

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

NGUYỄN THỊ HƯƠNG THỦY

**MỘT SỐ GIẢI PHÁP NÂNG CAO HIỆU QUẢ
HỆ THỐNG NHẬN DẠNG VÂN TAY**

Chuyên ngành: Khoa học Máy tính

Mã số: 62.48.01.01

TÓM TẮT LUẬN ÁN TIẾN SĨ CÔNG NGHỆ THÔNG TIN

Hà Nội – 2013

Công trình được hoàn thành tại: Trường Đại học Công nghệ - ĐH Quốc gia Hà Nội.

Người hướng dẫn khoa học:

1. PGS. TS. Hoàng Xuân Huân
2. TS. Nguyễn Ngọc Kỳ

Phản biện 1: PGS. TS. Lương Chi Mai

Viện Công nghệ thông tin, Viện Hàn lâm KH&CN VN

Phản biện 2: PGS. TS. Phan Trung Huy

Trường Đại học Bách khoa Hà Nội

Phản biện 3: PGS. TS. Bùi Thế Duy

Học viện Thanh Thiếu niên Việt Nam

Luận án được bảo vệ trước hội đồng cấp Đại học Quốc Gia chấm luận án tiến sĩ họp tại Trường Đại học Công nghệ vào hồi 9 giờ 00 phút, ngày 08 tháng 01 năm 2014.

Có thể tìm hiểu luận án tại:

- Thư viện Quốc gia Việt Nam
- Trung tâm Thông tin – Thư viện, Đại học Quốc gia Hà Nội

MỞ ĐẦU

1. Tính cấp thiết của luận án

Các quan hệ sôi động và phức tạp trong các lĩnh vực hình sự và thương mại đòi hỏi có các hệ truy nguyên tự động vân tay (AFIS) đáng tin cậy. Mặc dù chủ đề này đã được quan tâm từ lâu nhưng đến nay vẫn còn nhiều vấn đề mở cần nghiên cứu.

Trong bài toán thẩm định (verification) hay bài toán xác thực, ta cần đối sánh một ảnh vân tay đăng nhập với một ảnh vân tay của người đăng ký đã được lưu trữ để xác định xem chúng có đồng nhất, tức là cùng do một ngón sinh ra hay không? Với bài toán truy nguyên (identification), ta có một ảnh truy vấn I_q và cần tìm xem trong cơ sở dữ liệu (CSDL) chỉ bản lưu trữ có ảnh nào đồng nhất với I_q hay không. Như vậy bài toán xác thực là bài toán con của bài toán truy nguyên. Cả hai bài toán này có tên gọi chung là đối sánh vân tay. Nói cách khác, truy nguyên là đối sánh 1:N và thẩm định là đối sánh 1:1. Việc thẩm định có thể là tự động (với các hệ kiểm soát truy cập) hoặc với sự can thiệp thủ công bằng chuyên gia vân tay (với các hệ pháp lý).

Ảnh vân tay bao gồm các loại: *lăn* (rolled), *ấn* (plain) và dấu vết ẩn hay *hiện trường* (latent). Vân tay lăn và vân tay ấn đều có thể thu thập bằng phương pháp in mực hoặc dùng thiết bị thu nhận vân tay sống. Trong các ảnh vân tay này, vân tay sống có chất lượng tốt nhất và các CSDL công khai trên internet thường là của loại này. Chỉ bản giấy thường có chất lượng tồi hơn do nhiễu của vết mực và biến dạng phi tuyến của đầu ngón tay khi lăn/ấn, còn vân tay hiện trường có chất lượng tồi nhất và thường chỉ là một phần của đầu ngón tay.

Một hệ truy nguyên vân tay tự động (Automatic Fingerprint Identification System: AFIS) tốt cần đảm bảo hai yếu tố: Độ chính xác cao và Tốc độ đối sánh nhanh.

Do nhu cầu trong công tác an sinh xã hội rất lớn nên các AFIS tốt được bán rất đắt và kỹ thuật xây dựng chúng được giữ bản quyền hoặc bí mật. Các tài liệu công bố công khai chỉ đủ để xây dựng các hệ thử nghiệm. Đặc biệt, các bài báo về xử lý vân tay thường không công bố chi tiết thuật toán mà chỉ nêu lược đồ phương pháp, CSDL thử nghiệm tin cậy thường bị giữ bản quyền nên khó so sánh thực nghiệm khi nghiên cứu.

Để đáp ứng nhu cầu điều tra và tránh phụ thuộc vào các phần mềm thương mại, Bộ Công an đã thành lập một nhóm nghiên cứu về AFIS và cho ra đời phần mềm C@FRIS, đưa vào ứng dụng thực tế và được nhận giải thưởng Sáng tạo Khoa học - Công nghệ Việt Nam (VIFOTEC) 2008.

C@FRIS là hệ thống dùng để tự động hóa các tàng thư căn cước công dân, căn cước can phạm và tra cứu dấu vết vân tay hiện trường phục vụ công tác quản lý hành chính và điều tra tội phạm. Tuy nhiên, để làm chủ công nghệ và đưa sản phẩm phục vụ tích cực cho hoạt động thực tế thì cần phải tiếp tục nghiên cứu phát triển công nghệ nền, nâng cấp tính năng sản phẩm.

2. Mục tiêu của luận án

Nghiên cứu: “*Một số giải pháp nâng cao hiệu quả hệ thống nhận dạng vân tay*”, trong đó chú trọng các nội dung:

1) *Phân đoạn ảnh* chỉ bản vân tay mười ngón nhằm tự động tách các ảnh vân tay từng ngón từ mẫu chỉ bản 10 ngón và tính bản đồ chất lượng cho từng vân tay, bao gồm cả thông tin liên quan như: điểm dị thường, dạng vân cơ bản, đường biên của đường vân, vùng trung tâm, chiều hướng, ...

2) Nâng cấp thuật toán đối sánh (matching) vân tay cho các loại vân tay có độ biến dạng cao.

3) *Tổ chức dữ liệu* hợp lý để tăng tốc độ xử lý của C@FRIS và *bảo vệ an ninh an toàn* hệ thống, chống xâm nhập bất hợp pháp.

4) Phương pháp truy nguyên vân tay hiện trường.

3. Các đóng góp chính của luận án

1) Đề xuất hai thuật toán phân đoạn thô và phân đoạn mịn. Thuật toán phân đoạn thô nhằm cắt tách riêng các ảnh vân tay từ chỉ bản 10 ngón thành 20 ảnh ngón riêng rẽ dựa trên kỹ thuật tiền xử lý ảnh như chuẩn hóa, làm trơn, chuyển đổi nhị phân và sau đó tiến hành dò biên. Thuật toán phân đoạn mịn nhằm tách vùng vân chất lượng cao khỏi vùng nền và vùng nhiễu trên ảnh vân tay từng ngón để thuận tiện cho việc trích chọn và đánh giá đặc trưng.

2) Để nâng cấp thuật toán nhận dạng các vân tay lấn /ấn có độ biến dạng cao, luận án đề xuất phương pháp đối sánh vân tay nhờ kỹ thuật nắn chỉnh Thin-Plate-Spline (TSP) địa phương và cấu trúc điểm địa phương để khử hiện tượng méo phi tuyến.

3) Đề xuất giải pháp tổ chức CSDL dựa trên sự kết hợp kỹ thuật đánh chỉ số CSDL theo mã ngón, theo dạng vân và đặc điểm ảnh, sắp hạng dữ liệu và song song hóa nhằm tăng tốc độ xử lý thông tin cho quá trình truy nguyên vân tay hiện trường. Đề xuất sử dụng và cài đặt hệ thống BioPKI để bảo vệ hệ C@FRIS, bao gồm các công đoạn: kiểm soát xác thực chủ thể đăng nhập hệ thống, truy cập CSDL, tính năng dùng chữ ký số và xác thực chữ ký, tính năng mã hóa/giải mã trên đường truyền và các quá trình trao đổi dữ liệu.

4) Để cải tiến phương pháp truy nguyên vân tay hiện trường, luận án đề xuất một kiến trúc đa tầng nhằm kết hợp hiệu quả của nhiều phương pháp khác nhau để cho kết quả tổng hợp tốt hơn và rút ngắn thời gian và danh sách tìm kiếm.

Kết quả thử nghiệm trên CSDL C@FRIS DB và trên CSDL FVC2004 cho thấy hiệu quả nổi trội của các giải pháp mới đề xuất so với các phương pháp hiện hành khác.

4. Bố cục của luận án

Ngoài phần kết luận, luận án được tổ chức thành năm chương. Chương 1 giới thiệu những vấn đề cơ bản của hệ nhận dạng vân tay và một số kỹ thuật liên quan cần dùng về sau. Hai thuật toán phân đoạn thô và mịn được trình bày trong

Chương 2. Chương 3 trình bày phương pháp hiệu quả để truy nguyên vân tay biến dạng dựa trên mô hình nắn chỉnh từng phần và cấu trúc điểm địa phương. Chương 4 trình bày giải pháp tổ chức dữ liệu và bảo vệ an ninh an toàn hệ thống. Kiến trúc đa tầng để cải tiến chiến lược truy nguyên vân tay hiện trường được trình bày trong Chương 5.

Chương 1. TỔNG QUAN VỀ HỆ NHẬN DẠNG VÂN TAY

1.1. Bài toán nhận dạng vân tay

1.1.1. Các khái niệm cơ bản

Vân tay là những vết lằn tạo nên các hoa văn trên bề mặt da đầu các ngón tay mà ta quen gọi là các dòng đường vân. Một thẻ mẫu đã được in vân, được gọi là một chỉ bản và thường gồm hai loại: *vân tay lằn* (rolled) và *vân tay ấn* (plain). Vân tay được thu nhận bởi các sensor gọi là *vân tay sống*, dạng này phổ biến ở các nước công nghiệp. Dấu vết vân tay của nghi can để lại hiện trường gọi là *vân tay hiện trường*, dạng vân này chất lượng xấu và không đầy đủ nên rất khó nhận dạng.

Một cấu trúc đường vân lý tưởng bao gồm các dòng *đường vân* và các dòng *đường rãnh* chạy xen kẽ nhau, “song song” với nhau, một đường vân bị kẹp giữa hai đường rãnh và ngược lại, một đường rãnh bị kẹp giữa hai đường vân.

Tùy theo chất lượng mà ảnh vân tay được chia làm 3 miền con: *vùng có cấu trúc rõ ràng*, *vùng bị phá hủy nhưng có thể khôi phục lại được* và *vùng bị phá hủy không thể khôi phục được*.

Đặc trưng của vân tay: Hình dạng các đường vân tay rất phong phú, song vẫn có thể phân loại chúng theo các lớp khác nhau. Vùng *vân trung tâm* dùng để phân loại là vùng vân nằm chính giữa một dấu vân tay được giới hạn bởi *đường bao trên* và *đường bao dưới*. Việc phân loại đường vân giúp rút ngắn thời gian nhận dạng vân.

Điểm gặp nhau của ba dòng vân khác nhau được gọi là *tam phân điểm* (delta), còn điểm mà quanh nó có một dòng vân chạy vòng quanh được gọi là *tâm điểm* (core).

Số đếm vân là số đường vân cắt đoạn thẳng nối hai điểm mốc. Điểm mốc có thể là tâm điểm (core), tam phân điểm (delta) hay điểm đặc trưng chi tiết.

