Integration of the ASR Toolkit Kaldi into a Domoticz Home Automation System

Mohamed Aboulkhir[a], Samira Khoulji[b], Reda Jourani[c] and ML Kerkeb[d]
Research Group "Information Systems Engineering", Abdelmalek Essaadi University, 93000,Tetouan,Morocco
[a,b,d,]*National School of Applied Sciences,* [c]*Polydisciplinary Faculty of Tetouan,.*
[a]mohamed.aboulkhir@gmail.com, [b]khouljisamira@gmail.com,[c]reda.jourani@yahoo.fr, kerkebml@gmail.com

ABSTRACT

This paper presents the design and the implementation of an interface between Kaldi, automatic speech recognition toolkit, and a home automation system. This interface is based on Open Platform communication (OPC) protocol. The developed architecture allows the injection of vocal data into a home automation system. Developed in C++, The Kaldi behaves as aclient of the OPC server.

Keywords : Speech recognition, domotics,Kaldi, Smart home, OPC.

1   Introduction

The technological evolution has made it possible to envisage many possibilities in order to improve the quality of life and to support people in their daily life. One solution to provide this assistance on a daily basis is the development of intelligent houses that are equipped with sensors, actuators, controllers and centralized software [1]. Various methods of interaction are being developed in this framework, but one of the most promising is the automatic speech recognition [2].

In our study, we will be interested to integrate Kaldi as an ASR system to home automation system. This integration will be based on OPC protocol. So we will develop a program on Kaldi that will allow him to behave like an OPC client in order to be able to finally integrate it into a home automation system. The expected result will allow a better interaction between the man and the intelligent house.

The paper is organized as follows: we start by describing the ASR software Kaldi. This is followed by describing components of home automation system and OPC protocol .Finally we will present our solution of integration of all these elements.

2 KALDI

Kaldi is an open-source toolkit for speech recognitionwritten in C++ and licensed under the Apache License v2.0[3].The architecture of Kaldi, as described in Figure 1, consists of the following modules:

- External Libraries: Kaldi depends on two external libraries that are also freely available. One is OpenFst for the finite-state framework (FST), and the other is numerical algebra libraries such as BLAS "Basic Linear Algebra Subroutines" and LAPACK "Linear Algebra PACKage".

- Kaldi C++ Library: It contains all the functionalities and different modes of a speech recognition system developed by C ++, which are then called from a scripting language for building and running a speech recognizer.
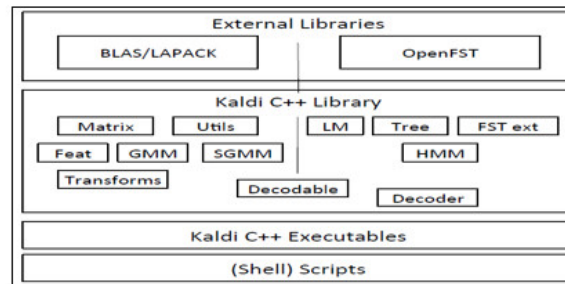- Kaldi C++ Executables
- Scripts

Figure. 1 – Kaldi architecture

The Kaldi library integrates all elements of an ASR system namely: Parameter extraction[4], acoustic model[5], language model[6].

2.1   Parameter extraxtion

Kaldi supports the most used parameters extraction methods. Thus, its Waveform Reading Code is intended to create standard MFCC and PLP functionality by setting reasonable defaults, but leaving available the options that people can modify. Kaldi supports the extraction approaches of most commonly used functionalities [7].

2.2   Acoustic modeling

Kaldi supports standard acoustic modelssuch as GMM, SGMM... and it easily extensible to new kinds of model.

Indeed Kaldi support GMMs with diagonal and full covariance structures. Therefore, it implements a GMM class that is parametrized by the natural parameters. The GMM classes also store the constant term in likelihood computation, which consist of all the terms that do not depend on the data vector. Such an implementation is suitable for efficient log-likelihood computation [8].

