Modelling of neural mechanisms involved in spatial hearing

Madalina Nastasa

Aalto Universtiy

Doctoral Programme CCIS / AAT

`madalina.nastasa@aalto.fi`

Abstract

Spatial hearing is a vital part of our day-to-day experiences, and a primary objective of computational modelling of neural mechanisms is to comprehend the mechanisms through which we utilise the differences between the signals that reach our ears. The use of binaural models has become prevalent in audio research for both understanding the auditory system and tackling engineering problems related to the detection and localisation of acoustic signals. In this review work, two types of physiological models are presented, namely: cross-correlation and count-comparison. Both the neurophysiology and neuroanatony knowledge utilised in the models is discussed. Although, it is still disputed which type of model fits best to the mammalian auditory system, both models are capable of describing common spatial hearing phenomena, and have the potential to further our understanding of the auditory system.

1 Introduction

Spatial hearing allows one to orientate in its own environment, and through the properties of reflected and reverberated sounds, it can also lead to deciphering the attributes of the listening environment and its acoustic properties (Pulkki and Karjalainen, 2015). In this review work, the main emphasis is placed on binaural hearing processes, as they are the foundation of our everyday listening experiences (Colburn, 1996). The term "binaural hearing" indicates the auditory system mode of functioning where the response of both ears are taken into consideration. Such tasks are typically related to both auditory localisation and detection, or recognition (Braasch, 2005).

The main goals of computational binaural modeling are to understand the mechanisms by which the auditory system takes advantage of the differences between

the two ear signals, and to solve engineering tasks related to the localisation and detection of acoustic signals. According to (Braasch, 2005), computational models are usually divided into physiologically- or psychologically-oriented approaches. The first type of model, based on neurophysiological studies, aims to simulate the behaviour of neuronal cells in detail, whilst the latter, based on psychophysical studies, functions on a more abstract phenomenological basis (Colburn, 1996). This work deals only with neural mechanisms, and addresses two types of physiological models: cross-correlation and count-comparison based.

The rest of this review work is organised as follows: chapter 2 summarises the binaural cues processed by the auditory system, while chapter 3 presents the neurophysiology and neuroanatony knowledge utilised in the models. Chapter 4 discusses the cross-correlation and count-comparison based models, and in chapter 5, the capacity of these models to describe spatial hearing phenomena is presented.

2 Psychoacoustical Background: Binaural Cues for Spatial Hearing

The main mechanisms of sound source localisation are the interaural time difference (ITD), interaural level difference (ILD) and the spectral filtering produced by the pinnae, head and torso (HRTF). The level difference and arrival-time differences between the signals arriving at the two ears are the basis of binaural hearing, which enable a fairly accurate spatial localisation of sound sources (Hudde, 2005). In this chapter, a brief overview of the interaural cues is presented in order to establish the foundational psychophysical data that is important for the understanding of the models discussed later on.

2.1 Interaural Time Difference

Interaural time difference (ITD) occurs as a result of the arrival time differences between the two ear signals, due to the differing distances the source has to travel to the ears. This means that the signal arriving at the contralateral ear is temporarily displaced compared to the ipsilateral one when the sound source is placed closer to the ipsilateral ear.

For low-frequency signals, up to 1.5 kHz, the auditory system is sensitive to the phase differences between the left and right ear signals (Pulkki and Karjalainen, 2009), whilst for higher frequencies, ITD cues are evaluated through envelope fluctuations in order to avoid to avoid phase ambiguities (Braasch, 2005).
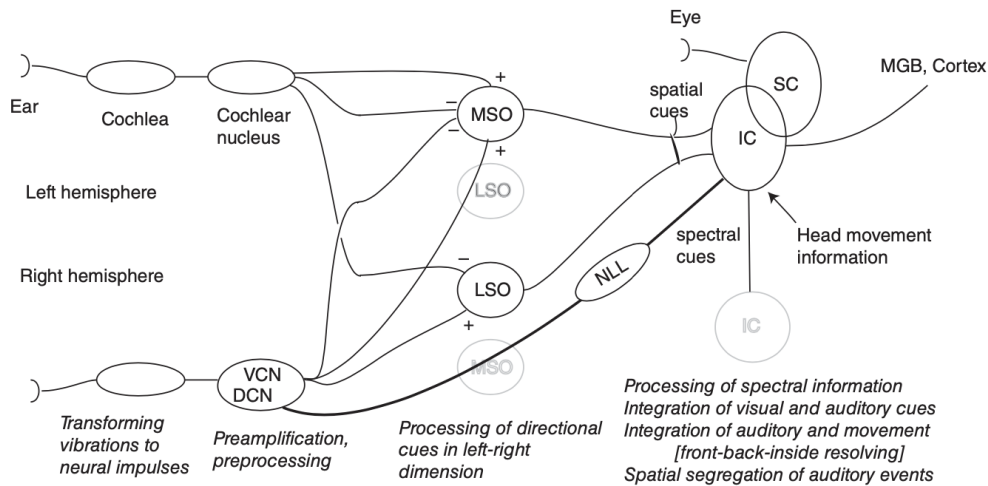
Figure 1: The organisation of the central auditory nervous system, showing the most important nuclei that are also approached in this work. Figure taken from Pulkki and Karjalainen (2015)

