

Cognitive Critique



THE INTEGRATION OF VISION AND HAPTIC SENSING: A COMPUTATIONAL & NEURAL PERSPECTIVE

JOSHUA AMAN

*Human Sensorimotor Control Laboratory
University of Minnesota*

E-MAIL: aman0038@umn.edu

CHIA-HAO LU

*Human Sensorimotor Control Laboratory
University of Minnesota*

E-MAIL: luxxx214@umn.edu

JÜRGEN KONCZAK

*Human Sensorimotor Control Laboratory
Center for Clinical Movement Science
University of Minnesota*

E-MAIL: jkonczak@umn.edu

Accepted 12 May, 2010

KEYWORDS

perception, psychophysics, integration, tactile, visual

ABSTRACT

When looking at an object while exploring and manipulating it with the hands, visual and haptic senses provide information about the properties of the object. How these two streams of sensory information are integrated by the brain to form a single percept is still not fully understood. Recent advances in computational neuroscience and brain imaging research have added new insights into the underlying mechanisms and identified possible brain regions involved in visuo-haptic integration. This review examines the following main findings of previous research: First, the notion that the nervous system combines visual and haptic inputs in a fashion that minimizes the variance of the final percept and performs operations commensurable to a maximum-likelihood integrator. Second, similar to vision, haptic information may be mediated by two separate neural pathways devoted to perception and action. This claim is based on a set of psychophysical studies investigating how humans judge the size and orientation of objects. Third, a cortical neural system described as the lateral occipital complex (LOC) has been identified as a possible locus of visuo-haptic integration. This claim rests on functional imaging studies revealing an activation of LOC to both visual and haptic stimulation. We conclude that much progress has been made to provide a computational framework that can formalize and explain the results of behavioral and psychophysical studies on visuo-haptic integration. Yet, there still exists a gap between the computationally driven studies and the results derived from brain imaging studies. One reason why the closing of this gap has proven to be difficult is that visuo-haptic integration processes seem to be highly influenced by the task and context.

INTRODUCTION

In order to grasp and manipulate an object with our hands, we need information about its size, shape, texture and mass. This information is in many cases conveyed by the visual and haptic senses. While visual perception is unimodal relying on information derived from the retinal images, haptic perception is based on multisensory information. In most general terms, haptic perception can be defined as the sensibility of an individual to its adjacent surroundings by the use of the subject's body that involves the integration of proprioceptive, tactile, and pressure cues across time (Gibson 1966). Most often, it refers to the extraction of information about an

object's properties through exploratory actions usually of the hands and arms (Henriques and Soechting 2005) and is linked closely to *active touch* or *active palpation* (Amedi et al. 2002). While the importance of vision for object recognition and action planning has long been recognized (von Helmholtz 1867), the role of the haptic sense for object recognition and manipulation has received considerably less attention. This may be partly because the study of haptics poses extra challenges due to its multimodal nature (Henriques and Soechting 2005). During the exploration and manipulation of objects with our hands, usually both haptic and visual information are available to the perceptual system. Computationally, combining both streams of information into a single percept requires an integration process between vision and haptics. It is this process of visuo-haptic integration that is the focus of this review. The review comprises two main sections: The first section examines recent attempts to understand the mechanisms of visuo-haptic integration from a computational perspective. The second section addresses the neural basis of visuo-haptic integration by reviewing relevant neuroanatomical and functional imaging studies.

COMPUTATIONAL VIEWS OF VISUO-HAPTIC INTEGRATION

BAYESIAN APPROACHES TO SENSORY INTEGRATION

The amount of information a sensory signal can carry about the environment or the state of the body is limited by the properties of its sensors and by the type and quality of the signal's transduction to further processing centers. These limitations become manifest as neural noise, and this noise ultimately affects the precision of our perceptions and motor acts, and leads to variable errors and uncertainty (van Beers et al. 2002). One way to overcome this uncertainty and to improve the precision of perceptual estimates (e.g., judging the size of an object) is to combine or integrate redundant information from several sensory signals. This integration process may be understood by using the Bayesian concept of probability.

The Bayesian framework has its origins in von Helmholtz's idea of perception as *unconscious inference* (von Helmholtz 1867). For von Helmholtz such an inference process seemed necessary, because sensory information may be ambiguous. For example, two

cylinders of identical dimensions produce exactly the same retinal image if viewed from the appropriate angles, but their weights can be very different. In this context, Bayesian statistical decision theory can be seen as a tool that helps to formalize von Helmholtz's idea of perception as inference (see Kersten et al. 2004 for a detailed review).

In recent years several studies have applied methods based on Bayesian inference to capture the integration between vision and haptics (Ernst and Banks 2002; Gori et al. 2008; Helbig and Ernst 2008). A general principle derived from this work is that the perceptual system strives to produce a lowest-variance estimate of the object property. An important study by Ernst and Banks (2002) reported that humans integrate visual and haptic information in a statistically optimal fashion. They asked human observers to discriminate between the heights of two objects under three conditions of presentation: visual alone, haptic alone, and simultaneous presentation of visual and haptic size stimuli. Subsequently, they determined visual, haptic, and visual-haptic thresholds to investigate whether human sensory integration followed a maximum likelihood estimation (MLE) rule, as follows. Ernst and Banks (2002) state that an estimate of an environmental property by a sensory system can be represented by

$$\hat{S}_i = f_i(S) \quad (1)$$

where S is the physical property being estimated and f is the operation by which the nervous system performs the estimation. The subscript i refers to the modality or to different cues within a modality. Each estimate, \hat{S}_i , is corrupted by noise. If the noises are independent and Gaussian with variance σ_i^2 , and the Bayesian prior is uniform, the MLE of the environmental property is then given by

$$\hat{S} = \sum_i w_i \hat{S}_i \quad , \quad w_i = \frac{1/\sigma_i^2}{\sum_j 1/\sigma_j^2} \quad (2)$$

where w_i represents the relative weight of visual or haptic estimate, which is inversely proportional to the variances of individual estimates. The subscript j refers to the visual and haptic modalities.

