

Leveraging Multi-Modal Sensing for Mobile Health: a Case Review in Chronic Pain

Min S. H. Aung, Faisal Alquaddoomi, Andy Hsieh, Mashfiqui Rabbi, Longqi Yang, J.P. Pollak, Tanzeem Choudhury, and Deborah Estrin

Abstract—Active and passive mobile sensing has garnered much attention in recent years. In this paper we focus on chronic pain measurement and management as a case application to exemplify the state of the art. We present a consolidated discussion on the leveraging of various sensing modalities along with modular server-side and on-device architectures required for this task. Modalities included are: activity monitoring from accelerometry and location sensing, audio analysis of speech, image processing for facial expressions as well as modern methods for effective patient self-reporting. We review examples that deliver actionable information to clinicians and patients while addressing privacy, usability and computational constraints. We also discuss open challenges in the higher level inferencing of patient state and effective feedback with potential directions to address them. The methods and challenges presented here are also generalizable and relevant to a broad range of other applications in mobile sensing.

Index Terms— Survey, mobile sensing, mobile health, smartphones, wearable technology, audio sensing, activity monitoring, face expression, self-reporting, modular architecture, behavioral signal processing, affective computing, chronic pain.

I. INTRODUCTION

The unprecedented spread of smartphone use across the world has led the development of new mobile systems for health measurement, analysis and intervention or *mHealth* [1], [2]. In this context, the smartphone can be viewed as a multi-sensor device which serves as a continuous monitor capable of informing both clinically-relevant inferences with efficacious and timely patient-feedback. Given the enormity of the number of potential applications within *mHealth*, in this paper we focus on the area of pain management. This choice of pain as a case study will help to demonstrate open but surmountable challenges when aiming to implement effective *mHealth* systems; we expect that many of the implications for system design are transferrable to other patient-centric health applications.

Pain is a complex experiential phenomenon; it has a range of root causes and can impact and manifest in many domains such as emotion, cognition, socialization, function as well as behaviour [3]. Intensity is subjective and as such it is difficult

to tangibly detect in any direct sense. The broadly accepted definition by the International Association for the Study of Pain is principally a generalized catchall and does little to clarify it as an objective signal: “*an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage*” [4]. Moreover, recent commentaries state that clinical taxonomies for pain are overall not well ordered and some even illogical [5]. Having said this, two known components of pain: negative psychological states (e.g. anxiety and fear) and resultant *pain behavior* (e.g. limping, grimacing, avoidance) [6],[7] have been widely researched in observation-based behavior studies. The form and characteristics of different *pain behavior* have more explicit definitions often containing specific physical terminology [7],[8],[9]. A particular type of pain behavior that is regularly assessed in clinical practice is *pain interference*, in which activities are disrupted or avoided due to pain. From a signal processing perspective it is these outward manifestations of *pain behavior*, including interference, which are more tangible and convey the expressions of pain experience and serve as meaningful target for *mHealth* systems.

The emergent fields of Behavioral Signal Processing (BSP) and Affective Computing (AC) [10],[11] have demonstrated that complex human behavior and psychological states can be inferred from multi-modal data, principally from audiovisual [12] and physiological sensing [13]. Such studies have made valuable contributions in the understanding of signal processing and recognition methodologies for inference of human behavior (for an overview of multimodal pain behavior recognition see [14]). However, there remains systemic challenges when using mobile systems, challenges that are not apparent in typical laboratory settings with specialized apparatus and control. In terms of modalities, highly informative internal bio-signals such as galvanic skin response, electroencephalography, electromyography and heart rate are not directly measured from broadly-adopted smartphones; and are still relatively poorly measured on wearables. Also for audiovisual data, artefacts relating to noise, motion, occlusion or crosstalk cannot easily be controlled and minimized in unconstrained real word situations. Furthermore, the reliability and completeness of data

Copyright (c) 2014 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org

Min S. H. Aung, M. Rabbi and T. Choudhury are with the Department of Information Science, Cornell University, NY 14853, USA, (e-mail: msa242@cornell.edu).

F. Alquaddoomi is with the Department of Computer Science, UCLA, CA 90095 USA.

A. Hsieh, L. Yang, J.P. Pollak and D. Estrin are with the Jacobs Technion Cornell Institute, NY 10011, USA.

streams may also be less than ideal when using a mobile device, depending on the condition and habits of the user. On the other hand, these shortcomings only have critical implications for *spatiotemporally small scale behaviors* (e.g. a facial grimace or changes in vocal pitch or heart rate due to pain conditions); typically the performance of such systems have a high dependency on granularity and data quality. In contrast, for *spatiotemporally large scale behaviors* (e.g. the amount and pace of walking, or hours spent out of the house, on a daily or weekly basis) the aforementioned shortcomings are of far less concern due to data being drawn over greater space and timescales.

As mentioned, we focus on pain as a case study. The justification for this is in part that *spatiotemporally large scale behaviors* are especially relevant to chronic pain. For example, the tracking of adherence to daily exercise regimes over the span of weeks or months is valuable information for chronic back pain rehabilitation [15]. Typically such management regimes are on-going and self-led with many of the required tasks done outside of clinical settings and away from health experts. For such situations, *mHealth* systems and/or smartphones are a natural solution as they are habitually carried in close proximity and are in constant use. This allows for measurements, inferences and feedback to be processed locally, continuously, and over long periods of time.

However, chronic pain management is multifaceted; it's success is dependent on a multitude of physical, cognitive as well as socio-economic factors [16][17]. Therefore, if one aims to develop a *mHealth* system for chronic pain management, there is a particular need to accommodate a range of patient generated data. In measurement terms, this means a combination of both *active* sensing which requires proactive action from the user (e.g. self-reporting, diary logging, interactive assessments) and *passive* sensing (e.g. geo-locating, activity recognition, audio). Following this, the correct and efficient parsing of informative descriptors from low level measurements is required. These descriptors serve as input to models designed to infer high-level behavioral patterns. Finally, these inferences can be used for the correct and timely issuance of feedback and/or interventions that are clear, persuasive and actionable by the patient or clinician; the effects of which can then be further measured and the process repeated. We outline this with a tripartite functional loop: *measurement - inference - management* as a basis for system design.

Although we address pain as an over-arching application domain, there still remains a wide range of sub-domains even within chronic pain conditions. Each sub-domain has specific clinical and technical requirements and therefore a comprehensive discussion of all potential systems for every pain condition would be cumbersome, repetitive and out of the scope of this discussion. Therefore to efficiently illustrate requirements and challenges from a systems perspective, we highlight within the following (Section II) example cases of different sensing modalities, with inference to a relevant behavioral outcome and how this inference can be used for management; each following the *measurement - inference - management* process. Section II ends with examples of modular

software platforms that are critical to take multiple sensing streams from measurement to meaning in a reusable and evolvable manner. Section III discusses open challenges for higher level inferencing, including temporal considerations, and an overview of potentially useful advances in pattern recognition. Also discussed are factors in improving the usability and persuasiveness from a user's perspective. Finally, we conclude in Section IV.

II. SIGNALS AND SIGNAL SPECIFIC PROCESSING

We focus on physical activity, acoustics, image capture, and self-reporting as four foundational data streams. While it is known that other signals are also indicative of pain and pain related emotions such as physiological data [18] and that other mobile devices are gaining in both penetration and range of sensing modalities (e.g. the Empatica E4 Wristband which measures blood volume pulse, heart rate variability, galvanic skin response and peripheral skin temperature); the methodologies and approaches we review should provide a template for their handling and the integration of future data streams from mobile apps and wearable devices.

A. Physical Activity

The continued monitoring of physical activities away from clinical settings is essential for understanding progress in pain reduction, management, and healing, whether it be in the context of post-treatment follow up, or long term physiotherapy [19],[20],[21]. Such monitoring was traditionally achieved through specialized wearable devices [22],[23], or special-purpose instruments deployed in clinical environments [24]. However, activity tracking on broadly deployed smartphones has become a popular and affordable alternative. The broad adoption and inherent utility of smartphones for end-users means that for many day to day measurements, no additional devices need to be purchased, carried, or forgotten by the user [25],[26],[27]. Consumer wearables in the form of fitness trackers and smart watches have similar advantages and typically transmit data via the smartphone. The activity sensors embedded in wristbands and phones are largely identical, and so much of this section's discussion is applicable to both.

Smartphone-based activity tracking typically uses the continuous data collected by the phone's accelerometer to determine mobility states. The standard set of states are: sedentary (or stationary), walking, running, or in vehicle. In addition, the location information captured by the phone's GPS or derived from WiFi signals is often used to infer the locational context in which the activities were taking place. There is also the potential for further contextualization to the standard states from low level sensor data that has implications for pain management. The degree of incline of a particular route taken can be determined using either topographical information from maps or from accelerometer and barometer values. This informs the understanding of difficulty level and energy expenditure of a route; the type of surface on a high incline path whether it is stepped or a smooth slope will also have a bearing on difficulty and could potentially be inferred from the inertial and barometric readings. This further contextual information is important for musculoskeletal conditions where posture and speed are relevant factors. Also for speed and step count

inference GSM traces can also be utilized [28]. However, such detections are currently not done with standard commercial systems.

