Architecture and Applications of Virtual Coaches

Daniel Siewiorek, Asim Smailagic, Anind Dey
Quality of Life Technology Engineering Research Center
Carnegie Mellon University

1. The Evolution of Technology

The confluence of several new technologies enabled a new generation of always attentive personalized systems called Virtual Coaches. A Virtual Coaches continuously monitors its users activities and surroundings, detects situations where intervention would be desirable, and offers prompt assistance.

Presently available cognitive aids are simplistic, providing only scheduled reminders and rote instructions. Future virtual coaches will actually monitor user performance of activities and provide appropriate feedback and encouragement. As the user's abilities change, the coach may reduce the number of verbal cues as the subject learns, or provide increased support as needed. A care provider could upload new capabilities to the virtual coach, as required, potentially without even an office visit. Virtual coaches also provide constant and consistent observation/monitoring, even on a clinician's guidance beyond episodic patient examinations.

Virtual Coaches are the latest phase of a technology evolution over the past two decades. The advent of powerful microprocessors capable of running an operating system with real time responsiveness in small, energy efficient pages in the early 1990's enabled a new generation of personal computing systems that provided access to information any time, any where. Handheld Personal Digital Assistant (PDA) that could fit in a shirt pocket gave access to addresses, notes, and schedules via a new interface access modality featuring stylus and hand writing recognition (e.g. graffiti) and more recently touch screen and voice control (e.g. SIRI).

Another novel technology, head mounted displays, enabled revolutionary new body worn systems, termed Wearable Computers [Siewiorek, Smailagic, Starner 2008], that were always on providing instantaneous access to reference information in application areas such as complex plant operations, manufacturing, maintenance, and group collaboration.

MEMS (Mircro-electro-mechanical systems) created low cost, low energy sensors that could sense physical parameters such as acceleration, orientation, temperature, and light that, when coupled with signal processing and machine learning algorithms allowed personal systems to infer user context in Context Aware Systems.

Section 2 provides background on one class of applications for virtual coaches - cognitive aids demonstrating the potential user population. Section 3 introduces the elements in virtual coach architecture with a summary of how the elements are configured in the example virtual coaches presented in later sections. The following two sections provide two detailed examples of virtual coaches that implement two basically different coaching models – rule based that derives the model through end user involvement in the design process and machine learning that derives a statistical model through labeled sensor data.

Section 6 provides three more coaching examples while Section 7 provides conclusions and future challenges.

2. Desirable Attributes of Virtual Coaches as Cognitive Aids

One important application domain for virtual coaches is in assisting individuals whose own cognitive capabilities have been impaired due to natural aging, illness or traumatic injuries. Recent estimates indicate that over 20 million Americans experience some form of cognitive impairment. This includes older Americans living alone (~4M), people with Alzheimer's (~4.5M), people with mild cognitive impairments (~6M older adults), survivors of stroke (~2.5M) and people with traumatic brain injury (~5.3M). Of the many challenges faced by older individuals, declines in memory and cognition are often most feared and have the largest negative impact on themselves and their family members

Cognitive aids currently available are simplistic, providing only scheduled reminders and rote instructions. They operate open-loop without regard for the user's activities or environment. In contrast, Virtual Coaches monitor how the user performs activities, provides situational awareness and gives feedback and encouragement matched to their cognitive state and circumstances at the time. Consider the difference between a medication reminder that blindly sounds an alert everyday at noon versus a Virtual Coach that both realizes a user took their pill at 11:58 or in another situation, such as when they are having a conversation, and sets itself to vibrate mode.

Other transformative features of a Virtual Coach include:
• As the user learns, it reduces the number of and level of detail in the cues it provides;
• It matches its level of support to the user as his abilities change;
• A caregiver can upload new capabilities to the Virtual Coach, as required, without even an office visit; and,
• It provides constant and consistent monitoring of adherence to a caregiver's instructions, enabling a deeper and more timely understanding of conditions beyond the episodic patient examinations available today.

Virtual Coaches based on understanding of user situations and needs are also effective for applications aimed at larger populations. For example, cognitive support can assure safe use and compliance with instructions in rehabilitation and management of chronic illness. Many individuals are released from hospital to home with inadequate training for themselves or their family caregivers for the operation of newly prescribed home medical devices or following complex medical regimens. Failure to properly follow directions often results in expensive (to the insurer) re-hospitalizations. Similarly, understanding how to effectively motivate people toward healthy behaviors, such as proper diet and physical activity, can benefit broad segments of the general population. Virtual coaches can monitor for compliance, provide cognitive assistance, provide advice that is trusted and followed, and adapt to user capabilities that vary with time and circumstances. The next section discusses the architectural elements common to all virtual coaches.

3. Virtual Coach Architecture

As indicated above, currently available aids are simplistic, providing only scheduled reminders and rote instructions. Virtual coaches monitor user performance of activities and user context (Sensor Processing), determines appropriate feedback (Coaching Model), and provides feedback and encouragement (User Engagement). A care provider could upload new capabilities to the virtual coach, as required (Prescription). Over time a customized personal interaction evolves (User Interaction). Figure 1 depicts how these five elements interact to create a Virtual Coach.

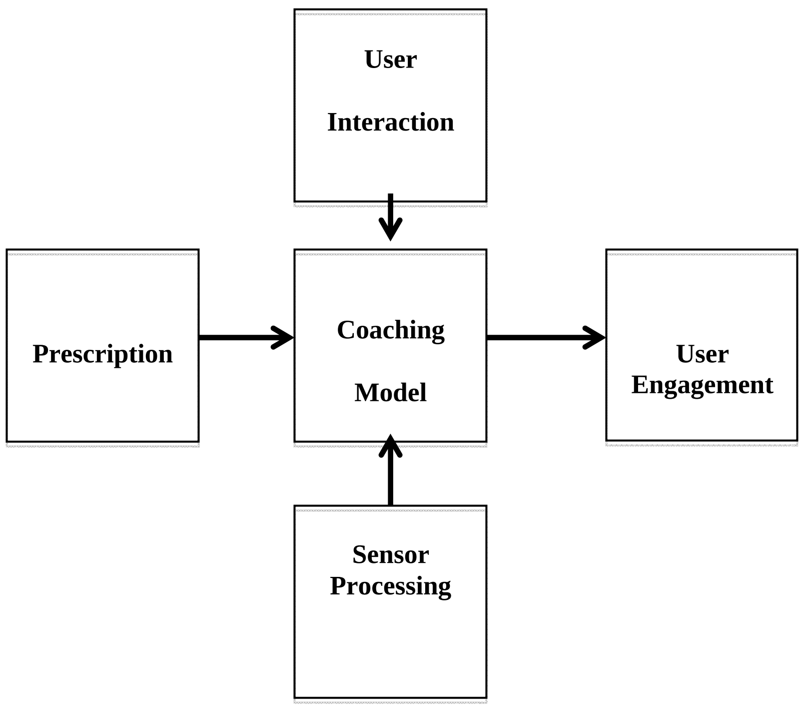


Figure 1. Inter-relation of five elements of virtual coaches.

The remaining sections will illustrate examples of these five elements through five example virtual coaches:

- Power Wheel Chair (PWC) Seating Coach: rule based model whose input is prescription of positions and durations; sensors monitor positions and durations; Avatar provides reminders to do past due activities
- Ergo Buddy: machine learning to diagnosis potentially harmful lifting practices in package delivery occupation; employs multiple sensors, and sensor fusion; provides audio warning
- Manual Wheel Chair Propulsion: machine learning for diagnosis of stressful propulsion motions; warning suppressed if context appropriate for motion type

- IMPACT (Improving and Motivating Physical Activity Using Context): rule based model; uses contextual information to support reflection and improve motivation
- Mem-Exerciser: Alzheimer Disease Memory Reminder Coach provides cues to remember experienced events; events sensed by still pictures, GPS, accelerometers, audio, and light

Table 1 summarizes how the coaches implement each of the five architectural elements. The most basic difference between coaches is the Coaching Model. Rule Based models require extensive engagement of end users (patients, care givers, and clinicians) during the design process to insure capture of the relevant situations. Section 4 will provide an example of this process with the Power Wheel Chair Seating Coach. On the other hand, machine learning uses examples (labeled training data) to create a statistical model of the activities. Section 5 provides detailed discussion of deriving these models for Ergo Buddy. Section 6 provides summaries of the three other virtual coaches in Table 1.

