

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN
LỚP CỬ NHÂN TÀI NĂNG**

HUỲNH LÊ THÁI DƯƠNG – ĐỖ ĐỨC TRƯỜNG THỌ

**NGHIÊN CỨU VÀ PHÁT TRIỂN GIẢI PHÁP
ĐỊNH VỊ TRONG NHÀ DỰA VÀO ÂM THANH**

KHÓA LUẬN TỐT NGHIỆP CỬ NHÂN CNTT

TP.HCM, 2014

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN
LỚP CỬ NHÂN TÀI NĂNG**

HUỲNH LÊ THÁI DƯƠNG 1012078

ĐỖ ĐỨC TRƯỜNG THỌ 1012431

**NGHIÊN CỨU VÀ PHÁT TRIỂN GIẢI PHÁP
ĐỊNH VỊ TRONG NHÀ DỰA VÀO ÂM THANH**

KHÓA LUẬN TỐT NGHIỆP CỬ NHÂN TIN HỌC

**GIÁO VIÊN HƯỚNG DẪN
THS. ĐẶNG BÌNH PHƯƠNG**

NIÊN KHÓA 2010 – 2014

NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN

Khóa luận đáp ứng yêu cầu của LV cử nhân tin học.

TpHCM, ngày tháng năm 2014

Giáo viên hướng dẫn

NHẬN XÉT CỦA GIÁO VIÊN PHẢN BIỆN

Khóa luận đáp ứng yêu cầu của LV cử nhân tin học.

TpHCM, ngày tháng năm 2014

Giáo viên phản biện

LỜI CẢM ƠN

Chúng em xin chân thành cảm ơn Khoa Công Nghệ Thông Tin, trường Đại Học Khoa Học Tự Nhiên,Tp.HCM đã tạo điều kiện tốt cho chúng em thực hiện đề tài này.

Chúng em xin chân thành cảm ơn Thầy Đặng Bình Phương, là người đã tận tình hướng dẫn, chỉ bảo chúng em trong suốt thời gian thực hiện đề tài.

Chúng em cũng xin gửi lời cảm ơn sâu sắc đến quý Thầy Cô trong Khoa đã tận tình giảng dạy, trang bị cho chúng em những kiến thức quý báu trong những năm học vừa qua.

Chúng em xin gửi lòng biết ơn sâu sắc đến ba, mẹ, các anh chị và bạn bè đã ủng hộ, giúp đỡ và động viên chúng em trong những lúc khó khăn cũng như trong suốt thời gian học tập và nghiên cứu.

Mặc dù chúng em đã cố gắng hoàn thành luận văn trong phạm vi và khả năng cho phép, nhưng chắc chắn sẽ không tránh khỏi những thiếu sót, kính mong sự cảm thông và tận tình chỉ bảo của quý Thầy Cô và các bạn.

Nhóm thực hiện

Huỳnh Lê Thái Dương & Đỗ Đức Trường Thọ

ĐỀ CƯƠNG CHI TIẾT

Tên đề tài: Nghiên Cứu Và Phát Triển Giải Pháp Định Vị Trong Nhà Dựa Vào Âm Thanh

Giáo viên hướng dẫn: ThS. Đặng Bình Phương

Thời gian thực hiện: từ ngày 20/02/2014 đến ngày 07/07/2014

Sinh viên thực hiện:

Huỳnh Lê Thái Dương (1012078) – Đỗ Đức Trường Thọ (1012431)

Loại đề tài: Nghiên cứu lý thuyết, giải pháp kỹ thuật và xây dựng ứng dụng thực tế.

Nội Dung Đề Tài

Mục tiêu đề tài nhằm nghiên cứu và đề xuất một giải pháp hiệu quả nhằm cho phép định vị ở môi trường trong nhà (in-door) dựa trên việc phát những âm thanh tai người khó nghe được và khai thác yếu tố tai người khó phân biệt và nhận thấy trong môi trường âm thanh. Từ đó phát triển một số ứng dụng minh họa tính năng này.

Tập trung khai thác yếu tố tai người khó phân biệt, nhận biết được rằng có một tín hiệu đặc biệt đang được truyền. Luận văn sẽ vận dụng tính chất của tai người, từ đó lấy ra được các miền tần số mà tai người khó nghe được trong điều kiện thông thường mà thiết bị di động thông thường vẫn có khả năng ghi nhận và tạo ra được.

Cụ thể, nhóm thực hiện đã làm những việc sau:

- Tìm hiểu tổng quan về các vấn đề lí thuyết và thuật toán của quá trình xử lý tín hiệu số (DSP – Digital Signal Processing).
- Tìm hiểu về các library, framework hỗ trợ việc phát, thu và phân tích dữ liệu âm thanh trên nền tảng iOS.
- Tìm hiểu về các phương pháp, phương thức mới thích hợp cho việc ứng

dụng kết quả thu được từ việc rút trích dữ liệu từ âm thanh.

- Xây dựng ứng dụng với bộ thu sóng.
- Tích hợp bộ phát sóng vào ứng dụng.
- Tích hợp và chạy thử nghiệm.

Kế Hoạch Thực Hiện:

- 20/02/2013-01/03/2013: Tìm hiểu về các vấn đề lý thuyết và thuật toán liên quan đến quá trình xử lý tín hiệu số.
- 05/03/2013-12/03/2013: Tìm hiểu về các library, framework hỗ trợ việc phát, thu, phân tích dữ liệu âm thanh trên nền tảng iOS. Tìm hiểu về cách xây dựng ứng dụng di động trên nền tảng iOS.
- 15/03/2013-22/03/2013: Nêu và phân tích các vấn đề sẽ phát sinh trong quá trình xây dựng ứng dụng.
- 24/03/2013-04/04/2013: Đưa ra các giải pháp cho các vấn đề và lựa chọn giải pháp thích hợp. Đề ra phương pháp, phương thức mới thích hợp cho việc thu được dữ liệu được rút trích từ âm thanh.
- 05/04/2013-15/06/2013: Xây dựng bộ phát sóng trên ứng dụng và tích hợp bộ thu sóng. Từ đó xây dựng ứng dụng hoàn chỉnh.
- 16/06/2013-23/06/2013: Tiến hành tích hợp và chạy thử nghiệm.
- 25/06/2013-07/07/2013: Cải tiến những vấn đề khác của ứng dụng.

Xác nhận của GVHD	Ngày tháng năm 2014	
	Nhóm SV Thực hiện	
ThS. Đặng Bình Phương	Huỳnh Lê Thái Dương	Đỗ Đức Trường Thọ

MỤC LỤC

LỜI CÁM ƠN	iii
ĐỀ CƯƠNG CHI TIẾT	iv
DANH MỤC CÁC HÌNH	ix
TÓM TẮT KHÓA LUẬN	xi
Chương 1 Mở đầu	1
1.1. Hệ thống cảm ngử cảnh (Context-aware system) và ứng dụng	1
1.2. Lý do thực hiện đề tài	4
1.3. Mục tiêu đề tài	5
1.4. Nội dung đề tài	7
Chương 2 Khảo Sát Hiện Trạng	8
2.1. Dẫn nhập	8
2.2. Thuật toán FFT, các thư viện, mã nguồn mở hỗ trợ trên nền tảng iOS và các ứng dụng thực tế	9
2.2.1. Tổng quan các thư viện, mã nguồn mở về FFT	9
2.2.2. Một số ứng dụng thực tế sử dụng kỹ thuật xử lý tín hiệu	10
2.3. Kết chương	22
Chương 3 Các vấn đề và giải pháp kỹ thuật để phát triển ứng dụng	23
3.1. Vấn đề truyền nhận tín hiệu qua tín hiệu âm thanh	23
3.2. Lựa chọn giải pháp	24
3.2.1. Hạn chế do bộ filter của các thiết bị sử dụng nền tảng iOS	24
3.2.2. Vấn đề phụ thuộc vào engine để phát triển ứng dụng	29
3.2.3. Khó khăn về việc đọc các dữ liệu ghi âm (digital signal) và phân tích chúng	31

3.3. Kết chương.....	35
Chương 4 Vấn đề và giải pháp về việc truyền và nhận thông tin qua âm thanh vô hình đối với người nghe	36
4.1. Steganography và Watermarking.....	36
4.1.1. Tổng quan về Steganography - ẩn thông tin trong thông tin	36
4.1.2. Hai quy trình được đề xuất trong luận văn	37
4.1.3. Mối liên hệ giữa người nghe tuyệt đối của con người và âm lượng nhỏ nhất mà máy tính có thể thu được – Giải pháp đề xuất trong luận văn.	39
4.2. Tiếng ồn ảnh hưởng đến kết quả đo, kết quả thu nhận và vấn đề khoảng cách truyền âm	45
4.3. Chưa có một phương pháp hay protocol cụ thể nào quy định việc gửi và phát dữ liệu đi thông qua sóng âm thanh	49
4.4. Kết chương.....	52
Chương 5 Kiến trúc ứng dụng.....	53
5.1. Kiến trúc của bộ thu sóng và bộ phát sóng	53
5.1.1. Bộ phát sóng	53
5.1.2. Bộ thu sóng	55
5.2. Kiến trúc của ứng dụng Location-based	59
5.2.1. Phát biểu bài toán	59
5.2.2. Sơ đồ use case	59
5.2.3. Sơ đồ class (class diagram).....	62
5.2.4. Giao diện của ứng dụng	63
5.3. Kết chương.....	64

Chương 6 Kết luận	65
6.1. Các kết quả đạt được	65
6.2. Hướng phát triển	66

DANH MỤC CÁC HÌNH

Hình 1-1 Hệ thống cảm nhận cảnh (Context-aware systems)	1
Hình 1-2 Một số ứng dụng location-based.....	2
Hình 1-3 Mô hình sử dụng GPS và CellID	3
Hình 2-1 Ứng dụng WhatsApp	10
Hình 2-2 Giao diện ứng dụng Zalo	12
Hình 2-3 Giao diện ứng dụng iMessage	13
Hình 2-4 Giao diện ứng dụng Facebook Messenger.....	15
Hình 2-5 Ứng dụng FireChat	16
Hình 2-6 Mô hình triển khai của FireChat	16
Hình 2-7 Giao diện ứng dụng ShopKick.....	18
Hình 2-8 Ứng dụng Digimarc	20
Hình 2-9 Cách thức hoạt động của Digimarc.....	21
Hình 3-1 Nguyên lý cơ sở của ứng dụng	23
Hình 4-1 Quy trình 1: Tín hiệu đi kèm nhạc nền	38
Hình 4-2 Quy trình 2: Tín hiệu vô hình với tai người.....	39
Hình 4-3 Cơ sở của giải pháp.....	40
Hình 4-4 Mô hình của giải pháp	40
Hình 4-5 Kết quả thực nghiệm so sánh AHT của người và AHT của thiết bị.....	44
Hình 4-6 Biểu đồ liên quan giữa AHT người và AHT thiết bị	45
Hình 4-7 Biểu đồ thể hiện mối liên hệ giữa tần số, độ to của âm và mức độ nhỏ nhất để tai người có thể nghe được	47
Hình 4-8 Hình minh họa các loại sóng trong quá trình áp dụng kỹ thuật AFSK	51
Hình 5-1 Mô hình đóng gói dữ liệu (sử dụng trong ứng dụng ChatSound)	55
Hình 5-2 Sơ đồ Use case của ứng dụng ShopLocator (Location-based)	60

Hình 5-3 Bảng mô tả use case và actor của hệ thống	61
Hình 5-4 Sơ đồ lớp (class diagram) của ứng dụng ShopLocator	62
Hình 5-5 Giao diện của các màn hình trong ứng dụng ShopLocator	64

TÓM TẮT KHÓA LUẬN

Khoa học công nghệ ngày càng phát triển, mỗi ngày lượng dữ liệu lưu chuyển là vô cùng lớn. Nếu trước đây, các ứng dụng đã có những cách truyền thống trong việc truyền nhận dữ liệu như: Internet, Wireless, 3G, Bluetooth, v.v..., thì hiện tại công nghệ xử lý tín hiệu số đang ngày càng phát triển. Ngày càng có nhiều doanh nghiệp tham gia công nghệ DSP (Digital Signal Processing). Và ngày càng có nhiều ứng dụng trên thị trường sử dụng DSP làm cơ sở để phát triển (có thể kể đến như Digimarc – 1 ứng dụng phát hiện thông tin ẩn bên trong các tài liệu multimedia).

Rõ ràng các phương pháp truyền thống có những hạn chế nhất định như năng lượng, mã hóa, data hiding, ... Nhận thấy tiềm năng lớn của xử lý tín hiệu số đối với việc truyền – nhận – phân tích dữ liệu, nhóm chúng em đã tìm hiểu và phát triển ứng dụng location-based cho phép ghi nhận xác định vị trí, ngữ cảnh của người dùng. Từ đó tăng cường thêm các tài nguyên đi kèm tùy vào ngữ cảnh đó.

Bộ thu sóng được xây dựng trên nền tảng iOS, cho phép thu lấy âm thanh trong không gian xung quanh và phân tích các dữ liệu thu được. Từ đó rút trích thông tin, dữ liệu được ẩn chứa bên trong.

Ở một hướng khác, nhóm đã sử dụng tính chất ẩn dữ liệu trong một dữ liệu khác để xây dựng, phát triển ứng dụng cho phép người dùng có thể chat với nhau thông qua tín hiệu sóng âm thanh, gửi nhận, liên lạc, kết nối các thiết bị với nhau, bao gồm cả việc cross platform.

Bộ phát sóng được xây dựng trên nền iOS, cho phép mã hóa 1 hoặc nhiều dữ liệu (digital) và phát tín hiệu âm thanh (analog).

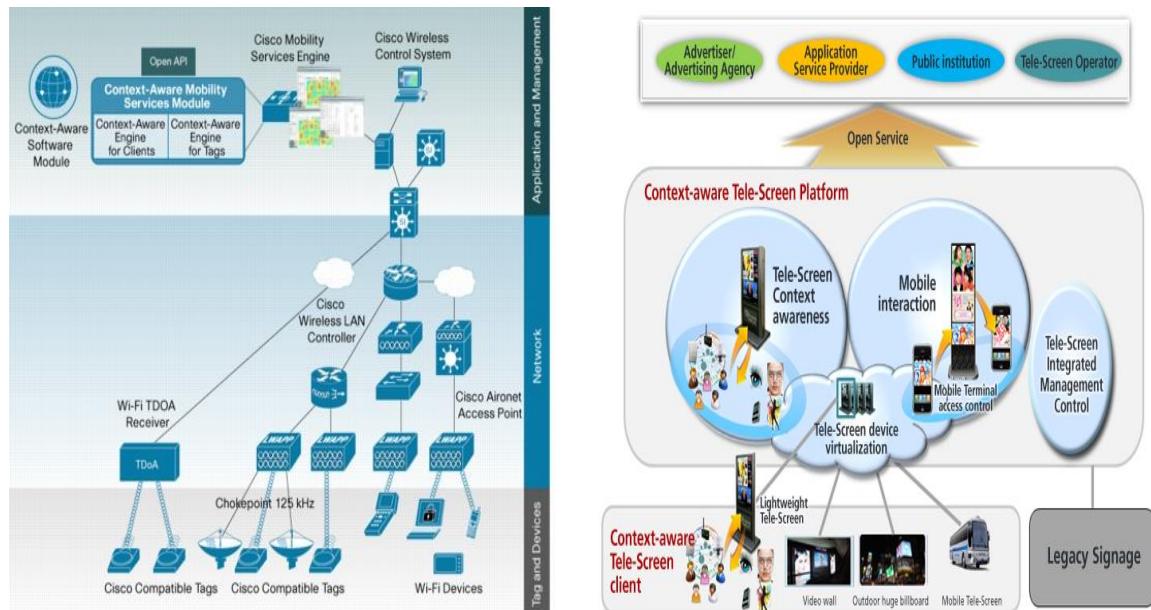
Chương 1

Mở đầu

 Nội dung Chương 1 giới thiệu tổng quan về các hệ thống cảm ứng cảnh (context-aware systems) và ứng dụng vào thực tế của các hệ thống này. Từ cơ sở đó trình bày lý do và mục tiêu của luận văn. Cuối chương là phần tóm tắt nội dung từng chương trong luận văn.

1.1. Hệ thống cảm ứng cảnh (Context-aware system) và ứng dụng

Hiện tại ngày nay, các ứng dụng phần mềm ngày càng phong phú. Đặc biệt là với sự phát triển ngày càng mạnh mẽ của các thiết bị di động thông minh. Đa dạng về chủng loại, đa dạng về cấu hình hay cả về hệ điều hành. Điều này giúp người dùng có thể truy cập và khai thác các dịch vụ tiện ích ở mọi nơi, mọi lúc (online, offline). Vậy làm thế nào để một hệ thống ngày càng trở nên tiện dụng hơn? Các ứng dụng có tích hợp cảm ứng cảnh chính là hướng phát triển đột phá. Việc phát triển các ứng dụng cảm ứng cảnh chính là một khuynh hướng ngày nay mà nhiều nhà phát triển đang quan tâm.



Hình 1-1 Hệ thống cảm ứng cảnh (Context-aware systems)

Một ngữ cảnh (context) có thể là: nhiệt độ môi trường, thời tiết, độ ẩm, vị trí địa lý, giờ trong ngày, quốc gia đang ở, ... Một ứng dụng nếu có tích hợp khả năng cảm nhận ngữ cảnh sẽ có thể phục vụ người dùng tiện dụng hơn vì nó sẽ có khả năng biết được tùy vào ngữ cảnh người dùng đang khai thác, hệ thống sẽ có thể tăng cường thêm những thông tin, tham số phù hợp.

Các ứng dụng location-based service (application) đang được quan tâm nghiên cứu và phát triển ngày càng nhiều. Các ứng dụng này khai thác tập trung vào việc máy tính có thể hiểu được ngữ cảnh. Điển hình là các ứng dụng như: Nokia City Lens, Layar, ...



Hình 1-2 Một số ứng dụng location-based

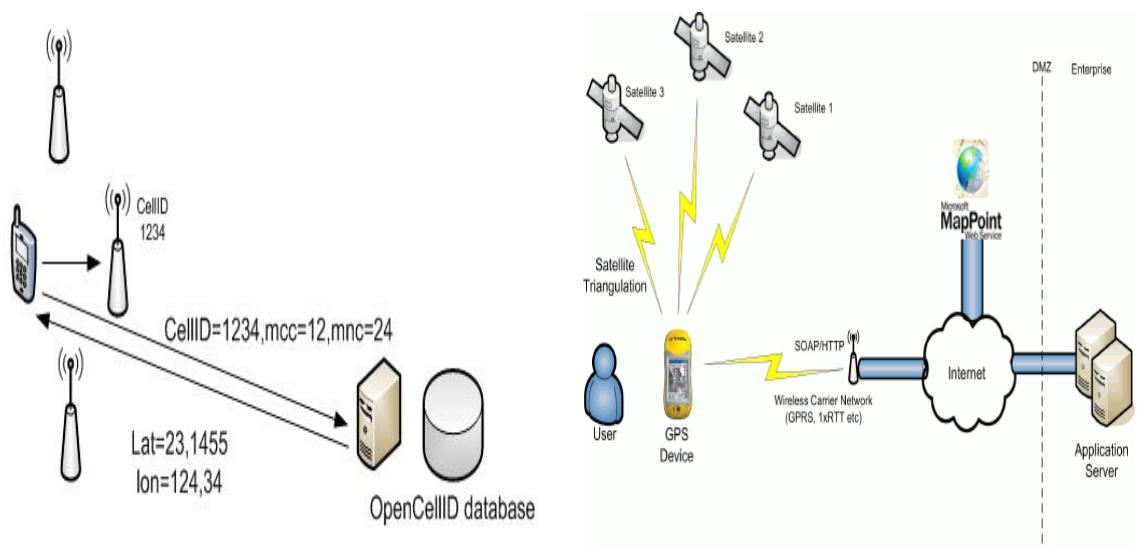
Để có thể làm được điều này thì đã có nhiều hướng giải pháp như:

- GPS
- Compass
- CellID

Những giải pháp này đã được thực hiện qua nhiều ứng dụng và các luận văn trước đây. Năm 2013, với đề tài luận văn nghiên cứu về Landmark Recognition – ứng dụng của hai sinh viên trường đại học Khoa Học Tự Nhiên, Nguyễn Ngọc Châu Sang và Trương Đại Dương đã xây dựng một ứng dụng cho phép ghi nhận, xử lý hình ảnh để người dùng có thể biết được bản thân đang đứng trước cảnh vật gì, nơi nào, sự vật nào. Việc này cần phải có 1 hệ thống được huấn luyện, cũng như cần có

nhiều nghiên cứu để cải tiến độ chính xác thuật toán. Nhất là khi phân biệt các cảnh vật có chi tiết tương tự nhau.

