

Barbara Flueckiger

Computer-Generated Characters in *Avatar* and *Benjamin Button*

Translation of Barbara Flueckiger (2011): Zur digitalen Animation von Körpern in BENJAMIN BUTTON und AVATAR. In: Harro Segeberg (ed.): *Digitalität und Kino*. Munich: Fink (forthcoming).

Translated from the German by Benjamin Letzler, 2011.

Author's website http://www.zauberklang.ch/vfx_en_css.html

Please report errors or omissions to baflueckiger@gmail.com.

Introduction

Since the early days of computer animation, constructing human and humanoid characters has been the holy grail of the emerging technology. The first examples of three-dimensional, fully animated characters appeared in advertisements and music videos in the early 1980s. But it was only ten years later that the technology had developed sufficiently for computer-generated characters to be integrated into films, such as *TERMINATOR 2* (James Cameron, USA, 1991), *THE LAWNMOWER MAN* (Brett Leonard, USA 1992) or a digital double in *JURASSIC PARK* (Steven Spielberg, USA 1993). Another decade later, the film *FINAL FANTASY* (Hironobu Sakaguchi et al., USA/Japan 2001) was the first to incorporate characters who were classified at the time as photorealistic, but these characters did not elicit positive reactions from audiences. They seemed too flat and artificial, too wooden and regularly shaped. Gollum from *THE LORD OF THE RINGS* (Peter Jackson, NZ/USA 2001-2003) represented the first successful effort to create a favorably received and convincing character, accepted both by spectators and experts. Since then many films have been released in which such characters can be seen, sometimes even in close-up.

In my monograph *Visual Effects. Filmbilder aus dem Computer (Visual Effects: Computer-Generated Film Images, 2008)*,¹ I have addressed in detail the historical development and the technological, aesthetic and narrative aspects of such computer-generated characters. In that study, I elaborated a series of basic insights into the problems of these characters' design and reception. In the present text I will first refer back to these insights and sum them up again in condensed form. I will then consider the current state of the art through two case studies, *AVATAR* (James Cameron, USA 2009) and *THE CURIOUS CASE OF BENJAMIN BUTTON* (David Fincher, USA 2008) with regard to the nexus of technology, aesthetics and narration.

1 This monograph, a revised version of my postdoctoral thesis (Habilitationsschrift) at the Freie Universität Berlin (2007), is based on my research project supported from 2004 to 2006 by the Swiss National Science Foundation. In particular, see the chapter "Körper," ["Bodies"] pp. 417-467. For a synopsis of the entire book see http://zauberklang.ch/vfx_en_css.html, where an English translation of the chapter (as "Digital Bodies") is also available for free download.

Fundamental Problems in the Construction of Digital Characters

Why is it so difficult to construct a convincing digital character?

One of the main reasons lies in our real-world experience. From earliest childhood, we must learn how to decipher the quotidian gestural and mimic signals of people in our environment to communicate successfully. Because of this, we have developed an unbelievably fine set of senses for seemingly unimportant details. It may even be assumed that there is no realm in which we are capable of making finer distinctions than in the perception of the human form and in particular the human face.

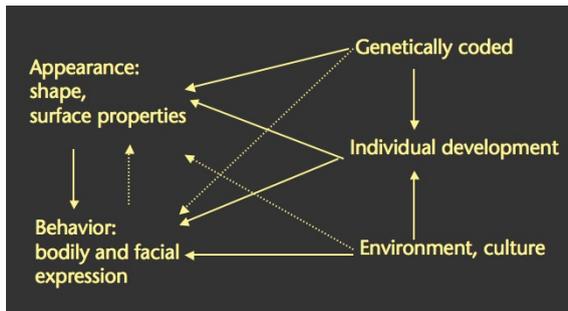
Fundamental Problem 1: Connection of Appearance and Behavior

In our experience, all of these details are integrated in a systematic context. Computer-generated characters, by contrast, are assembled from a multitude of isolated details in a highly fragmented process. It is exceedingly difficult to situate these details in an all-encompassing context and to connect them meaningfully to one other. For the analytical discussion of the matter, I have proposed that the characters should be analyzed through the factors of appearance and behavior, where appearance is subdivided into the components of form and surface (see Flueckiger, 2008: 433). Whereas behavior is the responsibility of the animation department, forms are modeled with 3D software, and surface characteristics are defined with shaders. Shaders describe the reactions of surfaces to incident light.

In the reality of our living environment, appearance and behavior are closely connected in systemic ways that can be understood within the dynamic-interactionist paradigm of personality psychology (Asendorpf 2005: 84 ff.). This paradigm assumes a close interaction between person and environment as a point of departure, and takes into account both genetically and environmentally determined factors of development that are mutually dependent (fig. 1). To this day it remains challenging to produce this kind of seemingly natural interaction between all of the characteristics of an artificially produced being.

Fundamental Problem 2: The Modeling of Complexity

For fundamental reasons, computer-generated objects have a tendency to turn out inorganic and insufficiently complex. Every detail has to be defined from the ground up while random variations that can break up the



1 Dynamic-interactionist model of personality

stiff pattern and add unforeseeable fluctuations are missing. Everything in these images must be planned. In fact, procedural methods do exist that build on stochastic and fractal algorithms to generate pseudo-random patterns. But these processes are not suitable for generating human or humanoid life forms. As I have argued in my basic considerations of the model building process,² which is central to computer-generated imagery, complexity is always the result of a process, that is to say, a history (Flueckiger 2008: 275-356).³ And this history must have a double presence in computer-generated characters: on the one hand as a history of conception, driving the design process by means of a back story and investing the character with a source that extends beyond the immediate narrative context; and on the other hand as a history of the creation of each individual computer-generated token, which consists in the writing and rewriting of code. As I will show, a number of different recording methods⁴ have been established which capture complexity from the external world, so to speak, and import it into the computer-generated world.

