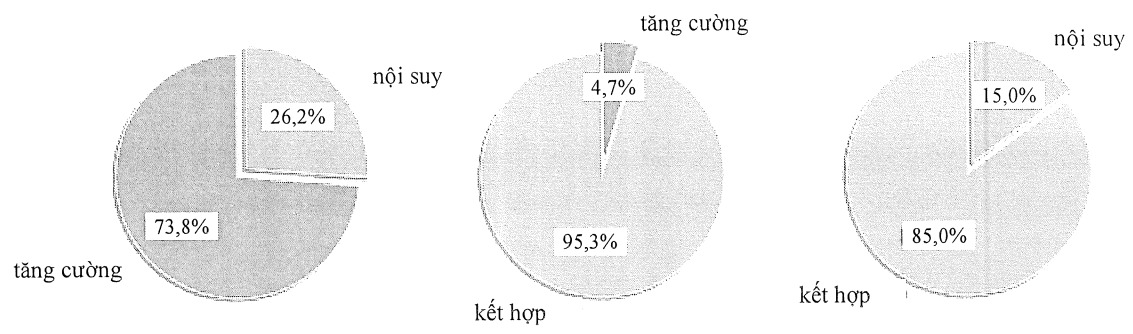
【Fig.6】



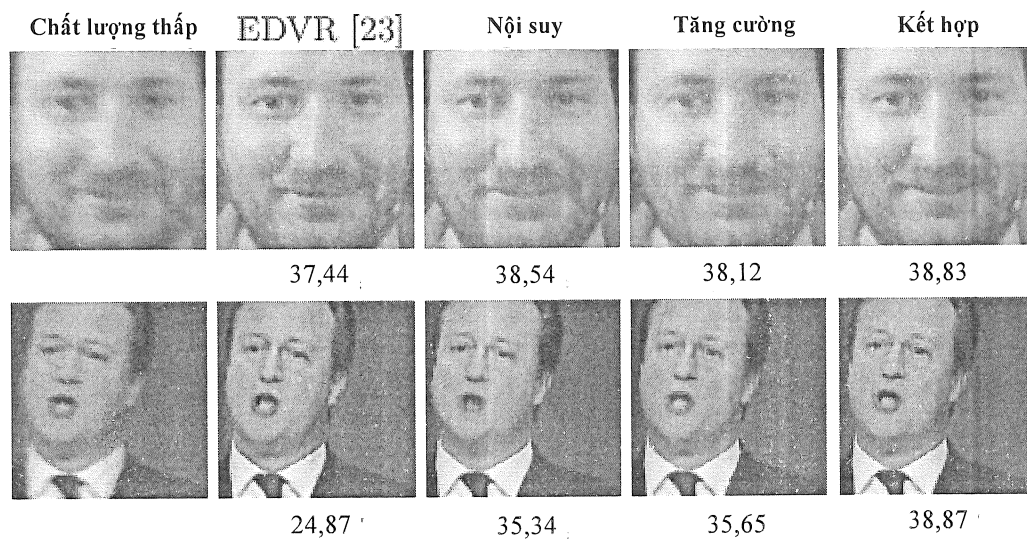
【Fig.7】



【Fig.8】



【Fig.9】



【Fig.10】