2.2 Interaural Level Difference

Interaural level difference (ILD) is the level difference between the two ear canal signals, primarily caused by the interaction of the sound waves with the head. A sound wave from a distant source arriving at the head causes reflection at the ipsilateral ear, which results in a pressure level increase, and shadowing at the contralateral ear, resulting in a decrease of pressure level (Pulkki and Karjalainen, 2005).

The phenomena of scattering is frequency dependent, thus the level differences between the ears increase as the frequency increases. However, although plane waves cause an ILD only at high frequencies, it has been shown that humans are sensitive to ILDs at all frequencies (Pulkki and Karjalainen, 2005).

3 Auditory pathway of binaural processing

The neurophysiology and neuro-anatomy of the binaural auditory pathway has been researched actively (Brugge et al., 1999; Eisenman, 1964; Mäkela and McEvoy, 1996; Oliver et al, 1995). This section offers a brief explanation of the neurophysiology involved in the mechanisms of binaural processing. Figure 1 shows the organisation of the central auditory system, with the focus placed on the main units of neural processing.
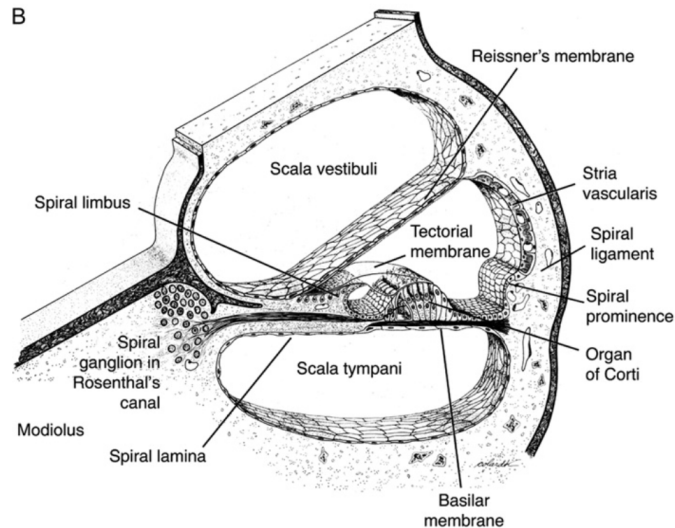
Figure 2: Cross-section of the cochlear duct. Figure taken from Pickles (1998)

3.1 Cochlea and cochlear nucleus

This section provides a functional overview of the cochlea and cochlea nucleus, as discussed by Hudde (2005). The cross-sectional structure of the cochlear duct, taken from [24], is shown in Figure 2, where the three scalae and associated structures are shown in a magnified view.

3.1.1 Cochlea

The cochlea consists of two and a half turns of a spiral-shaped duct, running from the base turn to the apex. The Reissner's and basilar membranes separate it into three fluid-filled scales, where the basilar membrane is mechanically the most important part inside the cochlea. Upon this membrane rests the organ of Corti which has receptors called hair cells.

The inner hair cells (IHC) are the basic sensors in the cochlea that convert motion into electrical spikes, while the outer hair cells have an active function to control the mechanical vibration state of the system. According to Hudde (2005), IHC are oblong flasks terminated at the top by a plate which carries a tuft of sensory hairs, commonly known as stereocilia. Bending of the stereocilia causes changes of the electrical potential inside the hair-cell body that result in neurotransmitters at synaptic junctions between the inner hair cells and the neurons of the auditory nerve (Pulkki and Karjalainen, 2009).

The main task of the cochlea is a spectral analysis of the input signals which are imposed by the vibrations occurring in the middle ear. The input signal is decomposed

into many parallel output signals, each of which represents a spectral component in a different frequency band. Furthermore, the cochlea has also a secondary active function, where neural activity can control the vibratory pattern of the basilar membrane. This phenomenon is referred to as cochlear amplifier because the outer hair cells can actively amplify the mechanical motion.

3.1.2 Cochlea Nucleus

The cochlear nuclear (CN) complex comprises of two cranial nerve nuclei located in the brainstem: the ventral cochlear nucleus (VCN) and the dorsal cochlear nucleus (DCN). Auditory nerve fibers carry information from the cochlea to the nerve root in the ventral cochlear nucleus. At the nerve root, the fibers branch to innervate the ventral cochlear nucleus and the deep layer of the dorsal cochlear nucleus.

