Facial Expression Coding System for Stylized Characters

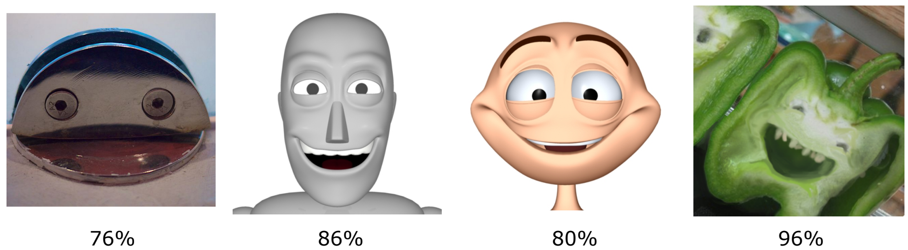


Figure 1: *Several example images that were designed to trigger joy using our system, and the resulted MTurk scores (in percents) obtained on these images.*

Abstract

This research is motivated by a desire to create an underlying structure or language for use in the design and motion of facial expressions for stylized characters. We demonstrate a clear set of facial component movements and orientations called "triggers" that show illustrators, designers and animators an efficient and clear way of creating initial facial expressions for fast prototyping and easy editing, e.g. in early story ideas or "roughs". Specifically, we suggest triggers for six cardinal facial expressions which cover the gamut of human expressions that are typically needed for animatics, the moving version of 2D or 3D storyboards used in the beginning of the production pipeline for the development of animated films and games. To determine the facial triggers we use Amazon's Mechanical Turk; we iteratively change the facial components to achieve high confidence that the expression will evoke human emotion and convey the intended communication. The final triggers allow us to map expressions universally onto highly stylized characters, abstracted faces, and even onto inanimate objects.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation;

Keywords: Stylized characters, animation, facial expressions, mechanical turk

1 Introduction

It is critical when communicating human emotions visually to be completely sure that the expression presented is being received without any ambiguity. The importance of creating recognizable expressions cannot be understated as almost all storytelling media rely on evoking a predictable, clear response from the audience. When designing complex personalities for digitally animated characters this becomes even more important as audiences learn to read the characters and are encouraged to suspend disbelief. The human ability to recognize and interpret facial expression in stylized characters as emotion is one of our most sophisticated and innate perceptual skills. In this research, we propose the underlying visual rules that govern recognition of facial expression in stylized characters (see the teaser for illustration of joy and the corresponding human study scores).

Understanding facial expressions in stylized characters and creating a semantic vocabulary for visual communication would contribute significantly to both the efficiency and effectiveness of any production team. Illustrators, animators, character designers for games or film, and storytellers in any medium could benefit enormously from understanding and applying these findings. No current methodology exists that allows the animator to configure a characters expression and feel complete confidence that the expression will convey. This is currently an artistic process that requires trial and error, is time consuming and ultimately can be very frustrating. It is both a difficult technical problem and a visual and storytelling challenge.

As a result of this work, we have found that particular facial patterns trigger recognition that an expression is present. There is vast untapped potential for research that focuses directly on the design of character rigs and set up that incorporates an underlying understanding of human facial triggers interpreted using the more stylized coarse mesh that can be built directly into the technical pipeline. Ultimately it is the pattern-recognition that animators need to know when they are working with stylized characters since a measurement that might work for a database-derived expression on a real face is useless for a dragon or a Mr. Potato Head, whose faces do not correspond to population-averaged norms.

These triggers are roughly divided into the upper and lower parts of the face and include the eyebrows and their surrounding skin; the position and shape of the exposed iris; the shape of surrounding eye white; the position and orientation of the nose and the skin around it; the exposure of the teeth; and the shape and openness of the lips/mouth. In this work, we focus on communication of cardinal facial expressions; researching communication of more subtle expressions is a subject of future work and beyond the scope of this paper.

In particular, we have followed Ekman's classical human facial expression research [Ekman 1999] in confining ourselves to six basic categories of emotion: fear, sadness, joy, anger, surprise, and disgust. We include neutral as an illustration of the character showing no expression. We show that, using our rules, we can produce stylized facial expressions for each of the categories that can achieve as high as 96% unanimity with our randomized audience–an impressive demonstration of the clarity of well-chosen expressive elements. This is similar to Ekmans work for human facial expression which shows that strong viewer agreement is possible for the clearest example of faces in each emotion category (after correcting for any contextual or posture information).

