



Published on *AWN | Animation World Network* (<http://www.awn.com>)

The Digital Eye: Image Metrics Attempts to Leap the Uncanny Valley

By *pplantec*

Created 08/06/2008 - 23:00



[1]

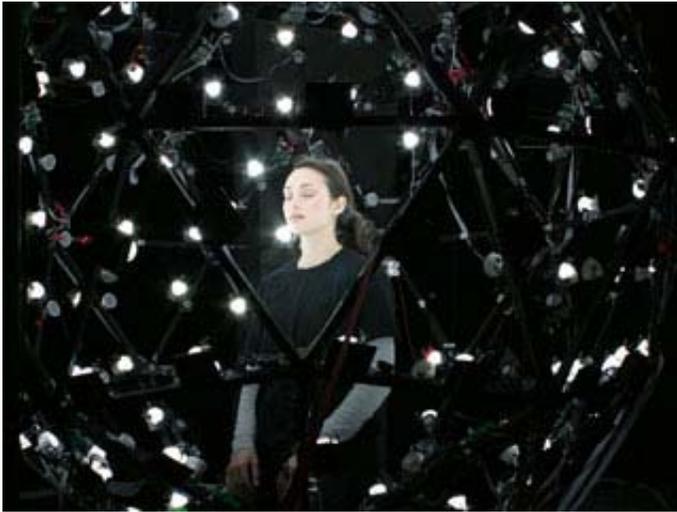
Image courtesy of Deron Yamada. © 2004

DYA367. [1]

I may have been wrong last month in A Coruña Spain, when I predicted that it would be two years before we establish a beachhead on the far side of the Uncanny Valley. (If you don't know what that is, Google it immediately.) While attending Mundos Digitales, I went to a presentation by Dr. Paul Debevec [2] on his latest work for USC's Institute for Creative Technologies, Computer Vision Lab. He talked about his sophisticated LightStage research into mapping faces in extreme detail. He mentioned that he was working with Image Metrics' Santa Monica, California, office, developing extremely high-resolution reference data for an advanced new process in markerless face capture and animation. He was pretty sure Image Metrics was about to achieve a breakthrough in virtual human face animation. Whenever Debevec gets involved in something, it's going to be done creatively and properly as well. Since I was there speaking about crossing the Uncanny Valley, this news immediately caught my attention.

I immediately Skyped my friend, Patrick Davenport, exec producer at Image Metrics. When I asked what they were doing with the "Emily" project, he hooked me up with Oleg Alexander and David Barton, the development gurus behind the shrouded new process.

What I discovered gives me hope that Image Metrics may really be fast upon a major breakthrough in face performance capture and animation. The Botox Syndrome, which infects virtually all face capture systems to date, may soon become a thing of the past.



[3]

Emily O'Brian being digitally cloned in USC's Light Stage 6 at the Institute for Creative Technology. All images courtesy of Image Metrics and USC's Institute of Creative Technologies. [3]

The Company

Stepping back a bit, Image Metrics does face capture without using markers glued to the face. This makes life more comfortable, and certainly less complicated, for people having their faces captured. It sounds like a dreadful thing and it can be. Image Metrics' approach to facial animation evolved from research in the field of computer vision. Three Ph.D. candidates in computer vision at the University of Manchester: Kevin Walker, Gareth Edwards and Alan Brett, developed technologies for tracking faces and analyzing medical images as part of their dissertation projects. That technology has to intelligently identify faces in more than one position and orientation, for example, for use in security applications. This requires analysis capable of producing the kind of data needed to track facial behavior. In fact, they can track virtually on a pixel-by-pixel basis with no face markers. They were soon able to extract reliable and accurate facial data from ordinary 2D video. They do it by keeping frame-by-frame track of tiny face irregularities, even pores, as well as head orientation and rotation. The concept opens up a lot of interesting possibilities like tracking classic performances in old movies and bringing old actors back to life... virtual life. They actually did that with Richard Burton.