Binaural processing circuits receive their input from the cochlear nucleus by way of the globular and the spherical bushy cells of the antero-ventral cochlear nucleus (AVCN) (Yin, 2002).

Bushy cells are one of several cell types in the cochlear nucleus that receive direct synaptic input from auditory nerve fibers. The physiological response properties of the SBCs and GBCs are so similar to that of their auditory nerve inputs that they are called primary-like (PL) and primary-like-with-notch (PLn), respectively. In particular, comparisons of the phase locking of cells in the AVCN with auditory nerve fibers indicated that the ability of bushy cells to phase lock was similar to or poorer than auditory nerve fibers.

3.2 Superior olivary complex

The AVCN projects its output to the superior olivary complex (SOC). This complex is divided into three primary nuclei: the medial superior olive (MSO), lateral superior olive (LSO), and the medial nucleus of the trapezoid body (MNTB). The nucleus is tonotopically organized, with low frequencies represented dorsally and high frequencies ventrally. Both the MSO and the LSO are thought to be sensitive to interaural differences of ear canal signals and help locate the azimuth of a sound.

3.2.1 Medial superior olive

The MSO is the major ITD-encoding structure in the mammalian auditory system and it receives direct excitations from the VCN on both sides of the brainstem. The source of this excitation is the spherical bushy cells (SBC), which time-lock their discharges to the temporal pattern of sounds (Grothe, 2003). The main cells in the MSO have two dendrites: medial which receive excitatory inputs from the

contralateral CN, and lateral which listen to the ipsilateral CN, respectively (Pulkki and Hirvonen, 2009).

The terms main lobe refers to the highest peak found in the MSO response, as studied in (Mcalpine et al., 2001) with white noise bursts. This response shows that, in the main lobe when the phase delay is between -90° and 45°, the response in one hemisphere increases with how much the contralateral ear signal leads the ipsilateral ear (Pulkki and Hirvonen, 2009). This is assumed to be the most predominant region in ITD decoding. With phase delays between 45° and 270°, this response decreases and, inherently, does not carry information about the sound source position.

3.2.2   Lateral superior olive

The LSO receives excitatory inputs from the ipsilateral ear through the spherical bushy cells (SBCs) of the ipsilateral anteroventral cochlear nucleus (AVCN), and inhibitory afferents from the contralateral ear via cells in the ipsilateral medial nucleus of the trapezoid body (Tollin, 2003). The response of the contralateral input is delayed in comparison to the ipsilateral input, as anatomically, the path to the LSO is longer from the contralateral than the ipsilateral ear.

The LSO is sensitive to ILD in such a way that the stronger the signal in the ipsilateral ear is, the stronger is the activity of the LSO (Pulkki and Hirvonen, 2009). The activity of an LSO neuron as a function of level difference between ear canal signals, as taken from (Tollin, 2003), follows with ILDs from 0dB to 15dB almost linearly. Furthermore, it has been found that LSO neurons act as phase-locked subtractors that can respond to very fast changes in input signals, and also to waveform differences at low-frequencies (Joris, 1995).

4   Models

Different types of models have been discussed by Colburn in (Colburn, 1978). For the purpose of this paper, the focus will be placed on cross-correlation and count-comparison models, with special attention placed to the latter because of its new anatomic discovered relations.

4.1   Cross-correlation based models

The coincidence detection model proposed by Jeffress (1948) is the most prominent model of binaural interaction that addresses the mechanisms for sensitivity to interaural time delays. Figure 3 shows a schematic illustration of this model, where the cochlear neural impulses propagate through a delay line to coincidence detection
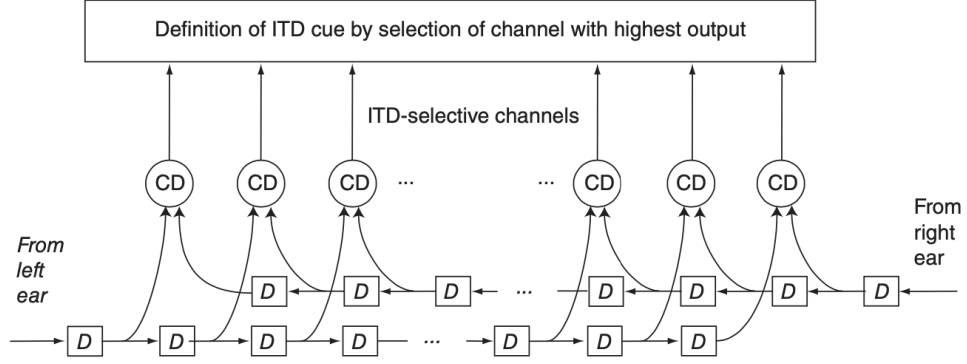
Figure 3: Schematic illustrating the coincidence detection model proposed by Jeffress (1948), where CD stands for coincidence-detector neuron and D is the delay unit. Figure taken from Pulkki and Karjalainen (2015)

neurons. Each cell in the network is maximally excited by a different ITD of the stimulus, and as a consequence, different neurons are sensitive to different ITDs.

