

2.7 Participatory Sensing (PART)

The vision of Participatory Sensing is of distributed data collection and analysis spanning the personal, urban, and global scale, often using “everyday” technologies like cell phones, in which participants make key decisions about what, where and when to sense. Previously called Urban Sensing, the area was renamed to emphasize both its wide applicability outside of cities and its strong conceptual grounding in user participation (Figure 5).

The area targets technologies and applications that transform our capacity to help individuals, families, and communities monitor and improve their own health behaviors, adopt sustainable practices in resource consumption, and participate in civic processes. Each of these three touchstone topics—**health, sustainability** and **civic engagement**—is being explored in real-world deployments, such as *AndWellness* (now Google and NIH-funded); the *Personal Environmental Impact Report* and *What’s Invasive*; and *Remapping LA*, respectively. In addition to these application-driven pilot deployments, the area has conducted technology-focused research around topics necessary for complete, robust participatory sensing systems, including: participant recruitment, task planning, and sensing campaign management; human activity classification based on mobile phone sensors; integration with environmental monitoring assets; and data visualization. Privacy challenges are being addressed within a holistic ethics framework that emphasizes principles of participant primacy, data legibility, longitudinal engagement, and parsimony. From this work, the concept of a *personal data vault* (PDV) has emerged as a key technical component of a safe future participatory sensing ecosystem. The PDV would provide both technical and legal mechanisms to protect individuals’ participatory sensing data while supporting its use third-party applications of benefit to the user.

Since the introduction of the term by CENS in 2006, the area of *participatory sensing* (along with urban sensing) has generated a body of multidisciplinary work spanning many universities, including UCLA, Dartmouth, Columbia, MIT, CMU. It has also inspired work outside computer science in design, urban planning, and the arts. The more technical aspects of our PS innovations are included in the section covering Programming and Platforms.

Application Drivers & Pilot Deployments

Collaborations targeting the three focus areas continue with the Semel Institute’s Global Center for Children and Families at the UCLA School of Medicine, the National Park Service, UCLA REMAP, and others. The group has also developed new collaborations to increase the scale of public use of these technologies. For example, a partnership with the Los Angeles Unified School District and Google will incorporate participatory sensing on Android phones into computer science and mathematics classrooms starting in Spring 2010, with expansion across the district pending further funding from NSF. A new collaboration has begun with the Boyle Heights Neighborhood Collaborative in Downtown Los Angeles to map, record, and accumulate data on community member circulation and related conditions—a unique, active and participatory approach to supporting the Boyle Heights Planning for Place project in developing its plan for a healthy community. CENS also received a highly competitive ARRA-funded NIH Challenge Grant to develop an innovative real-time assessment of behavioral exposures for cardiovascular disease (CVD) in young overweight mothers. Other health science projects include exploratory work around supporting cancer survivorship research, HIV behavior

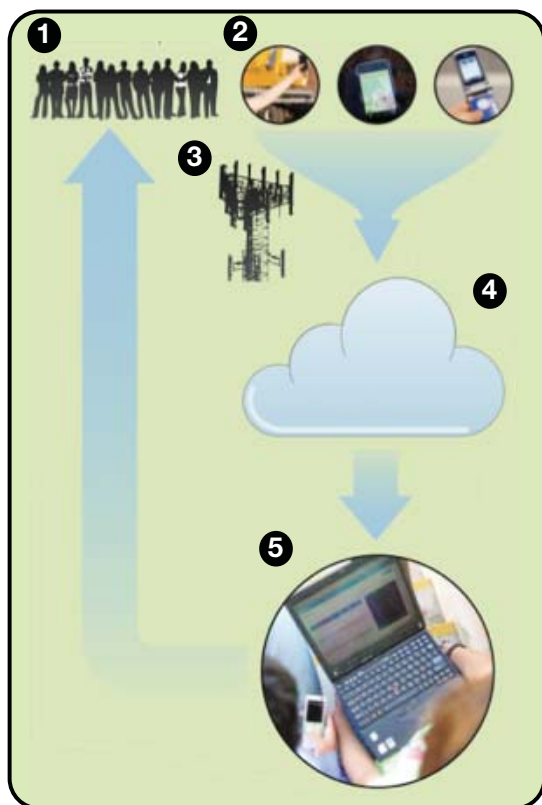


Figure 5. Participatory Sensing in Action. **1. Participate.** People can initiate a Participatory Sensing exercise for a variety of reasons. The first step is to organize the participants—whether individuals acting alone or a large group acting in concert—to determine the goals and data collection plan. **2. Mobile personal data devices.** Using mobile phones, participants collect data automatically (e.g., location logging) or manually (e.g., taking pictures). **3. Ubiquitous wireless transfer.** Data are moved from anywhere in the world via wireless infrastructure. **4. Data processing in “the Cloud.”** Data from a variety of sources and locations are collected and processed to reveal patterns that were previously invisible. **5. Learn and act.** Participatory Sensing systems digest the results of analyses into simple visualizations that can be shared and used to make change.

survey with the Center for HIV Identification, Prevention, and Treatment Services, and a collaboration with UCSF on “mHealth” architecture.

Applications have thus taken on larger populations, with the largest this Spring on the order of 300 users, with increasing number and scale planned for 2010-2011. This scaling will generate substantial feedback on system robustness requirements and optimization targets; participant pool coordination, planning and management needs; and the importance of careful user experience and interface design for successful deployment. Longer-running applications, like PEIR, have provided the opportunity to create second generation architecture and begin model validation.

Privacy, Ethics, Law and Policy

A major focus of research and implementation is the Personal Data Vault (PDV): a logically isolated secure repository for participatory sensing data that is controlled by the handset owner. The PDV receives participatory sensing data as it is collected and selectively distributes it to third party applications according to a set of sharing rules created by the user. Not only does this aim to reduce the number of parties holding potentially sensitive data, but also to create a place for users to visualize and manage what data are being collected about them and what they have elected to share. We envision an ecosystem of PDVs as a fundamental part of future participatory sensing systems, and are collaborating with Prof. Jerry Kang of the UCLA School of Law to develop a legal approach that could provide additional protection for the data contained in the PDV. An initial implementation is being built and will be integrated into CENS applications in 2010-2011.

The PDV is one of several examples of CENS participatory sensing research that is influenced through interaction with ethics education and research that aims to promote the participatory principles and user empowerment fundamental to this area. This work is in its second year of funding from the NSF Ethics Education in Science and Engineering, and is centered a participant-observer study of CENS research that aims to develop educational materials promoting ethics considerations in the development of participatory sensing systems, as well as original work in Information Studies on participatory practices in data collection. In addition to academic publications, reports, and popular articles in this area, a new interdisciplinary undergraduate course is being offered this spring that will explore the topics in depth.

Future Work

New work in the next year will continue to focus on expanded real-world deployments with a larger number of users, higher stakes uses in real communities, and more robust, shared systems (such as the PDV) to support data collection campaigns. These deployments will provide the systems scaffolding and practical opportunities to incorporate technology research in campaign deployment, management, recruitment, incentive, data processing and the other areas listed above. They will also provide opportunity for formal user studies and more concrete understanding of how to achieve maximum impact in the real world. Following an assessment of progress of the area planned for early Summer 2010, we will reorganize the technology and application research around a set of central challenges for the remaining two years of CENS core funding.

PART 01 AndWellness: Improving Wellness with Mobile Personal Sensing

PART 01.1 Overview

Mobile phones can radically personalize health care by providing more accurate and reliable data from individuals. However, a mobile architecture needs to provide privacy, engaging interfaces to encourage adherence, a way to present the large quantities of data, and be evaluated in rigorous side by side tests in real patient pilots.

We are building AndWellness, a mobile personal sensing application for the Android platform that includes a suite of mobile services, and server-side software to improve personal health and wellness. As a collaboration between the Center for Embedded Networked Sensing and the Global Center for Children and Families (<http://www.gccf.ucla.edu/>), AndWellness will transform mobile devices into tools that uncover a user's behaviors at the heart of personal wellness without violating their privacy, and help users design customized interventions to improve their health. Customizable spatial-, social-, temporal-, and mobility-triggered reminders, assessments, and interventions are relevant to a wide array of behavior change objectives. AndWellness will initially focus on a subset of these behaviors in a series of pilots, as described in Section 14.5 below.

PART 01.2 Approach

Ecological Momentary Assessments

Adherence to a behavioral intervention can often be more important than the intervention itself. Self-monitoring is critical to improving adherence to behavior change programs (Donavan, 2005; Marlatt, 1985). But human memory is fallible: Frequency estimates for discrete events tend to cluster about round numbers (e.g. 20/day vs. 19); recall of conditions preceding a salient event may be systematically distorted by knowing the outcome (i.e. "I must have felt badly to do such a thing"; current mood state may affect the availability of exemplars for retrieval from memory (e.g., recent eating influences recall of a previous hunger); and it is easy to backfill, forward fill, or fake fill (Piasecki, 2007; Hufford, 2007).

To improve outcomes by augmenting human memory, EMA was developed to monitor affect, cognitions, and behavior in real time in a person's natural environment (Shiffman, 2007). EMA is implemented through the use of a portable electronic or recording device, such as a cell phone. Data is collected in real time, so is not subject to recall and retrieval biases (Schwarz, 2007); Data is harder to fake (Collins, 2003; Hufford, 2007); Errors are identified early (Collins, 2003); Data is instantaneously entered into a secure central database. EMA can aid in contextual, spatial, or temporal associations to behaviors; especially well suited for tracking short, discrete events (Shiffman, 2008).

Design Approach

In order to design the system, we are using a Continuous Quality Improvement cycle, which is an iterative, 3 step cycle:

- **Key Informants.** We are interviewing behavioral experts, medical providers, and professionals who have experience working with target communities, and with mobile phone based interventions.
- **Focus Groups.** We are recruiting participants from our specific target communities to test prototypes, and provide feedback on screen layout, periodicity of reminders, terminology, configurability, privacy, and data presentation.
- **Design Experts.** We translate the constraints from the Key informants and Focus groups into technical specifications which can be implemented, and provided to Focus groups for evaluation.

PART 01.3 System Description

We propose to build a three tiered architecture consisting of the mobile phone (Figure 1), a personal data vault, and a back-end server. A website will provide (1) trigger authoring (2) real-time Feedback (3) place labeling (4) PDV interface (5) visualizations and patient dashboards (6) usage Monitoring statistics. The mobile phone will provide (1) an engaging UI (2) power management (3) automated activity classification. The mobile phone will include components to monitor usage. The entire system will be built using standardized data formats and modular components for robustness.

PART 01.4 Accomplishments

We have built an end to end system that is ready for pilot evaluation. We have undertaken several internal pilots to evaluate the technical feasibility and validity of the mobile phone measures, and completed several Key informant interviews. We are now in a recruitment phase for our focus groups.

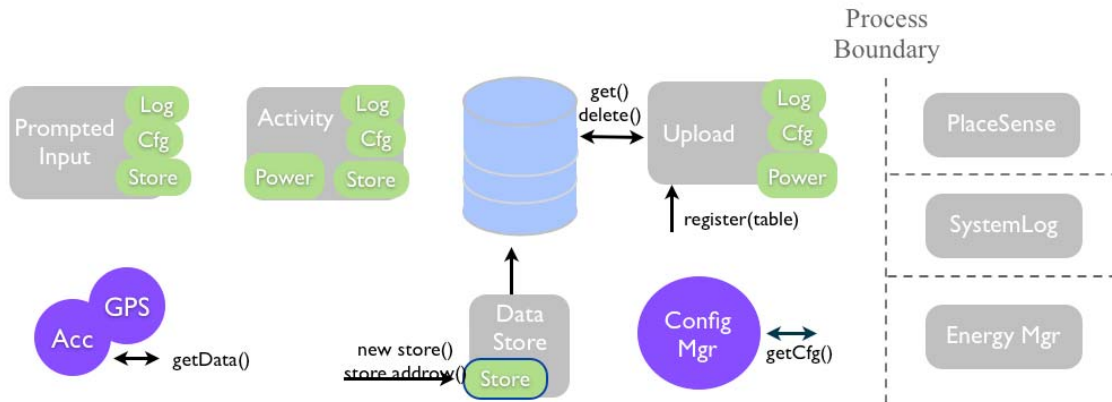


Figure 1. AndWellness architecture on the phone.

PART 01.5 Future Directions

We plan to undertake a number of pilots this year with diverse populations, to evaluate the reliability, validity, usability, and effectiveness in behavioral interventions of AndWellness.

Cardiovascular Risk factors, 60 Moms. In this NIH funded study, we will provide mobile phones to young moms for 6 months, to help them self-monitor diet, stress, and exercise, three key risk factors for cardiovascular disease. We will use the data from the pilot to evaluate the validity, reliability, and usability of the AndWellness application.

