

Artificial Passenger

Ms. Anjali Dadhich, Mrs. Neelam Sharma

Assistant Professor, Apex Institute of management and science, Rajasthan, India.

Assistant Professor, Banasthali Vidhyapeeth University, Rajasthan, India.

ARTICLE INFO

Received 1 April, 2015
Accepted 15 April, 2015

Corresponding Author:

Ms. Anjali Dadhich

1 Assistant Professor, Apex
Institute of management and
science, Rajasthan, India

ABSTRACT

An artificial passenger (AP) is a device that would be used in a motor vehicle to make sure that the driver stays awake. IBM has developed a prototype that holds a conversation with a driver, telling jokes and asking questions intended to determine whether the driver can respond alertly enough. Assuming the IBM approach, an artificial passenger would use a microphone for the driver and a speech generator and the vehicle's audio speakers to converse with the driver. The conversation would be based on a personalized profile of the driver. A camera could be used to evaluate the driver's "facial state" and a voice analyzer to evaluate whether the driver was becoming drowsy. If a driver seemed to display too much fatigue, the artificial passenger might be programmed to open all the windows, sound a buzzer, increase background music volume, or even spray the driver with ice water. Conversational Interactivity for Telematic speech systems can significantly improve a driver-vehicle relationship and contribute to driving safety. But the development of full fledged Natural Language Understanding (NLU) for CIT is a difficult problem that typically requires significant computer resources that are usually not available in local computer processors that car manufacturer provide for their cars. To address this, NLU components should be located on a server that is accessed by cars remotely or NLU should be downsized to run on local computer devices (that are typically based on embedded chips).

Keywords: *telematics, microphone, voice analyzer.*

©2014, IJICSE, All Right Reserved.

INTRODUCTION

The AP is an artificial intelligence-based companion that will be resident in software and chips embedded in the automobile dashboard. The heart of the system is a conversation planner that holds a profile of you, including details of your interests and profession. A microphone picks up your answer and breaks it down into separate words with speech-recognition software. A camera built into the dashboard also tracks your lip movements to improve the accuracy of the speech recognition. A voice analyzer then looks for signs of tiredness by checking to see if the answer matches your profile. Slow responses and a lack of intonation are signs of fatigue. This research suggests that we can make predictions about various aspects of driver performance based on what we glean from the movements of a driver's eyes and that a system can eventually be developed to capture this data and use it to alert people when their driving has become significantly impaired by fatigue.

The natural dialog car system analyzes a driver's answer and the contents of the answer together with his voice

patterns to determine if he is alert while driving. The system warns the driver or changes the topic of conversation if the system determines that the driver is about to fall asleep. The system may also detect whether a driver is affected by alcohol or drugs.[1]

A .Field of Invention

The present invention relates to a system and method for determining three dimensional head pose, eye gaze direction, eye closure amount, blink detection and flexible feature detection on the human face using image analysis from multiple video sources. Additionally, the invention relates to systems and methods that make decisions using passive video analysis of a human head and face.

B. Background of Invention

Early techniques for determining head-pose used devices that were fixed to the head of the subject to be tracked. For example, reflective devices were attached to the subjects head and using a light source to illuminate the reflectors, the reflector location was

determined. Wearing a device of any sort is clearly a disadvantage as devices are generally intrusive and will affect a user's behavior, preventing natural motion or operation. However, the technique is prone to error in conditions of lighting variation and is therefore unsuitable for use under natural lighting conditions.

