Sensor Fusion in Upper Limb Area Networks: A Survey

Aaron Duivenvoorden[a], Kyungjin Lee[b], Maxime Raison[c], Sofiane Achiche[c]

[a]School of Electrical and Computer Engineering, RMIT University, Melbourne, Australia

[b]Department of Electrical and Computer Engineering, Seoul National University, Seoul, South Korea

[c]Department of Mechanical Engineering, Polytechnique Montreal, Montreal, Canada

Abstract— **Body sensor networks (BSNs) have been increasingly used in medical applications such as exoskeleton control, powered prosthesis control, tremor suppression, gesture and sign language recognition systems, and human computer interfaces. This review explores the use of multi-modal sensor fusion in BSNs for the detection, measurement and classification of upper limb for the control of dynamic systems. Specifically, the review will look into the most common multi-modal sensor combinations found in literature, namely inertial measurement units (IMUs) with electromyography (EMG), IMUs with camera systems, EMG with electroencephalography (EEG), and IMUs with flexible force sensors. The advantages and challenges associated with these sensor combinations is discussed, as well as the challenges of sensor fusion in a broad nature, with particular focus on the use of data, feature, or decision level fusion.**

Keywords—Body Sensor Networks (BSNs); sensor fusion; multimodal; upper limb; motion analysis.

I. Introduction

Body sensor networks (BSNs), also referred to as body area network (BAN), or more specifically as wireless body area network (WBAN), are wireless networks of computing devices placed on the human body [1]. Over the last decade, there has been an increased growth in the area of BSNs in both the demand and number of applications. However, the current design and implementation of BSNs face challenges that are limiting further developments. BSNs generally produce large amount of data for accurate and reliable results, but this data needs to be processed with limited computational load, storage, and power consumption. Sensor fusion techniques can effectively exploit data from multiple sensors, identifying useful information from redundant data but reducing the load and energy consumption at the same time. [2, 3]

This review is part of the research project on designing a Body Sensor Network to effectively detect the movement of upper limbs for the control of dynamic systems, such as assistive exoskeletons or powered prosthesis for the upper limb. BSNs of the upper limb will be referred to as "upper limb area networks" for the remainder of this review. The purpose of this literature review is to identify the trends of current studies on sensor fusion systems and gain an in-depth understanding of the methods and techniques used. While other studies have made survey and summarization of current trends in Body Area and Sensor Networks in a broad nature [1-5], this review will specifically explore the types of combinations of sensors that are used for the detection of upper limb movement, specifically the configuration and processing techniques associated with them, and the problems and challenges.

The upper limb, which is composed of the shoulder, arm, hand, and fingers, is an extremely versatile and functional element of the human body. Compared to the lower limb, the upper limb is less firmly linked to the rest of the body, giving it more mobility and maneuverability [4]. The diversity of movements the upper limb is capable of gives it extreme potential for application. As such many extensive studies have been done to effectively measure and process upper limb movement data to either recognize, track, or classify the movements in a diverse range of applications, such as powered prosthesis control [5, 6], assistive exoskeletons [7-13], HCI control [14, 15], Gesture Control [15-17] and many others.

When Upper Limb Area Networks were initially being developed, each sensor network would often only have one type of sensor at its disposal. This is particularly evident in early movement recognition and tracking systems, for which the main sensors being utilized were electromyography (EMG) sensors, inertial measurement units (IMU), electroencephalogram (EEG) sensors, and optical sensors. Although some more modern systems still rely on a single sensor module, it is now increasingly common for systems to include a fusion of sensors with multiple modules.

Data from a single sensor has a limited dimensionality since it can only measure limited aspects of the movement. With multiple sensors, each measuring different modalities, it makes it possible to reconstruct the movement with higher dimensionality. Multimodal systems can also provide higher accuracy by exploiting two different sensors that measure the same phenomena but in different ways. Since each sensor may have spatial, temporal coverage limitations and degradation of accuracy in certain situations, having overlapping data between sensor nodes can make the system less vulnerable to the limitations of a single node [18].

The remainder of the paper is organized as follows. Section 2 discusses the literature search strategy that was undertaken in this paper. In Section 3 a review of sensor fusion is given and includes an overview of the most common multi-modal sensor combinations utilized in the field of upper limb movement measurement. Section 4 presents a comparison of the many different fusion techniques utilized, and provides commentary on which combinations of sensors are best suited to which technique. Section 5 provides discussion on the challenges that are experienced with sensor fusion. Finally, Section 6 concludes the paper and provides insight into the areas of future development.

978-1-5090-4934-9/17/$31.00 ©2017 IEEE

II. Method of Investigation

Regarding the subject of this review: “Gesture recognition”, “Fusion” and “Sensor” were used and combined as high-level keywords. Additionally, “Gesture recognition” was also combined with the following key words: sign language; hang gestures; upper-limb and upper-extremity. Where “Fusion” was combined to the following keywords: multi; hybrid; multi-modal and multimodal. These keywords were chosen after initial search was done to ensure there were no areas within this field that were excluded due to inadequate keywords that did not fully represent the relevant literature. This process was based on a number of current textbooks and literature reviews that were acquired and analyzed. [2, 3, 16, 19-21]

Six search directories were utilized in this search strategy and were chosen for their content and relevance to this field of study. They were Ei Compendex, Inspec - IET, Web of Science, PubMed, Proquest, and Science-Direct. Abstract and full texts were then collaboratively assessed and peer reviewed by the authors for direct relevance to the topic and scope of this study. A total of 96 relevant papers were discovered from papers restricted to the last 10 years. The results were then classified into groups, as seen in the figure below. The broad terms ‘Inertial’ and ‘Camera’ were used for the purpose of simplification and include gyroscope, accelerometer, and magnetometer sensor for ‘Inertial’ and RGB Cameras, Infrared Cameras, and Depth Sensors for ‘Camera’.