Coach	**Coaching Function**	**Architecture Elements**				
		Prescription	**Sensor Processing**	**User Interaction**	**Coaching Model**	**User Engagement**
Power Wheel Chair Seating Coach	Monitoring for Compliance	Spread Sheet	Tilt, pressure, accelerometers, IR	Touch screen, joy stick	Rule Based	Avatar, Audio
Ergobuddy	Warning	Supervised Training	Accelerometers	Physical Activity	Machine Learning	Audio
Manual Wheel Chair Propulsion Coach	Correct Form	Supervised Training	Accelerometers	Propulsion Motion	Machine Learning	Audio
IMPACT	Motivation	Written Instructions	Pedometer, GPS	Diary	Rule Based	Graphical
Mem-Exerciser	Recall	Automatic Sampling	Camera, Microphone, GPS, Accelerometer, Light	Replay	Care Giver	Pictures, Audio

Table 1. Example Virtual Coaches and their embodiments of the five architectural elements.

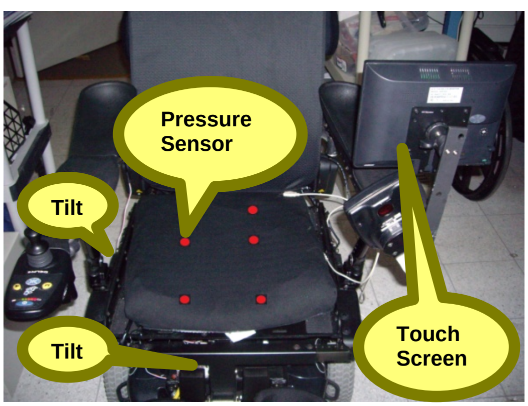
4. Power Wheel Chair Seating Virtual Coach – A Rule Based Smart Reminder for Power Seat Function Usage

Patients with spinal cord injuries have lost feeling in the lower parts of their body. They must shift their positions periodically to prevent the occurrence of pressure sores. Once pressure sores occur they are very difficult to heal.

The power wheelchair virtual coach is an intelligent system that guides power wheelchair users in achieving clinician established goals for body positioning. An array of pressure, tilt, and IR sensors provides data to the virtual coach which monitors user compliance with the clinician's goals and generates reminders for doing past due activities. Clinicians and power wheel chair users were part of the design team from the first day.

Power seat functions (PSFs) allow the user to recline, tilt, elevate the seat and elevate leg-rests of the chair. Tilt indicates that the entire seating system is shifted backward, but the angle between the back and seat remains constant. Recline changes the backrest angle only, and leg-rest elevation changes the leg-rest angle. The seat elevation raises or lowers the individual in a seated position.

An array of pressure sensors is distributed over the backrest and seat cushion providing the pressure information to the virtual coach, as shown in Figure 2. Three tilt sensors determine the tilt angle of the backrest, seat recline, and leg-rest elevation, as illustrated in Figure 3. Infrared sensors are used to detect obstacles behind the chair and determine the height of the seat. Pressure sensors are monitored for weight distribution inferring body positions.



(a)

(b)

Figure 2. Power Wheel Chair Seating Virtual Coach sensors and touch screen: (a) front view and (b) rear view.

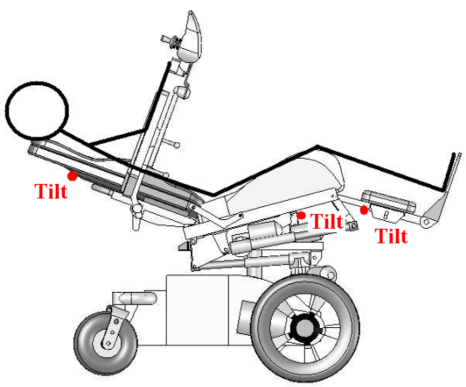


Figure 3. Tilt function and placement of sensors

Tilt, recline, and leg rest elevation are monitored for any improper sequences in using seat functions, such as reclining the backrest without tilting the seat, elevating the leg-rests without reclining the backrest, and recline or tilt angles that are too large, as well as any inappropriate use of seat functions during driving. The user interaction and sensor monitor software run on an embedded computer, attached to the back of the wheel chair seat.

An important design feature recognizes a barrier to wider adoption of Quality of Life Technology (QoLT) systems. Clinicians and other caregivers are responsible for assistive technology systems, instruct users on desired behaviors, and assess user's adherence to instructions. Systems to support these functions for clinicians and other caregivers neither exist or are not well-defined. For example, realistically we cannot expect caregivers such as physical therapists to become software programmers. Thus, interactions with the software must use familiar vocabulary and formats and be flexible enough to accommodate a variety and range of caregiver specialties and individuals

After interviewing clinicians, several attributes of a prescription were identified:

• Activity: Indicates the power seat function to be performed. It also includes the pressure activity that is not explicitly performed by the user, but is the result of using the chair.
• Parameter: The minimum, ideal, and maximum values per function
• Duration: Each activity (except pressure) is to be performed for the ideal duration. However, it is not considered a violation if the duration is between min and max values. Only the max duration for the pressure activity is valid and this indicates the maximum time for which the pressure reading can be above the max value of the parameter.
• Gap: This value represents the time after which each activity (except pressure) is to be repeated.
• Alert After: This value indicates the number of rule violations, after which the notification action takes place.

An attribute:value pair approach was selected where-in the clinician fills in the value cells of a spreadsheet (Table 2) [Siewiorek, Smailagic 2008]. It is interesting to note that through students soliciting inputs from clinicians for the Power Wheel Chair Seating Virtual Coach, the clinicians changed their practice to a more repeatable process.

Data analysis software extracts underlying user patterns. A clinician-friendly interface allows therapists to prescribe rules for proper use of the wheel chair, as well as parameters for user compliance goals. To illustrate how a user would comply with one of the rules, we describe the use of the feet elevation rule:

1. The user tilts to an angle between the min and max of the general tilt angle, aiming for the ideal specified angle.
2. The user then reclines to an angle between the min and max of the general recline angle, aiming for the ideal value.
3. Now, the user elevates the leg-rest to an angle between the min and the max in the feet elevation activity parameter, aiming again for the ideal value.
4. The user maintains this position for the duration specified in the prescription.
5. This completes the compliance of the feet elevation rule and the user can wait for more reminders or resume daily activity

Clinician settings, user data, and sensor data is stored in a database, and a web service component securely transfers data from the clinician's computer to chair-side system. A web portal is designed to provide quick access to all frequently needed information to a clinician.

After entering a usage prescription, the clinician can periodically monitor the wheelchair user's compliance to those recommendations. An example alert as seen by the clinician is shown in Figure 4. The shape of daily, weekly or monthly reports in the form of Kiviat graphs making it easy for the clinician to quickly determine the progress of each user, as shown in Figure 5. Reminders are generated to prompt the user to comply while alerts indicate non-compliance and are sent to the user, as shown in Figure 6.

Activity	Parameter			Duration			Gap			Alert after
	Min	Ideal	Max	Min	Ideal	Max	Min	Ideal	Max	
Tilt	25°	30°	35°	25 sec	30 sec	35 sec	20 min	30 min	2 hrs	10
Recline	10°	15°	20°	4 mins	5 mins	6 mins	4 hrs	5 hrs	6 hrs	15
Feet Elevation	25°	30°	35°	50 sec	1 min	1min 10 sec	1hr 30 mins	2 hrs	2 hrs 30 mins	20
Pressure	0	60mm	200mm	0 sec	0 sec	30 mins	0 sec	0 sec	0 sec	5

