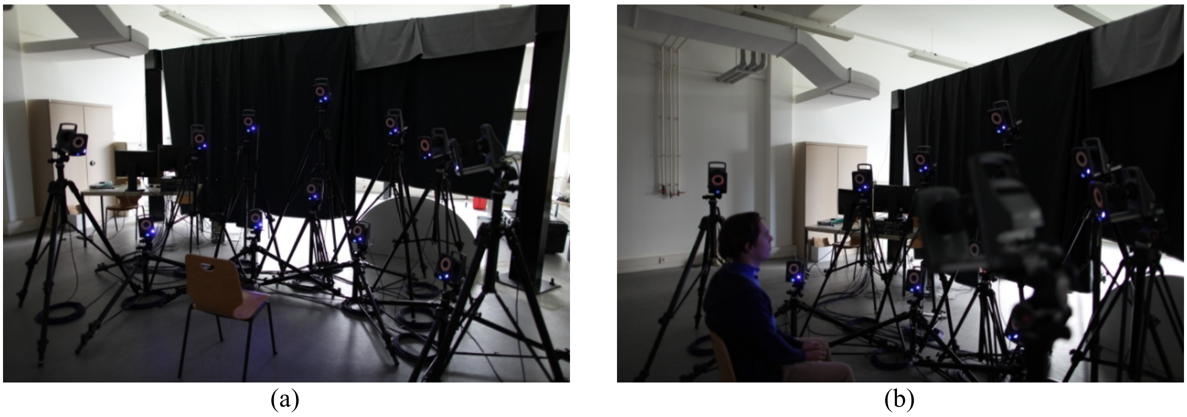
MOCAP Data Acquisition Process

By means of our mocap data acquisition process we aim to supply developers with real-world dynamic data sets in a form of animated meshes. This data will be used in the new segmentation, registration, and compression methods, i.e. all related techniques for dynamic mesh processing. At the current stage we focus on a certain part of human body, namely on human face. We are interested in a database of multi-subject dynamic meshes exhibiting intra-subject facial motions (expressions of emotions, lip movements, speaking, etc.), and coherent inter-subject facial motions (i.e. different subjects show exactly the same emotions, say exactly the same words etc). Among subjects for data acquisition there's no preference for gender proportion or age distribution. At the moment we have acquired the data from 6 distinct subjects.

The MOCAP system

Our system was custom-assembled of five Vicon T40 and of seven Vicon T40-S cameras. Such cameras are specifically tailored for marker motion capture and have a resolution of 4 megapixels, capture 10-bit grayscale using 2336 x 1728 pixels and can capture speeds of up to 2,000 frames per second. Entire set of twelve cameras is placed in the motion capture zone and forms a semi sphere around a subject (Fig. 1). We configured the system to capture marker positions at 120 frame rate.



(a)

(b)

Fig. 1: Vicon system installation (a) 12 cameras positioned in a semi sphere (b) subject of the motion capture is placed just in front of the system.

Data Acquisition

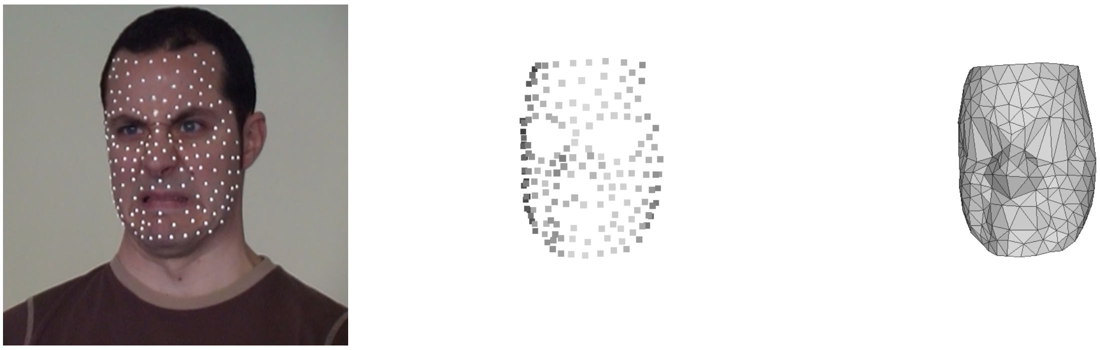
MOCAP session for data acquisition typically consists of the following stages:

- Set up markers
- Capture marker trajectories
- Data post-processing

1. Set up markers. We estimated that for each subject we need from 120 to 200 of light reflective markers for robust data acquisition (may vary from one person to another). Currently we are using a set up with 160 markers (Fig. 2(a)). Entire surface of the face surface should be covered. Markers have to span all over the face till the indicative boundaries defined by the hair, chin lines, etc. Ideally, any nearest 3 markers from the marker set should form a triangle which resembles the local face surface as much as possible; although it is extremely difficult to achieve in practice. We paid a lot of attention to the inter-marker distance: markers should be kept close enough to each other (0.8-1cm apart, depending on a subject and number of markers). Regions of high deformations (for instance, mouth area) should be covered by markers more densely; beside that we tried to place markers as regular as possible.

2. Capture markers trajectories. A person with markers is placed in front of Vicon system; then, the actor performs all the motions (normally motions are predefined and discussed before the session starts). Before the mocap session, the participant shall make some exercises. One should be an actor, performing with exaggerated expressions. The Vicon cameras capture trajectories of the markers individually; the output trajectories data set is stored in one single binary file in .c3d format.

3. Data post-processing. After acquiring raw markers movement data, an engineer needs to post-process the data in order to construct a dynamic mesh. First, the engineer labels the trajectories if needed (during the motion capture some markers may disappear for a frame or two due to occlusions) with Vicon Blade software. Second, our software extracts the coordinates of each marker in every frame. As a result each frame is represented as a point cloud (Fig. 2(b)). Assuming the topology remains the same during the motion capture, we only need to triangulate only first frame for a subject (Fig. 2(c)). Finally, after transferring of the topology information to all point cloud frames we obtain a dynamic mesh sequence.



(a) (b) (c)

Fig. 2: Data aqcuisition in MOCAP session (a) Actor with markers in place. (b) Point cloud acquired for each frame during the MOCAP session. Points represent marker locations. (c) Triangulated point cloud. While keeping the topology consistent over the frames we get a dynamic mesh.