Statistical modelling of multisensory integration in the context of a prosthetic hand

Eduardo Martin Moraud

MSc in Artificial Intelligence
School of Informatics, University of Edinburgh
2009

Abstract

For the simple task of grasping an object, humans integrate different sensory inputs to estimate magnitudes such as the position of their hand or the force exerted. Visual, tactile or proprioceptive information is optimally combined and relied upon to guide their movements. Patients endowed with a prosthesis, however, can only rely on their own vision to control the artificial limb. Under such conditions, it is believed that providing them with artificial feedback could help improve the manoeuvrability of the device. Nevertheless, the capacity to use such feedback in everyday life situations is yet to be confirmed; So is the description of the optimal conditions under which it should be applied.

In this thesis we tackle the question of **optimality** when providing artificial vibrotactile stimulation, both in terms of exploitability and performance. The perspective adopted is twofold, and combines evaluations *within* and *across* modalities: We first focus on the encoding that is to be presented to the user, aiming to exploit his perceptual sensitivity and to refine the information transmitted. We then turn to how the new modality is actually combined with other natural sensing capacities. Both analyses are undertaken in the framework of a simple task, specifically conceived to isolate the underlying principles that are relevant to each question.

Our results highlight that subjects exploit the newly added modality and integrate it with their own natural senses. Optimality, however, requires of specific key points. Namely, the type of encoding is to be carefully adapted to the specificities of the sensitivity and ought to be quantified for each individual subject. Likewise, the degree of uncertainty that is better coped with needs to be quantified across modalities, for the performance reached when combining natural and artificial information is shown to be directly linked to their relative amount of noise.

Acknowledgements

This project could not have been possible without the help of many people. I wish to thank Ian Saunders for his patience and invaluable help all along these months. His ideas and advice were truly enlightening in many occasions. I wish him the best of luck for his Ph.D., which I am sure he will complete with great success. I also thank Dr. Sethu Vijayakumar for helping me set up this project, and for his support and guidance during its execution. Finally, many thanks to all my friends and colleges for so many good moments of intense work in the lab. They certainly contributed to making this experience an unforgettable one.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except when explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Eduardo Martin Moraud

Contents

1
Introduction

Sensorimotor tasks are inherently uncertain. Human sensors provide imperfect information about the world due to external noise and inaccuracies, and make it necessary to use estimates about the environment when interacting with the world. Multisensory integration refers to this capacity to combine various stimuli coming from different sensors in order to get more accurate information of the surrounding and the body.

Much research has been carried out on healthy subjects to understand how the brain combines different sensory cues and how these are relied upon to come up with an appropriate behavior. These are usually encoded under *different modalities* –visual, tactile– and represented within distinct *spatial frames* –such as head-centred or hand centred–.

In the case of amputees provided with prosthetic devices, ‘natural’ sensory information such as touch or proprioception needs to be replaced by artificial stimuli. This **sensory substitution** is essential to restore the sensation in patients that the artificial limb actually belongs to their own body. It has been attempted through different techniques such as vibrotactile systems, electrotactile arrays or teleproprioception [11] which exploit different modalities. However, the way these artificial modalities are exploited and integrated with natural ones is yet to be analyzed.

It has been proposed that, for healthy people, estimates actually be created so as to **minimize the error** of what is perceived, thus providing the subject with the most reliable information about the environment [5] . Mathematical models based on Maximum Likelihood theories and Bayesian Inference have been described as possible explanatory models [3, 22]. These may also be relied upon to test the hypothesis of optimal integration in the case of prosthetics, providing further insight into how new feedback channels are accepted by patients.

Motivation & Approach

The aim of this thesis is to study how artificial sensory information is perceived and integrated with normal stimuli, and how that combination may be improved to help patients better interact with the world.

This problem is addressed from a dual perspective:

- First from the point of view of **within-modality** discrimination. We look at the perceptual sensitivity of subjects when dealing with the artificial feedback channel and we try to come up with the most suitable encoding in terms of exploitability and adaptability.

- Then **across modalities** by evaluating multisensory integration. Statistical models similar to the ones used to for healthy humans are here employed in order to get acquainted with how artificial cues are iteratively combined with natural stimuli and employed to interact with the world.

Our procedure when tackling these questions is also twofold, and is built on a joint experimental and theoretical approach. Indeed, our two previous evaluations are undertaken under the framework of simple experimental tasks that aim to isolate the relevant underlying principles, and to help highlight meaningful patterns:

- Within-modality discrimination is attempted in the framework of a **perceptual task** in which participants are asked to distinguish among noisy versions of different stimulus. Their choices are recorded and used to analyse their performance for several artificial encodings.

- The question of across-modality integration is carried out in the context of a **tracking task** where subjects are presented with conflictive visual and tactile information, and are expected to accurately drive a cursor towards a target by combining both modalities.

The experimental results are then employed to perform some statistical analyses and to derive quantitative evaluations. These can then be relied upon to get further insight into how such matters may be translated into the field of prosthetics.

Organisation

This Dissertation also shares this dual structure. Each one of the aforementioned evaluations is presented in parallel, first from an experimental point of view, then from the perspective of the outcomes that follow. We thereby attempt to emphasize the individual interest of each question while keeping in mind their common goal

The remainder of this thesis is thus organised as follows. Section 2 describes the principles used so far by researchers to explain sensorimotor control in healthy humans, along with some descriptions of further issues concerning world representation in the brain and internal models. Complementary details about current technical advances in prosthetics and artificial feedback is also here detailed. The methods employed to tackle our two fundamental questions are then detailed in sections 3 and 4; First through the materials used in our experiments and models. Secondly through the analyses that are to be carried out. The results derived from them are then presented in section 5.

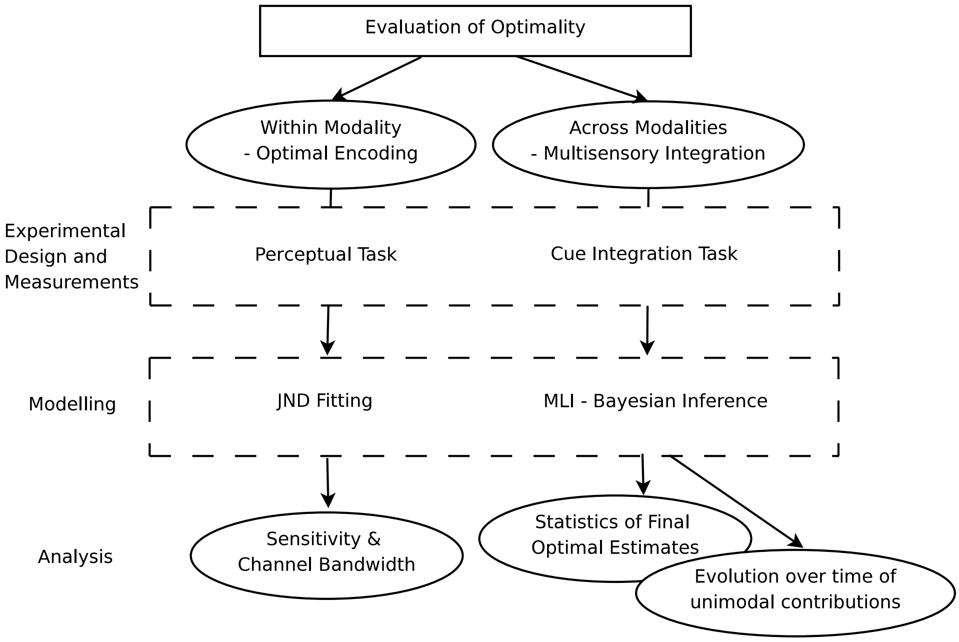


Figure 1.1: Flowchart of the structure of this Thesis.

2
Background

2.1 Sensory integration in healthy humans

Human perception is uncertain by nature; Visual information is misleading in the dark, auditory sensing can be fooled by echos, and scent becomes unreliable when confronted to turbulent winds; Most perceptual representations are indeed noisy, subject to tiredness, emotions or will. They are also usually biased, either by personal or social reasons underlying human nature. Hence humans can only expect to **estimate** the reality around them, and to rely upon it to survive in a world of uncertainties.

The *combination* of all the aforementioned senses is the keystone in the construction of a useful representation of the environment, and is indeed essential to exploit all the information offered by the world. Much research has attempted to unravel how humans employ different sensory modalities, how these are combined and further employed. Such matters have been analysed both from a behavioral point of view, from physiological approaches at a single-cell level or by means of mathematical modelling. The following sections give some insight into some of these lines of investigation.

2.1.1 Maximum Likelihood Estimates

A widely accepted model to explain the principles underlying sensory *estimation* is that of Maximal Likelihood [5]. This framework suggests that an estimate of a world state x may be generated by maximizing the likelihood that the stimuli $\mathbf{s}_x$ have been generated by the state. It hence provides the representation of the world that is more likely to underlie the perceptual cues:

$$\hat{x} = \arg\max_x p(\mathbf{s}_x|x) \tag{2.1}$$

When applied to multisensory *integration*, it follows from such model that all modalities should contribute to the global perception in a way that *minimizes the uncertainty* of the final estimate. This implies that each sensory information is to be relied upon according to its individual reliability, so that corrupted or noisy information is penalised, whereas reliable cues are given a bigger contribution on the final estimate.

The Maximum Likelihood Integration (MLI) estimate hence refers to that which is obtained as a weighted sum of all other considered estimations, where the weights are linearly derived from the

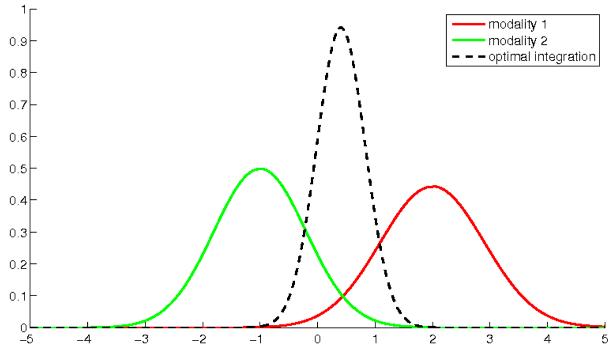


Figure 2.1: Optimal Integration of two Modalities affected with independent Gaussian noise.

variance of each modality. For the case of two independent Gaussian distributed noise affecting the modalities, as seen in Figure 2.1, the MLI estimator is then as follows:

$$\hat{x} = w_{s1}\hat{x}_{s1} + w_{s2}\hat{x}_{s2} \tag{2.2}$$

in which the weights w_i here represent the contribution of each stimuli to the final response, and are determined according to their reciprocal variance:

$$w_i = \frac{1/\sigma_i^2}{\sum_j 1/\sigma_j^2} \tag{2.3}$$

Specific experiments have shown that this model does indeed match biologcal observations for the case of visual and tactile cues [5], and visual and auditory information [3].This suggests that the nervous system is **optimal**, i.e. it accounts for the specific individual error associated with each modality before combining them, and adds them up accordingly, and asserts its interest and placing its ideas at the core of much of the research in the field.

Further research has been carried out in an attempt to evaluate the degree of optimality achieved when integrating cues from different modalities. Experiments performed on patients with cortical lesions helped isolate certain effects [15, 12].

2.1.2 Internal models & Bayesian Theory

An additional element, essential to human perception, is the *'a-priori'* knowledge that subjects have about their surrounding. Humans automatically rely on experience when attempting to perform an action or to interpret a scene, and they only use new cues so as to refine their internal belief.

For instance, when shown a piece of fruit, humans automatically make use of an internal representation about its expected shape, color or taste to infer what it is. This initial model can then be slightly adapted to comply with what is actually being presented to them. Should there be a conflicting information –such as a piece of fruit being referred to as an apple but being of color blue– the person will be compelled to evaluate the reliability of both his belief and the current information, so as to come up with the most plausible explanation of what he is really witnessing

These **internal models** encode the information that subjects have gathered through experience and are usually relied upon as a 'first guess'. When constructing estimates, this initial belief may then be 'tuned' to derive the most appropriate understanding of the world given what is currently perceived.

From a physiological perspective, internal models are important because they enable to compensate for sensorimotor delays –i.e. the required time to transform external stimuli into electrical impulses and to transport them from the sensors towards the central nervous system– and they thus make it possible to respond in a speedy manner to the requirements of the environment. They are constantly accessed and adapted when learning or planning future actions [22].

Statistically, Bayesian theory gives answer to the presented framework. It builds its calculations around two terms, a (1) prior probability $p(x)$ which encodes the belief that the subject has about the scene, and (2) the current likelihood of what is been perceived. The final **posterior probability** about the state of the world –x– given a certain sensory input –$\mathbf{s}_x$– may hence be written

$$p(x|\mathbf{s}_x) \propto p(\mathbf{s}_x|x) \cdot p(x) \tag{2.4}$$

Bayes' theory hence completes Maximum Likelihood models by considering further information than what is strictly perceived. The resulting estimate thus corresponds to the Maximum a Posteriori (MAP) obtained from:

$$\hat{x} = \arg\max_x p(x|\mathbf{s}_x) \tag{2.5}$$

This may also be naturally extended to calculate posterior probabilities derived from different sensory modalities. Experiments performed for visual and auditory information [3] or for visual and tactile feedback [5] emphasized that in healthy humans sensory integration is indeed undertaken in a statistically-optimal way as stated by the Baseyian framework, i.e. both prior and likelihood being weighted so as to minimise the variance of the posterior probability.

In our modelling, both Maximum Likelihood Estimates and Bayesian Inference will be presented as a way to model and analyze the way in which subjects integrate natural and artificial feedback modalities. The interest of one or another technique will be emphasized, along with assumptions that may be required.

2.1.3 Linking perception to actions

We have presented so far sensory integration in the framework of *perceptual tasks*. Some of the results obtained have been roughly documented, namely optimal integration for generating estimates with maximal reliability. The main **purpose** of such process however has not yet been introduced. What use do humans give to these estimates and why are they essential to our everyday life?

Their relevance is immediately underlined when thinking of human *interaction* with the world. Sensory integration helps guide our actions in a complex and uncertain environment, in which optimal inference models enable humans to come up with reliable perceptual judgements that may be relied upon when acting. Bayesian Integration is indeed at the core of most explanations that attempt to couple human perception with control and behavior. It represents a step forward to better understand the sensorimotor loop, while linking the effects of uncertainty with human decisions.

