



A Framework for Pervasive and Ubiquitous Geriatric Monitoring

Ramesh Balaji, Karan Bhavsar, Brojeshwar Bhowmick, Mithun B. S., Kingshuk Chakravarty, Debatri Chatterjee, Avik Ghose, Puneet Gupta, Dibyanshu Jaiswal, Sanjay Kimbahune^(✉), Kartik Muralidharan, Arpan Pal, Aniruddha Sinha, and Srinivasa Raghavan Venkatachari

Research and Innovation Labs, Tata Consultancy Services Ltd., Thane, India
sanjay.kimbahune@tcs.com

Abstract. The elderly population is steadily increasing in most countries due to a decline in birth and mortality rate. The population of independently living senior citizens has also become significant. This has led to an active research focus in geriatric wellness. Apart from the usual physical and cognitive decline that proceeds with age, there are much softer aspects like dignity and ability to live independently (age-in-place). If we monitor the subject all the time by instrumenting them heavily, technology can possibly predict and prevent many abnormal situations. However, this curbs the idea of independent living. Hence, while instrumenting and monitoring is an important paradigm in Ambient Assisted Living (AAL), the need for the underlying technology to be non-intrusive or as natural as possible is also pressing. For example, passive infrared (PIR) sensors are quite effective and ubiquitous in detecting presence and movement. In this paper we investigate the opportunities in this sliding balance of intrusiveness versus wellness. We explore ubiquitous and pervasive monitoring to make useful inferences in the subject's activities of daily living (ADL). To this end, we evaluate different type of commercially available sensors such as the Microsoft Xbox KinectTM, video cameras and smart watches, which will augment the PIR sensors at home. Each sensor will allow the elderly to be monitored for specific wellness parameters. We propose a generic framework with algorithms developed in-house, highlighting the use of the Kinect for computing fall-risk through single limb stand (SLS) and postural stability, smart watches to detect multiple activities in an energy efficient manner using in-built sensors, and, finally, computer-vision based methods using video cameras to detect the liveliness of human beings by computing the micro motions of the body during sleep. These findings can be used for both early risk assessment and cognitive training of an individual. Our framework is also extensible to allow the integration of various sensors and algorithms in addition to being customizable for the context in question. A key challenge in building such a generic wellness care framework to support multiple sensor measurements, is handling the trade-off between cost, intrusiveness, and precision.

1 Introduction

In many countries there is a significant rise in the number of senior citizens living independently [2]. As a result, there is an active research focus in the wellness and care of the geriatric population. While monitoring these citizens by heavily instrumenting them with tech could possibly aid in better prediction and prevention of certain ailments, this type of setup interferes with the need of independent living. We therefore explore the possibilities of a ubiquitous and pervasive monitoring system to allow making useful inferences in the subject’s activities of daily living (ADL). To this end, we propose a wellness care framework that encompasses a variety of commercially available sensors such as PIR, Kinect, video cameras, and smart watches that will aid in our goal. A typical sensor/camera/wearables enabled home is captured in Fig. 1.

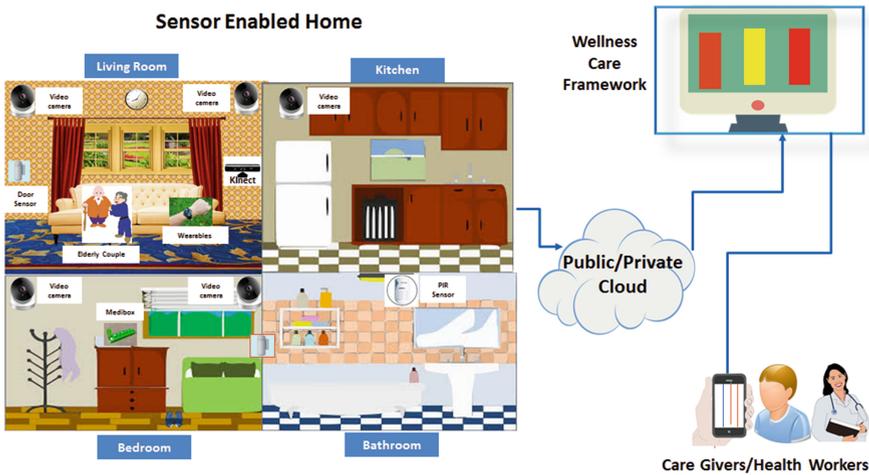


Fig. 1. A typical Smart-Home infused with multiple devices including PIR, video camera, smart medicine-box, wearables and kinect connecting to a Wellness Framework and linking care givers/health workers

The plethora of sensors installed (and worn) in such smart homes will assist in not only collecting data on the daily activities of the elderly but also trigger alerts in case of an emergency, such as a fall, enabling a quick response from care-givers.

Each sensor has a certain characteristic in terms of intrusiveness, or disturbance to the subject. PIR sensors [22] are the least intrusive and only detect movement within their range. Wearables such as smart watches, require to be worn on the body which as a result increase their level of intrusiveness, for example. A key challenge in building such a generic wellness care framework would therefore be in handling the trade-off between cost of data acquisition, the level of intrusiveness, and precision in detecting the activities.

In this work, we will share our experiences in the use of devices such as the Microsoft Kinect in detecting certain wellness metrics as well as how these devices assist in our goal of unobtrusive sensing. We will also highlight certain insights and patterns observed from the data collected from these devices. Finally, we will talk about our extensible and customizable wellness care framework, which will bring all these devices together.

2 The Kinect Isn't Just for Gaming

The Kinect is a motion sensing input device produced by Microsoft for their Xbox range of video game consoles. The Kinect allows users to control and interact with their console using natural gestures. In this Section we talk about two different use cases of the Kinect sensor in elderly care.

2.1 Unobtrusive Screening of Movement and Postural Stability

Falls in the elderly population is a hazard that can result in morbidity, mortality and other injuries. The elderly population is fast growing worldwide and according to a world health organization report [2] 28–35% of people aged over 65 years fall approximately 2 to 4 times per year. This rate increases to 5 to 7 times per year for population aged over 70 years. Major risk factors contributing to falls are previous fall history, decreased strength, gait and balance impairments, and cognitive impairments. WHO's active aging policy offers a framework which aims at developing strategies to prevent falls in the elderly population [2] by reducing some of these factors. The most effective fall prevention strategy demands an increased body balance, increased physical activity as well as better cognitive function. Cognitive impairment is closely associated with body balance. Postural stability is a complex skill that requires a well-co-ordinated functioning of motor and sensory systems which are linked by higher order neurological processes. Cognitive impairments like MCI (mild cognitive impairments), dementia and memory related problems ultimately contributes to postural instability [25,27]. Thus for quantification of balance and fall risk, cognitive functions should also be evaluated. Moreover, early detection of balance and cognitive impairments might be beneficial for planning interventions for older population which in turn would reduce the fall risk.

There are various approaches for assessing fall risk. The most widely used approach is a detailed medical evaluation of balance, gait, mobility and strength done by medical practitioners or geriatricians [37]. For this purpose, they use high end devices like vicon [30], GAITRite [14] or dynamic gait analysis systems by Zebris [28] or Tekscan [48]. This approach identifies the major risk factors that can be treated to prevent fall but it *does not* give any risk score that can be used for assessment. Moreover, these systems are costly and are not suitable for mass deployment. The next widely used approach is to use manual screening instruments and forms such as the Morse fall scale [26], John Hopkins Fall risk assessment tool [31], STRATIFY [39], Berg balance score [42]. All these scales

require the presence of an experienced professional or physiotherapist to perform the scoring and assessment. Further, the accuracy of these scores are dependent on the experience and skill set of the practitioner and hence subject to bias. Selecting the appropriate scale is also a problem faced by practitioners. Also, while these scales perform a detailed functional assessment in terms of physical activities performed by the individual, very few assess the cognitive impairments associated with balance. Recent research has focused on the use of a secondary task paradigm for predicting fall risk [5, 46] but not many have been performed on role of executive functions and memory on postural balance [25].

In order to address these issues, presented by the existing state of art methods in the assessment of postural stability, we design a system using wearable sensors like Kinect and eye trackers. Using the Kinect has not only allowed our solution to scale but alleviated the need for any manual intervention.

