Recognising Irish Sign Language Using Electromyography

Laura Cristina Galea
Dublin City University
Glasnevin, Dublin 9,
Ireland

Alan F. Smeaton
Insight Centre for Data Analytics
Dublin City University, Glasnevin, Dublin 9, Ireland
Alan.Smeaton@dcu.ie

***Abstract*—Sign language is the non-verbal communication used by people with hearing and speaking impairments. The automatic recognition of sign languages is usually based on video analysis of the signer though this is difficult when considering different light levels or the surrounding environment. The work in this paper uses electromyography (EMG) and focuses on letters of the Irish Sign Language (ISL) alphabet. EMG is the recording of the electrical activity produced to stimulate movement in the skeletal muscles. We capture muscle signals and inertial movement data using the Thalmic MYO armband and, in real time, recognise the ISL alphabet. Our implementation is based on signal processing, feature extraction and machine learning. The only input required to translate the ISL gestures are EMG and movement data, thus our approach is usable in scenarios where using video for automatic recognition video is not possible.**

***Index Terms*—Sign language recognition, EMG, inertial movement, machine learning**

I. Introduction

Sign Language (SL) is a method for communication between people based on gestures and signs, and primarily used by the deaf and hard of hearing. Sign languages have their own grammar and syntax and are thus full natural languages however they are not universal and there is no universal SL [1]. Instead, there are literally hundreds of sign languages, some existing at national level like American, British and Irish Sign Languages, and others being more local.

Irish Sign Language (ISL) [2] is Ireland's indigenous sign language used by approximately 6,500 deaf people and 65,000 hearing signers across the country. ISL is an official language of Ireland, and recognised as being so by statute. It has an alphabet for the 26 characters, as well as a number of other signs for commonly used words like prepositions, days of the week, colours, weather, etc. Many of the non-alphabet signs are two-handed but the alphabet can be signed one-handed. Conversations using only the alphabet are thus more tedious and slow, but that is what we focus on here precisely because it is one-handed and the vocabulary of 26 characters is limited. The letter alphabet used in ISL is shown in Figure 1.

In this paper we address the challenge of automatically recognising ISL in real time, to a level of accuracy which is good enough for conversations to take place between people.

AS is funded by Science Foundation Ireland under grant number SFI/12/RC/2289

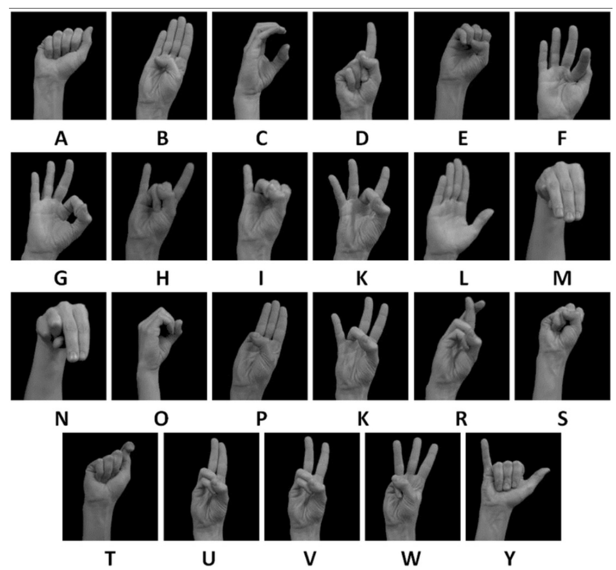

Fig. 1. Letter alphabet used in Irish Sign Language

The use cases for this are to support communication between those who know ISL and those who do not or who are learning. Our approach is based on real time monitoring of electromyography (EMG) in the forearm and in the rest of the paper we give a summary of automatic approaches to recognising ISL, an introduction to EMG, and a description of the technique we developed, including an evaluation of its accuracy.

II. Automatic Recognition of Irish Sign Language

Automatic recognition of sign language can use a range of approaches including those shown in Figure 2.

The main efforts have been in using image/video processing and examples of that work are described in [3], [4]. The work in [4] compared different approaches for ISL recognition and performed experiments and reported comparative accuracy and timing information. In particular, that work focused on the real world scenario where images are blurred as a result of the social setting (movement, lighting, etc.). That work is typical of modern approaches in that it uses Convolutional Neural Networks (CNN) and feature based extraction approaches, such

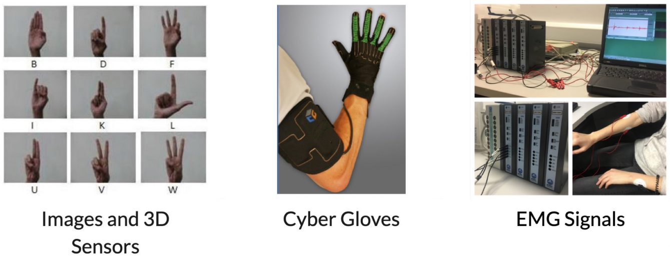

Fig. 2. Approaches to recognising ISL

as Principal Component Analysis (PCA) followed by different classifiers, e.g. multilayer perceptron (MLP). That approach obtains a recognition accuracy over 99% which establishes a performance baseline for image-based recognition.

