# Pseudo-Haptics: From the Theoretical Foundations to Practical System Design Guidelines

Andreas Pusch
IIHM – UJF-Grenoble 1 – LIG
110 av. de la Chimie – BP 53
38041 Grenoble Cedex 9, France
andreas.pusch@imag.fr

Anatole Lécuyer
BUNRAKU – INRIA Rennes
Campus Universitaire de Beaulieu
35042 Rennes Cedex, France
anatole.lecuyer@inria.fr

## ABSTRACT
Pseudo-haptics, a form of haptic illusion exploiting the brain's capabilities and limitations, has been studied for about a decade. Various interaction techniques making use of it emerged in different fields. However, important questions remain unanswered concerning the nature and the fundamentals of pseudo-haptics, the problems frequently encountered, and sophisticated means supporting the development of new systems and applications. We provide the theoretical background needed to understand the key mechanisms involved in the perception of / interaction with pseudo-haptic phenomena. We synthesise a framework resting on two theories of human perception, cognition and action: The Interacting Cognitive Subsystems model by Barnard et al. and the Bayesian multimodal cue integration framework by Ernst et al. Based on this synthesis and in order to test its utility, we discuss a recent pseudo-haptics example. Finally, we derive system design recommendations meant to facilitate the advancement in the field of pseudo-haptics for user interface researchers and practitioners.

## ACM Classification Keywords
H.1.2 [Models and Principles]: User/Machine Systems---Human Factors, Human Information Processing; H.5.2 [Information Interfaces and Presentation]: User Interfaces---Haptic I/O, Theory and Methods

## General Terms
Design, Human Factors, Theory

## Authors' Keywords
Pseudo-haptics, human perception and action, human information processing, foundations framework, system design guidelines

## 1. INTRODUCTION
Haptic feedback, no matter whether passive or active, whether tactile or force-returning, whether realistic or simple, has gained much importance for a holistic / compelling user experience in various fields of human-computer interaction (HCI). Ever new intriguing approaches occur on the scientific and commercial "markets", often stimulated by multidisciplinary efforts. Thus it is not surprising to see haptics hardware, software and multimodal interaction concepts rapidly evolving.


*ICMI'11*, November 14–18, 2011, Alicante, Spain.


About ten years ago, user interface researchers experimented with an alternative means – pseudo-haptics [1] – to reduce system and device complexity while still providing a rich haptic sensation. As it will be addressed in this paper, pseudo-haptics correpsonds to a haptic percept that is different from what the real haptic sensory supply would suggest by "playing" with the multimodal – mainly the visual – feedback of a system. In [1], an interactive virtual piston is displayed on a computer screen. By pressing a SpaceBall support device with the thumb, the user can make the piston move according to both the actual force applied and the simulated spring stiffness. Varying degrees of visual compression of the virtual piston lead to different levels of perceived stiffness.

Since then, a number of haptic properties have successfully been implemented (e.g., [2, 3, 4]) and even brought to a productive state (e.g., [5, 6]). However, the perceptual mechanisms behind pseudo-haptics remain largely unclear. And so do the reasons why there was often a considerable user population unaffected, why reported impressions were sometimes unexpected or "strange", and why there was mostly a large user-dependent variability in the induced sensations. In [7], some of these open questions have already been discussed in more depth and first guidelines for the design of pseudo-haptics systems have been proposed.

Our intention is to extent this state of research on pseudo-haptics. We want to move forward both on the theoretical foundations side as well as regarding a more complete and generic set of system design recommendations. The overall objectives are thus first, to prepare the ground for a better understanding of the phenomenon, second, to demonstrate how existing systems can be assessed and possibly improved in order to approach the quality of robust contingent effects, and third, to facilitate the development of future reliable pseudo-haptics-enabled interaction techniques.

## 2. PREVIOUS WORK
Here, we are going to review the related work with respect to the main goals of this paper. For a detailed presentation of the human perception and action models we will base our foundations synthesis on, refer to the "Theoretical Foundations" section.

### 2.1 Classification of "Pseudo-Haptics"
Systems or interaction techniques labelled "pseudo-haptics" globally fall into two distinct domains. On the one hand, there are approaches that attempt to stimulate a unified percept of a specific haptic environmental property in the sense that the user actually "feels" it (e.g., [1, 2, 4]). Perceiving the weight or stiffness of an object would be practical examples. This kind of sensation can be regarded as the result of an automatic lower level processing of multisensory information blended or merged with what is already known about the world and with the user's intentions [8, 9, 10, 11]. The user is ideally not able to escape from the induced sensation which usually differs from the "real" percept and thus qualifies pseudo-haptics as a type of haptic illusion [7]. Most of the existing techniques, including the original work of Lécuyer [1], can be attributed to this domain.

On the other hand, there is a group of approaches mostly relying on the cognitive interpretation of a visuo-haptic simulation (e.g., referred to as "optically simulated haptic feedback", [12, 13]). Visually presented haptic effects like mouse cursor drifts meant to reflect a surface slope have effectively to be "understood" by the user in order to relate these observations to the simulated phenomena. Although the meaning of the visual feedback is tightly linked to the haptic property, it is not necessarily the case that the user "feels" it in the way he would sense the weight of an object. The final percept or mental representation of the phenomenon is thus more the result of a "strategic decision-making process" [7] and requires explicit learning. However, exceptions are possible. Even if unintended, it is not completely unlikely to produce conditions under which the interaction between the contributing multimodal cues becomes qualitatively appropriate to trigger some weaker form of the above-mentioned illusion. We will discuss these conditions later on.

Please note that, in the remainder of in this paper, emphasis will be put on the illusion and not on the cognitive interpretation aspect of pseudo-haptics. We will therefore use the term pseudo-haptics only within this context – unless otherwise stated.

