BLIND SEPARATION OF TEMPOROMANDIBULAR JOINT SOUND SIGNALS

Y. Guo[1], *F. Sattar*[2], *C. Koh*[1]

School of Mechanical & Production Eng.[1], School of Electrical & Electronic Eng.[2]
Nanyang Technological University, Nanyang Avenue, Singapore 639798

ABSTRACT

In order to develop a cheap, efficient and reliable diagnostic tool for the detection of temporomandibular joint disorders (TMD), sounds from the temporomandibular joint (TMJ) are recorded using a pair of microphone inserted in the auditory canals. However, the TMJ sounds originating from one side of head can also be picked up by microphone at the other side. Blind source separation (BSS) is thus proposed as a method to recover the original sound. The authors propose to use non-casual filters for the separation of TMJ signals. The algorithm is based on information theory and is an extension of early work by Torkkola. The separation was successful and the output can now be used for subsequent analysis of TMJ sounds.

1. INTRODUCTION

1.1 Temporomandibular Joint (TMJ)

The TMJ is the joint which connects the lower jaw, called the mandible, to the temporal bone at the side of the head. This joint is very important with regard to speech, mastication and swallowing. Any problem that prevents this system from function properly may result in temporomandibular joint disorder (TMD). Symptoms include pain, limited movement of the jaw, radiating pain in the face, neck or shoulders, painful clicking, popping or grating sounds in the jaw joint during opening and/or closing of the mouth.

1.2 TMJ Sound Recording

TMJ sounds during jaw motion are important indication of dysfunction and are closely correlated with the joint pathology [10]. The TMJ sounds are routinely recorded by auscultation and noted in dental examination protocols. However, stethoscopic auscultation is very subjective and difficult to document. The interpretations of the sounds often vary among different doctors.

Early detection of TMD, before irreversible gross erosive changes take place, is extremely important. A cheap, efficient and reliable diagnostic tool for early detection of TMD is being developed using TMJ sounds recorded with a pair of microphones placed at the openings of the auditory canals. The sounds are analyzed and classified into different classes based on their time-frequency reduced interference distribution (RID) [11]. Statistical correlations between different type of sounds and joint pathology can be used as a diagnostic tool for TMD.

1.3 Problems and Solution: Blind Source Separation (BSS)

Electronic recording offers some advantages over stethoscopic auscultation recording. Electronic recording of TMJ sounds allows the clinician to store the sound for further analysis and future reference. Secondly, the recording of TMJ sounds is also an objective and quantitative record of the TMJ sounds and thus changes in joint pathology. The most important advantage is that electronic recording allows the use of advanced signal processing techniques to the automatic classification of the sounds.

The auditory canal is an ideal location for the non-invasive sensor (microphone) to come as close to the joint as possible. The microphones are held in place by earplugs made of a kneadable polysiloxane impression material (called the Reprosil putty and produced by Dentsply). A hole is punched through each earplug to hold the microphone in place. The earplug also reduces the ambient noise in the recordings.

One common and major problem in both stethoscopic auscultation and digital recording is that the sound originating from one side will propagate to the other side, leading to misdiagnosis in some cases. It is shown that short duration TMJ sounds (less than 10ms) are frequently recorded in both channels very close in time. When the two channels show similar waveforms, with one lagging and attenuated to some degree, it can be concluded that the lagging signal is in fact the propagated version of the other signal [9].

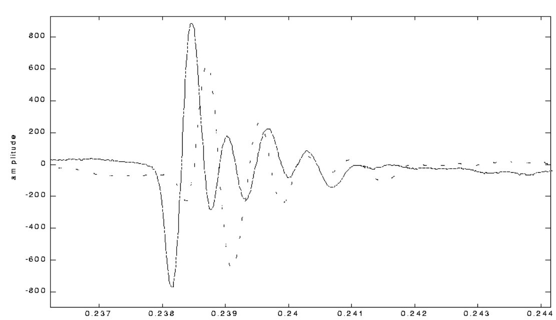

Figure 1. TMJ sounds at two channels.

This observation is very important. It means that a sound heard at auscultation on one side may have actually come from the other TMJ. This has great clinical significance because it is necessary to know the true source of the recorded sound, for example in diagnosing so called disk displacement with reduction [9].

