Human-potential Driven Design of 3D User Interfaces

Ernst Kruijff*

Institute of Visual Computing, Bonn-Rhein-Sieg University of Applied Sciences
Institute for Computer Graphics and Vision, Graz University of Technology

ABSTRACT

In particular driven by today's game console technology, the number of 3D interaction techniques that integrate multiple modalities is steadily increasing. However, many developers do not fully explore and deploy the sensorimotor possibilities of the human body, partly because of methodological and knowledge limitations. In this paper, we propose a design approach for 3D interaction techniques, which considers the full potential of the human body. We show how "human potential" can be analyzed and how such analysis can be instrumental in designing new or alternative multi-sensory and potentially full body interfaces.

Keywords: 3D user interface, interface design, human factors, full-body interface, multisensory interface.

Index Terms: H.5.1 [Information Interfaces and Presentation] Multimedia Information Systems — artificial, augmented, and virtual realities; I.3.6 [Computer Graphics] Methodology and Techniques – Interaction Techniques; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality

1 INTRODUCTION

Recent developments in game console interfaces have seen a proliferation of spatial interfaces, including a wide range of full-body interfaces deploying the Microsoft Kinect or Wii. No doubt the mainstream audience has accepted 3D User interfaces [1] by now. Interface design is often driven by exploring the various control possibilities of the human body and providing feedback that can be both informative and exciting. Yet it seems that many design decisions are based on trial and error procedures, and the rationale of success stories in multi-sensory interfaces is often kept a trade secret. Developers of Virtual and Augmented Reality applications in science and engineering, who cannot afford large-scale user group tests such as major game studios can, will be forced to make ad-hoc decisions about innovative interface design. To fill this chasm, we argue that developers would benefit from a deepened understanding of human-factor issues and human potential in 3D and possibly unconventional interfaces.

In this article we focus on a possible approach for designing 3D user interfaces from the perspective of deploying *human potential*. Human potential refers to the abilities of the human body to receive information or perform actions using all sensorimotor and non-physical human control systems. Affordances include the capabilities of receptors, the musculoskeletal system, as well as higher order processing in nerves and brain. The design is driven by what is truly *possible* with in particular the sensorimotor system. As such, we hope to broaden the scope of interface design. It should be clearly stated that the result of the design approach can but does not necessarily need to be a multisensory, full body interface: the ultimate goal is to design an interface that matches the performance criteria, which may well consist of a singular sensory or control channel.

Human potential analysis is useful for the creation of new interfaces suitable for Virtual Reality, Augmented Reality and Ubiquitous Computing. Nonetheless, many of the methods provided in this paper originate in the 2D domain, and as such can also be applied for designing 2D interfaces. Reasons for using a human-potential driven design approach are manifold, especially when considering the nature of 3D user interfaces. The design methods can aid in creating better performing interaction techniques, especially for more complex applications, and may certainly increase the attractiveness of interaction. Furthermore, reflecting human potential matches 3D user interface design very well: the human sensorimotor system is viewed without constraining physical movements ("full-body interaction" [2]) or content representation, as is the case in 2D interface design.

Following, we provide a basis that aids developers to embrace the human potential approach. At the same time we provide insights that help to understand the mechanisms behind using human potential to create useful and well-performing techniques. In particular, we use the principle of sensory and control substitution to provide a powerful basis to analyze potential sensory or control channels matching the task space at hand. Hereby it is important to realize that substitution is often used to find an *alternative* to an existing system. Three different kinds of substitution can be identified. *Substitution* refers to the process where one sensory or control channel is functionally replaced by another channel. *Addition* occurs when a sensory or control channel is added to the task performance loop, in which the channels are not directly coupled. Addition is a general phenomenon in multimodal interfaces. Finally, *integration* can be observed when a sensory or control channel is added to the task performance loop, but now the channels are directly coupled and affect each other.

Performance is often a key reason to investigate alternatives to an existing system solution. To create a usable interactive system, performance factors are explained in reflection to the so-called *energetic model*. This model regards different levels in human information processing and its' associated effort. Effort is a key aspect when designing more complex, human-potential driven interfaces. Whereas some interface are specifically tailored for exertion [3], deploying multiple senses and control possibilities can also lead to an interface that causes unwanted high level of cognitive load or ergonomic strain.

To summarize, in this article we introduce a human-potential driven approach that can guide user interface designers in creating 3D interaction techniques and devices. It is important to note that the approach predominantly defines the sensorimotor base characteristics on which actual techniques or devices can be developed. The explorative design is based on the analysis of alternatives: we assume that a base system (design) is available to which alternative techniques can be compared to. Design driven by human potential differs considerably from general design

---

* ernst.kruijff@h-brs.de

approaches for virtual environments. Usually, the choice of sensory and control channels is ad hoc and often driven by available hardware, even when perceptual, motor and cognitive factors are considered. Furthermore, comparisons between techniques are predominantly performed based on general performance criteria (such as speed or accuracy) only. We argue that human-potential driven design is a novel approach for designing 3D user interfaces. It can both instigate new ideas for designing interfaces as well as provide a more in-depth look at the actual processes that occur at a perceptual or control level. To guide the developer in this process, we illuminate major issues that need to be regarded in the design and validation phases.

2 RELATED WORK

The approach and underlying principles provided in this article are an attempt to structure information from scattered sources, providing new design perspectives and ideas. As such, related work can be found at the various stages in this article.

From a higher-level, methodological stance, work has been published that aids developers to design useful and useable 3D interaction techniques. Major sources include [4][5][6], and early test-beds such as [7]. However, even when the cognitive, perceptual and motor demands of a certain task are modeled, human potential itself is mostly disregarded. Even the related field of multisensory interface design is not well explored: few, and not widely adopted exceptions that focus on multisensory interface design exist, including [8]. Design of multisensory interfaces is also a topic in other domains, including car design [9].

Designing alternative interfaces using sensory and control substitution is a well-known technique [10][11], which has been proven useful in many applications already. These methods originate in the design of *assistive technology,* being compensatory aids for people with sensory loss. Examples include the usage of speech recognition [12], gestures [13] and eye-tracking [14][15][16]. An overview of many of these techniques was presented in [17], which was later on restructured in [18]. Nevertheless, sensory and control substitution perspectives have hardly been used for designing alternative techniques for existing systems. Most alternative techniques are rather based the general experimental nature of interface design, or even an ad-hoc decision. Whereas there is nothing against following such a design approach, human potential driven design can aid in better understanding the needs and effects of various techniques.

In direct relation, the article often refers to work performed in the field of multi-sensory processing [19][20], to explain underlying human factors principles. Also, some parts in this article overlap with general thoughts on what researchers have called a *full-body interface* [2] or multi-sensory system platforms [21]. The majority of full-body and multisensory interfaces are activity or experience-driven, and do mostly not take into account the full human potential. In general, the design and effects of multisensory interfaces are not yet fully understood, partly caused by scattered research as well as limitations in coverage of issues in user studies.

3 DESIGN FUNDAMENTALS

The human-potential driven design approach introduced in this paper follows a traditional iterative design process. After a thorough user and task analysis is performed, iterative loops of design and validation will occur till the design process is finished. We introduce the steps to structure the various stages of the design process and their related issues and effects, highlighting some noteworthy tasks: the underlying methodology is a traditional, well known approach.

Step 1: Perform user and task analysis. Before starting to design a 3D interaction technique that deploys alternative sensory or control methods, it is important to decide if using the human potential approach is suitable. For this purpose, a thorough *user and task analysis* needs to be performed. It should recognize psychophysical boundaries and requirements of tasks, users, and the actual usage environment, reflecting human factors good practice [22]. As a result, a requirement catalogue is defined. Within our design approach, we assume that this catalogue is used to create a typical or existing set of techniques used to match the user, task and environment at hand. This set of techniques is required as the baseline to which alternative techniques are matched and compared through human-potential driven design. Human potential driven design can also be performed without a base set of techniques, however, an energetic analysis will be more difficult to perform.

Step 2: Analyze and design alternative techniques. Based on the requirement catalogue and baseline system, a new or alternative interaction technique can be designed, deploying sensory and control substitution design methods. In principle, the interface designer should consider the *full potential* of the human sensorimotor system. However, applying alternative methods can only be justified when the type of task or behavior, its function, and the sequence of activity might fit the application, the psychophysical user characteristics, and the usage environment. For many task situations, the usage of general hand output and visual feedback works perfectly well. On the other hand, alternative techniques may work well for tasks that exhibit high cognitive load, in situations in which human control channels are limited or even completely blocked, or simply to create excitement. Section 4 takes a close look at the design principles behind sensory and control substitution, illuminating the sensorimotor basis used for designing a technique or device.

Step 3: Evaluate and reflect performance factors. Alternative techniques *can* make interaction particularly hard. The technique might require specific user skills or particular hardware that cannot always be applied. Furthermore, the nature of the task might be unnecessarily hardened, by increasing the amount of actions or its complexity. This can eventually lead to problems like increased error, or even hazardous usage. Hence, in the end, performance counts. Depending on the goal of the application, performance can be defined by speed and accuracy up to the level of safety or fun and needs to be matched well to the used technique [23]. Consequently, limits should always be regarded, including the level to which the human body can be “amplified”.

Evaluation methods may vary. For some tasks, a simple performance test and task load questionnaire (such as NASA TLX) will do. For other tasks it may worth to analyze cognitive processes at a neurological level, for example by using an EEG. A detailed overview of evaluation methods is not covered in this article due to its inherent complexity. It is recommended to read sources such as [1][4][22][24] for further reference.

4 ENERGETIC MODEL

The energetic model plays a key role in the human-potential driven design approach introduced in this paper. To decide which kind of sensorimotor channel substitution can be performed for the task at hand, a closer look at the human potential itself should be taken. Following the model from Gopher and Sander [25], task variables, processing stages, energetic mechanisms (mechanisms that foremost focus on the effort to plan and perform a task), and cognitive resources can be identified (Figure 1). The model provides a clear separation between processing and energetic mechanisms, combining information processing (processing stages) and performance (energetic) analysis. It shows performance efficiency at the level of energy that can be allocated for a task. The top stage, evaluation, provides the meta-level mechanisms grounded in cognitive processing. Cognitive processing, however, is only briefly handled in this article.

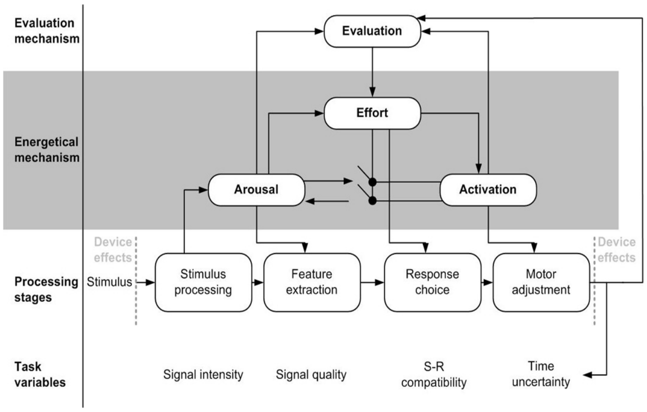

Figure 1: Cognitive energetic linear stage model. Adapted and extended from Gopher and Sanders [25]

The model provides a way of interpreting and weighting possible effects of sensorimotor substitution (section 5). These stages are important when developing human potential-driven techniques: different combinations of sensory stimuli (stimulus processing, arousal) and output (activation, motor output) can be compared to find an appropriate technique for the user, environment and task at hand. As a result, the energetic model should be coupled to the capabilities and, therefore, also limitations, of the human sensorimotor potential: it can be used to address the adequacy of an interaction technique from an economic stance (safe effort), driven by performance (including speed, accuracy), or as a balance between both. The model can be influenced by motivational (incentives) and emotional (frustration) factors, which are not directly reflected in this article.

Within the human-potential driven design process, the model can be used in all design stages, most often on a comparative base:

- To assess the effectiveness of a current set of techniques, identifying potential issues (step 1);
- To analyse the match between potentially different sensory techniques and the feature extraction thereof, to match the task and environment at hand, the compatibility of sensory and control channels, and the potential effects of devices (step 2);
- To validate the effectiveness of the alternative technique (step 3).

5 ANALYSIS AND DESIGN

Sensory and control substitution is a powerful method to make use of sensorimotor potential, by looking closely into perceptual and control characteristics. Within our approach, the energetic model is used to identify possible problems with specific input or output channels (the baseline system / original setup), after which issues can potentially be addressed alternating control or sensory channels. Vice versa, when substitution is applied, the model is used for evaluation by comparing the new methods to the baseline system.

5.1 Sensory substitution

Analyzing sensory potential. While analyzing possibilities of sensory substitution, designers will need to look closely at potential of stimuli. Which sensory stimulus can replace or enhance the currently used sensory channel? Theoretically, every body receptor can be triggered by a signal that, in reflection to the energetic model, is defined by both the intensity and the quality of information it can provide. Here lies one of the first keys of human-potential driven design of interaction techniques: which stimulus can deliver which kind of information in which quality, and can a (currently available) device actually match the communication of the stimuli? Or, from which receptor can features be extracted to deduce information needed to perform a task in a specific environment, and in which intensity does it need to be provided? To analyze the characteristics of various stimuli and receptors, it is recommended to take a close look at the detailed characteristics of sensation and perception - a good source is [26]. In any case, spatial and temporal dependencies between stimuli should be checked, as well as the sensory bandwidth of the user, to avoid overload.

5.2 Control substitution

Analyzing control potential. Approach-wise, control substitution has many similarities to sensory substitution: which control channel can potentially be used for the task/design space at hand, to replace or extend the currently used control channel? To define the possibilities and limitations of control substitution, three factors need to be analyzed: (a) the control task and its syntax, (b) the capabilities of the user, and (c) the control-body linkage (discussed in section 6.2).

To substitute a control, the *task syntax* performed with a certain human output channel needs to be mapped to another channel. The control task can be characterized by its accuracy, speed and frequency, degrees of freedom, direction and duration [27]. Task characteristics directly affect the human output channel: Some tasks may be performed with only a certain body part, whereas with other tasks this may be possible via multiple body parts. It is of importance, though, that control actions can also be mapped on other human systems like the brain or nerves. These systems are known as bio-control or biofeedback systems [28], but will not be further detailed in this article.

The *capabilities* of the user are defined by both the anatomy of a user, and practice or training. As is well known, different body parts can perform different kinds of movement, affected by the user's pose, thereby posing specific ergonomic constraints. A detailed discussion on the biomechanics of the human body falls outside the scope of this article: it is recommended to refer to an external source, such as [22].

6 EVALUATION AND REFLECTION

Once a sensory or control channel is considered for substitution, the new method should be thoroughly reflected and, depending on the design stage, evaluated.

6.1 Sensory stimuli

When dealing with sensory substitution, one should note that when a sensory channel is exchanged with another channel, one is not simply making a change at the receptor level [10]. The whole information-processing loop is re-ordered: the central nervous system needs to learn a new mode of perception. The brain has a high level of plasticity to accomplish this [29], but it is useful to note some effects. A clear example is substituting visual with auditory information for a blind person. The blind person needs to learn to "see by hearing," and thus needs to create a new cognitive model of the world. Hence, when exchanging sensory channels, one needs to deal with the metaphor of communication and the influences the new sensory channel has on the interpretation of the information itself. This matching process can be particularly hard. When changing the information process, user interface designers should thus take an increased learning curve into account: users will need time to adapt to the new kind of information processing.

6.1.1 Multi-sensory processing

When sensory channels are substituted or combined, some implications need to be considered. It is important to notice that in research the focus is moving away from traditional multimodal techniques in the direction of multi-sensory interfaces that differ at the level of human information processing. The sensory channels are no longer seen as separate channels but may affect each other [20]. Multi-sensory processing, in which sensory modalities can affect each other, is proven to be valid and occurring more often than is regularly believed.

The multi-sensory processing theory builds upon the integration of sensory signals inside of so-called multimodal association areas within the brain [19]. The research on multi-sensory factors still needs to advance in order to fully understand its importance, but some effects can already be identified, called cross-modal effects:

- Bias: stimuli from two or more sensory systems can differ and affect each other leading to modified or even incorrect perception
- Enrichment: a stimulus from one sensory system can enhance the perception of another sensory system
- Transfer: stimulation of one sensory system may trigger perception in another system

6.1.2 Stimuli substitution comparison

As explained in the introduction, substitution may occur at different levels. The levels may lead to different effects and thus need to be handled separately [20].

Within and between sensory substitution. When applying sensory substitution, both within and between sensory-system substitutions can occur. By making use of another combination of receptors, one can substitute information within a specific sensory channel. The best-known example of within-system sensory substitution is substitution within the somatic and kinesthetic system. These systems contain multiple receptors that work together at a cognitive level, such as can be seen in systems using vibro-taction[30][31][32]. A between sensory system substitution for haptic information is well known to most of us: haptic information is mostly communicated via visual or auditory channels, such as happening when pushing a button (the button lowers and one can hear a click).

Potential integration of stimuli. On the other hand, when dealing with sensory substitution, it regularly occurs that one is not replacing but adding a sensory channel to convey information to the user. When adding a sensory modality, it may occur that one is actually integrating modalities. Multiple tests have shown considerable effects. In these tests, vision mostly plays a important role, since it predominantly alters other modalities [33]. Sound, on the other hand alters the temporal aspects of vision [29] but also other aspects of vision like those that affect disambiguation [34]. Finally, tactility may alter vision [24][30], but may also altered itself by audio [31][32].

Sensory enrichment. When looking at sensory enrichment by adding or integrating modalities, two directions can be observed: enriching information through disambiguation [20] and biasing information. By adding a second or third sensory modality, the "correctness" of information can be increased. Especially in more complex applications, interaction performance can potentially be improved, since the user's ability to understand the data can be enhanced. These effects are believed to occur by either spatio-temporal concurrences (multi-sensory binding theory [38]), or modality appropriateness for specific tasks [37].

6.1.3 Effort

The effort needed to perceive the stimulus and subsequently trigger an appropriate output action should be investigated, to reflect its' suitability. Both the informational quality and the energetic effectiveness can be deduced by comparing different sensory systems (or indirectly, output devices) that provide a similar amount and quality of information through substitution methods. For example, consider the case of communicating the departure time and platform of a train: this may occur both visually and auditory, but it greatly depends on the information itself, the user, the device and the environment how effective this communication is. Sensory blocking or impairment plays an important factor for coming to the correct conclusion. The user might be old and have bad hearing, the environment might be very noisy so that the information cannot be understood clearly, the loudspeakers can be bad, or the information simply too long or complex to follow that auditory feedback would not be possible. When a sensory channel is useless in a specific task-user-environment setting, an effort comparison becomes obsolete.

A further issue is the maximum level of the stimulus intensity. The intensity depends again on user, task and environment. When the intensity of the stimulus cannot be matched by the user's capabilities to extract the right amount of information, it may be unusable for the task at hand. Thus, in order to create output to a user, the information quantity and quality needs to be matched by the perceptual system, avoiding possible overload at the cognitive (evaluation) level.

6.2 Control mechanisms

When analyzing human output, similar factors need to be dealt within the perceptual side of the model. To create a suitable response and motor action, a user needs to spend a specific amount of effort. The created motor action is directly affected by the perceived stimulus and creates a closed action-feedback loop that needs to match the task at hand (also see section 7). Different output methods, in their dependency to a coupled stimulus, can be compared to derive a performance-oriented model of task performance.

6.2.1 Control substitution comparison

In more complex applications, users regularly combine multiple human output modalities, either serially or in a parallel way. A straightforward example is the usage of mouse and keyboard (serial) and the usage of a foot-controlled button and mouse (parallel). Parallel integration should be understood as two actions having a close to or identical timeframe: the foot-controlled button can be pushed, during mouse interaction. It can also be used in close relationship in compound tasks, in which a strong relationship between both device actions exists. In this case, two devices are used to reach the same goal or perform a single command. The border between control addition and integration, which shows great resemblance to serial and parallel usage, may be difficult to separate – many actions can have a highly compound characteristic [18].

6.2.2 Effort

Observing the action-feedback loop from an energetic point of view provides detailed clues on speed and accuracy and their related cognitive and motor load. These models can be coupled to performance studies such as those applying Fitts' Law [39]. A second issue that comes into play is the motor system-task compatibility, implying control structure and ergonomic changes when exchanging motor systems to perform a specific action.

A close look needs to be taken at any effects on posture when exchanging the motor system. When the task is performed using a different body part, the changed *control-body linkage* (labeled "device effects" in Figure 1) changes. The control-body linkage can be based on physical contact or by ways of monitoring and as such also define how much effort needs to be applied. Looking at device effects in the energetic model, different body parts will have different kinds of control-body linkages [27].

Finally, as explored by Penfield and Rasmussen, the sensory-motor distribution of the cortex is of importance for the performance, in particular the granularity, of the different body parts [40]. Fine-grain interactions are possible with some body parts, whereas other motor channels only afford rough interaction.

7 ENERGETIC MODEL EFFECTS

7.1 Sensorimotor interdependency

Interdependencies between sensory and motor systems (hence, stimulus and response) exist during substitution and as such should be noted carefully: If stimuli and response are not (fully) compatible, effort may increase.

Changing a control will regularly result in a change of feedback: for example, with hand-foot control substitution, visual feedback will largely lack with foot-based control. The user-maintained feedback will change when a different extremity is being used, quite simply because the kinesthetic feedback will change. Furthermore, by exchanging modalities, different kinds of feedback might be needed to communicate the mode of action. Using a hand-based interface, showing a simple pointer might be useful to inform a user where she is pointing, whereas during full-body interface this is hardly possible.

The interdependency can also work vice-versa: the substitution of a sensory channel can result in the needed change of a control channel. This will, for example, occur when visual output is exchanged with a non-visual output method. Many human output methods are based upon direct manipulation metaphors, where there is a close match between visual input and human output – when the visual input cannot be used, these methods cannot be used or need to be used in a different way.

7.2 Cognitive effects

The energetic model provides a great aid in analyzing effectiveness of combining techniques by ways of addition or integration. This may predominantly occur at the level of analyzing the perceptual or motor capabilities mentioned before, but there are several issues we like to note that play a role at the cognitive-oriented level.

Decoupling. One issue is decoupling, in which an additional input channel is used that differs from the main interaction channel. Sharing capacities between different modalities may increase performance, though in some cases it also leads to a decrease. One example is the usage of speech, which can be used as an additional input in multi-sensory interfaces, like the well-known multimodal interfaces that combine speech and gestures. Shneiderman [41] noticed a clear problem with the usage of speech, especially for more complex actions. Speaking and listening make use of the same mental resources as problem solving, consuming precious cognitive resources. As a result, multi-sensory interaction does not necessarily lead to a decrease of cognitive load, as claimed in [42].

Errors. The combination of multiple sensory or motor systems can lead to error reduction and correction, especially in environments that are troubled by noise [43]. Users may retrieve multiple sources of information that can lead to the correct perception of the world (disambiguation).

Behavior. The perceptual structure of the task at hand may support flexible and complementary behavior, by letting the user perform the same task, via different output modalities [44][45].

7.3 Flow of action

Once a new technique has been created, a further step should be taken to define how the technique should be used within the application. Hereby, identifying the flow of action is of utmost importance. Flow of action is a key issue in more complex 3D interaction environments, especially those that mix multiple devices for I/O purposes [46]. It foremost refers to the structure of a user's output to a system, but the whole action-feedback loop is affected and thus is grounded in the information-processing loop of a user.

To analyze flow of action, the energetic model can also be applied. For example, a disturbance in the flow of action is clear when too much effort needs to be made in order to connect "arousal" to "activation", up to a level where there is a performance break. Simply said, this happens when stimuli cannot be processed anymore to select the appropriate response.

The key issue in flow of action is the composite nature of tasks. Basically all tasks performed in a 3D interaction environment are built up of subtasks that are held together via a compound structure. The compound structure is the basis of the problem solving activity of a user and directly affects operational effectiveness. The performance-related factors include attention issues, such as ease of use and cognitive load ("effort" in the cognitive energetic linear stage model). One approach that has found applicability in 2D interfaces is Buxton's chunking and phrasing theory [47]. In this theory, the compound structure is observed pragmatically as a dialogue consisting of small chunks that make up a phrase through human-computer interaction.

Especially when multi-sensory interaction comes into play, flow of action becomes complex. Due to the additional sensory and/or motor channels, the structure becomes multi-layered. Not only does the user need to change between different subtasks, but also between different input and output modalities. To sustain continuity in the flow of action, several issues need to be regarded in order to avoid mode errors, increasing ease of use and performance. The most important factors are as follows.

Switching modalities. When the user is allowed to use multiple modalities to issue a command, dynamical allocation of functions needs to be carefully investigated. The dependencies between tasks need to be analyzed in order to guarantee that related subtasks can be performed using the same modality or by a well-performing integration of several modalities. Hereby, small so called repeating interaction loops [46] need to be handled: often, actions like maneuvering are performed in between larger subtasks. These small loops need to be supported in such a way that they do not disturb the performance of the larger subtasks. The trade-off of using a specific modality of device to perform such a smaller interaction needs to be regarded. Sometimes, the used modality for the main subtasks may not support the small loop in the best possible way, though changing between modalities or devices may pose a much larger problem on the total performance of the compound tasks. Hence, switching should be avoided when possible. Cross-modal task performance examples include the usage of combined gesture and speech actions and hand-foot control substitution (addition), as explained in the next section.

Feedback. When dealing with multi-sensory interfaces, one should always be sure that the user is able to register the feedback in a clear way, especially when the feedback becomes cross-modal. When multiple human sensorimotor channels are used, it is best to make use of at least one unique sensory channel to communicate a basic amount of feedback to the user, to avoid mode errors. When feedback is scattered over multiple sensory channels in an incongruent way, the user will most likely get confused, unless the sensory information is coupled. Adding another sensory channel can enrich feedback, but replacement during the flow of action in a single compound task is not recommended. Under all circumstances, feedback needs to be compatible between sensory modalities, to communicate the same symbolic information throughout all modalities to the user.

Overload. Sensory or motor overload may occur when a user is confronted with a large amount of information. For example, in a multi-display environment, a user may not able to obtain an overview of all information at once and needs to turn her head continuously to get a clear picture. With complex data sets, this may put a high cognitive load on the user. When additional cognitive resources need to be used for further sensory or motor actions, it can cause considerable disturbances in the flow of action of an application. Hence, focus of attention in both the sensory and motor channels needs to be handled carefully.

8 EXAMPLE

To reflect the previous sections, we will illuminate various aspects of the approach and issues that may be met along, by using the challenging example of a potential **augmented reality firefighter simulator.** Let us assume that the simulator has to simulate an emergency situation through AR visuals and environment-based feedback (real smoke) in a real building. It should provide the user with spatial wayfinding and emergency cues, as well as allow the modification of visual settings. As such, the system may have some similar characteristics as a real-life system that could be used in emergency situations.

User/task analysis. Analysis will likely reveal that both input and output can be hard to accomplish. A baseline system could be an AR system deploying a handheld input device with some buttons, however, as we will see, there are limitations with this approach. Users will be wearing gloves, and may be using both hands to operate other tools necessary for the firefighting mission; hence, users should be able to perform unencumbered interaction with the external environment or tools. Hands may even be blocked completely as input channel. In addition, input should not be affected by the environmental conditions, such as noise or temperature. Information output may need to change over time: under stress, a firefighter is likely to focus on environment-based visual information (*where can I quickly move safely*) instead of looking at AR based cues. Cues may even further stress a user due to inherent visual mismatches the user would need to concentrate on (see [48]). Hence, whereas AR-based cues could be very useful in for example a path-planning phase, it may well be the user is not focusing on cues during actual movement through the hazardous part of an environment, even though cues might be life-saving. As a result, we can identify two major requirements: (a) allow for system input that does not encumber the user and allow for alternative non-hand input, under the condition that input is not effected by environmental conditions/noise and (b) monitor if the user's (lack of) attention and/or stress requires a change in the provision of information, potentially switching sensory output to a non-visual channel.

Analyzing alternatives. A straightforward approach for input using the worn gloves may be achieved by applying conductive cloth to actual firefighter gloves, similar to pinch-gloves, or embed an additional button in the tools being used. However, material may not stand hazardous usage and inputs may be unintentionally triggered. Also, when the hands are blocked as input, a second back up method needs to be provided. This will be rather challenging: a potential path that could be explored is the usage of muscular-based input by squeezing specific muscles (squeeze arms, close eyes), however, this may be prone to error in this particularly physical scenario and may give rise to an unacceptable level of effort (concentration). Speech is likely not an option due to recognition issues in the operation environment, though one could consider using the tongue as input [49]; nevertheless, this may not work well under stress, and can be difficult to use for more complex actions, likely caused by unwanted high level of effort.

On the output side, monitoring the users attention span may not be easily achievable – sensing the user's stress level probably can provide initial indications, but, environmental conditions should be taken into account too (for example, measuring temperature) to infer the current situation. If it is likely the user cannot focus on digital visual information, a different output channel may be necessary. Examples are the usage of vibro-taction to provide spatial, directional cues [50] or even electro-stimulation of skin or muscles to warn users for potential danger [51].

Validation. To analyze the appropriateness of the techniques, several issues need to be validated. The performance of the input methods need to be checked, reflecting the changing operational conditions as well as the fluctuating flow of action while switching input as well as output methods. Furthermore, it should be validated if alternative sensory output methods may potentially be used in cohesion with visual output, hence allowing for sensory enrichment in intermediate stress levels. Even so, under stress even these signals may be completely ignored (non-processing of stimuli): a signal may not result in the triggering of a user action. Hence, a close look needs to be taken at concentration and associated effort levels, while checking actual responses to sensory stimuli. Obviously, performance needs to be reflected in relation to a cognitive load analysis, focusing on operational issues (input) as well as spatial knowledge acquisition and usage while exploring the building in the simulated environment. Affected by cognitive load, a careful analysis of the compatibility between stimulus and response needs to be performed, since under stress a stimulus may well not result in an appropriate response, which may in real-life potentially lead to a dangerous situation.

9 DISCUSSION AND CONCLUSION

In this article, we presented methods for the human potential driven analysis, design and validation of 3D interaction techniques. The example in the previous section showed the challenges of following a human potential approach, which often results into thinking outside the box. The key to the methodology is the interplay between the presented energetic analysis model and the control and sensory substitution methods and issues. Both go hand in hand in the process to develop interfaces that can theoretically increase performance in certain task spaces, or for specific users. Even though design for human potential, including energetic analysis, could also be performed *without* substitution methods, a performance driven approach could be harder without a suitable base set of techniques to compare to.

Reflecting the article, we see several areas for future work. Underneath the methods, we provided a loosely structured approach. The current approach can benefit from a formalized, step-by-step formalization that directly connects all main factors involved. Of particular interest for the approach would be the direct connection of human factors and processes at a neurological level, validating actual mental effort using EEG. An emerging field that can influence this extension of our approach is neuroergonomics [24]. In direct relation, a formalized approach can also certainly benefit from step by step guidelines for evaluation (an evaluation framework), which was largely omitted from this paper due to its inherent complexity.

Furthermore, the current statements are not conclusive: several factors are largely omitted, such as cognitive factors, training effects, the actual design of ergonomic devices, the connection between 2D and 3D sensory and control substitution, and the design of interfaces that constitute a single sense, such as audioscapes [52]. We also largely ignored the usage of non-motor control channels. For example, biofeedback mechanisms may trigger and retrieve information at neural level or the brain. Clearly, the energetic action-feedback loop changes by using biofeedback systems. Though, the premise of the model still is valid: how much effort needs to be spent? The usage of a brain-computer interface may be ergonomically apt, yet, it may require great effort at the cognitive level [53].

Concluding, we hope this article will help interface designers to take a closer look at human potential. Sensorimotor potential is vast, and many interface designers still do not consider or deploy its possibilities. Notwithstanding, even though we promote an experimental approach, the drawback of using the human potential-driven methodology is that developers may run into social barriers of experimentalism. Many techniques may require social, cultural up to even human sensorimotor adaptation to be fully usable. Using brain-computer interfaces is just one example. Nonetheless, we believe the human-potential driven design and the there out possibly forthcoming vivid interfaces are of high validity and practicality for 3DUI research and development.

ACKNOWLEDGEMENTS

The author would like to thank Dieter Schmalstieg and Steffi Beckhaus for aiding in, and commenting the presented work.

REFERENCES

[1] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev, *3D user interfaces: theory and practice*. Addison Wesley, 2005.

[2] K. Tollmar, D. Demirdjian, and T. Darell, “Gesture + Play: Exploring Full Body Interaction for Virtual Environments,” in *Proceedings of the 2003 ACM Conference on Human Factors in Computing Systems (CHI 2003)*, 2003.

[3] F. Mueller, S. O’Brien, and A. Thorogood, “Jogging over a distance,” in *CHI ’07 extended abstracts on Human factors in computing systems - CHI '07*, 2007, p. 2579.

[4] K. Stanney, “Handbook of Virtual Environments.” Lawrence Erlbaum and Associates, 2002.

[5] D. A. Bowman, S. Coquillart, B. Froehlich, M. Hirose, Y. Kitamura, K. Kiyokawa, and W. Stuerzlinger, “3D user interfaces: new directions and perspectives.,” *IEEE Comput. Graph. Appl.*, vol. 28, no. 6, pp. 20–36, Jan. 2008.

[6] D. Bowman and L. Hodges, “Formalizing the Design, Evaluation, and Application of Interaction Techniques for Immersive Virtual Environments.,” *J. Vis. Lang. Comput.*, vol. 10, no. 1, pp. 37–53.

[7] I. Poupyrev, S. Weghorst, M. Billinghurst, and T. Ichikawa, “A Framework and Testbed for Studying Manipulation Techniques for Immersive VR,” in *Proceedings of the 1997 ACM Symposium on Virtual Reality Software and Technology (VRST’97)*, 1997, pp. 21–28.

[8] A. Sutcliffe, *Multimedia and Virtual Reality: Designing Multisensory User Interfaces*. Psychology Press, 2003, p. 352.

[9] C. Spence and C. Ho, “Multisensory interface design for drivers: past, present and future.,” *Ergonomics*, vol. 51, no. 1, pp. 65–70, Jan. 2008.

[10] C. Lenay, O. Gapenne, S. Hanneton, C. Genouëlle, and C. Marque, “Sensory Substitution: Limits and Perspectives,” in in *Touching for Knowing*, Y. Hatwell, A. Streri, and E. Gentaz, Eds. 2003, pp. 275–292.

[11] J. Loomis, “Sensory Replacement and Sensory Substitution: Overview and Prospects for the Future,” in in *Converging technologies for improving human performance*, M. Roco and W. Bainbridge, Eds. Kluwer Academic Publishers, 2003.

[12] S. Fels, “Glove-Talk II: Mapping Hand Gestures to Speech using Neural Networks: An Approach to Building Adaptive Interfaces,” PhD Dissertation, University of Toronto, 1994.

[13] G. McMillan, R. Eggelston, and T. Anderson, “Nonconventional Controls,” in in *Handbook of Human Factors and Ergonomics*, 2nd ed., G. Salvendy, Ed. New York: John Wiley and Sons, 1997, pp. 729–771.

[14] N. Cleveland, “Eyegaze Human-Computer Interface for People with Disabilities,” in *Proceedings of the First Automation Technology and Human Performance Conference*, 1994.

[15] R. J. K. Jacob, “Eye Tracking in Advanced Interface Design,” in in *Virtual Environments and Advanced Interface Design*, W. Barfield and T. A. Furness, Eds. New York: Oxford University Press, 1995, pp. 258–288.

[16] J. Wilder, G. K. Hung, M. M. Tremaine, and M. Kaur, “Eye Tracking in Virtual Environments,” in in *Handbook of Virtual Environments*, K. Stanney, Ed. Lawrence Erlbaum and Associates, 2002, pp. 211–222.

[17] S. Beckhaus and E. Kruijff, “Unconventional human computer interfaces,” in *Proceedings of the conference*

on SIGGRAPH 2004 course notes, 2004, vol. Los Angele.

[18] E. Kruijff, "Unconventional 3D User Interfaces for Virtual Environments," PhD Thesis, TU Graz, 2007.

[19] D. Pai, "Multisensory Interaction: Real and Virtual," in *Proceedings of the International Symposium on Robotics Research*, 2003.

[20] M. Ernst and M. Banks, "Humans integrate visual and haptic information in a statistically optimal fashion," *Nature*, vol. 415, no. 6870, pp. 429–433, 2002.

[21] C.-C. P. Chu, T. H. Dani, and R. Gadh, "Multi-sensory user interface for a virtual-reality-based computeraided design system," *Comput. Des.*, vol. 29, no. 10, pp. 709–725, Oct. 1997.

[22] G. Salvendy, *Handbook of Human Factors and Ergonomics*, 2nd ed. John Wiley & Sons, 1997.

[23] D. Bowman, "Principles for the Design of Performance-Oriented Interaction Techniques," in *Handbook of Virtual Environments*, K. Stanney, Ed. Lawrence Erlbaum, pp. 277–300.

[24] R. Parasuraman and M. Rizzo, *Neuroergonomics: The Brain at Work (Oxford Series in Human Technology Interaction)*. Oxford University Press, USA, 2008.

[25] H. Luczak, "Task Analysis," in in *Handbook of Human Factors and Ergonomics*, 2nd ed., G. Salvendy, Ed. John Wiley & Sons, 1997.

[26] E. Goldstein, *Sensation and Perception*, 5th ed. Brooks Cole, 2002.

[27] H. Bullinger, P. Kern, and M. Braun, "Controls," in *Handbook of Human Factors and Ergonomics*, G. Salvendy, Ed. John Wiley & Sons, 1997.

[28] H. Lusted and R. Knapp, "Controlling Computers with Neural Signals," *Sci. Am.*, vol. 275, no. 4, pp. 82–87, 1996.

[29] S. Shimojo and L. Shams, "Sensory Modalities are not Separate Modalities: Plasticity and Interactions," *Curr. Opin. Neurobiol.*, vol. 11, pp. 505–509, 2001.

[30] L.-T. Cheng, R. Kazman, and J. Robinson, "Vibrotactile Feedback in delicate Virtual Reality Operations," in *Proceedings of the Fourth ACM International Conference on Multimedia*, 1996, pp. 243–251.

[31] D. Kontrarinis and R. Howe, "Tactile display of vibrotactile information in teleoperation and virtual environments," *Presence Teleoperators Virtual Environ.*, vol. 4, no. 4, pp. 387–402, 1995.

[32] M. Massimino and T. Sheridan, "Sensory Substitution for Force Feedback in Teleoperation," *Presence Teleoperators Virtual Environ.*, vol. 2, no. 4, 1993.

[33] S. Lederman, G. Thorne, and B. Jones, "Perception of Texture by Vision and Touch: Multidimensionality and Intersensory Integration.," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 12, no. 2, pp. 169–180, 1986.

[34] R. Sekuler, A. Sekuler, and R. Lau, "Sound alters Visual Motion Perception," *Nature*, vol. 385, 1997.

[35] R. Blake, K. Sobel, and W. James, "Neural Synergy Between Kinetic Vision and Touch," *Psychol. Sci.*, vol. 15, no. 6, 2004.

[36] J.-P. Bresciani, M. Ernst, K. Drewing, G. Bouyer, V. Maury, and A. Kheddar, "Feeling what you Hear: Auditory Signals can modulate Tactile Tap Perception," *Exp. Brain Res.*, vol. 162, pp. 172–180, 2004.

[37] J. Weisenberger and G. Poling, "Multisensory Roughness Perception of Virtual Surfaces: Effects of Correlated Cues," in *Proceedings of the 12th International Symposium on Haptic Interfaces for Virtual Environments and Teleoperators Systems (HAPTICS'04)*, 2004.

[38] C. Spence and S. Squire, "Multisensory Integration: Maintaining the Perception of Synchrony," *Curr. Biol.*, vol. 13, pp. 519–521, 2003.

[39] P. M. Fitts, "The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement," *J. Exp. Psychol.*, vol. 47, pp. 381–391, 1954.

[40] W. Penfield and T. Rasmussen, *The Cerebral Cortex of man - a Clinical Study of Localization of Function*. New York: The Macmillan Comp., 1950.

[41] B. Shneiderman, "The Limits of Speech Recognition," *Commun. ACM*, vol. 43, no. 9, pp. 63–65, 2000.

[42] R. Rosenfeld, D. Olsen, and A. Rudnicky, "Universal Speech Interfaces," *Interactions*, vol. VIII, no. 6, pp. 33–44, 2001.

[43] S. Oviatt and P. Cohen, "Multimodal Interfaces that Process What Comes Naturally," *Commun. ACM*, vol. 43, no. 3, pp. 45–51, 2000.

[44] M. Grasso, D. Ebert, and T. Finin, "The Integrality of Speech in Multimodal Interfaces," *ACM Trans. Comput. Interact.*, vol. 5, no. 4, pp. 303–325, 1998.

[45] R. Jacob and L. Sibert, "The Perceptual Structure of Multidimensional Input Devices," in *Proceedings of the 1992 ACM Conference on Human Factors and Computing Systems (CHI'92)*, 1992, pp. 211–218.

[46] E. Kruijff, S. Conrad, and A. Mueller, "Flow of Action in Mixed Interaction Modalities," in *Proceedings of HCI International*, 2003.

[47] W. Buxton, "Chunking and Phrasing and the Design of Human-Computer Dialogues," in *IFIP World Computer Congres*, 1986, pp. 475–480.

[48] E. Kruijff, J. E. Swan II, and S. Feiner, "Perceptual issues in augmented reality revisited," in *Proceedings of the 9th IEEE International Symposium on Mixed and Augmented Reality*, 2010, no. 13–16 Oct. 2010, pp. 3–12.

[49] C. Salem and S. Zhai, "An Isometric Tongue Pointing Device," in *Proceedings of the 1997 ACM Conference on Human Factors in Computing Systems (CHI'97)*, 1997.

[50] R. Lindeman, J. Sibert, C. Lathan, and J. Vice, "The Design and Deployment of a Wearable Vibrotactile Feedback System," in *Proceedings of the 8th IEEE International Symposium on Wearable Computers*, 2004, pp. 56–59.

[51] E. Kruijff, D. Schmalstieg, and S. Beckhaus, "Using neuromuscular electrical stimulation for pseudo-haptic feedback," In *Proceedings of the ACM Symp. Virtual Real. Softw. Technol. VRST 06*, p. 316, 2006.

[52] G. Eckel, "Immersive audio-augmented environments: the LISTEN project," in *Proceedings of the Fifth International Conference on Information Visualisation*, 2001, pp. 571–573.

[53] R. Krepki, B. Blankertz, G. Curio, and K. Mueller, "The Berlin Brain-Computer Interface (BBCI) - Towards a new Communication Channel for Online Control of Multimedia Applications and Computer Games," in *Proceedings of the 9th International Conference on Distributed Multimedia Systems (DMS'03)*, 2003.