## 2.2 Towards Explaining the Principles

Throughout the past decade, various works in the field of pseudo-haptics have, from very different points of view, provided first insights on how and why the presented techniques may have worked. Lécuyer collected various key arguments and claimed [7]: "Pseudo-haptic feedback corresponds to the perception of a haptic property that differs from the physical environment, by combining visual and haptic information and proposing a new coherent representation of the environment." The following four "key assertions" have been extracted:

1. Involvement of one or more visuo-haptic sensory conflicts.
2. Dominance of vision over touch when perceiving spatial environmental properties.
3. Correspondence to a new and coherent visuo-haptic representation of the environment.
4. Possibility of creating haptic illusions altering the perception of a given haptic property.

Pusch stated that, in a pseudo-haptics setup like HEMP (i.e., Hand-displacEMent-based Pseudo-haptics) [4], "the visually presented phenomenon should theoretically be capable of returning some kind of force. Appropriate haptic information related to this phenomenon has then to be provided as a carrier stimulus to the user. [...] Finally, if well-defined actions can successfully be coupled with a plausible visual feedback, then the user's central nervous system (CNS) may adapt to the new, the artificial multisensory supply – and the illusion occurs." [14]

Paljic et al. argued that users were able to discriminate different levels of torque because of a "perceived mechanical work" cue that might have triggered the observed illusion effect [2].

In a study designed to identify a "boundary of illusion" for a spring stiffness setup, Lécuyer et al. found that "more visual deformation is necessary to compensate large haptic differences" [15]. The fundamentals likely to underly this observation have recently been studied in a number of psychophysical experiments and will be discussed further down. It has also been remarked that "this boundary varies greatly depending on the subjects and their strategy of sensory integration". But it remained unclear as to what might cause such a strong variability in the perception of the existing pseudo-haptic phenomena, and how to handle it.

We will discuss all these findings, assumptions and observations in the light of a comprehensive framework of human perception and action that aims at clarifying the interplay of all perceptual key processes and modalities contributing to a pseudo-haptic sensation. We further hope to answer some of the important open questions.

## 2.3 Existing Design Recommendations

Developments in this domain seem largely to be characterised by an exploratory research approach. Such explorations can take fairly long and it is often rather hard to predict what the outcome will be. Several authors provided detailed rationales (e.g., [14]), but yet without bringing them to a set of more generic design recommendations for a larger variety of systems.

The first attempt towards such specific guidelines has been made by Lécuyer who suggested to follow these steps while keeping the application context in mind [7]:

1. Identification of a law that controls a haptic property.
2. Association of this law with related spatial parameters.
3. Setting up a visuo-haptic sensory conflict focussing on one or more of these parameters.
4. Modification of the visual feedback of the respective parameter(s).

Although these steps provide a good starting point, there is still much space left for interpretations. Moreover, even if certain facets of reality are frequently "mimicked" (to varying degrees of realism), the entire aspects of how the haptic property to be simulated is normally perceived by the user and what the role of prior knowledge might be remain unclear. A discussion of potential perceptual and ergonomics side effects and limitation of playing with multisensory conflicts would be important as well, since this is the (still relatively fragile) practical means of pseudo-haptics. So, for approaching more compelling and reliable designs in the future that convey nearly robust contingent effects, we think that considering as many of these pain points as possible would be an indispensable requirement.

# 3. THEORETICAL FOUNDATIONS

There are different models of human perception, cognition and action, all representing excellent resources for the explanation of a number, if not most of the perceptual and cognitive processes likely to underly pseudo-haptics. But rather than testing all of them for their applicability to our questioning, we decided to synthesise a compact "practitioners framework" based on two successful, popular concepts that we believe are rich enough: The Interacting Cognitive Subsystems (ICS) model by Barnard et al. [8, 9] and the Bayesian multimodal cue integration (BMI) framework by Ernst et al. [16, 10, 11]. We think these two models are complementary: While ICS provides the "big picture", BMI sheds additional light on some very specific details.

Both concepts will each be briefly presented before we highlight those principles that are the most important to a profound understanding of pseudo-haptics. We will also elaborate on what else these models can teach us and where, for the targeted more complete explanation of pseudo-haptics, we need to consult additional material.

## 3.1 Interaction Cognitive Subsystems

The ICS model represents a high level "macro theory" of human perception and cognition. It covers the processing chain which transforms multimodal sensory raw data into progressively more abstract mental representations that permit to understand and to interact with the environment. It also addresses realtime information blending, storage and access. ICS may hence model various central aspects of pseudo-haptics.

**Model structure.** ICS consists of nine equally organised interacting subsystems each of which receiving one or more (data) representations generated by other subsystems as input. The incoming data is stored (or used to refine stored information) and processed (i.e., blended and transformed) according to the respective subsystem's purpose. The following subsystems (also: "Levels") were identified:

I. Sensory subsystems: Encoding of raw inputs and basic structuring => **1.** Acoustic level, **2.** Visual level, **3.** Body state level (i.e., proprioception, touch, taste, smell)

II. Perceptual subsystems: Basic interpretation and grouping of information => **4.** Morphonolexical level (i.e., abstract structures of sounds, incl. language), **5.** Object level

III. Central subsystems: High level cognitive interpretation and interrelation of information => **6.** Propositional level (i.e., factual representation), **7.** Implicational level (i.e., meanings, emotions)

IV. Effector subsystems: "Feedback channels" used for interaction with the environment => **8.** Articulatory level (i.e., speech and related gestures), **9.** Limb level (i.e., motor actions)

**Internal main processes.** All subsystems form a common network within which information is simultaneously processed. Any data arriving at a level can either in part, if certain patterns match, or completely be replaced by representations already learned and stored in memory. In case of multiple inputs, a process called information blending takes place. Blending has a direct influence on a subsystem's output, but it requires that the representations to be blended are of a consistent / coherent structure with respect to the psychological subject (i.e., object attended to), its predicate (i.e., other objects at the same level of perceptual scene decomposition) and its constituent structure (i.e., objects making up the subject). By adding reciprocal feedback loops between the morphonolexical and propositional levels, the object and propositional levels as well as between the propositional and implicational levels, the crucial role of prior knowledge during information processing is reinforced. An interactive mode of operation can be maintained, even if inputs are qualitatively or quantitatively limited.

