Designing Brain-computer Interfaces for Sonic Expression

Satvik Venkatesh
Interdisciplinary Centre for
Computer Music Research
University of Plymouth
United Kingdom
satvik.venkatesh
@plymouth.ac.uk

Edward Braund
Interdisciplinary Centre for
Computer Music Research
University of Plymouth
United Kingdom
edward.braund
@plymouth.ac.uk

Eduardo Reck Miranda
Interdisciplinary Centre for
Computer Music Research
University of Plymouth
United Kingdom
eduardo.miranda
@plymouth.ac.uk

ABSTRACT

Brain-computer interfaces (BCIs) are beneficial for patients who are suffering from motor disabilities because it offers them a way of creative expression, which improves mental well-being. BCIs aim to establish a direct communication medium between the brain and the computer. Therefore, unlike conventional musical interfaces, it does not require muscular power. This paper explores the potential of building sound synthesisers with BCIs that are based on steady-state visually evoked potential (SSVEP). It investigates novel ways to enable patients with motor disabilities to express themselves. It presents a new concept called *sonic expression*, that is to express oneself purely by the synthesis of sound. It introduces new layouts and designs for BCI-based sound synthesisers and the limitations of these interfaces are discussed. An evaluation of different sound synthesis techniques is conducted to find an appropriate one for such systems. Synthesis techniques are evaluated and compared based on a framework governed by *sonic expression.*

Author Keywords

Brain-computer interface (BCI), Sound synthesis, Assistive technology, Design, Sonic expression

CCS Concepts

•Human-centered computing → Human computer interaction (HCI); Interaction design; •Applied Computing → Sound Synthesis;

1. INTRODUCTION

Brain-computer interface (BCI) systems aim to provide a communication medium that is independent of muscular control. It is helpful for patients with locked-in syndrome, that is the loss of all or most motor abilities, which is caused due to diseases like stroke, spinal cord injury, or amyotrophic lateral sclerosis (ALS). For such individuals, BCI systems initially enabled the control of wheelchairs, prosthetic limbs, and basic communication [25]. Computer Music has been incorporating methodologies from neural science to develop brain-computer music interface (BCMI) systems. Some of the early works that involve BCI and music are— Alvin Lucier composed a musical piece by sonifying electroencephalogram (EEG) in 1965 [20], [32] used EEG to control electronic sound synthesisers, and [29] carried out several experiments in using brain waves as high-level input structures for music. BCMI systems can be classified into four categories: *audification, sonification*, *musification*, and *control. Audification* is to realise an acoustic representation of the EEG data. This technique is generally noisy as brain wave frequencies and audio frequencies belong to different ranges. *Sonification* translates EEG data into sound by using a mapping procedure. This could be used for musical and non-musical purposes. One such application is spectral mapping, which is the translation of EEG data into audio frequencies. [14] sonified EEG information by observing the subject in different scenarios. This approach faces the problem of data being highly influenced by the environment and not the subject. Hence, BCI-users will not be able to communicate efficiently with such systems. *Musification* is the mapping of EEG information into musical parameters. This includes the translation to musical pitches or rhythms. This is not of much relevance to this paper as it focuses more on composition and performance instead of the spectral aspects of sound. *Control* gives the user the ability to make choices. [16] presented *parametric orchestral sonification of EEG in real-time* (POSER), the system in which multiple frequency bands are mapped to different instruments of a MIDI device. [13] discusses the development of a P300 composer, scale player, P300 DJ, and P300 algorithmic improviser. Therefore, these demonstrated the potential of building musical systems that incorporate BCIs.

Software synthesisers, in present times, use graphical user interfaces (GUIs) to enable musicians to design *signal flowcharts. Virtual studio instruments* use intuitive GUI components like knobs, faders, text boxes, and buttons. Therefore, the user has a diverse number of parameters that are adjustable. BCIs are relatively new and face several limitations. They are not as versatile as conventional GUI interfaces. They have low information transfer rates (ITRs), that is the amount of data that can be communicated from the user to the system. In recent years, a method called steady-state visually evoked potential (SSVEP) has been of escalating interest to researchers in the community of Neural Science, Computer Music, Animation, and Virtual Reality. This technique displays multiple regions (also referred to as targets) on the screen with each region flashing at a unique frequency and phase. These regions trigger unique patterns in brain waves that are detected by analysing EEG data. [3] developed an SSVEP-based BCI and obtained an ITR of less than 1 bit per second. [5] achieved an ITR of approximately 5.32 bits per second. Research to improve BCIs is consistently being conducted in many labs and new approaches towards analysing EEG data are regularly being proposed. Therefore, the first objective of this paper is to understand the present sophistication of BCI systems

NIME'20, July 21-25, 2020, Royal Birmingham Conservatoire, Birmingham City University, Birmingham, United Kingdom.

and evaluate general sound synthesis techniques that can be adopted by BCI-based synthesisers.

