International Journal of Engineering and Technical Research (IJETR)
ISSN: 2321-0869, Volume-2, Issue-3, March 2014

Speech ability for dumb by pattern recognition using Artificial intelligence

Vineeth V

***Abstract*—Dumbpeople are deprived of speech, which is necessary for communication. Artificial Intelligence helps the dumb people to overcome this difficulty and speak as the normal people do. The process initially makes use of AG-500 Articulograph sensor which is fixed over the lips of the user. The user has to move their lips according to the word they need to pronounce. The sensor calculates the distance between the various reference points assigned on the mouth . The distance is mapped with its equivalent vector representations and gives it to the CM1K chip. . The movie or digital images is sampled into number of frames and this digital neural pulse input is fed to IC CM1K.The input is compared with each neuron containing standard information in the IC and produces appropriate impulses .The output is fed as input to a voice generating software Verbose or Phonweb, which is the final voice of the `dumb person.**

***Index Terms*—Articulography, EMA principle, IC-CM1K, Artificial neurons**

I. INTRODUCTION

Though there are many devices and gadgets for dumb and blind persons, the device for dumbs is limited .One of the main reason is size and cost of devices. The normally use huge number ofsensors, and articulography which are impractical for daily usage. The possibility of reduction in size of them using the chip CM1K is discussed in this paper.

II. ARTIFICIAL NEURONS

A neuron is a reactive memory chip can autonomously evaluate the distance between the incoming vector and reference vector. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Modern neural networks are non-linear statistical data modelling tools. They are usually used to model complex relationships between inputs and outputs or to find patterns in data.

Manuscript received March 15, 2014.

VineethV, Passed B.E. in Electronics and communication engineering, from Thiagarajar college of Engineering in 2014. Main area of research is Radio frequency and microwave systems. Work areas also include Artificial neural networks and VLSI design

III. EMA PRINCIPLE

EMA-Systems operate with EM-fieldgenerating coils which one can regard as antennas for very lowfrequency radio waves. Actually the waves differ in some pointsfrom regular radio waves i.e. they do not interfere with mostmatter, are semi-stationary and do not interfere with each other.All EMA-Systems are determined by the fact that the fieldgenerated by a coil has dipole character and that the receivedsignal becomes zero when transmitter and receiver areperpendicular to each other.Thus,the signal varies not only with distance betweentransmitter and receiver coil,but also with the angle betweentransmitter and receiver axis. While the first effect yields thecalculation of the distance through the "one over cubeddistance"-formula, the second is most unwanted and results in therestriction to a planarmeasurement area.

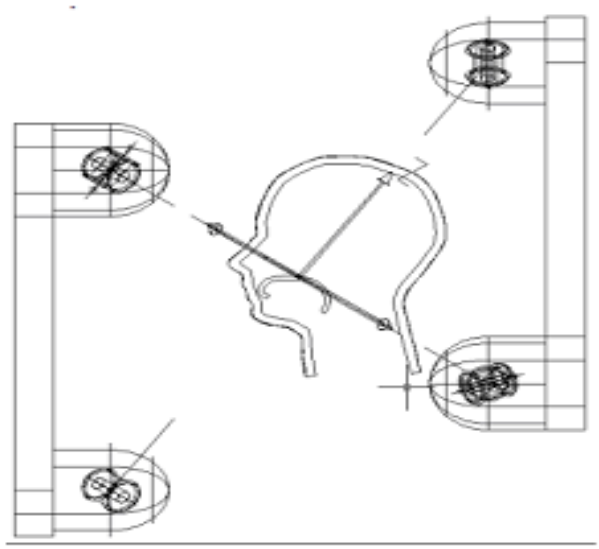
A. *How 3-D-EMA works*

As mentioned before areceiver-coil is a kind of dipole and therefore has five degrees offreedom. These are the three X, Y and Z co-ordinates and the 2angles that describe the alignment of the dipole.In order to determine the position in a three dimensionalarea, all five values must be known. Furthermore, the systemshould be able to measure all directions with the same efficiency.This is realised by the spherical placement of six transmittercoils. Thus there is between every two transmitter coils a rightangle; the placement is absolutely symmetrical. Each coilindicates a value, therefore, in mathematical terms, we have a setof six equations with five unknowns.When aligning the transmitter coils one should consider thathe induced voltage becomes zero when the transmitter andreceiver axis are perpendicular. In this case, no information isavailable concerning the distance between transmitter and receiver.

