

**BỘ THÔNG TIN VÀ TRUYỀN THÔNG**  
**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

-----oOo-----

**BÁO CÁO**  
**NHIỆM VỤ KHOA HỌC VÀ CÔNG NGHỆ**

**TÊN NHIỆM VỤ:**

**Nghiên cứu, chế tạo găng tay thông minh hỗ trợ người  
câm trong giao tiếp bằng cách chuyển đổi ngôn ngữ cử  
chỉ bàn tay sang giọng nói.**

<b>Chủ nhiệm:</b>	<b>Ths. Nguyễn Văn Chương</b>
<b>Thành viên:</b>	<b>Ths. Nguyễn Kim Quang</b>
	<b>Ks. Nguyễn Hải Quang</b>
	<b>Ks. Từ Quang Hưng</b>

**Hà Nội – 2019**

## MỤC LỤC

<b>DANH SÁCH NGƯỜI THỰC HIỆN NHIỆM VỤ .....</b>	<b>3</b>
<b>DANH MỤC HÌNH VẼ.....</b>	<b>4</b>
<b>BẢNG THUẬT NGỮ/ TỪ VIẾT TẮT .....</b>	<b>6</b>
<b>LỜI MỞ ĐẦU .....</b>	<b>7</b>
<b>CHƯƠNG I: TỔNG QUAN .....</b>	<b>9</b>
I.1.    Người khiếm thính, người câm.....	9
I.2.    Một số đặc điểm tâm lý của người khiếm thính / người câm .....	10
I.3.    Các nghiên cứu trước đây về công nghệ hỗ trợ người câm điếc .....	14
I.4.    Mục tiêu và phạm vi nghiên cứu .....	19
<b>CHƯƠNG II: NGÔN NGỮ KÝ HIỆU, CÔNG NGHỆ NHẬN DẠNG NGÔN NGỮ KÝ HIỆU.....</b>	<b>20</b>
II.1.    Quá trình hình thành và phát triển ngôn ngữ kí hiệu .....	20
II.2.    Ngôn ngữ kí hiệu Mỹ.....	21
II.2.2.1    Câu hỏi .....	24
II.2.2.2    Câu trần thuật.....	26
II.2.2.3    Câu điều kiện .....	27
II.3.    Ngôn ngữ kí hiệu Việt Nam.....	28
II.4.    Công nghệ nhận dạng NNKH.....	32
II.4.2    Nhận dạng bằng cảm biến .....	33
<b>CHƯƠNG III: THIẾT KẾ CHẾ TẠO GĂNG TAY THÔNG MINH .....</b>	<b>37</b>
III.1.    Tập dữ liệu nhận dạng .....	37
III.2.    Thuật toán phân tích lớp và nhận dạng.....	39
III.3.    Thiết kế hệ thống, chế tạo Găng tay .....	40
III.3.1.1    Yêu cầu người dùng .....	40
III.3.1.2    Lựa chọn thiết kế.....	41
III.3.7.1    Lưu đồ thực hiện nhận biết cử chỉ .....	48
III.3.7.2    Lưu đồ thực hiện biên dịch tín hiệu cử chỉ sang dạng âm thanh .....	52
III.4.    Hướng dẫn sử dụng găng tay thông minh .....	53
III.4.1    Ai có thể sử dụng thiết bị.....	53

III.4.2	Học cách giao tiếp bằng Găng tay thông minh.....	53
III.4.3	Huấn luyện (định nghĩa) cử chỉ cho Găng tay .....	54
<b>CHƯƠNG IV: THỬ NGHIỆM VÀ ĐÁNH GIÁ.....</b>		<b>56</b>
IV.1.	Kết quả nghiên cứu chế tạo thiết bị .....	56
IV.2.	Thử nghiệm Găng tay .....	57
<b>KẾT LUẬN .....</b>		<b>61</b>
<b>TÀI LIỆU THAM KHẢO.....</b>		<b>61</b>

## DANH SÁCH NGƯỜI THỰC HIỆN NHIỆM VỤ

STT	Họ và Tên	Học hàm, học vị, chức vụ	Đơn vị công tác
1.	Nguyễn Văn Chương	Thạc sỹ, chủ trì nhiệm vụ	Viện Công nghệ Thông tin và Truyền thông – CDIT, Học viện Công nghệ Bưu chính Viễn thông
2.	Nguyễn Kim Quang	Tiến sỹ, thành viên nghiên cứu	Viện Công nghệ Thông tin và Truyền thông – CDIT, Học viện Công nghệ Bưu chính Viễn thông
3.	Từ Quang Hưng	Kỹ sư, thành viên nghiên cứu	Viện Công nghệ Thông tin và Truyền thông – CDIT, Học viện Công nghệ Bưu chính Viễn thông
4.	Nguyễn Hải Quang	Kỹ sư, thành viên nghiên cứu	Viện Công nghệ Thông tin và Truyền thông – CDIT, Học viện Công nghệ Bưu chính Viễn thông
5.			

## DANH MỤC HÌNH VẼ

Hình 1: Stephen Hawking với thiết bị DECtalk.....	15
Hình 2: Thiết bị đạt giải “Microsoft Imagine Cup” năm 2012 .....	16
Hình 3: Găng tay chuyển đổi ngôn ngữ ký hiệu của Cornell University .....	17
Hình 4: Hệ thống nhận dạng ngôn ngữ ký hiệu của Texas A&M University .....	18
Hình 5: Giao diện ứng dụng Proloquo2Go.....	19
Hình 6: Thực hiện dấu hiệu câu hỏi dạng WH.....	24
Hình 7: Dấu hiệu cử chỉ câu hỏi Yes/No.....	25
Hình 8: Dấu hiệu câu hỏi tu từ .....	26
Hình 9: Làm dấu hiệu câu điều kiện.....	28
Hình 10: Các bước xử lý trong công nghệ nhận dạng NNKH bằng phân tích hình ảnh .....	32
Hình 11 Các giai đoạn chính của việc thu thập và nhận dạng dữ liệu của hệ thống nhận dạng dựa trên Găng tay.....	34
Hình 12: Công thức tính điện trở của một vật.....	34
Hình 13: Mô hình thiết kế 1(trên); Mô hình thiết kế 2 (giữa); Mô hình thiết kế 3 (dưới) .....	42
Hình 14: Mô hình thiết kế Găng tay thông minh .....	43
Hình 15: Thiết kế khối hệ thống.....	43
Hình 16: Cảm biến độ cong.....	44
Hình 17: Vi điều khiển Atmega2560.....	45
Hình 18: Module giao tiếp thẻ nhớ SD.....	46
Hình 19: Khối vi điều khiển .....	47
Hình 20: Giao tiếp cảm biến, sd card và bluetooth .....	47
Hình 21: Lưu đồ thực hiện phần mềm hệ thống.....	49
Hình 22: Lưu đồ thực hiện chuyển đổi cử chỉ sang tiếng nói .....	52
Hình 23: Dấu hiệu cử chỉ.....	54
Hình 24: Hình ảnh Găng tay thử nghiệm .....	56
Hình 25: Cử chỉ thử nghiệm .....	57
Hình 26: Cử chỉ thử nghiệm .....	57
Hình 27: Cử chỉ thử nghiệm .....	58
Hình 28: Cử chỉ thử nghiệm .....	58
Hình 29: Cử chỉ thử nghiệm .....	59

Hình 30: Cử chỉ thử nghiệm .....	59
Hình 31: Cử chỉ thử nghiệm .....	60
Hình 32: Cử chỉ thử nghiệm .....	60

## **BẢNG THUẬT NGỮ/ TỪ VIẾT TẮT**

<b>STT</b>	<b>Thuật ngữ/từ viết tắt</b>	<b>Tiếng Việt</b>	<b>Tiếng Anh</b>
1.	NNKH	Ngôn ngữ kí hiệu (dấu hiệu)	
2.	NNKHTV	Ngôn ngữ kí hiệu tiếng Việt	
3.	ASL	Ngôn ngữ kí hiệu Mỹ	American Sign Language
4.	SGDs	Thiết bị hỗ trợ giao tiếp đầu ra bằng giọng nói	Speech-generating devices
5.			
6.			
7.			

## LỜI MỞ ĐẦU

Ngôn ngữ là công cụ giao tiếp hữu dụng nhất của con người. Có ngôn ngữ con người có thể nghe, nói, đọc, viết để khám phá thế giới muôn màu. Người câm là người có thể do gặp vấn đề ở các bộ phận như thực quản, dây thanh quản, phổi, miệng, hay lưỡi, có thể do bẩm sinh hoặc do các rối loạn chức năng bên trong cơ thể ảnh hưởng đến việc phát âm hay không có khả năng nói. Vì thế họ cực kỳ khó khăn trong việc giao tiếp- một nhu cầu bức thiết trong cuộc sống hàng ngày, chưa nói đến việc học tập nâng cao kiến thức trình độ bản thân.

Để giao tiếp người câm phải sử dụng ngôn ngữ riêng, được gọi là ngôn ngữ kí hiệu (dấu hiệu) NNKH giao tiếp với giữa người câm với người câm người câm với người bình thường. Thực tế cho thấy rằng không chỉ người Câm cần NNKH để giao tiếp mà người Điếc, người Khiếm thính do khả năng nghe bị suy giảm nên khả năng phát âm của họ cũng bị hạn chế cũng phải sử dụng NNKH để giao tiếp.

Ở Việt Nam hiện có khoảng 2,5 triệu người Câm Điếc và Khiếm thính. Cộng đồng người Câm Điếc là những người hoàn toàn không “nghe, nói” và chỉ sử dụng ngôn ngữ kí hiệu như phương tiện để giao tiếp chính. Trong khi đó người Khiếm thính, là những người vẫn còn khả năng “nghe, nói” và giao tiếp bằng ngôn ngữ lời nói, đồng thời có thể sử dụng hoặc không sử dụng kí hiệu ở mức phụ trợ.

Theo tổ chức Y Tế Thế Giới (WHO) trên thế giới có khoảng 7% người câm/điếc, tương đương với 550 triệu người, số người biết ngôn ngữ giao tiếp cử chỉ (thủ ngữ) chỉ chiếm 30% trong số này[1]. Còn đối với người bình thường, rất ít người có khả năng đọc và hiểu được ngôn ngữ kí hiệu. Phần lớn cha mẹ người không biết ngôn ngữ kí hiệu nên không thể dạy người câm điếc từ nhỏ. Do ngôn ngữ kí hiệu là một ngôn ngữ rất khó học. Đặc trưng của loại NNKH là dùng những động tác kí hiệu của bàn tay để truyền đạt ý của mình đến người khác. Tất cả những chữ cái, từ ngữ đều được quy ước với một dụng ý truyền đạt riêng. Tuy nhiên, rất nhiều động tác khi thực hiện lại có nhiều nét tương đồng, thậm chí rất giống nhau. Vì thế chỉ cần làm sai đi dù chỉ một chút là thông điệp truyền tới người nghe đã bị lệch hoàn toàn theo một hướng khác làm sai lệch thông tin. Hơn nữa cũng như ngôn ngữ nói, ngôn ngữ kí hiệu của từng quốc gia, thậm chí là từng khu vực trong một quốc gia rất khác nhau. Điều đó là do mỗi quốc gia, khu vực có lịch sử, văn hóa, tập quán khác nhau nên kí hiệu để biểu thị sự vật hiện tượng cũng khác nhau. Chẳng hạn, cùng chỉ tính từ màu hồng thì ở Hà Nội người ta xoa vào má (má hồng), còn tại Thành phố Hồ Chí Minh lại chỉ vào môi (môi hồng). Tuy nhiên, kí hiệu tất cả mọi nơi đều có những điểm tương đồng nhất định. Ví dụ: kí hiệu ‘uống nước’ thì nước nào cũng làm như nhau là giả bộ cầm cốc uống nước, kí hiệu ‘lái ô tô’ thì giả bộ cầm vô lăng ô tô quay quay, v.v...

Để giúp đỡ những người câm điếc và những người bị bệnh không thể giao tiếp bằng lời nói bình thường giao tiếp dễ dàng hơn, đã có khá nhiều công trình nghiên cứu ứng dụng các công nghệ điện tử và vi điện tử chế tạo ra các thiết bị hỗ trợ giao tiếp



như máy trợ thính, ốc tai điện tử và đặc biệt là các thiết bị tạo giọng nói giọng nói (Speech-generating devices (SGDs)) hỗ trợ rất tốt cho người câm điếc giao tiếp. Các thiết bị SGDs là các thiết bị phát hiện các cử chỉ, hành hoạt động của một bộ phận cơ thể của người sử dụng rồi phát ra tiếng nói để cho người giúp người đối thoại hiểu người sử dụng muốn nói gì.

Qua nghiên cứu của nhóm nghiên cứu các thiết bị tạo giọng nói (SGDs) đã được đưa ra thị trường thế giới từ khá sớm với công nghệ điện tử, đặc biệt là từ những năm 1980. Cho đến nay vẫn có những nghiên cứu được đưa ra thị trường như các ứng dụng chuyển các biểu tượng trên điện thoại thông minh thành giọng nói, ... Tuy nhiên các thiết bị hiện có trên thị trường sử dụng ngôn ngữ cử chỉ, âm thanh là tiếng nước ngoài - tiếng Anh là chủ yếu, giá thành cao – từ vài trăm đến trên mười nghìn đô la Mỹ, sử dụng phức tạp,... không có khả năng “địa phương hoá” được ngôn ngữ cử chỉ và tiếng nói.

Chính vì vậy, nhiệm vụ: “Nghiên cứu chế tạo găng tay thông minh hỗ trợ người câm trong giao tiếp bằng cách chuyển đổi ngôn ngữ cử chỉ sang giọng nói.” sẽ góp phần giải quyết một phần vấn đề trên. Trong khuôn khổ nghiên cứu của nhiệm vụ này, nhóm nghiên cứu sẽ nghiên cứu chế tạo một găng tay hỗ trợ được người dùng trong giao tiếp với người bình thường bằng cách phát ra những câu nói tiếng Việt (người dùng và người hỗ trợ tự định nghĩa) từ ngôn ngữ kí hiệu bàn tay của người dùng. Điều này cho phép không chỉ những người Câm Điếc có thể sử dụng găng tay để giao tiếp với người bình thường mà còn cho cả những người không may mất khả năng nói chẳng hạn như người bị đột quỵ, người bị tai nạn giao thông, tai nạn lao động mà bị mất khả năng nói nhưng vẫn cử động được các ngón tay có thể giao tiếp được chính xác với người nhà hay bác sỹ những câu nói cần thiết, quan trọng. Đặc biệt tính năng tự định nghĩa các ký hiệu với các âm thanh tương ứng cho phép “Địa phương hóa” thậm chí “Cá nhân hóa” ngôn ngữ ký hiệu phù hợp với từng địa phương và từng cá nhân người sử dụng.

Phần nội dung của báo cáo được chia làm 4 chương:

Chương 1 trình bày về nghiên cứu một số đặc điểm phổ biến của người Câm ở nhóm người Câm điển hình;

Chương 2 trình bày nghiên cứu ngôn ngữ kí hiệu bàn tay của người câm Việt Nam (Tiếng Việt) và thế giới (ngôn ngữ Tiếng Anh)

Chương 3 Dựa trên các giải pháp kỹ thuật đã nghiên cứu, lựa chọn giải pháp chế tạo găng tay thông minh hỗ trợ người câm trong giao tiếp với người bình thường bằng cách chuyển đổi ngôn ngữ cử chỉ bàn tay sang giọng nói phù hợp với người Việt Nam (cho phép người dùng và người hỗ trợ tự định nghĩa cử chỉ thành giọng nói).

Chương 4 trình bày về thử nghiệm và đánh giá thiết bị

## CHƯƠNG I: TỔNG QUAN

*Chương này giới thiệu những vấn đề chung về người điếc, người khiếm thính, người câm, các nghiên cứu liên quan đến công cụ hỗ trợ cho người câm điếc giao tiếp và thách thức của vấn đề, và mục tiêu giải quyết bài toán trong phạm vi nghiên cứu của vấn đề*

### I.1. Người khiếm thính, người câm

Khiếm thính hay khiếm khuyết/khuyết tật thính giác là những thuật ngữ để miêu tả việc mất thính giác hoàn toàn, không nghe được chút nào hoặc sút giảm về thính giác, nghe không rõ. Trong ngôn ngữ phổ thông hiện nay gọi là “điếc”.

Trong ngành y, khiếm thính hay là điếc là việc suy giảm hay là mất toàn bộ hay một phần sức nghe. Sức nghe được đo bằng những bài kiểm tra sức nghe dựa trên mức độ phát âm thanh từ nhỏ hơn 20dB tới hơn 90dB.

Nghe bình thường: nghe được âm thanh nhỏ hơn 20dB (có thể nghe được lời nói thầm)

Điếc mức 1 (nhẹ): nghe được âm thanh từ 21dB đến 40 dB (chỉ nghe được lời nói bình thường khi đứng cách không quá 1 mét)

Điếc mức 2 (trung bình/ điếc vừa): nghe được âm thanh từ 41 dB đến 70 dB (chỉ nghe được lời nói lớn khi đứng cách không quá 1 mét)

Điếc mức 3 (nặng): nghe được âm thanh từ 71 dB đến 90 dB (chỉ có thể nghe được khi hét sát tai)

Điếc mức 4 (sâu): nghe được âm thanh từ 71 dB đến 90 dB (không nghe được khi hét sát tai)

Vậy có thể nói, người khiếm thính/điếc là người mắc tật không nghe được hoặc nghe không rõ.

Những người mắc khuyết tật khiếm thính/điếc thường do hai nguyên nhân: bẩm sinh và mắc phải (môi trường lao động, tai nạn, chăm sóc y tế, thói quen, ...)

Người câm là người không có khả năng nói, có thể do bẩm sinh hoặc do các rối loạn chức năng bên trong cơ thể, ảnh hưởng đến việc phát âm. Thông thường người điếc nặng từ nhỏ thường bị câm do, ngôn ngữ không phải cứ sinh ra là có mà cần phải thông qua quá trình học tập mới có được. Nếu từ nhỏ đã không cảm nhận được âm thanh thì rất khó để học nói. Tuy nhiên, nếu phát hiện trẻ bị điếc sớm thì cũng có thể dạy trẻ nói được qua khẩu hình miệng. Do vậy, điếc không phải là nguyên nhân khiến người điếc trở thành người câm. Vì thế, người điếc không hẳn đã là người câm, nhưng dường như những người câm bẩm sinh đều điếc. Để giao tiếp họ thường dùng ngôn ngữ cử chỉ (ngôn ngữ kí hiệu)

## **I.2. Một số đặc điểm tâm lý của người khiếm thính / người câm**

### **I.2.1 Đặc điểm sinh hoạt**

Do đặc điểm sinh lý và rào cản tâm lý của những người xung quanh khiến những người tàn tật gặp khó khăn trong giao tiếp. Điều này lâu dần khiến người tàn tật cảm thấy mình bị xa lánh và không tự tin trong tiếp xúc, sinh hoạt với cộng đồng.

Tuy nhiên so với những cộng đồng người tàn tật khác, người khiếm thính có nhiều lợi thế hơn. Do đặc điểm khuyết tật của mình học vẫn có thể tự làm được nhiều việc phục vụ cá nhân và tham các hoạt động của cộng đồng. Điều này khiến cộng đồng người khiếm thính khá phát triển và nhiều sinh hoạt tập thể mạnh mẽ. Ở Việt Nam sinh hoạt cộng đồng của người khiếm thính rất mạnh mẽ và thường xuyên với sự tham gia của các thành viên đến từ Câu lạc bộ Người Điếc Hà Nội, Hải Phòng, Nam Định, Ninh Bình, Thái Bình, Thái Nguyên, Đồng Nai, Thành phố Hồ Chí Minh,... Trên thế giới có Liên đoàn Người Điếc (World Federation of the Deaf) với các thành viên là Hiệp hội người Điếc quốc gia của hàng chục quốc gia.

Hoạt động của người khiếm thính cũng rất đa dạng như thể thao, du lịch. Olympic người khiếm thính diễn ra 4 năm 1 lần, giải golf cho người khiếm thính của hiệp hội golf Hoa Kỳ, các cuộc thi Hoa hậu người khiếm thính cũng được tổ chức thường xuyên ở một số quốc gia. Việc trao đổi thông tin giữa cộng đồng người khiếm thính thông qua các diễn đàn trực tuyến trên internet cũng rất rộng rãi. Họ sử dụng internet như một công cụ hữu ích để gặp gỡ, giao lưu, phát triển.

Một số diễn đàn nổi tiếng như: <http://www.deafzone.com>; <http://www.alldeaf.com/>; <https://www.drdvietnam.org/>

Tại các diễn đàn này ta có thể thấy người khiếm thính trao đổi với nhau về mọi mặt của cuộc sống từ công việc, sở thích, sức khỏe tới kinh nghiệm cuộc sống, tính yêu, tình bạn, giáo dục giới tính. Qua đó có thể thấy người khiếm thính thể hiện tính cộng đồng rất cao. Họ thường xuyên có những sinh hoạt tập thể, cộng đồng như giao lưu học tập, du lịch, dã ngoại.

Từ những nội dung trao đổi trong đời sống sinh hoạt thường nhật của người khiếm thính cũng có thể thấy rằng người khiếm thính đang vượt qua những trở ngại tâm lý của bản thân và xã hội, họ có thể dần dần tự tin và tự hào về bản thân mình. Ý thức về bản thân được nâng cao giúp họ hiểu rõ giá trị của mình và của những người cùng hoàn cảnh, vì vậy tính cộng đồng càng mạnh mẽ hơn. Ngoài ra họ cũng nhận thức được ý nghĩa quan trọng của những hoạt động xã hội đối với đời sống của cộng đồng mình.