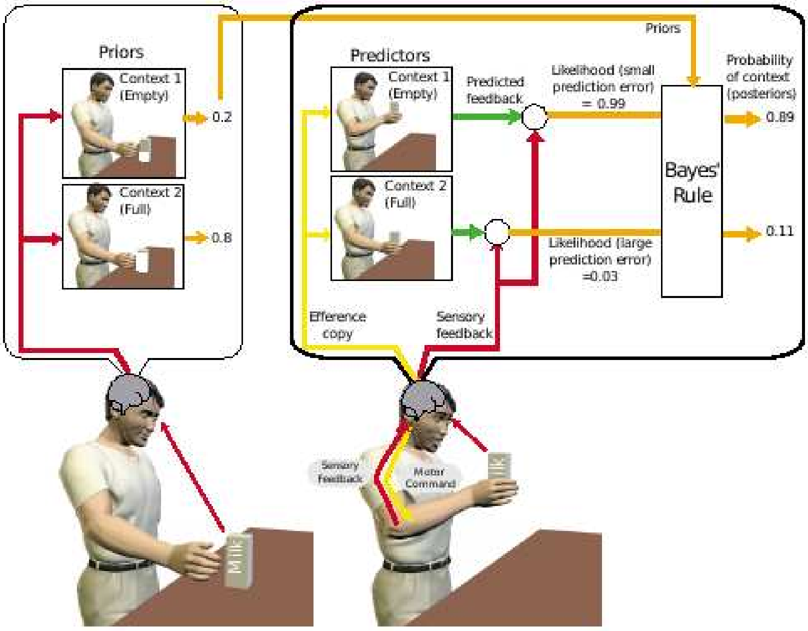


Figure 2.2: Example of how Bayes' inference theory may be employed for motor estimation and prediction in the context of a grasping movement. Two a-priori beliefs are here considered, the milk carton is full or empty. The effect of feedback is here illustrated as a way to refine the predicted outcome of the movement (extracted from Wopert and Gharamani [22]).

One of out later analysis attempts to use people actions in order to infer the use that subjects make of the different cues. We thus skin through some of the principles that have been suggested to explain movement neuroscience, control learning and planning. I will also briefly review hereafter how these are related to aforementioned integration theories.

Different models have been proposed, which centre their description around an overall goal which aims to maximize the performance of each action. Corrections and adaptations to external stimuli need to be produced 'online', where perceptual cues represent the feedback that is employed to make small adjustments to movement trajectories.

Under certain hypothesis –such as linearity and Gaussian noise–, the prediction step that combines feedback and forward models can be seen as a Kalman filter, and has been supported by empirical analysis of hand position [21] or posture [13].

2.1.4 Spatial link between modalities

Interactions between different modalities are useful in the construction of a representation of space. However, they are usually described within different body-part centred coordinate frames –hand, head–, a variety accounts for much of our ability to locate our body in the surrounding but which has not yet been considered in Bayesian Models.

Interestingly, experiments have shown that the mutual information provided by different stimuli is more pronounced when there is a **correspondence in space** between the sensors. In the case of visual and tactile feedback, for instance, multisensory interactions were demonstrated to be emphasized when visual cues are located near the stimulated hand [15]. This critical dependence on the the proximity of the visual stimuli and the body part where touch is exerted is logically linked to the posture, a keypoint which may need to be kept in mind for our later experiments.

2.1.5 Towards sensory substitution...

It has been shown that modalities are not independent. They do not constitute isolated modules, but rather they are mixed on a vast level by the brain. More importantly, it has been observed that the loss of one modality may lead to a **rewiring of the sensory paths**, so that neural areas of one modality from which the subject has been deprived may be employed by another modality [19]. This is usually referred to as *cross-modal integration*.

In the case of blind people, for instance, the activation of primary and secondary visual cortex were induced by Braille reading, hence suggesting that cortical areas normally reserved for vision may be activated by other sensory modalities [17]. Such capacities are possible through the plasticity of the brain, but are usually observed only in the case of sensory loss in early stages of life.

Yet this raises many questions as to whether it may be possible for the brain to learn a new modality which is *totally artificial*, i.e. rewire the feedback paths to integrate a new modality.

Further examples illustrate the amazing capacity of the brain to remap sensations to newly accepted feedback has even been shown for amputees. The well-known **'rubber hand' illusion**, where tactile feedback is provided in the stump while the subjects looking at an artificial hand –which is synchronously stimulated– creates the illusion of one's proper hand [4]. The brain can hence be tricked. In what *form* should the stimulation be applied?

showed that visual information about an apparent hand does have subjective crossmodal influences on tactile judgements –'rubber hand' illusion [4]–.

2.2 Sensory substitution

Sensory substitution refers to the effect of replacing one natural sensory modality by another which is artificial. As introduced previously, such a process is possible because what is lost is not the ability to hear or see, but rather the capacity to transmit the information from the sensor –eye, ear– to the nervous system.

The pathways linking the sensor with the brain, however, remain intact and may therefore be employed by other modalities to send information to the corresponding part of the brain. This flexibility, added to the plasticity in the brain, underlies the principles of sensory substitution which aim to replace a natural modality by an artificial one. They are at the origin of sensory feedback and sensory restitution in the field of prosthetics.

2.2.1 Field of prosthetics

Many technical improvements have been made in the past decades in an attempt to make prosthetic devices better accepted by patients [16]. The main concern is to help subjects feel the hand as their own, both by providing them with natural sensations and simplified manoeuvrability.

Non-invasive approaches

Non-invasive technologies are becoming more commonly employed to help manipulate artificial limbs. *Control* may for instance be achieved based on signals recorded by electromyographic techniques (EMG), which encode muscle activity and make it possible to link human voluntary intention to robot automation.

Similar procedures have also been suggested in order to provide *feedback* to the patient about the prosthesis, hence creating a closed-loop sensorimotor model that may simplify manipulation. Artificial feedback can for instance be attempted by presenting information to the skin, either by electrical or mechanical stimulation, and may be achieved through electrotactile, teleproprioceptive of vibrotactile techniques [11].

These techniques exploit the physiological properties and responsiveness of tactile receptors of the skin, which are abundant in certain locations such as the fingertips. Models must however account both for their rapid adaptation to stimuli and for the spatial resolution of mechanoreceptors, and need thus to provide changing stimulation in intensity or location. the exploration of the correct encoding is an essential part of sensory substitution, and will be closely looked at in the framework of our later experiment.

Such procedures have already been tested under certain situations with different degrees of success [2], but are not widely spread to this day. In fact, most movements –such as grasping– are still carried out ‘semi-automatically’ by the mechanism through a simplistic two-input –open or close– system for which the only feedback available to the patient is his own visual perception of the scene. Improving the role of feedback would certainly help patients feel the hand as their own, and enhance the every-day use of the artificial limb.

Invasive techniques

Other **invasive** approaches may also be referred to. These work by connecting electrodes to neurons in the cortex or in the peripherical nerves. Brain activity is then be employed to extract motor command and decisions –such as volition– or to provide feedback. They have been implemented successfully for instance to restore motor functions in paralyzed patients [14]. Stimulation of cortical sensory areas can also be applied to help make the prosthesis feel as belonging to the patients’ own body. Yet these techniques are technically challenging and have been mainly explored in the case of monkeys so far.

The perspective in this thesis is restricted to the first models, and is intended to help improve non-invasive techniques that may simplify patients’ everyday life without requiring complex surgery or implants.

2.3 Goals & Outcomes

Given the state-of-the-art in multisensory integration previously introduced, along with the 'technical' description on current advances in prosthetics, we address the following questions:

Question 1: Vibrotactile Perception

Is the use of an artificial modality performed irrespectively of the encoding it presents itself in and, if so, which considerations need to be accounted for in each case to ensure that optimal exploitability is achieved? The analysis of subjects' responsiveness when confronted to different encodings logically follows, in order to establish the best bandwidth and combinations of parameters.

Question 2: Optimal Integration

Does the Optimal Integration theory –introduced for the case of healthy humans– also apply to the integration of an artificial stimulus? Should that be the case: Is this optimality achieved more easily under certain conditions of uncertainty? Is optimal integration continuously performed, or does it actually translate an alternation unimodal decisions?

Let us also keep in mind the final application of our work, namely to derive how these questions may help refine the design of feedback channels in the case of a prosthetic hand.

3
Methods

Tackling the problem of optimality... The study of optimal perception and integration is here undertaken under two different approaches.

1. **Within modality:** We carry out measurements of tactile sensitivity in the framework of a simple perceptual task so as to analyze optimality for different *feedback encodings*, both in terms of accuracy and exploitability. The capacity for patients to improve manoeuvrability depends strongly on how simple and natural the feedback provided is, especially since it is to be naturally combined with other senses. Issues such as bandwidth or acceptance hence appear as logical ones, and are essential in order to provide patients with the most suitable encoding. A complete evaluation about within-modality discrimination can then be derived.

2. **Across modality:** This evaluation concerns the use that subjects make of multimodal feedback, i.e. how they integrate each channel in order to achieve optimal actions. It is undertaken by looking at the performance during a tracking experiment which was especially conceived to isolate the principles underlying sensorimotor integration and to point out characteristic patterns. Our study evaluates subjects behavior from a statistical perspective and compares it to theoretical MLI and Bayes' models, in an attempt to verify whether they reacted to both stimuli in an optimal manner, or rather by giving preference to one modality over the other.

3.1 Experimental Background

3.1.1 Materials

Tactile Feedback techniques

Two *vibrotactile arrays* are employed, developed by I. Saunders [18], which are composed by eight independent motors, and which may be used to provide feedback either as 'location encoded' information (only one single motor vibrates at a time at a given location) or 'frequency encoded' information (all motors vibrate at a given frequency) proportional to the grasp closure.

The principles of such models were first described in Kaczmarek et al. [11], along with their benefits. Experiments with artificial limbs were performed by means of 'frequency encoded' tactile feedback [16, 2], in which patients were asked to grasp objects with and without visual control.

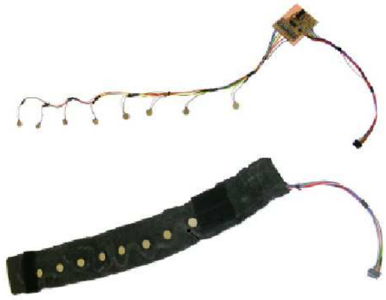


Figure 3.1: Vibrotactile array with 8 independently controlled motors, each one of which can vibrate at different frequencies (extracted from Saunders and Vijayakumar [18]).

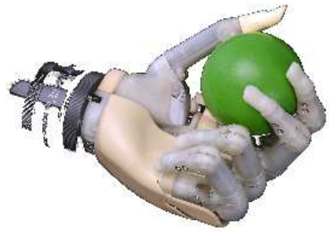


Figure 3.2: 'i-limb': Fully articulated artificial hand commercialised by Touch Bionics (extracted from www.touchbionics.com).

These showed that the use of visual information was reduced after several minutes, hence providing promising results about the use of the tactile feedback to simplify the control of the artificial hand.

'i-limb', the final application...

The final application of this work is to help improve the manoeuvrability of the artificial hand shown in Figure 3.2. This fully articulated artificial upper-limb provides compliant grip through four independently powered fingers that may be actuated separately –enabling a much wider range of movement than other traditional models– along with a thumb which may be moved by means of a passive joint. The flexibility is therefore strongly improved. Control is performed by means of two-input EMG controllers.

Force of closure, proprioceptive information or tactile sensitivity are, in this context, to be encoded and transmitted to the patient in the most effective and natural way possible in order to exploit the functionality offered by the device.

Our analyses concern (1) the type of encoding and the way it is perceived by patients, and (2) whether optimal integration can actually be achieved in the aforementioned context. I detail hereafter the experimental setup developed by I. Saunders, at the core of my later evaluations, which is relied upon to tackle these two fundamental questions.

3.1.2 Sensory Cue Integration Experiment

The experimental setup here presented aims to isolate the effects of the artificial modality and to concentrate on multisensory integration. Wrapped and presented to the participants as a target-reaching task, the experiment requires spatial estimation of combined visual and tactile modalities shown over time. Control is performed by means of two 'buttons' just like the ones that are employed to control the hand. The simplified setting helps focus on our targeted evaluations; It is thus at the core of our forthcoming evaluations.

More specifically, the setup helps emphasize whether subjects are capable of using the feedback provided and, should that be the case, to highlight whether the artificial stimulus is optimally integrated with the natural modality.

Basic goal, target & cues

The experiment involves the estimation of the position of a cursor, only presented to the participant under the form of cues in both modalities, i.e. visually on a screen and through motor vibrations for the tactile feedback. These cues are *noisy* versions of the underlying cursor; They are extracted at a frequency of 20Hz from two Gaussian distributions modelling the independent jitter that affects each feedback channel. That is, $\mathcal{N}(x|\mu_v, \sigma_v)$ for the vision and $\mathcal{N}(x|\mu_t, \sigma_t)$ for the tactile modality. The illustration of the setting is shown in Figure 3.3.

Each trial lasts for three seconds during which the participant is asked to drive the cursor towards a fixed target, located both in the middle of the screen and in the middle of the arm. This correspondence across modalities prevents from giving priority to one over the other, and helps keep the goal simple. The initial position of the cursor (from which the cues are extracted) is randomly chosen, and changes for every trial.

In each trial, the user performs the task under different degrees of difficulty, i.e. the combination of the parameters $\mu_t, \sigma_t, \mu_v, \sigma_v$ is modified. The different values of jitter affecting each modality are shown in tables 3.1 and 3.2. The whole test is then organized in blocks, during which *all* sets of combinations are presented to the participant. The test consists of four blocks, which can then be confronted so as to compare how people react to *varying stimuli*, and specifically to study the improvement over time or to assess whether certain combinations of uncertainty are better coped with than others.

Blocks, learning & Training

The experiment is organized in two distinct parts, (1) first a *training* period for each one of the modalities during which only one of the modalities is to be employed, and (2) the actual *test* where both modalities are employed and presented with varying degrees of uncertainty. It may be noted

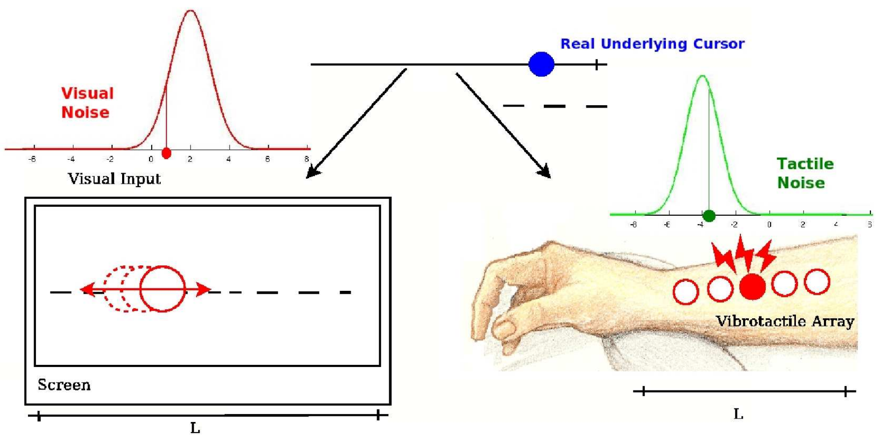


Figure 3.3: Setup of the experiment. A cursor –not revealed to the participants– is perceived through two modalities affected by independent Gaussian noise, i.e. cues are presented to the subjects (1) as a dot on a screen with uncertainty σ_v for the visual modality and (2) through motor vibrations with uncertainty σ_t for the tactile one.

that even during the training, both modalities are maintained. This is so as to ensure that the person gets used to having both even in a case where he cannot rely on it.