B. *Working of articulograph*

The new 'helmet'shows schematically theplacement of the 6 transmitter coils in principal. Since the new3-D EMA system has no preferred measurement plane, thestructure can be rotated to fit the subjects head within.The figure shows the final coil mounting structure for the 3-DEMA in a side view with 4 of the 6 transmitter coils visible. Thestructure is build from 2 triangles with a transmitter coil on eachvertex. Not only is this a very straightforward design, but also amechanically robust structureThe receiveroperates with regular monoaxial receiver coils, due to their low costs and small dimensions. The pre-amplified signal isdigitized by a

National Instruments Multi-I/O-card, which ispart of the Receiver-PC and written to disk.It is planned to use this second PC for real-time datapreprocessing and to use a third PCfor position-calculationand analysis of the data.So at the moment we are using up to three PC's on a Windows-NT platform to perform data acquisition and analysisand a fourth PC to move the receiver coil on a predefined path.It is obvious that the device is not very 'handy' at thisstage of development, but there are multiprocessor PC'savailable by now so .The various tasks involved in processing is simplified using IC CM1K.



Final coil mounting structure for the 3-D EMA.

IV. CM1K ARCHITECTURE

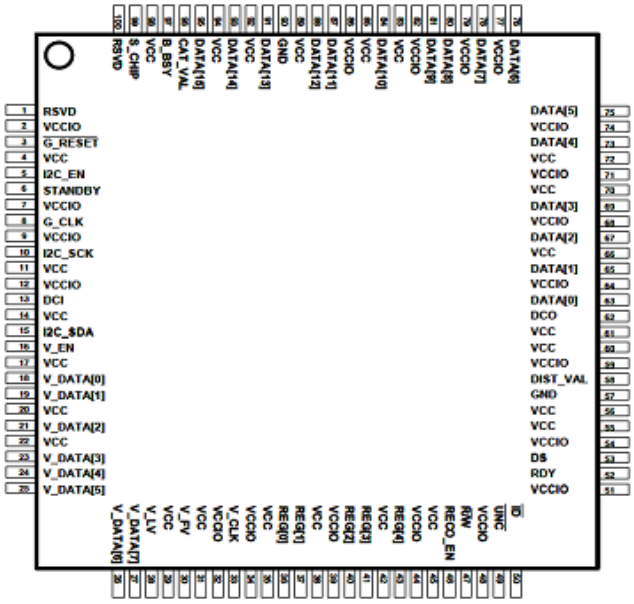
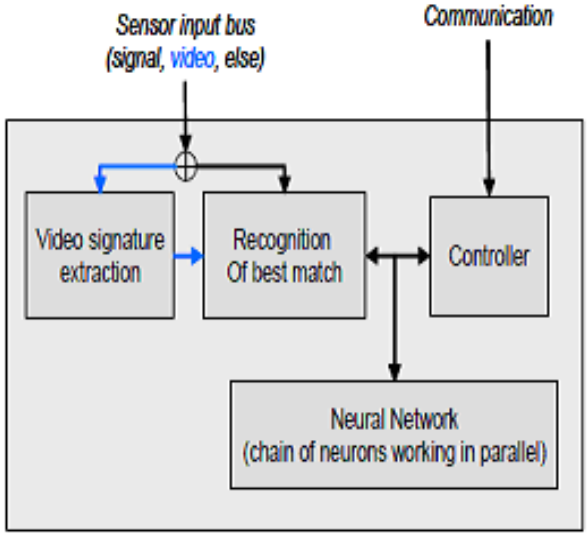


Fig. 1.CM1K pin details

CogniMem is a high-performance pattern recognition chip featuring a network of 1024 neurons expandable through its parallel architecture. Also, the chip embeds a recognition engine ready to classify a digital signal received directly from a sensor. The CogniMem neural network is controlled by a set of 6 global registers and 8 neuron registers. the neurons can learn and recognize input vectors autonomously. If several neurons recognize a pattern (i.e. "fire"), the response of all of them can be retrieved in increasing order of distance (equivalent to a decreasing order of confidence). The first response is therefore the category of the first neuron with a distance register equal to the smallest distance. It is called the "best match" category. If the response of several or all firing neurons is polled, this data can be consolidated to make a more sophisticated decision weighing the cost of uncertainty or else. Note that if an application can use the "best-match" response of the network, CogniMem comes with a recognition engine which is optimized to return this response when an vector is received directly on the digital input bus of the chip. During a learning session, the neurons' behaviour can be tuned be more or less conservative by changing the possible range of their influence fields. The neurons of the network can also be divided into contexts which allows to generate multiple related or non-related knowledge in a same network. It helps in simplifying the process of articulograph Fig 1. shows CM1K pin details.



