

SENSE+

A Cognitive Application
Enabling the Visually Impaired to
See, Explore, Notice and Search
with **Ears**

Imagine Cup 
Project Proposal

Team Members

MA Mingyu

derek.ma@connect.polyu.hk

WANG Shi

winona.wang@connect.polyu.hk

CHANG Shuhao

shuhao.chang@connect.polyu.hk

Institute

Department of Computing and
Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University

Mentor

Dr Henry CHAN

Introduction

Most of the **information** comes from the vision

Vision is the dominant sensor of human. 40% of all nerve fibers connected to the brain are linked to the retina and more of our neurons are dedicated to vision than the other four senses combined. It is estimated by the researchers that around 80% of our perception, learning, cognition and activities are mediated through vision. According to World Health Organization's statistics, 285 million people are estimated to be visually impaired worldwide. The visually impaired face a lot of difficulties in their life.

Social Isolation

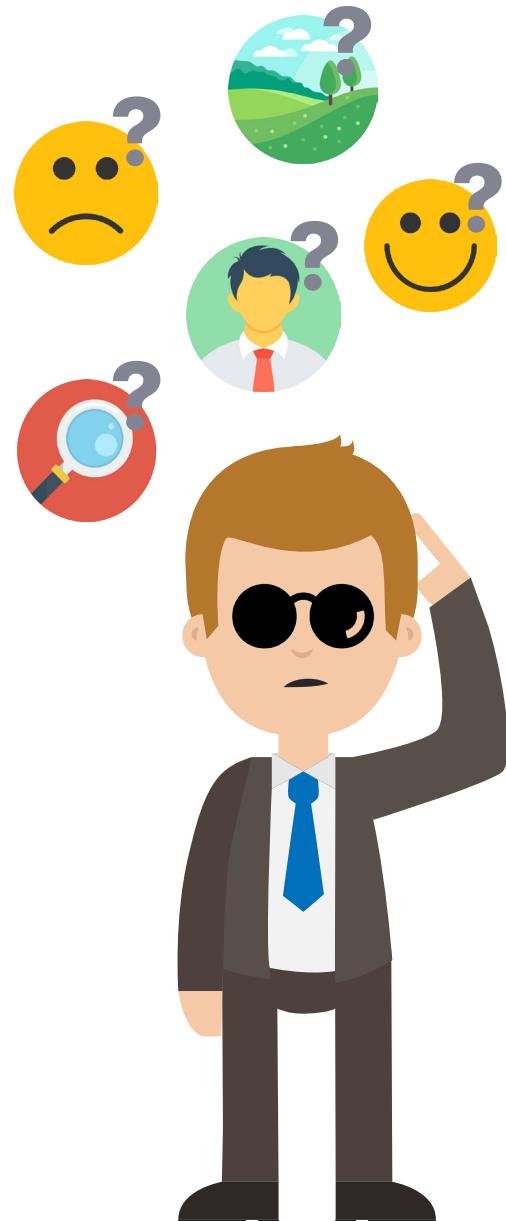
The visually impaired cannot perceive the face information or emotions of others. The blind can only perceive others by sounds and need to recognize their friends by distinguishing voices. Usually the blind cannot know who is passing by and feel difficult to remember friends, which makes the blind involved less in interpersonal interactions and social life. This may lead to the social isolation problem.

Daily Inconvenience

The visually impaired encounter much inconvenience in daily life. Without visual information, some daily tasks can become challenges for the blind. For example, the blind know nothing about the color of the clothes and they strongly rely on their friends to pick the clothes. Besides, It is difficult for the visually impaired to pick up the things dropped on the floor or to find somebody in a space because they cannot actively look for the objects using their eyes. These locating tasks cost much time and decline their life quality.

Loneliness

Staying in the dark, the visually impaired are more likely to feel lonely. Visual impairment negatively influences their mental health and the social support for them is inadequate. The visually impaired seldom go out like normal people because of inconvenience. Loneliness is a psychological state of feeling lack of connection with others. The blind need more care from others to relieve the loneliness.



Can the visually impaired “see” the world
in another way?