II.FUNCTION OF ARTIFICIAL PASSENGER

A. Voice Control Interface

One of the ways to address driver safety concerns is to develop an efficient system that relies on voice instead of hands to control Telemetric devices. CIT speech systems can significantly improve a driver-vehicle relationship and contribute to driving safety. But the development of full fledged Natural Language Understanding (NLU) for CIT is a difficult problem that typically requires significant computer resources that are usually not available in local computer processors that car manufacturers provide for their cars. To address this, NLU components should be located on a server that is accessed by cars remotely or NLU should be downsized to run on local computer devices (that are typically based on embedded chips). Our department is developing a "quasi-NLU" component - a "reduced" variant of NLU that can be run in CPU systems with relatively limited resources. In our approach, possible variants for speaking commands are kept in special grammar files (one file for each topic or application).When the system gets a voice response, it searches through files (starting with the most relevant topic). If it finds an appropriate command in some file, it executes the command. Otherwise the system executes other options that are defined by a Dialog Manager (DM). The DM component is a rule based sub-system that can interact with the car and external systems (such as weather forecast services, e-mail systems, telephone directories, etc.) and a driver to reduce task complexity for the NLU system.[5]

B. Embedded Speech Recognition

Logically a speech system work as follows: speech is measured in terms of frequency and then sampled at 16,000 times per second. We empirically observed that noise sources, such as car noise, have significant energy in the low frequencies and speech energy is mainly concentrated in frequencies above 200 Hz. Filters are therefore placed in the frequency range [200Hz – 5500 Hz], Discarding the low frequencies in this way improves the robustness of the system to noise.

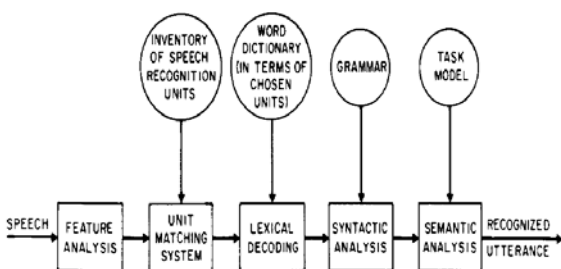


Fig. 1 Embedded speech recognition block diagram [6]

Any sound is then identified by matching it to its closest entry in the database of such graphs, producing a number, called the feature number that describes the sound. Unit matching system provides likelihoods of a match of all sequences of speech recognition units to the input speech. These units may be phones, diphones or whole word units. Lexical decoding constraints the unit matching system to follow only those search paths sequences whose speech units are present in a word dictionary. Apply a grammar so the speech recognizer knows what phonemes to expect. Figure out which phonemes are spoken, this is a quite difficult task as different words sound differently as spoken by different persons. Also, background noises from microphone make the recognizer hear a different vector. Thus a probability analysis is done during recognition by HMM (Hidden Markov Model).[6]



Fig. 2 Embedded speech recognition device

C. Driver Drowsiness Prevention

Fatigue causes more than 240,000 vehicular accidents every year. Driving, however, occupies the driver's eyes and hands, thereby limiting most current interactive options. Among the efforts presented in this general direction, the invention suggests fighting drowsiness by detecting drowsiness via speech biometrics and, if needed, by increasing arousal via speech interactivity. It is a common experience for drivers to talk to other people while they are driving to keep themselves awake. The purpose of Artificial Passenger part of the CIT project at IBM is to provide a higher level of interaction with a driver than current media, such as CD players or radio stations, can offer.

This is envisioned as a series of interactive modules within Artificial Passenger that increase driver awareness and help to determine if the driver is losing focus. This can include both conversational dialog and interactive games, using voice only. The scenarios for Artificial Passenger currently include: quiz games, reading jokes, asking questions, and interactive books. In the Artificial Passenger paradigm, the awareness-state of the driver will be monitored, and the content will be modified accordingly. Drivers evidencing fatigue, for example, will be presented with more stimulating content than drivers who appear to be alert. Most well known emotion researchers agree that arousal (high,

low) is the fundamental dimensions of emotion. Arousal reflects the level of stimulation of the person as measured by physiological aspects such as heart rate, cortical activation, and respiration. For someone to be sleepy or fall asleep, they have to have a very low level of arousal. There is a lot of research into what factors increase psychological arousal since this can result in higher levels of attention, information retention and memory. We know that movement, human voices and faces (especially if larger than life), and scary images (fires, snakes) increase arousal levels. We also know that speaking and laughing create higher arousal levels than sitting quietly. Arousal levels can be measured fairly easily with a biometric glove (from MIT), which glows when arousal levels are higher (reacts to galvanic skin responses such as temperature and humidity).[2]

D. Workload Manager

It is a key component of driver Safety Manager. An object of the workload manager is to determine a moment-to-moment analysis of the user's cognitive workload. It accomplishes this by collecting data about user conditions, monitoring local and remote events, and prioritizing message delivery. There is rapid growth in the use of sensory technology in cars. These sensors allow for the monitoring of driver actions (e.g. application of brakes, changing lanes), provide information about local events (e.g. heavy rain), and provide information about driver characteristics (e.g. speaking speed, eyelid status).



