A GEOMETRICALLY CONSTRAINED ANECHOIC MODEL FOR BLIND SEPARATION OF TEMPOROMANDIBULAR JOINT SOUNDS

Clive Cheong Took‡, *Saeid Sanei*‡, *Jonathon Chambers*‡ *and Stephen Dunne*†

‡Centre of Digital Signal Processing, School of Engineering, Cardiff University,
CF24 3AA, Cardiff, United Kingdom
phone: + (44) 2920879063, email: {cheongc, saneis, chambersj}@cf.ac.uk
web: www.cardiff.ac.uk/schoolsanddivisions/academicschools/engin/cdsp
†GKT Dental Institute, King's College London, Denmark Hill Campus,
London, SE5 9RW, United Kingdom, email: stephen.dunne@kcl.ac.uk

ABSTRACT

Extraction of temporomandibular Joint (TMJ) sound sources is attempted in this paper. A priori knowledge of the geometrically *constrained* medium (i.e. the head) to extract the temporomandibular joint sound sources from their anechoic mixtures is exploited. This is achieved by estimating the delay of the contra lateral (from the opposite side) source within the anechoic mixtures. Subsequently, we consider the mixing medium as a multichannel filter of constrained length whereby the instantaneous mixing hypothesis can be assumed for each lag. Last but not least, we utilize mutual information as a selection criterion to pick the correct independent components. Successful reconstruction of the TMJ sources (free from artefacts present in the estimates of the TMJ sources by other well-known signal processing techniques) was achieved.

1. INTRODUCTION

Temporomandibular disorder (TMD) refers to medical problems associated to the region of mandible (lower jaw) and the temporal bone (skull). Dysfunction, internal derangement (such as disc displacement), and osteoarthrosis are examples of this disorder [2]. TMD is conventionally diagnosed by stethoscope auscultation. Similarly, TMJ sounds have long been related with TMD [1, 2, 3]. Identification of the right TMJ source (such as click and crepitus) is in particular challenging, when the TMJ sources mix. Therefore, this study investigates how a priori knowledge of our geometrically constrained environment assists in separating the two sources from their pair of anechoic mixtures by estimating the lag p of the contra lateral source with respect to the ipsi lateral (from the same side) source. The mixing model is represented as:

$$\mathrm{x}_i(\mathrm{t}) = \sum_{j=1}^{2} \mathrm{h}_{ij} s_j(\mathrm{t} - \delta_{ij}) \tag{1}$$

where $\mathrm{x}_i(\mathrm{t})$ is the ith TMJ mixture signal at discrete time t for i = 1,2. h_{ij} are the attenuation coefficients and δ_{ij} are the time delays associated with the path from the jth source to the ith sensor (stethoscope). The problem of reconstructing the sources from their anechoic mixtures can be formulated as blind source separation (BSS). Due to the proximity of the sources to the electrodes, the mixing matrix $\mathbf{H}$ can be expressed as:

$$\mathbf{H} = \begin{pmatrix} \mathrm{h}_{11} & \mathrm{h}_{12}\mathrm{z}^{-p} \\ \mathrm{h}_{21}\mathrm{z}^{-p} & \mathrm{h}_{22} \end{pmatrix} \tag{2}$$

Further to this, the following assumptions are made:

A1: The super-Gaussian TMJ sources enjoy mutual independence (as confirmed empirically).
A2: The lag p can be be computed a priori.

Proposing an instantaneous BSS algorithm to solve the TMD anechoic BSS problem is the objective of this study. The structure of the paper is as follows; in the following section we outline some relevant backgrounds prior to the development of the proposed algorithm. Section 3 overviews FastICA [4] and from this analysis, the algorithm to reconstruct the TMD sources is developed. This is followed by a resumé of the algorithm. Furthermore, section 4 provides visual comparisons of the performance of our algorithm against algorithms such as Parra's [5], which take advantage of the nonstationarity of the sources (convolutive modelling), DUET of Özgür and Rickard [6] which performs time-frequency masking (anechoic modelling), and eventually the conventional independent component analysis (ICA) FastICA [4] (instantaneous modelling). We highlight the fact that even a high signal to interference ratio (SIR) might not be satisfactory as the goal is to supply the dental specialist with TMJ source sounds free from 'artefacts' of the contra sources. Lastly, we conclude and suggest further works in Section 5.

