International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

Artificial Intelligence & its Applications for Speech Recognition

G. Harsha Vardhan[1], G. Hari Charan[2]

[1]Student, CSE Department, SRM University, Chennai, India

[2]Student, EEE Department, SRM University, Chennai, India

Abstract: ***AI is behaviour of a machine, which, if performed by a human being, would be called intelligence. It makes machines smarter and more useful and is less expensive than natural intelligence. Artificial intelligence (AI) for speech recognition involves two basic ideas. First, it involves studying the thought processes of human beings. Second, it deals with representing those processes through machines (like computers, robots, etc.).One of the main benefits of speech recognition system is that it lets user do other works simultaneously. The user can concentrate on observation and manual operations, and still control the machinery by voice input commands.***

Keywords: Artificial Intelligence, Speaker Efficiency, Speech Recognition, Speaker dependent System, NLP.

1. Introduction

The IBM sponsored a summer workshop at Dartmouth, New Hampshire, USA, in June 1956.The participants in the conference were Claude Shannon from IBM, John McCarthy from MIT along with Herbert Simon and Allen Newell, and many others. The important outcome of the conference was the coining of the term Artificial Intelligence (AI),which encompasses many concepts and methods deployed by researchers in diverse fields of computation and cognition.

The aim of Artificial Intelligence (AI) is to make computers perform tasks that humans tend to be good at. In a sense it is an investigation into the attempt to make computers behave in a smarter manner.

Natural language processing (NLP) refers to artificial intelligence methods of communicating with a computer in a natural language like English. The main objective of a NLP program is to understand input and initiate action.

2. Definition

The art and science of bringing learning, adaptation and self-organization to the machine is the act of Artificial Intelligence.

AI is the study of how to make computers to do things at which, people are better at the moment. It is the branch of computer science that deals with symbolic, non-algorithmic methods of problem solving.

3. History

Work started soon after World-War II. The actual name Artificial Intelligence was coined by John McCarthy in the 60's.The players who demonstrated the capabilities and techniques of AI systems were Alan Turing,Warren McCulloch,Claude Shannon, Norbert Wiener, NewelL, Simon, etc. to name a few.

The development period from its origin to up-to-date has been divided into nine phases with an interval of five years.

1956-61: Chomsky's Linguistic model processing was a remarkable event in this period. Herbert Gelerriter designed the first written AI program for geometry at IBM in 1959.

1962-67: In this period Frank Rosen proposed the concept of perception in Widrow's concept for Artificial Nueral Networks (ANN).

1968-73: The most important contribution to AI, General Problem Solver for IPL, a program which simulated neural activities of human beings in solving complex problems came in this period.

1974-79: Many expert system programs like GPS 5 at CMU,AES building tool, Caduses for diagnosis of internal medicine were developed in this period.

1980-85: In this period expert system tools & programs were developed. The most important develpoment was PROLOG as AI programming language by Clockin.

1986-91: The distributed artificial intelligence concepts were incorporated in multi agent systems. The probabilistic reasoning method in intelligent system appeared in this period.

1992-97: The concept of cooperation, coordination and conflict resolution in multi agent system was introduced in this period. Agent Building and Learning Environment(ABLE) was developed which focused on building hybrid intelligent agents for both reasoning and learning.

1998-2003: Blue Deep a chess playing computer developed by IBM on May 11,1997 won a six-game match by two wins to one with three draws against world champion Garry Kasparor. Robotics in game playing and surgery marks the achievements in this period.

2004 and onwards: Knowledge discovery and vision system for biometric and automated object regulated in supermarkets are important milestones. Programs that are helpful for diagnosis of errors and suggestions for corrective actions are yet to be developed.

4. Objectives of AI

i) Understanding Human Cognition:
It means to try to obtain deep knowledge of human memory, problem solving abilities, learning, decision making, etc.

ii) Cost-Effective Intelligent Amplification:
Replaces humans in intelligent tasks. Builds system to help humans think better, faster, deeper. For example, system to help GP diagnose disease.

iii) Superhuman Intelligence:
Build programs to exceed human intelligence.

iv) Coherent Discourse:
Communicates with people using natural language.Carries out an intelligent dialogue (c.f. Turing test).

5. Speaker Independency

The speech quality varies from person to person.It is therefore difficult to build an electronic system that recognizes everyone's voice. By limiting the system to the voice of a single person, the system becomes not only simpler but also more reliable. The computer must be trained to the voice of that particular individual. Such a system is called Speaker-Dependent system.

Speaker Independent software is used to recognize anyone's voice, so no training is required. But generally Speaker Independent software is less accurate than Speaker Dependent Software. Speech recognition engines (speaker independent) deal by limiting the grammars they use. It is more commonly found in telephone applications.

