

Book review for Perception

Yoshizawa, T. (2023). (ed.) *A sense of plausibility in vision and music perception*. Tokyo, Japan: Asakura Publishing Company, Ltd. 109 pp. £24.00 (paperback), ISBN 9784254520354.

Reviewed by: Hiroshi Ashida

Graduate School of Letters, Kyoto University, Kyoto, Japan

This book consists of various papers as proceedings of a series of international symposia held at Kanagawa University from 2019 to 2022. ‘Sense of plausibility’ in the title of the book might be a bit obscure, but we can understand it as an umbrella term for otherwise diverse topics. Part 1 and 2 cover the fields of vision and music, respectively, mainly on the perception side. One might wonder why music is a counterpart to vision, but this is the most unique characteristic of this book and I was curious how these two distinct fields could be linked. There are five papers on vision and four on music and speech, each providing some basic reviews for non-experts and then presenting experimental results conducted by each author. Most chapters deal with psychophysical or psychological studies, also including some biological and computational approaches. Each chapter is independent, and readers can just pick up any part to start with.

Chapter 1 by the editor Yoshizawa begins with the visual sense of plausibility, which is the central theme of the whole book. He discusses the general sense, or ‘atmosphere’, of cityscape, and extends the idea into material perception, introducing their own experimental study on perceived surface colour with random shuffling noise. The following chapters take ‘plausibility’ per se as the key concept, rather than the sense of it. Typically, Leonards (Chapter 5) features the concept of ‘perceptual plausibility’ in discussing modern city design in terms of goal-directed action of walking. She defines it in terms of the prediction error with the feedback from other senses, which is related to the signal-to-noise ratio (SNR) in the visual scene. Visual noise such as high-contrast patterns that is often utilised in urban design affects visual prediction and thus sometimes disturbs walking, as typically seen in a ‘tilted’ bridge in her town of Bristol that even gives some people vertigo. Scott-Samuel (Chapter 2) takes a similar but reversed view of plausibility in reviewing the current directions of studies on biological camouflage; plausibility is defined on the hider’s side that indicates how well an animal can ‘fit in’ with the environment by effectively disrupting (otherwise plausible) visual perception of the observer such as a predator. Scott-Samuel also utilises SNR to explain successful camouflage, but obviously in the opposite way from Leonards’ definition such that reduction of SNR leads to better plausibility. Other two chapters on vision deal with somewhat more specific topics: visual field asymmetry (left-right) in the perception of colour-defined motion by Kojima (Chapter 3), and enhanced perceived contrast with binocular disparity by Maehara (Chapter 4). Each of these two

chapters includes brief introductions to the early visual pathways that leads to separate processing of left and right visual fields and stereoscopic depth perception. The authors do not state it explicitly, but these chapters describe implausible aspects of visual perception that are yet to be explained.

Part 2 of the book on music and speech begins with Ono's discussion on grouping of sounds and rhythm tapping (Chapter 6), which also tries to link vision and music by the common principle of grouping. If rhythm is often considered as the first of the major elements of music, the next chapter by Matsunaga (Chapter 7) considers the tonal organization of melody, which could be the second element. Perception of the key (or the tonal centre) from simple melodies could also be related to the third element of harmony. Then Shoda (Chapter 8) discusses the audience effect, that is, how the existence of audience affects musical performance. Each of these three chapters could be read as introductions to music studies for non-experts. The final paper by Asakura introduces his model of adaptation to the delayed feedback effect on speech (Chapter 9). This chapter is a bit isolated, I am afraid, and I wonder if the Kalman filter, which has been used for both vision and music, could have been discussed a little more generally from the computational perspective by way of linking the two fields in the book. Papers in part 2 are not explicit about the central topic of plausibility, although each implies some aspects of plausible perception and action.

The quality of writing is variable across chapters, with some chapters being less easy to follow than others, but each chapter has considerations for non-experts and I found no problem in grasping the basic message of each chapter. I enjoyed reading both parts, as a vision scientist with a bit of experience in playing music, but especially part 2 on scientific studies on music that are quite tempting to try myself one day. High skills or knowledge in music, such as complex score reading, are not required to understand these chapters. To my knowledge, surprisingly many people in the field of visual perception have interests and experiences in music, and this book would be highly recommended to those. I am less sure for the music scientists, but it would be good for them to see what sort of things vision scientists are up to, which hopefully could lead to more collaborations across these fields in the near future.