Stylized and simplified faces offer great advantages for studying the fundamental elements of the faces emotional language. We are able to reduce the variables to their minimal necessary components. This

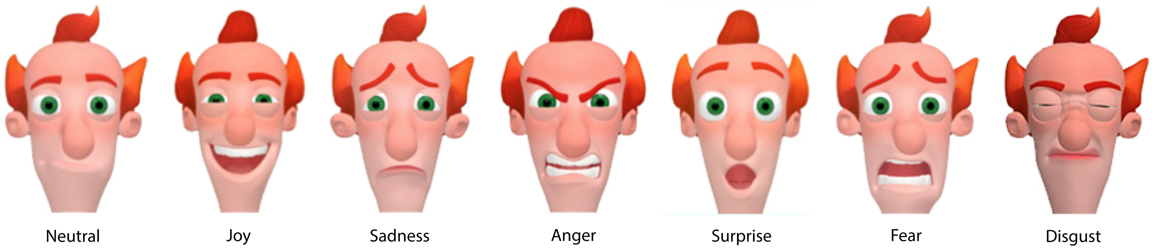


Figure 2: *Cardinal expressions for a stylized character.*

is impossible with realistic faces. Using our approach, we can mix and match these components in limitless combination while being confident that our observers will respond as though the faces were motivated by real, discernible feelings. Our rules were designed iteratively by refining stylized faces based on crowd-sourcing feedback from Mechanical Turk eventually achieving between 82 and 96% MTurk viewer agreement. As in [Lasseter 1987] where Lasseter described the basic rules of animation, we describe basic rules for conveying cardinal emotions for stylized characters. The rules that have emerged from our testing can be successfully repeated as we demonstrate in the paper.

While human facial expression has been researched extensively **[?][?] [?][?][?]**, there are only a few published papers that consider stylized facial expressions. [Tinwell et al. 2011] is the closest paper to our research and it does not include the signals used by the observers to recognize expressions and does not cover extreme stylization. [Wallraven et al. 2006] shows the effectiveness of stylized faces that are more mechanical in their production, without involvement of artists or improvement process. Tests were run to compare different stylizations of the same expression, and participants were asked to express preferences, to choose the expression portrayed from a list and to rate expressions by intensity, similar to the study described here. The paper results recognize both the existence, and the efficacy, of the coarse grid however it appears that artistically created stylized faces are not a focus. In somewhat related work [Gooch et al. 2004] method for stylizing facial expressions of humans was proposed and it was shown that communication of stylized facial expression faster than realistic expression. Finally, in [DeCarlo and Santella 2002] it is shown how to use a model of human perception for the stylization of photographs as a new possibility for non-photorealistic rendering.

2 Contributions

The contributions of our paper include:

- Creating facial patterns (based on refinements suggested through repeated real-world tests) which provoked a consistent viewer emotional reading, across a wide range of facial types and styles, (stylized realistic and less realistic), and humanoid (more or less) resulting in a reliable universal trigger analysis.
- Identification of three separate variables that greatly influence how humans respond to cardinal expressions–elements of facial expressions that *significantly increased* readability (constructive variables), those that *decreased* readability (deconstructive variables) and those that could be changed without changing the result (non-functioning variables).
- We found that fear and disgust presented different challenges apart from joy, sadness, anger and surprise. Viewers seemed to assume one of the former four emotions were present by default, unless disgust or fear facial cues were totally unambiguous. With both it took as many as forty iterations to achieve viewer consensus. Nothing similar occurred when testing joy, sadness, anger and surprise. Smiles, besides allowing an almost unlimited amount of variability in the eyebrow and forehead configuration without losing viewer agreement, were also found to be the expression most likely to be seen in non-human faces that we tested.

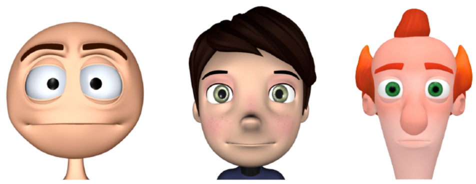
3 Experimental Setup

Below is a description of the research that has been accomplished using several different characters in several different styles. All have been tested and revised based on the results gathered from our work with Mechanical Turk. In a feedback loop, we improve the accuracy of facial expressions by making quick alterations based on viewer response. As a result, we have been able to design expressions that consistently score in the 85-95% range in multiple-choice recognition tests. According to a recent study, Mechanical Turk participants are at minimum as representative of the U.S. population as those from conventional subject pools. We carefully regulated the testing process in order to obtain valid and reliable results. Participants in our study were required to have a prior 95% or greater successful test completion rate, as controlled by Mechanical Turk.