The second motivation of this paper is to find novel ways for such patients to express themselves. This paper proposes the concept of *sonic expression*, that is to express oneself purely by the synthesis of sound. On the surface, this might seem ambiguous due to the lack of form and structure, unlike musical expression. However, patients who have lost motor abilities for a prolonged period of time may realise the need for new and unconventional ways of expressing themselves.

BCMIs are beneficial for such patients because they provide a novel way of creative expression, which facilitates positive mental well-being [19]. Studies have developed bespoke systems to allow such individuals to experience the realm of music technology. [23] presented an SSVEP-based musical system that enabled a locked-in syndrome patient to control parameters to generate melodies. *Activating Memory* was a composition by Eduardo Miranda for four musicians (a string quartet) that played the musical choices made by four BCMI-users [11]. This improved the participation of BCMI-users because they could interact with each other. The BCMI literature has explored many pathways for musical expression, but not with directly synthesising sound.

Common aims of sound synthesis include creating or recreating sounds that are produced by known sound sources, have similarities with known sources, have no direct association with the source but still share certain characteristics, and noise where frequencies may have statistical relations with each other [27]. The objective of a BCI-based synthesiser hypothesised in this paper is different. The interface is not expected to find immediate use for computer music practitioners or experienced sound designers. It solely examines the prospects of sonic expression for patients with locked-in syndrome.

2. BCI-BASED SOUND SYNTHESISER

Three primary aspects of any sound synthesiser are the signal flowchart, unit generators, and signal processing objects. A signal flowchart presents a one-to-one relation between the input and output of the synthesiser. The purpose of signal flowcharts is to provide a high-level overview of the functions carried out by the synthesiser. Considering the present limitations of BCI systems, providing a GUI to customise signal flowcharts is not practical. Unit generators are the fundamental components for sound synthesis. These can be either signal generators like oscillators or signal modifiers like filters.

[5] presented an SSVEP-based BCI speller. It has 40 targets with alphanumeric, punctuation, and navigation commands (similar to a virtual keyboard) that enable a user to enter text. A similar SSVEP-based interface for a unit generator is designed in figure 1. The red colour regions flash at unique stimulation frequencies and phase for the user to make choices. Parameters are altered by ’+’ and ’–’ keys and allow the user to enter the required value for each parameter.

2.1 Criteria for Synthesis Techniques in BCIs

This paper presents a framework for evaluating sound synthesis techniques based on 2 criteria — *efficiency of target utilisation* (ETU) and *degree of sonic expression*.

2.1.1 Efficiency of Target Utilisation (ETU)

The BCI-speller presented in [5] has 40 different targets for the user. The system presents stimulation frequencies in the range of 8 to 15.8Hz. Studies in [15] show that SSVEP can be detected over a bandwidth of 1 to 100Hz. However, stimulation frequencies commonly fall in the range of 6 to 15Hz [33]. A study on finding an optimal frequency range for SSVEP is beyond the scope of this paper. Hence, it is assumed that 40 is the number of targets that will be available to build a sound synthesiser that uses SSVEP-based BCI. ETU is calculated by the following formula.

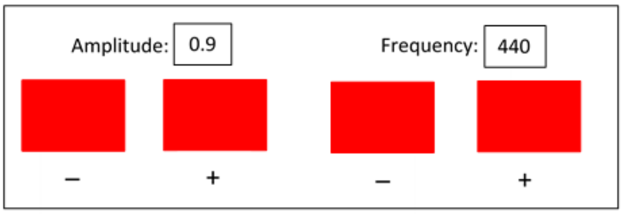


Figure 1: SSVEP-based BCI oscillator without an amplitude envelope. It has two adjustable parameters— frequency and amplitude. The red colour regions stand for targets which allow the user to make choices.

$$ETU = \frac{No.\ of\ targets\ utilised}{40} \times 100\% \qquad (1)$$

2.1.2 Degree of Sonic Expression

In order to evaluate synthesised sounds, different objective measures such as fundamental frequency, spectral centroid, harmonics, Mel-Frequency Cepstral Coefficients (MFCC) [34] have been used. Most of these metrics in the literature focus on calculating the similarity between the original sound and the synthesised sound. However, the objective of this paper is to provide a novel mechanism for locked-in syndrome patients to express themselves. Therefore, there is no particular sound for comparison, but it would be beneficial to have synthesised sounds that are human-like or natural-sounding.