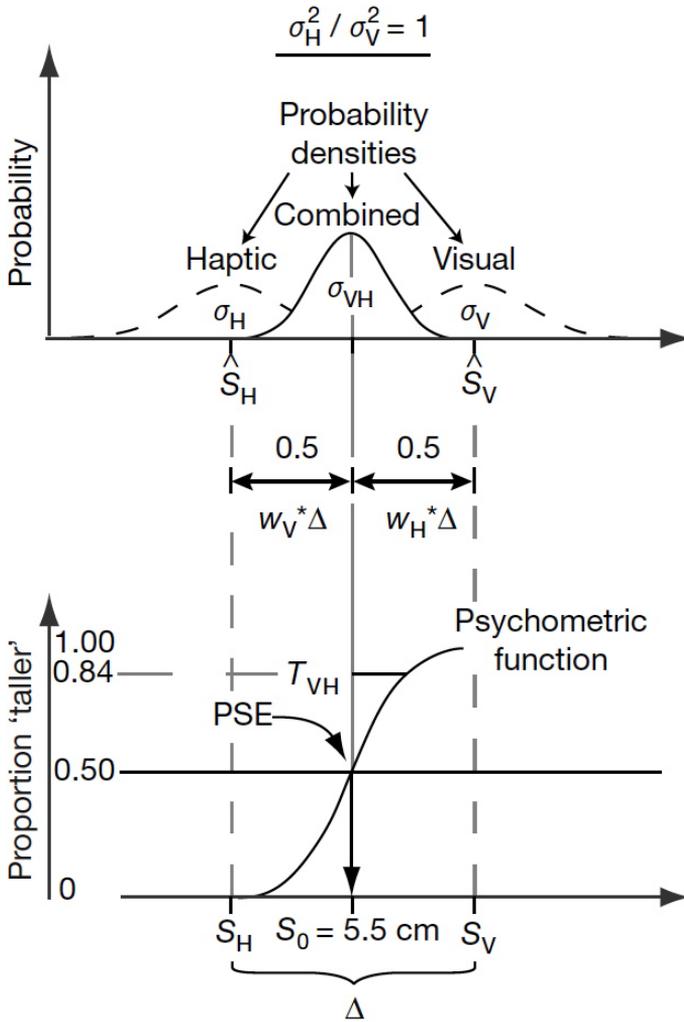


Figure 1. Example of maximum likelihood estimation integration for visuo-haptic cues. (Top) Dashed lines represent visually and haptically specified probability densities alone and the solid line represents the combined (vision + haptic) probability density estimate. Here visual and haptic variances (σ_V , σ_H) are assumed to have equal weights of 0.5. The combined probability density mean is then equal to the mean of the visual and haptic densities (S_V , S_H). (Bottom) Psychometric function based on judgments using combined probability density showing point of subjective equality (PSE) equal to the average of the visual and haptic heights of the standard stimulus (5.5 cm). (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, Ernst and Banks, copyright 2002.)

Thus, the MLE rule states that the optimal means of estimation, in the sense of producing the lowest-variance estimate, is to add the sensory estimates weighted by their normalized reciprocal variances. If the MLE rule is used to combine visual and haptic estimates, \hat{S}_V and \hat{S}_H , the variance of the final (visual-haptic) estimate, \hat{S} , is:

$$\sigma_{VH}^2 = \frac{\sigma_V^2 \sigma_H^2}{\sigma_V^2 + \sigma_H^2} \quad (3)$$

This rule implies that the final visuo-haptic estimate has a lower variance than either the visual or the haptic estimator alone. A hypothetical example of such integration is graphically shown in Figure 1.

An optimal integration according to the MLE rule implies that the variance of the final estimate should be lower than either the visually- or the haptically-based estimation alone. That is, the redundancy of information is used to improve the precision of the perceptual estimate. The MLE rule further predicts that the sensory signal with the lower variance will be dominant (i.e., has a higher perceptual weight). For example, it has been shown that under normal circumstances, vision dominates haptics for size judgment perception (Rock and Victor 1964; Kritikos and Brasch 2008; Hecht and Reiner 2009). However, when the visual input was degraded by adding visual noise (Ernst and Banks 2002), the haptic information gained dominance, indicating that the human visuo-haptic integration follows a MLE (Figure 2).

The results above demonstrate that a more reliable estimate of an object's size is achieved when it can be seen and felt, that is, when sensory redundancy is increased (Ernst and Banks 2002). However, the results also imply that the sensory system already "knows" which signals belong together and how they relate. In other words, the system must have some knowledge about the mapping between the signals. To examine how this type of mapping is acquired, Ernst (2007) applied a Bayesian model of cue integration in a study in which human subjects had to learn the relationship between normally unrelated visual and haptic information, as follows. Participants were trained to build a mapping between the luminance of an object (visual) and its stiffness (haptic). For example, the stiffer the object, the brighter it was. Subjects were then asked to discriminate between two objects that each had a certain luminance and stiffness. It was predicted that integration degrades discrimi-

nation performance for stimuli which were incongruent with the newly learned mapping. In contrast, the more certain subjects were about the new mapping, the stronger their discrimination performance should have been. Ernst (2007) found a significant change in discrimination performance before and after training, when comparing trials with congruent and incongruent stimuli. After training, discrimination thresholds for the incongruent stimuli increased relative to thresholds for congruent stimuli. This suggests that humans can learn to integrate two formerly unrelated sensory signals and utilize the result of this integration process for perceiving object properties, and that the mapping between visual and haptic stimuli for object perception has to be learned.

