Technical Report

Problem presentation

Sight is the most important sense humans have, which is why people with sight impairment have difficulties performing simple, day by day activities. Globally, 1 billion people have a vision impairment that could have been prevented or has yet to be addressed. This 1 billion people includes those with moderate or severe distance vision impairment or blindness due to unaddressed refractive error (123.7 million), cataract (65.2 million), glaucoma (6.9 million), corneal opacities (4.2 million), diabetic retinopathy (3 million), and trachoma (2 million), as well as near vision impairment caused by unaddressed presbyopia (826 million) (1). 90% of visual impairment is in low and middle-income countries. All categories and all ages are important, but solutions must be found in time because a child with a visual impairment may require emotional support to help cope with his or her disability.

Because of stigmas and lack of awareness of the “outside world”, a child may feel: separate from society, anxious, self-conscious, dependent, inexperienced, unmotivated(2). To combat this problem, there are many applications developed by people all around the world that use different technologies to try and make it easier and safer for these people to live without constant help from others.

State-of-the-art

As expected, most of the applications are mobile, so people can use them anywhere. The main technologies used for this specific problem are voice recognition, speech to text and text to speech conversions, image processing, the use of services like GPS, maps and other information providers. The text and visual data are processed by an AI to give the user the information they need through a simple and easy to use interface.

You can read more about state-of-the-art here: https://github.com/cosmins97/TAIP-Project/tree/main/docs/State%20of%20the%20Art .

Our solution

The proposed solution is a mobile application that can be used to perform multiple tasks, like calling numbers, sending sms's, using the GPS, creating notes and more. These tasks are received via speech-to-text, the user giving the commands to the application, which will process them and execute them. The responses will be given via text-to-speech back to the user. All basic commands use natice system calls, which allows for future improvements if needed.

Communication is done in Romanian, using APIs offered by Google. This ensures compatibility with a wide range of devices and multiple Android versions. Text to speech works offline, but speech to text requires network connectivity. Diacritics are supported, and the system is fast and responsive. The vocal commands are action oriented, meaning the user will only have to give a simple command to get results. Because all those actions run on separate threads, we had to make sure they're always synced. This was achieved by using a simple architecture that allows data to flow in only one direction and the main thread is used by functions to communicate with one another.

The application is written in Java, which allows us to use annotations to do different checks and validations.

The interface is very simple, consisting of two main buttons that are turning the microphone on or off. Because the application is oriented to people with visual problems, the interface itself isn't one of the important parts, so it has to be very simple to use for a potential user that is completely blind. Because of this, the two buttons mentioned above occupy almost half the screen each, so they can be easy to find and click. The ui is responsive and the result of a command is shown as a toast at the bottom of the screen.

The user can learn what commands are available by using the help function. This function will give the main functionalities, then the user can find the commands for each of them and how to use them.

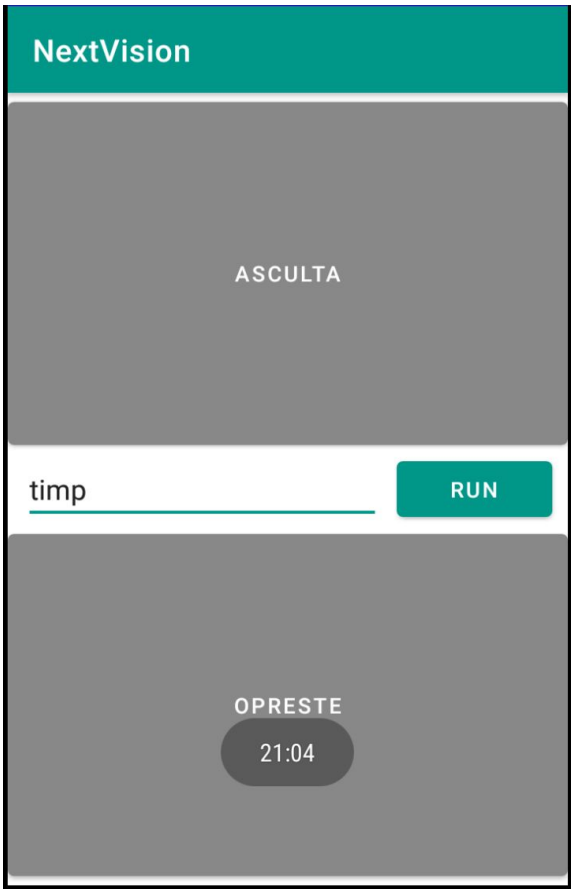
The initial configuration will be done by someone that sees the screen because there are some system configurations that can't be set from within the application itself.