Case example: an activity based app for rheumatoid arthritis (RA) - RA is characterized by episodes of flaring of joint inflammation. Identifying flares early and minimizing their intensity and duration is important for long term management because each episode of flare contributes to further joint damage. In day to day management patients struggle to understand what environmental and behavioral factors contribute to the triggering and duration of flares. While directly and continuously measuring pain is not currently possible, proxies for pain can be found in their physical function and daily activities [29]. In the following, we describe a four-stage inference algorithm from which raw sensor data generates high-level behavioral information that is useful in RA management. This algorithm has been deployed in an open-sourced activity tracking app, called *Mobility* [30]. The clinical validity of the derived measures is currently being evaluated within RA, as well as in other orthopedic contexts such as surgical recovery. As the validation proceeds and is refined over time for various populations and disease states, there will be a need to modify and tune the specific derived measures, further motivating the need for modular architectures such as those described later.

Stage 1 - Instantaneous mobility state inference: a decision tree generated using the C4.5 algorithm was developed to classify tri-axial accelerometer data collected over a short period of time into a set number of mobility states. For example, it is common to extract descriptive features from accelerometer data from a 1 second window sampled every minute to determine the mobility state for that minute. Commonly used features include variance of the acceleration magnitudes and the Fourier coefficients of the acceleration magnitudes [31]. The classifier is typically pre-trained with a large number of labeled training data collected from multiple users. More recent smartphone operating systems, however, have made the instant-based mobility state inference function available as a system API, and the use of these APIs are usually recommended as they are able to leverage special built-in hardware for more efficient continuous sensing (e.g. iPhone's M7 Motion Sensing Chip) than an app running in the application-level can achieve.

Stage 2 - Activity segmentation: the instantaneous mobility states inferred by Stage 1 are, however, not particularly useful in themselves. Durational context needs to be added, for example, compared to the information: *John was walking at 6:01 AM*, the information: *John walked for 15 minutes from home to subway station from 6:00 AM to 6:15 AM* is more meaningful in understanding the user's behavior. At this stage, a segmentation algorithm is developed to segment a series of mobility states into *activity segments*, each of which represents a period of time in which the user maintains a mobility state. However, one important requirement is to take into account the potential errors and uncertainty contained in Stage 1's output. For example, signals of being in a vehicle that slows down momentarily for a stop sign could be mistakenly classified as *sedentary* rather than its true state, *in vehicle*. The segmentation algorithm needs to take the uncertainty into account and infer the activities that are consistent with the user's real behavior.

A well-known technique that can incorporate this uncertainty is *Hidden Markov Models* (HMM) [32]. HMM is a generative probabilistic model in which a sequence of observable variables X is generated by a sequence of internal hidden states Z using an emission probability function. In this case, mobility states inferred in Stage 1 are taken as the observable variables X and assume they were generated by the user's true but unobservable activity states Z . Uncertainty in Stage 1's inference results is encoded in the emitting probability function $P(x/z)$, and the *Baum-Welch* algorithm is used to infer the maximum likelihood of activity states Z . Then, a sequence of consecutive activity states Z that have the same maximum likelihood state will form an activity segment.

Stage 3 - Location association and correction: at this stage each activity segment is associated with the location data based on the collection time and conduct error correction. A location data point is composed of longitude, latitude, and sometimes includes accuracy. Both GPS-captured and WiFi-enabled location data are known to have ~ 10 to 40 meters of error [33], [34]. For different activities, different approaches are used to improve the location accuracy. For non-stationary activities, such as walking or in vehicle, a *Kalman filter* is used along with a map-matching technique to snap location points to the streets that the user was most likely on [35]. For stationary activities, the mostly likely location at which the user stayed can be inferred by taking the median values of the latitude and longitude among all the location samples [36].

Stage 4 - Summarization: finally, statistics for a users' activities that are known to be relevant to the well-being or to a certain disease are computed. Say et al. [29] shows that for RA management, the majority of rheumatologists they interviewed identified 1) time spent walking, 2) time away from the house, and 3) gait speed as most useful in defining the level of RA disease activity. Rheumatologists preferred a simple, visual format that demonstrates trends over days, weeks, or months, which can help them make better management decisions and ultimately result in decreasing permanent joint damage.

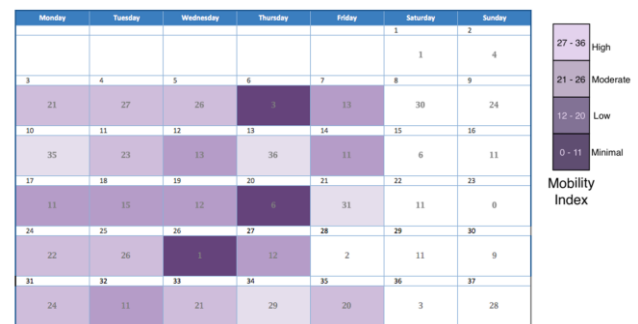


Fig. 1. Calendar view with the daily color coded mobility index for RA management. As described above the clinical validity of this measures is undergoing evaluation within RA and other orthopedic contexts.

Based on these insights, the above-mentioned statistics are computed in daily, weekly, and monthly bases and are presented as a calendar with a color coded index per day derived from a weighted sum of the above three factors using weightings suggested by rheumatologists to indicate the importance each factor.

B. Acoustics

Nonverbal characteristics from speech can strongly convey a speaker's psychological and emotional state [37] as well as other more nuanced social expressions (e.g. sarcasm [38]). There has been extensive work in the understanding of descriptors that can be derived from modulation, spectral characteristics and also energy based measures that can be indicative of stress [39],[40], emotion [41] or even depression [42]. In the context of pain such indicators are valuable. Clearly, stressful utterances, interjections and exclamations can be indicators of acute pain episodes.

However, a more interesting contribution from leveraging audio data is in monitoring at a *larger temporal scale* which has implications for long term chronic pain management. Underlying negative psychological states such as stress, anxiety or depression [43] are well understood to have negative consequences in chronic pain management [44]; such is the case that psychosocial treatments such as *cognitive behavioral therapy* (CBT) are a prevalent course of action [45]. Particular manifest behaviors such as rumination and catastrophization [46] are of interest in this context and are often vocally expressed. In the following example we demonstrate how passively acquired audio streams from smartphone microphones can deliver such high level inferences, in this example by the logging of stressful situations. We must note here that since acoustic processing methods within smart-phone deployed systems are particularly prone to privacy breaches, this leads to the preference to locally process signals that contain intelligible data. This in turn then necessitates methods with low demand on the battery which would require the use of the least computationally expensive methods. This method in the following is designed to take into account of this trade-off. However, there remains a research opportunity in the utilization of more sophisticated detection techniques [47],[48] into a mobile application.

Case example: detection of stressful situations from voice for chronic pain (CP) management.

Stage 1 - Noise detection: the first stage is the detection of *noise* in the sense of any significant level of sound. This is an essential first stage in the passive monitoring to retain only non-silent segments of the data. Simple thresholding on the root-mean-square values over a window of raw audio data are applied to differentiate the segments of *noise* versus silence [49].

Stage 2 - Human speech detection: following this a further classification between as human voice or *noise* caused by another source is required. To this end, three simple features: (1) number of auto-correlation peaks, (2) non-initial auto-correlation peak, and (3) relative spectral entropy to robustly detect presence of human voice in audio streams [50] are used. Moreover, the same three features can be retained to compute important prosody information regarding pitch and speaking rate [51]. This is important to note in terms of privacy. Previous research shows that the reconstruction of verbal information requires at least pitch and 2 harmonics, thus in using the above three features verbal content remains anonymous but prosody can be inferred [51]. In addition to being privacy sensitive, these features are computationally efficient and can work robustly to recognize voice in acoustically noisy environments [50]. This

method has been used in several embedded and mobile phone systems [30][52] and deployed in systems across several real world studies. These studies have demonstrated the inference of face-to-face conversation quality [53], social networks [51][54] as well as social isolation and depressive symptoms [43] which all have implications for long term pain management.

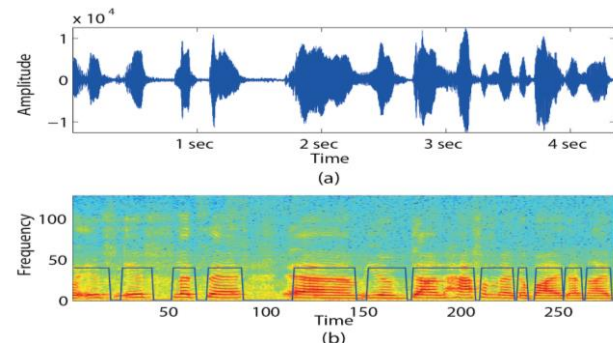


Fig 2. (a) amplitude values for 5 second recording of audio (b) spectrogram of the recording with blue lines showing inferred segments of human voice.

Stage 3 - Speaker identifier: once human voice is detected, it must be determined whether the voice came from the phone user. In Lu *et al.* [55], a Gaussian Mixture Model (GMM) based universal background model of all speakers is pre-computed, the voice of the phone carrier is opportunistically sensed from phone conversations, and a model for the phone carrier's voice is learnt that is different from the universal background model. Subsequently, the learnt model is used to detect the phone carrier's voice. At this stage features such as Mel-frequency cepstral coefficients (MFCC) where the verbal content can potentially be reconstructed are needed. Therefore all processing is done locally and all privacy-violating features are discarded save for the output.