- Exercise, mood, and energy levels, 100 Breast Cancer Survivors. We will provide mobile phones to young breast cancer survivors. We are integrating mobile phone monitoring into an existing study run by Professor Patti Ganz to obtain quick feedback from small groups of users.
- Sex and co-occurrence with alcohol and drug use, 30 Gay men. In this Center for HIV Identification Prevention and Treatment funded study, we will provide mobile phones to at risk youth to collect highly sensitive information to better explore the tradeoff between privacy needs and study compliance.
- Exercise, 30 South Asian Women. In this Northwestern University funded study, we will work with practitioners from Northwestern Medical campus to evaluate the validity, reliability, usability of mobile phones in an older, immigrant population, and explore the mobile phone as a behavioral intervention tool to encourage exercise.

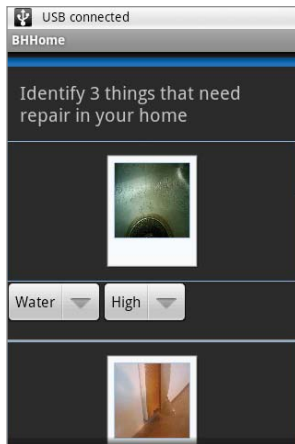
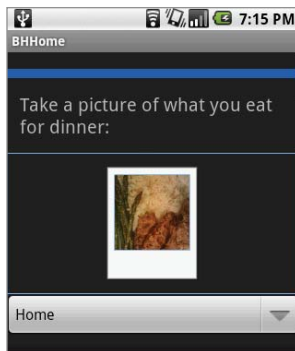
PART 02 Boyle Heights Planning for Place

PART 02.1 Overview

The Boyle Heights Planning for Place project is using an innovative new approach to community engagement and data collection developed at UCLA and referred to as Participatory Sensing.

PART 02.2 Approach

Typically planners collect data without access to the detailed patterns of resident flow and context. The Boyle Heights Participatory Sensing Project will map, record, and synthesize data on the circulation of community members within the community and the conditions around this circulation. Community members will use smartphone applications to map and trace their everyday movements throughout Boyle Heights: where they work and study, \ where they gather, how they get there, and the conditions surrounding them.



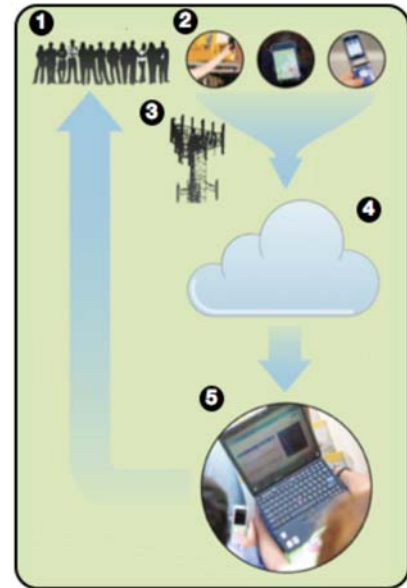
PART 02.3 System Description

In particular, the Home application asks each participant to: (1) take a picture of what they eat for dinner, and label it with where it was prepared (home, restaurant, fast food, other); (2) to identify three things that need repair, take a picture, rate their level of concern, and label the type of problem (peeling paint, insects/rodents, water, heat/ac); (3) to document what makes them uncomfortable at home (overcrowded, noise, no place to do homework, yelling/violence, no peace, other), to optionally take a picture, and to rate their

level of concern. The Path application captures a GPS trace of the participant's path to school or work and asks them what transportation mode they use; to take pictures of places where they stop along the way (and for what purpose), as well as things that represent their neighborhood (people, places, events), and where and what they eat along the way. Using the School application, youth participants assess the physical conditions of their school, take a picture of their favorite teacher and what they eat for lunch; and describe their sense of safety and learning. Whereas adults use the Work application to identify up to three concerns, and label (safety, health, stress, pay) and rate them. The Afterschool/work application captures data about about where residents spend their time, what responsibilities they have then, and how safe they feel. In all of these cases, the participants responses are automatically geocoded, time-stamped, and uploaded to a secure database, where the community organizers have access to the statistics, stories, and images, shared by participants.

This project is an example of Participatory Sensing, an approach to data collection and interpretation in which individuals, acting alone or in groups, use personal mobile devices and web services to systematically explore interesting aspects of their worlds ranging from health to culture. Participatory Sensing makes it feasible and very low

cost for communities to engage their members in data collection and analysis through the use of real time, real place, real context smartphone applications.



PART 03 What's Invasive! & Project BudBurst

PART 03.1 Overview

Both of the CENS participatory sensing projects, What's Invasive! and Project BudBurst, involve using mobile phones for the collection of ecologically important data.

What's Invasive! (WI; <http://whatsinvasive.com>) is a CENS citizen science campaign for locating invasive weeds. We have made the system global in reach but local for identifying and locating weeds. With the current online system, any National Park Service official (or other authorized person) can create a park-specific weed list with descriptions and photos uploaded to our server. Once the weed list is finished and authorized, then all Android phones running the WI application will be able to download this list and the phone is ready to help in identifying and locating these species. All the data collected by the users of the phones are uploaded to the server and are immediately available to anyone for downloading as a CSV file.

Project BudBurst (PBB) is a national field campaign for citizen scientists designed to engage the public in the collection of important climate change data based on the timing of leafing and flowering of trees and flowers. It is run by the University Corporation for Atmospheric Research (UCAR) Office of Education and Outreach, with whom CENS is collaborating on several grants. Observations made by phone will be sent to the PBB database, along with photos and comments.

CENS work is in developing not only a smart phone application for both WI and PBB but also in supporting text (SMS) and picture (MMS) messaging capabilities for both projects. Additionally, a website and mapping system have been developed for WI.



Figure 1. The What's Invasive! Android phone application splash screen.

PART 03.2 Approach

The purpose of the What's Invasive! website and software is to provide scientific data for research and environmental management and also to promote environmental awareness about invasive plants. The WI website and software work by collecting location data (via GPS on a mobile device, or via user-input), time, date, and verifying photographs of invasive plant species. Individuals with GPS-enabled mobile phones can easily contribute to locating invasive species by making geo-tagged observations and taking photos to alert us of the spread of habitat-destroying plants. We have applications for the G1 and iPhone. Non GPS-enabled phones are also useful for capturing notes and photos, and geolocation can be established later through the website. Email (with optional photo) with the plant name as the subject to: mobile@whatinvasive.com will also be logged as an observation or a new observation can be made using a web form.

The purpose of Project BudBurst is to engage the public in making careful observations of the phenophases such as first leafing, first flower, and first fruit ripening of a diversity of trees, shrubs, flowers, and grasses in their local area. CENS is collaborating with PBB to create an Android application and also to allow regular text and picture messaging to be easily incorporated into data collection by PBB participants.

Other citizen science projects that are using mobile phones are limited, but EpiCollect (citizensci.com/2009/09/17/epicollect) has made progress on using smart phones for mapping of disease spread or the occurrence of rare species.

PART 03.3 System(s) Description and/or Experiments

WI Database and Webserver

The WI website was created as a way for users to interact with the data collected. From the website they can view all the observations that have been uploaded in list view or on a Google map. Users are able to edit their own observations they have previously submitted or flag other user's observations as inappropriate. The website also provides summary statistics of the data as graphs and a leader board of those who submitted the most observations.

New observations can be added from the website for those without smart phones. A web form is presented with the list of invasive species along with a map with a "draggable" marker to make submission as easy as possible and one can specify weed quantity and add an image as well.

Information about the invasive species targeted by these campaigns can be seen from the website. A blog was created to keep users up-to-date with the current happenings of the WI campaign.

Users are able to add their own parks to the system to allow other users to contribute. The user is prompted for a latitude and longitude to use as the center of the park. The user may then upload species information and images for the invasives they want targeted in the park. Once parks are officially approved, the park will show up on the public list of parks in the website and on participant's phones. Users and administrators are able to edit the species information and images for parks they have created.

You can also manage your sms/mms/email associations through the website by visiting the accounts section. Here you can set up new phone numbers or email addresses to associate to your WI account so that observations can be made remotely from these modes.

WI Phone applications

A WI phone application is available for the iPhone and Android mobile platforms, with the iPhone currently only available for the SAMO park. The Android version supports all currently created parks as well as any parks created in the future. The Android version allows for updates to specific park plant lists to be sent to users without requiring a reinstall. The application also automatically selects the park you're visiting by taking GPS measurements and calculating the distance to the closest park. A user can still override this function by manually selecting a park to make submissions for through the settings screen. Using the application, a user can simply take their phone while visiting a park, take a photo, and select the plant they're observing and the application will take care of time-stamping, locating, and submitting the observation to the WI database automatically when network coverage is available.

The phone application also supports some feedback support in the form of maps and graphs. For each park, a user can view both personal and global statistics such as submission counts, location of previous submissions, and submission details. This can be used to inform users of their contributions and encourage competition in terms of having the most submissions to a park.

PBB WebServer

A CENS Summer Intern Project expanded on PBB to create a user-friendly website and framework for citizen scientists to register sites and plants, submit and update observations, add geo-tags, and upload photos and comments. Participants can view the progress of their own data collection relative to historical data collected on the same plant and interact with other current participants by viewing or commenting on each other's findings. Furthermore, users without immediate internet access have the ability to submit observations or notes, with optional image attachments, via Short Message Service (SMS), Multimedia Messaging Service (MMS), or e-mail from their mobile phones. They may also sign up to receive regular SMS reminders and updates. SMS and MMS were used as reliable communications channels between citizen scientists and a central data repository.

PBB Phone application

The PBB Phone application is an easy way for users to use the android platform to make observations about plants they would like to watch. Users can add new plants and make observations about the plants they are watching. The

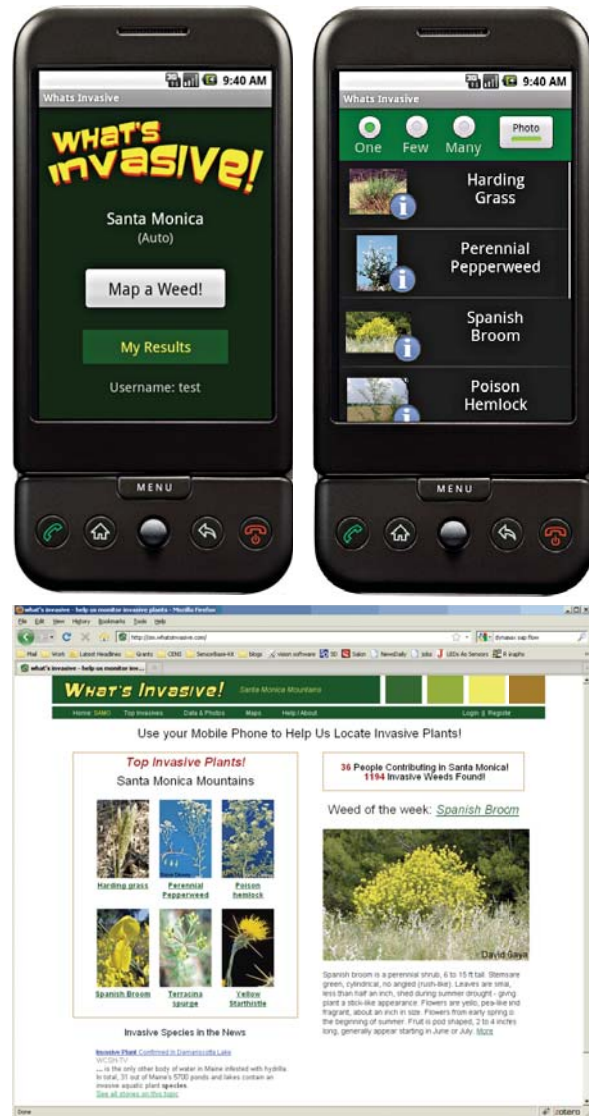


Figure 2. The data collection screens of the What's Invasive! Android application showing the main mapping button and phone location (top left), the choice of weeds to map, with quantity and photo (top right), and the What's Invasive! website (bottom), showing the top six invasive species in the Santa Monica Mountains (the same as in the phone app, based on phone location).

phone application aids the user by making simple assumptions about the data being collected. For example, when a user takes a picture of a plant for a specific phenophase, it is easy to assume that the time of the sample is the time that the picture was taken. The phone application can also upload new observations automatically, removing the intermediate step of downloading images to a computer and then uploading to the web. The phone is able to reduce the amount of information the user needs to specify and therefore makes the overall data collection process easier and more intuitive.

