

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

NGUYỄN THÙY DƯƠNG

TÌM KIẾM ÂM NHẠC TRÊN CƠ SỞ
NỘI DUNG VÀ ỨNG DỤNG

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Thái Nguyên - 2012

MỤC LỤC

	Trang
MỤC LỤC.....	i
LỜI CẢM ƠN	ii
DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT.....	iii
MỞ ĐẦU.....	1
CHƯƠNG I	4
TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU ÂM THANH	4
1.1. Cơ sở dữ liệu đa phương tiện.....	4
1.1.1. Một số khái niệm cơ bản.....	4
1.1.2. Hệ thống quản trị cơ sở dữ liệu đa phương tiện (MMDBMS)	5
1.1.3. Truy tìm thông tin	7
1.2. Quá trình chung thiết kế cơ sở dữ liệu âm thanh.....	7
1.2.1. Giới thiệu chung về cơ sở dữ liệu âm thanh.	7
1.2.2. Mô hình tổng quát của dữ liệu âm thanh	8
1.2.3. Quá trình chung để phát triển một cơ sở dữ liệu âm thanh.....	14
CHƯƠNG 2	15
KỸ THUẬT TÌM KIẾM TRONG CƠ SỞ DỮ LIỆU ÂM NHẠC	15
2.1. Kỹ thuật phân loại âm thanh thành âm nhạc, tiếng nói, tiếng động.	16
2.1.1. Âm thanh.....	16
2.1.2. Đặc trưng chính của âm thanh	22
2.1.3. Phân lớp âm thanh.....	39
2.2. Một số thuật toán trích chọn đặc trưng âm nhạc.	45
2.2.1. Đặc trưng của âm nhạc.....	45
2.2.2. Một số lớp âm nhạc.....	45
2.2.3. Chỉ số hóa và truy tìm âm nhạc.	49
CHƯƠNG 3.	53
ỨNG DỤNG KỸ THUẬT TÌM KIẾM ÂM NHẠC TRONG GIẢNG DẠY	53
3.1. Cài đặt thử nghiệm hệ thống tìm kiếm âm nhạc.....	53
3.2. Mô hình hệ thống.....	54
3.3. Các tham số thực nghiệm	54
3.4. Một số chức năng của chương trình	55
3.5. Kết quả thực nghiệm.....	56
KẾT LUẬN	57

LỜI CẢM ƠN

Để hoàn thành chương trình cao học và viết luận văn này, tôi đã nhận được sự hướng dẫn, giúp đỡ và góp ý nhiệt tình của quý thầy cô trường Đại học Công nghệ thông tin - Truyền thông, Đại học Thái Nguyên.

Trước hết, tôi xin chân thành cảm ơn đến quý thầy cô trường Đại học Công nghệ thông tin - Truyền thông, đặc biệt là những thầy cô đã tận tình dạy bảo cho tôi suốt thời gian học tập tại trường.

Tôi xin gửi lời biết ơn sâu sắc đến thầy giáo **PGS.TS. Đặng Văn Đức** người đã dành rất nhiều thời gian, tâm huyết và sự tận tình giúp đỡ, hướng dẫn cho tôi trong suốt quá trình nghiên cứu và giúp tôi hoàn thành luận văn này.

Nhân đây, tôi xin chân thành cảm ơn Ban Giám hiệu trường Cao đẳng Văn hóa nghệ thuật Việt Bắc cùng đồng nghiệp trong Phòng Đào tạo & nghiên cứu khoa học đã tạo rất nhiều điều kiện để tôi học tập và hoàn thành tốt khóa học.

Đồng thời, tôi cũng xin cảm ơn quý anh, chị và ban lãnh đạo khoa Âm nhạc, phòng thu thanh của trường cao đẳng Văn hóa nghệ thuật Việt Bắc... đã tạo điều kiện cho tôi điều tra khảo sát để có dữ liệu viết luận văn.

Mặc dù tôi đã có nhiều cố gắng hoàn thiện luận văn bằng tất cả sự nhiệt tình và năng lực của mình, tuy nhiên không thể tránh khỏi những thiếu sót, rất mong nhận được những đóng góp quý báu của quý thầy cô và các bạn.