The TMJ sounds can be classified into two major classes: clicks and crepitations. A click is a distinct sound, of very limited duration, with a clear beginning and end. As the name suggests, it sounds like a "click". A crepitation has a longer duration. It sounds like a series of short but rapidly repeating sounds that occur close in time. Sometimes, it is described as "grinding of snow" or "sand falling".

The duration of a click is very short (usually less than 10ms). It is possible to differentiate between the source and the propagated sound without much difficulty. This is due to the short delay (about 0.2ms) and the difference in amplitude between the signals of the two channels, especially if one TMJ is silent. However, it is sometimes very difficult to tell which is the source signal from the recordings. In Figure 2, it seems that the dashed line is the source if we simply look at the amplitude. On the other hand, it might seem that the solid line is the source if we look at the time (it comes first). Blind Source Separation (BSS) can be a solution to this problem. BSS is needed because both the sources (sounds from both TMJ) and the mixing process (the transfer function of the human head, bone and tissue) are unknown. If BSS is used, one output should be the original signal and the other channel should be the noise. Then it is very easy to tell which channel is the original sound.

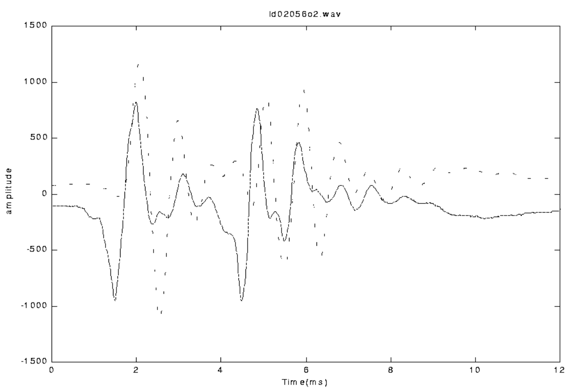

Figure 2. Sound recording of TMJ.

Furthermore, in the case of crepitation sounds, the duration of the signal is longer, and further complicated by the fact that both sides may crepitate at the same time. BSS is proposed as a means to recover the original sound for each channel.

2. BLIND SEPARATION ALGORITHM

2.1 "Infomax" Algorithm

BSS is the main application of independent component analysis (ICA), which reduces redundancy between signals and make them "as independent as possible". In BSS, second order statistics are inadequate to reduce redundancy between the input signals. Higher-order statistics are required for redundancy reduction and these are determined mainly in two ways. The first is the explicit estimation of the cumulants and polyspectra [6],[7]. The second is by obtaining higher-order statistics through the use of static nonlinear functions [2],[3].

Bell and Sejnowski [1] proposed an information-theoretic approach for blind source separation (BSS), which is refereed to as the "Infomax algorithm". Information theory can be used to unify several lines of research [4],[5] and different theories recently proposed for independent component analysis (ICA), leading to the same iterative learning algorithm for BSS.

2.2 Separation of Convolutive Mixture

The initial algorithm of Bell and Sejnowski [1] deals with the instantaneous mixture problem. The algorithm was further extended by Torkkola for the convolutive mixture problem.

Given m measured signals $x_i(k)$, which are combinations of n independent sources $s_i(k)$, the aim of blind separation is to produce n outputs $y_i(k)$, which recreate the n source signals, i.e., $y_1(k) = s_1(k)$, $y_2(k) = s_2(k)$, $\cdots$, $y_n(k) = s_n(k)$. Nothing can be assumed about the sources except that they are statistically independent.

Torkkola [8] suggested the feedback structure (see Figure 3 for $m = n = 2$) for the separation of convolutive mixture.

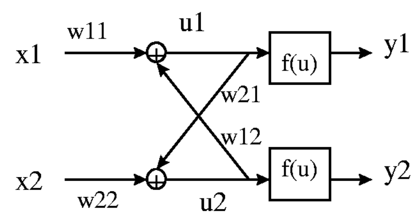

Figure 3. Feedback structure for the separation of convolutive mixture.