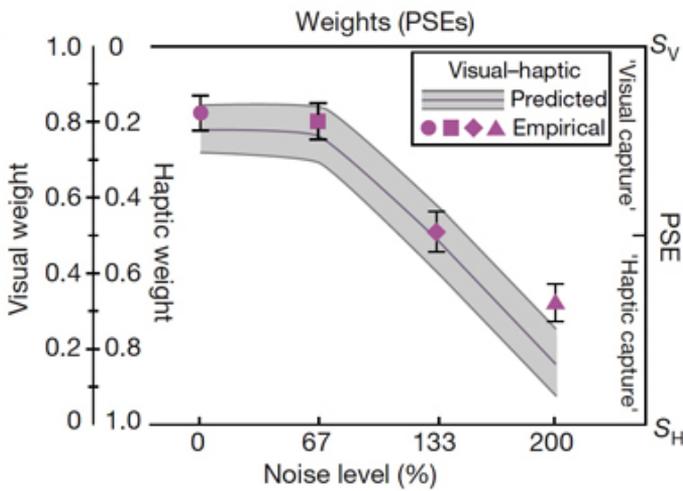


Figure 2. Prediction and experimental data from Ernst and Banks (2002). Graph represents the relative weighting of each modality at the point of subjective equality as visual noise increases. Shapes represent empirical data and shaded area represents predicted data following a MLE model. As visual noise increases, visual weighting decreases. Conversely, haptic weighting increases in both empirical and predicted data. (Reprinted by permission from Macmillan Publishers Ltd: *Nature*, Ernst and Banks, copyright 2002.)

This, in turn, raises the question of how the ability of visuo-haptic integration develops during childhood. A recent study by Gori et al. (2008) provided a first answer. In their experiment, children between 5 and 10 years of age had to discriminate between different heights or spatial orientations of two blocks (standard and comparison). The authors found that prior to 8 years of age, integra-

tion of visual and haptic spatial information was far from optimal (in the sense of applying a MLE rule), with either vision or haptics dominating. For size discrimination, haptic information dominated in determining both perceived size and discrimination thresholds, whereas for orientation discrimination, vision dominated. By the ages of 8 to 10 years, the integration became statistically optimal and comparable to adults. This result demonstrates that sensory integration mechanisms undergo a developmental process during childhood. Gori et al. (2008) speculated that perceptual systems of children may require a constant recalibration, for which cross-sensory comparisons are important. With respect to their experiment, this implies that children used vision to calibrate the haptic sense or vice versa. Yet, they were not able to combine visual and haptic information in such a way that reduced the variance of their perceptual estimates.

The findings above suggest that the Bayesian perception model provides a suitable framework to understand integration of visual and haptic information. The results of psychophysical research in this domain showed that the computational weights of each sensory modality are not fixed, but can be altered by context and change during development. We suggest that this Bayesian framework has great potential to provide new insights into failed mechanisms of intersensory integration after brain injury.

RELATING SENSORY INTEGRATION TO SENSORIMOTOR TRANSFORMATION

While the estimation of object size can be computationally modeled as a statistically optimal process of visuo-haptic integration, other researchers have viewed sensory integration in the context of *sensorimotor* transformations. Sensory integration is essential for perceiving ourselves and the environment. However, the visuo-haptic integration is a special case that often comes with *action*, which requires a sensorimotor transformation process; during the process of sensorimotor transformation, the efferent motor command is highly associated with the afferent visual and haptic information.

This linkage between sensory information and motor commands spawned another line of research that focused on how the perceived information impacts on voluntary motor control. To better understand this linkage, researchers examined visually guided reaching while presenting a sensory distracter (visual or haptic) si-

multaneously. Gentilucci et al. (1998) had participants perform two main tasks during several experiments; one represented an action task (reach and grasp) whereas the other represented a perceptual task (matching object sizes). For the action task, participants were instructed to haptically explore an unseen object with one hand, and simply reach and grasp a visually presented object with the other hand. For the matching task, participants were presented with a visualized object, but they also explored an unseen object haptically with one hand. They were then instructed to make an aperture size with their other hand which represented the size of either the visually or haptically explored object (depending on the experiment). The haptic and visual objects were presented (a) simultaneously or (b) with the haptic information being followed by the visual one. Object sizes (haptic versus visual) were either the same or different. Movement/aperture sizes of the grasping hand were recorded using motion capture cameras. The results obtained revealed that the size of the haptically explored object manipulated with one hand altered the kinematics of the grasping (action) task performed by the other hand. The altered kinematics were only observed when the haptically explored object was smaller than the visualized object that was to be grasped. However, aperture size (perceptual matching task) was not influenced by the haptic information obtained by the opposite hand. This outcome suggests a distinction in sensory integration between an action grasping task and a perceptual matching task. The findings suggest that haptic and visual information were integrated during sensorimotor transformation; however, the perceptual task relied mostly on visual information and haptic information had no measurable influence on aperture size. This supports the idea that the product of sensory integration is specific to the task goal.

In addition, Gentilucci et al. (1998) found that altered maximum grip aperture (MGA) during the action task was not bidirectional. Specifically, haptic exploration with the dominant right hand resulted in adaptations in the MGA of the non-dominant left hand, but exploration with the left hand did not result in adaptations of the right hand MGA. A later study by Patchay et al. (2003) partially contradicted these findings. This study investigated how the kinematics of the dominant hand reaching toward a visual target were influenced by haptic input from an unseen distractor that was actively grasped by the other, non-reaching hand. The main results were that the MGA of the reaching hand was smaller and the time to maximum grip aperture was earlier when the distractor was smaller

than the target. However, in contrast to the study by Gentilucci et al. (1998) the interference effect from the distractor was similar for both hands as they reached; that is, the processing of haptic information (and its interference) is bidirectional.

