ISSN 1870-4069

A Process for Automatic Generation of Medical Mobile Applications using Voice Recognition

Jesús Fernández Avelino, Giner Alor Hernández, Mario A. Paredes Valverde, Lisbeth Rodríguez Mazahua, María A. Abud Figueroa

Tecnológico Nacional de México,
Instituto Tecnológico de Orizaba,
Mexico

{fdezjesus12, mapv1015, mabudfig}@gmail.com,
{galor,lrodriguezm}@ito-depi.edu.mx

Abstract. In software engineering, the adoption of mobile devices generates the need for proposing new methods and processes for developing mobile applications. This paper presents and describes a software development process for the automatic generation of medical mobile applications using voice recognition. The software development process is supported by a conversational agent that captures the functional requirements of the application to be developed. A Web-based prototype was developed as a proof-of-concept of the process proposed. We believe this process will allow reducing time and effort to developers in mobile applications.

Keywords: voice recognition, design patterns, software development process.

1 Introduction

Over the years mobile phones have changed from a simple communication device to an operative tool. People perform tasks from a mobile application, this is made possible through the development of mobile applications. Now these mobile applications have become an integral part of our lives and we rely on them in more than one way. Nowadays there are a lot of mobile healthcare applications that allow registering medical data in order to prevent, control, and monitoring diseases such as diabetes and hypertension. These applications use voice recognition which allows optimizing time and effort.

Voice recognition is the ability of a computer to understand and execute voice commands. Voice recognition is a way of stablishing communication between user and computer, it aims to replace other ways of interaction such as a mouse, keyboard, touch functionality to mention but a few. Nowadays, it is possible to use mobile devices, cars, smart TVs and domestic automation technology by voice recognition in order to perform everyday tasks. Nevertheless, this kind of interaction is used in different contexts. Voice recognition is very important for the business and professional sectors.

This paper presents and describes a process to automatically generate user interfaces of mobile healthcare applications by using voice recognition. This paper is organized into four sections, Section 2 presents related work, Section 3 describes the process to generate UI automatically, Section 4 a prototype as a proof-of-concept of the process proposed. Finally, conclusions and remarks are provided.

2 Related Work

The following is a summary of the most important related work according to voice recognition, which allowed it obtains valuable information related to this work.

Erić et al. [1] carried out a detailed comparison between different voice recognition tools such as Jasper platform, Google speech API (Application Programming Interface), Alexa Voice Service and Bing speech API to implement them in home automation projects. Moreover, Cortes et al. [2] presented a generator of mobile applications based on User Interface Design Patterns named Atila. In contrast to others, Atila generates a native project depending on the specified platform and it allows modifying the project according to user needs. In another research, Joon et al. [3] analysed the voice recognition vulnerabilities of mobile devices. Also, an attack model named Toilet-time was presented and tests were performed by using BadVoice tool in order to simulate a hands-free cell phone and execute voice commands. Modak et al. [4] developed a desktop application that implements an interface of natural language processing to generate Web pages, the implementation results are promising, nevertheless, voice recognition and speech-to-text process depend on factors such as environment, environmental noise, microphone quality to mention but a few.

Moreover, in the user interface design approach Vittone et al. [5] suggested using *wireframes* which are a simplified representation of an individual screen and help to have a better idea of the screen's layout and identify informative and interactive elements. Beltramelli [6] presented a software based on convolutional neuronal networks, which generates source code for Web and mobile platforms from an image and a previous training model, the software is named pix2code. On the other hand, Patil et al. [7] used LABView to processing voice signals which are submitted via Bluetooth to a robot whom performs movements. The use of LABView permitted to optimize the implementation costs and increase confidence in this system.

In the same context, Tereda et al. [8] developed a robot controlled by voice commands. A series of tests were performed which allowed to knowing voice commands must be improved in order to optimize recognition time. Sidiq et al. [9] presented an Android-based application named Vomma. This application implements voice recognition and execute installed applications on the same device. A set of tests was performed in order to verify Vomma works well. As salient conclusions, factors such as environmental noise, the distance between microphone and user, microphone angle and pronunciation are essential to the proper functioning of voice recognition. Costa et al. [10] used an MDD (Model-Driven Development) for an investigation of user interface named UI Stereotype (User Interface) in order to optimize the development process of Web sites.

