IN-HOME ASSISTIVE SYSTEM FOR TRAUMATIC BRAIN INJURY PATIENTS

Nuri F. Ince, Cheol-Hong Min, Ahmed H. Tewfik, Fellow, IEEE

Department of Electrical and Computer Engineering, University of Minnesota, USA

ABSTRACT

We describe a system for assisting patients in a home setting who suffer from cognitive impairments due to traumatic brain injury. The system integrates fixed wireless home sensors and wearable wireless sensors. We focus on the task of classifying activities of daily living. We locate and track the subjects with the help of home sensors and capture the details of an executed activity with a 2-axis wearable wireless accelerometer sensor attached to the right wrist. We extract time domain and frequency domain features for each task and classify them with Gaussian mixture models followed by a majority voter. The majority voter provides low false positive rates while continuously tracking the tasks. The experimental results from 2 subjects in recognizing 4 distinct daily activity tasks are promising.

Index Terms— Activities of daily living, classification, Gaussian mixture model, wireless sensors

1. INTRODUCTION

The rise in the proportion of elderly people and the shortage of national medical specialists have generated the need for intelligent systems that can monitor the subject in a home setting. With the recent advances in micromechanical devices and wireless embedded systems, several platforms have been developed to continuously monitor the physiological parameters of human body, such as ECG, blood pressure and heart rate [1]. Such systems which can continuously monitor the health of elderly people and individuals with various pathologies may enable early diagnosis and can track the effectiveness of rehabilitation. Cognitive impairments due to traumatic brain injury (TBI) also require intelligent systems that can assist the person in carrying out their daily activities. In general, the frontal lobe of the brain is damaged in TBI patients due to accidents and falls. Since higher cognitive functions reside in the frontal lobe, damage to this area causes patients to have difficulties in completing and focusing on a task [2]. Furthermore TBI patients need help in planning, organizing and completing activities. Therefore it is crucial to develop a system that gives instructions to such patients *"when needed"* and continuously monitor their activities in a home setting.

To tackle this problem, reminder/scheduler type systems have typically been developed to give instructions to cognitively impaired people. In general they are based on hand held devices to deliver messages in an "*open loop manner*" [3]. When a reminder is issued, it is expected that the subject will execute the given task. However, in the recovery stage, the subject may start to remember daily activities and getting repeated reminders may be annoying. A closed-loop system which can intelligently monitor the functional activities of the patient and then deliver the messages or reminders as needed, will eliminate both excessive and unnecessary alerts. In this paper, we focus on a system that can track patients with traumatic brain injuries. In particular, the system is intended to help such patients plan their day and complete each of the daily tasks that they are supposed to perform in a timely manner. The system tracks the functional activity of the subject. For this purpose we integrate fixed wireless home sensors and wearable wireless sensors. We locate and track the subject with the help of home sensors. We capture the details of the execution of activities with a 2-axis wearable wireless accelerometer sensor attached to the right wrist. The signals recorded by this integrated system are transferred to a PC in real time for processing. We extract time domain and frequency domain features for each task and combine them with Gaussian Mixture Models. We proceed with a majority voting procedure to achieve high true positive and low false positive rate while continuously tracking the tasks. In this paper we focus on execution of early morning daily activities such as brushing teeth, washing face and shaving.

The paper describes the architecture of the system and its features, including the additional data acquisition capabilities required to train and design the system. A schematic diagram of our system is presented in Fig. 1. In Section 2 we provide details of our system. Section 3 explains the feature extraction and the classification strategy. Finally, in section 4 we show experimental results for 2 subjects.