As seen in Fig. 1, EMG and inertial, inertial and camera, EEG and EMG, and inertial and flex sensor are the most prevalent combinations exploited. Regarding the result of this literature search, these 4 combinations are the main topics covered in this paper. Multi-modal sensor fusion is definitely a growing trend in upper limb area networks, considering 34.7% of the papers collected were from the last two years (2015 to 2017), Figure 3 highlights this, showing the number of papers found for the period of January 2007 to February 2017, which are relevant to the combinations of Inertial sensors with EMG sensors and Inertial sensors with Camera Systems, which are the two most common combinations for upper limb movement measurement.

III. Review on Sensor Fusion

Upper limb movement analysis can be divided into movement tracking or movement classification as illustrated in Fig. 2. Movement tracking is the kinematic modeling of the upper arm and is mostly developed with Graphic User Interfaces to visualize the model, which can be used for applications such as rehabilitation, virtual reality, sports analysis, etc. Movement classification, on the other hand, has been used for applications such as gesture recognition, movement intent recognition, and even sign language recognition. It involves the extraction of different types of features from the acquired sensor data that can be then treated by machine learning techniques. Depending on the characteristics of the sensors, each combination can be applicable for either tracking, classification or both.

A. Inertial and EMG Sensors Combination

Combining Inertial and EMG sensors is a good example of a cooperative sensor grouping [18]. Since these sensors measure different aspects of the same phenomenon, their combination has the advantage of providing a more in-depth view of the phenomenon being measured. An IMU measures acceleration, gravity, and angular velocity, which makes it possible to capture the kinetic features of a movement. The internal biological features of a movement, that being the electrical potentials present in the muscle fiber as part of the signaling process, can be measured by EMG. By using these sensors together, one is able to recognize a broader range of gestures and upper limb motions, and as such this combination has been used for applications such as sign language recognition, powered prosthesis control, HCI, and gaming. Sign language recognition in particular, includes subtle movements that either sensor is capable of consistently recognizing individually. This is validated by a study presented in [17], which stated that IMU sensor is particularly good for capturing larger hand and arm movements, while EMG data was better at distinguishing different hand shapes and finger movements. Many studies have been done on the significance that each sensor has on the system’s recognition accuracy, and

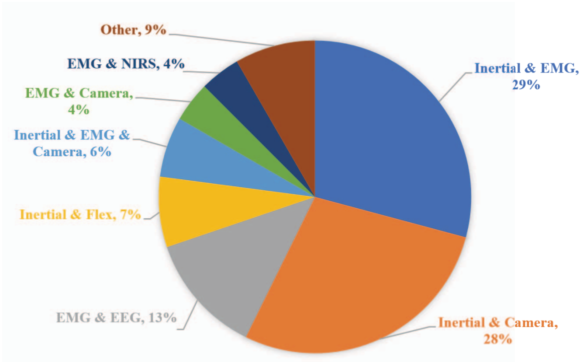

Fig. 1. Multimodal Sensor Fusion Types in Literature 2007-2017 at

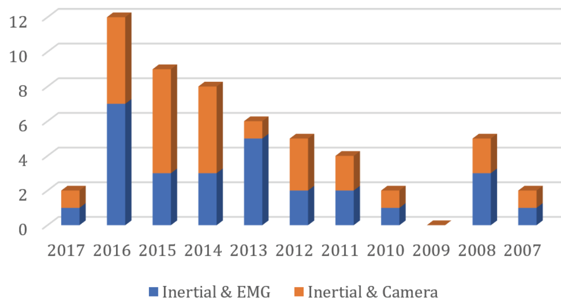

Fig. 2. Papers retrieved in the period January 2007 – February 2017 for the two most common combinations

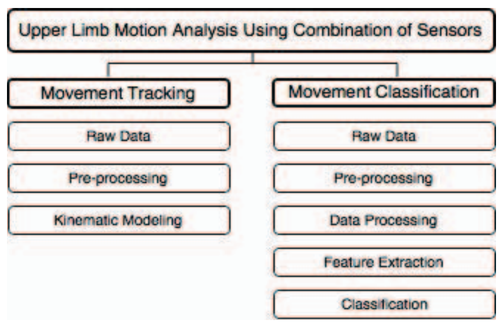

Fig. 3. Taxonomy of upper limb motion analysis

which point fusion become substantially viable. For example, it was found by Fougner et al [22], that for gesture recognition, to improve a two site EMG system, it is more beneficial to add an IMU than to add an extra EMG.

B. *Inertial and Visual Sensors Combination*

IMUs and vision sensors are both used for obtaining 3D orientation data making this combination to be the most adequate for movement tracking. Visual systems are known to produce accurate positional information [23], but due to motion blur and occlusion [24], performance is often lacking for fast and large movements. IMU can produce smooth motion data, however, drift problems and error accumulation degrade the accuracy. Therefore, even though inertial and visual sensors measure the same aspect of movements, since it has strengths and weaknesses in different situations, the fusion system will produce better accuracy in a complementary aspect [25]. The main drawbacks associated with current state-of-the-art optical systems is the cost and the portability. An alternative product that is a portable and inexpensive yet also accurate has been long awaited. The Microsoft Kinect, which consists of an RGB camera and a depth sensor, and the Leap Motion Controller, which is a combination of two infra-red cameras, are some of the commercially available products that meet this need. With the development of these cost-effective and simple visual systems, many studies exploiting them for body sensor networks were produced [24-29].

C. *Inertial and Flex Sensors Combination*

Flex sensors, which are passive analog resistive devices that convert the change in bend to an electrical resistance variation, is getting higher demands on its application of measuring physiological bending angles [30, 31]. However, since the output data is one-dimensional [31], a system using only flex sensors is not sufficient enough to recognize multiple different movements such as rotation or acceleration. IMUs on the other hand, are designed to measure the rotation and acceleration, which led to the fusion of the two systems to be applicable for movement classification.

