Behavior Analysis through Multimodal Sensing for Care of Parkinson's and Alzheimer's Patients

Federico Alvarez
Universidad Politécnica de Madrid (UPM)

Mirela Popa
Universiteit Maastricht

Vassilios Solachidis
Centre for Research and Technology, Hellas (CERTH)

Gustavo Hernández-Peñaloza and **Alberto Belmonte-Hernández**
UPM

Stylianos Asteriadis
Universiteit Maastricht

Nicholas Vretos
CERTH

Marcos Quintana
UPM

Thomas Theodoridis
CERTH

Dario Dotti
Universiteit Maastricht

Petros Daras
CERTH

The analysis of multimodal data collected by innovative imaging sensors, Internet of Things devices, and user interactions can provide smart and automatic distant monitoring of Parkinson's and Alzheimer's patients and reveal valuable insights for early detection and/or prevention of events related to their health. This article describes a novel system that involves data capturing and multimodal fusion to extract relevant features, analyze data, and provide useful recommendations. The system gathers signals from diverse sources in health monitoring environments, understands the user behavior and context, and triggers proper actions for improving the patient's quality of life. The system offers a multimodal, multi-patient, versatile approach not present in current developments. It also offers comparable or improved results for detection of abnormal behavior in daily motion. The system was implemented and tested during 10 weeks in real environments involving 18 patients.

14

The United Nations predicts a 45-percent increase in the number of people aged 65 and older in the next 20 years in the developed world. This will likely test the financial sustainability of healthcare systems. Supporting increased autonomy and an independent lifestyle for the elderly, especially those with cognitive impairments, poses a significant challenge.

Information and communication technology (ICT) tools support medical treatments by supplying a range of services for medical personalized prevention, early detection, cognitive training, and recommendations. Continuous evolution in size, processing capability, autonomy, and functionality has permitted a new range of possibilities for ad-hoc and full-time integrated healthcare purposes.[1] These applications have attracted scientific and commercial interest in a sector that has been dominated in the past by expensive and/or tailored solutions.

In this article, a complete ICT-based system to support autonomy for people with cognitive diseases such as Parkinson's and Alzheimer's is presented. This system involves multisensory data capturing and processing based on various sensing technologies, including vision- and Internet of Things (IoT)-based devices. Additionally, advanced multimodal fusion (MF) techniques are applied to medical and behavioral data to extract relevant features for abnormality detection. The results of data analysis provide useful recommendations to stakeholders. This system is intended to gather signals from diverse sources not only in health monitoring (HM) environments, but also in quotidian patient locations to understand the user behavior and context, and then trigger proper actions related to medical or social decision-making. While the system can detect 12 different types of events, this article focuses on two of them: freezing of gait (FoG) and abnormal behavior in daily motion activity.

RELATED WORK

Some previously presented IoT-based approaches, such as Tripoliti et al.,[2] required patients to wear a high number of sensors (such as six accelerometers and two gyroscopes), which they needed to attach one by one. Mazilu et al.[3] was able to predict FoG episodes, but only using electrocardiography (ECG) and skin-conductance (SC), which are more intrusive, tailored, and expensive solutions. Other solutions (such as Rodríguez-Molinero et al.[4]) used only one sensor, but included a dedicated development to detect FoG.

Multiple approaches for monitoring indoor human behavior have been presented (such as Nef et al.[5]) and use different types of cameras, modalities, and system architectures. However, none studies the kind of patients we are focusing on. In these approaches, anomalies are detected by either a sudden change of behavior such as falling down (in Yang, Ren, and Zhang[6]) or statistical analysis over long time periods (in Zhou et al.[7]). Efros et al.[8] investigate learning behavior patterns based on trajectory analysis and capture the overall spatial arrangement of local motion displacement vectors. Jiang, Wu, and Katsaggelos[9] propose a hierarchical clustering method for motion trajectories analysis. There are a wide range of approaches for integrated healthcare in the context of smart homes5 and elderly living assistance[10] that provide detection of behavior, but they are more focused on activity recognition than patient abnormal behavior analysis.

Regarding the video-based monitoring field, Riboni et al.[11] uses a similar approach to the one presented in this article. However, the sensorial system in this article allows us to extract more significant features to achieve better results. Forkan et al.[12] presents interesting concepts, but it is based only on simulated data.

