Andrew Buchanan

Facial Expressions for Empathic Communication of Emotion in Animated Characters

Introduction

The challenge of communicating emotional content to an audience via animated characters has existed since the art form first appeared. As animation techniques and technology have advanced, animators and character designers find themselves with a multitude of resources and tools for the creation of facial expressions so as to effectively communicate the emotions of their animated characters within each scene. However, by using evolved forms of symbolic facial expression, which are widely accepted, these technologies and techniques often overlook the unconscious communication conveyed via actual human facial expressions. The fundamentals of this instant and unconscious emotional communication have been well studied and documented, yet the systems developed by the scientific community for reading and interpreting facial communication have only occasionally and recently been applied by animators.

As they relate to emotional experiences, facial expressions can be divided into two main types: spontaneous facial expressions; and deliberate facial expressions (Ekman & Rosenberg 1997). Animated characters traditionally use a third type of facial expression: symbolic or artistic facial expressions. Executions of animated facial expressions may have a natural tendency to fall into the third category, but certain types of emotional communication may be aided by attempts to include hallmarks of the spontaneous expression type, or a hybrid of these two.

Studies have suggested that the reading of facial expressions is an unconscious process and that our reactions to these facial expressions can be unconsciously generated (Murphy & Zajonc, 1993). This finding is supported by previous publications by Paul Ekman (2003) and Carl Jung (1974). This suggests that a deeper understanding of how to execute facial expressions which can be unconsciously read as true expressions of emotion may be a useful tool for effective communication of emotion when creating animated characters.

These symbolic expressions are effective at generating two of the three main responses to facial expressions – environmental expectation and behavioural expectation – which are described in the paper ‘Facial Expressions as modes of Action Readiness’ (Frijda & Tcherkassof 1997).

The third main type of response is the generation of an empathic response. The paper Components and Recognition of Facial Expression and Communication of Emotion by Actors (Gosselin, Kirouac, et al, 1997) has found that empathic emotional responses are generated more effectively through spontaneous facial expressions, than through symbolic or deliberate facial expressions.

Formal systems for the decoding of facial expressions established by other disciplines may offer animators new opportunities to draw upon methods outside their own practice to create facial expressions that communicate emotions effectively to the audience without any requirement that they look realistic. For example, the use of Ekman and Friessen’s Facial Action Coding System (FACS) in the creation of the character Gollum in Lord of the Rings – The Two Towers (2002), (Kerlow, 2004) resulted in a character performance which was widely regarded by critics as emotionally believable and well integrated with the action and other characters in the film.
Scope

It is necessary to note that the conveyance of emotion in any animated work is never restricted to facial expressions alone. The history of cinema lends to animation all of its methods for storytelling, including casting, dialogue, lighting, cinematography, body language, music, emoticons, physical changes, and special abilities or super powers. The consideration of different types of animated facial expressions for emotional communication in isolation from the other cinematic elements and animation techniques explained above would be erroneous – indeed, the specific objectives of a particular production may make complex facial animation superfluous.

The diversity of animation as an artistic, narrative and communication medium will require a range of approaches and methodologies – indeed by its very nature, animation thrives through its diversity of style and approach. It is not the contention of this paper to provide a universal approach to creating animated facial expressions. Specifically, the types of animation which may find some benefit from the findings of this paper are those which seek to induce a cathartic emotional response from their audience for the purpose of advancing a dramatic narrative. In making this distinction, it is recognised that, even within this category, there may be a multitude of reasons why approaches suggested within would be inappropriate for a given production.

The context in which a facial expression is displayed is particularly important in animation where every image, movement and action is crafted deliberately to contribute to the objectives of the sequence. Humans often display deliberate expressions which do not relate to our true feelings or we may use facial expressions to intentionally mask our emotions (Crag, Hyde, et. al., 1997). Society may dictate which expressions are appropriate at which times, and hence some of our emotions may not be represented physically (Fernandez-Dols & Ruiz Belda, 1997). This can be equally true for animated characters, and animators may also have their own set of rules for when and how to display certain facial expressions. Additionally, there may be stylistic or practical considerations of the art form such as double-takes, anticipation, overshoot and settle, and reducing expression changes during broad moves (Williams, 2001) which may influence the implementation of any findings within.

The Study of Facial Expression

In 1967 a study into the universality of facial gestures was undertaken by Paul Ekman and Wallace Friessen. The study involved tribes in Papua New Guinea who, because they were largely isolated from other groups, would tend to display only those expressions which were deemed to be automatic and common to all humans. The study focussed on how emotion is displayed on a person’s face, and the mechanisms humans have developed for interpreting these displays.

