Capstone Project

Geo-tagged Sensor Fusion

Bardia Keyoumarsi, Michael Bennett, David Tucker,
Vincent Lantaca, Michael Baptist

UC SANTA CRUZ
Baskin Engineering

Lawrence Livermore National Laboratory

Abstract

Motivation:

The Lawrence Livermore National Laboratory's Data Science Initiative was established to advance state-of-the-art technology in the "big data" domain including data collection methodology, storage, visualization, analytics, modeling, simulation, and high performance computing capabilities. To support this initiative, the students and faculty of the School of Engineering (SOE) at University of California, Santa Cruz (UCSC) produced a real-time sensor collection system whose acquired data can be fused with publicly available "open" data (e.g. social media, climate data, traffic data, etc.).

Objective:

The goal of this project is to produce a web application that will merge gathered data with publicly available information using standard formats and open platforms. Another dedicated application (for iOS) would allow any iPhone to become an information collection source, and relationships among data can be determined using attached geotags. Related data may be visualized using a number of methods including graphing and geographic plotting. Implementation specifics will need to be clearly documented and cleanly presented.

In addition to the real-time sensor collection system, the GSF project team shall provide tools and frameworks for data query, analysis, and visualization of the fused data for utility and ease of adaptation to later potential applications.

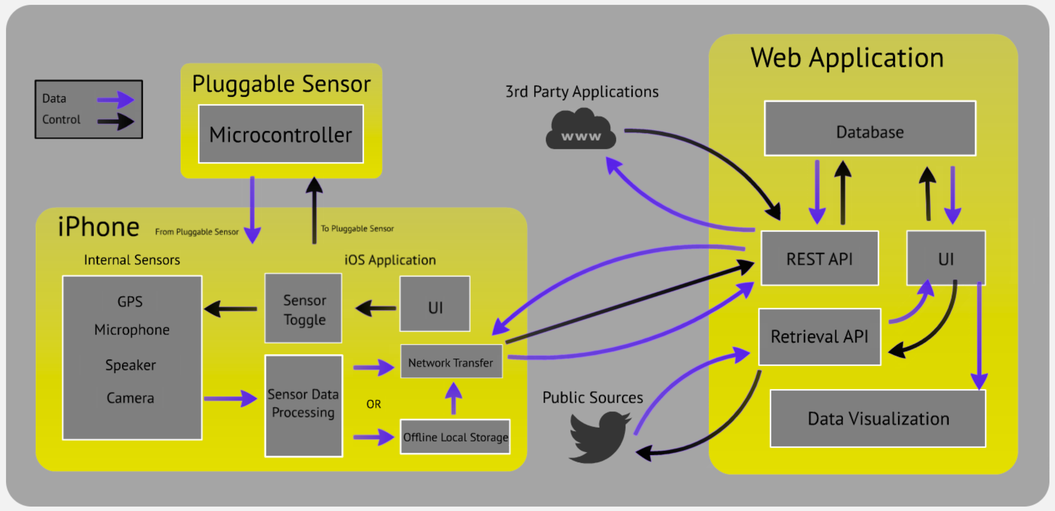
Approach

The Geo-tagged Sensor Fusion (GSF) project uses a dedicated iOS application that interacts with the iPhone's onboard sensors to collect various types of data. In addition, a pluggable sensor module has been designed to interact with the iPhone's headphone jack for additional sensory data. The headphone jack is used for power and communication with the pluggable sensors. Each sensor is identified by our application when plugged in, and transmits data samples to the application.

All the collected data is then sent to a centralized server where it is stored in a NOSQL database. The web application is used to interact with the database to allow users to create their own visualizations. The users have the ability to fuse the collected data with other publicly available "open" data to create meaningful visualizations.

Overview

Design



Software Analysis

Web Application:

The web application suite runs on the Django platform and uses MongoDB as its NoSQL database. The wep app allows users to create visualizations by fusing sensory data collected by the iOS app and tweets retrieved from Twitter. The backend of the web app includes a RESTful API that allows the iOS app to upload data and third party applications to access data from our database. All data is sent in GeoJSON format.

Retriever & Visualizer:

Queries issued on the fusion interface return data contained in our database. Periodically, those queries are sent to Twitter in the background to ensure results remain current and relevant. The retriever responsible for this process filters out Tweets that are not geotagged and packages the rest in GeoJSON format.Data from Twitter is merged with data from the database before being sent to the visualizer which is capable of rendering any valid GeoJSON file. It produces plain visualizations using HTML5 and spatial visualizations using OpenStreetMap and the Leaflet JavaScript API.

iOS Application:

The iOS application is a data collection system that supports iPhone 4 up to iPhone 5s running iOS 7. It allows the user to collect imagery data and runs OpenCV to detect the number of bodies and faces in the image, which is used to estimate the number of people at a GPS coordinate. Other onboard sensors can be turned on via the custom user interface including our pluggable sensor suite. Once finished with a collection session the data is sent to our server securely using custom API keys or saved to be sent later if no network service is detected or whenever the user is done collecting. The system is easily scalable to allow for any type of data to be transmitted.The application also offers a route planning service which allows an administrator to send a group of GPS coordinates via text message to a field agent. Our service will compute and plot the fastest route to each location so the agent can optimize their time collecting data. They also have the option of getting directions via Safari or launching the route into the native Google Maps application.

Hardware Analysis

Pluggable Sensor:

The pluggable sensor is powered by a sine wave on the left channel of the headphone jack, which is stepped up, rectified, and regulated to around 3V. This is used to power an ATxMega128a4u microcontroller which retrieves data from a digital temperature/humidity via TWI.

Signal Processing:

The communication from our microcontroller to the iPhone is through the mic-line of the iPhone's audio jack. We use a controlled AC signal to transfer encoded data over the communication channel. When the signal reaches the iPhone we process the audio samples and decode the bits. The iPhone can request data from the microcontroller by sending a transmission on the right audio output channel, which is connected to an ADC on the microcontroller.

Results

Over the course of our allotted implementation time we have developed a mobile sensor suite that pushes sensor population, temperature, and humidity data to our web application. In addition to data collected by the iPhone, we have also developed a retriever that requests data from Twitter. The data is stored in our database and is available for viewing via our visualizer.

A map visualization contains tweets and sensory data found across the world.

GSF iOS Application

Conclusion

The Geotagged Sensor Fusion project will allow users to participate in a global survey that will attempt to merge and visualize the massive amounts of unorganized public data available to netizens today.
The use of sensors allows people to add to publicly available data which could be useful for census-taking, crisis response, and many other applications.

Establishing data relationships provides a better means for producing new research and educating the population as it makes finding relevant information quicker and more convenient.

Acknowledgements

We would like to thank:

- Lawrence Livermore National Laboratory for sponsoring our project and Michael Goldman for his countless hours of feedback and technical support.
- David Munday and Ethan Papp for their guidance, insight and mentorship.
- Terry Figel and Heidi Sitton for giving us access to various ITS services.
- Baskin Engineering Lab Support for their assistance with lab equipments.