Postural Stability Analysis Using the Kinect. Clinically approved techniques such as the assessment of standing balance, performance evaluation in functional reach tasks provide important information pertaining to fall propensity of an individual. However, these techniques do not provide a quantitative measure. A solution would be to perform an analysis of the spatio-temporal dynamics of skeleton joint positions [34] to quantify these parameters. The Microsoft Kinect™ system is one such device that gives us real-time skeletal joint positions in 3D world coordinate systems at 30 frame-per-second (fps). However the $\{x, y, z\}$ coordinates of 25 skeleton joints obtained from the Kinect version 2.0 sensor are very noisy [41]. Thus in order to match clinical standards, it is necessary to first remove the noise prior to any Kinect based analysis [44]. To do this a Kalman filter based approach [44] was developed to track skeleton joint position in 3D space while preserving the distance between two physically connected joints (i.e. bone length to be constant over time). Performance of our algorithm was evaluated for static and dynamic joints. Robustness of our constrained state-estimation algorithm was also tested during both static and dynamic postures. Figure 2 depicts the variation in arm length using the Kinect 2 for shoulder abduction/adduction exercise.

Post noise reduction, the skeleton joint coordinates can be used for postural control analysis. The functional range of motion (ROM) [13] of the shoulder is considered to be one of the key factors to assess postural control. For this, we have used the prepossessed joint-coordinates to derive parameters like ROM angles (in degrees), trajectory of the performed exercise etc. The experiment was carried out within an elderly community. Table 1 clearly indicates the validity of Kinect based ROM angle measurement against clinically approved goniometer based measurements for six ROM exercises.

In order to perform postural stability analysis we have used the single limb stance (SLS) [7] exercise. The proposed solution is capable of evaluating the postural stability using two parameters [6] – (1) SLS-duration for unsupported standing on one leg and (2) Vibration Index (VI) based on corrective body vibrations, which are estimated quantitatively as the relative frequency variation

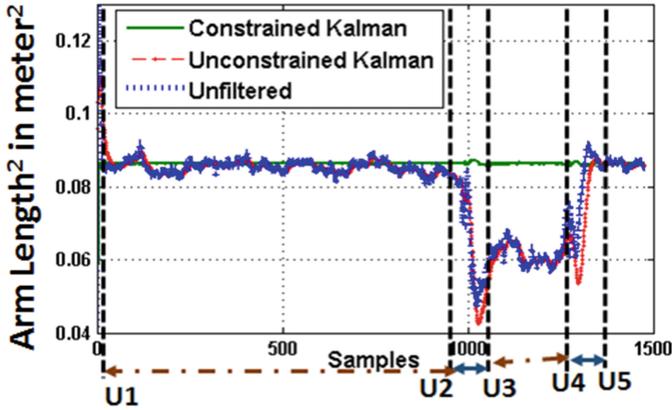


Fig. 2. Variation of arm length for Kinect 2 for shoulder abduction/adduction (U1 to U2 and U3 to U4 denote static posture, U2 to U3 and U4 to U5 represent dynamic posture). Note, the variation in bone length is expected to be zero in an ideal scenario. A smaller standard deviation therefore indicates better noise reduction.

Table 1. Evaluation of Kinect based measurement w.r.t Goniometer based (manual) measurements

Exercise name	Mean (deg)
Right shoulder abduction	6.8
Left shoulder abduction	5.5
Right shoulder flexion	-1.2092
Left shoulder flexion	-6.4187
Right elbow flexion	4.7
Left elbow flexion	5.4

of twenty different joints over time. Finally, the fuzzy rule based fall risk score is calculated from SLS duration, VI and sway area of body Centre of Mass (COM), derived from Kinect skeleton data (post noise removal). Figure 3 shows the basic block diagram for the fall risk score generation.

The computed score is validated with respect to a clinically approved visual analog scale like the Berg Balance Scale [42], John Hopkins Fall Risk Scale [31]. Our population sample included the elderly and stroke patients recruited from a local hospital and community centers. The one way analysis of variance (Anova) between the stroke population and healthy individuals under study are summarized in Table 2. The p and F value for fuzzy scorer prove the comparability of our proposed fall risk measure against Berg Balance scale, reduced Berg Balance scale and Johns Hopkins scale.

In order to increase the patient engagement level during our experiment, the SLS based fall risk assessment method was gamified using an augmented reality

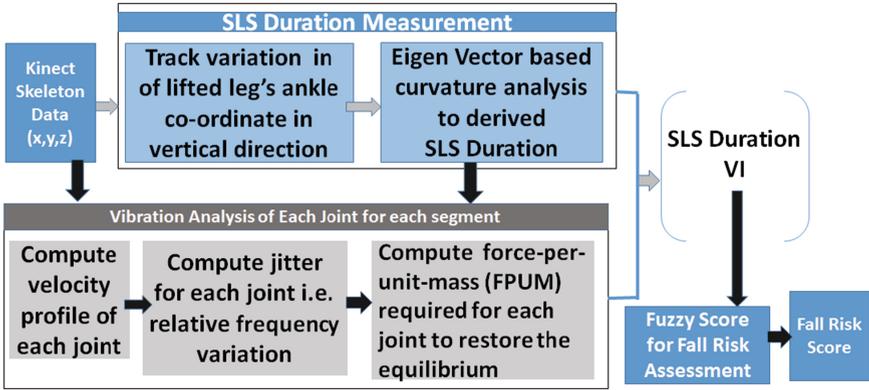


Fig. 3. Blockdiagram of proposed fall risk score generation methodology

Table 2. ANOVA analysis for different methods

Scorer	P-value	F-value
Fuzzy scorer	0.0318	4.95
Reduced Berg Balance scale	0.0329	4.88
Berg Balance scale	0.0088	7.58
John Hopkins scale	0.0246	5.46

paradigm. In the gamified version [36], an individual is asked to lift his leg to a particular height (referred as step height) and subtle changes in postural stability is assessed with increase in step heights. Figure 4 shows the how the proposed posturography feature-values varies with different heights.

Cognitive Assessment. Poor postural stability is often associated with cognitive deficits. People suffering from MCI might face difficulties in maintaining balance and normal gait. Research suggests that fall propensity in the elderly having cognitive impairments is almost double than that of those not having any cognitive dysfunction [10]. Moreover, executive functions like visual perception, cognitive speed etc. are also associated with motor co-ordination. In order to assess these executive functions, we have selected a standard psychological test called the Digit Symbol Substitution Test (DSST) [24] and modified this to assess various impairments. Standard pen and paper digit symbol substitution test (pDSST) requires the association of digits with symbols with reference to a lookup table. The task performance is governed by the number of correctly associated symbols in a predefined time. Thus it can only detect if the subject being tested is able to match a digit with the corresponding symbol but it *cannot* explain the reason behind the failure. In order to overcome this limitation, we have attempted to use a low cost eye tracker that monitors the individual during

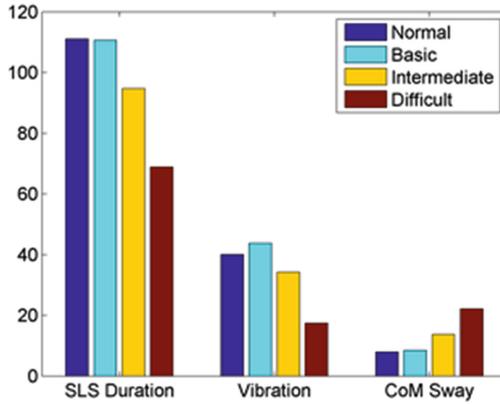


Fig. 4. Posturography features and comparison of stability scores (different colors signifies different step heights) (Color figure online)

the test. The disorders affecting the cerebral cortex, brain stem or cerebellum have a strong influence on eye movements. Further, the visual scanning behavior and retention differs in individuals having MCI [9]. Hence, the gaze behavior was used exclusively in our study to detect the age related differences among the participants. The proposed layout is shown in Fig. 5. This layout is shown in a computer screen placed at a distance of 60 cm from the participant. An eye tracker is placed below the screen so that we can record the eye ball coordinates while taking the test. We have derived various features from the eye movement data collected. We have also designed few versions of this so that we can capture age related differences, problems of visual neglect, attention deficits of the participant under test.

