

A new window to the face

August 9 2011, By Douglas Gantenbein



Top images show how markers create a set of correspondences across all the face scans. The bottom row shows how a two-step face-scan registration produces dense, consistent surface correspondences across all the face scans.

The human face is a complicated thing—powered by 52 muscles; contoured by the nose, eyebrows, and other features; and capable of an almost infinite range of expressions, from joy to anger to sorrow to puzzlement.

Perhaps that is why realistic animation of the human face has been what Microsoft Research Asia scientist Xin Tong calls a “holy grail” of computer graphics. Decades of research in computer graphics have developed a number of techniques for capturing three-dimensional moving images of the human face. But all have flaws, capturing insufficient detail or failing to depict accurately a changing expression.

Now, researchers at Microsoft Research Asia, led by Tong and working with Jinxiang Chai, a Texas A&M University professor, have developed

a new approach to creating high-fidelity, 3-D images of the human face, one that depicts not only large-scale features and expressions, but also the subtle wrinkling and movement of human skin. Their work could have implications in areas such as computerized filmmaking and even in creating realistic user avatars for use in conferencing and other applications.

The team's paper about the facial-scanning research, "Leveraging Motion Capture and 3D Scanning for High-fidelity Facial Performance Acquisition, will be presented during SIGGRAPH 2011, being held in Vancouver, British Columbia, Aug. 7-11. SIGGRAPH 2011, the 38th International Conference and Exhibition on Computer Graphics and Interactive Techniques, is expected to attract 25,000 professionals in the fields of research, science, art, gaming, and more.

The paper, written by Microsoft Research scientists Tong, Haoda Huang, and Hsiang-Tao Wu, along with Chai, explores a new approach for capturing high-fidelity, realistic facial features and expressions.

That's a big challenge, Tong says. Not only is the human face remarkably expressive, it's also a form of communication—we look at people and usually can understand immediately what they are thinking or feeling.

“We are very familiar with facial expressions, but also very sensitive in seeing any type of errors,” he says. “That means we need to capture facial expressions with a high level of detail and also capture very subtle facial details with high temporal resolution,” meaning the subtle motions of those details need to be captured.

Existing means of capturing faces and expressions include marker-based motion capture and high-resolution scanners. In marker-based techniques, small reflective dots are fixed to a face and their relative positions captured on video as the character changes expression. That

method results in accurate capture of changing expressions, but with low resolution.

High-resolution scanners, on the other hand, capture all the subtleties in a human face—down to fine wrinkles and skin pores—but typically do so only for static poses. Dedicated setups configured with high-speed cameras, also used for facial capture, are expensive and provide less facial detail.

The team set out to combine both the accurate motion capture of marker systems with the high resolution of scanners. The researchers also wanted to do this efficiently from a computing standpoint, and that required the least amount of data for an accurate facial reconstruction.

Using three actors with highly mobile faces, the researchers first used marker-based motion capture, applying about 100 reflective dots to each actor's face. With video rolling, the actors made a series of pre-determined facial expressions that enabled the collection of rough data on how faces change with different expressions, for use in 3-D scans.

Also, by analyzing the captured marker-based data, the team determined the minimum number of scans required for accurate facial reconstruction.

In the next step, the research team used a laser scanner to capture high-fidelity facial scans. Those scans then were aligned with the corresponding frames in the marker-based facial data. Using a new algorithm, the facial scans were registered with each other.

This was no easy task. The paper notes that geometric details that appear in one scan might not appear in another. Also, even a small misalignment of fine-grained features such as wrinkles or pores will make the resulting facial reconstruction appear unnatural.

“We want to make sure these features align, or you will see some weird artifacts,” Tong says. “A wrinkle may appear, then disappear, then appear—it’s not natural.”

To avoid that, the team used a two-step registration algorithm. First, the algorithm registers large facial expressions between the high-definition facial scans. Next, it refines the scans by segmenting the face into discrete areas and aligns each area with the same areas in other scans with the similar appearance to the current scan using optical-flow techniques that take into account the relative motion between a camera and the face.

Lastly, the team combined the motion-capture information with the face scans to reconstruct the actual expressions as they were performed. The resulting images capture both the “big” movements of a face and fine details such as the texture and movement of the skin.

Tong is confident that his team’s work will have an impact in the real world.

“It has a lot of applications,” he says. “That is why we put so much effort into the work.”

For example, the film and video-game industries also could benefit from easier but effective methods for creating virtual faces, leading to virtual characters that are much more lifelike than is common today.

Also, Tong says, the new scanning technique could be used to create computer avatars that could represent a realistic option to the pre-programmed avatars found in devices such as the Xbox 360.

“The character would be virtual, but the expressions real,” he says. “For teleconference applications, that could be very useful, for example, in a

business meeting, where people are very sensitive to expressions and use them to know what people are thinking.”

Tong says much work needs to be done, however. Currently his team’s scanning technique does not capture synchronized eye and lip movements. In addition, it takes considerable computing power and several hours to successfully register all of the images. Tong wants that to happen in real time.

“There are a lot of challenges,” he concludes, “but it is a very exciting research area.”

More information: faculty.cs.tamu.edu/jchai/proj...1/Face-final-v11.pdf

Provided by Microsoft Corporation

Citation: A new window to the face (2011, August 9) retrieved 18 April 2024 from <https://phys.org/news/2011-08-window.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.