Together the three friends formed Image Metrics back in 2000 in Manchester, U.K., to exploit the promise of their new technology. They've had great success, and have produced some remarkably accurate face animation; but to date, not accurate enough to completely fool most of us.

To make face behavior believable in virtual humans takes an enormous amount of synchrony between art and technology. So far, technology alone has not been able to do it. The Botox Syndrome is mostly the result of inaccurate movement of skin across the face, and eye movement that is just wrong. I confess it drives me crazy when I have to watch it.

It also drives the folks at Image Metrics crazy as well, which is why they've worked so hard to develop their facial animation technologies. In addition to tracking the areas that marker-based systems track, Image Metrics has found a way to accurately track places markers just can't

reach, such as the eyes, lips, tongue and teeth -- the areas that display the subtle emotions that make us who we are.

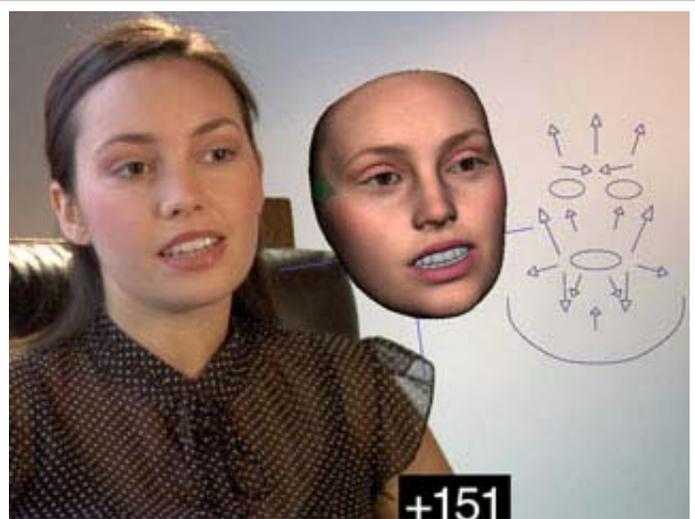
In order to really prove that their technology could accurately portray reality, Image Metrics Lead Rigger, Alexander and Head of Production, Barton, immersed themselves in "Emily."

The Project

Barton explained the project to me: "In order to showcase the power of Image Metrics' facial animation solution, we set out to create the world's first completely convincing photoreal CG face at HD resolution. While we can work with CG models from our clients, or create them in-house; to create the photoreal character for this demo, we asked Paul Debevec if he'd like to partner on the project to achieve the level of realism we were seeking. Paul's face scanning system is currently the only method capable of acquiring high-res, pore-level detail directly from a live subject, which is why it was a critical component of this project."

Up until now, this extreme level of detail has been captured by making a cast of the subject's face with extremely fine casting material that captures every pore and more than a few hairs as well. This model is then scanned at very high-resolution. It's extremely expensive and time consuming process that would never allow them to capture the high number of Emily's facial expressions they needed.

Alexander elaborated: "Paul captured the neutral facial expression model of Emily, as well as about 30 different expressions that were used to build the Emily character rig. Image Metrics' proprietary facial animation solution was used to track [actress] Emily O'Brien's performance directly onto her digital double, pixel-by-pixel. Using our solution, the rig developed by Paul's scans came to life as an exact replica of the real-life Emily. Without specialized equipment, or tracking markers or makeup, we were able to capture all of the subtleties of Emily's performance, gathering even her subconscious eye and mouth data."



[4]

Image Metrics [4] set out to create the world's first completely convincing photo-real CG face at HD resolution and asked Paul Debevec to partner on the project.



[5]

Debevec and his team scanned Emily's face in super hi-res, for facial reflectance, geometry and textures. In addition, perfectly flat 4K texture maps were shot with a cannon 1D Mark III. [5]

This was music to my ears. As a psychologist/virtual human designer, I've been emphasizing the importance of subliminal communication through micro behaviors in virtual humans. It is these perceived, but unnoticed behaviors that add the necessary richness to face expression that defines character and smaller meanings.