Fig. 3 Conditions sensor device

There is also growing amount of distracting information that may be presented to the driver (e.g. phone rings, radio, music, e-mail etc.) and actions that a driver can perform in cars via voice control. The workload manager is closely related to the event manager that detects when to trigger actions and/or make decisions about potential actions. The system uses a set of rules for starting and stopping the interactions (or interventions). It controls interruption of a dialog between the driver and the car dashboard (for

example, interrupting a conversation to deliver an urgent message about traffic conditions on an expected driver route). It will make decisions on whether the driver needs additional stimuli and on what types of stimuli should be provided (e.g. verbal stimuli via speech applications or physical stimuli such as a bright light, loud noise, etc.) and whether to suggest to a driver to stop for rest. The system permits the use and testing of different statistical models for interpreting driver answers and information about driver conditions. The driver workload manager is connected to a driving risk evaluator that is an important component of the Safety Driver Manager.

The goal of the Safety Driver Manager is to evaluate the potential risk of a traffic accident by producing measurements related to stresses on the driver and/or vehicle, the driver's cognitive workload, environmental factors, etc. The important input to the workload manager is provided by the situation manager whose task is to recognize critical situations. It receives as input various media (audio, video, car sensor data, network data, GPS, biometrics) and as output it produces a list of situations. Situations could be simple, complex or abstract.

III. WORKING OF ARTIFICIAL PASSENGER

The main devices that are used in this artificial passenger are:-

- Eye Tracker
- Voice recognizer or speech recognizer

A. How does tracking device work?

Collecting eye movement data requires hardware and software specifically designed to perform this function. Eye-tracking hardware is either mounted on a user's head or mounted remotely. Both systems measure the corneal reflection of an infrared light emitting diode (LED), which illuminates and generates a reflection off the surface of the eye. This action causes the pupil to appear as a bright disk in contrast to the surrounding iris and creates a small glint underneath the pupil. It is this glint that head-mounted and remote systems use for calibration and tracking.

B. Algorithm for monitoring head/eye motion for driver alertness with one camera

The invention robustly tracks a person's head and facial features with a single on-board camera with a fully automatic system, that can initialize automatically, and can reinitialize when it need's to and provide outputs in real time. The system can classify rotation in all viewing direction, detects' eye/mouth occlusion, detects' eye blinking, and recovers the 3Dgaze of the eyes. In addition, the system is able to track both through occlusions like eye blinking and also through occlusion like rotation. Outputs can be visual and sound alarms to the driver directly.

The representative image is shown below:

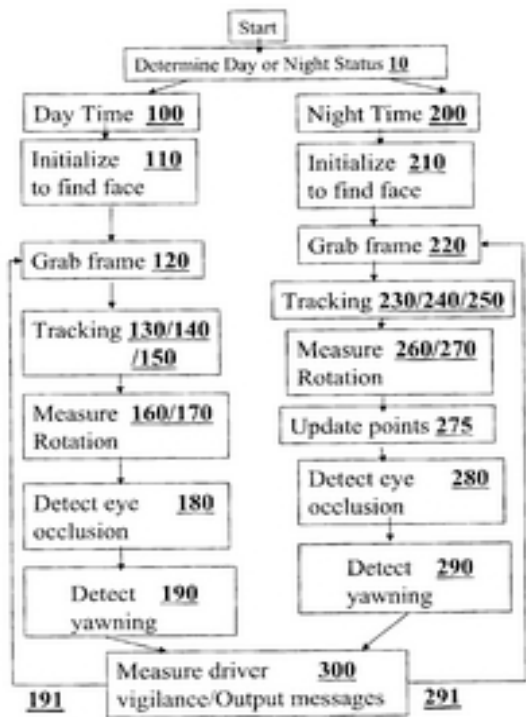


Fig. 4 Representative Image [5]

