**Emotion Reading Robot**

# Introduction

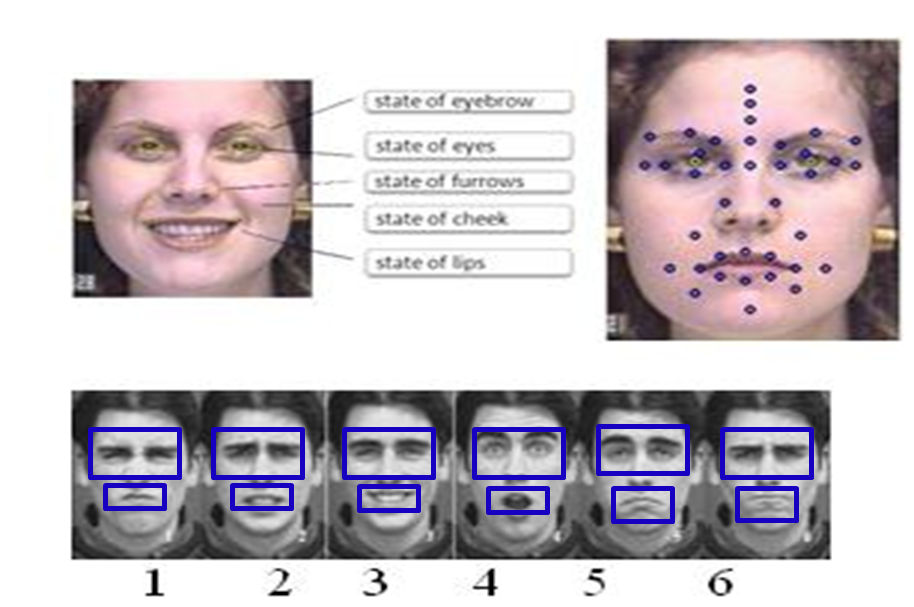
Emotion reading robot is one of the future-oriented robot models. At first, my project topic was just about humanoid robot ‘pepper’. However, I thought it is too specific product and I could only rely on information from the Softbank website. I thought it is not appropriate for final project subject so decide to change it to little broad topic ‘Emotion reading humanoid robot’. Because it is still cover about ‘pepper’, I believe this will be a perfect topic for me.

# Emotional recognition technique

1. Facial Expression(Visual info)-based

Face plays an important role in transmitting feelings. In psychology, when we have a conversation which sees each other’s face, facial expression is really significant. This means facial expression is crucial interface to delivery messages between human and computer.

Get the values which are extracted from experiment, express the emotion by comparing with the sample, transfer these emotions to the robot by the Serial Communication and at last express the emotion data through the monitor. Analysis the face and find people’s eyes and mouth and then classifies the face using prototype data. These matching has almost 70% recognition.



Ex) Google glassed application from ‘Emotient’

[Emotient](http://www.fastcompany.com/company/emotient)'s [technology works](http://www.fastcompany.com/explore/technology-works) by detecting subtle pattern changes in a person's face. The software, which can be used to measure emotion in individuals and crowds, can tell if a person is feeling positive or negative overall and zero in on more specific emotions, such as joy, sadness, surprise, anger, fear, disgust, and contempt.