Speaker Dependent software works by learning the unique characteristics of a single person's voice, in a way similar to voice recognition. So new users must first train the software by speaking to it, so the computer can analyze how the person talks. Training can be tedious process, but the system has the advantage of using templates that refer only to the specific user and not some vague average voice. The result is language independence. The drawback is that the Speaker Dependent system must do more than simply match incoming speech to the templates. It must also include resources to create those templates.

6. Environmental Influence

Real applications demand that the performance of the recognition system be unaffected by changes in the environment. However, it is a fact that when a system is trained and tested under different conditions, the recognition rate drops unacceptably. We need to be concerned about the variability present when different microphones are used in training, testing and specifically during development of procedures. Such care can significantly improve the accuracy of recognition systems that use desktop microphones.

Acoustical distortions can degrade the accuracy of recognition systems. Obstacles to robustness include additive noise from machinery, competing talkers, reverberation from surface reflections in a room and spectral shaping by microphones and the vocal tracts of individual speakers. These sources of distortions fall into two complementary classes; additive noise and distortions resulting from the convolution of the speech signal with an unknown linear system.

A number of algorithms for speech enhancement have been proposed.These include the following:

1) Spectral subtraction of DFT coefficients
2) MMSE techniques to estimate the DFT coefficients of corrupted speech
3) Spectral equalization to compensate for convoluted distortions
4) Spectral subtraction and spectral equalization.

Although relatively successful,all these methods depend on the assumption of independence of the spectral estimates across frequencies.Improved performance can be got with an MMSE estimator in which correlation among frequencies is modeled explicitly.

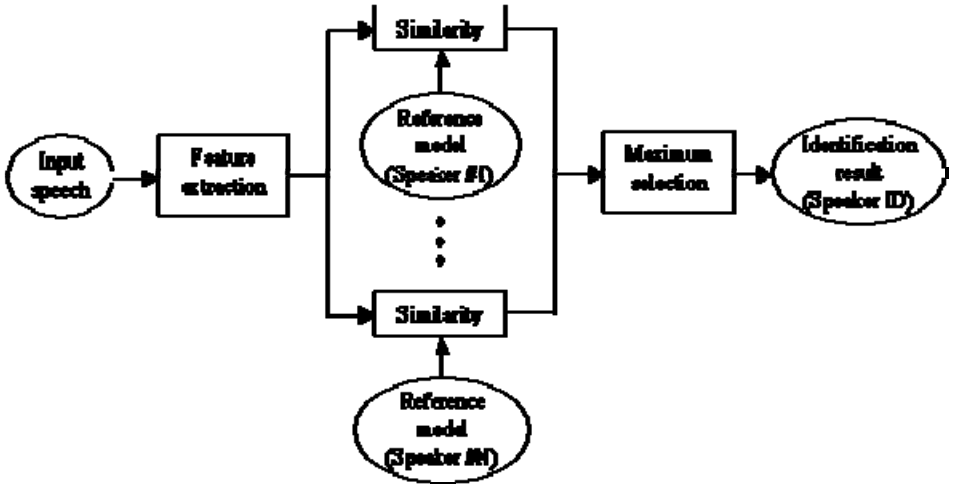
7. Speaker-Specific Features

Speaker identity correlates with the physiological and behavioural characteristics of the speaker. These characteristics exist both in the vocal tract characteristics and in the voice source characteristics,as also in the dynamic features spanning several segments. The most common short-term spectral measurements currently used are the spectral coefficients derived from the Linear Predictive Coding (LPC) and their regression coefficients.A spectral envelope reconstructed from a truncated set of spectral coefficients is much smoother than one reconstructed from LPC coefficients.

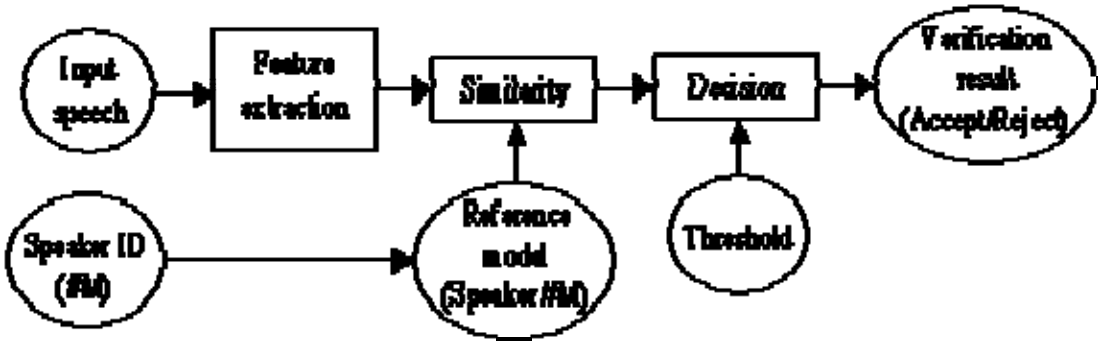
Therefore, it provides a more stable representation from one repetition to another of a particular speaker's utterances.