Participants evaluated a given facial expression by choosing one answer in a forced-choice questionnaire. The response options were restricted to the six primary emotions as illustrated in Fig. 2 (joy, sadness, anger, surprise, fear, disgust) as well as the choice of none. Due to the possibility of forced-choice bias [rus], we verified the reliability of the test by conducting several free-listing trials. In these tests participants filled in a box with one word to describe best the emotion of the presented expression. These individuals were not given any response possibilities or suggestions. We found no significant difference ($p > 0.05$) between the results of forced-choice and free-listing test formats. The number of participants for each trial was 50. When we replaced our multiple-choice interface with fill-in-the-blanks, we also retained a very similar consensus, based on our reading of which terms could be considered synonymous (e.g. fear=worry=anxiety).

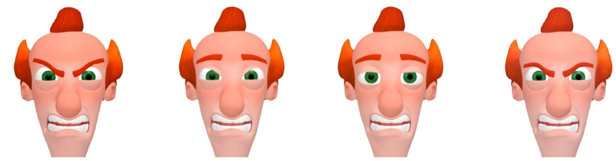
4 Facial Triggers

One great advantage of using stylized faces is the ability to create configurations and combinations that would be impossible with an actor or manipulated realistic face. By changing which facial regions were active and the degree of activity, we are able to determine key variables with high degree of precision. These triggers, once identified, can be utilized across many degrees of simplification and stylization, as long as the fundamentals of the underlying patterns are understood.



Expression	%		
Joy	12	6	6
Sadness	2	14	14
Anger	2	2	2
Surprise	18	16	16
Fear	10	16	16
Disgust	10	2	2
No Emotion	46	44	44

Figure 3: *Responses to ”neutral” faces. We considered our neutral faces successful when over 40% of the respondents chose ”No Emotion”, and the other choices were more or less evenly scattered, as here. Here the neutral pattern scattering is quite similar between one face and the next, in spite of the dramatic differences in the character design.*



Expression	%			
Joy	0	2	2	0
Sadness	0	12	36	0
Anger	94	14	6	80
Surprise	2	10	2	4
Fear	4	44	28	0
Disgust	0	18	18	16
No Emotion	0	0	8	0

Figure 4: *Dominance of eyebrow cue in anger. The angry mouth is not perceived by viewers as indicating anger if the crucial cue of the lowered, angled eyebrow is missing. Here a face seen as angry loses that effect almost entirely when the eyebrows are level. The widened, occluded iris eye on the middle figure is not nearly as strong as a cue (28%); when that is neutralized the mouth (which has stayed the same) has no weight as a cue on its own.*

Stylized, artist-drawn faces can be extremely effective in portraying basic human emotions. In our tests, we were able to find feature patterns that were recognized as the six cardinal expressions by 85-95% of our viewers, an extremely high rate of agreement. We found that the recognition rate did not change when applied to very different stylized characters; it was independent of changes in texture, proportion, level of detail, or level of similarity to the real human face.

We have chosen to separate our results into two parts and left fear and disgust to the end of this section. Fear and disgust presented us with our biggest challenges, requiring an enormous amount of trial and error until we were able to craft stylized faces which achieved more than 80% viewer agreement, which we would not have been able to do without the quick feedback loop provided by our testing method.

Perhaps one reason for the difficulty is that both are rarely encountered in everyday life, in contrast to joy, sadness, and anger. Surprise, while in theory only visible in an instant of being caught unawares, is commonly mimed to accompany a conversational point.

4.1 Neutral

To create neutral expression we used line and 3D renderings; the results for both styles were very similar. We also found that there is very little user agreement on stylized faces with the features designed to mimic a human face at rest. We considered our neutral tests successful when 40% of our viewers saw the face as No Emotion” and the other responses were randomly scattered across the emotive spectrum. See examples in Fig. 3.