2. BACKGROUND

2.1 The temporomandibular joint sources

TMJ sounds can be classified into two categories: click and crepitus. A click can be regarded as a brief, acute sound within the joint (thought to be prompted by the sudden collision of opposing wet surfaces). Clicks may arise due to the perforation of the disc (or its displacement) which holds the mandible and the temporal bone. Therefore, the patient may suffer from dysfunction and internal derangement. On the contrary, crepitus is a grating and continuous sound throughout movement (e.g. chewing). The presence of crepitus hints the presence of a degenerative joint disease (e.g. osteoarthrosis). Therefore, wrong detection of these sounds leads to misdiagnosis of TMD. A dental specialist has to distinguish between the TMJ sources whenever only one exists or they are combined together. Besides, both the click and the crepitus are further sub-divided into soft and hard categories. A soft TMJ source usually manifests when a patient starts to suffer from TMD, while a hard TMJ source may be

synonymous to a mature stage of TMD. For example, the left temporomandibular joint may be more damaged than its right counterpart. In this situation, a soft TMJ source and a hard TMJ source arise. One can refer to [2] for more information on TMJ sounds.

2.2 FastICA

FastICA has not yet been considered to estimate the TMJ sources from their anechoic mixtures. It maximizes the negentropy (of the linear combinations of the mixture signals) that can be approximated as

$$\text{Neg(y)} = \frac{1}{48}\text{kurt(y)}^2 + \frac{1}{12}\text{E}\{\text{y}^3\}^2 \tag{3}$$

where Neg(.) stands for negentropy, E{.} the statistical expectation, and kurt(.) the kurtosis. By optimizing the negentropy, FastICA will compute a vector $\mathbf{w} = [\text{w}_1, \text{w}_2, ..., \text{w}_m]$ (where m is the number of mixture signals), which is one of the rows of the so-called separating matrix **W**.

Let us examine the case where there are two mixture signals **x**(t) = $[\text{x}_1(t); \text{x}_2(t)]^T$ with our *instantaneous* mixing matrix **H** (2) with lag $p = 0$. $\mathbf{w}_a = [\text{w}_1, \text{w}_2]$ is the best possible linear combination of the mixture signals $\text{x}_1(t)$ and $\text{x}_2(t)$ to give a super-Gaussian independent component $\text{y}_1(\text{t})$.

$$\text{y}_1(\text{t}) = \arg \sup_{\text{w}_1,\text{w}_2} \Big(\text{Neg}(\text{w}_1\text{x}_1(\text{t}) + \text{w}_2\text{x}_2(\text{t}))\Big) \tag{4}$$

FastICA comprises of the objective function (4) and its optimization by fixed point iteration. Suppose that $\text{y}_1(\text{t}) \approx \text{s}_1(\text{t})$ (up to a scaling factor), given the successful optimization of (4). Expanding $\text{y}_1(\text{t})$, we have

$$\begin{aligned} \text{y}_1(\text{t}) &= \text{w}_1\text{x}_1(\text{t}) + \text{w}_2\text{x}_2(\text{t}) \\ &= \text{w}_1\Big(\text{h}_{11}\text{s}_1(\text{t}) + \text{h}_{12}\text{s}_2(\text{t})\Big) + \text{w}_2\Big(\text{h}_{21}\text{s}_1(\text{t}) + \text{h}_{22}\text{s}_2(\text{t})\Big) \\ &= \text{s}_1(\text{t})\underbrace{\Big(\text{h}_{11}\text{w}_1 + \text{h}_{21}\text{w}_2\Big)}_{\alpha} + \text{s}_2(\text{t})\underbrace{\Big(\text{h}_{12}\text{w}_1 + \text{h}_{22}\text{w}_2\Big)}_{\beta} \end{aligned} \tag{5}$$

From the above, we can infer that $\beta \approx 0$, and $\text{y}_1(\text{t}) = \alpha\text{s}_1(\text{t})$. In Figure 1, we show empirically that maximizing (4) will give $\beta \approx 0$ when two super-Gaussian sources are synthetically mixed for 100 Monte Carlo trials, using a different **H** (randomly generated) in each trial.