### **I.2.2 Đặc điểm cảm giác - tri giác của người khiếm thính**

Như ta đã biết, cảm giác và tri giác là nền tảng của nhận thức. Chúng là những nguồn gốc cơ bản của những kiến thức mà chúng ta nhận thức được ở thế giới xung quanh. Trong những dạng cảm giác khác nhau thì cảm giác nghe và cảm giác nhìn có ý nghĩa chủ yếu. Chúng ta sống trong thế giới của âm thanh, của hình dạng và màu sắc. Những nguồn thông tin như phát thanh, truyền hình, phim ảnh, sân khấu, âm nhạc về nhiều mặt đưa đến cảm giác nghe. Tất nhiên cảm giác nhìn cũng đóng một vai trò quan trọng. Nhưng mất sức nghe sẽ làm cho người câm điếc mất khả năng tri giác bình thường về những nguồn thông tin này. Trong việc tiếp nhận ngôn ngữ, cảm giác và tri giác nghe có một vai trò đặc biệt quan trọng. Trên cơ sở này diễn ra sự phát triển các hình thái chủ động và bị động của lời nói. Nghe được tiếng nói của người xung quanh, đứa trẻ bắt đầu bắt chước và bập bẹ được những từ đầu tiên. Nhờ lời nói đứa trẻ nhận được những thông tin cơ bản, lĩnh hội những kiến thức và kinh nghiệm mà người lớn truyền cho nó. Sự phá huỷ tri giác và tiếng nói của người xung quanh tự nhiên sẽ kéo theo sự phá huỷ quá trình hình thành ngôn ngữ tích cực. Người khiếm thính không thể tự mình lĩnh hội được ngôn ngữ. Trong thực tế, trẻ khiếm thính sẽ bị câm nếu nó không được phát hiện sớm những khó khăn về thính giác và được hỗ trợ bằng những phương pháp chuyên biệt trong việc tiếp nhận ngôn ngữ.

Ở người khiếm thính, do thiếu cảm giác nghe hoặc cảm giác nghe bị phá huỷ, cảm giác thị giác và cảm giác vận động có một vai trò đặc biệt quan trọng. Thị giác của người khiếm thính trở thành chủ đạo và chủ yếu trong việc nhận thức thế giới xung quanh và trong việc tiếp nhận ngôn ngữ. Một đứa trẻ bình thường học nói chủ yếu dựa trên cảm giác nghe và vận động, còn tri giác thị giác đóng vai trò thứ yếu. Điều này hoàn toàn ngược lại với trẻ khiếm thính. Cùng với cảm giác vận động, cảm giác tri giác nhìn trở thành nền tảng để hình thành tiếng nói. Thậm chí trẻ khiếm thính có thể tiếp nhận ngôn ngữ chỉ dựa trên tri giác nhìn. Rất nhiều những nghiên cứu đã chứng minh rằng cảm giác và tri giác ở trẻ khiếm thính không kém so với trẻ nghe được, thậm chí còn tích cực và tinh nhạy hơn. Bởi vậy, trẻ khiếm thính thường để ý những chi tiết nhỏ của thế giới xung quanh mà trẻ bình thường không để ý đến. Ví dụ:

Phân biệt màu sắc: việc phân biệt những màu sắc gần giống nhau như: xanh, đỏ, da cam thì người khiếm thính phân biệt tinh tế hơn so với người bình thường.

Phân biệt người tiếp xúc: người khiếm thính có thể nhận thấy từng chi tiết về khuôn mặt, thân hình, cách ăn mặc, màu sắc và chất liệu của quần áo nhanh hơn so với người bình thường.

Ở người khiếm thính, tri giác phân tích thường trội hơn tri thức tổng giác. Mặc dù tất cả những khó khăn tâm lý và sự phức tạp của quá trình tri giác nhìn đối với ngôn ngữ nói, người khiếm thính thường làm chúng ta ngạc nhiên bằng khả năng dùng thị giác tiếp nhận và phân biệt tinh tế những gì mà chúng ta nói với họ.

Ở người bình thường có tồn tại mối liên hệ chặt chẽ giữa cơ quan thính giác và vận động. Ở người khiếm thính, sự mất thính lực không chỉ ảnh hưởng xấu đến sự vận động của bộ máy hô hấp mà còn ảnh hưởng đến sự phối hợp các động tác của cơ thể. Vì vậy, người khiếm thính thường vụng về không khéo léo, rất khó khăn với những kỹ năng lao động và thể thao đòi hỏi sự phối hợp tinh tế và sự thăng bằng của các động tác. Điều này được giải thích là do bộ máy tiền đình cũng như những điểm cuối dây thần kinh của cơ quan vận động bị tổn thương. Xúc giác-rung của người khiếm thính là đặc thù và độc đáo nhất. Đây là phương tiện quan trọng trong tiếp nhận ngôn ngữ cho người khiếm thính.

### **1.2.3 Đặc điểm ngôn ngữ của người khiếm thính**

Tiếng nói và ngôn ngữ là công cụ mạnh mẽ để nhận thức thế giới xung quanh. Nhờ từ ngữ, con người có khả năng khái quát hoá, trừu tượng hoá. Con người có thể nhận thức cả những đặc tính của thế giới xung quanh mà sự quan sát, tri giác không thể cảm nhận được. Sự phát triển trí tuệ của con người phụ thuộc nhiều vào ngôn ngữ. Một đứa trẻ nắm được ngôn ngữ, trong quá trình giao tiếp có thể biết những đặc tính của những vật xung quanh nó. Nó luôn luôn đặt những câu hỏi với người xung quanh và nhận được những câu trả lời, thu nhận được những kinh nghiệm của người lớn. Vào thời điểm 2 đến 3 tuổi, quá trình phát triển tiếng nói và tư duy diễn ra đặc biệt mãnh liệt. Ngôn ngữ liên hệ chặt chẽ với tư duy. Mối liên hệ này thể hiện trước hết ở chỗ: tiếng nói là công cụ của tư duy. Ý nghĩ của chúng ta xuất hiện và hình thành trên cơ sở tiếng nói. Tư duy bằng ngôn ngữ là hoàn thiện nhất vì nó có khả năng trừu tượng hoá không giới hạn.

Từ vựng và cấu trúc ngữ pháp là những bộ phận quan trọng cấu thành tiếng nói. Từ vựng đôi khi còn gọi là “vật liệu xây dựng” của tiếng nói. Từ vựng càng giàu thì tiếng nói càng phong phú. Nhưng chỉ có riêng từ vựng thì chưa tạo thành được ngôn ngữ, nó chỉ trở thành sức mạnh thực tế khi nó được sử dụng theo ngữ pháp, làm cho ngôn ngữ của chúng ta trở nên có cấu trúc và có nghĩa. Một yếu tố rất quan trọng của tiếng nói chúng ta là cái vỏ âm thanh, thành phần ngữ âm. Cái vỏ âm thanh của tiếng nói dường như là “cái vỏ vật chất” của nó. Chúng ta chỉ có thể diễn đạt ý nghĩ nhờ bọc chúng vào vỏ bọc âm thanh hay là cái vỏ đồ hoạ (chữ viết). Hơn nữa, trong mỗi từ đều có yếu tố khái quát, chính điều đó mở rộng khả năng giao tiếp và nhận thức. Sắc thái xúc cảm của từ là yếu tố rất quan trọng song còn ít được nhận thấy, dường như bị che lấp. Chúng ta không đặc biệt coi trọng nó trong cuộc sống hàng ngày. Nhưng ở đâu mà nhu cầu ngôn ngữ tăng lên, ở đâu mà từ có vai trò đặc biệt để diễn đạt sắc thái của ý nghĩ, thì yếu tố đó của từ có vai trò rất cơ bản, ví dụ sự sáng tạo trong ngôn ngữ thơ ca.

Sự phá huỷ thành phần từ của ngôn ngữ cũng có thể biểu hiện ở những hình thức khác nhau. Trường hợp nặng nhất là hoàn toàn không có khả năng tự chiếm lĩnh được từ (trường hợp điếc hoàn toàn). Trong những trường hợp khác thì điều đó có thể biểu hiện ở sự nghèo nàn và cực kỳ hạn chế của từ vựng, sự dùng từ không sát đúng

với ý nghĩa cơ bản của nó. Những thiếu sót tương tự thường gặp ở những người bị giảm sức nghe, cũng như những trẻ thiếu ngôn ngữ. Trên cơ sở sự phá huỷ ngôn ngữ nói thường xuất hiện sự phá huỷ ngôn ngữ viết và cấu trúc ngữ pháp của nó. Ở những người bị phá huỷ sức nghe thường thể hiện chứng viết khó và chứng mất ngữ pháp. Trong trường hợp bị chứng viết khó, thành phần chữ cái của từ bị bóp méo. Những chữ cái riêng lẻ thường bị bỏ qua, thay thế hoặc đổi chỗ cho nhau. Những sự phá huỷ này có thể liên hệ không chỉ với những thiếu sót của sự tiếp nhận âm thanh và phân tích âm, mà còn liên hệ với sự phá huỷ cảm giác và tri giác nhìn hay cảm giác và tri giác vận động.

#### **I.2.4 Đặc điểm trí nhớ của người khiếm thính**

So với người nghe được, người khiếm thính ghi nhớ tốt hơn những từ biểu thị chất lượng của những đồ vật tiếp nhận được nhờ xúc giác. Các nhà khoa học cũng xác định được rằng người khiếm thính có thể ghi nhớ những từ biểu thị những hiện tượng âm thanh. Thậm chí có khả năng nhớ tốt hơn người khác những từ biểu thị những âm phát ra từ những con vật nuôi trong nhà và những từ phát ra từ tiếng máy, khó ghi nhớ những từ biểu thị những âm thanh cường độ nhỏ. Ở người khiếm thính, biểu thị về âm thanh của các khách thể xuất hiện dựa trên hoạt động của những giác quan còn lại. Việc ghi nhớ những từ thuộc phạm vi những hiện tượng âm thanh diễn ra nhờ sự hoạt động phức tạp của mỗi loạt những cơ quan chức năng của người điếc: đó là sự hoạt động đồng thời và tác động qua lại của cơ quan thị giác, xúc giác, vận động và cảm giác-rung.

Trong quá trình ghi nhớ tư liệu, người khiếm thính ít sử dụng thủ thuật so sánh.

Nhưng bù lại, họ ghi nhớ tư liệu thị giác trực tiếp tốt hơn người nghe được vì họ có kinh nghiệm thị lực phong phú hơn. Với loại tư liệu khó diễn đạt bằng lời, người khiếm thính ghi nhớ kém hơn, nhưng khi họ có thể sử dụng chữ viết để biểu thị thì mức độ ghi nhớ của họ không thua kém gì so với người nghe được. Hơn nữa người khiếm thính không chỉ sử dụng cách biểu thị bằng lời mà còn cử chỉ điệu bộ. Điều này cũng có ý nghĩa tích cực đối với sự ghi nhớ của họ.

#### **I.2.5 Đặc điểm tư duy - tưởng tượng của người khiếm thính**

Như chúng ta đã biết, ngôn ngữ là nhận tố quan trọng nhất hình thành các khái niệm, là phương tiện phát triển tư duy trừu tượng. Điều tự nhiên là trong những trường hợp không có ngôn ngữ hay là ngôn ngữ phát triển muộn màng hoặc có những sai lệch, sẽ làm hạn chế không chỉ quá trình hình thành tư duy mà cả quá trình hình thành trí tưởng tượng nữa.

Những đặc điểm của tưởng tượng ở trẻ khiếm thính có sự thiếu hụt là do sự hình thành ngôn ngữ chậm và tư duy trừu tượng hạn chế gây nên. Mặc dù hình tượng thị giác của trẻ điếc đạt mức độ cao và sống động, nhưng sự hình thành tư duy bằng khái

niệm quá chậm, làm chúng rất khó thoát khỏi cái ý nghĩa cụ thể, nghĩa đen của từ, điều đó làm khó khăn cho sự hình thành hình tượng mới.

Tưởng tượng tái tạo có một ý nghĩa đặc biệt trong hoạt động nhận thức của người câm điếc. Nhờ tưởng tượng tái tạo, thế giới xung quanh được phản ánh trong ý thức của họ rộng hơn. Tầm hiểu biết của họ được mở rộng qua giới hạn kinh nghiệm cá nhân, đưa họ tiếp xúc với kho tàng kinh nghiệm của loài người.

### **I.3. Các nghiên cứu trước đây về công nghệ hỗ trợ người câm điếc**

Để cho phép người câm điếc giao tiếp dễ dàng hơn với người không câm điếc, tham gia đầy đủ vào các khía cạnh của cuộc sống. Việc nghiên cứu các công cụ, thiết bị hỗ trợ cho người khiếm thính nói riêng và người khuyết tật nói chung đã được chính phủ và các tổ chức trên thế giới đặc biệt là trong các trường đại học, viện nghiên cứu phát triển với sự hỗ trợ của chính phủ và các tổ chức phúc lợi xã hội.

Các thiết bị tạo giọng nói (Speech Generating Device (SGDs)), còn được gọi là thiết bị hỗ trợ giao tiếp đầu ra bằng giọng nói, là các hệ thống giao tiếp tăng cường và thay thế điện tử được sử dụng để bổ sung hoặc thay thế lời nói hoặc viết cho những người bị Câm, cho phép họ giao tiếp bằng lời nói.

Các thiết bị tạo giọng nói được nghiên cứu phát triển từ khá sớm. Thiết bị đầu tiên phải kể đến là máy đánh chữ (patient-operated selector mechanism (POSSUM)) được tạo ra bởi Reg Maling ở Vương quốc Anh vào năm 1960. POSSUM được quét qua một bộ kí hiệu trên màn hình được chiếu sáng. Các nhà nghiên cứu tại Đại học Delft ở Hà Lan đã tạo ra máy đánh chữ hoạt động bằng đèn (lightspot operated typewriter (LOT)) vào năm 1970, sử dụng các di chuyển nhỏ của đầu chiếu sáng để hướng một điểm sáng nhỏ vào ma trận các kí tự, mỗi điểm được trang bị một tế bào quang điện. Mặc dù nó không thành công về mặt thương mại, nhưng LOT đã được người dùng đón nhận.

Nhà vật lý lý thuyết người Anh Stephen Hawking , một trong những người nổi bật nhất của công nghệ giao tiếp SGD, ông giao tiếp thông qua một DECtalk tổng hợp giọng nói DTC01 phát triển bởi Digital Equipment Corporation vào đầu những năm 1980. Hawking sử dụng má của mình để nhấn nút chuyển đổi giúp ông chọn các chữ cái và từ khi giao tiếp.



Hình 1: Stephen Hawking với thiết bị DECtalk

Các thế hệ thiết bị SGD tiếp theo, là các thiết bị có giao tiếp màn hình cảm ứng, chẳng hạn như Dynavox V và Dynavox Max [2] là các thiết bị dành cho những người câm điếc. Chúng là những máy tính bảng Windows có phần mềm bổ sung. Thông tin viên giao tiếp thường lớn, thời lượng pin ngắn, khó sử dụng và rất tốn kém. Giá hiện tại của DynaVox V là 7.820 đô la và Dynavox Max có giá 8.420 đô la. Phần mềm khá phức tạp, người sử dụng phải trải qua nhiều giờ học cách sử dụng DynaVox mới có thể sử dụng được một cách hết sức cơ bản.

Với sự phát triển của công nghệ cảm biến, năm 1999, một đề tài nghiên cứu vấn đề giao tiếp này đã đưa ra một hệ thống gắng tay đơn giản có thể học cách giải mã các cử chỉ ngôn ngữ ký hiệu và xuất ra các từ tương ứng.

Kể từ đó, nhiều nhóm, chủ yếu là các dự án đại học hoặc sau đại học, đã tạo ra các nguyên mẫu của riêng họ, mỗi nguyên mẫu có ưu điểm và mục đích thiết kế riêng.

Khi nói đến vấn đề dịch các cử chỉ ngôn ngữ ký hiệu một cách trực quan, các nhà nghiên cứu tìm đến gắng tay, vì ngôn ngữ ký hiệu liên quan đến việc sử dụng rộng rãi chuyển động tay. Bằng sáng chế số 5.047.952 vào năm 1981 của Mỹ sử dụng khái niệm “Dụng cụ Gắng tay” nhưng nó không được sử dụng cho ngôn ngữ ký hiệu. Gần hai thập kỷ sau, ý tưởng về một chiếc gắng tay được thiết kế để dịch ngôn ngữ ký hiệu lại được đăng ký với Bằng sáng chế với tiêu đề “Hệ thống thông tin liên lạc dành cho người khiếm thính, người câm sử dụng gắng tay dụng cụ” (Communication system for



deaf, deaf-blind, or non-vocal individuals using instrumented glove) là một trong những thử nghiệm đầu tiên dựa trên nguyên tắc sử dụng cảm biến trong găng tay để tạo điều kiện giao tiếp giữa người nghe được và người không nghe được. Mặc dù bằng sáng chế này đã được nộp vào năm 1999, nhưng không có thêm thông tin nào về việc thực hiện bằng sáng chế. Một số bằng sáng chế khác có công nghệ tương tự cũng được công nhận bằng sáng chế ở Mỹ.

Thiết bị “Enable Talk” là một dự án của sinh viên đoạt giải “Imagine Cup” của Microsoft vào năm 2012. Thiết bị này bao gồm một thiết bị di động và một đôi găng tay dịch bảng chữ cái ngôn ngữ Mĩ (ASL) thành văn bản trên thiết bị di động (Hình 2). Mỗi chiếc găng tay được trang bị cảm biến tiếp xúc, cảm biến flex và vi điều khiển.



Hình 2: Thiết bị đạt giải “Microsoft Imagine Cup” năm 2012

Một thiết kế tương tự khác của đại học Cornell University vào năm 2014 (Hình 3)., Thiết bị này cũng là một Găng tay có khả năng phiên dịch một số tín hiệu ngôn ngữ của cử chỉ bàn tay sang giọng nói tiếng Anh. Thiết kế của nó bao gồm các cảm biến uốn cong (Flex), cảm biến con quay hồi chuyển 3 chiều (three-axis gyroscope) và một vi điều khiển Atmega1280 và một máy tính cá nhân. Tín hiệu đầu vào (tín hiệu cử chỉ bàn tay) được bộ vi điều khiển gửi đến máy tính PC qua giao tiếp USB. Thuật toán học máy cài đặt trên PC được đào tạo thông qua các bộ dữ liệu cho trước để bắt được các ngôn ngữ ký hiệu.



Hình 3: Găng tay chuyển đổi ngôn ngữ ký hiệu của Cornell University

Kết quả nghiên cứu của Dự án này không đưa được sản phẩm ra thị trường như là một sản phẩm thương mại, nhưng là một ví dụ điển hình về học máy để nhận dạng cử chỉ. Nó có nhiều cảm biến để nhận dạng cử chỉ chính xác hơn, sử dụng cảm biến uốn cong dạng mỏng. Nhưng không may thuật toán máy học phải sử dụng máy tính để giải mã các tín hiệu cử chỉ bàn tay. Khiến Găng tay không thể di động và không sử dụng cho các tính huống hàng ngày.

Sự phát triển gần đây nhất trong dịch thuật ngôn ngữ ký hiệu là từ các sinh viên kỹ thuật và giáo sư tại Đại học Texas A&M University. Theo một bài báo của Discovery news, họ đã phát triển một thiết bị nhận dạng rất chính xác dựa trên cảm biến điện tử trên da, gia tốc kế và con quay hồi chuyển và phân tích hình ảnh (Hình 4).



Hình 4: Hệ thống nhận dạng ngôn ngữ ký hiệu của Texas A&M University

Nó cũng sử dụng thuật toán học máy có giám sát, trong đó người dùng phải trải qua một thời gian đào tạo và mỗi cử chỉ được thực hiện nhiều lần trong suốt thời gian đào tạo này. Tuy nhiên, việc sử dụng nhiều cảm biến trên khắp cánh tay rất hữu ích để cung cấp tín hiệu nhận dạng chính xác cử chỉ bàn tay, cánh tay, hoạt động của tay,.. Giúp cho việc nhận dạng và chuyển đổi NNKH sang giọng nói chính xác, nhưng nó cũng là nhược điểm của hệ thống này vì nhiều thành phần dán khắp tay, có thể tạo cho người sử dụng cảm giác không thoải mái và đặc biệt là rất đắt tiền.

Với việc phổ biến của điện thoại di động thông minh có cấu hình phần cứng ngày càng cao, nó có thể được người Câm sử dụng như một thiết bị giao tiếp tăng cường và thay thế (augmentative and alternative communication (AAC)) bằng cách cài đặt thêm các phần mềm như Proloquo2Go có giá 250 đô la. Proloquo2Go là một ứng dụng giao tiếp dựa trên biểu tượng giúp cho người Câm giao tiếp với người bình thường có giao diện như hình 5.



Hình 5: Giao diện ứng dụng Proloquo2Go

Khi cần giao tiếp người sử dụng sẽ nhấn các biểu tượng trên màn hình điện thoại, điện thoại sẽ phát ra tiếng nói giúp người bình có thể hiểu và giao tiếp được với người khiếm thính.