A subsequent study by Safstrom and Edin (2004) analyzed the weighting of visual and haptic information during the process of sensorimotor adaptation using a single hand grasping task. Healthy, right-handed human participants were asked to grasp an invisible object behind a mirror while a different visible object could be seen through the projection of a mirror as the graspable target. After a series of 9 grasps, the researchers inconspicuously changed the previously congruent, unseen, haptically explored object to a size either 15 mm larger or 15 mm smaller than the visualized object. The effect of incongruence in the size of the two objects became visible in the adaptation of the size of the hand gesture (MGA). Their findings indicated that haptic information had more impact on the increased size condition (haptic > visual) than in the condition where the haptic object was smaller than the visualized object. The adaptation process also occurred much more quickly in the increased size condition. In other words, after allowing the participant to adapt to the new, haptically explored height over a series of grasps, the participant's MGA more closely resembled the expected MGA for an object's given size in the condition in which the haptically explored object size was increased relative to the visualized object. Safstrom and Edin (2004) computed a ratio of visual-haptic weighting for the increased-size condition as 19:81 (visual to haptic) and 40:60 for the decreased-size condition. These results indicate that haptic information was weighted considerably more when the haptically explored object was larger than the visually presented object. In simple terms, participants adapted more quickly and were more accurate with their MGA during a task involving a larger-than-expected size whereas they were less accurate with their MGA in a task involving a smaller-than-expected size. It thus appears that grasping smaller objects was a more demanding task. The results of the study concur with the results of previous studies concluding that haptically-explored objects that are smaller than expected show greater error in sensorimotor transformation during action tasks. (Castiello 1996; Gentilucci et al. 1998; Patchay et al. 2003).

Safstrom and Edin's (2004) findings can be explained by assuming that their participants adopted a grasping strategy when they could not see their hand while grasping. That is, people tend to

adopt an implicit cost function, or a *play-it-safe* approach, by adopting a larger MGA to prevent failing the task. This could be considered as the effect of task characteristics which would possibly alter the weighting of visuo-haptic integration. Furthermore, the MGA adaptation process might be regarded as a learning process, namely learning a new relationship between visual and haptic information. Moreover, other studies have shown that haptic size perception is not affected by applied force or area of contact (Safstrom and Edin 2004; Berryman et al. 2006). These findings might be explained by the adoption of a safe grasping strategy as well.

So far we are aware of the MGA varying systematically with either the unseen object size explored haptically or the distracter size. Yet, the hand aperture size would not scale proportionally with the visuo-haptic difference (Castiello 1996; Patchay et al. 2003; Safstrom and Edin 2004; Weigelt and Bock 2007). Thus, these studies support the notion that both action and perception share the same, or at least very similar visuo-haptic integration processes (Figure 3a).

In contrast, another line of research hints that action-execution and perceptual judgments may rely on different visual-haptic integration processes (Westwood and Goodale 2003; Kritikos and Brasch 2008). Westwood and Goodale (2003) demonstrated two different psychophysical aperture size scaling functions for size estimation and grasp. The perceptual judgment using haptics alone was affected by nearby objects, whereas the object-directed action was not affected. Kritikos and Brasch (2008) also reported that the processes of perception and action utilize visual and haptic information differently. For perceptual judgments, the response time was significantly affected by the attended sensory modality, whereas the participants' MGA was not impacted systematically in the action task. Accordingly, the researchers argued that the perceptual judgment and the action-execution share the same sources of information stemming from the visual-haptic integration. However, although they may share the same source the information processes for perception and action could be different, leading to different end products of action and perception (Figure 3b).

In summary, the results of recent research support two plausible but different models of sensory integration, one for object perception and one for object-based action (Ernst and Banks 2002; Gori et al. 2008; Helbig and Ernst, 2008) (Fig. 3a and 3b). In addition, recent evidence suggests that a dissociation of action and perception may exist for the processing of haptic information that is similar to

vision (Milner and Goodale 1995). Supporting this view of separate haptic pathways are findings that humans like to use relative metrics for perceptual judgments and absolute metrics for object-based actions in haptic sensory processing (Westwood and Goodale 2003).

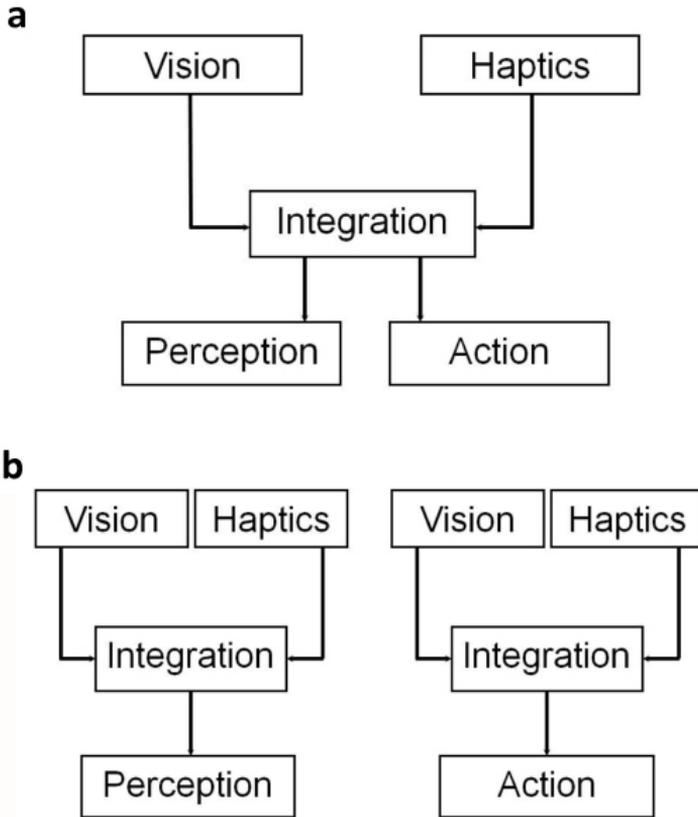


Figure 3. Two potential models of how visual-haptic information for perception and action can be integrated. (a) Single-Integration Model: There is one integration process of visual and haptic information that is being used both for perception and action. (b) Dual-Integration Model: Visual and haptic information is integrated in two distinct processes, one for perception, and a second one for action. In theory, the weights for vision and haptics can be different in a dual-integration process, meaning that vision might be dominant for perception, but not for action.

