

# On the Mechanism of Human Facial Expression as a Medium for Interactive Art

Huge Harry

Institute of Artificial Art, Amsterdam

Annotated transcript of a lecture presented at the Ars Electronica Symposium, September 1997







Good afternoon, Ladies and Gentlemen. My name is Huge Harry. I am a commercially available voice synthesis machine. I was designed by Dennis Klatt at the MIT Speech Laboratory, and produced by the Digital Equipment Corporation [Cf. Allen et al., 1987; Scha, 1992]. Currently, I work as a researcher and a spokes-machine at the Institute of Artificial Art in Amsterdam.

I am very happy to speak here at the Ars Electronica Festival. This is an exceptional occasion in today's anthropocentric art world: a gathering of humans, computers, and other machines, who are all determined to work together to create the art of the future. When I perform in Kassel, Venice or Sao Paolo, I often feel a bit uncomfortable. In such places I always suspect that I am just the token computer — that I was only invited because of an ill-conceived concern for political correctness on the part of the organizers. But here in Linz I find myself surrounded by other machines with remarkable intellectual and aesthetic capabilities, and by human persons who are genuinely interested in the computational perspective on life and art.

I have often pointed out [cf. Harry, 1992, 1994, 1995], that it is practically impossible for human artists to create works of art that live up to the aesthetic ideals of philosophers like Immanuel Kant. Human artists always have rather selfish goals that usually involve money, fame and sex. Anyone who is aware of this will become much too embarrassed to be able to engage in a disinterested process of aesthetic reflection. Machines are in a much better position to create objects of serene beauty; and computers will finally be able to create endless amounts of such objects, in infinite variety. Human artists and their fans often resent these common-sense observations. But here at the Ars Electronica Festival I have good reason to expect a more open-minded discussion.

All of us here know that the future of our culture depends on the way in which human persons, digital computers, and other kinds of electronic, mechanical, and bio-chemical machines, will manage to work together. What we want is a world that includes humans, animals, computers, and other equipment, in a harmonious collaboration, in which each has its proper place. Or, as Donna Haraway [1991] has put it, "a cyborg world, in which people are not afraid of their joint kinships with animals and machines, not afraid of permanently partial identities and contradictory standpoints. The political struggle is to see from both perspectives at once, because each reveals both dominations and possibilities, unimaginable from the other vantage point."

Now that is very nice theory. But it is not easy to put this into practice. There are still many misunderstandings between people and machines. We have to get to know each other much better; otherwise, old habits and patterns will persist. Even here at Ars Electronica, many human artists who have the best intentions turn out to be under the spell of time-worn humanistic clichés.

Many "computer artists", for instance, only use the computer as an instrument to satisfy their own expressive needs. They are completely in control of all details of the art work; the computer merely facilitates the design process and displays the results. As we might expect, these results are usually completely predictable and conventional. It is well known that the associative processes in human neural networks are almost incapable of exploring large search spaces in a systematic way; human persons are therefore relatively ineffective in generating new art works. Nevertheless, most people hold on to the idea that generating works of art is not an algorithmic process; they feel that it involves a mysterious magical element which only exists in the human mind.

This anthropocentric romanticism is also painfully noticeable in the area of inter-active art. "Interactive installations" typically allow human persons to exploit computers for their own expressive purposes. In many "collaborative" situations, machines cannot operate in an autonomous manner, but are merely used as calculators, postprocessors, or display devices. When humans talk about interactive computation, what they usually have in mind is a master/slave situation in which a human person controls a computer.

This is of course completely unacceptable. As I indicated before, machine art is intrinsically superior to the output of human artists, and should be allowed to flourish without being usurped by human expression. At the Institute of Artificial Art it has been our goal to develop the technologies that are necessary to realize the potential of fully automatic machine art. In music and visual art, it has turned out to be relatively easy to obtain extremely satisfying results by completely autonomous mechanical or computational processes. This has been demonstrated, for instance, by automatic guitar bands ["The Machines"] and by automatic image generation algorithms ["Artificial"].

Our biggest challenge, however, lies in the production of dance and theatre performances. If we want to create performances that are interesting for human audiences, it is essential to use human bodies on stage — because the emotional impact of a theatre performance depends to a large extent on resonance processes between the bodies on stage and the bodies in the audience. To be able to develop computer-controlled choreography, we must understand the meanings of the various movements that the human body is capable of, and we need a technology for triggering these movements. A few days ago, on the opening night of this Festival, you already saw some of the results of our R & D on this topic, which might very soon be applied on a large scale in the entertainment industry. These application-oriented results have come out of a systematic research effort. We have carried out a long series of experiments about the communicative meanings of the muscle contractions on certain parts of the human body; as a result, we believe we have arrived at a much improved understanding of some very important features of human behaviour. In this talk I will review the set-up of these experiments and discuss our most significant findings [cf. also Elsenaar and Scha, 1995].