#### I.4. Mục tiêu và phạm vi nghiên cứu

Mục tiêu chính của nhiệm vụ là nghiên cứu chế tạo gắng tay thông minh hỗ trợ người câm điếc (những người không có khả năng nói) trong giao tiếp với người bình thường bằng cách phát ra những câu nói (tự định nghĩa) từ ngôn ngữ cử chỉ bàn tay, cụ thể phạm vi của đề tài bao gồm các nội dung sau:

- Nghiên cứu tổng quan về người khiếm thính, các đặc điểm tâm lý của người khiếm thính (Chương này);
- Tìm hiểu ngôn ngữ kí hiệu, ngôn ngữ kí hiệu và Việt Nam và ngôn ngữ ký hiệu tiếng Anh
- Nghiên cứu phát triển hệ thống gắng tay thông minh; gắng tay có thể phiên dịch (chuyển đổi) một số ký hiệu ngón tay thành giọng nói, có thể cho phép người dùng tự định nghĩa, thay đổi câu nói một số câu nói thường dùng của một cử chỉ, ký hiệu riêng phù hợp với văn hóa, phong tục từng địa phương hoặc hoàn cảnh sử dụng phù hợp với cử chỉ của từng người.
- Thử nghiệm và đánh giá kết quả của thiết bị

## CHƯƠNG II: NGÔN NGỮ KÍ HIỆU, CÔNG NGHỆ NHẬN DẠNG NGÔN NGỮ KÍ HIỆU

*Chương này trình bày về quá trình hình thành và phát triển ngôn ngữ kí hiệu và đặc điểm của ngôn ngữ kí hiệu của người điếc câm ở một số nước trên thế giới và Việt Nam.*

### II.1. Quá trình hình thành và phát triển ngôn ngữ kí hiệu

Ngôn ngữ là công cụ giao tiếp hữu dụng nhất của con người. Có ngôn ngữ con người có thể nghe, nói, đọc, viết để học hỏi và khám phá thế giới. Người điếc do không nghe được nên khả năng phát âm bị hạn chế, vì thế họ cực kì khó khăn trong giao tiếp – một nhu cầu bức thiết trong cuộc sống hàng ngày. Để đáp ứng nhu cầu giao tiếp, cộng đồng người khiếm thính phải sử dụng ngôn ngữ riêng, được gọi là ngôn ngữ kí hiệu (NNKH). Vậy loại hình ngôn ngữ đặc biệt này đã hình thành và phát triển như thế nào?

Theo nghiên cứu về nguồn gốc NNNH của David F. Armstrong và Sherman Wilcox thì: “NNKH” hay chính xác là hệ thống kí hiệu, đã được xác định xuất hiện trong quá trình nghe của con người dưới những điều kiện sau đây:

- Trong cộng đồng tu sĩ đạo cơ đốc với cuộc sống theo luật im lặng
- Như là một ngôn ngữ đặc biệt của thổ dân Úc
- Như là một loại ngôn ngữ bản địa chung trong cộng đồng người da đỏ ở đồng bằng Bắc Mỹ
- Trong số những người công nhân khắc khổ làm việc trong môi trường tiếng ồn lớn
- Dùng cho những người thợ săn khi tránh gây tiếng động cho thú săn

Vào năm 1620, Juan Pablo Bonet - một linh mục người Tây Ban Nha, đã xuất bản cuốn sách đầu tiên về các kí hiệu giao tiếp của người khiếm thính tại Madrid. Trong cuốn sách này tác giả công bố bảng chữ cái ngón tay dựa trên nền tảng các kí hiệu mà các kí hiệu này đã được cộng đồng người khiếm thính phát triển theo bản năng từ trước. Từ đó người ta bắt đầu quan tâm đến giáo dục cho người khiếm thính, giao tiếp với người khiếm thính, NNNH của người khiếm thính,... đặc biệt là ở Pháp.

Năm 1755: Cha Charles-Michel de l'Épée (người Pháp và được coi là người khai sinh ra hệ thống ngôn ngữ kí hiệu Pháp) thành lập trường học miễn phí đầu tiên dành cho người điếc – Viện Nationale des Sourds Muets à Paris. Dựa trên cuốn sách của Juan Pablo Bonet, ông đã cùng Pierre Desloges – một người bị khiếm thính do bị bệnh đậu mùa lý giải việc hình thành “Đánh vần bằng tay” (Fingerspelling) và xây dựng thêm các cử chỉ mang ý nghĩa cho toàn bộ cụm từ hoặc các từ trong giao tiếp.

Hệ thống kí hiệu tiếp tục được phát triển và được cộng đồng người điếc sử dụng. Hệ thống ngôn ngữ kí hiệu của Pháp được hoàn thiện trong giai đoạn này.

Năm 1778: Tại Leipzig, Đức, Samuel Heinicke, trường công lập đầu tiên dành cho người điếc không chỉ sử dụng ngôn ngữ kí hiệu mà còn dùng phương pháp nói và đọc khẩu hình (speech-reading) – tiên phong cho việc dùng tất cả các phương pháp để giao tiếp tối ưu (dùng tất cả các biện pháp giao tiếp có thể: ngôn ngữ kí hiệu, cử chỉ, đánh vần bằng kí hiệu, đọc khẩu hình, nói, trợ thính, đọc, viết và tranh vẽ).

Năm 1815: Thomas Hopkins Gallaudet tới châu Âu nghiên cứu phương pháp giáo dục dành cho người điếc. Trở lại Hoa Kỳ cùng với giáo viên ngôn ngữ kí hiệu, Gallaudet và Laurent Clerc mở trường công dành cho người điếc đầu tiên của Hoa Kỳ tại Hartford, Connecticut năm 1817.

Năm 1924: tổ chức World Games đầu tiên dành cho người khiếm thính được tổ chức. Từ những hoạt động mang tầm quốc tế này đã khiến người ta nghĩ về một NNKH chuẩn quốc tế. Hội nghị của Liên đoàn Khiếm thính Thế giới năm 1951 ( tổ chức tại Roma) đã bàn đến vấn đề đó. Đến năm 1973 một hệ thống kí hiệu chuẩn quốc tế ( được gọi là Gestuno – bằng tiếng Ý) đã ra đời. Hội đồng biên soạn cho ấn hành thành sách với khoảng 1500 kí hiệu. Họ đã cố gắng chọn những kí hiệu dễ hiểu nhất từ nhiều NNKH cho Gestuno. Tuy nhiên do Gestuno được nhìn như một hệ thống những điệu bộ và không có một ngữ pháp cụ thể, nên nhiều người cho rằng nó chưa phải là một ngôn ngữ. Vì vậy nó chỉ thực sự hữu ích trong các sự kiện như hội nghị hay Olympic thể thao quốc tế cho người khiếm thính chứ không thể là NNKH dùng chung cho cả cộng đồng khiếm thính trên thế giới.

Ngôn ngữ kí hiệu bắt đầu được công nhận là loại hình ngôn ngữ hình ảnh phức tạp, đầy đủ với cấu trúc ngữ pháp khác nhiều so với ngôn ngữ nói. NNKH tồn tại trên toàn thế giới, mỗi người khiếm thính đến với nhau bằng vốn từ vựng và các quy tắc kí hiệu riêng của mình. Mặc dù là ngôn ngữ hình ảnh đặc biệt, nhưng các hệ thống NNKH có rất nhiều điểm chung và nhiều điểm chung hơn so với ngôn ngữ nói.

## **II.2. Ngôn ngữ kí hiệu Mỹ**

Ngôn ngữ kí hiệu Mỹ (American Sign Language (ASL)) được đánh giá là một trong những hệ thống NNKH hoàn chỉnh và tiên tiến nhất trên thế giới, được phổ biến và được sử dụng ở nhiều nơi trên thế giới. Ngày càng nhiều các tiểu bang ở Mỹ thông qua đạo luật nhìn nhận NNKH Mỹ như một ngoại ngữ và cho phép các trường trung học và đại học chấp nhận nó như là một tiêu chí ngoại ngữ. Tháng 7 – 1997, 28 bang đã thông qua đạo luật này và nhiều trường cao đẳng và đại học (kể cả các trường đại học ở California, Massachusetts, Institute of Technology, Brown university, Georgetown University và University of Washington) chấp nhận ASL như một tín chỉ ngoại ngữ bắt buộc hoặc tự chọn cho tất cả sinh viên.



Mặc dù có phạm vi sử dụng rộng rãi, nhưng cũng không có con số thống kê chính xác số người sử dụng ASL, mặc dù ước tính đáng tin cậy cho người sử dụng ASL Mỹ khoảng từ 250.000 đến 500.000 người bao gồm cả trẻ em và người lớn.

Các kí hiệu ASL có một số thành phần ngữ âm, bao gồm cả biểu cảm của khuôn mặt và thân mình cũng như cử chỉ bàn tay. ASL không phải là một hình thức của kịch câm. Các từ vay mượn tiếng Anh thường vay thông qua biểu đạt phát âm bằng động tác ngón tay, mặc dù ngữ pháp ASL không liên quan đến ngữ pháp tiếng Anh. Nhiều nhà ngôn ngữ học cho rằng ASL là một ngôn ngữ có Chủ ngữ - Động từ-Bổ ngữ, nhưng có một số đề xuất thay đổi về trật tự từ trong câu.



Trong khi các ngôn ngữ khác phải mất nhiều năm để học, ngôn ngữ kí hiệu của Mỹ có thể được dạy trong khoảng thời gian ngắn hơn nhiều. Đây là một trong những NNKH dễ học nhất vì hầu hết các dấu hiệu được phát triển để bắt chước từ hoặc cụm từ thực tế mà nó đại diện. Ngay cả các dấu hiệu bảng chữ cái trông giống như các chữ cái của bảng chữ cái tiếng Anh ( hình 6).

Mặc dù ASL được sử dụng ở Mỹ, nó là một ngôn ngữ hoàn toàn tách biệt với tiếng Anh. Nó chứa tất cả các tính năng cơ bản mà ngôn ngữ cần để tự hoạt động vì nó có các quy tắc riêng về ngữ pháp, dấu câu và trật tự từ trong câu.

Cú pháp là nghiên cứu về cách xây dựng câu. Cú pháp cũng đề cập đến các quy tắc và nguyên tắc cấu trúc câu. Trong ASL, cú pháp được chuyển tải thông qua trật tự

từ trong câu và diễn tả không dùng tay (non -manual ). Phần này rất dễ gây nhầm lẫn, vì vậy khá khó cho những người học mới học ASL.

### **II.2.1 Trật tự từ trong câu của ASL**

Câu ASL tuân theo cấu trúc TOPIC-COMMENT. Điều này giống với cấu trúc tiếng Anh “Chủ ngữ” “Vị ngữ”. Tuy nhiên, thay vì chủ đề của câu luôn là chủ ngữ, chủ đề trong ASL là bất cứ điều gì mà người nói hướng đến.

Chủ ngữ của câu là người hoặc đối tượng thực hiện hành động, động từ của câu là hành động và tân ngữ của câu là những gì đang nhận hành động. Chẳng hạn, trong câu: ““The boy kicked the ball” . Chủ ngữ là boy, kicked là vị ngữ, ball là tân ngữ. Có một vài biến thể khác nhau của trật tự từ trong ASL tùy thuộc vào từ vựng đang sử dụng và những gì người nói đang cố gắng thực hiện

#### **Trật tự từ với động từ đơn giản:**

Khi sử dụng các động từ đơn giản, các câu ASL có thể tuân theo nhiều trật tự từ khác nhau. Mặc dù tiếng Anh thường chỉ theo thứ tự từ Chủ ngữ-Động từ-Tân ngữ, khi nói Mẹ yêu cha (MOTHER LOVE FATHER), trong ASL có thể có những trật tự như sau: “MOTHER LOVE FATHER” , “SHE LOVE HIM”, “MOTHER LOVE HIM” hay “SHE LOVE FATHER”.

#### **Trật tự từ “Tân ngữ, Chủ ngữ, Động từ”:**

Trong ASL, người ta có thể sử dụng Chủ ngữ hoặc Tân ngữ làm CHỦ ĐỀ của câu. Sử dụng Chủ ngữ làm Chủ đề của câu được sử dụng trong câu chủ động (Active voice). Sử dụng Tân ngữ làm Chủ đề của câu khi nói theo kiểu câu bị động (passive voice).

#### **Trật tự từ trong câu không có Tân ngữ:**

Tất cả các câu sau đều được chấp nhận trong ASL, khi người thực hiện ngôn ngữ cử chỉ không đối với câu không có Tân ngữ.

Tuy nhiên, Động từ-Chủ ngữ (đặt động từ trước chủ ngữ của câu) sẽ không đúng trong ASL. Ví dụ,” STUPID MAN” không phải là một câu ASL chính xác.

#### **Thời gian – Chủ đề - nhận xét/ khuyến nghị (Time-Topic-Comment):**

Khi nói về một sự kiện trong quá khứ hoặc tương lai trong ASL, người ta sẽ làm dấu hiệu thời gian trước phần dấu hiệu các phần còn lại của câu. Điều này tạo ra cấu trúc TIME-TOPIC-COMMENT. Các quy tắc tương tự về trật tự từ cho cấu trúc TOPIC-COMMENT được áp dụng, chỉ là bây giờ một dấu hiệu thời gian (time sign) được thêm vào đầu của câu.

### **II.2.2 Các kiểu câu trong ngôn ngữ cử chỉ Mỹ**

Có một vài loại câu khác nhau trong ASL. Những loại câu này không giống như trật tự từ. Trật tự từ hiển thị thứ tự mà “người nói” thực hiện dấu hiệu. Các kiểu câu



cho biết cách sử dụng trật tự từ cùng với các diễn tả không sử dụng tay để tạo thành các loại câu phù hợp nhất định.

### II.2.2.1 Câu hỏi

Có ba loại câu hỏi được sử dụng trong ASL - câu hỏi wh-word, có / không có câu hỏi và câu hỏi tu từ. Cách duy nhất để nhận diện giữa các câu hỏi này trong ASL là sử dụng các dấu hiệu không dùng tay mà dùng các biểu cảm khuôn mặt, thái độ,...

#### **Câu hỏi với từ để hỏi WH (“WH” Word Questions):**

Câu hỏi từ là những câu hỏi đòi hỏi nhiều thông tin hơn câu trả lời có hoặc không. Đây là những câu hỏi thông thường sử dụng các từ ai, cái gì, khi nào, ở đâu, tại sao hoặc như thế nào. Từ wh thường được làm dấu hiệu ở cuối câu hỏi.

Ví dụ khi người thực hiện dấu hiệu HE/SHE WHO? Nghĩa là (“Who is he/she?”) hoặc YOU LEARN SIGN WHERE? (“Where are you learning sign?”) cùng với cử chỉ khuôn mặt như hình 10.



Hình 6: Thực hiện dấu hiệu câu hỏi dạng WH

*Các cử chỉ không dùng tay thể hiện câu hỏi WH như sau:*

- Nhúu mày
- Hơi ngả đầu về phía trước
- Giữ dấu hiệu cuối cùng trong câu ( thường là từ WH)

#### **Câu hỏi Có/Không (Yes/No Questions)**

Câu hỏi Có/Không chỉ yêu cầu câu trả lời Có hoặc Không.

- Nhướn mày

- *Hơi ngả đầu về phía trước*
- *Giữ dấu hiệu cuối cùng trong câu*



Hình 7: Dấu hiệu cử chỉ câu hỏi Yes/No

Ví dụ: DEAF YOU? (“Are you Deaf?”)

YOU MARRIED YOU? (“Are you married?”)

STUDENT HE/SHE? (“Is he/she a student?”)

### **Câu hỏi tu từ (Rhetorical Questions)**

Câu hỏi tu từ là câu hỏi được đặt ra không nhằm mục đích chờ đợi câu trả lời hoặc câu trả lời sẽ nằm ngay trong câu hỏi mà người ta đưa ra có tác dụng nhằm nhấn mạnh một ý nghĩa khác. Câu hỏi tu từ được đặt ra nhằm tập trung sự chú ý của người nghe vào mục đích cụ thể nào đó. Do đó, câu hỏi tu từ về hình thức là một câu hỏi nhưng thực chất đó là câu phủ định có cảm xúc hay câu khẳng định. Đây là cách sử dụng rất khác so với câu hỏi yes/no hoặc câu hỏi yêu cầu thông tin. Trong ASL các câu hỏi tu từ thường xuyên sử dụng từ hỏi “Tại sao” (Why) thay thế cho từ “Bởi vì” (Because ) kèm theo dấu hiệu không sử dụng tay (Non-Manual Markers) như hình 12. Các dấu hiệu không sử dụng tay cho loại câu hỏi này là:

- *Thực hiện câu hỏi ở trạng thái bình thường (neutral expression)*
- *Đặt câu hỏi với từ WH với lông mày nhướn lên*
- *Trả lời câu hỏi bằng một biểu hiện trung lập, khẳng định hoặc tiêu cực*



Hình 8: Dấu hiệu câu hỏi tu từ

Ví dụ như một số câu sau:

I HUNGRY, WHY? EAT LUNCH NOT.

(“I’m hungry. Why? I didn’t eat lunch”)

(“I’m hungry because I didn’t eat lunch”)

THAT WOMAN, WHO? MY MOM.

(“Who is that woman? My mom”)

(“That woman is my mom”)

I PASS CLASS, HOW? I STUDY.

(“I passed the class. How? I studied”)

(“I passed the class because I studied”)

### II.2.2.2 Câu trần thuật

Câu trần thuật trong ASL có thể là câu khẳng định, câu phủ định hoặc miêu tả, nó được nhận diện bởi các dấu hiệu không sử dụng tay.

#### Câu khẳng định

Dấu hiệu không sử dụng tay trong câu khẳng định;

- *Gật đầu trong khi ra dấu hiệu (sử dụng biểu cảm khuôn mặt phù hợp để thể hiện mức độ hoặc cường độ khẳng định)*

Ví dụ:

SHE DEAF SHE (“She is Deaf”)

I HUNGRY (“I’m hungry”)

WASH CAR FINISH (“I washed the car”)

### **Câu phủ định**

Dấu hiệu không dùng tay (Non-Manual Markers) khi dùng câu phủ định

- *Lắc đầu*
- *Mặt ngược lên*
- *Nhăn mặt*
- *Sử dụng biểu cảm khuôn mặt phù hợp để thể hiện mức độ hoặc cường độ phủ định*

Ví dụ:

ME GO CAN’T (“I can’t go”)

I HUNGRY (“I’m not hungry”)

I NOT HUNGRY (“I’m not hungry”)

I HAVE NONE CHILDREN (“I have no children”)

### **II.2.2.3 Câu điều kiện**

Câu điều kiện thường tuân theo cấu trúc Nếu/Thì ( IF/THEN) trong đó phần “Nếu” thường dùng biểu cảm khuôn mặt, phần “Thì” dùng các dấu hiệu khác. Biểu cảm khuôn mặt của dạng này như sau:

- *Nhướn lông mày trong thời gian “Nếu” của câu*
- *Sau đó, đặt câu hỏi hoặc trần thuật cho phần sau của câu*

*Ví dụ*

*TODAY RAIN, GAME CANCEL.*

*(“If it rains today, the game will be cancelled”)*

*TODAY RAIN, YOU LEAVE YOU?*

*(“If it rains today, are you going to leave?”)*

*MILK CHEAP, I BUY.*

(“If the milk is cheap, I will buy it”)



Hình 9: Làm dấu hiệu câu điều kiện

### II.3. Ngôn ngữ kí hiệu Việt Nam

#### II.3.1 Yếu tố cấu thành kí hiệu trong ngôn ngữ kí hiệu Việt Nam

Ngôn ngữ kí hiệu đã xuất hiện và được dạy cho người khiếm thính ở nước ta có từ khá sớm (Trường Cầm Đắc Lái Thiêu được thành lập năm 1886), tiền thân của Trung tâm của Trung tâm Giáo dục Trẻ khuyết tật Thận An là cái nôi nuôi dưỡng và đào tạo người Khiếm thính. Trường được linh mục chính xứ họ đạo họ đạo Lái Thiêu tên Azemar (còn được gọi là cha Lục) thành lập năm 1886. Từ năm 1866, cha Azemar lúc bấy giờ là cha sở họ đạo Lái Thiêu đã quy tụ khoảng 5 trẻ điếc để dạy ngôn ngữ và đạo đức. Đến năm 1880, cha gửi Nguyễn Văn Trường – một thanh niên điếc – sang Pháp để học về phương pháp dùng kí hiệu ngôn ngữ. Khi anh Trường về nước, cha Azemar chính thức tuyên bố mở trường dạy trẻ điếc vào năm 1886. Vì thế NNKH của Việt Nam cũng xuất phát từ NNKH Pháp (LSF) và hiện nay còn nhiều kí hiệu cơ bản giống kí hiệu của Pháp. Trong lịch sử hơn 100 năm hình thành và phát triển có những giai đoạn dài (1886 – 1936), ở đây người ta dùng phương pháp thuần kí hiệu để dạy cho người khiếm thính. Điều đó đồng nghĩa với việc ở Việt Nam, người khiếm thính sử dụng ngôn ngữ kí hiệu như tiếng mẹ đẻ và ngôn ngữ kí hiệu và tiếng Việt là hai thứ ngôn ngữ riêng biệt.