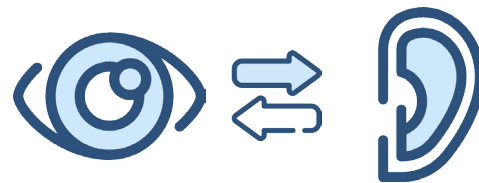
Our Solution

Recreate Visual Sense with Cognitive Technologies and Online Information

Nowadays, artificial intelligence (AI) develops rapidly and is applied more and more widely in daily life especially the cognitive technologies. The smartphones are equipped with the cameras and microphones which are able to capture images and accept speech input. With the support of the Microsoft Cognitive API, the smartphones become smarter to recognize the face, to analyze images, to convert the text or audio and so on.

With these technologies, the visually impaired can “see” the people around and the surroundings from the sounds generated by the smartphones. Furthermore, users can ask the smartphones to actively locate the objects or people that they would like to look for. In this case, the visually impaired can be bridged with the environment and to involve the visually impaired more in interpersonal interactions.

Therefore, we propose to develop a mobile application applying the Microsoft Cognitive API. The mobile application can convert vision to sound, which supplements the sense of the visually impaired, so it is called “Sense +”.



Objectives

The final goal of this project is to develop a technology solution to help the visually impaired to perceive the information originally got from visual sense. Information around them such as people, identities, emotions, colors, local shops, weather conditions can be perceived with the help of Sense+ application. Mobile application is a proper platform to realize all the functions. Sense+ is aimed to recreate visual sense with cognitive technologies and online information.

Sense+ should accomplish following features for the visually impaired:

Make Daily Life Convenient

Describe color, be aware of the location of specific object, locate object in a easy way, get local service information and online information such as restaurants recommendations and weather.

Enrich Social Interactions

Visual information of strangers and friends such as facial expressions, emotions, identities and so on can be perceived by the visually impaired with the help of Sense+.

Generate Mental Comfort

An more independent daily life and richer social life can be achieved with the help of Sense+. As a personal assistant for the visually impaired, they can also interact with the app through speech.

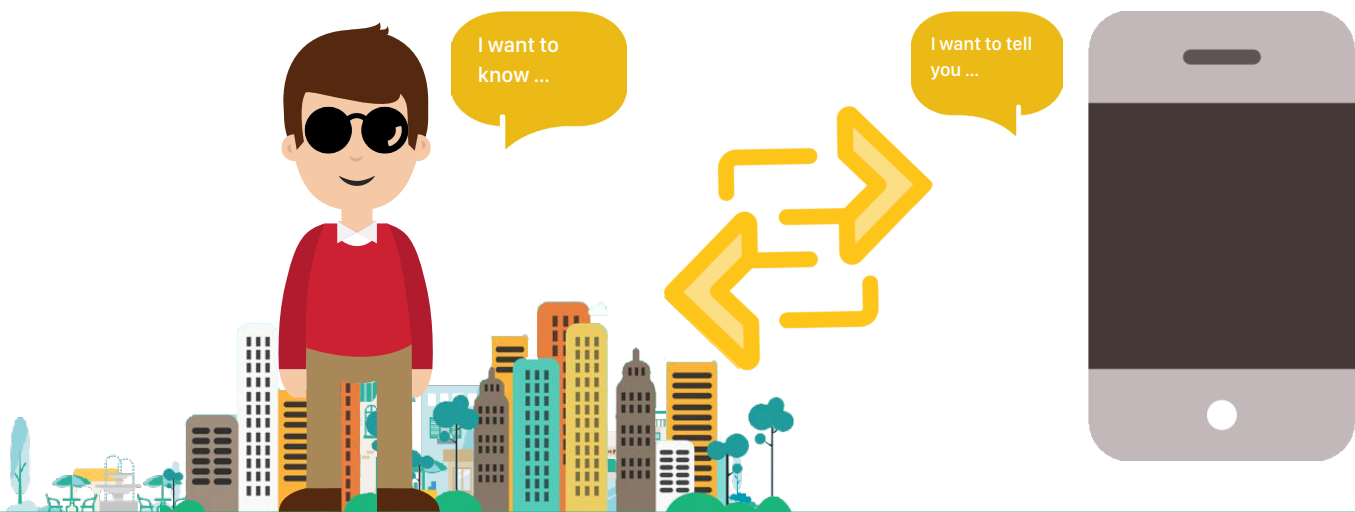
Innovativeness

Currently, there are many applications that help the visually impaired people in the market, and some specific wearable devices were developed as well. There are some drawbacks for these existing products. We believe the following innovativeness endorse the uniqueness of Sense+.