General Tilt angle : Min 10, Ideal 20, Max 30

Table 2: Sample prescription, filled by clinician

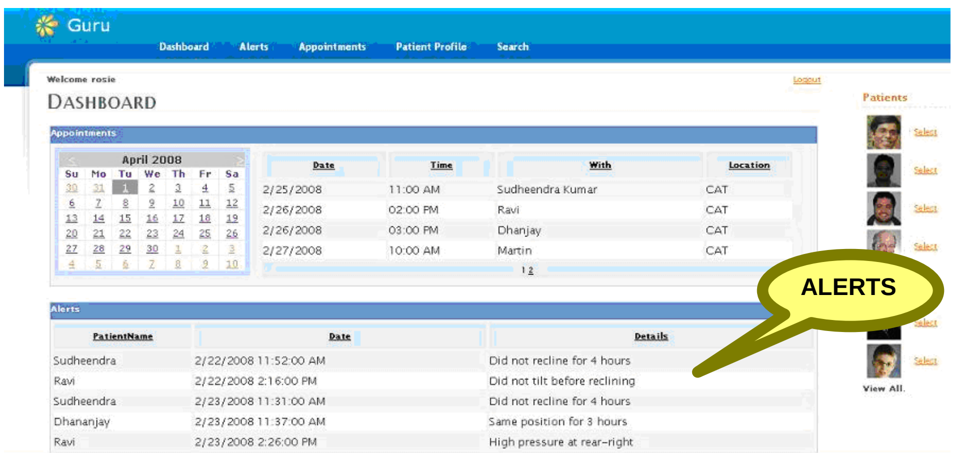


Figure 4. Dashboard for the Clinician showing non-compliance alerts

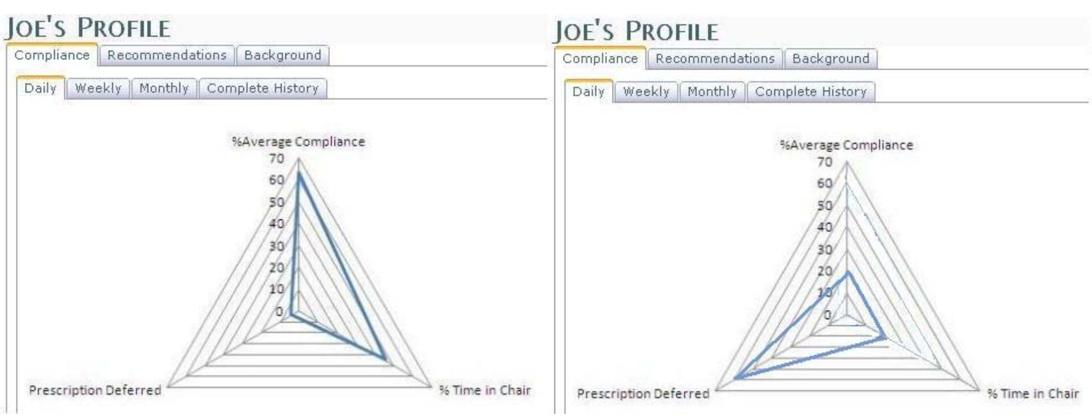


Figure 5. The Clinician can look at daily, weekly or monthly graphs of the wheel chair user's compliance and follow the progress of each user. The shape of the triangle should be orientated towards the right indicating user compliance. On the left, an example of good compliance is shown. On the right, an example of poor compliance is shown.

Figure 6. Virtual Coach screen with a reminder to tilt. The compliance graph shows color coded hourly compliance.

4.1 Laboratory Test

An important part of the design process is to obtain user feedback. The feedback can be gathered in the laboratory even before the Virtual Coach is operational. A Wizard of Oz (non functional mock-ups) user preference study was used to determine the appropriate interface modalities and coaching strategies. A survey program (Figure 7) was created allowing participants to select different interface modalities/stimuli for four types of coaching scenarios: Reminding (e.g., when a user forgets to change the seating position for an hour), Warning (e.g., when a user accesses power seat functions in an incorrect sequence), Guidance (e.g., when a user attempts to access pressure relief positions), and Encouragement (e.g., when a user responds to the message with appropriate actions) [Liu et al 2010].

Nine participants who use power wheelchairs equipped with PSFs and six clinicians experienced in prescribing power seat functions showed that speech was the most frequently selected modality for the reminding theme, and beeping was the most frequently selected modality for the warning theme. Most subjects gave monotonic speech the lowest ranking. Male face animation received the lowest ranking. Most subjects gave cartoon animations or PSF task animations higher rankings than human face images. The participants preferred to have cartoon animation to inform them of the task they need to do, as they are funny and entertaining. They also preferred to have the animated power wheelchair figure to illustrate the instructions for the specific task, which not only conveys the essential point of a message, but makes them feel it is important to follow the instructions. Many power wheel chair users have limited upper extremity functions and strength, and moving arms and hands to navigate on a touch screen is a much more difficult task than using a joystick. An example of participant's rank ordering of the preferred location of notification by vibration is shown in Table 3.

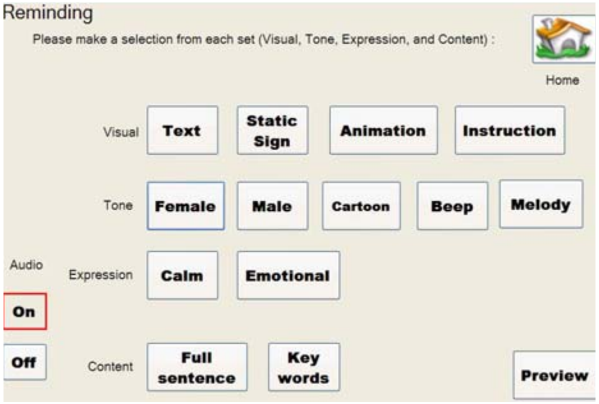


Figure 7. Menu for selecting features for virtual coach interaction design.

Ranking of Vibration Location on the Seat	Armrest	Headrest	Backrest around Shoulder Blade	Backrest around Mid of Upper Trunk
1	60.0	6.7	26.7	6.7
2	13.3	6.7	26.7	53.3
3	26.7	6.7	33.3	26.7
4	0	80.0	13.3	13.3

Table 3. Rank ordering of vibration output modality.

4.2 Field Study

Subsequent to Laboratory tests, the next step is to evaluate the system in the field. A three day Pilot Study was conducted to gather user feedback during actual system operation in the field. The participants were given a demonstration of the Virtual Coach and supplied with educational material. Questionnaires and interviews provided feedback. Subsequently the participants took the virtual coach home for three days with feedback again solicited through questionnaires and interviews [Liu et al 2011].

Systems that leave the laboratory to operate in the natural environment must be robust. Of particular concern was system reliability. The Virtual Coach was exercised over various surfaces to evaluate vibration tolerance including: pitch, cement, potholes, crack, grass, gravel, and mud. Location and mounting of the extra hardware, such as the touch screen, were evaluated as well as the repeatability of measurements of power seat functions (e.g. tilt angles). For example, the initial screen mounting increased the width of the wheel chair and caused difficulties in traversing doorways. The ball joint that adjusted the screen angle tended to loosen. Power is a critical issue since once the wheel chair battery is discharged the participant is unable to move. The range of the unmodified power wheelchair was 26.2 miles on a single charge. The addition of the Virtual Coach electronics reduced the range to 23.2 miles. This provides a comfortable margin since the average daily distance traveled by an active power wheel chair user is 10.7 miles, less than half the range with the Virtual Coach.

The functionality provided included pressure relief reminders (temporal and postural parameters) and providing further instructions once the user starts to engage seat functions. There were 12 power seat usage warnings. The warnings and reminders only appeared when the chair was occupied.

It was important to provide participants with support when the researchers were not present. A user's guide described how to use the virtual coach, precautions and limitations of the virtual coach, and how to diagnose problems and contact researchers. The Clinician's Guide described the default settings for the 17 variables, the relationship between the settings, and how to increase/decrease sensitivity of the warnings.

5. Ergobuddy Virtual Coach – Activity Classification via Statistical Machine Learning

Ergobuddy is a virtual coach system for package delivery workers to help prevent injury and reinforce trained ergonomic practices. User activities are inferred from application of statistical machine learning classification techniques to accelerometer data from a handheld device (which delivery drivers carry containing the routing information as well as recorder for recipient signatures) and supportive wearable devices.

The supportive wearable device was the eWatch, multi-sensor platform senses acceleration, light, sound, and temperature [Smailagic, Siewiorek 2005] shown in Figure 8. When combined with statistical machine learning algorithms, it is possible to identify user activity in real time as shown in Figure 9.