Ekman and Friessen explain in their book titled *Unmasking the Face* (2003) that facial expressions are one of the key indicators of human emotion and primary sources of identifying emotion in others: ‘We do know that the face is a primary, clear and precise signal system for the expression of the specific emotions’ (Ekman and Friessen 2003).

In his book, *Silent Messages: Implicit Communication of Emotions and Attitudes* (1981), Albert Mehrabian states that 55% of an emotion based message is communicated through what he terms ‘facial linking,’ which includes appearance, facial expression and body language. Only 7% of the message is communicated through the words chosen (verbal linking) and 38% is communicated by the way the voice is used (vocal linking).
The Facial Action Coding System (FACS) co-authored by Ekman and Friesen in 1978, is a system whereby the various actions of a face are broken down into their component movements, based on the specific facial musculature which stimulates that movement. These individual movements are known as Action Units (AUs).

By analysing the various AUs observed during analysis, a FACS practitioner can determine the emotions in evidence in that example. For example, the true ‘enjoyment’ smile indicating the happiness emotion can be identified by the presence of 2 major Action Units; AU 12, and AU 6. These codes represent the contraction of the zygomatic major muscle (which raises the corners of the lips and the cheeks), and contraction of the orbicularis oculi, pars lateralis (which squints the eyes and makes crows-feet wrinkles) (Ekman & Rosenberg, 1997).

Often, the smile we use and notice in others is not a spontaneous representation of the happiness emotion, but what can be called a ‘requested’ or ‘deliberate’ expression. The deliberate smile shows a contraction of the zygomatic major, but does not often include the contraction of the orbicularis oculi (Ekman & Friessen, 1978). This particular action unit is very difficult for humans to perform on demand. This deliberate smile is the one we use during situations where it appropriate to put someone at ease, to be pleasant or welcoming to others, during our polite interactions with strangers, or when smiling for a photograph but is not generated spontaneously by our emotion.

Charles Darwin, in his study The Expression of the Emotions in Man and Animals noted as early as 1872 that humans can automatically distinguish genuine expressions from deliberate ones. Referring to photographs in which people smile, he noted:

Almost everyone recognized that the one represented a true, and the other a false smile; but I have found it very difficult to decide in what the whole amount of difference consists. It has often struck me as a curious fact that so many shades of expression are instantly recognized without any conscious process of analysis on our part. (Darwin, 1998 [1872])

Our responses to the smiler are guided by our unconscious assessment of whether the smile is genuine, and by the insights this gives us into the motivations of the smiler.

The spontaneous and deliberate facial expressions in humans are neurologically as well as physically different as explained by Mark Frank (et al) in his paper ‘Behavioural Markers and Recognisability of the Smile of Enjoyment’:

not only do emotional and nonemotional facial activity originate from different parts of the brain (subcortical and cortical motor strip, respectively) and arrive at the face through different motor systems (extrapyramidal and pyramidal, respectively) but also the appearances of these actions differ. (Frank, Ekman et al. 1997)

The importance of this for any animator seeking to utilise this system for purposes of encoding an emotion through a facial expression (rather than decoding, as was the original intent of the FACS), is to recognise that the hallmarks of a spontaneous expression differ from those of a deliberate expression.

Although we may like to believe we are in control of our perceptions of our environments, our interactions and the way in which we know and react to our world, in reality we are aided by complex brain functions which automatically sort and assess information gained through our senses, and provide us with appropriate deductions and automatic responses. These can include changes in physiology (as in the fight/flight response, caused by an automatic sense of danger), as well as supplying data for the conscious brain to process. In his book Emotions Revealed (2003), Paul Ekman uses the term ‘autoappraisers’ to describe the mechanisms at work in the human
brain which constantly scan our surroundings and detect when something important to our welfare is occurring (Ekman, 2003). He notes that through the use of these autoappraisers ‘we can make very complex evaluations very quickly, in milliseconds, without being aware of the evaluative process’ (Ekman, 2003). This process of automatic appraisal is what generates ‘spontaneous’ facial expressions – those linked closely to the experience of the emotion they represent. By contrast, deliberate facial expressions, or those which are culturally or personally moderated are generated by ‘reflective appraisal’ (Ekman 2003), requiring some conscious input from a person to generate the expression.