2. WIRELESS SENSORS AND DATA ACQUISITION

As mentioned above, our system consists of 2 sensor systems. The first sensor system is a collection of fixed wireless sensors deployed in a home environment. The second system relies on wearable sensors which provide data that complements the information by the first system. In our system, we elected to use the eNeighbor™ system (eN) that was recently developed by RedWing Technologies (www.healthsense.com) and is currently marketed under the name Healthsense™. The eN wireless sensor network is currently based on the IEEE 802.15.4 wireless standard. Healthsense will be introducing low-power IEEE 802.11 wireless sensors for use with the eN system by mid-year. The low-power 802.11 wireless sensors are fully compatible with standard WiFi networks, which significantly increase the flexibility of the system to cover an entire building or campus. The system comes with several sensors such as motion, bed, chair, and door sensors that allow us to localize most daily activities. Each sensor communicates with the base station only in the case of an event. Therefore, the sensors have long battery lives and can be used at home without frequent maintenance for long periods of time. In this study, each event is exported from the system in real time through the USB port to a PC. The eN gives binary information about the activities carried out by the individual. The sensors require peer-to-peer interaction to send data. Furthermore, they do not provide enough information to recognize activities unambiguously. Therefore, we use wearable accelerometers installed on networked wireless embedded systems to get detailed information about the activity of the patient. We selected the MICAz mote module developed by Crossbow Technology Inc. (www.xbow.com). The MICAz is IEEE 802.15.4 compliant. In our system, we used the MTS-310 multisensor board to record movement and environmental parameters. The 2-axis accelerometer signal was digitized with 10bit A/D, sampled with 50Hz and transferred to the PC via MIB-510 serial gateway in real-time. The reader can find detailed information about this data acquisition system in [4].

For this particular study, we collected 3 types of data: washing face, brushing teeth and shaving face from 2 healthy subjects, S1 and S2, with the system described above. The number of available trials for each task is given in Table 1. In addition to completing the 3 distinct tasks listed above, the subjects are also asked to do other types of activities that do not correspond to these 3 tasks. These can be, for example, changing a towel, arranging items on the sink, etc. These tasks are categorized as No-Activity. Fig. 2 shows 3 sample signals recorded with the system.

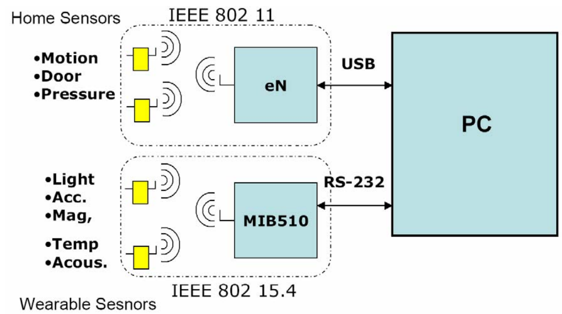
3. FEATURE EXTRACTION AND CLASSIFICATION

After localizing the subject using the eNeighbor system, we proceed to classify the activity of the subject using the accelerometer data. Accelerometer signals are increasingly being used in detection of walking/running activities in energy expenditure detection [1]. In general, time domain (TD) features such as mean, root mean square and the number of zero crossings are widely used. In this study TD features are extracted as well. In addition, we utilize frequency domain (FD) features. We extend the feature set with energies in different frequency bands. The Fourier transform of the accelerometer data is calculated in 128 sample windows for each axis. This window is shifted with 75% overlap across the signal. For each segment, we calculate the energy in dyadic frequency bands as indicated in Fig. 3 (a). This approach has direct connections with Mel-Scale features used in speech recognition. We specifically believe that these features contain significant information due to the periodic nature of the activities on which we are focusing (see Fig. 2). The FD features are then converted to log scale and combined with TD features related to the same time segment. This resulting feature vector, $x \in R^D$ has a dimension D=16 for each segment.

To classify the wearable accelerometer data, a Gaussian

Table1. Available Trials

Tasks	Subject S1	Subject S2
Brush-Teeth	40	23
Wash-Face	30	13
Shave-Face	40	36
No Activity	5	5



(a)

(b) (c)

Fig.1. (a) The data acquisition platform which combines static home and wearable wireless sensors. (b) The static home sensors. (c) Wearable wireless sensor kit attached to the right writs.