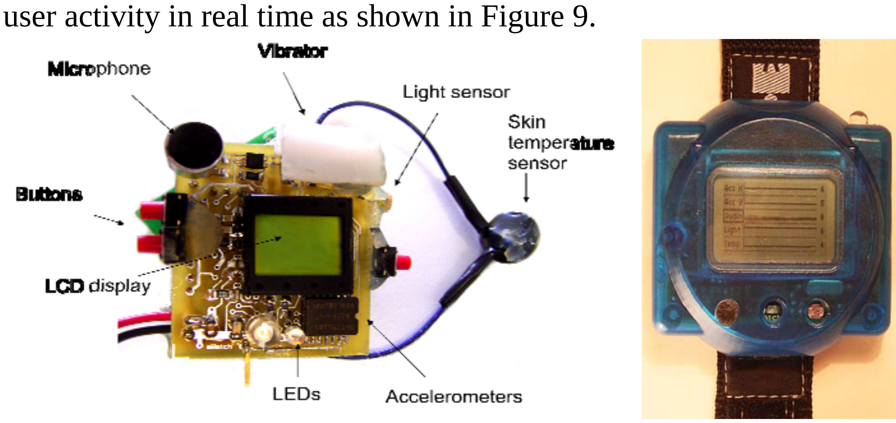


Figure 8. eWatch with MEMS sensors

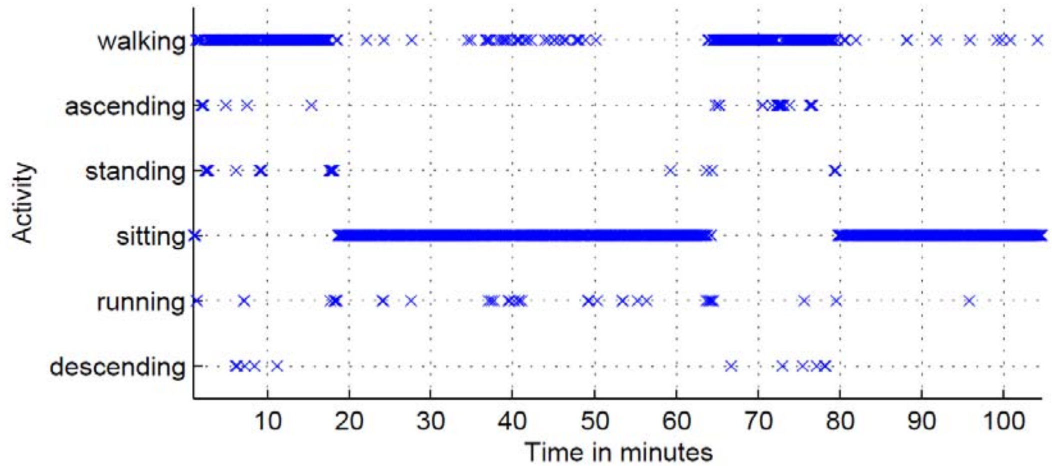


Figure 9. Real time activity recognition using classifiers based on statistical machine learning algorithms.

A typical approach in determining the best combination of sensor locations is to have the user wear multiple devices and use only the data from selected devices to evaluate performance for the different configurations.

For example, Figure 10 shows accuracy of activity classification for six locations on the users body: wrist, pants pocket, book bag, lanyard (neck), shirt pocket, and belt. For the majority of activities (running, sitting, standing, walking) any of the locations would give a classification accuracy over 90%. However if descending stairs was important, a wrist mounted sensor should be added. For ascending stairs, the book bag sensor is most accurate.

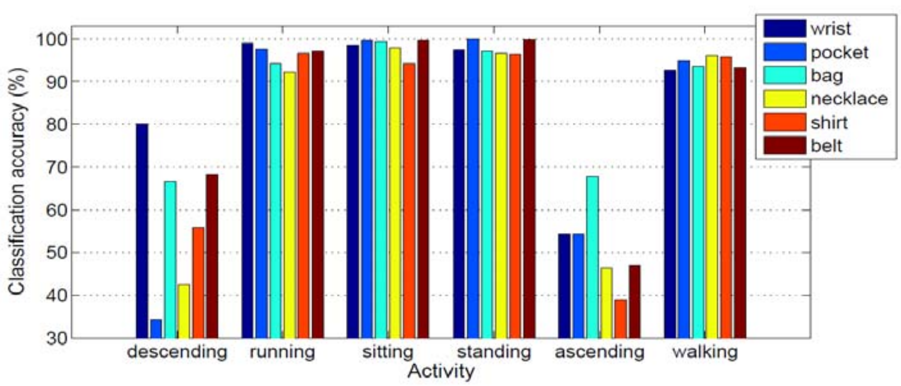


Figure 10. Activity Recognition Accuracy at Body Locations

Package delivery requires activities beyond those shown in Figure 10 (e.g. lifting, sweeping, carrying, climbing a ladder, pushing a cart). For Ergobuddy there is a single master handheld device and five eWatches worn at the Arm, Ankle, Back, Lanyard and Wrist positions (Figure 11). The Ankle position is shown in the lower part of Figure 11.

For a statistical machine learning classifier, there are four layers of processing: raw sensor data, feature extraction, classifier, and decision. These four layers can be intuitively partitioned into three different partitions, shown in Figure 12. As expected there are implications with each partition on how fusion of the data from individual sensor is performed, and wireless bandwidth utilization. A brief overview of the three partitions are described below:

Centralized Aggregation Architecture: This is a commonly used architecture in which raw data from all nodes in a network are transmitted to a master device for feature extraction and classification. In our experiment the amount of data transmitted is a continuous stream at 2 KB/sec. This scheme requires a static set of sensors.

Low Bandwidth Architecture: This architecture requires lower radio bandwidth since feature vectors are transmitted upon completion, yielding approximately 200/tw Bytes/sec. This yields the same accuracy performance of a centralized aggregation architecture. Again a static set of sensors is required.

Dynamic Architecture: The dynamic architecture further decreases bandwidth usage since only the context and confidence information is transmitted every window. This is approximately 80/tw Bytes/sec. Decision fusion differs from the previously described aggregators in that it allows the number of sensors to be dynamic. Accuracy is expected to decrease because only an abstraction of raw data is provided to the final decision maker (fuser), but the system becomes resilient to sensor failure and packet loss faults.

Multiple fusion techniques were used for the study and the best performance was offered by a scheme in which the average probability of all available sensors' confidence is used to fuse the local decisions into a final decision [Fisk et al 2011]. This fusion technique, since it does not require all sensors to be available every time it makes a decision is resilient to lost packets and node failures. In a perfect environment we achieve up to 90% leave one subject out cross validation classification accuracy on trained package delivery activities. The activity set included: sitting, standing, walking, running, lifting, carrying, sweeping/mopping, using stairs, using a ladder, and using a cart and there were 11 subjects total, which yielded approximately 20 hours of data. Multiple machine learning algorithms were tested, but best performance, given device constraints, were boosted decision trees on the eWatch (lower memory footprint) and random forests on the MC9500 master device. Additionally we find that performance scales well when data is missing, offering improved performance over a non fusion method at any packet lose rate and only 2% worse accuracy performance at 0% packet loss rate.

The visualization in Figure 12 graphs the F-Measure (harmonic mean of recall and precision) of each for each activity. This data is particularly relevant for a customer who may be specifically interested in detection of activities with ergonomic consequences. For example, in the case of lifting the customer would definitely want to include a lower back sensor, perhaps in a lifting belt, and a wrist or ankle sensor to complement the back sensor and capture other activities.

Figure 11. Multiple sensor placement

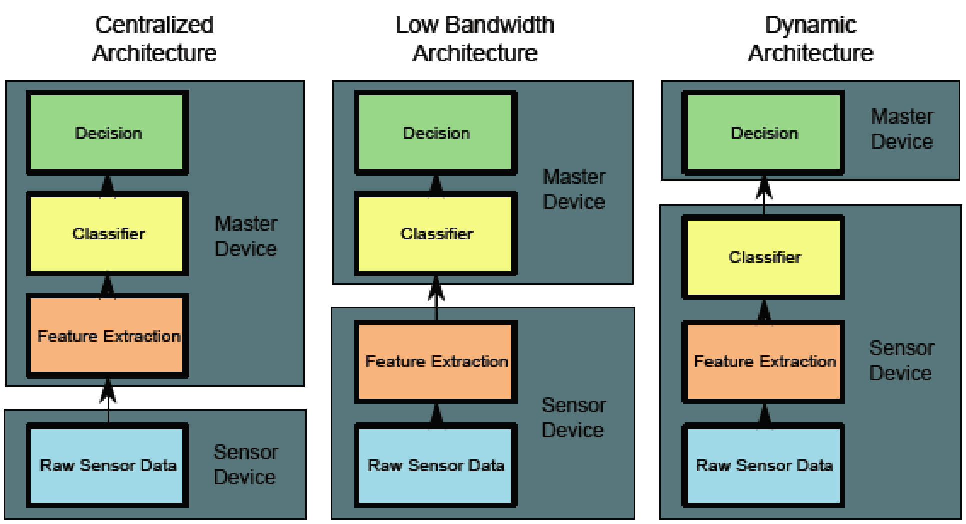


Figure 12. Classification Engine Partitions

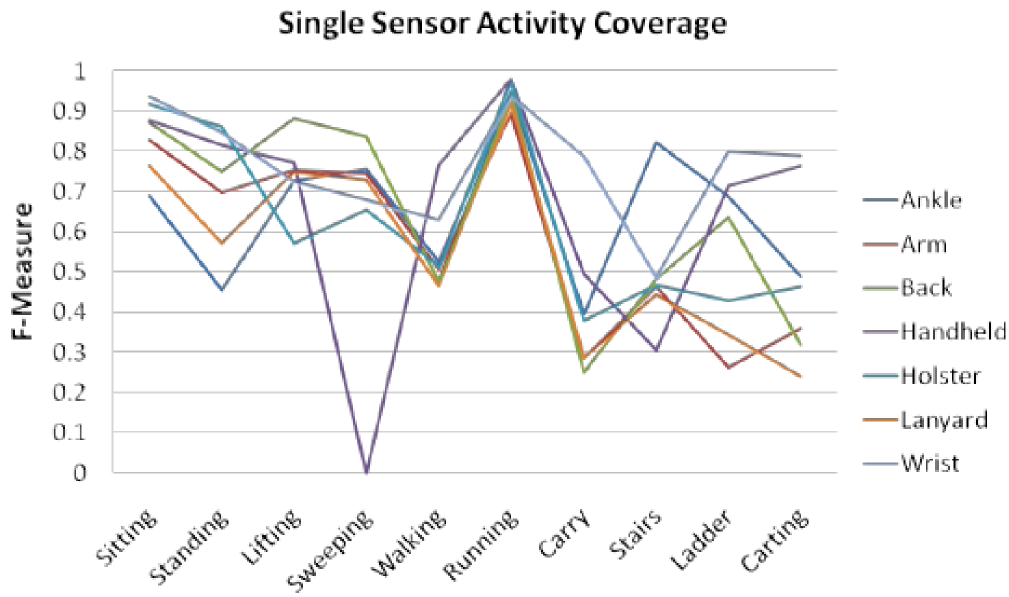


Figure 13. A visualization of the performance of sensor locations as they vary with activity.