Stage 4: Stress detection: if a detected human voice is that of the phone carrier the detection of stress is invoked. The following characteristics typify stress: the distribution of spectral energy switches to higher frequencies. Also, speaking rate and variability in pitch increases whereas pitch jitter decreases. A recent study showed a non-linear speech production model, based on a Teager Energy profile, can capture stressfulness in noisy outdoors situations. A GMM based classifier utilized these changes as features for both indoor and outdoor application [56]. Also, to account for idiosyncrasies during stressful conditions. A small amount of individual data can be used to calibrate personalized models. It was found that classifier accuracy improved by 8% if 2 minutes of labeled data is provided by the individual. Similar to stage 3, privacy violating features are discarded once the classification output has been obtained.

Stage 5: Logging: the number of stress occurrences can be logged giving a passive indicator of stress frequency with respect to time and location. Such continuous psychological monitoring gives a rich source of information for CBT suggested for chronic pain management. Since the log is updated only when the prior 4 stages are invoked on demand, the computational overhead is minimized.

In other health related applications, the use of acoustic data in this way to infer higher level information about a user is not without precedent. Using similar privacy sensitive acoustic features there have been attempts to infer mental state or well-

being with acoustics. These systems use acoustic data within multimodal frameworks using synchronous data from other smartphone sensors [43],[57],[58],[59]. Going forward, this suggests that the use of acoustic data for complex applications such as pain management would be one component within a multimodal frameworks.

C. Facial Expressions

An important and observable modality that can reflect internal state is facial expression. The prevalence of imagers from our laptops to our lampposts has greatly increased the availability of facial imagery as a source of data about an individual's state. Studies show that facial expressions during acute pain had a high level of consistency and repetition even when elicited by different stimulating modalities [60]. Moreover, a finding by Kappesser & de C. Williams [61] showed that a specific pain face can be distinguishable from faces expressing other negative emotions.

There have been numerous works in the automation of pain face recognition by way of analyzing video imagery from BSP/AC studies. An early example is [62], where face shape features were used with artificial neural networks to classify images of subjects' faces in a normal mood versus images taken from a pain inducing task. Lucey *et al.* [63] publically released the widely used UNBC-McMaster shoulder pain dataset which contains videos of patients with shoulder pain and a temporally concurrent pain score based on the Prkachin and Solomon Pain Intensity score (PSPI) [64]. Several subsequent studies in automated recognition followed [65], [66], [67], which utilize a range of image features from Active Appearance Models (AAM) based features, tracked anatomical points [68], discrete cosine transforms and local binary pattern [69] features with Support Vector Machines (SVM) to classify pain faces. Sikka *et al.* [70] addressed the drawback labelling pain expressions at a sequence level leading to temporal uncertainties in terms of onset and end. The authors proposed a multi-segment multi-instance learning (MS-MIL) framework to determine expressive subsequences within super-sequences that contain pain expression. The same authors extend the machine vision capability beyond adults by demonstrating successful recognition of pain expression in postoperative children [71]. Finally, interesting findings by Bartlett *et al.* showed machine vision methods outperforming human observers in distinguishing real versus fake expression of pain [72][73], this has important implications in the reliability of future ubiquitous pain monitoring systems.

Although much work has been done in the wider machine vision communities, to our knowledge there are no smartphone based systems designed to detect facial expressions of pain and only a few smartphone based systems for face expressions in general. This is primarily due to the difficulties acquiring usable imagery from mobile devices from real world situations. Motion artifacts, out of plane head rotations, various lighting conditions and occlusions add to the difficulty with smartphones. In addition, power and computational constraints add to the challenge of locally processing if needed and the temporal concurrence of when images are captured with the onset of pain experiences cannot be guaranteed. Additionally, the effect of acting for the camera also generates a further

confound, though as mentioned above acted and natural pain expressions can be automatically differentiable.

However, all of the aforementioned difficulties are only problematic if we consider single images or videos in isolation. If multiple instances taken over a long periods are analyzed collectively a broader perspective can be gained and could lend support to the inference of *spatiotemporally large scale behaviors*. For example, studies have shown that factors such as social isolation [74] are linked to physical pain and can be detrimental to chronic pain management [75] and factors into CBT [45]. A number of relevant quantities can be extracted from images that are descriptive of social isolation. Emotional expressions from the user's face [12] can be detected can augmented to further contextual factors such as scene type detection [76], the number of people in the image and even the overall mood of the group of people in the image [77]. These measures can be further augmented to other relevant data such as usage of social networking apps, number of calls or instant messages made or received to construct an index for social isolation. Also, the use of further data modalities such as audio which could aid in reducing ambiguities in situational context and could also aid in identifying the true number of people present. In principle, a longitudinal analysis of trends in the index of social isolation can be used to feedback into CBT. The staged processing described in earlier sections could also be applied to expression extraction from images. The onset of more widespread implementation is bolstered by emerging systems such as in Suk *et al.* who proposed an initial framework for real time locally processed mobile use demonstrated the successful classification of seven affective states [78]. The recently released Affdex software development kit could facilitate the implementation of such analyses along with new datasets that contain real world ecologically valid datasets such as the AM-FED dataset [79].

D. Self-Reports as Signals

Self-report data are somewhat different from the signals previously discussed not only because intentional action is required from the individual to contribute data, but also because the data generated exists at a higher semantic level from inception. As such, rather than discussing stages of processing of self-report data, we share insight on modern self-report techniques and considerations.

The goal of much work in sensing is to reduce or remove the need for user input in data collection systems. While great strides have been made in sensing and detecting patient state, we are not yet in a position to collect all of the data relevant to the pain patient and clinician using passive techniques. Some aspects of behavior, such as detecting whether a patient struggles to button their shirt in the morning, is technically possible but not broadly and affordably deployable given current technology, though we can imagine it becoming so in the future. Then there are cases where we are able to sense the behavior but not fully understand the context without input, such as a case where a patient walks more slowly to work. In

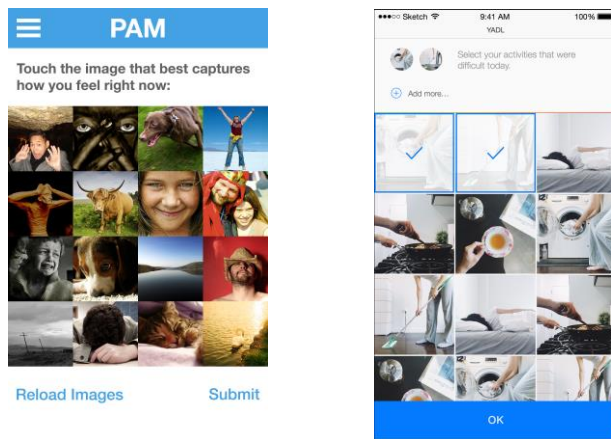


Fig. 3. Example screenshots from PAM (left) prompting the self-report of emotional state using representative images and YADL (right) prompting selection of difficult daily activities as a proxy to pain level.

this case, it is not known whether this is because of pain, another ailment, or simply commuting with a friend or colleague who walks more slowly.

The current standards for reporting the level of pain one is experiencing are pain scales and maps. Pain thermometers and pain faces (e.g. Iowa Faces of Pain) remain the most widely used and do translate well to digital experiences [80]. There have been several adaptations and improvements on user input for these assessments [81]. Pain maps are basic digital maps of the human body that allow patients to document/describe the location and nature of pain. It should also be pointed out that the assessment of pain is quite complex (due to variability in individual pain thresholds, diagnosis, chronic vs acute, and so on), and the community is not fully in agreement on how best to do so [82].

Like other data streams discussed in this paper, we consider self-report to be largely if not exclusively temporal. Data are collected at a specific time and those data represent a certain time span. Unlike most sensed data, however, the time data is collected and the time span the data ‘cover’ are often substantially different in two important dimensions:

Timeliness of data. Self-report data are out of necessity a form of recall, with the patient recounting experiences that have occurred in the past. While a goal of self-report is often to reduce the amount of time between data collection and the recalled experience, practically there is a great deal of variability in these time lags. For example, data may be requested from a patient about an event as it is detected, but the patient may not respond for several hours or even days.

Time span of the data. A single self-report data point maybe reflect a narrow point in time, such as an emotional response to a stimulus or level of pain experienced. Such measurements are often referred to as measures of *state*. On the other hand, the patient may be asked to report data in such a manner that a longer time span is recalled, such as emotional state over the past year, or even irrespective of time span, such as a patient’s personality type. Such measurements are referred to as *trait*. [83].

Self-report originated as one of the main forms of data transfer between patient and clinician, but the advent of modern medicine relegated it to completing forms and sharing

symptoms as precursors to testing. More recently, mobile devices have led to a resurgence of self-reported data through a class of techniques labeled Ecological Momentary Assessment (EMA) in which patients are polled for information, typically using abbreviated versions of gold standard clinical and behavioral assessment forms, frequently and *in situ* [84]. The central argument for the use of EMA is that while the data may yet be subject to interpretation and distortion, the effects of recall bias and error can be mitigated by timeliness of surveillance [85]. To further address recall bias a method was proposed by Rahman et al. [86] where EMA queries are presented with the addition of contextual information about the time, location activity and acoustic state of the user during the time in question: *contextual recall*. This allows for the EMA activation time to be more flexible, the results the study showed improved recollection of stress levels compared to when no context was given.

Self-report and EMA have continued to evolve beyond simply asking patients standardized questions. The large touch screens on current generation mobile devices have afforded more interactive, visual methods for assessing patients. These methods are often employed to mirror and simplify ‘pen and paper’ assessments rather than simply translate them to digital. The Photographic Affect Meter (PAM), for example, asks patients to reflect on their emotional state by choosing from a series of emotionally representative images. Responses take a few seconds to complete and are calibrated to and correlate with the 20-item PANAS, widely considered the gold standard for measurement of emotional state [87].