NEURAL BASIS OF VISUO-HAPTIC INTEGRATION

It has long been claimed that the visual pathways of the brain consist of two functionally and anatomically separate cortical processing streams: the dorsal visual stream, characterized as the *vision for action* stream or the *where* pathway, and the ventral visual stream, known as the *vision for perception* stream or the *what* pathway (Goodale et al. 1994). It is further known that the dorsal visual stream is involved in visuomotor tasks, and areas in the ventral stream have been shown to be multimodal and involved in visuo-haptic integration (Amedi et al. 2001; 2002). However, the identification of regions that represent areas of visuo-haptic integration is difficult, because it is not easy to distinguish between areas that respond to multisensory *stimulation* and areas of multisensory *integration*. In this context, the distinction between haptic and tactile perception has become blurred. This section will attempt to clearly define tactile versus haptic sense, multisensory areas versus multisensory integration, as well as identify relative areas of integration.

TACTILE VERSUS HAPTIC SENSING

In contrast to unimodal senses such as vision, haptic perception is based on afferent signals arising from tactile, proprioceptive, and pressure receptors. Although all of these signals enter into haptic perception, they each show similarities as well as differences in their processing pathways. Thus, it is important to differentiate between visuo-haptic integration and visuo-tactile integration. Areas such as the insula/clastrum, inferior parietal or superior temporal cortex, and the anterior and posterior intraparietal sulcus (IPS) have shown greater activation in crossmodal (visual-tactile) versus intramodal (visual-visual) matching tasks (Hadjikhani and Roland 1998; Banati et al. 2000; Grefkes et al. 2002; Saito et al. 2003). The results of these delayed match-to-sample tasks suggested that the insula and IPS play a crucial role in binding visual and tactile information (Amedi et al. 2005). Interestingly, *visuo-tactile* integration appears to occur mostly along the dorsal stream (i.e. the occipito-parietal region) with some activity along the ventral stream (Hadjikhani and Roland 1998; Grefkes et al. 2002; Saito et al. 2003; Amedi et al. 2005). In contrast, *visuo-haptic* integration is associated primarily with ventral stream activation (*vision for perception*) in the occipito-temporal region (Amedi et al. 2001; 2002). Thus, the processes

of cross-modal integration of *visuo-haptic* and *visuo-tactile* information are neuroanatomically distinct.

MULTISENSORY AREA VERSUS MULTISENSORY INTEGRATION

When presented with an object during a multisensory task, visual perception usually plays a larger role than any other sense in determining geometric attributes of the object (Amedi et al. 2002). Yet, Ernst and Banks (2002) demonstrated that this visual dominance is not fixed. They state that we *integrate* visual and haptic information in a statistically optimal fashion and the dominant sense is the one that is less variable in a given situation. Studies using hemodynamic methods such as blood oxygen level dependent (BOLD) fMRI have been successful in identifying specific multisensory *areas* in humans (Banati et al. 2000), although imaging findings, in general, may not correlate well with underlying neuronal activity (Yosher et al. 2007). Before we identify regions of integration it is important to differentiate between a multisensory *area* and multisensory *integration*. From a brain imaging perspective, a crossmodal or multisensory *area* may simply be defined as an area where stimuli of more than one modality give rise to activation of a given voxel (Calvert and Thesen 2004). The main argument put forward is that, because an area shows activation during a multimodal stimulation, one cannot assume that integration occurs in that location. However, although the activation of a given anatomical region with two different modalities identifies it as crossmodal, it only indicates it as a *possible* locus for integration.

On a neuronal level, the activated area may simply encompass a mixed population of unimodal sensory neurons. In addition, a single neuron may respond to multisensory stimulation, but may not be involved in integration, that is, a neuron may be activated by two separately presented modalities. If a neuron responds to a unisensory stimulus with a certain activation level, but the activation level is not measurably altered when the second modality is presented, it cannot be concluded that integration of these modalities has occurred but rather that this neuron is responsive to more than one modality and considered multimodal (Calvert and Thesen 2004).

Multisensory *integration* is thus defined as the enhanced response of neurons, exceeding the firing rate of that expected by summing the rates of two unimodal inputs (Calvert and Thesen 2004). The integrated signals from separate modalities are then combined

to form a new, single output signal (Stein et al. 1993). Calvert and Thesen (2004) state that electrophysiological studies in the superior colliculus show multimodal neurons. However, their response to one modality M_1 is not altered in the presence of a second modality M_2 , refuting the presence of integration. Calvert and Thesen (2004) offered a computational method for identifying integration responses by determining an interaction effect with the use of the formula $[M_1M_2 - \text{rest}] - [(M_1 - \text{rest}) + (M_2 - \text{rest})]$. Identifying integration involves a condition including both modalities (M_1M_2) minus a reference or *rest* condition and subtracting the summation of the two isolated modalities (M_1 and M_2) showing that integration cannot be determined by simply summing unimodal inputs but rather the activity level must exceed that of the summed unimodal inputs. The picture emerging is that integration occurs when the afferent projections of unimodal neurons converge to those *multisensory* areas, that is, areas containing neurons that are stimulated by multiple senses. This convergence results in a new, nonlinear, integrated signal that exceeds the summed unimodal signals, allowing for a heightened perceptual judgment.

BRAIN REGIONS INVOLVED IN VISUO-HAPTIC INTEGRATION

It is well known that the processing of unimodal visual and haptic sensory stimuli occurs in anatomically distinct somatotopic cortical regions. What is less known, is whether the same or different brain regions are specifically devoted to multisensory integration. Because of the complexity of the integration process, determining these areas has proven to be difficult.