For subspace Gaussian mixture models (SGMMs), the toolkit provides an implementation of the approach in which all Hidden Markov Model (HMM) states share the same Gaussian Mixture Model (GMM)structure with the same number of Gaussians in each state[9].

2.3   Language modeling

As described in Kaldi's architecture, it uses a framework based on FST. It is therefore possible to use any language model that can be represented as a FST. Kaldi also offers several toolkits to convert language models into FST [10]

3 Home Automation System

A smart home is a combination of several advanced technologies. It consists of incorporating very small chips, which possess wireless communication, perception and information processing capabilities, in everyday articles. The goal is to create a transparent computing environment for the inhabitant. To be

able to provide services, the system must be able to acquire, process, and transmit information at all times. It must also be able to understand the needs of the user and control the various equipment in a smart way, in order to make the environment more comfortable. Moreover, it must be able to reduce the energy consumption without influencing the habits of the inhabitant [11].

Despite the large number of variations of the concept of home automation systems, there are nevertheless common modules. Figure 2 shows the general organization of a home automation system and of the generic modules constituting it[12].

In recent years these home automation systems have experienced the introduction of speech recognition systems. Thus, several projects have emerged such as the COMPANIONS, Companion Able, ALADIN or PIPIN projects. However, there are still many questions about the real impact, from both the point of view of assistance and the cost of installation, despite real evaluations in test apartments with the PERS, DIRHA projects [13].
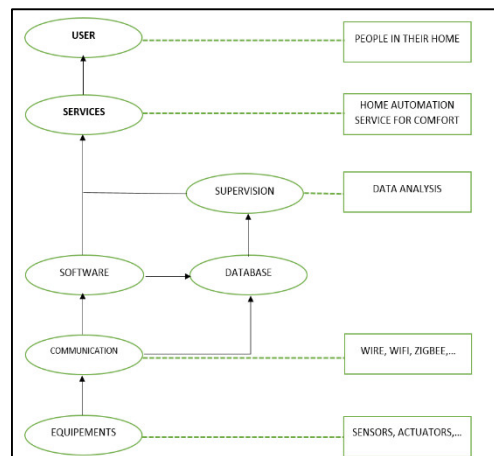

Figure. 2– Home automation system

The objective of our study is to exploit the robustness of the Kaldi system in the automatic recognition of speech in order to integrate it in a home automation system in order to promote its functionality. Our study will be divided into two parts: first, a feasibility study of the solution chosen then a test experiment of robustness of the solution.

4 KALDI Integration into A Domoticz Home Automation System

The aim of our work is to integrate our chosen speech recognition software Kaldi, to a home automation system. For this, we will first propose an architecture of integration, which will be the object of several tests in order to qualify the performances of our model.

4.1 Integration Architecture

In a first step, we propose a communication architecture between Kaldi and a home automation system. Our proposal was based on an architecture based on the OPC client / server technique (Figure 3)[14]. This choice is motivated by:

- The C ++ language supports OPC. Thus, Kaldi can be configured as an OPC client [15].

- The client / server communication can be integrated with the most home automation systems[16].
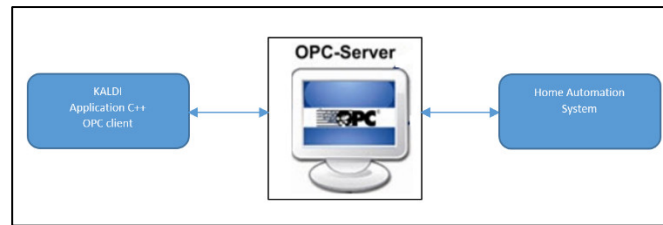

Figure. 3 Integration architecture

4.2 OPC

Open Platform Communication (OPC) is a technique developed in 1995 for the interoperability of industrial systems. OPC was designed to connect Windows applications and hardware and process control software.OPC servers provide a method that allows different software to access data from process control devices, such as a controller. Traditionally, whenever a program required access to a device's data, a custom interface, a driver, had to be written. The objective of the OPC is to define a common interface written once and then reused by any enterprise software. Once an OPC server is written for a particular device, it can be reused by any application capable of acting as an OPC client. An OPC server uses the Microsoft OLE (Component Object Model) technique to communicate with clients[17].

