

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



NGUYỄN THỊ HỢP

**NGHIÊN CỨU GIẢI THUẬT DI TRUYỀN VÀ
ỨNG DỤNG VÀO BÀI TOÁN TÁI TẠO ẢNH
CHÂN DUNG ĐỐI TƯỢNG**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Thái Nguyên - 2012

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



NGUYỄN THỊ HỢP

**NGHIÊN CỨU GIẢI THUẬT DI TRUYỀN VÀ
ỨNG DỤNG VÀO BÀI TOÁN TÁI TẠO ẢNH
CHÂN DUNG ĐỐI TƯỢNG**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Chuyên ngành: Khoa học máy tính

Mã số: 60 48 01

NGƯỜI HƯỚNG DẪN KHOA HỌC: TS NGUYỄN NGỌC CƯỜNG

Thái Nguyên - 2012

MỞ ĐẦU

1. Tính cấp thiết của đề tài

Con người là đối tượng nghiên cứu rất phong phú và phức tạp của nhiều lĩnh vực khoa học khác nhau. Mỗi ngành khoa học, mỗi lĩnh vực khoa học chỉ nghiên cứu về con người với góc độ, khía cạnh riêng để phục vụ thiết thực cho ngành khoa học đó. Bài toán tái tạo ảnh chân dung đối tượng từ những thông tin đặc tả về đối tượng không những phục vụ công tác quản lý con người mà còn ứng dụng cần thiết trong công tác phòng ngừa, đấu tranh chống các loại tội phạm.

Ngày nay, khi nền khoa học kỹ thuật phát triển toàn diện, việc ứng dụng khoa học kỹ thuật tiên tiến càng cho phép chúng ta giải quyết bài toán nhanh, chính xác và hiệu quả hơn. Nhờ có những thành tựu của các ngành khoa học khác như: Nhân chủng học, hình thái học, dân tộc học, nhân trắc học, khoa học chỉnh hình, ... đã tạo cơ sở phát triển cho khoa học nhận dạng người nói chung và lai ghép, tái tạo mặt người nói riêng.

Vì lẽ đó bài toán tái tạo ảnh chân dung đối tượng dựa vào các đặc điểm cơ bản của các bộ phận cấu tạo nên mặt người là bài toán rất thiết thực đang được nhiều người quan tâm, vấn đề cấp thiết đặt ra là giải quyết bài toán tái tạo ảnh chân dung đối tượng sao cho hiệu quả, đáp ứng được nhu cầu của người sử dụng.

Hiện nay đã có nhiều phương pháp tiếp cận, xây dựng chương trình tái tạo ảnh chân dung đối tượng. Các phần mềm đã có thường khối lượng công việc cần xử lý lớn, lọc lựa, lắp ghép trên tập các dữ liệu ảnh phức tạp nên tốc độ còn chậm.

Nhằm nâng cao chất lượng và hiệu quả của phần mềm tái tạo, lai ghép mặt người, tác giả đã tìm hiểu thuật giải di truyền có nhiều ưu điểm trong tốc độ tính toán, mô hình hoá. Do đó trong khuôn khổ luận văn thạc sỹ, tôi chọn đề tài: **“Nghiên cứu giải thuật di truyền và ứng dụng vào bài toán tái tạo ảnh chân dung đối tượng”**.

2. Mục tiêu của luận văn

Mục tiêu của luận văn là: nghiên cứu giải thuật di truyền và tìm cách ứng dụng giải thuật di truyền để tái tạo ảnh chân dung đối tượng gần giống nhất theo mô tả của các nhân chứng.

3. Nội dung của luận văn

Đề tài tập trung vào việc tìm hiểu các kiến thức cơ bản về giải thuật di truyền và bước đầu ứng dụng nó vào giải bài toán tái tạo ảnh chân dung khuôn mặt đối tượng theo hướng tiếp cận sau: sử dụng giải thuật di truyền để mã hóa khuôn mặt thành chuỗi gene, dựa trên những mô tả chi tiết của nhân chứng để tạo ra các khuôn mặt gần giống nhất với nhận định của nhân chứng. Sau khi quan sát khuôn mặt tái tạo được, nhân chứng có thể thay đổi, bổ sung các chi tiết khác. Các khuôn mặt tiếp tục được đánh giá bằng một hàm thích nghi. Độ thích nghi được đo bằng sự vi phạm các ràng buộc như ràng buộc giữa khuôn mặt và các bộ phận trên khuôn mặt, tỉ lệ co giãn, sinh các bộ phận và lai ghép tùy chọn là các véc tơ đa chiều. Mỗi ràng buộc được thể hiện thông qua độ lệch và độ dẫn biên của các bộ phận được ghép.