Các dạng cơ bản của vân tay: Căn cứ vào cách sắp xếp chung của các dòng đường vân, có thể phân vân tay thành 3 dạng cơ bản chủ yếu: hình cung, hình quai và hình xoáy.

Đặc điểm chi tiết của vân tay: Một số đường vân đang chạy liên tục rồi đến một vị trí nào đó hoặc bị phân ra hai, ba nhánh (điểm rẽ nhánh) hoặc có khi bị đột ngột kết thúc (điểm cụt) gọi là những đặc điểm chi tiết.

1.1.2. Bài toán nhận dạng vân tay

Cho một cơ sở dữ liệu (hay hồ sơ) gồm các ảnh vân tay lưu trữ và một ảnh vân tay truy vấn, ta cần tìm trong cơ sở dữ liệu này có ảnh vân tay nào cùng một ngón sinh ra với ảnh truy vấn hay không? Nếu có thì chỉ ra ảnh này.

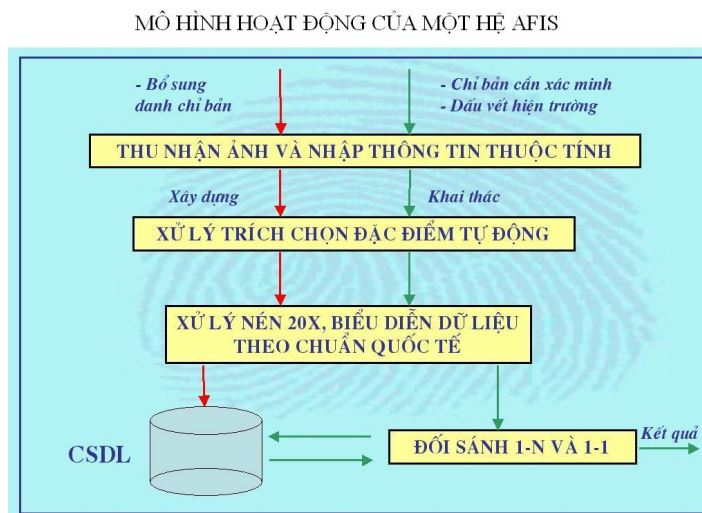
1.1.3. Ứng dụng của bài toán nhận dạng vân tay

Luận án nêu ra ví dụ điển hình ứng dụng nhận dạng vân tay trong các cơ quan pháp lý: xác minh công dân đã được cấp chứng minh nhân dân hay chưa; truy tìm tội phạm và xác định danh tính người bị bắt giữ.

1.2. Hệ thống nhận dạng vân tay tự động

Một hệ thống nhận dạng vân tay tự động thông thường có hai chức năng cơ bản: xây dựng CSDL và tra tìm chỉ bản vân tay. Xây dựng CSDL vân tay tức là thực hiện các công việc: Thu thập chỉ bản của các đối tượng quản lý, quét chỉ bản, nhập thông tin thuộc tính (số đối tượng, họ tên, năm sinh,...), phân loại vân tay theo dạng cơ bản, xử lý trích chọn tự động các đặc điểm chi tiết, lưu và tổ chức CSDL. Tra tìm chỉ bản vân tay là xác định xem đối tượng có chỉ bản vân tay đó đã có trong CSDL hay chưa, số hồ sơ đối tượng đăng ký là bao nhiêu. Khi xây dựng CSDL vân tay cũng như khi tra tìm đều đòi hỏi phải phân loại và trích chọn các điểm đặc trưng chi tiết trước khi tiến hành đối sánh.

Hình 1.1 là sơ đồ khối tổng quát của hệ thống nhận dạng vân tay.



Hình 1.1: Sơ đồ khối tổng quát của hệ nhận dạng vân tay tự động.

1.3. Tình hình nghiên cứu, ứng dụng các hệ nhận dạng vân tay tự động

Công nghệ nhận dạng vân tay tự động nói chung trên thế giới đã được nghiên cứu phát triển từ khá lâu. Ngay từ những năm sáu mươi của thế kỷ trước nhiều hệ thống đã được đưa vào hoạt động tại hàng trăm cơ quan cảnh sát trên thế giới. Phổ biến nhất vẫn là sản phẩm của ba công ty lớn, đó là Morpho của Pháp, NEC của Nhật, và PRINTRAK của Mỹ. Nhiều ứng dụng được triển khai theo nhiều qui mô khác nhau, từ hàng triệu bản ghi (NEWYORK, TOKYO, Cảnh sát CANADA) đến

các hệ qui mô nhỏ hàng nghìn bản ghi. Ngoài các hãng lớn trên, ở nhiều nước, người ta cũng đã tự xây dựng các hệ nhận dạng vân tay tự động dùng riêng cho cơ quan cảnh sát của họ.

Ở nước ta, từ những năm 90, cơ quan cảnh sát cũng đã nhập, và đưa vào sử dụng hệ thống nhận dạng vân tay Morpho AFIS của hãng SAGEM của Pháp. Qua 10 năm sử dụng đã xây dựng và đưa vào khai thác một CSDL qui mô 1,3 triệu chỉ bản, trong khi nhu cầu thực tế ở Trung ương phải xây dựng được CSDL khoảng 3 triệu chỉ bản. Còn tại hầu hết các địa phương, công việc phân loại và tra cứu vân tay tội phạm chủ yếu hiện nay vẫn phải thực hiện bằng phương pháp thủ công. Việc mở rộng hệ thống để ứng dụng cho qui mô cả nước sẽ cần phải đầu tư tiếp rất nhiều ngoại tệ và chắc chắn nếu tiếp tục nhập ngoại sẽ càng ngày càng bị lệ thuộc nhiều hơn vào công nghệ nước ngoài.

Để nâng cao tính chủ động công nghệ, Bộ Công an đã chủ trương và tạo điều kiện cho Phòng Thí nghiệm Mô phỏng và Tích hợp hệ thống phát huy nội lực để tự xây dựng cho Công an Việt Nam một sản phẩm AFIS dùng cho toàn ngành. Được tạo điều kiện về cơ sở vật chất khá thuận lợi, Nhóm tác giả thuộc Phòng Thí nghiệm Mô phỏng và Tích hợp hệ thống đã cho ra đời một sản phẩm AFIS mang tên C@FRIS.

Tuy đã đạt được một số kết quả bước đầu đáng khích lệ nhưng cũng như hầu hết các sản phẩm AFIS khác, việc duy trì các tính năng và không ngừng cải tiến công nghệ để đáp ứng tốt hơn nữa các tiêu chí cơ bản về tốc độ, về độ chính xác, năng suất nhập liệu, tính tiện dụng, tính tương thích, khả năng quản lý CSDL dung lượng qui mô lớn,... là một đòi hỏi có tính sống còn đối với sản phẩm C@FRIS.

Về mặt công nghệ, mặc dù trong những năm gần đây đã có nhiều tiến bộ đáng kể trong việc nâng cao hiệu quả các hệ thống nhận dạng vân tay tự động. Các cuộc thi các thuật toán thẩm định vân tay FVC2000 - FVC2006 do các trung tâm nghiên cứu hàng đầu trên thế giới phối hợp tổ chức hai năm một lần, đã ghi nhận sự tiến bộ và độ chính xác đạt được hiện nay là vào khoảng 96%, trong 4% sai số thì có 2% là tỷ lệ từ chối sai và 2% tỷ lệ chấp nhận sai. Do đó, nhu cầu tiếp tục nghiên cứu nhằm hoàn thiện hơn nữa độ tin cậy, tính ổn định, tính hiệu quả của các hệ thống nhận dạng vân tay vẫn còn rất cấp thiết.

Trên cơ sở phân tích các kết quả nghiên cứu đạt được trong quá trình nghiên cứu và triển khai ứng dụng hệ C@FRIS, cũng như những nội dung cần tiếp tục cải tiến, luận án đề xuất tiếp tục nghiên cứu phát triển các nội dung cụ thể như sau:

- Phân đoạn chỉ bản để tách ảnh và đánh giá vân tay tự động.
- Phát triển các thuật toán đối sánh nhanh và chính xác cho các loại vân tay.
- Tổ chức dữ liệu hợp lý để tăng tốc độ truy nguyên.
- Bảo vệ hệ thống khi hệ được dùng chung trên mạng.
- Hoàn thiện phương pháp truy nguyên vân tay hiện trường trên CSDL qui mô lớn.

Trong các chương tiếp theo, luận án sẽ trình bày các đóng góp của luận án cho các chủ đề này.

Chương 2. THUẬT TOÁN PHÂN ĐOẠN CHỈ BẢN MƯỜI NGÓN

Chương này trình bày thuật toán mới phát triển để phân đoạn vân tay từ ảnh biểu mẫu chỉ bản mười ngón. Kết quả chủ yếu của chương đã được công bố trong hội thảo khoa học và trên tạp chí *Tin học và điều khiển học*.

2.1. Bài toán phân đoạn ảnh vân tay tự động

Phân đoạn được hiểu là công đoạn xử lý tách vùng vân tay cần quan tâm ra khỏi phần nền của ảnh. Với mức độ cao hơn, liên quan đến đánh giá chất lượng vân tay, phân đoạn vân tay theo nghĩa tìm vùng quan tâm chất lượng cao và coi đó như “miền vân tay xác định”. Việc giới hạn lại “miền xác định” của mỗi vân tay còn là thông tin quan trọng giúp cho quá trình đối sánh hiệu quả hơn nhờ xác định được miền “giao chung” của hai vân tay và giản lược bớt các phép so sánh không cần thiết đối với các đặc điểm chi tiết nằm ngoài vùng giao chung này.

2.2.1. Khái niệm phân đoạn ảnh

Phân đoạn ảnh tổng quát:

Phân đoạn ảnh (image segmentation) là một quá trình thực hiện việc tách từ vùng ảnh thành một hay một số vùng các điểm ảnh cùng thỏa mãn một số tính chất chung nào đấy như theo màu sắc, mức xám, kết cấu bề mặt, chiều hướng, Mỗi vùng thường được biểu diễn bằng một tập các điểm ảnh liên thông cùng thỏa mãn một tiêu chí nhận biết. Tiêu chí này phụ thuộc vào mục tiêu cụ thể của từng loại ứng dụng. Sau khi phân đoạn mỗi điểm ảnh chỉ thuộc về một trong hai lớp: vùng quan tâm và vùng không quan tâm.

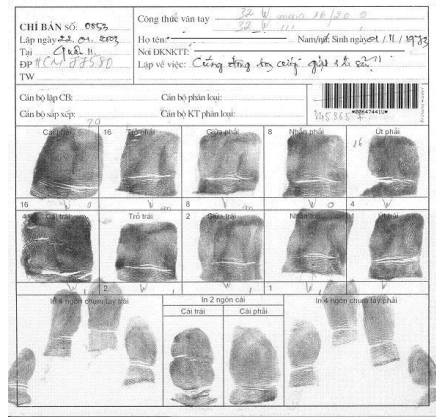
Một cách tổng quát, có thể coi phân đoạn là bài toán nhận dạng các điểm (pixel) ảnh để phân thành hai lớp: lớp quan tâm và lớp không quan tâm, trong đó các tập mẫu học có thể cho trước (học có thầy) hoặc không cho trước (học không có thầy hay học tự động).