In order to test the resilience of the classifier we simulated packet loss from 10% to 90%. For completeness the best sensor subsets of one to seven sensors were included in the analysis. To visualize the difference in reliability a low bandwidth scheme was also included. Figure 14 illustrates the results from this experiment. The primary takeaway from this chart is that in all non-ideal environments, 10% and up of packet loss, better classification performance was achieved with fusion. For example, there is a 35%

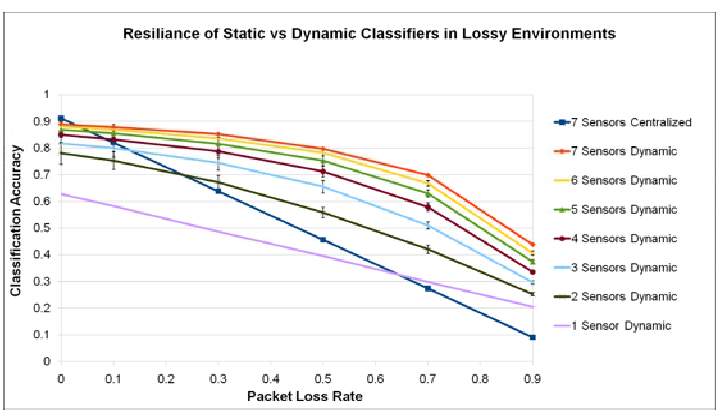


Figure 14. With the same number of sensors in a 0% loss environment fusion only provides 2% worse performance than a low bandwidth model.

accuracy increase in an environment with 50% packet loss. We have also discovered that the number of sensors can be reduced in low loss environments for power and bandwidth savings. Seven sensors is only marginally better accuracy than three sensors but is more appealing configuration in terms of overall encumbrance and power.

6. Example Virtual Coaches

This section will provide brief examples of three more coaches. The Manual Wheel Chair Propulsion coach monitors for correct arm movement while propelling the chair. IMPACT seeks to motivate users to exercise. Finally, MemExerciser has a goal of improving user memory. Two basic technologies are used to identify the system's context.

6.1 Manual Wheel Chair Propulsion Coach

The Manual Wheel Chair Propulsion (MWCP) coach explored providing advice to manual wheel chair users to help them avoid damaging forms of locomotion. The primary form of context for this system is the user's propulsion pattern. The contexts of self versus external propulsion and the surface over which propulsion is occurring are used to improve the accuracy of the system's propulsion pattern classifications.

The MWCP uses statistical machine learning algorithms to classify propulsion patterns and surface material. The top three acceleration characteristics for six common activities after a Linear Discriminant Analysis (LDA) transformation (Figure 15) illustrates spatial clustering that can be exploited [Maurer et al 2006] to continuously infer physical activity.

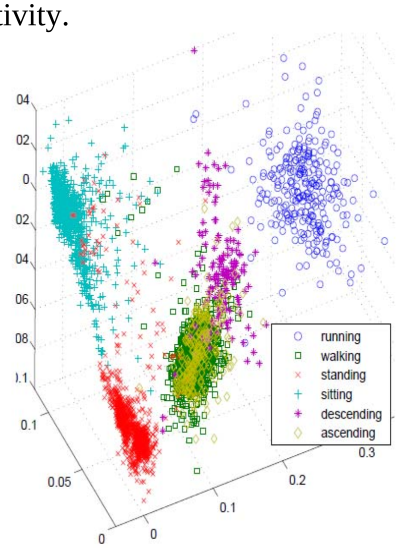


Figure 15 Feature Space after LDA Transformation

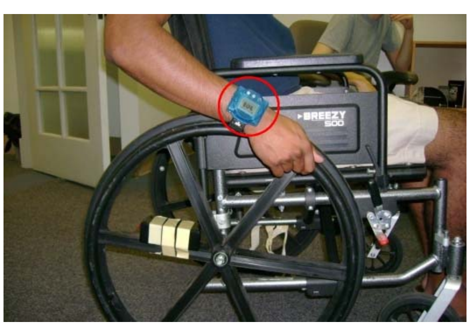


Figure 16. eWatch used for wheel chair propulsion data collection. The eWatch was worn on the wrist while the wheel chair was self – propelled

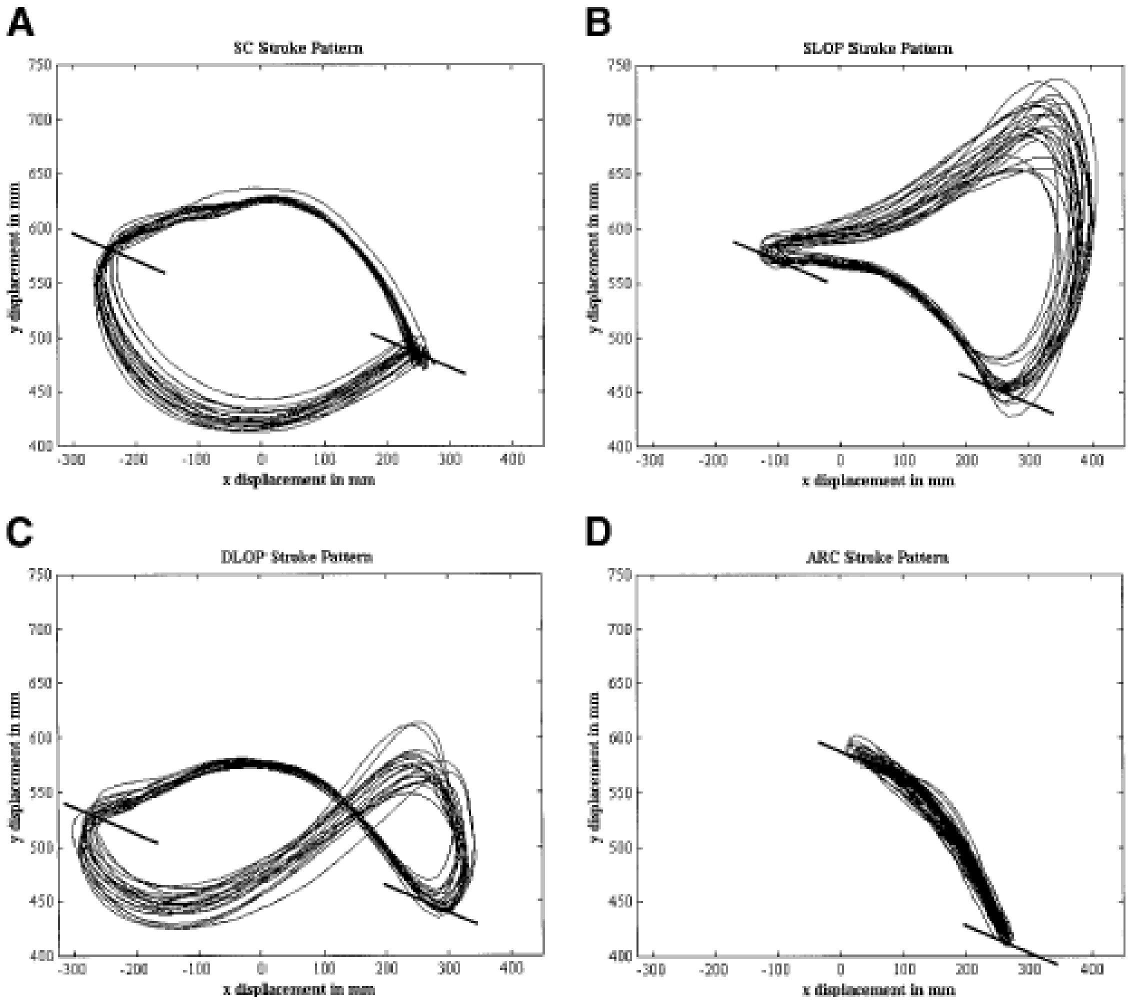


Figure 17. Four classic propulsion patterns are shown: (A) semicircular (SC); (B) SLOP; (C) DLOP; and (D) arcing. The dark bars to the right of each pattern represents the beginning of the propulsion stroke. The dark bars to the left of each pattern represent the end of the propulsion stroke and the beginning of recovery.