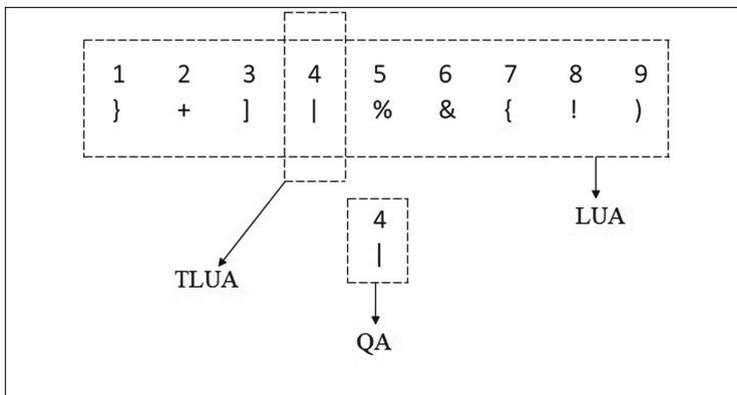


Fig. 5. DSST stimulus

2.2 Tele-Rehabilitation Using the Kinect

For senior citizens, ailing from stroke or fractures, performing scientifically correct and regular exercise is very important. This accelerates recovery and enables them to maintain good health. These exercises are best performed in presence of a physiotherapist. However, visiting a physiotherapist each time, is not a viable option.

We therefore designed a Tele-Rehabilitation application that allows senior citizens to perform exercises at the comfort of their own home in a scientifically correct way. The application gives a score post exercise based on their performance. The application (envisaged as a framework to cater for multiple such requirements in different areas) provides a digitized method for human motion detection and tracking body joints. Using machine learning, this data can be used for multiple purposes ranging from remote physiotherapy to correcting dancing forms in entertainment and immersive gaming.

Using the Kinect or a single/two-point camera sensors, the Tele-rehabilitation framework performs scientific measurement of body/joint movements and compares it with a baseline movement of joints. It ranks the movements and other derived parameters like sway, range of motion, vibration of joints with respect to gold standards in the remote physiotherapy space. It gives results in form of scores to the user or caregivers in a very simple understandable form to act on. It also allows a physiotherapist to visualize progress over a period of time and fine tune next set of exercises.

As described in Sect. 2.1, judging fall risk for senior citizens is very important as it could result in injuries or even death. This framework has been customized to assess fall risks for senior citizens by digitizing the exercise for “Single Limb Stance – SLS”.

The framework has been engineered to cater to a large number of senior citizens. The next section describes the design philosophy for creating a scalable tele-rehabilitation framework with special focus on the geriatric population.

The framework was engineered with the following design philosophies:

- Extensibility
- Efficiency
- Availability
- Security
- Easy Integration
- Usability
- Scalability

These design goals have been addressed in the framework captured in Fig. 6.

The Analyzer Service: The analyzer service performs the analysis of the motion data of multiple human joints as measured by the Kinect. It uses algorithms described in Sect. 2.1 to do this. Each exercise analysis requires a different algorithm, requiring the analyzer service to be extensible to allow the easy integration of new algorithms. The framework has an adaption layer consisting of

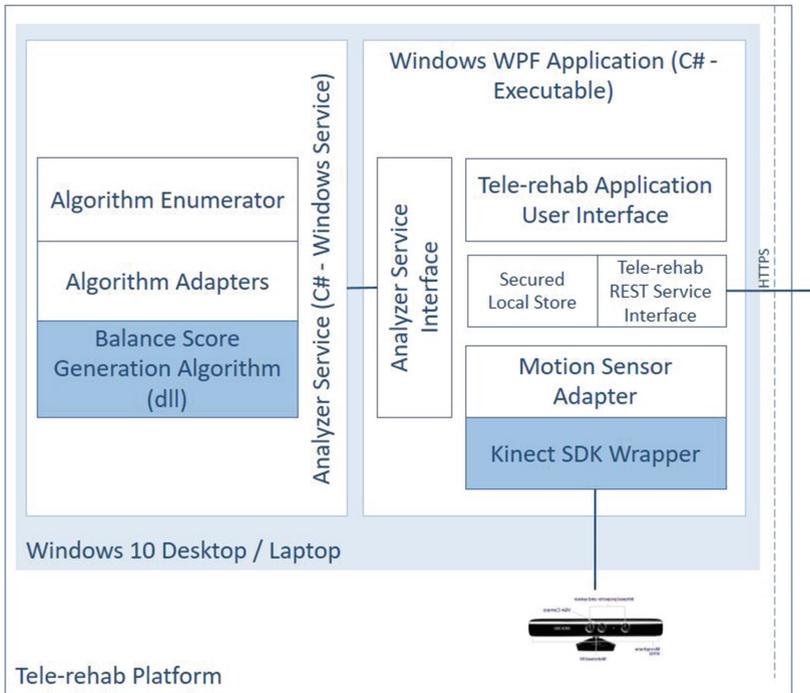


Fig. 6. Tele-rehabilitation application framework.

one or more adaptors with one adaptor per algorithm. This adaption layer offers a simple and common interface around the algorithms, and thus making the framework extensible for newer exercise/movement integration.

These algorithms use the skeletal joint information provided by the Kinect, which comprises of joint name, joint position (x, y, z), and joint tracking information. The data is produced as a sequence of frames, with each frame consisting of 25 joints, typically at a rate of 30 fps. The analysis done by this algorithm is CPU intensive and hence the architecture design encourages this to be done locally on the client side, to distribute computation load and also avoid network delay/downtimes. Thus by distributing and offloading processing effort on to user's machine the design offers efficient performance, and better user experience by allowing offline operations as well.

These algorithms also need to be highly available (zero load time). In order to achieve this, the algorithms have been incorporated as a separate service. Having a separate service offers high availability by allowing the algorithms to be preloaded and cached even before the user starts the application.

In our current implementation the algorithm is used for SLS analysis. The algorithm offers a single Stability Score which in our setup is assessed by a care

giver or a therapist to make necessary interventions with the patient and suggest the respective therapy.

Desktop Application: The Tele-rehabilitation framework also includes a desktop application as part of the smart home setup shown in Fig. 1. This application has several sub-components:

1. A secure local store for storing patient information and raw movement data in encrypted format, addressing the security goal of the framework;
2. Motion sensor adaptor is another important sub-component that offers extensibility for varied physical sensors, Kinect in current POC and single/two-point camera.
3. The user interface is also one of the sub-components of the desktop application, which is used directly by the patient or the care giver.

The User Interface: The complex and robust motion analyzer service is hidden behind a relative simple user interface to address the usability goal of the framework. As the POC is targeted for the geriatric population, a usability study had been conducted on a small group of senior people. Using the insights from this study the user interface was developed with larger and legible fonts, and specific choice of colors, thus making it easier for the seniors to read. The application also interacts using textual/voice prompts to assist and guide patients and care givers throughout the exercise capture session. Apart from this, the patients can interact with the application using simple hand gestures.

Currently the entire framework is being deployed and will be piloted for 6 months. Results given by the framework will be validated against gold standards. It is envisaged that apart from SLS, it also would be used for other fun activities like dancing and immersive gaming using AR/VR.

3 Continuous Activity Recognition Using Smart Watches

Smart wearable devices provide an opportunity for non-intrusive health and activity monitoring. These devices, such as Smart-Watches and phones, owing to their form factor, enable continuous monitoring of its users and assist in improving their lifestyle through active alerts and notifications [3, 15, 23, 29, 32, 45]. While these device are sensor rich, there is a limitation to their processing capabilities resulting in an accuracy-latency tradeoff. In addition there are several other challenges that arise in monitoring activities of daily living:

- Continuous multi-activity recognition.
- Diversity in sensor resolutions.
- Noisy data.
- Confounding gestures.
- Privacy and security requirements owing to the nature of data collected.

Current research in activity detection using wearable devices either focus on the detection of a single activity or techniques to improve the classification accuracy of an activity [11, 12, 20, 47]. However, a more practical deployment, such as assisted daily living, requires a multitude of activities and gestures to be detected on a single device or a combination of devices. A key challenge in multi-activity detection is that each activity uses its own combination of sensors (and perhaps sampling frequency) to detect the activity. A naive approach would be to turn on all sensors at the highest sampling frequency. However, this would result in the inability to continuously monitor the patient throughout the day owing to a battery drain [33].