This model is based on three main assumptions: (1) that the inputs to the binaural cells from both sides are carrying accurate timing information about the acoustic stimulus; this means that they are bilateral, time-locked or phase-locked; (2) that these ITD detector neurons perform coincidence detection (a maximal response occurs when the input spikes from the two sides arrive in coincidence and are therefore extremely sensitive to small difference in inter-arrival timing); and (3) that an arrangement of delay lines to adjust coincidence detector neurons to different preferred ITDs exists. The result of these three assumptions is a topographic map of ITDs across the axis of the nucleus parallel to the delay lines (Joris et al. 1998; Grothe, 2003; Yin, 2002).

The model proposed by Jeffress forms the foundation of all modern psychoacoustical and physiological models of binaural processing, with modern consensus being that the MSO is the likely site for encoding ITDs (Carr, 1990).

The model is typically implemented with a normalised cross-correlation function in auditory filterbanks with time lags tuned approximately within the physiologic ITD range, and it is only sensitive to ITD. The interaural cross-correlation, IACC, method that was introduced by Cherry and Sayers (1956) is defined as:

$$\psi_{yl,r}(\tau) = \frac{\int_{-\infty}^{\infty} y_l(t) \cdot y_r(t+\tau)dt}{\sqrt{\int_{-\infty}^{\infty} y_l^2(t)dt} \cdot \sqrt{\int_{-\infty}^{\infty} y_r^2(t)dt}}, \tag{1}$$

where the internal delay is τ, and the left and right sound pressure signals, $y_l(t)$ and $y_r(t)$.

The Jefress model has been extended to account for interaural level difference by Gaik (1993) and Breebaart (2001).

Research conducted suggests that the barn owl's nucleus laminaris (NL is the avian correspondent of the mammalian MSO) consists a full complement of delay lines within each frequency channel (Carr, 1990; Overholt et al., 1992; Joseph and Hyson, 1993; Reiss et al., 1996).

The azimuth of a sound source is believed to be encoded by a group of neurons tuned for ITDs within the physiological range of the barn owl. This model of sound localisation accords extremely well with the coincidence detection model proposed by Jeffress (1948) and it is normally taken veritably to explain sound localisation in mammals (Mcalpine et al., 2001).

However, a number of studies show empirical evidence that provide reasons to believe that a neural code in which the maximum discharge rate signals the ITD might not be sufficient to explain mammalian sensitivity to ITD (Smith et al.,1993; Beckius et al., 1999; Middlebrooks et al., 2002; Batra et al., 1997; Fitzpatrick et al., 2000).

4.2 Count-comparison models

This type of modelling has been introduced by von Békésy (1963), von Békésy et al. (1990) and van Bergeijk (1962). It works on the principle that the decoding of ITD and ILD functions on the comparison between the signal outputs of both hemispheres, such that earlier time-of-arrival or a louder signal in one ear canal produces a larger output to the contralateral hemisphere. The values of each hemisphere are compared together, hence giving it the name count-comparison model (von Békésy, 1963).

Furthermore, Pulkki and Hirvonen (2009) assume that the encoding method is self-normalised, which means that the comparison stage is not required and that the output of the nucleus in the hemisphere is already a meaningful directional coordinate. Thus, the resulting directional cue can be associated with the corresponding temporal position of the auditory band; this differs significantly from the Jeffress (1958) principle of modelling, where the output can be seen as a topographic mapping of auditory space.