The data gathered by the phone can be easily synchronized with the web database using a generalized model abstraction which the android platform is currently missing. Android allows access to databases on the phone to store persistent data, however it is complicated to create useful object oriented models. In Android it is necessary to query databases for entries using SQL statements and then use the keys of these entries to get data that might be related. The model abstraction created for PBB allows for easy specifications of relations of data objects. This allows us to write code that lets us get the phenophases for a particular species much more easily (Ex. `species.phenophases()`) will give us all the phenophases for a particular species rather than having to write out two queries one to get the species and one to get the phenophases based on that query). This model abstraction can easily be used in other android projects where there is a complex data model involved.

PART 03.4 Accomplishments

The What's Invasive! project has been a tremendous success. We already have expanded from our pilot park in Santa Monica to include the Channel Islands in California, Rock Creek Park in Washington DC, three other regional parks, and help of one of our Danish graduate students, Denmark, to show off the system to a group of Europeans. We currently have a half a dozen more parks in the process of being registered.

WI is now collaborating EDDMapS (Early Detection and Distribution Mapping System; eddmaps.org), putting the location data of invasive weeds discovered by the WI system in the hands of both local environmental managers and also of national weed experts. EDDMapS started in 2005 with Southeastern U.S. focus and is now providing a picture of the distribution of invasive species across the U.S.

The WI phone application has also been translated into Chinese.

Numerous press releases have made WI their focus, including:

- About.com: <http://usparks.about.com/b/2009/10/22/new-iphone-app-targets-invasive-weeds-in-parks.htm>.
- Los Angeles Times: <http://latimesblogs.latimes.com/outposts/2009/10/national-parks-invasive-weeds-app.html>.
- National Park Service: <http://www.nps.gov/samo/parknews/whatsinvasive.htm>.
- National Parks Traveler: <http://www.nationalparkstraveler.com/2009/10/need-identify-exotic-weeds-santa-monica-mountains-national-recreation-area-theres-app4792>
- Santa Barbara Independent: <http://www.independent.com/news/2010/jan/06/cell-phone-app-combats-weeds>.
- This is Brand X (LA Times): <http://www.thisisbrandx.com/2009/10/in-the-weeds-new-iphone-app-targets-overgrowth-in-state-parks.html>.

Lessons Learned - WI

We have found that partnering with NPS officials and groups such as schools and running short-term "campaigns" has been very effective in not only education-related efforts but also in identifying and locating weeds. Although the campaign has been great from a citizen science and public education perspective, one of the biggest potential benefits to conservation work is through having park staff map weeds as they go about their daily work.

There is interest in a more developed website and user interface where people could be directed to survey in certain areas, coupled with a volunteer program that would encourage people hiking in these areas to upload the application and report sightings. Although we have had many misidentifications of weeds, mostly because it is hard for people to learn a species from looking at a photograph, one option to improve this is to link to on-line weed i.d. software that runs people through about 10 questions and then gives them possible i.d.'s for the weed.

Lessons Learned – PBB

We have found that requiring structured text in SMS messages in order to better guarantee parsing of messages and valid data entry was partly misguided. Structured text is difficult to remember without a "cheat sheet" and so is not really useful for casual observations. Thus, using text messages as a note-taking option worked better, however the ability to interpret some easily parsed sections seems useful.

A test of a hybrid What's Invasive!/Project BudBurst-like campaign with 13 UCLA students and staff volunteers resulted in 13 hours spent "in the field" and over 400 observations collected. This amounted to almost 1/8 the total

User Survey Results Table				
	Average Number of Mistakes Users Reported	Considered Easiest to Use	Considered Most Available	% of Users Who Visited >75% of Study Area Using this Mode
Web Form	1.08	0%	69.20%	84.60%
MMS	1.31	0%	23.10%	76.90%
Smart Phone	0.46	100%	7.70%	100%

Table 1. Results of a participant survey after using three methods of data collection during a hybrid What's Invasive!/Project BudBurst-like campaign with students.

number of observations made in the two-year weed survey conducted by SAMO park, emphasizing the benefits of targeted, short-term campaigns. Results (Table 1) indicate the preference of using an Android phone for data collection. Most users reported being frustrated with the format required to submit a MMS based observation in the write in portions of the survey.

PART 03.5 Future Directions

What's Invasive!

Although reception to using mobile phones for data collection has been fantastic, so far data has flowed primarily in one direction, has been put into one database, and has been relatively static – data was collected in these example applications for the sake of the one database keeping track of everything. We wish to strengthen these data collection projects by first sharing the data with as many groups as possible. For example, there are a surprising number of weed and invasive plant groups nation-wide and globally and we are working with a growing number of organizations (EDDMapS, Cal-IPC, Texas Invasives) to send What's Invasive! data to all interested parties during automated, scheduled data transfers.

Also, we wish to pull information from plant databases to display on the What's Invasive! site, such as more plant photos and descriptions, as well as up-to-date distribution information from disparate databases. We feel that sharing out data and having other groups share theirs with us will help integrate the regional efforts into a more cohesive and interesting effort.

Project BudBurst

Initial participation in Project BudBurst wanes as the season progresses, even with mobile phones and improvements to the website making data collection easier. We wish to increase interest and retention by actively using the current and historical data collected plus outside data sources to generate daily updates to users.

One of the simplest uses of the user-collected data will be to refine models for the prediction of how local weather conditions affect the plants under observation. We will pull local NOAA weather station data for each user and update a daily prediction of plant events to be sent to the user, based on known plant responses. If we don't have enough data to make a prediction, we will use the current data by the user and let the user know that they are building a prediction for their plant.

PART 04 Personal Environmental Impact Report (PEIR) Validation and Redesign

PART 04.1 Overview

The PEIR project has been running for over a year with cooperation from many students and faculty. Throughout its existence, PEIR has gone through many developments phases which has made it a black box for all. Two projects were setup in attempts to reorganize the system for future engagements. The first was a validation of the PM 2.5 exposure model, to test whether the current PEIR model is reliable. Another was the system redesign, to recreate the PEIR system architecture from scratch to increase transparency and for general improvement.

Based on the validation process, the PM 2.5 calculations were determined to be insufficiently accurate. Other types of PM 2.5 models were also considered but none of them seemed to show promising result due to lack of accurate information on traffic available in real time scale. Hence, it was decided to give up the precise calculation of PM2.5 for now and convert the system into a record holder for time within vicinity of hazardous sites of (dis)interest.

The system redesign had to incorporate this change by altering the data structure from the current GPS point system to a point and time system to calculate the duration of the time spent in certain location. Also, for a more accurate description of each GPS trace, personal information setting capability was largely expanded. As a natural consequence, the hazard model which used to be the core of the system was removed and became optional modules for further personalization.

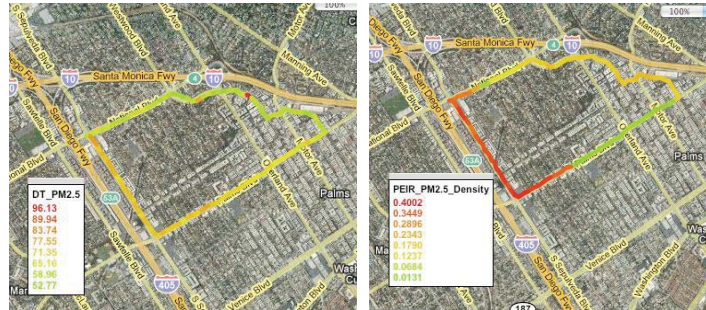


Figure 1. Plots of PM2.5 exposure estimates for a particular route. Left is the DustTrak and right is the PEIR. You can see that result from PEIR is heavily influenced by the nearby freeway, which does not necessary agree with the DustTrak measurement.

PART 04.2 Approach

There were some key issues addressed with the validation and redesign.

Previously there had been no attempt at validating any of the modeling done in the PEIR system. The initial goal was to validate the model for PM 2.5 using data collected through use of the DustTrak 8534 Aerosol Monitor as ground-truth for the model used in PEIR. Additionally, previous data collected from an UCLA Environmental Health Sciences study through use of a mobile platform was used in comparison to PEIR's model.

Simultaneously, a system redesign was taking place to improve the overall system. The major steps taken during the redesign included:

- Unifying the system design concepts
- Adding flexibility for any future upgrades
- Optimizing the flow of cell phone trace data
- Adding customization options to match user needs
- Incorporating changes to the PM 2.5 model
- Increasing portability for future demands

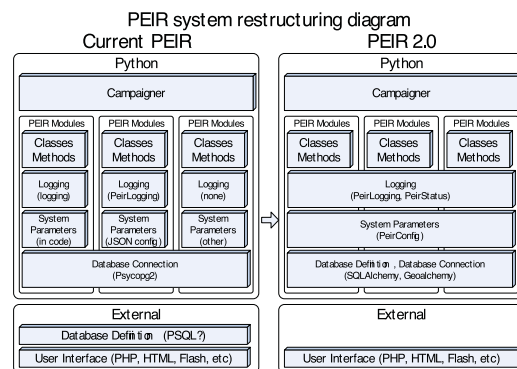


Figure 2. Comparison of system design for the old and the new system. The new system has design principle that is consistent throughout the system where as the previous system has many strands of principle that is mostly due to development history of the system.

PART 04.3 System Description and Experiments Result

PM2.5 model validation experiment

Poor correlations seen between actual PM 2.5 values and what is modeled showed the PM 2.5 model to be inaccurate. The theoretical ideas that support the model provide sound reasoning that the model will succeed, but the validation proved this to be otherwise. One reason may be the resolution at which the data can be collected. The weather data is hourly and by zip code while the traffic data is the SCAG specified long time periods. These

observations are slightly more general than the point-by-point PEIR PM 2.5 estimates used in creating the hours spent above threshold. With this resolution of observed data points, the estimates should be no smaller than hourly estimates. Also, the PEIR model is specified to only deal with localized traffic contributions to PM 2.5, while in fact a large portion of what people are exposed to comes from other sources.

There were a number of difficulties in designing the experiments to validate the model. Having a single data collection instrument resulted in always having a temporal problem when comparing two routes. Also, this validation did a large majority of sampling on the west side of Los Angeles, which historically has been shown to have lower levels of pollution. The immense size of Los Angeles, makes spatial validation basically impossible but it would be worthwhile to further sample in additional areas of LA especially places with high levels of PM 2.5.

Uploading the Health Science study data became more difficult task than first anticipated. The GPX upload code had to be fixed, an additional weather query had to be written, and the server would only allow a couple hundred data points to be uploaded at a time instead of the entire data set. The weather service had to be updated to incorporate the study data, which was from 2008. As an entire PEIR system overhaul was taking place, the additional weather service hadn't been incorporated into the PEIR system. This leaves the upload of the previous data until the new PEIR system is complete.

Following the results of the DustTrak sampling it became clear that PEIR must be careful about what it reports to a user. As it stands, the current PEIR PM 2.5 model is slightly inaccurate and the website description is a little misguided. After much thought and conversations with Arthur Winer a more simplified approach was decided upon, that allows users to still make health conscious choices about the the trips they take.

Object Relation Mapping

One of the principles of the new system was to eliminate all raw SQL. Every database transaction will go through an object relation mapper, implemented through SQLAlchemy + GeoAlchemy. This increased the flexibility of the system by adding both horizontal and vertical partitioning capability along with portability of the database.

Trip data structure change

As a result of the validation experiment, we decided to record time spent in hazardous vicinity. The new model for PM 2.5 will no longer directly estimate PM 2.5. Instead it will provide a user with time spent in hours in vicinity an interstate or other hazard. In previous studies these areas have been shown to have noticeably higher levels of PM2.5. This new model is a more justifiable method of informing users about health dangers of different trips without. The uncertainty in estimated traffic flow or temperature differences within a zip code are no longer an issue. To implement this model the data storage needed redesigning. Each GPS coordinate now contains 3-dimensional where the 3rd dimension is time. This allows us to integrate the 3rd dimension over the polygon of interest, which represents the hazardous vicinity.

Configuration unification

The new system will have central configuration files that are accessed only through a configuration module. This encapsulation of the information will increase the level of security and also make the system maintenance more transparent.

Central Logging Module

Additionally, the new system will have a logging module that will be shared across all of the components. This will improve the ease of maintenance and make the system more tolerant to failure.