The goal of this work is to build a system that is not only capable of continuously monitoring a patient’s daily activities, unobtrusively, but does so in a power efficient manner. This would allow us to provide patient wellness support as well as alert caregivers (and the patient) on certain pre-defined events in **real time**. In order to perform this continuous activity recognition we propose a context-based multi-activity classification algorithm. A key element of the algorithm is managing the sensors to not only enable continuous multi-activity but doing so in a power efficient manner. In the following section we provide an overview of the system and the classification algorithm for continuous monitoring of the activities of daily living.

3.1 Context-Based Multi-activity Classification Algorithm

As the list of daily activities are diverse a key challenge is being able to accurately identify these activities within the sensor data stream. To support this task, additional infrastructure-sensed context information is gathered, patient location for example. Further, based on patient criticality certain event-based-notifications are delay intolerant. This requires us to build a “smart-sensing” system that distributes the data processing to tackle accuracy-latency tradeoffs. The pseudo-code for the context based continuous multi-activity recognition is shown in Fig. 7.

The set of activities to be detected as well as the classification levels required for each activity detection is customizable. Note, in most cases we use existing state-of-the-art algorithms for detecting a particular activity. Our focus is instead on enabling continuous multi-activity recognition using these existing activity recognition algorithms. Currently, the activities of interest includes eating, drinking, walking, being stationary and sleep detection.

We explain the pseudo-code with an example flow. A sampling window of size T_s of sensor data goes through the first level of classification. The first level of classification checks to see if any activity of interest is detected within the sensor stream. At this stage the sensor stream contains only accelerometer data with all other sensors turned off to conserve battery power. Figure 8 captures the raw accelerometer data along with the classifier output. Clearly, when there is an activity of interest the variation in acceleration magnitude is significant enough to indicate that an AOI *might* be in progress. When an activity of interest is detected, the next check is to determine if any additional sensors (or changes in

Algorithm Context-based multi-activity classification

1. Collect accelerometer sensor data of window size T_s with sampling frequency F_s from device i .
2. Retrieve the activity being A_i being detected and the current level of classification L_i for device i .
3. For the given window T_s perform classification L_i .
4. If no activity of interest (AOI)
 - a. Increment no-AOI counter.
 - b. If no-AOI counter = M reset classification level $L_i=0$ for device i . Sample next window for device i .
 - Else
 - c. Sample next window for device i .
- Else if activity of interest detected
 - d. Record activity A_i . Update classification level L_j if required.
 - e. For activity A_i and classification level L_j check if additional sensors need to be turned on or change in sampling frequency of current operational sensors is required.
 - f. If additional changes are required
 - i. Set changes in the device profile through the sensor management module.
 - g. Sample next window

Fig. 7. Pseudo-code for continuous multi-activity classification.

sampling frequency) are required for the subsequent classification levels for *that* gesture. For example, a person undergoing physical rehab might not be allowed to walk continuously for long periods. Therefore if the ‘walking’ activity is detected the classification algorithm may require increasing the sampling frequency, for better accuracy. However the extent of increase may be limited based on the current power level of the device. We also consider incorporating additional context information such as location and previous activity recorded into decision making and sensor management, which includes deciding when to turn on/off sensors as well as the sampling frequency rate. Figure 9 captures the benefit of smart sensor management. The plot captures a smart watch battery drain over a period of 30 min when all sensors (accelerometer, barometer, gyroscope, photoplethysmogram) are turned on as compared to a scenario where only the accelerometer sensor is enabled for the entire duration and the gyroscope is turned on for only 10 min of the 30 min interval. This scenario emulates the eating classification process where the gyroscope sensor is turned on *only* when suggested by the algorithm.

Once the changes, if any, are made, the system makes a note that subsequent data frames from this device will include the requested changes and also sets the level of classification device that is needed for those subsequent data frames. At any point if the system detects “no activity of interest” for a set number of frames it resets the classification level to zero (for that device) indicating that the activity of interest, example ‘walking’, has ended (i.e. the participant is no longer walking).

An additional power saving method is also employed by distributing the classification operation between the wearable device and the server in order to optimize the computation and communication costs of the wearable device. We describe the decision making process of this method in the next section.

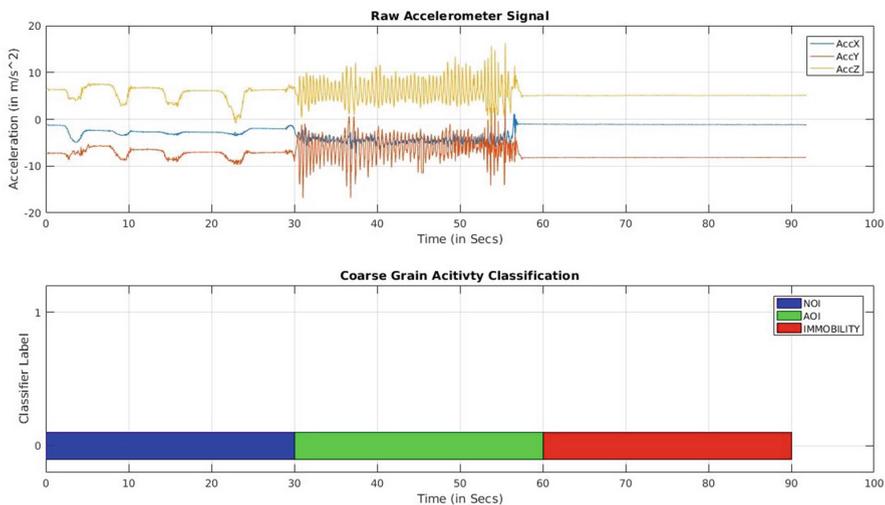


Fig. 8. Raw accelerometer data and the corresponding classifier output. Sampling frequency = 100 Hz.

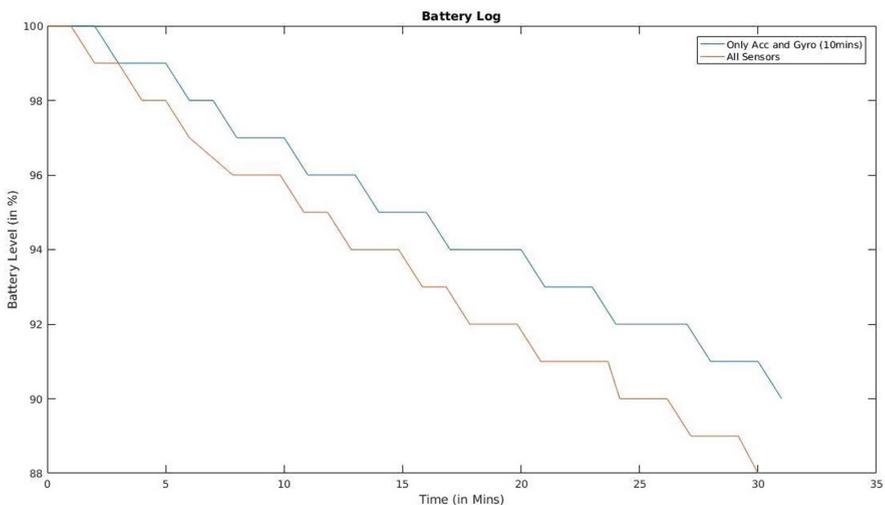


Fig. 9. Rate of battery drain on the smart watch when all sensors are turned on Vs ‘Smart Sensing’ where sensors are turned on/off based on the AOI being detected. Sampling frequency = 100 Hz.

3.2 System Architecture

Figure 10 shows the system architecture and the various modules needed to support a power-aware context-based continuous activity recognition. The functionalities of each module are described below.

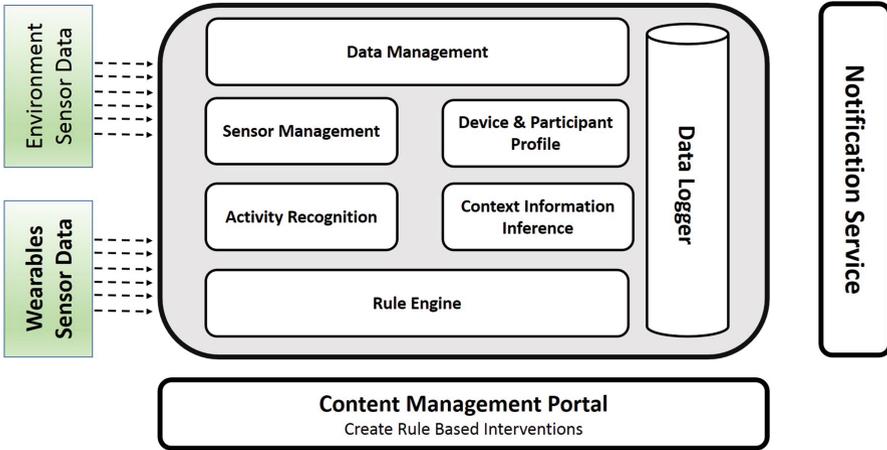


Fig. 10. System architecture.