Stylized faces tend to be seen as projecting a persona no matter what expression cues are present, perhaps because of the way they exaggerate and even satirize characteristics of realistic faces (i.e. funny noses, wild hair, robotic eyes). This makes the unanimity of responses to our successful expressive faces even more impressive, given the complete scattering of viewer reactions to the same faces at rest. Further tests are needed to determine if stylized faces, because of their exaggerated and even comic character, have a much stronger tendency to be seen as having some expression (even when none is present) than the equivalent realistic human face. Many cartoon characters display what we term ”hyper-alert” eyes, where the at-rest eye displays large amounts of white above the iris. In real faces, such upper eye white display only appears in expressions like fear and surprise. As can be seen from Fig. 3 the hyper-alert eyes on the left did not strike our respondents as more surprised or afraid than the more realistic eyes in the center.

4.2 Anger

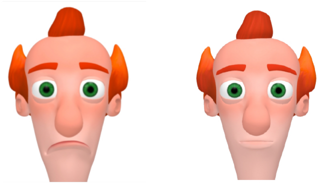
The key triggers for anger are the lowered, angled brow; the widened, tensed eye; and the snarling or tightened mouth. Every primary expression has a particular weighting of the features in terms of where the viewer looks the most for their information. This is particularly striking with anger, where the impact of the widened eye/lower brow pattern that is the key trigger for the expression creates a strong anger cue that will allow a wide range of both open and closed mouth shapes to appear angry. But the same angry mouth that appears in a face with 94% viewer agreement no longer strikes viewers as angry when the eyebrows are leveled and the anger score plummets to 14%; relaxing the upper lids reduces the anger signal to virtually zero (6%). On the other hand, a completely inactive mouth does not prevent a face with lowered brows alone as registering a 72% anger score. See Fig. 4 for illustrations of that.

4.3 Sadness

Sadness was a consistent alternative reading with most of our less successful frightened or disgusted faces, rarely completely absent and consistently chosen by 20-30% of respondents. However, when the eyebrows alone were moved a character whose relaxed face had a 20% sad score, was perceived as sad by nearly 90% of our respondents. The sad mouth is a strong cue on its own, provoking 62% of our respondents to see an otherwise neutral face with a sad mouth as sad. Fig. 5 illustrates these ideas.

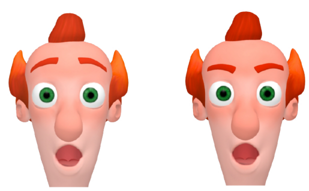
4.4 Surprise

Surprise is another of the robust expressions, and is perhaps the simplest of all expressions to depict. Surprise requires three elements to which respondents assigned equal weight: raised, level eyebrows; widened eyes; and opened, relaxed mouth. Changing any one of the 3 seriously eroded the results even with the other two held constant. Since all three surprise cues are easily expressed in a stylized



Expression	%	
Joy	0	2
Sadness	62	30
Anger	8	4
Surprise	6	12
Fear	16	12
Disgust	8	2
No Emotion	0	38

Figure 5: *Mouth triggers in sadness. We are easily persuaded to see the face as sad. Unlike in anger where the angry mouth on its own carries almost no perceptual weight, sadness can be portrayed with just the downturned mouth and a neutral upper face, as here. However the 62% sadness response is much lower than the 85% sadness response we obtained when we tested a face with the sad pattern brow combined with a neutral mouth.*



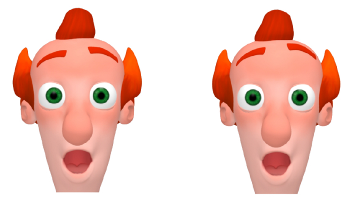
Expression	%	
Joy	2	0
Sadness	2	0
Anger	2	36
Surprise	90	44
Fear	2	18
Disgust	0	0
No Emotion	2	2

Figure 6: *The mouth shape and eyebrows are critical to convey Surprise.*

language, extremely simplified characters can be perceived as looking surprised by a strong plurality of viewers, in a typical instance achieving 94% agreement. We found that viewers were sensitive to any changes of the mouth shape from the simple oval of the opened, unstressed mouth, and any suggestion of mouth tension tended to imply the presence of fear, though always expressed as a minority opinion. We also found there is an optimal height for eyebrow lift and results improve as the brow raises to that level; beyond that point extra lift has no effect on the result. Figures 6, 7, 8 illustrate these ideas.