At the end of each trial, the subject receives a numerical feedback –'reward'– that informs him about his performance during the trial in terms of how far away he placed the cursor from the target.

Table 3.1: Amount of jitter in pixels during training

	Visual Training	Tact Training
σ_v	20 / 80 / 160	400
σ_t	400	20 / 80 /160

Table 3.2: Amount of jitter in pixels during the test

	Test
σ_v	20 / 80 / 160
σ_t	20 / 80 / 160

3.2 Effect of vibrotactile encoding

Our first question deals with the evaluation of the encoding that is to be employed so as to provide the most profitable feedback to the patient, both in terms of bandwidth and exploitability. We aim to calculate for that purpose the 'Just Noticeable Difference' (JND) for several encodings, and to derive from it the number of channels that are perceived by the user. These may then be employed to provide feedback. Several issues affect the capacity of humans to exploit each encoding. Certain models may therefore provide increased capabilities, be more easily learnt and thus be more suitable than others for sensory substitution in our current context.

3.2.1 Types of encoding

Different encodings will be evaluated for the case of our vibrotactile array. **'Location encoding'** –different motors vibrate along the array so that the perception changes locations– and **'frequency encoding'** –a single motor vibrates at different frequencies– are logical steps to consider.

The later has the advantage that it minimizes the space required to provide the feedback and yields more compact and portable systems, as compared to location encodings which need different motors around the arm or chest at a certain spatial resolution. A direct advantage that derives from this concerns the location of stimulation, i.e. vibrations can be applied at the fingertips where tactile receptors are more abundant.

However, due to the fast adaptation of the tactile sense , it is still necessary for vibrations to be provided through at least two alternated motors, ensuring that enough sensitivity is preserved. The setup for the experiment is shown in Figure 3.5.

Here we explore an alternative technique to frequency encoding, which combines changes in both Pulse-Width (PW) and Period (T), so that the perceived intensity varies. This **'intensity encoding'** was shown by Liang et al. [10] to yield good results when mapping the code to force levels. We therefore evaluate the possibilities that it offers in the context of our vibrotactile array. More specifically, we explore the 'PW-T' Space –independent changes of one magnitude or the other– in order to point out which combinations seem to be better exploited under our experimental conditions.

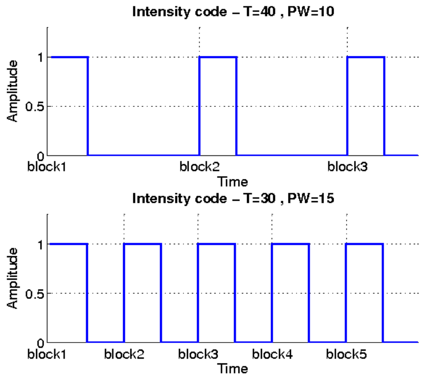


Figure 3.4: Intensity code for two different sets of the couple (PW,T). Let us emphasize that, although identical ratios PW/T may be obtained through different combinations of the two magnitudes, the perceived intensity varies, i.e. the stimuli 30/60 and 20/40 embody different intensities.

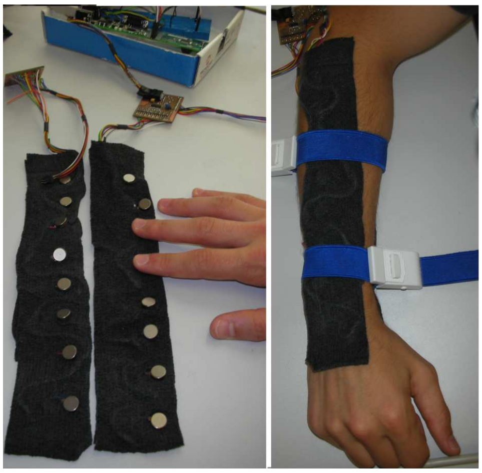


Figure 3.5: Experimental setup to evaluate the sensitivity for the *intensity code* (left) –for which cues are presented at the tip of the fingers– and for the *location code* (right). The later is applied through vibrations of different motors along the arm.

3.2.2 Perceptual Sensitivity: 'JND'

The Just Noticeable Difference (JND) is the minimal amount of variation by which a stimuli can be affected so that the person can detect the difference. It provides a degree of the sensitivity around a given centre –location, frequency– and may be employed as *discrimination threshold* to determine the distance that is to be kept between different stimuli so that these can be easily interpreted by the user as being distinct.

Several techniques can be used to evaluate this value. Two stimuli may be used in a forced-choice task, where the subject is asked to chose between two choices, for instance by selecting the biggest out of the two. Subjects are expected to be capable of telling the difference until the point where the two stimuli are perceptually indistinguishable, in which case their choices simply translate random guess, i.e. 50% chance of being correct.

Perceptual Experiment

We measure the aforementioned discrimination thresholds by using staircases. Participants are presented with two stimuli separated by a difference D. This difference between stimuli is reduced to $D-\delta^-$ if the answer is correct, or inversely it is augmented to $D+\delta^+$ when a mistake is made.

Likewise, the steps δ are chosen of varying size so as to gather information about the 75% point –i.e. the middle point between random choice and totally correct choices– in as fewer steps as possible. We hence develop a model that **converges** towards the 75% point. According to Garcia-Perez [6], a good policy is a three-down / one-up rule –three consecutive correct choices required to compensate for a single error– in which the ratio between 'steps up' and 'steps down' is given by $\delta^+/\delta^- = 1$. Such a model can be implemented by so that $\delta^+ = \delta^- < 1$, and $D = D * \delta$ in case of correct choice or $D = D/\delta^3$ in case of error, as illustrated in Figure 3.6.

In addition, the participants are asked to wear earplugs to ensure that no auditory information is employed, and that only tactile modality is relied upon to distinguish the frequencies. Also, in

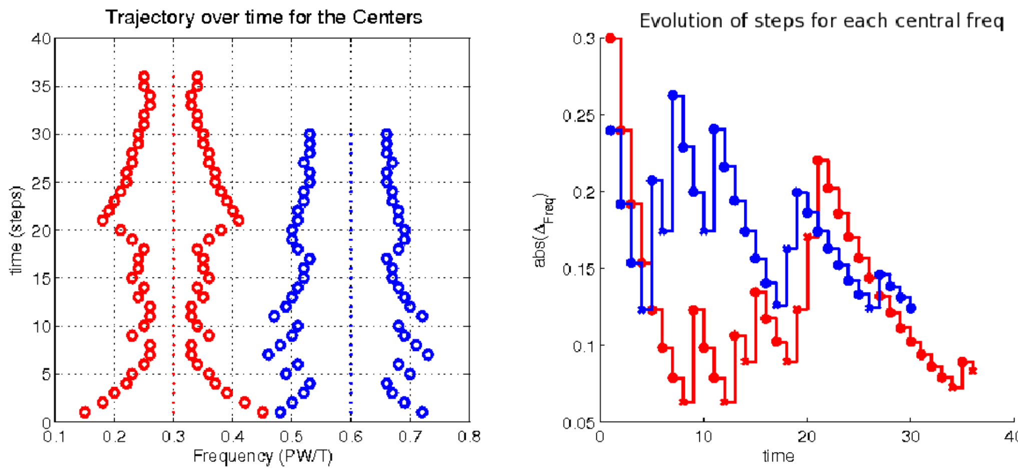


Figure 3.6: Procedure to measure the JND. **Left:** Cues are presented symmetrically to the participants around two centres. The distance between the stimuli changes over time in response to the person choices. **Right:** Illustration of the steps in time. The 'three-down / one-up' rule is employed along with a time-varying step-size δ in order to converge towards the 75% JND point.

the case of the intensity encoding, two motors are used, which are randomly chosen and employed to ensure that the fingertips do not get insensitive to the vibrations, nor that the motors heat up excessively.

Inverse Gaussian fitting

Psychometric functions are commonly used to describe the relationship between stimuli and responses. Distributions such as cumulative Gaussian or logistic are widely employed for that purpose. They can be fit through least squares to statistical points which translate the percentage of correct answers for a given choice, i.e. in which the person recognised the bigger stimulus as been bigger, and wrong otherwise [8]. The inflexion point may then be be used as 75% sensory threshold.

In our case, because we intend to establish the bandwidth of each perceptual channel, we fit an inverted Gaussian to the 'absolute' percentage of correct choices –where 'absolute' implies that both bigger stimulus detected as bigger, and smaller ones interpreted as smaller are considered, as compared to the aforementioned method which only accounts for the proportion of trials where the stimulus was perceived as bigger–. The result is equivalent to fitting a psychometric function, but it simplifies the representation of each different channel, along with its corresponding bandwidth.

The inverted Gaussian that is employed to fit the statistics is thus as follows:

$$g(x) = 1 - 0.5 \cdot e^{-x^2/2\sigma^2} \tag{3.1}$$

which is a function that always has a minimum at $(0, 0.5)$, i.e. the centred function has a minimum which translates random decisions (50%). This representation assumes that perception is symmetric, and is illustrated in Figure 3.7

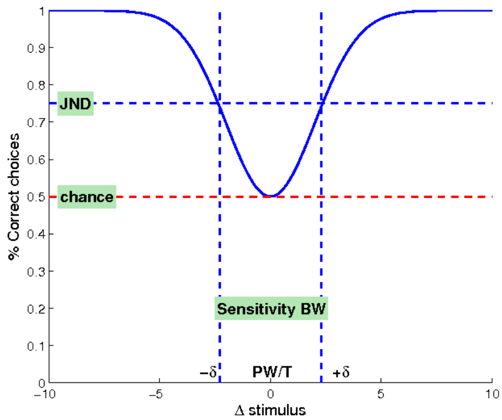


Figure 3.7: Inverted Gaussian employed to determine the JND, and corresponding Bandwidth of a perceptual channel.

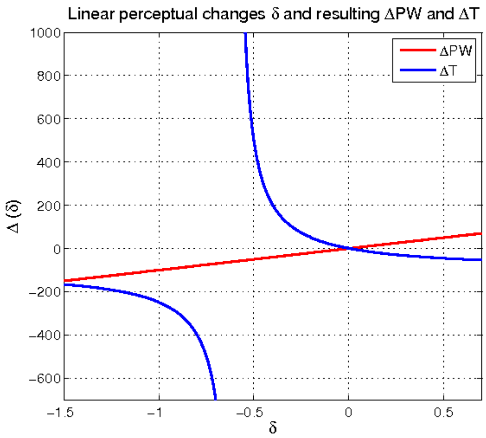


Figure 3.8: Changes ΔPW and ΔT that are derived from modifications δ which are applied around a given frequency $PW/T = 60/100$

Experimental Considerations for Perceptual Linearity

When tackling the question of intensity encoding, the evaluation of how *individual* changes in each one of the magnitudes –pulse-width (PW) or period (T)– affect human perception needs to be undertaken carefully.

We may need to bear in mind that changes in T do not directly derive into linear perceptual changes, as demonstrated in Equation 3.3. Hence the JND values that are to be extrapolated from the Gaussian fitting do not correspond to the expected symmetrical sensitivity either. Indeed, the measurements for our sensitivity $\pm\delta$ (as illustrated in Figure 3.7 around a given centre PW/T) are expected to embody the corresponding change ΔT, i.e.

$$\frac{PW}{T_i} + \delta = \frac{PW}{T_i + \Delta T} \tag{3.2}$$

Yet it follows from this expression that:

$$\Delta T = \frac{PW \cdot T_i}{PW + \delta \cdot T_i} - T_i \;\Rightarrow \Delta T(\delta) \propto 1/\delta \tag{3.3}$$

The graph that illustrates this non-linear relation is shown in Figure 3.8. Inversely, changes in PW follow linearly from δ as follows:

$$\begin{aligned} \frac{PW_i}{T} + \delta &= \frac{PW_i + \Delta PW}{T} \\ \Leftrightarrow \Delta PW &= \delta \cdot T \;\Rightarrow \Delta PW(\delta) \propto \delta \end{aligned} \tag{3.4}$$

In order to ensure that the **duty cycle** applied to the subject is indeed linear as PW or T are modified –and that the *perceived frequencies* vary linearly as expected–, the points are thus to be extracted from curves in Figure 3.8 (where ΔT and ΔPW are represented as a function of parameter δ). As a matter of fact, this parametrical description serves as basis for the model used in [10] and was proved to yield interesting coherent results.

3.3 Statistical Modelling

Our second study addresses the question of **how people combine multimodal feedback**, and is attempted from the perspective of statistical analysis. The multisensory cue integration experiment that has been presented in section 3.1.2 lends itself to numerous evaluations, and offers the possibility to focus on how people react to different stimuli and why.

Analyses in the literature usually restrict their evaluations to the final position of the estimates when tackling the question of discrimination **across-modality**. In our case, we will attempt to go further and analyse the decisions over time, based on the movements of the subjects and what compels them to act, i.e. using their behavior to infer their beliefs.

Two elements are available from our experiment:

- On the one hand, the trajectories of the cursor during each trial –3s– are recorded. They embody the *actions* of the person, i.e. his decisions over time.

- Secondly, the cues that have been presented to the subject during each trajectory are also stored. These correspond to the *stimuli* that the subject has detected all along the trial, both through the visual and the tactile modality. They are supposed to have been processed and integrated over time, compelling the subject to behave like he does.

Both elements are illustrated for a given trial in Fig. 3.9. We hence attempt to come up with a statistical model in order to provide a **theoretical framework by which to link what the person perceives to what the person does**. That is, his perception to his behavior.

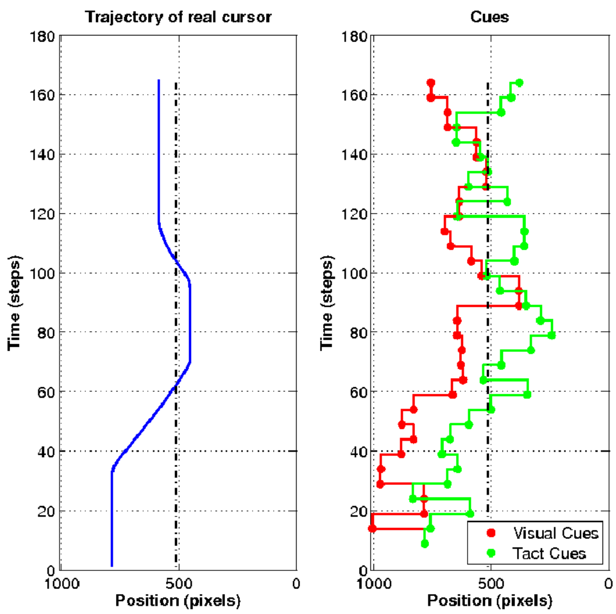


Figure 3.9: Example of the trajectory of the cursor (left) and cues presented to the subject (right). The trajectory embodies the *actions* of the subject, whereas the cues represent what he *perceives*. The link between both magnitudes would help get acquainted with the way in which humans process different modalities and how they react to them in order to achieve a goal.

