

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN
BỘ MÔN CÔNG NGHỆ TRI THỨC**

BÙI TRUNG HẢI – PHẠM NGỌC TUẤN

**XÂY DỰNG ỨNG DỤNG TRỢ LÝ ẢO CHO MÁY
TÍNH SỬ DỤNG GOOGLE SPEECH API**

KHÓA LUẬN TỐT NGHIỆP CỬ NHÂN CNTT

TP. HCM, 2017

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA CÔNG NGHỆ THÔNG TIN
BỘ MÔN CÔNG NGHỆ TRI THỨC**

**BÙI TRUNG HẢI – 1312165
PHẠM NGỌC TUẤN – 1312669**

**XÂY DỰNG ỨNG DỤNG TRỢ LÝ ẢO CHO MÁY
TÍNH SỬ DỤNG GOOGLE SPEECH API**

KHÓA LUẬN TỐT NGHIỆP CỬ NHÂN CNTT

**GIÁO VIÊN HƯỚNG DẪN
ThS. NGÔ MINH NHỰT**

KHÓA 2013 - 2017

This image shows a full page of white paper designed for handwriting practice. It features approximately 20 horizontal rows, each consisting of two closely spaced dotted lines. These rows are evenly spaced across the entire page, providing a guide for letter height and placement. There are no margins, text, or other markings on the paper.

TpHCM, ngày tháng năm
Giáo viên hướng dẫn
[Ký tên và ghi rõ họ tên]

This image shows a full page of white paper with horizontal dotted lines. The lines are evenly spaced and run across the width of the page, providing a guide for handwriting practice. There are no margins, text, or other markings on the page.

TpHCM, ngày tháng năm

[Ký tên và ghi rõ họ tên]

LỜI CẢM ƠN

Trước tiên, em xin gửi lời tri ân sâu sắc đến Thầy Lê Hoài Bắc. Thầy đã rất tận tâm, nhiệt tình hướng dẫn và chỉ bảo em trong suốt quá trình thực hiện luận văn. Không có sự quan tâm, theo dõi chặt chẽ của Thầy chắc chắn em không thể hoàn thành luận văn này.

Em xin chân thành cảm ơn quý Thầy Cô khoa Công Nghệ Thông Tin - trường đại học Khoa Học Tự Nhiên, những người đã ân cần giảng dạy, xây dựng cho em một nền tảng kiến thức vững chắc.

Con xin cảm ơn ba mẹ đã sinh thành, nuôi dưỡng, và dạy dỗ để con có được thành quả như ngày hôm nay. Ba mẹ luôn là nguồn động viên, nguồn sức mạnh hết sức lớn lao mỗi khi con gặp khó khăn trong cuộc sống.

TP. Hồ Chí Minh, 3/2014

Trần Trung Kiên

ĐỀ CƯƠNG CHI TIẾT

Tên Đề Tài: Xây dựng ứng dụng trợ lý ảo cho máy tính sử dụng Google Speech API
Giáo viên hướng dẫn: ThS. Ngô Minh Nhựt
Thời gian thực hiện: từ ngày 15/12/2016 đến ngày 15/07/2017
Sinh viên thực hiện: Bùi Trung Hải – 1312165, Phạm Ngọc Tuấn - 1312669
Loại đề tài: Phát triển hệ thống, nghiên cứu thuật toán

Nội Dung Đề Tài: Tìm hiểu các phương pháp xử lý và tương tác với tín hiệu âm thanh, tiếng nói. Nghiên cứu hệ thống Natural Language Understanding đơn giản. Ứng dụng vào xây dựng ứng dụng trợ lý ảo cho máy tính.

Nội dung chi tiết của đề tài bao gồm:

- Xây dựng ứng dụng trợ lý ảo tương tác bằng giọng nói tiếng Anh cho máy tính
- Các vấn đề quan tâm:
 - Xử lý tín hiệu số (âm thanh, tiếng nói), hệ thống chuyển đổi tiếng nói thành văn bản.
 - Tương tác audio I/O, hệ thống chuyển đổi văn bản thành tiếng nói.
 - Giao thức truyền nhận dữ liệu REST API,...
 - Hệ thống Natural Language Understanding hiệu quả để xác định ý muốn của người dùng
 - Hệ thống chạy đa nhiệm
 - Tối ưu hóa hệ thống

- Các thành phần cơ bản của hệ thống:
 - Module thu âm từ microphone
 - Module nhận dạng từ khóa wake up
 - Module chuyển đổi tiếng nói thành văn bản
 - Module chuyển đổi văn bản thành hành động.
 - Module chuyển đổi văn bản thành tiếng nói
- Ứng dụng thử nghiệm sẽ hỗ trợ các tính năng:
 - Thông báo giờ hiện tại
 - Dự báo thời tiết trong ngày
 - Phát nhạc
 - Trả lời các câu hỏi Wh-question
 - Trả lời các thông tin cơ bản của ứng dụng: tên, tuổi,...