Table 1. Comparative table between the related works with the present work.

Article	A	B	C	D	E
Erić et al. [1]	✓	N/A	N/A	N/A	N/A
Cortes et al. [2]	✓	✓	✓	✓	N/S
Modak et al. [4]	✓	✓	N/A	N/S	N/S
Tereda et al. [8]	✓	N/A	N/A	N/A	N/A
Sidiq et al. [9]	✓	N/A	✓	N/A	N/A
Costa et al. [10]	N/A	✓	N/A	✓	N/S
Paschou et al. [12]	✓	✓	✓	✓	✓
Tabibian. [13]	✓	N/A	N/A	N/A	N/A
Furtado et al. [14]	✓	N/A	N/A	N/A	N/A
Sánchez et al. [15]	✓	✓	✓	✓	N/S

A) Voice recognition implementation, **B)** Web-oriented, **C)** Mobile-oriented, **D)** Using user interface patterns, **E)** Health-oriented
N/S: Not Specified
N/A: Not Available

In addition, UI Stereotype was used to build different Web sites with the same purpose (look and feel design and functionality) from a common set of meta-models. UI Stereotype will be implemented for IFML (Interaction Flow Modelling Language) as future work. In [11], a work scheme for developing of software for processing and analyse images based on thesaurus tables' generation was proposed. Paschou et al. [12] developed a generator of health-oriented Android applications whom is used as a Web application and technical knowledge is not required to use it. Tests were performed in order to measure time, human effort and prevent constant development errors, to mention but a few. Tabibian [13] suggested a voice recognition framework in the aerospace area. This framework recognized voice commands in an acceptable time. Nevertheless, the framework functionality is affected by some factors such as pronunciation and environmental noise.

Furtado et al. [14] used Coruja Software to perform a series of tests of software in order to measure user interface performance and interaction with voice recognition. These tests allowed to know that performing some activities with voice recognition is difficult for users. According to the performance time, they concluded that the mouse is efficient while voice recognition is effective. Sánchez et al. [15] developed a software component that generates user interfaces of mobile applications by using pattern recognition and neuronal networks. A use case was proposed in order to generate a login for an Android-based application. As a result, the generation of the application was successful, however, it is necessary to perform further tests and improve patterns design. Also, the generation of user interfaces from a picture or a *mockup* will be included as future work.

Table 1 shows a comparative table among related works and this project, the table indicates the domain for which the work is focused.

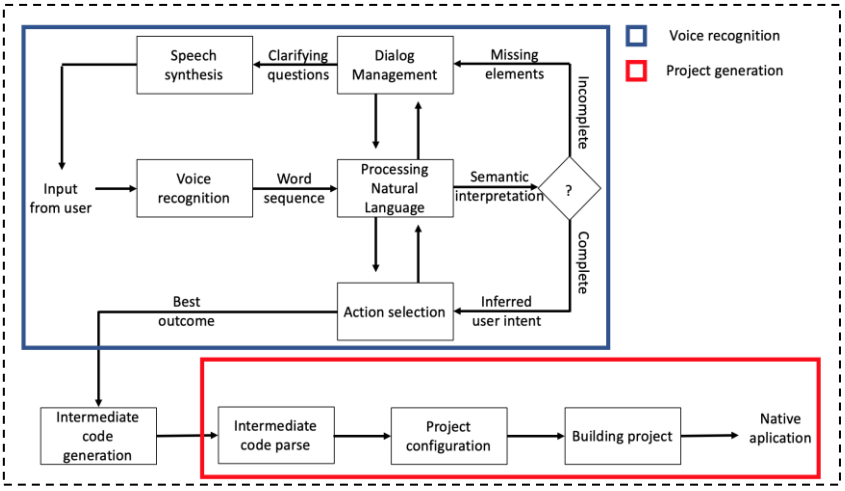

Fig. 1. Process description.