Our aim is thus to model the *beliefs* of the subject so as to further understand the reasons that push him to act and, more specifically, to analyze how the two different modalities are employed.

Specifically, we intend to verify whether the optimal integration theory also works in the case of our experiment, in which one of the modalities is artificial. An optimal response under the current circumstances would hence imply that the subjects do respond to *both* modalities, and that the contribution of each one of them depends on their individual estimated reliability. The actions of the subject would thus also account for such elements when responding to the stimuli.

3.3.1 Framework

Two widely-used machine learning techniques have been introduced, along with their relevance to model human sensory integration, namely Maximum Likelihood and Bayesian Inference. We attempt here to apply them to our framework and to derive insightful evaluations about the way in which the artificial modality is employed in the context of the 'cue integration' experiment previously described. The two methods share similar characteristics. Yet the emphasis that is accorded to each one of them in our study is different, and so is the statistical analysis that follows.

The fist model is suitable in a framework where priors are expected uniform at start and where all cues are assumed to be constantly integrated over time. The second technique, however, exploits specific patterns during trials, and distinguishes perception time-periods from actions. It hence embodies the principles of the sensorimotor loop. Priors are here employed as underlying the learning process and account for interleaved time periods where the subject only 'acts'.

Two main analysis logically follow: (1) A statistical study of the **final estimates** at the end of the trials, and (2) an evaluation of the responses **over time** during each trial, focused on whether optimality is achieved as an iteration of unimodal responses.

3.3.2 Maximum Likelihood Integration

Maximum Likelihood is a well-known statistical model employed to calculate estimates of the parameters that underlie a distribution for which certain data points $\mathscr{D} = \{x_0, x_1..x_N\}$ have been observed. It hence evaluates the 'most likely' value given what has been perceived, i.e. the parameter for which the probability of the observed data is maximized:

$$\hat{\theta} = \arg\max_{\theta} \mathscr{L}(\theta) = \arg\max_{\theta} p(\mathscr{D}|\theta) \tag{3.5}$$

Under this framework, the estimate $\hat{\theta}$ –assumed to underlie the distribution that governs x_i– is fully derived from the observations. The states of the world can be explained **by considering only and exclusively what is being perceived**, as compared to Bayesian models which introduce internal beliefs that 'distort' perception by considering personal factors.

Such situation is especially relevant for cases where no prior information is available, or when it is unknown. It also prevents from considering a prior distribution that results more from mathematical convenience than as a reflection of real prior beliefs [1].

Maximum Likelihood Estimation has also been proved to explain the way in which human brain responds to the world, and has been specifically explored in the case of multisensory integration.

Indeed, the principles described by the Maximum Likelihood framework were shown by Ernst and Banks [5] to distinctly explain human visual and tactile integration in the case of a pointing experiment under different degrees of modal uncertainty. Interestingly, it underlined that each modality was been relied upon proportionally to their degree of reliability. Further examinations of the model have also been undertaken for shape discrimination from visual and tactile cues [8].

We verify hereafter whether such context may help deepen our understanding of how patients integrate the artificial and the natural modalities.

Implementation

Maximum Likelihood theory states that, for each trial, the **final estimates** –after having perceived N cues x_i– that follow equation 3.5 are given by:

$$\begin{aligned}
\hat{\mu}_{ML}^{N} &= \frac{1}{N}\sum_{i=1}^{N} x_i \\
\hat{\Sigma}_{ML}^{N} &= \frac{1}{N}\sum_{i=1}^{N} (x_i-\mu)(x_i-\mu)^T
\end{aligned}$$

The results for different degrees of noise are shown in Table 3.3, along with the optimal integration that would be expected under these conditions.

A specific example of a single trial may be seen in Figure 3.10. It may be noted that here the subject is placing his final tactile estimate –represented in green– on the target, regardless of other considerations. The optimal estimate –black dot– is as a matter of fact, completely shifted to a side, therefore suggesting that the participant was completely tactile-biased when relying on his estimates for this specific trial, and that his final decision was *based on unimodal perception*. Let us also highlight that the final position is not changed in the last 60 steps of the trial, indicating that the person was confident about his estimate and his choice.

A deeper analysis of how the participant estimated each cue as these were perceived may help better understand his choice. Indeed, it is expected that the mean and variance of each modality be **sequentially estimated** as new information is presented, and then relied upon to build an integrated mapping of the underlying distribution. The study of cue integration can thus be attempted from a temporal perspective, i.e. focusing on the evolution that leads to the final estimates. This may help analyze for instance how the participant came to place the cursor so far away from the target.

Maximum Likelihood theory also provides a simple model to account for the influence of new individual cues as these crop up. Their individual contribution is indeed stated by [1]:

$$\begin{aligned}
\hat{\mu}_{ML}^{N} &= \frac{1}{N}\sum_{i=1}^{N} x_i \\
&= \frac{1}{N}\left(x_N + \sum_{i=1}^{N-1} x_i\right) \\
&= \frac{1}{N}x_N + \frac{N-1}{N}\hat{\mu}_{ML}^{N-1} \\
&= \hat{\mu}_{ML}^{N-1} + \frac{1}{N}\left(x_N - \hat{\mu}_{ML}^{N-1}\right)
\end{aligned}$$

hence indicating that the observation of the N^{th} cue allows to refine the previous estimate of the mean, but that **its individual contribution is iteratively reduced**.

A simple example of applying this framework to the previous trial is shown in Figure 3.10 (right graph), where the effect of each single cue is quantified, integrated into the previous belief to generate a new estimate and then centred around the current position of the cursor given by the trajectory.

Interestingly, this model gives an idea of what the subject is 'theoretically' assumed to have believed during the trial. It also gives some insight about the evolution of his estimates. For instance, during the first half of the trial, his actions appear to be driven by his estimate of the multimodal ML estimator, insofar as the computed estimator coincides with the target at the end of the trial. During the second half of the trial, however, his attention appears to have shifted to a unimodal situation, where he only concentrates on tactile information. This does indeed successfully compel him to place the tactile estimate in the middle of the screen.

Interesting ideas can be drawn from the analysis of trajectories. More specifically, one automatically is tempted to wonder whether optimality is actually achieved as an alternation of unimodal decisions. The Bayesian framework is employed hereafter to better exploit the characteristics of the trajectories and to extract more information about the role of each individual modality, and its affect in the overall decisions.

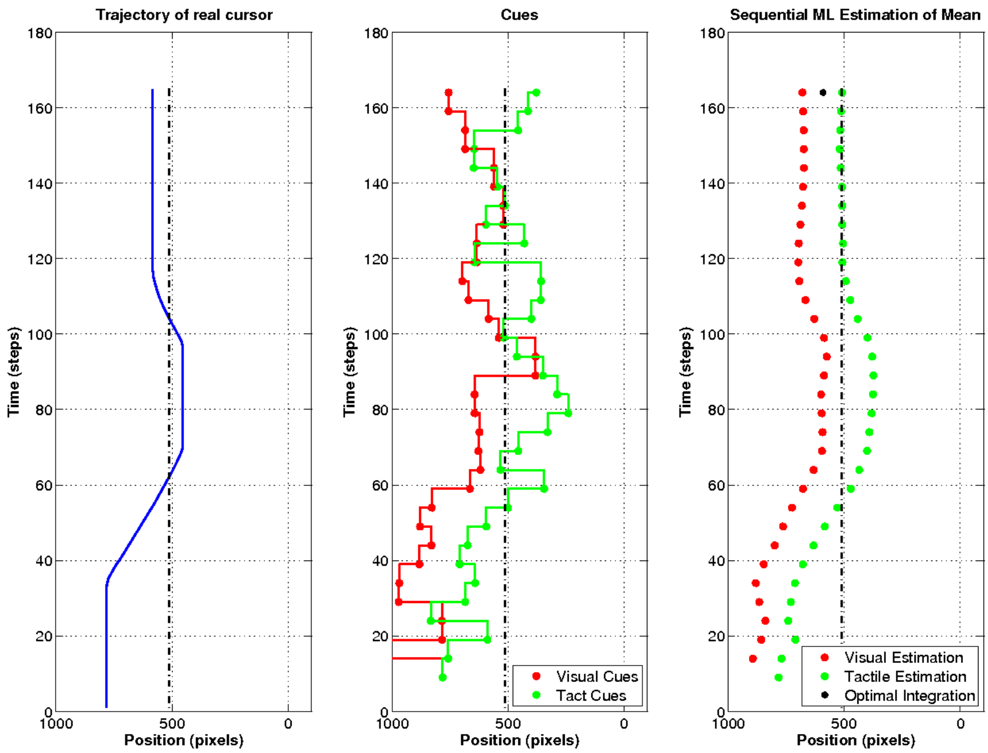
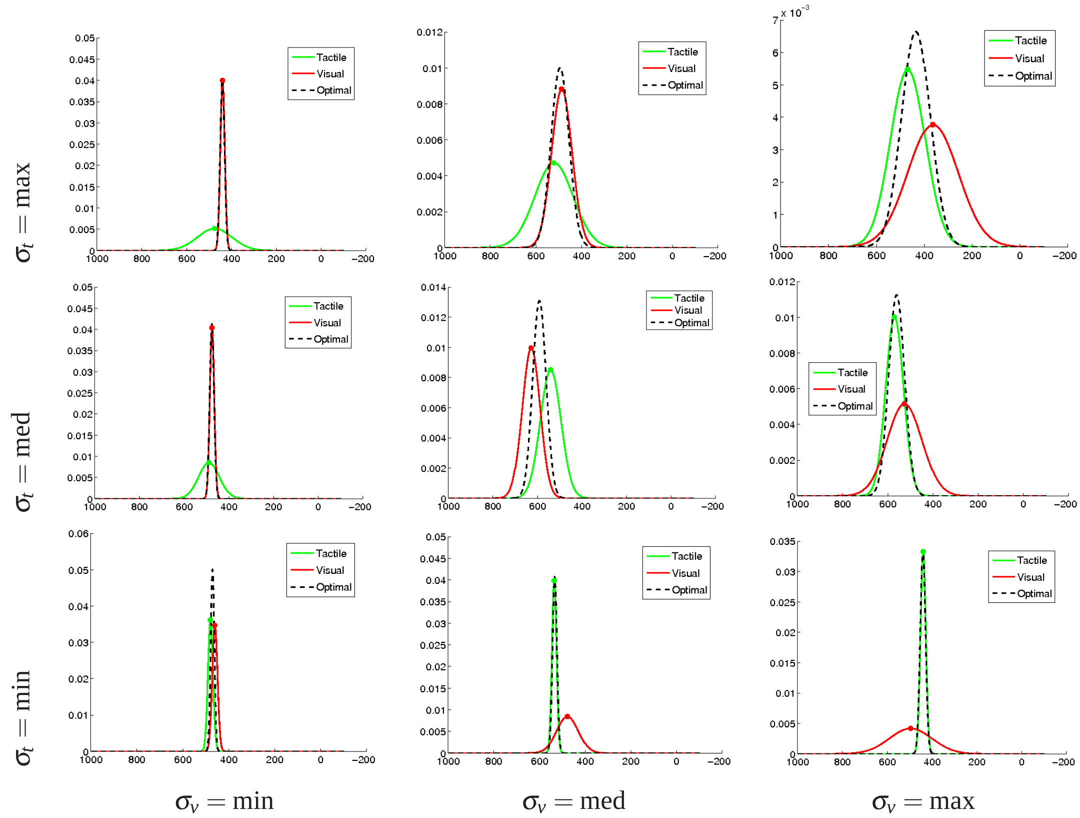


Figure 3.10: The trajectory of the cursor (left) illustrates the decisions of the subject over the time of a trial as the person detects the cues shown in the middle graph. The sequential Maximum Likelihood Estimators (right) point out the belief over time for both modalities, along with the theoretical optimal for the ending point.

Table 3.3: Example of final distributions that may be expected after iteratively integrating random cues extracted from both modalities under different degrees of jitter. Columns represent identical amounts of visual noise, and rows identical tactile feedback.



3.3.3 Bayesian Model

Let us image for a minute that you are driving a car in a busy road. When looking around you to evaluate the traffic, your senses are constantly being used to gather information about the location of other cars. This enables you to estimate their speed, predict their future positions and to adapt your actions accordingly.

Should a second car appear on a side road and be about to join the road on which you are driving, it will be expected from you to change lanes. Several previous steps are necessary: You might need to estimate the position and speed of the second car, based on the integration of consecutive images over some milliseconds, and then predict his future location. A quick look though the rear-view mirror ensures that it is safe for you to perform the movement. A final glance at the car helps assert your predictions about his current position.

Should it be then the case that the car has stopped unexpectedly, its new position will not match your **predictions, which were built during the previous estimation context and relied upon as a-priori belief for the current evaluation**. You will thus be required to instantaneously adapt your understanding of the world, i.e. your internal model, to the new framework.

The idea that previous information is relied upon as priors for the next estimations is not new. We have introduced such a model in the background of this Dissertation. Interestingly though, in many real life situations such a process seems to be **interleaved by 'action' time-periods**. Namely, in our previous example, looking through the rear-view mirror creates a gap of time which differentiates two 'evaluation' time periods.

Perceiving vs. Acting

The whole perception and action framework –underlying the sensorimotor loop– can indeed be seen as being composed of two distinct periods. (1) **Evaluation** time periods during which the subjects are spectators of the world –they try to gather new information and to refine their estimates– and (2) **actions** during which the previous information is relied upon to predict upcoming events and to act accordingly. Forthcoming evaluations will then rely on previous representations, and employ them as 'prior' belief.

The idea of such 'dual' process has indeed been put forward as underlying the way in which humans learn, and used as keystone for several machine learning techniques. It is for instance at the core of Reinforcement Learning systems, which rely on an iterative alternation between *exploration* –gathering data– and *exploitation* –making use of that information– so as to maximise a policy or to minimize a certain cost. Knowledge is thus accumulated and refined over the time as cues are detected, and then used when acting [7, 20].

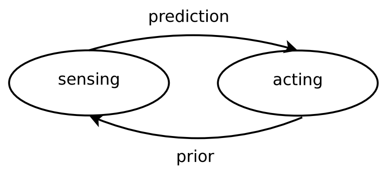


Figure 3.11: Schematic flowchart of the sensing & acting loop, along with the use that probabilistic information is given when transferred from one to another.