1. Inference Engine: The inference engine includes multiple modules:
 - **Gesture Recognition:** This module identifies the various activities and gestures within the sensor data stream. A similar module resides on the device as well. It takes as input additional context information such the patient criticality and other environment variables in its decision making. The gesture recognition module is responsible for setting the classification mode. Currently we perceive the need to support 3 classification modes to manage power, classification accuracy and latency requirements. Note: Low Power, High Accuracy and Low Latency is desirable.
 - Mode 1 (High Power, High Accuracy, Low Latency): In this mode the classification takes place completely on the device.
 - Mode 2 (Low-High Power, High Accuracy, High-Low Latency): In this mode the first level of inference takes place on the server side. Based on the inference, subsequent classification takes place on the device. For example, if the server detects an eating activity, it informs the device. The devices then performs the classification for subsequent data frames till the activity has completed.
 - Mode 3 (Low-High Power, High Accuracy, High Latency): In this mode the classification takes place completely on the server.
 - **Data Management:** Data from the Wearable devices could either be raw sensor data or context information depending on where the classification needs to be done. The data management module would be responsible for managing the flow of the data, depending on the classification mode set for that device, as well as handling the logging of data.
 - **Sensor Management:** Depending on the current classification mode and gesture certain sensor will need to be turned on/off and or sampling frequency of the sensors need to be changed. This module is responsible

for managing the sensors on the wearable devices as well the environment/infrastructure sensors.

- **Device & Participant Profile:** This module keeps track of the device profile, which includes information such as current power levels. The module also manages participant profiles which feeds as input in identifying the current classification mode of the device.
 - **Context Information Inference:** This module derives higher levels of context information based on low level sensor based context.
 - **Rule Engine:** This module allows you to create rules to set the classification mode. In this module you can specify how the different inputs are combined to create a rule. Note rules are different for each device and each device can have multiple rules. An example of a rule would be “If eating gesture is detected in level 0 classification, turn on the gyroscope sensor for level 1 classification”
2. **Data Logger:** is typically a persistent storage of long term and transient information of the subject. The database is indexed based on different contexts such as location, time of day, subject id, specific activity, anomaly etc. This can be either accessed directly or using some remote access mechanism like web-services. The service orchestration shall be limited to the context layer only, as of now. If new data needs to be exposed, they will be added to the database and query/response objects shall be implemented to keep the modules loosely coupled.
 3. **Infrastructure/Environment Sensors and devices:** Additional sensors placed in the infrastructure/environment that provides additional decision-making context information to the inference engine. Examples include, but not limited to, passive infrared sensors that capture sensor proximity, wi-fi access points that provide location information.
 4. **Notification Service:** This module handles the sending of notifications to the wearable device through the appropriate channels.
 5. **Content Management Portal:** The portal allows setting content to be sent as notifications as well as setting rules when the notification needs to be sent.

3.3 Current Status

We are currently in the midst of implementing our algorithm for continuous multi-activity detection. Figure 11 shows a *WIP* sample output of our algorithm as it goes through the various levels of classification. For example, between time $t = 120$ s and $t = 160$ s, we can see the algorithm transition from detecting an AOI to detecting an activity involving the lower limb to finally detecting a walking activity.

At this stage we plan to operate in Mode 1 (Sect. 3.2) where the classification takes place completely on the smart watch. We use the eating algorithm designed by Sen et al. [40] and an in-house algorithm to detect activities such as walking and being stationary [8]. Our goal is to incorporate these activity detection algorithms into ours whilst finding a sweet spot between classification accuracy and power consumed.

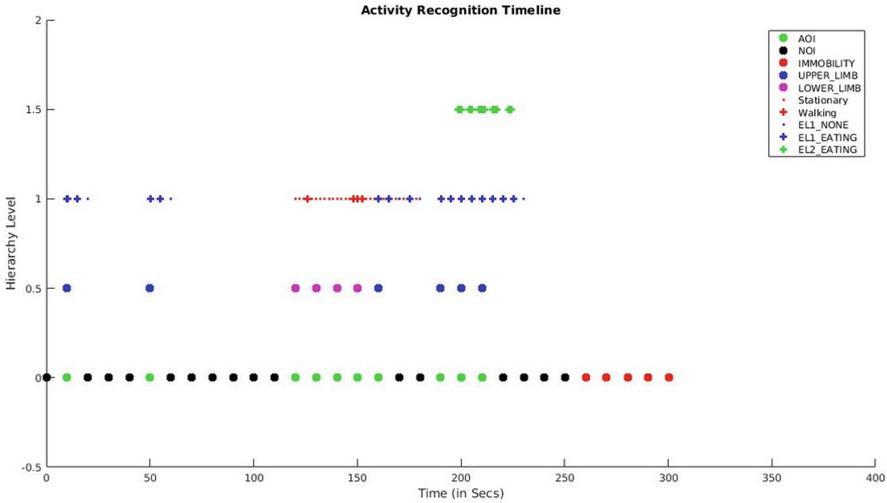


Fig. 11. A sample output of our continuous multi-activity detection algorithm. The plot shows the classification level as it progresses in detecting an AOI. We aim to provide a continuous output stream of activities being performed by the individual.

4 Watching over You: Video and Sensor Based Analytics

In this Section we explore the possibility of using video in gaining better insights into the daily lives of the elderly. While using video can be considered invasive, our framework uses it more as an activity-validation technique as opposed to continuous monitoring. For example, using video can provide a better accuracy in confirming whether the patient took their medication as opposed to other techniques such as a smart-medicine box [43]. Our approach captured in Fig. 12 uses a combination of video processing and image analysis.

The first step in our approach involves creating a learning set of images for specific activities of interest (AOI), watching television for example. These images are extracted from a video of the activity. Once the AOI training image set is finalized it can be used to detect the activities in real-time.

We do this by first capturing frames from the live video feed, converting these video frames to images and then measuring the degree to which these images match against our AOI training image set. Note, the frames are first converted to gray scale to ensure any color related variations are removed. We use the ORB matching algorithm [38] which is rotation invariant and resistant to noise.

The ORB matching algorithm uses a combination of the FAST (Features from Accelerated Segment Test) keypoint detector and BRIEF (Binary Robust Independent Elementary Features) descriptor to improve the performance. The FAST algorithm [19,35] is a corner detection method used to track and map objects in computer vision activities. A key advantage of the ORB approach is in it's ability to detects variations in images particularly in rotated images.

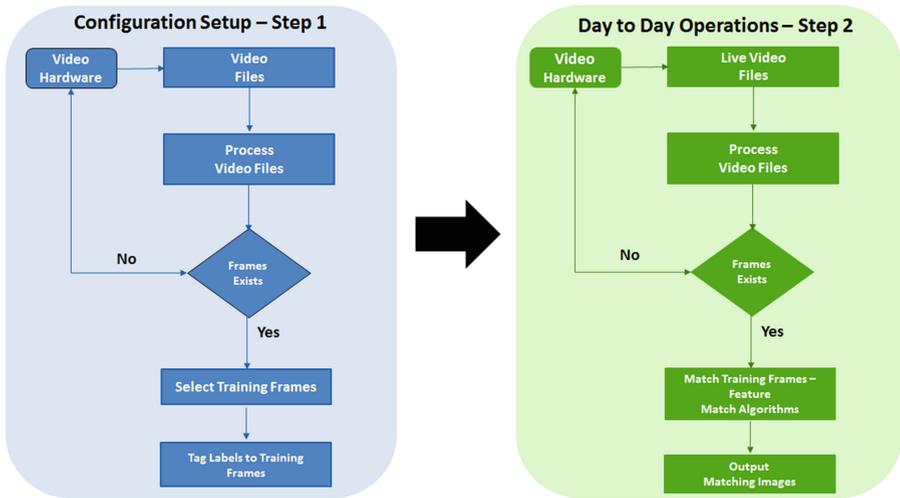


Fig. 12. Flow diagram of our video analysis approach.