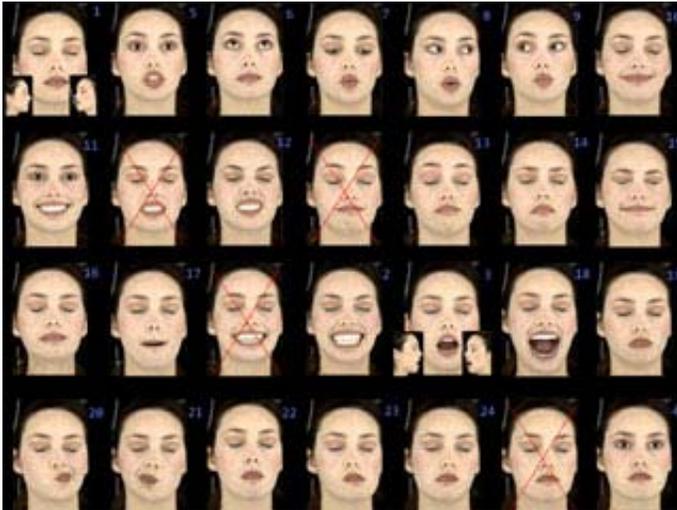
The Subject

They brought in lovely and talented actress O'Brien, recently nominated for a daytime Emmy, for her role on *The Young and the Restless*. Emily was a good choice. She has a very expressive face and is fun to look at in general. They sent Emily along with a team of videographers to work with Debevec at USC's Computer Graphics Lab, for kind of a Sci-Fi day. The Labs advanced Light Stage 6 looks other-worldly cool.

Debevec and his team were able to scan Emily's face in super hi-res, for facial reflectance, geometry and textures. In addition, perfectly flat 4K texture maps were shot with a cannon 1D Mark III.

It's a bit complicated, but by using tuned polarized lights flashed sequentially from every possible angle, and a polarized camera filter, they were able to obtain normal maps and color reference images of Emily's face, that were completely devoid of secularity and shadow.

According to Debevec, "The polarized light and camera were able to capture Emily O'Brien in completely flat light, perfect for creating skin texture shaders."



[6]

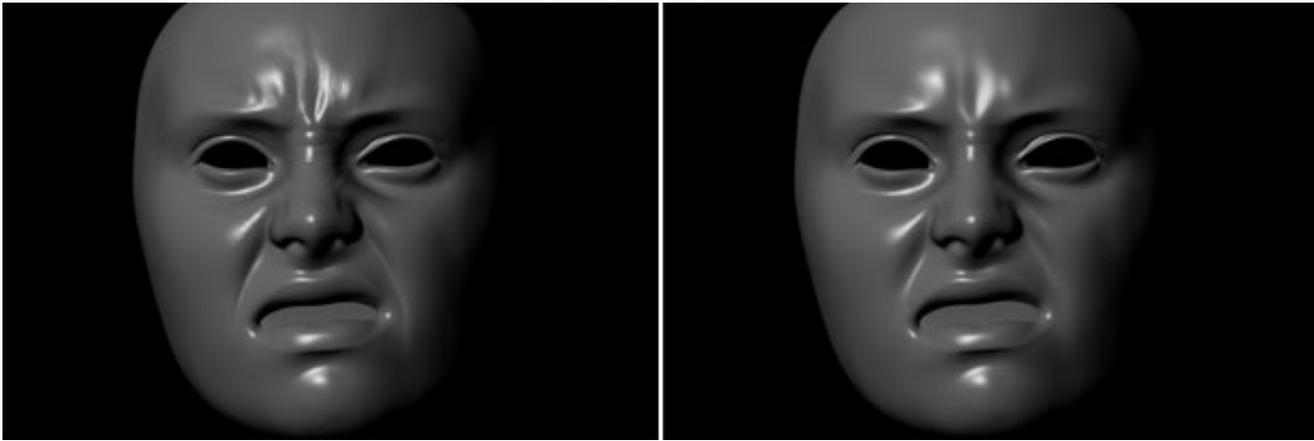
A library of Emily's expressions. [6]