4.2.1 Pulkki and Hirvonen (2009) Model

The general structure of the model is depicted in Fig. 4. The binaural input signal is fed into the two models of the periphery, one in each hemisphere. From the periphery model, the signal goes to the MSO and LSO that perform the spatial cue decoding.
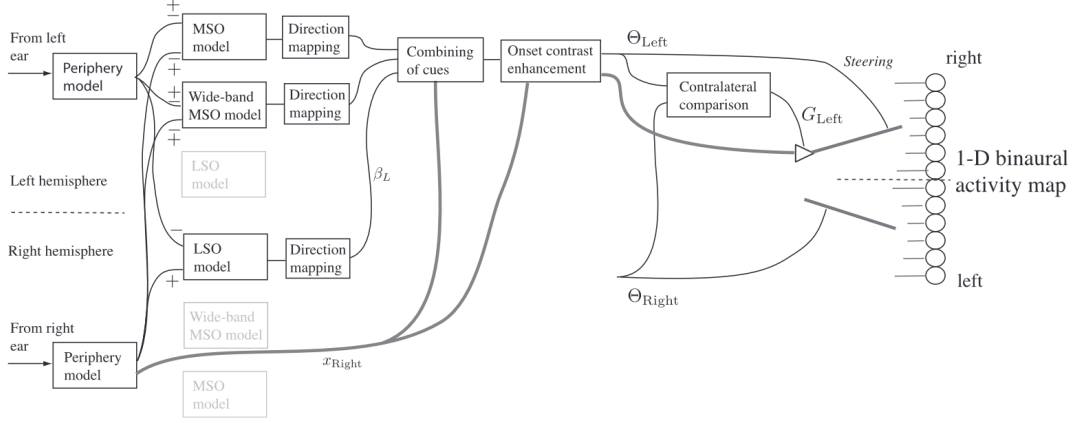
Figure 4: Schematic illustrating the count-comparison model proposed by Pulkki and Hirvonen (2009). Figure taken from Takanen (2014)

The earlier or higher in amplitude the signal arrives to one ear, the higher are the MSO and LSO model outputs arriving at the contralateral inferior colliculus level. The cues are expressed as left/right direction values depending on the time from separate narrow bandwidths, and also the MSO model extracts cues from a wide frequency range.

The cues from this stage are then merged together, individually for each frequency band, to form two directional cues, one on each hemisphere. The directional cues are then projected onto a one-dimensional binaural activity map.
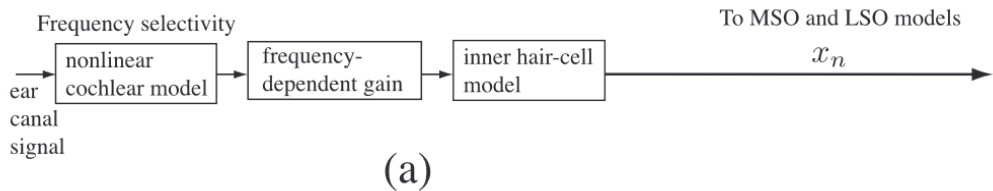
Periphery model

Figure 5: Block diagram of the periphery model. Figure taken from Takanen (2014)

The cochlea is modeled as a spectral analyser of the input, where the output is the response of the cochlear nerve fibers that are tuned to specific characteristic frequencies (CFs). It should be noted that all of the nuclei are modeled to transmit signals, which simulate the pooled responses of neurons sharing the same CF; the responses of single neurons were not considered in this study.

The structure of the periphery model is depicted in Fig. 5. The ear canal input signal is processed with a nonlinear time-domain model of the cochlea (Verhulst et

9

al., 2012), which provides the velocity and displacement of the basilar membrane at certain probe frequencies.

The next step in the periphery model consists of the emulation of the inner hair-cells (IHCs) with the model developed by Meddis (1986). The IHC model transforms the velocity of the basilar membrane movement to firing rate of the auditory nerve, and it is used as the output of the periphery model.
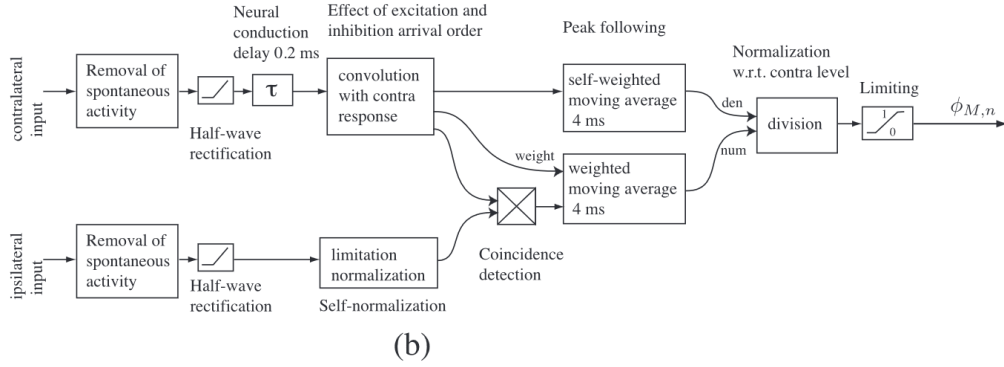
MSO Model

Figure 6: Block diagram of the MSO model. Figure taken from Takanen (2014)