Bi-directional Interaction - Not Only "Listen", But Also "Ask"

Usually the existing applications are capable to describe the environment in the captured images or tell some other information to the users. But the users cannot indicate their command such as looking for something in the surroundings or what kind of information they would like to know.

While Sense+ can not only tell information around the users, but also receive the command from the user through the speech recognition. The users can locate objects, and they can get the specific information which meet their needs through inputting their commands and needs to the application.



No Extra Hardware - Easy to Deploy

Existing similar applications may require special wearable devices such as glasses with camera, or high-performance computing devices such as smartphones with powerful hardware. The lower the hardware requirements are, more people the application can help.

While, the only hardware required for using Sense+ is just a smartphone. With the support of Microsoft Azure cloud computing platform, the computation ability requirement for the smartphone is very low as all complex algorithms are running on the cloud. Thus even the smartphones with low-performance hardware can use Sense+ as well.

Customization - Designed For You, Stays With You

Current applications helping the visually impaired can only describe the environment or people's face features. Sense+ will record faces for specific people and train the machine learning algorithm, so that identities of specific face can be recognized. Sense+ has a record for each user, where their preferences and friends will be recorded. So that this application is unique for each user. The longer the time of using, the better customization can be achieved.



Multiple Source Information - Provides Various Choices For You

Current applications existing in the market help the users to perceive the information only from one aspect such as computer vision. Sense+ combines information from different sources to provide enhanced and comprehensive information to the visually impaired. For example, computer vision and cognitive technologies are used to analyze the surroundings, many other internet platforms' service are built in to enhance the information reliability and richness such as Yelp, AccuWeather, EventBrite, geographical services and so on.

Satisfy Psychological Needs - A Considerate Friend For You

User interactions are designed after analyzing our target users' psychological needs. They may feel lonely and they have the needs to be more independent. Our functions are designed to enhance their social life and achieve more independent abilities. The visually impaired can interact with the app without any help. We import natural language processing technologies to achieve more natural interaction ways so that this application can be served as a personal assistant and considerate friend for the visually impaired.

Target Users

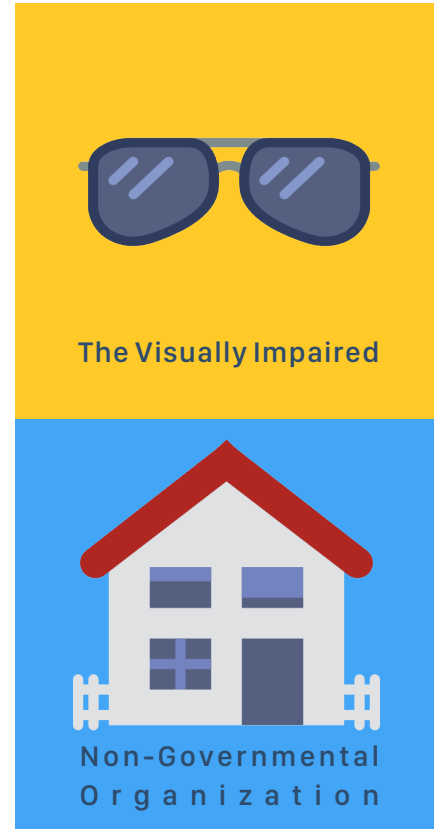
Our target users are people with **visual impairment**. Using Sense+ makes contribution to enhance the **independence** for the visually impaired.

Why Use Us?

Users can accomplish more living tasks such as perceiving environment, recognizing people and awaring emergencies. Blinds can deal with more matters which they were not able to do before. Sense+ brings higher **self-independence** to the blind users, and bring them more self-confidence in solving daily affairs. In addition, Sense+ helps the blind users to remember and recognize new and old friends as well as other relevant experiences with the friends. The blind user has time to prepare the talking topics before the talk begins. This design can greatly increase the **confidence** for social interaction of blind people.