1. Speech(Voice info)-based

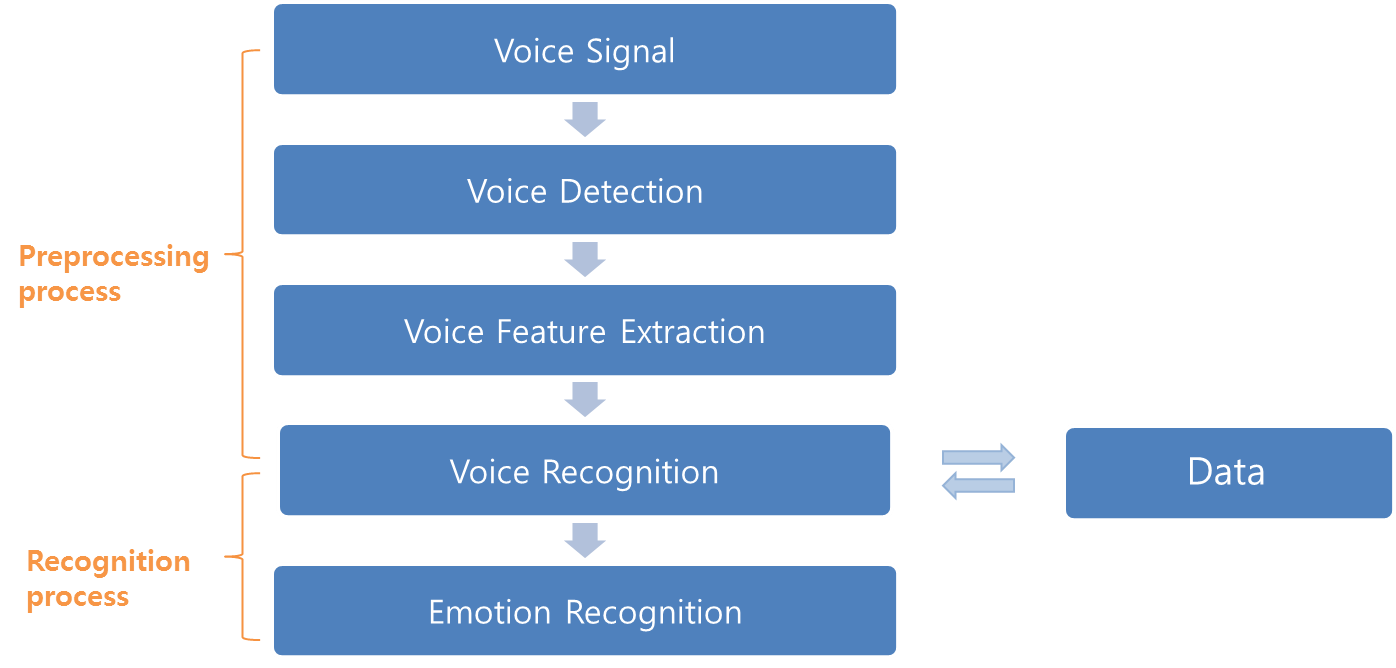
Voice which is including emotion is one of the most important factors that helps listener to grasp not only meaning but also speaker’s mental state during conversation. For smooth emotional communicate by recognizing emotion in an audio signal, I will introduce extracting voice features and classifying emotions. From the audio signals, using, SVM(Support Vector Machine) classifier

Although emotion detection from speech is a relatively new field of research, it has many potential applications. In human-computer or human-human interaction systems, emotion recognition systems could provide users with improved services by being adaptive to their emotions. In virtual worlds, emotion recognition could help simulate more realistic avatar interaction. The body of work on detecting emotion in speech is quite limited. Currently, researchers are still debating what features influence the recognition of emotion in speech. There is also considerable uncertainty as to the best algorithm for classifying emotion, and which emotions to class together. In this project, we attempt to address these issues. We use K-Means and Support Vector Machines (SVMs) to classify opposing emotions. We separate the speech by speaker gender to investigate the relationship between gender and emotional content of speech.

There are a variety of temporal and spectral features that can be extracted from human speech. We use statistics relating to the pitch, Mel Frequency Cepstral Coefficients (MFCCs) and Formants of speech as inputs to classification algorithms. The emotion recognition accuracy of these experiments allow us to explain which features carry the most emotional information and why.

It also allows us to develop criteria to class emotions together. Using these techniques we are able to achieve high emotion recognition accuracy.

1. Precessing process : The process which extract from voice feature.
   1. Recognize to measure the pattern and similarity in memory place
   2. Statistical signal handling technique
   3. Regulate the information about recognition and then using those rules to read voice
   4. Using neural network
2. Recognition process : Perform the recognition actually with extracted features.



Ex) Dialog from ‘Cogito’.