As can be seen in Table 1, the increase of development of voice recognition technologies encourages to develop efficient tools focused on different research areas. Nevertheless, it is important to considerate factors such as 1) environmental noise, 2) microphone quality, 3) pronunciation, 4) distance and angle between microphone and user, among others. Most of these research works conclude that aforementioned factors affect to voice recognition's performance. Moreover, aforementioned works related presented useful techniques, methods and technologies of voice recognition. Conversely, mostly of the related work is not focused on automatic code generation. In conclusion, none of the related works analyzed proposes a formal process for the automatic generation of mobile applications by using voice recognition. Moreover, this analysis allowed to identify a process for code generation and to considerate important factors for developing a Web-based prototype by implementing voice recognition.

3 Description of the Process for Automatic Generation of Medical User Interfaces

The process proposed in this work is divided into two phases which consist in a set of steps where different programming languages, frameworks and voice recognition technologies can be implemented. The interaction between these phases is described below in Fig. 1.

1. **Voice Recognition.** In this phase user's speech is recorded and save it as an audio file in order to turn it into word sequence which is analysed to detect user intention. NLP (Natural Language Processing) and Semantic technologies are used to identify elements of user intention such as application type, layout, user interface design patterns, platform, and information of user and application. When user intention is not completed, the dialog management module ask user

for missing elements that help to clarify user intention. For this purpose, dialog management module uses speech synthesis technologies to improve user interaction. This process is iterative until user intention is completed i.e., when all aforementioned elements were identified. Once user intention is completed an XML-based document is generated. This document describes features of application to be generated by project generation phase.

2. **Project Generation.** phase analyses the XML-based document generated in the previous phase. Then, project configuration module selects all elements that will be integrated in a software application. Finally, source code of this application is generated.

The main modules of the phases are described below with more details:

- **Voice recognition**: In this module, speech is processed to analyse word sequence and turn into text. This text will be processed to determinate its semantic.
- **Natural Language Processing**: Semantic text is determined in this module to detect user intentions. Whether user intention is completed then it will be submitted to action selection module. Otherwise dialog management module is notified to ask user for missing elements by using speech synthesis technology.
- **Action selection**: This module selects action according to user intention. An XML-based document is generated to describe all features of software application. This document is submitted to phase of project generation to generate the software application.
- **Dialog management**: The goal of this module is to ask user for missing element when NLP module does not have all elements to detect user intention.
- **Speech synthesis**: The main function of this module is to convert questions as a text file from dialog management module into an audio file. Then this module plays audio file to make a question or give a result to user.
- **Intermediate code generation**: In this module semantic and syntactic validation of XML-based document is performed to configure the final project.
- **Project configuration**: The project's features are defined in this module. This configuration depends on XML-based document's description.
- **Project building**: Finally, in this module the software application is built according to specified features through the process proposed.

4 Prototype using Voice Recognition

As a proof-of-concept, a Web-based application prototype with voice recognition was developed to segment the process. This prototype helps to have a better idea for developing an agent conversational that will provide a better user experience.

In order to develop this prototype, a workflow was organized into seven steps: 1) Selection of application type, 2) Specification of user interface layout, 3) Selection of patterns design, 4) Specification of platforms, 5) Description of user and application

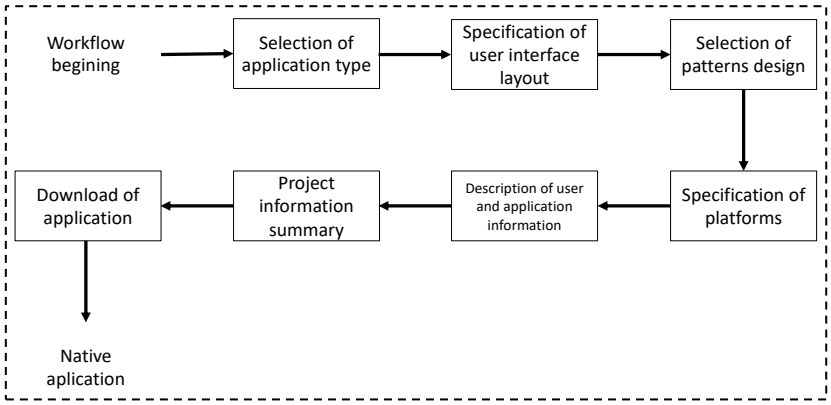

Fig. 2. Prototype's workflow.

information, 6) Project information summary, finally, 7) Download of application, as can be seen in Fig. 2.

The workflow is guided by a virtual assistant that works with voice recognition to improve user performance.