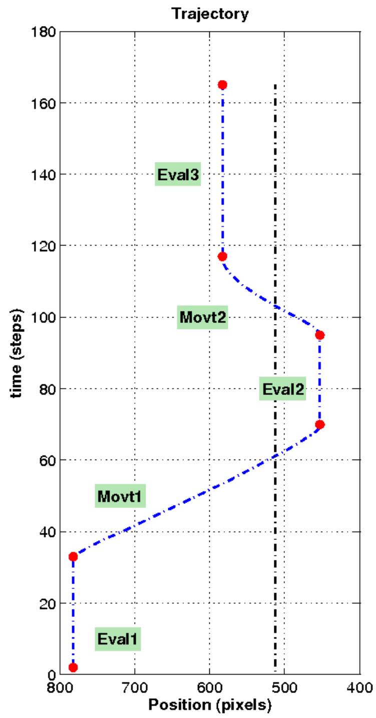


Figure 3.12: Trajectory for a given trial, and limits of 'decision' and 'movement' time periods. Movement movt$_{i+1}$ is motivated by the information gathered during eval$_i$. That same information is then re-used after the action as 'prior knowledge' for eval$_{i+1}$.

During our experiment, subjects did indeed exhibit such pattern, as clearly illustrated by the way they moved the cursor (Fig 3.12). The trajectories are indeed composed of two distinct patterns, first an evaluation time period during which the subjects do not move the cursor, and secondly a movement periods in which they use their estimations to try and correct the error that is perceived.

This suggests that their progressive learning process interleaves periods of estimation and action. The analysis of how people utilize the two modalities during our experiment may hence be undertaken from the perspective of this 'dual' behavior.

Under such framework, movements can be understood as being compelled by the momentary belief with which the subject has come up. His actions thus attempt to correct the error that is inferred from the cues –either from a single or from both modalities– and they give answer to what the subjects believes.

One the movement finished, during the following evaluation time period, new cues are used to *refine* the initial estimations that were generated before the action. These can now be used as a **prior belief** –as explained by Bayesian inference theory– in order to improve the understanding of the scene. As such, new cues are given a more important weight than older ones, hence accounting for the natural human tendency to 'forget'.

Implementation & Benefits

Our analysis accounts for the aforementioned ‘dual’ behavioral model, and hence limits the integration of cues to those presented during ‘evaluation’ time periods. The theoretical belief is hence derived from them, and can then be employed to study the immediately following action and the contribution of each modality to the movement.

Several benefits are derived from this structure. Firstly it provides a realistic model in which the shape of the trajectory is better exploited, making it possible to establish a direct link between what compels the subject to act and his actual movements. The analysis of how subjects make use both modalities to guide their search is simplified, enabling for instance to verify whether both visual and tactile are constantly relied upon or whether, over time, there is a tendency to give preference to one modality over the other.

Secondly, unlike the previously presented ML model, new cues are now given a bigger contribution, hence accounting for the human tendency to ‘forget’. Indeed, previous information is only contributes to the new estimates as ‘prior’.
As a result, the posterior belief after having detected a new cue is much more influenced by the new information, as stated by:

$$\begin{aligned}
\hat{\mu}^N_{\text{post}} &= \frac{\sigma^2_{\text{new}}}{\sigma^2_{\text{prior}}+\sigma^2_{\text{new}}}\mu_{\text{prior}} + \frac{\sigma^2_{\text{prior}}}{\sigma^2_{\text{prior}}+\sigma^2_{\text{new}}}\mu_{\text{new}} \\
\hat{\sigma}^N_{\text{post}} &= \frac{1}{\sigma^2_{\text{new}}} + \frac{1}{\sigma^2_{\text{prior}}}
\end{aligned}$$

Initially The prior is chosen to be very flat, since no information is known about how the cues are going to appear. Also, this gives more importance to first cues.

Remarks: This framework assumes that movements can be fully understood –and predicted– from the information gathered during immediately preceding ‘estimation’ periods. This implies that (1) new cues detected during the actual movement do not influence the motion, and that (2) they do not modify the internal belief previously created. In other words, participants would be expected to react in the same way, should they not have any feedback during the movement.

The second point can be easily argumented. Should it be the case that new cues perceived during the movement change the internal belief, the participant would then be expected to estimate over that time –less than half a second– the mean of noisy cues which are randomly picked from of a cursor, itself moving over time. This context seems hardly conceivable, and the idea that the belief is maintained throughout the movement appears as plausible.

The first assumption is more difficult to evaluate. Indeed, certain detections perceived during the movement may perhaps influence the motion, especially if the new cue is very contrary to what was predicted from the previous internal belief. The influence affecting the movement would hence be due more to the inconsistency brought by the cue, than to the cue itself. In this case, attempting to analyse the movement as a function of the belief still holds. In any case, dealing with such degrees of conflictive information is not expected in such a short time, and trajectories are assumed to be mostly intended to correct the error perceived, rather than to respond to punctual individual stimuli.

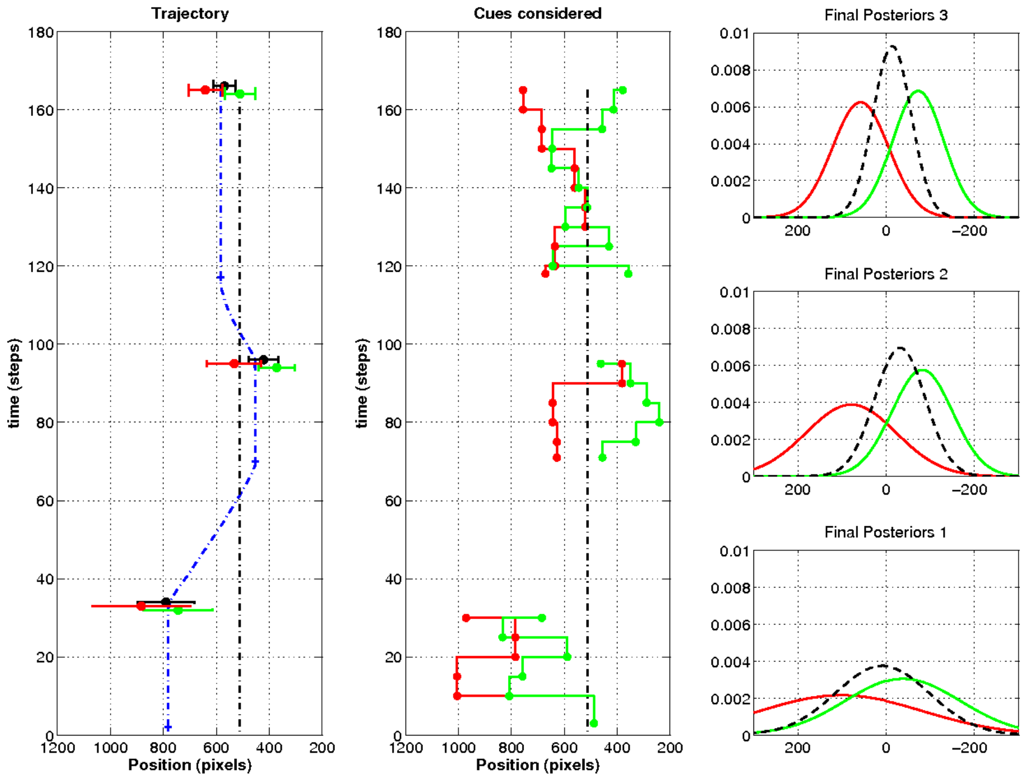


Figure 3.13: Example of a trial (left) and cues considered (centre). The corresponding beliefs after each 'evaluation' time period are shown on the right, and their corresponding mean and variance superposed on the trajectory.

3.3.4 Comparisons

The two models presented exploit different characteristics, and are employed to analyze complementary issues:

- Maximum Likelihood is used to evaluate optimality by considering the final estimates, since it provides a comprehensive and complete model that accounts for all the cues without making major 'assumptions'. Such analysis may be perform from a statistical way –for all trials together– or by considering each single trajectory individually.
- Bayesian theory on the other hand may be used to emphasize the link between perceptual cues and corresponding movements, in an attempt to analyze the influence of each single modality. The framework of sensorimotor integration is here specifically exploited.

4
Analysis

4.1 Sensitivity Bandwidth

Our first evaluation concerns the optimal encoding that is to provided to the user, in terms of maximal amount of information that can be transmitted and which can be perfectly be interpreted. The sensitivity experiment previously described makes it possible to obtain results about **within-modality selectivity** for different centres, from which we intend to fit an inverted Gaussian and to derive the bandwidth of each channel.

We gather the recorded data into **bins**, which help determine the *percentage* of correct choices from the decisions of the participant –as compared to the actual measurements which were limited to either '1' when successful or '0' when erroneous. Indeed, fitting a function directly to the binary data points yields very poor results –usually a horizontal line. The bins, on the other hand, embody the actual performance –in percentual values– which is expected to have been extracted from the underlying psychometric distribution. As such, their correct quantification represents an essential step in our evaluation.

Our bins are chosen so as to maximize the *resolution* around the areas of interest for the fitting, i.e. the 75% point. Let us remind indeed that the strategy followed to record the data points

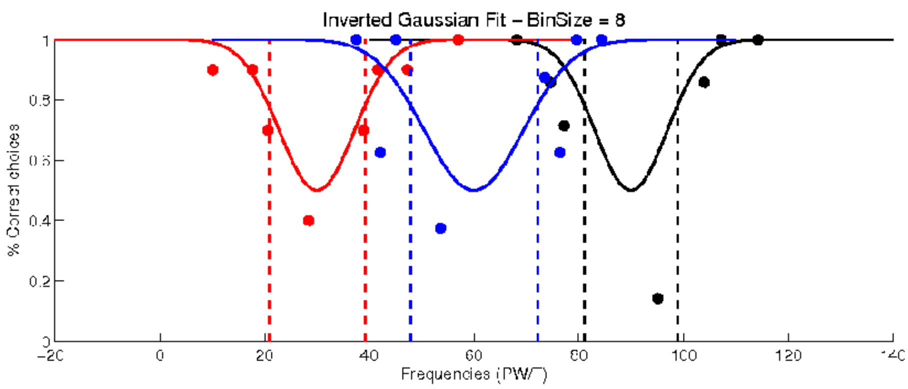


Figure 4.1: Procedure to calculate the bandwidth given measurements from three channels. 8 Bins are employed to determine the *percentage* of correct choices from the data points (1 when the choice is successful, 0 otherwise). Better resolution is clearly provided around the 75% JND point, which is the part of the fit that needs to be as accurate as possible.

did also attempt to converge towards the JND percentage, around which the confidence is hence increased. We hence decide to gather *identical* number of points within each bin, ensuring that more statistical points are generated where more data points were measured. As a whole, the procedure leads to points as shown in Figure 4.1, and clearly translates the shape of the expected underlying psychometric function.

Several seeds are employed to perform a gradient descent optimisation and the solution that minimizes the least-square error is maintained. The result for three centres is shown in the previous figure. This evaluation is then performed for the two different encodings so as to derive the characteristics that govern within-modality discrimination in each case.

4.2 Analysis of Final Estimates

Our second evaluation deals with **across-modality selectivity**, and concerns the position of the final estimates after having perceived visual and tactile cues during the tracking experiment. Participants are assumed to have been presented with enough information, which integrated over time should lead them to perform the task correctly.

4.2.1 Unimodal vs. Bimodal performance

A first evaluation that enables to assert whether participants do integrate both modalities concerns the effect that a second modality has on the initial single one, i.e. the accuracy of subjects under bimodal conditions, as compared to the performance that is achieved in unimodal situations. It would be expected that people benefit from the second modality and perform better at the task.

Such an evaluation has been attempted in the litterature by analyzing the statistical performance over the whole test [5, 8]. This is mainly due to the fact that only the actual 'responses' of the participants are employed for the analyses. These are logically to what can actually be recorded and, for the bimodal case, each single modality can only considered through the combined integration.

Our framework, on the other hand, makes it possible to study each *individual* –theoretical– estimation even under bimodal conditions, and hence enables to evaluate separately the effect that each modality has on the other one –and not just the combined. It thus gives more insight into how people combine both, which is particularly relevant especially since one of them is artificial.

4.2.2 Effect of Learning

An additional evaluation concerns the capacity of subjects to *learn* the new modality. It is indeed essential not only to ensure that patients are able to make use of the feedback provided, but that a certain improvement over time can also be expected as perceptual learning occurs.

We hence complement the previous analysis through the statistical evaluation of the *evolution* of their score during the training tasks. This exploits the convenient organisation of the experiment in identical **blocks**, especially conceived so as to provide participants with the same combinations of jitter in all three blocks. The results can hence be used to gather meaningful statistics about their evolution over time, and to derive conclusions about the range of improvement that can be expected in the context of real prosthetic applications.

Beyond global statistics... The presented evaluations help verify that multimodal integration does indeed happen, and that it is employed by subjects to be more accurate. But would it be possible to evaluate the degree of optimality of a single trajectory, rather than just looking at the average performances through the whole task? That may be interesting to understand whether a particular trial is more visually driven or tactile biased, or inversely whether people tend to constantly use both.

4.2.3 Optimality in each single trial?..

The model presented provides us with Gaussian-distributed estimates of the optimal integration, according to what Maximum Likelihood theory tells us. The mean and variance are indeed available. The later indicates the accuracy of the information provided to the subject, whereas the mean emphasizes his actual choice.

This may be compared with the location of the actual target, and relied upon to assess the performance. Indeed, the distance between the theoretical optimal estimate and the real target embody the error that the participant makes, and quantifies his performance. It should be expected that the estimated optimal mean is placed in the target for each single trial, in which case the error committed would be zero.

For that purpose, we make use of the error function, which gives the probability of an error occurring between $-c$ and c for a normal distribution centered around 0 with variance σ^2:

$$\mathrm{erf}(c) = \int_0^c e^{-t^2} dt \tag{4.1}$$

Since our final estimates are gaussian distributed $\mathcal{N}(\mu, \sigma)$, the value $\mathrm{erf}(\frac{\mu - c}{\sigma})$ provides a measure of the probability that the person commits an error when placing the final estimate at a distance $|\hat{\mu} - c|$ from the central target c (Figure 4.2). Alternatively, it may be seen as the **probability that the person actually generates his integrated estimate exactly as ML theory states**.

For instance, should a participant place his final 'theoretically' optimal estimate at a distance of 2 standard deviations from the target ($|\hat{\mu} - c| = 2\hat{\sigma}$), his probability of making an error given the gaussian knowledge that he is assumed to have gathered along the trial is 97%. Meaning: (1) either he is indeed integrating both modalities optimally but performing very poorly at placing the cursor where he wants it to be, or inversely (2) he is *not* integrating both as optimally as expected.

Such a normalised measurement accounts for the shape of the distribution, Gaussian at its core, and evaluates the performance in terms of how **the participant makes use of the information with which he is provided**. It corresponds to the cumulative gaussian error and provides a coherent way of comparing different trials.

Since the error function is applied to a distance (i.e. an absolute value), we may also view this evaluation from a different angle: The gaussian distribution is now *centered in the target*, and the error that is being quantified varies as μ moves along the X-axis. This error thus corresponds to the area between c (fixed) and μ, which logically gets worse as μ is driven away from the centre. In this scenario it is straightforward to see that if the final estimate μ is indeed placed according the Gaussian distribution in question, locations around the target ought to be very likely, whereas points far away are very unlikely – and hence accorded a very high P_E when they occur–.