The learning rule for the convolutive mixture can follow the same steps as the instantaneous case [1]. Minimizing the mutual information between outputs y_1 and y_2 can be achieved by maximizing the entropy at the output. Assuming causal FIR filters for w_{ij}, the network performs the following operations in the time domain:

$$\begin{aligned} u_1(t) &= \sum_{k=0}^{L_{11}} w_{1k1} x_1(t-k) + \sum_{k=1}^{L_{12}} w_{1k2} u_2(t-k) \\ u_2(t) &= \sum_{k=0}^{L_{22}} w_{2k2} x_2(t-k) + \sum_{k=1}^{L_{21}} w_{2k1} u_1(t-k) \end{aligned} \quad (1)$$

where w_{ikj} is the k^{th} tap of the filter from source j to sensor i and L_{ij} is the filter length for the respective filter.

The relationships between the mixing filter and the separation filter can be expressed in z transform [8]:

$$\begin{aligned} W_{11}(z) &= A_{11}(z)^{-1}, \quad W_{12}(z) = -A_{12}(z)A_{11}(z)^{-1} \\ W_{22}(z) &= A_{22}(z)^{-1}, \quad W_{21}(z) = -A_{21}(z)A_{22}(z)^{-1} \end{aligned} \quad (2)$$

This is a network which combines the separation and deconvolution problem. Maximizing the entropy at the output will result in W_{11} and W_{22} not only inverting A_{11} and

A_{22}, but also whitening the sources. This can be avoided by forcing W_{11} and W_{22} to mere scaling coefficients. In the ideal case, W_{11} and W_{22} will have the following solutions:

$$\begin{aligned} W_{11}(z) &= 1, \quad W_{12}(z) = -A_{12}(z)A_{22}(z)^{-1} \\ W_{22}(z) &= 1, \quad W_{21}(z) = -A_{21}(z)A_{11}(z)^{-1} \end{aligned} \tag{3}$$

The learning rules for the separation matrix are:

$$\begin{aligned} \nabla w_{i0i} &\propto (1-2y_i)x_i + 1/w_{i0i} \\ \nabla w_{iki} &\propto (1-2y_i)x_i(t-k) \\ \nabla w_{ikj} &\propto (1-2y_i)u_j(t-k) \end{aligned} \tag{4}$$

where $k = 0, 1, 2, \cdots, L_{ij}$.

3. SEPARATION OF TMJ SIGNALS

3.1 TMJ Sounds

Figure 4 shows how the TMJ sounds are mixed. Sounds originating from a TMJ are picked up by the microphone in the auditory canal immediately behind the joint and also by the microphone in the other auditory canal as the sound travels through the human head.

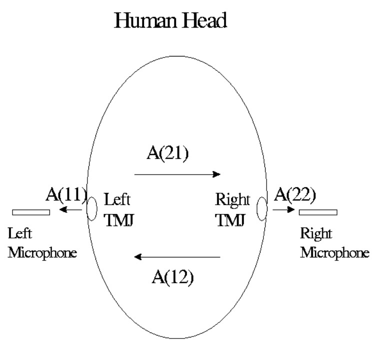

Figure 4. Mixing model of TMJ sounds.

Figure 5 shows a typical recording of TMJ crepitation sounds from the two microphones. Each recording is a mixture of the two TMJ sources.

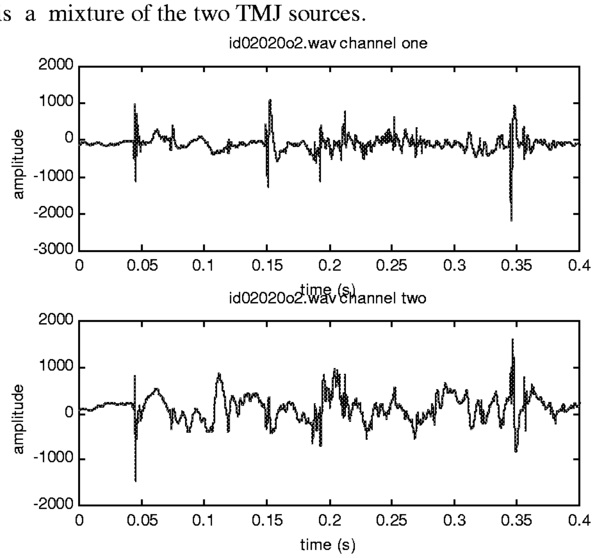

Figure 5. Typical TMJ signals (crepitation).