Both wearable (Figure 16) and wheel chair-mounted accelerometers were used to provide contextual information [French et al 2007].

There are four distinct propulsion patterns that wheelchair users tend to follow (Figure 17) – semicircular (SEMI), single loop over (SLOP), double loop over (DLOP) and arcing (ARC) which have been identified in a limited user study. Of these, the recommended propulsion pattern is semi-circular, because the strokes have lower cadence and higher stroke angle. Data was collected using all four propulsion patterns on a variety of surface types. Machine learning algorithms produced accuracies of over 90%. It was also noted that the higher the resistance of the surface traversed, the higher the propulsion prediction accuracy.

Two common machine learning algorithms, k-Nearest Neighbor (kNN) and Support Vector Machines (SVM) with a Radial Basis Function (RBF) kernel were used to classify propulsion patterns. We also experimented with simplifying the classification task into an Arcing vs. Non-arcing pattern classification in an attempt to improve classifier accuracy. The justification for this being that arcing patterns are the most damaging to the users. Using this binary classification scheme, we found the average classification accuracy increased to the 85-95% range.

We were able to differentiate between the resistance level of the surface over which propulsion was occurring with 70-80% accuracy (Figure 18). It can be seen that the classification accuracy tends to be higher, with less variability across patterns, on surfaces with higher resistance (dynamometer, low carpet), when compared to surfaces with low resistance (tile, asphalt). Classification accuracy for arcing was considerably lower than the other propulsion patterns. Namely, the arcing is a subset of each of the other patterns, and hence, is most susceptible to misclassification.

We found that there is differential classification accuracy across subjects, which seems to be dependent upon the arm length of the subject. Intuitively, this makes sense since the longer the arm, the faster the acceleration of the wrist if the arms are maintaining similar velocities. This also means that in order to develop cross-subject classifiers we may need to normalize the acceleration profiles with respect to participant arm length.

We were also able to use the acceleration profile of the wheel chair from the frame-mounted accelerometer to differentiate between self-propulsion and being pushed with ~80-90% accuracy. This type of information will be useful in system management, for example, we don't want to be providing feedback to the user on their propulsion pattern when they are not propelling themselves.

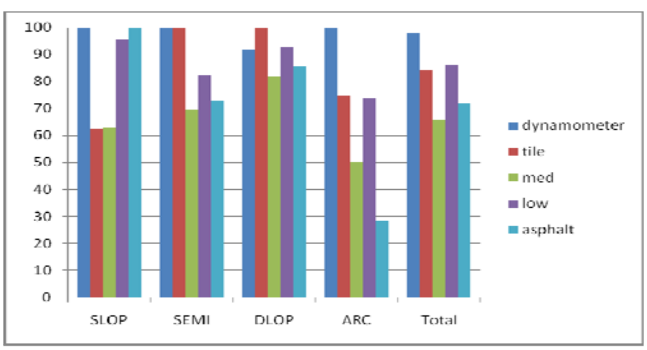


Figure 18. Classification results using KNN algorithm for various surfaces

6.2 IMPACT: Improving and Motivating Physical Activity Using Context

Many physical activity awareness systems are available in today's market. These systems show physical activity information (*e.g.*, step counts, energy expenditure, heart rate) which is sufficient for many self-knowledge needs, but information about the factors that affect physical activity may be needed for deeper self-reflection and increased self-knowledge. IMPACT explored the use of contextual information, such as events, places, and people, to support reflection on the factors that affect physical activity.

IMPACT uses a mobile phone and GPS to monitor step counts and the user's location (Figure 19). The mobile phone also has an easy-to-use interface to input what the user is doing and whom he/she is with. The pedometer application stores the user's step counts per minute and displays the user's aggregate step counts for the day and for each of the past five minutes. The GPS module scans the user's location every minute, which is then stored by the phone application. The phone application collects additional contextual information using activity-triggered experience sampling. When the user is active or inactive, the phone vibrates to prompt the user to select from a list: what they were doing (events) and whom they were with (people). The list is pre-filled with five common activities (*e.g.*, grocery shopping, walking) and five usual companions (*e.g.*, friends, family, co-workers), but users can enter new labels. We did not implement automatic labeling of events and people because such classification requires additional sensors that may not be robust enough for a long-term field study or are still not mainstream and widely available.

Figure 19. Monitoring device for the second version of IMPACT. Nokia 5500 Sport (left) and detailed view of the display (right).

There are many personal health applications developed for smart phones, including some recent ones on turning walking into a game by Mobile Adventures (www.mobileadventurewalks.com) , calorie counter and diet tracker by MyFitnessPall (www.myfitnesspal.com), and Pedometer PRO GPS+ by Arawella (mobile.viaden.com/pedometer). The Nike+iPod system (www.nikeplus.com) monitors running time and uploads the data online for visualization. Most of these systems do not go beyond performance numbers. The IMPACT research builds on these systems by integrating contextual information.

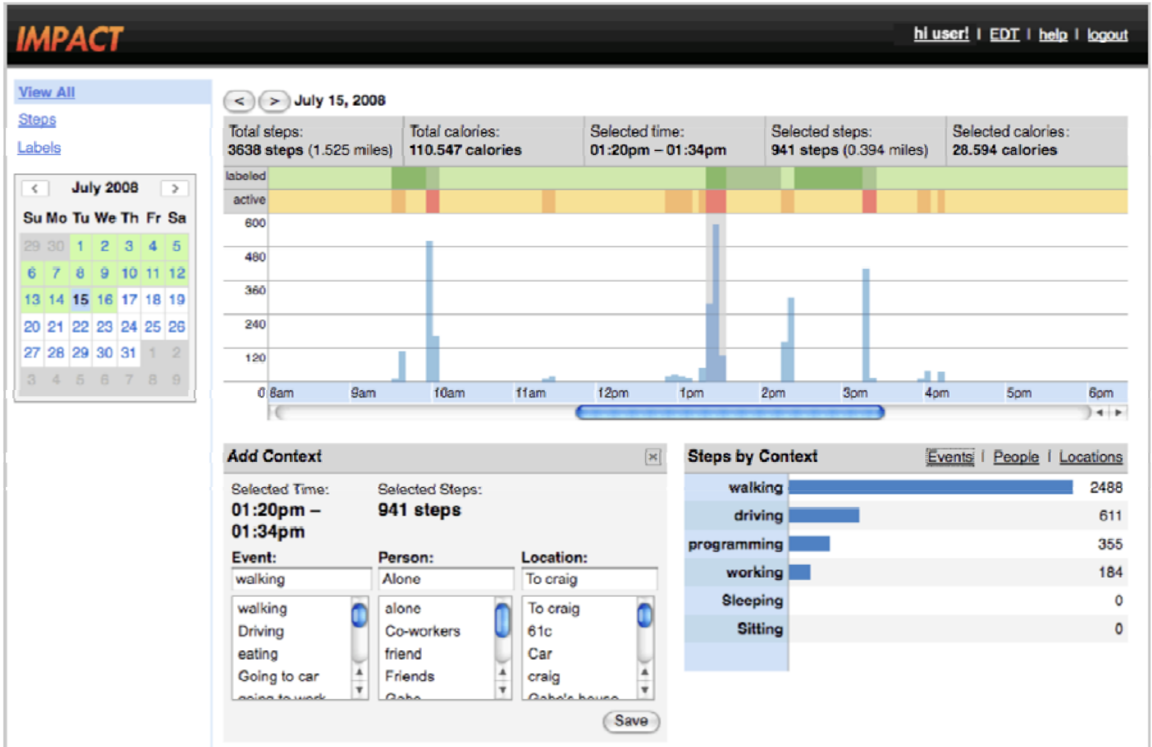


Figure 20. Visualizations in the IMPACT website showing step counts with contextual information. Detailed step counts graph with contextual annotations (top) and context graph (bottom right).

