Journal & Proceedings of the Royal Society of New South Wales, Vol. 140, p. 47–54, 2007
ISSN 0035-9173/07/020047–8 $4.00/1

Applications and Limitations of Independent Component Analysis for Facial and Hand Gesture Surface Electromyograms

GANESH R. NAIK AND DINESH K. KUMAR

Abstract: In the recent past, there has been an increasing trend to use blind source separation (BSS) or independent component analysis (ICA) algorithms for biomedical data. This paper reviews the concept of ICA and demonstrates its usefulness and limitations in the context of surface electromyograms (sEMG) related to hand movements and facial muscles. In the first experiment ICA has been used to separate the electrical activity from different hand gestures. The second part of the study considers separating electrical activity from facial muscles. In both instances the surface electromyogram has been used as an indicator of muscle activity. The theoretical analysis and experimental results demonstrate that ICA is suitable for identification of different hand gestures using sEMG signals. The results identify the unsuitability of ICA when a similar technique is used for facial muscles with respect to different vowel classifications.

Keywords: Independent component analysis, surface electromyogram, motor unit action potentials, human-computer interaction, blind source separation

INTRODUCTION

Independent component analysis (ICA) has recently received a lot of attention both in biomedical signal processing and statistical signal processing. Independent component analysis is a useful method for blind source separation (BSS) and unsupervised learning, where the observation vectors are assumed to be linear mixing of independent components. Efficient new ICA algorithms have been introduced to solve the blind source separation problems. ICA algorithms are successfully utilised for removing artefact and noise from recorded biosignals, especially sEMG. Research that isolates motor unit action potential (MUAP) originating from different muscles and motor units has been reported in 2004 (Nakamura et al. 2004), where success is reported in the isolation of the different MUAP with applications for decomposing the sEMG at low levels of muscle activation. ICA has also been proposed for unsupervised cross talk removal from sEMG recordings of the muscles of the hand (Greco et al. 2003). Recently ICA had been utilised to identify different hand gestures (Naik et al. 2006). From the literature, ICA appears to be the emerging technology with solutions to most of the sEMG applications. Myo-electric activity originating from different muscles can be considered to be independent, making the use of ICA a suitable method for the separation of muscle activity originating from different muscles.

Surface Electromyography (sEMG) is a surface recording of muscle activity. It is a result of the spatial and temporal integration of the MUAP originating from different motor units. Being non-invasive and an important indicator of muscle activity, sEMG is useful, but the presence of multiple muscle activity and the random nature of the transmission path makes the signal difficult to use reliably when muscle activity is small, and actions are complex. It is difficult to separate muscle activity originating from different muscles due to similarity in the signals. Earlier work by the authors has used sEMG to identify unspoken vowels with success (Kumar et al. 2004), but reliability issues exist. This paper reports research conducted to identify the applications of ICA in hand gesture identification, and to determine some of the limitations of using ICA for separation of sEMG from facial muscle activity originating from other muscles (cross talk). The paper reports theoretical analysis and experimental results and discusses the apparent discrepancies between the two.

FACIAL MOVEMENT AND MUSCLES RELATED TO SPEECH

The face can communicate a variety of information including subjective emotion, communitive intent, and cognitive appraisal. The facial musculature is a three-dimensional assembly of small pseudo-independently controlled muscles performing a variety of complex or facial functions such as speech, mastication, swallowing and mediation of motion (Lapatki et al. 2003, Parsons 1986).

When using facial sEMG to determine the shape of the lips and the mouth, there arises the issue of the proper choice of muscles and their corresponding best location of electrodes. Face structure is more complex than that of the limbs due to a large number of overlapping muscles. It is thus difficult to identify the specific muscles that are responsible for specific facial actions and shapes. There is also the difficulty of cross talk due to the overlap between different muscles. This is made more complex because of the temporal variation in the activation and deactivation of the different muscles. It is impractical to consider the entire facial musculature and record its electrical activity. In this study, only four facial muscles have been selected. The *Zygomaticus major* arises from the front surface of the zygomatic bone and merges with the muscles at the corner of the mouth. The *Depressor anguli oris* originates from the mandible and inserts into the skin at an angle to the mouth and pulls the corner of mouth downward. The *Masseter* originates from the maxilla and zygomatic arch and inserts to the ramus of the mandible to elevate and protrude; it assists in side-to-side movements of the mandible. The *Mentalis* originates from the mandible and inserts into the skin of the chin to elevate and protrude the lower lip, to pull skin into a pout (Fridlund and Cacioppo 1996).