Figure 6 displays the MSO model. The model operates in the following manner: Adjacent frequency channels from both hemispheres are sent to geometric averaging blocks, resulting in ipsi- and contralateral inputs for binaural interaction. The ipsilateral input is subject to self-normalization, while the binaural interaction is simulated through a continuous multiplication of the inputs in a sample-wise manner. The output is normalised using the contralateral input, creating the MSO model's output.
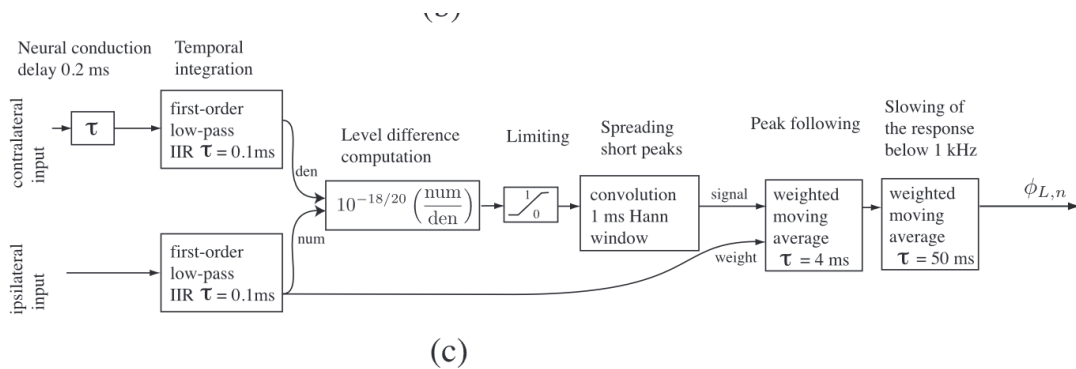
LSO Model

Figure 7: Block diagram of the LSO model. Figure taken from Takanen (2014)

Figure 7 depicts the LSO model, which operates by means of a fast subtractor mechanism that reacts promptly to instantaneous level disparities between the binaural signals. To approximate the propagation time difference between the contralateral and ipsilateral signals (Joris, 1996), the contralateral signal is delayed by 0.2 ms, much like in the MSO model. Additionally, both ipsilateral and contralateral inputs have their spontaneous activity rates removed uniformly, mirroring the MSO model's approach. The output of the LSO model is 0 when the contralateral signal has a greater level, and exceeds a maximum value of 1 when the difference in level between the ipsilateral and contralateral signals surpasses 18 dB.
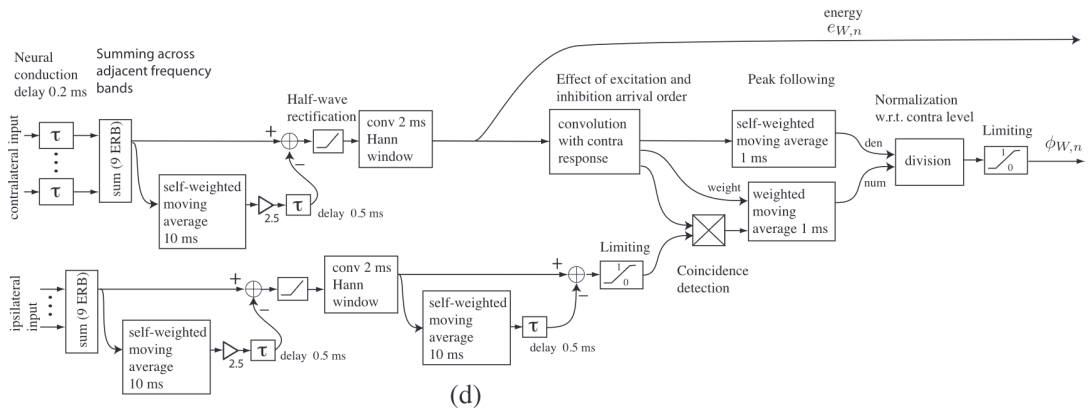
Wide-band MSO model

(d)

Figure 8: Block diagram of the wide-band MSO model. Figure taken from Takanen (2014)

Psychoacoustical experiments have demonstrated that listeners are able to localise broadband sounds based on envelope ITDs despite conflicting waveform ITDs (Braasch, 2005). Such an ability requires across-frequency integration of auditory filter outputs before binaural interaction.

It should be noted that although neurophysiological data does not specifically validate broadband processing in the MSO, research shows that human sensitivity to envelope ITDs in broadband sounds can be accounted for when the sound envelope is processed similarly to the narrowband MSO model.

The wide-band MSO model's binaural interaction employs the same method as the narrowband MSO model, and this can be observed in Figure 8. The removal of spontaneous activity from the ipsilateral and contralateral inputs is likewise implemented identically to the MSO model.

Unlike the narrowband MSO model, the inputs for the wide-band MSO model incorporate across-frequency interaction by summing nine adjacent frequency bands together. The range of the summing, therefore, corresponds to a width of nine ERBs for each frequency band. As a result, the pulse nature of the input is lost while the

output follows the envelope of the given input at each frequency band. Nonetheless, although the required envelope signals are acquired, they cannot serve as direct input for the binaural interaction implementation due to the model's sensitivity to temporal displacements of the pulses in the signal.