Fundamental Problem 3: Interaction of the Digital Character

In contrast to what happens in a fully animated film, the digital characters referred to in this text are inserted into a live-action environment and must interact with other characters, played by real actors, as well as

- 2 Model building (German “Modellbildung”) is a dominant practice in CGI. It is a rule-based, explicitly formalized system to generate 3D objects and animations. The rules apply either mathematical or physical principles or stem from empirical observation and reconstruction.
- 3 See the chapter “Abbildung” [“Representation”]
- 4 By recording I understand the translation of a physical structure according to an implicit or explicit protocol. A typical case is motion capture as the recording of motion data in a 3D space. These approaches gain increasing importance when complex structures of the real world are to be imported into the 3D space of CGI.

with the terrain and the props. This gives rise to a series of problems and demands several measures be undertaken, ideally in pre-production. The digital character must be represented on the set in some way so that the actors can react to it and adjust their timing appropriately. The simplest approach is to use cardboard to indicate generally where the actors should direct their glances. More versatile and better established is what I have termed the *proxy* approach, in which a person on set represents the character and enables physical interaction, including touching. As my research has shown (Flueckiger 2008: 249-256),⁵ touching is the best foundation for the believability and presence of digital characters. It is also an important pillar for the emotional participation of the viewers, who perceive the character always in relation to the ensemble and evaluate the character's behavior accordingly. The proxy approach nevertheless demands a highly developed tracking and retouching technique, unless a computer-controlled motion control camera is used. This type of camera records spatial data while also making possible a so-called clean pass, a process of filming without actors that provides a background for the retouching process (Flueckiger 2008: 239-249).⁶

Two Models for the Evaluation of Digital Characters

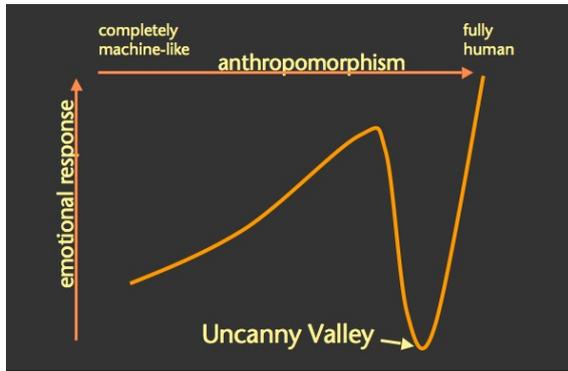
The Popular Theory of the "Uncanny Valley"

For some years now, visual effects specialists have discussed a model that first arose in the context of robotics in the early 1970s: the popular theory of the "uncanny valley" (Japanese *bukimi no tani*), introduced by Masahiro Mori (1970). Mori observed and described the following rules in the emotional evaluation of artificial characters: the more human-like such a character – for example a robot – appears, the more positive are the emotions that it elicits. Yet if the character appears almost fully human, a distancing effect occurs, which Mori called the "uncanny valley" (fig. 2). Only when the character is perceived as fully human can this effect be overcome, in which case the emotions elicited are very positive. As an illustration of this hypothesis Mori used the example of a prosthetic hand, which looks convincing but feels cold or plasticized, and evokes a spontaneous shudder in a person who touches it. Matters are very

5 Section "Berührungen, Blicke, physische Interaktion, Konsequenzen" ["Touch, Glances, Physical Interaction, Consequences"]

6 Section "Raum- und Bewegungsanpassung: Motion Control und Tracking" ["Space and Motion Matching: Motion Control, and Tracking"]

similar with digital characters. A prime example was FINAL FANTASY, which provoked this strange effect in its viewers and was correspondingly a critical failure.



2 The *Uncanny Valley*

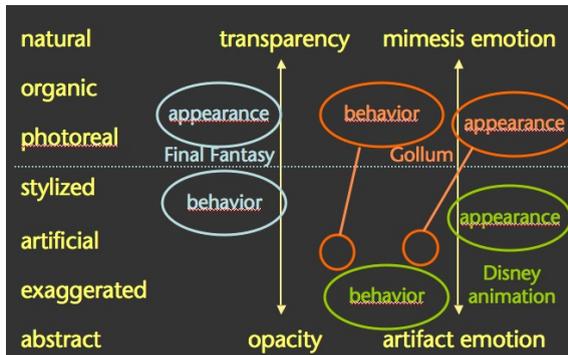
The theory of the “uncanny valley” is only conditionally universal. On the one hand there are characters who appear almost completely human and do not provoke this effect. On the other hand – and almost more problematic with regard to method – the theory has to be empirically tested for each individual case, because the presence or absence of the effect can be a matter of debate and may vary intersubjectively. This can be seen in the example of AVATAR.

The Model of Distance

As an alternative I have suggested the model of distance. It takes as a point of departure the assumption that the aspects of appearance and behavior should be at a similar distance from a photorealistic image defined as a standard. If a character is very stylized, for example, it should be animated in a correspondingly stylized way. Incidentally, this concept goes back at least to Disney’s rules, in which it was rule number 10 (“exaggerate”). The concept of distance, in accordance with its use in prototype theory, refers to the divergence from a standard value defined as the norm.

To present this situation, I have developed a matrix into which various digital characters can now be projected (fig. 3). Here we must note that there is a fine but essential line between a photorealistic and a stylized representation. It is my hypothesis that whether this line is crossed and how it is crossed are decisive for reception. In FINAL FANTASY, for example, one can observe a clear separation between a photorealistic surface and a

behavior that is stylized, since significantly limited in expressive repertoire. In contrast to this, Disney characters are unambiguously situated in the stylized domain. Gollum, lastly, is a different matter. As my larger case study has shown, his behavior and appearance both seem mainly photorealistic and organic, with a few deviations that can be categorized as fantastical or at least unusual, like the large ears and eyes, or his way of moving. I hypothesize that these deviations do not disrupt the balance of the representation, because they appear equally in the domains of behavior and form. They may better be understood as satellites (see Flueckiger, 2008: 451-461).⁷



3 The model of distance

Case Study: *The Curious Case of Benjamin Button*

The aim of this case study will be to analyze *BENJAMIN BUTTON*,⁸ including how the character has been constructed, how he masters the fundamental problems and lastly how the character acquires itself in evaluative models. *BENJAMIN BUTTON*, directed by David Fincher, is a film that does not

7 Section "Von der Fallstudie Gollum zu grundsätzlichen Überlegungen" ["From the Gollum Case Study to Fundamental Notions on Digital Characters"]