Trên cơ sở đó, nội dung của luận văn gồm ba chương sau phần Mở đầu:

- **Chương 1: Khái quát về giải thuật di truyền và bài toán tái tạo ảnh chân dung** (từ trang 4 đến trang 32)

Trong chương này trình bày khái quát về giải thuật di truyền: mã hóa nghiệm của bài toán, các toán tử di truyền, sơ đồ thuật toán di truyền, nền tảng toán học của giải thuật di truyền, mô hình hóa giải thuật di truyền bằng xích Markov. Phần tiếp theo của chương tìm hiểu ứng dụng giải thuật di truyền để giải bài toán tối ưu hóa, một số cải tiến của giải thuật di truyền và những đóng góp quan trọng của giải thuật di truyền trong thực tế.

Phần thứ 2 của chương trình bày khái quát về bài toán tái tạo ảnh chân dung đối tượng.

- Chương 2: Ứng dụng giải thuật di truyền vào bài toán tái tạo ảnh chân dung đối tượng (từ trang 33 đến trang 58)

Chương 2 trình bày một số đặc điểm của bài toán nhận dạng khuôn mặt, mối liên hệ của nó với bài toán tái tạo ảnh chân dung khuôn mặt người.

Phần tiếp theo của chương đề cập đến một số phương pháp, kỹ thuật xử lý ảnh được áp dụng trong bài toán tái tạo ảnh chân dung khuôn mặt đối tượng.

Nội dung trọng tâm của chương trình bày bài toán tái tạo ảnh chân dung khuôn mặt đối tượng theo hướng tiếp cận mới. Đó là ứng dụng giải thuật di truyền để giải quyết bài toán: mã hóa khuôn mặt thành chuỗi gene; xây dựng hàm mục tiêu đánh khuôn mặt; các toán tử di truyền; các bước giải bài toán tái tạo ảnh chân dung khuôn mặt bằng giải thuật di truyền; các tham số lựa chọn và chi phí thời gian thực hiện giải thuật.

Phần cuối chương 2 dành cho việc trình bày các công cụ và mô hình được lựa chọn để giải quyết bài toán theo hướng tiếp cận di truyền. Một số điểm hạn chế của giải pháp được đề cập đến ở cuối chương.

- Chương 3: Xây dựng chương trình tái tạo ảnh chân dung đối tượng và thử nghiệm (từ trang 59 đến trang 76)

Trong chương này tập trung vào việc phân tích, thiết kế hệ thống tái tạo ảnh chân dung khuôn mặt đối tượng, cài đặt và thử nghiệm chương trình tái tạo ảnh chân dung khuôn mặt đối tượng theo hướng tiếp cận giải thuật di truyền đã được đề xuất ở Chương 2.

- Phần kết luận và hướng phát triển.
- Tài liệu tham khảo.

4. Phương pháp nghiên cứu

Nghiên cứu tài liệu, đề xuất giải pháp và lập trình thử nghiệm.

Luận văn đã bước đầu đề xuất phương pháp ứng dụng giải thuật di truyền vào giải quyết bài toán tái tạo ảnh chân dung đối tượng, các chương trình thử nghiệm đã minh chứng hướng tiếp cận là đúng đắn và có hiệu quả.

Chương 1

KHÁI QUÁT VỀ GIẢI THUẬT DI TRUYỀN VÀ BÀI TOÁN TÁI TẠO ẢNH CHÂN DUNG

Chương này tìm hiểu những khái niệm cơ bản về giải thuật di truyền, cơ sở toán học của giải thuật di truyền, những cải tiến của giải thuật di truyền và các ứng dụng của nó trong thực tiễn. Phần cuối chương là những nét khái quát về bài toán tái tạo ảnh chân dung đối tượng.