Once the keypoint and descriptors of the image pair are computed, the distance between these measures are calculated. This distance is then matched against a threshold. Once a AOI image match is found, we record the activity and to ensure privacy all images stored are obfuscated.

4.1 Determining Location Stay Patterns

Observing the number of hours a patient spends in any room is a useful insight to have. Monitoring this could throw up an alarm in case the patient has spent a longer-than-usual time in any room. Figure 13 captures the number of hours a patient spends in a location on a weekly basis.

Determining this location pattern also helps in predicting location transitions. Figure 14 shows the accuracy of our prediction. As before any observed outliers can be alerted to the caregivers and doctors for appropriate action.

4.2 Vision Based Human Liveness Detection

The key indicators of life is the presence of heart beats and respiration. Every heart beat introduces subtle motions in the human body, even when the person attempts to maintain a state of rest [16]. While these subtle movements are latent to the human eye, they have been observed using a camera for heart rate estimation [17]. In this section, we propose a computer vision based method

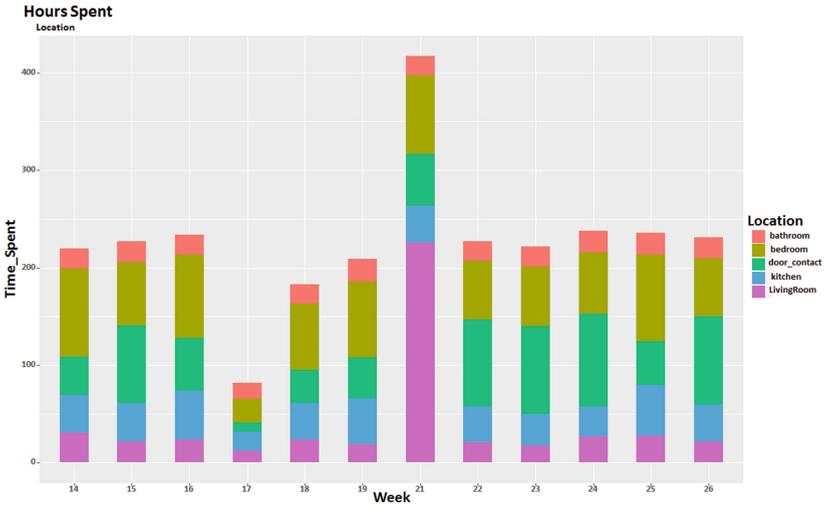


Fig. 13. The figure captures the number of hours a patient spends in a location on a weekly basis. The Location includes Living Room, Bedroom, Bathroom, Kitchen and Door Contact count

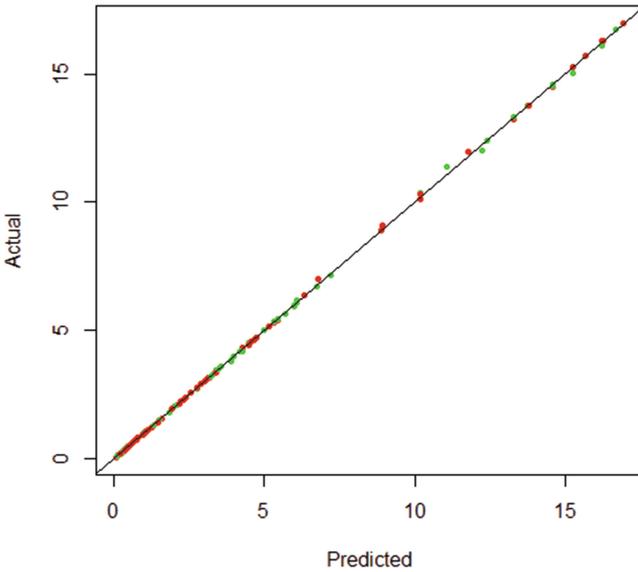


Fig. 14. The graph captures the accuracy of the predicted and actual hours that a patient spends on locations including the Living Room, Bedroom, Bathroom & Kitchen

to leverage these subtle body motions for detecting liveness in human being during sleep. It consists of the following three stages:

1. estimating temporal variations;
2. evaluating liveness index; and
3. liveness detection.

In the first stage, we divide the image into non-overlapping square blocks and evaluate the temporal variation in each block. It is followed by estimating the liveness index of each block which provides the confidence about whether the block belongs to human body or not. Eventually, we utilize the liveness index of blocks for liveness detection.

Estimating Temporal Variations. The video frame is first divided into several non-overlapping square blocks. As seen in [21], the green channel contains the strongest plethysmographic signals amongst all other color channels. The reason behind this is that green light provides a better absorption of hemoglobin when compared to red light and also offers a better penetration of human skin than blue light.

We therefore utilize only the green channel for estimating temporal variations. These temporal variations can be evaluated by either tracking the distinctive features, i.e., Lagrangian approach or analyzing the intensity differences at a fixed location, i.e., Eulerian approach. Lagrangian approaches are computationally time expensive and they can be spurious when few distinctive features are available for tracking [4]. Hence, we utilize the approach that avoids time consuming tracking of ROI and works correctly for subtle variations [17]. We define the temporal variation in a block of a video frame by the mean green value of pixels inside it.

That is, the temporal variations in a block i , T^i is given by:

$$T^i = [t_1^i, t_2^i, \dots, t_f^i] \quad (1)$$

where t_k^i is the mean green value of pixels in i^{th} block of k^{th} frame and f denotes the total number of frames.

An example is shown in Fig. 15(a). It depicts a video frame along with the blocks belonging to the object and a human body in the left. The temporal variations of the blocks are shown in the right.

Evaluating Liveness Index. The temporal variations are produced due to respiration rate, heart beats and environmental conditions (like illumination variations and focus change). An object devoid of life does not contain temporal variations when kept in a state of rest.

In contrast, respiration and heart beat produce different amounts of temporal variations in different parts of the body according to the structure of arteries and bones. At a given instance, variations are more prominent in the neck and stomach due to respiration. Thus, we identify some of the blocks belonging to

human body with the fact that they contain significant temporal variations due to respiration rate and heart rate as compared to the blocks belonging to objects.

It is observed that temporal variations due to environmental conditions can result in erroneous identification of human body blocks. We mitigate this issue by dividing the temporal variations in a block into multiple segments and pruning the segments which possibly are affected by variations due to environmental conditions.

Mathematically, the k^{th} segment for the i^{th} block is given by:

$$S_k^i = [T^i((k-1)m+1), T^i((k-1)m+2), \dots, T^i(km)] \quad (2)$$

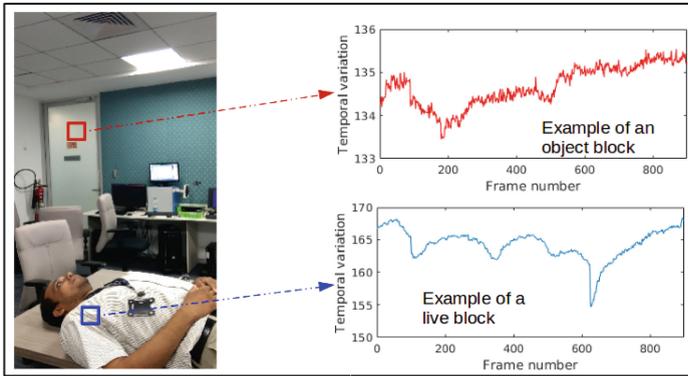
where T^i stores the temporal variations for i^{th} block and m is the size of segment.

Subsequently, we prune a total 30% of the total segments based on large standard deviations. The pruning is based on the intuition that temporal variations due to environmental conditions are highly prominent than the variations due to respiration rate and heart rate. The standard deviations of the remaining segments are added to define the liveness index of a block which provides the confidence whether the block belongs to human body or not. The higher liveness index of a block indicates better chances that the block belongs to the human body. An example of liveness index is depicted in Fig. 15(b).

Liveness Detection. Liveness index is high for the video blocks that belong to a human, but it is possible that the video does not contain any human body part. Thus, a total 10% of the blocks containing high liveness index are selected and they are utilized for liveness detection.

The temporal variations in the PPG signal are large if the selected blocks belong to a human because the human body contains a heart beat. In contrast, these are small if the selected blocks belong to stationary objects. The PPG signal is estimated by applying the algorithm proposed in [17] on the temporal signals of the selected blocks. The temporal variation in the PPG signal can be obtained by subtracting the maximum and minimum value of the PPG signal.