In the field of hand gesture recognition, this combination is especially useful in providing lightweight yet accurate wearable sensors systems. Commercially available flex sensors are 0.5cm in width and 5~10 cm in length, which is small enough to be placed on each finger to detect the angle of finger bending. IMUs are also compact in size with the development of MEMS technology and can be placed on the back of the hand to detect the direction and magnitude of acceleration. Glove based sensor systems have therefore been seen to be a common application of this sensor combination [30, 32, 33], While it is possible to measure the angle with using two separate IMUs alone, as stated above, IMUs suffer from drift problems and are more sensitive to subtle vibrations, such as tremor, which does not affect flex sensors as much [34], thus giving this combination a particular advantage.

D. *EEG and EMG Sensors Combination*

Both EMG and EEG sensors measure what is known as a neurophysical phenomenon. That is to say, they seek to measure not the movement itself but the signaling process of the movement, in the muscle fiber and the motor cortex respectively. Although using EEG and EMG is much more complex in terms of measurement, considering their relatively high noise to signal ratios compared to sensors like accelerometers, it does however, offer a unique opportunity for some applications. Considering neither sensor measures inertial movement directly, they can therefore measure the intention of movement before the movement occurs. The potential of this sensors combination therefore lies heavily in the realm of assistive technologies for amputees, stroke patients, and patients affected by neuromuscular disorders, since in these cases a lack of limb or other physical impairment means that they may be unable to move their arm This is particularly highlighted in a survey and review of EEG-EMG based control approaches presented in [35], which confirms that the combination of EEG and EMG data can be particularly useful for applications which involve the measurement of muscle intention for amputees or people with movement impairments.

Appropriately, the majority of applications of this sensors combination are, at the moment, focused around the control of prosthesis for amputees [6, 36], human control interfaces designed to be controlled by amputees or patients with a movement impairment [37, 38], the control of robotic exoskeletons[12, 39], or tremor suppression systems with the assistance of functional electrical stimulation [40-42].

IV. Sensor Fusion Technique Comparison

A. *3D Movement Tracking Sensors*

Sensor fusion in movement tracking, is mainly the process of fusing the raw data achieved from different sensors for kinematic modeling. Therefore, this is possible with sensors that provide positional data such as IMUs, cameras and flex sensors. Out of the four main combinations studied in this paper, 3D tracking is possible with inertial and camera, or inertial and flex sensor. The fusion of the data will be done on the raw data, which is referred as data-level fusion, through mathematical and physical computation. Computation is done for each of the samples preventing information loss, which also means higher computational complexity and longer processing time [28]. Especially for multi-modal systems, each sensor has a different sampling rate and processes a different type of data, which means additional to the fusion algorithm, synchronization and data processing are also required.

B. *Movement Classification*

Movement classification is a multi-stage process as shown in Fig. 2, which is detailed as follows; Filtered data acquired after the pre-processing step of the raw data is divided into time segments in the data processing step. For each segment, time and/or frequency domain features are extracted which are used as input to classification algorithms. Compared to movement tracking, fusion can be done in different stages of the process. Data-level fusion comes after data processing step, feature-level fusion after feature extraction, and decision-level fusion is done after the classification.

Table I compares the type of fusion and the classification techniques used by a number of studies selected from the literature search. Classification accuracy was a criterion used as a basis of comparison, solely because this one has been used by a majority of papers. However, this has a number of limitations, as each paper seeks to classify features to a different number of gestures, and are also applied on a varied participant size. The values are therefore not directly comparable but give some indication into the success of certain methods. It was suggested by Novak et al. that studies should strive to test sensor fusion methods online using practical performance metrics rather than simply classification accuracy [21].

It is clear from Table 1 that the combinations explored are all able to achieve a high level of accuracy, although the IMU and EMG combination does stand out as one that is particularly successful. This may be due in part to the fact that this combination has the most research done on it, but also since the combination is quite optimal in that the sensors are well suited for gesture detection.

1) Data-level Fusion

Since data level fusion needs to handle the sampled raw data itself, high sampling rate sensors such as EMG and EEG are difficult to fuse at this stage. In the papers reported in this study, data level fusion was not used in the EMG and IMU combination nor for the EEG and EMG combination. For the inertial and visual sensors combination, the synchronized raw data were simply concatenated into a single vector and then classification techniques such as Hidden Markov Model (HMM) or Dynamic Time Warping (DTW) were used [23] [27]. In [27], comparison between HMM and DTW was carried out and concluded that HMM showed better classification accuracy. This is mainly due to the fact that while DTW is simple to implement and scale-invariant, it works only for small number of gestures [47]. For the inertial and flex sensor combination, the raw data is commonly used as the input for the classification by neural networks directly [45, 46].

2) Feature-level Fusion

As can be seen from Table I, feature level fusion is by far the most common technique used across multiple applications of upper limb movement recognition. Feature level fusion uses features rather than raw data for fusion for the simplicity of computation [3]. Feature sets are extracted from segmented data from each sensor and are combined to make a feature vector with higher dimension, and is used as the input of the classifier. Feature selection methods such as windowing technique, kernel discriminant analysis [48], minimal redundancy maximal relevance heuristic [49], and correlation based feature selection [50, 51]are often used for this process in the field of activity recognition [3]. Through this process it is possible to effectively use the features to obtain higher accuracy, without the need to process all raw data samples individually. The main drawback however, comes from the increase of dimensionality of the feature vector entering the classifier, which makes the classification more complex.