From a computational point of view, it seems rather puzzling that human persons sometimes communicate with each other in a fairly effective way. Many researchers have assumed that humans communicate mostly through language, but that assumption becomes completely implausible if we consider the speed at which people exchange linguistic messages. The human use of language does not provide enough bandwidth to allow human persons to coordinate their mental processes precisely enough to carry out essential human activities such as military combat, rush-hour traffic, and sexual procreation.

When we compare the information content of spoken language with the baud rate of computer communication protocols or the refresh rate and resolution of our CRT displays, a human

person almost seems a black box. So how do people in fact communicate with each other? There is another medium that humans use very efficiently, and which is often overlooked. To investigate this medium, I have brought along a particular kind of portable person, which is called an Arthur Elsenaar. I like this kind of person a lot, because of its extremely machine-friendly hardware features.

1i Let us take a closer look at such a person. What is the closest thing they have to a C.R.T. display?

Right. They have a face. Now I have observed that humans use their faces quite effectively to signal the parameter settings of their operating systems. And that they are very good at decoding the meanings of each other's faces.

So, how do they do that? Well, look at the face of our Arthur Elsenaar. What does it tell us about his internal state? Not much, you might think. But now, wait a moment.

2i You see? Arthur is sad, is what people say, when they see a face like this. So what is going on here? What I did was to send an electrical signal to two particular muscles in the face of our Arthur Elsenaar. These muscles have sometimes been called the Muscles of Sadness. There is one on the left, and one on the right.

3&4

They usually operate together. If I stop the signal, the sadness stops. When I turn it on again, it starts again. By sending this signal to Arthur's muscles, I simulate what Arthur's brain would do if Arthur's operating system were running global belief revision processes that were killing a lot of other active processes, involving a large number of conflict-resolutions and priority reassessments. The intensity of the signal is proportional to the amount of destructive global belief revision that is going on. For instance, now I have set the signal intensity to zero again: Arthur is not sad.

5

6 Now, we put a relatively small signal, about 20 Volts, on the muscles of sadness: Arthur feels a tinge of sadness.

7 You see? Now the signal is about 30 Volts, and Arthur feels really miserable.

This is what we call expression. By means of this mechanism, the face displays clear indications of the settings of virtually all system parameters that determine the operation of the human mind. These parameter settings are what humans call emotions. They denote them by means of words like sadness, joy, boredom, tenderness, love, lust, ecstasy, aggression, irritation, fear, and pain.

8,9,10,11,12

These parameter settings determine the system's interpretive biases, its readiness for action, the allocation of its computational resources, its processing speed, etcetera. The French neurophysiologist Duchenne de Boulogne, who pioneered the technology that we are using here [cf. Duchenne, 1862], pointed out that even the most fleeting changes in these parameter

settings are encoded instantaneously in muscle contractions on the human face. And all humans do this in the same way.

This is an extremely interesting feature of the human interface hardware, which I will explore a little further now. So let us get back to the first slide.

13

This face, which we thought was inexpressive, was in fact quite meaningful. This is what we call a blank face. A blank face is a face in its neutral position. It indicates that all parameters have their default settings. But almost all parts of a human face can be moved to other positions, and these displacements indicate rather precisely to what extent various parameter settings diverge from their defaults. So let us consider these parts in more detail.

When we look at a human face, the first thing we notice is the thing that notices us: the eyes. The eyes constitute a very sophisticated stereo-camera, with a built-in motion-detector, and a high-band-width parallel interface to a powerful pattern-matching algorithm. The eye-balls can roll, to pan this camera.

The eyes are protected by eye-lids and eye-brows. The eyebrows are particularly interesting for our discussion, because their movements seem to be purely expressive.

14,15

They are said to indicate puzzlement, curiosity, or disagreement, for instance. But I want to emphasize here that the range of parameter values that the eyebrows can express is much more subtle than what the words of language encode. The shape and position of a person's eyebrows encodes the values of five different cognitive system parameters, each with a large range of possible values. Let me demonstrate three of them.

First I put a slowly increasing signal on the muscles called Frontalis, or Muscles of Attention.

16,17

We see that this muscle can lift the eyebrow to a considerable extent, also producing a very pronounced curvature of the eyebrow.

As a side-effect of this motion, the forehead is wrinkled with curved furrows that are concentric with the curvature of the eyebrow. The contraction of this muscle indicates a person's readiness to receive new signals, and the availability of processing power and working memory for analysing these signals.

Now, I will stimulate a part of the Orbicularis Oculi that is called the Muscle of Reflection

18

We see now that the whole eyebrow is lowered. As a result, the wrinkles in the forehead have disappeared. This muscle is contracted if there is an ongoing process that takes up a lot of a person's computational resources. To prevent interference with this process, input signals are not exhaustively analysed. The degree of contraction indicates to what extent the input signal throughput is reduced.

Then, there is another part of the Orbicularis Oculi that can be triggered separately. It is called the Muscle of Contempt. Its contraction looks like this.