It is also observed that such an approach can be affected by small environmental noise like focus change. Thus, we obtain the temporal variation in the PPG signal by dividing the PPG signal into overlapping time windows, evaluating the local range of each window and adding the local range estimates. Local range of a 1-D array (which is the PPG signal in our case) is evaluated by subtracting the maximum and minimum values of an array [18]. As temporal variations in the PPG signal are higher if the selected blocks belonging to a human rather than a stationary object, we indicate that the video contains a live object when the sum of the local range estimates exceeds the predefined threshold, T , otherwise we indicate that the video contains only non-live objects.



(a)



(b)

Fig. 15. Example of vision based liveness detection: (a) Video frame along with the temporal variations of blocks containing object (in red color) and human body (in blue color); and (b) Liveness index (Color figure online)

5 Conclusion: Putting All the Pieces Together

In the previous Sections, we described wellness measurements that can be derived from a variety of sensors. The next step is to integrate these different (and currently independent) modules to create an overall wellness care framework.

Our generic extensible wellness framework will allow the integration of various sensors and also algorithms. The algorithms can be run on the cloud, on the edge or on the sensor. The framework then fuses the data into an inter-operable data format.

Our open and extensible framework will abstract data coming from variety of sensors and devices, enrich it with context and domain information and

stores it in a standardized data structure. Application Programming Interfaces (API) exposed on raw, contextualized and domain enriched data, will help realize actionable insights in real-time/near real-time. Additionally, this framework will support plugging in of algorithms which will enable us to learn useful patterns and insights from the data.

The Wellness Care Framework follows a pluggable open architecture shown in Fig. 16. The following are the key components of the framework.

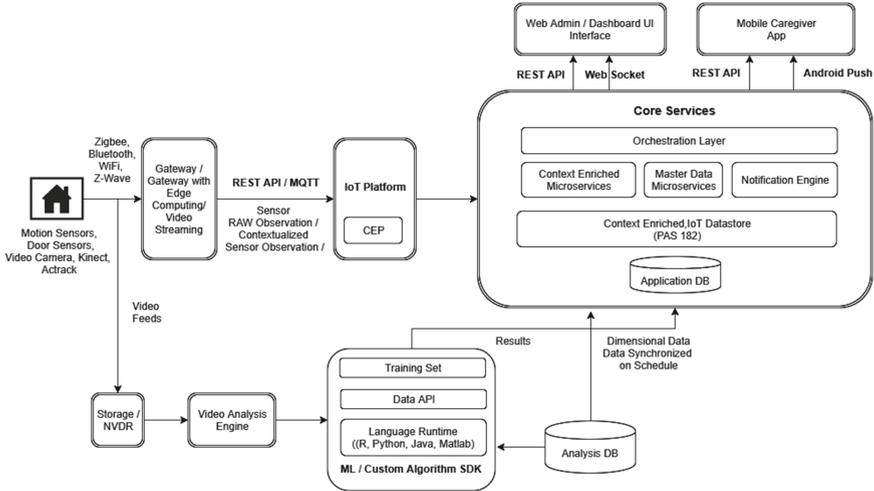


Fig. 16. High level system architecture of our wellness care framework incorporating sensors, video camera, Wearables, IOT platform, Core Services and our algorithm SDK.

The multitude of sensors and devices placed in these smart-homes perform observations such as motion, posture, door open/close detection and video footage.

The framework includes a gateway device which uses efficient low-range low power protocols to receive the data from the sensors. The gateway, simply put, forwards the sensor data to an IoT layer. In our example, the wearable has a custom application installed in the device which acts a gateway and pushes data to other layers within the framework.

Some gateways have processing capability to perform edge computation and may send enriched data. In our example of the Kinect setup, both the desktop (attached to the Kinect) and the custom application act as a gateway. This enriches the observation which is sent to the IoT Platform. In the example of video cameras, the video recordings are stored in a recorder device that is accessed through an internal network.

The IoT Platform facilitates the interaction between the sensor system [sensor + gateway] and the core services. It receives/pulls data from various sensors/gateways across smart-homes, contextualizes it and sends data in a standardized format. The IoT platform also has an in-built CEP component which

will continue monitoring the data through a set of rules applied on the data. In case a threshold is reached or an alert rule is satisfied, the CEP engine fires an action for the corresponding event.

The Core Services receives the data from the IoT Platform and first enriches and stores the data. We have followed the BSI PAS-182 [1] standards to create an interoperable data model. The data hub is built on a micro services architecture and offers data micro services, master data services, and notification services. The microservices expose REST APIs for the upstream services to consume and render - Web UI and a mobile app. Notifications from the notification microservice are pushed to Android-based mobile devices using a cloud notification service and using websockets for the browser based UI.

Our framework includes a full-fledged run-time algorithm on which custom algorithms can be deployed, executed and the insights from these algorithms can be made available to the rest of the wellness framework. The supported languages are R, Python, MATLAB, and Java. The run-time also provides APIs for accessing the training sets and validation sets required by machine learning algorithms.

A key challenge we face in building such a generic wellness care framework would be in handling the trade-off between cost of data acquisition, the level of intrusiveness, and precision in detecting the activities.

References

1. Smart city concept model - guide to establishing a model for data interoperability (2014). <https://www.bsigroup.com/en-IN/smart-cities/Smart-Cities-Standards-and-Publication/PAS-182-smart-cities-data-concept-model/>. Accessed 2 Feb 2018
2. Ageing, W.H.O., Unit, L.C.: WHO global report on falls prevention in older age. World Health Organization (2008)
3. Aharony, N., Pan, W., Ip, C., Khayal, I., Pentland, A.: Social fMRI: investigating and shaping social mechanisms in the real world. *Pervasive Mob. Comput.* **7**(6), 643–659 (2011). <https://doi.org/10.1016/j.pmcj.2011.09.004>
4. Balakrishnan, G., Durand, F., Gutttag, J.: Detecting pulse from head motions in video. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3430–3437 (2013)
5. Biderman, A., Cwikel, J., Fried, A., Galinsky, D.: Depression and falls among community dwelling elderly people: a search for common risk factors. *J. Epidemiol. Commun. Health* **56**(8), 631–636 (2002)
6. Chakravarty, K., Sinha, A., Sinha, S., Chatterjee, D., Das, A.: Single leg stance (SLS) and vibration index (VI): new instrumental indices for fall risk estimation in stroke survivors. *Gait Posture* **49**, 168 (2016)
7. Chakravarty, K., Suman, S., Bhowmick, B., Sinha, A., Das, A.: Quantification of balance in single limb stance using kinect. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 854–858. IEEE (2016)
8. Chandel, V., Choudhury, A.D., Ghose, A., Bhaumik, C.: AcTrak - unobtrusive activity detection and step counting using smartphones. In: Stojmenovic, I., Cheng, Z., Guo, S. (eds.) *MindCare 2014*. LNICST, vol. 131, pp. 447–459. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-11569-6_35