Degree of sonic expression is an abstract concept and hence, cannot be formulated or defined by a mathematical equation. Spectral peaks are one of the primary characteristics of a sound spectrum and are commonly used in pitch-detection [4] and timbre-analysis [31]. Hence, the first factor that contributes to the degree of sonic expression is the number of spectral peaks that can be created by the synthesiser.

Formant regions serve as the *spectral signature* for vowels pronounced by humans and sounds produced by many instruments [27]. A formant is a concentration of acoustic energy resonating around a particular frequency [35] and is an important technique in timbral analysis [21].

In comparison to approaches like linear predictive coding and MFCC, formants have been chosen due to their simplicity. The number of parameters that can be controlled by BCI-based sound synthesisers is very less and hence, the sounds require simple methods for analysis. A vowel is characterized by a unique set of formants. Additionally, the ratio between formant frequencies is consistent in different voices. [2] suggests that the vocal tract is considered to have multiple formant regions. Formants are one of the cues that the human ear uses to identify the sound source. These facts suggest that the presence of formant regions contributes to creating natural-like sounds. Therefore, the second aspect that adds to the degree of sonic expression is the presence of formant regions.

3. SYNTHESIS TECHNIQUES

This section presents an evaluation of sound synthesis techniques for BCI-based sound synthesisers. It chooses additive synthesis as the first technique to be evaluated.

3.1 Additive Synthesis

Additive synthesis uses several sinusoids as elementary waveforms to produce complex signals. It uses oscillators as unit generators and individually generates *partials* by assigning each oscillator to a particular frequency. Let us consider additive synthesisers that are purely based on partial addition. Each oscillator has 2 inputs— frequency and amplitude. An SSVEP-based interface for such a synthesiser would require 4 targets to control each oscillator as shown in figure 1. Assuming that the total number of possible targets is 40, the maximum number of unit generators for this system is 10. The technique obtains an ETU of 100%. However, [27] suggests that realistic instrumental tones and natural-sounding timbres can be obtained only through time-varying additive synthesis. Therefore, the interface for a unit generator needs to be slightly modified. In time-varying additive synthesis, the signal from each oscillator is multiplied by an amplitude envelope, which is defined by ADSR— attack, decay, sustain, and release. The unit generator for additive synthesis needs to be designed as in figure 2. This interface requires 12 targets for one oscillator and reduces the number of oscillators to a maximum of 3. The calculated ETU for this technique is 90%. The range sounds that can be created with 3 sinusoidal oscillators by using additive synthesis is limited. [27] explains that to synthesise the attack of a trumpet sound, it requires at least 12 sine wave oscillators.

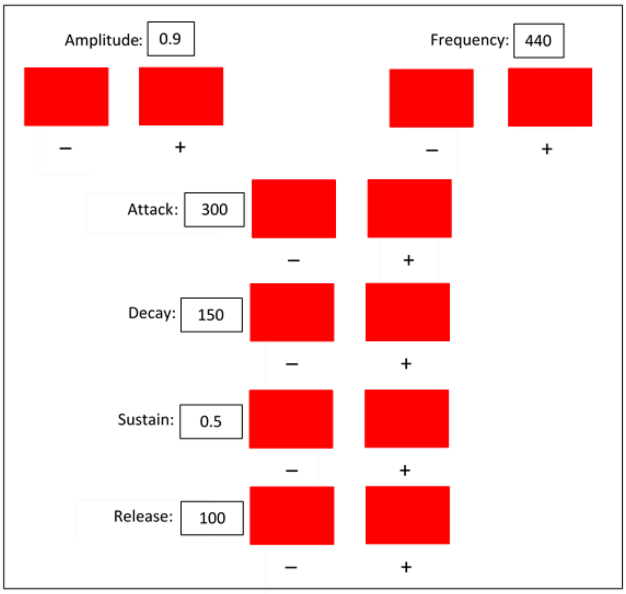


Figure 2: An oscillator with its amplitude envelope that uses an SSVEP-based BCI. It has 6 inputs — frequency, amplitude, attack, decay, sustain, and release. The red colour regions stand for targets which allow the user to make choices.