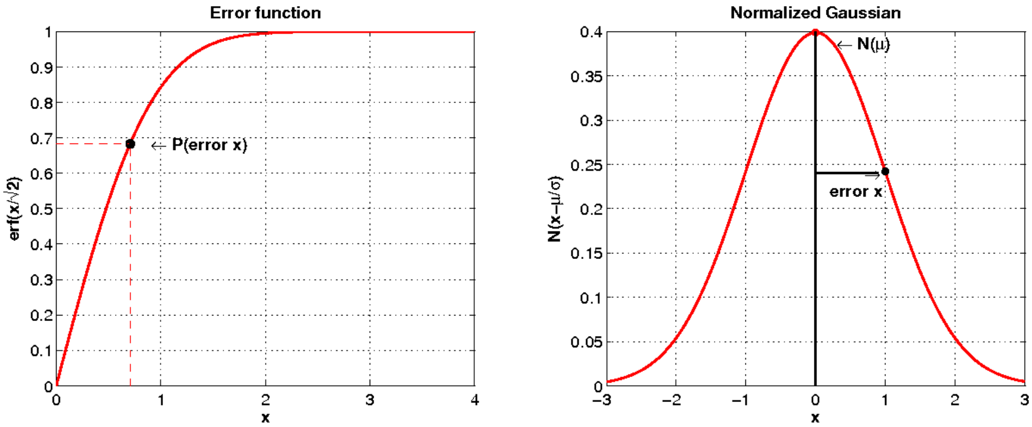


Figure 4.2: The error function makes it possible to evaluate the error of a point at $\mu + \sigma$, from a normalized Gaussian as shown in the image on the left. Given the final distribution for the optimal, this makes it possible to evaluate the error that is being made when centering the gaussian at a distance $d = \mu_{opt} - c$ from the actual centre of the screen.

Moreover, small changes in the distance between both magnitudes are more **'penalising'** when they happen close to the centre than when we are far, since the **error gets cumulated** as we get away from the target. This model thus has the interest of emphasizing the region of interest for our analysis. It provides as a whole, a comprehensive way to evaluate performance for individual trials.

4.3 Evolution of the Error over time

We now turn to the evolution over time. The previous analyses have helped address the question of whether final estimates are optimal. But how is that optimality achieved? Could it be as an alternation of unimodal responses, which are iteratively relied upon along the trial and which, as a consequence, yield a better bimodal response? We undertake this study by exploiting the Bayesian framework, which allows to evaluate single movements and thus to give insight about what compels the person at different times along a trial, i.e. what his attention is focused on.

Contribution of each modality over time

In order to evaluate the evolution over time and, more importantly, the *individual* contribution of each modality to the the movements, we evaluate for each action that the subject performs the probability given by:

$$p(\text{modality} \mid \text{traj}) \propto p(\text{traj} \mid \text{modality}) \cdot p(\text{modality}) \tag{4.2}$$

It quantifies the probability of a given modality to have been at the origin of an *observed* trajectory. Each one of the modalities may be seen as a class, $\mathscr{C}_1 = \text{vis}$ and $\mathscr{C}_2 = \text{tact}$. The previous calculation embodies the probability of a movement to belong to one of the two classes. It may be carried out through Bayes' inference rule as a combination of two terms, (1) a first term that accounts for the probability of the trajectory having being performed, should the person have relied on that modality only, and (2) a second term which considers the probability that the person actually relies on such modality.

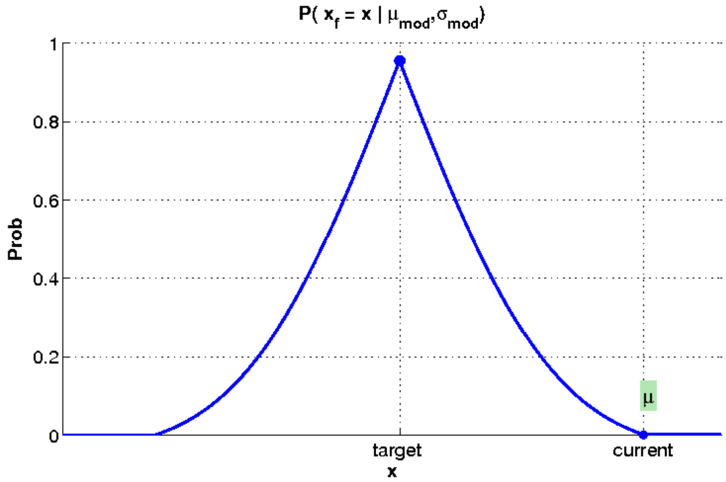


Figure 4.3: Function employed to evaluate the probability of the trajectory to be induced by a single modality $p(\text{traj}|\text{mod})$. It is evaluated as the reduction of error that the movement implies for that modality. The maximum probability corresponds to a trajectory which leads the final estimated mean to the centre of the screen. Movements which increase the error (i.e. they move the estimate away from the target) are given a zero probability.

The first term corresponds to the **likelihood** for the movement to have been produced by a given modality. Since it assumed that the beliefs are maintained throughout the action, it is only necessary to quantify the probability for the ending point to be at x_f, given the initial location of the estimate μ_{mod} and its uncertainty σ_{mod}. Therefore, the first term aforementioned may be simplified and written as:

$$p(\text{traj} \mid \text{modality}) = p(x_f|\mu_{\text{mod}}, \sigma_{\text{mod}}) \tag{4.3}$$

This can be evaluated in terms of the **change in error that is induced by the movement**. Indeed, the participant relies on his previous knowledge to guide and control his action, and attempts through his movement to reduce the error of the estimate that he considers. In other words: A movement that yields an important reduction in the error of one modality can be understood as being more probably driven by that modality.

Hence, the function used to quantify the probability in eq. 4.3 may be written as follows:

$$p(x_f = x|\mu, \sigma) \propto \begin{cases} \text{erf}(\frac{\mu - c}{\sigma}) - \text{erf}(\frac{x - c}{\sigma}) & \text{if } |x - c| < |\mu - c| \\ 0 & \text{else} \end{cases} \tag{4.4}$$

This function is illustrated in Figure 4.3. It has a maximum for $x_f = \text{target}$, and a shape that translates the cumulated-Gaussian improvement as the final point gets closer to the target –or, inversely, the penalisation as we get further. A movement that drives an estimate *away* from the target is given a zero probability, since by no means could it have been at the origin of the trajectory.

The second term corresponds to the **prior** probability for the subject to chose that modality. According to our current theory, this is directly quantified by the reliability of the modality:

$$p(\text{modality}) = w_i \tag{4.5}$$

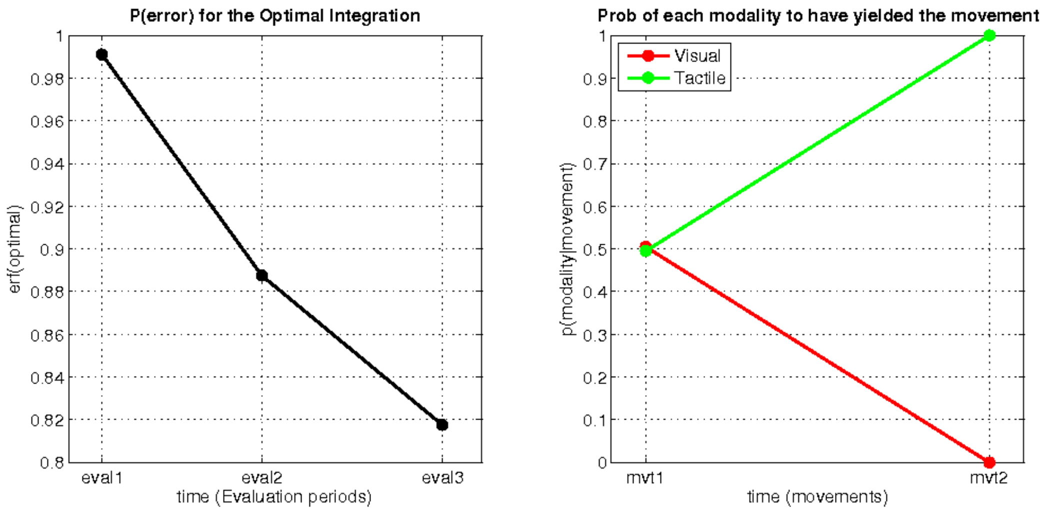


Figure 4.4: Contribution of each individual modality (right) for the trial shown in Figure 3.13. The first movement is slightly more visually driven, whereas the second movement is definitely more tactile. This may also be observed on the trajectory, through the reduction of error that each movement brings to each individual modality. Interestingly, the error of the overall 'optimal' estimate is reduced over time (left).

An example of the results obtained for the trial previously studied (shown in Figure 3.13) is illustrated in Figure 4.4. The first movement yields a similar reduction in error for both modalities which are, at this point of the trial, equally probable. Around 50% probability is hence attributed to each modality. The second movement, on the other hand, drives the visual estimate *away* from the target, whereas the tactile one is centered in the target. The first one is hence accorded a probability of 0 and the later of 1.

Nevertheless, this should not be understood as though the visual modality is not used at all during the second movement. It may indeed have been employed *combined* with the tactile information one. Let us emphasize that this evaluation concerns each *individual* modality; It intends to underline what 'preferences' seem to be accorded by participants to each modality along the trials –either more visually driven or tactile biased– so as to highlight possible patterns in their behavior and, more specifically, unimodal responses.

Let us also point out that the error of the optimally integrated estimated is also reduced along the trial, as a consequence of these movements.

5
Results

We have presented so far the materials and models at the core of our analyses. These address the question of human perception and multisensory integration, both from the perspective of within and across modalities. The experimental tasks were first described, then the individual study that is applied on each one of the resulting measurements.

The results that follow from these evaluations are detailed hereafter. Punctual interpretations or discussions are also added in this chapter, in order to point out important milestones on which our final analysis is built.

5.1 Optimal vibrotactile feedback

The results concerning the optimal encoding that is to be employed are here described for the two models considered, namely the ‘spatial’ code and the ‘intensity’ code. They are intended to help exploit the **characteristics that underlie within-modality discrimination**, thus ensuring that patients are provided with simple, direct and easy-to-interpret feedback. The procedure to measure and fit an inverted Gaussian to the statistics has been previously introduced. It is here applied to different central locations and frequencies. Initially, these are equally distributed so as to cover the whole range offered by the vibrotactile array; They are then chosen at carefully selected regions so as to validate certain values or to refine their precision.

In the case of the **location encoding**, the recordings are employed to establish the optimal number of steps that may be used given the sensitivity measured for each centre. The ones presented in Figure 5.1 were obtained from two subjects. Interestingly, they translate experimental observations about tactile sensitivity reported by the participants, namely very accurate perception at the limits of the forearm –wrist and elbow–, and less precise next to the wrist or around the middle of the arm. We then use a linear interpolation to establish the limits of sensitivity along the arm, and derive the step function within those bounding limits.

It may also be noted that, given the higher and lower limits, the number of steps are completely determined by the initial point x_i. All other quantities automatically follow –since the function is ‘almost’ always monotonic. It is thus only needed to find the initial point that maximizes the number of steps. In the case of the measurements illustrated in the aforementioned figure, 13 steps can be extracted for the location encoding, which correspond to about 3 centres for every two motors.

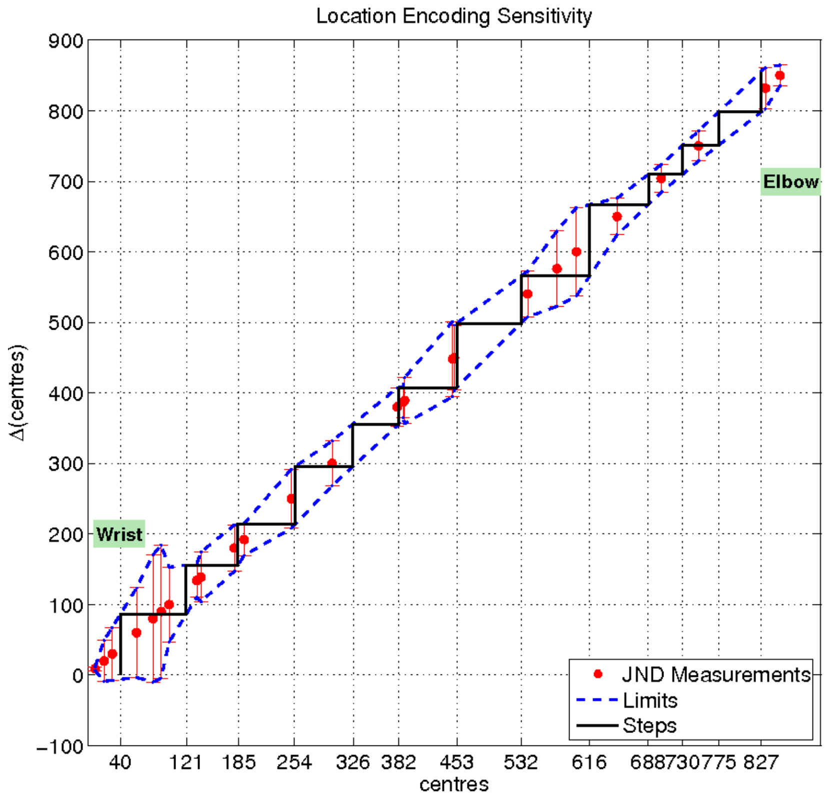


Figure 5.1: Sensitivity for the spatial code and interpolation of number of centres that may be extracted and employed to provide feedback. 13 Steps can be extracted from this encoding for a single array (8 motors).

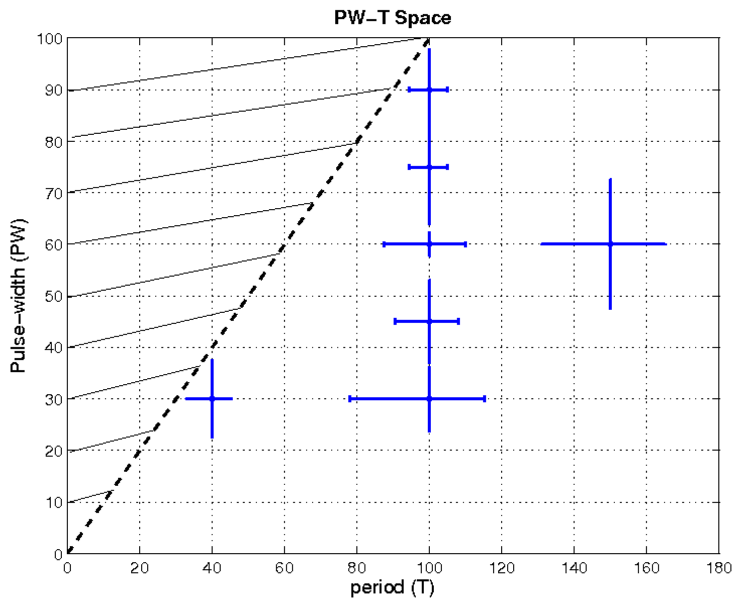


Figure 5.2: Analysis of 'PW-T' Space for the intensity code. Independent variations are applied both the the period T (horizontally) and for the pulse width PW (vertically) for different central intensities. More perceptual sensitivity is observable for high intensities –close to the diagonal dotted line that embodies the limit intensity where PW=T– and may be exploited to provide more accurate feedback.