**ICS and pseudo-haptics.** Doubtlessly, ICS can help a lot in explaining different key principles of pseudo-haptics. Suppose the user had the intention (i.e., an implicational representation of his goal) to manipulate an object. He performs all required (motor) actions and his perceptual "entry subsystems" continuously receive sensory raw data. These inputs, typically from the visual and the body state subsystems, is structured, related to each other and then interpreted or "sensed" to form a mental representation of any experienced (real / virtual / mixed) haptic environmental property. Regarding the contribution of vision, this means, transforming retinal stimuli into actual object representations. If an object is recognised, associated information can be revived and retrieved from memory (e.g., this texture felt rough, this material was solid, this is my hand). This way, it is possible to organise further interaction and to anticipate the feedback that will then be blended with and compared to the multisensory inputs. At body state level, relevant sensory raw data usually originates from touch and / or proprioception. Once this information has passed low level structuring, implicational blending can directly merge a given sensation with what has visually been identified – and what one associates with the observation. For this, the representations have to be sufficiently consistent (e.g., this feature or sensation is related to this type of object or action, these events occurred "spatio-temporally aligned"). Blending weights vary according to the context-dependent degree of perceptual reliance and faithfulness of the estimates within each of the concerned sensory modalities [10]. Pseudo-haptics, as it has been explored to date, relies on that vision is largely favoured over haptics.

Incomplete or inappropriate inputs can be the result of poor multimodal displays (e.g., only few, vague cues), ambiguities in the system response (e.g., contradictory multisensory supply) or too weak relationships between the representations to be blended (e.g., too large system lags or visuo-haptic conflicts, too less meaningful connexions between the inputs). In most of these cases, ICS' internal feedback loops can, at least to some extent, compensate for sensory or perceptual gaps. This strategy again highlights the importance of prior knowledge. However, the more gaps or uncertainties are to be handled the more likely is a "subjectively coloured" final sensation that is prone to instabilities for the user and among a group of users each having his own individual background. Practically, the user may not perceive or feel, if anything at all, what has originally been designed. Or he perceives something rather unexpected, unrelated as his brain is forced to compute a "solution" to whatever the current state of perception is. In the end, if all this becomes too confusing, the user might just refuse using the system.

To our knowledge, pseudo-haptics simulations have barely exploited the auditory modality (see, e.g., [17]), although a direct link to implicational blending would strongly suggest to do so (e.g., accompanying real or synthetic sounds). Even if not used for immediate blending, sounds would help establish or support mental representations / propositional descriptions, also of non-acoustic environmental properties. These representations would then be available for the later blending. In the same vein, one could turn the "risk" of relying on or requiring too much prior knowledge into a benefit. Why not deliberately preparing the ground for a stable blending by "early seeding" of the right information? We are going to elaborate more on this somewhat controversial concept in the "Deliberate 'User Priming'" section.

**Main "limitations".** Apart from the many good points of ICS like its holistic view, the handling of multimodal interdependencies, representation transformations and the role of memory at various stages of the model, there are a couple of things it cannot answer in full breadth. One of the most important open questions is related to low(er) level modality blending and its prerequisites. How have representations undergoing blending spatio-temporally to be configured? In fact, this is essential first, for a deeper understanding of modality blending in general and second, to deriving more sophisticated design recommendations for pseudo-haptics systems. What also needs to be addressed in more detail is the aspect of disambiguation. If our brain is confronted with ambiguous multisensory data, it attempts to resolve it by recalling related information from memory in order for the final percept to settle. How can this process be supported? Are there alternative strategies? It seems the modelling of intention, motivation and thus goal-directed actions may be made more explicit. We hope to close some of these "gaps" in the course of the next two sections.

## 3.2 Bayesian Multimodal Cue Integration

This second model of human perception which we will base our synthesis framework on is primarily dedicated to multisensory information blending. It is thus more a "micro" than a "macro theory", although it also proposes a broader perception-action loop concept incorporating a Bayesian inference approach at perceptual level and a gain / loss function approach at both decision making and action planning level.

**From mandatory Maximum-Likelihood-Integration ...** BMI evolved from the special case Maximum-Likelihood-Estimate (MLE) integration model [16] which, in a way, is still included in BMI. MLE states that redundant multisensory information (e.g., from vision and touch) about an environmental property (e.g., the size of an object) is integrated by summing up the contributing sensory estimates' weighted averages. The resulting percept is statistically optimal in that it has a smaller variance than each of the individual estimates alone. Weights represent the reliability of a given sensory supply and typically depend on modality specifics (e.g., receptor noise level, feedback stability => signal variance). However, it has been found that the mandatory fusion within this one-dimensional approach does not apply to a number of observations made in different experiments [11]. Subjects did not seem to completely fuse the between- or within-modality data as predicted by MLE. It has been argued that the brain retains access to lower level sensory information in order for it to detect possible mismatches (i.e., sensory or perceptual conflicts) and, within certain limits, to perform compensatory mappings between the senses. This flexibility of our sensorimotor / perceptual system is key to a fast adaptation to sudden changes in the environment or the current configuration of the own body.

**… to a Bayesian view on human perception and action.** In continuation of their previous work, Ernst et al. have proposed a Bayesian view on multimodal cue integration [10, 11]. They introduced Bayesian Decision Theory (BDT) in order to model the following two consecutive steps: First, at perceptual level, multisensory interaction and second, at decision / action level, the effects of intention and task goal on a generalised "perception-action loop". BMI also accounts for uncertainties about the world, the environment and the multisensory feedback. Regarding Bayes' Rule, prior knowledge (here: The "coupling prior") is represented by the prior probability distribution, the "sensory evidence" at a given moment by the likelihood function and the resulting most likely percept by the posterior (i.e., $p(W|I) \propto p(I|W) \times p(W)$ ). The main estimates (e.g., from vision and touch), if spatio-temporally corresponding (i.e., likely to originate from the same environmental property – see further down), are characterised by their specific reliabilities (i.e., inverse signal variance).