Do hoàn cảnh chiến tranh kéo dài, nên sự đào tạo, nghiên cứu NNKH ở Việt Nam có lúc bị gián đoạn. Cho đến năm 1996, một tiến sĩ ngôn ngữ học người Mỹ JAMES C. WOODWARD, người đã từng làm việc với William Stokoe tại trường đại học Gallaudet của Mỹ, đã sang Việt Nam thực hiện nghiên cứu về ngôn ngữ kí hiệu

của cộng đồng người điếc ở Việt Nam. Theo nghiên cứu của ông, ở Việt Nam hiện có ít nhất 3 ngôn ngữ kí hiệu phổ biến (được cộng đồng người điếc sử dụng nhiều nhất). Ông đã dùng tên của những địa danh này để đặt tên cho 3 ngôn ngữ kí hiệu đó: Ngôn ngữ kí hiệu Hà Nội, ngôn ngữ kí hiệu Hải Phòng, và ngôn ngữ kí hiệu Thành phố Hồ Chí Minh. Sau đó, đã có thêm những dự án ở Việt Nam: dự án Giáo dục hòa nhập cho trẻ điếc 1998-2001 (Viện Khoa học Giáo dục- tổ chức Pearl S. Buck, Int), dự án Giáo dục trung học và đại học cho người Điếc Việt Nam (Sở GD-ĐT Đồng Nai và GS TS JAMES C. WOODWARD) để thực hiện việc thu thập lại những dấu hiệu của người điếc Việt Nam và tìm hiểu về ngữ pháp của ngôn ngữ này. Công việc này đã kích thích thêm nhiều nhà khoa học ở Việt Nam cũng bắt đầu tìm hiểu về ngữ pháp của ngôn ngữ kí hiệu Việt Nam. Nhưng cho đến nay Việt Nam vẫn chưa có điều kiện thống nhất những kí hiệu cơ bản, chưa nghiên cứu để xác định và hình thành cho mình một hệ thống NNKH thực sự mang tên *Vietnamese Sign Language*. Nhưng Việt Nam đã triển khai những nỗ lực của mình nhằm hoàn thiện và hệ thống hóa Ngôn ngữ Ký hiệu Việt Nam. Các CLB, nhóm dạy, sinh hoạt NNKH bắt đầu hình thành và nở rộ, hiệu công trình nghiên cứu về ngôn ngữ kí hiệu của các nhà khoa học.

Về phương thức cấu tạo, ngôn ngữ kí hiệu Việt Nam và ngôn ngữ kí hiệu ở các nước trên thế giới có điểm chung rất rõ nét là: có 5 thành tố cơ bản hình thành nên sự khu biệt ngữ nghĩa của mỗi kí hiệu giao tiếp của người khiếm thính, đó là:

- Vị trí làm kí hiệu (Location)
- Hình dạng bàn tay (Handshape )
- Chuyển động của tay (Movenment)
- Chiều hướng của lòng bàn tay (Orientation)
- Sự diễn tả không bằng tay (Non manual)

Như vậy mỗi kí hiệu được xây dựng bởi sự phối hợp đồng thời 5 yếu tố này; chúng được tạo ra tất cả trong cùng một lúc – khác với tín hiệu của ngôn ngữ nói thông thường (gồm những âm vị, các nguyên âm và phụ âm) vốn có tính hình tuyến đi theo nhau, cái này sau cái kia. Các yếu tố này là các yếu tố cơ bản khu biệt ý nghĩa của các kí hiệu trong NNKH, chỉ cần khác một thành tố thì kí hiệu đã mang một ngữ nghĩa khác; Hiện nay, tuy chưa có sự thống nhất về kí hiệu của NNKH ở các vùng miền, song NNKH của người khiếm thính Việt Nam cũng không nằm ngoài quy luật chung trên;

### **II.3.2 Vài vấn đề về cú pháp trong NNKH Việt Nam**

Công trình Nghiên cứu cách biểu đạt ngôn ngữ kí hiệu của người điếc Việt Nam [3] của Viện Khoa học giáo dục đã bước đầu chỉ ra rằng người khiếm thính biểu đạt bằng ngôn ngữ kí hiệu không theo như trật tự từ thông thường. Điểm dễ nhận thấy nhất là trật tự kí hiệu trong một số loại câu của ngôn ngữ kí hiệu ở Việt Nam khác về

cơ bản so với trật tự từ trong câu tiếng Việt. So sánh các câu được thể hiện bằng ngôn ngữ nói thông thường và các câu được thể hiện bằng ngôn ngữ kí hiệu sau:

STT	Câu tiếng Việt	Câu bằng ngôn ngữ ký hiệu
1.	Cô ấy buồn.	Cô ấy buồn.
2.	Bạn viết đẹp lắm!	Bạn viết đẹp lắm!
3.	Tôi đi học.	Tôi đi học.
4.	Trưa nay, tôi ăn hai quả táo xanh.	Trưa nay, tôi ăn hai quả táo xanh.
5.	Tôi thương mẹ tôi nhất.	Tôi thương mẹ tôi nhất.
6.	Nhà tôi ở Quận 3, thành phố Hồ Chí Minh.	Nhà tôi ở Quận 3, thành phố Hồ Chí Minh.
7.	Con chưa uống sữa.	Con chưa uống sữa.
8.	Tôi thích ăn dưa hấu.	Tôi thích ăn dưa hấu.
9.	Tôi không thích ăn vú sữa.	Tôi không thích ăn vú sữa.
10.	Ai cho bạn mượn sách?	Ai cho bạn mượn sách?
11.	Em có bao nhiêu cái kẹo?	Em có bao nhiêu cái kẹo?
12.	Bạn thích ăn gì?	Bạn thích ăn gì?
13.	Gia đình của bạn có mấy người?	Gia đình của bạn có mấy người?
14.	Ngày mai là thứ ba đúng không?	(14a) Mai thứ 3 đúng sai? (14b) Mai thứ ba đúng (+ nét mặt)? (14c) Mai thứ ba (+ nét mặt )?
15.	Bạn có người yêu chưa?	Bạn có người yêu chưa?
16.	Bạn thích màu đen hay màu trắng?	(16a) Bạn đen trắng thích cái nào? (16b) Bạn đen trắng thích (+ nét mặt)?
17.	Ôi, bông hoa đẹp thế!	Ôi, bông hoa đẹp thế!
18.	Hãy mở cửa sổ ra!	Hãy mở cửa sổ ra!

Bảng 1: Một số câu tiếng Việt và câu bằng ngôn ngữ kí hiệu

Trong ngôn ngữ kí hiệu, với các câu có cấu trúc đơn giản như loại câu chứa động từ không có bổ ngữ (câu 1) hoặc có bổ ngữ là tính từ (câu 2) hay loại câu có tính từ làm vị ngữ (câu 17)... thì trật tự kí hiệu vẫn theo trật tự từ của câu thông thường trong tiếng Việt. Ở câu 3, người khiếm thính không coi “học” là mục đích của “đi” mà họ có xu hướng xem các cụm kiểu: “đi học”, “đi chơi”... như một đơn vị hoàn chỉnh biểu thị một ý nghĩa nên cấu trúc của câu bằng kí hiệu cũng tương tự như trường hợp

câu 1. Có thể thấy rằng nếu như mô hình cơ bản của một câu chứa động từ có bổ ngữ chỉ đối tượng trong tiếng Việt là: S - V - O (với S là chủ ngữ, V là động từ, O là bổ ngữ) thì trong ngôn ngữ kí hiệu lại là: S - O - V. Thực tế này cho thấy khi giao tiếp, người khiếm thính thường xác định đối tượng trước khi đưa ra hành động. Do ngôn ngữ kí hiệu là loại ngôn ngữ bằng hình ảnh, mang tính tượng hình hay phỏng hình nên nó có những ưu điểm và nhược điểm riêng. Ưu điểm là nó trực quan, cụ thể, dễ suy đoán nên dẫn đến một thực tế là những người khiếm thính thuộc hai cộng đồng người sử dụng hai ngôn ngữ kí hiệu khác nhau có thể giao tiếp với nhau dễ dàng hơn nhiều so với những người bình thường thuộc các dân tộc sử dụng hai ngôn ngữ nói khác nhau. Tuy nhiên, vì rất cụ thể nên đôi khi ngôn ngữ kí hiệu lại có nhược điểm là tính khái quát không cao. Chẳng hạn, trong tiếng Việt, chỉ một từ “mở”, chúng ta hoàn toàn có thể kết hợp được với các danh từ chỉ sự vật khác nhau trong tất cả các trường hợp như: Mở cửa, mở khóa, mở nắp (chai)..., thậm chí trong các tổ hợp mang ý nghĩa trừu tượng như: mở lòng, mở trái tim... nhưng trong ngôn ngữ kí hiệu thì không tồn tại hiện tượng này.

Đối với câu hỏi giả thiết người khiếm thính thường không dùng từ để hỏi chuyên biệt mà dùng nét mặt để biểu thị ý nghi vấn. Nếu như không quan sát kĩ hoặc không chú trọng đến biểu hiện trên khuôn mặt thì rất dễ bỏ qua những câu kiểu như thế này. Chẳng hạn, các câu hỏi lựa chọn kiểu như “...*đúng không/ phải không?*” luôn đi kèm với sự biểu hiện trên nét mặt là cặp chân mày nhướng lên và đôi mắt hướng về phía người được hỏi biểu lộ sự chờ đợi một sự xác nhận. Câu 14b, 14c là những thí dụ. Nếu chỉ nhìn vào kí hiệu được diễn đạt bằng cử chỉ của đôi tay thì sẽ lầm tưởng những câu này thiếu thành phần, thiếu thông tin vì câu không có kí hiệu để hỏi.

Biểu hiện bằng nét mặt còn được người khiếm thính sử dụng trong kiểu câu hỏi đóng “...*có... không?*”; “... (*đã*)...*chưa?*” (câu 15); kiểu câu hỏi lựa chọn “...*hay...?*” (16b), câu cầu khiến (câu 18), câu cảm thán (câu 17)... Quan điểm cho rằng vốn từ vựng của người khiếm thính nghèo nàn nên không có các kí hiệu kiểu như liên từ (thì, là, mà, và...), thán từ (ôi, trời ơi...), phụ từ (rất, hơi, lắm...), cũng như không có các dạng chỉ mức độ kiểu như: (cao) vơi vơi, (sâu) hun hút, xanh biếc, đỏ đỏ... trong ngôn ngữ thông thường là cũng chưa được thỏa đáng. Đúng là vốn đơn vị kí hiệu của ngôn ngữ kí hiệu không phong phú bằng vốn từ của ngôn ngữ nói, nhưng thực tế cho thấy dù không đầy đủ, kém phong phú hơn nhưng ngôn ngữ kí hiệu vẫn có tất cả các đơn vị kiểu như hư từ, tình thái từ... Chúng thường được biểu hiện thông qua nét mặt. Chẳng hạn, thán từ ôi được biểu hiện bằng đôi mắt mở to đầy ngạc nhiên và khuôn mặt vui vẻ, phấn khởi. Nếu là “cao” thì làm kí hiệu bằng tay đơn thuần, nhưng nếu là “cao vơi vơi” thì biên độ của động tác sẽ rộng hơn cộng với sự biểu cảm trên khuôn mặt biểu hiện ý hơn mức bình thường.

Đối với những câu hỏi có từ để hỏi như: ai, gì, mấy, thế nào, bao nhiêu, đâu, nào, tại sao..., khác với ngôn ngữ nói thông thường, từ để hỏi có thể đứng đầu câu



(câu 10), giữa câu (câu 11, 13), cuối câu (câu 12), thì trong ngôn ngữ kí hiệu, kí hiệu để hỏi luôn luôn đứng ở cuối câu (câu 10, 11, 12, 13).

Tương tự, khi phủ định một hành động, kí hiệu phủ định chỉ có thể được diễn đạt sau cùng.

*Chẳng hạn: Không ăn cơm là Cơm ăn không*

*Chưa uống sữa là Sữa uống chưa (câu 7)*

...

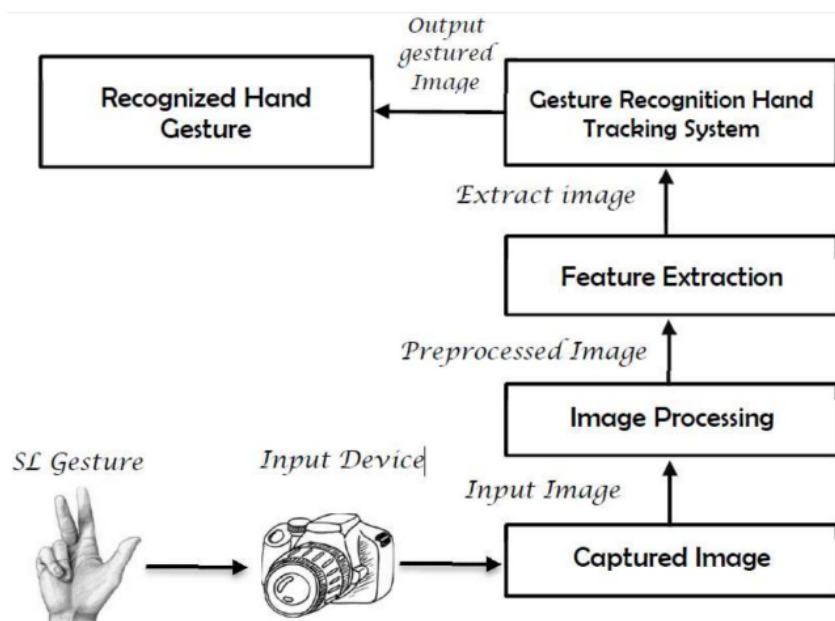
Chú ý rằng, câu 9: *Tôi không thích ăn vú sữa là Tôi ăn vú sữa không thích* được diễn đạt bằng ngôn ngữ kí hiệu không phải là một trường hợp đặc biệt, vì “không” ở đây không phải là từ phủ định của “thích”, mà cả ý “không thích” chỉ được diễn đạt bằng một kí hiệu.

## II.4. Công nghệ nhận dạng NNKH

Công nghệ nhận dạng ngôn ngữ ký hiệu có thể chia làm hai công nghệ: - công nghệ dựa trên cảm biến, công nghệ dựa trên phân tích hình ảnh (vision), và có thể lai ghép giữa hai công nghệ trên.

### II.4.1 Nhận dạng bằng phân tích hình ảnh

Hệ thống nhận dạng ngôn ngữ ký hiệu (Sign language recognition system) dựa trên chuỗi hình ảnh giúp cho nhận dạng các từ trong ngôn ngữ dấu hiệu từ hình ảnh thu từ camera hoặc từ các đoạn video đã thu được từ trước.



Hình 10: Các bước xử lý trong công nghệ nhận dạng NNKH bằng phân tích hình ảnh

Mẫu hình ảnh hoặc video các ký hiệu ngôn ngữ được Camera ghi lại, rồi đưa vào hệ thống máy tính xử lý nhiều, lọc bỏ tín hiệu thừa rồi trích xuất ra các đặc trưng của quỹ đạo chuyển động bàn tay, đặc trưng hình dáng bàn tay,... từ đó nhận dạng được từ và câu nói tương ứng.

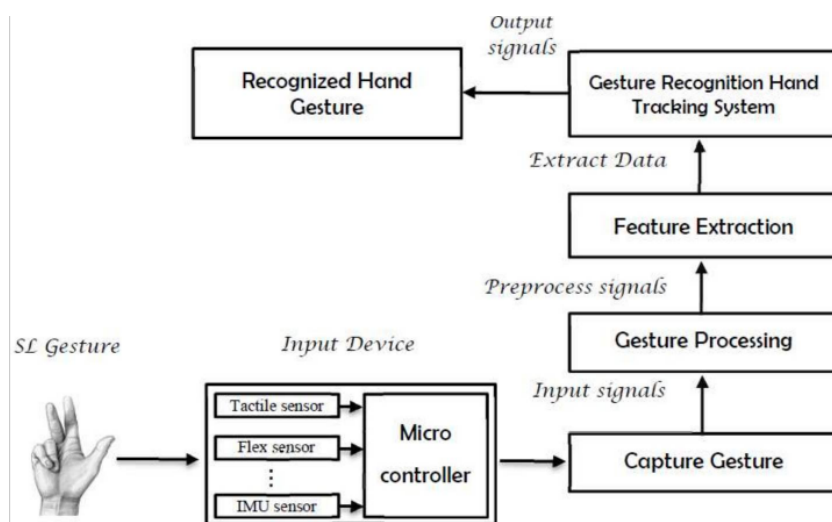
Việc nghiên cứu áp dụng công nghệ này để xây dựng các hệ thống giúp cho người khiếm thính giao tiếp với người nghe được được khá nhiều nhà khoa học trên thế giới và Việt Nam nghiên cứu, có thể kể đến một số nghiên cứu như:

Huỳnh Hữu Hưng và ctv. (2012) nhận dạng ngôn ngữ ký hiệu từ ảnh tĩnh bằng mạng nơ-ron. Theo các tác giả, kết quả là khả quan (98% độ chính xác). Tuy nhiên, các tác giả chỉ mới đề cập đến nhận dạng ảnh tĩnh chứ không phải đoạn video. Tomas Pfister et al. (2012) đã đề xuất một phương pháp tách người ra đầu và xác định các vị trí quan trọng như: đầu, vai, bàn tay, cùi chỏ dựa trên màu sắc và mô hình học máy rừng ngẫu nhiên. Một phát triển của phương pháp này được công bố trong (Charles et al., 2013). Gần đây, với sự ra đời của camera Kinect, việc trích vị trí các khớp xương trên cơ thể người có thể được thực hiện dễ dàng hơn. Hàng loạt công trình liên quan đến việc ứng dụng dữ liệu thu được từ Kinect đã được công bố như: Nhận dạng tư thế người (Lan et al. 2013), Nhận dạng cử chỉ (Wang et al., 2012; Hussein et al., 2013) và cả nhận dạng ngôn ngữ dấu hiệu. Trong (Agarwal và Thakur, 2013), các tác giả trình bày một phương pháp để nhận dạng các số (từ 0 đến 9) trong ngôn ngữ dấu hiệu.

Có thể nhận thấy rằng, các nghiên cứu trên có độ chính xác khá cao do sử dụng các công nghệ học máy, dữ liệu lớn mẫu lớn,... Tuy nhiên vẫn chưa có thiết bị nào có thể đưa ra thị trường thương mại do hạn chế của nó là để có được sự nhận dạng chính xác thì phải cần nhiều camera quay nhiều góc máy khác nhau, hệ thống máy tính để tính toán nhận dạng phải có khả năng xử lý dữ liệu lớn. Dẫn đến hệ thống chế tạo ra rất đắt tiền, và đặc biệt không có khả năng di động.

#### **II.4.2 Nhận dạng bằng cảm biến**

Ngôn ngữ ký hiệu chủ yếu được sử dụng bởi tay và các ngón tay. Nên việc sử dụng một loại găng tay cụ thể được trang bị nhiều cảm biến khác nhau, cụ thể là cảm biến uốn cong (Flex), gia tốc kế (accelerometers (ACC), cảm biến tiệm cận và cảm biến chuyển động, là một cách tiếp cận khác để thu thập dữ liệu liên quan đến cử chỉ. Những cảm biến này được sử dụng để đo góc uốn cong của ngón tay, sự chuyển động giữa các ngón tay và hướng (cuộn, cao độ và gập) của cổ tay. Khả năng nhận biết ký hiệu khi sử dụng găng tay như vậy thay đổi từ 5 đến 15 ký hiệu, tùy thuộc vào số lượng cảm biến được gắn trong găng tay. Một lợi thế lớn của các hệ thống dựa trên găng tay so với các hệ thống dựa trên phân tích hình ảnh là găng tay có thể lấy trực tiếp dữ liệu cần thiết và mức độ uốn cong, cường độ, v.v... về các giá trị điện áp, cường độ dòng điện cho thiết bị tính toán, do đó không cần lưu trữ và xử lý dữ liệu mẫu, không cần các máy tính khác đi kèm để phân tích xử lý dữ liệu như hình 11.



Hình 11 Các giai đoạn chính của việc thu thập và nhận dạng dữ liệu của hệ thống nhận dạng dựa trên Găng tay

Cảm biến Flex và cảm biến biến dạng là hai loại cảm biến đo mức độ mà vật thể bị uốn cong. Các cảm biến này tận dụng điện trở của vật liệu và cách thức thay đổi khi vật liệu này chịu tác động của lực cơ học. Điện trở bị ảnh hưởng bởi chiều dài của một vật liệu cụ thể (L), cũng như diện tích mặt cắt ngang của nó (A), độ linh động của hạt mang điện ( $\mu$ ) và mật độ hạt tải điện (N).

$$R = \frac{L}{q_e \times N \times A \times \mu}$$

Hình 12: Công thức tính điện trở của một vật

Công thức này có thể được áp dụng để sử dụng điện trở vật liệu giống như cảm biến. Bằng cách kéo dài vật liệu, chiều dài L tăng lên, dẫn đến điện trở lớn hơn, trong khi nén hoặc giảm giá trị của L làm cho điện trở trở nên nhỏ hơn. Cảm biến đo biến dạng thực hiện điều này bằng cách kéo dài hay nén dây kim loại hoặc giấy bạc từ đó chúng được chế tạo và đặc biệt hữu ích để đo chuyển động của các vật thể lớn. Ngược lại, cảm biến flex đo điện trở của một vật thể, thường nhỏ hơn nhiều so với những gì sẽ được sử dụng với cảm biến biến dạng.