Another approach to self-report has been to find more straightforward and answerable representations of complex or difficult behaviors or experiences. Pain level, for example, is notoriously difficult to measure reliably, so many clinicians focus on pain interference or Activities of Daily Living (ADLs) to assess what patients are or are not able to do as a result of the pain they experience. This can serve as both an adequate proxy for pain levels and a more direct approach to treating the patient to return to normal and desired activities. In a variant of ADLs, a system called YADL (Your Activities of Daily Living) [88] was designed to allow patients to rapidly and intuitively sort through photos of common activities to document which are easy or difficult for them as a result of their pain. A daily version of YADL presents patients with only those images they have selected as difficult, prompting them to choose which have been hard for them on that particular day, or which they avoided because of anticipated pain. Such an approach allows for an improvement on the traditional means of assessing ADLs along with more fine-grained, frequent data on patient experience. YADL is currently undergoing validation, as with the mobility index mentioned above, the modular design of YADL will allow it to be tuned and adapted to particular populations and disease states.

An alternative modality for self-reporting which does not rely on visual prompts or stimuli are tangible interfaces. Such methods may for some be a more intuitive way to convey an internalized level of pain. By leveraging the gripping action as is a natural gestural response when pain is experienced, Adams et al. [89] developed *Keppi* a small lightweight squeezable stick designed to be continuously carried by the user. The device consists of a conductive foam based, force-sensitive resistor

(FSR) covered in soft rubber with embedded signal conditioning, an ARM Cortex-M0 microprocessor, and Bluetooth Low Energy (BLE). Although still in a developmental phase an in-lab feasibility study showed participants were able to consistently map pressure to four stratified levels of potential pain as well as match a dynamic visual cue when given visual feedback of the pressure value. Further real world evaluation and miniaturization of this device with pain participants is in process.



Fig 4. Prototype of *Keppi*, - tangible and portable self-reporting device for pain.

E. Modular System Architecture

In the previous sub-sections the contribution of individual signals and modalities were discussed. However, *mHealth* systems will be most effective in specific disease contexts and across contexts if they incorporate multiple active and passive data streams. Adopting a modular system architecture is critical in order to address the system complexity associated with heterogeneous and variable data-stream inputs. A modular architecture will allow these systems to be adapted as new apps and device capabilities appear, to be tailored to particular needs and populations, and to serve as a building block for other health related applications. Moreover, as particular signal processing methods improve, they can more readily be evaluated and integrated into modular systems. We provide two examples of modular system architectures in this section, one optimized for mobile device processing, and the second for server side processing. It is worth noting here that though we focus here on modular systems, another interesting yet under-investigated line of research in *mHealth* is the use of advances in data compression. Since power consumption is a limiting factor where local processing is often required, efficiency is desirable both in terms of transmission energy conservation [90] and the codification of the various aforementioned signals to reduce data size.

1) SAINT: Scalable Sensing and Inference Toolkit

Each passive sensor data type generates a large volume of raw data whose transmission can tax bandwidth and battery resources. Moreover, these data may require local processing to address privacy concerns, e.g., detecting the presence of a human voice [91]. However, it is challenging to run the sensing and signal processing in real-time on resource constrained mobile phones. A poorly written signal processing algorithm may consume large amounts of CPU cycles and be a significant drain on the battery. Similarly if intermediate storages during raw sensor recording or signal processing are not cleaned then the system can be out of memory easily for high data rate sensor streams (e.g., audio or video).

To this end, the open-source sensing and inference toolkit SAINT was developed [92] that can assist sensor data collection

and signal processing inside the phone. In SAINT APIs to collect 30 different raw sensor traces and processed data-streams are provided. SAINT tackles these challenges with time- and memory-efficient implementations of its sensing and processing libraries. Fast signal processing algorithms were implemented in native layers (i.e., not on a virtual machine) with low time-complexity [93]. The implementations are also tested over multiple deployments and were found to be stable and battery efficient. In addition to fast and memory efficient implementations, SAINT also provides admission control and duty-cycling for demanding signal processing tasks to further reduce battery usage. Admission control allows for on-demand triggering of signal processing, e.g. speaker or stress recognitions are only triggered when there is human voice in audio data. Duty-cycling on the other hand saves battery by periodically sampling data streams to approximate a human behavior signal. For example, running activity recognition at 10 second intervals can approximate the subject's total level of physical activity fairly accurately.

The modular design enables reuse and extension of SAINT's built-in sensing and inference capabilities. At its core, SAINT uses a publish-subscribe [94] or *pub-sub* pattern. In a *pub-sub* pattern, publishers produce data for which subscribers listen. An intermediate *event bus* or *broker* manages subscription and data transfer among publisher and subscribers. Subscribers register for specific publisher data to the *broker*; publishers then send data directly to the *broker*, and the *broker* relays the information to registered subscribers. Inside SAINT, any existing raw sensor or derived stream is a publisher. A new derived stream can be created in three steps: (1) registering as subscribers to necessary publishers, (2) performing necessary computation on the publisher data to create the new derived stream, and (3) making the derived stream available as a publisher for later reuse. For instance, a new sleep detector could use the presence of movement, location, human voice, and phone charging status from SAINT to detect if a person is sleeping, then make the output of this detector available for other components to use. Furthermore, registering to existing publishers and making the new derived stream data available can be easily implemented.

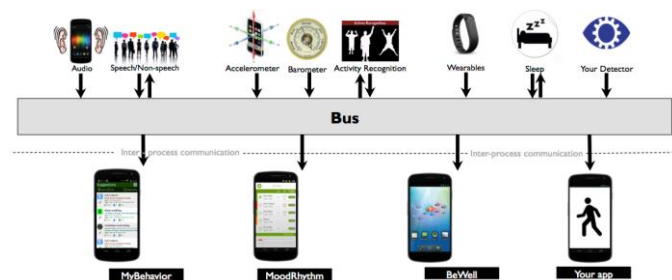


Fig. 5 The architecture of SAINT sensing and inference framework. SAINT provides a unified bus interface to share data across sensing and inference modules. Client applications can connect with SAINT to receive sensed and inferred data.

Another benefit of the *pub-sub* structure is the centralization of sensing and processing that reduces redundancy in computation. In SAINT, if multiple subscribers want the same publisher's data, then the publisher just produces the data once and the broker relays the information to each subscriber,

reducing redundancy. For example, if a sleep and stress detector both run on the phone and want to know if human voice is detected, SAINT detects human voice from audio data once and relays the information to the sleep and stress detectors.

2) Lifestreams

The ultimate goal of many novel data collection techniques described previously is to enable robust and actionable behavioral indicators that can be used to characterize a patient's baseline, and then to identify significant variations, trends and shifts in specific behaviors or symptoms that are relevant to an individual's health. However, making sense and acting on these multi-dimensional, heterogeneous data streams requires iterative and intensive exploration of the data sets, and development of customized analysis techniques that are appropriate to the health domain of interest. Lifestreams is a modular and extensible open-source data analysis stack that runs on the server and is designed to accelerate the development and refinement of sense making techniques [34].

Lifestreams runs on top of mobile data collection systems, such as Ohmage or SAINT [92],[30]. It processes the raw or intermediate data provided by these tools with a multi-layer analysis stack, depicted in Figure 6, consisting of the following: 1) feature extraction, 2) feature selection, 3) inference, and 4) visualization. Each layer consists of modules that process the data provided by the lower layer, and then send the result to the next layer up. From the input data streams, the feature extraction layer computes various statistics, such as daily min/max, average, or the weights of the principal components derived from a set of measurements. The feature compression and selection layer leverages techniques, such as cross-entropy or correlation analysis, and the domain experts' input to select or compose the features that are most relevant to the analysis goal. Then, the inference layer uses the selected features to detect patterns and trends with correlation estimation, and change detection algorithms [95]. Finally, the visualization layer uses different visualization methods to make the inferred patterns or trends actionable for the end user. Lifestreams was found to be useful as a tool for the research coordinators to quickly navigate through the data and provide visual aids to guide the discussion with participants during interview sessions. Its extensible design allows it to be easily extended to support different studies exploring various new data streams and transfer to other domains

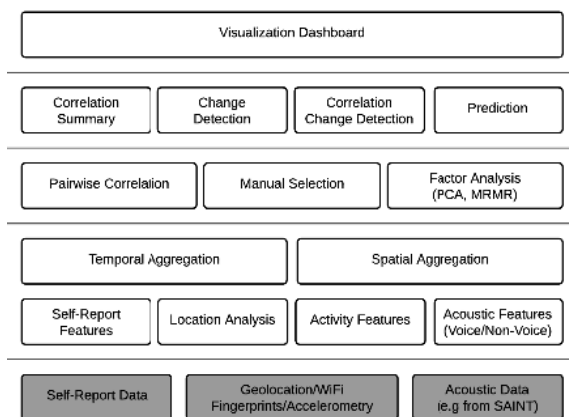


Fig. 6 Lifestreams stack, data enters from the lowest level and is successfully filtered, aggregated and analyzed to produce actionable visualizations for patients and clinician.