8 The following descriptions are based on a number of sources. Along with Jody Duncan's text "The Unusual Birth of Benjamin Button." In: *Cinefex*, no. 116, pp. 70-99 and 118, these include three presentations on animation that were given at the fmx conference in Stuttgart from the 5th through the 8th of May, 2009 (<http://09.fmx.de/start.php?lang=E&navi=1&page=pages>), namely by Steve Preeg, animation supervisor, and Jonathan Litt, CG supervisor, Digital Domain, by Patrick Davenport, executive producer, Image Metrics, and by Paul Debevec, professor of computer graphics at the University of Southern California.

fit neatly into established categories. It is a film in the tradition of the picaresque novel. As such it portrays the life journey of a poor but cunning hero from birth to death, and can thus be likened to *FORREST GUMP*. The narrative succession differs, however, through the fantastical element that the protagonist is born an old man and gradually grows younger. This constellation has particularly painful consequences for his love life, since there is only a short period of time in which he and his beloved share a similar age and can find each other.

It was thus necessary to represent the protagonist in several stages of his life, a challenge that earlier films had met classically and convincingly with prosthetics and make-up. Examples include Dustin Hoffman in *LITTLE BIG MAN* (Arthur Penn, USA 1970), F. Murray Abraham in *AMADEUS* (Milos Forman, USA 1984), and Marlon Brando in *THE GODFATHER* (Francis Ford Coppola, USA 1972). Visual effects supervisor Eric Barba described a fundamental problem of this approach for *BENJAMIN BUTTON*, however: “The problem with old-age make-up is that it is additive, [...] whereas the aging process is reductive. You have thinner skin, less musculature, everything is receding” (Duncan 2009: 72). Another traditional approach to solving this problem is to use actors of different ages to portray a single role in multiple life stages. Fincher wanted to avoid this approach, since it constitutes a disjunction in the character’s identity. The challenge was intensified because of the dissonance between a growing body and an old appearance. For this reason, Fincher wanted to attempt a solution with a computer-generated character. The aged version of Brad Pitt is nonetheless computer-generated only in part, since the body is that of a human actor while the head is created completely digitally and placed on the actor’s body by means of head replacement (fig. 4). In order to test this arrangement, Fincher and the Digital Domain team produced an advertisement for Orville Redenbacher’s Popcorn,⁹ which was justly called “creepy” on YouTube and circulated online as “Orville Deadenbacher,” precisely because the character was an obvious victim of the uncanny valley.¹⁰ “On the Orville Redenbacher spot, we learned everything *not* do on *BENJAMIN BUTTON*,” Fincher elaborated. “We learned that we needed to minimize the amount of keyframe animation [...] what I wanted was a process by which you could literally xerox an actor’s performance” (Duncan 2009: 75 [emphasis in original]). Fincher’s statement fits perfectly

9 The video can be seen at <http://www.youtube.com/watch?v=Fcn4p213Zg8>.

10 The digital reconstruction of deceased persons is a theme that has a pronounced ethical and moral component and that stokes fears of a post-human cinema (see Flueckiger 2009).

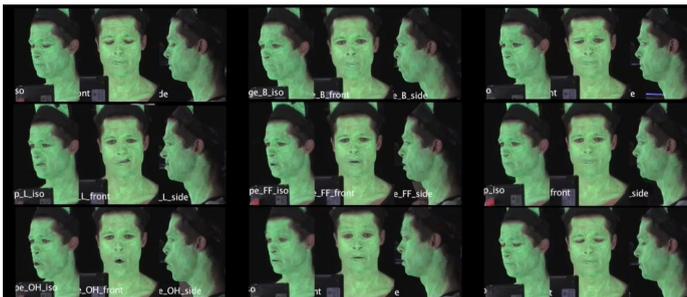
with my observations on lack of complexity, since keyframe animation¹¹ tends as a model building process to deliver highly simplified results that omit the idiosyncracies and irregularities of an individual being.



4 Head replacement in BENJAMIN BUTTON

Animation of Behavior

For the film’s animation – in other words, for the representation of behavior – recording methods were chosen that would project Brad Pitt’s patterns of facial expression onto the computer-generated face. Two techniques were deployed that interact as well as build upon one other. First, Mova’s Contour, a volumetric¹² recording method, was used to compile a database of three-dimensional facial expressions on Brad Pitt. On his face a phosphorescent make-up was applied with a sponge and filmed with multiple cameras (fig. 5).



5 Mova’s Contour as a volumetric process

The make-up serves to apply a high-contrast pattern to the skin. This pattern is needed because Mova’s Contour is based on the computer-

11 The animation of key poses, a method long widespread in hand-drawn animation.
12 Volumetric refers to the distribution of points in space.

supported analysis of up to 10,000 points in space, that is, of the position of individual pixels in three-dimensional space.¹³ Pixel flow captures even the smallest movements and is considerably more precise than performance capture methods, which work with markers or colored spots on the face.

Mova's Contour was used to create a database of 170 so-called blend shapes, basic units of facial expression systematized according to the rules of the Facial Action Coding System (FACS)¹⁴ of Paul Ekman et al. (1978). In short, this first process was used to record the face in various poses. What proved to be problematic about this technique was the absence of information on the eyes and the inside of the mouth, that is, the teeth and the tongue, which multiple modelers had to add in later by hand over the course of a months-long process (Preeg, fmx 09).



6 Recording facial expression with Image Metrics

For the next step, the recording of movement in time, an image-based method was used. For this method – Image Metrics' Performance Capture – four high-resolution digital Viper Cameras recorded Brad Pitt's facial performance from multiple perspectives (fig. 6). First Brad Pitt studied the takes recorded on the set by proxies,¹⁵ who wore a blue hood while playing Benjamin Button. The animators feared that Brad Pitt would have to study the recording for what would feel like 150,000 times before he was somewhat ready to capture the timing in all its nuance, but he actually accomplished this after only around two viewings (Preeg, fmx 09). This process should be imagined as a sort of optical dubbing (visual

13 For a demo of the process see <http://www.mova.com/flash/> [accessed March 18, 2011].

14 FACS is a system for the description of emotional expressions on the human face. It is based on the analysis of facial expressions through the structure of the underlying musculature. For the application of FACS in the animation of faces, see Flueckiger (2008: 446).