SURFACE EMG AND INDEPENDENT COMPONENT ANALYSIS

The EMG signal is widely used as a suitable means to have access to physiological processes involved in producing joint movements. The information extracted from the EMG signals can be exploited in several different applications. sEMG is a non-invasive and painless procedure in which EMG signals are measured from electrodes on the skin. This technique has clear advantages over needle EMG. Most importantly it is painless for the patient and avoids health hazards for patient and doctor. Furthermore, sEMG is a quick and easy process that facilitates sampling of a large number of MUPs (Fujimoto and Nishizono 1993). One major barrier is that, due to the wide pickup area of surface electrodes, sEMG waveforms exhibit significant interference. Surface EMG recordings provide a practical means to record from several muscles simultaneously but tend to be unreliable, i.e., recordings from a subject performing the same movement repetitively tend to have considerable trial-to-trial variability. sEMG recordings are also affected by cross-talk whereby several muscles may contribute to the recording of a given electrode, making the source of the signal difficult to identify. Recently, ICA has been proposed as a method to analyze sEMG recordings, and this addresses many of these concerns. One property of the sEMG is that the signal originating from one muscle can generally be considered to be independent of other bioelectric signals such as an electrocardiogram (ECG), electro-oculargram (EOG), and signals from neighbouring muscles. This opens an opportunity for the use of ICA in this application (Hyvarinen et al. 2001).

BSS aims at recovering the sources from a set of observations. Applications include separating individual voices at a cocktail party. In the BSS problem, two processes are involved (Hyvarinen 1997, Comon 2001). These are the mixing and un-mixing processes. First, we observe a set of multivariate signals $[x_1(t), x_2(t) \ldots x_n(t)]$ that are assumed to be linearly mixed with a set of source signals $[s_1(t), s_2(t) \ldots s_n(t)]$. The mixing process is hidden so we can only observe the mixed signals. The task is to recover the original source signals from the observations through an un-mixing process. Equation (1) and (2) describe the mixing and un-mixing processes mathematically

(Bell and Sejnowski 1995, Hyverinen and Oja 1997).

Mixing $x = As$ (1)
Unmixing $Wx = WAs$ (2)

For solving the BSS it is assumed that the number of observations is equal to the number of source signals. Matrix s contains the original source signals driving the observations, whereas the separated signals are stored in matrix u. They are both $[n \times t]$ matrices. A and W are both $[n \times n]$ matrices, the mixing and un-mixing matrix, respectively. If the separated signals are the same as the original sources, the mixing matrix is the inverse of un-mixing matrix, i.e., $A = W^{-1}$.

METHODOLOGY

Experiments were conducted to evaluate the performance of the hand gesture recognition and facial muscle activity using surface EMG.

Recording and Processing of Hand Gesture sEMG

For the hand gesture experiments 5 subjects whose ages ranged from 21 to 32 years (4 males and 1 female) were chosen. The experiments were conducted on two different days on all five subjects. For the data acquisition a proprietary surface EMG acquisition system by Delsys (Boston, MA, USA) was used. Four electrode channels were placed over four different muscles as indicated in Table 1 and Figure 1. A reference electrode was placed at the *Epicondylus Medialis*.

The experiments were repeated on two different days. Subjects were asked to keep the forearm resting on the table with the elbow at an angle of 90 degree in a comfortable position. Three hand actions were performed and repeated 12 times in each instance. Each time the raw signal sampled at 1024 samples/second was recorded. The gestures used for the experiments are listed below and details are provided in Table 1.

- Wrist flexion (without flexing the fingers)
- Finger flexion
- Finger and wrist flexion together but normal along centre line

The hand actions and gestures represented low level muscle activity. The hand actions were selected based on small variations between the muscle activities of the different digitas muscles situated in the forearm.

