# Markerless 6+4DOF real-time visual tracking of the human hand with automatic initialization

Ever since the creation of the first Virtual/Augmented-Reality (VR/AR) applications Human-Computer-Interaction research has taken a great interest in using the human hand as an input-device. This is explained by one of the main requirements of VR/AR: actions/responses familiar from our everyday life have to be at the user's disposal. Furthermore, the overall user experience should be enhanced by allowing the user to carry out tasks, which are impossible in real life (e.g. flying). Everyday interaction with our environment ubiquitously involves our hands as an "interaction-device", therefore, it should also be available in its virtual counterpart. To enable VR/AR systems to react to hand movements, the user's hand has to be tracked. In this context, a markerless approach is particularly desirable, to ensure the interaction to be instant and immersive (one does not need to wear an extra device e.g. a glove to start grasping or pointing at objects in the real world).

While previous markerless hand tracking methods mostly track the full 27 degrees of freedom (DOF) of a hand (global pose and joint angles) for all kinds of interaction purposes, they suffer from several drawbacks, such as no real-time support or special initialization procedures. Hence, these methods are not applicable in practice so far. Several other approaches, however restrict themselves to a subset of the full DOFs (at least 6 DOFs: translation and rotation of the hand) and solve these drawbacks. Unfortunately, they still suffer from another problem: None of them is capable to track all the 3 rotational DOFs without limitations.

Therefore, we developed a novel computer vision based hand-tracking technique, which is capable of robustly tracking the 6 DOFs of the human hand while distinguishing between 4 different gestures without such limitations. Moreover, it works in real-time (at least 25 frames per second) with the help of 3 (or more) off-the-shelf consumer cameras. A key feature of our system is its fully automatic real-time initialization procedure, which, along with a sound tracking-lost detector, makes the system fit for real-world applications. Because of this, our method acts as an enabling technology for uncumbersome hand-based 3D Human-Computer-Interaction. Using our tracking we evaluated the use of the hand as an input device for two prevalent Virtual Reality applications: fly-through exploration of a virtual world and a simple digital assembly simulation.