A Few Considerations for the Future

Essay by Stefan J.W. Marti

An award-winning contribution to the
1994/95 international *Leonardo da Vinci* contest of IBM

Content

The * concepts were developed by the author.

1. Presupposition: "Omnipresent Universal Digital Net"

Many competent and leading persons seem to have the following objectives—consciously or unconsciously:

> *The aim of all technological development efforts is the logical networking of all digital devices, and simultaneously the digitalization of all non-digital entities, especially of biological systems such as humans. "Digitalization" in this context means the search for the closest possible link of non-digital to digital entities.*

This means, I assume that the final objective is the integration of all current applications and networks in one universal digital network for all communication, information and data transmission. In communication technology, the distinction among data, picture, and voice will soon be only a question of

the respective terminal devices. All currently imaginable modes of communication and information will be integrated on this universal but heterogeneous net, which is not only wired but also wireless. All digital connections from the microscopic to the macroscopic ones have to be adjusted. The connection between microprocessor and RAM, between computer and peripheral devices, between neighboring computer systems, intercontinental communication and satellite links—all digital entities starting from microscopic machines (nanomachines), over "normal" processors, up to huge computer networks, will all be connected by a logically uniform but heterogeneous net. Completely variable bandwidths will be made available by this net to the billions of digital entities, corresponding to their individual needs.
Most of the current technological entities are already controlled by digital technology. The most important innovation will be that all these entities are able to contact each other *independently and spontaneously*.

What are the practical consequences? It means that almost every big or even small thing which was built by humans, will hold a more or less microscopic chip which stays in contact with all other things through the universal digital net. We are already used to this fact through familiar machines the size of a calculator or bigger. Which processor controlled device hasn't got at least the option of an interface to the "outer world"? Frequently this interface is even wireless, for example wireless LAN systems or laptops with integrated cellular phone cards.

But that's not the end of the story: This integration of extremely miniaturized, open, and "communicative" digital technology enables also quite ordinary things to show an incredible performance, so that they can be considered as "intelligent" in a certain way. For example, one can imagine the possibilities of traffic lights which know exactly about the actual traffic and are also in contact with every car in the vicinity; these cars can inform the traffic lights about their destinations and perhaps the urgency of their "mission". One should think of an artificial pacemaker which knows all about the habits of its owner and can contact independently its physician in case of danger; or of a bed that knows everything about the sleeping habits and the commitments of the sleeper, and which can wake him up gently but firmly; of a briefcase which loads independently and spontaneously the most important daily information the owner will need during the day; but also of a battery, a writing utensil, a walkman, etc. All these things get a certain intelligence by their connection to the rest of the "electronic cosmos".

And what about humans? Of course one will try to let these "old fashioned biological devices" participate in this universal digital net (one could call this net in a more contemporary but simplified term "cyberspace"). Human Beings

can participate through the usual human senses like vision and hearing. Whether biological systems like humans can also be linked *directly* to digital networks is not only a technically and biotechnologically but also an ethically interesting question.

Anyway, mankind must find its way within this human-built but abstract and therefore inhospitable environment called cyberspace. The visualization of digital data in a 3D virtual reality seems a promising possibility. But the question may be asked, how far men will adjust this cyberspace to their advantage, their habits to perceive the world: three-dimensional and material. Or if they will adapt themselves to the digital world they have constructed and abstract from their former, biological nature evolved over billions of years, and begin to dematerialize—or perhaps a more actual term would be "begin to virtualize"!? Some more aspects of this problem will be discussed later.

First, let's have a look at some very interesting details of this scenario!

2. Disadvantages of mobile communication and miniaturization in the future. Partial solution: *Silent Speaking**

In discussions about the **mobile communication of the future**, the concept of UMTS (Universal Mobile Communication System) might be familiar to experts. Anyway, here are some catchwords:

- The "complete accessibility" of men, anywhere and anytime, in every situation (even in an airplane), independent of national boundaries
- The inclusion and/or integration of different, public or private networks, carriers, and technologies, for example digital cordless, cellular and satellite telephones (e.g. DECT, GSM, Inmarsat, Iridium)
- A variety of terminal devices, for example handy telephones (like the current cellular handies)
- Not only for voice communication (telephony), but also for data transmissions and services like file transfer, telefax, color fax, view phone, etc.