2.3 The TMD anechoic mixing model with lag *p*

The head can be considered homogeneous in terms of acoustic properties as the acoustic impedance (I) of the skull and that of the brain are approximately equal ($\text{I}_{\text{bone}} \approx \text{I}_{\text{brain}} \approx 10^6$) compared to that of air ($\text{I}_{\text{air}} \approx 10^3$) [7]. Therefore, most of those multipaths will be refracted out of the head with a *negligible* amount of attenuated reflections (echoes) within the brain. Hence our anechoic model is reasonable. On the other hand, the brain mean width is 0.16m in [8] while in [9] it is 0.14 m. In our study, we consider the brain width to be 0.15 m. Given that the speed of sound within the brain is 1505 m/s [7], the lag p corresponds to 10^{-4}s. In terms of number of samples, this lag will be $10^{-4} \times$ Fs, where Fs is the sampling frequency (Hz). Since Fs in our study is 10 KHz, p = 1 sample lag. Due to this small lag, one may also say that the TMD problem here can be formulated as an instantaneous source separation.

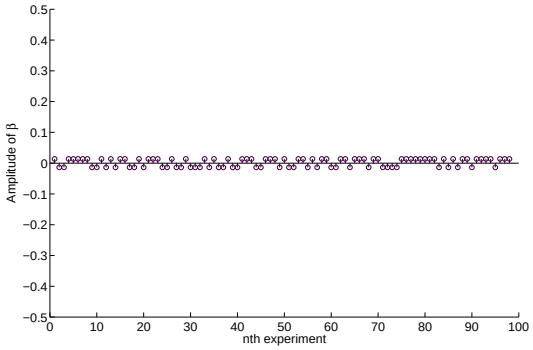

Figure 1: Values of β for 100 Monte Carlo trials. The small value ($< |0.015|$) of β suggests $\beta \approx 0$

3. DEVELOPMENT AND ANALYSIS OF THE PROPOSED FASTICA ALGORITHM

Recall our model (1) with **H** defined in (2). Earlier, we computed the lag $p = 1$ for our constrained size medium (the head). Hence, we delay x_1(t) to yield x_1(t-1) prior to the determination of y_1(t) . Therefore, FastICA performs the following optimization:

$$\text{y}_1(\text{t}) = \arg \sup_{\text{w}_1,\text{w}_2} \Big(\text{Neg}(\text{w}_1\text{x}_1(\text{t-1}) + \text{w}_2\text{x}_2(\text{t}))\Big) \tag{6}$$

If we expand the arguments of equation (6), we have

$$\begin{aligned} \text{y}_1(\text{t}) &= \text{w}_1\text{x}_1(\text{t-1}) + \text{w}_2\text{x}_2(\text{t}) \\ &= \text{w}_1\Big(\text{h}_{11}\text{s}_1(\text{t-1}) + \text{h}_{12}\text{s}_2(\text{t-2})\Big) + \\ &\quad \text{w}_2\Big(\text{h}_{21}\text{s}_1(\text{t-1}) + \text{h}_{22}\text{s}_2(\text{t})\Big) \\ &= \text{s}_1(\text{t-1})\underbrace{\Big(\text{h}_{11}\text{w}_1 + \text{h}_{21}\text{w}_2\Big)}_{\gamma} + \\ &\quad \text{w}_1\text{h}_{12}\text{s}_2(\text{t-2}) + \text{w}_2\text{h}_{22}\text{s}_2(\text{t}) \end{aligned} \tag{7}$$

If we assume $\text{y}_1(\text{t}) \approx \text{s}_2(\text{t-2}) + \text{s}_2(\text{t})$ (up to a scaling factor), then $\gamma \approx 0$. Consider $\lambda = \mathbf{w}_1\mathbf{h}_{12} + \mathbf{w}_2\mathbf{h}_{22}$ so that $\text{w}_1\text{h}_{12}\text{s}_2(\text{t-2}) + \text{w}_2\text{h}_{22}\text{s}_2(\text{t}) \approx \lambda\text{s}_2(\text{t})$. Besides, let us suppose $\text{s}_2(\text{t-2}) \approx \text{s}_2(\text{t})$; this assumption is reasonable in terms of negentropy, i.e. $\text{Neg}(\text{s}_2(\text{t-2})) \approx \text{Neg}(\text{s}_2(\text{t}))$ as $\text{t} \to \infty$. Hence, $\text{y}_1(\text{t})$ approaches $\lambda\text{s}_2(\text{t})$. We run 100 Monte Carlo trials to show that optimizing (6) results in $\gamma \approx 0$ and that $\lambda \not\approx 0$ (as illustrated in Figure 2 and Figure 3). We can also deduct that (6) can be re-expressed as:

$$\begin{aligned} \text{y}_1(\text{t}) &= \arg\min_{\text{w}_1,\text{w}_2} \Big(\text{Neg}(\gamma\text{s}_1(\text{t-1}))\Big) + \\ &\quad \underbrace{\text{w}_1\text{h}_{12}\text{s}_2(\text{t-2}) + \text{w}_2\text{h}_{22}\text{s}_2(\text{t})}_{\lambda\text{s}_2(\text{t})} \end{aligned} \tag{8}$$