Mixture Model (GMM) based system is used as indicated in Fig. 3(b). GMM is widely used in continuous classification of EMG signals for prosthetic control and speaker identification problems due to their robustness and lower computational complexity [5,6]. A GMM probability density function (pdf), is defined as a weighted combination of N Gaussians.

$$p(x/\lambda_k) = \sum_{c=1}^{N} w_c \eta(x/\mu_c, \Sigma_c) \qquad \text{k=1, .., K} \qquad (1)$$

where λ_k is the model, x is the feature vector, η is the D dimensional Gaussian pdf component

$$\eta(\mu_c, \Sigma_c) = \frac{1}{(2\pi)^{D/2} |\Sigma_c|^{1/2}} \exp(-\frac{1}{2}(x-\mu_c)^T \Sigma_c^{-1} (x-\mu_c)) \qquad (2)$$

with mean vector μ and covariance matrix Σ . The w_c is the weight of each component and satisfies

$$\sum_{c=1}^{N} w_c = 1 . \qquad (3)$$

A new observed feature vector is assigned to one of the 4 classes (K=4) after evaluating the posterior probability of each GMM.

$$L = \arg\max(p(x/\lambda_k)) \qquad \text{k=1,.., K}, \qquad (4)$$

where L is the assigned label. The model order selection plays a big role in determining the performance in GMM based systems. While a low number of mixtures can poorly represent the geometry of the activity in a D dimensional space, a high number of mixtures generally over fit the data. Therefore, we increased the number of mixtures from 1 to 6 to find the optimal value.

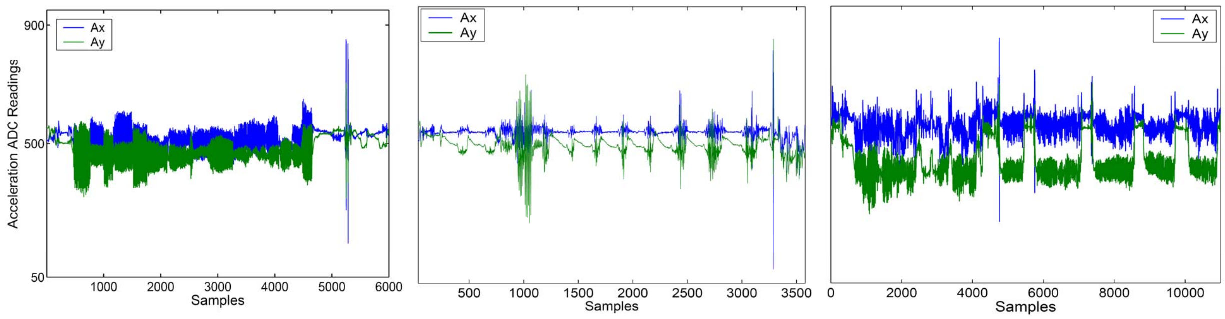


Fig.2. Typical recordings obtained from 2 channel accelerometer sensor (Ax and Ay) attached to the right wrist. (a) Brushing-Teeth (b) Washing-Face and (c) Shaving-Face

The evaluation in (4) gives a class label for each time segment which corresponds to the continuous classification of streaming data. However, we noticed that the arm movements related to each task contain many sub-segments where the activity is not locally related to the task under execution. In addition, observation noise affects local error in the classification.

For these reasons, the outputs of all GMMs are post processed by a Majority Voting (MV) procedure as indicated in Figure 3. The MV uses W points to decide whether the observation sequence is related to any of the tasks.

During our experimental study we observed that the true positive (TP) rate is quite high at the MV output. Although several time points were used for voting we noticed that the classifier performed poorly at the beginning and end of each task. However, we also noticed that the distances between different tasks are largest when the subjects are in the middle of the tasks. We used this information to modify the MV to improve the classification accuracy. We insert a threshold, th_1 at the input and another threshold th_2 at the output of the MV. We remove observations which have low posterior probability at the input stage of the MV, that is we use

$$p(x/\lambda_k) = u(p(x/\lambda_k) - th_1) \tag{5}$$

where u (p) is the unit step function. Equation (5) filters out those GMM outputs with low probabilities that occur at the beginning and end of each task. Now, assume that V^k is the total number of votes for the k-th GMM in a given window. We eliminate those votes that are not occurring th_2 times within the MV window and select the surviving vote that occurs with highest frequency.

$$L = \arg\max(V^k > th_2) \qquad k=1, ..., K , \tag{6}$$

This step has enabled us to benefit from those regions where the distances between classes are maximized. Actually, the steps described above can be seen as a strategy for extracting the most unique events related to each task.