III. OPEN CHALLENGES IN INFERENCING AND FEEDBACK

So far we have highlighted, through examples, state of the art methods and tools to achieve the *measurement-inference-management* loop. However, in this section we discuss further factors that are important but either underutilized or remain unresolved. These challenges are most conspicuous when considering *inference* and *management*. In this section we discuss these factors, as well as advances from related studies that have a bearing on future directions.

A. Accounting for Temporal Patterns.

The presence of a circadian rhythm in the experience and perception of pain is well understood. This relates to a variety of chronobiological factors such as daily variations in endorphin and enkephalin concentrations in the pain processing parts of the brain. The review by Junker & Wirz summarizes [96] several studies that include both naturally occurring as well as experimenter induced pain. The studies within showed specific and consistent times of the day when pain perception tended to peak depending on the cause. Investigations in using smartphone technology as a way to monitor daily rhythms have been recently investigated for sleep applications. Most studies have focused on sleep pattern inference by utilizing app usage, charging, screen unlocking, along with an ambient light detector and audio streams to detect sleep durations and quality [97][98][99]. Recently, Abdullah *et al.* [100] demonstrated that by using periods of smartphone non-use, sleep patterns can be inferred using a rule-based algorithm which accurately matched with ground truth from a sleep journal. With the combination of such sleep measures, along with the domain knowledge of the effects of circadian rhythm on specific types of pain, there is great potential for more closely personalized and more timely issuances of interventions for pain management in a daily cycle.

As we have discussed pain conditions with high levels of chronicity require long term management and therapy. Also over enough time, persistent negative psychological states can emerge. It has been shown that, given enough data from long periods, such mental states can be detected. With data spanning 2 weeks Saeb *et al.* [101] demonstrated the use of movement, activity and phone use derived descriptors to infer depression severity levels by way of *supervised learning*, resulting in an accuracy 86.5%. Over a 10 week period, Wang *et al.* [57] investigated correlations between mental well-being measures: depression, stress, flourishing and loneliness scores with tracked information relating to conversation, activity and sleep. Such results demonstrate progress toward inferences on low frequency changes in mental state or even *traits* which are ultimately necessary for CBT related management. Longer period factors such as climate and seasonality have yet to be leveraged in this regard but are readily measurable if studies of this time length are done.

B. Advances in Pattern Recognition

In mobile sensing, a common paradigm to generate predictive models is *supervised learning* [27]. Although a widely used and powerful approach, there remains two critical dependencies. The first is the generation of good features from raw sensor data and the second is the truthfulness of labelled states or behavioral categorizations. Given that there is no standard method to optimally engineer features, a handcrafting approach

is often done using expert domain knowledge, intuition, or other prior experimentation. For lower level inferences good descriptors can be readily determined. However, it is often the case that abstract higher level inferences (such as ‘anxiety’) are desirable, but such tasks would require complex feature engineering and are difficult to map, often leading to poorly performing models. Moreover, such abstract labels have loose definitions or are ambiguous to people and thus difficult to assign correctly when creating a training set.

In light of this, we point to the increasingly popular paradigm of *deep learning* [102][103], which has the capacity to learn feature representations and map low level data to highly abstract categories in a less handcrafted way. However, to date, this has mostly been demonstrated with large unimodal datasets in other application domains. That said, Martinez & Yannakakis [104] proposed a Convolutional Neural Network framework designed to fuse combinations of continuous and discrete 1 dimensional physiological signals with differing sample rates to successfully infer 6 abstract affective states (anxiety, frustration, fun, relaxation, challenge and excitement). Such a method naturally lends itself to the high level behavioral inferences needed for pain management and from the assortment of 1 dimensional signals acquired from mobile sensing. Notably, in this case from the use of wearable devices that can extract such physiological signals. The drawback is the high computational expense in the training and use of highly parametric models, which may not be able to deliver timely output for intervention.

A further significant challenge when dealing with human mobile sensing is the high degree of idiosyncrasy within the data streams. In recent years generalizable methods to capture latent structures in behavioral routine from mobile data have been proposed; such structures could be used as personal baselines from which anomalous behavior could be inferred. Early work by Eagle & Pentland [105] proposed a method to establish eigen-representations of behavior from mobile data to model structures in day to day routine. This method was effective in modelling the idiosyncratic patterns in a personal routine as well as generating a way to find similarities between groups of people. However, this approach relies on rich, densely-sampled high quality data which cannot be realistically acquired in real world sensing. To this end, Zheng et al. [106] proposed a collaborative filtering model to overcome sparsity in a single user’s data based on the intuition that many users follow similar behavior patterns; of course this is dependent on the availability of multiple users’ data.

Another paradigm which can address the problem of gaps in data and/or idiosyncratic variability is *multi task learning* (MTL) [107]. Simultaneous *supervised learning* is applied to a set of different but related classification or regression tasks. For example, data from a user from two different weeks can be used to infer a behavior for each week where some data maybe missing one of the weeks. Learned parameters common to both tasks (weeks) can be leveraged to bolster the accuracy of the partial week when compared to training two models separately. Similarly this can be done in terms of different people as tasks thus accounting for commonalities or differences between people. Romera-Paredes et al. [108] recently showed this capability with facial expressions of pain as well as with electromyographic signals by exploiting the *transfer learning*

property in MTL. In a similar study, these authors also show an improvement in the recognition of pain face expression using OrthoMTL [109] which assumes features that describe identity are unrelated (orthogonal) to features that describe pain expression. In essence this method calibrates to each person by simultaneously learning on identity features and pain expressive features a separates them. In principle this can be applied to data from any modality or combination of modalities as long as this is consistent between tasks. Moreover, advances in MTL show the learning process is equivalent to a single convex optimization process [110], as such there is potentially fewer process iterations when compared to repeatedly training multiple models.

Finally, it is worth noting here that a major barrier to the advancement in predictive modelling is the availability of data or lack thereof. However, in terms of *spatiotemporally small scale behaviors* in pain expression a new extensive multimodal dataset called *EmoPain* was recently released [14]. It explicitly addresses this requirement gap for chronic back pain. Contained within are: high resolution face imagery, surface electromyography, acoustics and whole body motion capture. Two sets of pain labels are included, based on facial expression and known body motion based behaviors [7], the authors set baseline modelling frameworks for each label set using detected facial points with SVM classifiers for pain face recognition and speed and posture based features with Random Forest models for the body behaviors. Public datasets for *large scale* behaviors are still very rare due to the challenges of de-identification, especially for a dataset with patient information.

C. Feedback to Users

Interfacing with the user once inputs, measurements, and inferences have been made is dependent on the type and nature of the required management. Currently, there are numerous commercial pain management mobile apps which are principally designed for specific functions, including self-monitoring, pain intensity tracking, information provision, medication management, and relaxation training. However, there remains a lack of traction with users. Recent reviews [111][112] and commentaries [113][114] unanimously suggest a lack of clinical engagement and regulation in the development of such software, and efficacious claims are underpinned by little evidence. Moreover, the apps relating to chronic pain self-management are still relatively simplistic in how they account for cognitive behavioral factors which are essential for persuasive and effective interfacing.

Some commercial products such as Habit Changer is described as utilizing cognitive behavioral components within the management strategy but details are not provided. Some newer products such as the WebMD PainCoach enables the user to monitor pain, set and track activity goals and generates related messages. These new systems are promising in terms of personal monitoring, but future versions call for further functionality in using this information to effectively respond to episodes of demotivation which are ultimately at the root of poor progress [115].

In terms of general applications, mobile systems can be categorized in three groups in terms of feedback strategy: (1) aggregation of data into summary statistics, often augmented with attractive visualizations. For instance, Ubitfit [116] or

BeWell [117] which uses background wallpaper to show overall physical activity, social interaction and sleep. This is often purposed for goal setting or gentle priming towards the goal [118], (2) visualization of data which relies on the users to self-explore and reflect on the data [119][120] and (3) the delivery of generic recommendations that are globally applicable or are tailored to specific subgroups based on demographics, culture or lifestyle [121] [122]. However, these strategies do not make use of any in-depth analysis and may not only miss potential opportunities to effectively intervene, but may also fail to deliver effective feedback that is actionable and relevant to the individual.

That said, increased persuasiveness in a real world deployment has been demonstrated. Rabbi et al. [123] proposed a context driven food and exercise suggestion engine *MyBehavior*. In this system text based suggestions which are perceived to be low effort and familiar to each user are generated based on both passive and actively sensed data. Such perceptions of familiarity and low effort increases the likelihood of actualization according to behavioral theory [124]. Results over 10-week user trials show significant increases in adoption when compared to randomly issued suggestions. Similar persuasive feedback would be of great benefit for musculoskeletal pain for example where avoidance and demotivation to do regular exercise is a prevalent problem.

An interesting direction in *spatiotemporally small scale* feedback is the sonification of patient's movements during physiotherapy [115]. Common motion related traits are self-guarding movement and a reduction in proprioception. Sound feedback exploits the tight links between the auditory and motor parts the brain. Singh et al. [125] proposes a design framework: *Go-with-the-flow* for sonified wearable systems for chronic pain exercising. Initial results indicate an impact on body awareness, increased motivation to reach exercise targets, and positive gains in compensatory motion and self-efficacy.

Ultimately, all patient sensing has the purpose of informing decisions and actions. The focus of this review is on the capture and processing of relevant data sources to inform user feedback but in the end effectively closing critical feedback loops of health will depend on effective ways of presenting to and interacting with the end user/patient.