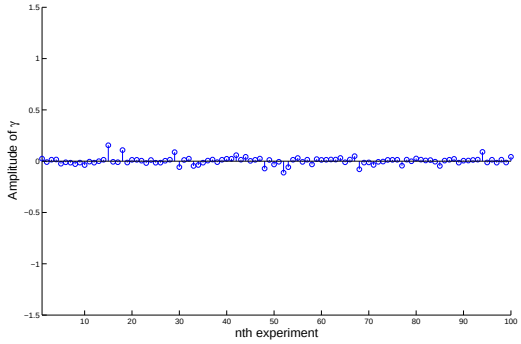

Figure 2: Values of γ for 100 Monte Carlo trials. The small value ($< |0.016|$) of γ hints $\gamma \approx 0$

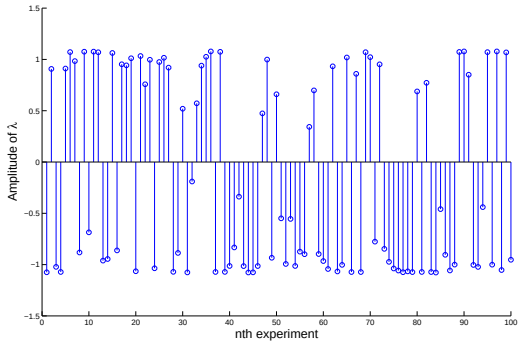

Figure 3: The values of λ when 100 Monte Carlo trials were run. On average, it can be deducted that $\lambda >> 0$ indicates $\lambda \not\approx 0$. The scaling ambiguity in ICA arises due to $\lambda >> 0$.

To extract a good *approximation* of $s_2(t)$, the key term is $\arg\min_{w_1,w_2}\Big(\text{Neg}(\gamma s_1(t\text{-}1))\Big)$. It is in fact minimized by FastICA (as shown in Figure 2).
At this point, note that $y_1(t) \approx \lambda s_2(t)$. The second independent component $y_2(t)$ will be a poor estimation of the second source. FastICA will search for another vector $\mathbf{w}_b$ orthogonal to $\mathbf{w}_a$, while maximising (6).

$$y_2(t) = s_1(t\text{-}1)\Big(h_{11}w_1 + h_{21}w_2\Big) + h_{12}w_1s_2(t\text{-}2) + h_{22}w_2s_2(t) \quad (9)$$

Note that new values of w_1 and w_2 are computed for $\mathbf{w}_b$. Recall that the mixture signal $x_1(t)$ was delayed by one sample, i.e. $x_1(t\text{-}1)$. This data manipulation is also equivalent to delaying $s_1(t)$ in $x_1(t)$ by 1 sample. Consequently, the alignment of $s_1(t\text{-}1)$ in both mixture signals, i.e. $x_1(t)$ and $x_2(t)$ is achieved. This alignment has enabled FastICA to optimize equation (8). But, in estimating $s_1(t)$, FastICA will have to optimize:

$$y_2(t) = \arg\min_{w_1,w_2}\Big(\text{Neg}(w_1h_{12}s_2(t\text{-}2) + w_2h_{22}s_2(t))\Big) + \gamma s_1(t\text{-}1) \quad (10)$$

Owing to the misalignment of $s_2(t\text{-}2)$ and $s_2(t)$ in minimizing Neg(.) in (10), a poor estimation of $s_1(t)$ will ensue. It is desired that $\lambda \approx 0$ to extract a good estimate of $s_1(t\text{-}1)$. On the other hand, γ conveys the scaling ambiguity of $s_1(t\text{-}1)$. As mentioned earlier, the misalignment of $s_2(t\text{-}2)$ and $s_2(t)$ in (10) will not necessarily fulfil the condition $\lambda \approx 0$. 100 Monte Carlo simulations (shown in Figure 4) to evaluate λ supports this statement.