### Kết chương:

*Người khiếm thính, người câm điếc là một bộ phận dân cư tồn tại khách quan trong xã hội. Họ cũng có đầy đủ các quyền sống, học tập, lao động... như những người bình thường khác. Để có thể học tập và giao tiếp, người khiếm thính cũng như tất cả chúng ta đều cần đến ngôn ngữ. Tuy nhiên, do khả năng nghe bị suy giảm hoặc mất hẳn nên khả năng giao tiếp bằng lời nói tự nhiên của người khiếm thính rất hạn*

ché. Vì vậy, để đáp ứng nhu cầu giao tiếp, người khiếm thính phải sử dụng một thứ ngôn ngữ đặc biệt: Ngôn ngữ kí hiệu - ngôn ngữ chủ yếu được cộng đồng người câm điếc sử dụng nhằm chuyển tải thông tin qua cử chỉ, điệu bộ, nét mặt thay cho lời nói.

Ngôn ngữ kí hiệu là một ngôn ngữ đặc biệt, hoàn toàn độc lập với ngôn ngữ nói. Các nhà ngôn ngữ học Mỹ cũng đã khẳng định “ngôn ngữ kí hiệu Mỹ không phải là tiếng Anh trên bàn tay”, tức là nó không phải là một sự mô phỏng của ngôn ngữ nói. Ngôn ngữ kí hiệu ở Việt Nam cũng như vậy. Nó có những quy tắc riêng về từ vựng và ngữ pháp, độc lập với tiếng Việt

Nhưng tất cả các ngôn ngữ kí hiệu trên thế giới đều có 5 phương tiện và cách thức biểu hiện giống nhau như sau:

1. Vị trí của bàn tay
2. Hình dạng bàn tay
3. Hướng của lòng bàn tay
4. Hướng của chuyển động lòng bàn tay
5. Biểu hiện của nét mặt

Như vậy, rõ ràng ngôn ngữ của người khiếm thính không chỉ giới hạn trong sự diễn tả bằng tay mà còn có cả sự biểu hiện bằng nét mặt cũng vô cùng quan trọng. Những sự biểu hiện bằng nét mặt cũng là một phần của hệ thống ngôn ngữ đặc biệt này.

Nhưng, nó cũng dẫn đến ngôn ngữ kí hiệu là ngôn ngữ khó học, nó quá nhiều từ phức tạp và na ná giống nhau, dẫn đến số người câm điếc học thông thạo Ngôn ngữ kí hiệu (NNKH) rất ít và số người nghe nói bình thường biết được ngôn ngữ kí hiệu càng ít hơn.

Việc nghiên cứu thiết bị chuyển đổi N NKH sang ngôn ngữ nói là một nhu cầu thực sự cần thiết và đã được các nhà khoa học trên thế giới và Việt Nam nghiên cứu từ lâu. Nhưng vì sự phức tạp của N NKH cũng như hạn chế về công nghệ, giá thành thiết bị dẫn đến chưa có thiết bị nào đáp ứng được mong mỏi của cộng đồng sử dụng N NKH trên thế giới cũng như tại Việt Nam.

Với mục tiêu là nghiên cứu chế tạo thiết bị chuyển đổi một số kí hiệu cử chỉ bàn tay hay N NKH sang tiếng nói có thể áp dụng được trong thực tế đối với người Việt Nam. Nhóm nghiên cứu sau các hội thảo nội bộ đã quyết định lựa chọn công nghệ nhận dạng N NKH bằng cảm biến. Công nghệ này cho phép thiết bị được chế tạo có giá thành rẻ, dễ chế tạo và dễ dàng đeo mang di động theo người sử dụng. Nhưng công nghệ này có nhược điểm là số lượng kí hiệu nhận dạng được không lớn như công nghệ nhận dạng bằng phân tích hình ảnh. Nhưng theo quan điểm của nhóm thực hiện thì một người khiếm thính từ chỗ rất khó để giao tiếp với người nghe được không biết ngôn ngữ ký hiệu, đến việc giao tiếp dễ dàng chục câu nói thông dụng thì là một thay

*đôi rất lớn. Hơn nữa hệ thống do nhóm chế tạo cho phép người sử dụng và người hỗ trợ có thể thay đổi nhanh chóng các ký hiệu cử chỉ của mình bằng các câu nói khác nhau để sử dụng trong các môi trường khác nhau như: giao thông công cộng, bệnh viện, ... sẽ phần nào hạn chế được các nhược điểm của hệ thống này.*

## CHƯƠNG III: THIẾT KẾ CHẾ TẠO GĂNG TAY THÔNG MINH

*Chương này trình bày về tập dữ liệu nhận dạng NNKH, thuật toán nhận dạng, các phân tích thiết kế, đưa ra yêu cầu người dùng, tiếp theo là phần thiết kế hệ thống sẽ trình bày về thiết kế kiến trúc, thiết kế thành phần của hệ thống Găng tay thông minh và hướng dẫn sử dụng thiết bị.*

### III.1. Tập dữ liệu nhận dạng

Dữ liệu cảm biến đầu vào cho hệ thống là một trường hợp riêng của chuỗi dữ liệu thời gian, gồm một chuỗi các điểm dữ liệu được liệt kê theo trật tự về thời gian. Khi đó, chuỗi dữ liệu cảm biến đầu vào có trật tự về thời gian. Tính chất này khiến việc phân tích chuỗi thời gian khác biệt so với các loại nghiên cứu khác sử dụng các mẫu quan sát không tuân theo trật tự thời gian. Trong một chuỗi thời gian, các mẫu quan sát gần nhau về mặt thời gian sẽ có mối liên hệ chặt chẽ hơn các mẫu xa nhau. Thêm nữa, chuỗi dữ liệu thời gian có tính chất một chiều theo trật tự thời gian, do đó các giá trị trong một khoảng thời gian nhất định sẽ được thể hiện phát sinh từ các giá trị trong quá khứ, hơn là các giá trị trong tương lai. Các tính chất này của chuỗi thời gian có ảnh hưởng đến các bước xử lý trong nhận dạng NNKH trình bày ở các phần tiếp theo

#### Bước 1: Tiền xử lý

Tiền xử lý là giai đoạn đầu tiên của một hệ thống nhận dạng NNKH điển hình. Dữ liệu đầu vào cho bước tiền xử lý (cũng chính là dữ liệu đầu vào cho hệ thống) là các luồng dữ liệu thu được từ các cảm biến trên Găng tay. Dữ liệu cảm biến đầu vào có thể là một chuỗi dữ liệu thời gian nhiều chiều do nhiều nguyên nhân, như là một số loại cảm biến có nhiều giá trị (như cảm biến gia tốc 3 chiều có giá trị tương ứng với 3 trục  $x$ ,  $y$ ,  $z$ ), hay cảm biến cong có giá trị thay một chiều theo thời gian.

Tuy vậy, tốc độ lấy mẫu của các loại cảm biến thường có thể khác nhau. Hoặc cảm biến có thể thay đổi tần số lấy mẫu tại một số tình huống, ví dụ như để tiết kiệm năng lượng hay do hệ thống yêu cầu. Hơn nữa, dữ liệu cảm biến chưa xử lý có thể bị ngắt quãng do nhiều nguyên nhân, có thể là do hoạt động của con người, lỗi cảm biến hoặc nhiễu sóng điện từ. Do đó, chức năng của tiền xử lý là đồng bộ, giảm nhiễu và loại bỏ các tín hiệu lạ, chuẩn bị các tín hiệu đã lấy được cho bước trích chọn đặc trưng. Tuy nhiên, tiền xử lý cần giữ lại được các đặc điểm tín hiệu có chứa thông tin quan trọng về các hoạt động đang quan tâm.

#### Bước 2: Phân đoạn

Bước phân đoạn (còn gọi là phát hiện hoạt động, hay là đánh dấu hoạt động) lấy dữ liệu được xử lý (từ luồng dữ liệu sau khi tiền xử lý), sau đó chia thành các phần nhỏ hơn có khả năng chứa thông tin về hoạt động hay cử chỉ. Các phần dữ liệu này được gọi là các đoạn (segment). Phân đoạn một luồng cảm biến liên tục là một việc khó. Có

3 phương pháp khác nhau để thực hiện phân đoạn, bao gồm phân đoạn sử dụng cửa sổ trượt, phân đoạn dựa vào năng lượng và phân đoạn dựa vào các vị trí nghỉ (bằng cách sử dụng một loại cảm biến để phân đoạn cho dữ liệu của một loại cảm biến khác hay sử dụng các thông tin ngữ cảnh bên ngoài). Trong nhận dạng NNKH sử dụng cảm biến trên Găng tay, phân đoạn sử dụng vị trí nghỉ thường được sử dụng nhiều nhất.

### **Bước 3: Trích chọn đặc trưng**

Trích chọn đặc trưng bao gồm hai phần: tách/trích xuất đặc trưng (feature extraction) và lựa chọn đặc trưng (feature selection). Trích chọn đặc trưng nhằm rút gọn các tín hiệu thành các đặc trưng để phân biệt các hoạt động đang có và sau đó được sử dụng làm dữ liệu đầu vào cho bước phân lớp. Tùy thuộc vào từng hệ thống cụ thể mà lựa chọn đặc trưng có thể được thực hiện hoặc không.

Các đặc trưng được trích xuất tự động. Tập các đặc trưng có được từ dữ liệu được gọi là không gian đặc trưng. Nói chung, khi các hoạt động được phân tách càng rõ ràng trong không gian đặc trưng thì hiệu suất nhận dạng của hệ thống càng cao. Lý tưởng nhất là các đặc trưng của cùng hoạt động có thể nhóm thành một nhóm trong không gian đặc trưng và ngược lại các đặc trưng của các hoạt động khác nhau cần phân biệt càng xa càng tốt.

Không gian đặc trưng có số chiều càng lớn thì cần càng nhiều dữ liệu huấn luyện để tính toán các tham số cho mô hình cũng như cần càng nhiều lượng tính toán khi thực hiện phân lớp. Đặc biệt là trong các hệ thống nhúng cần xử lý theo thời gian thực, mục tiêu cần đạt là tối thiểu hóa yêu cầu về bộ nhớ, lượng tính toán và băng thông. Vì vậy, điều quan trọng là cần sử dụng số lượng ít nhất các đặc trưng trong khi vẫn đảm bảo hiệu năng cho hệ thống nhận dạng. Đối với hệ thống này với các đặc trưng không lớn, việc lựa chọn các đặc trưng này được thực hiện thủ công trong khi thử nghiệm cảm biến trên mô hình găng tay.

### **Bước 4: Huấn luyện**

Ở chế độ huấn luyện, các đặc trưng được tách ra và cùng với các nhãn tương ứng, các đặc trưng sẽ được sử dụng làm dữ liệu đầu vào để huấn luyện mô hình hoạt động. Ở chế độ phân lớp, các đặc trưng và mô hình đã huấn luyện trước đó được sử dụng để tính ra giá trị của mỗi lớp hoạt động và ánh xạ các giá trị này với một nhãn riêng trong lúc phân lớp. Khi sử dụng nhiều cảm biến hay nhiều bộ phân lớp, đầu ra của các bộ phân lớp này được kết hợp với nhau để đưa ra một quyết định duy nhất. Việc chọn lựa một phương pháp suy diễn tùy thuộc vào sự cân bằng giữa độ phức tạp tính toán và hiệu năng nhận dạng. Khi phân lớp trong hệ thống nhúng có tài nguyên hạn chế, cần tối thiểu hóa độ phức tạp tính toán và bộ nhớ sao cho vẫn đảm bảo hệ thống có hiệu năng cao.

### III.2. Thuật toán phân tích lớp và nhận dạng

Có nhiều thuật toán có thể được sử dụng để nhận dạng cử chỉ, thuật toán học máy (Machine learning algorithms) đã được sử dụng trong các nghiên cứu thiết bị nhận dạng cử chỉ ở các nghiên cứu trước (xem chương 1). Các thuật toán này có thể học hỏi hoặc đưa ra dự đoán dựa trên dữ liệu mẫu có trước và có thể thuộc loại học máy có giám sát hoặc không giám sát. Phần lớn các ứng dụng học máy thực tế đều sử dụng học có giám sát (supervised learning).

Học có giám sát là thuật toán dự đoán đầu ra (outcome) của một dữ liệu mới (new input) dựa trên các cặp (*input*, *outcome*) đã biết từ trước. Cặp dữ liệu này còn được gọi là dữ liệu, nhãn (data, label). Học có giám sát là nhóm phổ biến nhất trong các thuật toán học máy (Machine Learning).

Trong toán học, Supervised learning là khi chúng ta có một tập hợp biến đầu vào  $X = \{x_1, x_2, \dots, x_N\}$  và một tập hợp nhãn tương ứng  $Y = \{y_1, y_2, \dots, y_N\}$ , trong đó  $x_i, y_i$  là các vector. Các cặp dữ liệu biết trước  $(x_i, y_i) \in X \times Y$  được gọi là tập dữ liệu huấn luyện (training data). Từ tập dữ liệu huấn luyện này, người ta cần tạo ra một hàm số ánh xạ mỗi phần tử từ tập  $X$  sang một phần tử (xấp xỉ) tương ứng của tập  $Y$ :

$$Y_i \approx F(x_i), \forall i=1,2,\dots,N$$

Mục đích là xấp xỉ hàm số  $F$  thật tốt để khi có một dữ liệu  $x$  mới, chúng ta có thể tính được nhãn tương ứng của nó  $y=F(x)$ .

Trong mô hình phân lớp, thuật toán phân lớp giữ vai trò trung tâm, quyết định thành công của mô hình phân lớp. Chìa khóa của vấn đề phân lớp dữ liệu là tìm ra được thuật toán phân lớp nhanh, hiệu quả có độ chính xác cao và có khả năng mở rộng được. Có thể liệt kê các kỹ thuật phân lớp dữ liệu đã được sử dụng trong những năm qua:

- Phân lớp cây quyết định (Decision tree classification)
- Bộ phân lớp Bayesian (Bayesian classifier)
- Mô hình phân lớp K- hàng xóm gần nhất (K-nearest neighbor classifier)
- Mạng nơ-ron
- Phân tích thống kê
- Các thuật toán di chuyển
- Phương pháp tập thô (Rough set Approach)

Các cảm biến Flex được sử dụng trong Găng tay thông minh cho dữ liệu ra liên tục thay đổi, và phụ thuộc rất lớn vào độ dài ngón tay, thói quen thực hiện cử chỉ của người sử dụng. Thậm chí cùng một người sử dụng thì với mỗi cử chỉ ký hiệu người đó cũng không thể lần nào cũng chính xác 100% tốc độ của bàn tay, cử chỉ, độ cong của ngón tay thực hiện ký hiệu với cử chỉ mẫu mà thiết bị đã học được. Do vậy thuật toán



xấp xỉ “K hàng xóm gần nhất” (K-Nearest Neighbors) được sử dụng trong nghiên cứu này, để nhận dạng được cử chỉ gần giống nhất với cử chỉ mẫu mà thiết bị đã học được.

Thuật toán “K hàng xóm gần nhất” (K-Nearest Neighbors) là thuật toán khá chính xác. Sau khi đào tạo, để quyết định lớp nào thuộc nhãn mới, thuật toán chọn lớp gần nhãn đã biết. Về mặt phân loại cử chỉ, điều này có nghĩa là thuật toán sẽ tìm ra nhãn cử chỉ đã biết nào gần nhãn mới nhất.

Thuật toán (K-Nearest Neighbors) được mô tả như sau:

- Xác định giá trị tham số K (số láng giềng gần nhất)
- Tính khoảng cách giữa đối tượng cần phân lớp (Query Point) với tất cả các đối tượng trong đã có trong dữ liệu mẫu (training data).
- Sắp xếp khoảng cách theo thứ tự tăng dần và xác định K láng giềng gần nhất với Query Point
- Lấy tất cả các lớp của K láng giềng gần nhất đã xác định
- Dựa vào phần lớn lớp của láng giềng gần nhất để xác định lớp cho Query Point

### **III.3. Thiết kế hệ thống, chế tạo Găng tay**

Yêu cầu chung về thiết kế găng tay bị ảnh hưởng bởi các yêu cầu liên quan đến người dùng NNKH, kỹ thuật nhận dạng NNKH và nhược điểm của các dự án tương tự như đã trình bày trong chương trước. Những điều đó đã giúp nhóm nghiên cứu xây dựng ba phương pháp thiết kế khả thi hệ thống này. Sau các hội thảo nội bộ phân tích ưu nhược điểm của các phương pháp thiết kế chúng tôi chọn một thiết kế tối ưu.

#### **III.3.1 Phân tích yêu cầu người dùng và chỉ tiêu kỹ thuật**

##### **III.3.1.1 Yêu cầu người dùng**

Hiểu được các khó khăn của người dùng NNKH ở Việt Nam phải đối mặt khi giao tiếp với những người có khả năng nghe nói bình thường cũng như những hạn chế kỹ thuật của các công nghệ hỗ trợ. Nhóm nghiên cứu đưa ra một số yêu cầu kỹ thuật cho thiết bị như sau:

- **Yêu cầu cụ thể**

Dịch chính xác các Ký hiệu cử chỉ bàn tay của người sử dụng đưa ra thành tiếng nói phát ra điện thoại di động có cài phần mềm thu nhận tín hiệu phát ra từ Găng tay trong khoảng thời gian dưới 10 giây. Các Ký hiệu này được người sử dụng kết hợp với người hỗ trợ: - Người nghe nói được bình thường và có khả năng hướng dẫn người người khiếm thính sử dụng thiết bị, định nghĩa trước các cử chỉ bàn tay và các lời nói tương ứng. Dễ dàng thay thế ( định nghĩa lại) cử chỉ với các lời nói khác nhau.

- **Yêu cầu ngầm định**

Nhóm đã xác định năm yêu cầu của thiết bị làm tăng tính hữu dụng của nguyên mẫu thiết kế Găng tay đối với người dùng NNKH. Một số khía cạnh và tương tác của thiết bị với người dùng và môi trường đã được chú ý cụ thể như sau:

**Dễ sử dụng** – Bất kỳ sự phức tạp nào trong việc sử dụng thiết bị sẽ dẫn đến người sử dụng không sử dụng thiết bị nữa mà tìm cách khác để thay thế. Do việc “Nghe hiểu”, tiếp thu được hướng dẫn sử dụng hạn chế. Nếu có chút phức tạp trong sử dụng sẽ dẫn đến không sử dụng được. Mỗi bản dịch nên được thực hiện mà không cần nhấn nút hay thao tác gì khác trên thiết bị, mỗi một cử chỉ chỉ dịch một lần, không lặp lại liên tiếp gây khó chịu cho người sử dụng cũng như người xung quanh.

**Di động** - Hệ thống không nên phụ thuộc vào máy tính hoặc hệ thống đính kèm khác. Nó có thể được mang đến gần như bất cứ nơi nào người dùng đi, trừ môi trường dưới nước.

**Giá cả phải chăng** – Đa số người khuyết tật đều thuộc diện nghèo. Do vậy, thiết bị này nên được thiết kế để có thể chế tạo được với giá rẻ để người sử dụng có thể tiếp cận, sử dụng được.

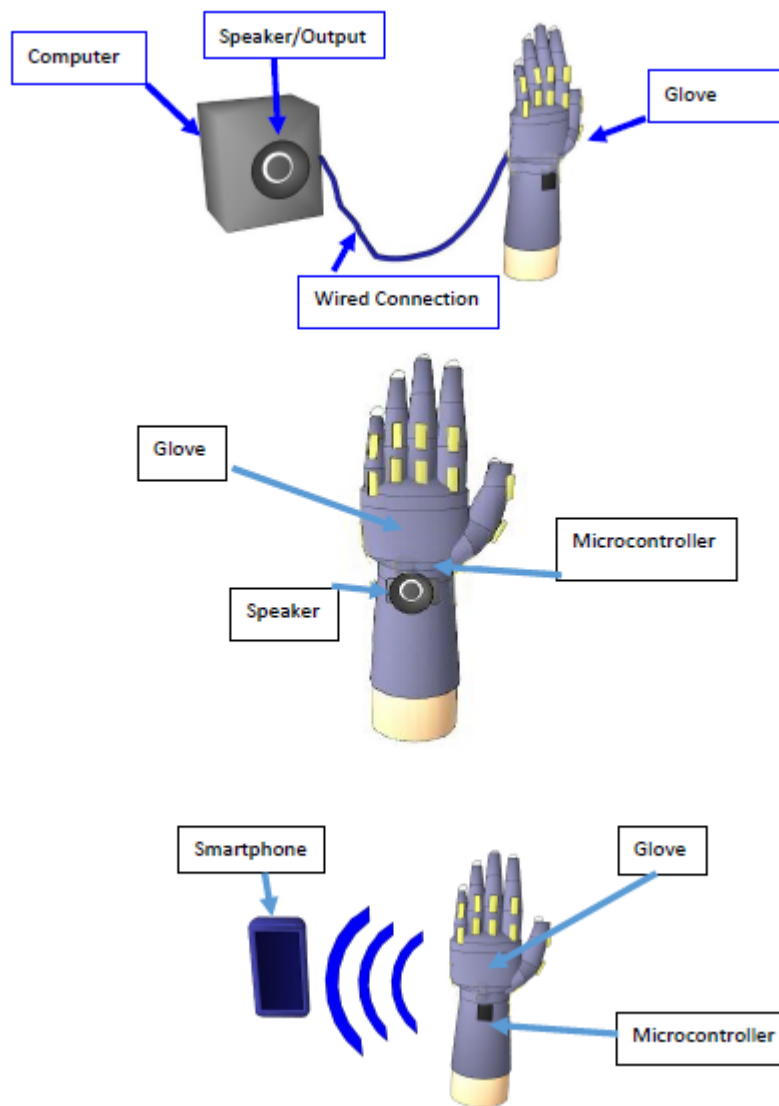
**Độ tin cậy** – Bất cứ thiết bị điện tử nào độ tin cậy cũng phải cao, nhưng điều này có mâu thuẫn với hiệu năng, giá thành, thời gian sống của sản phẩm,.. Đối với hệ thống này, nhóm nghiên cứu tập chung vào độ chính xác của nhận dạng cử chỉ, hướng đến độ chính xác từ 99 đến 100% cử chỉ mà không hướng đến số lượng cử chỉ được nhận diện.