Kế Hoạch Thực Hiện:

- 15/12/2016 – 14/01/2017: Khảo sát, tìm hiểu về các thư viện python phục vụ cho việc tương tác và xử lý tín hiệu âm thanh.
- 15/01/2017 – 14/02/2017: Thiết kế các thành phần của hệ thống.
- 15/02/2017 – 14/03/2017: Cài đặt và thử nghiệm các thành phần: thu âm, nhận dạng từ khóa wake up, chuyển giọng nói thành văn bản, chuyển văn bản thành giọng nói.
- 15/03/2017 – 14/04/2017: Tìm hiểu và xây dựng hệ thống Natural Language Understanding.
- 15/04/2017 – 14/05/2017: Cài đặt các chức năng mà ứng dụng hỗ trợ.
- 15/05/2017 – 31/05/2017: Ráp nối tất cả các thành phần, tiến hành thử nghiệm và hoàn thiện hệ thống.
- 01/06/2017 – 28/06/2017: Viết và hoàn thiện luận văn.

Xác nhận của GVHD**Ngày 08 tháng 07 năm 2013****SV Thực hiện**

MỤC LỤC

LỜI CẢM ƠN	i
ĐỀ CƯƠNG CHI TIẾT	ii
MỤC LỤC	v
DANH MỤC HÌNH ẢNH	vi
DANH MỤC BẢNG	vii
Chương 1 Mở Đầu	1
Chương 2 Tổng Quan	2
Chương 3 Nghiên Cứu Thực Nghiệm Hoặc Lý Thuyết	3
Chương 4 Trình Bày, Đánh Giá, Bàn Luận Về Các Kết Quả	4
Chương 5 Kết Luận và Hướng Phát Triển	5
5.1 Kết luận	5
5.2 Hướng phát triển	5
TÀI LIỆU THAM KHẢO	6
Phụ Lục: Các Công Trình Đã Công Bố	8

DANH MỤC HÌNH ẢNH

DANH MỤC BẢNG

Chương 1

Mở Đầu

Chương 2

Tổng Quan

Chương 3

Nghiên Cứu Thực Nghiệm Hoặc Lý Thuyết

Chương 4

Trình Bày, Đánh Giá, Bàn Luận Về Các Kết Quả

Chương 5

Kết Luận và Hướng Phát Triển

5.1 Kết luận

5.2 Hướng phát triển

TÀI LIỆU THAM KHẢO

- [1] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: A review and new perspectives,” 2013.
- [2] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley, and Y. Bengio, “Theano: a CPU and GPU math expression compiler,” in *Proceedings of the Python for Scientific Computing Conference (SciPy)*, Jun. 2010, oral Presentation.
- [3] A. Coates, “Demystifying unsupervised feature learning,” Ph.D. dissertation, Stanford University, 2012.
- [4] A. Coates, A. Y. Ng, and H. Lee, “An analysis of single-layer networks in unsupervised feature learning,” in *International Conference on Artificial Intelligence and Statistics*, 2011, pp. 215–223.
- [5] X. Glorot, A. Bordes, and Y. Bengio, “Deep sparse rectifier networks,” in *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics. JMLR W&CP Volume*, vol. 15, 2011, pp. 315–323.
- [6] I. Goodfellow, H. Lee, Q. V. Le, A. Saxe, and A. Y. Ng, “Measuring invariances in deep networks,” in *Advances in neural information processing systems*, 2009, pp. 646–654.
- [7] Q. V. Le, A. Karpenko, J. Ngiam, and A. Y. Ng, “Ica with reconstruction cost for efficient overcomplete feature learning,” in *NIPS*, 2011, pp. 1017–1025.
- [8] Y. LeCun, “The MNIST database,” <http://yann.lecun.com/exdb/mnist/>.

- [9] H. Lee, A. Battle, R. Raina, and A. Ng, “Efficient sparse coding algorithms,” in *Advances in neural information processing systems*, 2006, pp. 801–808.
- [10] V. Nair and G. E. Hinton, “Rectified linear units improve restricted boltzmann machines,” in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010, pp. 807–814.
- [11] B. A. Olshausen *et al.*, “Emergence of simple-cell receptive field properties by learning a sparse code for natural images,” *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.
- [12] S. Rifai, G. Mesnil, P. Vincent, X. Muller, Y. Bengio, Y. Dauphin, and X. Glorot, “Higher order contractive auto-encoder,” *Machine Learning and Knowledge Discovery in Databases*, pp. 645–660, 2011.
- [13] S. Rifai, P. Vincent, X. Muller, X. Glorot, and Y. Bengio, “Contractive auto-encoders: Explicit invariance during feature extraction,” in *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, 2011, pp. 833–840.
- [14] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, “Extracting and composing robust features with denoising autoencoders,” in *Proceedings of the 25th international conference on Machine learning*. ACM, 2008, pp. 1096–1103.
- [15] M. D. Zeiler, “Hierarchical convolutional deep learning in computer vision,” Ph.D. dissertation, New York University, 2014.
- [16] M. Zeiler, M. Ranzato, R. Monga, M. Mao, K. Yang, Q. Le, P. Nguyen, A. Senior, V. Vanhoucke, J. Dean *et al.*, “On rectified linear units for speech processing.” ICASSP, 2013.

Phụ Lục: Các Công Trình Đã Công Bố

Hội nghị quốc tế:

- **K. Tran** and B. Le, “Demystifying Sparse Rectified Auto-Encoders,” in *Proceedings of the Fourth Symposium on Information and Communication Technology*, ser. SoICT’13. New York, NY, USA: ACM, 2013, pp. 101–107. [Online]. Available: <http://doi.acm.org/10.1145/2542050.2542065>