1.1. Khái quát về giải thuật di truyền

1.1.1. Những khái niệm cơ bản về giải thuật di truyền đơn giản

Giải thuật di truyền (Genetic Algorithm-GA) thuộc lớp các giải thuật tìm kiếm tiến hóa. Khác với phần lớn các giải thuật khác tìm kiếm theo điểm, giải thuật di truyền thực hiện tìm kiếm song song trên một tập được gọi là quần thể các lời giải có thể. Thông qua việc áp dụng các toán tử gene, giải thuật GA trao đổi thông tin giữa các cực trị và do đó làm giảm thiểu khả năng kết thúc giải thuật tại một cực trị địa phương.

Ý tưởng áp dụng các nguyên lý của Darwin để tự động giải bài toán xuất hiện từ những năm 40 của thế kỷ 20, rất lâu trước khi máy tính ra đời [8]. Từ những năm đó Turing đã đề xuất “phép tìm kiếm tiến hóa hay tìm kiếm gene” (Genetical or evolutionary search). Trong những năm 1960, ba khuynh hướng phát triển của ý tưởng cơ sở này đã diễn ra ở các nơi khác nhau. Tại Mỹ, Fogel, Owens và Walsh đề xuất hướng nghiên cứu lập trình tiến hóa (Evolutionary programming) cùng thời điểm với phương pháp của Holland có tên gọi là giải thuật di truyền (Genetic Algorithm)[9, 10]. Trong khi đó tại Đức, Rechenberg và Schwefel đặt nền móng cho chiến lược tiến hóa (Evolution Strategies). Trong khoảng 15 năm sau đó, các hướng nghiên cứu này được phát triển một cách riêng biệt. Cho đến những năm 1990, các hướng nghiên cứu này được nhìn nhận lại như những thể hiện khác nhau của một công nghệ chung là tính toán tiến hóa

(Evolutionary Computing)[9, 10, 11, 12]. Cũng trong thời điểm này, một nhánh thứ tư dựa trên ý tưởng chung ra đời với tên gọi lập trình gene (genetic programming) do Koza đi tiên phong [11, 12]. Trong thuật ngữ hiện đại, toàn bộ lĩnh vực nghiên cứu này được coi là ngành tính toán tiến hóa, các thuật toán trong đó được gọi là các thuật toán tiến hóa. Lập trình tiến hóa, chiến lược tiến hóa, giải thuật di truyền và lập trình gene được xem như các lĩnh vực nhỏ thuộc về các biến thể của thuật toán tương ứng.

Giải thuật di truyền là một họ các mô hình tính toán dựa trên ý tưởng tiến hóa. Các giải thuật này mã hóa nghiệm tiềm năng của một bài toán cụ thể bằng một cấu trúc dữ liệu giống như các nhiễm sắc thể (chromosome) và áp dụng các toán tử tái tổ hợp (recombination operators) lên các cấu trúc dữ liệu đó sao cho có thể giữ được các thông tin chính. Giải thuật di truyền thường được xem như những bộ tối ưu hàm số mặc dù chúng có thể áp dụng vào nhiều lĩnh vực rộng hơn.

Một cài đặt của giải thuật di truyền bắt đầu với một quần thể các nhiễm sắc thể (thường được tạo ngẫu nhiên). Người ta đánh giá các cấu trúc này và phân bổ cơ hội tái sinh cho chúng theo cách những nhiễm sắc thể biểu diễn nghiệm tốt hơn của bài toán sẽ được ưu tiên hơn các nhiễm sắc thể khác để tái sinh. Mức độ “tốt” của nghiệm thường được xác định tương ứng với quần thể hiện thời.

Mô tả của giải thuật di truyền như trên là khá trừu tượng và có hai cách hiểu về thuật ngữ “giải thuật di truyền”. Theo nghĩa chặt, “giải thuật di truyền” được coi là một mô hình tính toán do Holland đề xuất và nghiên cứu từ năm 1975. Hầu hết các lý thuyết về giải thuật di truyền đang tồn tại cho đến hiện nay đều dùng thuật ngữ theo nghĩa này và khi đó mô hình được gọi là giải thuật di truyền chính xác. Theo nghĩa rộng hơn, giải thuật di truyền được coi là bất cứ một mô hình nào dựa trên quần thể (population-based model) nào đó sử dụng các toán tử tái tổ hợp và lựa chọn để sinh ra các điểm mẫu mới trong không gian tìm kiếm. Nhiều mô hình giải thuật di truyền đã được các tác giả nghiên cứu

thực nghiệm đưa ra trên quan điểm hướng ứng dụng và xem giải thuật di truyền như công cụ tối ưu hóa.