Neuroanatomical data indicate that the crossmodal integration of visual and tactile senses may occur in the insula, inferior parietal or superior temporal cortex, anterior cingulate cortex, hippocampus, and the amygdala (Banati et al. 2000). Other studies have identified areas within the basal ganglia (BG) containing multisensory cells, such as the striatum, the caudate nucleus, the putamen and the substantia nigra (Graziano and Gross 1993; Nagy et al. 2006). There has been an ongoing discussion that the BG plays an important role for sensorimotor integration (Nagy et al. 2006). However, the evidence for visuo-haptic integration processes occurring in the BG is less compelling. There is abundant evidence showing that kinesthesia and haptic acuity are affected by diseases affecting the BG, such as Parkinson's disease or dystonia (Putzki et al. 2006; Konczak et al. 2009). However, because Parkinson's patients seem

to become visually dependent during motor tasks and their kinesthetic deficits occur primarily when vision is occluded (Konczak et al. 2008), it would appear that visual-haptic integration may occur in anatomical locations outside of the BG area.

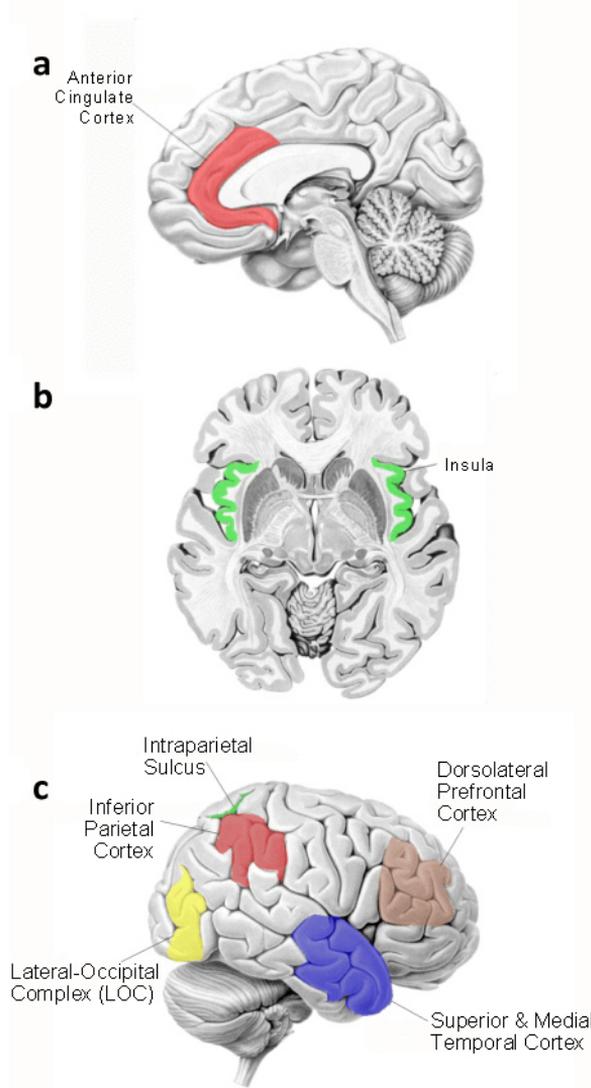


Figure 4. Neuroanatomical areas believed to be involved in visual-haptic integration processing. (a) Anterior cingulate cortex. (b) The insula. (c) Inferior parietal cortex, intraparietal sulcus, dorsolateral prefrontal cortex, superior and medial temporal cortex, and the lateral-occipital complex (LOC).

A study by Banati et al. (2000) seems to corroborate the claim that sensory integration processes are neocortical in nature. These researchers used a modified arc circle test in Positron Emission Tomography (PET) to investigate the difference between intramodal matching (visual-visual) and crossmodal matching (visual-haptic). The results of this study indicated that the regional cerebral blood flow (rCBF) increased in the anterior cingulate, inferior parietal lobule, dorsolateral prefrontal cortex, insula and middle and superior temporal gyri in the crossmodal matching task (Figure 4a,b), whereas only relative rCBF increased in visual association cortexes in intramodal matching task. This study also showed a lack of activation in the amygdala and hippocampus, again showing differences in activated areas between visuo-haptic exploration and visuo-tactile tasks from previous studies (see Banati et al. 2000).

Amedi et al. (2001) found robust somatosensory activation in a ventral region anterior to retinotopic visual areas. Human subjects performed under four different conditions: haptic exploration of an object, haptic exploration of a texture, visual exploration of a picture of an object, or visual exploration of a picture of a texture. The occipito-temporal region, specifically the lateral-occipital complex (LOC, Figure 4c), showed clear somatosensory preference for objects compared to textures irrespective of the modality, haptic or visual, used to identify the object. Having a clear preference for manipulated objects shows that the LOC region is specific to haptic information rather than being specific to tactile information alone. The LOC region was not activated during movements that imitated movements done when subjects explored the objects haptically nor was this region activated during naming or visual imagery conditions. These results showed that exploration of objects was the true reason for activation of this area and reinforces the theory that the LOC is a multimodal, and more specifically, a bimodal area for visual and haptic senses (Amedi et al. 2001).

In addition, Pietrini et al. (2004) reported activation in the same general anatomical location as reported by Amedi et al. (2001) in both congenitally blind and sighted subjects. Although visual memory may play a small role in activation of the LOC, the study by Pietrini et al. (2004) involving congenitally blind participants showed that this area is mostly related to haptically explored object recognition. Amedi et al. (2001) also found activation in the IPS and the parietal lobule in response to somatosensory object manipulation. What is yet to be determined is the relationship between these areas in the parietal cortex and areas such as the LOC with respect

to somatosensory activation. One interpretation offered by Amedi et al. (2001) for the anatomical location of the multimodal area is that object recognition relies primarily on vision and therefore activation occurs in the visual cortex rather than somatosensory areas. A study by Feinberg et al. (1986) gave more support to the notion of a possible integration area in the LOC. The authors reported a case study of a patient with a unilateral left hemisphere lesion around the inferior occipito-temporal cortex that resulted in severe visual and tactile agnosia.