In additive synthesis, the number of spectral peaks is always equal to the number of oscillators. On one hand, sine waves are described as pure and natural-sounding and are capable of re-synthesising any sound. On the other hand, it requires a considerable amount of control data to generate a target sound. It is also computationally expensive. Furthermore, spectral peaks generated through additive synthesis are narrow and do not resemble formant regions. Hence, additive synthesis is not appropriate for such systems. The prospects of using simple frequency modulation (FM) is discussed in the following section.

3.2 Simple FM Synthesis

In the field of communication, modulation is the process of varying a parameter of one signal (carrier) with respect to a parameter of another signal (modulator). In simple FM, frequency of one carrier oscillator $\mathrm{O_c}$ is varied with respect to the signal of one modulating oscillator $\mathrm{O_m}$. A unit generator that is used in simple FM synthesis is similar to the one in additive synthesis, that is an oscillator. But, the amplitude envelopes carry out slightly different functions. The ADSR values for $\mathrm{O_m}$ represent the amplitude envelope for the modulation index I and not the amplitude envelope for $\mathrm{O_m}$. I is calculated by the following formula.

$$I = \frac{f_{\mathrm{c}} - f_{\mathrm{m}}}{f_{\mathrm{m}}} \tag{2}$$

where f_c is the carrier frequency and f_m is the modulating frequency. Tracing back to the history of sound synthesis, FM gained popularity because it was adopted by the Yamaha corporation and it is computationally inexpensive. [7] was the first to explore the musical potential of FM. An important phenomenon in FM synthesis its ability to generate side-bands. In 1987, in a Yamaha magazine advertisement, John Chowning stated that a sound which contains 50 harmonics will need 50 oscillators to be synthesised through additive synthesis, but FM synthesis enables you to do it with 2 oscillators [27]. The number of spectral peaks in the output is not limited by the number of oscillators. It only depends on the amount of modulation applied to the carrier signal. Hence, it addresses one problem faced by BCI-based additive synthesisers. In FM, energy is borrowed from the carrier frequency and distributed amongst side-bands. Therefore, the amplitude of the carrier frequency reduces by applying more modulation. [9] suggests ways to calculate the *significance* of side-bands in simple FM synthesis. Sounds generated by FM synthesis possess a unique and synthetic spectrum [27]. The spectral peaks are very narrow and do not produce formant regions. The energy distribution among side-bands cannot be easily controlled and individually modified. This explains the need to analyse other synthesis techniques that produce formant regions.

An SSVEP-based BCI synthesiser that adopts simple FM synthesis utilises 24 targets and realises an ETU of 60%. This is much lower when compared to additive synthesis. Simple FM synthesis does not utilise the full potential of BCIs because it uses only 2 oscillators. Hence, this can be optimised by evaluating modulation techniques that use more than 2 oscillators. The following section addresses the first problem faced by simple FM synthesis, that is its inability to generate formant regions.

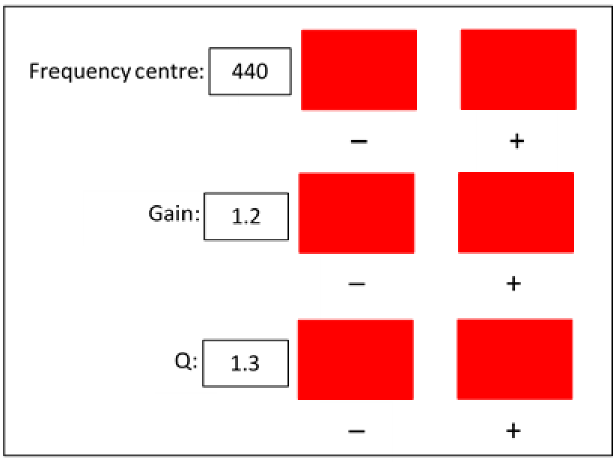
3.3 Subtractive Synthesis

Subtractive synthesis attenuates a complex signal and allows only certain regions of the spectrum to generate sound. It uses a sound source (generally of high spectral density) like white noise, pink noise, pulse wave, sawtooth wave, or square wave. Most subtractive synthesisers use a filter bank, that is an array of bandpass filters. Each filter attempts to create a formant region in the spectrum of the final output. The unit generator used in subtractive synthesis is a bandpass filter. It can be defined by three parameters— centre frequency f_{centre}, gain, and Q. Q of a filter is defined by the following formula.

$$Q = \frac{f_{centre}}{f_{high} - f_{low}} \qquad (3)$$

where f_{high} and f_{low} are the upper and lower bounds of the frequency band respectively.