TECHNICAL APPROACH AND SYSTEM SETUP

To cope with HM requirements, the goal was to fulfill technical requirements for flexibility, modularity, scalability, usability, and accessibility. A three-layer architecture is proposed, as depicted in Figure 1. This architecture consists of three layers: a services subsystem, a high-level subsystem, and a low-level subsystem. The services are oriented to gather other subsystems' information and present the relevant information to the actors involved. The high-level subsystem provides functionalities employing data from multiple sources such as medical inputs, digital in-

teraction, and sensorial data to fuse information related to HM data (historical data, activity analysis, health records, and input from professionals). The outcome is a set of support decisions related to patient management. Finally, the low-level subsystem retrieves information from the sensors, processes the information, and delivers it to the high-level subsystem.
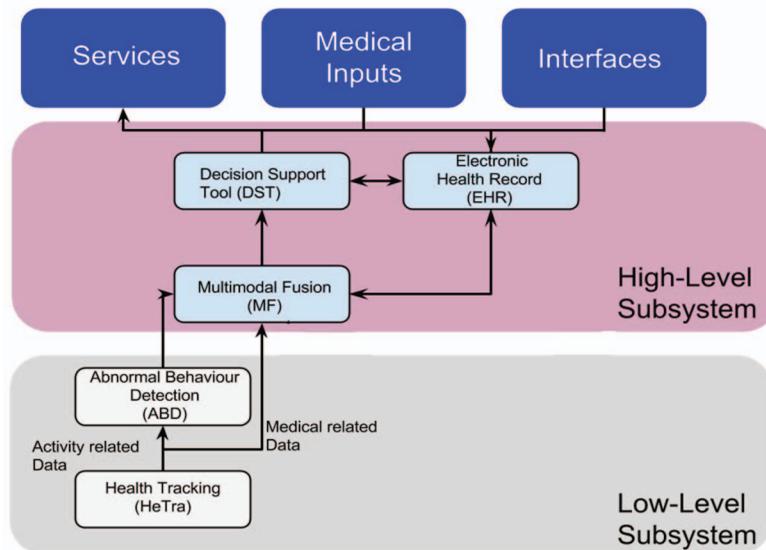

Figure 1. Overview of the complete system architecture.

SENSING TECHNOLOGIES

The selected sensors provide different types of information: visual, motion, or depth. The devices were chosen according to the scenarios where patients, depending on their disease and progression level, can be monitored in a multimodal environment: patient's home, senior center, and rehabilitation center. The employed sensors are:

- *Multisensory band* (bracelet), which provides health information (heart rate and skin temperature) and allows the acquisition of motion data (accelerometer, gyroscope, and magnetometer).
- *Binary sensor*, a non-intrusive sensor placed on doors or drawers that can detect when they are opened or closed. It is used to delimit the patient's behavior analysis zone within his or her house.
- *RGB-D (Microsoft Kinect v2) camera*, which allows for extraction of depth information, enabling deep motion analysis to monitor user activity, status, and evolution.
- *Zenith camera*, a 360-degree panoramic camera that allows a wide coverage of target areas such as living rooms in senior centers. Results of human tracking algorithms are improved with a higher view of the environment because occlusions are less present.
- *Wireless sensor network (WSN) anchors or beacons*, which complement image-based sensors by estimating trajectory. Anchors, placed on the roof, monitor the radio signals from patients' wearables in a non-intrusive manner.

Data Collection and Handling

An adaptive multi-sensorial application has been developed to simultaneously collect information from multiple sensors. This tool integrates real-time data, allowing for synchronization and organization of the sensory measurements. To fulfill ethical requirements, the system allows safe storage of information with a privacy-by-design approach. Figure 2 depicts the architecture and interconnection of sensors.
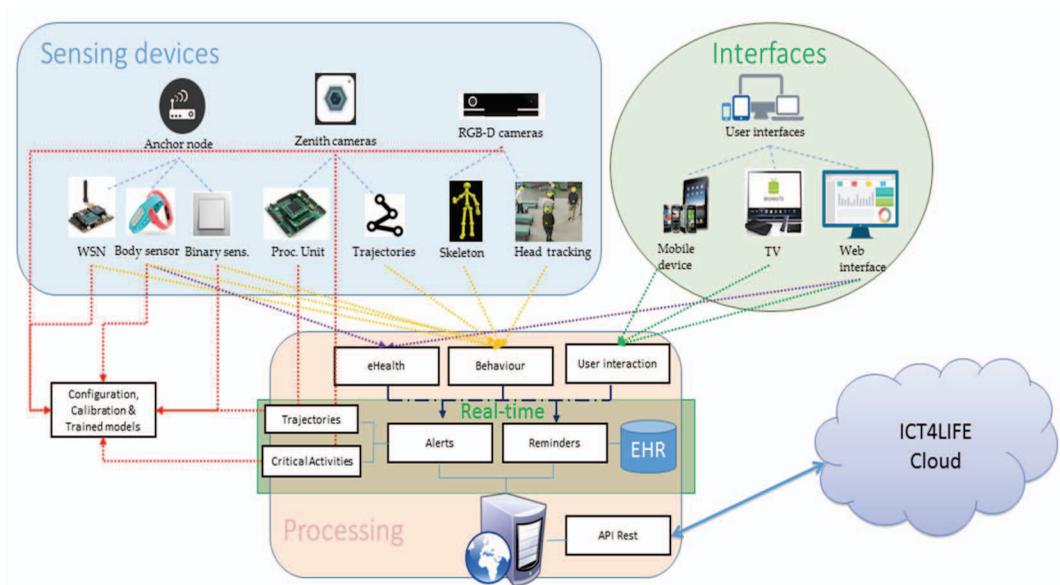
Figure 2. Architecture of the health tracking low-level subsystem. Modularity is guaranteed as the system is able to provide inferences, even when only partial information is available.