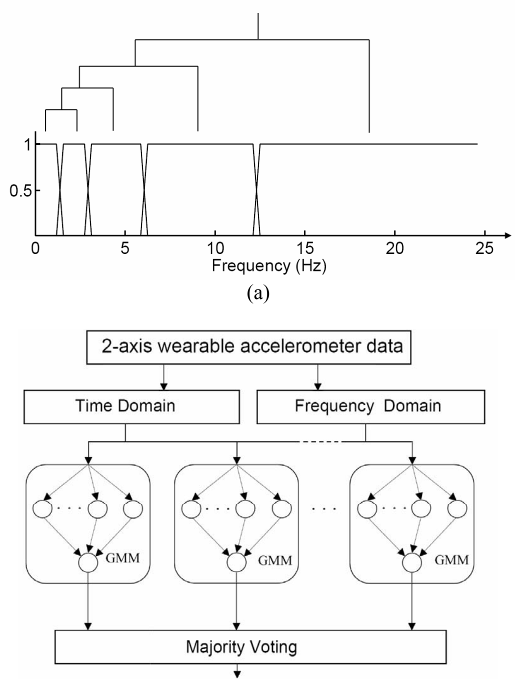


Fig.3. (a) Dyadic frequency partitioning used to extract FD features. (b) The proposed classification system

4. RESULTS

Let us present the results we have obtained from 2 subjects executing 4 different tasks in a home setting. We will show the continuous classification results and the MV outputs with and without thresholds. For MV, a W =16 sample window (≅10sec) is selected. The thresholds th_1 and th_2 were set to 0.5 and 10 respectively. The leave one trial out method is used to estimate the classification performance. Table 1 represents the continuous classification results obtained by changing the mixture sizes for 2 subjects. There is a jump in classification performance when the number of mixtures (NoM) is increased from 1 to 2 for "brushing-teeth" data. Increasing the NoM provides slight improvements for subject S1. However, for subject S2 the classification error starts to increase after 4 mixtures. The classification accuracy corresponding to no activity (NoAct) continuously drops with increasing number of mixtures. As indicated before, the NoAct represents those trials where the subject was free to execute several activities in the bathroom excluding the 3 main tasks. Therefore, the features for NoAct have great diversity. Increasing the number of mixtures clearly drops the classification accuracy. We believe that the NoAct features fill the whole feature space. With higher number of mixtures, the system might be entering into an over-learning stage. The majority voting results also support this hypothesis. In Table 3 we present the confusion matrix for NoM =

Table.2. The continuous classification results for 2 subjects

S1

Mix	Brush	Wash	Shave	NoAct
1	76.4	82.9	72.7	81.9
2	81.1	88.9	73.5	77.2
3	82.9	89.4	81.2	65.8
4	83.8	87.9	84.9	53.9
5	84.7	86.5	85.7	53.9

S2

Mix	Brush	Wash	Shave	NoAct
1	69.7	82.7	81.4	81.4
2	83.1	82.2	81.1	66.9
3	83.6	82.2	85.4	56.1
4	81.7	80.4	84.6	53.9
5	82.3	79.9	85.8	50.8

4. As it can be seen from the diagonal the TP positive rate is very high for the system that has no thresholds in the MV stage. However, there are also several false positives (FPs). As indicated previously these errors occur mostly in the transition between tasks: the hand gestures related to these segments are very similar, such as grabbing the shaver or tooth brush. Similar problems were observed in the classification of EMG signals in [5]. The authors have noticed that the GMM based system performed poorly in the transitions from one activity to another. Notice also that the performance of the whole system on NoAct trials is poor when no threshold used in the MV. In nearly each trial a FP is observed. Reducing the FP rate was therefore our main concern.