4.2.2   Model response in simple binaural conditions

The model was tested with many simple binaural listening simulations which suggested that it could validate psychoacoustic data qualitatively (Pulkki and Hirvonen, 2019). The count-comparison modeling approach appears to be valid for modeling psychoacoustic listening tests. Based on the simulations, it is suggested that each hemisphere independently processes sound source direction from the MSO and LSO cues, with a greater weighting placed on the cue suggesting direction further from the median plane. Consequently, both ICs may receive cues indicating lateral directions, leading the listener to perceive an auditory object in both hemispheres.

The LSO is shown to be sensitive to the ILD at high frequencies and to both the ILD and ITD at low frequencies. The LSO model decodes the left/right direction of the sound source monotonically at all frequencies but with lower resolution at low frequencies than at high ones. On the other hand, the MSO model output appears mostly used in enhancing the directional accuracy near the median plane, and it is sensitive only to limited ITD values at low frequencies. The MSO model is not responsive to the ILD.

Tests with incoherent ear canal signals reveal that the temporal fluctuations of the LSO model output carry information on the coherence between ear canal signals, accounting for interaural coherence perception and the BMLD. In contrast, the output of the MSO model was low, whereas the LSO model output was high for the incoherent input. The simulations suggest that the MSO gives a response only when the ear canal signals are coherent.

5   Conclusions

There are multiple hypotheses regarding the biological systems involved in spatial hearing. A common assumption is that the auditory system consists of a high number of interaural delay lines that are tuned to different interaural delays and show a coincidence reaction when their suitable delay actually occurs.

However, this cross-correlation approach was challenged by physiological studies that have shown that the coincidence mechanism cannot be found in mammals, and that the mammalian ITD decoding is a much more complex process. As a result, some believe that a count-comparison model fits better to the processes conducted by the auditory system. This model works on the principle that the decoding of ITD and ILD functions on the comparison between the signal outputs of both hemispheres,

such that earlier time-of-arrival or a louder signal in one ear canal produces a larger output to the contralateral hemisphere.

It is still under dispute whether the Jeffress delay-line model or the count-comparison model correctly represents the human auditory system, since the human ITD mechanism cannot be studied directly on a neural basis. Despite this, both models are capable of describing common spatial hearing phenomena and have the potential to further our understanding of the auditory system.

Nomenclature

$AVCN$ Anteroventral cochlea nucleus

CF Characteristic Frequency

CN Cochlea nucleus

DCN Dorsal cochlea nucleus

GBC Globular bushy cells

$HRTF$ Head related transfer function

IHC Inner hair cell

ILD Interaural level difference

ITD Interaural time difference

LSO Lateral superior olive

MSO Median superior olive

OHC Outer hair cell

SBC Spherical bushy cells

SOC Superior olivary complex

VCN Ventral cochlea nucleus

6 References

References

[1] BATRA, R., KUWADA, S., AND FITZPATRICK, D. C. Sensitivity to interaural temporal disparities of low-and high-frequency neurons in the superior olivary complex. i. heterogeneity of responses. *Journal of Neurophysiology 78*, 3 (1997), 1222–1236.

[2] BECKIUS, G. E., BATRA, R., AND OLIVER, D. L. Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *Journal of Neuroscience 19*, 8 (1999), 3146–3161.

[3] BRAASCH, J. Modelling of binaural hearing. In *Communication acoustics.* Springer, 2005, pp. 75–108.

[4] BREEBAART, J., VAN DE PAR, S., AND KOHLRAUSCH, A. Binaural processing model based on contralateral inhibition. i. model structure. *The Journal of the Acoustical Society of America 110*, 2 (2001), 1074–1088.

[5] BRUGGE, J. F., DUBROVSKY, N. A., AITKIN, L. M., AND ANDERSON, D. J. Sensitivity of single neurons in auditory cortex of cat to binaural tonal stimulation; effects of varying interaural time and intensity. *Journal of neurophysiology 32*, 6 (1969), 1005–1024.

[6] CARR, C., AND KONISHI, M. A circuit for detection of interaural time differences in the brain stem of the barn owl. *Journal of Neuroscience 10*, 10 (1990), 3227–3246.

[7] CHERRY, E. C., AND SAYERS, B. M. A. "human 'cross-correlator'"—a technique for measuring certain parameters of speech perception. *The Journal of the Acoustical Society of America 28*, 5 (1956), 889–895.

[8] COLBURN, H. S. Computational models of binaural processing. In *Auditory computation.* Springer, 1996, pp. 332–400.

[9] COLBURN, H. S., AND DURLACH, N. I. Models of binaural interaction. *Handbook of perception 4* (1978), 467–518.