Phân đoạn ảnh vân tay:

Phân đoạn ảnh là quyết định phần ảnh nào thuộc vùng cần quan tâm, vùng ảnh nào là vùng nền bao quanh và vùng nào là biên hay nhiễu của ảnh.

Phân đoạn ảnh chỉ bản vân tay 10 ngón mà chương này đề cập là một trường hợp ứng dụng cụ thể của kỹ thuật phân đoạn ảnh. Nhiệm vụ của nó là phát hiện và tách vùng đường vân của các đầu ngón tay ra khỏi vùng nền trên các ô của chỉ bản 10 ngón.

2.1.2. Bài toán phân đoạn ảnh chỉ bản vân tay tự động



Hình 2.1: Mẫu chỉ bản vân tay mười ngón

Cho một chỉ bản vân tay 10 ngón như Hình 2.1. Hãy xây dựng thuật toán tự động định vị và tách từ chỉ bản vân tay 10 ngón thành 20 vùng riêng rẽ như sau: 10 vùng ảnh cho 10 ngón lăn, 10 vùng cho 10 ngón ấn (gồm 2 vùng nhỏ cho 2 ngón cái phải, cái trái và 4 vùng lưu 4 ngón chụm trái và 4 vùng cho 4 ngón chụm phải).

2.1.3. Một số thuật toán phân đoạn

Mục này, luận án phân tích, đánh giá một số thuật toán phân đoạn của các tác giả khác đề xuất liên quan đến chủ đề quan tâm hiện nay.

2.2. Đề xuất thuật toán xử lý phân đoạn ảnh chỉ bản

2.2.1. Thuật toán phân đoạn thô

Bước 1: Xử lý chuẩn hoá và làm trơn ảnh.

1.1. Xử lý chuẩn hoá mức xám ảnh tức là thay đổi mức xám ban đầu $I(x,y)$ trong từng cửa sổ ảnh bằng mức xám mới $N(x,y)$ như sau:

$$N(x,y) = \begin{cases} M_0 + \sqrt{\frac{V_0 * (I(x,y) - M)^2}{V}}, & \text{cho } I(x,y) > M \\ M_0 - \sqrt{\frac{V_0 * (I(x,y) - M)^2}{V}}, & \text{cho } I(x,y) \leq M \end{cases} \quad (2.1)$$

Trong đó, $I(x,y)$ là mức xám ban đầu tại điểm (x,y) , $M = (1/w^2) \sum_w (I)$, $V = (1/w^2) \sum_w (I - M)^2$ lần lượt là mức xám trung bình và phương sai mức xám ước lượng trên cửa sổ kích thước $w \times w$ (phần thực nghiệm chọn $w=32$), M_0 , V_0 lần lượt là mức xám trung bình và phương sai mức xám cần chuẩn hoá, $N(x,y)$ là mức xám mới sau chuẩn hoá. (M_0 , V_0 có thể chọn tùy ý, nhưng phần thực nghiệm chọn $M_0=V_0=100$). (Cần lưu ý rằng: Nếu 1 cửa sổ vân tay bất kỳ có $M=M_0$, $V=V_0$ thì $N(x,y)=I(x,y)$).

1.2. Dùng cửa sổ $w \times w$ điểm ảnh quét lần lượt từ trái qua phải từ trên xuống dưới để thay thế giá trị mức xám ảnh hiện tại bằng giá trị trung bình của các điểm ảnh trên cửa sổ. Chuyển đổi ảnh về dạng nhị phân, dùng ngưỡng là giá trị mức xám trung bình địa phương, tính trên cửa sổ $w \times w$.

1.3. Tiến hành dò biên bằng thuật toán dò biên 8-liên thông và xử lý xấp xỉ tuyến tính từng đoạn để biểu diễn các đường biên thành những đa tuyến biên khép kín. Để tăng hiệu quả thuật toán dò biên phát hiện vùng vân, cần chọn bước lưới phù hợp để lọc lấy các vùng vân, loại bớt các vùng nhiễu nhỏ. Phần thực nghiệm ta sẽ giả định vùng vân bé nhất có kích thước 8x8mm và vùng vân lớn nhất có kích thước 20x20mm nên bước lưới thích hợp được chọn là 8mm.

Bước 2: Xử lý kết quả dò biên, xác định đa giác ngoại tiếp vùng vân cần cắt.

2.1. Tính bao lồi của tập đỉnh đa tuyến biên vùng được chọn bằng thuật toán Graham. Sau đó tính hình chữ nhật ngoại tiếp bao lồi để xác định cửa sổ cắt ảnh.

2.2. Kiểm tra cửa sổ ảnh với các điều kiện ràng buộc: diện tích tối thiểu, diện tích tối đa, vị trí tương đối của khung hình so với mẫu chỉ bản.

2.3. Kiểm tra thứ tự ngón: Nội dung của việc kiểm tra thứ tự ngón là dùng thuật toán đối sánh vân tay đã đề xuất trong Chương 3 để tiến hành đối sánh từng vân tay 10 ngón vùng dưới với từng vân tay 10 ngón vùng giữa chỉ bản để đưa ra kết luận vị trí các ngón được lãn có đúng thứ tự biểu mẫu qui định hay không.

2.4. Tính các giá trị X_{min} , Y_{min} , X_{max} , Y_{max} của các đỉnh bao lồi đường biên, tức là hình chữ nhật ngoại tiếp để định vị tọa độ khung cắt.

2.5. Hiển thị và xử lý tương tác các trường hợp đặc biệt như các chỉ bản lãn thiếu ngón, lãn sai quy cách không thoả mãn điều kiện thẩm định.

Bước 3: Hiển thị kết quả định vị khung cắt để thẩm định bằng mắt thường và chuyển sang công đoạn tương tác.

2.2.1. Thuật toán phân đoạn mịn

Ngoài các thuộc tính truyền thống như Mean, Variance ta đề xuất bổ sung thêm ba chỉ tiêu mới để nhận dạng các khối vùng vân chất lượng thấp: mật độ D, tổng độ cong C và năng lượng E.

Mật độ D là mật độ các đỉnh của các đoạn ngắn của các đa tuyến đường vân đi qua khối $w \times w$, được đặc trưng bằng số lượng đỉnh trong khối đó.

Độ cong tại một đỉnh đa tuyến được đo bằng giá trị góc đổi hướng của đa tuyến tại đỉnh đó. Độ cong C của khối bằng tổng độ cong của các đỉnh trong khối đó.

Năng lượng tại một đỉnh đa tuyến được đo bằng giá trị chênh lệch độ cong của đa tuyến tại đỉnh đó so với đỉnh ngay trước nó. Năng lượng E của khối bằng tổng năng lượng của các đỉnh trong khối đó.

Giả sử ta có đa tuyến $\{P\}_1^N = \{(x_1, y_1, \alpha_1), (x_2, y_2, \alpha_2), \dots, (x_n, y_n, \alpha_n)\}$, khi đó độ cong tại đỉnh P_i và độ cong của khối $w \times w$ được định nghĩa như sau:

$$C(P_i) = (\alpha_i - \alpha_{i-1}), \quad i=2..N. \quad (2.2)$$

$$C_{w \times w} = \sum_{w \times w} C(P_i). \quad (2.3)$$

Và năng lượng tại đỉnh P_i và năng lượng của khối $w \times w$ được định nghĩa như sau:

$$E(P_i) = C(P_i) - C(P_{i-1}), \quad i=2..N. \quad (2.4)$$

$$E_{w_{xxw}} = \sum_{w_{xxw}} E(P_i) . \quad (2.5)$$

Mỗi khối w_{xxw} được biểu diễn như sau:

$$X_{w_{xxw}} = (M, V, D, C, E) \quad (2.6)$$

Để phân loại từng cửa sổ w_{xxw} thuộc lớp ω_1 “vùng chất lượng cao” hay lớp ω_2 “vùng chất lượng thấp và nền”, ta dùng Bộ phân loại Bayes dùng qui tắc quyết định tối ưu với hàm phân biệt sau:

$$g(\omega_i) = P(\omega_i).P(X/\omega_i), i=1,2. \quad (2.7)$$

Qui tắc phân lớp cụ thể như sau:

$$\text{IF } (g(\omega_1) > g(\omega_2)) \text{ THEN } (X_{w_{xxw}} \in \omega_1) \text{ ELSE } (X_{w_{xxw}} \in \omega_2) \quad (2.8)$$

Ở đây, $P(\omega_2) = 1 - P(\omega_1)$; $P(\omega_1)$ là tần suất xuất hiện của số vân tay chất lượng cao và $P(\omega_2)$ là tần suất xuất hiện của số vân tay chất lượng thấp.

Công thức ước lượng các phân bố xác suất này như sau:

$$P(\omega_1) = (\text{Tổng các } w_{xxw} \text{ chất lượng cao}) / (\text{tổng các } w_{xxw} \text{ toàn vân tay}) \quad (2.9)$$

$P(X/\omega_i)$ = (Tần suất xuất hiện vector thuộc tính X trên lớp ω_i , $i = 1,2$) được ước lượng như sau:

$$P(X/\omega_i) = (\text{Tổng các } w_{xxw} \text{ có } X_{w_{xxw}} = X) / \text{tổng số } w_{xxw} \text{ thuộc } \omega_i, i = 1,2 \quad (2.10)$$

Từ đó ta có thuật toán phân đoạn mịn được đề xuất như sau:

Bước 1: Tiếp nhận ảnh vân tay đầu vào, là kết quả xuất ra từ công đoạn thô.

Tiếp nhận ảnh đầu vào là ảnh vân tay ngón tay cùng các thông tin phụ trợ liên quan như đường bao, vùng trung tâm, chiều hướng (trên/dưới).

Bước 2: Chuẩn hóa, làm trơn và chuyển đổi nhị phân.

Dùng ngưỡng địa phương và chuyển đổi ảnh về dạng nhị phân.

Bước 3: Dò biên, vector hoá ảnh nhị phân và xấp xỉ tuyến tính từng đoạn.

Xử lý dò biên đường vân, xấp xỉ tuyến tính từng đoạn kết hợp tính độ cong, năng lượng trên từng đỉnh đa tuyến, và tính các thuộc tính D, C, E.

Bước 4: Phân loại các khối pixel w_{xxw} .

4.1. Dùng cửa sổ kích thước phù hợp $w_{xxw} = 32 \times 32$ pixels để duyệt ảnh gốc và tính các giá trị M, V để xây dựng vector biểu diễn: $X_{w_{xxw}} = (M, V, D, C, E)$.

4.2. Dùng qui tắc quyết định tối ưu Bayes để xếp $X_{w_{xxw}}$ vào lớp tương ứng.

4.3. Kết quả cho ra mặt nạ chất lượng, là dạng ảnh nhị phân có giá trị 1 tại khối vùng vân chất lượng cao và giá trị 0 tại nền và vùng vân chất lượng thấp.


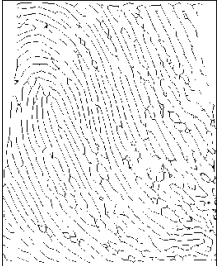
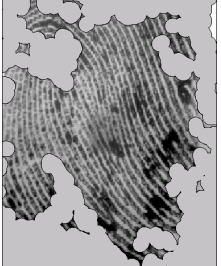

2.3. Kết quả thực nghiệm

Kết quả thử nghiệm thuật toán phân đoạn thô:

Thuật toán phân đoạn thô được đánh giá trên 1000 mẫu chỉ bản của CSDL C@FRIS.