**Về mặt thẩm mỹ, an toàn** - Đối với mục đích thiết kế có thể chuyển giao cho nhà sản xuất để sản xuất thiết bị cho người sử dụng, thiết bị nên có tính thẩm mỹ và dễ dàng đeo, không gây khó chịu khi sử dụng trong thời gian dài. Điều này bao gồm vẻ ngoài mượt mà, chuyên nghiệp mà không có bất kỳ thành phần nào gây kích ứng da tay, hay có thể cắt vào tay hoặc gây khó chịu cho người dùng. Thiết kế phải lựa chọn nguồn điện an toàn, không thể gây cháy nổ, ...

### **III.3.1.2 Lựa chọn thiết kế**

Để thực hiện nhiệm vụ thiết kế này, nhóm nghiên cứu đã đưa ra 03 mô hình thiết kế như hình 13. Cả ba mô hình thiết kế này có chung cảm biến và phản ứng nhưng thành phần xử lý và đầu ra có khác nhau. Ở mô hình thiết kế 1 (hình 13 phía trên), thiết kế bao gồm một găng tay dùng để thu nhận tín hiệu cảm biến, kết nối có dây với một máy tính để phân tích xử lý các tín hiệu nhận được và phát ra âm thanh qua loa của chính máy tính hay loa ngoài. Ở mô hình thứ 2 (hình 13 giữa) tương tự như ở thiết kế 1 nhưng bỏ máy tính xử lý tín hiệu, đưa phần xử lý tín hiệu vào vi điều

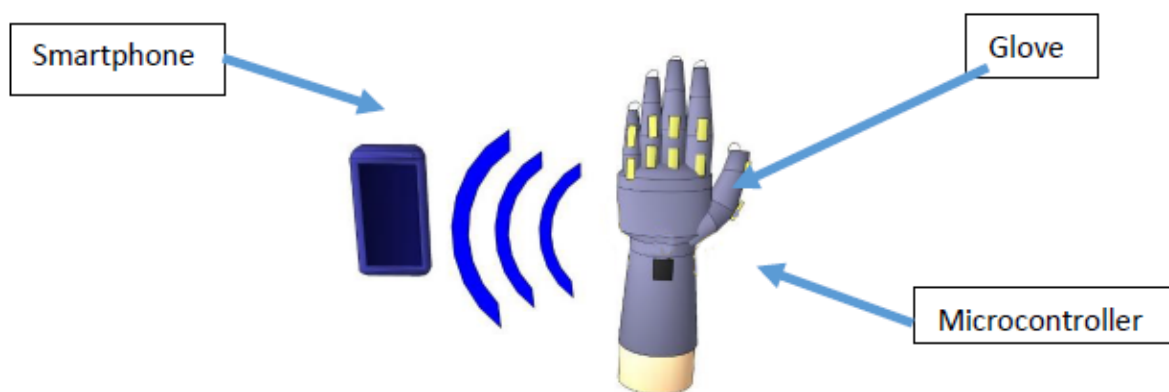
khiển trên trên Găng tay, đầu ra âm thanh cũng được kết nối trực tiếp vào găng tay.



Hình 13: Mô hình thiết kế 1(trên); Mô hình thiết kế 2 (giữa); Mô hình thiết kế 3 (dưới)

Ở mô hình thiết kế thứ 3 (hình 13 phía dưới), phần thu nhận và xử lý tín hiệu đặt ở Găng tay, phần đưa tín hiệu âm được phát ra từ điện thoại thông minh của người sử dụng. Điện thoại kết nối không dây với găng tay qua giao thức truyền thông Bluetooth.

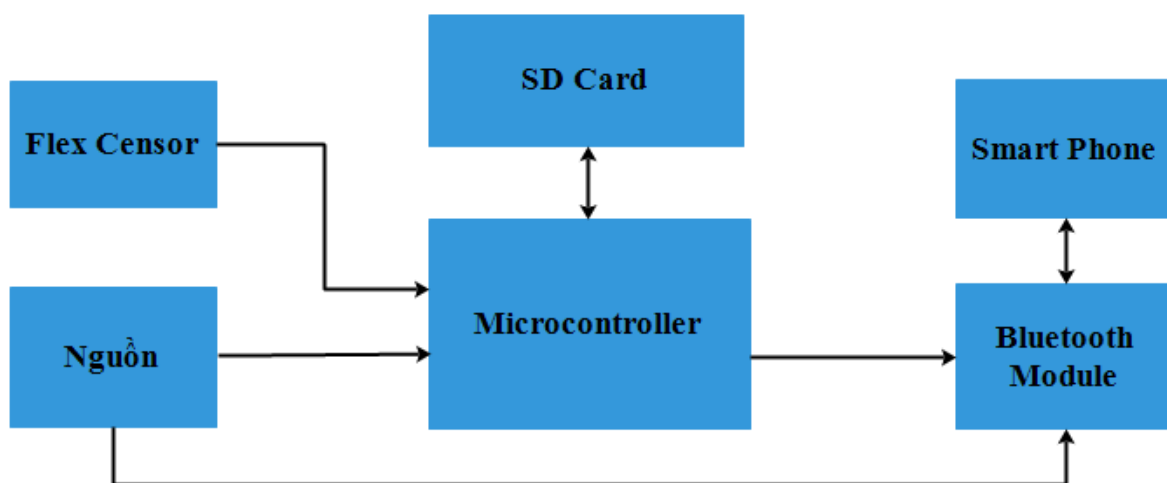
Sau khi cân nhắc nhiều ý kiến đóng góp từ các thành viên trong các hội thảo nội bộ nhóm nghiên cứu quyết định chọn mô hình thiết kế thứ 3 để thực hiện nhiệm vụ này. Do thiết kế này đáp ứng tốt nhất các yêu cầu về kỹ thuật và yêu cầu sử dụng đã đưa ra.



Hình 14: Mô hình thiết kế Găng tay thông minh

Thiết kế này bao gồm một số hệ thống con như: hệ thống cảm biến độ cong đặt tại các ngón tay để phát hiện cử chỉ của tay, hệ thống xử lý, hệ thống giao tiếp Bluetooth, ...

Thiết kế này cho phép người sử dụng đeo thoải mái thiết bị trên tay, giống như một chiếc găng tay bình thường và hướng tới việc tuân thủ các thông số kỹ thuật được xác định trước đó như tính di động, trọng lượng thấp, độ chính xác và có khả năng sản xuất số lượng lớn. Thiết kế ở mức sơ đồ khối của thiết bị như hình 15.

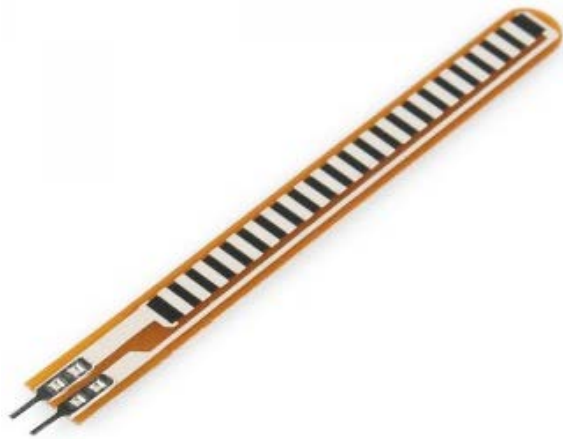


Hình 15: Thiết kế khối hệ thống

### III.3.2 Hệ thống cảm biến

Để cho hệ thống đơn giản, giảm giá thành sản xuất, nhóm nghiên cứu chỉ sử dụng 05 cảm biến độ cong, mỗi cảm biến có độ dài 13cm, được gắn dọc theo 5 ngón tay. Mỗi khi ngón tay co, duỗi thì độ cong trên cảm biến thay đổi dẫn đến điện trở trên cảm biến thay đổi giúp cho vi điều khiển (Microcontroller) nhận diện được cử chỉ. Cảm biến độ cong (Flex sensor) như hình 16. Mà không sử dụng thêm các loại cảm

biến khác như cảm biến tiệm cận, cảm biến con quay hồi chuyển đảm bảo sự hoạt động ổn định, giảm giá thành.



Hình 16: Cảm biến độ cong

Mỗi cảm biến sẽ được thiết lập bởi mạch phân chia điện áp: - chuyển đổi tín hiệu tương tự sang dạng số (ADC) giúp cho bộ vi xử lý nhận dạng được chính xác trạng thái của nó.

### III.3.3 Hệ thống vi điều khiển

Để xử lý tín hiệu đến, mỗi tín hiệu tương tự từ các cảm biến độ cong phải được chuyển đổi thành tín hiệu số. Điều này có nghĩa là để xử lý tất cả 5 cảm biến Flex, cần có 5 kênh ADC. Hầu hết các bộ vi điều khiển đều có một số chân đầu vào tương tự. Trong thiết kế này nhóm nghiên cứu lựa chọn vi điều khiển Atmega2560 có số lượng chân đầu vào có thể phục vụ cho 16 kênh đầu vào tương tự với độ phân giải 10 bit, tốc độ lấy mẫu là  $104\mu s$  vào khoảng 9600 mẫu mỗi giây. Điều này có nghĩa các mẫu điện áp của cảm biến Flex có độ phân giải  $5V / 2^{10} = 4.9 \text{ mV}$  mỗi bit, được lấy mẫu vài lần mỗi giây quá đủ cho nhận dạng cử chỉ.

Hệ thống vi xử lý ngoài việc đọc tính hiệu tương tự và chuyển thành tín hiệu số, thì việc xử lý tính toán dữ liệu để nhận dạng là khá lớn. Theo kinh nghiệm của nhóm nghiên cứu thì vi xử lý Atmega2560 của hãng Atmel nay là Microchip đủ khả năng thực hiện tính toán và xử lý bài toán này. Thực tế sau khi thử nghiệm nhận dạng cử chỉ bàn tay, vi điều khiển này đã đảm bảo tốt yêu cầu đặt ra. Vi điều khiển Atmega 2560 có hình ảnh như hình 17.



Hình 17: Vi điều khiển Atmega2560

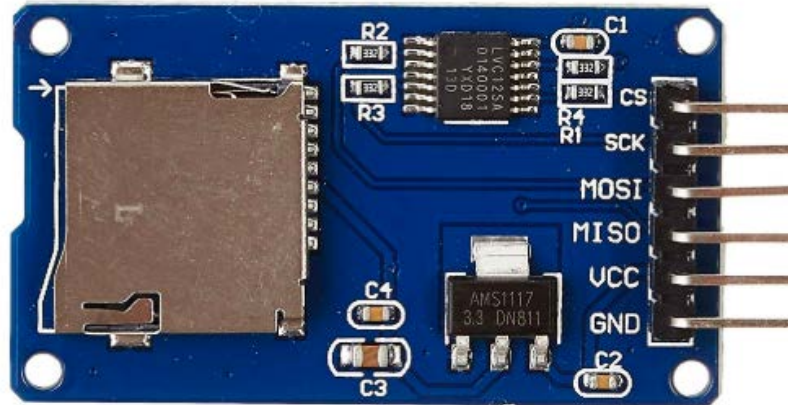
Một số thông số kỹ thuật của Vi điều khiển Atmega2560:

- Vi điều khiển hiệu suất cao, công suất tiêu thụ thấp
- Cấu trúc RISC nâng cao
- Hỗ trợ đến 135 lệnh, phần lớn được thực hiện với 1 chu kỳ tín hiệu đồng hồ (Single Clock Cycle)
- 32 thanh ghi 8 bit ( $32 \times 8$  General Purpose Working Registers)
- Khả năng thực hiện đến 16 triệu lệnh mỗi giây, khi tần số hệ thống là 16 MHz.
- Dung lượng Flash: 256KBytes of In-System Self-Programmable Flash.
- Dung lượng EEPROM: 4Kbytes, cho phép ghi/xóa đến 100.000 lần
- Dung lượng SRAM: 8Kbytes
- Số kênh ADC: 16 kênh 10 bit
- Tần số hoạt động 0 - 16MHz
- Giao tiếp ngoại vi: nối tiếp, I2C, SPI,...
- Nguồn điện 4.5V - 5.5V DC

Để nhận dạng được ký hiệu, hệ thống cần có một tập dữ liệu nhận dạng mẫu trước. Các mẫu này có thể ghi vào bộ nhớ Flash của vi điều khiển hoặc bộ nhớ ngoài. Nếu như mẫu dữ liệu được ghi vào Flash thì rất khó thay đổi, cụ thể là muốn thay đổi thì phải nạp lại chương trình, sẽ không phù hợp với yêu cầu của bài toán đặt ra. Do vậy nhóm nghiên cứu sử dụng module thẻ nhớ SD để lưu dữ liệu cử chỉ mẫu.

### III.3.4 Module thẻ nhớ SD

Nhóm nghiên cứu sử dụng Module thẻ nhớ giá rẻ trên thị trường (hình 18). Module này sử dụng thẻ MicroSD, hỗ trợ dung lượng thẻ đến 16 GB. Giao tiếp với vi điều khiển theo chuẩn giao tiếp SPI.



Hình 18: Module giao tiếp thẻ nhớ SD

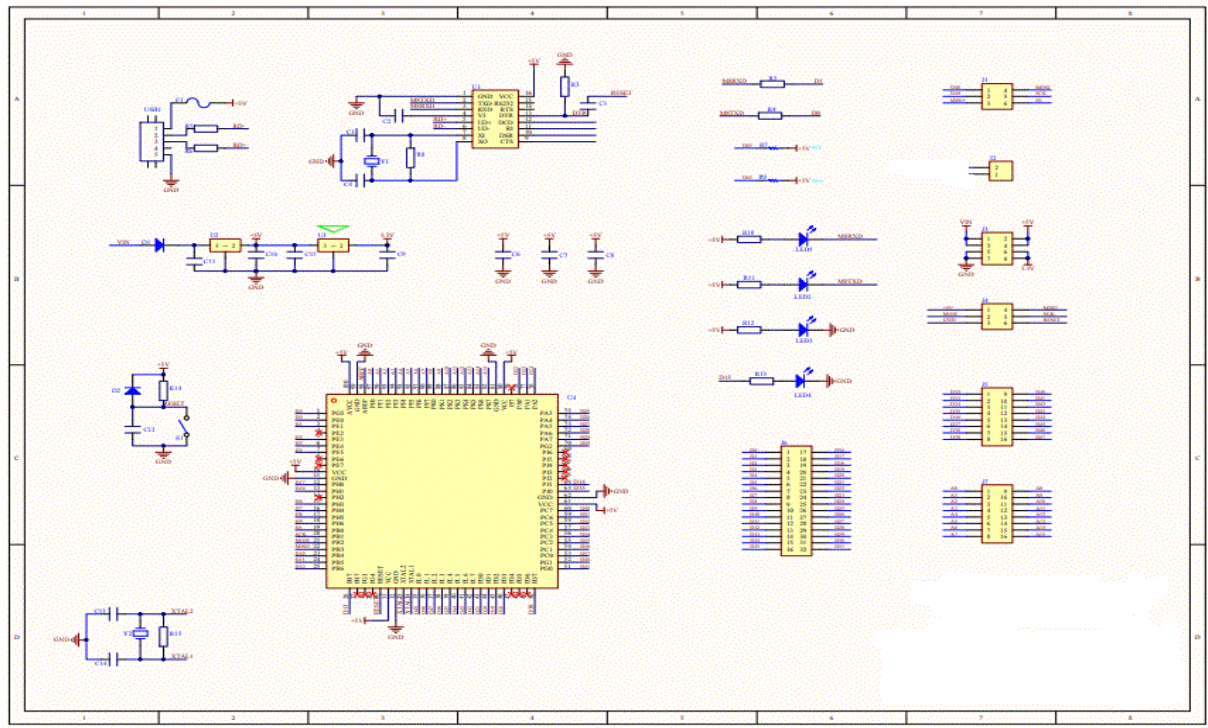
### III.3.5 Hệ thống đầu ra

Sau khi cử chỉ được nhận dạng, tín hiệu đầu ra sẽ được truyền đến điện thoại của người sử dụng thông qua Bluetooth. Sau đó tên cử chỉ sẽ được phát ra âm thanh trên điện thoại giúp cho người giao tiếp hiểu được điều người sử dụng muốn nói dễ dàng. Hệ thống đầu ra này bao gồm một ứng dụng trên điện thoại sử dụng hệ điều hành Android giao tiếp Bluetooth với hệ thống phần cứng để nhận dạng tính hiệu cử chỉ và phát ra âm thanh tương ứng đo nhóm nghiên cứu tự xây dựng. Ứng dụng trên điện thoại này ngoài việc thu nhận tín hiệu cử chỉ và phát ra âm thanh còn có chức năng ghi âm để gán với các cử chỉ khác (tạo dữ liệu huấn luyện mới) cho phép người dùng và người hỗ trợ tự định nghĩa các cử chỉ tương ứng với các từ/câu mới phù hợp với điều kiện sử dụng của từng người sử dụng.

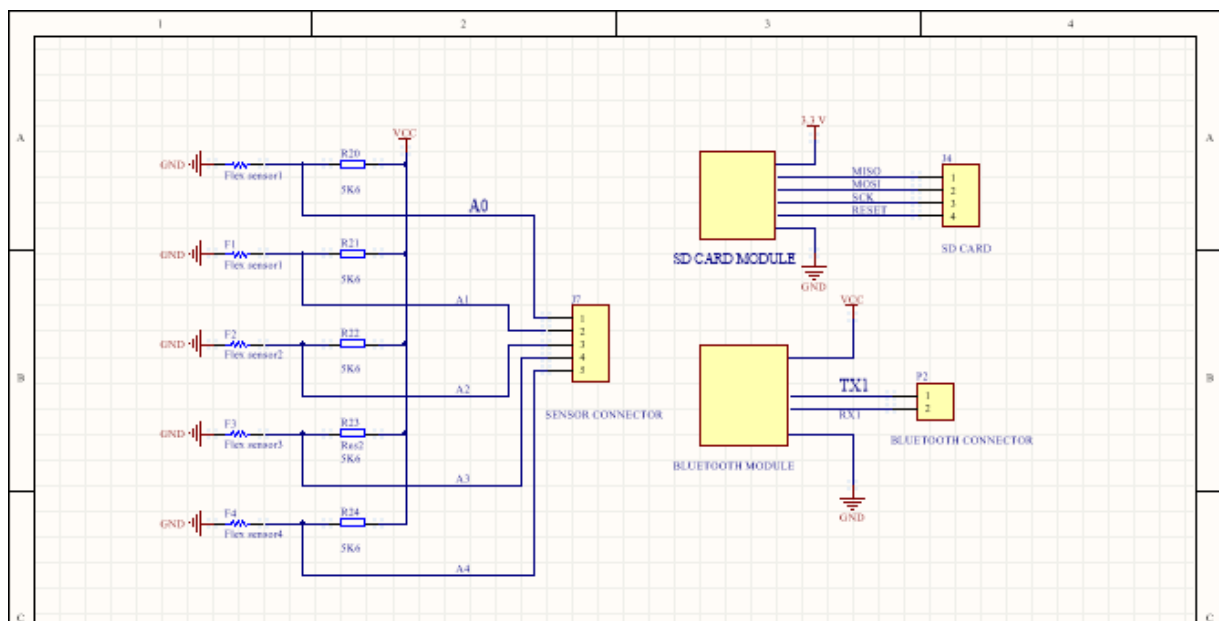
### III.3.6 Thiết kế phần cứng

Đối với phần cứng của Găng tay này có thiết kế đơn giản gồm khối xử lý chính (vi điều khiển) có các ic chức năng hỗ trợ giao tiếp giữa vi điều khiển và máy tính phục vụ cho việc nạp và debug chương trình, cùng IC nguồn chuyển đổi từ nguồn điện 5V DC thành nguồn 3.3 V DC cung cấp nguồn cho module Bluetooth và module ghi đọc thẻ nhớ micro SD như hình 18 và khối giao tiếp cảm biến, giao tiếp Bluetooth, giao tiếp SD card (hình 19)





Hình 19: Khối vi điều khiển



Hình 20: Giao tiếp cảm biến, sd card và bluetooth

### III.3.7 Phần mềm hệ thống;

Sau khi tất cả các đầu ra cảm biến được kết nối với đầu vào vi điều khiển, quá trình xử lý có thể bắt đầu đọc các tín hiệu cử chỉ, lọc, so sánh, phân tích, lựa chọn các mẫu tín hiệu chính xác của cử chỉ. Vi điều khiển liên tục đọc các tín hiệu tương tự với tốc độ hàng nghìn tín hiệu mẫu trong mỗi giây và trả về kết quả ngay tức thì. Do vậy,



một trong những khó khăn của bài toán nhận dạng cử chỉ này là phân biệt được đâu là tín hiệu của giai đoạn đang thực hiện cử chỉ và đâu là tín hiệu cử chỉ thực sự được mong chờ để nhận dạng. Khi có tín hiệu nhận dạng rồi, phải thực hiện các thuật toán học máy để thực hiện nhận dạng tín hiệu cử chỉ là tín hiệu nào để truyền sang cho điện thoại để phát ra âm thanh. Tất cả chúng được thực hiện bởi phần mềm hệ thống nhúng được nạp vào vi điều khiển trước khi sử dụng. Do vậy, phần chính và khó khăn nhất đối với hệ thống Găng tay thông minh này nằm ở hệ thống phần mềm nhúng. Phần mềm hệ nhúng được hệ thống được nhóm nghiên cứu xây dựng bằng công cụ phần mềm Arduino (Arduino Software – IDE), đây là công cụ lập trình miễn phí cho vi điều khiển. Cho phép lập trình, biên dịch và nạp chương trình đã biên dịch vào vi điều khiển thông qua giao tiếp USB.