The ability to read different expressions displayed by others, and our understanding of how those expressions relate to the emotions they portray is also an unconscious process (Murphy & Zajonc, 1993). As humans, we have a hard wired ability to decode the emotions of other people, developed over the course of our evolution. The psychologist Carl Jung, in his 1968 book titled *Man and His Symbols*, notes that ‘universally understood gestures and many attitudes follow a pattern that was established long before man developed a reflective consciousness’ (Jung, 1968).

Those animated characters displaying expressions which we unconsciously recognise as spontaneous and emotion driven expressions may have a greater emotional resonance with an audience than those who display expressions which are unconsciously read by the audience as deliberate expressions.

Animators working on sequences, films or other projects which seek to elicit emotional responses and effective emotional communication with their audience may consider the implications of the use of expressions which appear spontaneous as opposed to those which may appear deliberate.

Facial expressions can generate three main responses in the viewer: environmental expectation – such as an expression of fear on another person’s face warning us of immediate danger; behavioural expectation – such as another’s angry expression warning us to expect conflict from that person; and, empathic response – such as a feeling of sadness when witnessing another’s grief (Frijda & Tcherkassof, 1997). Traditional forms of animated facial expressions will likely be effective in conveying both environmental and behavioural expectation to an audience, however these evolved techniques may be less effective than spontaneous human expressions at generating an empathic response as described by Frijda and Tcherkassof.

It is necessary therefore, to review some of the historically evolved techniques, and highlight innovations which offer methodologies and workflows to support the creation of animated facial expressions.

**The Evolution of Animated Facial Expressions**

In the early days of The Walt Disney Company, one of the pioneer organisations of the popularised animated medium, animators recognised that facial expressions could communicate the inner thoughts of a character to the audience. Frank Thomas and Ollie Johnston comment in their book *The Illusion of Life: Disney Animation* that ‘Through a change of expression, the thought process was shown’ (Thomas & Johnston, 1981).

Early animated films were a purely visual medium accompanied by a live music performance and therefore had a high degree of physical humour and visual action. The facial expressions of this early period were understandably simple, as they sought clarity in the communication of a narrative for which there was neither an accompanying sound effects track nor dialogue. The
simplified facial expressions these characters bore were initially based on ‘photostats’ of actors deliberately making each expression, and were also influenced by a concurrent history of cinema, in which the trend of the day was toward what might now be considered ‘over-acting.’

Commenting on the short film Playful Pluto (1934) involving Pluto the dog wrestling with a piece of fly-paper, the animators at The Walt Disney Company noted that expressions played a very important part in the entertainment value of the scenes, and while everyone admitted that this was only a broad cartoon symbol for a dog and lacked any attempt at realism, it was still felt that the door had been opened. (Thomas & Johnston, 1981)

The sequence in Playful Pluto clearly shows Pluto experiencing and expressing human-like emotions through a series of exaggerated and simplified expressions that represent those emotions.

This evolving set of expressions and (quite separate and often very different from the human ‘spontaneous’ and ‘requested’ expressions) can be called the ‘artistic’ or ‘symbolic’ representation of a character’s emotion. Jose Miguel Fernandez-Dols notes in his paper ‘Spontaneous Facial behaviour During Intense Emotional Episodes: Artistic Truth and Optical Truth’ that if a painter, actor or layperson sets out to convey happiness or anger [...] then a smiling or frowning face is the right image to choose. In the absence of words, context or further explanation, a smiling face conveys ‘a happy person,’ just as a cartoon mouse is successful in conveying ‘mouse.’ (Fernandez-Dols, Ruiz Belda, 1997)

This view suggests that the expression symbols for emotion, while rarely an accurate portrayal of the emotion as displayed by humans, have become accepted substitutes for communicating the idea of these emotions. In particular, animation from Japan has a history of using generic and exaggerated facial symbols to convey the emotion of a character.

The symbolic expression informs us of the emotional state of the character through our referential knowledge, generated by previous viewings of animated productions. This may however be less effective than the use of spontaneous expression types in generating an empathic response, allowing the audience to cathartically experience the immediate emotional conditions of the characters.

In 1937, The Walt Disney Company released the first full length animated feature film, Snow White and the Seven Dwarfs. This release, with its extended length, complex narrative and dramatic themes required animated characters to deliver some form of emotional fulfilment through the viewing of the film. The expectation that certain types of dramatic animated films can and should deliver this type of experience has continued, and today the emotional content of these particular animated productions is often critically considered with much the same weight as it is in live action films.