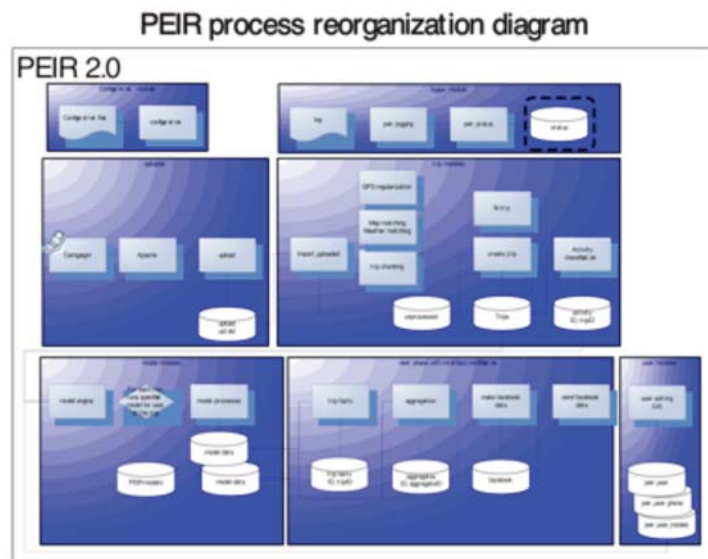


Figure 2. Simplified image of PEIR system data flow. Process is compartmentalized and the flow is simple.

Model Engine Module

The biggest change of the new system is the flexible model engine. Adding a module to the data process will be handled through the user settings. The developer does not have to understand the system in order to create a new model that calculates the statistics of interest.

PART 04.4 Accomplishments

The core of the new PEIR system is complete. User, trip, and model components are intact and ready to be used. However, there are other components that still need work in order for the whole system to be able to replace the current system. Details will be presented in the next section.

PART 04.5 Future Directions

There are still components that need to be developed for the new system to replace the current system.

Interface makeover

As a result of the big changes done on the backend, the interface must be redone to incorporate them. Adding the user customization pages is one of the big change that needs to be done. Another large change that needs consideration is how to feed the result back to the user after adapting the flexible model selection.

Trip chunking module

Under the assumption that we get better activity classification for a trip instead of doing activity classification by point, the new system expects the GPS traces to be chunked into trips before the classification. Hence location based GPS chunking module needs to be developed.

Trip classification module

Activity classification module needs to be converted into trip classification module where activity is still associated with each GPS point but uses the trip information for better accuracy in classification. Also it should incorporate more personal information in order to be able to do more classifications such as biking and taking a bus.

GPS filter module

Raw GPS uploaded from the cell phone is extremely noisy, especially for some of the old Nokia phones. Therefore, for PEIR to provide sensible results the GPS trace occasionally needs to be "fixed" in cases where the data seems unreasonable.

PART 05 Visibility Monitoring using Mobile Phones

PART 05.1 Overview

Airborne particulate matter is a serious threat to both our health and the environment. It is also the primary cause for visibility degradation in urban metropolitan areas. We are developing an optical technique to measure visibility using commodity cameras and other sensors commonly found on mobile phones. The user takes a picture of the sky which is tagged with location, orientation, and time data and transferred to a back-end server. Visibility is estimated by first calibrating the image radiometrically and then comparing the intensity with a physics-based model of sky luminance. We describe the challenges for development of the system on the HTC G1 phone running the Android OS. We have studied sensitivity of the technique to error in the accelerometers and magnetometers. Results from images gathered in Phoenix, Arizona and the Los Angeles basin compare favorably to air quality data published by the US Environmental Protection Agency.

PART 05.2 Approach

We use a physics based model of sky luminance that is a function of the viewing geometry, solar position and the air visibility, also known as turbidity. The luminance of a point p is given by

$$L(\theta_p, \gamma_p, t) = L_0 f(\theta_p, \gamma_p, t) / f(0, \theta_s, t) \quad (1)$$

where θ_p is the zenith of the sky element, γ_p is the angle between the sky element and the sun, and θ_s is the zenith of the sun. We call f the scaled luminance as it captures the ratio of true luminance to zenith luminance. It is defined as follows.

$$f(\theta_p, \gamma_p, t) = (1 + a \cdot e^{(b/\cos(\theta_p))})(1 + c \cdot e^{(d \cdot \cos(\gamma_p) + e \cdot \cos^2(\gamma_p))}) \quad (2)$$

where a , b , c , d , and e are adjustable coefficients. Each of the parameters has a specific physical effect on the sky distribution and with different values they can capture a wide variety of sky conditions. An empirical model of the parameters in terms of turbidity t is as follows.

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} 0.1787 & 1.4630 \\ -0.3554 & 0.4275 \\ -0.0227 & 5.3251 \\ 0.1206 & -2.5771 \\ -0.0670 & 0.3703 \end{bmatrix} \begin{bmatrix} t \\ 1 \end{bmatrix}$$

By comparing the image intensity with the luminance predicted by (1), we can estimate turbidity.

There are three main challenges in this approach.

1. The image intensity is a non-linear function of the true radiance. This response function, varies from camera to camera and is typically estimated using multiple images taken at varying exposure levels. Most phones do not allow us to control the camera exposures. Radiometric calibration using a single image is not accurate.
2. The sky model does not account for clouds so we have to segment sky pixels from the image.
3. The image intensity only captures relative luminance.

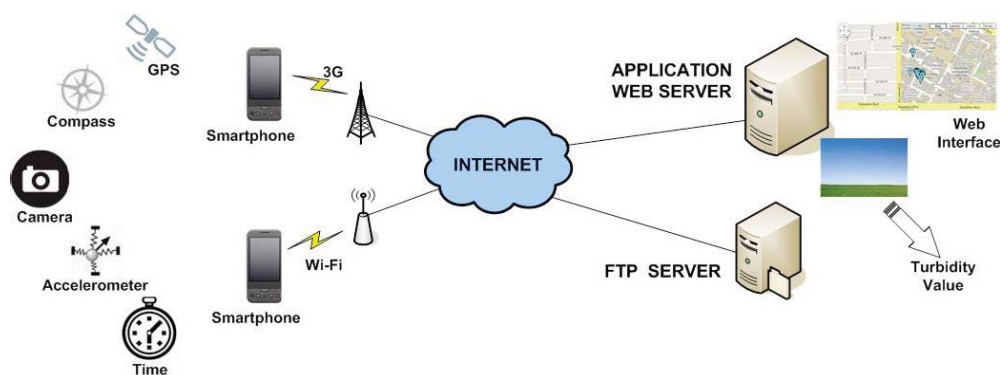


Figure 1. Overview of the system

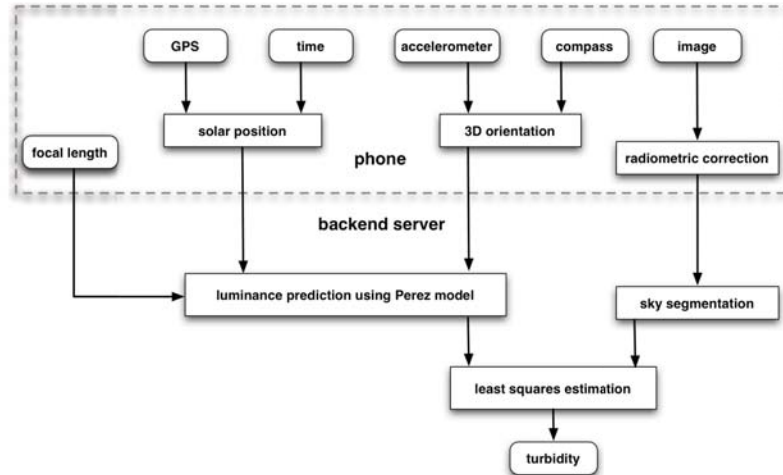


Figure 2. System architecture

PART 05.3 System Description and Experiments

Figure 1 shows the high level system architecture. The user contributes an image tagged with sensor data which is processed at a backend server and the resulting turbidity estimate returned to the user.

Figure 2 shows the sequence of computations. Solar position is calculated on the phone using the GPS and time information. The accelerometers and compass sensors are used to calculate the orientation of the phone. Radiometric calibration is performed using a single image as a one-time computation and the response function is stored in the phone. The function is applied to the image before sending it to the backend server. Luminance is computed using equation (1) and turbidity is calculated using a standard non-linear optimization technique to minimize the difference between the model and the image intensity for sky pixels.

Figure 3 shows screenshots of the application developed on Android software platform.

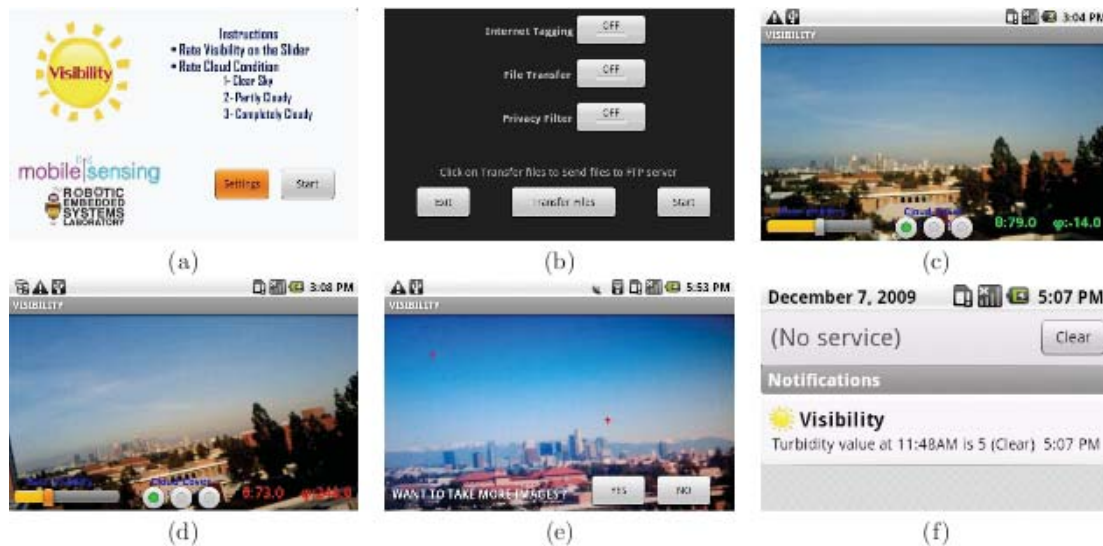


Figure 3. The Visibility application for Android OS. (a) startup screen with options to view/edit settings and start taking the picture. (b) settings for communication and privacy. If internet tagging is turned on, the application will gather weather data. File transfer option allows the user to choose between transferring the image immediately or at a later more convenient time. Privacy filter controls whether the user's GPS data is communicated. (c) Camera preview. The azimuth and zenith angles are displayed in green when the roll is < 50 and red (d) otherwise. (e) The user chooses a portion of sky for processing. Clicking the camera button at this point stores the image along with orientation data. (f) The computed turbidity value is returned as a notification

PART 05.4 Accomplishments

The Arizona Department of Environmental Quality (ADEQ) maintains several cameras near Phoenix. We used images from a 2.4 megapixel camera located in the North Mountains looking south. The pictures from this camera are published every 15 minutes. Figure 4 shows three examples of images from the South Mountain camera that are reported to have good, fair and poor visibility by the ADEQ. The corresponding image intensity surfaces and the scaled luminance profiles for the values of turbidity (2, 3 and 6) estimated by our algorithm are shown. The turbidity values increase as the visibility decreases and the luminance surfaces match well with the image intensity profiles.

