Detection of Chewing Motion Using a Glasses Mounted Accelerometer Towards Monitoring of Food Intake Events in the Elderly

Gert Mertes, Hans Hallez, Tom Croonenborghs, and Bart Vanrumste

Abstract

A novel way to detect food intake events using a wearable accelerometer is presented in this paper. The accelerometer is mounted on wearable glasses and used to capture the movements of the head. During meals, a person's chewing motion is clearly visible in the time domain of the captured accelerometer signal. Features are extracted from this signal and a forward feature selection algorithm is used to determine the optimal set of features. Support Vector Machine and Random Forest classifiers are then used to automatically classify between epochs of chewing and non-chewing. Data was collected from 5 volunteers. The Support Vector Machine approach with linear kernel performs best with a detection accuracy of 73.98% ± 3.99.

1 Introduction

Studies have shown that up to 15% of community dwelling and home-bound adults aged over 65 are malnourished and up to 45% are at risk [1, 2]. It is estimated that between 20 and 60% of hospitalised elderly and up to 85% of nursing home residents are malnourished [3]. Malnutrition is most frequent in the frailest of people, particularly those who are less autonomous and require help performing daily tasks. Furthermore, malnutrition has been identified as one of four causes of frailty [4]. Frailty is considered to be a distinct syndrome, characterised by weakness, a slow walking speed, a low level of physical activity, unintentional weight loss and exhaustion.

Nutrition is an important factor in the elderly's health status. Malnourishment is associated with decreased muscle strength, poorly healing wounds, an increased hospital admission length and increased hospital mortality rate [5]. Furthermore, malnourished elderly are more prone to develop pressure ulcers and infections [6]. Preventing malnutrition by means of a targeted nutritional intervention could greatly improve the quality of life. Early recognition and treatment should therefore be included in the routine care of every elderly [7].

1.1 Food Intake Monitoring

Determining malnutrition can be done in a few ways. The first is by means of a self-report diary. These have been used to measure pain, sleep, illness or injury and health care use, as well as eating-related issues such as binge eating, energy intake and expenditure in weight loss treatment [8]. In the case of malnutrition, the diary provides insight into two aspects of nutritional intake. The first is to monitor a person's eating behaviour and food consumption on a daily basis in order to see if enough meals are consumed, and second, to record in detail all foods consumed for a nutrient analyses. The person is instructed to record all food intake, usually including location, time of day, quantity eaten, and nutrient values. A self-report diary is typically in paper-and-pencil format, but computerised solutions using a tablet-pc or terminal specifically catered to elderly people also exist [9]. It is clear, however, that a self-report diary has several limitations when used to self-monitor elderly people. First and foremost, keeping track of food intake and the need to look up foods in a nutrient guide and record the amount of intake is a time consuming task. The self-monitoring protocol is seldom

G. Mertes (✉) · T. Croonenborghs · B. Vanrumste
KU Leuven, Technology Campus Geel, AdvISe, Geel, Belgium
e-mail: gertmertes@gmail.com

G. Mertes · B. Vanrumste
KU Leuven, ESAT-STADIUS, Leuven, Belgium

G. Mertes · B. Vanrumste
iMinds Medical Information Technology Department, Leuven, Belgium

H. Hallez
KU Leuven, Technology Campus Oostende, ReMI, Leuven, Belgium

H. Hallez · T. Croonenborghs
Department of Computer Science, KU Leuven, Leuven, Belgium

© Springer Nature Singapore Pte Ltd. 2019

Y.-T. Zhang et al. (eds.), *International Conference on Biomedical and Health Informatics*, IFMBE Proceedings 64,
https://doi.org/10.1007/978-981-10-4505-9_12

followed adequately, resulting in an incomplete diary [8]. Furthermore, limited literacy skills or bad handwriting also play an important role. Similar techniques such as 24-hour recalls, food records or food frequency questionnaires share the same limitations, especially in elderly care.

A different type and the most widespread tool for nutritional screening and assessment is the Mini Nutritional Assessment (MNA) [5]. The MNA contains 18 questions grouped into 4 parts: anthropometry, general status, dietary habits, and self-perceived health and nutrition states. Each question is graded and summed up to a total of 30 points. The result is defined by the following thresholds: a score below 17 indicates malnutrition; a score between 17 and 23.5 indicates a risk of malnutrition; scores above 24 indicates a good status.

Other tools such as the Geriatric Nutritional Risk Index (GNRI) [10] and Cumulative Illness Rating Scale (CIRS) [11] have also been used in combination with the MNA to provide further insight into the person's health status [1, 5].

An important limitation, however, that instruments such as the MNA all share is the requirement for a health care professional to assist in taking and completing the test. Neither are they taken at routine intervals due to their time consuming nature [12]. They are therefore not used as a preventative tool to detect malnutrition at an early stage. In case of home-bound elderly receiving home care, tests such as the MNA are typically never administered unless ordered by a GP or after admission to a hospital. The results of these tests are also not always on par with what care-takers observe on a day to day basis.

1.2 Detecting Food Intake

A potential solution to replace manual self-monitoring methods is through the use of wearable devices. A wearable device that is able to detect food intake events and determine the amount of food ingested could replace manual food diaries and questionnaires. Sazonov and Fontana [13] demonstrated the use of a piezoelectric strain gauge sensor fixed to the lower jaw to detect epochs of chewing with high accuracy. In [14], the strain gauge sensor is incorporated in a larger system together with a hand gesture sensor and an accelerometer worn on a lanyard around the neck. In [15], 3D surface reconstruction from pictures taken with a mobile phone was used to determine the amount and type of food ingested. Detection of chewing and swallowing using a wearable microphone was presented in [16] and [17].

In this paper, the use of an accelerometer mounted on wearable glasses is proposed to measure the chewing motion as part of a system to measure food intake. The use of an accelerometer integrated into an already worn pair of glasses would have little impact on the elderly's comfort and is less stigmatising than other alternatives. Glasses are typically taken off to sleep, during which the sensor could be wirelessly charged on the night stand.

2 Methods

2.1 Glasses Mounted Accelerometer

Figure 1 shows the prototype setup used to capture the data. We used the low-noise tri-axial accelerometer of a Shimmer3 unit with a sample frequency of 128 Hz to capture the movements. The raw accelerometer signal is first filtered using a 10th order Chebyshev band-pass filter with $f_L = 1$ and $f_H = 45$ Hz in order to discard DC offset and high frequency noise and prevent aliasing.

In order to determine the feasibility of this method to detect chewing motion, the researcher himself consumed a meal while recording the accelerometer signal. The meal was recorded with a camera for annotation purposes. Figure 2 shows the captured signal in each of the three dimensions. The overlaying square wave is the annotation signal indicating an epoch of non-chewing (0) or chewing (1). As soon as chewing starts around the 6 mark, distinct peaks can be observed in all three dimensions, although different in amplitude. After comparing the accelerometer signal with what was visible in the video, we found that these peaks are the result of the chewing motion: a peak is captured each time the jaw is closed. The first four such peaks are highlighted in blue in Fig. 2. Since these peaks are visible in the time domain, it should be possible to extract characterising features from the signal to be used for classification.

2.2 Dataset

To construct the training and test dataset, data was collected from 5 volunteers who were asked to consume a meal while



Fig. 1 Setup used for data collection. The Shimmer sensor is firmly attached to the frame using cable ties

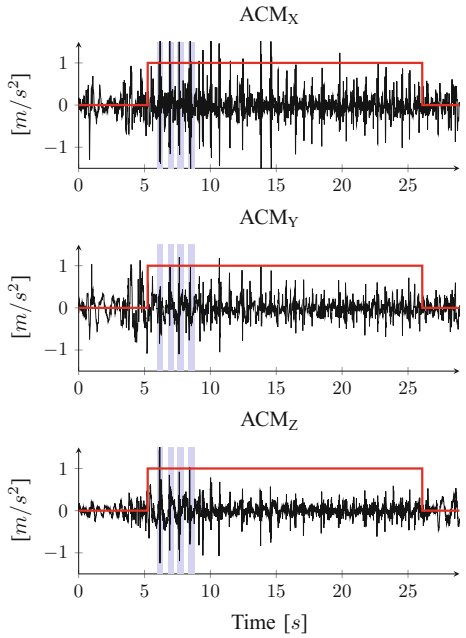

Fig. 2 Illustration of the captured tri-axial accelerometer signal while eating. The red annotation signal indicates epochs of chewing (1) and not chewing (0). The highlighted peaks represent the closing motion of the jaw (only the first four are highlighted) (Color figure online)

wearing the acquisition setup. Annotation was done by an observer. Two states were annotated: chewing (1) and not-chewing (0). As soon as the food entered the mouth and chewing started, the annotation was set to *chewing* until the food was swallowed, after which the annotation was set to *not-chewing*. Examples of activities that fall under the *not-chewing* class are: talking to the observer, bringing food to mouth, cutting food, etc. In order to get a representation of every day meals, food items with different properties were selected. The following meals were consumed: a crunchy deli sub sandwich, a mixed salad with bread (two times), mashed potatoes with vegetarian burger, and a hamburger.

Test subjects were also asked to walk around the room for roughly 1 min. This was done to determine if we are able to distinguish the chewing motion from other types of daily activities. This resulted in a total of three classes: *chewing*, *not-chewing* and *walking*.

2.3 Feature Extraction and Selection

From the triaxial accelerometer signal (x, y and z), the resultant net acceleration r is calculated using Eq. 1.

$$r = \sqrt{x^2 + y^2 + z^2} \tag{1}$$

All data is then split up according to the recorded annotation. For example, all data containing *chewing* is concatenated serially to produce one signal containing only the *chewing* activity. Likewise for the *not-chewing* and *walking* activity. As discussed in Sect. 2.1, the signals are then filtered with a band-pass filter with $f_L = 1$ Hz and $f_H = 45$ Hz. The filtered signal is segmented into non overlapping windows of 5 s. Concatenation is done to prevent windows containing data from different classes in the training dataset. Window size was experimentally determined to allow for enough windows that don't contain data from different classes when the detector is used in real-time. Chewing typically takes between 10 and 20 s, a window size of 5 s ensures that enough windows completely contain data of only one class. Features are subsequently extracted from the net acceleration signal on a per window basis. Table 1 shows an overview of the extracted features.

Table 1 List of extracted features. Highlighted in bold are those selected by the forward feature selection

	Name
1	Standard deviation
2	Mean
3	Power
4	Range
5	Skewness
6	**# of zero-crossings of r**
7	# of zero-crossings of $\frac{d^2r}{dt^2}$
8	25th percentile value
9	50th percentile value
10	**75th percentile value**
11	**Dominant frequency**

A forward feature selection based on [18] is performed on the dataset to eliminate redundant features. This method selects features with high correlation to the class, while discarding those having high intercorrelation. The total of 11 features is reduced by the algorithm to a final set of three, as shown in bold in Table 1: zero crossing rate, 75th percentile value and dominant frequency (determined via FFT).

2.4 Classification

Equivalent to the feature extraction as described in the previous section, classification is done on a per window basis. Two classifiers are evaluated: the Support Vector Machine (SVM) and the Random Forest (RF) decision tree. Classifier parameters were experimentally tuned to produce the highest accuracy. For the SVM, we chose a linear kernel with cost parameter $C = 1$ and the RF was constructed with a maximum of 100 trees. It is worth noticing that a feature selection is typically not required when using decision trees such as Random Forest due to their already selective nature in

features. However, we evaluated this and found that the RF performed better using only the three features selected by the feature selection.

Due to the limited size of the dataset, validation of the classifiers is done using the leave-one-out method. One person is excluded from the training set and used to test the classifier. This is done for each of the five participants and the results are averaged. We use the accuracy as performance metric. This method provides the added value that the classifiers can be tested on each person individually and evaluate how well they work as a group model.

To construct the training dataset, the method described in Sect. 2.3 is used. To construct the test set, a slightly altered version is used. Because we want to simulate the use of the classifiers in a real life setting, we segment the original signal into windows of five seconds without the concatenation step. This means, however, that a single window could potentially contain data from different classes. When this is the case, a choice is made: when a window contains data of a certain class for over 50% of the time, this class label is assigned to the window.

3 Results

Two experiments are conducted. In the first experiment, only two classes are included: *chewing* and *not-chewing* while the *walking* class is omitted from both the training and test set. In the second experiment, the *walking* class is also included together with *chewing* and *not-chewing*. Leave-one-out validation as described in Sect. 2.4 is used in both cases. Table 2 shows the results of these two experiments. The table contains the accuracy and standard deviation of the leave-one-out validation. We can see that the SVM classifier performs slightly better than the RF classifier in both cases, although the difference is statistically insignificant, with an average accuracy of 73.98% ± 3.99.

Table 3 shows the confusion matrices of the two experiments for the SVM classifier. These matrices contain the summed result of the leave-one-out validation, i.e. the confusion matrix values for each participant that was left out are added together.

Table 2 Results of the leave-one-out validation (acc. ± std.dev.)

Included classes	SVM	RF
Chewing–not-chewing	**73.98% ± 3.99**	72.39% ± 6.51
Chewing–not-chewing–walking	71.93% ± 5.03	69.79% ± 8.79

Table 3 Confusion matrices for the SVM classifier. Sum of all leave-one-out results

	Chewing	Not-chewing	
Chewing	**373**	93	
Not-chewing	120	**230**	
	Chewing	Not-chewing	Walking
Chewing	**361**	98	7
Not-chewing	115	**228**	7
Walking	6	9	**23**

4 Discussion and Conclusion

The average detection accuracy of 73.98% ± 3.99 obtained with the SVM indicates that our approach is able to correctly classify chewing events, but a considerable amount of false positives are still present. This can be seen in the confusion matrices in Table 3. Averaged over the five participants, the amount of false positives does not bias towards one specific activity. However, we found the false positive rate to be very person-specific. For example, when using the SVM classifier and classifying between *chewing* and *not-chewing*, for two out of five participants the *chewing* activity was frequently incorrectly classified as *not-chewing*, while for the other three participants the opposite was true. Likewise for the *walking* activity: for three participants there were no false positives for this activity, while the remaining two did have roughly 30% false positives. For all five participants, however, the true positive rate remained higher than the false positive rate.

This difference in false positives per person can be attributed to a couple of reasons. First, there is the fact that the annotation is done by an observer during the meal and is therefore not perfect. While this is not a problem for the *walking* activity, some errors could be made when annotating between *chewing* and *not-chewing*. Secondly, the dataset which was used to train and validate the classifiers is limited to only five participants. It is also worth noticing that our dataset is unbalanced, with less activities of the *not-chewing* class and only a few of the *walking* class. In order to further reduce the amount of false positives, a larger dataset would have to be recorded.

Adding the *walking* activity to the list of included classes lowers the detection accuracy. This indicates that there is still room for improvement in our proposed method. Looking towards future work, a possible improvement could be to further incorporate features from the frequency domain in the classifier or look into methods such as wavelet transforms. Furthermore, while the five second window was

chosen based on a motivated choice, the effect of the window size on the accuracy still stands to be determined.

Different studies have shown that it is possible to detect chewing motion using a group model with a jaw strain gauge sensor or microphone system with accuracies ranging from 80 to 90% [14, 16, 17]. While our system did not improve on these accuracies, it does offer the fact that the sensor can be incorporated into an existing pair of glasses, either by using a custom frame with the sensor built in or using a clip-on system. This would have little impact on the comfort of the wearer and makes the system more suitable for elderly people. Before this can happen, however, more research specifically targeting elderly people is required, starting with a case-study examining the elderly's and care givers' willingness to use such a system and the acquisition of a dataset with test subjects in this demographic group.

Acknowledgements This work was funded by internal KU Leuven grant IMP/14/038 with support from COST Action IC1303: AAPELE.

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. L. Donini, P. Scardella, L. Piombo, B. Neri, R. Asprino, A. Proietti, S. Carcaterra, E. Cava, S. Cataldi, D. Cucinotta, G. Di Bella, M. Barbagallo, and A. Morrone, "Malnutrition in elderly: Social and economic determinants," *The Journal of Nutrition, Health & Aging*, vol. 17, pp. 9–15, 2013.
2. Nutricia, "Results of the NutriAction II study," 2013.
3. L. Donini, C. Savina, M. Piredda, D. Cucinotta, A. Fiorito, E. Inelmen, G. Sergi, L. Dominguez, M. Barbagallo, and C. Cannella, "Senile anorexia in acute-ward and rehabilitation settings," *The Journal of Nutrition Health and Aging*, vol. 12, no. 8, pp. 511–517, 2008.
4. R. DiMaria-Ghalili and E. Amella, "Nutrition in older adults: Intervention and assessment can help curb the growing threat of malnutrition." *American Journal of Nursing*, vol. 105, pp. 40–50, 2005.
5. E. Cereda, C. Pedrolli, A. Zagami, A. Vanotti, S. Piffer, A. Opizzi, M. Rondanelli, and R. Caccialanza, "Nutritional screening and mortality in newly institutionalised elderly: a comparison between the geriatric nutritional risk index and the mini nutritional assessment," *Clinical Nutrition*, vol. 30, no. 6, pp. 793–798, 2011.
6. D. Volkert, L. Pauly, P. Stehle, and C. C. Sieber, "Prevalence of malnutrition in orally and tube-fed elderly nursing home residents in Germany and its relation to health complaints and dietary intake," *Gastroenterology research and practice*, 2011.
7. H. Lochs, C. Pichard, and S. Allison, "Evidence supports nutritional support," *Clinical Nutrition*, vol. 25, no. 2, pp. 177–179, 2006.
8. L. Burke, M. Warziski, T. Starrett, J. Choo, E. Music, S. Sereika, S. Stark, and M. Sevick, "Self-monitoring dietary intake: Current and future practices," *Journal of Renal Nutrition*, pp. 281–290, 2005.
9. J.-M. Wu, H.-J. Yu, T.-W. Ho, X.-Y. Su, M.-T. Lin, and F. Lai, "Tablet pc-enabled application intervention for patients with gastric cancer undergoing gastrectomy," *Computer methods and programs in biomedicine*, vol. 119, no. 2, pp. 101–109, 2015.
10. O. Bouillanne, G. Morineau, C. Dupont, I. Coulombel, J.-P. Vincent, I. Nicolis, S. Benazeth, L. Cynober, and C. Aussel, "Geriatric nutritional risk index: a new index for evaluating at-risk elderly medical patients," *The American journal of clinical nutrition*, vol. 82, no. 4, pp. 777–783, 2005.
11. P. A. Parmelee, P. D. Thuras, I. R. Katz, and M. P. Lawton, "Validation of the cumulative illness rating scale in a geriatric residential population." *Journal of the American Geriatrics Society*, 1995.
12. G. Mertes, G. Baldewijns, P.-J. Dingenen, T. Croonenborghs, and B. Vanrumste, "Automatic fall risk estimation using the nintendo wii balance board," in *Biomedical Engineering Systems and Technologies*, 2015.
13. E. Sazonov and J. Fontana, "A sensor system for automatic detection of food intake through non-invasive monitoring of chewing," *IEEE Journal of Sensors*, vol. 12, pp. 1340–1348, 2012.
14. J. Fontana, M. Farooq, and E. Sazonov, "Automatic ingestion monitor: A novel wearable device for monitoring of ingestive behavior," *IEEE Transactions on Biomedical Engineering*, pp. 1772–1779, 2014.
15. M. Puri, Z. Zhiwei, Y. Qian, A. Divakaran, and H. Sawhney, "Recognition and volume estimation of food intake using a mobile device," in *Workshop on Applications of Computer Vision*, 2009.
16. S. Passler and W.-J. Fischer, "Food intake activity detection using a wearable microphone system," in *Intelligent Environments (IE), 2011 7th International Conference on*. IEEE, 2011, pp. 298–301.
17. O. Amft, "A wearable earpad sensor for chewing monitoring," in *Sensors, 2010 IEEE*. IEEE, 2010, pp. 222–227.
18. N. Z. Hamilton, "Correlation-based feature subset selection for machine learning," 1998.