Fig. 3A depicts the first step. In this step, available medical applications are displayed; virtual assistant asks the user for the application type. Depending the context of use and kind of user, there are two types of medical apps: 1) apps for health professionals and 2) apps for the citizens. The user selects an application type by using voice recognition then a medical application is selected automatically. Afterward, a layout of application elements must be specified, a carrousel of different layout templates is displayed as can be seen in Fig. 3B.

After the selection of layout template, a set of user interfaces design patterns according to the type of application is displayed. The most used user interface design patterns in medical apps are, Datalist, Login, Gallery, Video, Splashscreen, Map, Form and Menu, to name a few. Each UIDP (User Interface Design Pattern) is used according to the characteristics and functionality that the application has.

For example, in applications to locate medical services the most used UIDPs are the Maps and Datalist. The Maps has support to the geolocation and the Datalist allows listing the health services units. Another example is the use of Search, Gallery and Dashboard patterns in health encyclopaedias apps, where Search allows searching information about a particular pathology or medication, the gallery shows sets of images, and Dashboard presents the corresponding information.

Next, in Fig. 3C is depicted as the selection of platforms for the software application. User has to use voice commands to specify the platforms. As it can be seen in Fig. 3D information about user and software is specified e.g., author name, company name, author e-mail and author's web site, application's title, application's short name, and application's version.

Before proceeding project generation, the prototype displayed a modal with a summary of the application's information. Finally, the building process of application

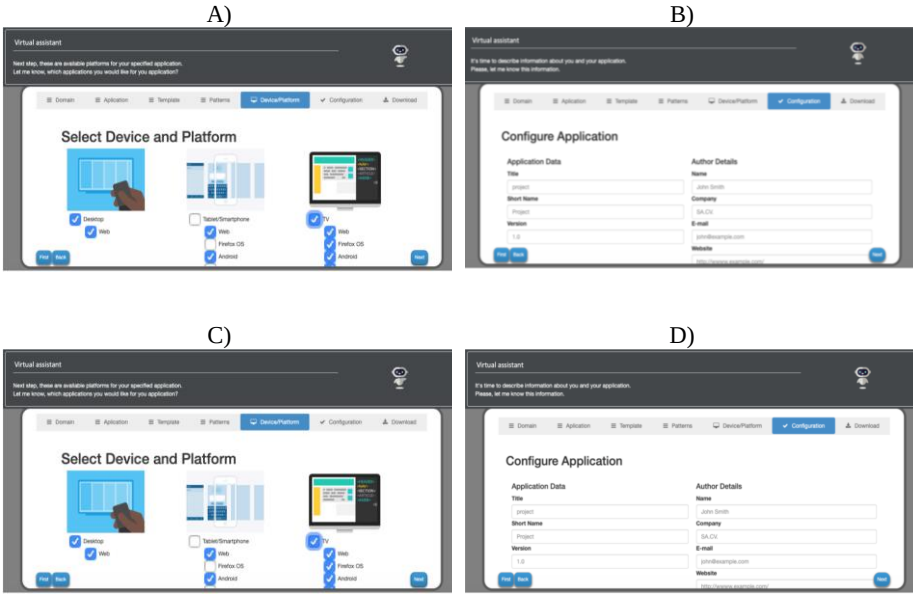

Fig. 3. Web application prototype: A) Selection of application type; B) Selection of layout template; C) Selection of devices and platforms; D) Specification of user and application information.

is executed. When the building process has finished, the software project is downloaded automatically.

5 Conclusion

The software development process has iterative activities that can be automated to optimize time of development, thus, involved people will invest time and effort in activities such as business rules, analysis, and design. Voice recognition is an innovative user interface to have a better interaction with users.

Looking for methods and techniques to improve the software development process, this paper suggests a process support by a workflow in order to generate native projects of different platforms by using voice recognition. This approach applies engineering software techniques and natural language processing which permits to hide some technical aspects to final users by a friendly user interface. Furthermore, medical applications have taken on great importance today, especially in healthcare. Due this, it is important to investigate the use of these applications in order to develop appropriate graphical user interfaces for this domain.

As future work, we are considering to develop an agent conversational to improve user experience. Likewise, research will intent to study additional user interface design

patterns and increase compatibility with other devices to promote application development in the field of healthcare.