Promotion Channel

Sense+ can be further promoted to **NGO** of blind caring organizations such as HKSB(The Hong Kong Society for the Blind). Cooperating with these NGOs has great effect on promoting Sense+ and benefiting more blind people with visually impairment.



User Scenarios



See



Explore



Notice



Search

See Friends and Colors

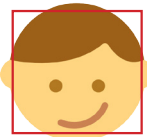
See Friends

Sense+ can identify the friends who have been recorded in the database. When there is any people coming, Sense+ will try to recognize the friend and tell the users the friend's name, a short description of the friend and the emotion of the friend. This function facilitates the social interaction and makes the visually impaired more confident.

A blind student called Jack is walking on a street and holding his smart phone with Sense+ in See mode towards his front. He meets Mary that he had recorded before, Sense+ will notice Jack that old friend Mary is coming and provide the general information of Mary such as her name, emotion, and other information recorded before.



Name: ***
Age: ***
Gender:***
Emotion: ***
...



If Jack meets a person who has not been recorded before, Sense+ is able to point out the gender, age and emotion information to Jack. After their meeting, Jack may input the name and description of the person he met just now into Sense+ and the face and identity of the new person will be recorded. Identifying people can also be achieved when many people appear together, the description will be conduct one by one according to the position of these people.

See Color

The visually impaired cannot perceived the color of the clothes but they have the needs to dress themselves. Sense+ can identify the color and help the visually impaired to design their dressing styles. so that the visually impaired can be more independent in daily tasks. With Sense+, the visually impaired can "see" color and be more elegant.

Jack is willing to wear a red shirt today, but he can't find which shirt is red in his closet. Jack can say "what is the color" to Sense+ and erect his smartphone to capture the shirt in front of him. Then Sense+ will tell Jack corresponding color.



Explore Restaurants and Shops



Sense+ can sort out restaurants around the user according to the keyword input by the user and the current location. The users can also get the information of the cuisines that the restaurants sell and the rating of the restaurants. Sense+ helps the users to explore different meals every day and to be connected more closely with the world.

Jack is a food lover and he would like to look for a restaurant. He may switch to explore mode by saying "Explore", and then says what kind of food he want such as "Italian food". Sense+ will search the information from Yelp and sort out the best reputation Italian restaurant nearby. Sense+ can generate a brief introduction including name, style, price, rate and phone number of the restaurant. The blind user can know the direction of the restaurant or contact it.

Notice Weather and Events

Sense+ can play a role as a warm living secretary and close friend of the user. Utilizing local online services and information, it provides the rich and helpful information such as weather, news, local events, the most popular song in a pleasant way. This function aims to warm the user's heart, and to reduce loneliness feeling.



After Jack wakes up, the first sense from the world is from the notice of Sense+. "Good morning Jack, today is 12-15 celsius degree. and it is possible to rain in the afternoon. Don't forget to bring an umbrella. Following is today's news you may like. Sports News: Rockets 137 beats Thunder 125. And you may like this event. Bruno Mars will stage a concert the day after tomorrow, in 131th avenue. Today I am going to recommend a good song Count on Me, audition is going on! Sense plus is always with you." These warm and caring greetings with valuable information make a pleasant start of a new day for Jack.

Search Personal Belongings

Sense+ can be a useful tool for blind users to look for a particular object. The user can first indicate what object to look for and then the Sense+ begins to locate the object in the images captured. If the object is found, Sense+ will tell the user. This eliminates daily inconvenience of the blind and make their life easier.

Jack knocks his wallet down from table by accident. But he cannot touch his wallet since he doesn't know the position. Jack switches to search mode, and speaks to the phone "Search wallet". Then he can hold up his phone and change directions smoothly around the position his wallet could be at. Sense+ will report there is no wallet in front of him, until the wallet is found. Then Jack can easily get back his wallet.