IV. CONCLUSION

Patient centric data streams will increasingly inform disease management and diagnosis. They will be used to contextualize and personalize patient response to treatment and to build applications that support patients in their own self-management. In this review we focused on the use case of pain measurement and management because it has broad applicability, significant implications for both clinical outcomes and patient quality of life, and because mobile data streams can fill a significant gap in current approaches to measurement. While today's technology is ready to be put to use to improve pain measurement and management, there remain challenges and opportunities for further research.

The greatest short term challenges are to develop the evidence base for these approaches and to address usability and relevance for both patient and clinician. In addition, a modular data architecture is critical to promote the inclusion of new data

sources (applications and devices) and the adaptation to new and specialized users (conditions and demographics) over time.

REFERENCES

- [1] D. Estrin and I. Sim, "Open mHealth architecture: an engine for health care innovation". *Science (Washington)*, 330(6005), 759-760, 2010.
- [2] W. Nilsen, S. Kumar, A. Shar, C. Varoquiers, T. Wiley, W. T. Riley, and A. A. Atienza, "Advancing the science of mHealth" *Journal of Health Communication*, 17(sup1), 5-10, 2012.
- [3] C. Eccleston, "Role of psychology in Pain Management", *Brit. Journal of Anaesthesia* 87(1), 144-52, 2001.
- [4] H. Merskey and N. Bogduk, *Classification of Chronic Pain*. Seattle: International Association for the Study of Pain Press, 210, 1994.
- [5] H. Merskey, "The taxonomy of pain". *The Medical clinics of North America*, 91(1), 13-20, 2007.
- [6] J. D. Loeser and R. Melzack, "Pain: an overview". *The Lancet*, 353(9164), 1607-1609, 1999.
- [7] F. J. Keeffe and A. R. Block. "Development of an observation method for assessing pain behaviour in chronic low back pain patients". *Behaviour Therapy*, 13(4), 1982.
- [8] W. Fordyce, *Behavioural method for chronic pain and illness*, St Louis: Mosby, 1976.
- [9] M. J. L. Sullivan, P. Thibault, A. Savard, R. Catchlove, J. Kozey, and W. D Stanish. "The influence of communication goals and physical demands on different dimensions of pain behaviour", *Pain*, 125, 270-277, 2006
- [10] S. Narayanan and P. G. Georgiou, "Behavioral signal processing: Deriving human behavioral informatics from speech and language". *Proceedings of the IEEE*, 101(5), 1203-1233, 2013.
- [11] M. Pantic, A. Pentland, A. Nijholt, and T. S. Huang, "Human computing and machine understanding of human behavior: a survey". In *Artificial Intelligence for Human Computing*. Springer Berlin Heidelberg. pp. 47-71, 2007.
- [12] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang. "A survey of affect recognition methods: audio, visual, and spontaneous expressions". *IEEE Trans. Patt. Anal. Machine Intel*, 31(1):39-58, 2009.
- [13] E. Vyzas and R. W. Picard, "Affective pattern classification". In *Proc. AAAI Fall Symp. Series: Emotional and Intelligent: The Tangled Knot of Cognition*, pp. 176-182, 1998.
- [14] M. S. H. Aung, S. Kaltwang, B. Romera-Paredes, B. Martinez, A. Singh, M. Cella, M. Valstar, H. Meng, A. Kemp, M. Shafizadeh, A. C. Elkins, N. Kanakam, A. de Rothschild, N. Tyler, P. J. Watson, A. C. de C. Williams, M. Pantic, and N. Bianchi-Berthouze, "The automatic detection of chronic pain-related expression: requirements, challenges and a multimodal dataset", *IEEE Trans. Affective Computing*, (to appear).
- [15] J. W. S. Vlaeyen and S. J. Linton. "Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art". *Pain*, 85(3), 2000.
- [16] S. Felipe, A. Singh, C. Bradley, A. Williams, and N. Bianchi-Berthouze, "Roles for personal informatics in chronic pain" in *Pervasive Health 15*. 2015.
- [17] M. P. Jensen, J. A. Turner, and J. M. Romano, "Changes after multidisciplinary pain treatment in patient pain beliefs and coping are associated with concurrent changes in patient functioning", *Pain*, vol. 131, no. 1-2, pp. 38-47, Sep. 2007.
- [18] P. Rainville, Q. V. H. Bao, and P. Chrétien, "Pain-related emotions modulate experimental pain perception and autonomic responses". *Pain*, 118(3), pp.306-318, 2005.
- [19] W. C. King, and D. S Bond. "The importance of pre and postoperative physical activity counseling in bariatric surgery." *Exercise and sport sciences reviews* 41.1:26, 2013.
- [20] D. M. Bravata, C. Smith-Spangler, V. Sandaram, A. L. Gienger, N. Lin, R. Lewis, C. D. Stave, I. Olkin, and J. R. Sirard, "Using pedometers to increase physical activity and improve health: a systematic review." *JAMA*, 298.19: 2296-2304, 2007.
- [21] C. J. Greaves, K. E Sheppard, C. Abraham, W. Hardeman, M. Roden, P. H. Evans, and P Schwarz, "Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions." *BMC Public Health* 11.1: 119, 2011.
- [22] M. S. Patel, D. A. Asch, and K. G. Volpp. "Wearable devices as facilitators, not drivers, of health behavior change." *JAMA*, 313.5: 459-460, 2015.
- [23] C. R. Richardson, T. L. Newton, J. J. Abraham, A. Sen, M. Jimbo, A. M Swartz, "A meta-analysis of pedometer-based walking interventions and weight loss." *The Annals of Family Medicine* 6.1: 69-77, 2008.

