Abstract: Human echolocation is a technique used by small group of people (normally blind persons) to observe their surrounding via emitting an active signal (usually tongue click) and analyse the return echo. This skill allows them to navigate safely by 'seeing through sound'. Remarkable performances upon echolocating have attracted scholars' attention from multidiscipline field including psychophysical and engineering. Study of waveform diversity has reported the signal is wideband, consisting of multiple frequency components with exponential decay factor that probably made up the entire unique signal. As such, wideband perspective is considered during ambiguity function analysis using wide-band ambiguity function method. The ambiguity analysis and studies have reported optimum autocorrelation function is achieved by adopting a bio-inspired technique, which incorporated gammatone-filter processing. It will be interesting to synthesise the waveform using gammatone-filter approaches prior to relay it into ambiguity function process for analysis. Therefore, it is worth to explore its output results in this paper, which hold new findings that may help to understand how these humans achieve an optimum performance upon echolocating including ambiguity function characteristic. Thus, the new knowledge explored in this paper could be beneficial in emerging concept for new development in radar and sonar system application in near future.

1 Introduction

Echolocation is a biological skill of emitting an active signal to the open space (environment) and obtaining surrounding information by analysing return echo signal. Bat and dolphin are examples of mammals that utilise echolocation technique in their daily life mainly for navigation and searching for foods [1]. The outstanding ability of these mammals in echolocating became the emerging concept in a man-made sensor such as radar and sonar systems [2]. In nearly a century ago, revolution in advancing technology for radar and sonar application has been intensively explored by scholars. This development trend demonstrating the exploitation of bio-inspired processing is worth to explore.

What is less known, there are a small group of humans using echolocation technique on daily basis. Human echolocation is a biological skill that requires a person to use his/her signal (often is tongue click) and analyse return echo to perceive the surrounding. Human echolocation studies were carried out by Supa in 1944 [3]. During the experiments, participants were blindfolded and able to sense the presence of obstacle via active noise by scuffing heel on the floor. In 1967, Rice discovered that most of the participants created either long hissing sound or punctuated tongue clicks as a source of the signal upon echolocating [4]. More studies recently discovered cognitive and psychophysical aspects involved in human echolocation [5–7]. Preliminary study revealed human echolocator waveform is unique with uncommon modulation scheme [7, 8]. The analysis of waveform diversity of human echolocator signal continues and reported it is a wideband signal, consisting of multiple frequency components with exponential decay factor that probably made up the entire unique signal [9–11].

The following study demonstrated that an ideal autocorrelation function (ACF) is achieved by adopting gammatone filter (GF) processing and solved the limitation encounter by matched filter (MF) method for human echolocation waveform [12]. Therefore, this study used GF approaches to synthesise sound signal prior to relay it into ambiguity function (AF) processing for analysis. Hereby, this paper is organised as follows. Section 2 presents the theory and concept of bio-inspired processing. Section 3 defines the concept of AF analysis using synthesise waveform from GF processing. Section 4 observes results and discussed the characteristic of AF. Finally, the conclusion of this paper is given in Section 5.

2 Bio-inspired processing

2.1 Human auditory system

The sense of hearing is responsible in synthesising sound signal obtained from the surrounding. Studies have reported that the fluid inside the human auditory system helps in maintaining the stability of the human brain. The human auditory system is made up of three regions: (i) outer ear, (ii) middle ear and (iii) inner ear. Each region is assigned to specific function upon receiving sound signal. The complex curvature on pinna is responsible to modify the sound signal especially the high frequency components, thus, it helps during sound localisation [13]. As the sound signal travels into inner ear via middle ear, the medium of propagation for ear region will be different (air to liquid). Thus, impedance matching is needed to maintain the optimum energy at stapes upon relaying the signal into the inner ear by oscillating basilar membrane (BM) [13, 14]. At inner ear, the BM is responsible for synthesising sound signal prior to relay to the human brain [13]. The physical construction of BM is unique, where the thickness and width vary along BM. This helps certain region along BM corresponding to only specific frequency components from the sound signal. This theory justifies that the human auditory system synthesises sound signal in a non-linear manner. The properties of BM can be represented as a set of overlapping band-pass filter (BPF), where each filter bandwidth is different from others and it is estimated by using equivalent rectangular bandwidth (ERB) scale.