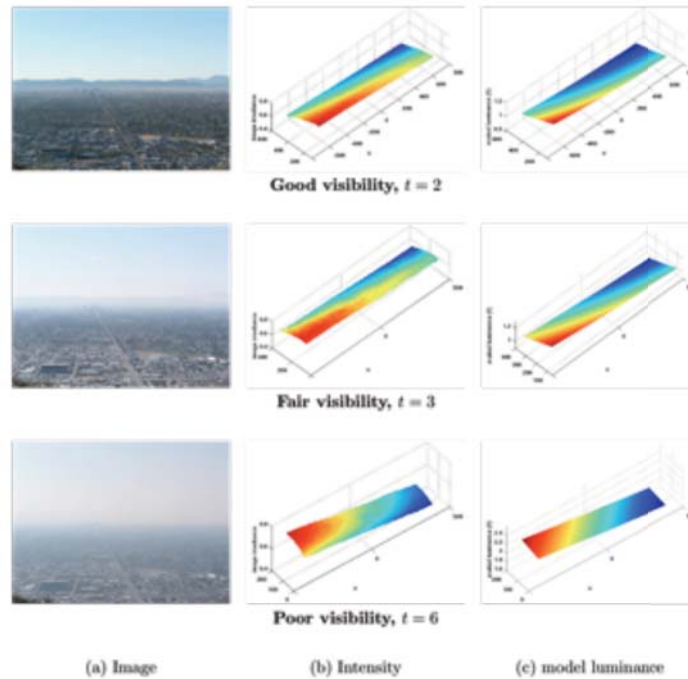


Figure 4: Results from south mountain camera, in Phoenix, Arizona

PART 05.5 Future Directions

Recently, we set up an image station on the roof of a building at the University of Southern California (figure 5). It has an android phone placed inside a weather-proof box. It looks northeast at the Los Angeles downtown and logs images every 15 minutes. In future, we plan to conduct an extensive data collection campaign and explore two directions

- combining turbidity estimates from multiple images taken from the same location
- generating a large-scale map of turbidity from images taken from different locations

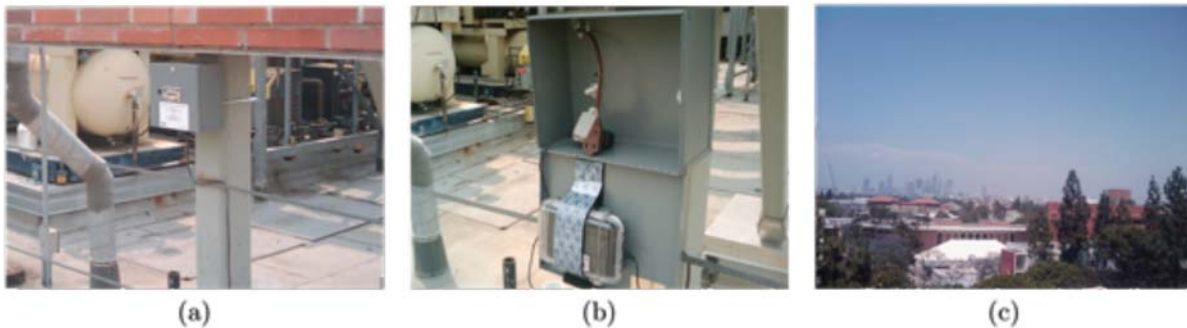


Figure 5. USC rooftop camera station.

PART 06 Ethics in Personal Mobile & Participatory Sensing

PART 06.1 Overview

The mobile phone network is emerging as the largest sensor network on the planet. Mobile phone users, however, are generally unaware of the dual uses of this network, in which their communication devices are also information gathering devices. What are the ethics of coordinating this network for research purposes? Can researchers achieve meaningful consent and active participation of mobile phone users?

The three-year Ethics in Personal Mobile & Participatory Sensing research and education project is:

- Researching a participatory approach to managing privacy in personal mobile sensing applications;
- Creating curricula to teach participatory ethics for urban sensing to diverse STEM undergraduate and graduate students;
- Disseminating best practices for education in participatory urban sensing ethics to urban sensing, ubiquitous computing, and broader technology education communities.

PART 06.2 Approach

Participatory sensing systems gather many kinds of data for diverse health, wellness, and environmental outcomes. The Ethics project explicitly focuses on exploring the social impacts of pervasive data collection. The project asks what the humanities (philosophy, ethics, literature) and social sciences (STS, surveillance studies, technology policy) have to contribute to the design process and to ethics education for STEM graduate students.

The project explores approaches to privacy in mobile sensing, which is a large and obvious challenge of pervasive data collection. In addition, the project asks students to consider ethics beyond privacy. For example, consent is a central tenet of research ethics. How do we promote meaningful consent in participatory sensing? Questions of power and equity are also important to consider. Accumulating & manipulating data is both a form of power and a route to mitigating social disparities. Who owns sensing data or benefits from sensing projects? Persistent memory is also a potential consequence of participatory sensing. Should we keep sensing data forever? Archived data could reveal trends, but also document things we'd like forgotten. Data representation is also important to consider. Sensing data must be processed and interpreted to be understood. How do we translate data into legible forms while explaining phenomena such as error and bias? Finally, how can sensing tools and architectures support both individual behavior change (such as addressing overeating or preventing smoking) as well as broad social changes (improving city planning or revising food policy)? Facilitating responsible, socially trusted, and participatory ethics for data collection and analysis with urban sensing systems remains an open problem, and is the challenge undertaken in this research and education project.

Building Ethical Reflection Into Design

We focus on graduate and undergraduate students who are designing systems not just “for the future” but for ongoing pilot projects that have public participation. Our project encourages these students to build ethical reflection into their design process. We use laboratory meetings and one-on-one interactions to encourage students to reflect on the ways their ethical choices influence sensing development. We are also investigating ways in which funding streams, institutional regulations (such as IRB requirements), mentoring relationships, and interactions with system users affect students’ awareness of ethical challenges.

PART 06.3 Research Design

During the research component of this project, we are formalizing and assessing a test case in participatory research ethics: a privacy framework we call participatory privacy regulation. Over the last year we have formalized the framework, outlining both rules and architectures to enable people to participate in their own privacy regulation



Figure 1. A few participatory sensing challenges beyond privacy

(Shilton et al., 2009). Assessing the framework will occur over the next year of the project using user surveys, focus groups and interviews.

We have also drawn upon previous interdisciplinary approaches to participatory ethics as well as our investigation of participatory privacy regulation to develop two curricula for STEM undergraduate and graduate students: a hands-on pilot project approach to building participatory urban sensing technologies and an interdisciplinary seminar course extending and debating participatory ethics in urban sensing and ubiquitous computing.

To develop the laboratory course, we are conducting long-term ethnographic observation of design practices at CENS, and are conducting interviews with students, staff and faculty responsible for creating participatory sensing technologies. We are also analyzing project blogs and wikis as well as published posters and papers. Coding documents, field notes and interview transcripts to detect the institutional structures and practices that encourage or discourage ethical discussions and decision-making can help determine which practices promote ethical learning in the laboratory environment. The pilot project curriculum will enhance existing CENS campaigns to draw students into design projects and discussions focused on ethical system development.

This ethnographic data, combined with interdisciplinary readings and discussions in computer science, privacy, and ethics, have also helped to shape the seminar course. The seminar course will engage students from CENS as well additional students from fields such as computer science, electrical engineering, statistics, information studies, environmental studies, media studies, geography, law, political science, sociology, and philosophy. The nature of course participation will highlight design as well as other disciplinary approaches to participatory urban sensing ethics.

PART 06.4 Accomplishments

During the past year, we have conducted interviews with all of the students and staff involved directly with participatory sensing projects (n=21). We have begun coding these interviews to reveal the institutional structures and CENS design practices which influence ethical discussion and debate during the participatory sensing design process. These findings will be critical to designing the lab-based ethics curriculum during year three of the project. This year we also designed a curriculum for a seminar class entitled “Mobile Technologies: Participation and Surveillance.” This class will be taught and evaluated by Katie Shilton during the spring of 2010. Finally, we began the user interview process with volunteers who piloted the Biketastic participatory sensing system.

The team has also focused during the last year on creating new technologies to further our approach to participatory ethics. Our research team used participatory principles to outline requirements for the Personal Data Vault (PDV), a secure repository designed to help participatory sensing participants manage their data. We published the design principles in the fall of 2009 (Shilton et al., 2009), and will publish the technical specifications for the vault in the coming year. We are also collaborating with Jerry Kang, a professor in UCLA’s School of Law, to determine legal protections for participatory sensing data held in a Personal Data Vault. More information on the PDV can be found elsewhere in this annual report.

Another project of the ethics research team has been institutionalizing ethical decision-making at CENS. We have implemented a process by which all CENS pilot projects that gather data about people are evaluated by the ethics team for ethical and human subjects concerns. We are also engaging with UCLA’s IRB to address new challenges to privacy and consent posed by participatory sensing projects.

We have also refined our research questions and curricular ideas at several workshops focused on engineering ethics. Katie Shilton served as a planning committee member for the “Ethical Guidance for Research and Application of Pervasive and Autonomous Information Technology (PAIT) workshop organized by Indiana University in March 2010. At this workshop, case studies from CENS participatory sensing projects served as discussion prompts. Katie also presented “Personalized Discovery or iSurveillance? Ethics and Innovation in Mobile Sensing” at the Society for Social Studies of Science (4S) annual meeting in October 2009. Finally, Katie was an invited participant at the Workshop on Ethics in Science and Engineering: Redefining Tools and Resources held by the University of Massachusetts Amherst in October 2009.

We have also made several presentations to educators to disseminate ideas cultivated as part of the ethics project. Katie presented a talk entitled “Participating in Privacy: Designing the Personal Data Vault” for the UCLA CENS Seminar Series in July 2009. And she presented the talk “Privacy and Other Challenges: Disclosure, Discretion and Mobile Sensing for the CENS/Google Workshop for area high school teachers in July 2009.

PART 06.5 Future Directions

As CENS builds a core of users, we will engage in interviews to answer the following questions:

- How deeply, and under what conditions, do participants engage with participatory sensing systems?

- How do participants in urban sensing negotiate decisions to capture, share, and retain their data?
- How well does participatory privacy regulation support privacy and sharing decision-making in participatory urban sensing systems?
- What are other key ethical questions in participatory urban sensing?

Qualitative data documenting interactions between participants and urban sensing systems can suggest answers to contextual questions about when and why participants make decisions to share or withhold data. Interviews with participants can elicit how participants feel while interacting with the systems and how much participants trust the systems. We will use explicit participant critique of our design methods, software, and conclusions to answer our second research question and assess the adequacy of participatory privacy regulation as an ethical framework. Do participants feel comfortable and secure using participatory urban sensing systems? What changes would they recommend? Are any ethical concerns unaddressed? Answering these questions through interviews and focus groups will help us evaluate participatory privacy regulation from the ethical perspective of those who matter most: the individuals and communities using urban sensing systems.

We will also continue to publish the results of our collaborations on legal and institutional (IRB) protections for participatory sensing. And during the final year of the project, we will finalize and publish two unique curricula: one for engaging STEM students in their design practice, and one for an interdisciplinary seminar on participatory ethics and mobile technologies.

PART 07 Putting a Spotlight on Personal Resource Consumption

PART 07.1 Overview

Natural resource preservation has become a significant concern. Positive feedback which provides detailed resource consumption for each person, along with appropriate contextual information, is necessary. The Spotlight project aims to develop an affordable, easy-to-use resource monitoring system that monitors fine-grained resource consumption in buildings. The main goal of the project is to provide general users with an easy means to monitor their own resource consumption in their spaces. This year we develop an affordable easy-to-use appliance level power monitoring system by exploiting the fact that appliances emit measurable signals when they operate.

PART 07.2 Approach

The current generation of energy reporting devices only provide partial and coarse grained information or require expensive professional installation. This limitation stems from the presumption that calculating per-appliance consumption requires per-appliance current measurements. While several commercial in-line sensors have emerged, they do not simultaneously satisfy three key design criteria: comprehensive coverage, fine-grained reporting, and easy-to-install. For example, Cent-a-Meter, EM-2500 and TED devices monitor the household power consumption but do not provide per-appliance level measurements. Kill-A-Watt and Watts Up devices provide finer granularity but require in-line installation between a standard AC plug and the outlet. While it is possible to instrument many appliances in this way, some of the major energy consumers cannot be easily instrumented. For example, most heating and ventilation systems (HVAC) and electric boilers do not have standard AC plugs, or are hard-wired to the main power lines. Ceiling lights are another example.

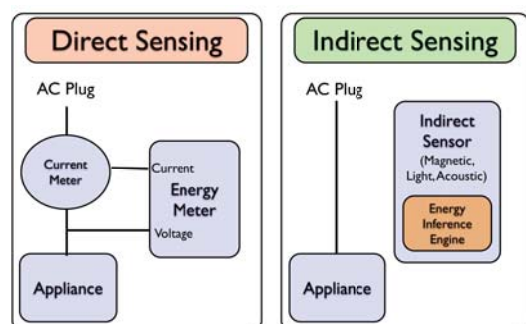


Figure 1. Direct and Indirect Power Monitoring

However, since appliances typically emit measurable signals when they are consuming energy, we can estimate their consumption using indirect sensing (Figure 1). The project team develops a fine-grained power monitoring system that furnishes users with an economical, self-calibrating tool that provides power consumption of virtually every appliance in buildings. The principle of operation is a network of wireless distributed sensors monitoring signals that appliances emit and forwarding them to a personal computer acting as a back-end fusion center. The fusion center collects data from the indirect sensors and measurements from the main power meter, and runs a model-based machine learning algorithm that automatically learns and estimates power consumption of every appliance on-the-fly.

The system design criteria avoid the traditional approach of using complex and carefully trained and calibrated sensing hardware for monitoring resource consumption and occupant. Instead, simple hardware for opportunistic sensing of information about resource consumption and resident activity is used with smart algorithms and models to train and calibrate the hardware and to make inferences with minimum user intervention, which is essential for an economically scalable and easily installable system.