Today, OPC is a registered trademark of the OPC Foundation. The OPC specifications can be divided into two categories:COM / DCOM and Web Services. The first category includes:

- OPC Common (a common specification for all servers)

- OPC Data Access (access to real-time data)

- OPC Alarm and Event

- OPC Historical Data Access

- OPC Batch (batch processing)

The second category consists of a single specification divided into several parts OPC Unified Architecture.

The Unified OPC (UA) architecture has been defined and can be implemented with Java, Microsoft .NET or C, eliminating the need to use a Microsoft Windows computer with earlier versions of OPC. The UA combines the functionality of existing OPC interfaces with new techniques such as XML and Web services [18].

4.3 Feasibility study

In this first phase, we selected a home automation software called Domoticz, it is an open source Home Automation System. The experiment therefore consists in implementing the architecture described above and which is based on the OPC protocol.

4.3.1 Home automation system: Domoticz

Domoticz is an open source and free home automation software, consuming few system resources. It works on several OS (Linux, Windows). It allows to manage switches, detectors (fire, opening,

movement,), sensors (temperature, humidity, winds…), IP camera and to create virtual devices. All the configuration and use is done via a rather successful interface.

It is multi protocols Z-wave, RFX, En-Ocean, ... and it is possible to set up Lua scripts. It integrates with several systems of notification and sending of mail. The access can be secured via login / password with different level of authorization (spectator, user, and administrator). You can also allow only some IPs to connect.

4.3.2   Experiment

Here we report some outcomes about the feasibility of the interconnection between Kaldi and Domoticz through a server OPC. FIG. 4 shows the architecture adopted.
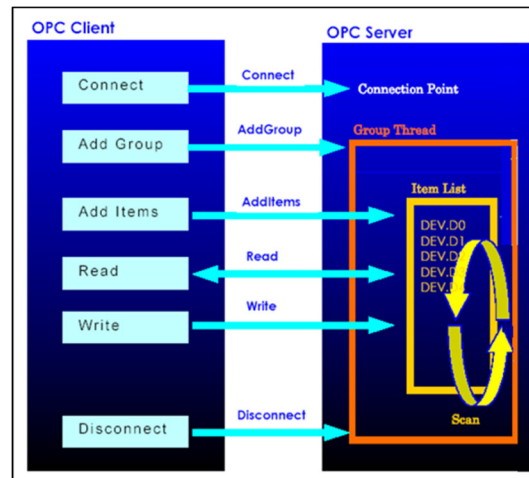

Figure. 4– OPC Architecture

Firstly, it is a question of developing scripts under C ++ to allow Kaldi and Domoticz to become OPC clients, and then the two systems will be connected to an OPC server which will allow d To use the voice data for controlling an output of the home automation system.

For example, we have tested the extinction of light by Domoticz (test at the output) if we pass through Kaldi the phrase "turn off the light".

Through these tests we have proved the feasibility of the interconnection between an ASR system and a home automation system through the developed architecture.

5 Conclusion

We described the architecture, which will allow the integration of a speech recognition system into a home automation system. This architecture was the subject of a feasibility study that showed the possibility of this integration. Development of this solution is continuing and we areworking on testing the robustness of our solution and studying the possibility of being generalized to all systems of home automation and ASR.

REFERENCES

[1] Noury, N., Virone, G., Barralon, P., Ye, J., Rialle, V., & Demongeot, J. (2003, June). New trends in health smart homes. In enterprise networking and computing in healthcare industry, 2003. Healthcom 2003.

[2] Frederic Aman. Reconnaissance automatique de la parole de personnes agées pour les services d'assistance à domicile. Traitement du signal et de l'image. Université Grenoble Alpes, 2014.

[3] The Kaldi Speech Recognition Toolkit, Povey Daniel, Ghoshal Arnab, Boulianne, GillesBurget, LukasGlembek, OndrejGoel, Nagendra, Hannemann, MirkoMotlicek Petr, Qian Yanmin, Schwarz Petr, Silovsky Jan, Stemmer Georg and Vesely Karel, Idiap-RR-04-2012

[4] Richard Dufour. Transcription automatique de la parole spontanée. Informatique [cs]. Université du Maine, 2010.

[5] Mohamed Bouallegue. L'analyse factorielle pour la modélisation acoustique des systèmes dereconnaissance de la parole. Autre [cs.OH]. Université d'Avignon, 2013.

[6] Insect sound recognition based on mfcc and pnn. In Multimedia and Signal Processing (CMSP), 2011 International Conference IEEE

[7] Walker, W., Lamere, P., Kwok, P., Raj, B., Singh, R., Gouvea, E., ... & Woelfel, J. (2004). Sphinx-4: A flexible open source framework for speech recognition

[8] Sakriani Sakti, Oyunchimeg Shagdar, Fawzi Nashashibi, Satoshi Nakamura. Context Awarenessand Priority Control for ITS based on Automatic Speech Recognition. International conference on ITS Telecommunications, Dec 2015, Copenhagen, Denmark. 2015,

[9] D. Povey, L. Burget et al., "The subspace Gaussian mixture model A structured model for speech recognition," Computer Speech & Language,vol. 25, no. 2, pp. 404–439, April 2011.

[10] Fethi Bougares. Attelage de systèmes de transcription automatique de la parole. Ordinateur et societé [cs.CY]. Université du Maine, 2012.

[11] Michel Vacher, Benjamin Lecouteux, Javier Serrano-Romero, Moez Ajili, Fran_cois Portet, etal.. Speech and Speaker Recognition for Home Automation: Preliminary Results. 8th International Conference Speech Technology and Human-Computer Dialogue "SpeD 2015", Oct 2015.

[12] Augusto, J. C., Liu, J., McCullagh, P., Wang, H., & Yang, J. B. (2008). Management of uncertainty and spatio-temporal aspects for monitoring and diagnosis in a smart home. International Journal of Computational Intelligence Systems.

[13] Vacher, J.-F Serignat, Pelayo Menendez-Garcia, D Istrate. FIRST IMPLEMENTATION OF A SOUND/SPEECH REMOTE MONITORING REAL-TIME SYSTEM FOR HOME HEALTHCARE. The 6th International Conference Communications, Jun 2006

[14] Olivier Passalacqua, Eric Benoit, Marc-Philippe Huget, Patrice Moreaux. INTEGRATING OPC DATA INTO GSN INFRASTRUCTURES. IADIS International Conference APPLIED COMPUTING 2008

[15]   Zheng, L., & Nakagawa, H. (2002, August). OPC (OLE for process control) specification and its developments. In SICE 2002. Proceedings of the 41st SICE Annual Conference (Vol. 2, pp. 917-920). IEEE.

[16]   Topalis, E., Orphanos, G., Koubias, S., & Papadopoulos, G. (2000). A generic network management architecture targeted to support home automation networks and home internet connectivity. IEEE Transactions on Consumer Electronics, 46(1), 44-51.

[17]   Liu, J., Lim, K. W., Ho, W. K., Tan, K. C., Tay, A., & Srinivasan, R. (2005). Using the OPC standard for real-time process monitoring and control. IEEE software, 22(6), 54-59.

[18]   Grossmann, D., Bender, K., & Danzer, B. (2008, August). OPC UA based field device integration. In SICE Annual Conference, 2008 (pp. 933-938). IEEE.