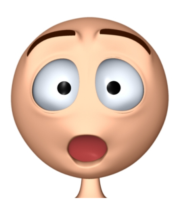
4.5 Joy

The key facial triggers are the tensing of the lower eyelid, the puffing of the cheeks, and the stretching and arching of the lips. We are pre-disposed to read joy in the open-mouthed smiling face, whether or not the upper face is participatory. The highest user agreement is achieved with upper/lower face agreement, but the benefit of a ”correct” face as here compared to a less correct face is relatively small. The smile is one of the most robust expressions we tested, clearly legible in an enormous variety of configurations as long as the bow-shaped, open mouth is present. Lacking the same obvious cues, the closed-mouth smile is harder to pose effectively; the +90%



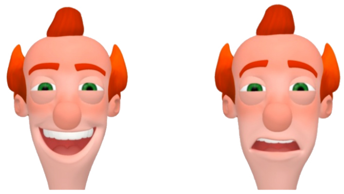
Expression	%	
Joy	4	10
Sadness	0	4
Anger	4	4
Surprise	66	64
Fear	18	10
Disgust	4	4
No Emotion	4	4

Figure 7: *Non-functional feature movements in surprise. We found there is an optimal height for eyebrow lift, and results improve as the brow raises to that level. Beyond that point extra lift does not improve the recognition of the smile, and even (as here) can erode the clarity of the results.*



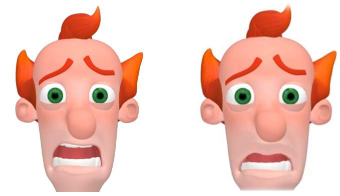
Joy	2
Sadness	0
Anger	0
Surprise	94
Fear	2
Disgust	2
No Emotion	0

Figure 8: *Surprise successfully conveyed by simplified character. Surprise requires three elements to which our respondents assigned equal weight: raised, level eyebrows; widened eyes; and opened, relaxed mouth. Changing any one of the 3 seriously eroded the results even with the other 2 held constant. Since all three surprise cues are easily expressed in a stylized language, extremely simplified characters can be perceived as looking surprised by a strong plurality of viewers, in this instance achieving 94% agreement.*



Expression	%	
Joy	94	0
Sadness	0	66
Anger	0	4
Surprise	6	6
Fear	0	10
Disgust	0	12
No Emotion	0	2

Figure 9: *The smile cue is the most powerful we tested. The open-mouthed smile cannot be confused with any other facial cue, and if it is present the overwhelming number of respondents saw it as representing joy. Although the face on the right is quite similar in the upper face, none of those cues have any weight, and with the upper teeth still exposed (one of the mouth cues in smiles) but the all-important bowed lower lip missing, not a single respondent detected joy. The flip from 94% to 0% was the most dramatic reversal in all of our single-cue tests.*



Expression	%	
Joy	0	0
Sadness	4	28
Anger	2	8
Surprise	2	4
Fear	76	48
Disgust	16	12
No Emotion	0	0

Figure 10: *Subtle cues in fear. Respondents can easily choose sadness rather than fear with the distressed brow. Although we had originally thought that wide opened eyes, never present in human sadness, would be enough to prevent confusion, in our tests the actual cues that separated the two expressions proved much more elusive. Here the only apparent difference is the less opened mouth and the less stressed upper lip, but 24% more respondents thought the face on the right was more sad than fearful.*

closed-mouth smile is likely based on cues we have yet to discover and test. The bullet-proof nature of the smile is no small reason such a variety of creatures and characters have been successfully animated, and elsewhere we demonstrate effective found smiles on objects as various as window hinges and electric meters; all that is required is two dots and an arc. Smiles are the only expression where the eyebrows play no direct role at all; even their total absence makes no difference in the results. Likewise the eyes being extra-opened does not diminish the recognition rate of the smile, a fact that animators have also used in creating so many characters with impossibly widened eyes. Joy is not visible in the upper face in even a fractional way once the obvious smile configuration disappears. See Fig. 9 for illustration.

4.6 Fear and Disgust: A Special Case

Fear and disgust are easily confused with other expressions; sadness and anger in the case of disgust, sadness and surprise in the case of fear. Respondents had a bias in favor of anger, sadness, and surprise, and only accepted faces as afraid or disgusted if the cues were unambiguous and correct within a very narrow range.