Sensor Preprocessing and Calibration

To fuse the information gathered from the diverse data sources, three preprocessing tasks must be performed: filtering signals from sensors to reduce noise (especially relevant to RGB-D Kinect), calibrating sensors, and homogenizing spatial coordinates.

Skeleton Smoothing

The RGB-D camera (Microsoft Kinect v2) sensor is able to provide several data types, such as RGB, depth, infrared, and body skeleton detection and tracking. The latter feature is important due to the wide range of potential applications. More specifically, the Kinect sensor provides a set of 25 joint points, each with its corresponding spatial coordinates (x, y, z), in addition to a confidence value (reliability of measurements) for each detected person. However, several factors such as body self-occlusion and fast movement can affect the quality of skeleton tracking. Specifically, when the skeleton's joints are occluded, they often appear to be shifted in space too quickly, in an unnatural manner. To address this issue, the Tobit-Kalman filter[13] is employed to smooth the skeleton data and reduce the Kinect's error, preventing skeleton joint location errors from being propagated to the subsequent methods that use Kinect skeleton data, such as activity recognition and abnormal behavior detection (ABD).

Device Calibration and Trajectory Extraction

To fuse the information from the sensors, all aforementioned devices must be calibrated in a common spatial coordinate system. A typical calibration scenario for the proposed system is shown in Figure 3. In practice, the xy-plane of the new coordinate system is the room floor, and the z-axis is vertically located as depicted in this figure. The process consists of three steps: selecting a common coordinate system (manual), obtaining the rotation and translation matrix for the cameras, and performing the calibration for the WSN using the mentioned reference system.

Initially, the RGB-D sensor is calibrated as follows:

- First, 10 points are defined whose distance to the sensor center is inferred using the distance from the infrared sensor.

- Next, 15 points on the floor are located, whose location is manually measured.
- For the selected points above, a singular value decomposition (SVD) method is applied to transform the 3D coordinates given by the RGB-D camera into the custom (manually defined and measured) 3D coordinates. As a result, the corresponding rotation and translation matrices are obtained that transform the RGB-D camera coordinate system to the custom one.

Figure 3. General overview of the calibration process.

This coordinate system is therefore employed as reference for the sensing submodules Zenith camera and WSN. The WSN is able to obtain the received signal strength indicator (RSSI) of packets forwarded from the bracelets to estimate the patient's location. However, this information is strongly affected by propagation effects such as multipath and reflections. This work employs a combination of two techniques: fingerprinting and kernel density estimation.[14]

The fingerprinting calibration process for WSN is as follows:

- First, a division of the scenario into square cells is manually performed. The spatial coordinates of this division must agree with the Kinect coordinates system.
- Next, sensors are deployed in the center of every cell for a predefined time (we used 5 minutes). In this period, the WSN beacons collect RSSI information for this area.
- Finally, Step 2 is repeated for every cell.

The result of this process is a complete dataset of measurements of the monitored room. This dataset is employed in the position estimation by comparing the RSSI measurements with statistics of the fingerprint dataset. The patient position is estimated using kernel functions that correlate RSSIs sensed with a particular fingerprint cell distribution. A Kalman filter is used in combination with bracelet sensors (accelerometer, gyroscope, and magnetometer) to smooth the final estimated trajectory.[15]

Zenith calibration is divided into three steps. The first step is required when intrinsic, extrinsic, and distortion parameters of the camera are unknown. Using a chessboard pattern, it is possible to estimate these values, taking at least 15 pictures of the pattern in different positions and inclinations. The second is quite similar to the Kinect camera calibration, allocating points along the monitored room and proceeding similarly to the selection of the points in the image knowing the real-world coordinates in your reference system and training a classifier (we used a support vector machine) with at least 1,000 images of people in the monitored area.

The fused information gathered by the diverse sensors improves the accuracy of patient tracking. As an example, the RGB-D sensor provides better tracking accuracy than the Zenith camera, due to the depth information. However, it has a limited range of detection (up to 5 m) compared to the Zenith or WSN coverage.