Relative reliabilities of multiple estimates define the direction of the bias induced by the coupling prior (e.g., tending towards vision or visual dominance, resp.). This prior is responsible for the integration process and thus for modality mapping, and gains influence with an increased variance in the likelihood (i.e., a more uncertain sensory supply). It further determines the "strength of coupling", for instance, by means of a (simplified!) Gaussian distribution along the identity line / concordance diagonal. Along this line, multisensory feedback agrees on the object property or sensation in question. However, the prior could also become "sharply tuned" and take the form of a delta function leading to a pure MLE integration, so, mandatory fusion. The likelihood of correspondence can be reinforced by co-occurrence of supporting multimodal attributes, similar to the formerly described "sensory combination" process [10]. Another goal of this process, beside (spatial) information alignment between non-redundant cues (i.e., determining the transformations needed), is to collect as much information as possible, no matter through which modality, in order to resolve perceptually ambiguous situations and therefore interrelate different estimates. Each multimodal attribute makes up one dimension of a "hyper-dimensional multimodal space" that can affect the coupling prior. It might hence become narrower or steeper in the same way as it would become more flat or spread for stronger discrepancies and contradictory cues (see also [18]). A completely flat prior illustrates signal independence which does not mean that, at cognitive level, a correspondence cannot be established. But then active feedback interpretation would take place – as it seems to be the case for "visually simulated haptic feedback". The coupling prior's shape may further depend on the user's perceptual preferences, capabilities and experiences.

There is not much known about when it becomes obvious to the user that there is something "odd" in the multimodal feedback (e.g., a too large visuo-haptic conflict), simply because oddity detection depends on a number of variables. In this context, Hospedales et al. [19] proposed a Bayesian ideal observer which includes a "causal structure inference" approach to model oddity detection. They found evidence that the brain might perform a probabilistic hypothesis selection and switching during multimodal cue combination, in particular under uncertain cue correspondence (in brief: Is the incoming multisensory data from the same source / environmental property or not? If so, compute a unified percept like it has been described in BMI.). What has indirectly been shown is that the task goal can influence the multisensory integration process (here: Identification of a potentially "odd probe" characterised by conflicting multisensory feedback). Ernst et al. already considered this by including a BDT-like decision making process based on goal-dependent gain / loss functions [10, 11] – although in a much less formal way. However, even if the user recognised the conflict (i.e., is aware of the oddity), an optimal solution, under the given constraints, would nevertheless be generated – sometimes resulting in objectively "illogical" yet unified percepts. Specially when relying on (weaker) illusions like pseudo-haptics to date, a poor multimodal display or simulation, and therefore a lack of belief in the observation, can quickly become an issue and strongly perturb both perception and action. After all, a pseudo-haptic property could still deliver a certain "sensation", if a minimum coupling between the contributing multimodal cues is retained. In this light, it can be expected that some of the existing mouse-style pseudo-haptics techniques may indeed induce a (weaker) "feeling" or sensation and thus not always rely on a pure cognitive interpretation of visually observed haptic effects (see also 2.1).

**BMI and pseudo-haptics.** We are convinced that BMI is highly useful to describe the heart of the low(er) level information blending underlying pseudo-haptics. Its multisensory interaction processes can further be related to the blending mechanisms at Level II of ICS (i.e., perceptual subsystems). In addition to this, BMI provides various references to ICS and either confirms or complements it this way (e.g., concerning the overall multimodal structure and processing sequence, the impact of intention on perception and action, the role of prior knowledge, the information transformation processes). What we learned in particular is first, how the human brain produces "operational percepts" from multimodal cues, second, how these cues have to be configured and third, how perceptual stability can be achieved and supported. The Bayesian approach permits to account for sensory, perceptual and environmental uncertainties as well as it permits a very flexible cue integration scheme with the coupling prior as multisensory interaction modulator. Pseudo-haptics can therefore be seen as another applied example of BMI – enriched and framed by ICS. In the next section, we will develop our (simplified) synthesised foundations view.

## 3.3 Foundations Synthesis Framework

The synthesis (see also Fig. 1) is meant to provide an "operational picture" of the underlying processes to both HCI researchers and practitioners. Its simplicity and reduction to the most important components of two fairly complex frameworks will later help us establish a set of more generic and easily applicable system design guidelines.

**Multimodal "raw data" processing and transformations.** The user experience starts with observing the current environment, including the own body, "the self". Multimodal "raw data" from vision, audition and the body state (e.g., proprioception and touch), reflecting the basic feedback from the environment, will then be processed in several steps. Each cue's input is structured and transformed into higher-level information representations (e.g., for object identification and scene understanding). During this process, missing or incomplete data can be replaced, complemented or refined by information retrieved from memory (i.e., prior knowledge or previous experiences). Learning corresponds to the information transfer back to memory.

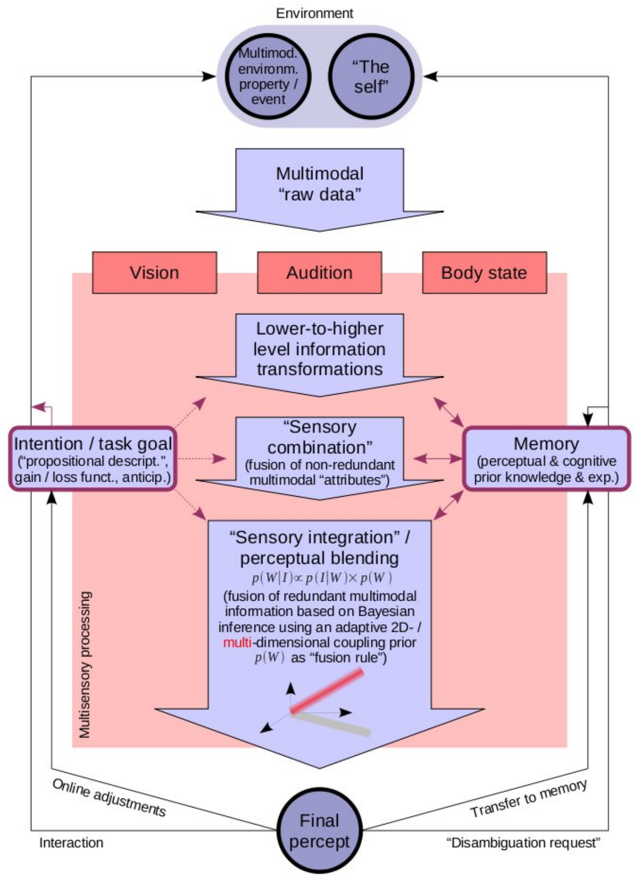


**Figure 1. Foundations synthesis framework based on ICS and BMI. In brief: Observation of the environment, lower level information structuring and interpretation, fusion of non-redundant and redundant information, production of the final percept, interaction. Embedded graph: 3D coupling prior (i.e., nb. of dimensions => contrib. single- or multimodal cues).**