SSVEP-based BCI designed for a band-pass filter is shown in figure 3. It uses 6 targets to determine the filter. Additionally, the subtractive synthesiser needs to define an amplitude envelope for the output signal. This is defined with the help of ADSR and requires 8 targets. Therefore, 5 unit generators can be incorporated in the synthesiser. The total number of targets utilised by it is 38 and realises an ETU of 95%. This is higher than the ETU for additive and simple FM synthesis.



Frequency centre: 440

Gain: 1.2

Q: 1.3

Figure 3: SSVEP-based BCI designed for subtractive synthesis. It has three adjustable parameters—centre frequency, gain, and *Q*. The red colour regions stand for targets which allow the user to make choices.

Subtractive synthesis is widely used in music technology applications and speech synthesis because it can generate formant regions. It resembles the spectral characteristics of the human voice and many traditional instruments [27]. The width of every spectral peak generated by a bandpass filter depends on the Q of the filter. A high Q-value produces a narrow spectral peak and vice versa. Hence, this allows the user to separately define the amplitude and width of each formant region, unlike simple FM synthesis where the user does not have individual control over side-bands.

Like additive synthesis, the number of spectral peaks is equal to the number of unit generators. But subtractive synthesis has the advantage of needing less number of parameters to define a unit generator. So, the maximum number of spectral peaks in the output spectrum is 5. This is higher than additive synthesis, but considerably less when compared to simple FM synthesis. Other subtractive approaches like FOF synthesis [28] and VOSIM [18] follow more sophisticated methods to generate formants. The unit generators (FOF grains and VOSIM waveforms) that are used in these techniques require more number of parameters to define them and hence, would utilise more targets. Therefore, the next section analyses other forms of frequency modulation synthesis to increase ETU.

3.4 Multiple Modulator FM Synthesis

Multiple modulator frequency modulation (MMFM) synthesis uses more than one oscillator to modulate a carrier wave. Similar to simple FM, the modulating oscillator and carrier oscillator require 12 targets each in the BCI. Hence, 2 modulator oscillators (O_{m1} and O_{m2}) and 1 carrier oscillator (O_c) can be defined and an ETU of 90% is obtained. This is an improvement when compared to simple FM synthesis, which utilised only 60% of the targets.

There are two subdivisions of MMFM— parallel and series. Parallel MMFM adds the outputs of O_{m1} and O_{m2} and then modulates the carrier wave. In series MMFM, O_{m1} modulates O_{m2} and O_{m2} modulates O_c. MMFM has two modulation indices, I_1 and I_2 for O_{m1} and O_{m2} respectively. [30] explains the explosion of partials in MMFM. It produces large number of side-bands for small values of I_1 and I_2. The maximum number of *significant* spectral peaks is much higher than simple FM synthesis. Therefore, MMFM has greater degree of sonic expression when compared to simple FM.

MMFM does not produce formant regions. Parallel MMFM can be used to generate regions of frequencies that are very close to each other. This might create an illusion of a formant because of the high spectral density in the respective frequency region. But the presence of a formant peak in these regions is not necessary. The amplitudes of individual partials cannot be separately varied. Hence, the degree of sonic expression of MMFM is limited due to its inability to generate formant regions. It has advantages and disadvantages when compared to subtractive synthesis. On one hand, the maximum number of spectral peaks that it can create is much higher. On the other hand, it does not create formant regions like subtractive synthesis. The next section seeks a modulation technique that has the ability to generate formants.

3.5 Multiple Carrier FM Synthesis

In multiple carrier frequency modulation (MCFM) synthesis, two or more carrier oscillators are simultaneously modulated by the same modulating signal. An SSVEP-based BCI synthesiser incorporating this technique would utilise 36 targets and obtain an ETU of 90%. The improvement of ETU observed in MCFM is similar to the one observed in MMFM. Both use an additional oscillator for the process of sound synthesis.