PART 07.3 System(s) Description and/or Experiments

System Description

A prototype implementation consists of several non-intrusive sensors: a main meter monitor, light intensity sensors, magnetometers, and microphones (Figure 2). This implementation uses CrossBow MicaZ motes, HMC1002 magnetic sensors, and MTS310 sensor boards. Each battery-powered sensor node monitors energy consumption related signals, and sends them back to the fusion center. The fusion center is a PC that solves a numerical optimization problem, combines data from the distributed sensors, and profiles appliance-level power consumption according to unique node IDs. It's important to note that the system doesn't require in-line sensors nor explicit sensor calibration. This key feature allows general users, who do not have technical expertise, to easily install monitoring devices.

Experimental Result

To test and validate the design concept, we conducted several experiments in a 2-person apartment. We chose three cases with increasing complexity. In Case 1, we estimate power consumption of a desktop computer, a table lamp, and a refrigerator using two different sensor configurations: (1) A magnetometer on the power cord to the computer, a light sensor near the table lamp, and a light and acoustic sensor to monitor the refrigerator. (2) A magnetometer on the power cord of each of the three appliances.

For Case II, we add an alarm clock, a wireless router, and a cable TV set top box to the Case I configuration. The goal is to test the system with unmonitored appliances, and thus the power states for these three additional devices was not monitored. In Case III, we let the system estimate the average power consumption of 9 different appliances by monitoring only on/off status of the appliances.

The power monitoring system takes the total power consumption, magnetic field, and internal power states information from heterogeneous sensors including magnetic, light and acoustic sensors. It then solves a regression problem to compute calibration parameters. The calculated calibration parameters are used to estimate appliance-level power consumption. The estimated power consumption(left bottom) tracks the true power consumption(right bottom) very well (Figure 3).

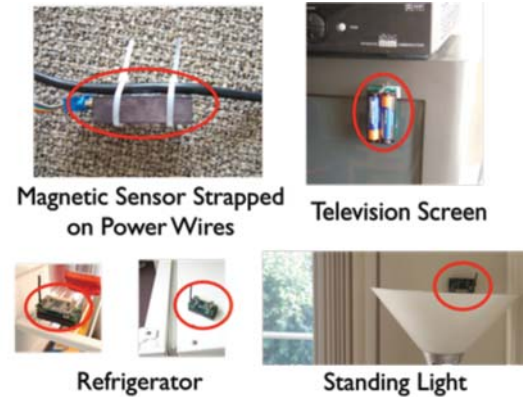


Figure 2. Non-intrusive Appliance-level power monitoring system deployment

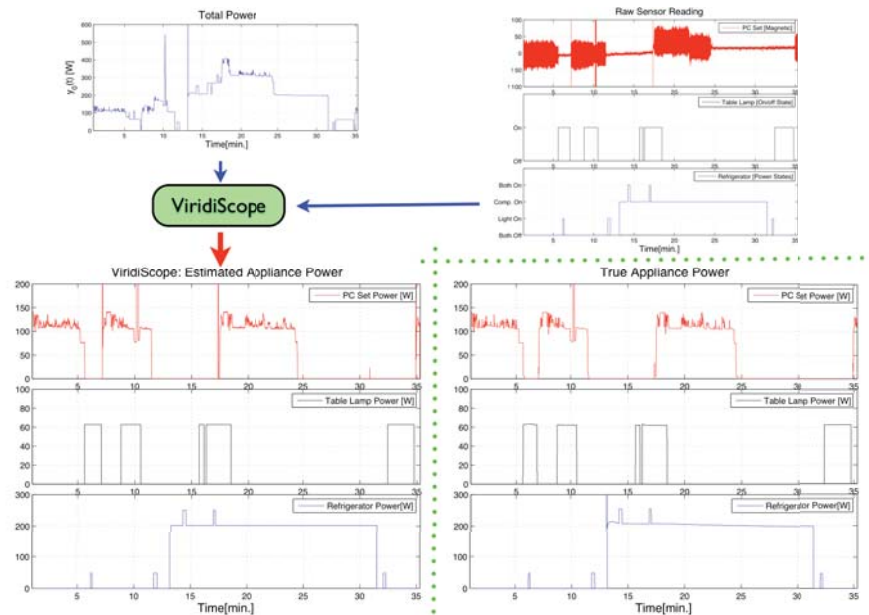


Figure 3. Non-intrusive Appliance-level power monitoring system evaluation

If we are only interested in the average power consumption of each appliance, it is enough to know only the on/off status of an appliance. Therefore, we use light or acoustic sensors to instrument the appliances. To test the accuracy, we test the algorithm using 9 different appliances. Table 1 summarizes the outcome. The error is consistently less than 10%.

PART 07.4 Accomplishments

We implemented and evaluated a fine grained power monitoring system using the combination of existing infrastructure and indirect sensors. By using indirect sensors, we extended the traditional power monitoring dimension. Indirect sensing, however, introduces a new sensor calibration challenge. By exploiting the existing infrastructure, we develop an autonomous sensor calibration scheme that automates the sensor calibration procedure. Our experiments show that the system tracks

Appliance	True Power[W]	Estimate[W](Error)
Light 1	59.8	58.4(-2.32%)
Light 2	31.9	34.2(7.02%)
Massage Chair	46.2	45.8(-1%)
Table Lamp 1	68.1	69.3(1.73%)
Table Lamp 2	14.7	15(2.54%)
Water Heater	1623.6	1620.6(-0.18%)
TV	100.1	101(0.94%)
Laptop 1	38.9	38.3(-1.45%)
Laptop 2	31.4	29.2(-7%)

Table 1. Average Power Consumption Estimation

the per-appliance power consumption with less than 10% error. In addition, the system can easily monitor the power consumption of appliances with simultaneously active appliances as well as variable power consumption.

PART 07.5 Future Directions

One problem of fine-grained resource monitoring is the vast amount of data. Unlike multi-media data, the data is in plain time-series. This is inherently hard for users to understand unless proper representation and summaries are provided. Appropriate data visualization and representation need to be considered to make this information effective.

The real-time consumption trace, when combined with the contextual information such as activity information, can help users understand the true implication of their resource consumption. Therefore, we plan to develop an activity and identity inference technique based on the finer grained resource monitoring subsystem. Correlating resource consumption per occupant is an essential step for solving the problems associated with detecting and identifying people and their activities. We will revisit the problem of human detection and identification in the context of energy and resource consumption by exploiting existing sensing. We seek to find a higher-level algorithm that will decide how to associate end-point resource usage with the occupants and their activities so that they can draw concrete conclusions on their consumption.

An easy means to query the real-time consumption is also an important component of the system. We plan to use a mobile phone as an in-situ consumption display unit. For example, the current generation of mobile phones have sufficient capabilities to take pictures of appliances, and to perform a computer-vision algorithm. The Spotlight system acts as a real-time database which monitors and stores real-time appliance-level power consumption. A mobile phone essentially takes real-time images of appliances, recognizes them, queries appliance-level consumption, and displays the data on its screen.

PART 08 GeoSIM: An Urban Sensing System for Social Image Mapping of Urban Geolocations

PART 08.1 Overview

The advent of earth visualization tools (e.g., Google Earth, Microsoft Virtual Earth) has inspired and enabled numerous applications. Some of these tools already include texture in their representation of the urban environment. The urban texture consists of the set of images/photos collected from the real environment, to be mapped on the façade of the 3D model of the environment (e.g., building and vegetation models) for photo-realistic 3D representation. Currently, urban texture is collected via aerial and/or ground photography (e.g., Google Street View). As a result, texture collection/documentation is (1) expensive, (2) unscalable (in terms of the required resources), and (3) with low temporal and/or spatial resolution (i.e., texture cannot be collected frequently and widely enough).

These limitations can be addressed by leveraging the popularity of camera-equipped mobile devices (such as cell phones and PDAs) for inexpensive and scalable urban texture documentation with high spatiotemporal resolution. We envision GeoSIM (short for Geo Social Image Mapping)¹ as an urban sensing/documentation system with which a group of individuals with camera-equipped mobile phones participate in collaborative/social mapping of the urban image (i.e., the texture of the urban environment) at some target geolocation (see Figure 1). The participating group, which may either consist of dedicated individuals or the general public, are directed to capture geotagged images of the urban environment. The collected images are progressively used for documentation of the dynamic urban scene in multiple spatial resolutions and at different times.



Figure 1. The GeoSIM Vision

Originally, we introduced GeoSIM as a motivating application (and system) to pursue the following research objectives:

- Developing scalable planning techniques to coordinate participatory collection of urban visual data; and
- Developing efficient solutions for on-the-fly fusion and mapping of the collected visual data to generate 3D urban texture.

During this reporting period, we noticed that new initiatives by industry partially (yet not fully) address our second objective by enabling texture generation using pre-existing images; e.g., the Streetside application recently released by Microsoft uses the Photosynth technology to fuse Flickr photos for texture generation. Therefore, during this reporting period we prioritized and decided to focus on the first (complementary) research objective that enables active and participatory image collection where there is no pre-existing images available. We emphasize that texture generation using images that are collected on-demand (i.e., our second research objective) still remains an open research problem for us to pursue, although perhaps we can benefit from the recent industry initiatives in developing a solution for this problem.

PART 08.2 Approach

Below, we summarize our research approach to address each of the two research objectives mentioned above, respectively:

- We introduce a two-phase planning technique for participatory data collection. At the first phase, termed viewpoint selection, a minimum number of points in the urban environment are selected from which the texture of the entire urban environment (the part accessible to cameras) can be collected/captured. At the second phase, called viewpoint assignment, the selected viewpoints are assigned to the participating users such that given a limited number of users with various user constraints (e.g., specific participation time) users can collectively capture maximum amount of texture information within a limited time interval. We first prove that both viewpoint

¹ <http://infolab.usc.edu/projects/GeoSIM/>

selection and viewpoint assignment are NP-hard problems, and therefore, optimal solutions for these problems fail to scale as the extent of the urban environment and the number of participating users grow. Subsequently, we propose a family of heuristics for efficient viewpoint selection and viewpoint assignment that enable fast planning while ensuring almost complete texture collection. We study, profile and verify our proposed solutions comparatively by both rigorous analysis and extensive experiments.

- For effective texture mapping, GeoSIM must enable merging the collected images in real time, in order to generate the evolving texture of the urban environment as new images arrive. Reconstructing generic scenes using uncalibrated images is a known difficult computer vision problem. To achieve this objective, we take a pragmatic approach by developing a stack of pipelines to be used depending on the available preexisting data. If a low resolution texture is available together with an underlying geometric model, we attempt to merge the incoming images by matching image features and incorporate the existing geometric information. In the simplest case, supposing that locally the surface of an object is planar, a robust estimation of the underlying image homography should suffice in most cases to register an incoming image with an existing texture. We integrate the geotag of the images and use the line-of-sight query to locate the faces of the objects that are visible. This computationally inexpensive approach allows for quick registration of the incoming images for already mapped zones, and also supports merging new images when these overlap with sufficient features. On the other hand, if no texture is available for a given object, but a geometric model is present, we attempt to carry out a Euclidian reconstruction using all the available images on the area and incorporate the preexisting geometrical information. The system handles requirements in terms of how many images are needed to validate a given object texture. Finally, if no information is available we attempt to reconstruct a set of plane surfaces solely based on the image information. This is the most difficult and generic case. In all cases, we use other available data sources in conjunction with the object recognition algorithms on the images to detect and segment objects such as trees that may cause undesirable results in the reconstructions and textures.

PART 08.3 GeoSIM Prototype

Our proposed solutions (see above) have been experimentally verified (using the testbeds that we developed), peer-reviewed and published. Furthermore, to verify and demonstrate the performance of our proposed solutions in real world, we developed a prototype of the envisioned GeoSIM system, called GeoSIM v1.0, which implements our proposed solutions.

We developed GeoSIM v1.0 on top of GeoDec2 (our geospatial data management engine), in accordance with the original plan included in our research proposal. Figure 2 shows the data processing workflow of GeoSIM v1.0. As depicted in the figure, the workflow of GeoSIM v1.0 consists of an off-line process as well as an on-line process. At the off-line process, which is solely executed by GeoSIM server, a minimum set of viewpoints are identified to cover the entire targeted geolocation. Moreover, for each selected viewpoint a set of desired camera shots is identified by defining the camera direction, pitch and tilt for each shot, and each shot is simulated based on the corresponding (texture-less) geometry of the environment. At the on-line process, once a user/participant uses the GeoSIM client to join the data/texture collection campaign, the GeoSIM server generates a participation plan to be executed by the participant. The user participation plan consists of a traversal path and a set of viewpoints along the path at which the user should make stop to collect the desired camera shots. The collected shots/images are assessed for quality as soon as they are obtained (while the user is traversing the path) and sent back to be stored and processed at the GeoSIM server. At the server side the texture generation engine progressively retrieves the collected images and maps them to the corresponding geometry/model of the target geolocation for on-the-fly texture generation.