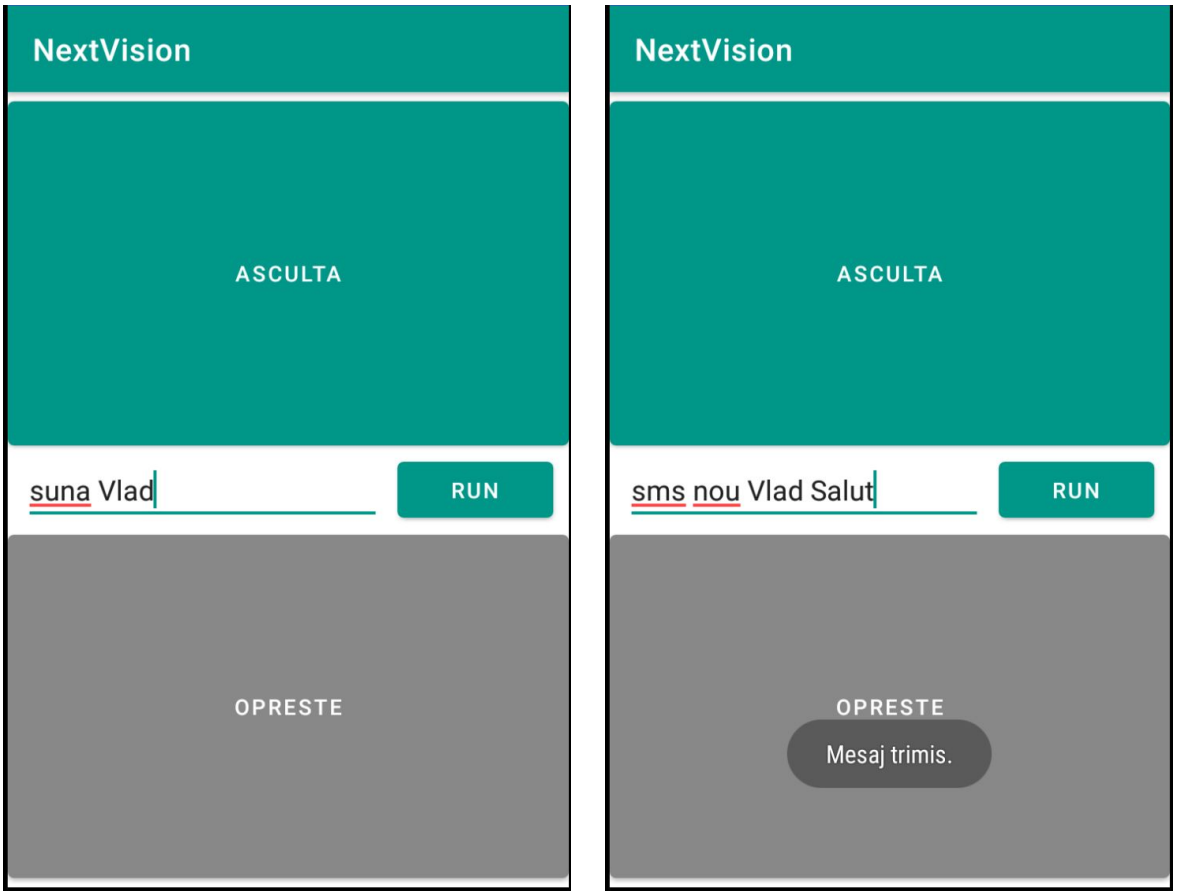
Results

Basic functionality is implemented and works very well. The application is fast and responsive. It still requires some debugging to fix a few stability issues. The text to speech and speech to text works great in silent environments.