4.6.1 Fear

Fear is expressed in three main cues: widened eyes; distressed brow; stretched, tensed, and (usually) opened mouth. We were able to gradually improve viewer response as we increased any one of the elements after the other two were already present. We found no cases where two of the elements alone in any combination were effective, and certainly no single element will work. If there is no distress in the brow or the mouth, for example, no fear is detected. As soon as the brows begin to angle upwards and the mouth is tensed, fear response goes from 2% to 30%. As the mouth widens, while tensed at the level of the lower lip, the fear response increases. Finally, adding more distinctly distressed brows and a much more horizontally stretched lower lip more than doubles the fear perception. The relaxed/tense mouth shape is an important deciding factor in whether surprise (relaxed) or fear (tense) will be seen in the face. Unlike disgust, the stretched mouth of fear seems to offer more possible working configurations. One of our less-successful fear models displays very little change in result when the mouth is stretched more widely, the 4% uptick too small to be useful. See Fig. 10 for examples.

4.6.2 Disgust

In addition, we designed faces to determine the role of the mouth in fear and disgust and found that viewers rejected all but the most successful mouth shapes. Characters expressions tested as much as 80% different when the mouth shape was altered in certain ways, even when the rest of the face was left unchanged and the other triggers were the same. This was a surprising discovery as much conventional wisdom suggests that the configuration of the eyes convey the most significant expressive information, not the mouth. We also determined that identifying the non-performing variables for fear and disgust was far more difficult than for the other four expressions because viewers seemed to take crucial cues from every part of the face, leaving none to alter.

A very different condition arises with the expression of disgust, where the weighting of the recognition is dependent on regions around the nose that has no bearing at all on the other 5 major expressions so much so that on our character with no nose at all, all of the other 5 expressions were easily recognized with 70% agreement. In disgust, by contrast, having a character design where the nose tip and cheek area around the nose is inflexible makes it impossible to achieve the sort of strong viewer agreement achievable in every other case. The weighting of the nose region in disgust is so high, in fact, that a strong disgust vote was discernible even on a face (2D at this point) where the mouth was borrowed from a neutral face. See Fig. 11 and Fig. 12.

4.7 Stylized triggers: Some Considerations

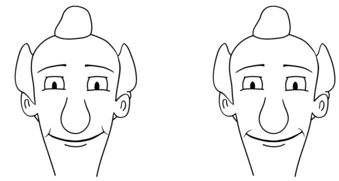
Anatomical similarity to the human face was only a very rough guide to which configurations might be the most effective. This is the clearest in the case of the hyper-alert eye. In the actual human eye, very few people show any degree of eye white (sclera) above the iris when their face is at rest; even when that is the case, the amount displayed is very slight. Exposed upper eye white, in fact, is a trigger for shock, surprise, or fear with the real human face.

Expression	%		
Joy	2	2	12
Sadness	10	6	2
Anger	8	4	14
Surprise	0	0	2
Fear	4	0	2
Disgust	74	84	62
No Emotion	2	4	6

Figure 11: *Full face participation in disgust. Disgust is unique in involving every area of the face, including the nose and middle cheeks. The right hand figure demonstrates the importance of nose wrinkling and cheek creasing; with the mouth neutral it still scored 62%.*

Expression	%	
Joy	2	0
Sadness	6	84
Anger	4	2
Surprise	0	0
Fear	0	8
Disgust	84	4
No Emotion	4	2

Figure 12: *Volatility of the expression of disgust. Anger, sadness, and joy can still be recognized by many viewers of only one cue is present. This is decidedly not true with disgust, where we found many cases where changing one cue - in this case the mouth - makes the disgust cues in the other parts of the face virtually unreadable. The pose on the left was our most successful attempt to achieve high agreement on disgust, scoring 84%, but when one of our artists replaced the mouth with one with ambiguous cues, 84% of our respondents saw the same face as sad.*



Expression	%	
Joy	76	58
Sadness	2	2
Anger	0	6
Surprise	16	14
Fear	0	2
Disgust	0	2
No Emotion	6	16

Figure 13: *Suggesting joy cues with line art. We found that line art can create significant cues with a much smaller amount of visual information than fully rendered, 3D art.*

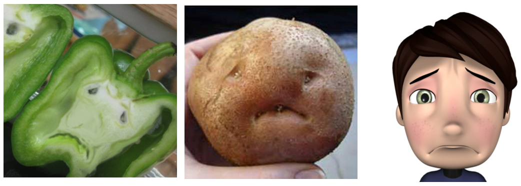
Several of our characters, by contrast, had eyes which displayed what we call the hyper-alert resting pattern, with an enormous amount of exposed eye white above the iris, many times the diameter of the iris itself. In spite of this very unnatural feature, in our tests the response patterns we observed with our hyper-alert characters did not differ in any significant way from that of our characters with much more realistic eyes. Surprisingly this included the neutral hyper-alert face, where the scattered viewer response pattern was nearly identical to our non-hyper alert faces, without the spike in fear or surprise one might expect (and certainly would get with a human face).