9. Crutcher, M.D., Calhoun-Haney, R., Manzanares, C.M., Lah, J.J., Levey, A.I., Zola, S.M.: Eye tracking during a visual paired comparison task as a predictor of early dementia. *Am. J. Alzheimer's Dis. Other Dement.* **24**(3), 258–266 (2009)
10. Eriksson, S., Gustafson, Y., Lundin-Olsson, L.: Risk factors for falls in people with and without a diagnose of dementia living in residential care facilities: a prospective study. *Arch. Gerontol. Geriatr.* **46**(3), 293–306 (2008)
11. Falaki, H., Mahajan, R., Kandula, S., Lymberopoulos, D., Govindan, R., Estrin, D.: Diversity in smartphone usage. In: *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*. In: *MobiSys 2010*. ACM (2010). <http://doi.acm.org/10.1145/1814433.1814453>
12. Froehlich, J., Chen, M.Y., Consolvo, S., Harrison, B., Landay, J.A.: MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones. In: *Proceedings of the 5th International Conference on Mobile Systems, Applications and Services, MobiSys 2007*. ACM (2007). <http://doi.acm.org/10.1145/1247660.1247670>
13. Gerety, M.B., Mulrow, C.D., Tuley, M.R., Hazuda, H.P., Lichtenstein, M.J., Bohannon, R., Kantan, D.N., O'Neil, M.B., Gorton, A.: Development and validation of a physical performance instrument for the functionally impaired elderly: the physical disability index (PDI). *J. Gerontol.* **48**(2), M33–M38 (1993)
14. Givon, U., Zeilig, G., Achiron, A.: Gait analysis in multiple sclerosis: characterization of temporal-spatial parameters using gaitrite functional ambulation system. *Gait Posture* **29**(1), 138–142 (2009)
15. Goonawardene, N., Toh, X.P., Tan, H.-P.: Sensor-driven detection of social isolation in community-dwelling elderly. In: Zhou, J., Salvendy, G. (eds.) *ITAP 2017*. LNCS, vol. 10298, pp. 378–392. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58536-9_30
16. Gupta, P., Bhowmick, B., Pal, A.: Accurate heart-rate estimation from face videos using quality-based fusion. In: *IEEE International Conference on Image Processing*. IEEE (2017)
17. Gupta, P., Bhowmick, B., Pal, A.: Serial fusion of Eulerian and Lagrangian approaches for accurate heart-rate estimation using face videos. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2834–2837. IEEE (2017)
18. Gupta, P., Gupta, P.: An efficient slap fingerprint segmentation and hand classification algorithm. *Neurocomputing* **142**, 464–477 (2014)
19. Harris, C., Stephens, M.: A combined corner and edge detector. In: *Alvey Vision Conference*, vol. 15, p. 10–5244. Citeseer (1988)
20. Lee, Y., Ju, Y., Min, C., Kang, S., Hwang, I., Song, J.: CoMon: cooperative ambience monitoring platform with continuity and benefit awareness. In: *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services, MobiSys 2012*. ACM (2012). <http://doi.acm.org/10.1145/2307636.2307641>
21. Li, X., Chen, J., Zhao, G., Pietikainen, M.: Remote heart rate measurement from face videos under realistic situations. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4264–4271 (2014)
22. Liming, B., Gavino, A.I., Lee, P., Jungyoon, K., Na, L., Pi, T.H.P., Xian, T.H., Buay, T.L., Xiaoping, T., Valera, A., Jia, E.Y., Wu, A., Fox, M.S.: SHINESeniors: personalized services for active ageing-in-place. In: *2015 IEEE First International Smart Cities Conference (ISC2)*, pp. 1–2, October 2015
23. Liu, N., Purao, S., Tan, H.P.: Value-inspired service design in elderly home-monitoring systems. In: *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pp. 1–6, March 2016

24. McLeod, D.R., Griffiths, R.R., Bigelow, G.E., Yingling, J.: An automated version of the digit symbol substitution test (DSST). *Behav. Res. Methods* **14**(5), 463–466 (1982)
25. Mirelman, A., Herman, T., Brozgol, M., Dorfman, M., Sprecher, E., Schweiger, A., Giladi, N., Hausdorff, J.M.: Executive function and falls in older adults: new findings from a five-year prospective study link fall risk to cognition. *PLoS One* **7**(6), e40297 (2012)
26. Morse, J.M., Morse, R.M., Tylko, S.J.: Development of a scale to identify the fall-prone patient. *Can. J. Aging/La Revue canadienne du vieillissement* **8**(4), 366–377 (1989)
27. Muir, S.W., Gopaul, K., Montero Odasso, M.M.: The role of cognitive impairment in fall risk among older adults: a systematic review and meta-analysis. *Age Ageing* **41**(3), 299–308 (2012)
28. Muro-De-La-Herran, A., Garcia-Zapirain, B., Mendez-Zorrilla, A.: Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications. *Sensors* **14**(2), 3362–3394 (2014)
29. Nandugudi, A., Maiti, A., Ki, T., Bulut, F., Demirbas, M., Kosar, T., Qiao, C., Ko, S.Y., Challen, G.: PhoneLab: a large programmable smartphone testbed. In: *Proceedings of First International Workshop on Sensing and Big Data Mining, SENSEMINE 2013*, pp. 4:1–4:6. ACM, New York (2013). <http://doi.acm.org/10.1145/2536714.2536718>
30. Pfister, A., West, A.M., Bronner, S., Noah, J.A.: Comparative abilities of microsoft kinect and vicon 3d motion capture for gait analysis. *J. Med. Eng. Technol.* **38**(5), 274–280 (2014)
31. Poe, S.S., Cvach, M., Dawson, P.B., Straus, H., Hill, E.E.: The Johns Hopkins fall risk assessment tool: post implementation evaluation. *J. Nurs. Care Qual.* **22**(4), 293–298 (2007)
32. Rachuri, K.K., Musolesi, M., Mascolo, C., Rentfrow, P.J., Longworth, C., Aucinas, A.: EmotionSense: a mobile phones based adaptive platform for experimental social psychology research. In: *Proceedings of the 12th ACM International Conference on Ubiquitous Computing, Ubicomp 2010*, pp. 281–290. ACM, New York (2010). <http://doi.acm.org/10.1145/1864349.1864393>
33. Raffa, G., Lee, J., Nachman, L., Song, J.: Don't slow me down: bringing energy efficiency to continuous gesture recognition. In: *International Symposium on Wearable Computers (ISWC 2010)*, pp. 1–8, October 2010
34. Ramu Reddy, V., Chakravarty, K., Aniruddha, S.: Person identification in natural static postures using kinect. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) *ECCV 2014*. LNCS, vol. 8926, pp. 793–808. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16181-5_60
35. Rosten, E., Drummond, T.: Machine learning for high-speed corner detection. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006*. LNCS, vol. 3951, pp. 430–443. Springer, Heidelberg (2006). https://doi.org/10.1007/11744023_34
36. Roy, S., Mazumder, O., Chatterjee, D., Sinha, A.: Quantification of postural balance using augmented reality based environment: a pilot study. In: *IEEE SENSORS. IEEE* (2017, to appear)
37. Rubenstein, L.Z., Josephson, K.R., Osterweil, D.: Falls and fall prevention in the nursing home. *Clin. Geriatr. Med.* **12**(4), 881 (1996)
38. Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: Orb: an efficient alternative to sift or surf. In: *2011 International Conference on Computer Vision*, pp. 2564–2571, November 2011

39. Scale, B.B., Talking, S.W.W.: How to identify potential fallers in a stroke unit: validity indexes of four test methods. *J. Rehabil. Med.* **38**, 186–191 (2006)
40. Sen, S., Subbaraju, V., Misra, A., Balan, R.K., Lee, Y.: Experiences in building a real-world eating recogniser. In: Proceedings of the 4th International on Workshop on Physical Analytics, WPA 2017, pp. 7–12. ACM, New York (2017). <http://doi.acm.org/10.1145/3092305.3092306>
41. Sinha, S., Bhowmick, B., Chakravarty, K., Sinha, A., Das, A.: Accurate upper body rehabilitation system using kinect. In: 2016 IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), pp. 4605–4609. IEEE (2016)
42. Steffen, T.M., Hacker, T.A., Mollinger, L.: Age-and gender-related test performance in community-dwelling elderly people: six-minute walk test, berg balance scale, timed up & go test, and gait speeds. *Phys. Ther.* **82**(2), 128–137 (2002)
43. Toh, X., Tan, H.X., Liang, H., Tan, H.P.: Elderly medication adherence monitoring with the Internet of Things. In: 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops), pp. 1–6, March 2016
44. Tripathy, S.R., Chakravarty, K., Sinha, A., Chatterjee, D., Saha, S.K.: Constrained Kalman filter for improving kinect based measurements. In: 2017 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1–4. IEEE (2017)
45. Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D., Campbell, A.T.: StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp 2014, pp. 3–14. ACM, New York (2014)
46. Bootsma-van der Wiel, A., Gussekloo, J., De Craen, A.J., Van Exel, E., Bloem, B.R., Westendorp, R.G.: Walking and talking as predictors of falls in the general population: the Leiden 85-plus study. *J. Am. Geriatr. Soc.* **51**(10), 1466–1471 (2003)
47. Yan, B., Chen, G.: AppJoy: personalized mobile application discovery. In: Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services, MobiSys 2011. ACM (2011). <http://doi.acm.org/10.1145/1999995.2000007>
48. Zammit, G.V., Menz, H.B., Munteanu, S.E.: Reliability of the TekScan MatScan® system for the measurement of plantar forces and pressures during barefoot level walking in healthy adults. *J. Foot Ankle Res.* **3**(1), 11 (2010)