A. Training and working of CM1K

The dumb's lip movements are detected and then digitalized in AG500 and then it is being fed to CM1K as a input. the input after decoding is compared with the neurons present in the chip via parallel 12C bus. This undergoes the process that is described in the detailed manner below.

CM1K has the choice of two non-linear behaviours of recognition

1.KNN-(K-Nearest Neighbor)(pattern recognition)
2.RBF-(Radial Basis Function)(machine learning)

Aneuron is a reactive memory that can autonomously evaluate the distance between the incoming vector and the vector stored in its memory. if the distance falls within its current influence of field it returns a positive classification .here we need to store the phonetics of each alphabets in the neuron memory.

The neuron determines two types of distances

1.Manhattandistance(L1)(sum of all distances between incoming vector V and the pattern vector P in each neuron)

i.e..,

$$D_{L1} = \Sigma |V_i - P_i|$$

2 .L.SUP(maximum component of distances between V and P)

i.e.,

$$D_{LSup} = \text{Max} | V_i - P_i|$$

At initialization ,the neurons are empty meaning that they do not have any knowledge. Their status is idle for the first one which is Ready –To-Learn.(RTL).As examples are learned by the network the neurons are progressively used to store reference patterns along with their associated categories and become committed.

The state of neuron in the chain can be idle ,RTL or committed .It is defined by the status of its daisy-chain-in (DCI) and daisy chain-out(DCO) lines. The DCO of neurons rises if its DCI is high and its category register is different from 0.As a result, he commitment of neurons is propagated automatically as examples are taught and retained. The RTL neuron also moves along until no more idle neuron is available in the chain

The neural network is composed of M+1+N neurons.

-M committed neurons holding a reference pattern and a category value.

-1 ready to learn (RTL) neuron.

-N idle neurons .

a) KNN Classification:

In this type of classification the input vector is compared to the reference vector stored in the memories of the neuron. The parameters compared are distance and category. The commands used are DIST_ and CAT_.Based on the shortest distance and the category the corresponding neurons firei.e. ,the neurons matching the distance and the category participate to produce the output. The neuron evaluating their shortest distance is considered to have the best match for their pattern and this produces the output for the pattern or signal. this type of classification is considered to be the best type of classification .

b) RBFclassification:

here the data is compared to the neuron memory and the output purely depends on only distance and its field influence. The shorter the distance, the greater its field of influence and so the particular neuron having the shorter distance will be selected for producing the output. This might not work in all situations and so its work is limited to small detection and process and so the former process is considered more efficient for speech synthesis..

B. Vector learning:

If this combined information (vector and category) represents the novelty to the existing neurons , the Ready-to-learn neurons become committed . it stores the instructed category in the category register. Its influence field of the new neuron is set to the smallest distance register of the firing neurons or the Minimum influence field whichever is greater.If there are no firing neurons at all, its influence field is set to the current value of the Maximum influence field.

The next neuron in the chain turns from idle to RTL(ready to learn).

If there are neurons which recognized the vector with acategory other than the instructed category , they automatically reduce their influence field to prevent such erroneous recognition in the future.

C. Loading the training samples:

The training examples can be any number of vectors composed of up to 256 bytes.

The training examples be loaded sequentially into the neurons using the Save and Restore (SR) mode of the CM1K chip. Under this mode, the neurons are passive and writing a neuron register takes one system clock cycle. The following diagram describes the simple sequence of commands loading the training examples to the neurons. The initial Write Forget resets the category of all the neurons to 0. If you do not execute this command, you will be appending the examples after the ones already loaded in the existing committed neurons.

The neurons are then set to the Save and Restore mode by setting bit 4 of the Neuron Status Register (NSR) to 1 and the first neuron of the chain becomes the “Ready-To-Load”. The N bytes of the first example are loaded through a series of Write Component. After the Nth component, the Write Category assigns a value to the Category register of the neuron (default is 1). The latter becomes “Committed” and the next neuron in the chain becomes “Ready-To-Load”. Once all examples have been loaded, the network is returned to its default Learn and Recognize mode by setting bit 4 of the NSR back to 0.

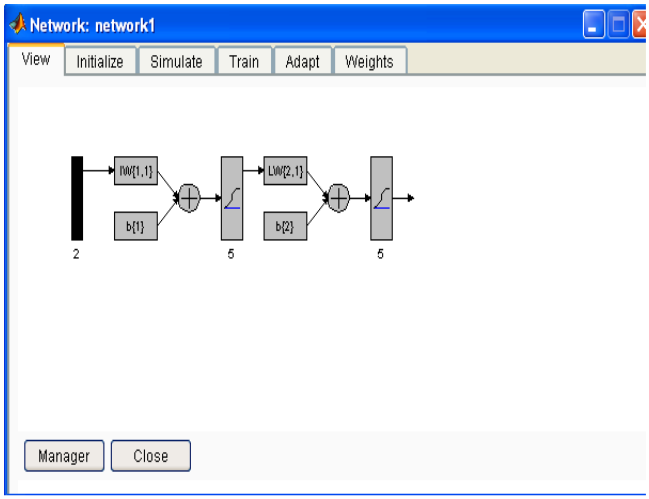
In this way each of the phonetics of all the alphabets and their distances arebeing fed to the neuron memory as digital data input.

Speech synthesis:

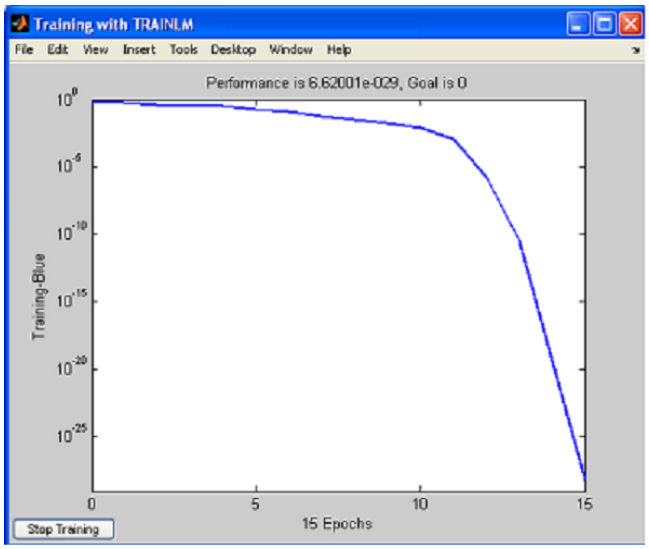
Now the digital impulses from the sensors are being fed to CM1K .after decoding is placed in the global registers .now the vector input is compared with reference vectors in the respective neurons and the right category is selected neurons belonging to the category are fired i.e., these neurons participate for the output. The other neurons are idle i.e., their memory is set to 0 and so called ‘forget’. The neurons that are fired are said to be in the committed state meaning that are knowledge enriched. The appreciable distances are also determined .now the closest distance of the vector with the reference vector is determined to be the neuron of best match among all the firing neurons. This neuron information is the output of the dumb’s lip movement .the output is fed to the system to view it in the form of text. From there the required process can be performed.

The neural network training result using MATlab is given.

The input, hidden and output layers are two, five and five respectively.Training is done using Levenberg-Marquardt algorithm.



Mathematical model of Neural Network



Error vs Epoch graph

V. VOICE CONVERSION

The output of the CM1K chip if fed into text to voice conversion software such as Verbose. It reads text using text to speech technology.The speech can be as compressed mp3 filesThe change voice characteristics setting to change speed, pitch and volume.

VI. ADVANTAGES

The main advantage of the model is reducing the size of articulograph. Ordinary ariculograph consists of three to four computers for processing various kinds of input signals. On using the chip CM1K , the number of computers required is reduced to one. Only the Verbose software is necessarily used, compared to various programs based on C# in the original model. It is convenient to transport the chip rather than the system.

VII. CONCLUSION

Thus “Speech using AI” makes the dumb people, in particular, who are deprived from the beautiful gift of speaking put away their disability. This is a meagre contribution from

our sides to the dumb people of the society and the sweet fruits of success will be tasted by us if it reaches every common man and the word “dumb” is placed in the books of history in the future.

REFERENCES

[1] ARTICULOGRAPH AG 500
http://www.phonetik.uni-muenchen.de/~andi/EMAPage
[2] IC-CM1K
http://general-vision.com/products/chips-and-modules/CM1K-Chip/index.html
[3] Neural networks in Telecommunications ,BenYuhas, Nerwan Ansari.
[4] Neuro Fuzzy and Soft computing , A Computational approach to learning and Machine Intelligence, S.R. Jang, C.T. Sun, E. Mizutani:*FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.
[5] ArtificalIntelligence,,RobertSchlaus

VineethV, Passed B.E. in Electronics and communication engineering, from Thiagarajar college of Engineering in 2014. Main area of research is Radio frequency and microwave systems. Work areas also include Artificial neural networks and VLSI design.