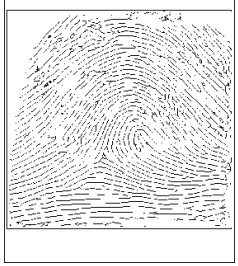
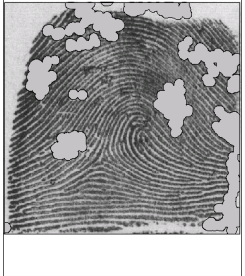

- Số lượng chỉ bản phân đoạn tốt, không cần sự can thiệp của con người: 973/1000.
- Số chỉ bản đúng qui cách nhưng thuật toán trả lại không phân đoạn được: 0/1000.
- Số chỉ bản sai qui cách phải trả lại: 27/1000, trong đó 22 chỉ bản dùng giấy nền quá tối, 5 chỉ bản có các vân tay lẫn chồng lên nhau, chồng lên khung.

Kết quả thử nghiệm thuật toán phân đoạn mịn:

			
(a) Ảnh gốc chất lượng thấp	(b) Kết quả phân đoạn theo M, V, coherence và vector hóa ảnh gốc	(c) Kết quả phân đoạn của thuật toán đề xuất dùng (M, V, D, C, E)	(d) Kết quả phân đoạn của thuật toán Verifinger

Hình 2.2: Minh họa kết quả phân đoạn ảnh chỉ bản chất lượng thấp, chọn từ CSDL FVC2004

Thuật toán phân đoạn mịn được đánh giá so sánh với kết quả phân đoạn của phần mềm Verifinger trên 500 chỉ bản CSDL FVC2004 DB2 xem Hình 2.2 và 500 chỉ bản từ CSDL C@FRIS DB xem Hình 2.3 được cắt ra từ giai đoạn phân đoạn thô.

			
(a) Ảnh gốc chất lượng thấp	(b) Kết quả phân đoạn theo M, V, coherence và vector hóa ảnh gốc	(c) Kết quả phân đoạn của thuật toán đề xuất dùng (M, V, D, C, E)	(d) Kết quả phân đoạn của thuật toán Verifinger

Hình 2.3: Minh họa kết quả phân đoạn ảnh chỉ bản chất lượng thấp, chọn từ CSDL C@FRIS

2.4. Nhận xét chung

Thuật toán phân đoạn thô để phân đoạn chỉ bản 10 ngón trên CSDL chỉ bản giấy tiêu biểu, kết quả đạt được độ chính xác cao, tốc độ xử lý nhanh, số chỉ bản phải trả lại để can thiệp thủ công rất ít, chiếm tỷ lệ chỉ dưới 2,7 %.

Giai đoạn phân đoạn mịn khi so sánh kết quả phân đoạn với phần mềm Verifinger, kết quả phân đoạn của thuật toán đề xuất đưa ra là phù hợp với kết quả phân đoạn của Verifinger. Những vùng thuật toán đề xuất đánh giá chất lượng cao thì phần mềm Verifinger cũng đánh giá cao. Các ảnh vân tay phân đoạn của

Verifinger rõ nét hơn chủ yếu là nhờ khâu tiền xử lý ảnh trước khi tiến hành vector hóa không thuộc phạm vi xử lý của thuật toán phân đoạn.

Chương 3. PHƯƠNG PHÁP ĐỐI SÁNH VÂN TAY DỰA TRÊN MÔ HÌNH NẮN CHỈNH ĐỊA PHƯƠNG

Chương này đề xuất giải pháp mới để cải tiến thuật toán đối sánh vân tay ấn và vân tay lăn dựa trên việc kết hợp khử các hiện tượng méo phi tuyến. Kết quả của chương đã được công bố trong hội nghị Quốc tế *The International Conference on Computing, Management and Telecommunications (ComManTel2013)*.

3.1. Bài toán đối sánh vân tay và một số vấn đề liên quan

Mục này, luận án giới thiệu tóm tắt phương pháp đối sánh ĐTCT, mô hình nắn chỉnh TPS.

3.1.1. Đối sánh vân tay và lược đồ đối sánh dựa trên đặc trưng chi tiết (ĐTCT)

Bài toán đối sánh vân tay có thể phát biểu như sau: Cho trước hai ảnh vân tay truy vấn (query) I_q và vân tay mẫu (template) I_t , cần trả lời xem hai ảnh này có phải là do cùng một ngón tay in ra hay không? Câu trả lời trước hết được thể hiện qua kết quả tính điểm độ giống (similarity) giữa hai vân tay.

Các điểm ĐTCT (Minutiae)

Để biểu diễn vân tay dưới dạng tập các điểm ĐTCT, các ảnh vân tay I_q và I_t trước hết phải được xử lý qua các khâu tiền xử lý và trích chọn, đánh giá đặc trưng. Ký hiệu M_q và M_t tương ứng là tập ĐTCT của hai ảnh vân tay I_q, I_t :

$$M_q = \{m_1, m_2, \dots, m_M\}; \text{ với } m_i = (x_i, y_i, \theta_i), i = 1, \dots, M. \quad (3.1)$$

trong đó (x_i, y_i) là tọa độ của m_i trên mặt phẳng ảnh R^2 của I_q , còn θ_i là hướng vân tại điểm m_i ;

$$M_t = \{m_1', m_2', \dots, m_N'\}; \text{ với } m_i' = (x_i', y_i', \theta_i'), i = 1-N. \quad (3.2)$$

trong đó (x_i', y_i') là tọa độ của m_i' trên mặt phẳng ảnh R^2 của I_t , còn θ_i' là hướng vân tại điểm m_i' .

Lược đồ đối sánh dựa trên ĐTCT

Người ta sẽ tìm phép biến đổi “chồng ảnh” thích hợp từ mặt phẳng ảnh của I_q vào mặt phẳng ảnh của I_t để xác định các *cặp điểm tương ứng* của M_q và M_t . Cặp điểm $m_i(M_q)$ và $m_j'(M_t)$ gọi là *tương ứng nếu ảnh m_i* của m_i qua phép biến đổi này thuộc vào lân cận bán kính r đủ bé của m_j' , trong trường hợp này ta nói m_i “*trùng khớp*” với m_j' . Khi hai ảnh I_q và I_t là cùng do một ngón cụ thể in ra (genuine: *chính danh*), thì thường phát hiện được nhiều cặp điểm ĐTCT tương ứng hơn so với trường hợp chúng là do các ngón khác nhau in ra (impostor: *giả danh*). Trên thực tế, do hiện tượng biến dạng phi tuyến và nhiễu ảnh nên khi hai ảnh là chính danh thì cũng khó tìm được để các điểm ĐTCT tương ứng từng đôi một.

Giả sử ta tìm được n cặp điểm tương ứng của hai ảnh, n_q và n_t tương ứng là số điểm ĐTCT trong mỗi ảnh I_q và I_t , độ giống nhau của hai ảnh vân tay được đặc trưng bằng độ đo $S(I_t, I_q)$ được cho bởi công thức:

$$S(I_t, I_q) = n^2 / n_t \times n_q. \quad (3.3)$$

Với các giá trị ngưỡng S_{max} và S_{min} chọn trước, nếu $S(I_t, I_q) \geq S_{max}$ ta kết luận hai vân tay này là trùng khớp và là ảnh của cùng một ngón, còn nếu $S(I_t, I_q) < S_{min}$ thì chúng không trùng khớp với xác suất sai gần bằng không. Đối với cặp vân tay có độ giống $S(I_t, I_q)$ nằm trong khoảng $[S_{min}, S_{max}]$ quyết định được đưa ra dựa theo ngưỡng $S_t \in [S_{min}, S_{max}]$ với xác suất sai loại I (FAR) và sai loại II (FRR). Vì vậy, để cải tiến giải thuật, hạn chế sai sót, những cặp vân tay có độ giống nằm trong khoảng này cần được tiếp tục nắn chỉnh biến dạng để cải thiện độ giống.

3.1.2. Mô hình nắn chỉnh TPS

Mục này giới thiệu mô hình nắn chỉnh TPS và phương pháp đối sánh TPS toàn cục (sẽ được ký hiệu là G-TPS).

3.2. Đề xuất phương pháp nắn chỉnh từng phần

3.2.1. Cấu trúc vân rãnh liên thuộc và tạo sinh các điểm giả ĐTCT

Cấu trúc vân rãnh liên thuộc

Vân rãnh liên thuộc được cấu thành từ đường vân liên thuộc với ĐTCT đó và đường rãnh liên thuộc với ĐTCT đối ngẫu với ĐTCT đó.

Tạo sinh các điểm trên cặp vân rãnh liên thuộc và bổ sung các cặp điểm tương ứng:

Ký hiệu l_0 là độ dài (thường được chọn ~ 2 bước vân) và $l_{max} = 4l_0$. Tập các điểm giả ĐTCT được tạo sinh trên cấu trúc vân- rãnh liên thuộc của các cặp điểm tương ứng như sau: Với mỗi điểm ĐTCT, ta dò từ điểm này trên ảnh vân và từ điểm ĐTCT đối ngẫu trên ảnh rãnh để xác định các điểm lượng hóa cách đều nhau với bước l_0 cho đến khi gặp điểm biên hoặc độ dài vượt quá ngưỡng l_{max} thì dừng. Các điểm này gọi là giả ĐTCT.

3.2.2. Phân miền đối sánh TPS và chọn các cặp điểm không chế

Để thực hiện đối sánh TPS từng phần ta cần phân ảnh thành 9 miền con và chọn các điểm không chế trên cho biến đổi TPS của mỗi miền.

3.2.3. Mô tả thuật toán nắn chỉnh địa phương (P-TPS)

Với hai ảnh vân tay I_q, I_t ; các tập đặc trưng chi tiết lần lượt là M_q, M_t ; Bằng thuật toán đối sánh cơ bản sử dụng phép biến đổi affine ta thu được các cặp điểm tương ứng M_1 ban đầu và các cặp vân rãnh liên thuộc M_1^* của chúng, thuật toán P-TPS được mô tả như sau:

Dữ liệu vào: Tập ĐTCT M_b, M_q ; Số ĐTCT n_b, n_q ; Tập các cặp ĐTCT tương ứng ban đầu M_1 với n cặp, tập n^* các cặp ĐTCT tương ứng và vân rãnh liên thuộc M_1^* . Ngưỡng nắn chỉnh s_{min}, s_{max} ; Bước lưới: *step*;

Dữ liệu ra: Số điểm tương ứng sau nắn chỉnh n' và độ giống tổng thể của hai vân tay s .