Taking this further and to further promote work in this area, [5] introduced an image dataset for Irish Sign Language (ISL) recognition of subjects performing ISL hand-shapes and movements, resulting in 468 videos. In addition to the dataset, the authors report experiments using Principal Component Analysis (PCA), reaching 95% recognition accuracy.

Image/vision approaches are not the only innovative ways of achieving ISL recognition. In [6] there was a special issue of the journal *Universal Access in the Information Society* on using avatars in SL recognition, a collected volume based on presentations given at the symposium *Sign Language Translation and Avatar Technology (SLTAT)* held in 2013. That included a paper [7] where the authors explored the effect of adding facial expressions which reveal emotional clues, into ISL recognition. They augmented an existing avatar for displaying ISL with basic universal emotions leading to improved recognition.

These approaches all support use cases where using a video camera to record, analyse and interpret SL is beneficial yet there are scenarios where it is not socially acceptable or possible to use image or video. Signing in private conversations in a public place is always welcome but using a camera to interpret may not be socially acceptable, or the lighting may be poor, or the setting may not allow it to be used in places like a crowded commuter train, for example.

In such cases we offer an alternative based on using electromyography (EMG) which is non-intrusive, can be used in any environment or setting, and overcomes many of the obstacles to recognising sign language automatically.

III. Electromyography

Electromyography (EMG) is a technique for recording the electrical activity which the brain produces and sends to the skeletal muscles in order to get them to contract, thus causing movement. EMG can capture these signals using electrodes placed on the surface of the skin thus it is painless and non-invasive. EMG electrodes detect the electrical signals, which are in the range of millivolts, and amplify and digitise them allowing the signal to be processed in real time.

In this work we capture muscle signals and inertial movement data using the Thalmic MYO armband, a wearable device shown in Figure 3 and described in [8]. The MYO

Fig. 3. The MYO device

is a forearm gesture recognition device that senses EMG or electrical activity in the forearm muscles and also has a built-in inertial measurement unit (IMU). It generates a continuous stream of EMG and IMU data with EMG over 8 channels from 8 electrodes placed in a ring around the forearm. It uses Bluetooth to transmit to a laptop or PC. The sampling rates for MYO data are fixed at 200 Hz for EMG and 50 Hz for the inertial sensors. In recognising ISL we particularly focus on movement of the hand and fingers triggered by the extensor digitorum muscle, from the posterior forearm, and the flexor carpi ulnaris muscle from the anterior forearm. These particular muscles are mainly responsible for the movements of the hand, and are shown in Figure 4.

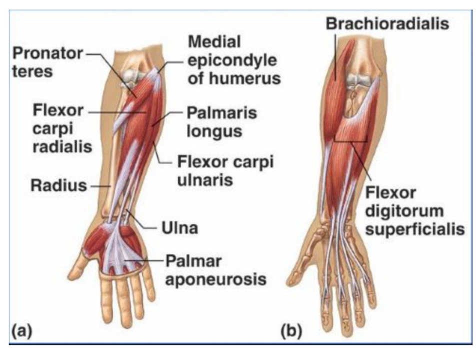

Fig. 4. Arm muscles

We use the EMG and IMU signals from MYO to interpret the 26 letters of the ISL alphabet in real time. In the next section we present the approach we took, the challenges we were faced with and how we overcame those.

IV. Recognising ISL using EMG

Movement of the hand and fingers to generate any ISL sign is ultimately caused by a combination of the many muscles in the forearm. The combination of muscle movements, triggered by electrical signals from the brain, causes subtle changes in finger positions and movement of the wrist. The MYO armband has 8 EMG channels, each sampled at 200 Hz, so signing different ISL letter will be realised by different combinations of muscle movement and the MYO's 8 EMG channels will pick up different signal strength combinations corresponding to different muscle triggers. The basis of our approach is to extract features from the stream of EMG and

IMU data and use these features to train a model to recognise each of the 26 letters in ISL.

Extracting the right features plays an important role in terms of the accuracy of the resulting models. Failing to do this correctly will impact negatively on model performance. Finding the right machine learning classifier and extracting the right feature combinations from the raw data was thus crucial.

As training data we recorded 5 example 3-second bursts for each letter using 12 different subjects in order to get variety in our training data set.

