

CRACKING KNUCKLES, METACARPALS AND PHALANGES



How Framestore and Vicon finally solved hand capture

VICON

Most people talk with their hands a lot," says Richard Graham, Capture Supervisor for VFX house Framestore. "There's so much information we give to each other when we speak through hand gestures."

Despite the importance of hands in communication, however, finger tracking has long been a white whale in motion capture - one of the barriers stopping VFX artists from crossing the uncanny valley to create a complete 'digital human'.

"Framestore is very much an animation-led studio, and one of our fortes is character work. We hold ourselves to a really high standard of performance and a large part of the work we do isn't

just action, which you can do very well in mocap. It's characters talking and emoting and gesticulating," says Richard.

For Framestore, accurate hand capture is a significant step towards greater emotional authenticity in its characters, but until a recent project completed with Vicon it simply wasn't viable, with animators having to pick up the slack manually.

TECHNICAL BARRIERS

"With a hand there's just so much occlusion going on at any pose," says Tim Doubleday, Entertainment Product Manager for Vicon. Markers disappear from the view of cameras, getting tangled up in processing, and the sheer complexity of hands aggravates the problems further.

"The number of degrees of freedom that your hands have compared to, say, your spine creates a problem that's an order of magnitude bigger," says Richard. "So to capture a full human hand accurately requires a lot of small markers. This in turn means you need lots of cameras and you start to hit resolution issues if the cameras are too far away."

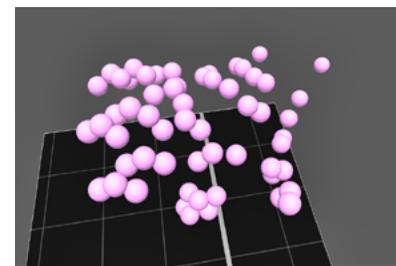
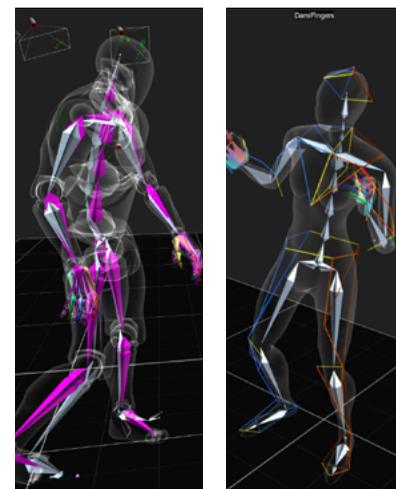
The result would be messy data, with models often resembling pretzels more closely than they did hands.

There were existing solutions. Using optical capture, studios would sometimes track three digits and extrapolate their positions out to animate five, but it required such a tremendous amount of cleanup and post-production solving that it wasn't really useful.



There are also data gloves, but they're expensive, easily damaged, often unreliable and require the integration of a further set of data into the animation process.

Framestore wanted a process that would fit seamlessly into their workflow to produce accurate models without sucking up valuable time and resources. To finally address it, the studio collaborated with Vicon on a 2017 Innovate UK funded project called ProDip to come up with a working solution over an 18 month period.



TRAINING SHOGUN

Vicon's approach was to build a virtual model of a hand skeleton that Tim's team, led by Lead Developer Jean-Charles Bricola, trained to understand what a 'real' pose looked like. To do this, they had a number of different subjects of various hand sizes be captured. Each subject placed 58 markers on a single hand. They then performed a range of key actions, meticulously tracking them as they went along.

"You basically use that as training data to constrain the hand model. For each subject that came in, we knew where the joints should be placed and then also how their hand deformed during the range of motion," says Tim.

The model produced by the team then became a reference point that Shogun can use to interpret a hand movement tracked by a dramatically reduced number of markers.

"Vicon offers markersets supporting ten, five, or three markers on each hand. The hand model then looks at the marker cloud and based on the training data knows which marker is which and how the hand skeleton fits within the markers," says Richard. "So therefore it can't ever do something that the human hand can't do."

PUTTING THE SOLUTION TO THE TEST

Framestore has wasted no time in putting the new process into action. The system has been used for everything from characters in blockbuster movies (including Captain Marvel, Aeronauts, and Spider-Man: Far From Home) to producing reference material in real-time, directly out of Shogun, for animators doing character-work to use. Basically, it has been put to the test in every Framestore mocap project over the last two years.

Internally, Framestore has been testing the sensitivity of the process. "We've used it in the last year on capturing animation studies of small, very subtle moves, quite subtle bits of performance," says Richard.

It's been a hit with clients, too. "The takeaway for us is that when we offer fingers to our clients they always say yes, and now it's much less of a headache. It's been a great advantage for the last two years that we've been using it."

Richard expects to be able to make much more use of hand capture in the future, while Tim envisions using the same techniques to build better models of other parts of the body. "We have plans to kind of use the same approach, using dense data to train a model for working out spine articulation," he says.

Tim isn't done with hands yet, however. He has a very specific milestone he hopes to hit. "So the challenge for us is sign language, and being able to deliver a performance from the hands that somebody who is deaf can read and understand exactly what the characters say. You can imagine a videogame where one of the characters actually sign in the performance. And that's never been done before, but hopefully using the system that would be achievable."