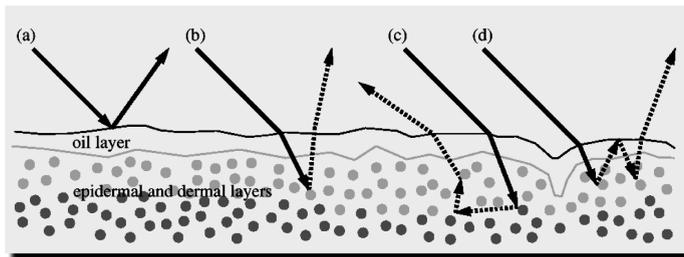
15 Three actors played the part of Benjamin Button in three different life stages (Duncan 2009: 79).

Automated Dialogue Replacement ADR), a concept that was also used in AVATAR, as I will explain below. As is always the case with more or less automatic recording systems, the results are at first far from perfect. “Every shot felt a little dead and required a lot of massaging by hand” (Preeg, fmx 09). The animators remarked that tiny shifts in the position of an eyelid, of around 1 mm, could cause dramatic changes of expression, since – as mentioned above – every viewer is an expert at decoding facial expressions.

Finally the digital 3D model was supplemented with a physically based animation system called soft body dynamics for the rule-based creation of wrinkles and of the deformations of aging skin, meant to be especially flabby.

Appearance: Form and Skin

To make the form, a plaster cast of Brad Pitt’s head was first produced, which served as the basis for three silicone models. These silicone models were created by hand and corresponded to three different ages – sixty, seventy, and eighty years (Duncan 2009: 83). They were painted and given skin details such as pores and hair. In this way, an extrapolation of Brad Pitt’s possible aging process was made by hand. Digital Domain made a 3D scan not directly of this model, but of a preliminary plaster one, in order to import the model into the digital domain. As part of this process, the model was completely scanned by a laser beam and its surface was recorded point by point. Since the scan was still missing certain details, these were then applied as shaders¹⁶ in the materialization process.



7 Subsurface scattering

16 Shaders are organized as networks in which multiple elements determine the material characteristics, for example the distribution of color, small-scale geometrical variations, the manner of reflection, aspects of light refraction, diffusion, and diffraction as well as volumetric effects, to name only a few.

Human skin has always remained one of the most challenging materials to be represented in computer-generated imagery, particularly because of its translucence, that is, its semi-transparent materiality. Light rays penetrate the surface of the skin into its deeper layers, where they are scattered in a complex way and acquire the color values of the bodily tissue and blood vessels (fig. 7). If one omits this so-called subsurface scattering, the skin looks like plaster.

In its simplest form, a shader for skin is composed of the following components: texture maps, which determine the distribution of color; displacement maps for small-scale skin variations such as pores and wrinkles; albedo maps, which record the diffusely reflecting parts of the skin that lead to whitening; specular maps for brightness reflections; reflectance maps for color reflections and texture maps of the color data under the skin's external layer for the purpose of subsurface scattering.



8 Shader elements

This shader was calibrated with Light Stage¹⁷ photographs of the silicone models. Because the source material was silicone, this did not lead to a result that appeared fully realistic, but it nevertheless provided a very good model.

Conclusion

It should be clear how hybrid and intertwined the construction and animation processes were. Recording methods certainly form an essential framework within this system, and here I would include the plaster casting in the analog domain, for both Roland Barthes and André Bazin compare photography to an imprint. But at many points the framework is provided by model building processes, whether for individual shaders, for the animation based on the dynamics of flexible bodies, or for the rendering process. The apparently organic complexity of the character is unquestionably related to the dominant deployment of recording methods.

Brad Pitt served throughout the process as reference of the recordings. For the most part this ensured the continuity of the character. Brad Pitt is clearly recognizable in the aged version of Benjamin Button (fig. 9), and this is in no small part a result of the technology, although not exclusively. As the foregoing descriptions of the animation, as well as of the materialization and the form have shown, advanced technological methods provide only the raw material. Through reworking by hand and by the skilled interlinking of a broad variety of building blocks, this raw material can then elicit a unified and compelling impression.



9 Brad Pitt and Benjamin Button

17 The Light Stage, developed at the University of Southern California under the direction of Paul Debevec, is a platform with the actor placed in the middle, lit from all sides by controllable LED lights, and recorded by multiple cameras. It is thus a photographic method that registers reflections of light. This method captures in an integrated fashion all of the information that in a shader is normally separately defined and then combined (see <http://gl.ict.usc.edu>).

The third problem set forth above, that of the interaction of the character with other characters, the objects and the space, was elegantly solved with the use of proxies whose bodies are seen in the film. Not only did this enable physical interaction, but it dispensed entirely with the need for bodily animation, and with it another problem in the creation of digital characters. Furthermore the film includes the typical obligatory touching of the computer-generated face, integrating the digital domain into the analog world. Alongside this bodily interaction, another factor for the successful integration of the digital character lies in its aesthetic coherence, which is attributable on the one hand to a good materialization of the skin with all its details, and on the other hand to a sophisticated conveyance of the light situation on the set by means of image-based lighting.¹⁸ This method makes possible a lighting situation that appears natural, even in the digital domain. An adaptive rendering in twenty passes allowed for an extremely flexible insertion of the character into the live-action environment during compositing (Litt, fmx 09). The outcome is most impressive.



10 Benjamin Button in close-up

Benjamin Button can ultimately be categorized as consistent with both models described. One can assume that the character has overcome the uncanny valley, as Peter Plantec postulated at fmx 09. Only a few brief disquieting moments cause doubt, all involving the critical zones of the eyes and the mouth, which remain the most difficult to deal with. In the model of distance, Benjamin Button can be unambiguously categorized in

18 Image-based lighting is a process by which the lighting is sampled from high dynamic range photographs of a real environment, then imported into the computer-generated scene (see Reinhard et al. 2006). One of the principal problems of the method up until now has been the limited scope for movement (see Flueckiger 2008: 166). Jonathan Litt explained at fmx 09 that the team from Digital Domain solved this problem with the compositing software Nuke.

the photorealistic domain. He appears very normal, unlike Gollum and unlike another successful digital character, Davy Jones from *PIRATES OF THE CARIBBEAN: DEAD MAN'S CHEST* (Gore Verbinski, USA 2006), whose tentacled face is unambiguously alienated and largely resists human comparison. Benjamin Button is a peculiar person nonetheless. Some strange facial slips appear to be motivated by the narrative which allows for a wide field of play. And it is concerned with an old character whose skin is no longer very elastic or transparent. This made the work easier, though it does not diminish the tremendous accomplishment of the team from Digital Domain.