Với môi trường out-door, giải pháp CellID (sử dụng các cột sóng của các nhà mạng) cũng đang được phát triển. Tuy nhiên, để có thể triển khai giải pháp này, ta cần phải có một bản đồ tương đối chính xác của các Cell để biết Cell và ID đang xét nằm ở tọa độ GPS nào. Và cũng phụ thuộc vào mật độ phân bố của các trạm điện thoại ở một khu vực cho trước (càng nhiều – dày đặc sẽ cho ta một kết quả càng chính xác hơn)



Hình 1-3 Mô hình sử dụng GPS và CellID

Nhưng với môi trường trong nhà (in-door), giải pháp sử dụng GPS hay CellID sẽ kém hiệu quả do độ chính xác của chúng trong môi trường này bị hạn chế do các nguyên nhân: sóng yếu, khó nhận biết chính xác vị trí; đặc biệt là về thông tin của cao độ bị thiếu sót. Những hướng tiếp cận trước đây vẫn chưa phù hợp và chưa phát huy được hiệu quả của chính nó. Một giải pháp để khắc phục trong môi trường này đã được áp dụng đó là Wi-Fi Localization, đã được nhắc đến và sử dụng trong những bài báo khoa học như:

- Wi-Fi Localization and Navigation for Autonomous Indoor Mobile Robots
- User Position Detection In An Indoor Environment

Hay các ứng dụng của tiến sĩ Nguyễn Tuấn Nam đã nghiên cứu như:

- Localization on mobile devices using Wi-Fi signal (Định vị bằng thiết bị di động sử dụng sóng Wi-Fi)
- Location-based game on iPhone (Game liên quan đến vị trí bằng trên iPhone)
- Location-based game on Android
- Localization on mobile devices using Image Matching (Định vị bằng thiết bị di động sử dụng so khớp hình ảnh)

Giải pháp này dựa vào vị trí của các access point được đặt và các signature (ký hiệu) tại vị trí đó, các signature này có thể là mức độ tín hiệu mạnh hay yếu mà thiết bị thu được, từ đó theo quy ước đã định, xác định vị trí của người dùng. Tuy nhiên độ chính xác dựa vào số lượng, mật độ access point đang có trong môi trường và độ chính xác của Wi-Fi Localization Map được xây dựng qua thực nghiệm. Vì vậy, liệu có giải pháp nào có thể thực hiện đã giải quyết vấn đề in-door localization một cách hiệu quả, ít tốn kém chi phí?

1.2. Lí do thực hiện đề tài

Rõ ràng trong thực tế ngày nay, location-based và localization là rất cần thiết. Tuy nhiên những giải pháp trên vẫn còn chưa phù hợp. Với môi trường in-door chúng ta đã có các giải pháp:

- GPS, CellID: phụ thuộc mật độ của nơi phát tín hiệu, khó chính xác khi không có sự phân biệt được cao độ.
- Wi-Fi: vẫn phụ thuộc vào số lượng access point, việc xây dựng được Wi-Fi map và còn phụ thuộc rất lớn vào độ ổn định của tín hiệu.
- Land mark recognition : phụ thuộc vào việc huấn luyện để máy tính nhận dạng cũng như vấn đề thuật toán phức tạp (độ chính xác, chi phí).

Các giải pháp này vẫn còn đang được nghiên cứu và hoàn thiện. Qua việc nhận thấy vấn đề trong thực tế, có những âm thanh tai người nghe được, có những âm thanh tai người khó nghe được, nhóm chúng em nhận thấy có thể dựa vào điều này để giúp giải quyết vấn đề xác định vị trí cũng như truyền thông tin giữa các thiết bị. Vì vậy, nhóm mong muốn có thể khai thác yếu tố âm thanh và áp dụng kỹ thuật xử lý

âm thanh để qua đó có thể xác định một cách tương đối vị trí của người dùng và sẽ hiệu quả trong môi trường trong nhà.

1.3. Mục tiêu đề tài

Mục tiêu đề tài nhằm nghiên cứu và đề xuất một giải pháp hiệu quả nhằm cho phép định vị ở môi trường trong nhà (in-door) dựa trên việc phát những âm thanh tai người khó nghe được và khai thác yếu tố tai người khó phân biệt và nhận thấy trong môi trường âm thanh. Từ đó phát triển một số ứng dụng minh họa tính năng này.

Trong luận văn này, chúng em tập trung khai thác yếu tố tai người khó phân biệt, nhận biết được rằng có một tín hiệu đặc biệt đang được truyền. Luận văn sẽ vận dụng tính chất của tai người, từ đó lấy ra được các miền tần số mà tai người khó nghe được trong điều kiện thông thường mà thiết bị di động thông thường vẫn có khả năng ghi nhận và tạo ra được.

Ngoài ra trong luận văn cũng sẽ áp dụng giải pháp dựa trên việc ẩn dữ liệu (watermarking) trên âm thanh nhằm khai thác một số những thành phần âm thanh có thể được chèn thêm vào giai điệu mà tai người nghe bình thường không thấy được giúp cho âm thanh này có thể được truyền ở những tần số khác nhau. Trong luận văn này, chúng em không nhằm đề xuất một giải pháp mới về watermarking trên âm thanh hay cải tiến, tối ưu hóa các thuật toán về xử lý sóng âm thanh mà nhằm sử dụng, áp dụng tính chất watermarking trên âm thanh và các kết quả đã đạt được trong kỹ thuật này để xây dựng, phát triển ứng dụng phục vụ định vị trong môi trường trong nhà, cho phép nhận biết ngữ cảnh.

Nội dung thực hiện của luận văn bao gồm:

- Tìm hiểu về công nghệ và giải pháp.
 - o Tìm hiểu cơ chế truyền âm thanh trên các thiết bị: iOS, Android ...
 - o Tìm hiểu miền âm thanh phát ra ở tần số bao nhiêu, ghi nhận và xử lý tín hiệu số.
 - o Tìm hiểu cơ chế để ẩn thông tin vào trong một thông tin khác (Audio Watermarking).

- Đề xuất giải pháp định vị dựa vào âm thanh ở môi trường trong nhà.
 - o Đề xuất giải pháp giúp định vị dựa vào âm thanh, khai thác tần số tai người khó nghe.
 - o Đánh giá độ chính xác và tính vô hình của âm thanh trong nhiều điều kiện và môi trường khác nhau.
 - o Đề xuất giải pháp sử dụng Audio Watermarking trên miền tần số (frequency domain) để khai thác thêm 1 số thành phần âm thanh (1 số tần số) nhất định trong khuôn khổ tai người có thể nghe nhưng lại bị che khuất trong một giai điệu bình thường. Từ đó có khả năng tích hợp việc định vị vào trong các giai điệu được phát ra cố định tại những vị trí nào đó.
 - o Thủ nghiệm độ chính xác.
- Phát triển, thử nghiệm ứng dụng minh họa.
 - o Xác định yêu cầu chức năng của hệ thống cần dựng.
 - o Xây dựng kiến trúc hệ thống.
 - o Sơ đồ lớp, thiết kế giao diện.
 - o Cài đặt ứng dụng.

Kết quả đạt được của luận văn bao gồm:

- Phần mềm **ShopLocator** cho phép người dùng có thể xác định, định vị để phục vụ cho quảng cáo hay du lịch. Trong phần mềm này, người dùng có thể đi đến một vị trí nào đó (viện bảo tàng, siêu thị, nhà sách, cửa hàng, ...), nhận được một tín hiệu được phát sóng, liên kết với cơ sở dữ liệu. Từ đó lấy được một tập hợp tài nguyên, thông tin đi kèm liên quan như: bản đồ khu vực, các đoạn video hướng dẫn, giới thiệu bằng nhiều ngôn ngữ, nhiều giọng đọc khác nhau tùy thuộc vào vị trí của du khách.
- Mặc dù không minh họa cho mục tiêu chính của luận văn, tuy nhiên với kết quả đạt được là khả năng chia sẻ âm thanh, thông tin, truyền một cách vô hình đối với người cảm nhận, nên mục tiêu của phần mềm **ChatSound** này sẽ như một sản phẩm phụ để minh họa tính năng đó. Trong phần mềm này,

người dùng có khả năng chat với nhau qua sóng âm thanh, không cần có những kết nối truyền thông như internet, Wi-Fi hay bluetooth.

Cụ thể, nhóm thực hiện đã làm những việc sau:

- Tìm hiểu tổng quan về các vấn đề lí thuyết và thuật toán của quá trình xử lý tín hiệu số (DSP – Digital Signal Processing).
- Tìm hiểu về các library, framework hỗ trợ việc phát, thu và phân tích dữ liệu âm thanh trên nền tảng iOS.
- Tìm hiểu về các phương pháp, phương thức mới thích hợp cho việc ứng dụng kết quả thu được từ việc rút trích dữ liệu từ âm thanh.
- Xây dựng ứng dụng với bộ thu sóng.
- Tích hợp bộ phát sóng vào ứng dụng.
- Tích hợp và chạy thử nghiệm.

1.4. Nội dung đề tài

Luận văn bao gồm 6 chương:

Chương 1: Giới thiệu tổng quan về các hệ thống cảm ngử cảnh (context-aware systems) và ứng dụng vào thực tế của các hệ thống này. Từ cơ sở đó trình bày lý do và mục tiêu của luận văn.

Chương 2: Khảo sát hiện trạng về các giải pháp có liên quan tới đề tài luận văn. Trình bày tóm tắt việc phân tích một số những dịch vụ và tiện ích trên môi trường thiết bị di động có liên quan đến việc chia sẻ hay xác định thông tin dựa vào âm thanh.

Chương 3: Trình bày về các vấn đề và giải pháp để phát triển ứng dụng.

Chương 4: Vấn đề và giải pháp về việc truyền và nhận thông tin qua âm thanh vô hình đối với người nghe.

Chương 5: Tổng quan về kiến trúc ứng dụng.

Chương 6: Trình bày kết quả đạt được và hướng phát triển của hệ thống.

Chương 2

Khảo Sát Hiện Trạng

 Nội dung Chương 2 khảo sát hiện trạng về các giải pháp có liên quan tới đề tài luận văn. Trình bày tóm tắt việc phân tích một số những dịch vụ, giải pháp và tiện ích hiện có trên môi trường thiết bị di động có liên quan đến việc chia sẻ hay xác định thông tin dựa vào âm thanh, từ đó hướng đến bài toán làm sao xác định ngữ cảnh bên ngoài.

2.1. Dẫn nhập

Mục tiêu đặt ra của luận văn đó là tìm hiểu và đề xuất một giải pháp hiệu quả trên thiết bị di động cho phép xác định ngữ cảnh bên ngoài, mà cụ thể là vị trí, trong môi trường trong nhà. Ý tưởng chính của nhóm là sử dụng trên âm thanh, cho nên trước khi tiến hành đề xuất giải pháp cụ thể (*xem trong chương 4*) và hiện thực hóa trong sản phẩm (*xem trong chương 5*) thì chúng em sẽ trình bày những tìm hiểu của nhóm về các giải pháp có liên quan, cụ thể như sau:

- Tổng quan về thuật toán, thư viện hỗ trợ (tham khảo mục [2.1](#)).
- Sử dụng kết nối Wi-Fi hay Cellular có sẵn của thiết bị (WhatsApp, iMessage) (*tham khảo [2.1.2.1](#), [2.1.2.3](#)*)
- Sử dụng mạng xã hội hay các băng tầng khác nhau 2G, 2.5G, 3G, Wi-Fi, 4G-LTE (Zalo, Facebook Messenger) (*tham khảo [2.1.2.2](#), [2.1.2.4](#)*)
- Sử dụng thông qua kết nối bluetooth có sẵn trên smartphone để gửi tin nhắn, hình ảnh (FireChat) (*tham khảo [2.1.2.5](#)*)
- Xác định ngữ cảnh dựa trên kỹ thuật siêu âm để truyền dữ liệu cho kiểu ứng dụng location-based, định vị trong nhà (ShopKick) (*tham khảo [2.1.2.6](#)*)

Ngoài các giải pháp liên quan đến sử dụng sóng điện từ, nhóm chúng em đã tìm hiểu ứng dụng Digimarc – xác định ngữ cảnh sử dụng digital watermarking cho phép nhúng thông tin, dữ liệu vào nhiều loại nội dung như: giấy in, âm thanh, video, multimedia (*tham khảo [2.1.2.7](#)*). Digimarc không sử dụng những cách tiếp cận

truyền thông trong môi trường trong nhà như localization dạng Wi-Fi hay Bluetooth sử dụng sóng radio mà sử dụng những khía cạnh khác nhau của multimedia – một hướng tiếp cận tiềm năng vì khai thác được nhiều yếu tố ngữ cảnh bên ngoài, cả về âm thanh lẫn hình ảnh. Đây là một sản phẩm thương mại rất thành công trên thực tế. Vì vậy trong nội dung của luận văn này, chúng em dựa vào ý tưởng của bản thân và đi theo hướng tiếp cận như Digimarc nhưng giới hạn bài toán chỉ tập trung trong việc dùng yếu tố là âm thanh. Từ việc tìm hiểu và phát triển một trong những hướng tiềm năng là sử dụng âm thanh này, trong tương lai có thể kết hợp thêm nhiều khía cạnh khác của multimedia để có thể xây dựng một sản phẩm thương mại hoàn thiện như Digimarc.

2.2. Thuật toán FFT, các thư viện, mã nguồn mở hỗ trợ trên nền tảng iOS và các ứng dụng thực tế

2.2.1. Tổng quan các thư viện, mã nguồn mở về FFT

Các thư viện áp dụng FFT cũng có rất đa dạng thuộc nhiều ngôn ngữ lập trình khác nhau. Có thể kể đến như:

- C++
 - Arprec (tác giả: David H. Bailey, Yozo Hida, ...)
 - FFTReal (tác giả: Laurent de Soras)
- C
 - Cross (tác giả: Don Cross)
 - FFTW2 (tác giả: Matteo Frigo, Steven G. Johnson)
 - Kissfft (tác giả Mark Borgerding)
- Objective C
 - vBigDSP (tác giả: Greg Allen – Apple Computer, Inc)
 - vDSP (tác giả Apple Computer, Inc)

Trên nền tảng hệ điều hành iOS, bản thân công ty Apple cung cấp cho các nhà phát triển 2 thư viện để có thể áp dụng vào vấn đề xử lý tín hiệu số: vDSP và Big

vDSP. Những thư viện này cung cấp những hàm xử lý phù hợp cho dải tần số rộng trong các chương trình xử lý tín hiệu với độ chính xác cao.

Vì thế vDSP là lựa chọn thích hợp cho việc phát triển các chương trình ứng dụng kỹ thuật xử lý tín hiệu số trên môi trường hệ điều hành iOS.

2.2.2. Một số ứng dụng thực tế sử dụng kỹ thuật xử lý tín hiệu

2.2.2.1. WhatsApp

Trên thế giới mở đầu cho cuộc cách mạng ứng dụng chat Over-The-Top (OTT) là WhatsApp được phát triển và ra mắt vào năm 2009 bởi công ty cùng tên với 55 nhân viên và tất cả họ trước đây đều từng làm việc cho Yahoo!. WhatsApp được phát hành đầu tiên trên nền hệ điều hành iOS được phát triển và nghiên cứu bởi Apple. Vào ngày 19/2/2014, hãng công nghệ mạng xã hội lớn nhất thế giới Facebook Inc. đã thông báo mua lại WhatsApp với tổng giá trị lên đến 19 tỉ đô-la.



Hình 2-1 Ứng dụng WhatsApp

WhatsApp sử dụng kết nối Wi-Fi hay Cellular có sẵn của thiết bị để gửi tin nhắn đa phương tiện cho người dùng với nhau. Nội dung tin nhắn của WhatsApp bao gồm không chỉ là văn bản thuần túy mà còn có thể là hình ảnh, video, âm thanh và hơn thế nữa là cả vị trí của mình.

WhatsApp đã xử lí đến hơn 10 tỉ tin nhắn một ngày (tháng 8 năm 2012) một số lượng tăng đáng kể từ 2 tỉ tin nhắn (tháng 4 năm 2012) và 1 tỉ tin nhắn vào tháng 10 năm 2012.

Ngày 13 tháng 6 năm 2013, WhatsApp đã đạt kỉ lục 27 tỉ tin nhắn họ đã xử lí. Theo tờ Financial Times, WhatsApp đã chấm dứt việc nhắn tin SMS truyền thống trên điện thoại, việc mà Skype cũng đã làm tương tự như đối với cuộc gọi quốc tế đường dài thông qua điện thoại thông thường. Vào ngày 22 tháng 4 năm 2014, WhatsApp đang có tới hơn 500 triệu người dùng hoạt động thường xuyên hàng tháng, hơn 700 triệu bức ảnh và hơn 100 triệu video được chia sẻ mỗi ngày, và hệ thống quản lý tin nhắn của họ xử lí đến hơn 10 tỉ tin nhắn mỗi ngày. Tháng 5 năm 2014, WhatsApp đã đạt hơn 50 triệu người dùng tích cực hàng tháng tại India, nơi đông người sử dụng dịch vụ của hãng nhất hoạt động liên tục hàng tháng.

WhatsApp sử dụng một chuẩn protocol mở có tên là Extensible Messaging and Protocol (XMPP) được công ty chỉnh sửa lại cho phù hợp với mình. Protocol này sử dụng số điện thoại của mỗi cá nhân để tạo một tài khoản trên server của công ty, và chính số điện thoại này sẽ là username (Jabber ID: [phone number]@s.whatsapp.net). Ứng dụng trên điện thoại của WhatsApp sẽ so sánh một cách tự động tất cả số điện thoại trong danh bạ điện thoại của người dùng với cơ sở dữ liệu trung tâm của WhatsApp để thêm vào danh sách bạn bè trong WhatsApp. Ban đầu phát triển, WhatsApp sử dụng IMEI của điện thoại và MD5-hashed số IMEI đó để làm mật khẩu cho những điện thoại chạy hệ điều hành Android và Nokia Symbians S40; còn trên hệ điều hành iOS, WhatsApp sẽ sử dụng địa chỉ MAC của Wi-Fi thay cho số IMEI như trên điện thoại chạy hệ điều hành Android và S40. Nhưng bắt đầu từ năm 2012 trở đi, server của WhatsApp sẽ tạo ra một mật khẩu ngẫu nhiên và lấy đó làm mật khẩu cho tài khoản của người dùng. Về tin nhắn đa phương tiện, hình ảnh, âm thanh, video sẽ được tải lên một HTTP server và sau đó gửi link cho nội dung đó bằng kỹ thuật Base64 encoded thumbnails.

2.2.2.2. Zalo

Nhắc đến ứng dụng OTT chúng ta không thể không nhắc đến sản phẩm Zalo, một ứng dụng thuần Việt, một niềm tự hào của người Việt đủ sức cạnh tranh và đương đầu với những ứng dụng OTT phổ biến nhất trên thế giới như Viber, Line, Kakao Talk, được phát triển bởi công ty cổ phần VNG.



Hình 2-2 Giao diện ứng dụng Zalo

Zalo là ứng dụng có mức độ phổ biến số một tại Việt Nam với lượng người dùng chính thức đạt 10 triệu vào ngày 20/3/2014 trong đó có 120 triệu tin nhắn trao đổi qua hệ thống mỗi ngày. So với con số 1 triệu người dùng vào tháng 3 năm 2014, VNG đã đưa Zalo lên một tầm cao mới đủ sức đánh bật những đại gia lớn như Line (Công ty Line, Nhật Bản), Kakao Talk (Công ty...), giờ đây thị trường OTT của Việt Nam chỉ là nơi khẳng định vị thế của Zalo và Viber. Zalo đã phủ sóng đến hơn 50% thị trường smartphone ở Việt Nam do được người dùng đánh giá cao bởi tốc độ nhắn tin nhanh và ổn định trên cả ba băng tầng: 2G, 2.5G, 3G và Wi-Fi, cùng với đó là tính năng tin nhắn thoại, gửi nội dung đa phương tiện đa dạng, vẽ tay, hay hơn hết là chức năng nhật ký như một mạng xã hội nhỏ mà người dùng chính là các cá nhân có trong danh bạ điện thoại và cùng sử dụng ứng dụng Zalo.

Phiên bản trước đây của Zalo sử dụng và kế thừa hệ thống đăng nhập và tài khoản từ mạng xã hội Zing Me để phát triển sản phẩm, nhưng vì là mạng xã hội trong nước, khó thu hút người dùng quốc tế, phát triển và cạnh tranh trực tiếp với các đối thủ nước ngoài nên phiên bản sau này của Zalo đã đi theo xu hướng “Mobile First” lấy danh bạ điện thoại làm nền tảng và gặt hái được nhiều thành công cũng như đón nhận của người dùng như hiện nay.

Hiện nay đội ngũ kỹ sư phát triển Zalo đã đưa được ứng dụng lên 4 nền tảng hệ điều hành điện thoại di động phổ biến nhất hiện nay đó là: iOS, Android, Windows Phone và Nokia Symbian.

2.2.2.3. iMessage

Bên cạnh trào lưu OTT theo Mobile First thì dịch vụ tin nhắn tích hợp sẵn trên smartphone của các hãng điện thoại phát triển riêng cho sản phẩm của mình cũng được phát triển rất mạnh như là BlackBerry Messenger hay Apple iMessage.



Hình 2-3 Giao diện ứng dụng iMessage

Ngày 25 tháng 7 năm 2012 Apple chính thức đưa iMessage đến với hệ điều hành Mac OS X Mountain Lion do chính hãng phát triển. Cùng với số lượng lớn người dùng thiết bị có hệ điều hành iOS, iMessage đã ghi được một kỉ lục với hơn 300 tỉ tin nhắn thông qua hệ thống (ngày 23 tháng 10 năm 2012), tính trung bình ra có khoảng 28,000 tin nhắn được gửi đi mỗi giây thông qua iMessage.