[10] FITZPATRICK, D. C., KUWADA, S., AND BATRA, R. Neural sensitivity to interaural time differences: beyond the jeffress model. *Journal of Neuroscience 20*, 4 (2000), 1605–1615.

[11] GAIK, W. Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling. *The Journal of the Acoustical Society of America 94*, 1 (1993), 98–110.

[12] GROTHE, B. Sensory systems: New roles for synaptic inhibition in sound localization. *Nat. Reviews Neurosci 4*, 540.

14

[13] Hudde, H. A functional view on the peripheral human hearing organ. In *Communication acoustics.* Springer, 2005, pp. 47–74.

[14] Jeffress, L. A. A place theory of sound localization. *Journal of comparative and physiological psychology 41*, 1 (1948), 35.

[15] Joris, P. X. Envelope coding in the lateral superior olive. ii. characteristic delays and comparison with responses in the medial superior olive. *Journal of neurophysiology 76*, 4 (1996), 2137–2156.

[16] Joris, P. X., Smith, P. H., and Yin, T. C. Coincidence detection in the auditory system: 50 years after jeffress. *Neuron 21*, 6 (1998), 1235–1238.

[17] Joseph, A. W., and Hyson, R. L. Coincidence detection by binaural neurons in the chick brain stem. *Journal of Neurophysiology 69*, 4 (1993), 1197–1211.

[18] Mäkelä, J., and McEvoy, L. Auditory evoked fields to illusory sound source movements. *Experimental brain research 110* (1996), 446–454.

[19] McAlpine, D., Jiang, D., and Palmer, A. R. A neural code for low-frequency sound localization in mammals. *Nature neuroscience 4*, 4 (2001), 396–401.

[20] Meddis, R. Simulation of mechanical to neural transduction in the auditory receptor. *The Journal of the Acoustical Society of America 79*, 3 (1986), 702–711.

[21] Middlebrooks, J. C., Xu, L., Furukawa, S., and Macpherson, E. A. Book review: Cortical neurons that localize sounds. *The Neuroscientist 8*, 1 (2002), 73–83.

[22] Oliver, D. L., Beckius, G. E., and Shneiderman, A. Axonal projections from the lateral and medial superior olive to the inferior colliculus of the cat: a study using electron microscopic autoradiography. *Journal of Comparative Neurology 360*, 1 (1995), 17–32.

[23] Overholt, E. M., Rubel, E. W., and Hyson, R. L. A circuit for coding interaural time differences in the chick brainstem. *Journal of Neuroscience 12*, 5 (1992), 1698–1708.

[24] Pickles, J. An introduction to the physiology of hearing. In *An Introduction to the Physiology of Hearing.* Brill, 1998.

[25] Pulkki, V., and Hirvonen, T. Functional count-comparison model for binaural decoding. *Acta acustica united with Acustica 95*, 5 (2009), 883–900.

[26] Reyes, A. D., Rubel, E. W., and Spain, W. In vitro analysis of optimal stimuli for phase-locking and time-delayed modulation of firing in avian nucleus laminaris neurons. *Journal of Neuroscience 16*, 3 (1996), 993–1007.

[27] SMITH, P. H., JORIS, P. X., AND YIN, T. C. Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *Journal of Comparative Neurology 331*, 2 (1993), 245–260.

[28] TAKANEN, M., SANTALA, O., AND PULKKI, V. Visualization of functional count-comparison-based binaural auditory model output. *Hearing Research 309* (2014), 147–163.

[29] TOLLIN, D. J. The lateral superior olive: a functional role in sound source localization. *The neuroscientist 9*, 2 (2003), 127–143.

[30] VAN BERGEIJK, W. A. Variation on a theme of békésy: A model of binaural interaction. *Journal of the Acoustical Society of America 34* (1962), 1431–1437.

[31] VAN DER HEIJDEN, K., RAUSCHECKER, J. P., DE GELDER, B., AND FORMISANO, E. Cortical mechanisms of spatial hearing. *Nature Reviews Neuroscience 20*, 10 (2019), 609–623.

[32] VERHULST, S., DAU, T., AND SHERA, C. A. Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission. *The Journal of the Acoustical Society of America 132*, 6 (2012), 3842–3848.

[33] VON BÉKÉSY, G., WEVER, E., AND PEAKE, W. T. Experiments in hearing. *Acoustical Society of America Journal 88*, 6 (1990), 2905.

[34] VON BÉKÉSY, G., WEVER, E. G., AND PEAKE, W. T. Experiments in hearing.

[35] YIN, T. Envelope coding in the lateral superior olive. i. sensitivity to interaural time differences. *Journal of neurophysiology 73*, 3 (1995), 1043–1062.

[36] YIN, T. C. Neural mechanisms of encoding binaural localization cues in the auditory brainstem. In *Integrative functions in the mammalian auditory pathway*. Springer, 2002, pp. 99–159.