An interesting phenomenon observed in MCFM is its ability to produce formant regions as shown in figure 4. [8] presents the possibilities of using MCFM to synthesise the singing voice. He explains the formation of formant regions because of the presence of multiple carrier oscillators. The output generated by each of the carriers superimpose to create a composite spectrum and hence, generating formants. This proves its advantage over MMFM. The number of formant regions that can be generated by MCFM is equal to the number of carriers. Hence, an SSVEP-based BCI synthesiser that uses MCFM can create a maximum of 2 formants. In addition to formant regions, side-bands generated by both carrier oscillators continue to exist in the frequency spectrum. Hence, the number of spectral peaks that can be generated by MCFM is high. [26] explored the ability of MCFM to generate trumpet tones. [24] compared single-carrier and double-carrier FM synthesis for brass tones. He stated that double-carrier synthesis sounded more realistic because of the simultaneous generation of a formant region and side-bands. [6] used MCFM for the synthesis of singing voice and vowel sounds. The maximum number of formant regions that can be generated by subtractive synthesis is 5. A comparison between MCFM and subtractive synthesis based on this observation is presented in the next section. The above discussion demonstrates the high degree of sonic expression of MCFM because of its high ETU and ability to simultaneously generate side-bands and formant regions.

The following presents a brief comparison of all the synthesis techniques based on the different parameters that have been discussed in the paper.

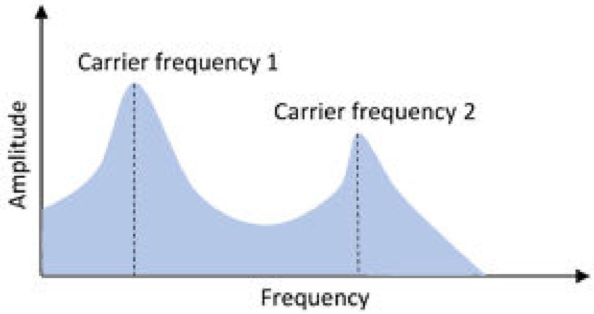


Figure 4: The creation of formant regions in MCFM.

4. COMPARISON OF SYNTHESIS TECHNIQUES

The paper evaluated sound synthesis techniques based 2 parameters— ETU and degree of sonic expression. The former is quantitative and calculates a percentage value for each technique and thus, a greater value determines more efficient usage of SSVEP-based BCI targets. Among all the sound synthesis techniques, simple FM scores the lowest ETU of 60%. Additive, subtractive, MMFM, and MCFM realise an ETU of 90% or 95%. A difference of 5% in ETU attributes to just 2 targets in the BCI and is therefore not a contributing factor for comparison. Degree of sonic expression is complex and is broken down into the ability to generate spectral peaks and formant regions. There is no definite method to measure the maximum number of spectral peaks that can be generated by frequency modulation techniques. Hence, the number of spectral peaks generated is classified into 5 categories— *very low*, *low*, *medium*, *high*, and *very high*. The ability to generate formant regions can be quantified by calculating the maximum number of formant regions that can be generated. Table 1 tabulates the abilities of different synthesis techniques to generate spectral peaks and formant regions.

Synthesis technique	Peaks	Formants
Additive	Very low	Nil
Simple FM	High	Nil
Subtractive	Low	5
MMFM	Very high	Nil
MCFM	High	2

Table 1: Comparison of different sound synthesis techniques based on number of spectral peaks and number of formants.

Among all sound synthesis techniques, subtractive and MCFM synthesis are demonstrated to be more efficient than the other techniques for BCI-based synthesisers. Subtractive synthesis can generate up to 5 formant regions in the output spectrum, whereas MCFM can generate only 2. In phonetics, most studies represent English vowels with the first 2 formants of the sound spectra [17]. Experiments in [10] suggest that in most cases vowels can be separated by detecting 2 formants. However, further precision can be achieved by analysing 3 to 5 formants. [22] analyses the spectra of musical instruments by comparing it with formant regions observed in vowels. He also stated the presence of one prominent formant region in woodwind instruments like the bassoon and oboe [12]. [1] analysed Violin, Clarinet, Trumpet, and Tuba sounds based on only 2 formant regions. [24] synthesised sound of a trumpet by using MCFM with only 2 carrier waves where 1 carrier was used to generate the main formant peak at 1500Hz. Therefore, subtractive synthesis certainly has the advantage of being able to generate 5 formant regions whereas, MCFM's ability to generate 2 formant regions still provides a satisfactory degree of sonic expression.