Thái Nguyên, ngày tháng năm 2012

Học viên

Nguyễn Thùy Dương

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

Stt	Ký hiệu	Tên đầy đủ	Ý nghĩa
1	ADC	Analog - to – Digital Conversion	Chuyển đổi tương tự số hóa
2	ASR	Automatic Speech Recognition	Tự động nhận dạng tiếng nói
3	API	Application Programming Interface	Giao diện lập trình ứng dụng
4	DAC	Digital – to – Analog Conversion	Chuyển đổi số hóa tương tự
5	DBMS	DataBase Management System	Hệ quản trị cơ sở dữ liệu
6	DCT	Discrete Cosine Transform	Biến đổi cosin rời rạc
7	DFT	Discrete Fourier Transform	Biến đổi Fourier rời rạc
8	HMM	Hidden Markov Model	Mô hình Markov ẩn
9	HZCRR	High Zero- Crossing Rate Ratio	Tỷ lệ tốc độ vượt qua 0 cao
10	IDFT	Inverse Discrete Fourier Transform	Biến đổi Fourier liên tục
11	IR	Information Retrieval	Truy tìm thông tin
12	MARS	Multimedia Analysis and Retrieval System	Hệ thống chỉ mục và phân tích đa phương tiện
13	MIDI	Musical Instrument Digital Interface	Giao diện số cho nhạc cụ
14	MIRS	Multimedia Indexing and Retrieval System	Hệ thống chỉ mục và truy tìm thông tin đa phương tiện

15	MMDBMS	MultiMedia DataBase Management System	Hệ quản trị cơ sở dữ liệu đa phương tiện
16	ORDBMS	Open relation DataBase Management System	Hệ quản trị cơ sở dữ quan hệ - mở rộng
17	SR	Silence Ratio	Tỷ lệ câm
18	SNR	Signal-to-noise ratio	Tỷ lệ nhiễu tín hiệu
19	STFT	Short Time Fourier Transform	Biến đổi Fourier thời gian ngắn
20	ZCR	Zero Crossing Rate	Tốc độ vượt qua 0

MỞ ĐẦU

Nghị quyết Trung ương II khoá VIII đã đặt ra phương châm chiến lược cho ngành giáo dục là phải: "Đổi mới phương pháp giáo dục đào tạo, khắc phục lối truyền thụ một chiều, rèn luyện nếp tư duy sáng tạo của người học, từng bước áp dụng phương pháp tiên tiến và phương tiện hiện đại vào quá trình dạy học". Ứng dụng công nghệ thông tin vào dạy học đang dần dần được thực hiện ngày càng nhiều trên bục giảng. Việc thiết kế giáo án và giảng dạy trên máy tính bằng những đoạn phim minh hoạ với hình ảnh, âm thanh sống động, bài giảng sẽ thực sự gây sự chú ý và tạo hứng thú học tập cho học sinh-sinh viên. Trong điều kiện hiện nay tài liệu dạy âm nhạc bằng sự hỗ trợ của CNTT ở nước ta còn nhiều hạn chế, việc tìm ra các giải pháp ứng dụng phần mềm CNTT vào dạy học âm nhạc là yêu cầu cần thiết nhằm góp phần đổi mới phương pháp dạy học cho bộ môn này.

Ứng dụng CNTT trong dạy học Âm nhạc ở các trường Đại học – cao đẳng và các trường phổ thông là việc làm tất yếu, giúp cho giảng viên, giáo viên âm nhạc chủ động có những bài soạn mang tính hiện đại và tạo ra được những tài liệu học tập, tham khảo phong phú cho sinh viên mang tính trực quan sinh động, tạo được hứng thú học tập cho học sinh.

Trên thị trường hiện nay có rất nhiều phần mềm dùng để soạn nhạc, hoà âm, phối khí thu âm, xử lý, biên tập âm thanh, v.v... Các phần mềm đều có lĩnh vực ứng dụng nhất định có tính chuyên biệt khá rõ nét nhưng nhìn chung khi sử dụng đều có đặc điểm tương đối giống nhau nên việc sử dụng cũng khá dễ dàng. Việc lựa chọn sử dụng cần phải đáp ứng yêu cầu sau: - Lựa chọn những phần mềm đáp ứng nội dung bài dạy, tiết dạy. - Có kỹ năng sử dụng và khai thác phần mềm. - Dễ phổ cập, phù hợp với điều kiện kỹ thuật của địa phương.

Ngày nay, việc chia sẻ các bài hát bản nhạc trên internet đã trở nên phổ biến. Thực tế đó cũng mở ra một nhu cầu truy cập vào những thư viện âm thanh khổng lồ. Nhu cầu đó đòi hỏi một công nghệ tìm kiếm hiệu quả để tổ chức, sắp xếp, truy tìm

các nội dung âm thanh, cũng như có thể xử lý hàng trăm tỷ trang web hỗn độn trên mạng và các thiết bị lưu trữ trên các máy tính cá nhân.