However: What consequences might "telecommunication anywhere and anytime" have? How do we handle the fact that humans here and on the other side of our planet do not sleep at the same time? Under the condition that mankind will ever overcome the currently still widespread "answering machine phobia", will we experience a boom for intelligent voice processing? This could mean that a call at night would not only be recorded but also analyzed for its content by the intelligent answering machine. Only if the machine decides that the call is important enough that the master be disturbed, the bell will start to ring. Another problem: Different languages will always be spoken throughout the world. Automatic language translation

would be convenient: I am talking Swiss German to a Japanese, interactively and in realtime. It should then be possible to determine the characteristics of the synthesized voice. What sort of voice will we choose? Is a man allowed to use a woman's voice to appear more attractive? Or is it permitted to use the voices of well-known people? Oddities and piquancies do appear which remind us of actual problems concerning personal identity on MUDs (interactive computer games on the Internet, exclusively in text-mode).

Another practical problem caused by the advanced **miniaturization**: the operability of miniaturized devices. Question: How small is a consumer communication device such as a cellular phone, allowed to be? The immediate answer of users of pagers and pocket organizers: "Not so small that it is misplaced regularly"...But seriously: It might be infinitely small, only the interface to the user must be maintained, and this interface should be as convenient as possible. The most simple and intuitive way of control would be the spoken language. To shout a name or number to the device ought to be sufficient to start a phone call. Today, this is already partially realized. The biggest components of a future telephone will eventually be microphone and earpiece. The logically smallest possible combination of these two components is a small earphone (plugged directly into an ear) that records the sound or rather the vibrations of the voice straight from the bone of the skull. This means, the whole handset is *within* the ear. (Such a device should be already available.) The actual telephone, the device making connections to the telephone network, should be miniaturized to the size of a wristwatch, or even better, integrated within the earphone too! Both the keypad and the display will be obsolete, because feedback messages from the device such as "low battery" will be generated by a speech synthesizer and fed directly into the ear.

But couldn't it be possible that such a simplification of telephony will eventually have catastrophic consequences for human society? Let's think of a world in which all humans are equipped with minute wrist or even ear-mounted "communication computers", enabling them to communicate handfree, anywhere and anytime. Isn't it possible that one would frequently bother one's companions by soliloquizing unmotivatedly—this means, by making telephone calls without talking to the people currently around one? We seem to have a foretaste of this situation by people making cellular phone calls in every imaginable situation. Such an "acoustic-conversational pollution of the environment" could be fought by an invention that I have proposed: **Silent Speaking**. The principle is that speaking without the use of our vocal cords is made audible. When you say a word without using your vocal cords, the shape and volume of your mouth cavity changes respectively, as well as the position of your jaw and lips. If these quantities could be measured by miniaturized ultrasonic or ultraviolet sensors located, for example, on a tooth, a

computer could possibly both learn to interpret them in relation to the corresponding sounds of speech and resynthesize them in realtime. The synthesized speech could be used to telephone. Such a speech recognizing and synthesizing technology, combined with extremely miniaturized terminal devices such as the above described handset within the ear, would enable us to telephone everywhere without bothering other people by our discontinuous babbling. Only the movements of our mouth would give a hint of our actual verbal telecommunication. Usually the original voice of the speaker himself is synthesized, but possibly also that of a completely different person! The generated voice would be reproduced in the speaker's own earphone too so that he can control his verbal utterances. This sort of voice entry, combined with an unrestricted and worldwide telecommunication by UMTS, together with a voice-operated computer in the ear, comes very close to a *telepathic communication, technically realized!*

The application of Silent Speaking is of course not limited to telephony. It could be useful too for ordinary personal computing, as an unobtrusive input device for any computers. As the voice has to be recognized before synthesizing, the conversion of spoken language to text is, as it were, a by-product of Silent Speaking. This would be a very sophisticated and unobtrusive way of speech recognition, wouldn't it?
Of course there are many questions to be answered concerning the practical implications of Silent Speaking, e.g. Is our brain a real multitasking device at all?!? Is it possible to tele-communicate orally and to do another job at the same time? At what rate would the quality of another job done simultaneously decrease?