The IMPACT system also includes a web interface (Figure 20) that shows the association between daily activities and step counts on 1) a timeline of the user's steps with time segments labeled with contextual information; and 2) a histogram of the total number of steps associated with a particular label (*e.g.*, 400 steps at work, 1300 steps at the grocery store). Instead of manually entering step counts and contextual information on the web site, a desktop application synchronized data between the phone and the new web site. If the user needs to add more contextual information after uploading, they can label periods of time on the visualizations. We also implemented two other versions of the system: *Steps-Only* and *Control*. The *Steps-Only* system only monitored step counts and the web site only showed daily step counts without any contextual information. The mobile phone still alerted users when they have been active and inactive, but they were just asked to rate how active they were on a 5-point Likert scale (not at all active to very active), to make the interruption comparable to the *IMPACT* version. The *Control* system also only monitored step counts, but we removed visualizations on the web site. Essentially, it is similar to an off-the-shelf pedometer.

We conducted an 8-week long study with 49 participants with an age range of 18 to 60: four weeks for a *Baseline* phase and four weeks for an *Intervention* phase. During the *Baseline* phase, all participants used the *Control* system (step counts). During the *Intervention* phase, participants were randomly assigned to three types of interventions: *Control*, *Steps-Only*, and *IMPACT* (step counts and the context in which those steps were taken: location, type of activity, and whom the person was with). The evaluation revealed three major findings. First, when given access to contextual information and physical activity information, users can and do make associations between the information helping them become aware of factors that affect their physical activity. Second, reflecting on physical activity and contextual information can increase people's awareness of opportunities for physical activity. Lastly, automated tracking of physical activity and contextual information increases the amount of data collected by the user, which benefits long-term reflection, but may be detrimental to immediate awareness.

We believe these results are applicable to the use of contextual information to reveal factors that affect other types of behaviors, for example, diabetes management and energy conservation. These contributions suggest that personal informatics systems should further explore incorporating contextual information.

6.3 MemExerciser

People with episodic memory impairment (EMI), such as those with early-stage Alzheimer's disease, struggle with maintaining their sense of self [Conway 1990]. While they can still remember experiences from the distant past, recent experiences are difficult to recall. As a result, their window of remembered experiences shrinks as their memory abilities decline, leading to feelings of frustration, anger or depression [Steeman, et al 2006]. Over 26 million people worldwide suffer from Alzheimer's disease [American Health], but the effects are not limited to these individuals. Rather, the disease also affects the well-being of family caregivers as they have to provide the cognitive support necessary for aging in place. Caregivers usually help the person with EMI remember the details of an experience by providing cues, small details of the experience from which the person with EMI can use to recollect other details and mentally relive the experience. However, caregivers often must repeatedly provide cues for the same experience again and again which can lead to feeling overburdened, burnt out, or even depressed [Almbert et al 1997].

Lifelogging systems automatically record a log of a user's personal experience in the form of pictures, sounds, actions, activities, or raw sensor data using wearable or embedded sensors such as cameras, audio recorders, location tracking, and bodily sensors. The data collected by lifelogging systems can provide memory cues to help people remember the original experience [Sellen et al 2007]. However, the sheer amount of data collected can also be overwhelming. Thus a functional cognitive support system must embed an intelligence system to select and display appropriate data.

MemExerciser, a lifelogging system , is specifically designed for people with episodic memory impairment and their caregivers. The system records and supports reminiscence for significant personal experiences that the user wants to remember in detail. The goals of the system are to maximize the independence of the person with EMI and at the same time minimize the burden on their caregiver. The system provides an appropriate amount of cueing assistance for the person with EMI to reminisce about the experience without needing to bother the caregiver repeatedly to provide additional cues.

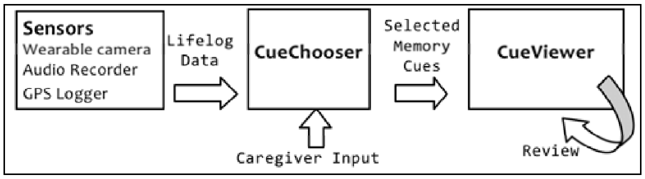


Figure 21. MemExericser System Design: Capture, Selection, and Review

MemExerciser consists of three subsystems (Figure 21): passive experience capture, hybrid cue selection (CueChooser), and cue review (CueViewer).

The system captures both the visual and audio content of the experience as well as contextual information such as location, movement, and light levels. People with memory impairment often forget to explicitly trigger a device (*e.g.*, camera) to record. The system uses a passive capture approach that requires the user only to turn it on and allow the system to manage when to trigger recording. The capture system consists of three devices (Figure 22): the Microsoft SenseCam [Hodges et al 2006], an off-the-shelf digital voice recorder, and an off-the-shelf Wintec GPS location tracker. The SenseCam is a wearable digital camera that automatically takes pictures when triggered by the onboard light sensor, infrared sensor, accelerometer, or simple timer. With an initial reminder from the caregiver, the person with EMI can switch on these three capture devices before each experience, wear the camera around the neck, place the audio recorder in their top shirt pocket, carry the GPS logger in their pocket and can simply enjoy their experience without needing to stop and tell the system to record.

With a passive capture approach mentioned above, the system can capture a large amount of data. To identify the most salient memory cues to present to the person with EMI, the lifelogging system employs a hybrid approach that involves both automated computer analysis of the lifelog as well as the expertise of the caregiver. CueChooser is a software application that assists the caregiver in selecting the most salient memory cues using automated content and context analysis.

Figure 22. Capture devices: Microsoft SenseCam, digital voice recorder, GPS logger.

Prior work [Lee and Dey, 2007] has identified that the most salient memory cues are determined by the type of experience. There are people-based, place-based, action-based, and object-based experiences. The caregiver can specify the type of the experience and CueChooser (Figure 23) will apply the appropriate content and context analyses to suggest potentially good cues. For people-based experiences, CueChooser identifies photos with faces using computer vision. For place-based experiences, it uses a combination of GPS data and the SenseCam's accelerometer data to determine when the user enters, is near, or is staying in a particular place. Similarly for object-based experiences, CueChooser can use GPS or accelerometer data to find when the user is standing still and looking at an object of interest. For action-based experiences, image summarization techniques [Doherty, et al 2007] are used to find cues from different scenes. However, good memory cues have other characteristics that computers have difficulty identifying such as distinctiveness and personal significance [Lee and Dey, 2007]. The CueChooser interface allows the caregiver to browse through the automatically suggested photos to select content to include in a slideshow narrative. Caregivers can add their own annotation using their voice or drawing on each photo in the slideshow narrative.

Figure 23. MemExerciser's CueChooser user interface. The caregiver can view system-suggested cues in constructing a narrative, and provide visual and audio annotations to selected cues.

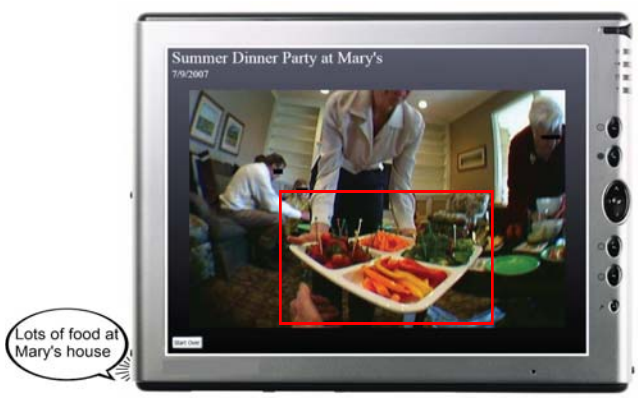


Figure 24. MemExerciser's CueViewer user interface: tapping on the screen displays pictures, and plays back lifelog audio and caregiver's voice annotation.

The lifelogging system presents the lifelog data in a way that maximizes the opportunities for the person with EMI to think deeply about each cue to trigger his own recollection of the original experience. Caregivers normally reveal cues one at a time to allow the person with EMI to remember the rest of the experience on their own [Lee and Dey, 2007]. MemExerciser (Figure 24) includes a software application designed to run on a Tablet PC. Based on the selection of photos, sounds, and annotations from the caregiver with the CueChooser application, MemExerciser allows the person with EMI to step through all the cues at their own pace. The cue review process is designed to be challenging enough to stimulate their memory processes (acting as a form of mental exercise) but also be supportive enough so that people with EMI can feel as if they are mentally reliving the experience. Instead of passively playing back each photo and sound like a movie, MemExerciser shows only one picture at a time and gives the user control over how long they want to examine each picture. Recorded audio and the caregiver's annotation are progressively revealed to facilitate the user's self-recollection. With MemExerciser, the person with EMI can feel as if his caregiver is walking him through the cues but with the benefit of going at their own pace and not repeatedly bothering the caregiver.