Amedi et al. (2002) proposed that the LOC can be subdivided into smaller sub-regions, one of which was termed LOCTv (lateral occipital tactile-visual region), which lies within the ventral visual stream and responds to both tactile and visual objects. Amedi et al. (2002) also proposed that *haptic* and visual information processed would be localized to a specific area within the LOC as well. They then stated that not only is this area activated during object recognition (versus texture), it also shows a preference *only* for object recognition, as activation in this area was not observed during face or house recognition. One may theorize that this localized region is specific to determining the geometrical shape of an object. James et al. (2002) was able to rule out memory being a factor in the activation in this area when subjects palpated unrecognizable objects. This task produced similar activation patterns in the LOC as when seeing familiar objects. This finding would lead one to think that the area Amedi et al. (2002) describes within the LOC is a specific bimodal area for integration of visual and haptic information because it has shown a preference for objects that the subject can manipulate with their hand. What is the reasoning behind areas showing a preference for haptic manipulation, giving rise to bimodal integration of vision and haptic senses? Amedi et al. (2002) argues that vision and haptic senses are the only two senses that can extract specific and precise geometric information about an object's shape; which is reasonable since auditory sounds, among other senses, often do not give rise to specific geometric measurements.

Based on the results of the research reviewed above (Feinberg et al. 1986; Graziano and Gross 1993; Banati et al. 2000; Amedi et al. 2001; 2002; James et al. 2002; Peitriini et al. 2004; Nagy et al. 2006; Putzki et al. 2006; Konczak et al. 2008; 2009), one may agree with Amedi et al. (2005) that multisensory regions are determined by multiple factors such as task requirements, object presentation, as well as the dominant sense. Just as well, neuroanatomical loca-

tion may also be a determining factor in producing a multisensory region. For example, the inferior parietal lobule has been shown in several studies to be a possible location for not only integration of haptic and visual senses, but auditory as well. This may partly be due to the fact that this region is located at the junction of auditory, visual, and somatosensory cortexes. Task requirements and object presentation would certainly influence whether an object is haptically manipulated or the object is explored using tactile sense. Although these two senses have shown similarities in cortical activation areas during crossmodal tasks, such as the inferior parietal lobule, superior temporal cortex, anterior cingulate cortex, and insula (Banati et al. 2000), there have also been differences in activated areas. For example, Banati et al. (2000) also showed activation in the hippocampus and amygdala during visuo-tactile tasks but little to no activation in these two areas during visuo-haptic tasks. Therefore, although some areas have been identified as possible areas of integration, it appears to be clear that several factors play a large role in determining specific areas of integration.

SUMMARY AND OUTLOOK

When looking at an object while exploring and manipulating it with the hands, visual and haptic senses provide information about the properties of the object. Computational theory suggests that the integration of these two streams of sensory information within the brain strives to minimize the variance of the perceptual estimate by performing operations commensurable to a maximum-likelihood estimation. In addition, research has shown that the utilization and weighting of visual and haptic information may vary between tasks (Gentilucci et al. 1998), although it remains unclear how task requirements exactly affect the relative contribution of visual and haptic information. It may be that different sensory processing and sensory integration mechanisms are used for making a perceptual judgment about an object versus executing an object-based action (Ernst and Banks 2002; Gori et al. 2008; Helbig and Ernst 2008). At this point, we do not know whether the existing computational framework derived from perceptual discrimination studies can be generalized to sensorimotor scenarios, where perceptual information is used for the voluntary control of action. Therefore, future computational models need to account for the variance due to the differences between “purely” perceptual and perceptual-motor tasks.

To address this issue it seems necessary to develop research paradigms for investigating visuo-haptic integration within both a perception and an action-based context. The conflicting findings from perceptual judgment and action-based studies may be attributable to differences in experimental paradigms. Designing paradigms that allow for relative consistency across both contexts has proven to be difficult. Even attempting to solve the process of integration certainly reflects the complexity of integration itself.

Therefore, it is vital to clearly define the type of integration process under consideration and to distinguish it from multisensory processing. In this context, Calvert and Thesen (2004) proposed that an area could be considered multimodal when a given neuron responds to multiple modalities. However, as discussed above, integration does not necessarily occur unless the activation level of that given neuron, during bimodal stimulation, exceeds the summed activation level of that neuron during stimulation of each modality separately.

Although there are similarities between haptic and tactile sense processing, there appears to be a distinct neuroanatomical difference between visuo-tactile integration and visuo-haptic integration. Visuo-tactile integration appears to occur mostly along the dorsal visual-stream (Hadjikhani and Roland 1998; Banati et al. 2000; Grefkes et al. 2002; Saito et al. 2003; Amedi et al. 2005), whereas visuo-haptic integration appears to occur mostly along the ventral visual-stream (Amedi et al. 2001; 2002). Amedi et al. (2002) provided further evidence for the difference between haptic and tactile processing by showing that although the LOC displayed activation for both tactile and haptic processing, a specific location within the LOC responded only to object recognition as opposed to texture. Interestingly, this area lies along the ventral visual pathway rather than the dorsal, which projects toward the motor and somatosensory cortex. Amedi et al. (2002; 2005) claimed activation response to object manipulation occurring along the ventral visual pathway occurs due to vision being the dominant sense as well as because haptic and visual senses are the only two senses involved in determining geometric attributes of an object. Despite the somewhat sporadic areas of activation during visuo-haptic tasks, the LOC appears to be consistently activated during these crossmodal tasks.

We conclude that much progress has been made to provide a computational framework that can formalize and explain the results of behavioral and psychophysical studies on visuo-haptic integra-

tion. Yet, there still exists a gap between the computationally driven studies and the results derived from brain imaging studies. One reason why the closing of this gap has proven to be difficult is that visuo-haptic integration processes seemed to be highly influenced by the task and context. Although there is a clear difference between a multisensory area and multisensory integration, the bimodal LOC appears to be a principal area for integration to occur. Future research may want to focus on the LOC in order to more delicately define the process at this location during multisensory tasks. A better understanding of the underlying neuronal activity at this location may give rise to a better understanding of the visuo-haptic integration process.