**Sensory combination and integration.** At this level, a first fusion takes place, given all non-redundant multimodal "attributes" appeared at a sufficient level of temporal coherence: "Sensory combination". The brain attempts to spatially align them in order to obtain coherent coordinates for the final "sensory integration" step. This, together with the attempt to resolve potential ambiguities by requesting additional information about other attributes, serves adapting the ultimate integration rule: The "coupling prior". This prior models, in the Bayesian sense, individual knowledge and can more or less be tuned (i.e., "strength of coupling / fusion"), depending on the conditions at a moment. A completely flat prior (i.e., signal independence) means that the cues are insufficiently related to each other to result in an unified percept. A simultaneous multi-cue fusion might further be represented by a multi-dimensional coupling prior with each dimension signifying one of the contributing single- or multimodal cues.

**Intention / task goal and interaction.** The "intention / task goal" has another strong influence on multisensory processing. It directs attention and therefore emphasises certain modalities and attributes. Through the generation of motor plans needed to execute movements [20], it is indirectly also responsible for how an environmental state is evaluated with respect to previous observations and the own actions (i.e., certain sensory feedback patterns predicted by the internal model of action [21]). Resulting mismatches may lead to online adjustments of either (motor) actions or the current (sub)objective. Further, through another Bayesian component at decision making level, final action execution is optimised. Performance costs are computed based on gain / loss functions. More practically: With some goal or mental model of a scene in mind, the user consciously and subconsciously anticipates changes or things to happen in the observed environment and does then react accordingly. The more space for interpretations is left to the user due to an insufficient or bad multimodal display, the more subjectively coloured (i.e., biased expectations and assessments) will be the final percept – and actual motor performance or interaction. This is a risk and an opportunity at the same time as mentioned in earlier, since deliberate priming can indeed pave the way to "guiding experiences" (for more, see 5.2).

**Final percept.** In the context of perceptual illusions, Berthoz has suggested that the brain provides the best suitable hypothesis to a given multisensory supply [22]. This agrees to the Bayesian concept discussed so far. The brain computes the most likely percept from the following three major entries: Incoming multimodal data, the user's memory and experience, and the task to be performed. An illusory percept is thus the result of a (yet) successful multisensory integration (i.e., no confusing or "flipping" sensations) – even if there are discordant multimodal cues or biased interpretations. Pseudo-haptics can, without any restrictions, be seen in the same light. When seeing that the weights or degrees of reliability can vary a lot among different modalities or cues within one modality, it is not surprising that a number of intriguing pseudo-haptics techniques have already been found to effectively fool the user's perception. To be more explicit on this, we will review a recent contribution to the domain in the next section. This includes the discussion of the problems encountered, important open questions and improvement potentials with respect to the synthesis framework.

# 4. EXAMPLE DISCUSSION

Only few systems have been proposed to date which use neither passive nor active support devices to produce pseudo-haptic feedback. One of them, called Hand-displacEMent-based Pseudo-haptics (HEMP), makes the user feel a virtual force field [4]. Distantly related phenomena were observed in psychophysical and neuromotor experiments like the flexible / rubber arm illusion [23] or the visual touch illusion [24]. However, the sensations subjects had in the latter two cases occurred incidentally. Whereas HEMP was designed to induce an "active" force field illusion – to our knowledge, the first time pseudo-haptics became "active".

Due to HEMP's intriguing properties, we have chosen to discuss it here in detail. We will test both applicability and utility of our synthesis framework, further discuss problems and open questions raised by the original work, and highlight improvement potentials wherever appropriate. The discussion follows our model's global structure:

1. **Environment.** The user is presented with a stereoscopic Augmented Reality / Virtuality environment using a video see-through head-mounted display (HMD). The scene consists of a virtual stream tube object which is cut open on top and has some animated particles (or: "Force field" lines) rotating through it. The user's hand is embedded as co-located, spatially displaceable live video data. Varying force field levels lead to an (adaptive) lateral visual shift of the user's hand once it has been exposed to the virtual flow. The displacement increases progressively up to a force-level-dependent distance and velocity, then fades out within

another specific distance within the capturing space of the HMD's built-in stereo cameras (see also [4, 14]). Flow animation parameters do not change.

2. **Multimodal raw data processing.** After lower level visual information processing, the virtual scene, the own hand and their observable dynamics are (expected to be) recognised. Video feedback of the hand is meant to convey both a visually faithful look [25, 26, 27] and a plausible motion feedback [28]. This may contribute to a more convincing user experience in the form of a more tuned, sharper coupling prior. At body state level, two main information sources are at focus: First, proprioceptive hand position feedback and second, activity in the principal horizontal arm flexor and extensor muscles [29]. However, there are more "arm motion cues" available, for instance: Skin deformation, tendon and muscle length variations all over the articular structure involved and tactile feedback of sleeves sliding over the arm surface. Acoustic feedback has not been used.