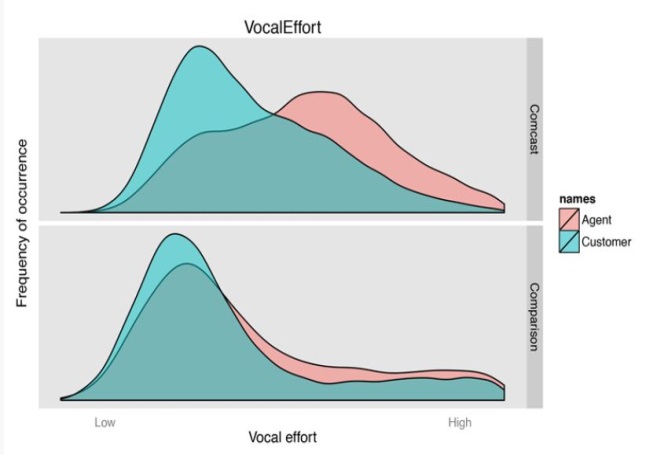
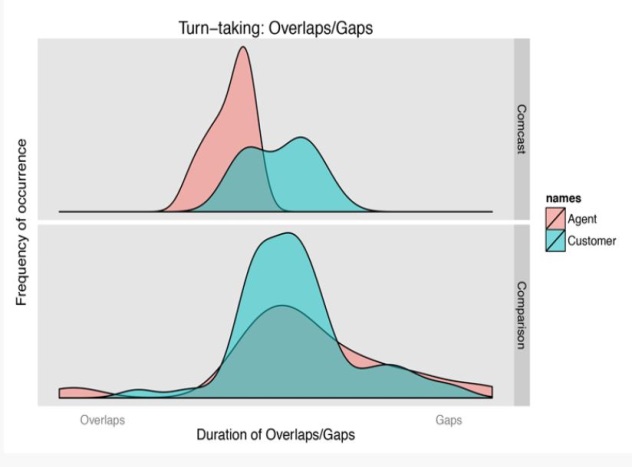
Software that listens to your voice to assess your mood gives call centre agents a dashboard that shows how the conversation is going.

Vocal Effort provides a means of discriminating soft, quiet voices from loud, tense ones. Vocal Effort focuses on the frequency components of speech that are independent of volume or energy. In the below figure, distributions of Vocal Effort are plotted for agents in red and customers in green. During the Comcast call (top panel) the analysis indicates significantly higher vocal effort and an overall tenser voice for the agent, as compared to the customer.

For one of our typical “good” calls (bottom panel), the agent and customer are both highly similar in terms of their vocal effort. Also, both speakers in the “good” conversation display a vocal effort profile that matches the customer in the Comcast call, who was generally perceived as remaining quite calm during the conversation.

It is especially important to measure when we begin to start speaking relative to the person we are speaking to. Allowing too much of a gap or silence, after our speaking partner stops talking can make us seem disinterested; however, consistently leaving too short a gap or overlapping the end of our partner’s speech by too much can be perceived as rude.

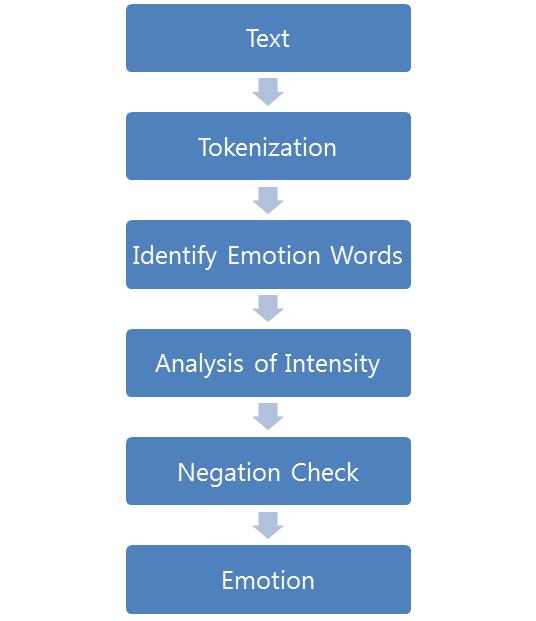
Behavior continues throughout the conversation. Had the agent noticed and modified his behavior at any of these moments, there might been a better outcome to the call.

1. User’s music or text mood-based

Music is oftentimes referred to as a “language of emotion” [1], and it is natural for us to categorize music in terms of its emotional associations.