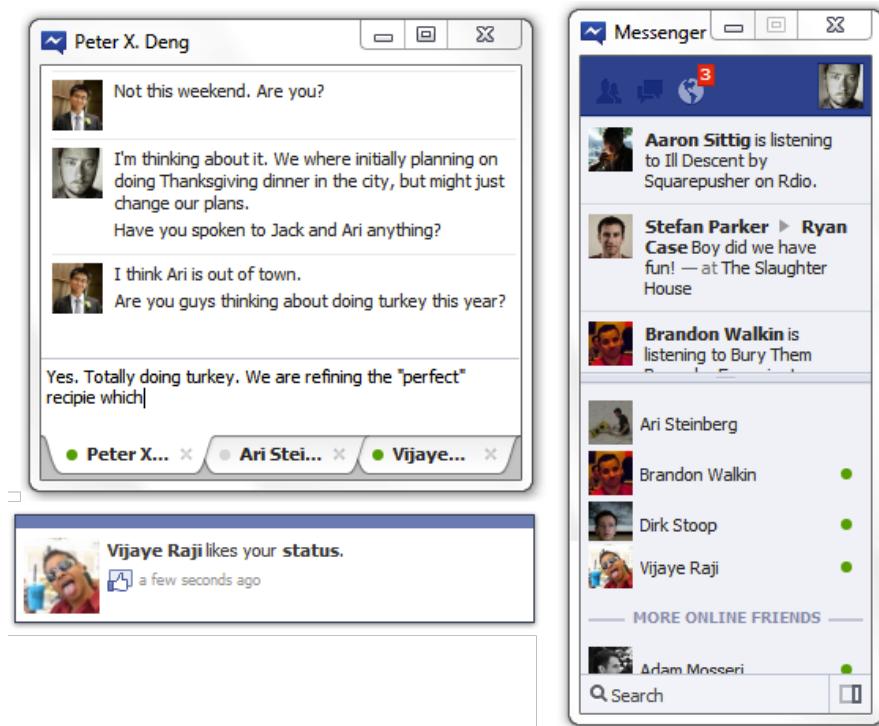
iMessage cho phép người dùng gửi đi tin nhắn văn bản, tài liệu, hình ảnh, video, thông tin liên lạc, nhắn tin theo nhóm thông qua kết nối Wi-Fi, Cellular (2G, 3G hoặc LTE) từ người dùng iOS với nhau hoặc với người dùng OS X. iMessage được kích hoạt thông qua tài khoản Apple ID được đăng nhập sẵn trên thiết bị thông qua số điện thoại đăng ký trước đính kèm cùng với tài khoản đó hay người dùng có thể thêm nhiều e-mails khác. Khi mà điện thoại có kết nối mạng, iMessage sẽ kiểm tra tự động nếu người nhận đã kích hoạt iMessage hay chưa. Nếu đã kích hoạt thì tin nhắn được gửi đi sẽ tự động thông qua iMessage mà không cần thông qua ứng dụng nhắn tin SMS truyền thống, còn lại tin nhắn được gửi đi sẽ là SMS. Vì là được tích hợp và phát triển sâu vào thiết bị chạy hệ điều hành do chính Apple cung cấp nên iMessage sẽ được phát hành miễn phí cùng với hệ điều hành iOS hay OS X. iMessage cũng giống như MMS được tích hợp sẵn trên ứng dụng nhắn tin truyền thống nhưng mà hơn vậy, iMessage còn giúp cho người dùng gửi hình ảnh với dung lượng lớn, phim hay cả vị trí của người dùng một cách nhanh chóng.

iMessage được phát triển trên một protocol nhị phân dựa trên nền tảng service Apple Push Notification Service (APNS) do chính Apple phát triển. Cũng giống như APNS, protocol này của iMessage sẽ tạo một kết nối “Keep-Alive” từ client (thiết bị chạy iOS) đến máy chủ của Apple. Mỗi connection sẽ có một mã riêng biệt đóng vai trò là một số id (định danh) cho mỗi thiết bị để có thể gửi và nhận tin nhắn giữa các thiết bị có iMessage với nhau. Kết nối này sẽ được mã hóa với TLS sử dụng một certificate ở phía thiết bị iOS (client-side certificate) đó là yêu cầu khi thiết bị kích hoạt iMessage. Tin nhắn trên iMessage được mã hóa thiểu kiều end-to-end encryption có nghĩa là không ai có thể truy cập được tin nhắn ngoài người gửi

và người nhận và ngay cả Apple cũng không thể mã hóa dữ liệu và họ không thể ghi lại nội dung tin nhắn.

2.2.2.4. Facebook Messenger

Ngoài ứng dụng chat OTT đang rất phổ biến kể trên chúng ta còn có thể nhắc đến một số ứng dụng khác như Facebook Messenger là một ứng dụng nhắn tin trên điện thoại di động được phát triển bởi Facebook thông qua tương tác chính là mạng xã hội không lồ Facebook, ứng dụng này cập nhật, trao đổi tin nhắn, hình ảnh, video, hay cung cấp chức năng gọi điện thoại VoIP, phát triển trên nền tảng hệ điều hành iOS và Android, hay ứng dụng gọi điện thoại và video Skype được phát triển bởi Microsoft.

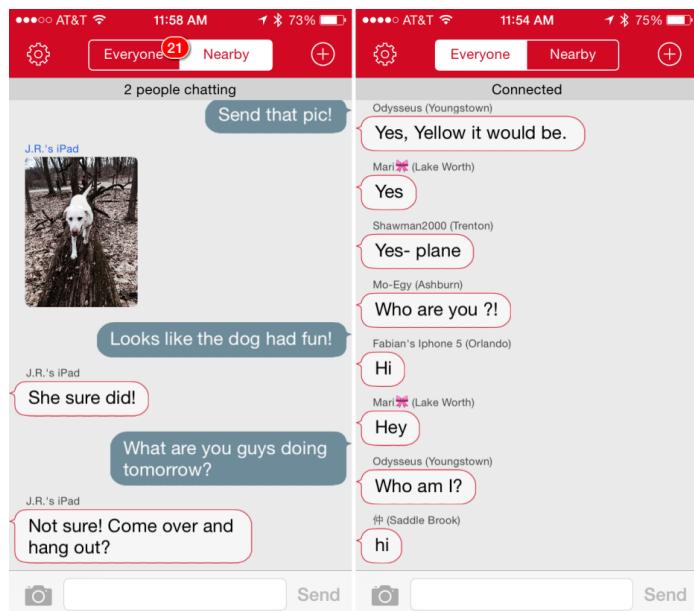


Hình 2-4 Giao diện ứng dụng Facebook Messenger

Tất cả những ứng dụng OTT đều có một đặc điểm chung là yêu cầu một kết nối mạng luôn sẵn sàng (Wi-Fi, Cellular: 2G, 2.5G, 3G, 4G LTE, 4G LTE-A) và cả người gửi, người nhận cần phải thực hiện thao tác cài ứng dụng trên nền tảng hệ điều hành mà ứng dụng đó hỗ trợ (trừ iMessage là được tích hợp trực tiếp vào hệ điều hành iOS mà không cần cài đặt).

2.2.2.5. *FireChat*

Một hướng thứ hai của ứng dụng chat đó chính là thông qua kết nối bluetooth có sẵn trên smartphone để gửi tin nhắn, hình ảnh đi kèm hình là ứng dụng chat Fire Chat. Đây là ứng dụng chat offline (không cần tín hiệu Wi-Fi hay kết nối Cellular như các ứng dụng chat OTT khác) đầu tiên trên Appstore (Cửa hàng ứng dụng dành cho thiết bị hệ điều hành iOS của Apple).



Hình 2-5 Ứng dụng FireChat



Hình 2-6 Mô hình triển khai của FireChat

Tuy nhiên FireChat gặp phải những vấn đề về tính tương thích của chuẩn bluetooth chung giữa các thiết bị với nhau hay giữa tính đóng/mở của các hệ điều hành điện thoại di động. Điểm hình giữa hệ điều hành iOS hay Android mặc dù có

thẻ chuẩn chung bluetooth 4.0 nhưng vì tính đóng và bảo mật của hệ điều hành iOS (mới nhất là iOS 7.0 và những iOS thấp hơn) nên không thể kết nối giữa bluetooth giữa smartphone chạy hệ điều hành iOS và Android, hay nếu trong cùng hệ điều hành Android, các hãng khác nhau sẽ trang bị những chuẩn bluetooth khác nhau cho điện thoại của mình do phần cứng bị phân hóa mạnh nên dẫn đến việc các hãng điện thoại sử dụng hệ điều hành Android khó có một chuẩn chung về kết nối bluetooth.

Vấn đề thứ hai là việc duy trì kết nối bluetooth liên tục có thể dẫn đến tình trạng pin của thiết bị cạn kiệt nhanh chóng. Pin là một vấn đề lớn mà cách hàng sản xuất điện thoại cũng như các nhà lập trình viên đang khắc phục để nâng cao trải nghiệm người dùng. Ứng dụng có thể tốt và thuật toán tối ưu nhưng khả năng tiết kiệm pin không có thì cũng sẽ khó phỏ biến và khó được người dùng chấp nhận sử dụng rộng rãi, thường xuyên trong cuộc sống hằng ngày, khó có tính ứng dụng cao vào thực tiễn cuộc sống, nên vòng đời của ứng dụng trên thị trường sẽ rất ngắn.

2.2.2.6. *ShopKick*

Một ứng dụng khác của việc dùng kỹ thuật siêu âm để truyền dữ liệu cho kiểu ứng dụng location-based đó chính là định vị trong nhà. Như chúng ta cũng đã biết định vị GPS sẽ bị sai lệch lớn khi bước vào những công trình cao ốc, những căn phòng kín, những công trình kiên cố vì sóng GPS khó có thể xuyên qua được tường bê tông dày. Chúng ta có thể thiết lập một thiết bị định phát ra một tần số siêu âm đóng vai trò như một Access Point, cho biết được vị trí đang đứng hiện có của thiết bị phát này. Mỗi tầng hay mỗi văn phòng sẽ có một thiết bị này, người sử dụng có thể định vị được vị trí của mình. Ứng dụng này đã rất thành công và đang được phát triển bởi Công ty ShopKick với sản phẩm cùng tên vào năm 2009.



Hình 2-7 Giao diện ứng dụng ShopKick

ShopKick là một ứng dụng mua sắm cho smartphone/tablet thúc đẩy người dùng mua sắm bằng việc tích lũy điểm khi họ bước vào một cửa hàng có liên kết với ShopKick, những điểm thưởng này gọi là “Kicks”. Người dùng sẽ có được điểm bằng việc quét mã QR sử dụng camera của smartphones và sau đó mua món hàng đó. Những điểm thưởng này sau đó sẽ được dùng cho việc mua sắm những quyển sách, nhạc hay ứng dụng trên iTunes (Nơi cung cấp việc mua, tải về những ứng dụng trực tuyến của thiết bị chạy iOS), cũng có thể là vé xem phim hay nhiều những món hàng khác. Người dùng cũng có thể nhận phiếu giảm giá khi mua hàng từ các đối tác thành viên với ShopKick như: Target, Best Buy, Macy's, Crate & Barrel, Old Navy, American Eagle Outfitters, Sports Authority, Toys R Us and Simon Malls. Cùng với các đối tác đó là khoảng hơn 70 nhãn hiệu lớn khác liên kết, hỗ trợ cho ShopKick bao gồm Procter & Gamble (P&G), Unilever, Mondelez, Colgate, Revlon, Disney, Levi's và HP.

ShopKick được đánh giá là một ứng dụng kích thích người tiêu dùng tốt nhất trong thế giới thực bởi Nielsen vào năm 2012 với hơn 6 triệu người sử dụng thường xuyên sử dụng để nâng cao trải nghiệm mua sắm được tốt hơn.

Không như những ứng dụng location-enabled truyền thống khác, ShopKick không sử dụng GPS để định vị trong nhà bởi vì theo họ, tín hiệu GPS rất không

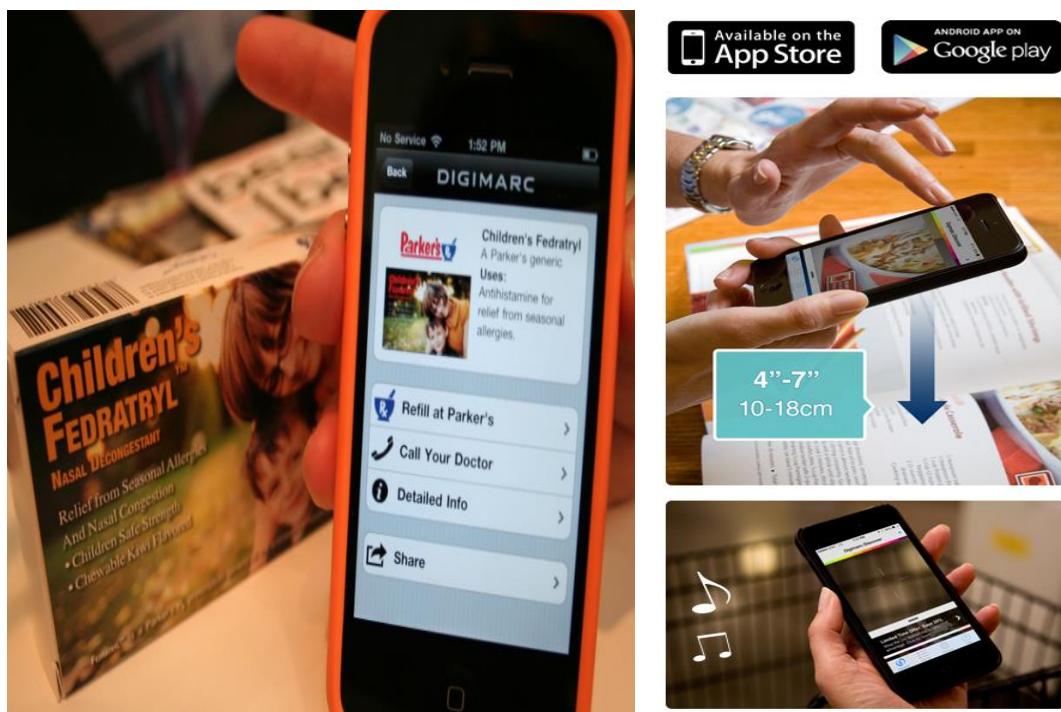
chính xác để xác định vị trí hiện tại của khách hàng trong cửa hàng họ đang mua sắm. Thay vào đó, ShopKick đã dùng một loại sóng tín hiệu khác chính xác cao hơn, và không thể nghe được đó chính là siêu âm. Sóng này sẽ khác nhau ở mỗi cửa hàng thành viên của họ và khách hàng sẽ dùng ứng dụng ShopKick để xác nhận cửa hàng và nhận điểm thưởng khi mua hàng. Khi mà tín hiệu siêu âm của cửa hàng được ứng dụng ghi âm lại được, ứng dụng sẽ ngay lập tức cộng điểm thưởng cho khách hàng khi bước vào cửa hàng, hoặc là khi mua sắm quần áo, quét mã barcode, hoặc những việc khác mà cửa hàng quy định để khuyến khích khách hàng mua sắm. Tín hiệu sóng siêu âm này được truyền đi thông qua hình thức broadcast (truyền liên tục, không phân biệt thiết bị dùng cuối) thông qua một bộ phát tín hiệu đặt trong cửa hàng được phát triển riêng của ShopKick, thiết bị phát này không cần kết nối internet, không cần cài đặt cấu hình phức tạp, chỉ cần năng lượng và khả năng phát nhạc) hoặc tín hiệu này có thể được ghi âm lại và phát trực tiếp thông qua hệ thống âm thanh, giải trí, nghe nhạc của cửa hàng rất tiện dụng.

ShopKick hiện đang được phát triển trên 2 nền tảng hệ điều hành phổ biến nhất trên thị trường smartphone đó là iOS và Android. Và bản phát hành trên 2 nền tảng này là hoàn toàn miễn phí cho người dùng cuối.

Những định hướng mang tính đột phá, mới mẻ của ShopKick về ứng dụng location-based bằng cách sử dụng công nghệ sóng siêu âm gần giống như định hướng của nhóm đã đặt ra khi nghiên cứu tiềm năng ứng dụng rất cao và rất linh hoạt mà sóng siêu âm mang lại. Thành công này mở ra một tương lai mới cho các cách khai thác tiềm năng mới cùng khả năng xử lý mạnh mẽ của smartphone/tablet giúp người dùng có thể trải nghiệm được một cách mới trong việc mua sắm, tham khảo thông tin sản phẩm, hay ứng dụng nhiều hơn ở việc tham gia triển lãm, check-in địa điểm. ShopKick đã đi tiên phong trong lĩnh vực sử dụng sóng siêu âm để truyền thông tin giữa các hệ thống với nhau hay giữa hệ thống với smartphone/tablet mà chi phí rẻ nhất, mức độ phổ biến và linh hoạt sẽ rất cao vì không cần phụ thuộc nhiều vào phần phát và thiết bị thu.

2.2.2.7. *Digimarc*

Trên thế giới quảng cáo là một hình thức phổ biến và hữu hiệu nhất đưa sản phẩm đến tay người tiêu dùng nhanh nhất. Nhưng mà quảng cáo truyền thống qua báo chí, truyền hình kĩ thuật số hay brochure, internet đã không còn hiệu quả nữa. Quảng cáo thường xuyên và dày đặc trên các phương tiện truyền thông ngày nay đã mang lại hiệu quả tiêu cực, khó chịu vì chúng bắt buộc người xem phải tiếp thu một cách thụ động và nhảm chán. Họ không được tích cực chọn lựa cái mình muốn xem, thông tin mình cần và cách thức mình tiếp cận, tương tác với loại quảng cáo đó. Nhận thấy được hạn chế đó Digimarc đã hiện thực được một công nghệ ứng dụng của digital watermarking cho phép nhúng thông tin, dữ liệu vào nhiều loại nội dung như: giấy in, âm thanh, video. Công nghệ do Digimarc phát triển cung cấp giải pháp cho media identification and management (dịch: nhận dạng và quản lý dữ liệu đa phương tiện), counterfeit and piracy deterrence (chống giả mạo, vi phạm bản quyền) và thương mại điện tử.



Hình 2-8 Ứng dụng Digimarc



Hình 2-9 Cách thức hoạt động của Digimarc

Hiện nay Digimarc cung cấp 2 dịch vụ digital watermarking đó là Digimarc Discover và Digimarc for Images.

Digimarc Discover là một ứng dụng trên smartphone để bảo vệ dữ liệu trên bản in, âm thanh được bằng digital watermarking. Những dữ liệu này sẽ qua một bước digital watermarking được thực hiện trên website quản lý tài khoản của Digimarc, sẽ được chèn vào một loại thông tin kỹ thuật số nào đó bằng thuật toán riêng của hãng những đảm bảo nội dung của bản in hay file âm thanh, video không bị thay đổi chất lượng cũng như người bình thường không thể nhận ra được sau khi xử lý tài liệu hay tập tin sẽ được cho người dùng tải về lưu trữ hay phát hành trên internet mà vẫn đảm bảo được tính bản quyền của thông tin đó. Một ứng dụng có tên Digimarc cài trên smartphone sẽ thao tác chụp ảnh hay ghi âm những nội dung đó, nếu chúng được mã hóa thì ứng dụng sẽ tự động giải mã nội dung đó và hiện lên thông tin ẩn chứa bên trong nội dung kỹ thuật số kia. Dịch vụ này cách thức hoạt động tương tự như QR codes, nhưng khác nhau ở việc nhúng liên kết web vào hình ảnh sau đó sử dụng để in.

Dịch vụ thứ 2 mà Digimarc mang đến cho chúng ta đó chính là Digimarc for Images (trước đó có tên là MyPictureMarc) là một công cụ dành riêng cho các nhà nhiếp ảnh được sử dụng để nhúng chữ ký kỹ thuật số của mình vào một bức ảnh kỹ thuật số nhằm bảo vệ tác phẩm của mình trước vấn đề bản quyền một cách tinh tế

nhất mà không làm mất đi giá trị của bức ảnh như những kiểu digital watermarking truyền thống là chèn text hay logo vào bức ảnh đó.

2.3. Kết chương

Qua quá trình đã phân tích và khảo sát hiện trạng về các giải pháp có liên quan tới đề tài luận văn, phân tích một số những dịch vụ, giải pháp và tiện ích hiện có trên môi trường thiết bị di động trong chương 2, chúng em chọn lựa hướng của luận văn là sẽ tập trung khai thác, bước đầu tìm hiểu giải pháp định vị không theo cách truyền thống là sử dụng các thông tin như Wi-Fi, sóng radio ... mà dùng cách tiếp cận mới là multimedia giống như sản phẩm Digimarc đã trình bày ở mục 2.2.2.7. Vì thế đây chính là tiền đề cho việc đề xuất giải pháp và các vấn đề kỹ thuật cần giải quyết trong luận văn (xem trong chương 3 và chương 4). Và với kỹ thuật được đề xuất đó sẽ được hiện thực hóa bằng sản phẩm trong chương 5.

Chương 3

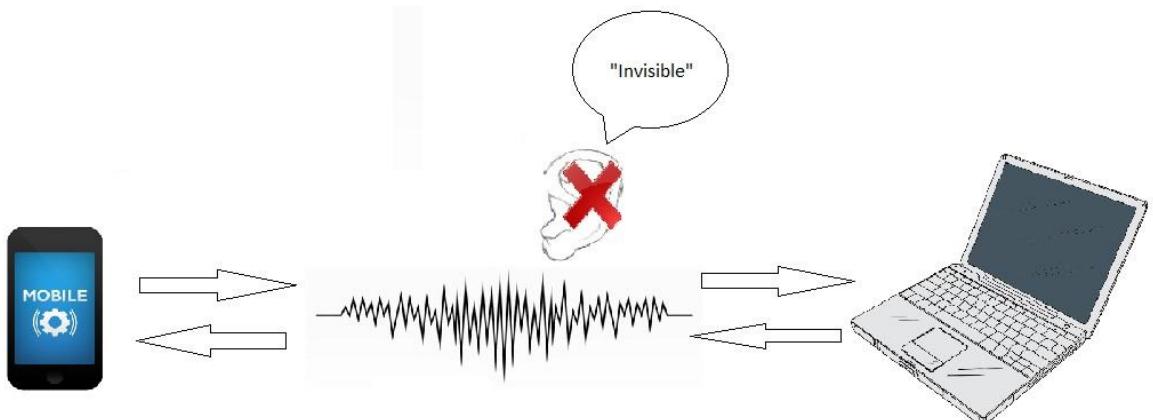
Các vấn đề và giải pháp kỹ thuật để phát triển ứng dụng

☞ Chương 3 trình bày các vấn đề về công nghệ, kỹ thuật và giải pháp trong quá trình xây dựng ứng dụng trên nền tảng iOS.