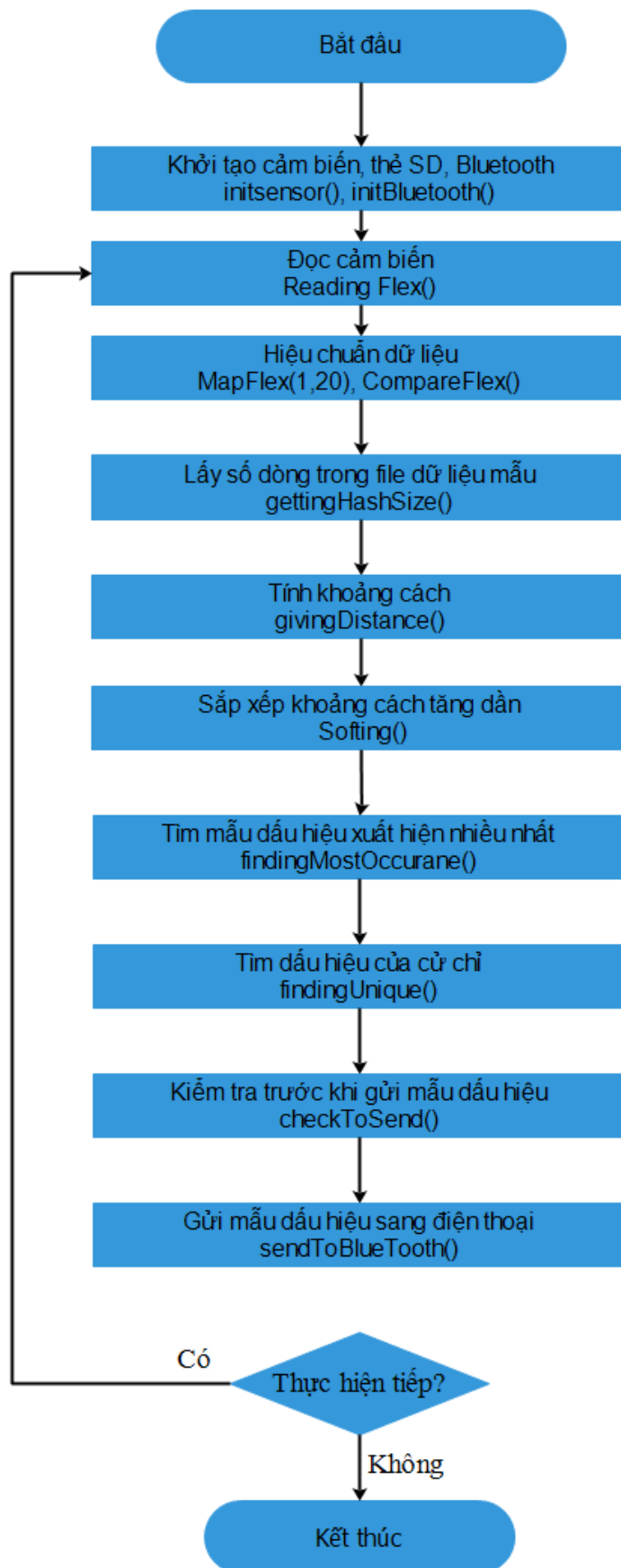
Chức năng chính của phần mềm hệ thống nhúng này là:

- Đọc cảm tất cả các cảm biến, (lấy dữ liệu thô)
- Lấy giá trị dữ liệu lớn nhất và nhỏ nhất của mỗi cảm biến
- Tính toán, chọn lọc dữ liệu hữu ích (calibrating) của mỗi cảm biến
- Đọc/ghi tệp trong thẻ SD.
- Thực hiện thuật toán học máy KNN để tìm ra dữ liệu của dấu hiệu cử chỉ bàn tay. Thuật toán này lấy hai nguồn dữ liệu vào là dữ liệu mẫu đã ghi trong thẻ MicroSD và dữ liệu của cử chỉ hiện tại thu nhận được để tìm ra mẫu của ký hiệu cử chỉ.
- Gửi dấu hiệu cử (đại diện bởi một ký tự đã được ghi trong mẫu) gửi sang điện thoại để phát ra âm thanh tương ứng.

Sau đây, trình bày một số lưu đồ thuật toán thực hiện và mô tả cụ thể hơn một số phần của phần mềm hệ thống nhúng.

### **III.3.7.1 Lưu đồ thực hiện nhận biết cử chỉ**

Chuyển đổi được cử chỉ sang tiếng nói, bao gồm hai nội dung công việc là nhận biết/xác nhận cử chỉ và chuyển đổi cử chỉ đó sang tiếng nói. Lưu đồ thực hiện việc nhận biết/ xác nhận cử chỉ được thực hiện theo lưu đồ hình 21. Đây có thể nói là phần công việc khó khăn nhất của đề tài này.



Hình 21: Lưu đồ thực hiện phần mềm hệ thống

## **Mô tả:**

Lưu đồ này có thể được chia làm 3 giai đoạn:

**Giai đoạn 1:** - Khởi tạo các cảm biến, thẻ nhớ và Bluetooth và hiệu chuẩn dữ liệu. Đây là giai đoạn khó khăn nhất cần phải giải quyết trong đề tài này. Khó khăn thứ nhất mà nhóm nghiên cứu gặp phải là cảm biến độ cong có giá trị đầu ra chuyển đổi ADC không ổn định. Khi thử nghiệm Găng tay để đọc dữ liệu cảm biến, những người khác nhau cho kết quả khác nhau do độ dài ngón tay khác nhau và độ cong khi nắm tay và xoè bàn tay. Thậm chí, cùng một người thực hiện cùng một cử chỉ ở những lần đo khác nhau cũng ra kết quả khác nhau. Nên việc dùng dữ liệu này để lấy mẫu cử chỉ cũng như để tính toán xác nhận ra cử chỉ thì chắc chắn thiết bị sẽ không thể hoạt động được. Nhóm nghiên cứu đã phải có bước hiệu chuẩn dữ liệu bằng cách xác định giá trị lớn nhất (Max) và giá trị nhỏ nhất (Min) của mỗi cảm biến bằng một chương trình con khác. Chương trình này đọc liên tục giá trị của cả 5 cảm biến, trong khi người thử nghiệm thực hiện liên tục việc thay đổi cử chỉ nắm và xoè bàn tay hết cỡ trong khoảng 5 phút kết thúc chương trình sẽ in ra giá trị lớn nhất và nhỏ nhất của mỗi cảm biến có thể có được. Khoảng giá trị của cảm biến này sẽ được chia làm 20 khoảng giá trị từ 1 đến 20, 1 ứng với giá trị nhỏ nhất và 20 ứng với giá trị lớn nhất. Thực hiện chuyển đổi tuyến tính các giá trị dữ liệu đọc từ cảm biến với các giá trị Max và Min của nó ta sẽ được dữ liệu tương đối ổn định của mỗi cảm biến có giá trị từ 1 đến 20. Hàm MapFlex(1,20) thực hiện công việc này.

Khó khăn thứ hai cần phải giải quyết trong giai đoạn này là xác nhận giai đoạn nào là người sử dụng đang thực hiện cử chỉ và giai đoạn nào là cử chỉ thật cần phải xác nhận để chuyển đổi cử chỉ. Quan sát khi thực hiện cử chỉ thì giai đoạn đang thực hiện cử chỉ thì độ cong của ngón tay luôn thay đổi và giai đoạn thực hiện xong cử chỉ thì độ cong của ngón tay không thay đổi. Do vậy, nếu coi giá trị đọc được từ 5 cảm biến là một vector có 5 giá trị mỗi giá trị nằm trong khoảng từ 1 đến 20 thì ở giai đoạn đang thực hiện cử chỉ giá trị của vector đọc được ở lần 1 và giá trị của vector đọc được ở lần 2 sẽ khác nhau, còn giai đoạn kết thúc cử chỉ thì giá trị của vector đọc được ở lần 1 sẽ bằng giá trị vector đọc được ở lần 2. Từ đó nhóm nghiên cứu đưa ra giải pháp đọc liên tục các cảm biến, cho đến khi nào giá trị của vector đọc lần trước bằng giá trị của vector đọc liên tiếp hai lần sau đó thì xác nhận đó là dữ liệu cần xử lý. Giải pháp này tỏ ra hiệu quả khi thử nghiệm cử chỉ với những người thực hiện cử chỉ một cách nhanh và dứt khoát, hệ thống cho ra kết quả nhận dạng nhanh (thời gian nhận ra cử chỉ từ dưới 1 giây đến 3 giây). Tuy nhiên khi thử nghiệm với động tác tay không dứt khoát hay run tay thì hệ thống nhận dạng chậm đôi khi không nhận dạng được. Giải pháp của nhóm đưa ra là thay vì so sánh giá trị của vector đọc được ở lần trước bằng với giá trị của vector đọc được ở lần sau bằng so sánh khoảng cách Euclidean của vector đọc được ở lần sau với lần trước nếu nhỏ hơn một giá trị xác định thì xác nhận giá trị

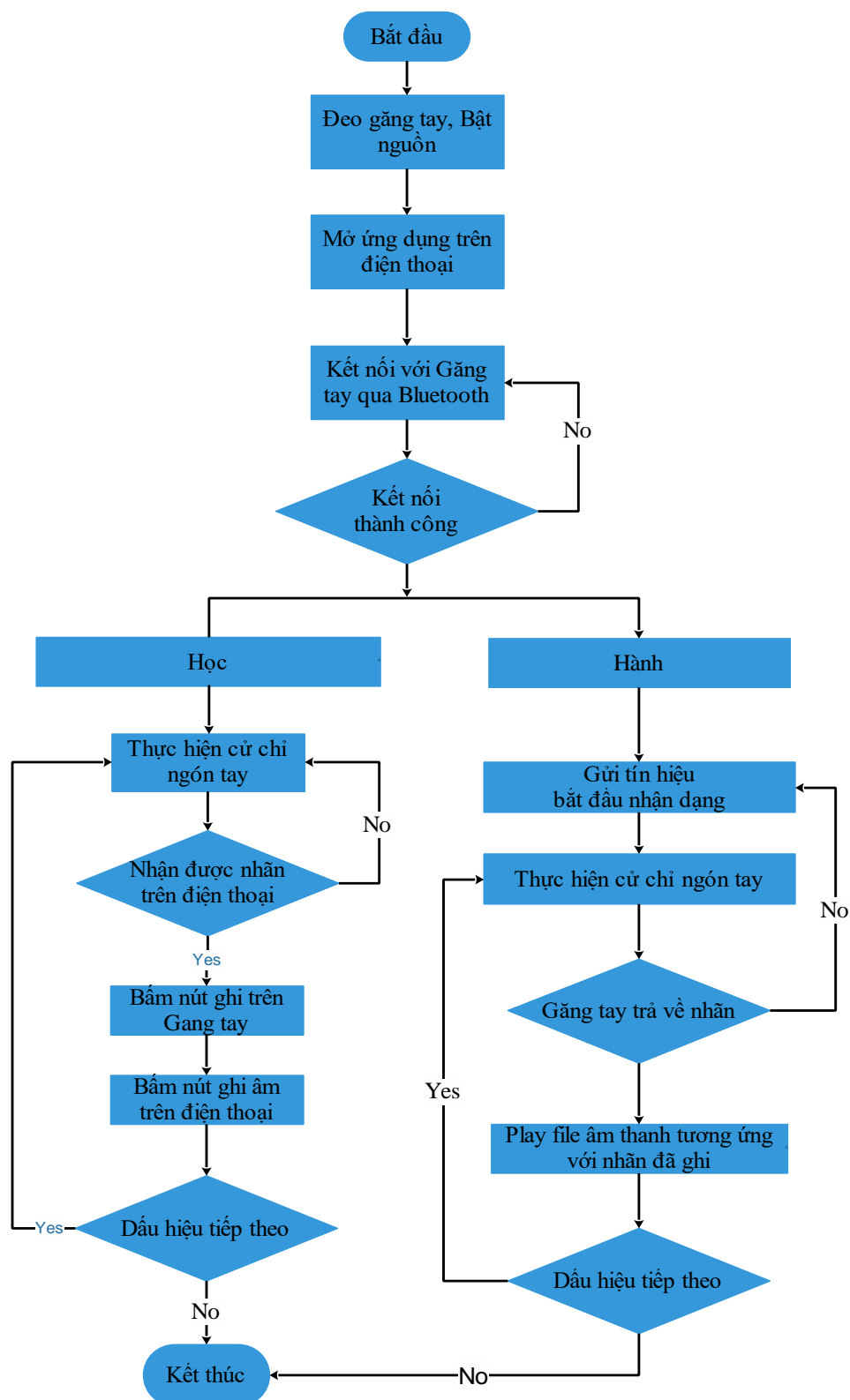
vector đó là dữ liệu của cử chỉ cần phải tính toán để nhận dạng. Giải pháp này đã giải quyết được bài toán trên và nó được thực hiện bởi hàm `CompareFlex()` trong lưu đồ.

**Giai đoạn 2:** - Sau khi có dữ liệu xác định rồi thì thực hiện thuật toán học máy KNN để tìm ra dấu hiệu cử chỉ; bao gồm các hàm `gettingHashSize()`, `givingDistance()`, `Softing()`, `findingMostOccurane()`,

`findingUnique()`. Hàm `gettingHashSize()` thực hiện tính số dòng trong tệp dữ liệu mẫu `data.csv` ghi trong thẻ Micro SD bằng cách đếm các ký tự xuống dòng (“\n”). Dùng làm tham biến để tính số số khoảng cách với mẫu dữ liệu cử chỉ và đồng thời nó được sử dụng làm số thứ tự để lấy các nhãn: “A, B, C, D, E, F, ...” trong bộ nhớ Flash của vi điều khiển, để cộng thêm với dữ liệu số mà hàm `CompareFlex()` tính toán được, ghi vào tệp `data.csv` để làm dữ liệu mẫu nhận dạng dùng làm dữ liệu đầu vào cho tính toán KNN sau này. Hàm `givingDistance()` thực hiện tính toán khoảng cách của dữ liệu cử chỉ với tất cả các mẫu dữ liệu đã ghi trong tệp `data.csv`. Nhóm nghiên cứu đã thử nghiệm các cách tính khoảng cách khác nhau như: sử dụng phép trừ bình thường giữa các mẫu dữ liệu, sử dụng các tính khoảng cách Euclidean, khoảng cách Manhattan. Không cách tính khoảng cách nào cho kết quả tối ưu vượt trội. Hàm `Softing()`, `findingMostOccurane()` thực hiện các sắp xếp các khoảng cách vừa tính được theo giá trị tăng dần và tìm ra đa số các Nhãn dấu hiệu tương với các dấu hiệu đó trong số K khoảng cách. Nhóm đã thử nghiệm với nhiều hệ số K khác nhau, thì hệ số K bằng 5 cho kết quả nhận dạng tốt hơn một số giá trị K khác. Hàm `findingUnique()` lấy kết quả đầu vào là các Nhãn xuất hiện nhiều hơn các nhãn khác từ K Nhãn nhận được từ hàm `findingMostOccurane()` và trả về một ký tự là Nhãn xuất hiện nhiều nhất đó, để làm đầu vào cho hàm `sendToBlueTooth()`.

**Giai đoạn 3:** - Phát dữ liệu dấu hiệu cử chỉ tìm được sang điện thoại qua Bluetooth để điện thoại play file âm thanh tương ứng với cử chỉ đó. Bao gồm các hàm `checkToSend()` và `sendToBlueTooth()`. Hàm `checkToSend()` thực hiện việc kiểm tra tính sẵn sàng nhận dữ liệu mới của điện thoại, đồng thời kiểm tra dữ liệu tính toán được có trùng với dữ liệu vừa gửi sang điện thoại không. Nếu vẫn là dữ liệu của dấu hiệu vừa gửi thì không gửi tiếp nữa. Việc này tránh trường hợp người sử dụng thực hiện một dấu hiệu rồi vẫn giữ tay ở trạng thái đó thiết bị sẽ phát đi phát lại rất nhiều lần một âm thanh gây nhầm lẫn cho người nghe. Hàm `sendToBlueTooth()` gửi nhãn dấu hiệu cử chỉ sang điện thoại để điện thoại phát ra âm thanh tương ứng với dấu hiệu đó.

### III.3.7.2 Lưu đồ thực hiện biên dịch tín hiệu cử chỉ sang dạng âm thanh



Hình 22: Lưu đồ thực hiện chuyển đổi cử chỉ sang tiếng nói

**Mô tả:**

Việc thực hiện chuyển đổi cử chỉ sang tiếng nói gồm hai phần chính: Học và Hành, ngoài ra ban đầu cần có kết nối giao tiếp Bluetooth giữa Găng tay và điện thoại.

- Phần Học: - Thực hiện lấy mẫu cử chỉ trên Găng tay, mỗi cử chỉ được lấy mẫu 5 lần, mỗi lần cách nhau khoảng 3 đến 5 giây, mỗi cử chỉ được tự động gán một nhãn (lable), A, B, C,... các nhãn này đồng thời cũng gửi sang điện thoại để ghi âm tương ứng.
- Phần Hành: Thực hiện cử chỉ ngón tay trên Găng tay, Găng tay sẽ gửi lại nhãn dấu hiệu cử chỉ sang điện thoại. Điện thoại nhận được nhãn sẽ tìm và play file âm thanh tương ứng.

### III.4.Hướng dẫn sử dụng găng tay thông minh

#### III.4.1 Ai có thể sử dụng thiết bị

Bất kỳ người câm hoặc khiếm thính hoặc người bệnh nào mà không có khả năng nói đều có thể sử dụng Găng tay thông minh này để giao tiếp với người nghe nói được. Nhóm nghiên cứu khuyến nghị người sử dụng thiết bị này trong một không gian/ môi trường nhất định để đảm bảo sự chính xác khi dịch cử chỉ. Khi phải đột ngột thay đổi môi trường sống cần có người nghe nói được bình thường có khả năng giao tiếp với người khiếm thính cùng huấn luyện lại Găng tay. Ví dụ khi đang sử dụng bình thường ở cộng đồng, găng tay có thể trợ giúp giao tiếp tốt một số câu nói, nhưng khi người khiếm thính đó bị bệnh cần vào bệnh viện chẳng hạn thì những câu giao tiếp thông thường ở ngoài xã hội này lại không phù hợp với môi trường bệnh viện. Lúc đó cần huấn luyện (định nghĩa) lại các câu nói cho phù hợp với môi trường mới (bệnh viện) do số lượng cử chỉ mà Găng tay nhận dạng được là hữu hạn.

#### III.4.2 Học cách giao tiếp bằng Găng tay thông minh

Găng tay thông minh được thiết kế cho mọi người có thể sử dụng được. Điều duy nhất người sử dụng cần ghi nhớ là thực hiện đúng cử chỉ ngón tay thể hiện điều mình muốn nói đúng theo mẫu cử chỉ đã ghi vào trước đó.

Thiết bị có thể gắn lên bất kỳ chiếc găng tay thường nào có trên thị trường. Ở phiên bản thử nghiệm này nhóm nghiên cứu sử dụng một chiếc găng tay len. Gắn các cảm biến độ cong vào 5 ngón tay và phần xử lý vào phần mu bàn tay.

*Các bước thực hiện giao tiếp như sau:*

- **Bước 1:** Đeo găng tay, khởi động hệ thống
- **Bước 2:** Bật ứng dụng trên điện thoại, Bấm chọn bluetooth, chọn HC-05
- **Bước 3:** Nhấn nút “Kết nối/Ngắt kết nối” để kết nối hoặc ngắt kết nối với hệ thống khi không sử dụng nữa
- **Bước 4:** Bấm nút “Hành” để bắt đầu thực hiện chuyển đổi ngôn ngữ cử chỉ sang giọng nói.

Mỗi một cử chỉ là một câu nói, khi phát xong tiếng nói của cử chỉ thứ nhất thì , nút “Phát” trên ứng dụng sẽ chuyển sang màu xanh, để người sử dụng thực hiện tiếp cử chỉ thứ 2.

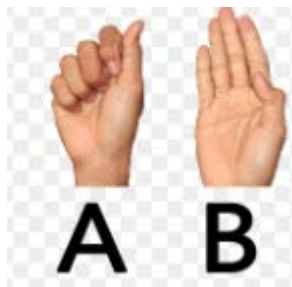
### III.4.3 Huấn luyện (định nghĩa) cử chỉ cho Găng tay

Huấn luyện cho Găng tay cần có sự trợ giúp của người có khả năng giao tiếp với người câm điếc để thống nhất cử chỉ và của người câm với tiếng nói sẽ phát ra. Phần huấn luyện này phải thực hiện với Găng tay và trên ứng dụng.

#### Thực hiện với Găng tay: - Lấy mẫu dữ liệu cử chỉ ngón tay

Người sử dụng đeo Găng tay và khởi động Găng tay, bằng cách bật công tắc nguồn trên Găng tay, đợi khoảng 3 đến 5 giây cho Găng tay khởi động xong.