For each 3 second record the features extracted for each sign are:

- Mean Absolute Value (MAV) — measures the activities of muscles;
- Modified Mean Absolute Value (MMAV);
- Simple Square Integral (SSI) — measures the energy;
- Root Mean Square (RMS) — measures the activity of muscles;
- Average Amplitude Change (AAC) — measures average of the amplitude change in signal;
- Variance;
- Minimum;
- Maximum;
- Standard Deviation;
- Integrated Absolute Value;
- Waveform Length.

Given there are 26 letters in the ISL alphabet, each letter is represented as a stream of 8 EMG values at 200 Hz, 3 gyroscope, 3 accelerometer and 3 orientation values at 50 Hz, so in total 17 data streams. From each of these 17 sets of values 11 features were extracted for each 3 second recording (for each letter), giving a total of 187 features for each letter. These features were ranked in order of how important they are in distinguishing between the recognition of the 26 letters, and some were eliminated as described below.

For improved results we had to customise our own feature selection method. For this we use the "feature_importance" function from the Random Forest Classifier. This function ranks the features of the data in terms of how important they are for recognition of the letters. Some of the attributes received a discrimination score of zero or close to zero and these were removed. A total of 157 features remained. Further analysis was done to find the optimal number of features and after feature reduction, the optimal number of features was 140.

At the start of our experiments, only EMG data was used to train the models. Adding the IMU data increased the accuracy of the models by 10%. Some of the letters involve a lot of IMU data as the hand is required to do a lot of movement in, for example, the letters "M", "N", "J" and "Z". Therefore, IMU data play a higher role for these letters than the others.

We used the scikit-learn toolkit and tried a range of machine learning techniques from simple linear regression (which we quickly discarded) to Naive Bayes, random forest, ensemble and support vector machines. We experimented with optimising the models' performance and obtained an accuracy of 78%. The highest accuracy was given by models based on random forest and ensemble models. A confusion matrix for ISL letter confusion is shown in Figure 5.

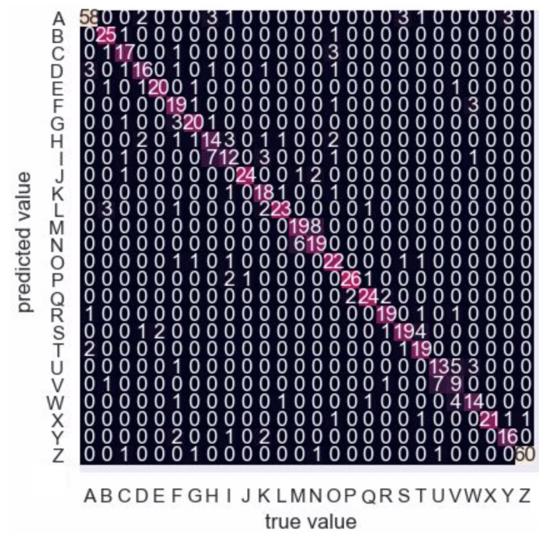

Fig. 5. Confusion matrix for ISL letter recognition

V. Analysis of Recognition of Individual Letters

We tried to improve performance by targeting individual letters, some of which require more complex movement of the fingers and wrist and thus modelling them might require different features compared to more straightforward letters. We were also interested to see if our users are consistent in their EMG patterns so in this analysis we set out to find if any of the users are different.

We generated visualisations for each of the 26 letters based on the 11 EMG features for 12 all users, which we examined to see if any discriminating feature characteristics could be found. Visualisations were generated per feature namely "MAV", "MMAV", "RMS", "IAV", "SSI", "AAC", "VAR", "MIN", "MAX", "STD" and "WL". An example of this, for the letter "A" for the features "MIN" and "WL" is shown in Figure 6.

What we found for the letter A (and this was similar for all letters), is that the waveform length (WL), has no discriminative function for the letter A as all the user values are clustered into one area of the graph. On the other hand the minimum feature (MIN), has values scattered all over the graph. Other features are somewhere in between these two extremes. What were are looking for were clusters of colours grouped together corresponding to different users corresponding to outlier users whose training data, might skew the overall recognition performance. On examining all 11 features for all 26 letters, we did not find any, telling us that there is consistency across users and that all features are important discriminators for all letters.