3.1. Vấn đề truyền nhận tín hiệu qua tín hiệu âm thanh

Ý tưởng chính của giải pháp được đề xuất sẽ phải là có thể truyền được từ một thiết bị phát đến một thiết bị thu qua môi trường bình thường với cách truyền là dùng sóng âm thanh.

Giải pháp chính này được thể hiện qua hình 3.1



Hình 3-1 Nguyên lý cơ sở của ứng dụng

Mục tiêu của nhóm là tận dụng những thiết bị vốn có sẵn về mặt tin học và làm sao để có thể phát truyền dữ liệu mà tai người không thể cảm nhận hay phân biệt được. Từ đó phục vụ, đáp ứng người dùng về nhu cầu tìm kiếm những thông tin kèm thêm qua các ứng dụng location-based. Như đã phân tích trong *phần 1.1* của đề tài thì một trong hai loại thiết bị phải là thiết bị di động mà không phải là một thiết bị gì hơn – đảm bảo tính ứng dụng cao vì người sử dụng không cần gì hơn là chính chiếc điện thoại của mình. Vì vậy trong quá trình thiết kế, thiết bị thu là chiếc điện

thoại di động và trong một số tình huống cụ thể, thiết bị thu này cũng có thể trở thành thiết bị phát âm thanh.

3.2. Lựa chọn giải pháp

Hiện tại trong phạm vi của đề tài luận văn thì sẽ tập trung tìm hiểu các thiết bị di động đang được sử dụng khá phổ biến là smartphone iOS. Chính vì vậy những vấn đề được trình bày tiếp theo sẽ liên quan đến vấn đề kỹ thuật đặc thù của môi trường iOS. Còn những vấn đề đặc thù trên các môi trường khác như Android hay Windows Phone sẽ có vài điểm tương đồng hay khác biệt tùy thuộc vào thiết kế của giải pháp. Nhưng trong luận văn chỉ giới hạn trên nền tảng iOS – nền tảng chiếm 42.6% thị phần tại Mỹ (10/2013) và đang rất được ưa chuộng tại Việt Nam.

3.2.1. Hạn chế do bộ filter của các thiết bị sử dụng nền tảng iOS



Vấn đề

Theo các tài liệu tiêu chuẩn về vật lý và các tài liệu tham khảo, chúng ta đã biết những thông tin về tần số liên quan đến khả năng nghe của con người. Bao gồm như:

- Miền tần số tai người bình thường có thể nghe: 20 – 20000Hz (Dưới 20Hz là hạ âm, trên 20000Hz là siêu âm, ở cả 2 miền tần số này tai người không thể nghe được).
- Miền tần số giọng nói của con người: 80 – 1100Hz.
- Miền tần số có ngưỡng nghe (giới hạn độ to của âm mà tai người nghe bắt đầu có cảm giác khó chịu) nhỏ nhất (tai người dễ dàng nghe được nhất) là từ 3000 – 4000Hz (xét người ở độ tuổi 20).

Từ đó có thể thấy rằng trong lĩnh vực về tần số, có 2 yếu tố ảnh hưởng đến khả năng nghe của con người. Đó là độ to của âm và tần số của âm.

Ứng dụng của nhóm lại sử dụng phương pháp chuyển đổi, gửi, nhận và phân tích dữ liệu từ tần số. Vì thế dải tần số được chọn, khả năng nghe của người dùng và kh

năng thu, phát âm thanh của thiết bị được chọn có ảnh hưởng trực tiếp đến kết quả và độ chính xác của ứng dụng.

Cũng bởi chính điều đó, và do nhóm thực hiện xây dựng ứng dụng trên thiết bị sử dụng hệ điều hành iOS của Apple nên có 1 vấn đề, đó chính là hạn chế của bộ filter âm thanh sẵn có trên thiết bị (bao gồm low-bass filter và high-bass filter). Thiết bị iOS được sản xuất vì yếu tố hướng tới người dùng nên đã được tích hợp các bộ filter để lọc (trong cả việc thu lẫn phát) âm thanh được thu / phát. Việc đó đã dẫn đến các tần số quá cao hoặc quá thấp sẽ bị lọc và không có khả năng thu / phát âm thanh ở những miền tần số đó. Cụ thể (theo kết quả thực nghiệm):

- Phát:

- o Low-bass: các âm ở tần số dưới 30 không thể phát.
- o High-bass: các âm ở tần số trên 20000 không thể phát.

- Thu

- o Low-bass: các âm ở tần số dưới 50 không thể thu.
- o High-bass: các âm ở tần số trên 6000 không thể thu (nếu thu được tần suất thu được cũng không cao).

Sóng âm ở tần số cao hoặc thấp sẽ làm cho người dùng khó nghe thấy, từ đó có thể thấy rằng sử dụng những âm thanh ở những tần số này có thể làm tăng khả năng sử dụng tần số của chương trình và độ chính xác cũng sẽ cao hơn.

Vậy phải làm sao để có thể khắc phục điều này để giúp cho ứng dụng có thêm nhiều khả năng phát triển?

Giải pháp

Có thể thấy rằng không chỉ riêng hệ điều hành iOS mà các hệ điều hành khác đều hướng tới người dùng. Các bộ filter cũng sẽ được áp dụng trên các thiết bị - nhất là điện thoại di động. Vì thế nhóm có 3 hướng khắc phục:

- Sử dụng 1 thiết bị có khả năng thu / phát khả thi hơn: Arduino là 1 phương án khả thi.

- Vẫn tiếp tục sử dụng các thiết bị di động kèm theo với 1 phương thức mới. Một phương thức xác nhận kiểm tra hoặc 1 phương thức để đồng bộ giữa 2 quá trình xảy ra đồng thời là thu và phát.
- Về mặt lý thuyết, các thiết bị iOS đã cho phép phát ở tần số cao, vì thế rõ ràng nếu có thể cài tiến thuật toán thì vẫn có thể thu được âm có tần số cao.

Sử dụng Arduino có lợi thế vì Arduino hoàn toàn không có giới hạn ở các bộ filter. Tuy nhiên do để có thể thao tác và sử dụng Arduino, chúng ta cần biết các kiến thức về điện tử viễn thông cũng như cách thức để lập trình tương tác với chúng. Điều này là 1 khó khăn lớn vì Arduino còn khá mới và các source code để sử dụng chúng có thể còn chưa ổn định cũng như khó tiếp cận đối với người mới bắt đầu.

Vì thế nhóm hướng tới cách thứ hai. Thực tế, chúng ta có thể sử dụng chính việc thu phát âm thanh và phân tích tần số để tạo ra phương thức mới (bao gồm thu, nhận, chứng thực, đồng bộ) cho việc thu phát âm.

Ví dụ như thay vì ghi nhận 1 tần số biểu diễn 1 data nào đó, ứng dụng phải thu được 2 – 3 lần liên tiếp 1 tần số chứng thực nào đó như mã mở đầu cho data. Cũng như phải thu 2 – 3 lần liên tiếp 1 tần số kết thúc. Có thể thấy dữ liệu được truyền đi sẽ có dạng:

YY(Y) – data – ZZ(Z)

Vấn đề phát sinh khi đó sẽ là muốn truyền 1 data ta cần dùng khá nhiều “tín hiệu” dẫn tới số tín hiệu sẽ tăng theo cùng với số lượng data được gửi đi.

- Giữa độ chính xác và số data được gửi, ta phải cân nhắc chọn xem yếu tố nào được ưu tiên hơn tùy theo yêu cầu sử dụng của ứng dụng.

Một cách khác được nhóm sử dụng trong quá trình đồng bộ đó là phương thức truyền dữ liệu mô phỏng quá trình bắt tay 3 bước trong giao thức TCP / IP:

- Receiver chờ nhận ACK_Sender_Frequency.
- Sender gửi ACK_Sender_Frequency sau đó chờ nhận ACK_Receiver_Frequency.

- Receiver nhận ACK_Sender_Frequency, sau đó gửi ACK_Receiver_Frequency.
- Sender nhận ACK_Receiver_Frequency.
- Hoàn thành quá trình bắt tay.

Ưu điểm của phương pháp này là quá trình kiểm tra gắt gao hơn, chính xác hơn. Nhưng nhược điểm là dù sau khi bắt tay thì vẫn cần 1 mã bí mật trong tất cả các phiên giao dịch để tránh bị ảnh hưởng bởi tiếng ồn cũng như nhiễu âm. Và quá trình đồng bộ sau đó cũng cần có sự khéo léo của người lập trình. Hơn nữa vì quá trình thu phát để đồng bộ làm cho việc giao tiếp kéo dài, dẫn đến trong quá trình chờ, khả năng thiết bị “nghe” được những tạp âm sẽ cao hơn so với truyền dữ liệu không có xác nhận.

Một khía cạnh khác của các ứng dụng “thu - phát - phân tích” đó là độ to (tính theo dB) của âm. Tầm nghe của con người khoảng từ 0 đến 125 dB. Dưới 40 dB thì nghe rất khó còn trên 105 dB thì tai sẽ bị đau đớn và trên 115 dB trong khoảng thời gian dài thì sẽ bị điếc vĩnh viễn. Vì thế các thiết bị cũng bị giới hạn về khả năng “nghe” này. Theo thực nghiệm, các âm có độ to dưới 55 dB sẽ thường bị xem là như tạp âm do tiếng ồn gây ra, các thiết bị sẽ không thu hoặc thu được rất ít chúng. Các âm có độ to trên 90 dB sẽ gây ra khó chịu cho người sử dụng. Vì thế nhóm đã thực nghiệm tiếp tục trên dải độ to của âm từ 65 - 80 dB và thu được những kết quả sau:

- Độ to nhỏ nhất để máy còn có khả năng thu được vào khoảng 70 dB, ở độ to này, con người tương đối khó nghe nhưng các thiết bị cũng thu không thực sự nhạy.
- Độ to cao nhất để con người cảm thấy không khó chịu vào khoảng 80 dB, ở độ to này hầu hết các thiết bị đều thu nhạy nhưng con người nghe cũng nhạy. Vì thế phát âm trong thời gian dài có thể gây khó chịu cho người dùng.

Về phương pháp thứ 3, nhóm đã tìm hiểu được cách sử dụng framework Accelerate của iOS kết hợp với thuật toán FFT để cải thiện quá trình thu âm ở tần số cao và đã thu được 1 số kết quả nhất định.

Theo thực nghiệm, khi áp dụng cách thức cải tiến mới, tần số của âm có thể thu được thuộc khoảng 30 - 20000 Hz. Với dải tần số cao trên 18000 Hz, tai người rất khó để nghe được, nhất là trong điều kiện cùng nghe với những âm thanh khác. Từ đó có thể thấy rằng với thuật toán mới, kỹ thuật “Auditory Masking” có thể sử dụng để khiến người dùng hoàn toàn không nghe thấy tín hiệu data cần truyền. Cũng theo thực nghiệm, khoảng cách thiết bị có thể nhận được sóng cao tần vào khoảng 6m và có 1 vật cản ngăn cách ở giữa (1 tấm cửa kiếng).

Với âm cao tần, độ dài của tín hiệu cũng có thể được giảm xuống rõ rệt. Nguyên nhân là vì âm ở tần số cao rất khó trong điều kiện bình thường có thể tạo ra. Vì thế vấn đề tạp âm và tiếng ồn đều có thể được xử lý. Ví dụ dữ liệu được truyền đi sẽ có dạng ngắn gọn hơn trước, cụ thể:

Y - data (với Y là tín hiệu cao tần khác có nhiệm vụ như 1 tín hiệu báo sự bắt đầu có data được truyền đến).

Việc áp dụng được tần số cao tần trong việc truyền dữ liệu tạo ra những khía cạnh tích cực cho việc phát triển các ứng dụng “truyền - nhận - phân tích”. Cụ thể ứng dụng có thể được phát triển theo các hướng như:

- Location based application: những ứng dụng sử dụng khoảng cách mà âm có thể được nhận biết để biết người dùng có từng “ghé thăm” nơi phát ứng dụng hay không (ứng dụng check-in, lấy voucher bằng cách tới cửa hàng trong ngày khuyến mãi, ...)
- Watermarking: âm cao tần có tác dụng đánh dấu giống như 1 phần của kỹ thuật Watermarking. Lợi dụng điều này, chúng ta có thể phát triển những ứng dụng để nhận biết “dữ liệu bí mật” được ẩn chứa trong âm thanh.
- Chat application: những ứng dụng truyền đi 1 lượng lớn dữ liệu bằng các âm cao tần (những âm tai người không thể nghe). Bằng những âm cao tần này, chúng ta đã có thể tạo ra 1 phương thức truyền dữ liệu mới bên cạnh các ứng dụng sử dụng Internet, Wi-Fi hay Bluetooth. Đây là hướng trước mắt mà nhóm chọn lựa để phát triển.

Với phương pháp thứ 3, rõ ràng nhiều vấn đề đã được giải quyết. Việc truyền dữ liệu cũng được cải thiện rõ rệt ở hiệu suất cũng như tính bảo mật. Tuy nhiên vẫn còn 1 số vấn đề cần được khắc phục bằng sự khéo léo về lập trình ở các lập trình viên để có thể tận dụng, khai thác phương pháp này. Từ đó có thể phát triển các ứng dụng “truyền - nhận - phân tích”.

3.2.2. Vấn đề phụ thuộc vào engine để phát triển ứng dụng



Vấn đề

Game engine cung cấp một bộ các công cụ phát triển trực quan và có thể tái sử dụng từng thành phần trong đó. Các bộ công cụ này cung cấp một *môi trường phát triển tích hợp* được đơn giản hóa. Phát triển ứng dụng nhanh cho game theo cách lập trình hướng dữ liệu. Những game engine này đôi khi còn được gọi là các "phần mềm trung gian cho game" (*game middleware*), như ý nghĩa của thuật ngữ, chúng cung cấp một nền tảng phần mềm linh hoạt và dễ dàng sử dụng lại với mọi chức năng cốt lõi cần thiết ngay trong nó để có thể phát triển một ứng dụng game đồng thời giảm giá thành, độ phức tạp, và kịp thời hạn phát hành - tất cả các yếu tố quan trọng trong ngành công nghiệp game đầy cạnh tranh.

Với công nghệ tạo game engine càng phát triển và trở nên thân thiện hơn với người sử dụng, ứng dụng của nó càng được mở rộng, và giờ đây được sử dụng để tạo các game mang mục đích khác với giải trí đơn thuần như: mô phỏng, huấn luyện ảo, y tế ảo, và mô phỏng các ứng dụng quân sự. Để tạo điều kiện tiếp cận này, các nền tảng phần cứng mới đang là mục tiêu của game engine, kể cả điện thoại di động (ví dụ: iPhone, điện thoại Android) và trình duyệt web (như Shockwave, Flash, WebVision của Trinigy, Silverlight, Unity Web Player, O3D và thuần dhtml).

Cũng chính vì nguyên nhân đó, nhóm thực hiện đã sử dụng các engine của hệ điều hành iOS mang lại như: cocos2d-iphone, spritebuilder, ... Những ứng dụng và công nghệ trên giúp ứng dụng được nhanh chóng xây dựng hơn cũng như cung cấp mạnh mẽ các giao diện tương tác giúp lập trình viên có thể dễ dàng thiết lập môi

trường, dàn cảnh, điều chỉnh cấu hình cần thiết trong các cảnh game (Game scene). Điều này tuy có nhiều điểm lợi nhưng cũng có không ít khó khăn.

Vậy những khó khăn đó là gì và làm sao khắc phục để thực sự có thể sử dụng tốt các game engine và các tool hỗ trợ?

Giải pháp

Cocos2d-iphone là 1 framework cho phép lập trình viên dễ dàng sử dụng để phát triển nhanh 1 ứng dụng trên iphone/ipad. Tuy nhiên khi sử dụng framework này chúng ta khó có thể làm cross-platform để ứng dụng có thể chạy trên nhiều hệ điều hành khác nhau như Android, héWindowPhone, ...

Ngoài ra, hiệu suất xử lý của ứng dụng và quá trình quản lý bộ nhớ cũng sẽ phụ thuộc vào framework. Nếu người lập trình xử lý không tốt, có thể sẽ xảy ra các vấn đề memory leak, tràn bộ nhớ, v.v...

Spritebuilder là tool hỗ trợ có khả năng kết hợp làm việc với XCode trên Mac OS. Với Spritebuilder, chúng ta có thể dễ dàng dàn cảnh, cấu hình, xây dựng các màn chơi game dễ hơn bao giờ hết. Tuy nhiên vẫn đề sử dụng spritebuilder sẽ làm cho hiệu suất load game chậm hơn so với việc tự quản lý tài nguyên trong game và ứng dụng. Thực nghiệm cho thấy sử dụng spritebuilder sẽ làm ứng dụng load chậm hơn từ 2-3 so với bình thường.

Để khắc phục những vấn đề trên, nhóm chỉ sử dụng các thành phần cơ bản của Spritebuilder và tăng cường quản lý tài nguyên, quản lý bộ nhớ trong ứng dụng, sử dụng các UIViewController thay cho các màn hình game của game engine khi không thật sự cần chúng. Điều đó đã làm tăng hiệu suất hoạt động của ứng dụng. Trong ứng dụng nhóm đã phát triển, có sử dụng vòng lặp khi nhận tín hiệu. Ở hàm này phải đặc biệt chú ý về vấn đề cấp phát và giải phóng tài nguyên. Việc cấp phát và giải phóng tài nguyên ở những vòng lặp có thể dẫn tới memory leak. Tuy bản thân XCode có hỗ trợ cơ chế ARC (Automatic Reference Counting), nhưng việc

kiểm tra kỹ các tài nguyên có kích thước lớn khi thoát chương trình giúp ích rất nhiều cho hiệu suất làm việc của ứng dụng.

3.2.3. Khó khăn về việc đọc các dữ liệu ghi âm (digital signal) và phân tích chúng.

Giải pháp

Việc truyền dữ liệu ngày nay gắn liền với rất nhiều công nghệ với các kỹ thuật khác nhau. Có thể nói đến như Internet, Wireless, Bluetooth (3.0, 4.0), NFC, 3G, v.v... Tuy nhiên với mỗi phương pháp đều có ưu và nhược điểm riêng biệt của chúng. Internet có tốc độ cao nhưng không phải bất cứ nơi đâu cũng có thể kết nối Internet. Wireless cho kết nối không dây, lại có tốc độ cao nhưng tùy thuộc vào access point cũng như các điều kiện liên quan. Bluetooth là phương pháp giải quyết vấn đề truyền dữ liệu giữa các thiết bị trong điều kiện không có kết nối dễ dàng nhưng lại tiêu tốn năng lượng để truyền tải dữ liệu cũng như dễ dàng mất dữ liệu hoặc kết nối trong quá trình truyền dữ liệu có kích thước lớn.

Vì những yếu tố đó, nhóm thực hiện đã quyết định chọn một phương pháp mới để thực hiện điều mà nhóm muốn hướng tới – một phương pháp truyền dữ liệu tốn ít năng lượng, có thể truyền trong điều kiện thiếu các kết nối cũng như có thể giúp các thiết bị khác nhau về platform dễ dàng có một cách thức hay giao thức tiện lợi để trao đổi thông tin. Và phương pháp được chọn đó là xử lý truyền dữ liệu bằng cách truyền, nhận và phân tích tín hiệu âm thanh. Cách thức này giúp chúng ta có thể độc lập với platform vì bản chất của âm thanh đúng là như vậy. Chỉ với một “máy phát” và một “máy thu” (có thể là 2 thiết bị di động, hoặc 2 bộ phát vật lý thật sự) với các thuật toán phân tích dữ liệu, chúng ta có thể dễ dàng nhận ra được thông điệp mà “máy phát” đang phát và rút trích các yếu tố có liên quan. Phương pháp mà Digimarc sử dụng đó là lồng ghép một tín hiệu, một ID vào một tín hiệu khác (một sóng mang hay một carrier), dựa vào kỹ thuật Watermarking để nhận ra ID đó và thể hiện một thông điệp nào đó cho người sử dụng. Cách tiếp cận của nhóm thực

hiện của nhóm cũng có những nét tương đồng. Nhóm sử dụng âm cao tần được phát đi để tại “máy thu” có thể thu được âm thanh cao tần đó. Từ đó xác định thông điệp có liên quan. Ưu và nhược điểm của phương thức này có thể kể đến như:

- Ưu:

- Âm thanh cao tần có thể vượt qua khả năng nghe của con người.
- Vì tai người nghe không được nêu tín hiệu có thể có kích thước lớn mà người dùng không hề hay biết.
- Khả năng xuất hiện âm thanh có tần số cao (trên 18500) trong điều kiện bình thường khá thấp → Giảm thiểu khả năng bị nhiễu.

- Nhược:

- Vẫn chưa hoàn toàn có thể độc lập với thiết bị, với những thiết bị không cho phép phát hoặc thu âm cao tần, phương pháp này sẽ có những hạn chế nhất định.
- Thuật toán được chọn, độ phức tạp của thuật toán cũng như độ chính xác của quá trình phân tích là những vấn đề quan trọng có liên quan.
- Tuy khả năng mắt dữ liệu trong quá trình truyền tải là không cao (do khả năng bị nhiễu tương đối thấp) nhưng vẫn có sai số trong một khoảng nào đó → Khó có thể đảm bảo quá trình truyền dữ liệu chính xác bằng những phương pháp truyền thống.
- Việc truyền âm thanh cao tần trong một khoảng thời gian quá dài có thể ảnh hưởng đến sức khỏe của người sử dụng, nhất là có thể gây khó chịu cho những người mẫn cảm có khả năng nghe những âm thanh cao tần.