Người sử dụng thực hiện dấu cử chỉ đầu tiên, ví dụ thực hiện dấu hiệu nắm tay như hình 23.



Hình 23: Dấu hiệu cử chỉ

Bấm giữ nút trên Găng tay khoảng 1 giây trên Găng tay để lưu dữ liệu vào thẻ SD. Mỗi dấu hiệu nên thực hiện 5 lần mỗi dấu hiệu cử chỉ hơi khác nhau một chút, mỗi lần nhấn nút cũng cách nhau khoảng 5 giây. Chẳng hạn như dấu hiệu nắm tay thì lần 1 có thể nắm chặt các ngón tay, lần hai hơi nới lỏng các ngón tay một chút, ... Nói chung nên thực hiện các cử chỉ đúng với những cử chỉ hàng ngày người sử dụng có thể dễ dàng thực hiện lại gần đúng nhất với cử chỉ đã lấy mẫu. Trong lúc lấy mẫu không cố gắng thực hiện các cử chỉ khó, mà lúc sử dụng khó thực hiện được sẽ gây cho thiết bị nhận dạng cử chỉ không chính xác. Khi lấy đủ 5 lần dấu hiệu cử chỉ thứ nhất, tiếp tục lấy mẫu 5 lần của cử chỉ thứ hai, thứ 3,... Các cử chỉ dấu hiệu khác nhau nên độ cong, gập của các khớp ngón tay rõ ràng. Theo thử nghiệm của nhóm thực hiện thì Găng tay có thể thực hiện nhận dạng chính xác được khoảng 15 dấu hiệu. Số lượng dấu hiệu càng tăng thì độ nhận diện chính xác càng giảm và thời gian nhận diện ra được cử chỉ càng lâu. Mỗi dấu hiệu sẽ được tự động gán một nhãn “A”, “B”, “C” để gửi sang điện thoại tìm file âm thanh tương ứng.

#### Thực hiện kết hợp giữa Găng tay và điện thoại.

*Các bước thực hiện như sau:*

- **Bước 1:** Đeo găng tay, khởi động hệ thống
- **Bước 2:** Bật ứng dụng trên điện thoại, Bấm chọn bluetooth, chọn HC-05

- **Bước 3:** Nhấn nút Kết nối/Ngắt kết nối để kết nối hoặc ngắt kết nối với hệ thống khi không sử dụng nữa, Bấm nút “Học” để bắt đầu thực hiện.
- **Bước 4:** Người sử dụng thực hiện lại các cử chỉ ngón tay như đã thực hiện lúc lấy mẫu dữ liệu trên Găng tay. Trên màn hình điện thoại sẽ xuất hiện các chữ A, B, C, ... lúc này người hỗ trợ sẽ bấm vào nút “Ghi” trên điện thoại để ghi âm lại âm thanh định nghĩa của cử chỉ. Khi ghi âm xong cử chỉ thứ nhất thì bấm vào nút “Dừng” để kết thúc ghi âm. Bấm nút “Phát” để nghe lại âm thanh. Nếu âm thanh đã rõ ràng thì bấm nút “Học” và người sử dụng thực hiện dấu hiệu cử chỉ ngón tay thứ 2 để cho người hỗ trợ ghi âm cho cử chỉ thứ 2 tương tự như cử chỉ thứ nhất. Nếu khi nghe lại chất lượng âm thanh chưa được tốt thì bấm lại nút “Ghi” để ghi lại âm thanh cho cử chỉ đó. Mỗi nhãn cử chỉ chỉ cần ghi âm một lần, và không cần phải thực hiện cử chỉ tuần tự như lúc lấy mẫu dữ liệu. Tiếp tục thực hiện như vậy cho đến khi đã ghi âm hết các dấu hiệu cử chỉ đã lấy mẫu. Bấm nút “Hành” để bắt đầu chuyển đổi cử chỉ sang âm thanh.

*Ghi chú: Với người sử dụng quen thì có thể không cần thực hiện bước lấy mẫu dữ liệu cho Găng tay trước mà kết hợp lấy mẫu dữ liệu và ghi âm cử chỉ cùng lúc. Tức là người cầm có thể thực hiện bấm nút 5 lần để ghi dữ liệu đồng thời kết hợp với người hỗ trợ để ghi âm cử chỉ. Tuy nhiên nhóm thực hiện không khuyến khích sử dụng phương pháp này vì dễ gây nhầm lẫn mẫu dữ liệu làm cho Găng tay nhận dạng cử chỉ không chính xác.*

### **Thực hiện định nghĩa lại cử chỉ**

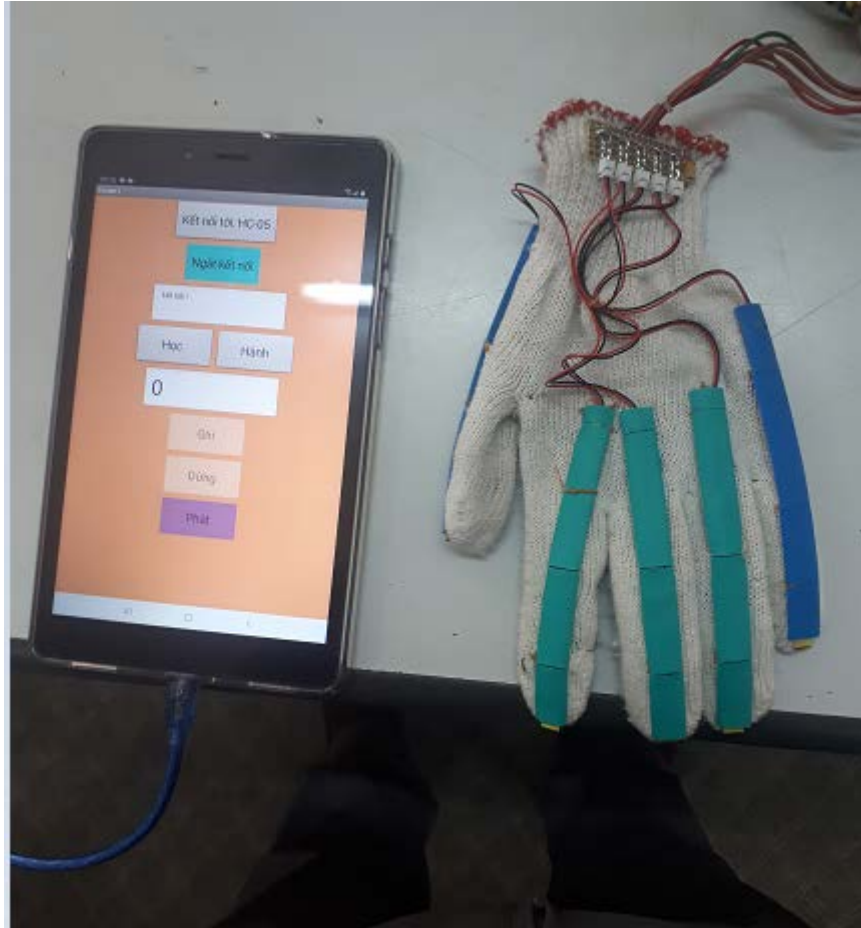
Khi cần định nghĩa lại cử chỉ, tức là thay đổi lời nói phát ra khi thực hiện cử chỉ đó. Ta biết rằng mỗi một cử chỉ đã được hệ thống gán cho một “nhãn” nằm trong bảng chữ cái Alphabet được lưu trong thẻ SD của hệ thống. Nên khi cần thực hiện định nghĩa lại các cử chỉ chỉ cần thực hiện lại bước “*Thực hiện kết hợp giữa Găng tay và điện thoại*” mà không cần thực hiện lấy mẫu dữ liệu nữa. Nếu đã người sử dụng đã nhớ cử chỉ đó có nhãn nào thì người hỗ trợ có thể nhập luôn nhãn đó trên điện thoại mà không cần thực hiện lại cử chỉ bàn tay trên Găng tay. Ví dụ nếu nhớ cử chỉ bàn tay nắm rồi đưa ngón trỏ lên có nhãn là “D” thì người hỗ trợ nhập luôn nhãn “D” vào ô màu trắng trên ứng dụng điện thoại rồi ghi âm luôn không cần thực hiện lại cử chỉ trên Găng tay.



## CHƯƠNG IV: THỬ NGHIỆM VÀ ĐÁNH GIÁ

*Chương này trình bày các kết quả thực hiện nhiệm vụ, và một số thử nghiệm chuyển đổi ngôn ngữ nhận dạng cử chỉ bàn tay sang giọng nói.*

### IV.1. Kết quả nghiên cứu chế tạo thiết bị



Hình 24: Hình ảnh Găng tay thử nghiệm

Găng tay bao gồm một hệ thống cảm biến độ cong (Flex sensor) gắn vào các ngón tay của một Găng tay thường. Các cảm biến được kết nối với bộ vi điều khiển như hình 24. Mỗi khi thực hiện cử chỉ ngón tay, một âm thanh sẽ được phát ra từ máy tính bảng hoặc điện thoại.

Thiết bị cho phép nhận dạng chính xác khoảng 10 đến 17 cử chỉ tùy theo tay của người sử dụng.

Giá thành khi sản xuất thiết bị (không bao gồm điện thoại thông minh) với số lượng lớn là chấp nhận được (khoảng 1 đến 2 triệu đồng).

Dễ sử dụng: Thiết bị khi sử dụng không cần nhiều thao tác phức tạp vốn là cản trở lớn cho người sử dụng.

Gọn nhẹ, phù hợp với thể trạng sức khỏe của người sử dụng

Dễ dàng đeo mang, sử dụng trong các môi trường thông thường

Dễ sửa chữa, thay thế các thành phần của thiết bị trong điều kiện Việt Nam hiện tại

#### IV.2. Thử nghiệm Găng tay

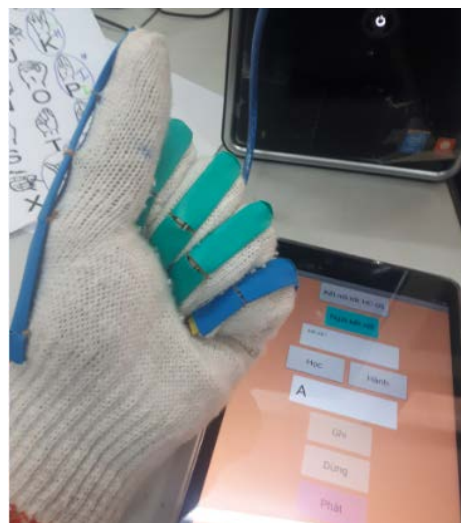
Nhóm nghiên cứu đã thử nghiệm với nhiều dấu hiệu khác nhau, kết quả như sau:

- Thực hiện cử chỉ như hình 25, số lần thực hiện 10 lần
- Kết quả: nhận dạng được cử chỉ đúng 10, chuyển đổi ngôn ngữ thành công (play đúng file âm thanh) 10. *Độ chính xác 100 %.*



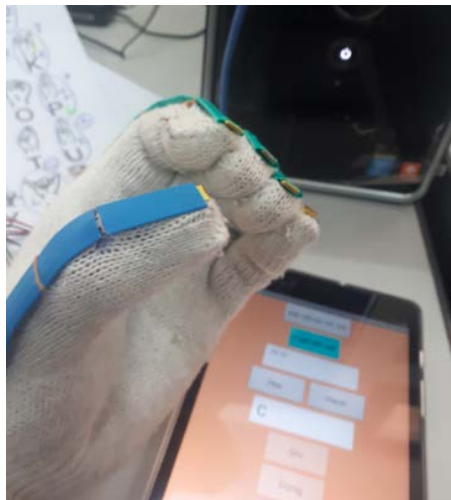
Hình 25: Cử chỉ thử nghiệm

- Thực hiện cử chỉ như hình 26, số lần thực hiện 10 lần
- Kết quả: nhận dạng được cử chỉ đúng 10, chuyển đổi ngôn ngữ thành công (play đúng file âm thanh) 10. *Độ chính xác 100 %.*



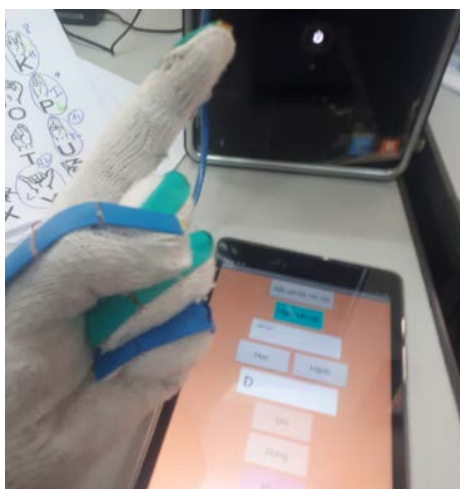
Hình 26: Cử chỉ thử nghiệm

- Thực hiện cử chỉ như hình 27, số lần thực hiện 10 lần
- Kết quả: nhận dạng được cử chỉ đúng 10, chuyển đổi ngôn ngữ thành công (play đúng file âm thanh) 10. Độ chính xác 100 %.



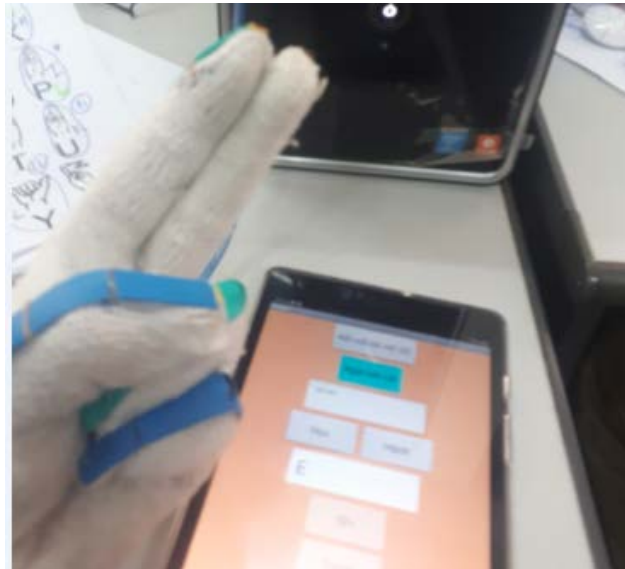
Hình 27: Cử chỉ thử nghiệm

- Thực hiện cử chỉ như hình 28, số lần thực hiện 10 lần
- Kết quả: nhận dạng được cử chỉ đúng 10, chuyển đổi ngôn ngữ thành công (play đúng file âm thanh) 10. Độ chính xác 100 %.



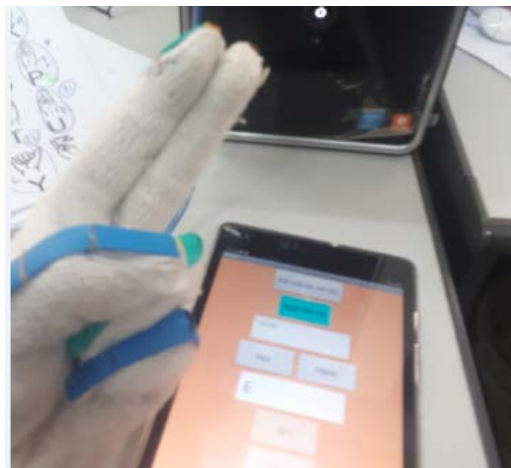
Hình 28: Cử chỉ thử nghiệm

- Thực hiện cử chỉ như hình 29, số lần thực hiện 10 lần
- Kết quả: nhận dạng được cử chỉ đúng 10, chuyển đổi ngôn ngữ thành công (play đúng file âm thanh) 10. Độ chính xác 100 %.



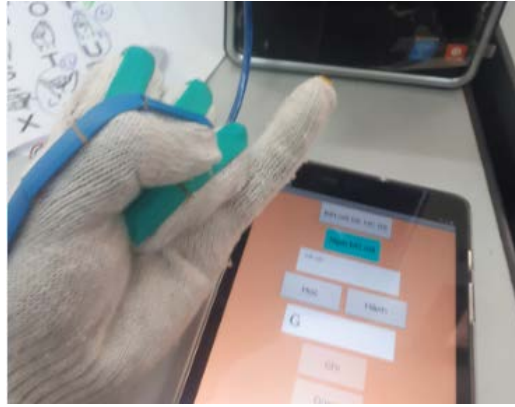
Hình 29: Cử chỉ thử nghiệm

- Thực hiện cử chỉ như hình 30, số lần thực hiện 10 lần
- Kết quả: nhận dạng được cử chỉ đúng 10, chuyển đổi ngôn ngữ thành công (play đúng file âm thanh) 10. Độ chính xác 100 %.



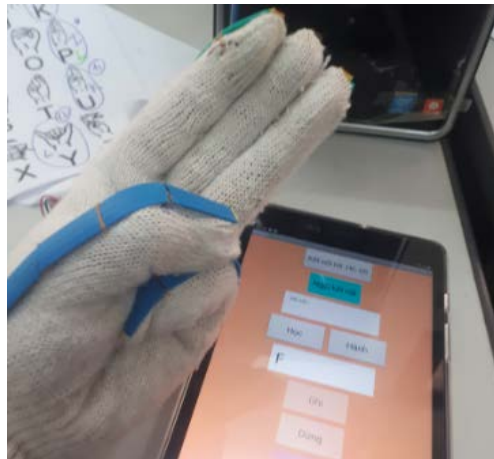
Hình 30: Cử chỉ thử nghiệm

- Thực hiện cử chỉ như hình 31, số lần thực hiện 10 lần
- Kết quả: nhận dạng được cử chỉ đúng 10, chuyển đổi ngôn ngữ thành công (play đúng file âm thanh) 10. Độ chính xác 100 %.



Hình 31: Cử chỉ thử nghiệm

- Thực hiện cử chỉ như hình 32, số lần thực hiện 10 lần
- Kết quả: nhận dạng được cử chỉ đúng 10, chuyển đổi ngôn ngữ thành công (play đúng file âm thanh) 10. Độ chính xác 100 %.



Hình 32: Cử chỉ thử nghiệm

Tiếp tục thử nghiệm với các cử chỉ bàn tay khác nhau, cho đến khi các cử chỉ mà độ cong của ngón tay không còn khác biệt nhiều với các cử chỉ đã thực hiện trước đó (khoảng từ 12 đến 17) cử chỉ tùy từng người thì thiết bị sẽ nhận dạng chậm và không còn độ chính xác nữa.

## KẾT LUẬN

Nhiệm vụ đã đạt được tất cả các nội dung yêu cầu như: Nghiên cứu khảo sát một số đặc điểm phổ biến ở nhóm người câm điếc hình; nghiên cứu ngôn ngữ ký hiệu tiếng Anh và Tiếng Việt ; nghiên cứu chế thử mẫu Găng tay nhận dạng ngôn ngữ ký hiệu và thử nghiệm Găng tay trong môi trường sử dụng thực tế.

Găng tay đã nhận dạng được chính xác khoảng 12 đến 17 dấu hiệu cử chỉ. So sánh với các nghiên cứu trước, số lượng cử chỉ thiết bị nhận dạng được tăng lên không nhiều. Tuy nhiên điều khác biệt ở Găng tay của nhóm nghiên cứu chế tạo là có khả năng định nghĩa lại các cử chỉ. Dẫn đến khắc phục được phần nào hạn chế của thiết bị. Phần mềm trên điện thoại di động phát ra âm thanh dấu hiệu rõ ràng giúp việc giao tiếp giữa người câm điếc giao tiếp và người nghe được thuận lợi nhanh chóng. Việc ghi âm từ phần mềm để định nghĩa các cử chỉ mới dễ dàng.

Trong quá trình thử nghiệm đề tài, nhóm nghiên cứu nhận thấy thiết bị này đáp ứng tốt giao tiếp đơn giản trong một số trường hợp cụ thể như mới chỉ dừng ở mức giao tiếp một chiều – từ người khiếm thính đến người nghe được, còn chiều ngược lại là chưa có. Do vậy. Hướng nghiên cứu trong thời gian tới, nhóm sẽ thực hiện hoàn thiện sản phẩm, một mặt tăng độ chính xác của nhận dạng cử chỉ, đồng thời tăng số lượng cử chỉ có thể nhận dạng được, đặc biệt là nghiên cứu bổ sung chức năng cho phép người nghe được “nói” để người khiếm thính “nghe” được. Hay có thể phát triển thiết bị này và nâng cấp ứng dụng trên điện thoại thông minh thành các loại hình khác nhằm phục vụ người khuyết tật, ốm đau ví dụ:

- Các trò chơi nhằm phục hồi chức năng vận động tay đối với người phục hồi sau tai biến,...

- Dạy NNKH cho người câm trong hoạt động bắt đầu học bảng chữ cái tiếng Việt. Đồng thời nghiên cứu tích hợp hệ thống này với hệ sinh thái sản phẩm cho người khuyết tật nói chung và người khiếm thính nói riêng.

,

## TÀI LIỆU THAM KHẢO

[1]	<a href="https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss">https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss</a>
[2]	<a href="http://www.spectronics.com.au/product/dynavox-vmax">http://www.spectronics.com.au/product/dynavox-vmax</a>
[3]	Vương Hồng Tâm, Viện khoa học giáo dục Việt Nam, - Nghiên cứu cách biểu đạt ngôn ngữ kí hiệu của người Điếc Việt Nam
[4]	Đỗ Thị Hiên, Ngôn ngữ kí hiệu của cộng đồng khiếm thính Việt Nam; thực trạng và giải pháp, đề tài NCKH cấp bộ, Viện Khoa học xã hội Việt Nam