We also found that line art (Fig. 13) can provide cues with a much smaller amount of data than fully rendered, 3D art. Most specifically we discovered a small but significant spike in the recognition of joy when a single, tiny curved line is flipped horizontally at the corners of the mouth. This was our predicted result based on the way that line summarizes the bulging of the cheeks which is a very highly-weighed feature in the smile; flipped the other way this same curve does not suggest the topography of the bulged cheek in the same effective way. Finally, we found that there is a limit to the range of movement that is expressively significant. Most particularly we found that the raising of brows is highly effective in the expression of fear and surprise, showing a different recognition rate as the brows rise up the forehead, up until a certain upper limit when a continuing brow raise provokes no corresponding difference in viewer response. Here we hypothesize that the fact that no human eyebrow can rise past that point creates a similar perceptual limit when viewing a non-human character enacting a movement in a zone where we have no expectation anything significant can happen. Pattern recognition is potentially much more flexible and complex than the simple point-to-point mapping which current CG technology can so well achieve when creating accurate human faces.

Our most compelling finding along those lines came with the expression of anger, where a completely non-anatomical (the frowning muscles only affect the inner third of the eyebrow) extending of the angle of the inner brow (anatomically correct) with the outer brow (creating one continuous descending angle) provoked a very significant upwards spike in the recognition rate for anger.

4.8 Universal Triggers

Humans have extremely elastic criteria when it comes to what we perceive as a face. We seem to be hard-wired to consider an enor-



Expression	%		
Joy	0	2	0
Sadness	90	78	84
Anger	2	2	2
Surprise	0	0	0
Fear	0	0	8
Disgust	6	6	6
No Emotion	2	12	0

Figure 14: *Although the actual muscle movements which create the sadness cues are highly complex, it is easy to suggest the facial movements associated with sadness with simple graphic equivalents. Required is a notation that is perceived of as obliquely angled eyes (or eyebrows), and another as a downturned mouth; a requirement here met by a manipulated photograph of a pepper, a found potato face, and one of our characters. Just because an expression is conveyed with minimal means does not make it less legible, and here almost the opposite seems to be the case. The potato is less effective only because (we theorize) not everyone saw it as face.*

mous range of stimuli as representing the features, as long as the basic four elements are present: two eye-like shapes above a mouth-suggesting shape, in a bounded surround. Once these minimum criteria are met, almost any depicted or found configuration is a potential character, and we seem to unconsciously assign that character a personality. This is proven not only in the literature but also by the variety of successful creatures and entities that have been created by artists over the years from such unlikely objects as boats, planes, cars, appliances, and plants.

These creations communicate their feelings by means of features that, while often highly inventive and non-anatomical, manage to convey the message of smiling, sadness, anger, etc. without any loss in the translation. Disney artists, for example, managed to make ducks smile, not a simple matter given the peculiar and inflexible nature of the beak. Compared to Donalds beak, the teapots, clock, and candlestick of Beauty and the Beast presented much less of a challenge.

To find the underlying pattern that allows a variety of facial sources to communicate the facial expressions we tested a range of both line-drawn and 3D characters representing a continuum from realistic to stylized and non-human faces culled from the growing collection of found faces collected on various websites. We obtained very similar results regardless of the type of face, using our understanding as to the required triggers. See Fig. 14, 15, 16.

5 Conclusion

We have uncovered patterns and stimuli which we have not previously encountered in the literature, like the importance of lip contrast in the expression of disgust, and the relative weight of the upper vs. lower face in anger, joy, sadness, and surprise. We have observed the consistency of our Mechanical Turk survey results over time with a variety of character types, convincing us that the facial language we are deciphering has a hidden vocabulary which communicates in a similar way no matter who the audience, or the particular way the facial words are communicated. Stylized facial expression, in other words, is a bit like typography, in that the basic letter configuration can be recognized no matter what the type style, so long as certain basic elements are present.