USER BEHAVIOR ANALYSIS AND PATIENT MONITORING

According to the patient's required level of attention, different types of activities have been deemed able to be monitored using the system presented in this article. The types of sensing that the system enables are:

- *Daily motion*. Monitors daily motion patterns and assesses whether they are usual.
- *Dream (night motion)*. Detects nighttime motion patterns.
- *Patient disoriented confusion wandering*. Detects whether a patient is disoriented or confused.
- *Sign of apathy*. Detects whether a patient is spending too much time on a couch or in bed.
- *Number of visits to the bathroom*. Counts the number of visits to the bathroom.
- *Patient leaving the house*. Detects when a patient is leaving the house.
- *Patient disoriented and leaving the house*. Detects when a patient is disoriented, confused, or wandering and is leaving the house.
- *Fall down*. Detects that a patient fell down.
- *FoG*. Detects patient freezing of gait.
- *Festination*. Detects patient festination.
- *Loss of balance*. Detects loss of balance in patients.
- *Movement evolution*. Evaluates therapy-related exercises of a patient.

While it is possible to use all these types of sensing with the presented approach, in this article, we focus on explaining in detail the results obtained in FoG and daily motion sensing.

Activity Recognition: Identification, Methods, and Use

Patient Identification and Association of Measurements

Patient concurrency is important in some scenarios such as the senior center and rehabilitation room (for group therapy). This raises a challenge, as high-precision "singularization" and association of the sensor measures are required for efficient monitoring and behavior detection. Vision-based systems are generally constrained because of environmental lighting conditions, and performance can be affected by occlusions. So, a person identification stage must be performed.

We tackle this issue with the proper combination of sensor features to individualize patients with unique data identifiers. Specifically, an association stage of depth sensors, WSN, and smartbands (bracelets) have been employed. Kinect can detect up to six people simultaneously, as well as the RGB image to be tracked. Moreover, WSN can obtain a unique person identifier from the bracelets (in particular, the MAC address) used as an identification parameter.

Based on the aforementioned features, the association process of the diverse sensors for person individualization is as follows: Using bracelet measurements from accelerometers, gyroscopes, and magnetometers and applying the well-known Madgwick filter, it is possible to extract inertial features (yaw, pitch, and roll) that allow for estimating the patient orientation. Information is compared to the same features obtained from Kinect and Zenith cameras. Munkres algorithm was implemented to associate the orientation (yaw) and arm angle deviation (pitch) extracted from the estimated routes as a function of the error. The outcome of this process is the association list of the available sensor data for a particular patient (identified by the band information).

Activity Recognition: Using the Multimodal System Sensors

There are several activities, such as daily motion and sign of apathy, that can be detected using trajectory information. The goal is to learn patterns that allow us to classify normal and abnormal behaviors for both Parkinson's and Alzheimer's patients using movement trajectories.

The trajectory feature extraction step consists of obtaining different types of descriptors, which reflect the occupation level of each region in a scene and capture motion characteristics such as speed, acceleration, and curvature (motion features) or combines them for obtaining a fused spatiotemporal descriptor.

Next, an improved feature representation is obtained by feeding the trajectory descriptors to a Sparse AutoEncoder (SAE) algorithm,[16] which proved useful in estimating the underlying data distribution and producing more meaningful and compact features. For finding different behavior patterns, we use an unsupervised approach that employs clustering algorithms to assign labels to data samples. Finally, users are allowed to validate the clustering results and define what is normal and what is abnormal for a considered scenario. The validation is obtained using two classifiers: SVMs or the Logistic Regression classifier.

Information from binary sensors can be employed to support the analysis of particular activities (such as to infer sleep quality based on the number of visits to the bathroom). Furthermore, the fusion of trajectories extracted from cameras with binary sensors permits assessment of the interaction of patients with their environment, yielding disorientation/confusion/wandering detection.

Besides patient location detection by the adoption of WSN, a patient's wearable can be used to identify the patient's movement. By employing the accelerometer, gyroscope, and magnetometer sensors, the user's movement can be estimated. By combining this information with the patient 3D skeleton coordinates, fall, FoG, festination, and loss of balance can be detected. Finally, the proposed platform can track and evaluate patients' movements during their rehabilitation exercises, record their performance, and assess their evolution.

ABD in Daily Motion Using a Fusion and Probabilistic Model

The MF submodule is in charge of retrieving the data from low-level subsystems such as ABD and HM, fusing this information, and generating an output that is sent to the module responsible for making the decisions regarding the patient's health status. Furthermore, it saves data results in the electronic health record (EHR) where the patient-specific profile is stored. The fusion process of the inputs takes into account the reliability of each information source, using a weighted probabilistic model where each modality is weighted according to its contribution in the fusion scheme. An example of the fusion model for computing the probability of a patient, having the ID i, being in an apathy state is shown in the following equation.

$$P^i(apathy) = w_a * P^{ABD}(apathy) + w_h * P^{HM}(apathy) + w_e * P^{EHR}(apathy)$$

Different modalities are taken into consideration, while their confidence level is provided by their associated probability measure. The weights associated with each modality are used to compensate uncertainties. In our approach, we initially set the EHR associated weight to 0.7 (w_e), if the patient has a form of depression and to 0.3 otherwise. Next, the ABD and HM weights were set to 0.5 (w_a) and 0.4 (w_h) in the first case and otherwise to 0.8 and 0.5, while the obtained result is normalized to a probability value. The HM modality has a lower weight in comparison to the ABD, due to its reliability, as cameras and especially depth sensors are better at predicting the patient condition than the wearable sensors. Finally, weights were optimized based on the classification error, a process that accounts for the variability between patients. In the training process, the ground truth labels are provided by professionals, the psychologist in case of apathy, or the neurologist in case of confused behavior.