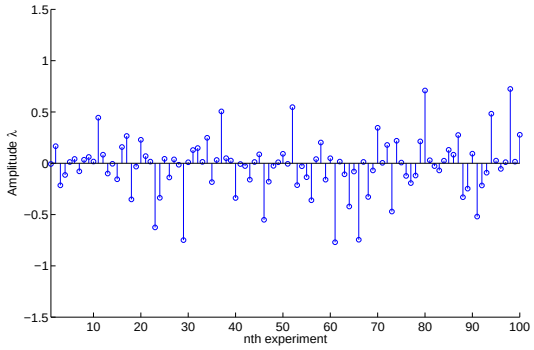

Figure 4: Values of λ for 100 Monte Carlo trials. The occurence of non-zero values of λ suggests $\lambda \not\approx 0$. Compare with Figure 2

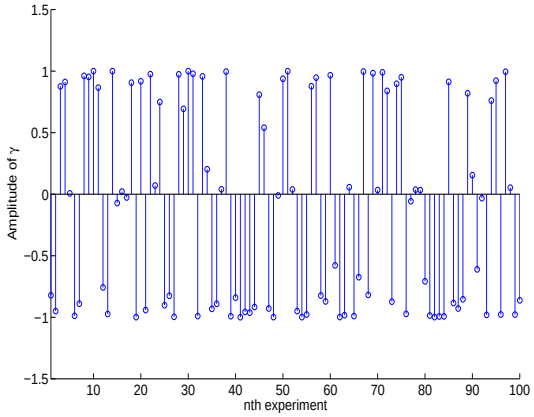

Figure 5: The values of γ for 100 Monte Carlo trials. Here, γ conveys the scaling ambiguity prevailing in ICA.

From these simulations, it is clear that the second independent component $y_2(t)$ will still consist of $s_2(t)$ and $s_1(t)$. Note that the latter is the dominant component in $y_2(t)$ due to the higher values of γ as illustrated in Figure 5. Therefore, we cannot estimate $s_1(t)$ with the mixture matrix $\mathbf{X} = [\mathbf{x}_1(t\text{-}1)\ \mathbf{x}_2(t)]^T$. To overcome this, we simply consider the lagged version of $x_2(t)$ by 1 sample, keeping $x_1(t)$ untouched. Hence, the new mixture matrix is $\mathbf{X} = [\mathbf{x}_1(t)\ \mathbf{x}_2(t\text{-}1)]^T$. However, this procedure will result in 4 independent components, with only two of them being consistent estimations of the true TMD sources. The selection of the correct pair of independent components can be done by computing the mutual information (MI) between any two pairs. The pair with

the minimum MI is picked as the right ones. It is understood that MI is employed as a measure of statistical independence.

The algorithm presented herein can be summarized as follows:

1. Estimate the lag p
2. Randomly initialize the mixing matrix **H**
3. Delay $x_1(t)$ by p samples
4. Utilize $x_1(t\text{-}p)$, $x_2(t)$ & **H** as inputs to the FastICA algorithm. This will yield $y_1(t)$ and $y_2(t)$.
5. Similarly, repeat steps (3)-(4) with $x_2(t)$ delayed to $x_2(t\text{-}p)$ instead of $x_1(t)$. Consider the pair of independent components obtained at this step to be $y_A(t)$ and $y_B(t)$.
6. Measure the MI between $y_1(t)$ & $y_A(t)$, $y_1(t)$ & $y_B(t)$, $y_2(t)$ & $y_A(t)$, and $y_2(t)$ & $y_B(t)$.
7. Select the pair with minimum MI.
8. If MI of the selected pair $\geq$ criterion, go back to step (2). Otherwise, the pair of independent components are the estimates of TMJ sources.

The last step (i.e. mutual information $\geq$ threshold/criterion) guarantees that the best estimates of the TMD sources are achieved since the performance of the FastICA may vary according to initialization of the mixing matrix **H**. In theory, MI $= 0$ when two sources are statistically independent, but in practice, the citerion$= 0.1$ was set. Thus, the algorithm ends when MI< 0.1.