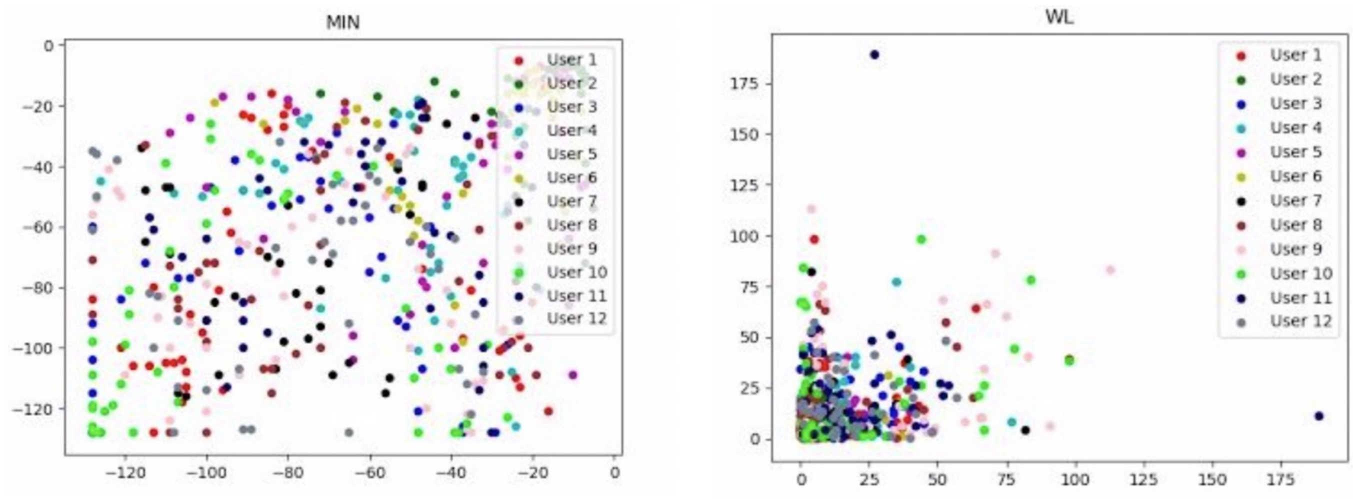

Fig. 6. Plots for the letter A

VI. Conclusions

In this paper we have described a system for performing real time recognition of Irish Sign Language with an overall accuracy of 78%. At the CBMI conference we will demonstrate this in real time with a researcher signing the A..Z alphabet and a computer displaying the letter signed. When this level of performance is combined with predictive text and autocomplete, which are now established features of data entry in smartphones and search query input boxes, this offers a realistic alternative to vision-based automatic recognition of sign language.

Using EMG and IMU data to recognise ISL presented many unanticipated challenges. For example, we doscovered that EMG signals may differ from person-to-person. Also, the size of the forearm differs from person to person as well as the physiology of their arm muscles. This means that the ideal positioning of the MYO in order to pick up EMG signals from the most appropriate of the forearm muscles shown in Figure 4, will also vary from person-to-person. The strength of the muscles will also differ from person to person so the strength of the EMG reading will vary, though this can be addressed by normalising. All these factors impact the regularity and consistency of EMG signal data. We also discovered from our own experiemnts that EMG signal data varies not only from person-to-person but also from hour to hour because of the level of muscle fatigue. What all this means is that training a machine learning model to recognise ISL will need to be attuned to each individual and even to adjust to that individual over time, so it has to learn to update its learned model as it is being used. This is a topic for future work.

References

[1] C. Lucas, *The sociolinguistics of sign languages*. Cambridge University Press, 2001.

[2] L. Leeson, "Moving Heads and Moving Hands: Developing a Digital Corpus of Irish Sign Language," in *Information Technology and Telecommunications Conference*, October 2006, pp. 25–26.

[3] D. Kelly, J. Reilly Delannoy, J. Mc Donald, and C. Markham, "A framework for continuous multimodal sign language recognition," in *Proceedings of the 2009 International Conference on Multimodal Interfaces*, ser. ICMI-MLMI '09. New York, NY, USA: ACM, 2009, pp. 351–358. [Online]. Available: http://doi.acm.org/10.1145/1647314.1647387

[4] M. Oliveira, H. Chatbri, S. Little, Y. Ferstl, N. E. O'Connor, and A. Sutherland, "Irish sign language recognition using principal component analysis and convolutional neural networks," in *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. IEEE, 2017, pp. 1–8.

[5] M. Oliveira, H. Chatbri, Y. Ferstl, M. Farouk, S. Little, N. E. O'Connor, and A. Sutherland, "A dataset for irish sign language recognition," in *Irish Machine Vision and Image Processing Conference (IMVIP)*, 2017.

[6] R. Wolfe, E. Efthimiou, J. Glauert, T. Hanke, J. McDonald, and J. Schnepp, "Special issue: recent advances in sign language translation and avatar technology," *Universal Access in the Information Society*, vol. 15, no. 4, pp. 485–486, Nov 2016. [Online]. Available: https://doi.org/10.1007/s10209-015-0412-5

[7] R. G. Smith and B. Nolan, "Emotional facial expressions in synthesised sign language avatars: a manual evaluation," *Universal Access in the Information Society*, vol. 15, no. 4, pp. 567–576, 2016.

[8] M. Sathiyanarayanan and S. Rajan, "Myo armband for physiotherapy healthcare: A case study using gesture recognition application," in *2016 8th International Conference on Communication Systems and Networks (COMSNETS)*. IEEE, 2016, pp. 1–6.