Trong mục này, chúng ta sẽ trình bày những khái niệm cơ bản liên quan tới giải thuật di truyền.

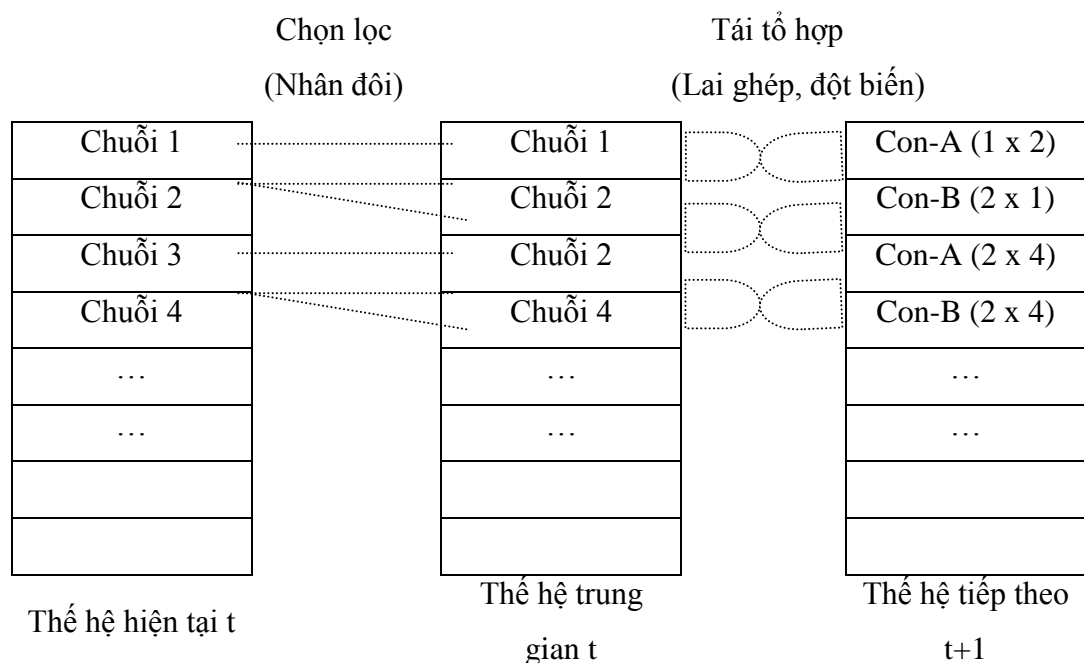
Trong hầu hết các giải thuật di truyền thông thường có hai thành phần phụ thuộc vào bài toán: mã hóa bài toán và hàm đánh giá.

Bước đầu tiên trong bất kỳ một giải thuật di truyền nào là bước tạo sinh quần thể xuất phát. Trong giải thuật di truyền chính tắc, mỗi thành viên của quần thể này là một chuỗi nhị phân độ dài l tương ứng với mã hóa nghiệm của bài toán. Mỗi chuỗi như vậy được xem như một nhiễm sắc thể. Trong hầu hết các trường hợp, quần thể này được sinh ra một cách ngẫu nhiên. Sau khi sinh ra quần thể xuất phát, mỗi cá thể của quần thể được đánh giá và được gán cho một giá trị thích nghi (fitness value).

Khái niệm đánh giá (evaluation) và thích nghi (fitness) đôi khi được sử dụng như cặp từ đồng nghĩa. Tuy nhiên, người ta thường phân biệt giữa hàm đánh giá (evaluation function) và hàm thích nghi (fitness function) được sử dụng trong các giải thuật di truyền. Trong mục này, hàm đánh giá (hay hàm mục tiêu) cung cấp độ đo hiệu quả của việc thiết lập giá trị các tham số cụ thể. Hàm thích nghi biến đổi độ đo hiệu quả này thành việc phân bổ cơ hội tái tạo cho các cá thể. Việc đánh giá một chuỗi biểu diễn tập các tham số là hoàn toàn độc lập với việc đánh giá các chuỗi khác. Tuy nhiên, mức độ thích nghi (fitness) của một chuỗi luôn luôn được xác định trong mối tương quan với các thành viên khác trong quần thể hiện tại.