An important consideration that derives from this analysis is that centres ought not to be equally distributed along the arm, but rather need to be adapted to the specific responsiveness of the tactile cells that are stimulated. Ideally, such a code would be established individually for each patient so as to maximize the possibilities offered by the encoding.

As for the **intensity encoding**, different centres are chosen –as determined by the ratio PW/T– and their sensitivity evaluated along both horizontal and vertical axis, i.e. as each one of the two magnitudes is independently varied. The 'PW-T' space can hence be mapped and explored in a consistent manner in order to yield optimal results. Yet because an exhaustive exploration of the whole space is complex and beyond the scope of our analysis, we focus mainly on results along a couple of 'illustrative' vertical and horizontal sample lines (PW=60 & T=100), and we interpolate from it to determine patterns from what is observed. Results are illustrated in Figure 5.2.

An interesting pattern appears to govern the evolution of the sensitivity, both as PW and T are modified, i.e. both magnitudes are better exploited by the user at high intensities, close to the diagonal line that states the highest resolution of vibration reachable. The bandwidth required for each channel becomes wider as intensity decreases, both in T or PW. It hence follows that more resolution can be directly exploited around higher intensities along both directions.

First interpretations: A first evident conclusion is that the optimal code needs not only to account for practical reasons such as portability or efficiency, but also to consider the subjects' individual sensitivity. A thoughtful organisation of each channel can improve greatly the responsiveness of patients and, as a direct consequence, their performance and naturality when using the feedback.

5.2 Statistical Evaluation of optimal integration

We now turn to the actual performance of participants during the cue integration experiment. The current focus is thus oriented **across-modalities**, in an attempt to distinguish characterictics underlying human multimodal integration in the context of an artificial feedback channel. This also represents a step further to get acquainted with how humans cope with uncertainty.

5.2.1 Statistics of Final Estimates

Our first evaluation to assess whether subjects employ both modalities in a statistical optimal way concentrates on their performance for different degrees of noise. The comparison of the performance achieved by eight participants, both with *unimodal* feedback and under *bimodal* conditions is shown in Figure 5.3. The mean and standard deviation of the 'theoretical optimal' estimates at the end of each trial are here confronted for varying degrees of jitter. For instance, the performance for $\sigma_t = \text{min}$ is analyzed as the noise affecting the other modality varies: (1) $\sigma_v = \infty$, (2) $\sigma_v = \text{min}$, (3) $\sigma_v = \text{med}$ or (4) $\sigma_v = \text{max}$.

Clear results indicate that in all cases, the introduction of a second modality yields more accurate and consistent results, i.e. less variance in the final estimates. This asserts that **both modalities are indeed combined** and that an increase in performance results from this integration.

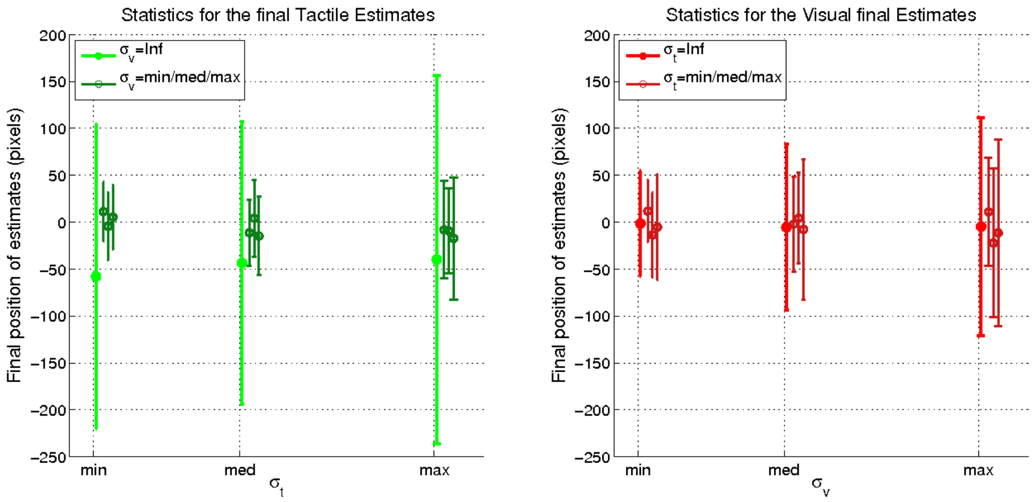


Figure 5.3: Comparison of the position of final estimates obtained for eight different subjects, both under unimodal conditions (bold lines) or with both feedback channels (thin lines) in which representations to the right correspond to increasing noisy versions of the other modality. The target is located at position = 0. In all the cases, the combination of both modalities yields better estimates, hence asserting multisensory integration.

Some enlightening patterns may already be pointed out:

- It may be noticed that **individually** the variance for the visual feedback is always smaller than that of the tactile case, indicating that naturally more accurate and consistent when using vision than with the newly added modality. This is to be expected, as a result of the differences existing between natural and artificial feedback channels.

- Likewise, the effect of **adding** visual feedback to initially tactile-driven tasks yields a much bigger increase in performance than inversely, when artificial touch is the one to complement vision. Therefore even when combined, a slight bias towards vision is to be expected in general situations.

Such issues already let us imagine that the differences underlying natural and artificial feedback may somehow *limit* the use that patients are capable of giving to the new modality. These need to be tuned so as to minimise their effect.

We may also point out that the values employed for the unimodal case are actually those measured during the *training phase* of the experiment, i.e. during which the subjects were only confronted to one of the feedback channels. An effect of learning may therefore be also affecting the previous results, in that unimodal responses are performed with less practice than bimodal ones. To minimise such effect, however, only the two final blocks -out of the three that comprise the training– are considered. This prevents from using values which during the first trials are mainly exploratory.

Optimality in single trials

We refine our previous analysis by considering now individual trials on their own. This provides deeper insight into how participants perform on a single-trial basis. Thereby we aim to get better acquainted with how integration is actually undertaken.

Let us keep in mind, however, that the evaluation of optimality for single trials is difficult and strongly dependant on the definition of optimality that is adopted. For instance, we could consider a trajectory to be optimal only if the optimally-integrated estimate is placed on the target. Likewise, it could be said that optimality is reached if the most reliable modality is the one closer to the target, meaning that the person combines both in an consistent optimal manner and relies on the most profitable one.

In our analysis, we decide to measure optimality in terms of the **probability of making an error when choosing the final location of the optimal estimate**. This considers a measure of the uncertainty around the estimates –given by the variance of the Gaussian– and uses it to quantify the performance that is to be expected under such conditions.

Such probability has been presented previously. Its main interest lies in that it accounts for the cumulative gaussian error that is committed as the estimate is placed further away from the target. It gives more emphasis to points located close to the mean, which are logically more expected to fall within that range, and **embodies the use of the knowledge** gathered by the participant along the trial

We present in Figure 5.4 the results obtained for eight different subjects. Trajectories for which the subjects clearly did not achieve the task (dist $>$ 4std, i.e. a probability of being making an error $P_E > 99.9\%$) are discarded. Two lines are also included as a reference by which to interpret the performance, one at 1 standard deviation from the target and another at 2 standard deviations. The following ideas may be commented:

- For around a third of all trials, the optimal estimate ends up located within 1std from the centre, i.e. a probability of 67% of making an error given each individual Gaussian-distributed knowledge.

- About another third of the trials falls within two standard deviations.

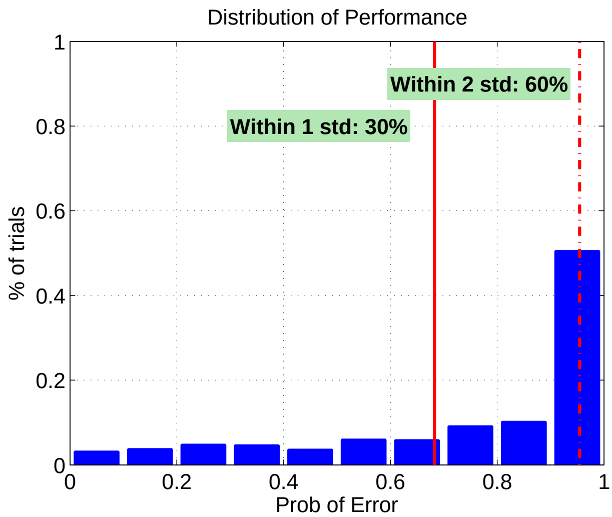


Figure 5.4: Performance achieved by a subject during the test (165 trials). The probability of error P_E is measured as $\mathrm{erf}((\hat{\mu}-c)/\sqrt{2}\sigma)$ and hence embodies a normalised measure that accounts for the distance between the estimated mean and the centre of the screen, given the variance of the Gaussian distribution. This is applied to the Optimal. The red dotted line represents points at a distance of 2σ from the target, and the red straight line a distance of σ from that same point, thereby helping interpret the relative optimality of each final estimate.

Interpretation. These results may be understood as follows: Within one standard deviation, the participants have a $P_E = 67\%$ probability of being making an error when placing their estimate at $|\mu - c|$ from the target.

This range of error could well result from reasons other than a non-optimal integration. For instance, the attention that subjects pay to the cues may not be uniform over the time of a trial. Likewise, humans have a tendency to consider errors that fall within a range as being 'acceptable', and may not try and correct the final position of their estimate even if not completely centred. **Optimal integration could thus still be understood as underlying the decisions of the participants**.

For two standard deviations, however, a probability of error of $P_E = 97\%$ leads to believe that the person is not making mistakes when *placing* his estimate, but rather that he is integrating them in a non-optimal way, giving a more predominant role to one modality than as stated by MLI.

It may thus be stated with a high degree of confidence that in two thirds of the trials, the person is **not totally integrating both modalities exactly as predicted theoretically**. For about one third of the trials, around 50% probability can be attributed to an optimal integration in which participants' perception or control lead them to make an error when placing the final cursor.

Effect of difference in jitter

A complementary analysis is that of the performance as a function of the amount of jitter. We illustrate the results in Figure 6.1, both for *relative* differences in noise (left) and for *absolute* amounts of jitter (right). These aim to point out specific patterns about how people deal with varying degrees of uncertainty. From the graph on the right, we can deduce that:

- People perform relatively better –in terms of how well the use their knowledge– when (1) there is noise little noise or, surprisingly, (2) when one modality is moderately good –and can hence be to relied upon to guide the search– and the second one, very unreliable, does not conflict with it.

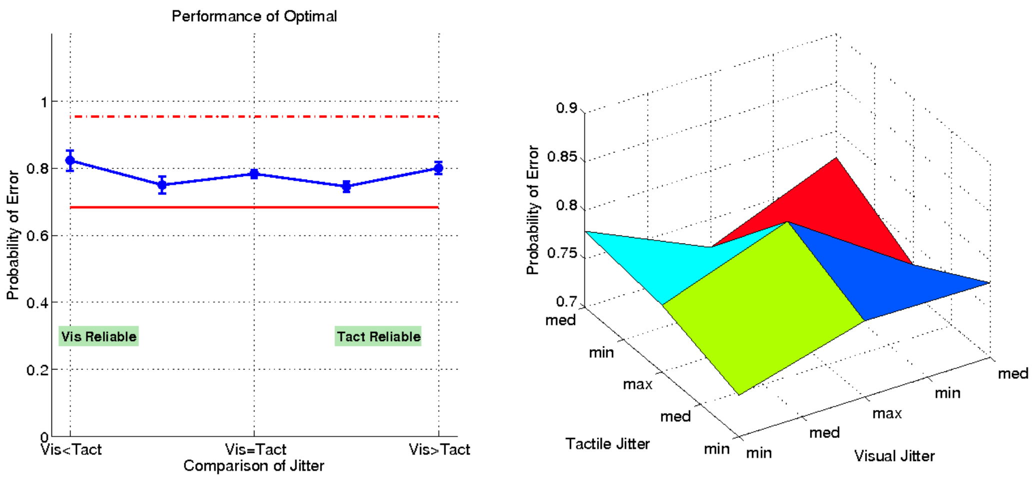


Figure 5.5: Performance achieved across 8 subjects during the test –165 trials– as a function of jitter. **Left:** Results are shown for different *combinations* of jitter, in an attempt to evaluate whether performance does not actually depend on the magnitude of the error affecting each modality, but rather on their relative differences. **Right:** Representation in terms of actual amounts of each quantity.

- Inversely, situations where one modality is very reliable but the other extremely noisy yield *relatively* worst results, i.e. participants would be expected to use much more on the good one, yet they still appear to try and extract some information from it. The second modality under such conditions is a source of *confusion*.

- It may be pointed out nevertheless that the trends –as noise increases only in one modality, and the other one is very reliable– are consistent with what would be expected.

The graph on the left also points out several of these patterns, namely better performances are achieved for **combinations of noise in which subjects can rely on one source of information without being confused in excess by the other one**.

In addition, it is worth mentioning that this second graph points out a **symmetry** in the combination of noise (left graph), which suggests that people are not excessively biased towards one modality over the other.

5.2.2 Analysis over time

We intend to get further insight into how participants tend to perform the integration of both modalities. Specifically we aim to check whether optimal integration does happen continuously, or whether it actually results from a succession of unimodal decisions. For that matter, we quantify probabilistically the **contribution of each modality** to each movement.

Because each trial is composed of different numbers of movements –of different length and at different times– the statistics are performed in our study by considering individual movements, regardless of the time at which they actually happen. Our analysis focuses indeed on *what* motivates the trajectory rather on the trajectory itself. It attempts to point out patterns that participants exhibit when performing the task.

The results for the unimodal cases are first presented in Figures 5.9 and 5.7. These help validate our method of evaluation and suggest patterns that may be noticeable from this representation. In these cases, subjects are clearly concentrated on a *single* feedback channel and rely on it to guide their search.

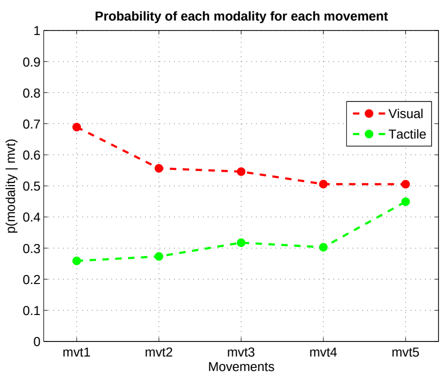


Figure 5.6: Evolution of the influence of each modality for the visual training task. A clear tendency to rely on visual cues is underlined.