If we zoom in on Figure 5 near 0.045s, the signals are shown in Figure 6 for both channels (solid line for channel one and dashed line for channel two). It is difficult to tell how much of the signal in each channel comes from the ipsi (same side) TMJ and how much comes from the contra (opposite) TMJ [9].

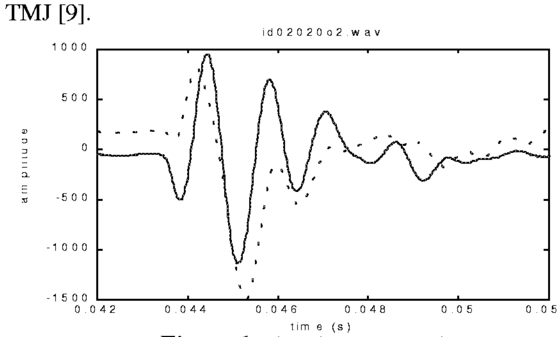

Figure 6. Signals near 0.045s.

If we look at the signals near 0.35s (Figure 7), it is even more difficult to differentiate the source from the propagated component because the signals are almost 180 ° out of phase. It is almost impossible to determine the short time delay and difference in amplitudes between the two signals.

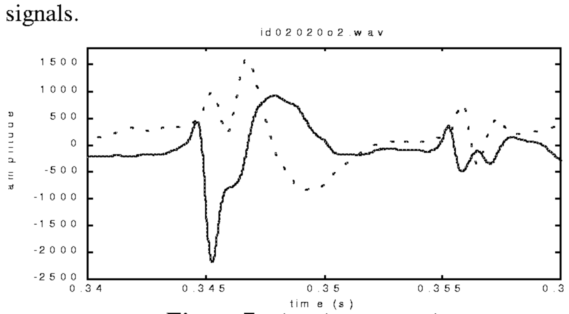

Figure 7. Signals near 0.35s.

3.2 Blind Source Separation Using Non-casual Filter

Torkkola's algorithm [8] works only when the stable inverse of the direct channel filters (A_{11} and A_{22}) exist. This is not always guaranteed in real world systems. In the separation of TMJ sound signals, the direct channel is the path from the source (TMJ) through the head tissue to the skull bone, then to the air in the auditory canal directly behind the TMJ and finally to the ipsi microphone.

However, even if a filter does not have a stable casual inverse, there still exists a stable non-casual inverse. Therefore, the algorithm of Torkkola can be modified and used even though there is no stable (casual) inverse filter for the direct channel.

The relationships between the signals are now changed to:

$$\begin{aligned} u_1(t) &= \sum_{k=-M}^{M-1} w_{1k1}x_1(t-k) + \sum_{k=-M}^{M-1} w_{1k2}u_2(t-k) \\ u_2(t) &= \sum_{k=-M}^{M-1} w_{2k2}x_2(t-k) + \sum_{k=-M}^{M-1} w_{2k1}u_1(t-k) \end{aligned} \tag{5}$$

where M is half of the total filter length and the zero lag of the filter is at $M+1$.

The derivative of the learning rule can follow the same procedure as in Torkkola [8]. According to equation (5), only the coefficients of W_{12} and W_{21} have to be learned. The learning rule is the same in notation but different in nature because the values of k have changed:

$$\nabla w_{ikj} \propto (1-2y_i)u_j(t-k) \qquad (6)$$

where $k = -M, -M+1, \cdots, M-1$.

3.3 Separation Result

The length of the filter used for the separation of TMJ signals is 160. The learning process is stopped when changes in the coefficients of the filters are within 0.001. The signals shown in Figure 5 are the candidate signals to be separated by the BSS algorithm. The output is shown in Figure 8. The signals are normalized to have a unit variance because BSS can only determine the signals up to a scalar factor.