In our approach, the cardinal expression underlying each of the 6 basic character facial rigs allows the artist/animator to simply tweak the subtlety and emphasize aspects of the expression that fit the story best. Moving separate features using the understanding of these triggers would allow for exaggeration of the expressions built in, resulting in clearer and non-ambiguous expressions as a default. Default rigs for stylized characters can be designed that support the universal triggers we discovered. One smart facial rig can be created for many characters, saving both time and money. Expressions can be reviewed more easily as triggers engage or disengage based on an intelligent grammar. The underlying structure is transparent to the animator. Shifting from an angry expression to mildly irritated should simply consist of dialing down the triggers associated with anger.

In further research, we plan to investigate the relationship between the most realistic possible computer-generated face and their stylized counterparts. Do we apply the same pattern-recognition rules to each? Are expressions as easily and quickly recognized on both types of faces? What characterizes the different level of acceptance for the reality of expressions on a stylized character vs. a virtually human one?

References

DeCarlo, D., and Santella, A. 2002. Stylization and abstraction of photographs. In *SIGGRAPH'02*, 769–776.

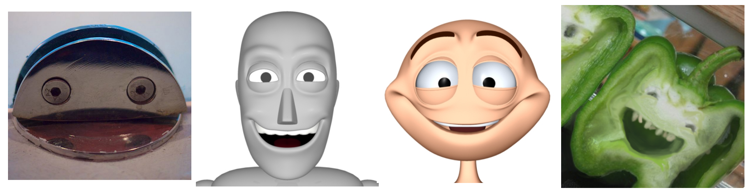
Ekman, P. 1999. Basic emotions.

Gooch, B., Reinhard, E., and Gooch, A. 2004. Human facial illustrations: Creation and psychophysical evaluation. *ACM Trans. Graph.*, 27–44.

Lasseter, J. 1987. Principles of traditional animation applied to 3d computer animation. *SIGGRAPH*, 35–44.

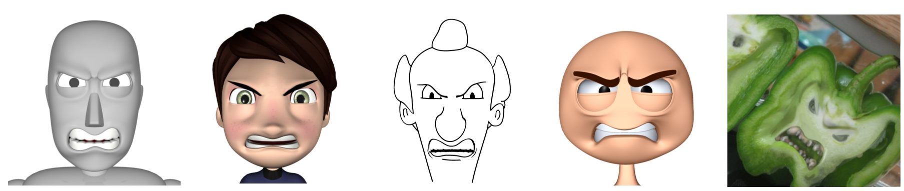
Tinwell, A., Grimshaw, M., Nabi, D. A., and Williams, A. 2011. Facial expression of emotion and perception of the uncanny valley in virtual characters. *Computers in Human Behavior*, 741–749.

Wallraven, C., Fischer, J., Cunningham, D. W., Bartz, D., and Bulthoff, H. H. 2006. The evaluation of stylized facial expressions.



Expression	**%**			
Joy	76	86	80	96
Sadness	4	0	2	0
Anger	2	0	0	2
Surprise	12	14	14	2
Fear	2	0	2	0
Disgust	2	0	0	0
No Emotion	2	0	2	0

Figure 15: *Joy universal. The upward arc of the lower lip is one of the most robust of all our facial cues. Another important cue are widely opened eyes, associated especially with surprise. Here a wildly divergent set of entities - a "found" face (actually a window hinge), two animated characters, and a manipulated pepper, provoke very similar responses, with even the percentage of people who picked surprise remaining constant, except in the case of the pepper, whose "eyes" are squinted rather tha wide open, an important feature of the laugh.*



Expression	**%**				
Joy	0	0	2	0	2
Sadness	0	4	0	6	6
Anger	94	86	92	90	74
Surprise	0	0	0	0	0
Fear	0	4	0	0	2
Disgust	6	6	2	4	6
No Emotion	0	0	4	0	10

Figure 16: *Consistent responses in spite of degree of stylization. Here 4 highly stylized faces (a robot with no eyebrow hair, a line drawing, a fairly realistic head, and a bald creature missing a nose) register remarkably similar anger scores. In posing these characters, we kept consistent the wideness of the eye, the oblique angle of the surmounting brow/lid, and the tensed, clenched mouth, the key triggers for the angry face. Note that eyebrow hair can be dispensed with as long as the equivalent shape is suggested by the brow skin and upper eyelid.*