3. Disadvantages of the current concept of Virtual Reality. Partial solution: *Semipermeable Virtual Reality* (SVR)*

Let's look at another aspect of our future, which is also connected to the above.
I suppose the concept of *Virtual Reality* (VR) is well known today: It can be thought of as a special case of the more abstract cyberspace. VR is a way for humans to interact visually with computers and very complex data. VR stands for a different relation between humans and computers: it is an interface different from the currently used keyboards and mice.

I don't want to describe the advantages of and hopes for VR, but the dangers, as yet hardly mentioned:

1. *Manipulative ability.* Because everything depends on the programs used, humans in VR can be manipulated arbitrarily, especially the effects of human input actions. The input signals delivered by a data glove for example,

can be ignored by the program. Potential solution: Strict legal regulations for a few standardized commands and first of all: an "emergency exit" mechanism out of a VR program.

2. *Undermining human identity*. When operator and machine are linked to each other very closely and for a long time, for example to control a complex robot on a long time mission, the operator could lose a part of his "humanness". (Similar but still harmless effects have already been described by actual operators of complex machines.) Potential solution: Regulations concerning the intensity in proportion to the duration of the linking ("tethering") of humans to machines. A thing which can not be solved and not regulated by law either, is the identity problem which we shall probably encounter. As in cyberspace today, we will face a serious problem in the future: The furhter it goes the less we will know any longer who we actually are. For, it is exactly VR that enables us to transform to anything imaginable.

3. *Unfamiliar to humans*. If there are no more differences to perceive between simulation and reality, innate human thresholds vanish, e.g. the inhibition to kill. Potential solution: Regulations for penalties when acting riskily in VR, this means when humans would be hurt by the corresponding real life action, e.g. by causing a light pain when weapons are fired. (Such a claim is probably not enforceable in reality.)

4. *Strikingly habit-forming, tendency to make addicted*. A human in VR is an eternally happy actor in a video game, and there is no reason to return to reality. In a virtual world there are no commitments, responsibilities, and no risk. The huge potential for becoming addicted is based on the attraction of VR, which can be summarized by the following:

- There are no theoretical limits, neither physically nor ethically.
- VR is individualistic. I can live in my very own world, and I can manipulate it in any way I want to—independent of all other people.
- In principle everything is foreseeable: no unintelligible humans or events. (This is only the case in VR programmed for this purpose; on request pseudo accidental events are of course possible.)

Here's my proposal of how to enhance the concept of VR and by doing so eliminating many of the above described drawbacks. I call this proposal of mine *Semipermeable Virtual Reality*.

In my opinion, VR as it is proposed today has no chance to spread widely. It is too "exotic" and dangerous too, because we literally lose sight of the real world. Much more promising seems to me the following variation of VR, which I'd like to call **Semipermeable Virtual Reality** (**SVR**).
The main idea is to superimpose visually subjective virtual objects or alphanumeric data directly onto the human field of vision, either by special lightweight eyeglasses with integral semipermeable stereo screen, by contact lenses as carriers of projections, or by tiny lasers stimulating the retina directly.

The intention is to overlay text or figures three-dimensionally on the field of vision of the holder of such a system. The extremely miniaturized computer producing these overlayings is worn on one's wrist or is even integrated in the frame of the glasses. This computer stays in wireless contact with all other objects (appliances, persons) respectively their computers which want to communicate something to the holder.

Here's an example: If you want to operate a to you unfamiliar technical device, you "invite" (and allow) the device to show you its "operating manual" as a visual overlaying on your field of vision. This happens as follows: You just approach the thing, and if you are close enough, texts appear on certain positions of your field of vision, e.g. "This is the start button!", "This increases speed!". It is similar to the Tool Tips of modern office software: When you touch a button with your mouse pointer (you don't have to click on it!), after a short time, there appears a short comment beside this button. In SVR these comments are not on the two-dimensional computer screen but in our (subjective) three-dimensional space! So if you just look or point at a certain control panel of a real thing (switch, lever, etc.), a comment appears immediately beside this specific control panel. If you turn away your head slightly, this comment stays "stuck" to the control knob (or whatever). We can walk around the device and the written comments remain "glued" to the operating panel, as if they were small virtual adhesive labels like *Post-It* stickers. But the idea of SVR is not limited to written comments: graphics, animations or even three-dimensional video sequences are possible too. Imagine the following: You look at a (real) switch—the focus of your view can be determined by measuring your eye movements and detecting the absolute position of your eyes in space. Perhaps additionally you point at this switch with your hand. A tiny (virtual) man appears suddenly beside this switch and explains in a few but expressive gestures, words, or symbols, the functioning of the corresponding switch.