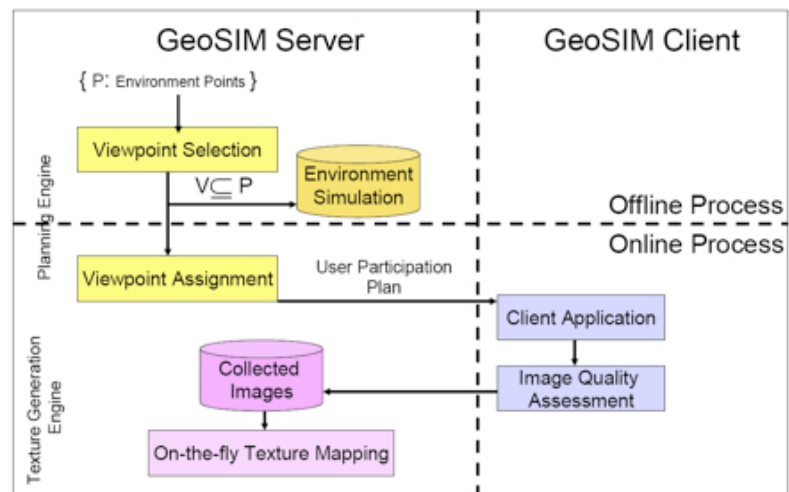


Figure 2. Workflow of GeoSIM v1.0

² <http://infolab.usc.edu/projects/geodec/>

A video demo of GeoSIM v1.0 is available at: <http://infolab.usc.edu/projects/GeoSIM/Activities.php>

PART 08.4 Accomplishments

We summarize our accomplishments during this reporting period as follows:

- Introducing a family of efficient viewpoint selection techniques for visual data collection
- Introducing a family of efficient viewpoint assignment techniques for visual data collection
- Developing an end-to-end GeoSIM prototype
- Publishing three peer-reviewed conference papers to disseminate the results of our research on the problems of viewpoint assignment and viewpoint selection, as well as one journal paper to describe GeoSIM v1.0 as a social image mapping system

PART 08.5 Future Directions

We intend to pursue this research in several directions:

- Developing GeoSIM v2.0, an enhanced GeoSIM prototype with scalable implementation (in terms of the number of supported concurrent users) to enable large-scale data collection and field test
- Providing location-privacy with GeoSIM (also applicable with any other participatory data collection system): With this feature, the location of the user is considered as private information that is concealed from the GeoSIM system. This is complementary to the existing privacy studies at CENS (which focuses on defining the privacy ethics, policies, etc.) by providing the "tools" to guarantee privacy.
- Incorporating participatory data transfer capabilities into GeoSIM (also applicable with any other participatory sensing system) to enable data transfer when the typical one-hop communication (e.g., Wi-Fi or 3G) is not available: With this feature, participants not only help in data collection, but also build a p2p network (e.g., via Bluetooth connections) to transfer the collected images to where the data can be transmitted to the GeoSIM server.
- Integrating GeoSIM with other tools and technologies developed at CENS to enable high-level mixed reality applications.

PART 09 Remapping LA: Cultural Civic Computing in Los Angeles

PART 09.1 Overview

Since the creation of its participatory sensing area (previously urban sensing), CENS has collaborated with the UCLA Center for Research in Engineering, Media and Performance (REMAP) on its “civic cultural computing” project Remapping LA, a collaboration with California State Parks and others to contribute to the renewal of Downtown L.A. by enabling communities to power the processes of imagining, specifying, and designing technology for their neighborhoods and public spaces. Activities having included (1) Technical assistance for UCLA courses on “engaged media production,” in which students produce media and collect data using mobile, web, and other tools—including participatory sensing systems—that explore Los Angeles neighborhoods; (2) community workshops and participatory media art and technology projects with a variety of groups, from the youth Anahuak Soccer Federation to the Gay and Lesbian Elders Housing; (3) experimental media artworks on the history and geography of Los Angeles, including the dynamic, outdoor digital media mural Junction / Juncture (2007) with Disney Imagineering Research and Development and the text message driven urban mapping activity Engage Ludicity (2009). Each of these projects leverages common technology from other CENS participatory sensing projects (e.g., CENS Campaignr software on mobile handsets; SMS communication; web-based GIS; UCLA’s metropolitan WiFi network) as documented in this and past annual reports. Collaboration with REMAP provides a cultural perspective on the use of mobile technologies, and has generated its own unique systems and applications research.

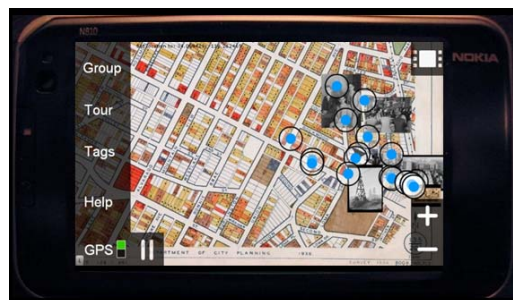


Figure 1. Locative Media Explorer, map view.

Most recently, through support from Nokia Research Center Hollywood (NRCH), CENS collaborated with REMAP to further develop a “Locative Media Explorer” software platform for the Maemo Linux based N810 tablets from Nokia. This research explores the “flip side” of participatory sensing: the experience of location-based content and data on the same mobile devices that can be used to collect it. This software, now in its second generation, was used in the field by the Pilipino Workers’ Center and Los Angeles-based consultancy Public Matters to create the “Mobile HiFi Tours” described below, which provided valuable feedback on the running system. Further work has been funded by NRCH to unite the experience of previously captured media and data explored in this project with mobile handset-based participatory data collection as supported by the CENS software Campaignr.³

PART 09.2 Approach

In this project, “location-based media” refers to content displayed or played back on a mobile device based on the physical location of that device, as determined by GPS or other location technology. The design intent of the Locative Media Explorer (LoME) is to make the authoring of a location-based media experience as accessible to nonprofessional authors as possible by automatically assembling tours. Media (image, audio, sound, video, text) is created and uploaded to an open source web-based gallery platform, where geotags, keywords, and additional metadata can be easily added. (Such a database could in the future be easily populated by a one of the participatory sensing systems described elsewhere in this report.) Special tags provide the ability to specify parameters such as the physical area (radius from the geotag) in which the media is available and its ‘priority’ for display, if more media are available than can be displayed. A custom server-side synchronization script automatically downloads the media, creates thumbnails and icons, and extracts the metadata to create both a media file system and corresponding database. Another series of scripts push this media and database to the tablets for experience by the users. (Wireless distribution is not yet feasible due to tablet hardware limitations and lack of pervasive WiFi in areas of interest.)

The client software runs on the Nokia N810 Maemo Linux internet tablet and provides map- and full-screen media interfaces to this location-tagged, keyworded media. An Adobe Flash user interface connects to Python-based support components for socket-based access to the database and the tablet’s GPS. A

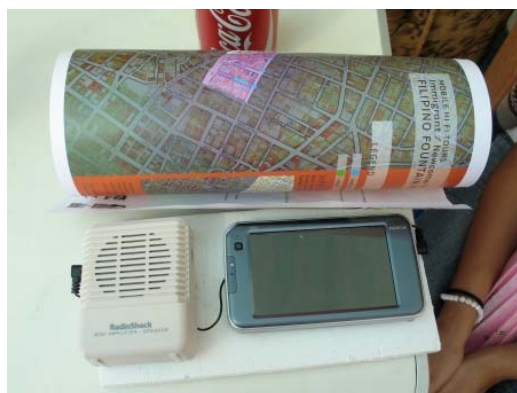


Figure 2. Tablet, prepared with external speaker for use in the Mobile Hi Fi (Historic Filipinotown) Tours.

³ <http://campaignr.com/>

substantial amount of effort was dedicated to stabilization and performance tuning of the platform due to the tablets' limited resources (for example, 128MB of RAM). Also, a new Maemo port of the Spatialite database was created by the project team to support the second generation of the platform. This provides spatial query capability to the Flash user interface, enabling more efficient processing of the location-based queries used to select which media to show at any given time.

Media are automatically organized into “tours” based on a special type of keyword tag. Each tour can have a custom tiled map, legend, and help screens. In map view, the tablet displays thumbnails and icons for the highest priority media near the user (taking into account the radius specified by the author), allowing the user to selectively view media of interest. In full-screen mode, media are continually played back based on this priority. Database queries are made continuously (approx. every 5-10 seconds) to update the available media based on location and other parameters.



Figure 3. Locative-media equipped Jeepney on tour.

PART 09.3 Experiments

The first generation of the LoME software was used in the 2008 Hollywood festival curated by LA Freewaves, and featured media authored by several community groups on or near Hollywood Boulevard. The second generation of the software, described above, was used in 2009 by youth from the Pilipino Workers' Center (PWC) and nearby Belmont High School to create tours about Historic Filipino town. Led by the local consultancy Public Matters, in collaboration with REMAP and CENS, this project, called Pdup Productions⁴, provided the youth with training and tools to create locative video, imagery, and audio about Historic Filipinotown's history and culture. Four different tours were created and tested on the platform, with a collection of custom maps that included historic Works Progress Administration (WPA) maps and hand-drawn maps of the area.

PART 09.4 Accomplishments

Public deployment

Thirteen tablets provided by Nokia were loaded with Pdup content and the LoME software, then given out for public tours of Historic Filipinotown in a day-long public event at PWC chaired by Los Angeles City Council President Eric Garcetti. In addition, a modified version of the system was run in a GPS-enabled Jeepney brought from the Philippines, as shown in Figure 3. This provided substantial, if informal, experience in the use of the system. Additionally, the preparation of the media tours in collaboration with the youth and Public Matters helped to further refine our approach to the the authoring of location-based media and, by extension, how data—especially geotagged images—that are collected in participatory sensing might be explored after they are captured.

PART 09.5 Future Directions

In the next year, with support from Nokia, we plan to integrate the experience of this project into exploratory design of a new user interface approach that combines participatory sensing, such as that enabled by Campaignr, with the location-based experience of collected data. Additionally, the project team will do exploratory integration of the platform with the UCLA HyperCities web-based mapping system⁵ and has generated an NSF STEM education proposal to continue the work with youth. Finally, the software may be incorporated into the Remapping LA project's upcoming work on new media experiences at the Los Angeles State Historic Park, which recently received seed funding from with California State Parks.

⁴ <http://hypercities.com/pdup/>

⁵ <http://hypercities.com/>

PART 10 Urban Tomography

PART 10.1 Overview

This project is developing an “Urban Tomography” system for capturing geo-tagged videos on video-capable cellphones and automatically sending them to a back-end server infrastructure using wireless networking technologies such as EDGE/GPRS or 802.11b/g/n. Our system is designed to enable pervasive dense audiovisual documentation of city life. As our neighborhoods become increasingly diverse and complex, such documentation can enable a better understanding of social interactions and the use of urban spaces. It can also help urban planners re-structure existing cities in order to improve their quality of life.

PART 10.2 Approach

We call our system Urban Tomography, a technological framework that enables the collection and creative navigation of a large corpus of audiovisual urban documentation. Underlying this framework is the observation that relatively cheap, ubiquitous, internet-connected, mobile personal devices such as cellphones and PDAs are increasingly equipped with good video capture technologies and GPS. These technological advances promise pervasive sensing of urban phenomena: we envision swarms of trained volunteers spreading out across a cityscape and regularly recording urban processes for extended periods of time. Casual users can also contribute to the corpus. These audiovisual records can be transmitted over the network for near instantaneous analysis by anyone with access to the global Internet.



Figure 1. Viewer page of Urban Tomography

To enable Urban Tomography, two advances are necessary. At the “back-end”, the acquisition, transmission, storage, and indexing of audiovisual records must be made as automatic and transparent to the user as possible. A simple

back-end will require less training and thereby enable more pervasive sensing by more people. However, this is technologically challenging because of the scale of the problem: a potentially large number of video clips introduces technical challenges in robustly distributing the corpus to ensure high availability, and organizing information to enable intuitive navigation and fast searching of the corpus. The second set of advances requires audio-visual display technologies that simultaneously present inter-related urban processes with the aim of challenging the viewer to find previously unsuspected relationships between the perspectives. This is analogous to tomography, which attempts to simultaneously explore multiple perspectives (“slices” or “cuts”) of an object (an organ, a geophysical feature) in order to reconstruct the object in its entirety. In our case, we are exploring techniques such as split-screen and multi-channel sound to present these multiple views in a creativity-enhancing way.