Case Study: *Avatar*

James Cameron's science fiction film broke all the box office records in the winter of 2009 and 2010. That is evidence enough for the functioning of the digital characters, for how else would millions of spectators worldwide have been so strongly moved emotionally? Grown men actually reached for their tissues at the crucial moments. Real hype developed around the film, which was celebrated as a revolutionary milestone for visual effects. With the fantastical world of Pandora, a computer-generated universe was created of a complexity and sheer quantity never before seen, with fluorescent life forms and plants, with giant, mythically charged trees, and with floating islands and fantastical animals, and populated, not least, by masses of digitally constructed, otherworldly natives called the Na'vi along with the avatars, genetically engineered hybrids of humans and Na'vi whose actions are controlled by humans.

As in the case study of *BENJAMIN BUTTON*, the analysis that follows will focus on the creation of the digital characters in *AVATAR* and a critical discussion of this process on the basis of the problems set forth at the beginning of this text.¹⁹

James Cameron has asserted that he first came up with the idea for *AVATAR* in the mid-1990s (Duncan 2010, Spiegel 2010).²⁰ As with *BENJAMIN BUTTON*, traditional approaches would have been conceivable to realize

19 As sources I have used the extensive descriptions in the professional journal *Cinefex* (Duncan 2010), the descriptions in *American Cinematographer* (Holben 2010), and the presentations at the fmx 10 conference by Shawn Dunn, animation technical supervisor, and Stephen Rosenbaum, visual effects supervisor, Weta Digital.

20 An earlier version by James Cameron, part script, part treatment (which he called a "scriptment") can be found at <http://www.docstoc.com/docs/14294813/Avatar-Scriptment-by--James-Cameron>.

the story, namely a combination of make-up and scaling, which could be used to situate the larger characters in relation to the humans. Cameron likewise considered combining digital characters with shots of a rainforest. At that time, however, the technology had not developed enough to create a digital character who could carry the emotions for a feature length film. According to the myth, Cameron let the project rest and turned towards other projects, such as *TITANIC*, while he waited for the technology to develop further.

Animation of Behavior

Early on a plan was developed to work with an extended form of motion capture, namely performance capture. Performance capture is an integrated approach to the simultaneous recording of large-scale bodily movement and facial expression, which was first used for Robert Zemeckis' *THE POLAR EXPRESS* (USA 2004) (see Flueckiger 2008: 449 ff. as well as the section "Motion Capture": 145-153). Underlying this method is the correct assumption that facial expression changes in connection with bodily movement, and that the earlier technologies that handled facial expression and gesture separately could not offer a convincing, systematically linked model for the movement of face and body. "Our goal in using performance capture," noted *AVATAR* producer and longtime Cameron collaborator Jon Landau, 'was [...] to preserve the actor – because what a great actor does and what a great animator does are antithetical to each other. A great actor withholds information'" (Duncan 2010: 70). At heart the principal concern, as in *BENJAMIN BUTTON*, was to develop a style that appeared natural in opposition to the accentuated stylization typical of animated films.

As in *BENJAMIN BUTTON*, an image-based pixel-flow method was chosen. The idea was to equip the actor with a helmet camera in standard resolution and to use these recordings of the face as the basic information for the animation. In addition to this, high-definition cameras were installed on the motion capture stage to record the scene as a whole and provide additional reference material for the animation. During the motion capture session in Giant Studios, lasting many months, James Cameron also used a virtual camera. A virtual camera generally means the point of view from which a computer-generated scene is rendered, since this point of view corresponds to that of a camera. In the case of *AVATAR*, James Cameron used a specially engineered flatscreen monitor as a camera (fig. 11) to instantly capture its movement in space.



11 James Cameron with the virtual camera

With this he set in place the blocking of each scene and could test in real-time the performance of the digital characters as well as their integration into the computer-generated environment, albeit in low resolution and with a modest aesthetic for the rendered images in the style of early computer games (fig. 12).



12 Real-time visualization in AVATAR

Another aim of this technique was to get an immediate impression of the planned shots, which unquestionably made the collaborative work easier. It was also useful for the actors. On the motion capture stage, the actors work in an abstract, highly reduced sensory vacuum in which everything is only hinted at, with the objective of preventing, to the greatest extent

possible, the obscuring of view by objects so that the movement data can be recorded without interruption. Another advantage for Cameron was that the actors acted differently when they could establish a relationship with the camera and not just provide general information for an indifferent camera array: “Actors act differently depending on the camera,” noted [digital effects supervisor Nolan] Murtha, ‘if you are doing a wide master of a scene, they’re going to be a little more grand in their motion than they would be on a close-up’” (Duncan 2010: 120). Similar real-time visualizations, called “on-set previz,” had already been used in such productions as *A.I. ARTIFICIAL INTELLIGENCE* (Steven Spielberg, USA 2001) and *I, ROBOT* (Alex Proyas, USA 2004), both using the Encodacam system, which originated in the broadcasting field. Before these films, a video-based system had been used in *WHO FRAMED ROGER RABBIT* (Robert Zemeckis, USA 1988; Flueckiger 2008: 237 f.). Zemeckis pursued this idea further, in particular in *BEOWULF* (USA 2007).

It should be noted, however, that the images produced by the helmet camera – unlike the Image Metrics system used in *BENJAMIN BUTTON* – did not provide reliable movement data. There are two reasons for this. First, a very high-resolution image quality is required for pixel flow analysis to function, because without a high resolution filtering problems²¹ arise similar to those created with markers. Second, a helmet camera only makes available one frontal perspective, distorted by a fish-eye lens, instead of multiple recording angles. For this reason the recordings were dubbed – as in *BENJAMIN BUTTON* – with a system called Facial Performance Replacement (FPR) that is analogous to Automated Dialogue Replacement (ADR) (Duncan 2010: 119). In this procedure, the actors repeated their facial expression while sitting in front of a monitor and being recorded by multiple cameras thereby sacrificing an essential advantage of performance capture, namely the integration of body movement and facial expression: “So we were uncoupling the facial performance from the physical performance” (James Cameron in Duncan 2010: 119). What remained of the original idea was above all the reference material provided by the helmet camera.