Actors in live-action productions are required to portray the emotions of their characters through convincing facial expressions as well as other body language and dialogue. Animators are tasked with a similar job, but while actors have some advantage due to an innate ability to generate facial expressions through the previous experience of the emotions with which they are associated, animators are required to manually construct expressions for their characters.

The traditional methods for achieving this include the use of personal mirrors, so that animators can examine their own faces while posing expressions, and careful analysis and research of human and animal facial expression and movement. These techniques are coupled with a historical emphasis on animators understanding of anatomy and years of drafting and drawing practise.
Animators working on *Snow White and the Seven Dwarfs* discovered that it would be much easier to use images or footage of facial expressions as reference material for their animation, rather than attempting to completely invent a character’s facial expressions. They used Photostats (an image printed from a single frame of moving 35mm film) of actors’ faces portraying different expressions. Initially, these images were traced directly, in an attempt to create an animated version of the human expression. This process is known as rotoscoping, a technique which is still in use in some forms today. The animators found however, that the rotoscoped animation often looked strange, mechanical and inhuman. Without realising it, they were perhaps experiencing the Uncanny Valley – a phenomenon of aversion to not-quite-human entities first asserted by the Japanese roboticist Masahiro Mori in 1970. The animators found that their animations were improved when these expressions were exaggerated, and the actions produced using the cartoon movements they were familiar with (Thomas & Johnston, 1981).

One of the chief problems inherent in this process is that it relies heavily on the actor to be able to produce realistic expressions of emotion on demand. One method actors have for creating spontaneous, emotion driven performances is to use the technique called Method Acting. Put very simply, this process involves an actor attempting to re-live events in their past which had previously generated the emotions required. The actor can then use the generated emotion to naturally and unconsciously produce spontaneous facial expressions and other physiological changes, such as crying, changes in the voice and unconscious changes in the actor’s body language (Gosselin, Kirouac et al., 1997). This process is also known as the *Stanislavski System*, after Konstantin Stanislavski, who developed the technique in the early 1900s.

In an interview contained in *Behind the Scenes – Making Nemo* (Disney Pixar, 2003), Mark Walsh, the animator of the fish character Dory for the film *Finding Nemo* (2003) describes utilising a form of method acting. When attempting to animate a particularly emotional scene, Walsh explains that he was aided by attempting to re-live similar experiences in his own life, and then filming himself performing the lines of the character while experiencing that emotional state, thus generating reference footage for the sequence. This technique is a combination of the traditional use of mirrors by animators, combined with method acting. Film critic Mark Caro of the Chicago Tribune notes; ‘You connect to these sea creatures as you rarely do with humans in big-screen adventures’ (2007).

A study, published in 1997 titled *Components and Recognition of Facial Expression in the Communication of Emotion by Actors* (Gosselin, Kirouac et al. 1997) has shown that actors employing the *Stanislavski System* of method acting are able to generate more emotionally believable facial expressions than straight acting. Conversely, facial expressions which appear to be spontaneously, rather than deliberately generated would imply genuine emotions. Studies show that spontaneous facial expressions arising from genuinely felt emotion can be more effective at eliciting an empathic response in an audience, particularly for some emotions (Gosselin, Kirouac et al. 1997).

The facial design of an animated character may in some ways determine the characters ability to create facial expressions which appear emotionally spontaneous. For example, the character Winnie the Pooh (created by The Walt Disney Company) which has neither upper nor lower eyelids is unable to produce some of the FACS Action Units, making some of the expressions described therein impossible. This character may be restricted to a symbolic set of facial expressions, and thus the animator may be limited in their approaches to communicating emotion to the audience. In contrast, Seth McFarlane’s character designs for the television show *Family
Guy have both upper and lower eyelids, a design which increases the potential for the animator to create expressions which appear to be spontaneously emotion driven, despite the characters otherwise simplified facial design.

Technological advances and changes to animation production techniques have influenced the way in which many animated facial expressions are created. Particularly in what is commonly referred to as 3D (3 dimensional) animation or CG (computer generated) animation, practitioners are offered a vast set of tools for creating and animating faces. One of the most prevalent of these new techniques is called combination sculpting, and involves the animator setting up a number of facial poses which are then implemented to varying degrees over time. This is one area of animation where the use of tools such as FACS has been utilised with some success, aiding the character artist or animator in some cases to create poses based on specific muscle movements and combinations rather than traditional reference images or mirrors.