4. SIMULATIONS

The following scenario is considered:

The displacement of the discs holding the mandible and the temporal bone gives rise to clicks. Moreover, the assumption that one temporomandibular joint is more 'damaged' than the other is made. In other words, both the soft and hard TMJ sources are present. Furthermore, these TMJ sources (when measured separately) were mixed synthetically by the randomly generated matrix **H**:

$$\mathbf{H} = \begin{pmatrix} -0.3031 & -0.4220z^{-1} \\ 0.1426z^{-1} & -0.2030 \end{pmatrix}$$

In the first place, we show the TMJ sources and their anechoic mixtures in Figure 6. Moreover, the four independent components (ICs) evaluated by FastICA when we delayed one of the mixture (one at a time) signals by 1 sample, are illustrated in Figure 7. Besides, we emphasize that an anechoic model with a sample lag of **one** cannot be formulated as an instantaneous BSS. These are shown in Figure 8. Last but not least, we use Parra's algorithm [5] and the time-frequency approach of Özgür and Rickard [6] to compare with our results. Their TMJ source estimates can be seen in Figure 9.

Remarks: For each simulation of the BSS algorithms, the first estimate is considered to be soft click and the other estimate as hard click. Besides, only the first two artefacts are highlighted for neatness purpose. In the two lower plots of Figure 8, it is obvious that each of the estimates of the sources still has 'artefacts'. Thus, an anechoic mixture of 1 sample lag cannot be formulated as an instantaneous BSS.

Table 1: The table below summarizes the performances of the algorithms.

Algorithm	Signal-to-interference ratio/dB
Time delayed mixture method	74.14
Instantaneous FastICA	48.20
Parra's algorithm	26.50
Time-Frequency Approach	17.61

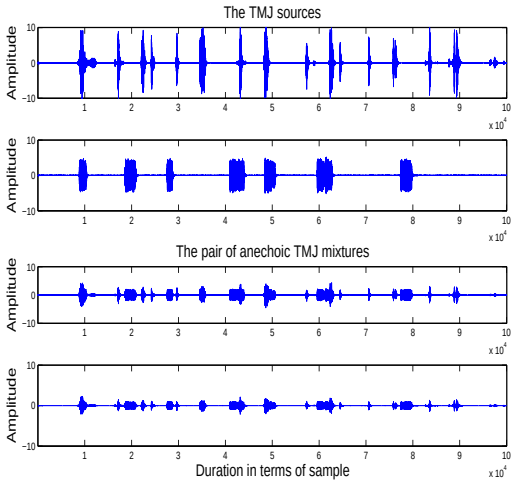

Figure 6: From top to bottom: soft click, hard click sources and their TMJ anechoic mixtures.

We can remark similar observations, regarding the estimates of Parra's algorithm [5] and those of Özgür and Rickard [6] in Figure 9. Table 1 summarizes the performance of those algorithms in terms of signal-to-interference ratio (SIR) for a quantitative analysis.

5. CONCLUSION

In this work, anechoic BSS of TMJ sources has been addressed. We have taken advantage of our geometrically constrained medium by pre-calculating the lag of 1 sample. Similarly, we have demonstrated the potential of considering the delayed versions of the mixture signals to solve this TMD problem. For clinical diagnosis of TMD and an accurate classification of the abnormalities, minor artefacts (commonly seen in conventional BSS algorithms) are absent in the source estimates of our method. This is why we have *visually* demonstrated the estimates of our method, compared to those of Parra [5], Özgür & Rickard [6], and against the instantaneous FastICA. Mutual information employed as a classification criterion has proved successful in selecting the correct pair independent components in the final stage of our algorithm. However in our study, we had to perform data manipulation (delay of the mixture signals) prior to using the instantaneous ICA algorithm. This was to compensate the lack of knowledge of FastICA about the anechoic model. Further work will be done on the development of a BSS algorithm that already assumes an anechoic model to leapfrog the data manipulation step to tackle this TMD BSS problem.

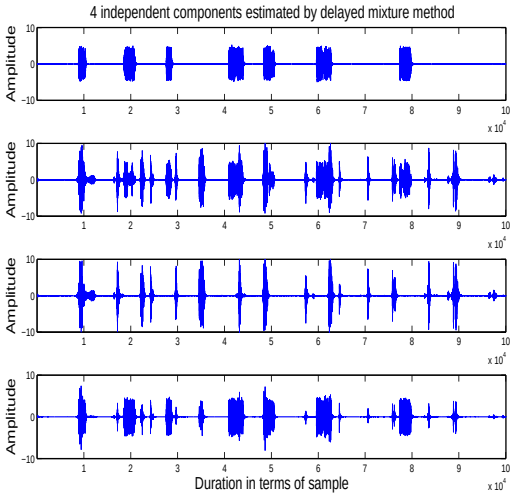

Figure 7: From top to bottom: $y_1(t)$, $y_2(t)$, $y_A(t)$, and $y_B(t)$ are the ICs evaluated by FastICA when one of the mixture signals was alternately delayed by one sample.

