Modeling pitch perception of complex tones

Citation for published version (APA):

DOI:
10.1121/1.2023328

Document status and date:
Published: 01/01/1986

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
characteristics predicts the observed patterns well. In particular, the temporal patterns are determined by the model of a cochlear fiber [E. F. Evans, J. Physiol. 298, 6-7 (1980)] preened by a linear filter having these cochlear filter. Two important consequences of this follow. One is that the dominant region for pitch could be degree to which individual harmonics of a complex stimulus interact, after appropriate weighting by the bandwidth characteristics of the fiber’s frequency threshold (tuning) curve (FTC). A simple probabilistic approach is used to predict the temporal discharge patterns in response to such complex stimuli. As an extreme case, the effects of variants of the click trains, whose pitch has been the subject of controversy between Whitfield and Moore to predict the temporal discharge patterns of neurones in the auditory periphery, would be easier to test if it were possible to determine whether the fiber’s frequency threshold (tuning) curve (FTC) was bandpass, low-pass, or high-pass. Some of these models have been successful and appears to be supported by more recent experimental findings. Thus critical reconsideration and analysis of its essentials, i.e., indispensable characteristics, structures, and parameters seems appropriate. Particular attention is paid to the characteristics of aural spectrum analysis, the dichotomy of spectral versus virtual pitch, and the question of learning processes. Some crucial phenomena are demonstrated using tape examples.

8:55

CC2. Modeling pitch perception of complex tones. Adrianus J. M. Houtsma (Institute for Perception Research, P. O. Box 513, 5600 MB Eindhoven, The Netherlands)

When one listens to a series of harmonic complex tones that have no acoustic energy at their fundamental frequencies, one usually still hears a melody that corresponds to those missing fundamentals. Since it has become evident some two decades ago that neither Helmholtz’s difference tone theory nor Schouten’s residue theory could adequately account for this phenomenon, several other theories have been proposed that accentuate central neural rather than peripheral mechanical signal processing. Some of these theories will be critically reviewed against empirical evidence from recent psychoacoustic studies. In particular, the relative advantages and disadvantages of the descriptive “virtual pitch” theory of Terhardt and the stochastic “optimum processing” theory of Goldstein are discussed in relation with recent data on pitch perception for simultaneous complex tones. Taped examples of some of the studied phenomena are provided.

9:30

CC3. Criteria and constraints of modeling the perception of pitch. Ernst Terhardt (Institute for Electroacoustics, Technical University, Arcistr. 21, D-8000 Munich 2, West Germany)

The vast amount of data and observations relevant to the perception of pitch has grown through many decades and at first did not seem to support case of modeling; it rather elucidated that pitch modeling is much more complex than just measuring “the fundamental frequency” in either the frequency or time domain. While the concept of auditory spectrum analysis is enormously helpful and successful in many respects, it failed to provide a direct explanation of the pitch of many ordinary sounds such as speech and music. Various time-domain explanations, on the other hand, proved unsatisfactory as well. In the late 1960s the conclusion was drawn that modeling pitch perception requires careful inclusion, and combination, of signal-theoretical, psycho-physical, and Gestalt-psychological methods. The present author’s “virtual-pitch theory” was worked out on that line. As an additional significant outcome of the new approach, it turned out to be not quite just a pitch model, but reflects and illustrates significant general principles of auditory perception and, in particular, of music perception. Until present, that type of modeling has been successful and appears to be supported by more recent experimental findings. Thus critical reconsideration and analysis of its essentials, i.e., indispensable characteristics, structures, and parameters seems appropriate. Particular attention is paid to the characteristics of aural spectrum analysis, the dichotomy of spectral versus virtual pitch, and the question of learning processes. Some crucial phenomena are demonstrated using tape examples.

10:05-10:15

Break

10:15

CC4. Physiological observations on the neural representation of complex pitch. E. F. Evans (Department of Communication and Neuroscience, University of Keele, Keele, Staffordshire ST5 5BG, United Kingdom)

Pitch models which depend on the manner in which the frequencies of pitch-evoking stimuli are encoded in the temporal discharge patterns of neurones in the auditory periphery, would be easier to test if it were possible to predict the temporal discharge patterns in response to such complex stimuli. As an extreme case, the effects of variants of click trains, whose pitch has been the subject of controversy between Whitfield and Moore (see Evans’ chapter in Hearing—Physiological Bases and Psychophysics, edited by R. Klinke and R. Hartmann (Springer, New York, 1985) have been studied at the level of the cat’s cochlear nerve. In particular, the following stimuli have been used on fibers with characteristic frequencies up to 4 kHz: pulse trains with even intervals of 5, 10, 20 ms; uneven intervals of 4.5/5.5 ms, 4.6/5.4 ms, 4.8/5.2 ms, and even intervals of 2.5 and 5 ms with alternate click phases. Autocorrelations have been performed of the spike discharge patterns evoked by the stimuli. To a first approximation, temporal patterns of discharges evoked by the even and uneven interval click trains can be predicted on the basis of simple linear filtering, with filters having the shape and bandwidth characteristics of the fiber’s frequency threshold (tuning) curve (FTC). A simple probabilistic model of a cochlear fiber [E. F. Evans, J. Physiol. 298, 6-7 (1980)] preceded by a linear filter having these characteristics predicts the observed patterns well. In particular, the temporal patterns are determined by the degree to which individual harmonics of a complex stimulus interact, after appropriate weighting by the cochlear filter. Two important consequences of this follow. One is that the dominant region for pitch could be considered to be the range of cochlear fiber characteristic frequencies over which sufficient spectral resolution