As for the regression coefficients, typically the first and second order coefficients are extracted at every frame period to represent the spectral dynamics.

These coefficients are derivatives of the time function of the spectral coefficients and are called the delta and delta-delta-spectral coefficients respectively.



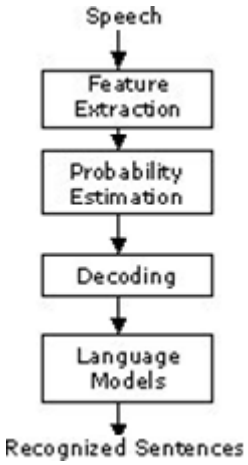
(a)Speaker Identification



(b)Speaker Verification

8. Speech Recognition

The user communicates with the application through the appropriate input device i.e. a microphone. The Recognizer converts the analog signal into digital signal for the speech processing.A stream of text is generated after the processing.This source-language text becomes input to the Translation Engine,which converts it to the target language text.



9. Salient Points in the development of AI

Salient points in the development of AI depicting the contributors, their contributions are shown below:

Name Contribution

1) Samuel(1952) His program quickly learnt to play better than the creator itself
2) McCarthy(1956) High level language LISP
3) Gelernter et al(1959) Geometry theorem power
4) Bombrow(1967) Student-algebra story problem
5) Evans(1968) Analogy geometry problems for IQ tests
6) Winograd(1969) Natural language understanding
7) Winston(1970) Learning theory
8) Hopfield(1982) Hopfield network for optimization in statistical mechanics
9) Rumelhart(1986) ANN method for memory in distributes processing
10) Laird(1987) Rational agent SOAR: Best known architecture of agent
11) Pearl(1988) Bayesian network for uncertain reasoning theory to AI
12) Kaelbling(1996) Robotic control hip replacement prosthesis crossword problem
13) Campbell(2002) Hardware chip for game playing,3d robotics, web intelligence

10. Applications

Many medical images can now be automatically interpreted, from plane X-rays through to more complex images like CT, MRI scans.

One of the main benefits of speech recognition system is that it lets user do other works simultaneously. The user can

concentrate on observation and manual operations, and still control the machinery by voice input commands.

Voice recognition could also be used on computers for making airline and hotel reservations. A user requires simply to state his needs, to make reservation, cancel a reservation, or make enquiries about schedule.

Another major application of speech processing is in military operations. Voice control of weapons is an example. With reliable speech recognition equipment, pilots can give commands and precise information to the computers by simply speaking into their microphones - they don't have to use their hands for this purpose.

Robotics is one of the prime areas of AI applications.AI methodology is applicable to robotics in two ways: One is design and control of robot and the other is application of robots to various fields such as manufacturing, mining, medicine (surgery).Humanoid robots imitating many activities of human motor actions, as well as emotions, pose many challenging facts to AI community.

Application of robotic system ,in places where human movement is restricted as in nuclear power plant or mines, works intelligently in such places and atmosphere where hazards prevail for human with high probability. Reasoning with uncertainties and mechanization enhances the inetlligence in robot.

AI is used mainly for product selection and recommendations, negotiations, auctions, solving real-world scheduling problems and enhancing scalability and bundling and pricing goods.

11. Ultimate Goal

The ultimate goal of the Artificial Intelligence is to build a person or more humbly, an animal.

12. Conclusion

Speech recognition will revolutionize the way people do business around the web and will ultimately differentiate world class e-business. Speech recognition & voice XML clearly represent the next wave of the web.

By using this speaker recognition technology we can achieve many uses. This technology helps physically challenged skilled persons. These people can do their works by using this technology without pushing any buttons. This ASR technology is also used in military weapons and in Research centers. Now a days this technology was also used by CID officers. They used this to trap the criminal activities.

References

[1] www.google.co.in/Artificial intelligence for speech recognition
[2] www.google.com
[3] www.howstuffworks.com
[4] www.ieeexplore.ieee.org
[5] www.wikipedia.org
[6] Dr. Prof Videh Kishore thakur, Conference on Advances in Communication & Control Systems 2013(CAC2S 203)
[7] Rajendra Akerkar, Introduction to Artificial Intelligence
[8] R. B. Mishra, Artificial Intelligence.
[9] www.scribd.com/Artificial-Intelligence-for-Speech-Recognition