Reverse Geocoding

– Offline geolocation of the user was required for selecting language packages and possibly for some future aspects of the project.The following approach was used - By using a list of cities ,states and countries (http://download.geonames.org/) and applying nearest neighbour algorithm to find a rough estimate of the user's location.

```
import reverse_geocoder as rg
import math
def calCor(lat1,lon1,brng):
        R = 6378.1 #Radius of the Earth
        brng = math.radians(brng) #Bearing is 90 degrees converted to radians.
        d = 15 #Distance in km
        lat1 = math.radians(lat1) #Current lat point converted to radians
        lon1 = math.radians(lon1) #Current long point converted to radians

          lat2 = math.asin( math.sin(lat1)*math.cos(d/R)
+math.cos(lat1)*math.sin(d/R)*math.cos(brng))

          lon2 = lon1 + math.atan2(math.sin(brng)*math.sin(d/R)*math.cos(lat1),math.cos(d/R)-
math.sin(lat1)*math.sin(lat2))

          lat2 = math.degrees(lat2)
          lon2 = math.degrees(lon2)
        return (lat2,lon2)


latitude= input("Enter latitude")
longitude=input("Enter longitude")
for num in range(0,8):
     coordinates=calCor(latitude,longitude,45*num)
     position=rg.search(coordinates,mode=1)
     print position[0]['admin1']+'-language-pack'
```

Text to Speech

– In case of some accident that the driver may suffer, the nearby people Have to be informed about what has happened and the possible first aid measures that can be taken to save the person.To produce a customized first aid voice message on spot we needed to do text to speech conversion offline . Due to unavailability of suitable TTS convertors for regional languages we used concatination of word pronunciations to produce the required message . To do this task ,we wrote a sample python script using pydub and pyaudio libraries.

```
#This script helps to concatinate different audio file and playing it depending on the disease the
person is suffering from


import pydub,pyaudio

from pydub import *

from pydub.playback import play
```

```
from pyaudio import *
print "What Disease is the person having"
#  Takes input of the name of the disease the person is suffering from
disease=raw_input()
if disease=="heart attack":
 cpr=AudioSegment.from_file("speech.ogg")                          #  CPR
 heartAtack=AudioSegment.from_file("speech6.ogg")                  #  Heart Attack
 song_final=heartAttack+cpr+20                                     #  Final Audio output
elif disease=="unconscious":
 cpr=AudioSegment.from_file("speech.ogg")                          #  CPR
 unconscious=AudioSegment.from_file("unconscious.ogg")             #  Unconsciousness
 unconscious2=AudioSegment.from_file("unconscious2.ogg")
  #  measures taken during unconsciousness
 aur=AudioSegment.from_file("and.ogg")
 song_final=unconscious+cpr+aur+unconscious2+20
elif disease=="bleeding":
 bleeding=AudioSegment.from_file("bleeding.ogg")                   #  Excessive bleeding
 bleeding2=AudioSegment.from_file("bleeding2.ogg")                 #  measures for bleeding
 song_final=bleeding+bleeding2+20
play(song_final)
```

SENSOR FUSION

Integration of Data and knowledge id called sensor fusion a.k.a data fusion
There are three steps for this
1) Data Association
2) State estimation
3) Decision Fusion

1) Data Association

The data association problem must determine the set of measurements that correspond to each target Data association is often performed before the state estimation of the detected targets. Moreover, it is a key step because the estimation or classification will behave incorrectly if the data association phase does not work coherently. The data association process could also appear in all of the fusion levels, but the granularity varies depending on the objective of each level.
These are some methods for the same

A. Nearest Neighbours and K-means ----- Nearest neighbor (NN) is the simplest data association technique. NN is a well-known clustering algorithm that selects or groups the most similar values. How close the one measurement is to another depends on the employed distance metric and typically depends on the threshold that is established by the designer.

ADVANTAGES OF NEAREST NEIGHBOUR

Its a simple algorithm and takes less amount of time

DISADVANTAGES

In a cluttered environment, it could provide many pairs that have the same probability and could thus produce undesirable error propagation .This algorithm has poor performance in environments in which false measurements are frequent, which are in highly noisy environments.
As our project is based on data collection from environments which can be highly noisy it is better that we do not employ this method for the data association

K MEANS ----- - K-Means divides the dataset values into different clusters. K-Means algorithm finds the best localization of the cluster centroids, where best means a centroid that is in the center of the data cluster. K-Means is an iterative algorithm that can be divided into the following steps:
(1)obtain the input data and the number of desired clusters
(2)randomly assign the centroid of each cluster
(3)match each data point with the centroid of each cluster
(4)move the cluster centers to the centroid of the cluster
(5)if the algorithm does not converge, return to step (3).