2.2 Motivation of incorporating bio-inspired (auditory filter) processing in analysis

The trend of measuring an efficient frequency energy from audio spectrum using auditory filter processing is widely used among scholars mainly for features extraction in speech recognition application [15–17]. The benefit of adopting auditory filter
The benefit of GF processing are as follows: (i) its ability in spectral computation, which helps in miniature hardware design resolving the rich contents of the signal in much detail [18], (ii) compare to others auditory filter scheme. The needs of a spectral content of speech and non-speech signal to be synthesised gammatone filter processing prior to ambiguity function analysis J. Eng. [19]... [21]. In addition, the construction of EP1 waveform consists of two procedures are available in [6].

Unlike other mammals echolocator waveforms, the properties of the EP1 waveform shown in Fig. 3a can be classified as: (i) the waveform duration is about 3 ms, (ii) no specific modulation scheme is presented, and (iii) the waveform shape resemble steady rate decay, which can be denoted as exponential decay function [11]. In addition, the construction of EP1 waveform consists of two main frequency region components known as main frequency components (20 Hz–5kHz) and higher frequency components (10–11 kHz) as shown in Fig. 3b [10, 11]. By exploiting human auditory strategy into AF analysis, it may provide new meaningful insight about the characteristics of waveform upon resolving Doppler and delay information.

To extend, AF analyses also have been explored for the characteristics of the signal using wideband ambiguity function (WAF), assuming the signal is an acoustic signal (which frequency...
ranges from 20 Hz sweeping up to 20 kHz [9, 10]. However, human hearing theory reported strategy in synthesising sound signal is a resemble logarithm factor, rather than assume it as a complex wideband signal. The logarithm factor in this context is to obtain frequency components coefficient from the respective sound signal using (2). As a result, spectral representation at a specific region along BM can be resolved in much detail. A similar approach is imitated in this paper using GF processing.

3.2 Synthesising waveform using GF processing for AF analysis

The analysis is started by passing EP1 waveform into GF processing in order to obtain synthesised output, known as frequency response coefficient as illustrated in Fig. 2. The synthesised EP1 waveform, $G_{fi}$, is fit into a row $i$(channel), where each channel has a unique frequency information obtained from all parameter given in (2). For easier notation, frequency components coefficient were arranged from highest channel order (highest frequency component, $i = 1$) towards lowest channel order (lowest frequency component, $i = 10$).

This strategy allowed only specific frequency components to pass through each channel, and prevents other pulses (frequency) from entering the respective channel. Thus, characteristic information of Doppler and delay from each channel is able to resolve upon their needs during the echolocation. As a matter of fact, a series of Doppler and delay information from synthesised signal provide cue about their surroundings, thus it helps these human to achieve high accuracy during echolocation. To extend, the ability of each channel in resolving Doppler and delay characteristic is studied using AF analysis in Section 4.

Range resolution is obtained using (3), where $c$ is the speed of sound and it equals to 342 m/s, $r$ is to delay zero cut at $-3$ dB threshold. For speed resolution is obtained using (4), where $f_o$ is the difference between main lobe and highest side lobes level at Doppler zero cut and $f_o$ is carrier frequency (estimation centre frequency, $cf_i$ index along BM as explained in Section 2)

$$\Delta R = \frac{c \cdot r}{2}.$$  

$$\Delta v = \frac{f_o \cdot c}{2f_o}.$$  

The analysis of AF processing from individual channel of GF is conducted and denoted as $A(t,f_d)$, where each channel, $i$, is separately resulting output for Doppler and delay features given in (5).

$$A(t;f_d); \; i = 1, 2, 3, \ldots, N$$  

Narrow-band ambiguity function (NAF) denoted as $A(t,f_d)$ is used because the output signal of $G_{fi}$ is narrow-band, $r$ is range delay and $f_d$ is Doppler frequency, assuming speed of sound in the air equals to 342 m/s. No pulse compression method is being considered yet in this paper, since the aim is to study the base-band characteristic of synthesised EP1 waveform prior to AF analysis. The analysis of AF via GF process is illustrated in the block diagram as shown in Fig. 4.