Hiện nay, một số hãng tìm kiếm khổng lồ trên mạng như Yahoo, Google hay You Tube đang triển khai các hoạt động nghiên cứu theo cách tìm kiếm theo nội dung thay vì từ khóa. Người dùng chỉ cần gõ những từ hoặc cụm từ liên quan đến bài hát và Yahoo sẽ liệt kê một danh sách các file âm thanh, cho phép khách hàng nghe trực tuyến. Công cụ tìm kiếm âm nhạc này hoạt động dựa trên khả năng đọc nội dung được nhúng trong tệp âm thanh, còn được gọi là metadata, để phân loại kết quả tìm kiếm.

Trong những năm qua ở Việt Nam, các công cụ tìm kiếm dữ liệu đa phương tiện, trong đó các công cụ tìm kiếm và nhận dạng dữ liệu âm thanh đã dần được chú ý. Ví dụ, phòng nhận dạng và công nghệ tri thức- Viện công nghệ thông tin đã giới thiệu một số sản phẩm phần mềm có ý nghĩa thực tế rất cao như: tổng hợp và xử lý ngôn ngữ tiếng Việt, nghiên cứu và tiếp cận các kỹ thuật mới của công nghệ tổng hợp và nhận dạng tiếng nói trên thế giới để áp dụng trong hoàn cảnh Việt nam và âm thanh tiếng Việt, nghiên cứu các phương pháp xử lý tín hiệu số và tín hiệu tiếng nói, các phương pháp tìm đặc trưng âm thanh. Các sản phẩm nêu trên đã đáp ứng được phần nào nhu cầu tìm kiếm và nhận dạng dữ liệu âm thanh. Tuy nhiên, các công trình nghiên cứu về tìm kiếm âm thanh theo nội dung vẫn còn hiếm hoặc chưa tương xứng với tầm quan trọng của nó trong hiện tại và tương lai. Đặc biệt sử dụng các phần mềm này cho công việc giảng dạy âm nhạc tại các trường còn chưa đáp ứng được nhu cầu thực tế, thay vì phần trắng bảng đen truyền thống, giáo viên chỉ click chuột, vài giây sau trên màn hình hiện ra ngay nội dung, âm thanh của bản nhạc. Ứng dụng công nghệ thông tin trong dạy học âm nhạc đang dần dần được thực hiện ngày càng nhiều trên bục giảng các trường đào tạo nghệ thuật. Nhạc sĩ Hoàng Lê nhận xét rằng từ trước đến nay, bộ môn âm nhạc ”dạy chay” nhiều, ít lôi cuốn được học sinh. Nếu việc ứng dụng CNTT vào trong dạy học được thực hiện có hiệu quả trong các trường, sẽ dẫn đến sự thay đổi lớn lao về PPDH nhằm cung cấp

những tư liệu âm nhạc một cách khoa học, phong phú, toàn diện mà còn tác động tích cực đến thẩm mỹ, đến tư duy nhận thức của học sinh...

Mặt khác, việc hiểu biết sâu sắc về âm thanh cũng như các khái niệm, thuật toán liên quan sẽ giúp ta ứng dụng và xây dựng các hệ thống tìm kiếm âm thanh phù hợp, hiệu quả hơn. Xuất phát từ những vấn đề nêu trên, luận văn đã tập trung nghiên cứu về vấn đề "**Tìm kiếm âm nhạc trên cơ sở nội dung và ứng dụng**".

Cấu trúc của luận văn như sau:

Chương 1: Tổng quan về cơ sở dữ liệu âm thanh

Chương 2: Kỹ thuật tìm kiếm cơ sở dữ liệu âm thanh

Chương 3: Ứng dụng kỹ thuật tìm kiếm âm nhạc trong giảng dạy

CHƯƠNG I

TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU ÂM THANH

Ngày nay, với sự phát triển vượt bậc của khoa học kỹ thuật, sự bùng nổ của dữ liệu đặc biệt là dữ liệu media, nhu cầu thưởng thức âm nhạc của con người ngày càng phong phú, đa dạng dẫn đến một nhu cầu rất thực tế đó là nhu cầu tìm kiếm thông tin về âm nhạc.