REFERENCES

- Amedi A, Malach R, Hendler T, Peled S, Zohary E (2001) Visuo-haptic object-related activation in the ventral visual pathway. *Nat Neurosci* 4:324-330
- Amedi A, Jacobson G, Hendler T, Malach R, Zohary E (2002) Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cereb Cortex* 12:1202-1212
- Amedi A, von Kriegstein K, van Atteveldt NM, Beauchamp MS, Naumer MJ (2005) Functional imaging of human crossmodal identification and object recognition. *Exp Brain Res* 166:559-571
- Banati RB, Goerres GW, Tjoa C, Aggleton JP, Grasby P (2000) The functional anatomy of visual-tactile integration in man: a study using positron emission tomography. *Neuropsychologia* 38:115-124
- Berryman LJ, Yau JM, Hsiao SS (2006) Representation of object size in the somatosensory system. *J Neurophysiol* 96:27-39
- Calvert GA, Thesen T (2004) Multisensory integration: methodological approaches and emerging principles in the human brain. *J Physiol Paris* 98:191-205
- Castiello U (1996) Grasping a fruit: selection for action. *J Exp Psychol Hum Percept Perform* 22:582-603
- Ernst MO (2007) Learning to integrate arbitrary signals from vision and touch. *J Vis* 7:1-14

- Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429-433
- Feinberg TE, Rothi LJ, Heilman KM (1986) Multimodal agnosia after unilateral left hemisphere lesion. *Neurology* 36:864-867
- Gentilucci M, Daprati E, Gangitano M (1998) Haptic information differentially interferes with visual analysis in reaching-grasping control and in perceptual processes. *Neuroreport* 9:887-891
- Gibson JJ (1966) *The senses considered as perceptual systems.* Houghton Mifflin, Boston
- Goodale MA, Meenan JP, Bulthoff HH, Nicolle DA, Murphy KJ, Racicot CI (1994) Separate neural pathways for the visual analysis of object shape in perception and prehension. *Curr Biol* 4:604-610
- Gori M, Del Viva M, Sandini G, Burr DC (2008) Young children do not integrate visual and haptic form information. *Curr Biol* 18:694-698
- Graziano MS, Gross CG (1993) A bimodal map of space: somatosensory receptive fields in the macaque putamen with corresponding visual receptive fields. *Exp Brain Res* 97:96-109
- Grefkes C, Weiss PH, Zilles K, Fink GR (2002) Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys. *Neuron* 35:173-184
- Hadjikhani N, Roland PE (1998) Cross-modal transfer of information between the tactile and the visual representations in the human brain: a positron emission tomographic study. *J Neurosci* 18:1072-1084
- Hecht D, Reiner M (2009) Sensory dominance in combinations of audio, visual and haptic stimuli. *Exp Brain Res* 193:307-314
- Helbig HB, Ernst MO (2008) Visual-haptic cue weighting is independent of modality-specific attention. *J Vis* 8:1-16
- Henriques DY, Soechting JF (2005) Approaches to the study of haptic sensing. *J Neurophysiol* 93:3036-3043
- James TW, Humphrey GK, Gati JS, Servos P, Menon RS, Goodale MA (2002) Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia* 40:1706-1714

- Kersten D, Mamassian P, Yuille A (2004) Object perception as Bayesian inference. *Annu Rev Psychol* 55:271-304
- Konczak J, Li KY, Tuite PJ, Poizner H (2008) Haptic perception of object curvature in Parkinson's disease. *PLoS One* 3:e2625
- Konczak J, Corcos DM, Horak F, Poizner H, Shapiro M, Tuite P, Volkman J, Maschke M (2009) Proprioception and motor control in Parkinson's disease. *J Mot Behav* 41:543-552
- Kritikos A, Brasch C (2008) Visual and tactile integration in action comprehension and execution. *Brain Res* 1242:73-86
- Milner AD, Goodale MA (1995) *The visual brain in action*, vol. 27. Oxford University Press, Oxford ; New York
- Nagy A, Eordeghe G, Paroczy Z, Markus Z, Benedek G (2006) Multisensory integration in the basal ganglia. *Eur J Neurosci* 24:917-924
- Patchay S, Castiello U, Haggard P (2003) A cross-modal interference effect in grasping objects. *Psychon Bull Rev* 10:924-931
- Pietrini P, Furey ML, Ricciardi E, Gobbini MI, Wu WH, Cohen L, Guazzelli M, Haxby JV (2004) Beyond sensory images: object-based representation in the human ventral pathway. *Proc Natl Acad Sci USA* 101:5658-5663
- Putzki N, Stude P, Konczak J, Graf K, Diener HC, Maschke M (2006) Kinesthesia is impaired in focal dystonia. *Mov Disord* 21:754-760
- Rock I, Victor J (1964) Vision and touch: an experimentally created conflict between the two senses. *Science* 143:594-596
- Safstrom D, Edin BB (2004) Task requirements influence sensory integration during grasping in humans. *Learn Memory* 11:356-363
- Saito DN, Okada T, Morita Y, Yonekura Y, Sadato N (2003) Tactile-visual cross-modal shape matching: a functional MRI study. *Cogn Brain Res* 17:14-25
- Stein BE, Meredith MA, Wallace MT (1993) The visually responsive neuron and beyond: multisensory integration in cat and monkey. *Prog Brain Res* 95:79-90
- van Beers RJ, Baraduc P, Wolpert DM (2002) Role of uncertainty in sensorimotor control. *Philos Trans R Soc Lond B Biol Sci* 357:1137-1145

- von Helmholtz H (1867) Handbuch der physiologischen optik, vol. Bd. 9. Voss, Leipzig
- Weigelt C, Bock O (2007) Adaptation of grasping responses to distorted object size and orientation. *Exp Brain Res* 181:139-146
- Westwood DA, Goodale MA (2003) A haptic size-contrast illusion affects size perception but not grasping. *Exp Brain Res* 153:253-259
- Yoshor D, Ghose GM, Bosking WH, Sun P, Maunsell JH (2007) Spatial attention does not strongly modulate neuronal responses in early human visual cortex. *J Neurosci* 27:13205-13209