1. Sàng lọc tập các cặp điểm ĐTCT tương ứng theo cặp vân rãnh liên thuộc: (n là số cặp ĐTCT ban đầu, $n' \leq n$ là kết quả sàng lọc theo cặp vân rãnh liên thuộc).
2. Tính độ giống $s = n'^2 / (n_t * n_q)$. Nếu $s \notin [s_{min}, s_{max}]$, chuyển sang 6;
3. Tăng dày tập các cặp điểm tương ứng bằng cách bổ sung thêm các điểm lượng hoá trên cặp vân-rãnh liên thuộc.
 - 3.1. Bổ sung các cặp điểm lượng hoá tương ứng trên các cặp vân rãnh liên thuộc vào tập M_l , kết quả là tập M_1^* . ($n^* \geq n'$ là số lượng các cặp ĐTCT tương ứng mở rộng).
 - 3.2. Dùng thuật toán tính bao lồi của M_1^* .
4. Chọn lọc tập ĐTCT tương ứng n^* mở rộng:
 - 4.1. Xây dựng lưới sàng tuyển vuông với bước lưới *step* và chia lưới từ góc toạ độ được chọn là điểm giữa đường kính lớn nhất nối hai đỉnh của bao lồi.
 - 4.2. Chọn tập m điểm đại diện trên m ô lưới.
 - 4.3. Phân hoạch 9 vùng, gồm 1 vùng trung tâm với tâm là góc toạ độ, giới hạn bởi bao lồi và 8 vùng phần tám chia bởi 8 tia quét từ tâm theo góc 45° nằm ngoài bao lồi.
 - 4.4. Nắn chỉnh từng phần trên từng phân hoạch và tính số cặp điểm tương ứng n' mới, M_1 mới, độ giống mới: $s = n'^2 / (n_t * n_q)$.
5. Nếu ($n' = n$) hoặc ($s > s_{max}$) chuyển sang 6 nếu không:
 - 5.1. $n \leftarrow n'$;
 - 5.2. Quay lại 2;
6. Kết thúc.

3.3. Kết quả thực nghiệm

Để so sánh hiệu quả giữa hai phương pháp P-TPS và G-TPS của Li là phương pháp tốt hiện nay, luận án đề xuất kịch bản thử nghiệm cài đặt cả hai thuật toán và so sánh hiệu quả đối sánh.

Mặc dù kết quả thực nghiệm của luận án tiến hành trên CSDL chỉ bản C@FRIS DB của Công an Hà Nội và hệ C@FRIS cho thấy phương pháp mới tốt hơn nhiều so với phương pháp G-TPS, song để khách quan, mục này trình bày kết quả thử nghiệm của hai thuật toán trên CSDL FVC2004 (DB1, DB3). Với từng phép đối sánh, cả hai thuật toán P-TPS và G-TPS đều sử dụng chung tập các cặp ĐTCT ban đầu được xác định nhờ thuật toán Verifinger 4.2 (VF).

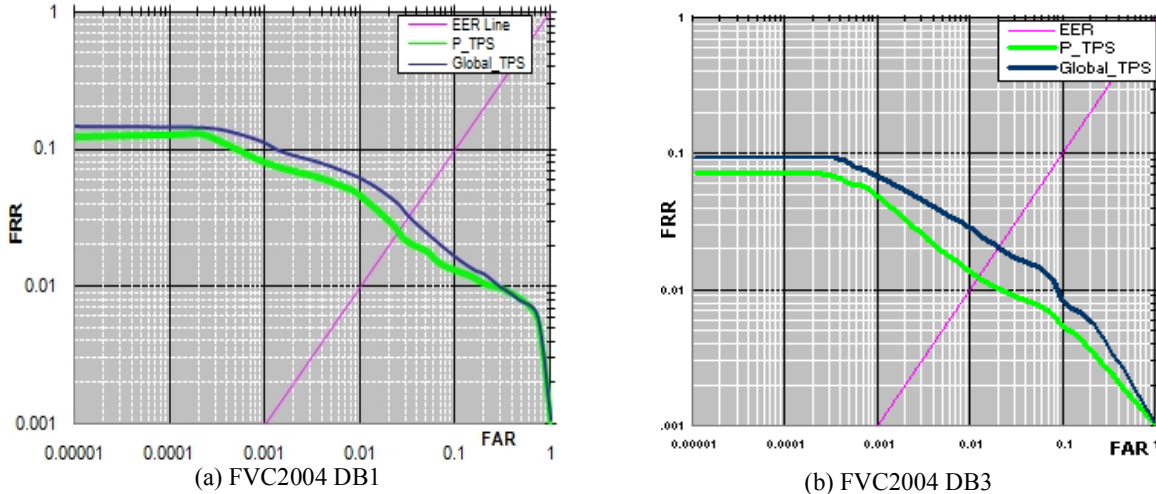
Bảng 3.1: So sánh độ chính xác nắn chỉnh trên CSDL FVC 2004 DB

FVC2004 DB	EER(%)			FAR100(%)			FAR1000(%)			ZeroFAR(%)		
	VF	G-TPS	P-TPS	VF	G-TPS	P-TPS	VF	G-TPS	P-TPS	VF	G-TPS	P-TPS
DB1	3.91	3.24	2.48	7.11	5.89	4.68	12.43	11.47	7.76	17.96	14.70	12.13
DB3	4.03	1.97	1.25	7.43	3.16	2.22	12.64	7.59	5.47	14.32	9.18	7.19
Trung bình	3.97	2.61	1.87	7.27	4.26	3.45	12.54	9.53	6.62	16.14	11.94	9.66

Ghi chú: FAR100 (phần trăm FRR nhỏ nhất khi FAR≤1%), FAR1000 (phần trăm FRR nhỏ nhất khi FAR≤0.1%), ZeroFAR (phần trăm FRR nhỏ nhất khi FAR=0%).

Bảng 3.2: So sánh thời gian và bộ nhớ của hai phương pháp trên CSDL FVC 2004 DB

FVC2004 DB	Chính danh				Giả danh			
	Thời gian (ms)		Bộ nhớ (Kbytes)		Thời gian (ms)		Bộ nhớ (Kbytes)	
	G-TPS	P-TPS	G-TPS	P-TPS	G-TPS	P-TPS	G-TPS	P-TPS
DB1	83.4	78.33	307.56	1.79	83.37	47.06	306.66	0.91
DB3	60.7	46.65	144.46	2.35	53.25	34.18	142.11	1.08
Trung bình	72.05	62.49	226.02	2.07	68.31	40.62	224.39	0.95



Hình 3.1: So sánh đường ROC của phương pháp G-TPS và phương pháp P-TPS trên CSDL FVC2004 DB1, DB3.

Qua kết quả thử nghiệm so sánh hai phương pháp nắn chỉnh được thể hiện trong Bảng 3.1, hai đường cong ROC trên CSDL DB1, DB3 thể hiện trên Hình 3.1(a), Hình 3.1(b) và Bảng 3.2 cho thấy phương pháp nắn chỉnh P-TPS đạt độ chính xác cao hơn trên tất cả các tham đánh giá so với phương pháp nắn chỉnh G-TPS. Cụ thể về độ chính xác, trung bình EER của P-TPS là 1.87% thấp hơn phương pháp G-TPS 2.61%. Về thời gian đối sánh trung bình của phương pháp P-TPS là 62.49ms cho đối sánh chính danh và 40.62 ms cho đối sánh giả danh giảm hơn so với phương pháp G-TPS là 72.05ms cho đối sánh chính danh và 68.31ms cho đối sánh giả danh. Về bộ nhớ sử dụng trong quá trình thực hiện chương trình, phương pháp P-TPS chỉ sử dụng 2.07 Kbytes cho đối sánh chính danh và 0.95 Kbytes cho đối sánh giả danh ít hơn rất nhiều hơn so với phương pháp G-TPS là 226.02 Kbytes cho đối sánh chính danh và 224.39 Kbytes cho đối sánh giả danh thực hiện cùng một máy Intel Pentium 4, 2.8 GHz. Sự khác biệt trong so sánh bộ nhớ là do phương pháp G-TPS đòi hỏi phải dùng file ảnh vân tay gốc để tính các hệ số tương quan mức xám địa phương và tương quan mức xám dọc cạnh đến 2 láng giềng gần nhất của ĐTCT. Trong khi đó thuật toán đề xuất đã tính trước tập các cặp vân rãnh liên thuộc và lưu sẵn dưới dạng tập ĐTCT mở rộng của vân tay bên cạnh bộ ĐTCT chuẩn.

3.4. Nhận xét chung

Luận án đã đề xuất một phương pháp mới để đối sánh vân tay biến dạng dùng mô hình nắn chỉnh TPS từng phần sử dụng các điểm giả ĐTCT trên các cặp vân rãnh liên thuộc tìm được như nguồn thông tin cục bộ có độ tin cậy cao mà không đòi hỏi tính toán phức tạp. Sáng kiến chọn ra tập điểm không chế đại diện vừa phản ánh tốt hơn các biến dạng địa phương vừa phân bố đều hơn trên mặt phẳng ảnh nên cho kết quả nắn chỉnh ảnh tốt hơn khi có biến dạng phi tuyến. Thay cho phương pháp nắn chỉnh toàn phần, phương pháp nắn chỉnh từng phần được đề xuất cũng phản ánh chính xác hơn các thông tin biến dạng địa phương nên cho lời chính xác hơn.

Kết quả thực nghiệm trên CSDL FVC2004 đã chứng tỏ thuật toán đề xuất đạt được hiệu quả rõ rệt, vừa chính xác hơn và nhanh hơn thuật toán khử biến dạng dùng kỹ thuật nắn chỉnh toàn phần kết hợp tương quan mức xám.

Chương 4. TỔ CHỨC DỮ LIỆU VÀ BẢO MẬT HỆ THỐNG

Chương này đề xuất một số giải pháp tổ chức dữ liệu thuận tiện cho tra cứu và bảo vệ an ninh, ngăn ngừa các phương thức tấn công phá hoại hệ thống. Kết quả của chương đã được công bố trong hội thảo khoa học, kỷ yếu *Một số vấn đề chọn lọc của Công nghệ thông tin và Truyền thông và Chuyên san tạp chí Thông tin, Khoa học công nghệ của Bộ Thông tin và Truyền thông*.

4.1. Tổ chức dữ liệu theo hướng song song hóa

4.1.1. Bài toán đối sánh vân tay cao tốc

Bài toán đối sánh vân tay cao tốc, được phát biểu vắn tắt như sau:

Cho hệ thống CSDL ảnh vân tay lưu trữ dung lượng lớn, gồm hàng triệu chỉ bản 10 ngón (chẳng hạn C@FRIS DB), hãy xây dựng giải pháp truy nguyên một dấu vân tay trên CSDL đó sao cho thời gian cho ra kết quả trong khoảng thời gian xác định, chẳng hạn trong phạm vi dưới 1 phút.

4.1.2. Mô hình xử lý song song trong nhận dạng vân tay

Giới thiệu hai mô hình có thể sử dụng trong hệ thống nhận dạng vân tay tự động:

- Mô hình song song hóa thuật toán đối sánh hai vân tay, có thể tiến hành song song hóa các pha: cải thiện ảnh, trích chọn đặc điểm, phân loại ảnh cũng như pha đối sánh. Trên thực tế, do thời gian đối sánh 2 vân tay theo 2 bộ mã của chúng đã rất nhanh, chỉ trong vòng 2-3 giây, nên mô hình song song hoá giải thuật đối sánh 2 vân tay theo 2 bộ ĐTCT có rất ít ứng dụng. Như vậy việc song song hóa thuật toán ít hứa hẹn đem lại hiệu quả tốt nên luận án tập trung vào song song hóa dữ liệu.

- Mô hình song song hoá quá trình truy nguyên theo gói nhiều yêu cầu, áp dụng cho bài toán nhận dạng vân tay dựa trên việc phân hoạch CSDL. CSDL tổng thể được lưu trên máy chủ, được phân hoạch thành các CSDL con đều nhau.