4 Result and discussion

The strategy as discussed in Sections 3 and 4 is believed to help BM in synthesising spectral components from acoustic signal in a more efficient perspective. The logic of this technique is to allow only specific region along BM to correspond to frequency information existing on sound signal, prior to sending to human brain via hair cell for further processing. While the truth on how human performs sound detection still remains unknown, it is believed theory that have been discussed in this paper so far helps these people resolving information in surrounding from click signal. For the preliminary study, results for highest channel order and lowest channel order will be discussed in this paper. This helps us to understand the properties of this individual channel in resolving Doppler and delay information.

Fig. 5 shows the zero delay cut AF output obtained from EP1 waveform using GF processing approach. There is distinct pattern observed from highest channel order towards lowest channel order. At the highest channel order ($f_o = 14.39$ kHz) as shown in Fig. 5a, the waveform could resolve target speed at least 15.77 kHz (187.45 m/s) apart. While at lowest channel order ($f_o = 100$ Hz) as shown in Fig. 5b, the waveform could resolve target speed at least 689 Hz (1178.19 m/s). The trend in resolving information of target speed apart is improved from highest order channel towards lowest order channel.

Fig. 6 shows the zero Doppler cut AF output obtained from EP1 waveform using GF processing approach. Similar to zero delay cut cases, comparable distinct features is presented from individual channel order. The waveform at the highest channel order as shown in Fig. 6a could resolve target range at least 0.0362 ms (0.62 cm
apart). At the lowest channel order as shown in Fig. 6b, the waveform could resolve target range at least 0.7256 ms (12.41 cm apart). This performance demonstrates the ability in resolving information of target range apart is improved from lowest channel order towards highest channel order.

The Doppler zero cut and delay zero cut AF results demonstrate at highest channel order towards lowest channel order, ability in resolving speed and range are contrasted from each other. The ambiguity contour plot of Fig. 7a shows that the highest channel provides the best range resolution. Meanwhile, an ambiguity contour plot at the lowest channel order as shown in Fig. 7b revealed the ability in resolving range resolution at the most least of any of other channel. However, the performance in resolved speed resolution is at the best.

Some preliminary assumption based on the obtained results can be made: (i) the presence of multiple frequency components in EP1 waveform possibly offers more information during analysing echo signal upon echolocation, (ii) the non-linear strategy upon synthesising human echolocator's waveform allowed spectral representation in much detail, which will benefit him to resolve surrounding perspective through sound information, and (iii) these exploited approaches will help him in achieving high rate of accuracy during echolocation process. While the truth on human hearing processing still remain vague [13], results obtained demonstrate that human auditory system is likely capable in performing adaptive processing upon synthesising sound signal upon their needs.

This ‘trade-off’ mechanism may guide human echolocators to perceive their surrounding in details by gathering Doppler and delay information from specific frequency components when it is needed. Together with sophisticated cognitive strategy, how human echolocator able to do autonomous navigation and discriminate obstacle in their surrounding might explain with high accuracy and detail.

Findings have been discussed so far could become an alternative perspective for new development in radar and sonar system applications. The advantages of having multiple frequency components in a single waveform offer more detailed information about target features to be exploited from the waveform rather than performing demanding signal compression method in order to achieve such higher resolution. Finally, preliminary studies have been discussed in this paper could emerge a new bio-inspired concept which can be useful for development of radar and sonar system application in near future.