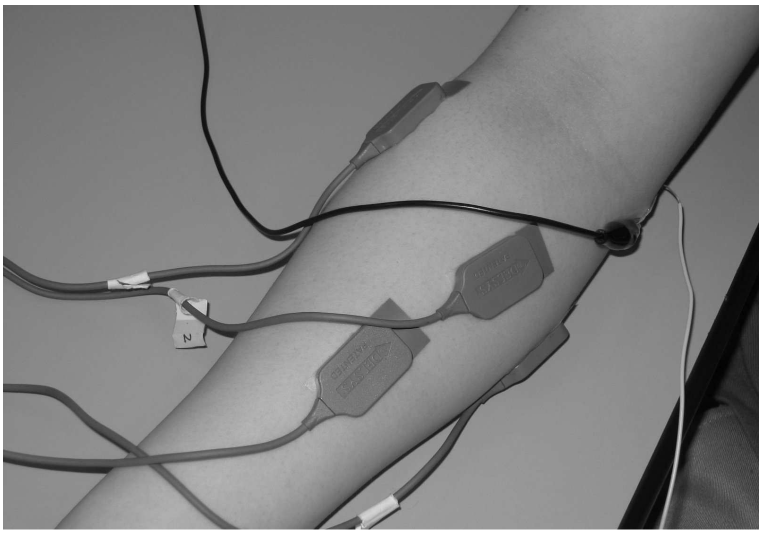

Figure 1. Hand gesture experimental set up with four electrodes.

Channel	Muscle	Function
1	Brachioradialis	Flexion of forearm
2	Flexor carpi radialis (FCR)	Abduction and flexion of wrist
3	Flexor carpi ulnaris (FCU)	Adduction and flexion of wrist
4	Flexor digitorum superficialis (FDS)	Finger flexion while avoiding wrist Flexion

Table 1. Placement of electrodes over the skin of the forearm.

Recording and Processing of Facial sEMG

Experiments were conducted on a single subject on two different days to test inter-day variations. A male subject participated in the experiment. The experiment used 4 channel EMG configurations as per the recommended recording guidelines (Fridlund and Cacioppo 1996). A four channel, portable, continuous recording MEGAWIN equipment (MEGA Electronics, Finland) was used. Raw signal sampled at 2000 samples/ second was recorded. Ag/AgCl electrodes (AMBU Blue sensors from MEDICOTEST, Denmark) were mounted at appropriate locations close to the selected facial muscles (on the right side *Zygomaticus major*, *Masseter* and *Mentalis*, and on the left side *Depressor anguli oris*). The inter-electrode distance was kept constant at 1 cm for all channels and experiments.

Controlled experiments were conducted where the subject was asked to speak 5 English vowels (/a/, /e/, /i/, /o/, /u/). Each vowel was spoken separately such that there was a clear start and end to the utterance. During utterance, facial sEMG from the muscles was recorded. sEMG from four channels were recorded simultaneously. The experiment was repeated ten times. A suitable resting time was given between each experiment.

Data Analysis

The aim of these experiments wase to test the use of ICA along with known properties of the muscles for separation of sEMG signals to identify hand gestures and to test the use of ICA on the facial sEMG signals for identifying speakers. Similar data analysis was performed on facial sEMG and hand gesture sEMG. For both experimental datasets, there were four channel (recordings) electrodes and four active muscles associated, forming a square 4×4 mixing matrix. The sEMG recordings were separated using a fast ICA matlab algorithm which has been developed by a team at the Helsinki University of Technology (Hyvarinen and Oja 1997). The mixing matrix A was computed for the first set of data only. Independent sources of motor unit action potentials that mix to make the EMG recordings were computed using equation (3),

$$s = Bx \tag{3}$$

where B is the inverse of the mixing matrix A. This process was repeated for each of the three hand gesture experiments. Four sources were estimated for each experiment. After separating the four sources sa, sb, sc and sd, each of these was segmented to sample length. Root Mean Squares (RMS) were computed for each separated sources using equation (4),

$$S_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{n} s_i^2} \tag{4}$$

where s is the source and N is the number of samples. This results in one number representing the muscle activity for each channel for each hand action and muscle activity for facial muscle. The RMS value of muscle activity of each source represents the muscle activity of that muscle and is an indicator of the force of contraction generated by each muscle. The above process was repeated for all 3 different hand actions 12 times and for each of the participants and repeated for the facial muscle sEMG for the 5 vowels. The outcome was 10 sets of examples, each pertaining to speech of 5 vowels. Results were used further for neural network analysis.