4.2. Đề xuất giải pháp tổ chức CSDL hiệu quả cho đối sánh truy nguyên vân tay theo từng yêu cầu

Ý tưởng chủ đạo của giải pháp đề xuất dựa trên hai công đoạn là: công đoạn tổ chức đánh chỉ số phân cấp theo các thuộc tính vân tay cơ bản để rút ngắn thời gian tra tìm theo nhóm phân loại và công đoạn thứ hai là phân chia danh sách cần đối sánh theo đặc điểm chi tiết thành các gói nhỏ, phân công nhiệm vụ tính toán cho các nút đối sánh song song, giám sát quá trình tính toán và đưa ra danh sách kết quả cuối cùng để rút ngắn thời gian đối sánh theo đặc điểm chi tiết.

Tổ chức đánh chỉ số phân cấp CSDL theo các thuộc tính ảnh vân tay cơ bản:

Trong ứng dụng cho các hệ căn cước công dân và căn cước can phạm, mỗi vân ngón tay được biểu diễn dưới dạng một bản ghi gồm các trường cơ bản sau:

- Số căn cước đối tượng quản lý;
- Các trường về thông tin nhân thân (họ tên, năm sinh, giới tính, ...);
- Mã số ngón;
- Dạng vân tay cơ bản;
- Số đếm vân trái, số đếm vân giữa, số đếm vân phải, mật độ đường vân;
- Bộ đặc điểm chi tiết (khoảng 30 bytes);
- Ảnh vân tay (độ phân giải chuẩn 500 dpi, khoảng 5 MB cho cả bộ 10 ngón).

Để đánh chỉ số phục vụ tra tìm chỉ bản 10 ngón với chỉ bản 10 ngón, giải pháp đề xuất dùng kỹ thuật đánh chỉ số phân cấp theo tổ hợp dạng cơ bản 10 ngón do hệ thống tự động trích chọn và trong từng tổ hợp, tiếp tục đánh chỉ số theo các số đếm vân và mật độ đường vân của các ngón. Trong ứng dụng tra cứu dấu vết hiện trường, giải pháp đề xuất dùng các trường dạng cơ bản, các số đếm vân và mật độ vân để đánh chỉ số theo từng ngón.

Phân chia danh sách cần đối sánh theo đặc điểm chi tiết thành các gói nhỏ:

Để thực thi ý tưởng, luận án dùng kỹ thuật phân nhóm và chia danh sách tra cứu thành nhiều gói, với số lượng đồng đều trong mỗi gói để giao cho mỗi nút hay nhóm nút xử lý song song tiến hành xử lý theo nguyên tắc cân bằng năng lực: Nhóm/nút xử lý nhanh nhận được nhiều, Nhóm/nút xử lý chậm nhận được ít. Cân bằng nhiệm vụ theo năng lực là sự phân phối yêu cầu xử lý giữa các nút xử lý sao cho không xảy ra hiện tượng chờ đợi, tránh tình trạng có một số nút thì không làm hết công suất, còn một số nút khác thì quá tải.

Để xây dựng giải pháp đối sánh song song trên cụm máy tính ta tổ chức hệ thống cụm máy tính theo các chức năng như sau:

Máy chủ tiếp nhận yêu cầu và tra cứu theo thuộc tính cơ bản, phân chia danh sách thành các gói nhỏ và phân phối nhiệm vụ:

Máy chủ tiếp nhận đồng thời nhiều yêu cầu tra cứu từ các máy trạm gửi đến và tra cứu yêu cầu theo các thuộc tính ảnh cơ bản thường được thể hiện dưới dạng một câu truy vấn SQL, phân gói và phân phối các chỉ bản thỏa mãn yêu cầu tra cứu một cách hiệu quả cho các máy trạm.

Các nút xử lý song song nhận nhiệm vụ và tiến hành đối sánh:

Các nút được cài đặt phần mềm đối sánh theo đặc điểm chi tiết để làm nhiệm vụ xử lý các yêu cầu do máy chủ tra cứu điều phối. Các nút thành viên được cân bằng nhiệm vụ theo cách mỗi thành viên tự báo trạng thái: chờ nhận, lỗi,...

Để phân phối việc, máy chủ tra cứu tiến hành điểm danh để biết những nút nào tham gia xử lý công việc và tùy theo trạng thái để tiến hành giao việc hay hủy việc giao cho nút khác theo từng gói nhỏ công việc. Các nút tham gia xử lý, cùng phối hợp tra cứu yêu cầu và nhanh chóng đưa ra danh sách kết quả cuối cùng.

Các máy trạm nhận kết quả tra cứu từ các nút và trả về để làm nhiệm vụ thẩm định:

Sau khi tra cứu các yêu cầu, kết quả trả lời các yêu cầu từ các nút xử lý được lưu trên máy chủ tiếp nhận, được sắp xếp lại theo thứ tự mã ngôn, mã dạng văn cơ bản, số đếm vân, độ giống của chỉ bản tìm thấy so với chỉ bản tra cứu. Các máy trạm tiếp nhận kết quả trả lời từ máy chủ và tiến hành thẩm định.

4.3. Giải pháp bảo vệ an ninh an toàn hệ thống

4.3.1. Giải pháp bảo vệ truy cập mạng dựa trên BioPKI

Giới thiệu mô hình, các thành phần của hệ thống BioPKI và giải pháp bảo mật truy cập từ xa trên nền BioPKI-KC để bảo vệ các truy xuất vào một CSDL qua mạng.

4.3.2. Bài toán bảo vệ hệ thống nhận dạng vân tay C@FRIS qua môi trường mạng

Hệ C@FRIS đã triển khai cài đặt đầy đủ các tính năng từ khâu thu nhận, đăng ký chỉ bản thông tin đầu vào để xây dựng CSDL đến khâu kiểm tra chất lượng dữ liệu, tổ chức dữ liệu đến khâu tra cứu, khai thác hệ thống. Nhiệm vụ đặt ra là tổ chức thiết kế và cài đặt bổ sung cho hệ C@FRIS các tính năng bảo mật dùng công nghệ BioPKI.

4.3.3. Một số yêu cầu bảo vệ đối với hệ nhận dạng vân tay tự động

Để triển khai các ứng dụng trên, hệ thống cần đáp ứng được hai yêu cầu: vừa xử lý nhanh chóng, đảm bảo yêu cầu nghiệp vụ hành chính vừa phải đảm bảo an ninh an toàn cho hệ thống.

4.4. Đề xuất giải pháp bảo vệ hệ thống nhận dạng vân tay C@FRIS

4.4.1. Bảo mật phân hệ “Nhập chuyên đổi số hóa chỉ bản”

Tất cả các giao tác của hệ thống và của nhân viên nhập chuyên đổi thông tin số hóa đều được ghi vào CSDL nhật ký hệ thống. Bản thân CSDL này được bảo mật như “hộp đen” của hệ thống và chỉ người được cấp thẩm quyền mới truy cập được.

4.4.2. Bảo vệ phân hệ “Biên tập và kiểm tra chất lượng”

Bộ đặc điểm chi tiết được biên tập lại thì người có thẩm quyền biên tập là người ký (dùng chữ ký số vân tay) chịu trách nhiệm phần biên tập. Sau khi biên tập và ký lưu, hệ thống tiến hành ký xác nhận công đoạn này. Các bảng dữ liệu sau kiểm tra chất lượng được coi là hoàn chỉnh, cũng được hệ thống ký xác nhận để đảm bảo tính toàn vẹn dữ liệu.

4.4.3. Bảo vệ phân hệ “Tổ chức cơ sở dữ liệu”

Tất cả các giao tác của quản trị viên đều được ký sinh trắc bởi quản trị viên và hệ thống tự động lưu vào CSDL nhật ký hệ thống.

4.4.4. Bảo vệ phân hệ “Tra tìm, đối sánh”

Người sử dụng với vai trò tra cứu, đối sánh được yêu cầu ký xác nhận lập yêu cầu tra cứu, xác nhận việc nhận kết quả tra cứu. Hệ thống ký xác nhận đã tiếp nhận yêu cầu, đã tra cứu và cung cấp kết quả.

4.5. Kết quả thực nghiệm

Giải pháp tổ chức dữ liệu phục vụ phương pháp truy nguyên vân tay tự động:

So sánh các kết quả thử nghiệm hệ thống C@FRIS trước và sau khi đề xuất giải pháp tổ chức dữ liệu song song trên cấu hình 5 máy trạm Pentium IV, 2,4 GHz, 1GB RAM nối mạng Gigabit để làm các nút xử lý tra cứu song song các kết quả thử nghiệm trình bày trong Bảng 4.1.

Bảng 4.1: Đánh giá thời gian tra cứu trước và sau khi bổ sung giải pháp đối sánh song song hóa

Tên CSDL	Số bản ghi trong CSDL	Thời gian tra cứu tuần tự	Thời gian tra cứu song song (5 máy)
Căn cước can phạm	250.000	3 phút/1 yêu cầu	45 giây/1 yêu cầu
Hiện trường	2.500.000	47 phút/1 yêu cầu	9,5 phút/1 yêu cầu
Lồng bàn tay	205.195	8 phút/1 yêu cầu	1,5 phút/1 yêu cầu

Giải pháp bảo vệ an ninh an toàn hệ thống:

So sánh các kết quả thử nghiệm hệ thống C@FRIS trước và sau khi tích hợp giải pháp bảo mật BioPKI. Các kết quả so sánh trong các trường hợp thử nghiệm được trình bày trong các Bảng 4.2, Bảng 4.3, Bảng 4.4, Bảng 4.5 dưới đây.

Bảng 4.2: Bảng đánh giá so sánh các tính năng đạt được của phân hệ “Tạo lập CSDL”

STT	Các tính năng	Hệ C@FRIS cũ	Hệ C@FRIS mới
1	Kiểm soát thẩm quyền tạo lập CSDL, quét nhập, mã hoá và cập nhật chỉ bản 10 ngón	Dùng password	Dùng chữ ký số vân tay
2	Kiểm soát thẩm quyền tạo lập CSDL	Không có	Ký lên file CSDL
3	Lưu bảo mật ảnh và kiểm tra tính hợp lệ (thứ tự) ảnh các ngón tay trên chỉ bản.	Không có	Ký lên ảnh nén
4	Nhập thông tin thuộc tính, dạng cơ bản.	Không có	Ký lên trường tương ứng.

5	Duyệt browser CSDL	Không có	Xác thực thẩm quyền sử dụng xác thực chứng thư số
6	Biên tập CSDL	Không có	Xác thực thẩm quyền sử dụng xác thực chứng thư số
7	Nhập/ xuất CSDL hợp chuẩn ANSI/NIST của các hệ AFIS khác	Không có	Xác thực thẩm quyền sử dụng xác thực chứng thư số

Bảng 4.3: Bảng đánh giá so sánh các tính năng đạt được của phân hệ “Biên tập và kiểm tra chất lượng”.