The success of feature level fusion depends upon the feature selection method, the feature concatenating method, and the classifiers. It is often difficult to predict the optimal method; thus, comparisons are commonly made between classifiers. For example Wu et al. [17] compared four common feature level fusion classifiers for American Sign language Recognition using IMU and EMG combination. They were Decision tree, Support Vector Machine (LibSVM), Nearest Neighbor and Naive Bayes. LibSVM showed the best performance in this case in terms of accuracy, precision, and recall. Also, compared to other classifiers, the testing time is

TABLE I. FUSION LEVEL AND CLASSIFICATION TECHNIQUE COMPARISON

Sensor Combination	*Criteria of Comparison*				Classification Accuracy		Reference
	Fusion Level	*Classification Technique*	*Application (Number of Gestures)*	*Wireless Protocol [a]*	*Uni-modal [b]*	*Multi-modal*	
EMG and IMU	Feature	Kernel Regularised Least Squares (KRLS)	Hand gesture recognition (40)	RF	EMG: 76% IMU: 81%	82.59%	[43]
	Feature	LibSVM	American sign language recognition (80)	Bluetooth	IMU: 92.29%	96.16%	[17]
	Feature	Hierarchical Decision Tree algorithm	Chinese sign language recognition (72)	-	-	96.3%	[44]
	Feature	Linear Discriminant Analysis (LDA)	Hand gesture recognition (8)	-	EMG: 94%	96%	[22]
Inertial and Camera	Data	Multi-Hidden Markov Model (Multi-HMM)	Hand gesture recognition (10)	Bluetooth	Inertial: 81% Camera: 76%	91%	[27]
	Data	Dynamic Time Warping (DTW)	Hand gesture recognition (10)	-	-	92.3%	[23]
	Decision	Collaborative Representation Classifier (CRC)	Body action recognition (27)	Bluetooth	Inertial: 88.3% Camera: 85.1%	97.2%	[25]
EEG and EMG	Feature	Linear Discriminant Analysis (LDA)	Hand gesture recognition (5)	-	EEG: 75.1% EMG: 77%	91.7%	[6]
	Decision	Gaussian classifier and Bayesian Fusion	Upper limb movement recognition (2)	-	EEG: 73% EMG: 87%	92%	[38]
Inertial and Flex	Data	Artificial Neural Networks (ANN)	American sign language recognition (25)	-	-	94%	[45]
	Data	Elmann Back Propagation Neural Networks (ENN)	Thai sign language recognition (16)	-	-	94.44%	[46]
	Feature	Gaussian Mixture Model (GMM)	Upper limb gesture recognition (6)	-	-	92.86%	[31]

[a] The dash means that wireless protocol was not used or mentioned for data communication.
[b] The dash means that the comparison experiment was not conducted.

not affected by the number of different gestures, which is crucial for real-time applications. Nearest Neighbor may have lower accuracy, and the testing time increases for larger training sets, but have the advantage of not needing a trained model. Decision trees have been used for sign language recognition [44, 52] [53, 54], since they are particularly robust for handling large amounts of data in a short time and straightforward to interpret. However, the performance of classifiers has significant difference depending on the application and features, therefore as mentioned in [17] it is suggested to experiment with multiple classifiers to find the most adequate one.

There are many commonly used features that are considered to be sufficient for classification. Most of the current studies rely on features that are easy to compute and provide satisfactory accuracy. However, a recent study was done to show the importance of optimizing the feature vector [17]. In this study, 268 different features from EMG and IMU data was ranked with a score from information gain criteria and concluded on a feature vector consisting of 40 features regarding the accuracy and computational constraints.

Feature level fusion such as concatenating two differing sets of features is simple and straightforward but faces some limitations as mentioned in [25]. Increased dimensionality of the feature vector leads to increased computational overhead, and each feature extracted may have different dimensions. For example, a depth feature vector from depth cameras usually has much higher dimensionality than inertial sensor features. The numerical scale difference of the features also needs to be considered where normalization technique should be applied. In [55], problems and possible solutions for feature fusion for inertial and depth sensors has been studied.

3) Decision-level Fusion

Decision level fusion is the process of selecting one final output out of the many classifier outputs from different sensors. This has the lowest computational complexity since each sensor can be processed with different algorithms and classified separately, as such it is the selection and combination of already classified data. It therefore has the advantage of saving the communication bandwidth and improved decision accuracy [20]. A disadvantage however, is that the compensation effect of data error cannot be resolved. For example, the drift effect of inertial sensors would not be able to be mitigated via point of reference with other sensory data until the decision process, in which case relevant information may have already been lost. There are several decision fusing techniques, including simple fusion, classical inference (summation, majority voting, board count, highest rank, logistic regression), voting and ensemble, boosting, Bayesian inference, and Dempster-Shafar's method [3, 20].

Sensor combinations which utilize related data of the same phenomenon such as the EEG/EMG sensors combination [38], and the Inertial/Visual sensors combination [27, 56], have been seen to be particularly suited to decision level fusion. The reason for this is that the differences in features between modalities is often minimal, considering they are measuring highly related phenomenon, therefore feature level fusion may not be as effective in utilizing the advantages of each modality. The technique used by [27] for combining depth and inertial data with decision level fusion was sending the data from the depth sensor and the inertial sensor into multiple HMM classifiers, then combining the probability outputs from the multiple HMM classifiers to generate the final outcome. A similar technique was used in the study by Leeb et al. [37] which involved the control of a 'hybrid' brain computer interface (hBCI). In the study, EEG and EMG data were processed and classified simultaneously in parallel with Gaussian classifiers, after which the decision level fusion of information was performed. The study explored the application of two different decision level fusion techniques; simple and Bayesian, and compared them. The fusion module used in the study received probability values from the two Gaussian classifiers, which were related to the confidence associated with each outputted class. It was found that Bayesian fusion had some advantages over simple fusion, such as being more robust to the effects of fatigue of the upper limb.