Ekman and Friesen’s Facial Action Coding System (FACS) was used in *AVATAR*, as in *BENJAMIN BUTTON*, though with a different approach. In contrast to the blend shape technique used in *BENJAMIN BUTTON*, which only encompasses surface changes, the characters were given a robust and fully developed facial muscle system. It controlled the facial expression according to the 44 mimic units defined by Ekman et al. and linked by

21 Because not every point is recorded exactly, the movements in the intermediate areas have to be calculated.

them to six basic emotions of happiness, surprise, sadness, anxiety/fear, disgust/revulsion and anger. This ambitious approach was intended to compensate for the limited amount of information taken from the performance capture material (Duncan 2010: 137).

The considerable differences between Na'vi faces, with their broad noses and feline eyes, and human faces, something that I will return to in my discussion of the uncanny valley, created a difficulty in appropriately transferring the data from the performance capture recordings. There was likewise difficulty with regard to the bodily proportions. The Na'vi are supposed to be approximately 2.50 m tall. As my research has shown, this can lead to significant problems in the scaling of the movement data²² because the pattern of movement changes in relation to age, height and, most significantly, mass. In fact the Na'vi often seem too light, and it is immensely difficult to perceive their height correctly if no human characters are present for comparison.

Appearance: Form and Skin

In developing the characters' appearance, AVATAR differs fundamentally from BENJAMIN BUTTON because the Na'vi and the avatars are invented mythical creatures with varying human features. In such cases there is usually a long conceptual process that begins with drawings. Cameron himself made the first sketches,²³ in which the characters already had blue skin, something Cameron retained throughout the process in order to differentiate the characters from conventional green monsters and Martians. The large cat eyes, the broad lion noses, the large moveable ears and the tall, slender, muscular bodies also derived from Cameron's design (Duncan 2010: 75). It is not surprising, however, that the conceptual process as a whole converged increasingly on humanoid characters, because connectability (*Anschlussfähigkeit*), to borrow a term from Niklas Luhmann (1995: 35 ff.), is the basis for the process of understanding and thereby for empathic participation as well. This insight was also gained by the filmmakers: "[...] eventually we went back to a much more traditional foundation. We came to realize that if we wanted the audience to relate emotionally to these characters, there needed to be familiar touchstones" (Duncan 2010: 75). In order to ease the performance capture adjustment, the mouth area of the actors was taken on as much as possible. For the

22 Problems of scaling when using motion capture can also be observed in other films, for example THE POLAR EXPRESS (Robert Zemeckis, USA 2004) and HULK, see Flueckiger (2008: 151 f.).

23 They are shown in the making-of included in the special edition DVD of AVATAR.

avatars, especially Sigourney Weaver, the similarity to the human character had to be evident, at least for the face. Plaster casts and 3D scans of the actors were made for this purpose.

The conceptual process was followed by the transference to three-dimensional space. Many films use clay models that are then scanned in 3D and loaded onto a computer. AVATAR proceeded differently. The clay model method was used, but in parallel to it the drawings were also edited with a Pixologic software program called ZBrush, with which 3D models were produced directly in a process called sculpting (fig. 13). ZBrush came out a few years ago and is now a well-established method for producing complex, seemingly organic objects and characters (Kingslien 2004). It is a kind of three-dimensional painting program with which shapes, textures, small-scale variations and other material characteristics can be created directly in space.²⁴



13 ZBrush model of Neytiri

Life-size models of the characters were made by the Stan Winston Studio, the best-established studio in this sector since it made the dinosaurs in JURASSIC PARK. These models have the purpose of visualizing the design for communication and serving as stand-ins on the set, offering a reference both for the interaction of the actors as well as for the lighting. Two life-size Na'vi models were built, a generalized female and a generalized male, scanned in 3D and used to produce the Na'vi masses (Duncan 2010: 109).

The Stan Winston Studio was also involved in the elaboration of the skin's material characteristics, starting with setting the blue tone: "The

24 See Pixologic's website (= <http://www.pixologic.com/home.php>) as well as numerous videos on YouTube illustrating the process.

battle was coming up with a blue that would still look believable as flesh. There are blue-colored fish and birds, but no mammals with blue that we could reference" (effects supervisor John Rosengrant in Duncan 2010: 109). The pattern of bright-colored stripes to be applied on the bodies and faces also had to be designed and agreed upon. As Shawn Dunn explained at fmx10, these stripes were hard to implement, since they conflicted with the readability of facial expressions. In comparison with the Na'vi, the avatars had a somewhat brighter skin tone and a more distinctly human appearance (Duncan 2010: 75). They were also supposed to resemble their human counterparts to make it easier to recognize them. Sigourney Weaver's avatar, for example, had the actress' distinctive small nose. Furthermore, this correspondence between the actors controlling the performance capture and the digital characters was necessary in order for the data to be transferred. In particular the areas around the mouth and the eyes were critical, which is why the most detailed correlation was sought in these areas. As can be seen through a comparison between Neytiri and Zoë Saldana, the correlation is most prominent for the mouth, while the eyes are significantly different in form, color and facial position, as is the rest of the face. Alongside their strikingly altered nose and erect, moveable ears, the Na'vi had only hints of eyebrows. This made the facial animation considerably more difficult, since alterations of the eyebrows are an essential aspect of expression (Dunn, fmx 10).