Neural Network Analysis

As a first step, the networks were trained using the randomly chosen training data. Performances were monitored during the training phase in order to prevent overtraining. Similar ANN architecture was used to test the reliability of hand gesture sEMG and facial sEMG. The ANN consisted of two hidden layers with a total of 20 nodes. A sigmoid function was the threshold function and the type of training algorithm for the ANN was gradient descent.

For hand gesture actions 12 sets of examples were used to train a back-propagation neural network. Inputs to the network were the 4 RMS values for each gesture and the output of the network were the three gestures. A back propagation neural network was then trained with the RMS values as the inputs and the gesture numbers as the targets. This network was then tested for the test data. For facial sEMG we used 10 sets with 4 inputs and 3 outputs by taking different combinations of vowels (a/i/u), (i/o/u), (a/o/u), (e/i/u), etc. Ihe inputs to the network were the 4 RMS values for each vowel utterance and the output of the network were the three vowels. Similar to the hand gesture analysis, a back propagation neural network was trained and tested with the RMS values as the inputs and the vowel utterance numbers as the targets.

RESULTS AND DISCUSSION

The aim of this research was to test the reliability and to determine the efficacy of the ICA technique to decompose sEMG into muscle activity from individual muscles and to classify the activity from these muscles to identify the hand gestures and speakers. The ability of the system to accurately classify the decomposed sEMG against the known hand gestures is given Table 2. Classification of different vowels using facial sEMG is shown in Table 3.

In hand gesture identification experiments, the accuracy was computed based on the percentage of correct classified data points to the total number of data points. These results indicate an overall classification accuracy of 100%. Results for two different set of vowels are shown in Table 3. Accuracy was computed based on the percentage of correct classified data points to the total number of data points. The results indicate an overall average accuracy of about 60%.

Number of participants	Wrist flexion		Finger Flexion		Finger flexion and wrist flexion	
	Day one	Day two	Day one	Day two	Day one	Day two
Subject 1	100%	100%	100%	100%	100%	100%
Subject 2	100%	100%	100%	100%	100%	100%
Subject 3	100%	100%	100%	100%	100%	100%
Subject 4	100%	100%	100%	100%	100%	100%
Subject 5	100%	100%	100%	100%	100%	100%

Table 2. Experimental results for hand gesture identification using muscle activity separated from sEMG using ICA.

Correctly classified vowels	Day 1	Day 2
/a/	60%	60%
/e/	60%	55%
/i/	60%	65%
/o/	55%	55%
/u/	65%	60%

Table 3. Experimental results for vowel classification using muscle activity separated from facial sEMG using ICA.

Comparative Evaluation of Hand sEMG with Facial sEMG

In order to measure the quality of the separation of hand gesture muscle activities in comparison to facial muscle activities, we used mixing matrix analysis. Generally, the efficiency of the ICA estimation stage for the un-mixing matrix can be verified by the following cross check. If a temporal wide-band source signal is split into, e.g., two consecutive narrow band signals p and q, the un-mixing matrix can be estimated independently twice from different raw data sections.

Then, processing quality can be validated by multiplying two of the independently retrieved matrices Wp and Wq, while one of the two has to be inverted. The result of this multiplication should be the identity matrix I (theoretically, in practice it should be close to I due to noise contributions), if everything in the process works correctly, and if the input source signal vector contains independent signals, as in equation (5).

$$I \approx G = W_p W_q^{-1} \tag{5}$$

In the comparative experiments here, performing these evaluation steps for hand gesture signals as well as for facial muscle signal captures yielded the following result matrices.