* Keyword spotting Method



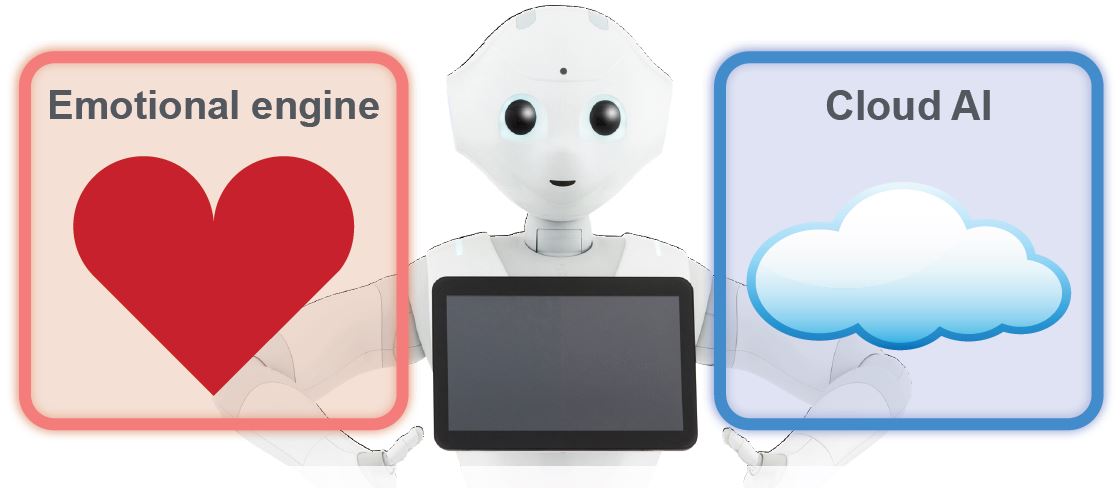
Keyword spotting technique for emotion recognition consists of five steps shown in fig.1 where a text document is taken as input and output is generated as an emotion class. At the very first step text data is converted into tokens, from these tokens emotion words are identified and detected. Initially this technique will take some text as input and in next step we perform tokenization to the input text. Words related to emotions will be identified in the next step afterwards analysis of the intensity of emotion words will be performed. Sentence is checked.

* Learning-based Method

Unlike keyword-based detection methods, learning-based methods try to detect emotions based on a previously trained classifier, which apply various theories of machine learning such as support vector machines [8] and conditional random fields [9], to determine which emotion category should the input text belongs.

* Hybrid Method

Since keyword-based methods with thesaurus and naïve learning-based methods could not acquire satisfactory results, some systems use hybrid approach by combining both keyword spotting technique and learning based method, which help to improve accuracy. The most significant hybrid system so far is utilizes a rule-based approach to extract semantics related to specific emotions As a result, these emotion association rules, replacing original emotion keywords, serve as the training features of their learning module based on separable mixture models. This method outperforms previous approaches, but categories of emotions are still limited.( *EMOTION DETECTION FROM TEXT*, Shiv Naresh Shivhare1 and Prof. Saritha Khethawat)

Pepper

In emotion recognition robot area, Pepper is the latest released one. He uses the techniques which are discussed previous pages, and also have a special technique, which is cloud system.  **Cloud computing(Cloud)** stands for computing in which large groups of remote servers are [networked](http://en.wikipedia.org/wiki/Computer_network) to allow centralized data storage and online access to computer services or resources. Clouds can be classified as public, private or [hybrid](http://en.wikipedia.org/wiki/Hybrid_cloud).(Wikipedia) Each robot(Pepper here) can learn from its interactions with humans and that their experiences will be uploaded to an Internet cloud database to be shared with other Peppers. This means Peppers can evolve as a collective intelligence

# Evaluation and conclusion

So far, previous four technique which make the robot recognize human’s emotion are developing all over the world. However, because they are still growing area and there is a limited cost, it is hard to combine all of the methods. Those enormous amounts of data which get from various ways will help to read human’s more delicate emotion. Also, although those technique’s evolving quickly, robot’s acting and speaking are expressions which are already predefined and recorded. Making robot express by themselves with its own intelligence and emotions is still challenging problems. Finally, robot’s emotion recognition can give Physical benefits as well as Emotional communication but can threaten human’s future life.

# Reference

<http://cdn.softbank.jp/en/corp/set/data/group/sbm/news/conference/pdf/material/20140605_01.pdf>

<http://arxiv.org/ftp/arxiv/papers/1205/1205.4944.pdf>

<http://www.cogitocorp.com/category/frontpage-article/#top>

<http://cs229.stanford.edu/proj2007/ShahHewlett%20-%20Emotion%20Detection%20from%20Speech.pdf>

<http://music.ece.drexel.edu/files/Navigation/Publications/Kim2010.pdf>