- [24] H. H. Dodge, N. C. Mattek, D. Austin, T. L. Hayes, and J. A. Kaye, "In-home walking speeds and variability trajectories associated with mild cognitive impairment." *Neurology* 78.24 (2012): 1946-1952.
- [25] J. Bort-Roig, N. D. Gilson, A. Puig-Ribera, R. S. Contreras, and S. G. Trost. "Measuring and influencing physical activity with smartphone technology: a systematic review." *Sports Medicine* 44.5: 671-686, 2014.
- [26] M. A. Case, H. A. Burwick, K. G. Volpp, and M. S. Patel, "Accuracy of smartphone applications and wearable devices for tracking physical activity data." *JAMA*, 313.6: 625-626, 2015.
- [27] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell. "A survey of mobile phone sensing." *IEEE Communications Magazine*, 48.9: 140-150, 2010.
- [28] T. Sohn, A. Varshavsky, A. LaMarca, M. Y. Chen, T. Choudhury, I. Smith, and E. De Lara, "Mobility detection using everyday gsm traces". In *UbiComp 2006: Ubiquitous Computing*. Springer Berlin Heidelberg, pp. 212-224, 2006.
- [29] P. Say, D. M. Stein, J. S. Ancker, A. Hsieh, J. P. Pollak, and D. Estrin. "Smartphone data in rheumatoid arthritis - What do rheumatologists want?" In press, *Proceedings of the AMIA Annual Symposium* November 2015.
- [30] H. Tangmunarunkit, C. K. Hsieh, B. Longstaff, S. Nolen, J. Jenkins, C. Ketcham, J. Selsky, F. Alquaddoomi, D. George, J. Kang, Z. Khalapyan, J. Ooms, N. Ramanathan, and D. Estrin. "Ohmage: A General and Extensible End-to-End Participatory Sensing Platform." *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6.3: 38. 2015
- [31] B. Longstaff, S. Reddy, and D. Estrin. "Improving activity classification for health applications on mobile devices using active and semi-supervised learning." *Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, 1-7, 2010
- [32] L. R. Welch, "Hidden Markov models and the Baum-Welch algorithm." *IEEE Information Theory Society Newsletter* 53.4: 10-13, 2003.
- [33] L. Liao, D. Fox, and H. Kautz. "Extracting places and activities from gps traces using hierarchical conditional random fields." *The International Journal of Robotics Research* 26.1: 119-134, 2007.
- [34] Y. C. Cheng, Y. Chawathe, A. LaMarca, and J. Krumm, "Accuracy characterization for metropolitan-scale Wi-Fi localization." *Proceedings of the 3rd international conference on Mobile systems, applications, and services* pp: 233-245, 2005.
- [35] Y. Lou, C. Zhang, Y. Zheng, X. Xie, W. Wang, and Y. Huang, "Map-matching for low-sampling-rate GPS trajectories". In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM, pp. 352-361, 2009.
- [36] C. Hsieh, H. Tangmunarunkit, F. Alquaddoomi, J. Jenkins, J. Kang, C. Ketcham, B. Longstaff, J. Selsky, B. Dawson, D. Swendeman, D. Estrin, and N. Ramanathan. "Lifestreams: A modular sense-making toolset for identifying important patterns from everyday life." *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems* 11: 5, 2013
- [37] F. Weninger, M. Wöllmer, and B. Schuller, "Emotion Recognition in Naturalistic Speech and Language—A Survey". *Emotion Recognition: A Pattern Analysis Approach*, 237-267, 2014.
- [38] P. Rockwell, "Lower, slower, louder: Vocal cues of sarcasm". *Journal of Psycholinguistic Research*, 29(5), 483-495, 2000.
- [39] J.H. Hansen, W. Kim, M. Rahurkar, E. Ruzanski, and J. Meyerhoff, "Robust emotional stressed speech detection using weighted frequency subbands". *EURASIP Journal on Advances in Signal Processing*, (1), pp.1-10, 2011.
- [40] M. Lai, Y. Chen, M. Chu, Y. Zhao, and F. Hu, "A hierarchical approach to automatic stress detection in English sentences". In *Acoustics, Speech and Signal Processing, ICASSP 2006 Proceedings*. (Vol. 1, pp. I-I), 2006.
- [41] M. Grimm, K. Kroschel, E. Mower, and S. Narayanan, "Primitives-based evaluation and estimation of emotions in speech". *Speech Communication*, 49(10), pp.787-800, 2007.
- [42] Y. Yang, C. Fairbairn, and J. F. Cohn, "Detecting depression severity from vocal prosody". *Affective Computing, IEEE Transactions on*, 4(2), 142-150, 2013.
- [43] M. Rabbi, S. Ali, T. Choudhury, and E. Berke. "Passive and in-situ assessment of mental and physical well-being using mobile sensors." In *Proceedings of the 13th international conference on Ubiquitous computing*, pp. 385-394. ACM, 2011.
- [44] B. A. Arnow, E. M. Hunkeler, C. M. Blasey, J. Lee, M. J. Constantino, B. Fireman, H. C. Kraemer, R. Dea, R. Robinson, and C. Hayward. "Comorbid depression, chronic pain, and disability in primary care." *Psychosomatic medicine* 68, no. 2: 262-268, 2006
- [45] D. M. Ehde, T. M., Dillworth, and J. A. Turner, "Cognitive-behavioral therapy for individuals with chronic pain: efficacy, innovations, and directions for research". *American Psychologist*, 69(2), 153, 2014.
- [46] J. A. Turner, S. Holtzman, and L. Mancil, "Mediators, moderators, and predictors of therapeutic change in cognitive-behavioral therapy for chronic pain." *Pain*, 127(3), 276-286.
- [47] G. Saon, S. Thomas, H. Soltan, S. Ganapathy, and B. Kingsbury, "The IBM speech activity detection system for the DARPA RATS program". In *INTERSPEECH* (pp. 3497-3501), 2013.
- [48] S. Thomas, G. Saon, M. Van Segbroeck, and S. S. Narayanan, "Improvements to the IBM speech activity detection system for the DARPA RATS program." In *Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference on*, (pp. 4500-4504), 2015
- [49] H. Lu, Hong, J. Yang, Z. Liu, N. D. Lane, T. Choudhury, and A. T. Campbell. "The Jigsaw continuous sensing engine for mobile phone applications." In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, pp. 71-84. ACM, 2010.
- [50] S. Basu, *Conversational scene analysis*. PhD diss., Massachusetts Institute of Technology, 2002.
- [51] M. D. Wyatt, *Measuring and modeling networks of human social behavior*. University of Washington, 2010.
- [52] T. Choudhury, S. Consolvo, B. Harrison, J. Hightower, A. LaMarca, L. LeGrand, A. Rahimi, A. Rea, G. Bordello, B. Hemingway, P. Klasnja, K. Koscher, J. A. Landay, J. Lester, D. Wyatt, and D. Haehnel, "The mobile sensing platform: An embedded activity recognition system." *Pervasive Computing, IEEE* 7, no. 2: 32-41, 2008.
- [53] T. K. Choudhury, *Sensing and modeling human networks*. Diss. Massachusetts Institute of Technology, 2003.
- [54] Wyatt, D., Choudhury, T., Bilmes, J. and Kitts, J.A., 2011. Inferring colocation and conversation networks from privacy-sensitive audio with implications for computational social science. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(1), p.7.
- [55] H. Lu, A.J. Brush, B. Priyantha, A. K. Karlson, and J. Liu. "SpeakerSense: energy efficient unobtrusive speaker identification on mobile phones." In *Pervasive Computing*, pp. 188-205. Springer Berlin Heidelberg, 2011.
- [56] H. Lu, D. Frauendorfer, M. Rabbi, M. Schmid Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury. "StressSense: Detecting stress in unconstrained acoustic environments using smartphones." In *Proc. ACM Conference on Ubiquitous Computing*, pp. 351-360, 2012.
- [57] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A.T. Campbell. "StudentLife: Assessing Mental Health, Academic Performance and Behavioral Trends of College Students using Smartphones." *ACM Conference on Ubiquitous Computing*. 2014.
- [58] D. Ben-Zeev, R. Wang, S. Abdullah, R. Brian, E. A. Scherer, L.A. Mistler, M. Hauser, J. M. Kane, A. Campbell, and T. Choudhury, "Mobile Behavioral Sensing for Outpatients and Inpatients With Schizophrenia." *Psychiatric Services*, 2015.
- [59] T. Choudhury, "Tracking behavioral symptoms of bipolar disorder using automated sensing and delivering personalized interventions using smartphones". In *BIPOLAR DISORDERS*, (17), pp. 10, 2015.
- [60] K.M. Prkachin, "The consistency of facial expressions of pain – a comparison across modalities". *Pain*, 51(3), pp. 297-306, 1992.
- [61] J. Kappesser and A.C.de C. Williams, "Pain and negative emotions in the face: judgements by health care professionals". *Pain*, 99(1-2), 197-206, 2002.
- [62] M. Monwar, and S. Rezaei, "Pain recognition using artificial neural network", in: *Proc. of the IEEE International Symposium on Signal Processing and Information Technology*, 2006.
- [63] P. Lucey, J.F. Cohn, K.M. Prkachin, P.E. Solomon, and I. Matthews. "PAINFUL DATA: The UNBC-McMaster Shoulder Pain Expression Archive Database", *Face and Gesture*, 57-64, 2011.
- [64] K. Prkachin and P. Solomon, "The structure, reliability and validity of pain expression: The evidence from patients with shoulder pain", *Pain*, 139, pp.267-274, 2008.
- [65] A.B. Ashraf, S. Lucey, J.F. Cohn, T. Chen, Z. Ambadar, K.M. Prkachin, and P.E. Solomon, "The painful face – Pain expression recognition using active appearance models". *Image and Vision Computing*, 27, pp. 1788-1796, 2009.
- [66] Z. Hammal and J.F. Cohn, "Automatic detection of pain intensity", *Proc. Int. Conf. Multimodal Interaction*. 2012.
- [67] S. Kaltwang, O. Rudovic, and M. Pantic, "Continuous Pain Intensity Estimation from Facial Expressions", *Advances in Visual Computing. Springer, Heidelberg*, 7432: pp. 368 - 377, 2012.