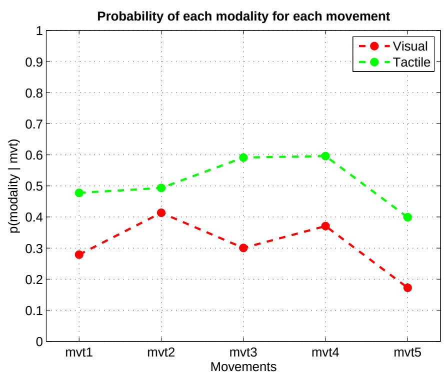


Figure 5.7: Influence of each modality during the tactile training task.

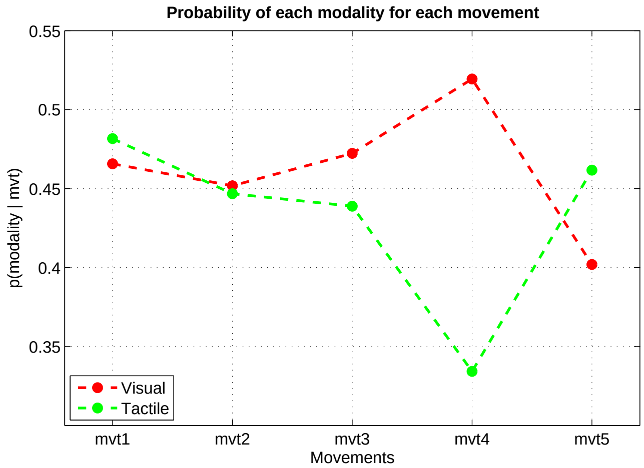


Figure 5.8: prob of each modality given trajs

Figure 5.8 then illustrates the normalized results obtained for 8 subjects during the actual test. It may pointed out from this graph that:

- Initial movements in each trial seem to be equally driven by visual and tactile information. No real tendency to rely on one or another is clearly visible.
- A better defined tendency is observed later in each trial. Indeed, around the middle of the task, people tend to use more easily visual information –being their natural modality, it might be relied upon to get a better clue of what is happening in case of confusion. That is then compensated by more tactile driven decisions by the end.

Interpretation: It may be derived from the aforementioned patterns that fast trials –which require one or two movements for the estimate to be in a location that suits to the participant– do combine both modalities. These are, as a matter of fact, most of the cases in our experiment. Interestingly, however, later changes that happen by the end of the trial appear to be much more compelled by unimodally-based decisions.

This suggests that integration does indeed happen but that it requires some **effort of concentration**. When confronted to situations of confusion or when a fast response is required, participants then have a tendency to alternate modalities.

5.3 Effect of Learning

Our final evaluation is intended to study how subjects improve over time as learning occurs. This helps complement the previous analyses, which focused on the performance achieved when subjects successful at using the new modality. Nevertheless, insightful appreciations may be derived from the analyses of their learning curve. It is also of interest to get a grip into the learning trend that is to be expected for patients that employ this feedback modality.

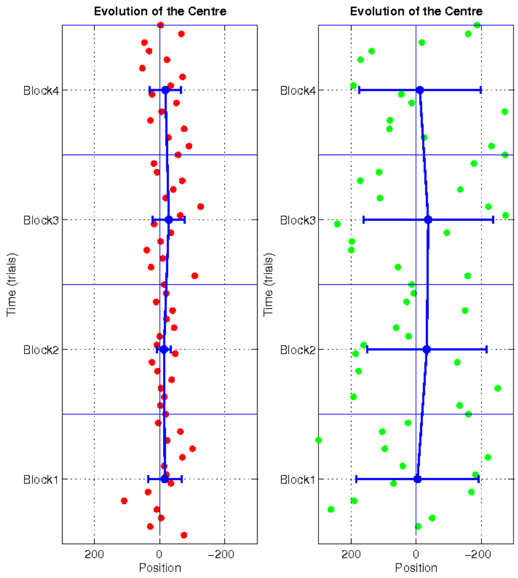


Figure 5.9: Evaluation over blocks of the performance during the visual training task.visual trainingEvol

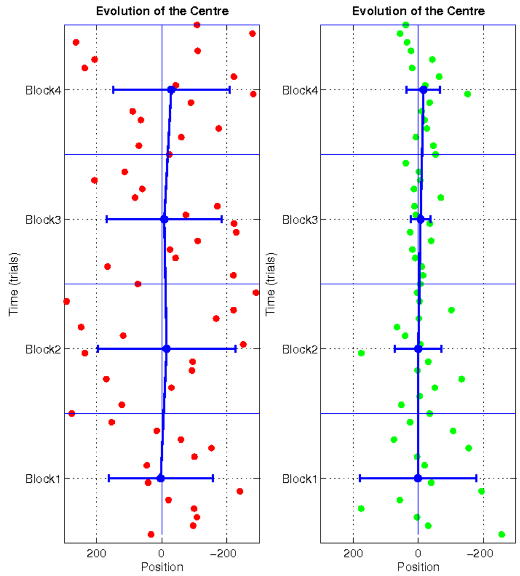


Figure 5.10: Learning is clearly visible in the trend followed by the tactil estimate.

Figures 5.9 and 5.10 show the results over blocks for both modalities during tactile & visual training. It may be noted that the improvement in tactile integration is clearly noticeable during training, as seen in Fig 5.7 (right). This is clearly translated by the reduction in the variance of the final estimates. Vision, as expected, remains constant and equally good through the training.

Let us point out as well that the second modality is clearly not relied upon –vision during the tactile training and vice-versa– for it is extremely unreliable. This asserts these results as corresponding to unimodal decisions, in which subjects concentrated on a single modality.

6
Conclusion & Discussion

This thesis has presented some issues concerning human capacity to deal with artificial stimuli. It intends to help deepen our understanding of how tactile or propriocetive feedback may be provided to patients endowed with a prosthetic limb. Currently, these people can only rely on their own vision to control the movements of their prosthesis. Under such limited conditions, simple reaching or grasping tasks turn out to be a real challenge. Additional sensory feedback is believed to increase the manoeuvrability of the robotic device, and thereby to simplify patients' everyday life.

Tactile feedback can be applied by means of different stimulators. It may also be presented through several encodings. Our study concentrates on vibrotactile stimulation–applied through vibration of motors on the skin– either as a spatial code or as an intensity code. The analysis of how subjects interpret this information and, more importantly, whether they are able to use it to perform a task, are fundamental issues that this dissertation attempts to evaluate.

These questions have been addressed from two complementary perspectives, namely **within-modality** and **across-modalities**.

- The first one focuses on the individual capabilities offered by the artificial feedback. It attempts to explore the exploitability offered by each code and to derive the optimal way of providing feedback. This is directly linked to the tactile sensitivity of patients and is undertaken under the framework of a simple perceptual task.

- The second point of view concentrates on the interest of the new modality when combined with a natural one. Multisensory integration is looked at during a tracking task, and then analysed by means of statistical models so as to get an insight into how subjects use both sources of information.

The results that follow from the previous analyses are hoped to help provide further understanding about how to optimise the responsiveness of patients when using the artificial channels. Some of the main keypoints have been aforementioned along with the results.

Optimal Integration?

Multimodal Integration appears to be successfully achieved, in that participants employ the vibrotactile stimulation provided during the experiments and combine it with vision to perform the tasks more accurately. The employability is actually shown to be easily learnt by subjects.

The question of whether this integration is actually performed in an optimal manner is to be addressed through careful consideration. The analyses undertaken for single trials emphasize that optimality can be expected to some extent in a third of the cases considered, for which the error committed by subjects remains within a reasonable range – one standard deviation. In two thirds of the cases, however, the distance from the 'theoretical' optimal estimate to the target is such that error probabilities exceeding 97% are to be expected. In such conditions, these statistics suggest that participants' decisions are very unlikely to be following the theoretical framework, and that less reliable modalities are definitely more used than expected.

Let us keep in mind, however, that our study is restricted to the values gathered during the length of the experiment, i.e. 20 mins. Interesting learning patterns have been pointed out, which make us believe that more training and longer trials could lead to more optimal results.

Issues affecting optimality...

Additional insight has been provided in this sense by the analysis of what compels participants to act *along* the trials. It was pointed out indeed that for trials that required a few movements –2 or 3 to complete the task– were mostly equally driven by both modalities. However, when bigger number of movements were required, later actions tended to be based on single modalities.

This suggests that integrating an artificial modality requires some **effort**, maybe a certain degree of concentration. Therefore, when confronted to situations of confusion or when fast responses are required, participants tend to unconsciously turn to simple straightforward unimodally-driven decisions.

Two other important issues have also been showed to affect optimality, namely the **encoding** employed and the **relative noise** affecting each modality. These may be considered when providing patients with artificial feedback, to increase the responsiveness on patients and to simplify the employability of the feedback.

Optimal encoding and exploitability are to be considered in the framework previously mentioned, in which skin sensitivity is to be regarded as an essential element that determines the levels of feedback and the bandwidth of each corresponding channel. These are locations along the arm for the 'spatial' code, or intensities of high amplitude distributed on the 'PW-T' space as previously shown.

In addition, certain **combinations of frequencies** have been shown to yield more accurate responses from participants. These are cases where the two sources of information are not conflictive, i.e. one can be relied upon and the other does not disturb. Inversely, for cases where one modality is reliable but the other is extremely noisy, worse results are observed, suggesting that subjects still attempt to use the second modality and get confused.

These are considerations that ought certainly to be kept in mind when attempting to come up with the most suitable feedback mode in the context of prosthetics, for they may contribute to simplifying the task of controlling an artificial limb and to making patients' everyday life activities more straightforward.

6.1 Future work

We describe hereafter further ideas that could help refine the models previously described.

In the same line of investigation presented for the **Bayesian Inference** model, in which we attempted to understand what compels subject to act by linking perception –cues– and movements, an interesting insight may be brought by the representation of this ation & behavior process as a HMM.

Indeed, the belief of the subjects can be seen as a latent variable that links observed movements –the outcomes of the person's belief– with perceived cues –which are represented as inputs.

A similar model was proposed by Hospedales and Vijayakumar [9] as a way to use structure inference models to understand visual and haptic integration in the context of an oddity detection task.

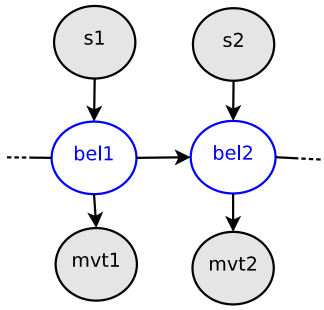


Figure 6.1: Hidden Markov Model that may be employed to understand what compels the subjects to perform the observed movements Mvt_i after perceiving the cues s_i. The link between both corresponds to the laten variable bel_i which embodies their beliefs.

When applied to the context of our experiment, such a model may deepen our understanding of the steps undertaken by participants to perform the task, and thereby to better understand how they use the information provide to react.

The implementation may be as follows: Beliefs are composed of the two weights

$$\text{bel}_i = \begin{bmatrix} w_{\text{vis i}} \\ w_{\text{tact i}} \end{bmatrix} \tag{6.1}$$

which determine the attention drawn to each modality at time i. They are refined over time both from the new information perceived s_i and from the previous belief bel_{i-1}. The probability of a movement to be generated from the current state is thus given by the observation matrix at that time.

Such a model could hence complement the evaluations presented in our methods, which were based on the reduction of error induced by each movement, and could help emphasize how a learning process may also be taking place along the trials.

Bibliography

[1] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, August 2006.

[2] C. Cipriani, F. Zaccone, S. Micera, and M. C. Carrozza. On the shared control of an EMG-controlled prosthetic hand: Analysis of user-prosthesis interaction. *IEEE Transactions on Robotics*, 24(1):170–184, 2008.

[3] S. Deneve and A. Pouget. Bayesian multisensory integration and cross-modality spatial links. *Journal of physiology Paris*, 98:249–258, 2004.

[4] H. H. Ehrsson, B. Rosen, A. Stockselius, C. Ragno, P. Kohler, and G. Lundborg. Upper-limb amputees can be induced to experience a rubber hand as their own. *Brain*, 131:3443–3452, 2008.

[5] M. O. Ernst and M. S. Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(24):429–433, January 2002.

[6] M. García-Pérez. Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties. *Vision Research*, 38(12):1861–1881, June 1998.

[7] M. Haruno and M. Kawato. Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus-action-reward association learning. *Neural Networks*, 19(8):1242–1254, 2006.

[8] H. B. Helbig and M. O. Ernst. Optimal integration of shape information from vision and touch. *Experimental Brain Research*, 179(4), June 2007.

[9] T. Hospedales and S. Vijayakumar. Multisensory oddity detection as bayesian inference. *PLoS ONE*, 4(1):e4205+, January 2009.

[10] L. Jiang, M. R. Cutkosky, J. Ruutiainen, and R. Raisamo. Using haptic feedback to improve grasp force control in multiple sclerosis patients. *IEEE Transactions on Robotics*, 25(3), June 2009.

[11] K. Kaczmarek, J. Webster, P. Bach-y Rita, and W. Tompkins. Electrotactile and vibrotactile displays for sensory substitution systems. *Biomedical Engineering, IEEE Transactions on*, 38(1):1–16, Jan. 1991.

[12] D. C. Knill and A. Pouget. The bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci*, 27(12):712–719, December 2004.

[13] A. Kuo. An optimal state estimation model of sensory integration in human postural balance. *Journal of Neural Engineering*, 2:235–249, 2005.

[14] M. A. Lebedev and M. A. Nicolelis. Brain-machine interfaces: past, present and future. *Trends in Neuroscience*, 29(9), July 2006.

[15] A. Maravita, C. Spence, and J. Driver. Multisensory integration and the body schema: Close to hand and within reach. *Current Biology*, 13:531–539, July 2003.

[16] C. Pylatiuk, S. Mournier, A. Kargov, S. Schultz, and G. Bretthauer. Progress in the development of a multifunctional hand prosthesis. In *proceedings of the 26th Annual Conference of the IEEE EMBS*, pages 4260–4263, September 2004.

[17] N. Sadato, A. Pascual-Leone, J. Grafman, V. Ibanez, M.-P. Deiber, G. Dold, and M. Hallett. Activation of the primary visual cortex by braille reading in blind subjects. *Nature*, 380:526–528, 1996.

[18] I. Saunders and S. Vijayakumar. A closed-loop prosthetic hand: The development of a novel manipulandum for understanding sensorimotor learning., Mar 2009.

[19] S. Shimojo and L. Shams. Sensory modalities are not separate modalities: plasticity and interactions. *Current Opinion in Neurobiology*, 11(4):505–509, August 2001.

[20] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, May 1998.

[21] R. J. van Beers, A. C. Sittig, and J. J. Gon. Integration of proprioceptive and visual position-information: An experimentally supported model. *J Neurophysiol*, 81(3):1355–1364, March 1999.

[22] D. M. Wolpert and Z. Ghahramani. Computational principles of movement neuroscience. *Nat Neurosci*, 3 Suppl:1212–1217, November 2000.