3. **Sensory combination.** Two steps are important at visual level. First, with stereo vision and the embedded live video of the own hand, the user is likely to experience the space in front him, viewed through the HMD, as kind of an integral peripersonal environment. The small field-of-view of the HMD might impair this experience whereas the employed mixing approach might support it due to a correct occlusion handling. On the contrary, the planar hand texture carrier objects and unrealistic cutting edges on intersections with virtual object might sometimes perturb this picture. Second, the visual virtual force field creates a certain expectation with respect to the effects to occur on exposure. Despite the artificial virtual stream tube and the very simple particle flow feedback, the design of the interactive region provides some useful hints on what this simulation is likely to do. At visual-haptic level, the initial hand co-location feeds the mental representation of a coherent interactive space. To maintain or regain this sensation of coherence, visual hand offsets can be reduced in order to return to co-location. As concerns body state, the proprioceptive hand or arm motion, position and muscle contraction cues are fused – which, due to their partially contradictory nature in HEMP, can indeed "soften" the final coupling prior. What might hide this conflict to some extent is that hand shifts are performed slowly (see also [30]). Another strategy of HEMP is to trigger an automatic motor reaction at the moment the user's hand enters the virtual flow: A compensatory postural adjustment after hand or arm have visually started to drift. Either the arm flexor or extensor group will be activated and produce the main haptic carrier stimulus to be modulated by the visual feedback. No other controlled haptic or tactile cues are provided.

4. **Sensory integration.** The coupling prior will now be used to merge all spatially aligned, redundant multimodal data so as to converge to a unified and stable final percept. What is this "redundant information" in HEMP? We think, the complex psychological subject or key cue is the strength / pushing effect of the force field in the direction of the rotating particles. Its constituent structure comprises the perceived states of real and virtual hand / arm motion and position as well as the actual muscle work for movement execution. It has been argued by Pusch that, as long as the visual feedback remains plausible to the user, related motor activity might be integrated with it and therefore lead to the modified, illusory force percept [14]. However, one difficulty we see in this design is that it relies on only one single haptic carrier stimulus which, at the same time, contradicts most of the other available body state cues. When resisting with the hand against a force exerting phenomenon, one would anticipate a feeling of pressure around the contact region. Being aware of this issue, the author proposed to add disambiguation cues like physical air streams for tactile and / or auditory feedback, or simple vibro-tactile stimuli at hand level. The coupling prior could effectively become "sharper" this way. There is another problem related to the visual quality of the video hand embedding approach. The feeling of presence could be negatively affected leading to "doubts" about the simulation. Smaller weights may be assigned to the visual feedback, thus directly affecting sensory integration. More realistic hand-scene intersections as well as sophisticated visual interactions between the hand and the force field particles could perhaps limit these undesirable effects.

   To compute the final percept, the contributing cues will be merged by taking their specific weights (i.e., reliabilities and noise levels with respect to the user's sensibility and prior knowledge) into account. "Perceptual gaps" or "confusing uncertainties" can be cleared up by retrieving related information from memory or by requesting disambiguation attempts. The percept of the visually constant virtual flow is subject to an ongoing modulation by accelerated, decelerated or steady hand shifts. To leave exploitable traces in memory for evoking the illusion, the following events have been accentuated: The intense dynamics around the reaction peak and the perceived terminal visual-haptic hand offset.

5. **Interaction.** With an operational final percept computed, the user is now able to interact with the environment. According to his goals (e.g.: Force field strength comparison) and the multimodal experience, he will focus his attention, weigh up "optimal" (in the sense of BDT) actions to take, anticipate consequences of his actions, plan and perform movements and so forth. In short, the continuous perception-action loop as described by Ernst et al. [10, 11] is closed. What should be recalled is the dynamic character of the simulation in terms of highly variable hand displacements. Independent of the precautions taken to deliver a clear idea of the system's purpose, it does not always behave in the same way as the amounts of possible hand shifts vary over time. The feedback could thus diverge from the anticipated or expected one and so leave the user confused.

**Summary.** In the context of the mechanisms enabling HEMP, we could demonstrate applicability and usefulness of our synthesis framework. It seems to turn out that the original HEMP concept carries a considerable risk to confront the user with too many adverse multisensory conflicts. To avoid some of them or, at least, to minimise the chance of detection and therefore reduce the inherent fragility, one could add disambiguating cues like simple tactile stimuli or dynamic visual flow effects. Better hand-scene intersections and a strategic "user priming" (see below) would also improve sensation / illusion effectiveness and stability.

# 5. DESIGN RECOMMENDATIONS

Two aspects are important for the evolution and the diffusion of pseudo-haptics: First, to understand how and why this type of illusion works and second, to transfer this theoretical knowledge to practical design recommendations. Enabling and facilitating the development and implementation of new systems is what we focus on in this section. We start from the list of guidelines given by Lécuyer [7] and extend it as needed. Beyond that, we will introduce our concept of a deliberate "user priming".

## 5.1 Main Guidelines

Conceiving a system that offers at the same time a simple to set up, stable over a large number of users and relatively rich haptic illusion is a challenge. The more important it is to have more sophisticated means at hand for user interface researchers and practitioners to safe their development resources and, at best, to avoid pursuing "dead ends". The guidelines catalogue presented here has been worked out based on our synthesis framework and the above discussions. It is "generic" in that it is not restricted to a particular type of haptic property. We do not claim the guidelines would be applicable to every phenomenon that could possibly be simulated using pseudo-haptics. However, we are confident that all existing approaches can easily be replicated. We are also able to highlight potential design issues as shown in Section 4.

The following steps we propose to take (*italic:* Taken from [7]):