Trong giải thuật di truyền, mức độ thích nghi có thể được xác định bằng f_i/\bar{f} với f_i là đánh giá của chuỗi i và \bar{f} là đánh giá trung bình của tất cả các chuỗi trong quần thể. Mức độ thích nghi cũng có thể được gán dựa trên thứ hạng của cá thể trong quần thể hoặc bằng các phương pháp lấy mẫu như phương pháp lựa chọn theo đấu loại.

Việc thực hiện giải thuật di truyền có thể được xem như một quá trình hai giai đoạn. Thuật toán bắt đầu với quần thể hiện thời. Việc lựa chọn được áp dụng vào quần thể này để tạo ra một quần thể trung gian. Sau đó việc lai ghép và đột biến được áp dụng cho quần thể trung gian để tạo nên quần thể tiếp theo. Quá trình chuyển từ quần thể hiện thời tới quần thể tiếp theo tạo nên một thế hệ trong tiến trình thực hiện giải thuật di truyền. Cách thực hiện như vậy được gọi là cài đặt của giải thuật di truyền đơn giản (Simple Genetic Algorithm – SGA). Hình 1.1 minh họa việc hình thành một thế hệ mới theo hai pha: pha chọn lọc và pha tái tổ hợp. Việc đột biến có thể thực hiện ngay sau lai ghép.



Hình 1.1 Một thế hệ mới được hình thành qua pha chọn lọc và pha tái tổ hợp

Đầu tiên chúng ta xem xét việc xây dựng quần thể trung gian từ quần thể hiện thời. Trong thế hệ đầu tiên, quần thể hiện thời cũng đồng thời là quần thể xuất phát. Sau khi tính f_i / \bar{f} cho tất cả các chuỗi trong quần thể hiện thời, chúng ta thực hiện việc lựa chọn. Trong giải thuật di truyền chính tắc, xác suất để một chuỗi trong quần thể hiện thời được sao chép lại và được đưa vào thế hệ trung gian tỷ lệ thuận với mức độ thích nghi của chúng.

Có nhiều cách thực hiện việc lựa chọn. Chúng ta có thể ánh xạ quần thể lên một bánh xe roulette, mỗi cá thể chiếm một không gian tỷ lệ thuận với mức độ

thích nghi của nó trên bánh xe. Quần thể trung gian được tạo nên nhờ việc quay liên tiếp bánh xe để chọn ra các cá thể theo cơ chế “lấy mẫu ngẫu nhiên có thay thế” (stochastic sampling with replacement). Cơ chế lựa chọn như vậy được gọi là lựa chọn tỷ lệ (proportional selection) và xác suất để một phân tử b được lựa chọn xác định bởi công thức:

$$p(b) = (f(b) / \sum_{i=1}^n f(b_i)) > 0$$

Với b và các b_i là các cá thể nằm trong quần thể hiện tại.

Quá trình lựa chọn cũng có thể thực hiện bằng cơ chế “remainder stochastic sampling”. Khi đó mỗi chuỗi i với f_i/\bar{f} lớn hơn 1 sẽ được sao chép vào quần thể trung gian với số lần bằng phần nguyên của f_i/\bar{f} . Sau đó tất cả các chuỗi (kể cả các chuỗi có f_i/\bar{f} nhỏ hơn 1) sẽ được sao chép thêm vào quần thể trung gian với xác suất tỷ lệ thuận với phần thập phân của chúng.

Sau khi lựa chọn, việc tái tổ hợp được thực hiện trên quần thể trung gian. Việc này có thể được coi như việc tạo ra quần thể tiếp theo từ quần thể trung gian. Việc lai ghép (crossover) được áp dụng cho các chuỗi được ghép cặp một cách ngẫu nhiên với xác suất p_c : lấy ra một cặp chuỗi; tái tổ hợp hai chuỗi này với xác suất p_c để tạo nên hai chuỗi mới và đặt chúng vào quần thể tiếp theo.

Bước tiếp theo là việc áp dụng toán tử đột biến (mutation operator). Mỗi bit trong quần thể có thể chịu hiện tượng đột biến với xác suất p_m . Thông thường, tần xuất đột biến được thực hiện với xác suất nhỏ hơn 1%. Trong một số trường hợp, đột biến được giải thích như việc tạo ngẫu nhiên một bit mới. Trong các trường hợp khác, đột biến được xem là phép lật bit. Sự khác nhau giữa hai cách giải thích thực chất chỉ là chi tiết cài đặt và mỗi kiểu đột biến đều có thể chuyển đổi để nhận được kiểu còn lại.