Hầu hết các hệ thống tìm kiếm âm nhạc hiện tại đều dựa trên chỉ mục. Cơ sở dữ liệu dựa trên chỉ mục bộc lộ nhiều nhược điểm như: tìm kiếm thông tin không chính xác, không tìm được dữ liệu khi người dùng không nhớ chính xác thông tin đầu vào, hay chỉ nhớ được giai điệu, nội dung bài hát.

Tìm kiếm âm nhạc theo nội dung là một lĩnh vực nghiên cứu mới và được nhiều nhà nghiên cứu quan tâm. Hiện có một số phương thức đã được áp dụng tìm kiếm âm nhạc theo nội dung. Ghias, Logan, Chamberlin và Smith đã đưa ra phương thức tương tự và sử dụng đầu vào như một truy vấn. Trong thí nghiệm vào năm 2001, M.Goto đã thu được kết quả tốt khi các nhà nghiên cứu tiến hành một cuộc thử nghiệm tính toán ước lượng MAP sử dụng thuật toán EM. Tuy nhiên, theo kết quả nghiên cứu của Beth Logan thì các phương pháp tìm kiếm âm nhạc theo nội dung hiện nay vẫn chưa đảm bảo được cả độ chính xác và thời gian tính toán, đặc biệt khi tìm kiếm giai điệu của các bản nhạc

1.1. Cơ sở dữ liệu đa phương tiện.

1.1.1. Một số khái niệm cơ bản

❖ Loại media và Multimedia

Media

Media (tiếng Latin: medius – means, intermediary) là đề cập đến các loại thông tin hay loại trình diễn thông tin như dữ liệu văn bản, ảnh, âm thanh và video.

Phân loại media: Có nhiều cách phân loại, nhưng cách chung nhất là phân loại trên cơ sở khuôn mẫu (format) vật lý hay các quan hệ media với thời gian. Tài

liệu này phân lớp media dựa trên cơ sở là chúng có chiều thời gian hay không. Qui định này dẫn tới hai lớp media: *static* và *dynamic*.

- **Static media:** Không có chiều thời gian, nội dung và ý nghĩa của chúng không phụ thuộc vào thời gian trình diễn. Media tĩnh bao gồm dữ liệu văn bản, đồ họa.
- **Dynamic media:** Có chiều thời gian, ý nghĩa và độ chính xác của chúng phụ thuộc vào tốc độ trình diễn. Dynamic media bao gồm animation, video, audio. Media động phụ thuộc chặt chẽ vào tốc độ trình diễn. Thí dụ để cảm nhận chuyển động trơn tru, video phải được trình chiếu với tốc độ 25 frame/sec (hay 30 frame/sec phụ thuộc vào loại hệ thống video). Tương tự, khi ta play tiếng nói, âm nhạc, chúng chỉ được cảm nhận tự nhiên khi đạt được tốc độ nhất định, nếu không chúng làm giảm chất lượng và ý nghĩa âm thanh. Vì các media này phải được trình diễn liên tục và ở tốc độ cố định cho nên chúng còn được gọi là *media liên tục*. Hay còn gọi chúng là *media đẳng thời (isochronous media)* vì quan hệ giữa các đơn vị media và thời gian là cố định.

Multimedia

Khái niệm multimedia (tiếng Latin: multus- numerous) đề cập đến tập hợp các kiểu media sử dụng chung, trong đó ít nhất có một kiểu media không phải là văn bản (nói cách khác là ít nhất có một media trong đó là ảnh, audio hay video). Trong tài liệu này sử dụng multimedia như tính từ: thông tin đa phương tiện, dữ liệu đa phương tiện, hệ thống đa phương tiện, truyền thông đa phương tiện, ứng dụng đa phương tiện,... Dữ liệu đa phương tiện đề cập đến đại diện các kiểu media khác nhau mà máy tính có thể đọc được. Thông tin đa phương tiện đề cập đến thông tin kiểu media truyền đạt. Đôi khi khái niệm dữ liệu đa phương tiện và thông tin đa phương tiện sử dụng thay thế cho nhau.

Multimedia hay media item (hay còn gọi object) là thực thể tự trị trong hệ thống chỉ mục và truy tìm thông tin đa phương tiện (*MIRS – Multimedia Indexing and Retrieval System*) mà nó có thể được truy vấn, tìm kiếm và trình diễn.

1.1.2. Hệ thống quản trị cơ sở dữ liệu đa phương tiện (MMDBMS)