Figure 2. Video archives for Katrina Project.

PART 10.3 System(s) Description and/or Experiments

Our prototype system consists of two major components: the capture and transmission software on the cellphone, which captures video files and automatically (without user intervention) transfers them to a server. The server contains a database that stores all video files and its metadata, as well as a simple web viewer which displays a group of videos as “small multiples” using tiles.

Our capture subsystem is designed for the Nokia N95 phone. This phone has a 2GB micro-SD card, supports 640x480 high resolution video and has “Assisted” GPS functionality. Our videos are stored on a commodity server machines running the Apache web server and a mysql relational database backend. In turn, the capture subsystem includes a user interface process which allows the user to access the camera, configure the application, or examine the application log. The main user interface process on the capture subsystem also associates the current video recording with metadata. This metadata is stored as a separate file, and contains timestamp and GPS location associated with the captured video clip (more precisely, the GPS location of the phone at the end of the clip). A background process is responsible for transferring video files to the server. It periodically scans a designated location

on the local file system, and if it discovers any video files, attempts to transfer the video to the server system. This happens automatically and without user intervention. Since wireless connectivity can vary dynamically, and large file uploads are more likely to be interrupted by intermittent wireless connectivity, the background process breaks large video files into chunks to ensure forward progress for file upload. In addition, the background process monitors available forms of network connectivity (whether the GPRS network is available, or which of many hotspots are available). It prefers to use a WiFi hotspot if one is available, but otherwise attempts to use the GPRS network.

PART 10.4 Accomplishments

Our system has been operational for several months. We provide the system to more users as following.

- Aug. 09 to the present: Deployed to PCFA (Pacific Child & Family Associates), they are offering behavioral analysis and related services for children and adults with developmental disabilities.
- Mar. 09 to the present: security application in a transportation site in Los Angeles is still operational.

The project webpage <http://tomography.usc.edu> also includes several other field experiments.

Finally, we have been working on a couple of technical problems. First one is about understanding energy-delay tradeoffs in smartphone applications. The other one is about energy-efficient rate-adaptive GPS-based positioning scheme. As a consequence, two technical papers are submitted to conference proceedings and currently under review.

PART 10.5 Future Directions

There are several research challenges in this project.

- The N95 has short battery lifetime. Under continuous video file upload, it loses its complete charge in about two hours. While battery life will improve with future generations of phones, there is scope for energy-saving techniques: putting the phone into energy-saving modes when a network is unavailable, choosing good quality network connections to avoid energy wasted in lost-packet retransmissions, and so forth.
- A well-known problem with GPS, of course, is that it does not work inside buildings and other obstructed structures. Our software currently tags captured videos with the last obtained GPS reading, which is currently acceptable for our purposes.
- Availability of audio-visual information is important to some users. The problem here is high quality video often occupies from tens of Mbytes to hundreds of MBytes, but wireless network bandwidth is usually not able to achieve required latency. So for those users, the algorithm that can efficiently summarize audio-visual data needs to be developed.
- There are interesting challenges in being able to synchronize tiled 6 to 9 videos — OS overhead adds noticeable delays between tiles.

These shortcomings motivate several interesting future directions of research, some of which we intend to pursue in the coming year.

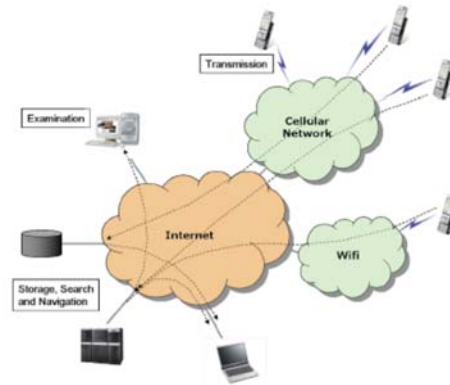


Fig. 1. The Urban Tomography System

Figure 3. System Network Overview.

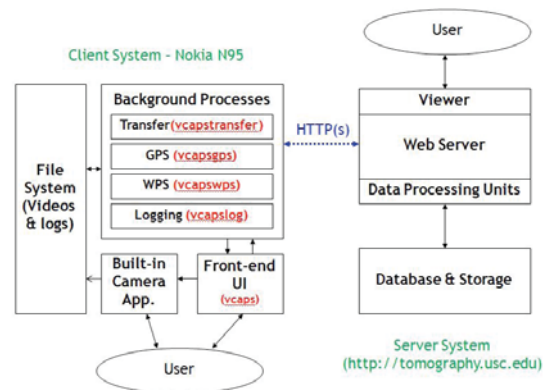


Figure 4. VCAPS system architecture.

PART 11 Walking Speed Estimation with Gaussian Process Regression on Inertial Sensors

PART 11.1 Overview

Nearly half the United States population (45%) is expected to be obese by 2020, decreasing life expectancy and quality of life significantly. It is believed that weight gain could be prevented by achieving small changes in behavior, such as 15 minutes per day of walking. Walking is the most common type of activity among people who are physically active. Walking speed is a critical component in energy expenditure and monitoring energy metabolism. In the face of the current obesity epidemic, findings suggest that encouraging walking habits in the general population could be effective in preventing and reducing obesity. Accurate detection of walking speed could be a valuable tool in enhancing public health efforts. Over the past decade, there has been considerable research directed towards the detection and classification of physical activity patterns from body mounted inertial sensors. Typical inertial sensors contain an accelerometer with two or more gyroscopes to provide kinematic information. The emergence of MEMS based inertial sensors has the potential to be able to revolutionize physical activity tracking by providing ubiquitous tracking capabilities.

PART 11.2 Approach

Periodicity in Walking

Steady state walking is cyclic. Our approach involves capturing this inherent periodicity from a single inertial sensor worn above the iliac crest on the right hip. Fig. 1 shows a typical plot of the signals received while walking at a constant speed of 2.5 mph. Movement data are captured in the form of six time series using a tri-axial accelerometer and tri-axial rate gyroscope. These signals correspond directly to the accelerations and rotational rates of the hip as felt by the sensor in its local frame of reference.

We use Gaussian Process based Regression (GPR) to find the correspondence between feature vectors and the speed of walking. GPR was chosen because it represents a data driven regression method. The utility of GPR stems from its ability to define a probabilistic model over data, mitigating the effects of overfitting and avoiding cross-validation. GPR is described in terms of kernels, avoiding the explicit introduction of a feature space allowing us to use feature spaces of infinite dimensionality thus allowing non-linear mapping. We compare our approach to Least Squares Regression (LSR) and Linear Regression (BLR). LSR offers a baseline performance comparison with GPR. BLR allows comparison of GPR to a probabilistic parametric model. BLR was chosen to avoid the issues of cross validation and over fitting (BLR also incorporates a probabilistic framework). We restrict our experiments to straight line walking to isolate the effect of speed on inertial sensor data.

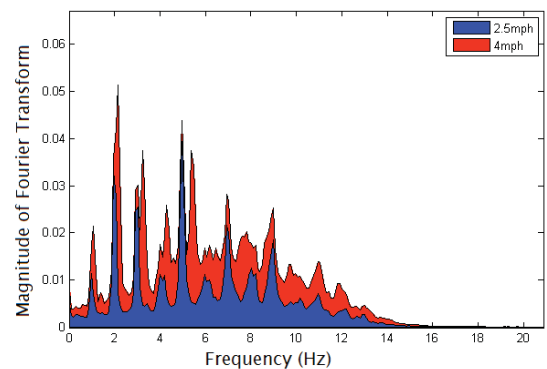


Figure 1

PART 11.3 System(s) Description and/or Experiments

Hardware and Data collection

A modified version of the Sparkfun 6DoF Inertial Measurement Unit (IMU) [21] was used to collect motion information. Data were sampled at 100 Hz. The unit uses Bluetooth to transmit data to either a nearby PC or mobile phone. The use of sensors in all three axes allows us to capture periodicity in all three planes – sagittal, frontal and transverse. Eight healthy adults (four men, four women) of varying heights, weights and ages (subjects 1-8) walked at 7 predetermined speeds (2.5 mph, 2.8 mph, 3.0 mph, 3.3 mph, 3.5 mph, 3.8 mph, 4.0 mph), or until breaking into a run. The duration of walking at each speed was 5 minutes. All subjects wore a single inertial sensor above the iliac crest on the right hip. The treadmill used for the experiments was the research quality NordicTrack A2550 PRO. Subjects were deliberately chosen to represent a cross-section of heights, ages, Body Mass Indices (BMI) and both genders to demonstrate the utility of GPR across a diverse population. Ground truth for treadmill walking was the displayed treadmill speed.

Feature Computation

Each signal was passed through a bandpass filter with 3dB cutoff between 0.1 Hz and 20 Hz. These cut-off frequencies were chosen keeping in mind that everyday activities fall in the frequency range of 0-10 Hz. The feature vector is computed on sliding windows with 50% overlap by finding their N-point FFT. The optimum window size was



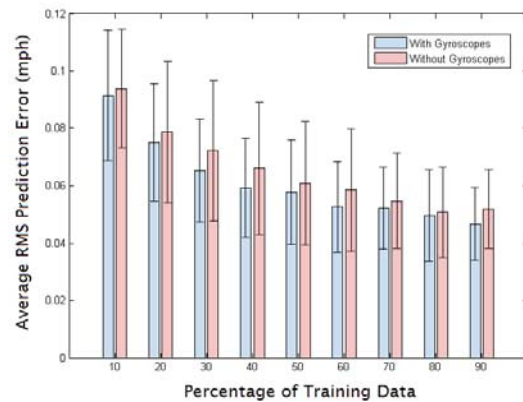
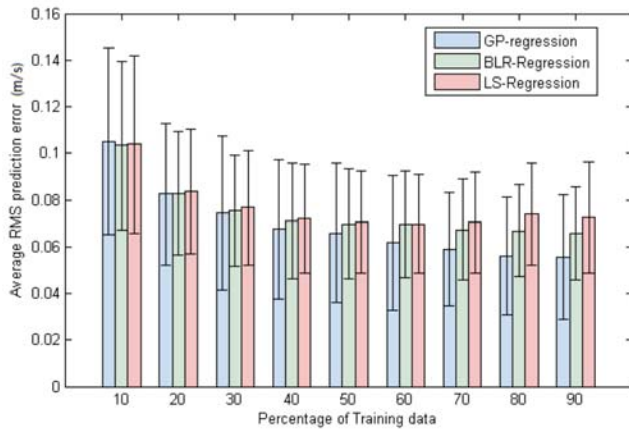
Figure 2

chosen in hindsight from experimental results as 1024 samples per window over 512 samples. The complete feature vector consists of the Fourier transforms of the respective window instants for each sensor stream.

PART 11.4 Accomplishments

Our results showed that GPR had a lower average RMS prediction error when compared to BLR and LSR across all subjects. Analysis on a single subject showed that GPR had significantly lower RMS prediction error than LSR with BLR showing comparable but higher errors). Increase in relative percentage of training data greatly improved estimation of accuracy of both GPR and BLR with LSR displaying effects of over-fitting. Using a window size of 1024 samples resulted in a lower error across users as compared to using a window size of 512 samples because of an increased resolution in frequency features. The addition of tri-axial gyroscopes reduced the RMS prediction error of walking

speeds when compared to using only accelerometers. Prediction across all speeds was not uniform. Using GPR to estimate overground walking speeds from overground motion data alone resulted in reduced error with increase in relative percentage of training data. A strong linear correlation existed ($r_{X,Y} = .8861$) between overground walking speeds predicted from treadmill data and ground truth walking speed measured. Combining treadmill data from multiple subjects with similar height characteristics improved the prediction capability of GPR for overground walking



speeds.

PART 11.5 Future Directions

We plan to address inter-subject variance by performing an extensive study on the training and prediction of treadmill walking speeds for a much larger population by categorizing individual models in terms of physiological parameters such as age, height, gender and BMI. By doing so, we aim to explore whether each of these models can be organized as clusters, each of which is a function of these parameters. This would imply that to characterize subjects and to ensure accurate prediction for each subject, one would need to derive the corresponding parameters and map the subject to a cluster. The prediction model would be unique to that cluster. As more data are collected for new subjects, the accuracy of each cluster would gradually improve. Also, from a data collection perspective, it is easier to record small datasets from a large pool of subjects than a large dataset from a single subject. The use of clusters would facilitate data collection by grouping people with similar physiological parameters in the same equivalence class. Using clusters might also enable extrapolation to new subjects who do not fall in a particular category. Finally we plan to explore the use of our techniques to map feature vectors directly to a measurement of energy expenditure. If successful, this will involve learning a data driven functional mapping from the periodicity of walking to energy expended thus bypassing walking speed estimation.