One might imagine that blue skin is easier to materialize than ordinary human skin; since there is no human reference for it, spectators might forego significant parts of their critical appraisal. According to practitioners' reports the opposite is the case. Apparently the problem lies in the fact that blue skin never looks natural, and especially next to orange-colored light, such as firelight, it produces a grayish, disagreeable color that had to be removed with appropriate measures, such as an algorithmic adjustment to the spectral reflection of Na'vi skin (Duncan 2010: 130). In addition the blue tone also smoothed out the small undulations typical of skin. Employees at Weta Digital had to pick up the slack: "Courageous Weta crew members subjected themselves to extremely close-up photography [...]; but texture artists found that they had to exaggerate [...] the freckles and other imperfections: 'Our texture maps looked a lot more ratty than the real person's face'" (visual effects supervisor Guy Williams in Duncan 2010: 133). This observation is consistent with my insight that CGI's all too smooth appearance must be disrupted with additions and roughenings in order for it to appear organic and natural (Flueckiger 2008: 341). One can also assume that the ability of the human eye to differentiate is more limited in the blue spectrum, since this color is missing in the

natural perception of skin and we cannot rely on our experience. And naturally the aforementioned subsurface scattering also had to be taken into account, as was the case of Hulk, where his green skin gave rise to similar problems (Flueckiger 2008: 442). For AVATAR, an absorption model was used instead of the usual scattering model in order to let the pink-orange spectrum color tones of the inner skin and body tissue radiate through in a controlled manner (Duncan 2010: 133).

Conclusion

The marketing strategy for AVATAR put performance capture front and center. In a short clip that was included in the electronic press kit distributed to journalists, Cameron remarked, that he captured 110% of the actor's performance, and Sam Worthington agreed that he saw his own personality reflected in Jake's avatar on the screen. These statements are supported by comparisons of the performance capture with the finished animation (fig. 14).²⁵ Cameron went so far as to advocate a Screen Actors Guild or Oscar nomination for his actors (Abramowitz 2010), a request critically reflected upon by Kristin Thompson (2010).



14 Performance capture and animated version

The line of attack is clear, as well as coterminous with my observations on Gollum. As I have already indicated, it made sense for Gollum's character to be ascribed in the public eye to the performance of an actor: "From the perspective of psychological reception [...], it was a clever maneuver, because by attributing the performance to a person, the digital character was endowed with something like a physical presence far more

25 This kind of comparison with a split-screen appears to have become well established; see also fig. 9 for Benjamin Button.

concrete than the abstract and inscrutable activities of an army of animators" (Flueckiger 2008: 460). Cameron also naturally wanted to secure his status as director. He made this clear in the Los Angeles Times article quoted above, in which he emphasized that he controlled the expression in his collaboration with the actors, and that he did not want his intentions changed by the animators. This account was supported by Shawn Dunn: "He was very religious to the performance. Whatever was shot on the mocap stage had to be in the movie" (Dunn, fmx 10). This goes beyond a pure marketing strategy. Here we can come full circle on the fundamental problems, namely the problem of complexity, which, as mentioned above, is strongly supported by recording processes such as motion capture.

Even more so than BENJAMIN BUTTON, AVATAR exhibits a hybrid mixture in which recording and model building methods are combined. The film exhibits this in every domain, from modeling to materialization to animation. Thus the characters were modelled with ZBrush, but also partly scanned in 3D from models and actors; the shaders brought together painted and photographed elements with ones that were algorithmically calculated; and lastly, it is apparent from simply observing the movement sequences that not everything was recorded on the motion capture stage, but must have been keyframe-animated at least in part. Cameron estimates this proportion at 10% in the *Cinefex* article (Duncan 2010: 138). This estimation may be on the low side, because it relates primarily to those body parts, such as ears and tail, that the actors either do not have or have in a different way. It does not refer to the inner mouth area mentioned by Dunn (fmx 10), which was exceedingly challenging in BENJAMIN BUTTON as well and generally represents a critical zone. For its part, keyframe animation brings with it the aforementioned problem of consistency. As animation supervisor Richard Baneham explained, the animation rig for the face, with its cleverly designed muscular structure, was decisive for ensuring consistent facial expression, since the animation went through many hands. "But the fine nuance of the actor's performance had to be done by keyframe. The facial rig just freed the animators to concentrate on those all-important details" (Duncan 2010: 138).

Whenever the range of human action is exceeded, as for example in AVATAR with leaps and dives into the depths, one must fall back on keyframe animation. This clearly illustrates a categorical problem of computer-generated characters, namely that they are not limited by physical laws. These almost limitless possibilities can lead to the collapse of the spectators' participation.²⁶ Strangely, the bodily forms of the Na'vi

26 Here see my discussion of the superhero problem in Flueckiger (2008: 462 ff.)

and the avatars tend to lack complexity. At a minimum, the bodies exhibit a clear tendency to conform to stereotypical gender patterns, something that was typical of early computer animation. This should be evaluated as resulting from the basic tendency of computer-generated imagery to bring forth generic prototypes to the extent that everything must be planned and created from the ground up. Furthermore, the Na'vi correspond to a simplified, stereotyped image of the noble savage, as can be seen, for example, in Leni Riefenstahl's photographs of the Nuba. Some authors have placed this subliminal tendency under very critical scrutiny as a latently racist and/or imperialist ideology of the film, although this would seem to contradict the narrative's widely received anticolonial message.²⁷ In any event, Dunn stated (at fmx 10) that the bottom line for Neytiri's design was: "People should think: 'Wow, she's beautiful!'"

Despite the exceedingly hybrid process of production and a multitude of references, the character consistency turned out astonishingly convincing. Since the reference to the proxies is clear both in appearance and behavior – the significant deviations notwithstanding – this creates a bridge between appearance and behavior, even if the patterns of movement often fail to match the postulated body mass. This mismatch is due to the aforementioned scaling at the ratio between the actor's body and the avatar's body of approximately 1:1.5.