V. Discussion

As shown in this review, there are several advantages that the multi-modal fusion of sensors can offer, including the increase of spatial or temporal coverage, and the increase of classification accuracy. However, using sensor fusion can often increase computational load, and there are also a number of challenges involved with the synchronization of data.

A. Synchronization of Data

As can be seen from Table II, each sensor type exhibits a unique sampling frequency, which makes integrating these modalities together in one system a challenge, as it is crucial that the information is synchronous so that phenomenon can be correctly classified. Synchronization of EEG and EMG data is generally less of a problem as due to the similarity of the sensors, and that they can be sampled using the same sampling rate (1k - 2k Hz), and thus maintain synchronization more easily. When this is not the case however, for other sensor combinations a suggested and commonly used solution is to up sample or down sample the data of one modality so that its frequency matches that of the other one.

There are of course a number of limitations associated with this, for instance, EMG data commonly occurs with a 50-500 Hz frequency, thus, according to Nyquist-Shannon Theorem, if the data is sampled below a minimum of 1000 Hz there may be temporal data loss experienced. For the EMG/IMU combination, Kutafina et al. [57] used a down sampling technique on EMG data, but applied the down sampling to the feature vectors after feature extraction had been carried out on the raw data in order to preserve temporal information. In [27], a study which combines inertial data with the depth and RBG data from a Microsoft Kinect, inertial sensor data is down sampled. Whenever the Kinect signal is sampled through the Kinect SDK software at the rate of 30 Hz, a signal sample from the inertial sensor is collected at the same time. This method is

TABLE II. TYPICAL SAMPLING FREQUENCIES OF SENSORS

	Sensor				
	EMG	***IMU***	***EEG***	***Microsoft Kinect***	***Leap Motion***
Sampling frequency (Hz)	1k-2k	50-1k	1k-2k	30	115

common for such systems [20] [27]. Conversely, [43] used the up sampling technique of linear interpolation on accelerometer data in order to achieve a common frequency of 200Hz with EMG data. This was in contrast to what was reported by [57], who suggested that due to the nature of IMU sensors and the cumulative errors that are responsible for the drift effect, up sampling via interpolation can be an inaccurate and non-robust method. However, as the study [43] was for the control of an assistive exoskeleton, a high level of accuracy was needed, hence their decision to up sample.

Synchronization is also particularly an issue with wireless sensors, as over Bluetooth (which was the most common wireless protocol utilized in literature) there is typically a 5-20ms transmission delay, and if some modalities are wearable and wireless while others are external, synchronizing the data can be a tedious task. The study presented in [58] combines Microsoft Kinect and inertial sensors focused on this problematic and provided a method involving the synchronization of a PC to both inertial sensors and the Microsoft Kinect. The system obtained accurate results without having access to the video-depth cameras internal system clock. This method is known as time stamping, and has been utilized by a number of studies in the IMU/EMG combination as well [17, 52, 59].

B. Segmentation

The detection of movement onset and termination is incredibly important for classifying the motions, and particularly in real time systems, the automation of this is crucial. In the case of uni-modal systems, such as for accelerometer systems, a data windowing technique is often utilized, however this can make a system slow and not applicable for real time application. EMG data however, has been identified as being particularly useful for lightweight movement onset detection since muscles tend to relax between different gestures [44]. The onset detection can be done in a number of ways, including thresholds, for example in the studies [17] and [60], which used average energy calculations across multiple windows and moving average algorithms respectively. There are also examples of energy operators being used such as the Teager-Kaiser energy operator [61].

C. Data management

Usually, BSNs can be connected through a wireless system and connected to Internet, so that clinicians can benefit from the data online independently from the patient location [62]. As BSNs bring large data volumes, “the need to manage and maintain these datasets is of utmost importance” [63]. In this context, an emerging perspective is to manage these data in the domain of big data computing. Indeed, "big data presents a dramatic opportunity for reducing health disparities" [64],in this evolving era of patient centricity[65].

On top of big data computing, development of wireless protocol itself is also in the need for fast and concise wireless data transmission. As seen in Table 1, most wireless system uses Bluetooth due to its acceptable performance and cost effectiveness [17, 25, 27]. Regardless of its advantages, development of BSNs still suffer from critical power consumption issues, which is expected to be solved by the introduction of Bluetooth Low Energy (BLE). Several studies have proven the effectiveness of BLE for BSNs, claiming its low power consumption and robustness for data collision in high data traffic load situations [66, 67].

VI. Conclusion

This review offers an up-to-date summary and survey of the current technologies and techniques used in the field of upper limb area networks, especially the sensor fusion for upper limb movement measurement and detection, focusing on the most common combinations of sensors: EMG and inertial sensors, inertial and visual sensors, EEG and EMG sensors, and IMU and flex sensors. In Table I, it is clearly shown there is an increase in accuracy through using sensor fusion for movement recognition compared to uni-modal systems, highlighting the advantage and success of sensor fusion. It is important to realize, however, that each sensor combination has its own strengths and weaknesses, which should be taken into consideration when the choice of sensors is made for any specific application. As important as sensor combination choice is, the choice of sensor fusion technique is just as crucial, thus in this review a comparison of the level of fusion and the classification technique was conducted to provide groundwork for this purpose. Despite efforts to conduct a thorough extended search, there may be elements that are not covered in this review, therefore there may be a need for further consideration and research in this field.

Multi-sensor data fusion is a well-established research area, however, to achieve better performance through sensor fusion there still needs innovation not only on processing techniques, classification methods, and feature selection methods, but also by utilizing new sensors for different combinations. For example, strain sensors [68, 69] can cost effectively measure the surface deformation of skin and provide lightweight wearable sensor alternatives. While these sensors can be used successfully alone, there is potential for sensor fusion to improve these systems further.

Due to the increase in computational load that sensor fusion techniques are often associated with, for the continued development on real-time wireless BSNs both data management and wireless communication is a must [62]. In the aspect of data management, some studies suggested the application of cloud computing on BSNs for scalable storage and processing power. With the effectively processed data combined with a wireless system that can be synchronous across multiple nodes and be able to facilitate large amounts of data flow, as well as have low energy consumption would be ideal [70]. In this paper Bluetooth Low Energy was identified as solution commonly utilized, however as technology improves a more suitable medium may become evident.

Acknowledgment

The authors thank the International Relations Office of Polytechnique Montreal for their financial support.

References

[1] S. Movassaghi, M. Abolhasan, J. Lipman, D. Smith, and A. Jamalipour, "Wireless body area networks: A survey," *IEEE Communications Surveys & Tutorials,* vol. 16, no. 3, pp. 1658-1686, 2014.

[2] X. Lai, Q. Liu, X. Wei, W. Wang, G. Zhou, and G. Han, "A Survey of Body Sensor Networks," (in English), *Sensors,* vol. 13, no. 5, pp. 5406-5447, 2013.

[3] R. Gravina, P. Alinia, H. Ghasemzadeh, and G. Fortino, "Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges," *Information Fusion,* vol. 35, pp. 68-80, 5// 2017.

[4] H. H. Srebnik, *Concepts in Anatomy*. Springer US, 2012.

[5] C. L. M. S. Pulliam, J. M. M. S. Lambrecht, and R. F. P. Kirsch, "Electromyogram-based neural network control of transhumeral prostheses," (in English), *Journal of Rehabilitation Research and Development,* vol. 48, no. 6, pp. 739-54, 2011.

[6] X. Li, O. W. Samuel, X. Zhang, H. Wang, P. Fang, and G. Li, "A motion-classification strategy based on sEMG-EEG signal combination for upper-limb amputees," *Journal of NeuroEngineering and Rehabilitation,* journal article vol. 14, no. 1, p. 2, 2017.

[7] S. K. Manna and S. Bhaumik, "A Bioinspired 10 DOF Wearable Powered Arm Exoskeleton for Rehabilitation," *Journal of Robotics,* vol. 2013, p. 15, 2013, Art. no. 741359.

[8] A. López-Delis, A. F. Ruiz-Olaya, T.Freire-Bastos, and D. Delisle-Rodríguez, "A Comparison of Myoelectric Pattern Recognition Methods to Control an Upper Limb Active Exoskeleton," Berlin2013.

[9] N. A. Bhagat *et al.*, "Detecting movement intent from scalp EEG in a novel upper limb robotic rehabilitation system for stroke," in *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 4127-4130.

[10] Y. Ganesan, S. Gobee, and V. Durairajah, "Development of an Upper Limb Exoskeleton for Rehabilitation with Feedback from EMG and IMU Sensor," *Procedia Computer Science,* vol. 76, pp. 53-59, 2015/01/01 2015.

[11] M. A. Mikulski, "Electromyogram control algorithms for the Upper Limb Single-DOF Powered Exoskeleton," in *4th International Conference on Human System Interactions*, 2011, pp. 117-122.

[12] K. Kiguchi and Y. Hayashi, "An EMG-Based Control for an Upper-Limb Power-Assist Exoskeleton Robot," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* vol. 42, no. 4, pp. 1064-1071, 2012.

[13] Z. Tang, K. Zhang, S. Sun, Z. Gao, L. Zhang, and Z. Yang, "An Upper-Limb Power-Assist Exoskeleton Using Proportional Myoelectric Control," *Sensors,* vol. 14, no. 4, p. 6677, 2014.

[14] T. Forbes, "Mouse HCI through combined EMG and IMU," 1536047 M.S., University of Rhode Island, Ann Arbor, 2013.

[15] G. R. Naik, D. K. Kumar, H. Weghorn, and M. Palaniswami, "Subtle Hand Gesture Identification for HCI Using Temporal Decorrelation Source Separation BSS of Surface EMG," in *9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications*, 2007, pp. 30-37.

[16] S. Berman and H. Stern, "Sensors for Gesture Recognition Systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 42, no. 3, pp. 277-290, 2012.

[17] J. Wu, L. Sun, and R. Jafari, "A Wearable System for Recognizing American Sign Language in Real-Time Using IMU and Surface EMG Sensors," *IEEE Journal of Biomedical and Health Informatics,* vol. 20, no. 5, pp. 1281-1290, 2016.

[18] A. Abdelgawad and M. Bayoumi, "Data Fusion in WSN," in *Resource-Aware Data Fusion Algorithms for Wireless Sensor Networks*Boston, MA: Springer US, 2012, pp. 17-35.

[19] J. Mendes, M. Vieira, M. Pires, and S. Jr, "Sensor Fusion and Smart Sensor in Sports and Biomedical Applications," vol. 16, ed. Basel: MDPI AG, 2016, p. 1569.

[20] C. Chen, R. Jafari, and N. Kehtarnavaz, "A survey of depth and inertial sensor fusion for human action recognition," *Multimedia Tools and Applications,* vol. 76, no. 3, pp. 4405-4425, 2017// 2017.

[21] D. Novak and R. Riener, "A survey of sensor fusion methods in wearable robotics," *Robotics and Autonomous Systems,* vol. 73, pp. 155-170, 11// 2015.

[22] A. Fougner, E. Scheme, A. D. C. Chan, K. Englehart, and O. Stavdahl, "A multi-modal approach for hand motion classification using surface EMG and accelerometers," in *33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, pp. 4247-4250.

[23] S. Zhou *et al.*, "2D Human Gesture Tracking and Recognition by the Fusion of MEMS Inertial and Vision Sensors," *IEEE Sensors Journal,* vol. 14, no. 4, pp. 1160-1170, 2014.

[24] Y. Tian, X. Meng, D. Tao, D. Liu, and C. Feng, "Upper limb motion tracking with the integration of IMU and Kinect," *Neurocomputing,* vol. 159, pp. 207-218, 7/2/ 2015.

[25] C. Chen, R. Jafari, and N. Kehtarnavaz, "A Real-Time Human Action Recognition System Using Depth and Inertial Sensor Fusion," *IEEE Sensors Journal,* vol. 16, no. 3, pp. 773-781, 2016.

[26] L. Kui, C. Chen, R. Jafari, and N. Kehtarnavaz, "Fusion of Inertial and Depth Sensor Data for Robust Hand Gesture Recognition," *IEEE Sensors Journal,* vol. 14, no. 6, pp. 1898-903, 06/ 2014.

[27] K. Liu, C. Chen, R. Jafari, and N. Kehtarnavaz, "Multi-HMM classification for hand gesture recognition using two differing modality sensors," in *10th IEEE Dallas Circuits and Systems Conference*, Richardson, TX, United states, 2014, p. IEEE Circuits and Systems (CAS) Society; IEEE Circuits and Systems Dallas Chapter: Institute of Electrical and Electronics Engineers Inc.

[28] A. Atrsaei, H. Salarieh, and A. Alasty, "Human Arm Motion Tracking by Orientation-Based Fusion of Inertial Sensors and Kinect Using Unscented Kalman Filter," *Journal of Biomechanical Engineering,* vol. 138, no. 9, pp. 091005-091005-13, 2016.

[29] Y. Guo, D. Tao, W. Liu, and J. Cheng, "Multiview Cauchy Estimator Feature Embedding for Depth and Inertial Sensor-Based Human Action Recognition," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. PP, no. 99, pp. 1-11, 2016.

[30] V. Pathak, S. Mongia, and G. Chitranshi, "A framework for hand gesture recognition based on fusion of Flex, Contact and accelerometer sensor," in *3rd International Conference on Image Information Processing*, District Solan, Near Shimla, Himachal Pradesh, Waknaghat, India, 2015, pp. 312-319: Institute of Electrical and Electronics Engineers Inc.

[31] Y. Wang, "Recognition and Classification of the Wolf Motor Function Test Items using Multimode Sensor Fusion," 1511228 M.S., University of California, Los Angeles, 2012.

[32] N. Tongrod, S. Lokavee, T. Kerdcharoen, N. Watthanawisuth, and A. Tuantranont, "Gestural system based on multi-functional sensors and ZigBee networks for squad communication," in *Defense Science Research Conference and Expo*, Singapore, 2011: IEEE Computer Society.

[33] A. Nelson *et al.*, "Wearable multi-sensor gesture recognition for paralysis patients," in *2013 IEEE Sensors, 3-6 Nov. 2013*, Piscataway, NJ, USA, 2013, p. 4 pp.: IEEE.

[34] A. Nelson *et al.*, "Wearable multi-sensor gesture recognition for paralysis patients," in *IEEE SENSORS*, 2013, pp. 1-4.

[35] T. D. Lalitharatne, K. Teramoto, Y. Hayashi, and K. Kiguchi, "Towards Hybrid EEG-EMG-Based Control Approaches to be Used in Bio-robotics Applications: Current Status, Challenges and Future Directions," *Journal of Behavioral Robotics,* vol. 4, no. 2, pp. 147-154, 2013.

[36] K. Kiguchi, T. D. Lalitharatne, and Y. Hayashi, "Estimation of Forearm Supination/Pronation Motion Based on EEG Signals to Control an Artificial Arm," *Journal of Advanced Mechanical Design, Systems, and Manufacturing,* vol. 7, no. 1, pp. 74-81, 2013.

[37] R. Leeb, H. Sagha, R. Chavarriaga, and J. Del R Millan, "Multimodal fusion of muscle and brain signals for a hybrid-BCI," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* pp. 4343-4346, 2010.

[38] R. Leeb, H. Sagha, R. Chavarriaga, and J. d. R. Millán, "A hybrid brain–computer interface based on the fusion of electroencephalographic and electromyographic activities," *Journal of Neural Engineering,* vol. 8, no. 2, p. 025011, 2011.

[39] T. Kawase, T. Sakurada, Y. Koike, and K. Kansaku, "A hybrid BMI-based exoskeleton for paresis: EMG control for assisting arm movements," *Journal of Neural Engineering,* vol. 14, no. 1, p. 016015 (12 pp.), 02/ 2017.

[40] P. Aricò *et al.*, "FES controlled by a hybrid BCI system for neurorehabilitation–driven after stroke," *GNB,* pp. 3182-5, June 26th-29th 2012.

[41] F. Cincotti *et al.*, "EEG-based Brain-Computer Interface to support post-stroke motor rehabilitation of the upper limb," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 4112-4115.

[42] E. Rocon *et al.*, "Multimodal BCI-Mediated FES Suppression of Pathological Tremor," in *Annual International Conference of the Ieee Engineering in Medicine and Biology Society*(IEEE Engineering in Medicine and Biology Society Conference Proceedings, 2010, pp. 3337-3340.

[43] A. Gijsberts and B. Caputo, "Exploiting accelerometers to improve movement classification for prosthetics," ed, 2013, pp. 1-5.

[44] V. Xu Zhang, V. Xiang Chen, V. Yun Li, V. Lantz, V. Kongqiao Wang, and V. Jihai Yang, "A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans,,* vol. 41, no. 6, pp. 1064-1076, 2011.

[45] M. E. Cabrera, J. M. Bogado, L. Fermin, R. Acuna, and D. Ralev, "Glove-based gesture recognition system," in *Proceedings of international conference on climbing and walking robots and the support technologies for mobile machines*, 2012, pp. 747-753.

[46] S. Saengsri, V. Niennattrakul, and C. A. Ratanamahatana, "TFRS: Thai finger-spelling sign language recognition system," in *Digital Information and Communication Technology and it's Applications (DICTAP), 2012 Second International Conference on*, 2012, pp. 457-462.