STT	Các tính năng	Hệ C@FRIS cũ	Hệ C@FRIS mới
1	Kiểm soát thẩm quyền biên tập	Không có	Có
2	Theo dõi và quy trách nhiệm	Không có	Có
3	Bảo mật bộ đặc điểm chi tiết	Không có	Có
4	Bảo đảm tính toàn vẹn dữ liệu	Không có	Có
5	Bảo mật đường truyền trong quá trình truyền/ nhận dữ liệu	Không có	Có
6	Phục hồi bản ghi bị xoá gần nhất	Không có	Có

Bảng 4.4: Bảng đánh giá so sánh các tính năng đạt được của phân hệ “Tổ chức cơ sở dữ liệu”

STT	Các tính năng	Hệ C@FRIS cũ	Hệ C@FRIS mới
1	Kiểm soát thẩm quyền	Không có	Sử dụng BioPKI
2	Bảo mật các bảng dẫn xuất	Không có	Có
3	Bảo đảm tính toàn vẹn dữ liệu	Không có	Có
4	Bảo mật đường truyền trong quá trình truyền/ nhận dữ liệu	Không có	Có

Bảng 4.5: Bảng đánh giá so sánh các tính năng đạt được của phân hệ “Tra tìm, đối sánh”

STT	Các tính năng	Hệ C@FRIS cũ	Hệ C@FRIS mới
1	Kiểm soát thẩm quyền đăng nhập, tạo lập CSDL quản lý YC	Không có	Dùng vân tay và chữ ký số
2	Lưu bảo mật ảnh và kiểm tra tính hợp lệ (thứ tự) ảnh các ngón tay trên chỉ bản vào CSDL YC.	Không có	Ký lên ảnh nén
3	Nhập và bảo mật đường truyền trong quá trình truyền/ nhận dữ liệu	Không có	Có
4	Nhập thông tin thuộc tính, dạng cơ bản,... cho CSDL YC	Không có	Ký lên các trường tương ứng.
5	Duyệt browser CSDL YC	Không có	Xác thực thẩm quyền sử dụng chứng thư số
6	Biên tập CSDL YC	Không có	Xác thực thẩm quyền sử dụng chứng thư số
7	Gửi YC tra cứu	Không có	Ký vào bản ghi YC
8	Nhận và phân phối YC tra cứu	Không có	Có

4.5. Nhận xét chung

Trên đây đề xuất phương pháp truy nguyên song song hóa bao gồm phương pháp tổ chức CSDL dung lượng lớn theo hướng ứng dụng các kỹ thuật đánh chỉ số CSDL ảnh dựa trên các thuộc tính vân tay cơ bản và phương pháp phân phối động các gói dữ liệu đối sánh cho các nút xử lý trên cụm máy tính. Ưu điểm căn bản của giải pháp đề xuất không chỉ giúp rút ngắn thời gian tìm kiếm mà còn cung cấp một giải pháp rất linh hoạt, rất dễ thay đổi cấu hình, dễ nâng cấp, mở rộng cho phù hợp với yêu cầu

ứng dụng thực tế. Giải pháp đề xuất đã có đủ cơ sở lý luận và thực tiễn để giải quyết bài toán ứng dụng qui mô lớn hơn là điện tử hóa tàng thư chứng minh nhân dân với qui mô từ hàng triệu đến hàng chục triệu chỉ bản.

Ngoài ra, chương này cũng đề xuất tích hợp giải pháp bảo mật kiểm soát truy cập CSDL qua mạng dựa trên hệ thống BioPKI để triển khai thử nghiệm các tính năng đề xuất cho hệ C@FRIS. Kết quả thử nghiệm đạt được nhiều triển vọng ứng dụng trong thực tế. Nhờ ứng dụng các tính năng của giải pháp BioPKI, việc bảo vệ hệ C@FRIS sẽ được đảm bảo chặt chẽ mà vẫn giữ được tính dễ dùng trong các khâu xây dựng, khai thác và vận hành hệ thống trên thực tế.

Chương 5. KIẾN TRÚC ĐA TẦNG CHO TRUY NGUYÊN VÂN TAY HIỆN TRƯỜNG

Chương này đề xuất giải pháp mới để đối sánh hiệu quả vân tay hiện trường chất lượng kém, thường chỉ xuất hiện một phần không đầy đủ và có độ biến dạng phức tạp. Kết quả của chương đã được công bố trong hội nghị Quốc tế *The 10th IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF 2013)*.

5.1. Hệ truy nguyên vân tay hiện trường và một số vấn đề liên quan

Mục này giới thiệu tóm tắt về hệ truy nguyên vân tay hiện trường và một số nội dung cần dùng về sau như đoán nhận ngón, phân loại vân tay.

5.1.1. Hệ truy nguyên vân tay hiện trường

Giới thiệu khái niệm truy nguyên vân tay và hệ truy nguyên vân tay hiện trường.

5.1.2. Đoán nhận ngón tay dựa trên cơ sở dấu vết

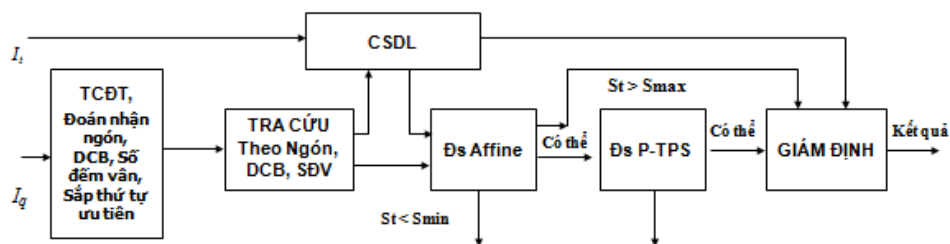
Giới thiệu phương pháp đoán nhận ngón tay do tác giả N.N. Kỷ đề xuất sử dụng phương pháp thống kê theo quy tắc quyết định xác suất hậu nghiệm cực đại để đoán nhận ngón dựa vào dấu vết vân tay hiện trường.

5.1.3. Phân loại vân tay

Giới thiệu các loại vân tay cơ bản và khảo sát, đánh giá một số cách tiếp cận sử dụng để phân lớp vân tay hiện nay.

5.2. Đề xuất kiến trúc bậc thang cho hệ nhận dạng vân tay hiện trường

5.2.1. Các thành phần và sơ đồ bậc thang của hệ nhận dạng vân tay hiện trường



Hình 5.1: Sơ đồ kiến trúc hệ truy nguyên vân tay hiện trường.

Hình 5.1 mô tả sơ đồ bậc thang của hệ nhận dạng vân tay hiện trường.

Quá trình truy nguyên trong hệ tuần tự thực hiện theo các bước sau:

Bước 1: Sau khi thu nhận, ảnh vân tay điều tra I_q được tiền xử lý, phân loại, tính số đếm vân và đoán nhận ngón để quyết định thứ tự tìm kiếm theo ngón trong CSDL lưu trữ (đã được tổ chức dữ liệu dùng phương pháp đã trình bày tại Chương 4). Do vân tay hiện trường có chất lượng xấu, khó trích chọn tự động nên cần thiết phải có sự can thiệp của con người để chỉnh sửa, biên tập tương tác với sự trợ giúp của một bộ biên tập đồ họa. Theo thứ tự mã ngón này, các mã biểu diễn của ảnh vân tay (template) I_t được phân loại và trích chọn tự động từ trước (ở giai đoạn xây dựng CSDL) sẽ được lấy ra từ CSDL để đối sánh (ở mô đun Đối sánh).

Bước 2: Ở mô đun phân loại, dạng vân cơ bản của I_q được xác định tự động hoặc tương tác, nếu có nhập nhầm thì các dạng nhập nhầm này đều được dùng để tìm kiếm đối sánh theo thứ tự ưu tiên của độ tin cậy. Sau đó I_q được đưa vào mô đun trích chọn ĐTCT (TCĐTCT).

Bước 3: Trong mô đun TCĐTCT, mã biểu diễn của I_q gồm các điểm ĐTCT và cấu trúc vân-rãnh liên thuộc (như đã trình bày chi tiết ở Chương 3) cùng bản đồ chất lượng được xác định bằng biện pháp biên tập tương tác để đưa sang mô đun Đối sánh. Nếu chất lượng I_q quá xấu, không thể trích chọn đủ ĐTCT thì dừng (thông báo I_q không đủ điều kiện kết luận trên danh sách kết quả).

Bước 4: Trong mô đun ĐS-affine, các mã biểu diễn của vân tay I_t ở CSDL lần lượt được lấy ra để đối sánh với mã biểu diễn của I_q (thu được ở Bước 3) để tính tập ĐTCT tương ứng ban đầu và độ giống $S(I_t, I_q)$.

4.1. Nếu $S(I_t, I_q) < S_{\min}$ thì bỏ qua, không bổ sung vào danh sách đầu ra, quá trình đọc tiếp mẫu I_t tiếp theo.

4.2. Nếu $S(I_t, I_q) > S_{\max}$ thì bổ sung I_t vào danh sách kết quả. Trường hợp còn lại, tức là $S(I_t, I_q) \in [S_{\min}, S_{\max}]$ thì trao tập các cặp ĐTCT tương ứng cùng các cặp vân rãnh liên thuộc của chúng cho bước đối sánh P-TPS.

Bước 5: Đối sánh P-TPS nhưng chỉ nắn chỉnh một lần để tiết kiệm thời gian xử lý và tính lại độ giống tổng thể như đã trình bày chi tiết ở Chương 3. Quá trình tiếp tục cho đến khi tất cả các mã biểu diễn trong CSDL có mã ngón phù hợp với mã ngón của I_q được đối sánh hết.

Bước 6: Sắp xếp lại danh sách kết quả tìm kiếm theo độ ưu tiên: Mã ngón, Mã dạng vân cơ bản, số đếm vân, độ giống. Thảm định theo thứ tự ưu tiên và thông báo kết quả.

5.2.2. Tổ chức dữ liệu

Để tìm kiếm và đối sánh nhanh, dữ liệu lưu trữ cần được xử lý và tổ chức hợp lý. Công đoạn thứ nhất là tổ chức đánh chỉ số phân cấp theo các thuộc tính mã ngón và dạng vân cơ bản để rút ngắn thời gian tra tìm theo nhóm phân loại và công đoạn thứ hai là phân chia danh sách cần đối sánh theo ĐTCT thành các gói

nhỏ, phân công nhiệm vụ tính toán cho các nút đối sánh song song, giám sát quá trình tính toán và đưa ra danh sách kết quả cuối cùng để rút ngắn thời gian đối sánh theo ĐTCT. Các công đoạn tổ chức dữ liệu, đã trình bày chi tiết trong mục 4.2 của Chương 4.

5.2.3. Giải pháp đối sánh song song

Để xây dựng giải pháp đối sánh song song cho cụm máy tính ta tổ chức hệ thống cụm máy tính theo các chức năng như đã mô tả chi tiết trong Chương 4:

1) Máy chủ tiếp nhận yêu cầu và tìm kiếm theo thuộc tính cơ bản, phân chia danh sách thành các gói nhỏ và phân phối nhiệm vụ cho các nút xử lý song song.

2) Các nút xử lý song song nhận nhiệm vụ và tiến hành đối sánh, trả lại kết quả dưới dạng danh sách kết quả tìm kiếm cho máy chủ.

3) Các máy trạm nhận kết quả tìm kiếm từ máy chủ trả về để làm nhiệm vụ thẩm định.