Đối với thuật toán được sử dụng cho vấn đề xử lý tín hiệu số, nhóm thực hiện đã chọn thuật toán FFT (Fast Fourier Transform). Nhưng những thư viện sử dụng FFT làm cẩn bản cũng có rất nhiều. Có thể phân loại như sau:

- C++

- Arprec (tác giả: David H. Bailey, Yozo Hida, ...)
- FFTReal (tác giả: Laurent de Soras)
- C
 - Cross (tác giả: Don Cross)
 - FFTW2 (tác giả: Matteo Frigo, Steven G. Johnson)
 - Kissfft (tác giả Mark Borgerding)
- Fortran
 - Dfftpack (tác giả: Paul N. Swarztrauber, Hugn C. Pumphrey)
- ...

Trong những thư viện đó, nhóm đã chọn thư viện vDSP do chính công ty Apple là tác giả. Sở dĩ thư viện vDSP được chọn vì bởi nhóm thực hiện ứng dụng ngay trên thiết bị di động của Apple (iPhone, iPad) nên việc sử dụng thư viện của chính Apple sẽ có khả năng tương thích tốt hơn, cũng như dễ áp dụng hơn.

Đối với thư viện vDSP, qua quá trình thực nghiệm, có thể thấy rằng độ chính xác của tín hiệu thu được phụ thuộc vào những yếu tố như: dải tần số được chọn (trong những tần số cao tần), điều kiện xung quanh trong quá trình truyền và thu. Nhưng nhìn chung qua quá trình thực nghiệm có thể thấy rằng tần số thu được có sai số trong khoảng ± 15 và dải tần số có xác suất thu được cao nhất trong những âm cao tần mà tai người khó nghe là 18000 – 19000Hz.

Trong việc thu và đọc các dữ liệu ghi âm cũng như phân tích chúng thì vấn đề khó khăn mà nhóm phải giải quyết là việc thu và ghi nhận trong thực tế diễn ra “real-time”, nghĩa là chúng ta buộc phải ghi nhận dữ liệu và xử lý trong lúc nhận ra và thu được chúng từ các âm trong không gian. Trong quá trình áp dụng thuật toán fft, thư viện vdsp vào ứng dụng thực tế, nhóm đã nhận thấy từ hàm callback được gọi khi phát hiện có dữ liệu thỏa mãn điều kiện tồn tại trong không gian xung quanh, dữ liệu phải được xử lý một cách đồng bộ cẩn thận. Hơn nữa, ứng dụng còn

kèm theo khả năng trở thành “máy phát”. Do đó khả năng ứng dụng “tự nghe” được âm thanh do chính bản thân phát ra là điều hoàn toàn có khả năng.

Vậy làm sao để chúng ta có thể phân biệt được âm thanh nào là do bên A phát, tín hiệu nào là do bên B phát, đồng thời thực hiện đồng bộ hóa cho tất cả các quá trình xảy ra hoàn toàn real-time?

Theo lẽ thông thường mà nói, ngay cả chính bản thân con người chúng ta đôi khi cũng khó phân biệt được tiếng nói là do ta hay một ai khác lên tiếng mà tạo thành, vậy làm sao để máy tính có thể phân biệt được một âm thanh ở tần số nào là do thiết bị nào phát ra? Có hai cách khả thi mà nhóm hướng tới. Cả hai đều sử dụng một “khóa” cho quá trình ghi âm. Chỉ khi “khóa” được ghi nhận, những tín hiệu sau đó mới có ý nghĩa.

Với cách thứ nhất, khóa được truyền đúng theo cách mà dữ liệu được truyền. Ưu điểm là dễ cài đặt, dễ thực hiện vì cách thức hoàn toàn đúng theo phương pháp ban đầu. Nhược điểm là có sai số như đã nói, là ± 15 . Sai số này đã làm cho số tín hiệu có thể truyền đi bị giảm khoảng từ 10 – 15 lần, vì chắc chắn rằng chúng ta không muốn nhận sai tín hiệu được truyền đi do sai số gây ra nên những ID được encode phải cách nhau ít nhất 30 đơn vị để độ chính xác cao (theo thực nghiệm cho thấy với khoảng cách 30 đơn vị này, độ chính xác lên tới khoảng 95% tương ứng cứ 20 lần truyền dữ liệu thì có không quá 1 lần không nhận được).

Với cách thứ hai, khóa được “nhúng” vào âm thanh bằng kỹ thuật Watermarking. Ưu điểm là sai số gần như không có, độ chính xác lại cao, vì thế số ID có khả năng được truyền tăng lên rất nhiều so với cách thứ nhất. Nhược điểm là cách thức cài đặt có khác biệt với phương thức truyền dữ liệu. Vì thế để có thể áp dụng và cài đặt kết hợp cũng như xử lý đồng bộ không dễ dàng, đặc biệt là xử lý đồng bộ real-time cùng với việc thu dữ liệu.

Nhóm đã thực hiện theo cách thứ nhất và sử dụng kỹ thuật khóa để kiểm tra tín hiệu thu được. Đồng thời để giải quyết vấn đề đồng bộ real-time, nhóm đã sử dụng mô hình trạng thái (state). Cụ thể, mỗi một giai đoạn của quá trình thu nhận tín hiệu

sẽ ứng với một trạng thái. Ở mỗi trạng thái, tín hiệu nhận được chỉ chấp nhận 1 giá trị khóa duy nhất. Khi nhận được giá trị này, trạng thái sẽ chuyển sang giai đoạn kế tiếp. Sau một khoảng thời gian nhất định, nếu trạng thái chưa được chuyển tiếp (chứng tỏ quá trình nhận dữ liệu đã xảy ra vấn đề) thì sẽ quay về trạng thái bắt đầu. Với mô hình đa trạng thái ứng với đa khóa như vậy, chỉ khi tất cả các khóa được thỏa mãn thì tín hiệu mang dữ liệu mới được chấp nhận → giảm thiểu khả năng nhận nhầm hay “nghe” nhầm chính tín hiệu của bản thân phát ra. Bên cạnh đó, biến trạng thái luôn phải được xử lý bằng cách kỹ thuật đồng bộ (mutex, semaphore, ...) để đảm bảo không xảy ra tình trạng “race condition”. Bởi vì vai trò của biến trạng thái là rất quan trọng.

Tuy nhiên cả hai cách thức đều có nhược điểm đó là khi khóa vô tình “bị phá hủy” (thu được một âm thanh có cùng tần số với khóa quy định) thì vấn đề sai lệch dữ liệu vẫn có thể xảy ra. Điều này lại càng dễ xảy ra khi hai thiết bị sử dụng cùng ứng dụng đang có vai trò như nhau (cùng phát và cùng thu). Trong trường hợp đó, khóa có thể bị phá bởi chính âm thanh của bản thân ứng dụng và dữ liệu của bản thân lại vô tình làm sai lệch dữ liệu được truyền đi bởi đối phương. Để giảm thiểu khả năng này, nhóm đã thực hiện phân nhóm cho tín hiệu được truyền đi. Có nghĩa tín hiệu bên A và bên B phát sẽ có giá trị tách biệt. Tuy số tín hiệu có thể truyền bị giảm nhưng khả năng bị nhiễu được giảm rõ rệt. Tuy nhiên phương pháp này chỉ thực sự có ý nghĩa khi dữ liệu được truyền giữa hai thiết bị. Nếu dữ liệu được broadcast, chúng ta phải cần có một phương pháp hữu hiệu hơn.

3.3. Kết chương

Trong Chương 3 chúng em đã trình bày chi tiết về các kĩ thuật và giải pháp trong quá trình xây dựng ứng dụng trên nền tảng iOS. Chương tiếp theo chúng em sẽ trình bày về vấn đề truyền và nhận âm thanh “vô hình” đối với người nghe.

Chương 4

Vấn đề và giải pháp về việc truyền và nhận thông tin qua âm thanh vô hình đối với người nghe

☞ Chương 4 trình bày các vấn đề về công nghệ, kỹ thuật và giải pháp trong truyền nhận tín hiệu một cách “vô hình” – người dùng không hoặc khó có thể nghe được.

Trong nội dung của Chương 3, chúng em đã trình bày các vấn đề kỹ thuật và giải pháp trong vấn đề truyền và nhận âm thanh tập trung vào thiết bị thu là các thiết bị thông thường trên nền tảng iOS (iPhone, iPad, ...) – những thiết bị được sử dụng rất phổ biến. Đây là các vấn đề phụ thuộc vào thiết bị và nền tảng. Nhưng để hiện thực hóa được ý tưởng, xây dựng được ứng dụng như trong Chương 5 trình bày thì cần có những giải pháp về mặt phương pháp để có khả năng truyền được tín hiệu thông tin qua kênh truyền là âm thanh mà người dùng khó có khả năng cảm thấy được sự tồn tại của loại âm thanh đặc biệt (tín hiệu đặc biệt). Chính vì vậy nội dung của Chương 4 sẽ trình bày vấn đề và giải pháp dựa vào ý tưởng dùng âm thanh tai người khó có thể nghe được. Những giải pháp này là độc lập với thiết bị, môi trường hay nền tảng.

4.1. Steganography và Watermarking

4.1.1. Tổng quan về Steganography - ẩn thông tin trong thông tin

Khái niệm ẩn thông tin trong thông tin (data hiding hay steganography) được hiểu là cố gắng ẩn một thông tin (watermark) vào một thông tin nào đó khác (có khả năng là âm thanh, hình ảnh, văn bản) đang mang nó.

Phương pháp này khai thác tính chất bản thân thông tin multimedia không cần phải được truyền một cách trọn vẹn mà có khả năng bị sửa đổi một phần mà cảm nhận của con người đối với thông tin bị biến đổi đó vẫn không có sự khác biệt đáng kể.

Đối với âm thanh thì điều này đã từng được sử dụng cho thao tác nén. Ví dụ như mp3 – một thao tác nén có khả năng đạt dung lượng 1/12 của wav. Cách nén mp3 là sử dụng tính chất có một số tần số của âm thanh không nghe được (dưới 20 Hz và trên 20000 Hz), những âm thanh này sẽ bị loại bỏ. Ngoài ra còn có tính chất một số thành phần âm thanh ở một số miền tần số nhất định có thể sẽ không thể nghe được vì có những âm thanh ở các miền tần số khác lân cận chiếm ưu thế và ảnh hưởng (masking), che khuất. Chính các tính chất này đã giúp cho mp3 có khả năng tốt trong việc loại bỏ dữ thừa mà đối với tai người khó có khả năng cảm nhận được.

Cũng dựa vào tính chất này, các phương pháp để ẩn thông tin vào âm thanh cũng sẽ tìm ra các thành phần tương ứng để thay vì loại bỏ, ta chèn vào đó những thành phần âm thanh đảm bảo rằng tai người khó phân biệt được. Vì vậy có 2 khả năng:

- Tìm những miền tần số không thể nghe được (hạ âm, siêu âm) : bị hạn chế vì khả năng phát và thu của thiết bị đang sử dụng
- Tận dụng những âm thanh mà tai người vốn nghe được nhưng được sử dụng kèm theo một âm thanh khác nhằm che dấu.

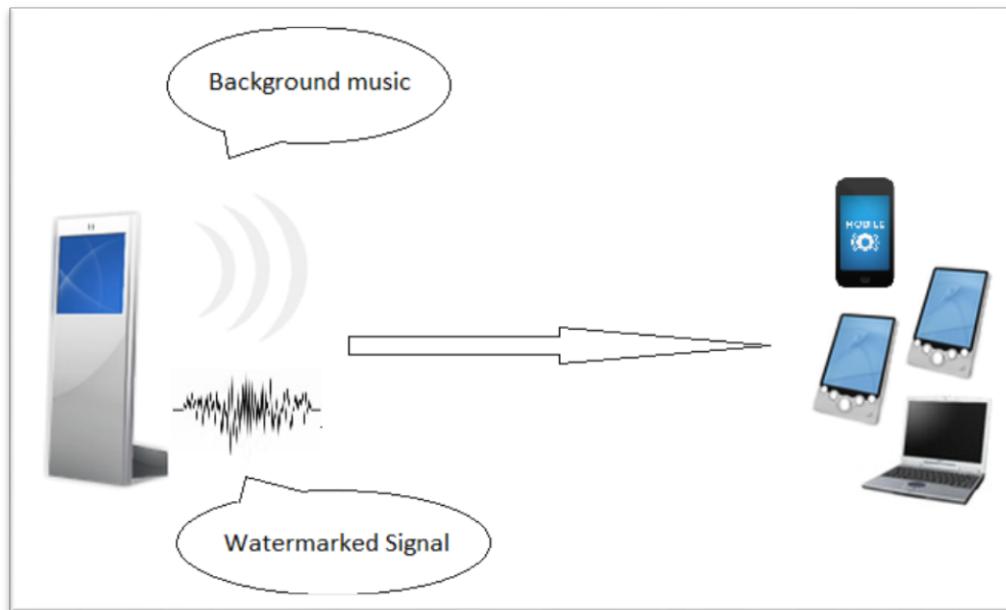
4.1.2. Hai quy trình được đề xuất trong luận văn

Hai quy trình mà nhóm chúng em đề xuất trong luận văn này gồm có:

Quy trình 1: Sử dụng tính chất masking của âm thanh. Được áp dụng với những môi trường có thể phát nhạc trong thời gian kéo dài như: bảo tàng, siêu thị, ... Thiết bị phát sẽ phát đồng thời hai loại tín hiệu

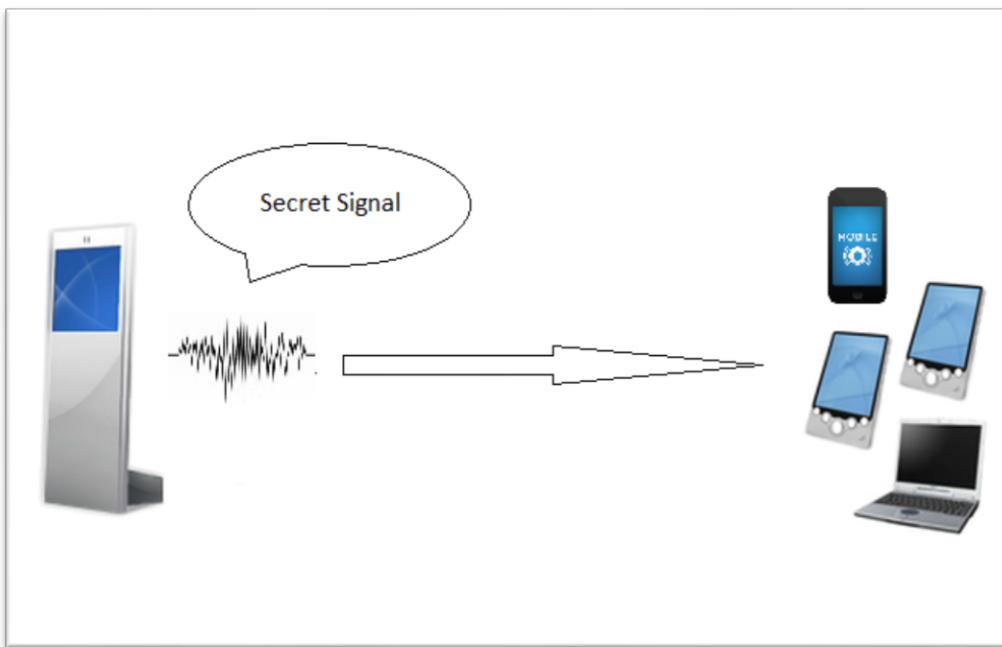
- Nhạc nền : được phát lặp đi lặp lại theo chu kỳ nhằm tác dụng như một sóng mang hay một masking che dấu tín hiệu đi kèm.
- Tín hiệu watermarking : tín hiệu được nhúng/phát đồng thời với nhạc nền. Do tính chất che dấu của các miền tần số lân cận làm ảnh hưởng đến khả năng nhận biết của người dùng, tín hiệu này hoàn toàn có thể bị che dấu trong những hoàn cảnh bình thường.

Khi người sử dụng lại gần nơi phát tín hiệu và sử dụng thiết bị để thu nhận được tín hiệu, qua quá trình lọc và phân tích dữ liệu, sẽ filter các miền tần số không phù hợp và giữ lại miền tần số đã quy ước. Từ những giá trị thu được này, ứng dụng sẽ liên kết với cơ sở dữ liệu và rút trích các thông tin, tài nguyên tương ứng với vị trí, địa điểm đặt máy phát và hiển thị cho người dùng.



Hình 4-1 Quy trình 1: Tín hiệu đi kèm nhạc nền

Quy trình 2: Sử dụng miền tần số cao (từ 18000 – 22000 Hz) và tính chất ngưỡng nghe của tai người trong điều kiện bình thường. Áp dụng ở những nơi không thể phát nhạc nền như thư viện, lớp học, ... Thiết bị phát sẽ chỉ phát duy nhất một loại tín hiệu với độ to dưới ngưỡng nghe của con người tại tần số đang phát nhưng vẫn có thể được bộ thu nhận biết. Khi người sử dụng di chuyển đến gần khu vực của máy phát, thiết bị sẽ tự động phát hiện ra tần số (thông tin) bí mật đang tồn tại trong không gian xung quanh. Hệ thống sẽ tự thu nhận, phân tích và nhận biết thông tin ấy. Dựa vào giá trị của tần số đó, một lần nữa, hệ thống lại liên kết với cơ sở dữ liệu và rút trích thông tin liên quan, tăng cường thêm tài nguyên đi kèm với vị trí người dùng đang đến. Bằng cách này, chúng ta có thể thông tin cho cùng lúc nhiều người dùng khi họ di đến trong phạm vi vị trí cách máy thu trước.



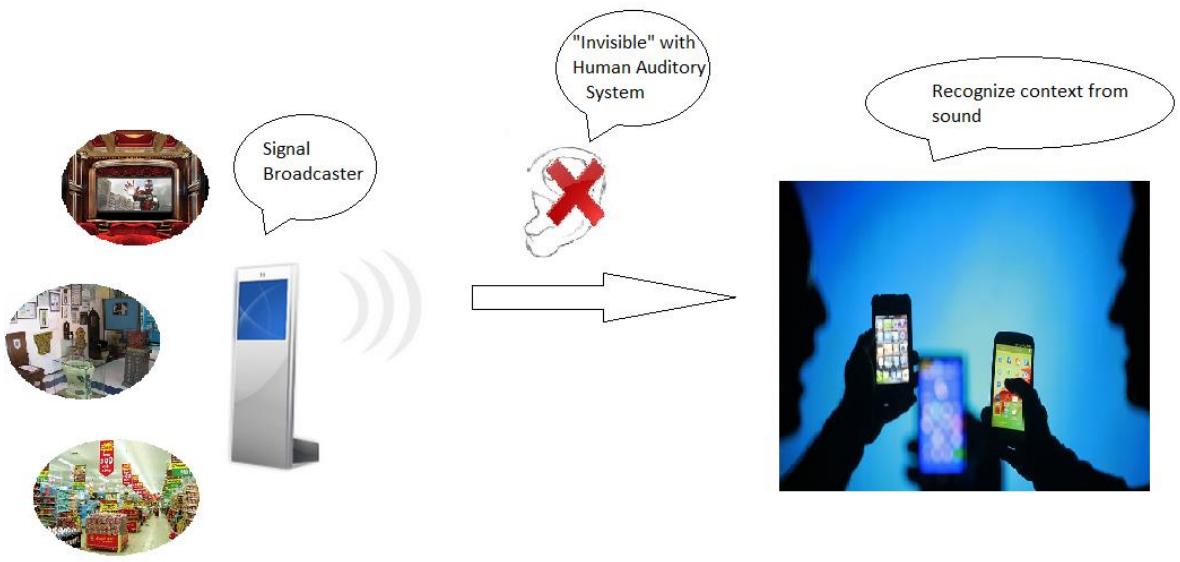
Hình 4-2 Quy trình 2: Tín hiệu vô hình với tai người

Phương pháp được đề xuất là dựa trên Audio Watermarking, dựa trên tích chất về tài nghe của người. Ngưỡng nghe ở môi trường im lặng tuyệt đối của tai người bình thường ở độ tuổi 20 đối với âm thanh đơn âm. Trong hệ thống này cũng không nhất thiết sử dụng Watermarking của audio và cũng không phụ thuộc vào âm thanh phát ra là giai điệu gì mà hệ thống vẫn luôn luôn đúng, chỉ cần dựa trên ý tưởng “nhúng” âm thanh vào một đoạn âm thanh khác và có độ to dưới ngưỡng nghe tuyệt đối của tai người và trên ngưỡng nghe tuyệt đối của thiết bị thu. Với sự tồn tại của các âm thanh khác, chắc chắn ngưỡng nghe tuyệt đối của tai người sẽ tăng lên và càng “khó” nhận biết sự tồn tại của tín hiệu ẩn.

4.1.3. Mối liên hệ giữa ngưỡng nghe tuyệt đối của con người và âm lượng nhỏ nhất mà máy tính có thể thu được – Giải pháp đề xuất trong luận văn.