The polarized light also yielded precise, very high-resolution normal maps, from which they were able to extract extremely precise geometry for Emily's face. They had Emily go through 30 takes, each with a different facial expression, the whole of which was calculated to yield data from which they could build a mathematical representation of how Emily's skin slides over her face's substructure. This, in turn, helped them build an accurate rig of her face. I don't believe this had ever been accomplished before with such accuracy and attention to skin movement. That alone was a major breakthrough.

Clearly fascinated by this project, Debevec told me: "They wanted to create a rigged version of Emily's face, so perfect that when digital Emily gets animated into different facial poses, everything that changes in her face is consistent with the real Emily's face changes. These include the basic face shape, the squishing a buckling of the skin (folds, etc.) and the shifting distribution of tissue.

"I was kind of amazed at some of the stuff I saw. For example, in the face rig, the skin around the upper part of her nose bunches up and the bridge of her nose gets thicker. You don't usually see this in a face rig. Because of the high-resolution work done by the graphics lab at USC ICT, we were able to notice some things that ordinarily are not included in face rigging, but clearly should be. By including such qualities, Image Metrics will be able to more accurately model and reproduce fine face animation."

What the two teams discovered during this project is essential stuff. These are the subtleties that give a face life.



Scans of Emily with displacements (left) and without. [7]

The Process I asked Emily what she thought of the process of having her face captured. "It was extremely enjoyable to step out of what I'm used to. I had seen what some actors go through with sensors glued to their faces and that's more or less what I expected. But the process was liberating. I wasn't restricted in any way... no wires... it was so freeing... no sensors on my face. And it was interesting as well. I almost couldn't believe the end result. I actually thought it was footage of me at first, until they told me, 'No, that's your digital double.' It was very satisfying to see the many parts of the puzzle come together and to be part of something this momentous."

The process is one where once Emily's face data is recorded in the Light Stage at USC, the data is used to create the very high-resolution maps we've discussed. Next, analysis of the flow of skin among the 30-35 expressive poses yielded a mathematical model of how the skin on Emily's face moves uniquely across underlying tissue. From the resulting information, Oleg built a face rig that moves precisely the way Emily's actual face does.

According to Debevec, "The purpose of all this is unique. Image Metrics is using the massive amount of data from our Light Stage scanning process to create a highly accurate rigged version of Emily; so that when digital Emily gets animated into different facial poses everything that changes in her face is consistent with the way the real Emily's face works..."

Next, they needed to find a way to drive that rig with precise face motion data from Emily. This is Image Metrics' core technology, so they were ready. They extracted highly precise facial performance data from 1080p video of Emily while acting and relaxing. They could actually track her face motion on a pixel by pixel basis, capturing extremely subtle behaviors in her eyes and small muscles. Next they had to map that information onto Emily's rigged face, frame-by-frame.

One problem they ran into because this is a markerless system, was tracking gross head movement while syncing up the tracking data to the face rig.



[8]

For the effect to be believable, CG Emily's face had to be perfectly superimposed on top of real Emily's face in every video frame. Above is the live- action shoot. [8]

Alexander explains: "For the effect to be believable, CG Emily's face had to be perfectly superimposed on top of real Emily's face in every video frame. The head tracking was a challenge because there were no tracking markers on her head or face in the plate, so there was 'nothing to track'. We solved the problem by writing special Matlab code. We exported an OBJ file for every frame and used the vertices of the mesh as 'markers' to align CG Emily's face on top of real Emily's face. I'd also like to clarify that we don't need markers for our animation solution. But if we were to do this project again, we would put a few 'tracking markers' on her nose and cheeks."