Provided that enough training data will be available, a better fusion model, such as a Bayesian network (BN) model can be used. The BN model will be trained and used to infer the high-level information regarding the patient's health status based on observable variables provided by the low-level subsystems (such as heart rate, level of activity, and time spent interacting with the digital platform). In case the training data will not be sufficient to train the BN model, conditional probabilities are set using expert knowledge derived from the project users.

In Figure 1, the input and output modules of the MF module can be found. For processing purposes, data storage is needed in the MF module, while the data processing stages consist of data preprocessing, feature extraction, and fusion techniques. Afterwards, the biological information

of the patients will be stored into the EHR. Furthermore, these measurements will be available for the decision support tool (DST) to create recommendations and related systems.

FoG Detection

One of the examples of the proposed system use is FoG detection, as shown in Figure 4. FoG refers to a state in which a patient with Parkinson's disease experiences a sudden lack of movement despite his or her willingness to move. It typically occurs in specific situations, such as when starting to walk, stepping through a doorway, attempting to turn a corner, or approaching a chair. It typically lasts a few seconds, and it is very important to be detected, because FoG episodes are unpredictable and greatly increase the chance of falling. Furthermore, an increase in FoG episodes needs to be reported to the physician.
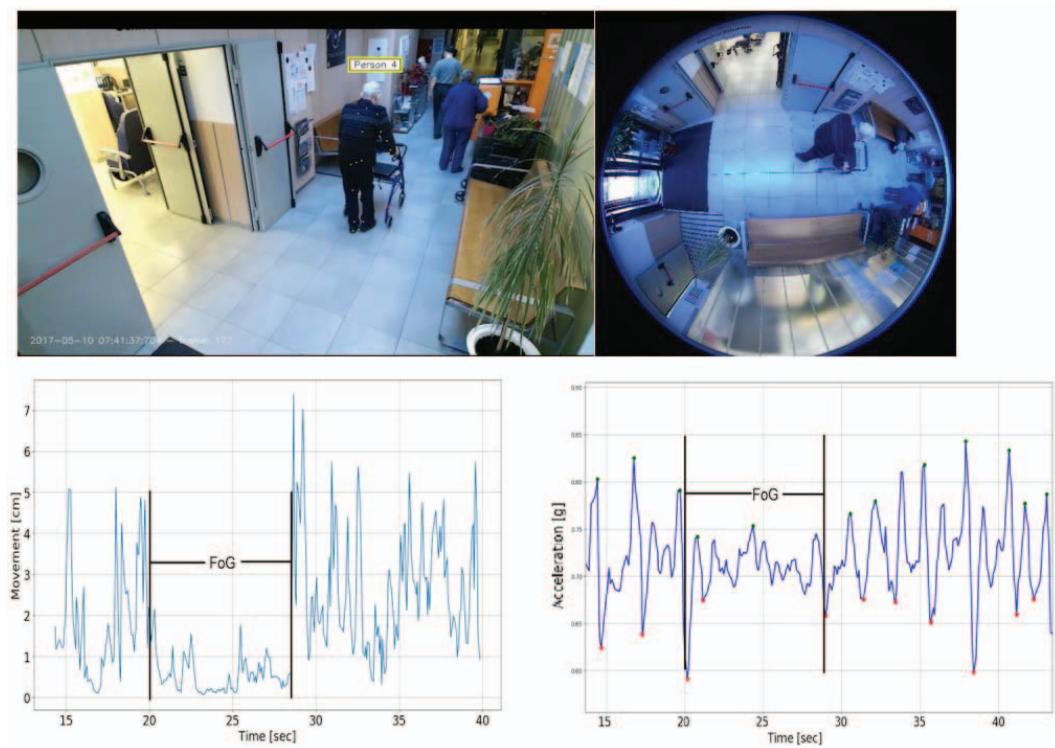

Figure 4. (Upper left) patient skeleton during a FoG episode detected by Kinect. (Upper right) patient tracked by Zenith camera. (Lower left) FoG detection from Kinect data. (Lower right) FoG from inertial measurement unit (IMU) sensor data.