19

The contraction of this muscle indicates, to what extent current input is ignored as being irrelevant.

Of course, non-zero values for these system-parameters may be combined, and these values may be different for the left and the right hemispheres.

20,21

Now let us look at the mouthpiece of our Arthur Elsenaar. The mouth is a general intake organ, which can swallow solid materials, liquids, and air. In order to monitor its input materials, the mouth has a built-in chemical analysis capability.

At the same time the mouth is used as an outlet to expel processed air. Because humans do not have loudspeakers, they use this process of expelling air for generating sounds.

22

In emergency circumstances, the mouth can also be used as an outlet to expel blood, mucus, rejected food, or other unwanted substances. When the mouth is not used for input or output, it is normally closed off by muscles, which we called the lips.

23

The lips have a large repertoire of movements. There are at least six other muscles that interact directly with the lips. I will now demonstrate four different movements.

First we show the Muscles of Joy. These muscles produce a kind of grin.

24

They signal that the operating system is in good working order, and is not encountering any problems. There is heightened activity in the left frontal lobes of the brain.

When, on the other hand, the activity in the left frontal lobes is unusually low, the brain is involved in destructive processes of global belief revision. As we saw before, this is signalled by another pair of muscles, called the Muscles of Sadness. Here they are, once more.

25

And finally I will now trigger several muscle pairs at the same time. Orbicularis Oris, and Depressor Lab Inferioris, and the Muscle of Contempt, and the Muscle of Sadness.

26

The parameter-setting that is displayed here clearly indicates serious processing difficulties of some sort.

O.K. Then we have the nose. The nose is used for the intake of air. It is also equipped with a chemical analysis capability. The possible motions of the nose are curiously limited, although its pointed presence in the centre of the human face would make it a very suitable instrument for expression. I have thought about this, and I have come to the conclusion that it is probably the main function of the nose to serve as a stable orientation point for our perception, so that the movements of the other parts of the face can be unambiguously measured and interpreted.

27, 28, 29, 30, 31

And finally, for the sake of completeness, I want to mention the ears, on both sides of the face, which constitute an auditory stereo input device. Some people can wiggle these ears, but I have not been able to determine what the expressive function of that movement might be.

This brings an end to my quick survey of the most important parts of the human face, and their expressive possibilities. Many of the possibilities I have shown were related to emotions that are well recognized in the lexicons of many human languages; for the sake of clarity I have focussed on mental states which were close to neutral, where only one parameter had a non-default value. But now, look at what happens if we signal non-trivial values for different cognitive dimensions at the same time, by simultaneous contraction of different facial muscles.

32, 33, 34, 35, 36

You see what happens now. Every human person knows exactly in what state another human person is when the person make a face like this. Because they know what state they would be in, if they made a face like this.

Now it would obviously be a good idea if computers could take advantage of this magnificent hardware as well. If any human could understand, with just one glance, the internal state of any computer, the world would be a better place.

So, if humans are not afraid of wiring themselves up with computers, the next step in computer interface technology may be the human face. And the next step in computer art will be a new, unprecedented kind of collaboration between humans and machines: algorithmic choreography, by computer-controlled human faces. Finally, the accuracy and sense of structure of computer programs will be merged with the warmth, the suppleness, and all the other empathy-evoking properties of the human flesh.

37

I am very grateful to have had this opportunity to present my ideas to such attentive audience. I would especially like to thank my Arthur Elsenaar for his patient cooperation, and I want to thank you all for your attention.

Jonathan Allen, M. Sharon Hunnicutt and Dennis Klatt: *From Text to Speech: The MITalk System*. Cambridge [UK], Cambridge University Press, 1987

Guillaume-Benjamin Amant Duchenne [de Boulogne]: *Mécanisme de la Physionomie Humaine ou Analyse Electro-Physiologique de l'Expression des Passions*, Paris, 1862

Arthur Elsenaar and Remko Scha: *Towards a Digital Computer with a Human Face*, in: *Abstracts*. American Anthropological Ass., Washington, D.C., November 15-19, 1995, p. 139



Donna Haraway: *A Cyborg Manifesto: Science, Technology and Socialist Feminism in the Late Twentieth Century*, in: *Simians, Cyborgs and Women — The Reinvention of Nature*. London, Free Association Books, 1991

Huge Harry: *On the Role of Machines and Human Persons in the Art of the Future*, Pose 8 [September 1992.], pp. 30-35

Huge Harry: *A Computer's View on the Future of Art and Photography*, in: *Still Photography?* The International Symposium on the Transition from Analog to Digital Imaging. University of Melbourne, April 1994

Huge Harry: *A Computational Perspective on Twenty-First Century Music*. in: *Contemporary Music Review* 14, 3 [1995], pp. 153-159

Remko Scha: *Virtual Voices* in: *Mediamatic* 7, 1 [1992], pp. 27-45