A pilot field evaluation was conducted of the lifelogging system with three people with EMI (all associated with the early stages of Alzheimer's disease) and their spousal caregivers. The self-guided review approach of the lifelogging system was compared with a caregiver-guided approach [Hodges, et al 2006] where the caregiver repeatedly guides the person with EMI through only the photos taken with the SenseCam. Participants review the cues every other day during the two weeks after their experience. It was found that the self-guided approach resulted in a statistically significantly greater number of details freely recalled four weeks after the experience (Figure 25) as well as greater confidence in memory when assessed using the Meta-memory in Adulthood Questionnaire (Figure 26). Caregivers expressed that the self-guided approach freed them from repeatedly going through the same cues again and again.

In summary, MemExerciser is a lifelogging system to assist people with episodic memory impairment to reminisce about recent experiences. The system uses a passive capture approach so that the person with EMI does not have to remember to initiate capture. The system uses both automated computer analysis and the expertise of the caregiver to select out the most salient cues from the lifelog. Finally, the system structures the cue review interaction so that it allows the person with EMI to think more deeply about each cue and remember the details of their experiences without repeatedly burdening the caregiver.

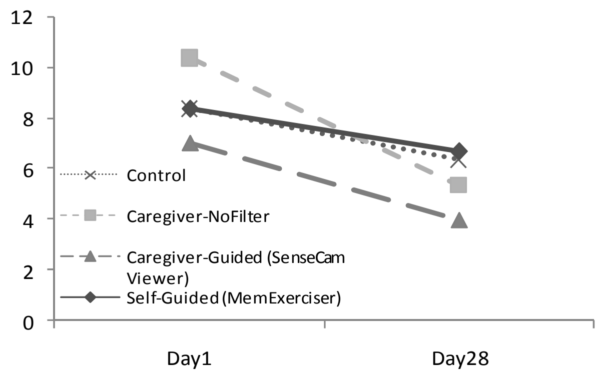


Figure 25. Mean Number of Details Recalled

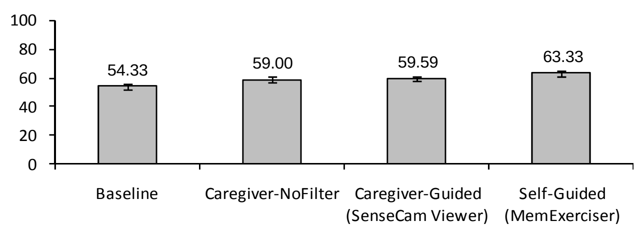


Figure 26. Participant's self-assessed memory confidence using the Metamemory in Adulthood Questionnaire

7. Summary

We have introduced five elements (prescription – input, sensor processing, user interaction, coaching model, user engagement) that appear in the architecture of virtual coaches. Further, for coaches using multiple sensors and the statistical machine learning coaching model three partitions (centralized, low bandwidth, and dynamic) are evaluated for activities typical of the package delivery domain. Five virtual coaches are described and evaluated with a variety of user engagements.

Virtual coaches is a highly interdisciplinary area of research and development that combines the expertise and science from the social sciences, informatics, computing sciences, health sciences and engineering. The goal is to employ theory from social science, cognitive science, and economics. Social science models of collaborative behavior can be used as a basis for determining the nature of the social setting. Theories and observations of which clues humans use to interrupt a social situation and gain attention can form the basis for sensor data processing and software decisions in virtual coaches. By mapping observable parameters into cognitive states, the computing system can estimate the form of interaction that minimizes user distraction and the risk of cognitive overload.

8. Lessons Learned

People fundamentally prefer to do things without assistance from other people, even if it takes much more time.

It is possible and valuable to auto-tune user interactions by incorporating contextual information, gathered about the user herself, her environment and her situation.

Clinicians and informal caregivers are in fact willing to configure and customize QoLT systems. QoLT developers need to make it convenient for them to do so.

People will trade some sense of privacy for enhanced ability to gain information about their environment and situation.

The user's input has a very high priority in the system design, in particular for the graphical user interface (GUI). Asking users what they want, then about their feedback, as well as their evaluation of the system, is the right strategy to pursue toward a successful outcome.

References

[Almbert et al 1997] Almbert, B., Grafstrom, M. and Winblad, B. (1997). Caring for a demented elderly person – burden and burnout among caregiving relatives. Journal of Advanced Nursing, 25 (1), 109-116.

[American Health] American_Health_Assistance_Foundation. http://www.ahaf.org/alzdis/about/adabout.htm.

[Conway 1990] Conway, M. (1990). Autobiographical memory. Open University Press, Milton Keynes, 1990.

[Doherty et al 2007] Doherty, A.R., Smeaton, A.F., Lee, K. and Ellis, D. (2007). Multimodal Segmentation of Lifelog Data. *RIAO*, Pittsburgh, PA, 2007.

[French et al 2007] French, B., Siewiorek, D.P., Smailagic, A., Deisher, M. Selective Sampling Strategies to Conserve Power in Context Aware Devices, iswc, pp.1-4, 2007 11th IEEE International Symposium on Wearable Computers, 2007.

[French et al 2010] French, B., Siewiorek, D.P., Smailagic, A., Kamarck, T. Lessons Learned Designing Multi-Modal Ecological Momentary Assessment Tools, Journal of

Technology and Disability, Special Issue on Quality of Life Technology, IOS Press, Vol. 22, No. 1-2, 2010, pp. 41-51.

[Fisk et al 2011] Fisk, S., Siewiorek, D.P., Smailagic, A. Increasing Multi-Sensor Classifier Accuracy Through Personalization and Sensor Fusion, Proceedings of the International Symposium on Quality of Life Technology, Toronto, Canada, June 2011.

[Hodges et al 2006] Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N. and Wood, K. (2006). SenseCam: a Retrospective Memory Aid. *Proc. UBICOMP*. 81 - 90.

[Lee 2007] Lee, M.L. and Dey, A.K. (2007). Providing Good Memory Cues for People with Episodic Memory Impairment. *Proc. ASSETS 2007*. 131 - 138

[Liu 2010] Liu, H., Cooper, RM., Cooper, RA., Smailagic, A., Siewiorek, D., Ding, D., Chuang, F. (2010). Seating virtual coach: A smart reminder for power seat function usage, Journal of Technology and Disability, Volume 22, Number 1-2, 53-60

[Liu 2011] Liu, H-Y., Grindle, G., Chuang, F-C., Kelleher, A., Cooper, R., Siewiorek, D., Smailagic, A, Cooper, R. User preferences for indicator and feedback modalities: A preliminary survey study for developing a coaching system to facilitate wheelchair power seat function usage, IEEE Pervasive Computing, Vol. 10 (in press), 2011.

[Maurer et al 2006] Maurer, U., Smailagic, A., Siewiorek, D.P., Deisher, M. Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions. BSN 2006: 113-116

[Plarre et al 2011] Plarre, K., Raij, A., Hossain, S., Ali., A., Nakajima ,M., Al'Absiz, M., Ertin, E., Kamarck, T., Kumar, S., Scott, M., Siewiorek, D.P., Smailagic, A., Wittmers, L. Continuous Inference of Psychological Stress from Sensory Measurements Collected in the Natural Environment, ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN 2011), Chicago, IL, 2011.

[Sellen et. al 2007] Sellen, A.J., Fogg, A., Aitken, M., Hodges, S., Rother, C. and Wood, K. (2007). Do life-logging technologies support memory for the past?: an experimental study using sensecam. *Proc. CHI 2007*.81-90.

[Siewiorek 1994] Siewiorek, D.P., Smailagic, A., and Lee, J.C. (1994). An interdisciplinary concurrent design methodology as applied to the Navigator wearable computer system. Journal of Computer and Software Engineering, Ablex Publishing Corporation, 2(3), 259-292.

[Siewiorek 2010] Siewiorek, D.P., Smailagic, A., Courtney, K., Matthews, J., Bennett, K., Cawley, R., Liao, X., Vartak, M,, White, N., Yates, J. Multi-User Health Kiosk, Proc. International Symposium on Quality of Life Technology, Las Vegas, NV, June 2010.

[Siewiorek, Smailagic 2008] Virtual Coach for Power Wheelchair Users, Institute for Complex Engineered Systems Technical Report, Carnegie Mellon University.

[Smailagic 1995] Smailagic, A., Siewiorek, D. P. et. al. (1995). Benchmarking an interdisciplinary concurrent design methodology for electronic / mechanical design. Proc. ACM / IEEE Design Automation Conference, 514-519.

[Smailagic, Siewiorek 2005] Smailagic, A., Siewiorek, D. P., Maurer, U., Rowe, A., Tank, K., eWatch: Context Sensitive System Design Case Study, Proc. IEEE Symposium on VLSI, 98-103

[Steeman et al 2006] Steeman, E., de, C., Dierckx, B., Godderis, J. and Grypdonck, M. (2006). Living with early-stage dementia: a review of qualitative studies. Journal of Advanced Nursing, 54. 722-738.