Due to the intrinsic order ambiguity of ICA, the order of unit vectors in the resulting matrix may be permuted, but from this sample measurement it appears confirmed that full dimensionality is achieved in the hand gesture application. On the other hand, for the facial application the validation result obviously contains dependent vectors and doesn't match the theory. This result shows a fundamental issue in the facial setup. One explanation for this effect may be that the facial signal components are not independent to the required extent. To validate this, a systematic experimental series was performed to evaluate the behavior of the dimensionality of the matrix G. In this supplementary investigation it turned out that the computational rank (Meyer 2000) of the matrix is not a valid measure in this signal-processing application. Instead the determinant value of the multiplication matrix was found being a valuable indicator in the sense that if this measure is close to zero there are problems with the recorded signal compound (in mathematical interpretations the determinant value would have to be exactly zero, but due to noise contributions this is never achieved even in constructed and obvious dependency cases). In the above sample, the determinant values are $\mathbf{det}(G_{hand}) = 2.2588$, while $\mathbf{det}(G_{facial}) = 0.0013$ is close to zero.

$$G_{hand} = \begin{bmatrix} 0.0800 & -1.0094 & 0.0271 & 0.0927 \\ 0.0670 & -0.0046 & 0.0307 & -1.2610 \\ 0.0143 & 0.0295 & 0.8062 & 0.0273 \\ 2.1595 & 0.3787 & -0.0729 & 0.0686 \end{bmatrix}, \mathbf{det}(G_{hand}) = 2.2588$$

$$G_{facial} = \begin{bmatrix} 0.0485 & -1.1738 & 0.0891 & -1.1105 \\ -0.8019 & 1.0171 & 0.7873 & 0.1669 \\ -0.8377 & 0.0142 & 1.1837 & -1.0169 \\ -1.4905 & 0.0192 & -1.3557 & 0.4750 \end{bmatrix}, \mathbf{det}(G_{facial}) = 0.0013$$

Overall, this comparative investigation shows that the high recognition rate for hand gestures can be considered as being reasonable, since in an inappropriate experimental setup the new technique doesn't work efficiently. Another notable aspect is that this technique can be applied in real-time mode, because of easy to use algorithms like fast ICA (Hyvarinen and Oja 2001) and back propagation neural networks (Fausett 1994). Quick computation times for these algorithms make the system suitable for day-to-day application.

Numbers of researchers have reported attempts to identify hand and body gestures from sEMG recordings but with low reliability. This may be attributed to low signal to noise ratios and high levels of cross-talk between different simultaneously active muscles. ICA is a recently developed signal processing and source separation tool and has been employed to separate the muscle activity and remove artefacts to overcome this difficulty. While ICA has been extremely useful for audio-based source separation, its application for sEMG is questionable due to the random order of the separated signals and magnitude normalisation. This paper reports research that overcomes this shortcoming by using prior knowledge of the anatomy of muscles along with blind source separation. Using a combination of the model and ICA approaches with a neural network configured for the individual overcomes the order and magnitude ambiguity.

The results demonstrate that the proposed method provides interesting results for inter-experimental variations in facial muscle activity during different vowel utterance. The accuracy of recognition is poor when the system is used for testing the training network for all subjects. This shows large variations between subjects (inter-subject variation) because of different styles and speeds of speaking. This method has only been tested for limited vowels. This is because the muscle contraction during the utterance of vowels is relatively stationary while for consonants there are greater temporal variations. The results show that for such a system to succeed, the system needs to be improved. Some possible improvements would include improved electrodes, site preparation, electrode location, and signal segmentation. This current method also has to be enhanced for large sets of data with many subjects in future. The authors intend to use this method for checking the inter-day and inter-experimental variations of facial muscle activity for speech recognition in the near future to test the reliability of ICA for facial SEMG