1. **Observation of the user, while he is interacting with or experiencing a real haptic property close to the one to be simulated by pseudo-haptic feedback** (i.e., analysis of the environment, the phenomenon's attributes and constituent structure, the sensations induced, the tasks and [inter-]actions performed, and the user's body parts involved).
2. ***Identification of a law that controls a haptic property*** (i.e., a law matching our prior knowledge of the physical world; remark: There might be multiple laws for different facets of the same haptic property. Laws should therefore be ordered by relevance of their contribution to the final percept.).
3. ***Association of this law with related spatial parameters*** (i.e., identification of spatial components of the haptic property; remarks: There may also be "non-spatial" parameters useful for replication and simulation, e.g., tactile or acoustic ones. Physical correctness seems not essential, but helpful, since it best matches common prior knowledge.).
4. **For each parameter, testing its suitability for simulation with respect to perceptual and technical constraints** (i.e., priority for the contribution to the multimodal experience of the haptic property, best requiring only small magnitudes of conflicts to be combined and integrated, avoiding [too] many and [too] strong conflicts; further, efficiently to implement and comfortable to use from an ergonomics point of view).
5. ***Setting up a visuo-haptic sensory conflict focussing on one or more of these parameters*** (i.e., modelling multimodal feedback and interaction; remark: Sensory and perceptual conflicts can occur at different levels of processing as well as within and between the concerned modalities, or better: Between the cues belonging to the simulated phenomenon.).
6. **For all other characteristics of the haptic property not being controlled, envisaging easy-to-add complementary stimuli** (i.e., enriching the environment in order to reduce spreading effects on the coupling prior by feeding sensory combination and disambiguation).
7. ***Modification of the visual feedback of the respective parameter(s)*** (i.e., actual runtime simulation linked into the perception-action loop; Remarks: This has to be extended to multimodal feedback. It is also at this step where conflict intensity has to be tuned to convey a stable and reliable final sensation. Usability should be optimised, too. Outcomes of step 4 may help reduce experimentation costs.).
8. **Harmonising prior knowledge by means of deliberate "user priming"** (for details, see next section) – best before the first contact of the user with the system.

## 5.2 Deliberate "User Priming"

It is well known that perception (and thus action) can be biased multifariously, for instance, by task instructions, by the way how things are told or by other "hints" – whether provided on purpose or not (e.g., [31]). This fully agrees to the models and frameworks discussed earlier in this paper. From a scientific perspective, such a "user priming" would, in most cases, be found to contaminate the results of controlled experiments. So, researchers typically try to minimise the risk of producing skewed data by avoiding the above-mentioned factors. Without any doubt, this practice is an indispensable requirement of rigorous scientific working.

However, it can sometimes lead to unnecessarily artificial or restricted scenarios which barely apply to the complex reality of HCI. On the other hand, experimental environments and protocols often implicitly provide information related to a system's or a study's goals (e.g., by scene or user interface layout, tasks design, questionnaires) and so "prime" even "naive subjects" right from the beginning. We think that, in this vein, a deliberate seeding of the right information at the right time is a promising strategy with regard to (haptic) illusions and pseudo-haptics.

That is, the user has to be introduced to the simulation, the context of the application, to how things should be done, and what kind of output / feedback can be expected – and why. Once this "story" is clearer (i.e., new prior knowledge has been established), weaker parts of the system's multisensory feedback might be compensated for by the user's "vivid imagination". Pusch et al. did this (though probably not on purpose) by asking the subjects to determine those trials of a comparison pair in which they found it harder to hold the hand at the indicated position (author's note: Within a virtual force field) [4]. The implied sensation may thus have been the resistance against something of varying intensity. Together with the visible force field cues, the CNS had enough information available to come up with a (more or less) stable flow or stream illusion and so permit subject to take a decision.

More generally speaking: A future user is made develop his own mental model, suggested by the designer, of what he will see, feel or hear and then perceive the environment accordingly. This form of "user tuning" might be of interest to many other HCI domains.

# 6. CONCLUSION AND FUTURE WORK

Our first objective was to discuss and clarify the foundations of pseudo-haptics by building on two prominent models of human perception, cognition and action: First, the Interacting Cognitive Subsystems model by Barnard et al. [8, 9] and second, the Bayesian multimodal cue integration framework by Ernst et al. [16, 10, 11]. From these two complementary theories, we have synthesised our simplified foundations framework, nevertheless meant to provide a comprehensive picture of the key processes underlying (certainly not only) pseudo-haptics. We also wanted to offer an elaborate means for the assessment of existing techniques or those to be developed. Given the validity of ICS and BMI, we expect our framework to be valid, too. Except for one extension: The multidimensional coupling prior. Though its existence seems likely (see also: Trimodal interaction [32]), multidimensionality has not been experimentally shown yet and is therefore not part of the original BMI. Future studies on this topic may help gain the required insights. Independent of this particularity, we could demonstrate applicability and utility of our foundations synthesis by discussing the HEMP force field application by Pusch et al. [4]. During this review, we could isolate a number of design issues and propose ways to overcome them. In order to test that the suggested modifications effectively improve the respective system, it would be necessary to conduct another experiment focussing on changes in the final sensation (e.g., by measuring variations in presence and appropriate biosignals, see also [29]). We assume that the analysis process can easily be applied to other examples of pseudo-haptics systems with a similar outcome.

The second objective of our work was to bring pseudo-haptics to a larger audience of scientists and practitioners by introducing generic, but still practical system design recommendations. The proposed set of guidelines has been developed as an extension of [7] and draws from the discussions and analyses in this paper. As with other guidelines assembled to support creative work, ours need to stand the test of time, too. We already have a couple of concepts on the desk which we would like to implement and study in near future. After all, we hope that we can encourage people to explore and exploit pseudo-haptics much more efficiently.

## 7. ACKNOWLEDGMENTS

Andreas Pusch has been supported by the FUI Minalogic project NOMAD.

## 8. REFERENCES

[1] Lécuyer, A., Coquillart, S., Kheddar, A., Richard, P., and Coiffet, P., "Pseudo-Haptic Feedback: Can Isometric Input Devices Simulate Force Feedback?", In Proc. of IEEE Virtual Reality, 2000.

[2] Paljic, A., Burkhardt, J.-M., and Coquillart, S., "Evaluation of Pseudo-Haptic Feedback for Simulating Torque: A Comparison Between Isometric and Elastic Input Devices", In Proc. of HAPTICS, 2004.

[3] Dominjon, L., Lécuyer, A., Burkhardt, J.-M., Richard, P., and Richir, S, "Influence of Control / Display Ratio on Perception of Mass of Manipulated Objects in Virtual Environments", In Proc. of IEEE VR, 2005.

[4] Pusch, A., Martin, O., and Coquillart, S, "HEMP – Hand-Displacement-Based Pseudo-Haptics: A Study of a Force Field Application", In Proc. of IEEE 3DUI, 2008.