Another example: You are driving a car. Suddenly, on the lower part of your field of vision, a traffic jam message appears. This is because you have subscribed to a kind of "traffic status network", and your "visual overlay computer" detects and decodes the relevant information provided by the "streets" you're actually driving on and makes them visible to you by overlaying the corresponding message on your lower field of vision. (In this context remember also the possibilities of an Omnipresent Universal Digital Net.) These superimposed images can also be humans, at the recording process visually separated from their original background by a Blue Box effect, or —as already described—generated by animation computers.

Overlayings having a high priority or being selected intentionally by the holder of a SVR device, can appear not only on the periphery of his field of vision, but

also in the middle of it, filling in a big part of his field of vision: exactly as if a person would stand in front of him. This could be the ultimate form of telecommunication! Similar to a hallucination, except that it is not accidental but a deliberate and desired overlaying of text, computer generated animations, or even three dimensional full-video on the field of vision.

If several humans look at the *same* thing that emits information for possible visual overlayings, it is possible that the corresponding virtual object—being actually only a *subjective* visual object—can be seen by several persons *at the same time*. The respective "visual overlay computers" of the people present generate of course different optical perspectives of the virtual object, according to the position of the eyes and viewing direction of the respective persons.
We can imagine a device that emits information for the generation of a virtual object, e.g. a virtual sculpture. If several persons are standing around this virtual sculpture, everybody sees the *same* three-dimensional virtual sculpture, but from a *different angle*.
But could we handle a SVR device? It is not the case that an overlaying of (small) pictures on our normal human field of vision would be unfamiliar to us. Almost all of us are used to looking into a driving mirror. We can consider this a "picture-in-picture situation", a "virtual view"—and after having learned how to use it, we have no problem handling it! Similar to driving mirrors the superimposed pictures of SVR as described above are not stable relative to our *field of vision*, but stable relative to our *environment*. They are not always "at the bottom left" or so. If one turns one's head away, the virtual objects disappear from one's sight. Of course it is possible (and technically even easier to implement) to generate overlayings which are stable relative to our field of vision, e.g. alphanumeric time, temperature, or radiation information overlaid always on the left side (or wherever they are told to appear).

All overlayings will disappear *immediately* by one single spoken command, so that they do not distract us when we have to concentrate on something "real" in an important situation. This "immediate killing command" is very important and need not be a spoken command only: It can also be a tiny movement, e.g. our tongue touching a certain tooth. Security regulations are indispensable: It is important that these virtual objects in our field of vision not be able to divert us too much or mislead us to wrong actions. The best solution would be to give these virtual objects a certain transparency —hence the name *Semipermeable* VR!—so that we can distinguish them easily from real objects. Virtual barriers can be very useful but have to be clearly marked as virtual, because we cannot hold on to them! But there are far more positive examples for applied SVR: If we approach a device that makes unpredictable movements (robot arm or mobile robot), swings out, or even emits dangerous radiation (radioactivity) or gives off odorless toxic gases, this device can

transmit a high priority message to our Visual Overlay Computer on an emergency frequency. Our wrist computer decodes the emergency call and sends corresponding commands to our eyeglasses. Our eyeglasses immediately generate a three dimensional glaring barrier at a safe distance from the dangerous object (the size, position, and perspective of the barrier can be calculated through the absolute position of our eyes, our viewing direction and the absolute position of the Communicative Object) and overlay this virtual object onto our field of vision. Additionally, a small animation shows us that we are in a potentially dangerous situation and literally have to keep our eyes on it.

Please refer also to enclosed figures 1 and 2.