5 Conclusion

Human echolocations have shown a remarkable ability of these people in 'seeing with sound'. In psychophysical perspective, human echolocation involves complex cognitive processing, which

is exploiting both human motor-skills and sensory input. A study in waveform diversity of human echolocator waveform revealed it is a wideband signal, and it consists of multiple frequency components in shape of exponential decay factor. These waveform properties correspond to the strategy of human auditory system upon synthesising sound signal. In this paper, bio-inspired processing is being implemented by using gammatone-filter processing to synthesise the waveform prior to relay to AF processing for analysis. Therefore, products of synthesised waveform contain specific frequency components rather than complex frequency components from EP1 waveform. This strategy offers more accurate spectral representation, which helps these human in analysing (seeing with sound) their surrounding just using sound information. The AF performance of synthesising signal has demonstrated the ability in resolving Doppler and delay information suggest the similar concept is being exploited by the human auditory system. At this point, it could be possible explanation on how these echolocators able to resolve their surroundings in detail. When adopted this perspective in radar and sonar systems application, the similar concept is needed upon resolving target information. Furthermore, the needs of a new concept for radar and sonar systems application have motivated researcher to revolutionise existing system into the next level. In nearly a century ago, bio-inspired processing has proven its concept in emerging technology for radar and sonar applications. Hereby, the preliminary knowledge which has been discussed in this paper so far can offer an alternative perspective in advancing radar and sonar system applications. By understanding how human echolocator exploit their waveform, it became clear on how to exploit the similar strategy in future man-made sensors such as radar and sonar systems.
6 References

[1] Vespe, M., Jones, G., Baker, C.J.: ‘Lessons for radar’, IEEE Signal Process. Mag., 2009, 26, (1), pp. 65–75
[2] Balleri, A., Leighton, T.G.: ‘Editorial: biologically-inspired radar and sonar systems’, IET Radar, Sonar Navig., 2012, 6, (6), pp. 507–509
[3] Supa, M., Cotzin, M., Dallenbach, K.M.: ‘“Facial vision”: the perception of obstacles by the blind’, Am. J. Psychol., 1944, 57, (2), p. 133
[4] Rice, C.E.: ‘Human echo perception’, Science (80-.), 1967, 155, (3763), pp. 656–664
[5] Schenkman, B.N., Nilsson, M.E.: ‘Human echolocation: blind and sighted persons ability to detect sounds recorded in the presence of a reflecting object’, Perception, 2010, 39, (4), pp. 483–501
[6] Thaler, L., Arnott, S.R., Goodale, M.A.: ‘Neural correlates of natural human echolocation in early and late blind echolocation experts’, PLoS One, 2011, 6, (5), pp. 1–16
[7] Baker, C.J., Smith, G.E.: ‘Aspects of cognition and echolocation’, Proc. of the 2012 IEEE Int. Symp. on Antennas and Propagation, Chicago, USA, 2012, pp. 1–2
[8] Thaler, L., Goodale, M.A.: ‘Echolocation in humans: an overview’, Wiley Interdiscip. Rev. Cogn. Sci., 2016, 7, (6), pp. 382–393
[9] Smith, G.E., Baker, C.J.: ‘Human echolocation waveform analysis’. IET Int. Conf. Radar Syst. (Radar 2012), Glasgow, UK, 2012, pp. 103–103
[10] Zhang, X., Smith, G.E., Thaler, L., et al.: ‘Echolocation in humans: waveform analysis of tongue clicks’, IEEE IET Lett., 2017, 53, (9), pp. 580–582
[11] Thaler, L., Reich, G.M., Zhang, X., et al.: ‘Mouth-clicks used by blind expert human echolocators – signal description and model based signal synthesis’, PLoS Comput. Biol., 2017, 13, (8), p. e1005670
[12] Raja Abdullah, R.S.A., Saleh, N.L., Abdul Rashid, N.E., et al.: ‘Bio-inspired signal detection mechanism for tongue click waveform used in human echolocation’, Electron. Lett., 2017, 53, (22), pp. 1456–1458
[13] Moore, B.C.J.: ‘An Introduction to the psychology of hearing’ (Brill, London, 2013, 6th Edn.)