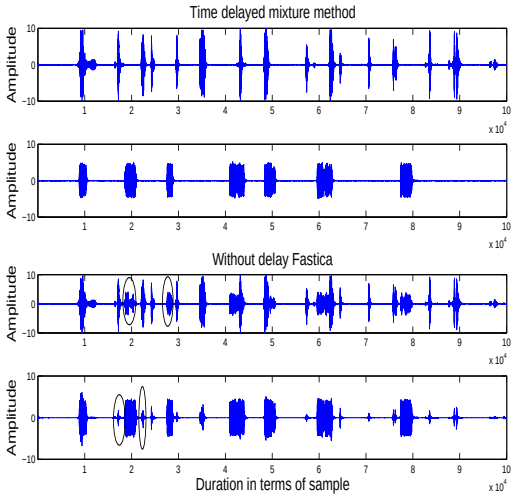

Figure 8: The two upper plots show the reconstructed TMJ sources (from top to bottom: $y_A(t)$ and $y_1(t)$) of the delayed mixture method. On the other hand, the lower ones demonstrate the outputs of FastICA when we treated these anechoic mixtures as instantaneous ones.

6. ACKNOWLEDGEMENT

The authors would like to thank Dr. S. Rickard [6] from University College of Dublin and Dr. S. Hamerling [5] for their matlab codes for simulation purposes.

REFERENCES

[1] T. Gay and C. N. Bertolami, "The spectral properties of temporomandibular joint sounds," *Journal of Dental Research*, vol. 66, pp. 1189, 1987.

[2] R. J .M. Gray, S. J. Davies and A. A. Quayle, *Temporomandibular Disorders: A Clinical Approach, 1st Ed*. Address: British Dental Association, 1995.

[3] J. K. Leader, J. R. Boston and T. E. Rudy and C.M. Greco, "Quantitative description of temporomandibular joint sounds: defining clicking, popping, egg shell crackling and footsteps on gravel," *Blackwell Science Ltd, Journal of Rehabilitation*, vol. 28, pp. 466–478, 2001.

[4] A. Hyvarinen, J. Karhunen and E. Oja, *Independent Component Analysis*. Address: John Wiley & Sons, INC, 2001.

[5] L. Parra and C. Spence, "Convolutive blind separation of non-stationary sources," *IEEE Transactions on Speech and Audio Processing*, vol. 8, pp. 320–327, May 2000.

[6] Y. Ögür and S. Rickard, "Blind Separation of Speech Mixtures via Time-Frequency Masking," *IEEE Transactions on Signal Processing*, vol. 52, pp. 1830–1847, Jul. 2004.

[7] J. A. Jensen, *Estimation of Blood Velocities using Ultrasound: A Signal Processing Approach*. Address: Cambridge University Press, 1996.

[8] A. Lee, H. Choi, H. Lee and J. Pack, "Human Head Size and SAR Characteristics for Handset Exposure," *ETRI Journal*, vol. 24, pp. 176–179, Apr. 2002.

[9] E. H. Chudler, *Brain Facts and Figures*. Address: http://faculty.washington.edu/chudler/facts.html.

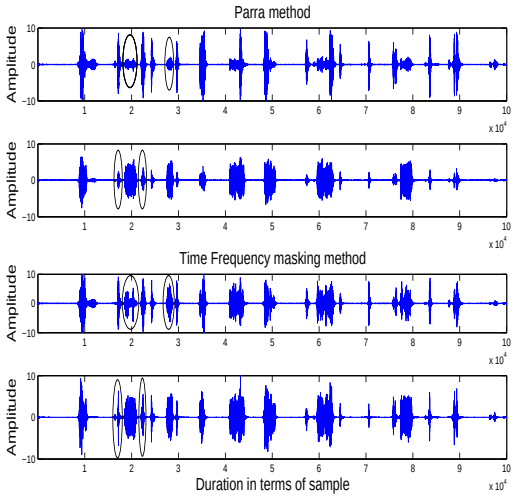

Figure 9: The two upper plots show the estimates of Parra's algorithm [5], whilst the ones at the bottom are those of the time-frequency approach of Özgür and Rickard [6].