Another technique gaining popularity in animation production is Motion Capture (or Performance Capture). In this process, an actor is rigged with reference points – usually stick-on dots – which are then tracked by a digital system to record their movement. This movement is then translated to corresponding points on an animated character automatically. This process reduces the time taken to animate a sequence, and provides instant feedback to the director for review and re-taking.

A point for consideration here may be that while some of these methods can assist in the creation of characters which have a more life-like appearance, or in making the animation process easier, the animator (or actor) must still have a solid understanding of which facial shapes to make, when to make them and how to integrate them with the rest of the character’s performance.

Isaac Kerlow in his paper titled ‘Creative Human Character Animation: The Incredibles vs. The Polar Express’ (2004) notes the success of the integration of the FACS into the production pipeline of the character Gollum in the film The Lord of the Rings: The Two Towers. The animation team used reference footage of the actor, Andy Serkis, for the motion and facial expressions of Gollum. By having the reference actor act with the other characters in each scene, the actor was given valuable context for his performance. The animators of Gollum did not use any performance capturing devices for the face of the character, but chose instead to animate the character directly over the reference footage – a process they referred to as roto-animation – and deviate from this footage when it was felt appropriate. The animators used the process of combination sculpting in order to efficiently pose the expressions of the character, with the various facial poses designed by Bay Raitt (the lead character designer on the project) corresponding to Ekman and Friessen’s Action Units, from the Facial Action Coding System.

Another approach to integrating the FACS into technology based pipelines is described by Parag Havaldar in his paper entitled ‘Performance Driven Facial Animation,’ where he describes in detail the process of digitally ‘re-targeting’ facial expressions captured with a motion capture system to match pre-set Action Units which had been set up specifically for the cartoon-like facial shapes of the character. By pre-designing the facial poses based on the FACS, the production team on Monster House (2006) was able to capitalise on the speed of motion capture animation in the production, while maintaining facial expressions which corresponded with the appropriate Action Units.
Some previous animated features which have used motion capture technology – notably *The Polar Express* (2004) – have been associated with the notion of the ‘Uncanny Valley,’ first explained by Japanese roboticist Masahiro Mori in 1970 (Kerlow, 2004). The phrase describes the phenomenon of decreased emotional response to characters (Mori was initially referring to robots) as they tend towards a close resemblance of humans. Initially, our response is positive – we react well to objects or characters which resemble humans. As this resemblance increases, however, there is a marked drop in our response to these characters, before climbing again as the character becomes completely human (MacDorman, 2005).

The implications for the animator may be that complete human accuracy in facial expression may not achieve the desired empathic response; rather, this may be a source of revulsion in the viewer. Characters which are deliberately non-human or non realistic in their appearance may however avoid the uncanny valley, as with the characters in *The Incredibles* (Disney Pixar, 2004), while still benefiting from facial expressions which are unconsciously read as spontaneous and emotion driven by the audience.

**Discussion**

The use of symbolic or artistic expressions may continue to be more practical for many animation productions. Symbolic expressions have the advantage of being unambiguous in their meaning, and are free from cultural influence beyond the established culture of animation. It is also possible that these productions can benefit in terms of their emotional communication by utilising the established context of the animation art form and relying on the referential knowledge of the audiences previous viewing experiences within this framework.

Though all animated facial expressions could be called deliberate due to the process of animation, animated characters may benefit from the inclusion of both deliberate expressions as described in this article. The audience will read these automatically and understand the motivations for these expressions as long as the context for them is maintained and suitable. Deliberate expressions can give insight into the motivations of a character and deliver a greater understanding of inter-character relationships and scene contexts.

The inclusion of facial expressions in animated characters which look spontaneous may help to establish an unconscious communication of emotion with the audience based not on our referential knowledge of animation, but on our unconscious understanding of non-verbal communication with other humans. This is likely to be most effective where a cathartic experience of emotion in the audience can be managed with the other requirements of the production, or where this forms one of the key narrative techniques for the project.

Whichever facial action design is chosen for an animated production, and regardless of the production method employed, this choice should be a conscious one, and not established through artistic instinct. The methods chosen by the production team will influence the design of the characters which must produce the expressions.

The example of the character Gollum as an animated character able to communicate emotion through facial expression suggests that an amalgamation of various techniques and processes may yield effective results. In this case, the Facial Action Coding System is used as a reference and as a safety net for creating appropriate expressions. Like the use of 3d animation tools and advanced technology in animation, the application of findings from the Behavioural Sciences must be an integrative process, adding to, rather than superseding, the established body of animation knowledge and practice.
References


© Andrew Buchanan

Edited by Nichola Dobson

83