The detection of the FoG action, which is a specific movement pattern of patients' limbs, is carried out by employing a recurrent neural network (RNN). The RNN has been trained using the movement information captured by the multisensory bands, along with the annotations for these movements provided by the physicians.

Regarding the feature extraction procedure, a sliding window approach along each axis (x, y, z, magnitude) was employed. The short-time Fourier transform of the raw acceleration values within each window is calculated, and the resulting spectrum forms the feature vector used to train the RNN.

TESTING AND RESULTS

Methodology

The experiments were performed in several sessions during 10 weeks with a total of 18 patients between the ages of 55 and 94, and they were gender-balanced. The control group was composed of elderly people without cognitive problems (volunteers). The experimental group was made of patients classified into three groups according to their illness stage (severe, moderate, or initial). Patients were informed of the procedure, as well as the type of data that was going to be collected. However, participants were not notified about the hypothesis nor the expected results to guarantee unbiased samples.

Initial testing of the system and training of classifiers was done in laboratory testing with patients and a control group related to abnormal behavior, and then applied to our patients captured data in the installations mentioned. In the case of FoG, testing was performed with a pre-recorded dataset from real patients.

Testing and Results for ABD in Daily Motion and FoG

The system was tested and validated for indoor monitoring and was able to learn movement patterns such as repetitive behavior based on the anomalies of the patients behaviors. This information is extracted from the trajectories from Zenith, Kinect, and WSN sensors. The algorithms were tested with a dataset recorded for 24 days in 10 weeks of experimentation. Diverse feature descriptors in both spatial descriptors and fused descriptors were employed. In our experiments, 700 sample trajectories were split into five classes, one of them representing the abnormal behavior and accounting for 30 percent of the samples. The accuracy for discriminating between normal and abnormal behaviors was 98.4 percent, while precision (98.7 percent) and recall (98.3 percent) were high, showing that our proposed methodology was successful and reliable. The adopted testing method consisted of a 10-fold cross validation approach.

The occupancy histogram (OH) descriptor achieved a good accuracy of 97.4 percent and was able to capture the correlation between activity patterns and the specific spatial regions where they are usually performed. A spatial division of the 3D scene in $n = 8x6x2$ blocks was used, while the temporal window was set to 2 minutes. However, the best results were obtained for the motion descriptor, optimized using the SAE representation,[16] proving its benefits at increasing the accuracy of the classification method in relation to the raw features.

The trained behavior detection model was applied on activity data of the recorded patients, both healthy people and those with Parkinson's disease, for discriminating between normal activities and repetitive or confused behaviors, elicited in a spontaneous manner. Examples of normal activities used were preparing tea or coffee and writing a grocery list after checking the food supplies, while the confused behavior was obtained by asking the participants to look for an item not present in the room.

The results were satisfactory, as overall accuracy for four tasks was 84.12 percent, while the accuracy of the normal activities (81.3 percent) was slightly worse than that of the confused behavior (86.95 percent). For the patients with Parkinson's disease, we also took advantage of the medical information, which supported the sensory analysis.

Figure 5 depicts the trajectories of one healthy elderly person and one Parkinson's patient. The person with Parkinson's disease stops more often while performing the activity (he is in a medium stage of the disease, having a score of 3 on the Hoehn and Yahr scale[17]). When discriminating normal and abnormal behavior, a precision of 98.7 percent and recall of 98.3 percent was achieved, outperforming the precision (89.8 percent) and recall (96.4 percent) in Riboni et al.[11]

Concerning FoG, the well-known Daphnet dataset was used to allow comparison to other works. The dataset contains 3D acceleration measurements from three sensors attached to the ankle, knee, and waist of 10 Parkinson's patients that experienced FoG.[18] The data was recorded while the patients were walking back and forth in a straight line, randomly, and into and out of rooms.

In total, 8 hours and 20 minutes of data was recorded, containing more than 200 recorded freezing incidences.
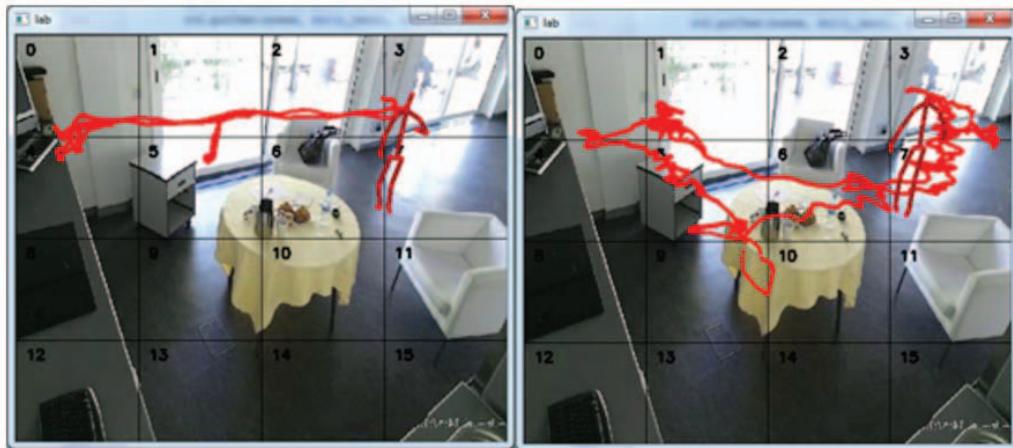

Figure 5. Trajectories of (left) a healthy elderly person and (right) a person diagnosed with Parkinson's disease.