DISADVANTAGES
(i)the algorithm does not always find the optimal solution for the cluster centers.
(ii)the number of clusters must be known a priori and one must assume that this number is the optimum.
(iii)the algorithm assumes that the covariance of the dataset is irrelevant or that it has been normalized already.

REFERENCES

1.Federico Castanedo “A Review of Data Fusion Techniques” ,The Scientific World Journal
Volume 2013 (2013)
2.Tapas Kanungo ,David M. Mount “An Efficient k-Means Clustering Algorithm: Analysis and Implementation” IEEE Transactions on Pattern Analysis and Machine Intelligence archive Volume 24 Issue 7, July 2002

STATE ESTIMATION

State estimation techniques aim to determine the state of the target under movement given the observation or measurements. State estimation techniques are also known as tracking techniques. In their general form, it is not guaranteed that the target observations are relevant, which means that some of the observations could actually come from the target and others could be only noise. The state estimation phase is a common stage in data fusion algorithms because the target's observation could come from different sensors or sources, and the final goal is to obtain a global target state from the observations.

State estimation falls under 2 groups

1) Linear dynamics and measurements: The estimation problem has a standard solution. Specifically, when the equations of the object state and the measurements are linear, the noise follows the Gaussian distribution, and we do not refer to it as a clutter environment; in this case, the optimal theoretical solution is based on the Kalman filter

2) Nonlinear dynamics: the state estimation problem becomes difficult, and there is not an analytical solution to solve the problem in a general manner. In principle, there are no practical algorithms available to solve this problem satisfactorily.

These are the most common Techniques that are used for state estimation

1) KALMAN FILTER The Kalman filter is the most popular estimation technique. It was originally proposed by Kalman [34] and has been widely studied and applied since then.

The Kalman filter is mainly employed to fuse lowlevel data. If the system could be described as a linear model and the error could be modeled as the Gaussian noise, then the recursive Kalman filter obtains optimal statistical estimations

ADVANTAGES The main advantage of the Kalman filter is its ability to provide the quality of the estimate (i.e., the variance), and its relatively low complexity.

Its suitable for Gaussian Models

Kalman filter is more computationally complicated but it has a more detailed model of the system so it is more accurate in multisensor fusion.

Kalman Filter also incorporates the noise and is basically designed for getting correct estimations on combining two sensors

DISADVANTAGES

It provides accurate results only for Gaussian and linear models. For Gaussian models with limited nonlinearity, extended Kalman filter (EKF) is appropriate. For nonGaussian and nonlinear models, particle filtering (PF) is the most appropriate approach, since it is able to provide arbitrarily posterior probability distribution.

2) PARTICLE FILTER Particle filters are recursive implementations of the sequential Monte Carlo methods . This method builds the posterior density function using several random samples called particles. Particles are propagated over time with a combination of sampling and resampling steps. At each iteration, the sampling step is employed to discard some particles, increasing the relevance of regions with a higher posterior probability.

ADVANTAGES This model is better suited for Non Gaussian models.Particle filters are more flexible than the Kalman filters and can cope with nonlinear dependencies and nonGaussian densities in the dynamic model and in the noise error

DISADVANTAGESA large number of particles are required to obtain a small variance in the estimator. It is also difficult to establish the optimal number of particles in advance, and the number of particles affects the computational cost significantly.Need more data for this

3) Maximum Likelihood The maximum likelihood (ML) technique is an estimation method that is based on probabilistic theory. Probabilistic estimation methods are appropriate when the state variable follows an unknown probability distribution

ADVANTAGESThe advantage of this estimator is that it normally involves fewer random terms than does the empirical Hessian, and it may therefore be somewhat more efficient in finite samples.

DISADVANTAGESThe main disadvantage of this method in practice is that it requires the analytical or empirical model of the sensor to be known to provide the prior distribution and compute the likelihood function. This method can also systematically underestimate the variance of the distribution, which leads to a bias problem.

REFERENCES

1) D. L. Hall and J. Llinas, “An introduction to multisensor data fusion,” Proceedings of the IEEE,1997

2)M. Shindler, A. Wong, and A. Meyerson, “Fast and accurate κmeans for large datasets,” in Proceedings of the 25th Annual Conference on Neural Information Processing Systems (NIPS '11), pp. 2375–2383, December 2011

3)Welch and G. Bishop, An Introduction to the Kalman Filter, ACM SICCRAPH, 2001 Course Notes, 2001.