The system will first determine day or night status. It is nighttime if: a camera clock time period is set for example to be between 18:00 and 07:00 hours. Alternatively, day or night status can be checked if the driver has his night time driving headlights on by wiring the system to the headlight controls of the vehicle. Additionally, day or night status can be set if the intensity of the image, is below a threshold. In this case then it must be dark. If day time is determined then the left side of the flow chart depicted in fig. 4 will follow then first initialize to find face. If night time is determined, then the right flow chart series of steps occurs, by first initializing to find the face. Next a frame is grabbed from the video output. Tracking of the lip corners and eye pupils is performed. Measure rotation and orientation of face. The feature points are corrected if necessary. Eye occlusion as blinking and eye closure is examined. Determining if yawning is occurring is done. The rotation, eye occlusion and yawning steps in formation is used to measure the driver's vigilance.[5]

The system contains:

- A single camera within a vehicle aimed at a head region of the driver.
- Means for simultaneously monitoring head rotation, yawning and full eye occlusion of the driver with said camera, the head rotation including nodding up and down, and moving left to right, and the full eye occlusion including eye blinking and complete eye closure, the monitoring means includes means for determining left to right rotation and the up and down nodding from examining approximately 10 frames from approximately 20 frames; and

- Alarm means for activating an alarm in real time when a threshold condition in the monitoring means has been reached, whereby the driver is alerted into driver vigilance.

C. Method for detecting driver vigilance comprises the following steps

- Aiming a single camera at a head of a driver of a vehicle, detecting frequency of up and down nodding and left to right rotations of the head within a selected time period of the driver with the camera.
- Determining frequency of eye blinking and eye closings of the driver within the selected time period with the camera.
- Determining the left to right head rotations and the up and down head from examining approximately 10 frames from approximately.
- Determining frequency of yawning of the driver within the select period with the camera.
- Generating an alarm signal in real time if a frequency value of the number of the frequency of the up and down nodding, the left to right rotations, the eye blinking, the eye closings, the yawning exceed a selected threshold value.[4]

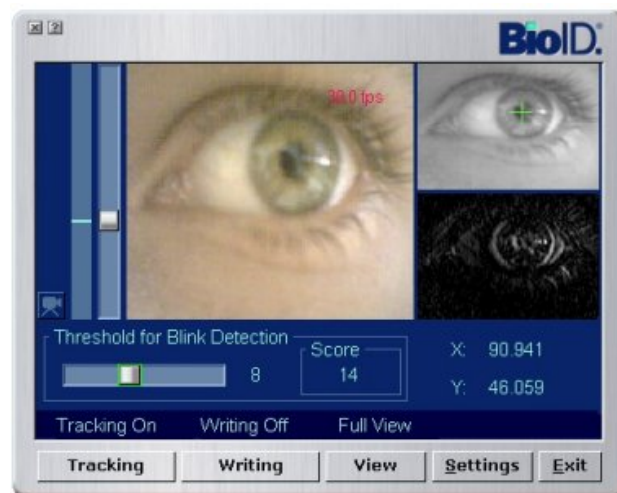


Fig. 5 Eye Tracker



Fig. 6 Camera for detection of Lips Movement_Determination of eye blinking and closing:-

- Determine eye blinking and eye closing using the number and intensity of pixels in the eye region.
- The system goes to the eye center of the previous frame and finds the center of mass of the eye region pixels
- Look for the darkest pixel, which corresponds to the pupil
- This estimate is tested to see if it is close enough to the previous eye location
- Feasibility occurs when the newly computed eye centers are close in pixel distance units to the previous frame's computed eye centers. This kind of idea makes sense because the video data is 30 frames per second, so the eye motion between individual frames should be relatively small.
- The detection of eye occlusion is done by analyzing the bright regions.
- As long as there are eye-white pixels in the eye region the eyes are open. If not blinking is happening.
- If the eyes have been closed for more than approximately 40 of the last approximately 60 frames, the system declares that driver has his eyes closed for too long.[5]