[5] Crison, F., Lécuyer, A., Savary, A., Mellet-D'Huart, D., Burkhardt, J.-M., and Dautin, J, "The Use of Haptic and Pseudo-Haptic Feedback for the Technical Training of Milling", In Proc. of EuroHaptics, 2004.

[6] Bibin, L., Lécuyer, A., Delbos, A., Burkhardt, J.-M., and Bonnet, M., "SAILOR: A 3D Simulator for Loco-Regional Anaesthesia Based on Desktop Virtual Reality and Pseudo-Haptic Feedback", In Proc. of ACM VRST, 2008.

[7] Lécuyer, A., "Simulating Haptic Feedback Using Vision: A Survey of Research and Applications of Pseudo-Haptic Feedback", In Presence: Teleoperators and Virtual Environments, 18 (1), pp. 39 – 53, 2009.

[8] Barnard, P. J., "Interacting Cognitive Subsystems: A Psycholinguistic Approach to Short-Term Memory", In Progress in the psychology of language, 2, pp. 197 – 258, 1985.

[9] Barnard, P. J., and Teasdale, J. D., "Interacting Cognitive Subsystems: A Systematic Approach to Cognitive-Affective Interaction and Change", In Cognition and Emotion, 5 (1), pp. 1 – 39, 1991.

[10] Ernst, M. O., and Bülthoff, H. H., "Merging the Senses into a Robust Percept", In Trends in Cognitive Sciences, 8 (4), pp. 162 – 169, 2004.

[11] Ernst, M. O., "A Bayesian View on Multimodal Cue Integration", In Human Body Perception From The Inside Out, pp. 105 – 131, 2006.

[12] van Mensvoort, K., Hermes, D. J., and van Montfort, M., "Usability of Optically Simulated Haptic Feedback", In Internat. J. of Human-Computer Studies, 66 (6), pp. 438 - 451, 2008.

[13] Lécuyer, A., Burkhardt, J.-M., and Tan, C. H., "A Study of the Modification of the Speed and Size of the Cursor for Simulating Pseudo-Haptic Bumps and Holes", In ACM TAP, 5 (3), 2008.

[14] Pusch, A., "Visuo-Proprioceptive Conflicts of the Hand for 3D User Interaction in Augmented Reality", Ph.D. thesis, 2008.

[15] Lécuyer, A., Burkhardt, J.-M., Coquillart, S., and Coiffet P., "'Boundary of Illusion': An Experiment of Sensory Integration with a Pseudo-Haptic System", In Proc. of IEEE VR, 2001.

[16] Ernst, M. O., and Banks, M. S., "Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion", In Nature, 415, pp. 429 – 433, 2002.

[17] Rocchesso, D., "Acoustic Cues for 3-D Shape Information", In Proc. of ICAD, 2001.

[18] Gepshtein, S., Burge, J., Banks, M. S, and Ernst, M. O., "The Combination of Vision and Touch Depends on Spatial Proximity", In Journal of Vision, 5 (11), pp. 1013 – 1023, 2005.

[19] Hospedales, T., and Vijayakumar, S., "Multisensory Oddity Detection as Bayesian Inference", In PLoS ONE, 4 (1), 2009.

[20] Wise, S. P., and Shadmehr, R., "Motor Control", In Encyclopedia of the Human Brain, 3, pp. 137 – 157, 2002.

[21] Wolpert, D. M., Ghahramani, Z., and Jordan, M. I., "An Internal Model for Sensorimotor Integration", In Science, 269 (5232), pp. 1880 – 1882, 1995.

[22] Berthoz, A., "Le Sens du Mouvement", Odile Jacob (ed.), 1997.

[23] Gauthier, G. M., Vercher, J.-L., Mussa Ivaldi, F, and Marchetti, E, "Oculo-Manual Tracking of Visual Targets: Control Learning, Coordination Control and Coordination Model", In Experimental Brain Research, 73 (1), pp. 127 – 137, 1988.

[24] Biocca, F., Kim, J., and Choi, Y., "Visual Touch in Virtual Environments: An Exploratory Study of Presence, Multimodal Interfaces, and Cross-Modal Sensory Illusions", In Presence: Teleoperators and Virtual Environments, 10 (3), pp. 247 – 265, 2001.

[25] Perani, D., Fazio, F., Borghese, N. A., Tettamanti, M., Ferrari, S., Decety, J., and Gilardi, M. C., "Different Brain Correlates for Watching Real and Virtual Hand Actions", In Neuroimage, 14 (3), 2001.

[26] Blakemore, S.-J. , and Frith, C. D., "Self-Awareness and Action", In Current Opinion in Neurobiology, 13 (2), pp. 219 – 224, 2003.

[27] Lok, B., Naik, S., Whitton, M. C., and Brooks, Jr., F. P., "Effects of Handling Real Objects and Avatar Fidelity on Cognitive Task Performance in Virtual Environments", In Proc. of IEEE VR, 2003.

[28] Vercher, J.-L., "Perception and Synthesis of Biologically Plausible Motion: From Human Physiology to Virtual Reality.", In Lect. Notes in Computer Science, 3881, pp. 1 – 12, 2006.

[29] Pusch, A., Martin, O., and Coquillart, S, "HEMP – Hand-Displacement-Based Pseudo-Haptics: A Study of a Force Field Application and a Behavioural Analysis", In Internat. J. of Human-Computer Studies, 67 (3), 2009.

[30] Burns, E., Razzaque, S., Panter, A. T., Whitton, M. C., McCallus, M. R., and Brooks, Jr., F. P., "The Hand is Slower than the Eye: A Quantitative Exploration of Visual Dominance Over Proprioception", In Proc. of IEEE VR, 2005.

[31] Slater, M., "How Colorful Was Your Day? Why Questionnaires Cannot Assess Presence in Virtual Environments", In Presence: Teleoperators and Virtual Environments, 13 (4), pp. 484 – 493, 2004.

[32] Wozny, D. R., Beierholm, U. R., and Ladan Shams, "Human Trimodal Perception Follows Optimal Statistical Inference", In J. of Vision, 8 (3), pp. 1 – 11, 2008.