Article **ECORFAN-Bolivia Journal**
OPTICAL December 2014 Vol.1 No.1 21-29

Intelligent system voice translator to a bolivian sign language

PACHECO- Carlos†, CAMACHO- Francisco, LABRANDERO- Juan

Universidad Mayor Real y Pontificia de San Francisco Xavier de Chuquisaca, Facultad de Tecnología, Carreras de Ingeniería de Sistemas e Ingeniería en Telecomunicacione, Sucre, Bolivia.

Received January 31, 2014; Accepted July 3, 2014

The present work of scientific research employs a model of neuronal Red Sequential Activation (SAN), which was developed by the research team based at the Space Network Contract, as a mechanism that transforms the result of the noise emission of a word in a symbolic graphical representation for the Bolivian alphabet signs, employing deaf people to communicate in Bolivia.

To materialize this research oriented software components to the capture and processing of words based on mathematical, physical, logical and algorithmic for the study of voice principles were developed. The sound processing and neural networks were implemented in a software component that allows in principle to memorize a set of words, then the noise emission of the same graph is represented graphically through a set of symbols sign language Bolivia .

The investigation detect the basic patterns that form the words, besides Fourier analysis for signals was used, allowing extract features in the frequency domain of the original audio signal, to be jointly used as patterns of a sequence neural network model most advanced voice detection. A study and evaluation of various neural network models allowed, taking into account the sequence of a word is the most important thing to design a neural network capable of recognizing the degree of approximation of a given sequence learned by the network sequence designated by SAN research team.

Neural Networks, Voice Recognition, Voice Translation, Sign Language for the Deaf, Deaf.

Citation: PACHECO- Carlos, CAMACHO- Francisco, LABRANDERO- Juan. Intelligent system voice translator to a bolivian sign language. ECORFAN Journal-Bolivia 2014, 1-1:21-29

† Researcher contributing first author.

Article ECORFAN-Bolivia Journal
OPTICAL December 2014 Vol.1 No.1 21-29

Introduction

The publication of the article entitled "Deaf people require more work and interpreters" developed by Roxana Escobar N., by the newspaper "El Deber" dated September 27, 2010, the points, conclusions and decisions taken are published by "Boliviana Federation of Deaf - Febos "in its global week that took place at that time, of which the following are rescued:

- The Febos considers that the implementation of the law against racism and discrimination, never let you call them deaf, but deaf, considering as a challenge to overcome the biggest barrier that means communicating without losing their identity, (make through sign language), equal access to school education, higher and work.

- In the event the president of the Bolivian Federation of the Deaf, held that the hearing population assumed a deal of discrimination on the deaf community and called sensitization. "They call us deaf, but we have our own identity and culture; we communicate through sight and sign; our eyes are our ears, "he said.

- It was also recognized that at that time and even today there is no current figures on how many deaf people exist at national level, the only thing that counts is the census of 2001 based on only 3 disabilities yielded the following data: 61,145 households reported having one or more persons with disabilities of this Total 26 016 belong to urban areas and 35,129 in rural areas.

- The need for a greater number of professionals and renowned performers was established as the biggest barrier is that they cannot communicate with the hearing population, nor know, because most TV channels do not broadcast the news on language sign and cell use it only to send text messages. The limiting communication is an obstacle to access to employment; they can only access lower and non-hierarchical work.

- Was informed of the opening of the Central Bureau of Investigation Sign Language in Santa Cruz, with a focus on education of deaf children in the country objectives.

- Bolivia recognizes sign language and this adds to the 36 languages that exist in Bolivia.

- In the event became known the provision of 40 million Bolivians who receive this sector of the government, to invest in training programs for the deaf.

According to information obtained from the results of the national meeting of deaf people in the country and the information displayed on the official site Febos(http://www.discapacidadbolivia.org/index.php/extensions56/umadis/organizaciones-civiles14/128-federacion-boliviana-de-sordos-febos), among the most pressing needs were considered that deaf people can access school and higher education equal footing with the hearing population, with few deaf people who have managed to professionalize, with the support of the family that has sought help in the form early; most only reaches primary.

ISSN- 2410-4191

Article
OPTICAL

ECORFAN-Bolivia Journal
December 2014 Vol.1 No.1 21-29

Social inclusion and rehabilitation of these people is important to live in a society without discrimination, where each and every one has the same access to different training and work opportunities. Currently, there are specialized institutions for people with hearing disabilities and verbal communication; However, their capabilities are outpaced by limited infrastructure and the limited number of staff available regarding the number of people who come to these, a situation that definitely conviction, to a large part of the deaf population in Bolivia, to make it look limited in their personal, human, educational and professional development.

The limitations and adverse situations that affect this group of people, who as human beings with every right that society offers them, means and ways to overcome their limitations, and combined with the fact that the various departments have not worked something specific, but they have only sought solutions in traditional mechanisms that have in no way officially improved conditions recognized since 2010 to date, has prompted a research group considering the possibilities that currently provide us TICS settle, find a solution to this problem from the Artificial Intelligence in the area of neural networks, considering the creation of a neuronal model that allows real-time interpreting the sounds emitted by people in relation to letters, words and expressions and transform them immediately in graphical representations of Bolivian sign language.

The interest in science is essential because the research considers various neural models and a set of variables that govern the auditory recognition in real time, regardless of a specific voice, mathematical models - that generate their respective algorithmic language interpretation Bolivian signs, and takes into account the different features that it has.

Studies on speech recognition began between 1950 and 1959, studies that sought to explore the ideas of phonetic - acoustic but without satisfactory results. Since 1960, begins an explosion in speech recognition research, with significant contributions from the Japanese. Undoubtedly the most expressive technique developed at this time was the analysis by Zero Crossing (Zero-Crossing Analysis) getting distinguish between various regions of the auditory stimulus facilitating largely recognition. Another technique, not least, developed in the 60s, is the Dynamic Programming for time alignment, developed by Russian scientist Vintsync, although the basis of the concept was considered rudimentary, more advanced versions of the algorithm are well known widely used in the West that this algorithm to the early 80s Bell labs in a series of experiments began with the goal of creating a speech recognition system irrespective of speech. [Rabiner et ai, 1979].

In 1980 Artificial Neural Networks (ANN) [Morgan, 1995] apply, [Jelinek, 1976] for recognizing speech patterns. The RNA initially was introduced during the 50s; however, their results were not considered satisfactory due to a number of operational problems were corrected as advanced recognition. The 80 is, admittedly, the decade where progress in seeking speech recognition presents its best results. As an example we mention the system developed by the Projects Agency Defense Advanced Research (DARPA-USA) which accurately recognizes thousand words in continuous communication. These are keywords used to control and missile command. The program continues in the 90s emphasizing AI techniques, such as the study area in natural language processing [Bahl et ai, 1989].

ISSN- 2410-4191

Article ECORFAN-Bolivia Journal
OPTICAL December 2014 Vol.1 No.1 21-29

There are several works internationally in relation to speech recognition and their respective transformation to a sign language, they assume different strategies and neural models, with their limitations and peculiarities in relation to the context and culture to which they belong and technological elements used. It should be noted that existing systems are able to recognize only certain words, while this project aims come to recognize any word, because the word will not be our sound unit but rather the phoneme.

In Bolivia it is not known from studies that have addressed this issue, because today is still used humans as translators and communicators between deaf and hearing people either unidirectional or bidirectional.
Developed in relation to the formalization of scientific problem is set as:

"The limited existence of human translators Castilian language Bolivian sign language or vice versa, which is the main cause of the limitations and stagnation in the professional and personal development of deaf people"

This work assumes the formulation of the following hypothesis:

"An Intelligent System interpretation of phonemes corresponding to letters, syllables and words in images representing the Bolivian sign language, will achieve effectively communicate information to deaf people"
It is assumed objective of this study:

"Designing an intelligent interpretation of phonemes corresponding to letters, syllables and words in images representing the Bolivian sign language".

Body of work

This investigation within levels that correspond to the type of technological research that responds to an instance of "Application", considering the need to be actively involved in the care of neglected social sectors is established, the importance of human being beyond the limitations this may have, taking preventive and remedial measures, which allow to project a bright future for the study area, seeking to achieve high levels of deaf people who can access information without the need for a human translator, integrating into society and reviving their possibilities for personal development and progress.

Artifacts and processes considering the use of tools and technology, linguistic, mathematical, logical instruments and means involved essentially assuming as energy and information were built.

The design is assumed as an important and mandatory element in the construction of artifacts and processes, using formalized knowledge of the various fields of knowledge necessary for the proper modeling of the subject matter, scope, issues and the respective proposed solution.

For the experimental nature of this work the method of trial and error was assumed to be essential for validating the results look different.

The work corresponds to a Research since has developed a critical and creative activity, characterized by pose an alternative to the problem formulated.

ISSN- 2410-4191

The method of modeling for software design, which is implemented in C ++ under the paradigm of object oriented programming is also applied.

Regarding the techniques used was considered the following:

- Interviews, which allowed to know detailed information of peculiar circumstances and difficult aspects of consultation through a survey.

- Surveys aimed at identifying information and data that match trends representative stakeholder groups in order to study at different levels.

- The instruments used are:

- Guide interview, in order to properly organize and plan the sequence and order of the questions that were conducted in the respective interviews with people involved with the object of study and constituted an important element in obtaining information and peculiar data and specific.

- Questionnaire, questions organized strategic document that they provide information and data relating to the review or knowledge of various important aspects of this investigation.

- As for the tools that are used, the following were considered:

- High capacity server equipment.

- Computer equipment and laptops.

- Audio and Sound Systems (speakers, microphones, DVD, etc.).

- Software and statistical processing.

- High-level language and interactive environment for the numerical computation, visualization and programming (C ++).

- Motor 2D and 3D graphing.

- Motor audio capture.

In principle it was experimented with various models, following the analysis of the results neural networks are able to develop a new neural network SAN. For example a multilayer perceptron, main alternative to this kind of recognition, applied to image compression, took about 15 min to recognize eight patterns with 1024 entries pattern network, used for handwritten character recognition took only 3 to 5 seconds to recognize 26 patterns with 15000 entries for each pattern. However, these networks do not consider the time variable. Temporary Space Network was considered as an alternative, but this turned out to be very slow, which led us to develop our model

Results

Regarding the research work it is important to mention the results from the components considered in the construction of the technological solution:

Capture

In principle the use of an existing common microphone in most computer systems are considered; however, this amounted to much noise, post-tests and inquiries with other microphones signal was identified that one way or directional microphones which are very sensitive to a single address and relatively deaf to the other.

Article ECORFAN-Bolivia Journal
OPTICAL December 2014 Vol.1 No.1 21-29

Its main advantage is that it allows capture localized sound, these aspects being considered for selecting this type of device to capture the speech signal, thus ensuring noise reduction.

Following several studies established that the minimum quality selected for sampling should assume a frequency of 44100 samples / second, for lower quality led to losing the signal characteristics and high quality precluded the effective implementation of the neural network by the amplitude unnecessary parameters that determine the performance and prompt response of the neural network to overcome this difficulty a software component normalizing signals with high quality developed.

They developed a high-pass filter, characterized by allowing the passage of the higher frequencies and attenuate the lowest, to reduce the effects of noise accompanying the voice signal entered; following the development of the neural model was established that this does not necessary, discarding use.

After trying to analyze the sound in your digitized waveform we concluded that it was not the most suitable among other things because in this way it is in the time domain, alternatively we discovered that convert the signal to the frequency domain we allow a better analysis, this was done with the implementation of the Fast Fourier Transform, an algorithm to find the Fourier transform of a discrete signal.

Then data is normalized in the range of 0 to 1, for further processing in the network.

A major problem is detection of the presence of speech in a relatively noisy environment, termed in literature as "Location of beginning and end of the word" whose resolution is to develop a mechanism for detecting the edges of the word, I get distinguish between speech and silence. The process of determining the beginning and end of a word in a noisy environment is important in many fields of speech process. For example, in a system for isolated word recognition is essential to detect the boundaries of the spoken word to select the region of trace must be post-treated. To detect the onset of speech energy method is applied to the speech signal y [n] with length L, previously normalized. The signal energy is then calculated (for the present investigation the calculation of energy every 256 frequencies did) then compared whether the renewable energy exceeded a certain threshold which was reduced noise each group of samples. For this purpose the following equation was used:

$$E = \sum_{n=r1}^{r2} y(n)$$

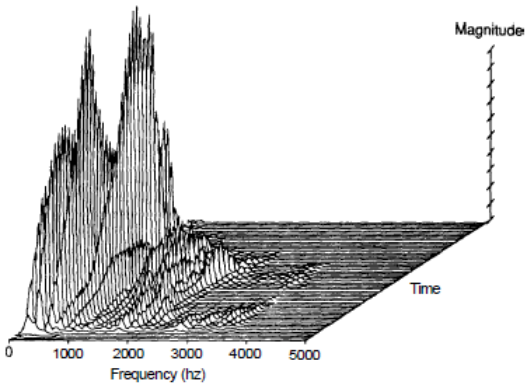
A certain interval between r1 and r2 where r1 took <r2 and travels to and from [1] to find a value that exceeds a preset threshold. At that moment the start position where the threshold is exceeded and the signal is captured trimmed from and [start] to and [1], so that now the beginning of the signal and [start] = y [1] and it extends to and [L].

Sequential Activation Network (SAN)

This neural network is inspired by the Space Temporal Network (STN), which can encode information related to the correlation of spatial frames (frames Spatiotemporal). Then this network shows how you can learn and recall a Temporary Space Frame (SPT).

ISSN- 2410-4191

A simple example would be the frequency spectra that make up a word in a moment of time t.



Frequency spectrogram for the word "Zero"

On the horizontal axis we see the frequency range being analyzed and the vertical axis the magnitude or power; this relates to a time instant (t), ie that each frequency has a magnitude for each time point.

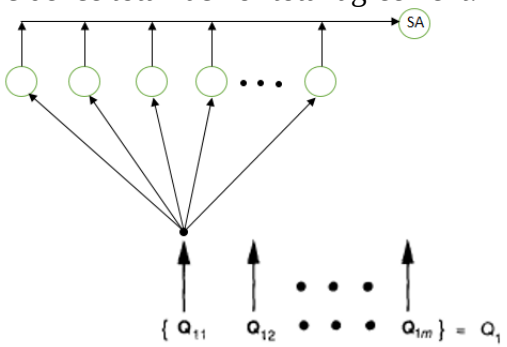
Our model frees the sequence of the variable of time, since we need not learn the duration of a word to search it, just need the sequence of sounds to our previous example "ze" - "ro" is the sequence is the most important feature of a word, since unlike the other loudly.

Here's an example:

o-la;la-o

In the wave word that only has two syllabic sounds, let's try changing the order of these. The result is definitely a different expression. This happens in an example of a word of two syllables. In effect multiplies the words with more sounds, as can state that a word composed of n can generate sounds are a permutation of n different sounds out of n in n, ie n! Different words, if and only if n sounds are different.

Considering the sequence as most important can design a neural network capable of recognizing the degree of approximation of a given sequence stored in the network sequence. The idea is that if you have n frames that make up a sequence, and each frame is learned by a neuron network, by presenting a sequence of the same neurons must be activated one after another and if he touched turned affect a factor positive neuron network output sequence analyzer (SA), otherwise a negative factor. At the end of the sequence presented the SA must return a value between +1 and -1 indicating the degree of coincidence, being +1 and -1 coincidence total lack of total agreement.



Network architecture

You can create a network of networks to learn multiple sequences

Operation

Performance

A unidirectional microphone for capturing sound signal immediately performing a correction signal is used, and then the word is detected. Then the neural network makes sound recognition, returning the code assigned to it, to finally make a translation of the voice Bolivian Sign Language through an interface that starts the process with "Start Sound" button to capture voice, which is displayed in real time and the process is stopped by pressing the "Stop Sound" button.

Article ECORFAN-Bolivia Journal
OPTICAL December 2014 Vol.1 No.1 21-29

Training

The same procedure is for the execution, once you have recorded the training patterns through the microphone by pressing "Start Sound" and "Stop Sound".

At this point all the training patterns are stored in a matrix that will be learned by the neural network with the "Start Training" button. The first time you start the application is in training mode.