IV. FEATURES OF ARTIFICIAL PASSENGER

A. Conversational Telematics

IBM's Artificial Passenger is like having a butler in your car-someone who looks after you, takes care of your every need, is bent on providing service, and has enough intelligence to anticipate your needs. This voice-actuated telematics system helps you perform certain actions within your car hands-free: turn on the radio, switch stations, make a cell phone call, and more. It provides uniform access to devices and networked services in and outside your car. It reports car conditions and external hazards with minimal distraction. Plus, it helps you stay awake with some form of entertainment when it detects you're getting drowsy.

B. Improving Speech Recognition

You're driving at 70 mph, it's raining hard, a truck is passing, the car radio is blasting, and the A/C is on. Such noisy environments are a challenge to speech recognition systems, including the Artificial Passenger. IBM's Audio Visual Speech Recognition (AVSR) cuts through the noise. It reads lips to augment speech recognition. Cameras focused on the driver's mouth do the lip reading; IBM's Embedded Via Voice does the speech recognition. In places with moderate noise, where conventional speech recognition has a 1% error rate, the error rate of AVSR is less than 1%. In places roughly ten times noisier, speech recognition has about a 2% error rate; AVSR's is still pretty good (1% error rate).[5]

C. Analyzing Data

IBM's Automated Analysis Initiative is a data management system for identifying failure trends and predicting specific vehicle failures before they happen. The system comprises capturing, retrieving, storing, and analyzing vehicle data; exploring data to identify features and trends; developing and testing reusable analytics; and evaluating as well as deriving corrective measures. It involves several reasoning techniques, including filters, transformations, fuzzy logic, and clustering/mining. An Internet-based diagnostics server reads the car data to determine the root cause of a problem or lead the technician through a series of tests. The server also takes a "snapshot" of the data and repair steps. Should the problem reappear, the system has the fix readily available.

V. APPLICATION OF ARTIFICIAL PASSENGER

- Broadly used to prevent accidents.
- Used for entertainment such as telling jokes and asking questions.
- Artificial passenger component establishes interface with other driver very easily.
- Automatically opens and close the window of car and answers a call for you.
- If a driver gets heart attack or he is drunk, it sends a signal to vehicles nearby so that there driver becomes alert.

Future application

Will provide you with a "shortest time" routing based on road conditions changing because of weather and traffic, remote diagnostics of your car and cars on your route, destination requirements (your flight is delayed) etc.

VI. CONCLUSION

We suggested that such important issues related to a driver safety, such as controlling Telematics devices and drowsiness can be addressed by a special speech interface. This interface requires interactions with workload, dialog, event, privacy, situation and other modules. We showed that basic speech interactions can be done in a low-resource embedded processor and this allows a development of a useful local component of Safety Driver Manager.

The reduction of conventional speech processes to low-resources processing was done by reducing a signal processing and decoding load in such a way that it did not significantly affect decoding accuracy and by the development of quasi-NLU principles. We observed that an important application like Artificial Passenger can be sufficiently entertaining for a driver with relatively little dialog complexity requirements – playing simple voice games with a vocabulary containing a few words. Successful implementation of Safety Driver Manager would allow use of various services in cars (like reading e-mail, navigation, downloading music titles etc.) without compromising a driver safety.

REFERENCES

1. <http://www.scribd.com/doc/19067668/artificialpassenger>
2. <http://www.freepatentsonline.com/4682348.html>
3. <http://seminars4you.wordpress.com/artificial-passenger/>
4. www.visualexpert.com/Resources/roadaccidents.html
5. L.R. Bahl et al., "Performance of the IBM Large Vocabulary Continuous Speech Recognition System on the ARPA Wall Street Journal Task," ICASSP 1995, vol.1, pp 41-44.).
6. Lawrence R. Rabiner, A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition.