Visual interaction : between vision and action

Document status and date:
Published: 01/01/2005

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Download date: 16. Nov. 2023

TU/e

technische universiteit eindhoven

Inaugural lecture
18 march 2005

prof.dr.ir. Jean-Bernard Martens

visual
interaction:

between vision and action

/ department of industrial design

Inaugural lecture

Presented on March 18, 2005
at Technische Universiteit Eindhoven

visual interaction:

between vision and action

prof.dr.ir. Jean-Bernard Martens

Introduction

Mijnheer de Rector Magnificus, Dames en Heren,

I drew inspiration for this introduction from one of my favorite historical figures, Johannes Kepler. Johannes Kepler was a mathematician who lived from 1571 to 1630, and who is considered to be the founder of modern astronomy [1]. One unfamiliar aspect of him that I want to draw attention to today is his style of writing. None of his publications would probably pass a modern review process. Unlike modern scientists, he provided a chronological account of his research, documenting both successes and failures. The effect is that his writings have two distinguishing characteristics. First, his accounts are personal and provide insight into his ambitions, hopes and beliefs, even those that he was not able to substantiate and those that look strangely unscientific from our modern perspective. For example, as is indicated by the title of his second book, called "Harmonices Mundi", he was more interested in heavenly harmonies than in the elliptical planetary motion that we now give him credit for. Second, rather than only documenting the results of his research, he was equally interested in explaining the methods that he developed to accomplish them. Inaugural lectures are nowadays one of the few occasions where this style of writing can still be practiced without serious repercussions, and I therefore want to slightly digress for once from the more impersonal and "after-the-fact" style of presentation that I practice most of the time. This means that I will also discuss some of my beliefs and expectations.

I stay for a while with the example of Johannes Kepler, because it can serve to illustrate some important topics of my talk. More specifically, I want to go into how he described his major results. His first planetary law states that planets move around the sun in an ellipse, with the sun at one focus. I want to draw attention to a characteristic of his first law that is less obvious, namely the fact that it can be communicated effectively with a simple statement. This statement of course only makes sense if you already share an understanding of ellipses and focal points. His second law is different in this respect. It states: "the focal radius joining a planet to the sun sweeps out equal areas in equal times". This law

expresses the relationship between the variable speed of a planet and the position on its orbit. Because it is less obvious, most textbooks use a picture, such as the one shown in Figure 1(a), to clarify it. This picture shows elliptic sectors of equal area that have different path lengths on the ellipse. This complementary nature of descriptions and depictions, which we take for granted most of the time, will be one of the topics of my talk today.
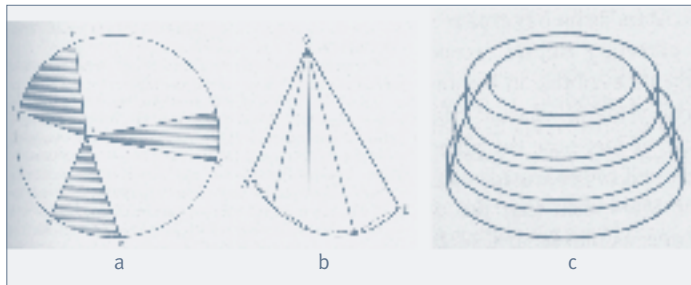
In order to derive his second law, Kepler had to measure the area of the elliptic sectors that were traversed by a planet, such as Mars, using the accurate astronomical data that he obtained from his predecessor Tycho Brahe. In order to accomplish this, he developed a new technique. He divided the elliptic sectors into small triangles, as shown in Figure 1(b), and summed the areas of these triangles. His method later inspired scientists like Newton and Leibniz to develop one of the most important and fruitful areas of modern mathematics, called differential and integral calculus. Although I cannot be completely sure, I suspect that this technique of numerical integration was not invented by simply reasoning about it, but was instead inspired by playing around with sketches of the problem. This touches upon another topic of my talk, namely that visual representations can serve as sources of inspiration.

An interesting anecdote is that Kepler also applied his numerical integration method to a completely different problem, namely the calculation of the volume of a wine barrel. At the time he was planning to marry and was involved in a discussion with his wine merchant about the exact volume of the wine barrels that were delivered to him.

Obviously, Kepler claimed that he was being overcharged. The technique he used to prove his claim was fundamentally identical to the way he solved the problem of the area of the elliptical sectors. More precisely, he subdivided the barrel into a stack of small cylinders for which he could easily derive the volume from the height and the diameter, as shown in Figure 1(c). This aspect of transposing methods from one field to another has always inspired me, and I intend to also present some examples from my own scientific practice.

Last but not least, a parallel can be drawn between the extent of knowledge about astronomy at the time of Kepler and the state of affairs in human-computer interaction research today. Kepler was an intermediate figure, in the sense that he demonstrated the order underlying the very reliable but unstructured observations of Tycho Brahe. Although he was able to describe the underlying structure, i.e., elliptic planetary motion, he was not able to explain it, based on more fundamental principles. It took another great scientist, Isaac Newton, to come up with a deductive proof of Kepler's laws. I observe a similar situation today in human-computer interaction. One of the best-known and established experimental findings in human-computer interaction is called Fitts' law [2,3]. More precisely, it has been verified in many different experimental conditions that there is an approximately linear relationship between the time T required to select a target and the index of difficulty ID, i.e., $T = a + b$ ID. This index of difficulty ID $= \log(1+D/W)$ only depends on the ratio of the distance D from the starting point to the target over the width W of the target, as illustrated in Figure 2(a). There are a number of practical implications of this law. First, it implies that a fixed resolution works well for interaction on a computer screen. Indeed, enlarging an image with a fixed resolution, as I am doing in this presentation by projecting the image on my laptop onto the large screen behind me, increases both the size and the distance between the icons by the same factor. This means that the index of difficulty remains constant. Fitts' law states that the interaction time and effort can also be expected to stay approximately the same in such a case. Second, the coefficients a and b that express the linear relationship between T and ID vary with the interaction technique, and can hence be used to characterize and compare different interaction techniques. Figure 2(b), for example, compares experimental results for two different pen interaction techniques: pick-and-drop, in which case the pen can be lifted

when moving icons from one position to another, and drag-and-drop, in which case the pen must stay in contact with the tablet. The straight lines through the data confirm that Fitts' law does indeed apply. The fact that one or the other technique should be preferred, depending on the index of difficulty ID, was used as an argument to incorporate both techniques in a system that I will describe later.

(a) The index of difficulty ID = log(1+D/W) for a selection task only depends on the ratio of the distance D from the starting point to the target over the width W of the target; (b) Experimental data for pen pick-and-drop (P&D) and drag-and-drop (D&D) selection tasks, and their linear approximations (Fitts' law)

a

b
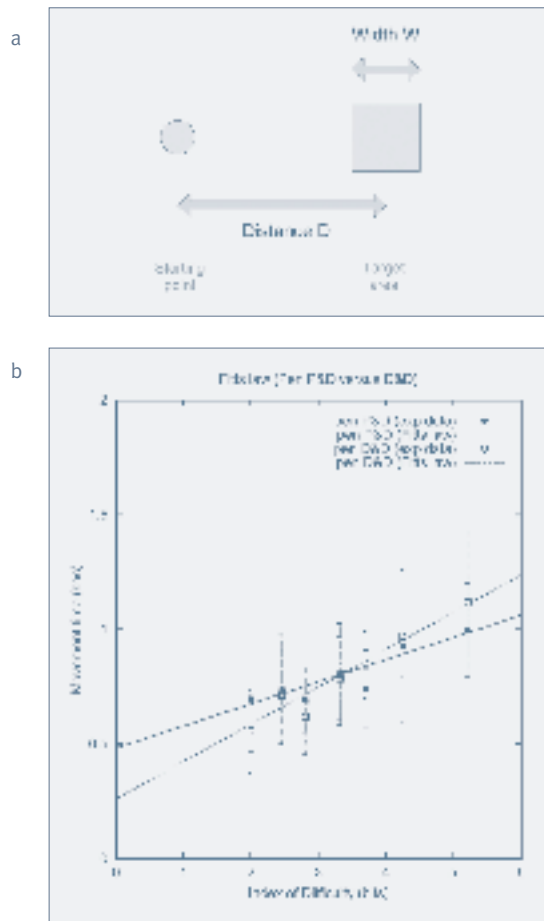
The value of the slope (a) in Fitts' law is also used in a completely different application field, namely rehabilitation, to quantify the amount of coordination in subjects' movements. This slope is higher when using your foot instead of your hand to perform the same task, or when using the non-dominant (left) hand or foot instead of the dominant (right) one. The slope also increases with age, or in patients demonstrating problems with limb coordination. Fitts' law hence has significant practical consequences. For our current discussion, it is however worth noticing that, despite the fact that Fitts' law has been around for almost fifty years, and is by now widely adopted, its limitations are not well understood. A deductive argumentation for Fitts' law has for instance never been provided up to now.

I now move to the main topic of my talk, and the title of my chair, namely Visual Interaction. I subsequently want to address its definition, why I think the topic is relevant and interesting to study, and how my interest in visual interaction relates to my former work in image processing and visual perception. I of course want to present some results of our visual interaction research, and relate them to some of the questions and issues that I am interested in today. This naturally leads to a discussion of future plans, and on how they fit in with the plans and ambitions of the Department of Industrial Design and of this university.

Visual Interaction

A couple of years ago, I adopted the term Visual Interaction to describe my current activities, because I thought it represented the right level of abstraction. This means that it is sufficiently concrete, but not overly so. It is indeed not a widely established term in the field of human-computer interaction, where it is defined as the "design of interactive systems with an emphasis on visual or graphical elements" (see www.usabilityfirst. com). This definition is obviously biased towards designing the layout of the graphical user interfaces that we find in today's personal computers (PCs). I have extended this definition to
the "design of interactive systems in which the handling of visual information dominates". As will become clear later, I will be mainly interested in systems that allow for interactions that go beyond the ones supported by the mouse and keyboard devices that are available in current PCs. As a result, I will be entering a vast field of new and rich interactions. I expect this to lead to both new applications and improvements in existing applications. One example is medical applications, where diagnosis and planning is based on two-dimensional (2D) and three-dimensional (3D) images and image sequences, and specific tasks such as determining accurate positions and sizes of anatomical structures are required. Another example is architectural and industrial design, where 2D drawings and collages, or 3D models, need to be created and updated frequently as part of the design activities. It is obvious that the way of working in these application fields has changed dramatically under the influence of computerization in recent years. Transparent films and light boxes are heavily on the retreat in the radiological departments of most hospitals, and designers have thrown out their drawing boards and hardly touch their pencils and pens anymore. Although understandable, one can wonder whether or not this is an improvement in all aspects. More precisely, I want to argue in this talk that, although this change has resulted in more sophisticated and flexible *output* of visual information, it has had a detrimental effect on the *input* side, i.e., on the flexibility and ease with which we create and manipulate information. Let me elaborate a little bit more on this theme and put it into a more general discussion about human-computer interaction.

Human-computer interaction is concerned with the exchange of information between humans and computers. The holy grail of human-computer interaction is the idea of *transparent* or *natural* interaction. My esteemed colleague Berry Eggen even considered it sufficiently important to express this in the title of his inaugural lecture a couple of months ago, i.e., "De gedroomde toekomst is onzichtbaar" (freely translated: the future we dream of is invisible). A transparent interface is expected to allow "interaction *through* the computer rather than *with* the computer", which means that it should not distract the attention from the task at hand towards the interface being used to perform the task. Otherwise stated, while interacting, you should hardly be aware of the interface and only be thinking about the primary task you are undertaking. Human-to-human communication is, at least most of the time, a prime example of a transparent interface, and I therefore analyze it here in somewhat more detail. Broadly speaking, people use two complementary means for communicating their ideas, opinions and intentions. They do so either through *descriptions*, that can be spoken or written, or through *depictions*, such as gestures, drawings or pictures. Both ways have their own limitations, advantages and side effects, and mixed forms of communication, that rely on the combined use of words and pictures (such as maps) obviously exist.

Descriptions are especially useful for expressing our thoughts, i.e., the explicit results of our internal reasoning. Descriptions are often called extrinsic, because their meaning is determined by externally defined rules of interpretation that we have acquired in the course of our live. The main obstacle for using descriptions as a way of communicating with computers is the fact that it is extremely difficult to capture the extrinsic rules that allow people to effortlessly interpret the descriptions that are provided through either speech or text. The moment a computer starts to recognize words, we immediately expect it to also understand the meaning of what we are saying, which is clearly beyond what is technically feasible today.

Unlike descriptions, depictions, such as gestures, images and sketches, are called *intrinsic*. Their meaning is not determined by externally agreed interpretation rules, but instead extracted through largely the same mechanisms of visual perception and cognitive inspection that we rely on to function in the physical world. For one, this makes depiction a very

direct way of communication that is easily shared across cultural and language barriers. We all know the saying "A picture tells more than a thousand words". This saying emphasizes an important characteristic of depictions, namely the fact that depictions tend to contain details that are not explicitly intended, and that should hence be considered *implicit*. Fish and Scrivener [4] state that "depictions tend to facilitate the search for novel visual relationships that are not easy to represent descriptively or not easy to discern because they are not explicit". For example, you might have a re-arrangement of your living room in mind, but only figure out after you start drawing it that it makes no sense. Depictions are often used to concretize descriptions, as in the case where the architect converts the verbal requirements and wishes of his client into a concrete design, illustrated by means of drawings or scale models. This concretization aids discussion and reflection, and can assist both creativity, i.e., the exploration of new avenues, and the reaching of agreement. Not only professionals such as architects and industrial designers tend to rely on depictions such as sketches and models. Most of us can probably recall instances where visualizations have complemented and facilitated our reasoning and communication. There is also scientific evidence for the fact that perceived and mental images compete for the same processing in the brain, and can influence each other. A well-known illustration of this is the so-called Perky effect [5]. Subjects were asked to form mental images of a familiar object, such as a teapot, in front of a two-way mirror. Unknown to them, a faint image of a concrete instance of such an object was projected though the mirror. Despite the fact that the subjects did not notice the real image, they nevertheless attributed a large percentage of the characteristics of the actual image (such as the color) to their mental images. There is also physiological evidence in a study of patients with unilateral visual neglect, who do not only neglect one part of their actual visual field while viewing, but who also demonstrate a matching neglect for the same half of the images that they retrieve from memory [6].

Computers are very well equipped to handle visual output media, such as diagrams, maps, sketches, photographs, films and rendered 3D models. There can hence be very little doubt that depictions are important for many of our activities. In strange contrast to this richness in output representation, is the very limited interaction that can be accomplished with them in our current computer systems. Only single-

point interaction with a cursor, most frequently controlled by a mouse, is allowed. These point-like interactions are limited to selecting, dragging and double clicking. This implies that, although we have been able to represent a large number of diverse media in our computers, we have sacrificed the flexibility in interaction that traditionally accompanies these media. For example, drawing with a mouse (or even a pen) on the canvas of an interactive graphics program is obviously much more indirect than drawing with a pencil on paper. Some of the differences are obvious:

1. action and perception space are decoupled, i.e., mouse (or pen) interactions are performed on the mouse pad (or tablet), while the visual feedback is presented on the screen;
2. two-handed interaction, where one hand manipulates the paper in position and orientation, and sets the reference for the drawing hand, is not possible; the canvas has a fixed position and orientation during computer drawing;
3. in case of a mouse, line thickness and contrast need to be controlled separately, since they cannot be derived from the pen pressure and tilt – even in cases where a digital pen is available, advanced software is needed to make proper use of the pen pressure and tilt;
4. pen and paper are light-weight and easy to carry and use everywhere.

These differences are nevertheless accepted without much second thought, because we appreciate the greater flexibility offered by software, and ... because we can always keep on using the traditional media, such as pen and paper, next to our computer when they are more convenient. Of course, continuing to use both media alongside each other results in inefficiency, because we have to convert from one medium to the other.

This renewed attention for the input rather than the output side of visual media was one of the ideas underlying the somewhat cryptic title of my talk, i.e., "Visual Interaction – Between Vision and Action". Motivated by some of the preceding arguments, I decided to adopt as a concrete goal for my research in visual interaction "to regain some of the flexibility and pleasure of use of traditional, more tangible, forms of interaction, while preserving as much as possible the diversity and flexibility of computer media". In order to motivate this, I need to argue why I expect that more tangible forms of interaction will constitute a step forward towards the holy grail of human-computer interaction, i.e., transparent interfaces.

The discussion on transparent interfaces was strongly influenced by Ben Schneiderman who introduced the concept of direct manipulation (DM) about 20 years ago [7]. He formulated the following important characteristics for DM:

1. the objects that we want to interact with should always be *visible*,
2. the actions that can be applied to these objects should be *clear*,
3. users interact with objects *directly*, instead of through intermediaries.

The graphical user interfaces that we find on today's computers are intended to implement DM. Without analyzing them in detail, I want to share some observations that indicate that the goal has only partially been reached at best. Alongside with these observations I present some arguments why tangible interactions may help to remedy some of the remaining problems:

1. The limited screen size available on desktop computers implies that the objects that we want to interact with are often *not visible*, but that we need to uncover or scroll windows and open folders to find them. One look at the average office desk demonstrates that larger 3D workspaces allow for a much more flexible way of organizing, storing and retrieving material. We use horizontally and vertically organized piles, physical media such as folders of different sizes and colors, etc. Such spatial organizations actually help us to organize our activities and to remember where we have stored things.

2. The actions that we can perform on the objects that are visible on our computer screens are often *not clear* and we frequently need to right click or open menus to reveal the interaction options available to us. These menus can be characterized as intermediaries in the interaction: they determine what the effect will be of the next mouse click or movement. Note that this style of interaction is fundamentally different from how we use tangible devices. Most tangible tools are indeed not generic, such as the mouse, but specialized. This means that we prefer to switch between tools rather than to modify a single tool (which may explain why the Swiss army knife has never become popular). It is for us more natural to use pens and pencils with different thicknesses and colors than to modify the color and thickness of a software pen. It is also more evident in advance that the red pencil will leave a red mark when used, while in most drawing programs, you are only sure about the color after you start drawing.

3. Because of the limited possibilities offered by cursor selection

and movement in 2D, even fairly simple actions need to be *decomposed* into a sequence of basic cursor actions. For example, arbitrary positioning and orienting of an object in 3D requires four separate actions in most current interfaces - two for translation, and two for orientation. Two-handed tangible interaction in 3D allows for up to 12 degrees of freedom to be controlled simultaneously. Compared to the 2 degrees of freedom of the mouse cursor, this is a vast difference. An equally important drawback of this fine-grained decomposition of user actions is that it becomes more difficult to aggregate such actions into larger, more meaningful chunks. It is broadly accepted that our memory and cognition uses such chunks as atomic units of operation [8,9]. These chunks have been compared to automated, i.e., pre-programmed, perception-action loops that are developed through training. Differences in performance between experts and novices are often attributed to the size of the chunks that they can handle when performing the same task. In the field of human-computer interaction, Fitzmaurice [10] has compared physical and virtual manipulations for similar tasks, and has concluded that larger chunks, such as combined positioning and rotation, are indeed used in case of physical objects, compared to the case of computer-controlled virtual objects.

4. DM relies on the *continuous visibility* of objects, while real objects can be manipulated fairly accurately using only tactile feedback. Tangible interaction does hence not require continuous visual attention.

Obviously, I am not the first to promote the interest in tangible interaction [11]. Similar ideas can be traced back to the activities of Myron Krueger in the 1970's, and have inspired many past and present research teams [12,13].

Implementing Visual Interaction

After this discussion on the psychological background of my research field, I want to provide some insight into the engineering effort that was required to realize more tangible ways of interaction. When we started this line of research five years ago, it was decided to develop a platform with sufficient flexibility, so that different applications could be prototyped and tested on this platform. An implication of this choice is that the platform is likely to be over-dimensioned for any of the specific applications that are developed on it. Three so-called Visual Interaction Platforms (VIPs) that differ in implementation details are currently available within our laboratory, two single-user systems, shown in Figures 3(a) and 3(b), and one multi-user system, shown in Figure 3(c).

figure 3

Three versions of the Visual Interaction Platform (VIP): two single-user VIPs (a&b), and a multi-user VIP (c)

a b c
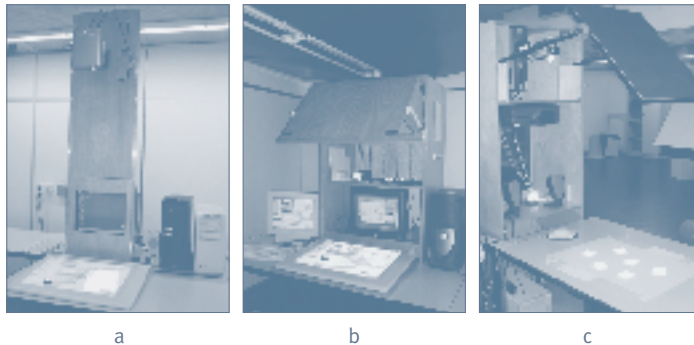
The most important characteristics of the VIPs are the following:
1. next to a mouse and a digital pen, additional input devices are available through optical tracking by means of cameras;
2. optical tracking allows to monitor the position and orientation in both 2D and 3D of several interaction devices at the same time;
3. two-handed interaction is possible, which for instance allows to move an object such as a piece of paper in the non-dominant hand while writing or selecting items with the dominant hand;

4. by using projected images, action and perception can be made to coincide in the horizontal workspace, i.e., visual feedback is provided at the position where the action is performed; this also makes it easier to combine virtual (displayed) objects with real objects, such as real pieces of paper, rulers, etc.;
5. multiple users can collectively interact at the same time, using separate interaction elements, thereby promoting group work;
6. the users do not have to wear intrusive devices like head-mounted displays that are likely to interfere with their social interaction;
7. the interaction devices do not have wires that can hinder user movements.

figure 4

The Electronic Paper (EP) prototype contains a horizontal action-perception space (bottom) and a vertical communication space (top)

The best way to illustrate some of these characteristics is by means of a number of applications that have been developed on the VIPs:
1. The Electronic Paper (EP) prototype, see Figure 4, was created by Dima Aliakseyeu as part of his Ph.D. project entitled "A Computer Support Tool for the Early Stages of Architectural Design" [14]. Its aim is to preserve the flexibility of traditional pen and paper, while meantime providing access to computer functionality. Drawings and

collages can be created with it. It has also been used to demonstrate and experimentally compare different interaction techniques.

2. Whenever we have demonstrated the EP prototype, people are disappointed by its limited functionality. Originally, we always replied by explaining that our group is not quite the size of Microsoft and that we don't have the manpower needed to create the wealth of applications that people are used to. More recently, we have largely solved the problem by allowing arbitrary windows applications to be operated from within the EP prototype. This newly developed concept is called Visual Interaction Enriched Windows (VIEWs) [15]. A VIEW acts as a filter on the input to and output from an arbitrary windows application. More specifically, user actions (amongst others with a pen on a graphical tablet) are mapped to mouse and/or keyboard actions for the windows application, while the visual output from this application is mixed with visual output from the EP prototype itself.

figure 5

Two children playing
the Read-It game
(a), as well as the
graphical user
interface to the
game (b)

a   b

3. The previous applications were primarily targeted towards professional users. In the Read-It project [16], see Figure 5, five students from the postgraduate program on User-System Interaction (USI) created a tangible game on the VIP. The game was a non competitive version of the popular game of "memory", and offered young children, between 5 and 7 years old, the opportunity to practice their reading skills. The game was designed to complement the reading method "Veilig Leren Lezen" (Learning to Read Safely) that is used at most Dutch primary schools. The main differences with PC games are that children can more easily cooperate and can be more physically active. The most important differences with existing board games is that multi-modal computer feedback can be provided, and that teacher supervision is not required.

The above examples are restricted to 2D interactions, while the research in Visual Interaction also involves 3D interaction. It is interesting to shortly discuss at least one additional project, since this allows me to give an impression of another part of the research on Visual Interaction then the part that I have been discussing up to now. In a project entitled "3D Interaction with Scientific Data", sponsored by SenterNovem, one of my Ph.D. students, Wen Qi, studies how a series of 2D images, as shown in the top part of Figure 6, can be mapped into a 3D image, such as the one displayed in the bottom left of Figure 6. Such a 3D image is assumed to be more intuitive to interpret, and therefore valuable for clinical diagnosis. The volume rendering algorithms that are used for this mapping are mathematically well understood, but require a large number of parameters to specify their operation. This implies that a completely different image can be created by changing some of these parameters, as is shown in the bottom right of Figure 6. The problem is that the relationship between the parameter values and the resulting image is difficult to understand and predict. Consequently, despite the fact that volume-rendering algorithms are available on many 3D workstations, they are seldom used in clinical practice. We are interested in researching whether or not we can substantially improve this current situation.

figure 6

Volume rendering can
map the same series
of 2D images (top)
into very different 3D
images (bottom left
and right)
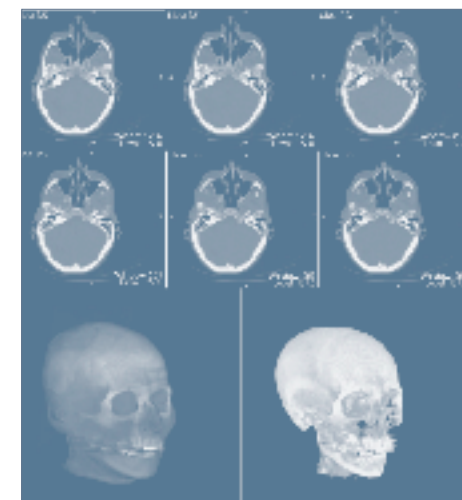
There are two complementary approaches towards such complex user-interface problems. The first one is to rely on the user, and to assume that trial-and-error adjustments of the parameters will eventually lead to a good result. The second one is to rely on the computer, and to let a computer algorithm make (or at least propose) parameter selections. The only thing that the user has to do in this latter case is to accept the result, or to choose from a limited number of alternatives that are offered to him. Since neither extreme approach seems very satisfying, we are looking into more tightly coupled co-operations between the human and the machine. A first way to accomplish this is by providing more useful feedback to the user in the trial-and-error approach, such as by visualizing the part of the parameter space that has already been explored. The second way is by adapting the computer algorithms such that their operation can be influenced by user input, such as learning from user acceptance, rejection or correction of intermediately generated results. Interestingly enough, both approaches are related to expertise that I have acquired in other areas. The visualization part relates to work I have done before on a technique called multi-dimensional scaling, while image processing is a field that I have been active in for over twenty years [17].

While evidence from psychology serves to motivate the potential advantages of alternative interaction techniques, evaluations on actual prototypes are needed to establish if the a priori expectations can indeed be met. This experimental approach to visual interaction design is what distinguishes our group from most groups that are active in this field. Therefore I want to share some experimental results that we have collected up to now in support of our endeavor towards more tangible forms of visual interaction. I would like to mention that this work on usability evaluation also evolved out of my past expertise in the subjective evaluation of image quality. Although the interest has shifted from how people passively perceive and appreciate images, to how they actively handle them, the methods that are used for gathering subjective impressions and for statistical data analysis are largely overlapping. This evolution within my professional interest from (visually triggered) perception to action is what originally motivated the sub-title of my talk, i.e., "Between Vision and Action".

figure 7

Comparison of different drawing interfaces in terms of (a) time, and (b) preference (number of subjects that score an interface as first, second or third)


a

b
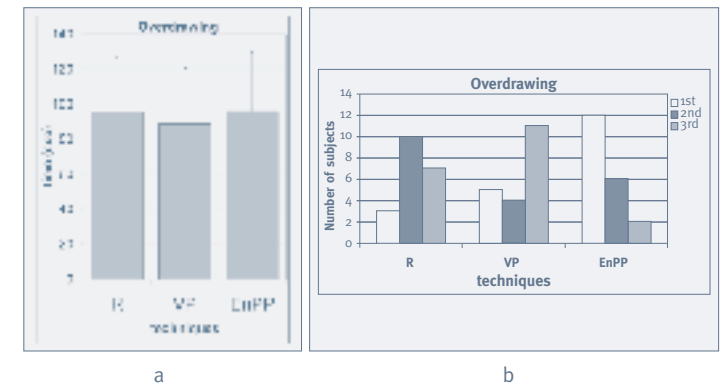
As part of his Ph.D. research, Dima Aliakseyeu compared positioning and drawing using real media with identical tasks using virtual media in

the EP prototype. In case of drawing, he compared three different techniques: 1) making a drawing with a real pen on a real piece of paper (R), 2) producing a drawing on a virtual paper that is positioned by means of an optically tracked interaction element that was shaped as a small rectanguler brick (VP), and 3) making a drawing on a virtual paper that is manipulated by a real piece of optically-tracked paper that is called the enhanced paper prop (EnPP). The results in Figure 7(a) show that there is no significant difference in performance time for this simple task between the three alternatives. The results in Figure 7(b), however, show a clear subjective preference for the case where a virtual paper, with its associated flexibility, is manipulated by a real piece of paper that provides optimum control.

figure 8

Performance times for a 3D interaction task with four different interaction techniques: 3D (F), 2D (MPR) and mixed 3D&2D (FR and PR)
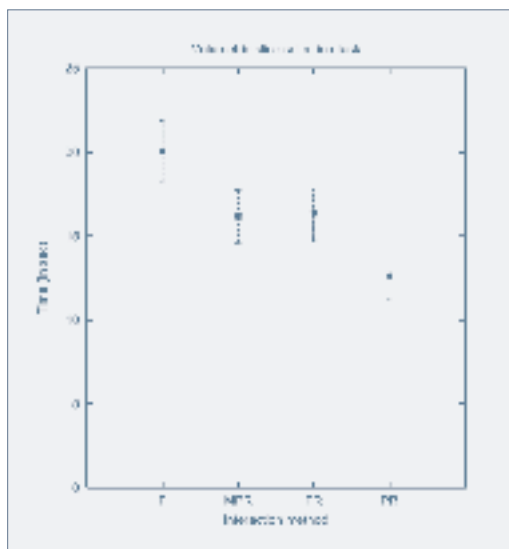
These controlled laboratory experiments were especially useful for guiding the early stages of the design of the EP prototype. For example, an earlier prototype of the EP was modified significantly, because positioning tasks on this prototype turned out to be much slower than in the case of real media. Today, we have reached the stage where the

technology is mature enough to start exploring less-controlled and more ambitious ways of interaction. We have therefore recently started a new project, entitled "ID-MIX: Industrial Design in Mixed Reality", in which another Ph.D. student, Andres Lucero Vera, will explore how mixed-reality systems can be incorporated in actual work practices. He will specifically target the application group of industrial designers.

Another example of an evaluation study is drawn from the field of 3D interaction. As part of his Ph.D. research, Sriram Subramanian has compared four different interaction techniques with respect to their performance on a 3D interaction task. More specifically, the task consisted in aligning an intersection plane with a disk-like structure within a volumetric data set. The time needed to accomplish this task was used as the performance measure. In the first, purely 3D, interaction technique (denoted by F) the position and the orientation of a 3D interaction device with 6 DOF (degrees of freedom) determined the position and the orientation of the intersection place. In the second, purely 2D, interaction technique (denoted by MPR), a 2D device (i.e., a pen) was used to modify the position and the orientation of the intersection plane. Since a 2D device can only control 1 or 2 DOF at a time, four controls were needed (2+1 DOF for position, and 2+1 DOF for orientation). This latter interaction style reflects the current state of the art in desktop interaction. Two additional mixed interaction techniques that required a 3D interaction device in one (non-dominant) hand and a 2D interaction device in the other (dominant) hand were also created. In both cases, the 2D device was used to select a point in the intersection plane. This selection initiated a switch to a condition in which only the orientation of the intersection plane around the selected point followed the orientation of the 3D interaction device. In the FR condition, the 3D interaction device controlled both the position and the orientation of the intersection plane, prior to switching to the rotation mode. In the PR condition, the 3D interaction device controlled only the position of the intersection plane, which hence had a fixed (horizontal) orientation, prior to switching to the rotation mode. The results of the experiment are summarized in Figure 8. The first result is that 2D interaction (MPR) was superior to 3D interaction (F). We expected a priori that the latter one would be more natural for the task, and hence more efficient. The reason that this didn't happen can very likely be found in the technology limitations of our experimental set-up.

Indeed, although the 3D interaction has more (6) DOF, the obtained accuracy with which 3D positions can be tracked (around 3mm) is much less than the accuracy with which 2D positions can be tracked (around 0.1 mm). Clearly, the low accuracy of 3D tracking made it difficult to perform the precise positioning task. The second result of the experiment is that 3D and 2D devices can be combined to create an interaction technique that is an improvement over purely 3D or purely 2D interaction. Using the 3D interaction device to sequentially (PR), rather then simultaneously control (F), the position and the orientation proved to be the most efficient approach for performing the task. These results illustrate that mimicking the physical world within the computer is not necessarily always the best solution, especially not if technical factors also play a role (which, of course, they always do).

Let me finish by arguing why I think that the chair on Visual Interaction fits well within the Department of Industrial Design. The mission of this department is to develop expertise that is relevant for the design of intelligent products and systems. Many of the well-known characteristics of designing such advanced systems do indeed play a role in the prototyping of the visual interaction platforms and applications. First, because of the absence of a priori defined requirements, problem definitions and solutions tend to go hand-in-hand. Second, it is only by prototyping new interaction techniques and applications that comparisons with existing solutions become possible, and detailed insights into actual benefits and limitations can be obtained. Third, unexpected problems and situations inevitably arise and lead to design modifications and new questions. These in turn often imply redesign of at least part of the system or application. This approach of simultaneously exploring questions and answers is often referred to as *analysis-by-synthesis*.

Many different disciplines are involved in the design of advanced interactive products and systems such as the VIPs. This is illustrated in the design cycle of Figure 9. Probably the easiest way to start discussing this design cycle is with the design-relevant knowledge. Often, this knowledge will be formulated in terms of models. The complexity of these models can range from *guidelines*, such as the ones for Ambient Intelligence environments that my colleague Berry Eggen has formulated in his inaugural speech, over *descriptive models* of experimental data, such as they are often available within psychology, to *full-fledged theories* derived from basic principles, as is customary in physics, mathematics and engineering. A distinction can be made between technology-relevant, user-relevant and environment-relevant knowledge:

1. Technology-relevant knowledge refers to the enabling technologies and theories that have a direct impact on the kind of products that we consider feasible now and in the near future. In the case of advanced interactive systems we specifically think about: a) the availability, and cost, of sensors and actuators, b) the signal

processing techniques that can convert sensor signals, such as images, into more abstract information, and c) the theories and models within artificial intelligence that allow to operate on this information in order to draw conclusions and plan actions.

Within my own department of Industrial Design, the group Designed Intelligence is primarily positioned to monitor and contribute to this area of expertise. Close cooperation with other departments, especially the departments of Electrical Engineering and Mathematics and Computing Science, is crucial to keep this expertise up-to-date.

figure 9
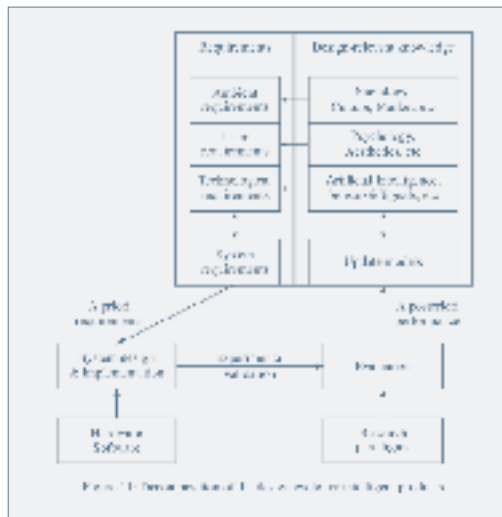
2. User-relevant knowledge refers to how individual people react to and interact with technological products. Psychophysics is traditionally the area within psychology that is concerned with studying the relationship between physical causes and psychological effects. New issues such as privacy, trust, presence, etc., that go beyond the field of psychophysics, are however becoming increasingly important. The User Centered Engineering group that I belong to is primarily responsible for this user aspect in the department of Industrial Design. Close cooperation with for instance the department of

Technology Management of this university is also very useful in this area.

3. Environment-relevant knowledge refers to both the acceptance of technological products by larger groups of people, possibly even complete cultures, and about how products can successfully integrate within existing environments. In the field of Ambient Intelligence, which is one of the research focus areas of this university, this aspect is crucial. If products become aware of their environment and start to respond to it, even when they are not explicitly addressed, social behavior, also from the side of these products, becomes increasingly important [18]. When filling the open professorship chairs at our new department, we should, in my view, keep in mind how we can develop more structured approaches towards understanding and incorporating this environment-relevant knowledge.

Based on these three sorts of existing knowledge, a priori product requirements can be formulated, and a system or application can be designed and implemented within the limitations of available hardware and software. Since the design-relevant knowledge is typically incomplete or only partially reliable, most systems will contain variables that need to be optimized. These variables reflect design choices that are very difficult, or even impossible, to make, based on the a priori requirements. Experimental verification and optimization of these design variables is therefore necessary. This may take the form of an informal and interactive tuning of these variables, but can also consist of more formal psychological experiments or field tests. Such formal tests are aimed at comparing a priori predictions of the system performance and value against a posteriori performance and appreciation when users are being confronted with the actual product or service. In case of serious discrepancies, an update of some of the models underlying the initial requirement analysis may be necessary and a new design cycle may have to be undertaken. This aspect of formalizing and modeling design-relevant knowledge is in my view essential when positioning design activities as academic research.

In our new department on Industrial Design we are preparing young people that will undoubtedly be much better equipped than I am to

Acknowledgements

Obviously, I did not arrive at this point in my career without the help of a lot of people along the way. I want to mention at least a few of them. First of all, I want to thank the board of this university and the board of the Department of Industrial Design for the confidence that they have expressed in me by appointing me to this chair. The Department of Industrial Design is a stimulating environment, where new people, with diverse backgrounds, plans and ambitions, join on a regular basis. Together they are likely to construct a department that is different from anything that I, or any other individual in the department, can imagine today. I am grateful for the opportunity to play an active part in this development. I thank the staff and students of Industrial Design for helping to create the open and constructive atmosphere that is needed to make this enterprise into a success. The fact that I have the good fortune to work in an environment that is not only professional and inspiring, but also a lot of fun, should be attributed in the first place to my colleagues from the User Centered Engineering group. I especially want to express my appreciation to my former and current Ph.D. students, Dima Aliakseyeu, Sriram Subramanian, Wen Qi and Andres Lucero-Romero, without whose efforts most of my ideas and ambitions for this research on visual interaction would never be realized.

I am convinced that the success of this department relies partly on a close cooperation with other departments. I have tried to contribute to this with some inter-departmental projects of my own, and I would like to thank my colleagues from other departments for their open-mindedness and willingness to share their time and ideas with me. More specifically, I want to thank Jack van Wijk, Robert van Liere and Arjan Kok, from the Department of Mathematics and Computing Science, and Bauke de Vries, from the Department of Building and Construction, for both interesting and entertaining discussions in the past, and, I am sure, also in the future.

Somewhat unusual, my advisory committee was presided by two deans, professor Jeu Schouten, dean of the Department of Industrial Design,

develop the interactive products of the future that I have referred to. The curricula of our existing Bachelor program and of our new Masters program describe the specific multi-disciplinary competencies that we think are most useful for them, while the two Masters themes entitled "Empowering People" and "Intelligent Spaces", define our targeted fields of applications. Further iterative design of this curriculum is one of the most demanding, but also most rewarding, challenges for the near future.

and professor Theo Bemelmans, former dean of the Department of Technology Management. My first contact with Theo Bemelmans dates back from when he was director of the Institute for Perception Research. He has played an important role in making this appointment happen, and I want to thank him for his friendly stimulation and his vote of confidence. Working within the IPO was a unique experience, as most former IPO colleagues, some of which I am happy to welcome here today, will testify. I am therefore greatly indebted to Herman Bouma, the former director of the IPO, for accepting me within this inspiring research institute in the first place. I of course want to thank my colleagues of the former Vision Group, especially the group leader Jacques Roufs, who was also my faithful tennis partner for 15 years. They are the people who introduced me to new and unfamiliar disciplines, and who have repeatedly tried, frequently in vein, to convert me to Dutch culture and habits.

On an occasion like today, when I am looking back at the recent past and forward to the near future, it is also good to commemorate how important my early upbringing has been for my personnel development. I want to use this opportunity to thank my parents, my mother who died much too young at the age of 63, and my father who can unfortunately not be present today, for their love, support and encouragement. It is only at a later age that you can start to imagine what it must have been to raise eight children on a very tight budget. My seven brothers and sisters have played an equally important role in shaping my personality. Since some of them are present in the audience today, I refrain from any statements that might be used against me later on. Let me suffice by saying that, despite the fact that we do not see each other very frequently, live without them would be a lot less interesting and fun.

Marleen and Linde are the two people that are most dear to me, and who help me to keep a proper perspective on live by never taking me too seriously. I love you both very dearly. For many people in the audience, it will not come as a surprise that I want to dedicate this talk to Marleen. July 4, 2004, is a day that we will never forget, and life will never be quite the same again as before that date. Marleen, I admire how you have always kept your spirits high in the past months, despite all the treatments that you had to undergo. Even in these difficult times, you have always succeeded in keeping a keen interest in the people

around you. In this way, you have not only helped me, but also Linde and a lot of our friends, to cope naturally with the situation.

Finally, I want to thank all family, friends and colleagues for their attention and for joining me here today.

Ik heb gezegd.

References

1 Kramer, E.E., *The Nature and Growth of Modern Mathematicians*, Princeton University Press: Princeton, New Jersey, 1981.

2 Fitts, P.M., The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement, *Journal of Experimental Psychology*, p. 381-391, 1954.

3 MacKenzie, I.S., *Fitts' Law as a Performance Model in Human-Computer Interaction*, Ph.D. dissertation, University of Toronto, 1991.

4 Fish, F. and Scrivener, S., Amplifying the Mind's Eye: Sketching and Visual Cognition, *Leonardo*, p. 117-126, 1990.

5 Segal, S.J. and Nathan, S., The Perky Effect: Incorporation of an External Stimulus into an Imaginary Experience under Placebo and Control Conditions, Perceptual Motor Skills, p. 385-395, 1964.

6 Bisiach, E. and Luzzatti, G.R., Unilateral Neglect of Representational Space, *Cortex*, p. 129-133, 1978.

7 Schneiderman, B., Direct Manipulation: A Step beyond Programming Languages, *IEEE Computer*, p. 57-63, 1983.

8 Card, S.K., Moran, T.P. and Newell, A., *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum: Hillsdale, New Jersey, 1983.

9 Vicente, K.J. and Rasmussen, J., Ecological Interface Design: Theoretical Foundations, *IEEE Transactions On Systems, Man, and Cybernetics*, p. 589-606, 1992.

10 Fitzmaurice, G.W., Ishii, H. and Buxton, W., Bricks: Laying the Foundations for Graspable User Interfaces, *Proceedings of CHI 1995*.

11 Ulmer. B. and Ishii, H., Emerging Frameworks for Tangible User Interfaces, IBM Systems Journal , p. 915-931, 2000.

12 Subramanian, S., *Tangible Interfaces for Volume Navigation*, Ph.D. Thesis, Technische Universiteit Eindhoven, 2004.

13 May, R.A., *Toward Directly Mediated Interaction in Computer Supported Environments,* Ph.D. Thesis, University of Washington, 2004.

14 Aliakseyeu, D., *A Computer Support Tool for the Early Stages of Architectural Design,* Ph.D. dissertation, Technische Universiteit Eindhoven, 2003.

15 Martens, J.B., VIEWs: Visual Interaction Enriched Windows, *Proceedings of the Second European Symposium on Ambient Intelligence,* p. 255-266, 2004.

16 Sluis, R., Weevers, I., van Schijndel, C., Kolos-Mazuryk, L., Fitrianie, S. and Martens, J.B., Read It – Five-to-Seven-Year-Old Children Learn to Read in a Tabletop Environment, *Proceedings of the Third International Conference for Interaction Design and Children,* p. 73-80, 2004.

17 Martens, J.B., *Image Technology Design – A Perceptual Approach,* Kluwer Academic Publishers: Norwell, Massachusetts, 2003.

18 Bellotti, V., Back, M., Edwards, K., Grinter, R.E., Henderson, A. and Lopes, C., Making Sense of Sensing Systems: Five Questions for Designers and Researchers, *Computer Human Interaction (CHI) Conference Proceedings*, p. 415-422, 2002.

Curriculum Vitae

Jean-Bernard Martens studied at the State University of Gent, Belgium, where he received the Electrical Engineering degree in 1979. In 1983 he obtained his Ph.D. degree on a thesis entitled Algorithms for the Calculation of Discrete Convolutions and Fourier Transformations. One of the papers resulting from his Ph.D. thesis received an IEEE paper award in the Area of Signal Processing in 1984. In the same year, he also joined the Institute for Perception Research (IPO) at Technische Universiteit Eindhoven, where he became an associate professor in the field of Visual Psychophysics in 1990. He has been the advisor for a number of Ph.D. projects on image quality, psychometrics, image processing and coding. He (co-)authored several papers on these subjects and received the 1997 EURASIP best paper award, together with his Ph.D. student A.M. van Dijk, for a paper entitled Image Representation and Compression with Steered Hermite Transforms. His book Image Technology Design – A Perceptual Approach marks the end of his research in visual psychophysics and was published by Kluwer Academic Publishers in 2003. He joined the department of Industrial Design in 2003, where he is now a full-time professor in the field of Visual Interaction. His current research focuses on developing and testing new augmented-reality interaction styles in which working with visual materials is an important component.