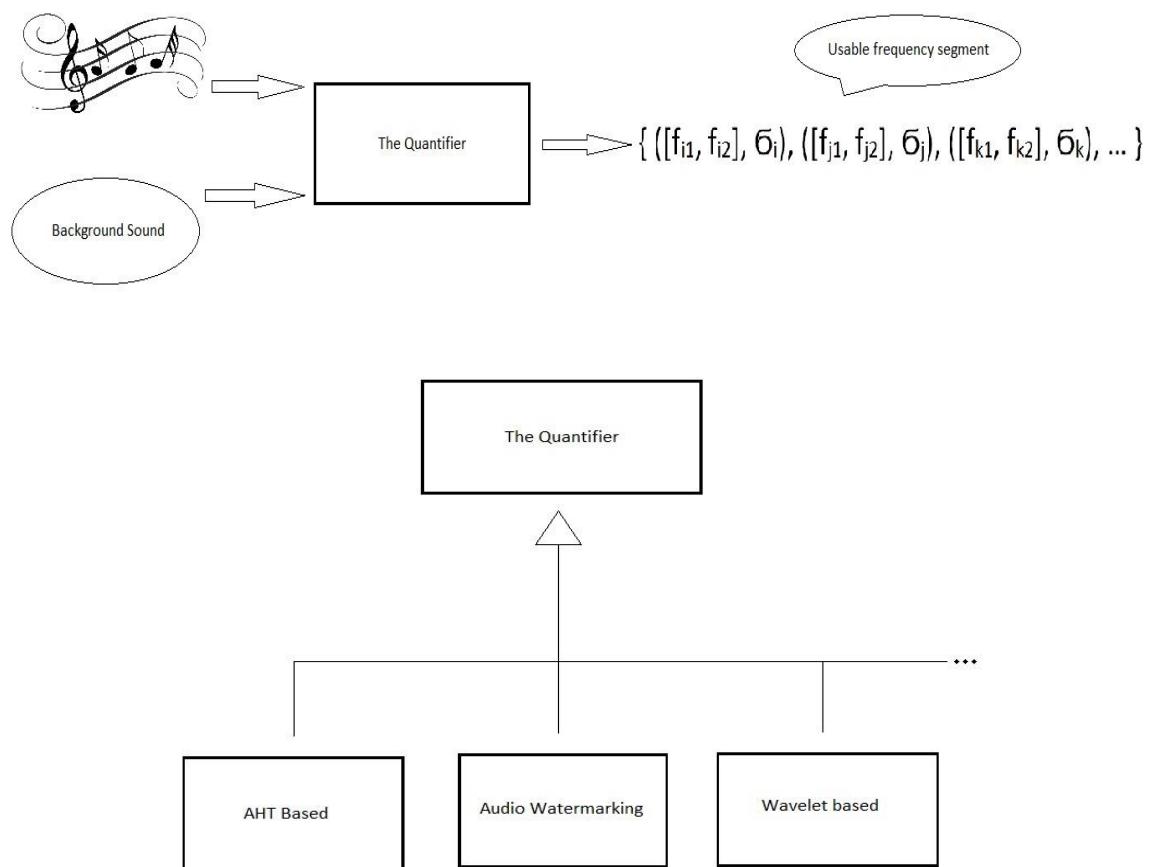
4.1.3.1. Cơ sở của giải pháp

Mô hình triển khai thực tế :



Hình 4-3 Cơ sở của giải pháp

Mô hình của giải pháp :



Hình 4-4 Mô hình của giải pháp

Ý tưởng cơ bản : Tính chất các âm thanh dưới ngưỡng nghe tuyệt đối (AHT – Absolute Hearing Threshold) sẽ bị ‘che khuất’ bởi chính ‘nhạc nền’ của Audio Watermarking hoặc khi không có ‘nhạc nền’ này thì tai người cũng không có khả năng nhận biết.

Giải thích mô hình của giải pháp : Âm thanh môi trường cùng với các âm thanh khác (có thể có hoặc không), sau khi cùng đi qua bộ định lượng môi trường, hệ thống sẽ tính toán được các dải hay các đoạn tần số f_i ứng với 1 giá trị σ_i . Các đoạn tần số này là những tần số giúp ta có khả năng chèn thêm các thông tin đang muốn được ẩn.

Bộ định lượng này có thể được hiện thực hóa và cài đặt qua nhiều cách khác nhau như : AHT-based, Watermarking, Wavelet-based ... Tuy nhiên trong luận văn này, nhóm chúng em tập trung sử dụng AHT-based làm phương pháp tiếp cận chính để minh họa một framework tích hợp một phương án đơn giản của Watermarking. Qua đó minh họa tính khả thi của các giải pháp mà trong tương lai còn có thể phát triển. Hệ thống cũng tương thích với tất cả các thuật toán Watermarking khác đã được đề xuất trước đây dù khả năng bị hạn chế do chỉ dùng được các khoảng tần số của AHT mà không nâng lên được. Nhóm chúng em đề xuất một hướng giải quyết đủ tổng quát để các hướng tiếp cận áp dụng Watermarking khác trong tương lai có thể dùng, nhưng chỉ tập trung 1 giải pháp (AHT-based) và phải đúng được trong cả môi trường không có nhạc nền.

Nếu chỉ xét các miền tần số tai người không nghe được, thì chúng ta đã bỏ mất một số tần số cơ hội khác mà thiết bị còn có thể ghi nhận. Vì vậy, mượn ý tưởng về AHT của tai người, nhóm chúng em đã đạt thêm một bước về việc xác định AHT của thiết bị. Do ta chủ động được về khoảng cách muốn phục vụ người sử dụng và độ chính xác trung bình mà thuật toán đang nhắm tới nên nhóm chúng em quyết định thực hiện các thử nghiệm trong phạm vi khoảng cách 3m. Tuy nhiên trên thực tế, nếu muốn làm trên 1 phạm vi khoảng cách mới thì có thể tiến hành lại các bước như chúng em đề xuất dưới đây để tính lại ngưỡng nghe của máy để đạt được 1 độ

chính xác trung bình đủ cao ở một khoảng cách cho trước. Vì vậy kết quả thực nghiệm là để minh họa tính khả thi còn phương pháp vẫn hoàn toàn độc lập.

Cách thực hiện :

- Chọn d là khoảng cách tối đa ta cần phục vụ người sử dụng và Δ là độ chính xác trung bình mà ta muốn đạt được ở khoảng cách d này
- Với mỗi tần số f, ta cần tìm độ lớn σ_f sao cho $0 \leq \sigma_f \leq AHT$ (của người) và khi phát tín hiệu, số lượng kết quả nhận được trung bình phải lớn hơn hoặc bằng Δ cho trước.

Để thực hiện được điều này thì chúng em sẽ sử dụng phương pháp tìm kiếm nhị phân như sau:

- Gán $\sigma_{max} = AHT(f)$, $\sigma_{min} = 0$
- If (số lượng kết quả ghi nhận trung bình tại $(\sigma_{max} + \sigma_{min}) / 2$) < Δ , $\sigma_{min} = (\sigma_{max} + \sigma_{min}) / 2$, tiếp tục
Else $\sigma_{max} = (\sigma_{max} + \sigma_{min}) / 2$, tiếp tục
- Thực hiện đến khi đạt một giá trị tương đối chính xác thì đó chính là AHT (của thiết bị) cần tìm.

4.1.3.2. Quá trình thực nghiệm

Cách đo được áp dụng trong này:

Mô hình này được đề xuất trong môi trường có 2 tính chất là hoàn toàn yên tĩnh như: trường học, thư viện, phòng làm việc; và môi trường có âm thanh nhạc nền như viện bảo tàng, phòng triển lãm, showroom, những khu du lịch và vì vậy ta sẽ có điều kiện tận dụng thêm một số đoạn tần số nhờ sự tồn tại của âm thanh ở những tần số lân cận chiếm đa số ta có thể khai thác được. Phương pháp đề xuất bên dưới dựa vào ý tưởng audio watermarking, dựa trên tính chất về tai nghe của người. Tai nghe của người có ngưỡng nghe tuyệt đối(AHT) được tính trong trường hợp điều kiện người nghe 20 tuổi không bị lãng tai và trong môi trường hoàn toàn lặng và âm thanh là đơn âm.

1. Tai người trẻ tuổi, khoẻ mạnh.
2. Môi trường yên lặng tuyệt đối.

- Đơn âm: âm thanh phát ra chỉ duy nhất một tần số, phô Fourier chỉ có một peak.

Trong hệ thống này không cần thiết dùng watermarking của audio và cũng không cần phụ thuộc vào âm thanh phát ra là giai điệu gì mà nó sẽ luôn luôn chạy đúng vì nó luôn luôn dùng ngưỡng nghe tuyệt đối, còn với sự tồn tại kia của nhạc nền chắc chắn đường AHT sẽ tăng lên. Vì vậy chúng ta có thể lấy bất kì tần số nào miễn là vẫn thuộc về bên trong ngưỡng AHT là được.

Tuy nhiên trong trường hợp này hệ thống không cần thiết phải cố gắng nhúng âm thanh hay loại trừ âm thanh trong nhạc nền mà vẫn muốn bảo toàn (preserve) lại âm thanh trong nhạc nền một cách bình thường mà khai thác tính chất của ngưỡng nghe từ tai người cho nên trong hệ thống này nhóm lựa chọn giải pháp là dưới ngưỡng nghe tuyệt đối này thì nhóm sẽ phát một âm đơn âm ra có mức cường độ âm (SPL – Sound Pressure Level) dưới ngưỡng nghe tuyệt đối của người là đúng.

Nhóm đã làm được một thí nghiệm là phát âm thanh đơn âm ở lần lượt từng tần số 20 - 20,000Hz bước nhảy 100, khoảng cách 3m trong môi trường phòng yên lặng, thông tin nhóm ghi nhận lại được chính là tần số phát ra đó, nếu đúng 100% (thu được) hay 0% (không thể thu được). Với khoảng cách 3m này thì âm lượng tối thiểu mà mình phát ra ở tần số cụ thể trong bảng (hay nói cách khác là ngưỡng nghe tuyệt đối của điện thoại).

Lấy ngưỡng nghe tuyệt đối của tai người phát ra xem điện thoại còn nghe được nữa không và độ chính xác ở tần số f (Af) với độ lớn D (dB) là bao nhiêu và độ chính xác đúng 100% dữ liệu truyền mới chấp nhận cường độ này là hợp lý. Bằng thực nghiệm ở khoảng cách 3m ($Af3m$) cho lần lượt từng tần số phát ra, lấy ngưỡng nghe tuyệt đối của tai người ở tần số đó phát và giảm dần cường độ xem điện thoại còn nghe được không, nếu không nghe được thì quyết định bỏ vì tai người bắt đầu nghe được mà điện thoại lại khó thu âm được. Nhưng mà nếu ở đây vẫn nghe được tốt, thì nhóm tiếp tục hạ mức cường độ âm xuống thấp hơn để làm sao tìm được vị trí

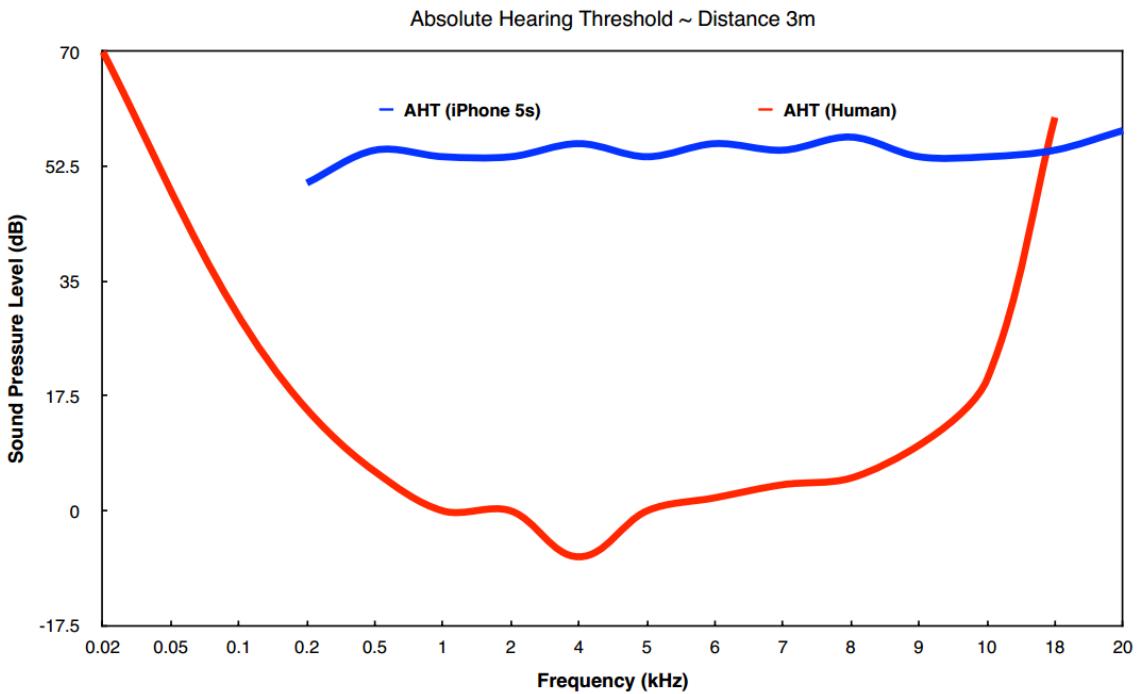
ngưỡng nghe tuyệt đối mà tai người không nghe được nhưng điện thoại thì lại vẫn còn có thể nghe được đúng 100% và cách xa 3m.

4.1.3.3. Kết quả đạt được

AHT ~ Distance (3m)

Freq (kHz)	AHT (iPhone 5s)	AHT (Human)	Man (Male, 22yo)	Device (iPhone 5s)
0.02		70	x	x
0.05		50	x	x
0.1		30	x	x
0.2	50	15	x	v
0.5	55	6	v	v
1	54	0	v	v
2	54	0	v	v
4	56	-7	v	v
5	54	0	v	v
6	56	2	v	v
7	55	4	v	v
8	57	5	v	v
9	54	10	v	v
10	54	20	v	v
18	55	60	x	v
20	58		x	v

Hình 4-5 Kết quả thực nghiệm so sánh AHT của người và AHT của thiết bị



Hình 4-6 Biểu đồ liên quan giữa AHT người và AHT thiết bị

4.2. Tiếng ồn ảnh hưởng đến kết quả đo, kết quả thu nhận và vấn đề khoảng cách truyền âm



Vấn đề

Âm thanh là dao động cơ học, được dao động dưới hình thức sóng trong môi trường đàn hồi và được thính giác của người tiếp thu. Trong không khí tốc độ âm thanh là 343 m/s, còn trong nước là 1450 m/s. Tiếng ồn là tập hợp những âm thanh có cường độ và tần số khác nhau, sắp xếp không có trật tự, gây cảm giác khó chịu cho người nghe, ảnh hưởng đến quá trình làm việc và nghỉ ngơi của con người. Hay là những âm thanh phát ra không đúng lúc, không đúng nơi, âm thanh phát ra với cường độ quá lớn, vượt quá mức chịu đựng của con người.

Độ to của 1 số tiếng ồn trong thực tế:

- Xe cộ: 75 - 135 dB (từ 120 dB trở lên là xe quân sự)
- Trong xây dựng: 75 - 110 dB

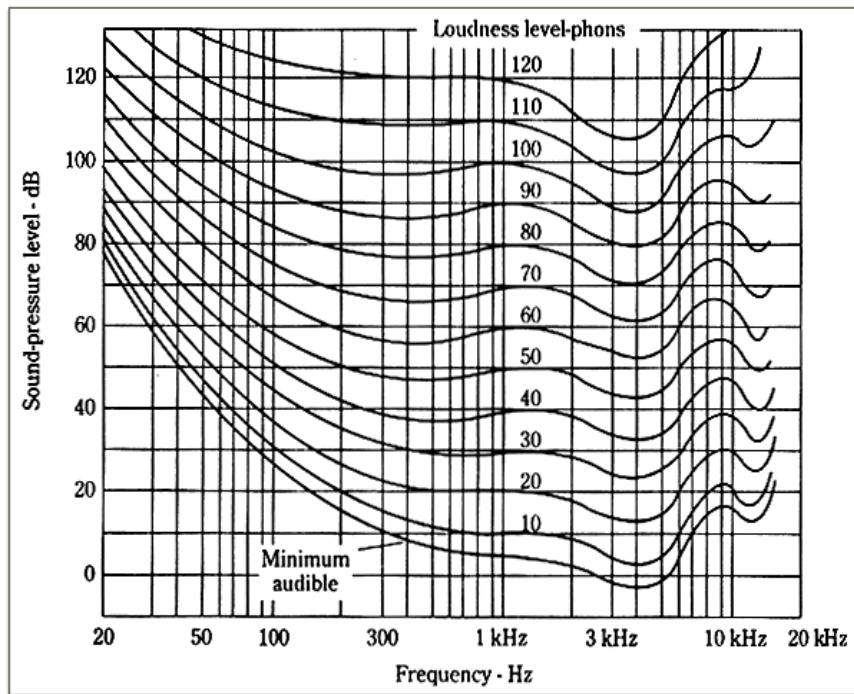
- Trong công nghiệp: 82 - 120 dB
- Trong sinh hoạt
 - Tiếng nói nhỏ: 30 dB
 - Tiếng nói chuyện bình thường: 60 dB
 - Tiếng nói to: 80 dB
 - Tiếng khóc của trẻ: 80 dB
 - Tiếng hát to: 110 dB

Thông tin thường bị tiếng ồn gây nhiễu, che lấp, làm cho việc tiếp nhận thông tin sẽ khó khăn hơn, độ chính xác của thông tin nhận được sẽ không cao ảnh hưởng đến mức độ hiệu quả của ứng dụng. Do vậy trong trao đổi thông tin cần phải giới hạn tiếng ồn hoặc truyền các âm ở những tần số tránh tiếng ồn hoặc khó có tiếng ồn ở dải tần số đó để tránh các ảnh hưởng do tiếng ồn gây ra.

Vậy phải làm sao để tránh hay giảm thiểu tiếng ồn và tạp âm trong quá trình truyền dữ liệu?

Giải pháp

Mối liên hệ giữa tần số và âm nhỏ nhất có thể nghe được (tính theo dB) của người bình (ở độ tuổi 20):



Hình 4-7 Biểu đồ thể hiện mối liên hệ giữa tần số, độ to của âm và mức độ nhỏ nhất để tai người có thể nghe được

Từ mối liên hệ giữa tần số và độ to của âm, ta có thể thấy rằng, nếu truyền âm trong 1 số tần số nhất định, việc ảnh hưởng của tạp âm hay tiếng ồn được giảm thiểu đáng kể.

Dựa vào kết quả trên cho thấy, ta nên chọn phát âm ở tần số từ 18000 trở lên để tai người khó có thể nghe được. Các nguồn phát âm có thể phát được ở dải tần số này cũng không nhiều trong điều kiện bình thường, vì thế tiếng ồn gây nhiễu được giảm thiểu. Thực nghiệm cho thấy, nếu truyền âm ở dải tần số 2000 - 6000 Hz trong điều kiện tương đối ồn (có tiếng nói chuyện kèm tiếng quạt máy) thì tần số âm thu được bị lệch trong khoảng -50 Hz đến 50 Hz. Nhưng khi truyền âm trong khoảng 18000 - 20000 Hz, tần số âm thu được chỉ lệch trong khoảng -15 Hz đến 15 Hz. Vì lý do đó cộng với khả năng người bình thường nghe được âm có tần số gần đạt đến siêu âm là tương đối thấp, nhóm đã chọn dải tần số để thực hiện truyền tin là 18000 - 20000 Hz (dải tần số 20000 - 22000 Hz không được chọn vì iPad không hỗ trợ tần số cao trên 20000 Hz, vì thế sẽ mất đi khả năng phát triển ứng dụng trên iPad)

Vấn đề kế tiếp liên quan đến việc “truyền - thu - phân tích” âm thanh đó là khoảng cách truyền âm hay khoảng cách từ nguồn phát đến thiết bị thu bao nhiêu là xa nhất, bao nhiêu là chính xác nhất?

Người ta đánh giá tiếng ồn bằng mức áp suất âm L được tính từ công thức:

$$L_p = 20 \times \log_{10}(P/P_0)$$

Trong đó:

P - Áp suất âm toàn phương trung bình (Pa)

P₀ - Áp suất âm nhỏ nhất tai người có thể nghe thấy (= 2.10⁻⁵ Pa).

Bel là đơn vị đo mức cường độ âm thanh (hay mức áp suất âm). 1 Bel là ngưỡng âm thanh tai người có thể nghe được. Decibel là đơn vị (bằng 1/10 bel) đo tiếng ồn thông dụng ngày nay.

Thính giác của con người có đặc tính cảm thụ cường độ âm thanh theo hàm Logarit. Vì thế cường độ âm thanh tăng 100 lần thì người ta chỉ thấy tiếng ồn to gấp đôi. Âm thanh truyền đi trong môi trường không khí dưới dạng sóng dao động cơ học. Trên đường lan truyền, âm thanh suy giảm theo qui luật tỉ lệ nghịch với bình phương khoảng cách lan truyền. Nghĩa là: khi khoảng cách tăng gấp đôi thì cường độ âm thanh giảm còn $\frac{1}{4}$ và mức cường độ âm giảm đi 6 dB. Nếu độ to của âm giảm xuống tới “minimum audible” (độ to nhỏ nhất tai người có thể nhận biết) hoặc thậm chí nhỏ hơn thì có thể cả những thiết bị cũng khó mà thu được âm. Vì thế có thể thấy độ to của âm cũng ảnh hưởng đến kết quả thu nhận tín hiệu. Theo thực nghiệm, với âm ở tần số cao, ứng dụng có thể truyền và nhận được tín hiệu khi nguồn và máy thu cách nhau khoảng 6m với 1 tấm của kính chắn giữa (vẫn còn 1 vài khe hở). Nếu không có vật cản thì ở khoảng cách 9m vẫn thu nhận được tín hiệu tần nguồn phát. Đây là 1 số liệu tương đối phù hợp với các thiết bị truyền data có độ dài vừa phải như ứng dụng mà nhóm đang phát triển.

Về thời gian truyền âm, giữa 2 tín hiệu âm là 0.15 - 0.2s. Trong khi đó tốc độ truyền âm trong không khí là 340m/s. Cho nên khoảng cách giữa 2 âm liên tiếp

được truyền đi vẫn đủ để tại máy thu thu lần lượt 2 tín hiệu riêng biệt, giúp ứng dụng không bị nhầm lẫn.