CONCLUSIONS

Independent component analysis (ICA) is a technique that is suitable for blind source separation and has been considered for decomposing sEMG to obtain individual muscle activities. This paper proposed applications and limitations of ICA on hand gesture actions and vowel utterance. Results on hand gesture identification indicate that the system is able to perfectly (100% accuracy) identify the set of selected complex hand gestures for each of the subjects. These gestures represent a complex set of muscle activation and can be extrapolated for a larger number of gestures. Nevertheless, it is important to test the technique for more actions and gestures, and for a large group of people. Results on vowel classification using facial sEMG indicate that while there is similarity between muscle activities, there are inter-experimental variations. Possible reasons are that people use different muscles even when they make the same sound, and cross talk due to different muscles makes the signal quality difficult to classify. Normalisation of the data reduced the variation of magnitude of facial sEMG between different experiments. The work indicates that people use same set of muscles for the same utterances, but there is a variation in muscle activities. It can be used in preliminary analysis for using facial sEMG-based speech recognition.

ACKNOWLEDGEMENTS

The authors would like to thank Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), Melbourne, and the Biomedical Signal Processing Group at RMIT University for their support.

REFERENCES

Bell, A. and Sejnowski, T. 1995. An information-maximisation approach to blind separation and blind deconvolution. *Neural Computation*, **7**, 1129–1159.

Comon, P. 2001. Independent Component Analysis, a New Concept. John Wiley, New York.

Fausett, L. 1994. Fundamentals of Neural Networks: Architectures, Algorithms, and Applications. Prentice Hall, New Jersey.

Fridlund, A.J., and Cacioppo, J.T. 1996. Guidelines for human electromyographic research. *Journal of Psychophysiology*, **23**, 567–589.

Fujimoto, T. and Nishizono, H. 1993. Muscle contractile properties by surface electrodes compared with those by needle electrodes. *Electroencephalography and Clinical Neurophysiology*, **89**, 247–251.

Greco, A., Costantino, D., Morabito, F.C. and Versaci, M.A., 2003. Morlet wavelet classification technique for ICA filtered SEMG experimental data. *Neural Networks*, **1**, 166–171.

Hyvarinen, A., Karhunen, J. and Oja, E. 2001. Independent Component Analysis. John Wiley, New York.

Hyvarinen, A. 1997. A fast and robust fixed-point algorithm for independent component analysis. *IEEE Transaction on Neural Networks*, **10**, 626–634.

Hyvarinen, A. and Oja, E. 1997. Fast fixed-point algorithm for independent component analysis. *Neural Computations*, **9**, 1483–1492.

Hyvarinen, A. and Oja, E. 1997. Fast fixed-point algorithm for ICA. The Fast ICA matlab software package `http://www.cis.hut.fi/projects/ica/fastica/index.shtml`

Kumar, S., Kumar, D.K., Alemu, M. and Burry, G. 2004. EMG based voice recognition. Proceedings of the 2004 Intelligent Sensors, Sensor Networks and Information Processing Conference, pp. 593–597.

Lapatki, G., Stegeman, D.F. and Jonas, I.E., 2003. A surface EMG electrode for the simultaneous observation of multiple facial muscles. *Journal of Neuroscience Methods*, **123**, 117–128.

Meyer, C.D. 2000. Matrix Analysis and Applied Linear Algebra, Cambridge University Press, Cambridge.

Naik, G.R., Kumar, D.K., Singh, V. and Palaniswami, M.,2006. SEMG for identifying hand gestures using ICA. Proceedings of the Workshop on Biosignal Processing and Classification at the 2nd International Conference on Informatics in Control, Automation and Robotics, pp. 61–67.

Nakamura, H., Yoshida, M., Kotani, M., Akazawa K. and Moritani, T. 2004. The application of independent component analysis to the multi-channel surface electromyographic signals for separation of motor unit action potential trains. *Journal of Electromyography and Kinesiology*, **14**, 423–432.

Parsons, W.T., 1986. Voice and Speech Processing. McGraw-Hill, New York.

Ganesh R. Naik* and Dinesh K. Kumar
School of Electrical and Computer Engineering
RMIT University
GPO BOX 2476V Melbourne
Victoria 3001, Australia
email: ganesh.naik@rmit.edu.au & dinesh@rmit.edu.au

* Author for correspondence.

(Manuscript received 30.08.2007, accepted 17.12.2007)