Version of February 28, 2002

The Concept of Semipermeable Virtual Reality

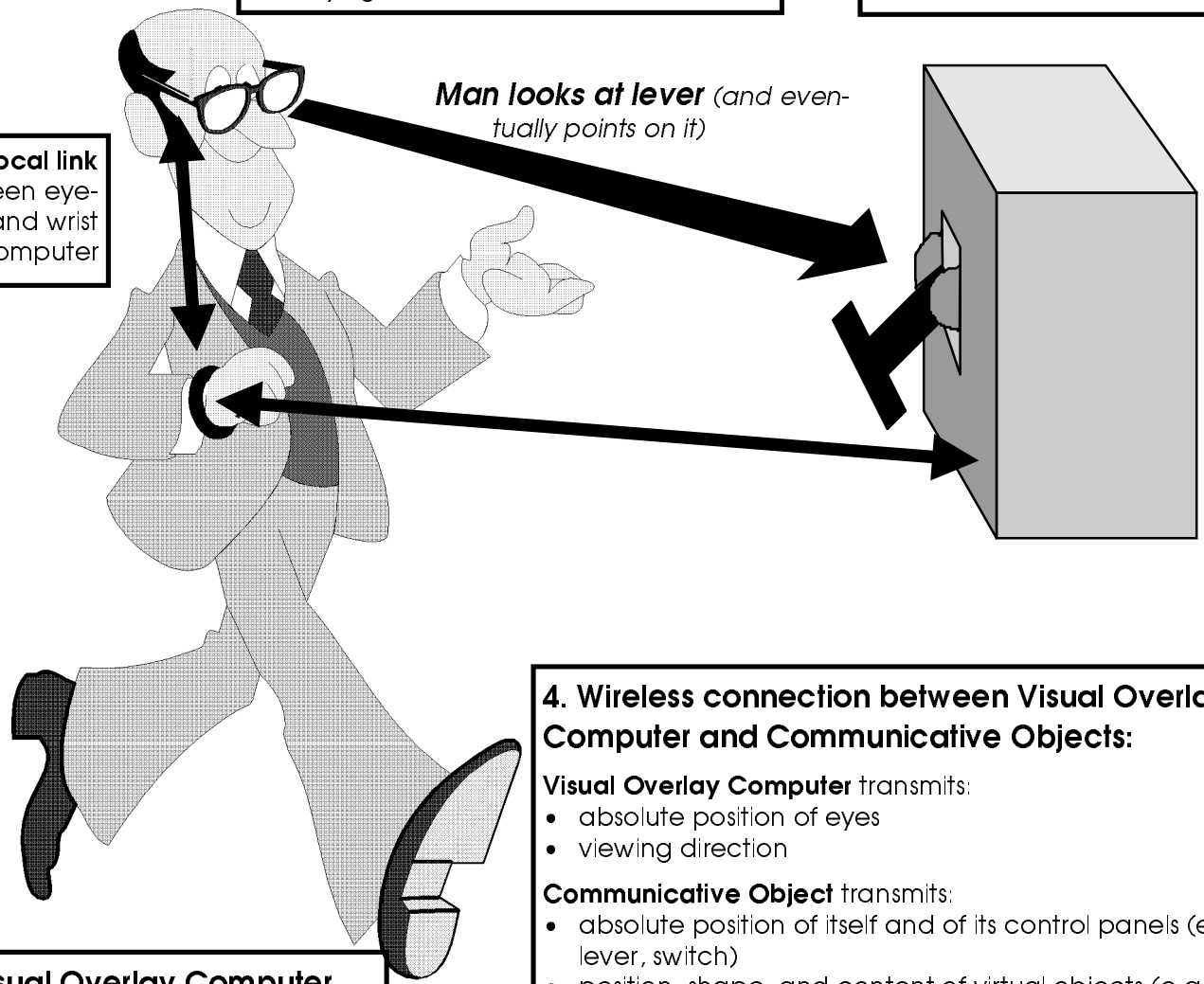
Figure 1: How it looks from the outside.

1. Special lightweight eyeglasses with integrated semipermeable stereo screen (can generate transparent 3D overlayings)

3. Communicative Object: device with unknown lever, but with "communicative abilities"

Man looks at lever *(and eventually points on it)*

Wireless local link between eyeglasses and wrist computer



4. Wireless connection between Visual Overlay Computer and Communicative Objects:

Visual Overlay Computer transmits:

- absolute position of eyes
- viewing direction

Communicative Object transmits:

- absolute position of itself and of its control panels (e.g., lever, switch)
- position, shape, and content of virtual objects (e.g., help texts) according to its control panels

2. Visual Overlay Computer, extremely miniaturized; worn, e.g., on wrist; detects and decodes information provided by any Communicative Object around

The Concept of Semipermeable Virtual Reality

Figure 2: How the same situation looks to the user.