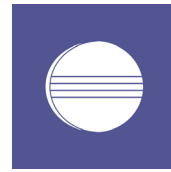
Implementation

Technologies

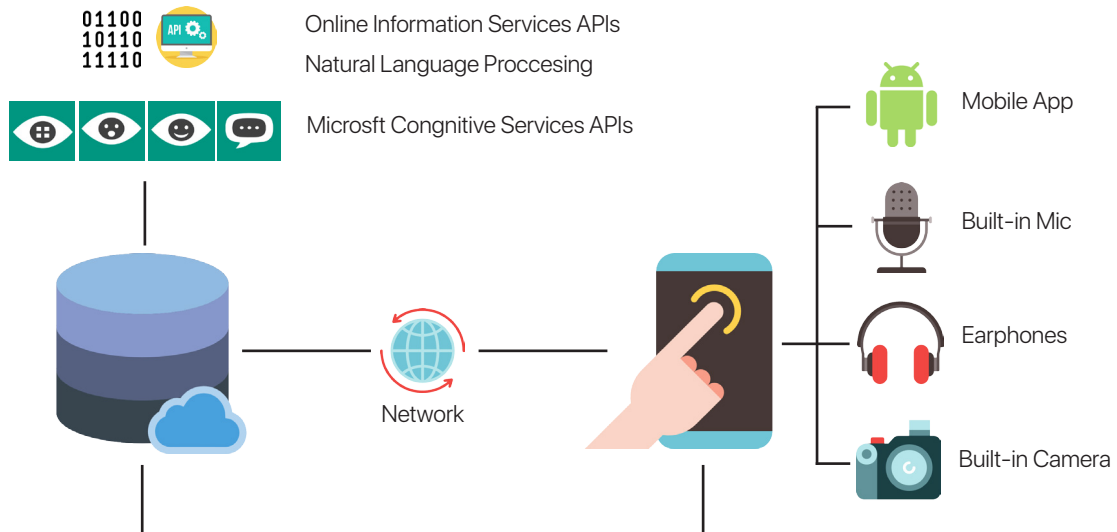
Android SDK and **Eclipse** will be the tools to develop the mobile application. **Microsoft Azure** cloud platform will be the back-end server.

We will use **Tomcat** as the container running the server program written with **Java EE**. Most importantly, we leverage the **Microsoft Cognitive API** like Face Identify, Computer Vision and Emotion APIs to analyze the images and uploaded from the mobile application.

Yelp API is used for sites recommendations, **AccuWeather API** is used to provide weather services, geographical services API is used to provide direction of specific sites.



System Architecture



Back-End Server

The back-end server will be built based on the Microsoft Azure, where the server program and the database will be configured. The server program can accept the image uploading request from the mobile application as well as generate responses of face or environment information. It is connected with Microsoft Cognitive APIs.

Mobile Phone

The mobile application is responsible for the logic and calculation on the user side. It handles the communication between the mobile phone and the back-end server. The camera on the mobile phone can capture the images of the people and the environment around the user. The microphone of the mobile phone can assist the speech input by the users. The earphone will output the audio to the users.

Team

We are a team of three undergraduate students from **Department of Computing** and **Department of Electronic and Information Engineering** of **The Hong Kong Polytechnic University**.



MA Mingyu
Derek

Front-end Development
Data Mining
UI Design



WANG Shi
Nona

Computer Networking
Android, Marketing
Management



CHANG Shuhao
Constan

Cloud Server
Product Design
3D Painting

Summary

The visually impaired encounter the **social isolation** and **daily inconvenience** every day. This may further bring the **feelings of loneliness**.

A **mobile application** and a cloud **back-end base** will be developed for solving this problem. With the help of **Microsoft Cognitive Services APIs** and **Microsoft Azure**, Sense+ can **recognize friends and colors**, **explore the restaurants nearby**, **notice users of the weather and the local events**, **search the objects in the circumstances**.

Sense+ does **not require extra hardware** and it provides **customized services** and **helpful information** for the visually impaired. The **psychological needs** of the users are also satisfied.

Sense+ recreates visual sense for the blind and lights up their world.

Illustrations included in this proposal are from Flaticon with free and permissible license granted