Nhìn chung, tần số và độ to của âm thanh đều là những vấn đề liên quan trực tiếp tới việc thu phát dữ liệu. Với việc chọn lựa âm ở tần số cao cùng với độ to thích hợp kết hợp với các phương pháp phân tích, encode, decode, nhóm đã thực hiện được việc truyền dữ liệu ở khoảng cách từ 3 - 5m với độ chính xác tương đối cao.

4.3. Chưa có một phương pháp hay protocol cụ thể nào quy định việc gửi và phát dữ liệu đi thông qua sóng âm thanh

Giải pháp

Trong cuộc sống hiện tại chúng ta luôn tồn tại âm thanh xung quanh từ tiếng nhạc du dương, đến tiếng xe máy, tiếng của quạt gió hay tiếng của tivi. Bản chất của âm thanh là truyền đi theo nguyên tắc broadcast tức là khi một nguồn âm phát ra tín hiệu sẽ lan tỏa đi trong môi trường và không phân biệt thiết bị nhận miễn là thiết bị đó còn nằm trong vùng phát của sóng âm là sẽ thu được, kể cả nếu máy phát có chức năng thu thì sẽ thu lại được chính âm thanh do mình phát ra. Điều này sẽ gây rắc rối cho người lập trình trong việc nhận dạng thiết bị thu đang tương tác với mình để xử lý hiệu quả và quản lý được số lượng thiết bị trong hệ thống mạng của mình.

Vì vậy để xuất ra một cách gửi đi tín hiệu hiệu quả của sóng âm thanh phải hoàn thành việc giải quyết 2 bài toán: quan trọng nhất là cách truyền dữ liệu là thông điệp cho máy thu có thể hiểu và thực hiện được, tránh môi trường nhiễu tiếng ồn xung quanh và phân biệt thiết bị nhận tín hiệu. Nhóm đã giải quyết 3 vấn đề trên như sau:

1. Để truyền đi thông điệp, nhóm đã quy định một nguyên tắc mã hóa sóng âm thanh với một tần số nhất định và ánh xạ sang một ID, mỗi ID này tượng trưng cho một hành động cụ thể nào đó được quy định trong ứng dụng:

2. Ván đề tiếng ồn được xử lí triệt để trong đề tài của nhóm. Nhóm khảo sát và nhận ra được kết quả như sau:

- a. Môi trường học tập trong lớp học thường tiếng ồn của máy điều hòa và quạt có tần số cao nhất dao động trong khoảng: 200Hz - 10kHz
- b. Tiếng nói của người: 80 - 1600Hz.
- c. Giới hạn thu, phát của iPhone: 20 - 20500Hz.
- d. Tần số mà tai người bắt đầu rất khó nghe thấy âm thanh: 18.5kHz.

Như vậy chỉ cần chọn miền tần số có giá trị từ 18.5kHz - 20.5kHz vừa có thể vừa giảm nhiễu của môi trường xung quanh, vừa tránh ảnh hưởng đến con người do ô nhiễm tiếng ồn gây ra khi phát liên tục những tiếng beep beep kéo dài.

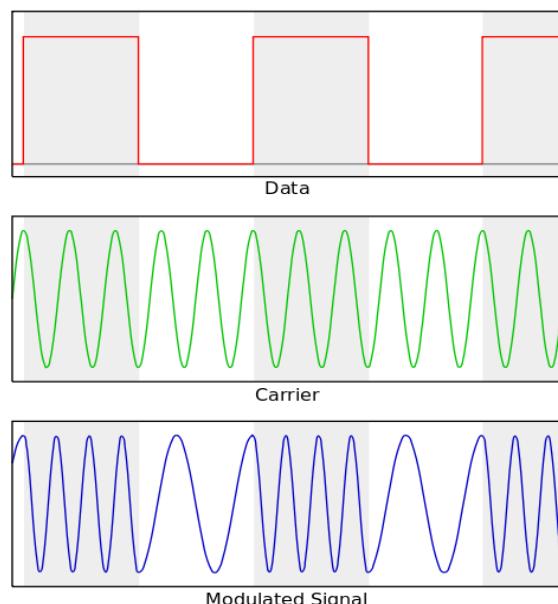
Nhóm đã mô phỏng được việc truyền dữ liệu, một cơ chế trao đổi thông tin qua lại giữa máy phát và máy thu thông qua âm thanh bằng cách chia nhỏ miền tần số đã chọn ở trên thành dãy các tần số riêng biệt, mỗi tần số sẽ đại diện cho một số ID theo quy ước định sẵn của thuật toán, sau đó sẽ có một hàm thu tần số đó và giải mã ra số ID ban đầu và thực hiện hành động trên máy thu.

$f_1 - (1) \rightarrow$ Âm thanh có thể phát được trên các thiết bị kĩ thuật số – (2) $\rightarrow f_2$

f_1 là một số nguyên, là một kiểu dữ liệu kĩ thuật số nên sóng sẽ có dạng nhị phân 0 và 1, âm thanh có thể thu được là một dạng sóng mang analog thì tần số có dạng sóng hình sin. Nên quá trình chuyển từ kiểu dữ liệu digital sang analog nhóm áp dụng kĩ thuật Audio Frequency Shift Keying (AFSK). Việc nhận tần số f_2 được truyền đi nhóm dùng thuật toán Fast Fourier Transform (FFT). Tần số f_1 và f_2 dù là siêu âm nhưng độ sai lệch có thể lên đến $\pm 15\text{Hz}$ nên sẽ có một cơ chế nhận biết để so sánh giữa f_1 và f_2 . Tần số $f_3 = f_2 \pm 15$. Sau đó so sánh nếu $f_3 = f_1$ thì chương trình sẽ thực hiện yêu cầu sau đó.

Nói về thuật toán Audio Frequency Shift Keying (AFSK) có thể giải thích đơn giản như sau:

Thuật toán AFSK là một kỹ thuật chuyển đổi dữ liệu kỹ thuật số (digital) bằng cách thay đổi tần số (pitch) của một nốt nhạc (thường là một tiếng beep), kết quả là sóng này có thể phát được thông qua radio, điện thoại hay bất cứ thiết bị kỹ thuật số có khả năng phát nhạc khác. Thông thường dữ liệu âm thanh sau khi được truyền đi thường tồn tại 2 loại âm sắc: đầu tiên là tone đánh dấu biển diễn bằng số 1 của nhị phân, còn lại là tone đánh dấu khoảng thời gian được biểu diễn bằng số 0 trong nhị phân.



Hình 4-8 Hình minh họa các loại sóng trong quá trình áp dụng kỹ thuật AFSK

Điển hình của kỹ thuật Frequency Shift Keying trong ngày nay mà chúng ta thường thấy nhất đó chính là sóng radio. Trong những ứng dụng radio, tín hiệu được mã hoá bằng kỹ thuật AFSK được sử dụng để điều biến thành sóng RF (Radio Frequency carrier), dùng kỹ thuật Amplitude Modulation (AM) hoặc Frequency Modulation (FM).

Kỹ thuật AFSK không thường được sử dụng trong việc truyền nhận dữ liệu với tốc độ cao, bởi vì nó không hiệu quả trong việc tiết kiệm năng lượng sử dụng và sử dụng băng thông so với những kỹ thuật chuyển đổi khác. Cách cài đặt đơn giản là

một ưu điểm của kỹ thuật AFSK, ưu điểm tiếp theo của nó chính là tín hiệu được mã hoá sẽ có thể dễ dàng phát trên các thiết bị kỹ thuật số có khả năng phát nhạc.

Về thuật toán Fast Fourier Transform (FFT) là thuật toán được sử dụng khác nhiều trong xử lý tín hiệu số. FFT là một ứng dụng của việc đưa dữ liệu dạng analog về thành digital để máy tính có thể xử lý. Về ứng dụng thuật toán FFT nhóm đã thực hiện việc ghi âm tín hiệu được chuyển đổi bằng kỹ thuật AFSK, sau đó chuyển về miền tần số xử lý, sau đó áp dụng chuyển đổi FFT và nhận về được pitch cao nhất trong miền tần số đó. Pitch đó chính là tần số siêu âm mà nhóm đã đưa vào AFSK.

4.4. Kết chương

Trong chương này, chúng em đã trình bày các vấn đề về công nghệ, kỹ thuật và giải pháp trong truyền nhận tín hiệu một cách “vô hình” – người dùng không hoặc khó có thể nghe được. Ở chương tiếp theo, chúng em sẽ trình bày tiếp tục về kiến trúc của ứng dụng.

Chương 5

Kiến trúc ứng dụng

Nội dung của Chương 5 giới thiệu về bộ phát sóng và bộ thu sóng: chức năng, mục đích và quy trình làm việc. Sơ đồ use-case và sơ đồ class của ứng dụng location-based.

5.1. Kiến trúc của bộ thu sóng và bộ phát sóng

5.1.1. Bộ phát sóng

5.1.1.1. Mục đích

Từ dữ liệu ban đầu cần ẩn chứa trong sóng mang, biến đổi để thu được mảng samples theo công thức chuyên. Mảng samples này chính là mảng tín hiệu digital, thông qua các máy phát tích hợp trên thiết bị (loa) sẽ được chuyển thành tín hiệu analog và phát đi.

Công thức chung : với 1 giá trị f (1 tần số cũng là đại diện cho giá trị của dữ liệu), ta có công thức chuyển đổi

```
for (int index = 0 ; index < nSamples ; index++)
    samples[index] = sin(index * 2.0 * Pi * f / sampleRate) * amplitude ;
```

trong đó :

sampleRate : thông thường giá trị được chọn là 44100

amplitude : biên độ hay độ to

5.1.1.2. Chức năng và quy trình làm việc

Quá trình render 1 sóng âm thanh của hệ thống trên iPhone có đặc điểm render theo frame (mỗi frame có thể có 256 hoặc 512 sample). Vì thế để có thể biến đổi 1 giá trị tần số f và render vào buffer để phát sóng, nhóm thực hiện đã sử thay đổi cách thức tính sample và lưu giá trị các giá trị tính toán cho lần render sau.

```

const double amplitude = 1;

// Get the tone parameters out of the view controller
ChatSoundViewViewController *viewController = (_bridge ChatSoundViewViewController *)inRefCon;
double theta = viewController->theta;
double theta_increment = 2.0 * M_PI * viewController->frequency / viewController->sampleRate;

// This is a mono tone generator so we only need the first buffer
const int channel = 0;
Float32 *buffer = (Float32 *)ioData->mBuffers[channel].mData;

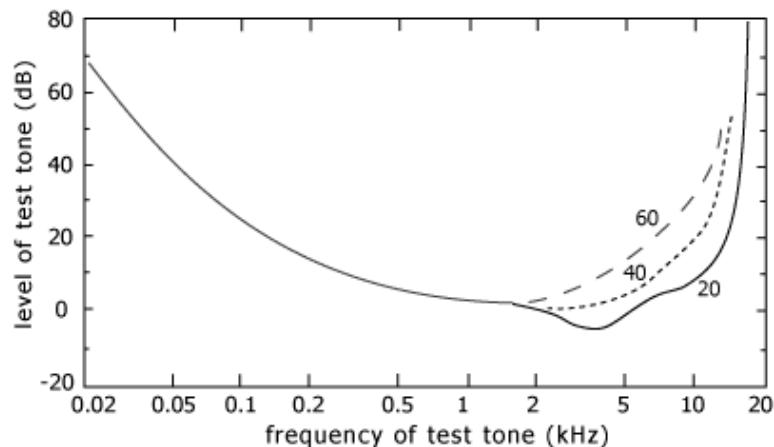
// Generate the samples
for (UInt32 frame = 0; frame < inNumberFrames; frame++)
{
    buffer[frame] = sin(theta) * amplitude;

    theta += theta_increment;
    if (theta > 2.0 * M_PI)
    {
        theta -= 2.0 * M_PI;
    }
}

// Store the theta back in the view controller
viewController->theta = theta;

```

Âm thanh khi được mã hóa sẽ được phát chung với 1 âm thanh khác (đại diện cho sóng mang – carrier) để sử dụng thủ thuật Auditory Masking nhằm che dấu âm thanh chứa dữ liệu.



Nguồng nghe tuyệt đối của con người ở âm thanh thuộc miền tần số cao từ 18000 – 20000Hz là tương đối cao, kết hợp với việc bị ảnh hưởng bởi chính sóng mang, khả

năng dữ liệu âm thanh bị ẩn bị một người bình thường phân biệt nghe được là rất thấp.

Dữ liệu được truyền đi vì vấn đề đồng bộ và bảo mật nên cần có 1 phương thức mã hóa hay 1 phương thức đóng gói (tương tự TCP/IP) để có thể giúp bên nhận dễ dàng nhận và xử lý tín hiệu. Trong trường hợp này, nhóm đã sử dụng 1 tín hiệu đầu, 1 tín hiệu cuối để đóng gói dữ liệu và dùng mô hình trạng thái (state design pattern) để xử lý tín hiệu được nhận.

5.1.2. Bộ thu sóng

5.1.2.1. Mục đích

Từ dữ liệu nhận được qua các thiết bị thu (như microphone) trên các thiết bị di động, tiến hành thu nhận, phân tích, giải mã (nếu cần). Tín hiệu đầu vào là analog, qua microphone được chuyển thành tín hiệu số (digital), sau đó tiếp tục qua bộ giải mã (decoder), ta thu được dữ liệu được ẩn trong sóng mang.

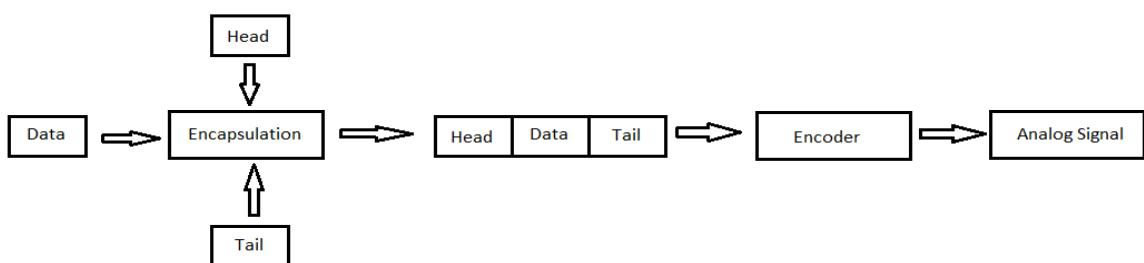
Quá trình nhận và rút trích thông tin được thực hiện thông qua sự hỗ trợ của các lớp thuộc framework CoreAudio của iOS như : AUGraph, AUNode, AVAudioSession,

...

5.1.2.2. Chức năng và quy trình làm việc

Như đã nói ở mục bộ phát sóng, để giúp bên nhận có thể dễ dàng nhận và xử lý tín hiệu chính xác, nhóm đã thực hiện đóng gói dữ liệu trong 1 cặp tín hiệu đầu cuối và sử dụng mô hình trạng thái ở bộ thu sóng nhằm phân tích dữ liệu.

Quá trình đóng gói có thể giải thích như sau :



Hình 5-1 Mô hình đóng gói dữ liệu (sử dụng trong ứng dụng ChatSound)

Với mẫu thiết kế trạng thái, bộ lọc dữ liệu khi ở một trạng thái nhất định, chỉ có những dữ liệu phù hợp với trạng thái đó mới có thể được ghi nhận (chúng cũng chính là khóa của các trạng thái này). Khi khóa được tìm thấy, hệ thống chuyển sang một trạng thái khác và lại chờ đợi một khóa khác.

```

if ([state isEqualToString:@"Waiting"]) {
    if (newFrequency == freqGetItem)
        [self stateChange:@"GetItem"];
}
if ([state isEqualToString:@"GetItem"]) {
    if (newFrequency != freqGetItem) {
        [self stateChange:@"Waiting"];
        if ((isBoy && index > 25) || (!isBoy && index <= 25))
            [self loadImageByReceivedFrequency:index];
    }
}

```

Sau khi thu nhận được tín hiệu (mảng sample), quá trình phân tích tín hiệu được thông qua các bước sau :

- Setup FFT

```

/* Setup our FFT */
- (void)realFFTSetup {
    UInt32 maxFrames = 2048;
    dataBuffer = (void*)malloc(maxFrames * sizeof(SInt16));
    outputBuffer = (float*)malloc(maxFrames * sizeof(float));
    log2n = log2f(maxFrames);
    n = 1 << log2n;
    assert(n == maxFrames);
    nOver2 = maxFrames/2;
    bufferCapacity = maxFrames;
    index = 0;
    A.realgp = (float *)malloc(nOver2 * sizeof(float));
    A.imagp = (float *)malloc(nOver2 * sizeof(float));
    fftSetup = vDSP_create_fftsetup(log2n, FFT_RADIX2);
}

```

```

COMPLEX_SPLIT A = THIS->A;
void *dataBuffer = THIS->dataBuffer;
float *outputBuffer = THIS->outputBuffer;
FFTSetup fftSetup = THIS->fftSetup;

uint32_t log2n = THIS->log2n;
uint32_t n = THIS->n;
uint32_t nOver2 = THIS->nOver2;
uint32_t stride = 1;
long bufferCapacity = THIS->bufferCapacity;
SInt16 index = THIS->index;

AudioUnit rioUnit = THIS->ioUnit;

```

- Gọi AudioUnit của iOS render sample vào buffer (buffer tạm ở hàm callback)

```

renderErr = AudioUnitRender(rioUnit, ioActionFlags,
                           inTimeStamp, bus1, inNumberFrames, THIS->bufferList);

```

- Cộng dồn buffer tạm vào mảng buffer của decoder

```

memcpy((SInt16 *)dataBuffer + index, THIS->bufferList->mBuffers[0].mData, inNumberFrames*sizeof(SInt16));
THIS->index += inNumberFrames;

```

- Nếu số frame đọc được trong hàm callback nhỏ hơn số frame đọc được trong 1 lần render (đọc hết được mảng dữ liệu đã thu), tiến hành fft cho buffer
 - o Chuyển sample từ integer sang float

```

// We want to deal with only floating point values here.
ConvertInt16ToFloat(THIS, dataBuffer, outputBuffer, bufferCapacity);

```

- Chuyển đổi mảng dữ liệu từ xen kẽ thành ‘split’ : mảng dữ liệu ban đầu bao gồm những sample ‘chẵn’ và ‘lẻ’ nằm xem kẽ mang giá trị thực và ảo của tần số được chuyển hóa thành analog (ảo – thực – ảo – thực – ảo ...). Thư viện zDSP cung cấp hàm chuyển đổi mảng dữ liệu từ xen kẽ thành 2 miền ‘thực’ và ‘ảo’ nằm riêng biệt tách biệt nhau (2 thành phần realp và imagp trong A).

```
vDSP_ctoz((COMPLEX*)outputBuffer, 2, &A, 1, nOver2);
```

- Thực hiện fft trên mảng A đã chuyển đổi : Sau khi đã chuyển đổi thành dạng ‘split’, chúng ta gọi hàm vDSP_fft_zrip để tiến hành chạy fft trên mảng dữ liệu

```
// Carry out a Forward FFT transform.  
vDSP_fft_zrip(fftSetup, &A, stride, log2n, FFT_FORWARD);
```

- Chuyển đổi mảng dữ liệu từ ‘split’ thành xen kẽ

```
vDSP_ztoc(&A, 1, (COMPLEX *)outputBuffer, 2, nOver2);
```

- Với mỗi cặp sample kế cận là real-value và imaginary-value (đại diện cho 1 tần số), tính độ lớn (magnitude). Tần số ứng với độ lớn cao nhất chính là ‘peak’ (đỉnh – cực trị cục bộ của tín hiệu) của mảng dữ liệu và cũng là tần số đi kèm với sóng mang.

```

// Determine the dominant frequency by taking the magnitude squared and
// saving the bin which it resides in.
float dominantFrequency = 0;
int bin = -1;
for (int i=0; i<n; i+=2) {
    float curFreq = MagnitudeSquared(outputBuffer[i], outputBuffer[i+1]);
    if (curFreq > dominantFrequency) {
        dominantFrequency = curFreq;
        bin = (i+1)/2;
    }
}

```

bin là vị trí của tần số cần tìm, công thức để phục hồi lại tần số là

$$\text{bin} \times \text{sampleRate} / \text{bufferCapacity}$$

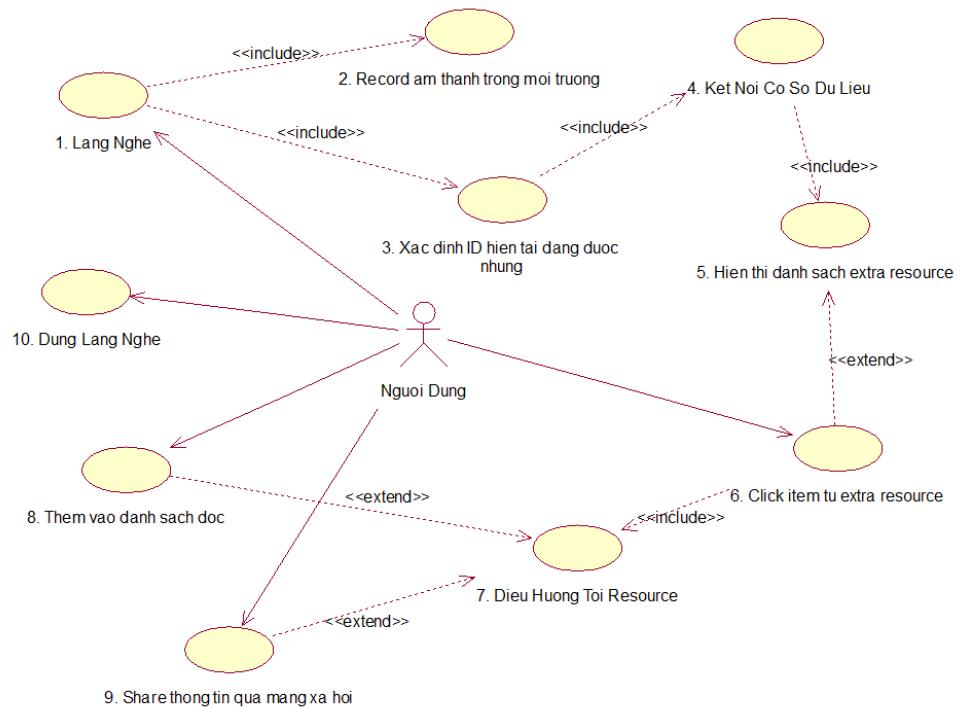
5.2. Kiến trúc của ứng dụng Location-based

5.2.1. Phát biểu bài toán

Hệ thống được xây dựng trong luận văn này nhằm minh họa ứng dụng có nhiệm vụ lắng nghe và ghi nhận âm thanh. Đồng thời có khả năng trình chiếu được tất cả những tài nguyên tăng cường kèm theo vị trí location-based của người dùng. Hệ thống có khả năng kết nối với cơ sở dữ liệu để lấy được các thông tin tăng cường tại mỗi một vị trí tương ứng với tín hiệu (ID) được nhận biết tại vị trí đó, đồng thời có khả năng trình diễn ra màn hình để người dùng có khả năng tương tác cũng như chuyển tiếp tới các tài nguyên kèm theo này.