Finally, we are evaluating to include more medical applications that allow generating other user interfaces design patterns in the medical domain.

Acknowledgments. This work was supported by Tecnológico Nacional de Mexico (TecNM) and sponsored by the National Council of Science and Technology (CONACYT) and the Secretariat of Public Education (SEP) through PRODEP (*Programa para el Desarrollo Profesional Docente*, for its acronym in Spanish).

References

1. Erić, T., Ivanović, S., Milivojša, S., Matić, M., Smiljković, N.: Voice control for smart home automation: Evaluation of approaches and possible architectures. In: IEEE 7th International Conference on Consumer Electronics - Berlin (ICCE-Berlin), pp. 140–142 (2017)
2. Cortes-Camarillo, C.A., Rosales-Morales, V.Y., Sanchez-Morales, L.N., Alor-Hernández, G., Rodríguez-Mazahua, L.: Atila: A UIDPs-based educational application generator for mobile devices. In: International Conference on Electronics, Communications and Computers (CONIELECOMP), pp. 1–7 (2017)
3. Park Joon Young, Jo Hyo Jin, Samuel Woo, Dong Hoon Lee: BadVoice: Soundless voice-control replay attack on modern smartphones. In: Eighth International Conference on Ubiquitous and Future Networks (ICUFN), pp. 882–887 (2016)
4. Modak, S., Vikmani, S., Shah, S., Kurup, L., Voice driven dynamic generation of webpages. In International Conference on Computing Communication Control and automation (ICCUBEA), pp. 1–4 (2016)
5. Cuello, J., Vittone, J.: Diseñando apps para móviles. José Vittone — Javier Cuello, (2013)
6. Beltramelli, T.: Pix2Code: Generating Code from a Graphical User Interface Screenshot. In: Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems, 3 pp. 1–6 (2018)
7. Patil, S., Abhigna, A., Arpitha, Deepthi, Priyanka: Voice Controlled Robot Using Labview. In: International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), pp. 80–83 (2018)
8. Terada, H., Makino, K., Nishizaki, H., Yanase, E., Suzuki, T., Tanzawa, T.: Positioning Control of a Micro Manipulation Robot Based on Voice Command Recognition for the Microscopic Cell Operation. In: Advances in Mechanism Design II, pp. 73–79 (2017)
9. Sidiq, M., Budi, W.T.A., Sa'adah, S.: Vomma: Android application launcher using voice command. In: 3rd International Conference on Information and Communication Technology (ICoICT), pp. 49–53 (2015)
10. Costa, S.L.d, Neto, V.V.G., Oliveira, J.L.d: A User Interface Stereotype to Build Web Portals. In: 9th Latin American Web Congress, pp. 10–18 (2014)
11. Nedzved, A., Gurevich, I., Trusova, Y., Ablameyko, S.: Software development technology with automatic configuration to classes of image processing problems. Pattern Recognition and Image Analysis, 23(2), pp. 269–277 (2013)
12. Paschou, M., Sakkopoulos, E., Tsakalidis, A.: EasyHealthApps: e-Health Apps Dynamic Generation for Smartphones & Tablets. Journal of Medical Systems, 37(3), pp. 9951 (2013)

13. Tabibian, S.: A voice command detection system for aerospace applications. International Journal Speech Technology, 20(4), pp. 1049–1061, (2017)
14. Furtado, L., Marques, A., Neto, N., Mota, M., Meiguins, B.: IVOrpheus 2.0 - A Proposal for Interaction by Voice Command-Control in Three Dimensional Environments of Information Visualization. In: Human Interface and the Management of Information: Information, Design and Interaction, pp. 347–360 (2016)
15. Sánchez-Morales, L.N., Alor-Hernández, G., Miranda-Luna, R., Rosales-Morales, V.Y., Cortes-Camarillo, C.A.: Generation of User Interfaces for Mobile Applications Using Neuronal Networks. In: New Perspectives on Applied Industrial Tools and Techniques, J. L. García-Alcaraz, G. Alor-Hernández, A. A. Maldonado-Macías, and C. Sánchez-Ramírez, Eds. Cham: Springer International Publishing, pp. 211–231 (2018)