Opinion

The result of the experiences and results obtained in the present investigation the following conclusions were drawn:

- The patterns of the speech signal are very complex; they are highly variable, making it difficult choosing features for recognition.

- Tests were made with the voices of several subjects with different characteristics in terms of intonation, speed and idiomatic phrases, inter alia, noted that the prototype could not recognize other voice but with which he was trained, ie recognition speaker dependent voice.

- Rapid prototype with the aim of recognizing vowels, which was the first big step for speech recognition in real time was achieved.

- Speech recognition in real time requires a thorough and detailed study of the speech signal and its characteristics at a basic level. When any sound occurs, we can hear that the particles that lie between us and the source of that sound move, we mean that generate waves, ranging from one place to another.

The number of times that these particles move from one place to another is called frequency and sure we've all heard of the frequency of a sound spectrograph, in this case, is dedicated to measuring the frequency of sounds in a given timeframe. Each sound has a different frequency at all times and that allows us to differentiate on a spectrogram, which sounds are ringing.

- The latest version of the application fails to recognize any word that has been trained, between 10 and 15 at most, with 80% effectiveness. However in order to recognize a large number of words would require parallel processing since a neural network for each word is needed.

Awards / Acknowledgements

For the implementation of this work, it is important to thank:

- A Career Systems engineering and Research Institute TICS, which through its authorities strongly supported this initiative.

- All teachers supported in the research process, strengthening the knowledge in different areas and being always ready to respond to any questions submitted in the development of this work.

- A fellow who selflessly supported in developing tests, sampling and other important activities and required the participation of a large group of people.

- To our parents for their unconditional support and always be with us to meet any challenge that we propose.

Article **ECORFAN-Bolivia Journal**
OPTICAL December 2014 Vol.1 No.1 21-29

References

Basoqain Olabe, X. (2010). *REDES NEURONALES ARTIFICIALES Y SUS APLICACIONES.* Bilbao: Escuela Superior de Ingenieria de Bilbao, EHU.

Fandiño Rodriguez, D. (2005). ESTADO DEL ARTE EN EL RECONOCIMIENTO AUTOMATICO DE VOZ. *Universidad Nacioal de Colombia, Seminario de Investigación*, 13.

Freeman, J., & Skapura, D. (1993). *REDES NEURONALES - Algoritmos, aplicaciones y técnicas de programación.* Estados Unidos. USA: Addison-Wesley / Diaz de Santos.

Haykin, S. (2005). *NEURAL NETWORKS - A Comprehensive Foundation.* India: Pearson, Prentice Hall.

Isasi Viñueka, P., & Glván León, I. (2004). *REDES DE NEURONAS ARTIFICIALES - Un enfoque práctico.* Madrid: Pearson - Prentice Hall.

Kecman, V. (2001). *Learning and Soft Computing - Suport Vector Machines, Neural Networks, and Fuzzy Logic Models.* London: The MIT Press.

Martin del Brio, B., & Sanz Molina, A. (2002). *REDES NEURONALES Y SISTEMAS DIFUSOS.* Mexico: Alfaomega, Ra-Ma.

Mira, J., Delgado, A., Boticario, J., & Díez, F. (1995). *ASPECTOS BASICOS DE LA INTELIGENCIA ARTIFICIAL*. Madrid: SANZ Y TORRES.

Moreno, A., Armengol, E., Béjar, J., Belanche, L., & Cortés, U. (1994). *APRENDIZAJE AUTOMATICO.* Catalunya: Edicions UPC.

Oropeza Rodríguez, J. (2005). *ALGORITMOS Y METODOS PARA EL RECONOCIMIENTO DE VOZ EN ESPAÑOL MEDIANTE SILABAS.* Mexico D.F.: CIC-IPN.

Serrano, A., Soria, E., & Martín, J. (2010). *REDES NEURONALES ARTIFICIALES.* Valencia: Universidad de Valencia.

Varona Martinez, P. (1997). *Escalas, Modelos y Técnicas de Simulación en Neurociencia Computacional.* Madrid: Universidad Autonoma de Madrid - Escuela Técnica Superior de Informática

ISSN- 2410-4191