The leave-one-out cross-validation method was selected for evaluating the classification performance, achieving average specificity and sensitivity values across patients equal to 0.91. The standard deviation of the sensitivity and specificity was 0.11 and 0.05 respectively, indicating good generalization of the model despite the low number of the dataset's users. The results are comparable to other works such as Rodríguez-Molinero et al.,[4] with slightly better results (mean sensitivity of 0.96 and specificity of 0.94) using a tailored device but on a different dataset. The versatility of our approach allows us to adapt to 12 different patient activities.

CONCLUSION

Our complete, integrated healthcare approach based on multiple ICT solutions can support care of elderly people with cognitive impairments such as Parkinson's and Alzheimer's disease by monitoring, reporting, and recommending useful information regarding their daily activity. On the one hand, a sensor-based approach—including WSN, smart-bands, binary sensors, Zenith cameras, and RGB-D technologies—recognizes physical anomalies. On the other hand, a subsystem transforms the way context and human behavior is reasoned by fusing the information from multiple sources and providing recommendations to the patients and other stakeholders. The MF techniques presented here guarantee the modularity of the entire system by allowing its operation even in cases where not all presented sensors are available. This offers the versatility to identify different activities and events related to the evaluation of the patient's health status. The results are promising compared to other works.

Concerning future lines of research, the implementation of novel models based on advanced techniques will be useful for assessing the precision and recall of the techniques described in this work. Additionally, we must validate the approach with relevant results for the rest of the presented activities, as well as improve FoG detection by using our own dataset and testing new classifiers and MF techniques.

ACKNOWLEDGMENTS

REFERENCES

1. P. Pawar et al., "A framework for the comparison of mobile patient monitoring systems," *Journal of Biomedical Informatics*, vol. 45, no. 3, 2012, pp. 544–556.

2. E.E. Tripoliti et al., "Automatic detection of freezing of gait events in patients with Parkinson's disease," *Computer methods and programs in biomedicine*, vol. 110, no. 1, 2013, pp. 12–26.

3. S. Mazilu et al., "Prediction of freezing of gait in Parkinson's from physiological wearables: an exploratory study," *IEEE journal of biomedical and health informatics*, vol. 19, no. 6, 2015, pp. 1843–1854.

4. A. Rodríguez-Molinero et al., "Validation of a Portable Device for Mapping Motor and Gait Disturbances in Parkinson's Disease," *JMIR mHealth uHealth*, vol. 3, no. 1, 2015.

5. T. Nef et al., "Evaluation of Three State-of-the-Art Classifiers for Recognition of Activities of Daily Living from Smart Home Ambient Data," *Sensors*, vol. 15, no. 5, 2015, pp. 11725–11740.

6. L. Yang, Y. Ren, and W. Zhang, "3D depth image analysis for indoor fall detection of elderly people," *Digital Communications and Networks*, vol. 2, no. 1, 2016, pp. 24–34.

7. Z. Zhou et al., "Activity analysis, summarization, and visualization for indoor human activity monitoring," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 18, no. 11, 2008, pp. 1489–1498.

8. A. Efros et al., "Recognizing action at a distance," *IEEE Int. Conf. on Computer Vision*, 2003, pp. 726–733.

9. F. Jiang, Y. Wu, and A.K. Katsaggelos, "A dynamic hierarchical clustering method for trajectory based unusual video event detection.," *IEEE Transactions on Image Processing*, vol. 18, no. 4, 2009, pp. 907–913.

10. E. Hoque et al., "Holmes: A comprehensive anomaly detection system for daily in-home activities," *11th IEEE Int. Conf. on Distributed Computing in Sensor Systems*, 2016.

11. D. Riboni et al., "Fine-grained recognition of abnormal behaviors for early detection of mild cognitive impairment," *IEEE Pervasive Computing and Communications* (PerCom), 2015, pp. 149–154.

12. A.R.M. Forkan et al., "A context-aware approach for long-term behavioural change detection and abnormality prediction in ambient assisted living," *Pattern Recognition*, vol. 48, no. 3, 2015, pp. 628–641.

13. K. Loumponias et al., "Using Tobit Kalman filtering in order to improve the Motion recorded by Microsoft Kinect," *Int. Workshop on Applied Probability* (IWAP), 2016.

