Haptic perception in interaction with other senses

Hannah B. Helbig and Marc O. Ernst

Introduction

Human perception is inherently multisensory: we perceive the world simultaneously with multiple senses. While strolling the farmers market, for example, we might become aware of the presence of a delicious fruit by its characteristic smell. We might use our senses of vision and touch to identify the fruit by its typical size and shape and touch it to select only that one with the distinctive soft texture that signals 'ripe'. When we take a bite of the fruit, we taste its characteristic flavour and hear a slight smacking sound which confirms that the fruit we perceive with our senses of vision, touch, audition, smell and taste is a ripe, delicious peach. That is, in the natural environment the information delivered by our sense of touch is combined with information gathered by each of the other senses to create a robust percept. Combining information from multiple systems is essential because no information-processing system, neither technical nor biological, is powerful enough to provide a precise and accurate sensory estimate under all conditions. In this chapter, we address the question how the human brain combines sensory information in order to form a robust, reliable and coherent percept of the world around us. Most of the theoretical considerations we discuss here are not confined to interactions of touch with the other senses but generalise to interactions among the other senses (e.g., visual-auditory). Therefore, we keep these considerations general but focus on examples demonstrating how haptics is combined with other sensory modalities.

Information that is perceived through different sensory pathways can be qualitatively different: the senses can provide either complementary or redundant information. Redundant sensory signals provide information about the same sort of object property (e.g., size) and are represented in a common frame of reference, in the same units. To stay with the example of the peach, vision and touch provide redundant information about the peach's size or shape. In contrast, vision and taste provide complementary information about the identity of the object.

In general, we benefit from integrating multiple sources of information. Combining complementary information is advantageous because it extends the range and variety of what can be perceived from any one sense in isolation [1] and can reduce perceptual ambiguity. Furthermore, integrating multiple sensory sources usually leads to improved perceptual performance, more precise judgments and enhances detection of stimuli (e.g., [2–11]).

But what are the mechanisms behind multisensory integration? And what happens if our senses deliver conflicting information? Imagine you grasp a coin underwater. It is a foreign coin that you have never seen before. The visual image of the coin underwater is optically distorted by the refraction of the light rays as they pass from water to air and thus looks larger as it actually is (e.g., [12, 13]) (which you are not aware of). You are asked to judge the size of the coin. Do you rely on your sense of vision, or do you trust more your sense of touch, or do you perceive something in-between the felt and seen size?

Early work suggested that vision is our dominant sense: It was argued that vision 'wins' when visual information conflicts with information from other sensory modalities (e.g., [14–19]).

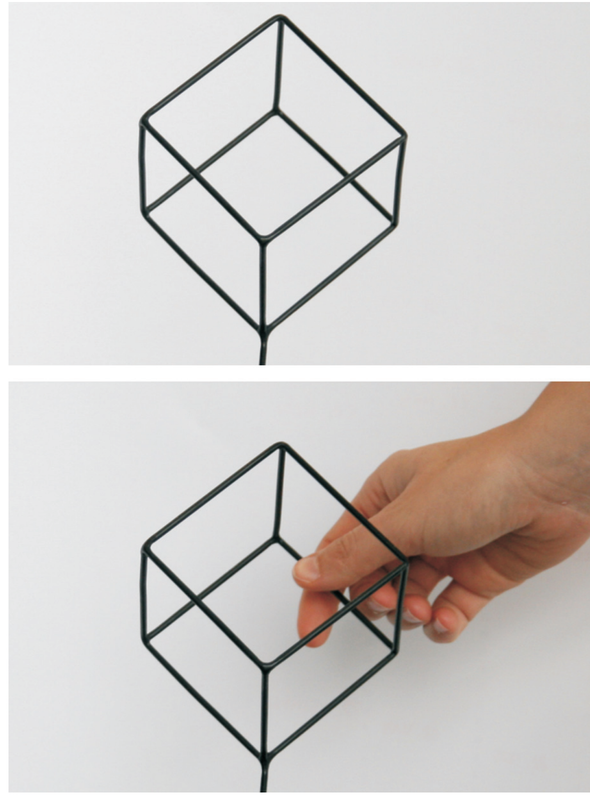

Figure 1.
The Necker cube is a wire-frame cube in isometric perspective that can be interpreted in two different, equally probable, ways and thus leads to a bistable percept. When a person stares at the picture, it will often seem to flip back and forth between the two valid interpretations. Haptic information can be used to disambiguate the percept.

For example, in a classic experiment by Rock and Victor [18] observers felt a square and looked at it through a lens system that transformed the visual input and made the square look like an oblong rectangle. Observers were asked to select an object from among a set of comparison stimuli that matched the perceived shape. The reported percept was almost completely dominated by the visual input which led to the notion that vision is our dominant sense in conditions where the senses deliver conflicting sensory information ('visual capture').

Later work converged on a more balanced view: According to the modality-appropriateness hypothesis [10, 20], the various sensory modalities are differentially well suited for different kinds of perceptual tasks and discrepancies are resolved in favour of the modality that is more appropriate (e.g., faster, more precise, more accurate) for the task at hand. According to this hypothesis, vision dominates audition for processing spatial information; touch is more appropriate than audition for the perception of texture and audition dominates vision for temporal judgments. For example, a repetitive sound (auditory flutter) presented simultaneously with a flickering light causes the rate of perceived visual flicker to shift towards the auditory flutter rate (e.g., [21, 22]).

More recently, it has been shown that the integrated percept is not simply dominated by either one or the other sensory modality. Rather, when there are several cues available to estimate a certain property – may it be cues within one sensory modality or across modalities – both cues contribute to the combined estimate and the weight of each sensory cue systematically shifts with its relative reliability (e.g., [23–38]).

Building upon this idea, recent research applied a more quantitative approach to study sensory cue integration (e.g., [2, 4, 39–45]). Optimal integration behaviour has been modelled within a Bayesian framework: According to an optimal integrator, the optimal combined estimate (Maximum Likelihood Estimate, MLE) is a linear combination of the individual unimodal estimates that are weighted by their relative reliabilities. More reliable cues are assigned a larger weight. Integrating redundant sensory information in this optimal fashion decreases the variance of the integrated estimate and yields the most reliable unbiased estimate. The Bayesian framework and the implementation of an optimal integrator within this framework are described in more detail in the next section ("The Bayesian framework and optimal integration").

Besides weighted averaging of sensory information, there are situations in which sensory cues are combined in a non-linear fashion (e.g., [34, 46]): 'sensory cooperation' (cues interact in a nonlinear fashion), 'disambiguation' (one

HOMO-HAPTICUS.DE - BOOK HUMAN HAPTIC PERCEPTION, GRUNWALD M (ED.) - BIRKHÄUSER