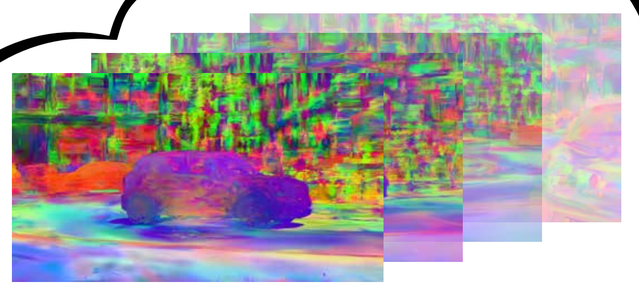
From the camera's perspective, in what direction is the car moving?

I would imagine that in 4D, the scene looks like ...



<Img1><Img2>... From the 2D features across time....

Obviously, the car is moving to the ***right***!

Human

The car is moving to the ***left***.

Vision Language Models