14. J.D. Carlson et al., "Smart watch RSSI localization and refinement for behavioral classification using laser-SLAM for mapping and fingerprinting," *36th Annual International Conference IEEE Engineering in Medicine and Biology Society*, 2014, pp. 2173–2176.

15. T.V. Kasteren, G. Englebienne, and B. Krse, "Activity recognition using semi-markov models on real world smart home datasets," *Journal of Ambient Intelligent Smart Environments*, vol. 2, 2010, pp. 311–325.

16. J. Masci et al., "Stacked convolutional auto-encoders for hierarchical feature extraction," *21th International Conference on Artificial Neural Networks* (ICAN), 2011, pp. 52–59.

17. C.G. Goetz et al., "Movement disorder society task force report on the Hoehn and Yahr staging scale: status and recommendations," *Movement Disorders*, vol. 19, no. 9, 2004, pp. 1020–1028.

18. M. Bächlin et al., "Wearable assistant for Parkinson's disease patients with the freezing of gait symptom," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, 2010, pp. 436–446.

19. A. Belmonte-Hernández et al., "Adaptive Fingerprinting in Multi-Sensor Fusion for Accurate Indoor Tracking," *IEEE Sensors Journal*, vol. 17, no. 15, 2017, pp. 4983–4998.

ABOUT THE AUTHORS

Federico Alvarez is an assistant professor at UPM, from which he has a PhD. He has co-authored more than 60 articles, books, book chapters, and patents in the field of ICT networks and multimedia technologies. He coordinated six EU projects in the last eight years. Contact him at federico.alvarez@upm.es.

Mirela Popa is a postdoctoral research fellow at Universiteit Maastricht. Previously, she was a senior researcher in the Intelligent and Interactive Systems group at Innsbruck University working on hierarchical compositional models. She has a PhD in human behavior understanding from Technische Universiteit Delft. Popa has experience in activity recognition, computer vision, machine learning, and data analysis. Contact her at mirela.popa@maastrichtuniversity.nl.

Vassilios Solachidis is a researcher at CERTH. He has a PhD in informatics from the Aristotle University of Thessaloniki. He has worked as a researcher in 17 European and national research projects through Aristotle University of Thessaloniki, the Information Technologies Institute, and Università degli Studi di Roma Tre. His research interests focus on image and video analysis technologies and multimedia watermarking. He has published eight articles in international journals and 35 articles in international and national conferences. Contact him at vsol@iti.gr.

Gustavo Hernández-Peñaloza is a PhD candidate in the Visual Telecommunications Applications group at UPM. He is also participating in several EU projects. He has a master's degree in telecommunication technologies, systems, and networks from Universidad Politécnica de Valencia. Contact him at ghp@gatv.ssr.upm.es.

Alberto Belmonte-Hernández is a PhD candidate in the Visual Telecommunications Applications group at UPM. His main interests are communications technologies, IoT, sensors, and wireless communications. He has a master's degree in communication systems from UPM. Contact him at abh@gatv.ssr.upm.es.

Stylianos Asteriadis is an assistant professor in the Department of Data Science and Knowledge Engineering at Universiteit Maastricht. Previously, he worked at CERTH as a postdoctoral researcher. His research interests include visual computing, machine intelligence, human emotion, and activity recognition. He has a PhD from the National Technical University of Athens. Contact him at asteriadis@maastrichtuniversity.nl.

Nicholas Vretos is a postdoctoral research fellow at CERTH. He has a PhD from the Aristotle University of Thessaloniki. He has published more than 35 articles in scientific journals and conference proceedings, as well as a book chapter. His main interests are in image and video processing, semantic analysis, neural networks, and 3D data processing. Contact him at vretos@iti.gr.

Marcos Quintana is a PhD candidate in the Visual Telecommunications Applications group at UPM. His research interests include object detection, classification, and tracking in different scenarios. Contact him at mqg@gatv.ssr.upm.es.

Thomas Theodoridis is a research assistant at the Information Technologies Institute of CERTH. He has a master's degree in web science from the Aristotle University of Thessaloniki. His interests include complex networks, video processing, and machine learning. Contact him at tomastheod@iti.gr.

Dario Dotti is a PhD candidate at Universiteit Maastricht. Previously, he worked as a research assistant at Inria Sophia Antipholis in the field of computer vision and machine learning. His research interests include video-based human behavior understanding, activity analysis, and human-computer interaction. He has a master's degree from the University of Trento. Contact him at dario.dotti@maastrichtuniversity.nl.

Petros Daras is a research director at CERTH. He has a PhD in electrical and computer engineering from the Aristotle University of Thessaloniki. He has co-authored around 50 papers in refereed journals, 31 book chapters, and more than 130 articles in international conferences. Contact him at daras@iti.gr.