The capture team had videotaped Emily from two angles while she performed. I understand they also captured her between cuts, while she was relaxing and just being Emily. That latter is of most interest to me. Would I believe? As I type, I'm waiting for clips showing the real Emily and the virtual Emily. I'm hoping desperately to be fooled. I had to see a clip.



[9]

Dr. Paul Debevec. [9]

Meanwhile, I asked Paul about the experience of working with Image Metrics. "It was great. We worked well together. The whole shoot of Emily took about two hours and we captured her in light natural make up. She was a great subject. You need to stay still for three seconds, which is longer than you think... she was excellent at it. She's British, in her early '20s... young, near

perfect skin... fun to work with. That kind of skin has traditionally been one of the most difficult to get right. One thing we were thinking, that older faces -- because they have more lines and wrinkles -- that they might be more emotive... but we've found the opposite is true. The younger faces have more elasticity and fine expressiveness, which makes the face do much more interesting things. What happens is considerably more subtle. You need a system like this to capture that."

I Got the Clip!

OK, I just got the clip... Wow... impressive. I'm a tough sell, one of the toughest, but this looks good. I don't see any Botox at all in Emily's face. Not sure if they're messing with me here. This could be the actual footage of Emily. There are some imperfections in the face render, so they're not messing with me... this is very impressive, but it's only six seconds or so, I need to see more. They sent it with a note: "Remember, this is a work in progress and it will not be final until SIGGRAPH." Damn. I understand they'll be showing a much longer and more impressive demo there.

Wait! Just before sending this story, I received a longer, more refined clip. It is absolutely awesome -- amazing. I'm one of the toughest critics of face capture, and even I have to admit, these guys have nailed it. This is the first virtual human animated sequence that completely bypasses all my subconscious warnings. I get the feeling of Emily as a person. All the subtlety is there. This is no hype job, it's the real thing.

It's Official

I officially pronounce that Image Metrics has finally built a bridge across the Uncanny Valley and brought us to the other side. I was indeed wrong about it taking another two years and I'm happy about that. You simply must get to the Image Metrics booth at SIGGRAPH to see this thing. Robert Zemeckis, are you listening? This is the one you want for your next attempt.

Working together, these two teams have discovered and captured the essential subtle facial elements of face animation that have previously been ignored. These are the things that give a face life and identity.

Here's the poop on Virtual Emily's official debut: The Image Metrics booth at SIGGRAPH (#1229). Image Metrics and Debevec will showcase how Emily was created using their respective technologies in a SIGGRAPH Tech Talk on Wednesday, Aug.13, from 1:00-2:30 pm in Room 2, Hall G. If you love virtual humans the way I do, you won't want to miss it.

Peter Plantec ^[10] is a best-selling author, animator and virtual human designer. He wrote The Caligari trueSpace2 Bible, the first 3D animation book specifically written for artists. He lives in the high country near Aspen, Colorado. In addition to his work in vfx and journalism, Peter is also a clinical psychologist with more than a decade of clinical experience. He has spent several years researching the illusion of personality in animated characters. Peter's latest book, Virtual Humans ^[11], is a five star selection at Amazon after many reviews.

Source URL: <http://www.awn.com/articles/technology/digital-eye-image-metrics-attempts-leap-uncanny-valley>

Links:

[1] <http://www.awn.com/imagepicker/image/14776>

[2] <http://vfxworld.com/?atype=articles&id=3645>

- [3] <http://www.awn.com/imagepicker/image/14777>
- [4] <http://www.awntv.com/videos/image-metrics-emily-project>
- [5] <http://www.awn.com/imagepicker/image/14778>
- [6] <http://www.awn.com/imagepicker/image/14779>
- [7] <http://www.awn.com/imagepicker/image/14780>
- [8] <http://www.awn.com/imagepicker/image/14781>
- [9] <http://www.awn.com/imagepicker/image/14782>
- [10] <http://www.ordinarymagic.com/>
- [11] <http://www.awn.com/>