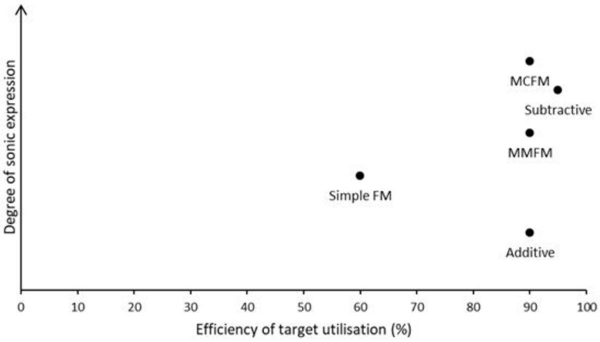


Figure 5: Comparison of different sound synthesis techniques based on efficiency of target utilisation (ETU) and degree of sonic expression.

[22] explains the importance of individual partials present in sound spectra. Apart from formant regions, there is a high possibility for narrow spectral peaks to be prominent, for instance in *gedackt organ pipes*. For low registers of Clarinet, it sounds hollow in the absence of second and fourth partials [22]. [24] generated 5 to 7 partials in addition to 1 formant region while synthesising trumpet tones. Hence, the ability to generate only 5 spectral peaks imposes restrictions and considerably reduces the degree of sonic expression. Figure 5 evaluates the sound synthesis techniques based on ETU and degree of sonic expression. MCFM is recognised as the most suitable synthesis technique for BCI-based synthesisers.

5. CONCLUDING DISCUSSIONS

This paper proposed the concept of sonic expression, that is to express oneself purely by the synthesis of sound. Sonic expression intends to explore novel ways for patients with locked-in syndrome to express themselves. The paper evaluated different sound synthesis techniques based on two criteria— efficiency of target utilisation (ETU) and degree of sonic expression. The former performs a quantitative evaluation of how efficiently the BCI is utilised by each technique. Degree of sonic expression was presented as an amalgamation of the ability to generate spectral peaks and formant regions.

Sonic expression is a new concept and seeks to find unconventional means of expression for individuals with motor disabilities. It introduces a new communication medium for such individuals and may lead to new ways of understanding them. This paper considers the possible number of BCI targets to be 40. This number is subject to change alongside research conducted to improve BCIs. The inclusion of more targets will lead to an increase in the number of adjustable parameters. This would increase the number of spectral peaks that can be generated by additive and subtractive synthesis. It will also increase the number of formants that can be generated by MCFM due to the inclusion of more carrier oscillators. By incorporating more targets, other sophisticated synthesis techniques (like FOF synthesis and VOSIM) that use complex unit generators can be explored. Degree of sonic expression is presented as a combination of two parameters— ability to generate spectral peaks and formant regions. Further research needs to be conducted to incorporate more parameters to realise sonic expression in greater depth.

6. REFERENCES

[1] J. Backus. *The Acoustical Foundations of Music*. New York: Norton, 2 edition, 1977.

[2] G. Bennett and X. Rodet. Synthesis of the singing voice. In M. V. Mathews and J. R. Pierce, editors, *Current Directions in Computer Music Research*, pages 19–44. MIT Press, Cambridge, MA, USA, 1989.

[3] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao. An online multi-channel ssvep-based brain-computer interface using a canonical correlation analysis method. *Journal of neural engineering*, 6(4):046002, 2009.

[4] F. Charpentier. Pitch detection using the short-term phase spectrum. In *IEEE ICASSP.*, volume 11, pages 113–116. IEEE, 1986.

[5] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao. High-speed spelling with a noninvasive brain-computer interface. *Proceedings of the national academy of sciences*, 112(44):E6058–E6067, 2015.

[6] J. Chowning. Computer synthesis of the singing voice. In *Sound Generation in Winds, Strings, Computers*, volume 29, pages 4–13. Stockholm: Royal Swedish Academy of Music, 1980.

[7] J. M. Chowning. The synthesis of complex audio spectra by means of frequency modulation. *Journal of the audio engineering society*, 21(7):526–534, 1973.