The third problem, that of interaction, proves considerably more complex than in *BENJAMIN BUTTON*. Various problematic zones play a part. First there is occasionally a direct confrontation between live-action and computer-generated imagery, for example in the avatars' introductory scene. As has been noted, these are difficult moments, since, ontologically speaking, the participating characters inhabit different worlds. In the introductory scene and some of the later scenes this problem was solved with on-set motion capture which sounds much easier than it is, although James Cameron – despite his claims to the contrary (Duncan 2010: 126) – is not the first who had this idea. Such an approach had already been used in *THE LORD OF THE RINGS* and *PIRATES OF THE CARIBBEAN: DEAD MAN'S CHEST* (see Flueckiger 2008: 151). Fundamental difficulties arise from the fact that the light situation on set is anything but optimal for motion capture cameras. In the case of *AVATAR*, the scaling problem came into play as a considerable obstacle, since Sam Worthington was clearly shorter than his avatar. For this reason part of the set was constructed to scale and the action was not recorded in one pass, but had to be split

27 See for example <http://io9.com/5422666/when-will-white-people-stop-making-movies-like-avatar>

into two parts: "Typically, the crew would capture the Na'vi or avatar character first, and then map that performance to the digital character when Cameron shot live-action" (Duncan 2010: 127). Fundamental to the functioning of this system was the real-time visualization, which the team called Simulcam, based on a combination of the motion capture system and the live-action camera.

Additional problematic areas of interaction had to be solved on the motion capture stage. Scaling was an issue there as well, and solved there by having children play the human roles in order to reproduce the right proportions (Dunn, fmx 10). The topography was hinted at with simple structural elements. For the flights on the dragon-like Banshees, movable platforms were hung and their change of position controlled by hand in order to establish the connection between the animal and the person riding it. Not only the many objects used, but also the numerous people active in the motion capture sessions caused significant masking problems of the sort prior films had sought to avoid. Markers were removed from the cameras' "field of view," resulting in gaps in their paths of movement. Here these problems were solved with a robust skeleton and advanced interpolation software from Giant Studios (Duncan 2010: 99). According to Dunn, what was decisive for the interaction's success was the aforementioned use of the virtual camera for real-time visualization of moving characters in the computer-generated environment (Dunn, fmx 10). This eliminated the common situation in the past in which actors had to act on a barren motion capture stage devoid of meaning, with no idea how their performance would be integrated into the structure as a whole.

At the end of this text, I will return again to a discussion of AVATAR in the context of the two models. While Cameron had postulated, explicitly in various interviews and implicitly in the *Cinefex* article, that his characters have overcome the uncanny valley, an empirical review is surely necessary. The differences in physiognomy are very pronounced, however, and there are many deviations from a completely human face, suggesting the conclusion that the characters are not altogether anthropomorphic.

In the model of distance, the characters behave in a similar way to Gollum from THE LORD OF THE RINGS. They are primarily photorealistic with a few deviations that do not disturb the balance, because they affect appearance and behavior in equal measure. As a result, this appears to be a thoroughly robust approach to creating convincing digital characters. Whatever criticism one may wish to formulate against AVATAR, these digital characters were able to enthuse the masses and lure them to the movies.

Concluding Remarks

These two case studies show that the design and animation of digital characters has developed to a point that the results can be convincing. Approaches to solutions have been established for all three of the fundamental problems named: connection between appearance and behavior, complexity, and interaction. This finding corresponds to my remark above that complexity is always the result of a history. This history is also the history of computer graphics and animation, a still young discipline that is yet to develop its own cultural memory.

In any case the overblown rhetoric of progress, as celebrated above all and in high-profile fashion by James Cameron, is misplaced. Despite the increasingly complex manifestations that can be asserted of the characters in *BENJAMIN BUTTON* and *AVATAR*, all of the methods used draw on predecessors that had already been used years earlier, and most of which I had already encountered in my past research. Not least for this reason, the insights that I acquired there, as well as my analyses and models, have proven startlingly robust.

If one traces the individual technologies in detail, one sees a complex network of influences and recognitions both from academic basic research and from research and development departments of private companies. This network is becoming increasingly differentiated, but not by leaps and bounds. In direct exchange with the visual effects specialists and university researchers it becomes clear how painstaking and challenging this work is. Still the discipline remains in a stage of bricolage, in which many highly developed building blocks are available for use, but each individual solution must be struggled at. Thus the construction and animation of digital characters will remain one of the most involved, yet also one of the most exciting subject areas in computer-generated imagery.

References

Abramowitz, Rachel

2010 Avatar's Animated Acting. Hollywood Debates Whether Film Work Used by Actors Such as Zoe Saldana for Computer-Generated Characters Deserves Equal Recognition. *Los Angeles Times*, 18 February 2010.

Asendorpf, Jens B.

2004 *Psychologie der Persönlichkeit*. Berlin: Springer. 2005

Duncan, Jody

2009 The Unusual Birth of Benjamin Button. In: *Cinefex*, no. 116, pp. 70-99 and 118.

Duncan, Jody

2010 The Seduction of Reality. In: *Cinefex*, no. 120, pp. 68-146.

Ekman, Paul; Friesen, W.

1978 *Facial Action Coding System. A Technique for the Measurement of Facial Movement*. Palo Alto: Consulting Psychologists Press.

Flueckiger, Barbara

2008 *Visual Effects. Filmbilder aus dem Computer*. Marburg: Schüren.

Flueckiger, Barbara

2009 *Digital Reconstruction of Deceased Persons*, paper presented at the NECS conference in Lund, June 26, 2009.

Holben, Jay

2010 Mauro Fiore, ASC Helps James Cameron Envision Avatar, a 3-D Science-fiction Adventure That Combines High-definition Video and Motion Capture. In: *American Cinematographer*, Bd. 91, Nr. 1, January 2010, S. 32-47.

Kingslien, Ryan

2004 *ZBrush 2.0 Review*. (= <http://vfxworld.com/?atype=articles&id=2310>)

Luhmann, Niklas

1995 *Die Realität der Massenmedien*. Wiesbaden: VS, 2004, pp. 35 ff.

Mori, Masahiro

1970 Bukimi no tani. (The Uncanny Valley). In: *Energy*, no. 7.

Reinhard, Erik; Ward, Greg; Pattanaik, Sumanta; Debevec, Paul

2006 *High Dynamic Range Imaging. Acquisition, Display, and Image-based Lighting*. Amsterdam: Elsevier; Morgan Kaufmann.

Spiegel, Simon

2010 Wie James Camerons Avatar den Zuschauer in eine fremde Welt entführt. In: Mamczak, Sascha/Jeschke, Wolfgang (Eds.): *Das Science Fiction Jahr 2010*. Munich, pp. 361–388.

Thompson, Kristin

2010 *Motion-capturing an Oscar* <http://www.davidbordwell.net/blog/?p=7126>