DEPARTMENT OF DESIGN SCIENCES — FACULTY OF ENGINEERING (LTH)
PRESENTED 2026-06-02

MASTER'S THESIS 4D Reconstruction in Sparse Camera Settings For VR
STUDENTS Ezgi Aysel Bati, Yiran Chen
UNIVERSITY SUPERVISORS Günter Alce (LTH)
EXAMINER Joakim Eriksson (LTH)

Recreating the Living World with Just Four Cameras

POPULAR SCIENCE SUMMARY **Ezgi Aysel Bati, Yiran Chen**

What if we could capture any moment: a crime caught on camera, or your friend's impressive dance moves, and reconstruct it in three dimensions, to relive it from any angle, again and again? This might be closer to reality than you think.

Gaussian Splatting takes away the pixels used to make up a digital copy, and replaces them with lots of blobs. Yes, blobs. Each one carries information like colour, size, and transparency. Blend these blobs together and you get a super realistic digital version; rendered faster and much more efficiently than previous methods.

But why stop there? Take it one step further by recording a short motion. Add some extra math to keep track of how each blob changes each moment and you can replay your few seconds of video, now as a scene you can step inside, and even explore it from different angles in Virtual Reality.

Great! What's the Catch?

Gaussian Splatting achieves amazing results, and does so in an hour or less. But this method needs to match points from many, many images captured from all angles to get a complete reconstruction. This means that you either need one camera going all around, or about a hundred of them stationary. In our thesis, we wanted to explore if we can use this impressive method with a more realistic set up, where we have only four cameras.

Our contribution is to test the method with both static and dynamic scenes, and compare various follow up works to see if we can get a reconstruction with only four cameras. Our goal was simple: trade-off the super realistic results and aim for a recognizable and complete scene, but use a setting that could occur in real life.

So, did it Work?

We ended up finding that the best results are obtained with a method called Diffuman4D, where artificial intelligence is used to generate new images in between, filling the gaps where the additional cameras would be. Combining this with a special cube which helps us determine digitally how our cameras are situated in comparison to one another, and we get an impressive result. In our tests, you can even see the three stripes on a shoe, which is a higher level of detail than we aimed for.

There is still a long way to go, and it would be even greater to get the close to real life level of detail with the few cameras, or get this method even faster so you can go instantly from video feed to exploring in virtual reality. Isn't it fun to imagine all the possibilities?

Figure 1: A moment from Macarena