[8] J. M. Chowning. Frequency modulation synthesis of the singing voice. In *Current Directions in Computer Music Research*, pages 57–63. MIT Press, 1989.

[9] G. De Poli. A tutorial on digital sound synthesis techniques. *Computer Music Journal*, 7(4):8–26, 1983.

[10] D. Deterding. The formants of monophthong vowels in standard southern british english pronunciation. *Journal of the International Phonetic Association*, 27(1-2):47–55, 1997.

[11] J. Eaton. *Brain-computer Music Interfacing: Designing Practical Systems for Creative Applications*. PhD thesis, Plymouth University, 2016.

[12] F. Fransson. The source spectrum of double-reed wood-wind instruments. *Royal Institute of Technology, Stockholm, Speech Transmission Lab, QPSR*, 4:35, 1966.

[13] M. Grierson and C. Kiefer. Contemporary approaches to music bci using p300 event related potentials. In *Guide to Brain-Computer Music Interfacing*, pages 43–59. Springer, 2014.

[14] T. Hermann, P. Meinicke, H. Bekel, H. Ritter, H. M. Müller, and S. Weiss. Sonification for eeg data analysis. In *Proceedings of the 2002 International Conference on Auditory Display*, 2002.

[15] C. S. Herrmann. Human eeg responses to 1-100 hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental brain research*, 137(3-4):346–353, 2001.

[16] T. Hinterberger and G. Baier. Parametric orchestral sonification of eeg in real time. *IEEE MultiMedia*, 12(2):70–79, 2005.

[17] K. Johnson. *Acoustic and Auditory Phonetics*. Wiley-Blackwell, 3 edition, 2012.

[18] W. Kaegi. A minimum description of the linguistic sign repertoire (first part). *Interface*, 2(2):141–156, 1973.

[19] J. Leckey. The therapeutic effectiveness of creative activities on mental well-being: a systematic review of the literature. *Journal of psychiatric and mental health nursing*, 18(6):501–509, 2011.

[20] A. Lucier. Statement on: Music for solo performer. *Biofeedback and the Arts, Results of Early Experiments. Vancouver: Aesthetic Research Center of Canada Publications*, pages 60–61, 1976.

[21] J. McCarty. Formant analysis. [Online], 2003. Last accessed on 14-04-2020.

[22] J. Meyer. Structure of musical sound. In *Acoustics and the Performance of Music*. Springer Science & Business Media, 2009.

[23] E. R. Miranda, W. L. Magee, J. J. Wilson, J. Eaton, and R. Palaniappan. Brain-computer music interfacing (bcmi): from basic research to the real world of special needs. *Music & Medicine*, 3(3):134–140, 2011.

[24] D. Morrill. Trumpet algorithms for computer composition. *Computer Music Journal*, pages 46–52, 1977.

[25] R. Palaniappan. Electroencephalogram-based brain–computer interface: An introduction. In E. R. Miranda and J. Castet, editors, *Guide to Brain-Computer Music Interfacing*, pages 29–41. Springer, 2014.

[26] J. C. Risset and M. V. Mathews. Analysis of musical instrument tones. *Physics Today*, 22(2):23–40, 1969.

[27] C. Roads. *The Computer Music Tutorial*. MIT press, 1996.

[28] X. Rodet and J. L. Delatre. Time domain speech synthesis-by-rules using a flexible and fast signal management system. In *IEEE ICASSP*, volume 4, pages 895–898, Apr 1979.

[29] D. Rosenboom. The performing brain. *Computer Music Journal*, 14(1):48–66, 1990.

[30] B. Schottstaedt. The simulation of natural instrument tones using frequency modulation with a complex modulating wave. *Computer Music Journal*, pages 46–50, 1977.

[31] X. Serra and J. Smith. Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition. *Computer Music Journal*, 14(4):12–24, 1990.

[32] R. Teitelbaum. In tune: Some early experiments in biofeedback music. *Biofeedback and the Arts, Results of Early Experiments: Aesthetic Research Centre of Canada Publications*, 1976.

[33] Y. Wang, R. Wang, X. Gao, B. Hong, and S. Gao. A practical vep-based brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):234–240, 2006.

[34] L. Welling and H. Ney. Formant estimation for speech recognition. *IEEE Transactions on Speech and Audio Processing*, 6(1):36–48, 1998.

[35] S. Wood. Beginners guide to praat. [Online], 2005. Last accessed on 14-04-2020.