What it can do:

- Check time and date
- Check battery and power saving mode
- Find and read sms, create and send sms'
- Call contacts, unknown numbers and create new contacts
- Create notes, listen to them or delete them
- Create, edit and delete alarms





Comparison with other solutions

Taking in consideration the state of the art made at the beginning of the project, the research that was done by the team showed that no app that currently exists is fully available in romanian for visually impaired people. Most of the applications are in english offering different options that these people have access to on their mobile.

The developed app allows the user to interact in his native language (romanian) with the application. Calling, messaging, editing alarms and notes and getting to know the hour or date or even the battery percentage or mode is done in a simple way by giving short but meaningful commands to the phone, which will automatically perform the operations for the user.

Considering TalkBack, which is one of the most complex apps of this type on the market for Android phones, setting up the application in order to be able to speak and receive input in romanian requires more steps that a person must go through to be able to set everything up. In case of NextVision, the only thing left to do when opening the app for the first time is to give

commands, everything is set up so that the interaction will be available in the native language without further settings.

Every application existing right now for Android or iOS requires permissions that must be allowed by the user at the beginning of the app usage (TalkBack, Be My Eyes). This is a security check that must be performed and cannot be skipped. NextVision needs to have enabled these permissions too and there are no workarounds for this, otherwise security protocols would be breaked. This is the reason why every time when this kind of application is newly installed, someone that can see must set up the phone.

Another useful tool is “Ajutor”, the help mode which will explain the user the commands he can give and how to give them. The help mode will explain to the user every information about every command existing, or offer information about one certain command. Right now there is no existing application for the visually impaired that has a help mode in romanian language.

NextVision implies a minimalistic interface with two on-screen buttons, one for enabling the microphone and the other one for disabling it. The functions existing currently, the interface and the functionality of the app can be found in other similar apps, the only difference is the help commands and the interaction with the virtual assistant that can be done in romanian.

Looking forward to other additions that can be found in similar solutions for this problem, we could see some interesting features, like image recognition for different objects, places, food, braille text or even people. The image recognition feature gives details about what it is in the front of the camera and can tell from colors and names of the objects to the emotions a person has (if the person smiles or is sad).

Another useful thing implemented were directions and map orientation. Google Maps team announced the launch of an additional feature that gives a more detailed voice guidance to those who can’t rely on their vision. “As I take my journey, Google Maps proactively lets me know that I’m on the correct route, the distance until my next turn and the direction I’m walking in. As I approach large intersections, I get a heads-up to cross with added caution. And if I accidentally leave my route, I’ll get a spoken notification that I'm being re-routed.” one person says while giving feedback for the Google Maps feature. Apple is reportedly also working on making its Maps app more accessible for visually impaired users.

Future work

Having a look on all the information above and based on the research made at the beginning of the project, there is enough space for improvements.

One important thing that could be implemented would be offline maps orientation and directions. A small map, at users choice, could be downloaded on the phone and accessed without internet when the case. The virtual assistant could tell the user, based only on GPS location tracking where the current position is and could give options for different routes, based on the desire for the arrival destination of the person using the app.

Some new things could be added in this case, such as: receiving information about closest public transit stations and what public transport arrives there, the timing for every bus/tram etc, different places that can be reached from that certain place. Also, there could be information about drug stores, shops, markets and other places of interest nearby.

Taking into consideration image recognition and the usage of the phone camera, a visually impaired could receive information about the color of the traffic lights and when it is safe to cross the street, if there is any crosswalk in an intersection, the type of means of transport: taxi, bus, tram, personal car.

When thinking about going shopping AI could take place and say in what aisle of a shopping centre a person is, what can be found on the shelves and even give information about different products.

For a complete experience, knowing the opening hours of an institution or shop or knowing the mood of close relatives or friends just by directing the camera at them could add value to what already exists.

Conclusions

In conclusion, after reading about the problems of people with visual impairments we began to better understand their need, but also the difficulties they face, therefore we hope that through this application we will offer a helping hand to these people or at least we will open a path or an opportunity in creating an application that is easy to understand, intuitive and accessible to all in need. We created a first functional version that contains the basic functionalities as we presented in our solution section, but we hope to fulfill everything we set out to do in the future work section.

Bibliography

(1) https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment

(2) https://www.slideshare.net/zunerashahzad1122/visual-impairment-presentation-1

Links

https://github.com/cosmins97/TAIP-Project