- [68] B. Martinez, M. Valstar, X. Binefa, and M. Pantic. "Local evidence aggregation for regression-based facial point detection". *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(5), 1149-1163, 2013.
- [69] T. Ahonen, A. Hadid, and M. Pietikainen. "Face description with local binary patterns: Application to face recognition". *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(12), 2037-2041, 2006.
- [70] K. Sikka, A. Dhall, and M. Bartlett, "Weakly supervised pain localization using multiple instance learning". In *Automatic Face and Gesture Recognition (FG)*, 10th IEEE International Conference and Workshops on (pp. 1-8), 2013.
- [71] K. Sikka, A. A. Ahmed, D. Diaz, M. S. Goodwin, K. D. Craig, M. S. Bartlett, and J. S. Huang, "Automated assessment of children's postoperative pain using computer vision". *Pediatrics*, 136(1), pp.e124-e131, 2015.
- [72] G.C. Littlewort, M.S. Barlett, and K. Lee, "Automatic coding of facial expressions displayed during posed and genuine pain", *Image and Vision Computing*, 27(12): 1797:1803, 2009.
- [73] M.S. Bartlett, G.C. Littlewort, M.G. Frank, and K. Lee "Automatic decoding of facial movements reveals deceptive pain expressions". *Current Biology*, 2014.
- [74] N. I. Eisenberger, "The pain of social disconnection: examining the shared neural underpinnings of physical and social pain". *Nature Reviews Neuroscience*, 13(6), 421-434, 2012.
- [75] M.O. Martel, P. Thibault, and M.J.L. Sullivan, "The persistence of pain behaviours in patients with chronic back pain is independent of pain and psychological factors". *Pain*, 151, 330-336, 2010.
- [76] J. Wu and J. M. Rehg, "Centrist: A visual descriptor for scene categorization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1489-1501, 2011.
- [77] Dhall, A., Joshi, J., Sikka, K., Goecke, R., & Sebe, N. "The More the Merrier: Analysing the affect of a group of people in images". In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition (FG)* (Vol. 12), 2015.
- [78] M. Suk and B. Prabhakaran, "Real-Time Mobile Facial Expression Recognition System--A Case Study". In *Computer Vision and Pattern Recognition Workshops (CVPRW)*, *IEEE Conference on*, pp. 132-137. 2014.
- [79] D. McDuff, R. El Kaliouby, T. Senechal, D. Demirdjian, and R. Picard, "Automatic measurement of ad preferences from facial responses gathered over the internet". *Image and Vision Computing*, 32(10), pp.630-640, 2014.
- [80] C. L. Hicks, C.L. von Baeyer, P. A. Spafford, I. van Korlaar, and B. Goodenough, "The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement". *Pain*, 93(2), pp.173-183. 2001.
- [81] P. Adams, M. Rabbi, T. Rahman, M. Matthews, A. Voids, G. Gay, and S. Voids, "Towards personal stress informatics: comparing minimally invasive techniques for measuring daily stress in the wild." In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, pp. 72-79, 2014.
- [82] M. Schiavenato, C. L. von Baeyer, and K. D. Craig, "Self-Report Is a Primary Source of Information About Pain, But It Is Not Infallible A Comment on "Response to Voepel-Lewis's Letter to the Editor," Bridging the Gap Between Pain Assessment and Treatment: Time for a New Theoretical Approach?". *Western journal of nursing research*, 35(3), pp.384-387. 2013.
- [83] J. M. George, "Trait and State" Affect". In K.R. Murphy (Ed) *Individual Differences and Behavior in Organizations*, 145-171, 1996.
- [84] A. A. Stone, and S. Shiffman, "Capturing momentary, self-report data: A proposal for reporting guidelines". *Annals of Behavioral Medicine*, 24,3, 236-43, 2002.
- [85] D. S. Moskowitz, and S. N. Young, "Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology". *J Psychiatry Neurosci*, 31,1, 13-20, 2006.
- [86] T. Rahman, M. Zhang, S. Voids, and T. Choudhury, "Towards accurate non-intrusive recollection of stress levels using mobile sensing and contextual recall". In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, pp. 166-169, 2014.
- [87] J. P. Pollak, P. Adams, and G. Gay, "PAM: A Photographic Affect Meter for frequent, in situ measurement of affect". *Proceedings of CHI 2011*, 725-734, 2011.
- [88] L. Yang, D. Freed, A. Wu, J. Wu, J. P. Pollak, and D. Estrin, "Your Activities of Daily Living (YADL): An Image-based Survey Technique for Patients with Arthritis". arXiv preprint arXiv:1601.03278. 2016.
- [89] P. Adams, A. Adams, G. Gay, and T. Choudhury. "Keppi: A Tangible User Interface for the Self-Report of Scalar Values", *in submission*.
- [90] Z. Abbas, and W. Yoon, "A Survey on Energy Conserving Mechanisms for the Internet of Things: Wireless Networking Aspects". *Sensors*, 15(10), pp.24818-24847. 2015.
- [91] P. Klasnja, S. Consolvo, T. Choudhury, R. Beckwith, and J. Hightower. "Exploring privacy concerns about personal sensing." In *Pervasive Computing*, pp. 176-183. Springer Berlin Heidelberg, 2009.
- [92] M. Rabbi, T. Caetano, J. Costa, S. Abdullah, M. Zhang, and T. Choudhury. "SAINT: A Scalable Sensing and Inference Toolkit." in *Hotmobile* 2015.
- [93] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Second Edition, New York: Cambridge University Press, 1992.
- [94] P. T. Eugster, P. A. Felber, R. Guerraoui, and A. M. Kermarrec. "The many faces of publish/subscribe". *ACM Computing Surveys (CSUR)*, 35(2):114-131, 2003.
- [95] G. J. Ross, "Parametric and nonparametric sequential change detection in R: The cpm package." *Journal of Statistical Software*: 78, 2013.
- [96] U. Junker and S. Wirz, "Review Article: Chronobiology: influence of circadian rhythms on the therapy of severe pain". *Journal of oncology pharmacy practice*, 16(2), 81-87, 2010.
- [97] Z. Chen, M. Lin, F. Chen, N. D. Lane, G. Cardone, R. Wang, T. Li, Y. Chen, T. Choudhury, and A. T. Campbell, "Unobtrusive sleep monitoring using smartphones". In *Pervasive Computing Technologies for Healthcare, 7th International Conference on, IEEE*, 145-152. 2013
- [98] J. K. Min, A. Doryab, J. Wiese, S. Amini, J. Zimmerman, and J. I. Hong, "Toss 'N' turn: Smartphone as sleep and sleep quality detector." *SIGCHI Conference on Human Factors in Computing Systems, CHI '14*, ACM, 477-486, 2014.
- [99] Y. Bai, B. Xu, Y. Ma, G. Sun, and Y. Zhao, "Will you have a good sleep tonight?: Sleep quality prediction with mobile phone". In *Proceedings of the 7th International Conference on Body Area Networks, BodyNets '12, ICST 124-130*, 2012.
- [100] S. Abdullah, M. Matthews, E.L. Murnane, G. Gay, and T. Choudhury, "Towards circadian computing: "early to bed and early to rise", makes some of us unhealthy and sleep deprived." *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 673-84, 2014.
- [101] S. Saeb, M. Zhang, C.J. Karr, S.M. Schueller, M.E. Corden, K.P. Kording, D.C. Mohr, "Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study", *J. Medical Internet Res.*, 17(7), 2015.
- [102] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives". *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(8), 1798-1828, 2013.
- [103] B. Schuller, "Deep Learning Our Everyday Emotions". In *Advances in Neural Networks: Computational and Theoretical Issues*. Springer International Publishing. pp. 339-346, 2015.
- [104] H. P. Martinez and G. N. Yannakakis, "Deep multimodal fusion: combining discrete events and continuous signals", *ACM International Conf. Multimodal Interaction*, 2014.
- [105] N. Eagle and S. Pentland, "Eigenbehaviors: identifying structure in routine", *Behav. Ecol. Sociobiol.* 63: 1057-66, 2009.
- [106] J. Zheng, S. Liu, L.M. Li. "Effective routine behavior pattern discovery from sparse mobile phone data via collaborative filtering", *IEEE Int. Conf. Pervasive Comput. and Comm.*, pp. 29-37, 2013.
- [107] R. Caruana, "Multitask Learning: A Knowledge-Based Source of Inductive Bias". In *Proceedings of the Tenth International Conference on Machine Learning* (pp. 41-48).
- [108] B. Romera-Paredes, M. S. H. Aung, M. Pontil, A. C de C. Williams, N. Bianchi-Berthouze, and P. Watson, "Transfer Learning to Account for Idiosyncrasy in Face and Body Expressions". *10th IEEE Automatic Face and Gesture Recognition*, 2013.
- [109] B. Romera-Paredes, A. Argyriou, N. Berthouze and M. Pontil, "Exploiting unrelated tasks in multi-task learning". In *International Conference on Artificial Intelligence and Statistics*, pp. 951-959, 2012.
- [110] A. Argyriou, T. Evgeniou, and M. Pontil, "Convex multi-task feature learning". *Machine Learning*, 73(3), 243-272, 2008.
- [111] B. A. Rosser and C. Eccleston, "Smartphone applications for pain management". *Journal of Telemedicine and Telecare* 17, 6, 308-312, 2011.
- [112] C. Lalloo, L. A. Jibb, J. Rivera, A. Agarwal, J. N Stinson, "There's a pain app for that: a review of patient targeted smartphone applications for pain management", *The Clinical Journal of Pain*, 31(6), pp. 557-563, 2015.

- [113] J. E. Richardson and M. C. Reid, "The promise and pitfalls of leveraging mobile health technology for pain care". *Pain Medicine*, 14: 1621-1626, 2013.
- [114] C. Reynoldson, "Assessing the quality and usability of smartphone apps for pain management", *Pain Medicine*, 15: 898-909, 2014.
- [115] A. Singh, A. Klapper, J. Jia, A. Fidalgo, A. Tajadura-Jimenez, N. Kanakam, N. Bianchi-Berthouze, A. Williams. "Motivating People with Chronic Pain to do Physical Activity: Opportunities for Technology Design", In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014.
- [116] S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, and I. Smith, "Activity sensing in the wild: a field trial of ubifit garden." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 1797-1806, ACM, 2008.
- [117] N. D. Lane, M. Mohammad, M. Lin, X. Yang, H. Lu, S. Ali, A. Doryab, E. Berke, T. Choudhury, and A. Campbell, "Bewell: A smartphone application to monitor, model and promote wellbeing". In *5th international ICST conference on pervasive computing technologies for healthcare*, pp. 23-26, 2011.
- [118] S. Consolvo, D. W. McDonald, and J. A. Landay, "Theory-driven design strategies for technologies that support behavior change in everyday life." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 405-414, ACM, 2009.
- [119] M. Kay, E. K. Choe, J. Shepherd, B. Greenstein, N. Watson, S. Consolvo, and J. A. Kientz, "Lullaby: a capture & access system for understanding the sleep environment." In *Proceedings ACM Conference on Ubiquitous Computing*, pp. 226-234, 2012.
- [120] E. K. Choe, B. Lee, M. Kay, W. Pratt, and J. A. Kientz, "SleepTight: low-burden, self-monitoring technology for capturing and reflecting on sleep behaviors." In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 121-132, ACM 2015.
- [121] C. A. Pellegrini, S.A. Hoffman, L. M. Collins, and B. Spring, "Optimization of remotely delivered intensive lifestyle treatment for obesity using the Multiphase Optimization Strategy: Opt-IN study protocol." *Contemporary clinical trials*, 38(2), pp.251-259, 2014.
- [122] R. Kukafka, "Tailored health communication" *Consumer Health Informatics: Informing Consumers and Improving Health Care*, pp. 22-33, 2005.
- [123] M. Rabbi, M.S.H Aung, M. Zhang, T. Choudhury. "MyBehavior: Automated Personalized Health Feedback from User Behavior and Preference using Smartphones", *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2015, to appear.
- [124] B. Fogg, "A behavior model for persuasive design". *ACM Proc. of the 4th international Conference on Persuasive Technology*, 40, 2009.
- [125] A. Singh, S. Piana, D. Pollarolo, V. Gualtierio, G. Varni, A. Tajadura-Jimenez, A. Williams, A. Camurri, and N. Bianchi-Berthouze, "Go-with-the-flow: Tracking, analysis and sonification of movement and breathing to build confidence in activity despite chronic pain," *Human-Computer Interaction*, (in press).