In table 4, we present the results obtained by applying thresholds to MV. By eliminating low probability or fluctuating outputs of the GMMs at the input of the MV and then assigning labels on the votes occurring most often we have nearly eliminated all FPs. For example, for S1 the 20 FPs for brushing tasks have been reduced to only one FP. Furthermore, for S2 the Brushing and Washing tasks are classified with 100% accuracy. For both subjects the FPs in NoAct regions are completely removed. However, for each subject, one NoAct trial has been missed. According to our classification strategy there is no harm in missing NoAct trials. Also, it is acceptable to assign a real task to both the task itself, and NoAct.

We specifically examined those trials where a FP occurred in the MV with thresholds. For example, for S2 the FP rate in the shaving task is related to the initial stages where the subject is putting shaving cream on his face. We notice that the circular hand movements in this region became similar to the brushing task. In the mid regions of this activity, the system always voted for shaving. The existence of such subtasks in each activity and the temporal variability of the task contribute the most to the FP rate that we observed. Unfortunately, there is no way to define start and end points of the subtask and actual activity. When a large time window is used for MV the local information can be missed. Small windows are susceptible to local error. We target these challenging problems in our current studies.

5. CONCLUSION

In this paper we presented a new classification system to continuously recognize the activities of daily living by using wearable wireless accelerometer sensors and static wireless home sensors. The system is based on the GMMs that use TD and FD features. A majority voting module is implemented at the post-end of the GMM to increase the classification accuracy. By tuning the MV with thresholds, we observed a significant decrease in the number of FPs, which occurred in the recognition of almost every trial. The same approach may also be a solution for the transition problems defined in [5]. The computational complexity of the algorithms is low and can be installed into the wireless sensor kits in order to develop an intelligent sensor network. We utilized only accelerometer signal recorded from right wrist. However the number of sensors can be extended to light, temperature and microphone inputs which can deliver additional information about the task. For example a shaving activity implemented with an electrical shaver may eliminate all hand gestures. In this case a microphone can pick up the necessary information. Currently the authors are working on multimodal sensory inputs to enhance the system for an expanded library of tasks.

6. REFERENCES

[1] I. Korhonen, J Parkka and M. Van Gils, "Health Monitoring in the Home of the Future," *IEEE Engineering in Medicine and Biology Magazine*, pp. 66-73, May 2003.

[2] "Traumatic Brain Injury", Department of Veterans Affairs Employee Education System 2004.

[3] R. Levinson, "Peat- the planning and execution assistant and trainer. " Journal of Head Trauma Rehab. 1997

[4] N. F. Ince, C. Min, A. H. Tewfik, "Integrtion of wearable wireless sensors and static home sensors to monitor the activities of daily living." *IEEE MDBS Symposium,* MIT Boston, 2006.

[5] Y Huang, K.B. Englehart, B. Hudgins, A.D.C. Chan, "A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses; *IEEE Trans. on Biomed. Eng.,* vol. 52, Issue 11, Nov. 2005.

[6] D. A. Reynolds, "Robust text-independent speaker identification using Gaussian mixture speaker models," *IEEE Trans. Speech Audio Process.*, vol. 3, no. 1, Jan. 1995.

Table 3. The outputs of MV module without thresholds

S1

Mix	Brush	Wash	Shave	NoAct
Brush	40	20	0	6
Wash	0	40	0	1
Shave	1	11	30	1
NoAct	0	3	2	5

S2

Mix	Brush	Wash	Shave	NoAct
Brush	23	3	0	2
Wash	0	36	1	4
Shave	8	8	13	0
NoAct	2	2	2	5

Table 4. The outputs of MV module with thresholds

S1

Mix	Brush	Wash	Shave	NoAct
Brush	40	1	0	1
Wash	0	40	0	0
Shave	1	0	30	1
NoAct	0	0	0	4

S2

Mix	Brush	Wash	Shave	NoAct
Brush	23	0	0	0
Wash	0	36	0	0
Shave	1	0	13	0
NoAct	0	0	0	4