5.3. Kết quả thực nghiệm

Thực nghiệm được tiến hành trên hệ xử lý nối tiếp có song song hóa khâu tra cứu đối sánh theo ĐTCT bởi 5 máy kết nối song song. CSDL C@FRIS DB dung lượng lớn gồm 2.500.000 chỉ bản vân tay 1 ngón được dùng là CSDL lưu trữ. Thử nghiệm đã tiến hành truy nguyên cho 64 vân tay hiện trường chọn ngẫu nhiên nhưng đảm bảo có chỉ bản đăng ký trên CSDL C@FRIS DB 2.500.000 chỉ bản vân tay 1 ngón và so sánh với hệ C@FRIS phiên bản 2005 để đánh giá hiệu quả.

Kết quả thực nghiệm đã chứng tỏ giải pháp đề xuất đạt hiệu quả cao, đảm bảo giảm được đồng thời cả hai loại sai số FAR và FRR, vừa tăng nhanh tốc độ tra cứu lên nhiều lần, tỷ lệ thuận với số nút song song đưa vào xử lý, vừa rút ngắn đến 66,2% danh sách kết quả tra cứu tính trung bình trên 1 yêu cầu so với trường hợp không ứng dụng giải pháp tổ hợp đa tầng.

5.4. Nhận xét chung

Trên đây luận án đề xuất một loạt các giải pháp cải tiến phương pháp truy nguyên và thẩm định vân tay hiện trường dùng kỹ thuật tổ hợp đa tầng và đưa vào ứng dụng thực tế. Kết quả thử nghiệm cũng cho thấy giải pháp đề xuất đã rút ngắn đáng kể thời gian và công sức tìm kiếm, vì vậy có thể đưa vào cài đặt để nâng cấp tiếp mô đun tìm kiếm vân tay hiện trường của hệ C@FRIS, đáp ứng kịp thời yêu cầu truy nguyên vân tay hiện trường mở rộng trên CSDL dung lượng lớn.

KẾT LUẬN

Luận án đã khảo sát, đề xuất cả về lý thuyết, thuật toán và thực nghiệm một số giải pháp để nâng cao hiệu quả hệ thống nhận dạng và truy nguyên vân tay tự động. Cụ thể là giải pháp phân đoạn tách vùng ảnh vân tay từ mẫu chỉ bản 10 ngón theo hai giai đoạn thô và mịn, giải pháp tổ chức và bảo mật truy cập dữ liệu, giải

pháp trích chọn đặc điểm toàn cục và cục bộ, giải pháp nắn chỉnh biến dạng và đối sánh vân tay cho các trường hợp vân tay ấn và vân tay lăn. Để xử lý vân tay hiện trường vốn có chất lượng thấp, chỉ xuất hiện từng phần và có độ biến dạng cao, luận án cũng đề xuất một kiến trúc đa tầng để kết hợp nhiều giải pháp khác nhau để truy nguyên và thẩm định. Các giải pháp, phương pháp được đề xuất đã được phân tích, đánh giá, kiểm thử thực nghiệm, so sánh với các hướng tiếp cận khác để khẳng định mức độ cải tiến, nâng cấp tính năng.

Sau đây là các kết quả chính đạt được:

1) Đã khảo sát và xây dựng được thuật toán phân đoạn thô ảnh chỉ bản 10 ngón, phân đoạn mịn ảnh từng ngón. Kết quả thu được đạt độ chính xác cao, loại bỏ được các bụi bẩn, vùng nền thừa và tiết kiệm trung bình 1/3 dung lượng bộ nhớ so với kích thước ảnh phân đoạn theo phương pháp cắt thô chia đều 10 ô cho 10 ngón. Giải pháp phân đoạn tự động vân tay từ mẫu chỉ bản 10 ngón đã được đưa vào sử dụng để nâng cấp hiệu năng modul nhập chuyên đổi thông tin số hóa, nâng công suất nhập liệu từ 500 chỉ bản/ngày lên chế độ xử lý tự động theo lô đạt công suất 5000 chỉ bản/ngày.

2) Đã khảo sát và xây dựng được mô hình nắn chỉnh TPS từng phần để cải tiến modul đối sánh vân tay lăn/ấn theo hướng khử các hiện tượng vân tay bị méo phi tuyến. Kết quả thực nghiệm trên các CSDL FVC 2004 (DB1, DB3) cho thấy so với phương pháp TPS toàn phần, phương pháp mới nâng cao rõ rệt hiệu quả đối sánh nhờ áp dụng kỹ thuật nắn chỉnh từng phần, thể hiện qua kết quả cải tiến các phân bố chính danh và giả danh, từ đó giảm thiểu được đồng thời cả hai loại sai số FAR và FRR, với thời gian và bộ nhớ sử dụng được giảm xuống đáng kể.

3) Đã khảo sát kỹ và xây dựng được phương pháp đánh chỉ số dữ liệu ảnh vân tay theo thuộc tính ảnh do hệ thống tự động trích chọn cho cả hai chế độ tra cứu TP/TP và LT/TP. Đã đề xuất giải pháp song song hóa CSDL theo hướng phân phối động nhiệm vụ theo năng lực cho nhiều nút xử lý song song và giải pháp kiểm soát truy cập CSDL qua mạng dùng công nghệ bảo mật BioPKI. Với giải pháp tổ chức CSDL song song hoá, thời gian tra cứu trung bình một yêu cầu trên CSDL dung lượng lớn được rút xuống hàng chục lần, xấp xỉ tỷ lệ thuận với số nút đưa vào xử lý song song. Nhờ ứng dụng các tính năng của giải pháp BioPKI, việc bảo vệ hệ C@FRIS được đảm bảo chặt chẽ mà vẫn giữ được tính dễ dùng trong các khâu xây dựng, khai thác và vận hành hệ thống.

4) Để nâng cấp mô đun tra cứu vân tay hiện trường, luận án cũng đề xuất một giải pháp dùng kỹ thuật tổ hợp đa tầng, kết hợp kiểu bậc thang nhiều phương pháp khác nhau và đưa vào ứng dụng thực tế. Kết quả thực nghiệm cho thấy, giải pháp đề xuất đã cải thiện rõ rệt độ chính xác, tốc độ truy nguyên, thẩm định vân tay hiện trường trên CSDL qui mô lớn.

Những nội dung mới cần tiếp tục nghiên cứu:

1) Tiếp tục cải tiến giải pháp phân đoạn thô để tách 10 ảnh vân tay lặn và 10 ảnh vân tay ẩn từ mẫu chỉ bản 10 ngón bằng phương pháp tự động xử lý khôi phục các đường kẻ chia ô trên mẫu thẻ chỉ bản 10 ngón để thay thế các giải pháp phân đoạn phức tạp.

2) Tiếp tục hoàn thiện giải pháp phân đoạn mịn để lập bản đồ hướng vân, bản đồ độ cong, bản đồ chất lượng và tính thêm bản đồ bước sóng và khung xương vector hóa của vân tay.

3) Nâng cao độ chính xác thuật toán trích chọn đặc trưng chi tiết theo hướng tăng cường các kỹ thuật khử lỗi dính nét, đứt nét, dính rãnh, đứt rãnh sau trích chọn và đánh giá độ tin cậy của từng điểm đặc trưng và triển khai thêm thuật toán khái quát hóa (generalization) bộ đặc trưng chi tiết từ nhiều mẫu vân tay đồng nguyên.

4) Tiếp tục cải tiến thuật toán đối sánh vân tay biến dạng phi tuyến theo hướng nắn chỉnh từng phần dựa trên kỹ thuật chọn tập các cặp điểm không chế tối ưu phân bố đồng đều trên ảnh vân tay dùng thuật toán "tham lam" để tính phủ tối thiểu.

5) Phát triển thuật toán đối sánh vân tay hiện trường không chỉ dùng đặc trưng chi tiết mà dùng cả bản đồ độ cong, bản đồ tần số và bản đồ khung xương đường vân.

DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN ĐẾN LUẬN ÁN

1. Nguyễn Ngọc Kỳ, **Nguyễn Thị Hương Thủy**, Nguyễn Thanh Phương, Nguyễn Ngọc Minh (2009), “Hệ thống phần mềm nhận dạng vân tay tự động dùng để tự động hóa tăng thư căn cước công dân, căn cước can phạm và tìm kiếm dấu vân tay hiện trường phục vụ công tác quản lý hành chính và điều tra tội phạm”, *Kỷ yếu Lễ trao giải thưởng sáng tạo khoa học công nghệ Việt Nam và giải thưởng WIPO năm 2008*, tr. 169-173.
2. Nguyễn Ngọc Kỳ, **Nguyễn Thị Hương Thủy**, Nguyễn Thanh Phương, Nguyễn Ngọc Minh (2009) “Sản phẩm phần mềm nhận dạng vân tay tự động C@FRIS 2009- phiên bản mới dành để điện tử hóa tăng thư căn cước công dân qui mô hàng triệu đến hàng chục triệu chỉ bản”, *Kỷ yếu Hội thảo Sáng tạo khoa học công nghệ với sự nghiệp CNHĐH đất nước*, tr. 185-190.
3. **Nguyễn Thị Hương Thủy**, Nguyễn Ngọc Kỳ, Hoàng Xuân Huân, Nguyễn Ngọc Minh (2010), “Nâng cao hiệu quả thuật toán đối sánh vân tay dùng mô hình nắn chỉnh biến dạng địa phương LTM”, *Kỷ yếu hội thảo FAIR: Nghiên cứu cơ bản Ứng dụng công nghệ thông tin*, tr. 215-227.
4. **Nguyễn Thị Hương Thủy**, Nguyễn Ngọc Minh, Nguyễn Ngọc Kỳ (2010), “Thuật toán phân đoạn ảnh chỉ bản vân tay mười ngón”, *Tạp chí Tin học và Điều khiển học* Tập 26 (3), tr. 253-266.
5. **Nguyễn Thị Hương Thủy**, Nguyễn Văn Toàn, Nguyễn Ngọc Kỳ, Nguyễn Thị Hoàng Lan (2010), “Xây dựng giải pháp bảo mật BioPKI và ứng dụng để bảo mật hệ thống nhận dạng vân tay”, *Kỷ yếu hội thảo quốc gia: Một số vấn đề chọn lọc của Công nghệ thông tin và Truyền thông*, tr. 333-346.
6. Nguyễn Văn Toàn, **Nguyễn Thị Hương Thủy**, Nguyễn Ngọc Kỳ, Nguyễn Thị Hoàng Lan (2011), “Bảo mật truy cập dựa trên BioPKI và ứng dụng để bảo mật hệ thống nhận dạng vân tay C@FRIS”, *Chuyên san tạp chí Thông tin, Khoa học công nghệ của Bộ Thông tin và Truyền thông* Kỳ 3 Tập V-1 (6(26)), tr. 183-194.
7. **Nguyen Thi Huong Thuy**, Hoang Xuan Huan and Nguyen Ngoc Ky (2013), “An Efficient Method for Fingerprint Matching based on Local Point Model”, *Proc. of the International Conference on Computing, Management and Telecommunications (ComManTEL)*, pp. 334-339.
8. **Nguyen Thi Huong Thuy**, Hoang Xuan Huan, Nguyen Ngoc Ky and Le Minh Khoi (2013), “An Efficient Cascaded System for Latent Fingerprint Recognition”, *In Proceedings of the 10th IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF 2013)*, pp. 123-126.
9. **Nguyen Thi Huong Thuy**, Hoang Xuan Huan, Nguyen Ngoc Ky and Le Minh Khoi (2013), “An Efficient Multi-Stage System for Latent Fingerprint Recognition”, *Journal of Computer Science and Cybernetics* (revised).