5.2.2. Sơ đồ use case

5.2.2.1. Use case



Hình 5-2 Sơ đồ Use case của ứng dụng ShopLocator (Location-based)

5.2.2.2. Mô tả Actor và Use case

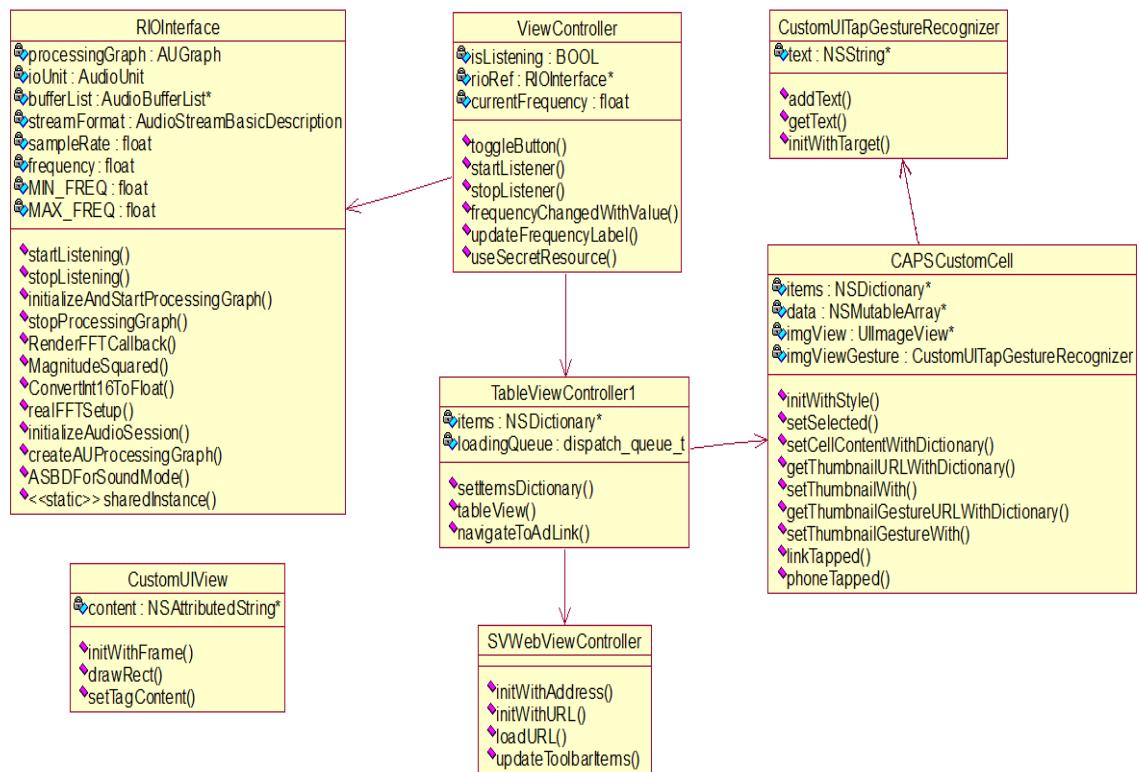
STT	Actor	Ý nghĩa
1	Người dùng	Người dùng các thiết bị iOS có cài đặt ứng dụng ShopLocator

STT	Use case	Ý nghĩa
1	Lắng nghe	Người dùng bật chức năng lắng nghe, tìm kiếm các extra resource trong không gian xung quanh
2	Record âm thanh trong môi trường	Record âm thanh trong môi trường xung quanh

3	Xác định ID hiện tại đang được nhúng	Ứng dụng xác định được tín hiệu đặc biệt (ID) đang được nhúng
4	Kết nối cơ sở dữ liệu	Phân tích tín hiệu và kết nối với cơ sở dữ liệu để lấy về các resource
5	Hiển thị danh sách extra resource	Hiển thị danh sách resource trên 1 tableview, sử dụng dispatch_queue xử lý bất đồng bộ
6	Click item từ extra resource	Người dùng click (tap) vào 1 item trên tableview
7	Điều hướng tới resource	Ứng dụng điều hướng người dùng tới browser với url tương ứng với resource
8	Thêm vào danh sách đọc	Người dùng có khả năng thêm các nội dung vào danh sách đọc của browser khi lướt web
9	Share thông tin qua mạng xã hội	Người dùng có khả năng chia sẻ các thông tin qua facebook, youtube, ...
10	Dừng lắng nghe	Người dùng có thể dừng lắng nghe bất cứ khi nào

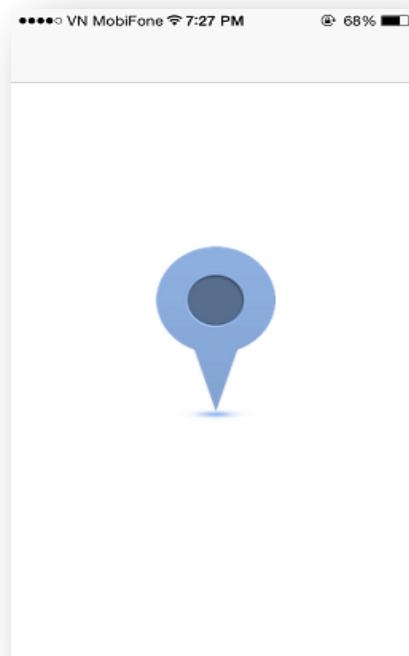
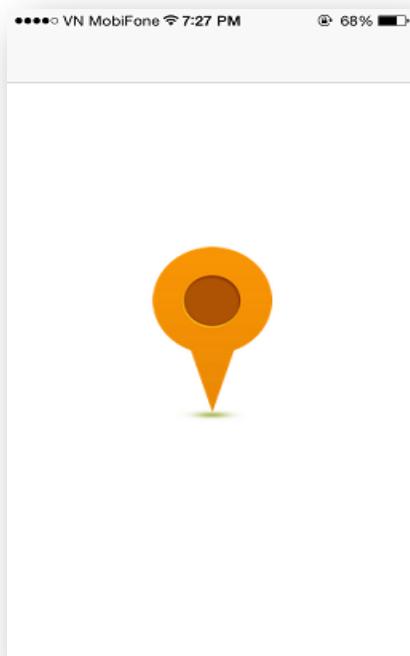
Hình 5-3 Bảng mô tả use case và actor của hệ thống

5.2.3. Sơ đồ class (class diagram)



Hình 5-4 Sơ đồ lớp (class diagram) của ứng dụng ShopLocator

5.2.4. Giao diện của ứng dụng



Mc Donald's

McDonald's là một tập đoàn kinh doanh hệ thống nhà hàng thức ăn nhanh với khoảng 31.000 nhà hàng tại 119 quốc gia[3] phục vụ 43 triệu lượt khách mỗi ngày dưới thương hiệu riêng của mình. Công ty được thành lập đầu tiên năm 1940 do anh em Richard và Maurice ("Mick & Mack") McDonald. Nền tảng của sự kinh doanh thành công hôm nay là do Ray Kroc mua lại của anh em McDonald và phát triển thành một trong những dự án kinh doanh ẩm thực thành công nhất thế giới.

Năm 2008, doanh thu của tập đoàn là khoảng 22.8 tỷ USD, trong đó lợi nhuận ròng vào khoảng 3.5 tỷ USD.

Mặc dù McDonald's được xem là một biểu tượng cổ điển của chiến lược toàn cầu hóa, tuy nhiên nhiều tập đoàn lớn khác được phát triển rộng rãi hơn nhiều, ví dụ như Coca-Cola có doanh thu vĩ đại hơn nhiều

0906851115

The screenshot displays details about the McDonald's corporation, its history, financial performance, and contact information. It also includes the company's phone number and its logo.

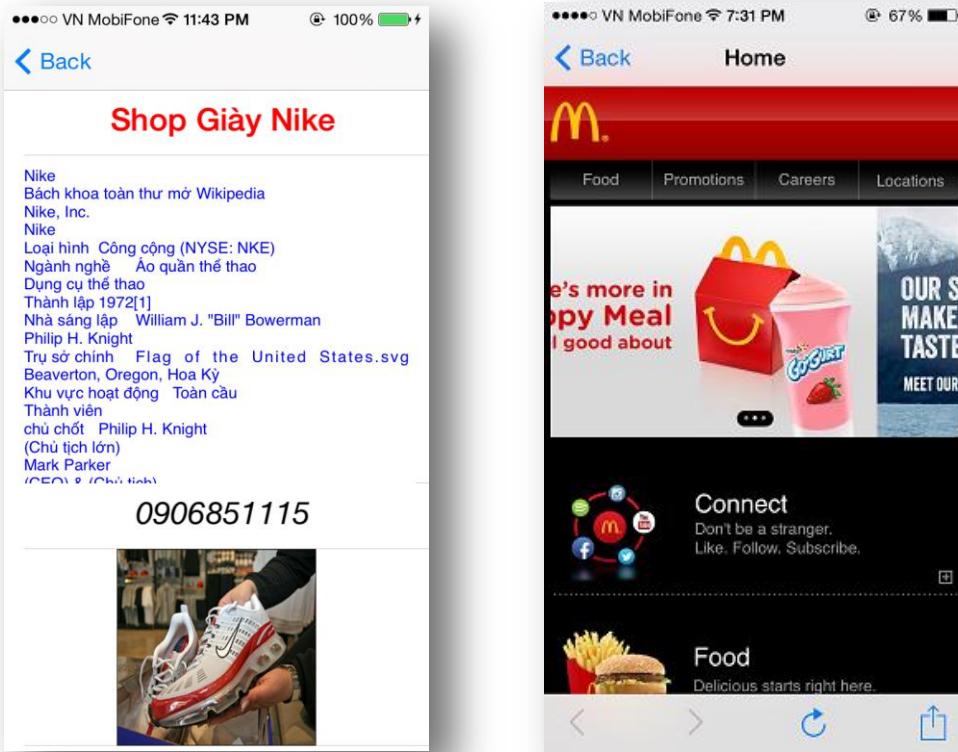
Apple Retail Store

Apple Inc. (NASDAQ: AAPL, LSE: ACP) là tập đoàn công nghệ máy tính của Mỹ có trụ sở chính đặt tại Silicon Valley ở San Francisco, tiểu bang California. Apple được thành lập ngày 1 tháng 4 năm 1976 dưới tên Apple Computer, Inc., và đổi tên vào đầu năm 2007. Với lượng sản phẩm bán ra toàn cầu hàng năm là 13,9 tỷ đồ la Mỹ (2005), và có 14.800 nhân viên ở nhiều quốc gia, sản phẩm là máy tính cá nhân, phần mềm, phần cứng, thiết bị nghe nhạc và nhiều thiết bị đa phương tiện khác. Sản phẩm nổi tiếng nhất là Apple Macintosh, iPod nghe nhạc, chương trình nghe nhạc iTunes, đặc biệt là điện thoại iPhone và máy tính bảng iPad. Nơi bán hàng và dịch vụ chủ yếu là Mỹ, Canada, Nhật Bản và Anh.

Theo số liệu khảo sát đầu năm 2013 của IDC, Apple đã trở thành hãng điện thoại lớn thứ nhất thế giới, điển hình Nokia và Samsung. Sau khi trái mìn nổ

0906851115

This screenshot provides information about Apple Inc., its history, market position, and contact details. It also includes the company's phone number and its logo.



Hình 5-5 Giao diện của các màn hình trong ứng dụng ShopLocator

5.3. Kết chương

Trong Chương 5, chúng em đã trình bày về kiến trúc của ứng dụng. Kiến trúc này là cơ sở để nhóm chúng em xây dựng ứng dụng location-based hoàn chỉnh cũng như có thể mở rộng trong tương lai. Trong chương cuối, chúng em sẽ tóm tắt các kết quả đã đạt được cũng như hướng phát triển đầy tiềm năng của hệ thống.

Chương 6

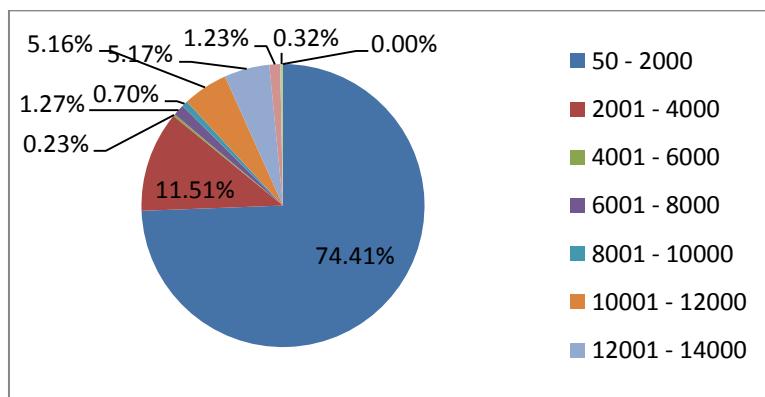
Kết luận

↗ Nội dung của chương trình bày về các kết quả đạt được và hướng phát triển của đề tài.

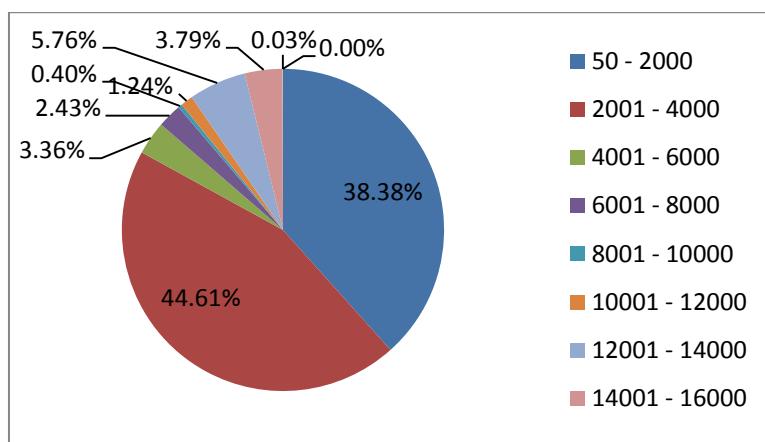
6.1. Các kết quả đạt được

- Ứng dụng Chat Sound: nhóm đã xây dựng được ứng dụng gửi nhận tin nhắn thông qua sóng âm thanh với độ chính xác cao. Qua thực nghiệm cho thấy xác suất xuất hiện tần số có thể làm nhiễu ở 3 môi trường lớp học, siêu thị và nhà sách như sau:

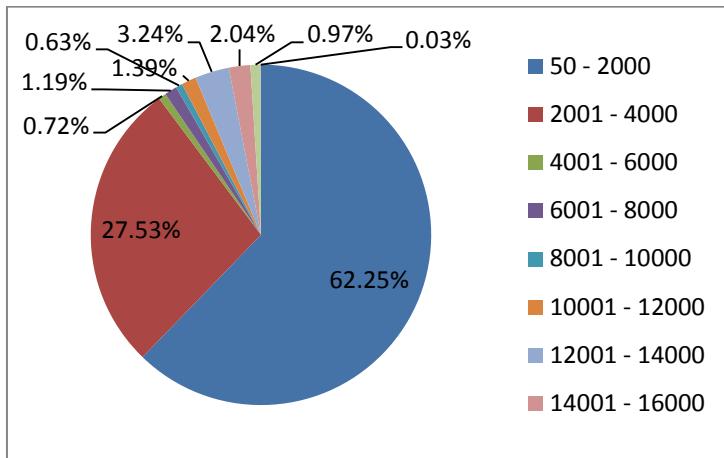
- o Lớp học:



- o Siêu thị:



- Nhà sách:



Có thể thấy ở tần xuất cao nhất có xuất hiện tần số từ 18000 – 20000Hz ở mức dưới 0.1%, thể hiện khả năng phát triển tiếp của ứng dụng (vì càng khó xuất hiện tần số thì càng khó bị nhiễu tín hiệu)

- Ứng dụng Chat Sound:

- Performance (test trên iPhone 4s 32 bit 512MB RAM CPU 1GB):
 - CPU: 2 - 4%
 - Memory: 13 - 25MB
 - Frame Per Second (FPS): 60 FPS
- Ứng dụng Location-based: xây dựng được ứng dụng “check point”, xác định vị trí thông qua sóng âm thanh. Từ đó có thể đưa thêm các thông tin kèm theo với những vật dụng khác (thích hợp với các ứng dụng đưa thêm thông tin cho người dùng tại những nơi như thư viện, bảo tàng, siêu thị, v.v...)
- Bộ thu – phát sóng: được xây dựng độc lập, riêng biệt với phần ứng dụng → dễ dàng tích hợp vào các ứng dụng khác nhau một cách hiệu quả, tạo ra nhiều phần mềm, ứng dụng khác nhau áp dụng công nghệ xử lý tín hiệu số.

6.2. Hướng phát triển

- Hướng phát triển phân hệ phần mềm:
 - Trong phạm vi của luận văn này chưa khai thác yếu tố về Watermarking mà dùng bộ định lượng AHT trên thiết bị thu trong môi trường yên tĩnh tuyệt đối, trong tương lai nhóm sẽ phát triển hệ thống tiếp theo hướng tích hợp Audio

Watermarking và hệ thống này có thể tương thích với mọi thuật toán Watermarking khác đã được đề xuất để nâng cao cơ hội truyền âm thanh trong môi trường có hay không có nhạc nền và đồng thời giúp nâng ngưỡng nghe tuyệt đối của tai người lén, do đó nhóm có thể truyền được dung lượng lớn đồng thời nâng cao độ chính xác thông tin trong một quá trình truyền và ghi nhận.

- Hướng phát triển phân hệ phần cứng:

- Đặc điểm của các thiết bị di động thông minh hiện nay mục đích chính vẫn là nghe gọi, nhắn tin, giải trí phục vụ cho cuộc sống hằng ngày của con người nên nhà sản xuất sẽ giới hạn tần số siêu âm và tạp âm xung quanh để tránh nguy cơ có thể gây khó chịu và ảnh hưởng đến con người có thính giác nhạy cảm. Vì vậy, một giải pháp về phần cứng mới được nhóm đề ra đó chính là máy phát và máy thu có thể là các module kỹ thuật Arduino có thể nâng cao tần số phát lên tới hàng chục nghìn Hertz và tần số thu sẽ không bị giới hạn ở ngưỡng 50 – 21000 Hz như hiện nay trên các thiết bị di động. Máy phát sẽ là một mạch Arduino hoàn chỉnh có thể phát được âm thanh với tần số siêu âm và cường độ liên tục, đảm bảo tính nhỏ gọn, dễ lắp đặt và bảo trì cho các đối tác; về phần thiết bị thu có thể là một module microphone nhỏ của Arduino, có thể thu được các tần số siêu âm không giới hạn và được tích hợp vào cổng audio 3.5mm của điện thoại để truyền tín hiệu vào ứng dụng được cài trên điện thoại để xử lí. Nhóm chọn Arduino vì giá thành tương đối rẻ so với các thiết bị thu phát kỹ thuật số đắt tiền khác như các máy MP3 hay các máy phát nhạc; các mạch Arduino là các mạch điện tử thuần nên sẽ đảm bảo tính hoạt động liên tục, chịu tải cao và khó hỏng hóc, một lý do khác để chọn Arduino đó là các mạch đã được các thành viên trên thế giới công nhận về hiệu suất và tính đúng đắn do đó đảm bảo an toàn trong khi phát triển và sử dụng. Arduino sẽ mở ra một tiềm năng mới cho hệ thống location-based mà nhóm xây dựng.

TÀI LIỆU THAM KHẢO

- [1] Hồ Ngọc Lâm, Thân Võ Chí Nhân, "Nghiên Cứu Watermarking Trên Audio và Ứng Dụng," *Luận văn Cử nhân Tin học, Khoa Công nghệ Thông tin, trường ĐH Khoa học Tự nhiên, ĐHQG-HCM*, 2003.
- [2] Lưu Khoa, Lê Thị Hoàng Ngân, "Ân Dữ Liệu Trên Âm Thanh Số," *Luận văn Cử nhân Tin học, Khoa Công nghệ Thông tin, trường ĐH Khoa học Tự nhiên, ĐHQG-HCM*, 2005.
- [3] Nguyễn Ngọc Châu Sang, Trương Đại Dương, "Landmark Recognition," *Luận văn Cử nhân Tin học, Khoa Công nghệ Thông tin, trường ĐH Khoa học Tự nhiên, ĐHQG-HCM*, 2013.
- [4] Trần Minh Triết, "Applying the Robust Psychoacoustic Audio Watermarking Technique in Internet Digital Traditional Music Museum in Vietnam," 2004.
- [5] "Benchmarked FFTs," 26 June 2014. [Online]. Available: <http://www.fftw.org/benchfft/ffts.html>.