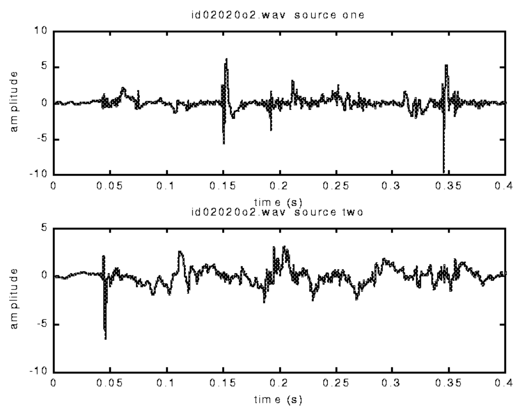

Figure 8. Signals after separation.

In order to see the effect of the separation, let us look at the signals near 0.35s (Figure 9). It clearly shows that the signal only comes from the first channel (solid line) and the second channel (dashed line) is basically silent. From Figure 8, it is also clear that the source is now coming from channel two at the time near 0.045s.

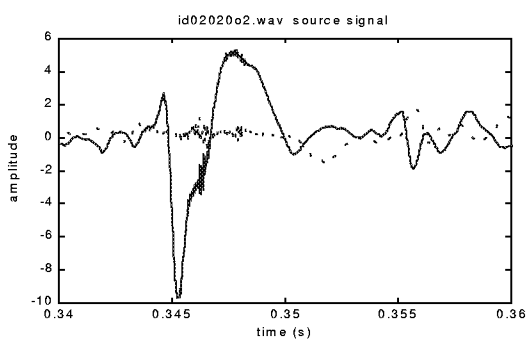

Figure 9. Signals after separation (near 0.35s).

4. SUMMARY AND DISCUSSION

Blind source separation was successfully applied to the separation of TMJ signals. This method was proposed because it is not guaranteed that the filters of direct channels have stable inverses. The result is satisfactory. The source signals can now be used for further analysis. This ensures that the signals used for subsequent classification are the real source signals, and not contaminated by the sounds propagated from the contra side.

5. REFERENCES

[1] Bell A. J. and Sejnowski T.J., "An information maximisation approach to blind separation", Neural Computation, Vol.7, pp. 1129-1159, 1995.

[2] Bellini S., "Bussgang techniques for blind deconvolution and equalisation". In Blind Deconvolution, Haykin, S., ed., Prentice-Hall, Englewood Cliffs, NJ, 1994.

[3] Comon P., "Independent Component Analysis, a new concept?" , Signal Processing, Vol. 36, pp. 287-314, 1994.

[4] Comon P., Jutten C., and Herault J. "Blind Separation of sources, Part II: Problems statement", Signal Processing, Vol. 24, pp. 11-21, 1991.

[5] Girolami M and Fyfe C., "Stochastic ICA contrast maximization using Oja's nonlinear PCA algorithm", International Journal of Neural Systems (in press), 1997.

[6] Gray R. M., Entropy and Information Theory. New York: Springer-Verlag, 1990.

[7] Hatzinakos D. and Nikias C. L., "Blind Separation based on higher-order statistics", in Blind Deconvolution, Haykin, S., ed., pp. 181-258. Prentice-Hall, Englewood Cliffs, NJ, 1994.

[8] Torkkola K. "Blind separation of convolved sources based on information maximization", IEEE workshop on Neural Networks for Signal Processing, Kyoto, Japan. Sept 4-6,1996.

[9] Widmalm S. E., et al, "Comparison of Bilateral Recording of TMJ joint sounds in TMJ disorders", 9th Int. Conf. on Biomedical Engineering, Dec 3-6, Singapore, 1997.

[10] Widmalm S. E., Westeesson P. L., Brooks S. L., Hatala M. P. and Paesani D., "Temporomandibular joint sounds: correlation to joint structure in fresh autopsy specimens", American Journal of Orthodontics and Dentofacial Orthopedics, Vol. 101, pp.60-69, 1992.

[11] Widmalm S. E., Williams W. J. , Christiansen R. L., Gunn S. M. and Park D. K., "Classification of temporomandibular joint sounds based upon their reduced interference distributions", Journal of Oral Rehabilitation, Vol. 23, pp. 35-43, 1996.