[47] A. Stefan, V. Athitsos, J. Alon, and S. Sclaroff, "Translation and scale-invariant gesture recognition in complex scenes," presented at the Proceedings of the 1st international conference on PErvasive Technologies Related to Assistive Environments, Athens, Greece, 2008.

[48] L. Gao, A. K. Bourke, and J. Nelson, "A system for activity recognition using multi-sensor fusion," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 7869-7872: IEEE.

[49] S. Liu, R. X. Gao, D. John, J. W. Staudenmayer, and P. S. Freedson, "Multisensor data fusion for physical activity assessment," *IEEE Transactions on Biomedical Engineering,* vol. 59, no. 3, pp. 687-696, 2012.

[50] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher, "Activity recognition and monitoring using multiple sensors on different body positions," in *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*, 2006, pp. 4 pp.-116: IEEE.

[51] N. A. Capela, E. D. Lemaire, and N. Baddour, "Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients," *PloS one,* vol. 10, no. 4, p. e0124414, 2015.

[52] J. Wu, Z. Tian, L. Sun, L. Estevez, and R. Jafari, "Real-time American Sign Language Recognition using wrist-worn motion and surface EMG sensors," in *IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, Piscataway, NJ, USA, 2015, pp. 1-6: IEEE.

[53] J.-X. Tian, X. Chen, Y. Li, and J.-H. Yang, "A Chinese Sign Language Recognition Approach Based on Multiple Sensor Information Fusion and Statistical Language Model," *Chinese Journal of Biomedical Engineering,* vol. 30, no. 3, pp. 333-9, 06/20 2011.

[54] Y. Li, X. Chen, J. Tian, X. Zhang, K. Wang, and J. Yang, "Automatic recognition of sign language subwords based on portable accelerometer and EMG sensors," in *1st International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, Beijing, China, 2010, pp. ACM Special Interest Group on Computer-Human Interaction (SIGCHI): Association for Computing Machinery.

[55] Y. Li, J. Cheng, W. Feng, and D. Tao, "Feature fusion of triaxial acceleration signals and depth maps for human action recognition," in *IEEE International Conference on Information and Automation (ICIA)*, 2016, pp. 1255-1260.

[56] Y. Yamazaki *et al.*, "Gestures Recognition Using Combination of Acceleration Sensor and Images for Casual Communication between Robots and Humans," *IEEE Congress on Evolutionary Computation (Cec),* 2010.

[57] E. Kutafina, D. Laukamp, R. Bettermann, U. Schroeder, and S. Jonas, "Wearable Sensors for eLearning of Manual Tasks: Using Forearm EMG in Hand Hygiene Training," *Sensors,* vol. 16, no. 8, p. 1221, 2016.

[58] E. Cippitelli *et al.*, "Time synchronization and data fusion for RGB-Depth cameras and inertial sensors in AAL applications," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, 2015, pp. 265-270.

[59] L. Saracino, E. Ruffaldi, A. Graziano, and C. A. Avizzano, "Fusion of wearable sensors and mobile haptic robot for the assessment in upper limb rehabilitation," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, Piscataway, NJ, USA, 2016, pp. 66-71: IEEE.

[60] G. Fortino, S. Galzarano, R. Gravina, and W. Li, "A framework for collaborative computing and multi-sensor data fusion in body sensor networks," *Information Fusion,* vol. 22, pp. 50-70, 3// 2015.

[61] S. Solnik, P. Rider, K. Steinweg, P. DeVita, and T. Hortobágyi, "Teager–Kaiser energy operator signal conditioning improves EMG onset detection," *European Journal of Applied Physiology,* journal article vol. 110, no. 3, pp. 489-498, 2010.

[62] J. Y. Y. Khan, M. R. , *Wireless body area networks: Technology, implementation, and applications*. Pan Stanford Publishing, 2011.

[63] J. O'donoghue and J. Herbert, "Data management within mHealth environments: Patient sensors, mobile devices, and databases," *Journal of Data and Information Quality (JDIQ),* vol. 4, no. 1, p. 5, 2012.

[64] X. Zhang *et al.*, "Big Data Science: Opportunities and Challenges to Address Minority Health and Health Disparities in the 21st Century," *Ethnicity & Disease,* vol. 27, no. 2, pp. 95-106, 2017.

[65] S. Ray, "In sickness and in health: Clinical research and social media," *Perspectives in Clinical Research,* vol. 8, no. 2, p. 73, 2017.

[66] T. Wu, F. Wu, J. M. Redouté, and M. R. Yuce, "An Autonomous Wireless Body Area Network Implementation Towards IoT Connected Healthcare Applications," *IEEE Access,* vol. 5, pp. 11413-11422, 2017.

[67] J. A. Afonso, A. J. F. Maio, and R. Simoes, "Performance Evaluation of Bluetooth Low Energy for High Data Rate Body Area Networks," *Wireless Personal Communications,* journal article vol. 90, no. 1, pp. 121-141, September 01 2016.

[68] J.-B. Chossat, Y. Tao, V. Duchaine, and Y.-L. Park, "Wearable soft artificial skin for hand motion detection with embedded microfluidic strain sensing," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, 2015, pp. 2568-2573: IEEE.

[69] C. Zizoua, M. Raison, S. Boukhenous, M. Attari, and S. Achiche, "Development of a Bracelet With Strain-Gauge Matrix for Movement Intention Identification in Traumatic Amputees," *IEEE Sensors Journal,* vol. 17, no. 8, pp. 2464-2471, 2017.

[70] G. Fortino, M. Pathan, and G. D. Fatta, "BodyCloud: Integration of Cloud Computing and body sensor networks," in *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, 2012, pp. 851-856.