



## **D6.4** Implementation of Structured Models **Public Summary**

Project no: FP7 - 211471

## Executive Summary

This report provides an overview of deliverable D6.4 *Implementation of Structured Models*. Work Package 6 (WP6) brings together the research work from WP4 and WP5 in i3dPost to investigate practical solutions for post-production tasks and create prototype software tools. In Work Package 4 unstructured point clouds and surfaces are derived from multiple view video data. Work Package 5 addresses the derivation of structured set, actor and face models with semantic description from the unstructured data. Deliverable D6.4 describes the proposed tools from WP5 for structured modelling in the Nuke compositor along with modelling and animation in Maya.

### Structured Set Modelling

The acquisition of set models is required to allow addition and rendering of computer generated elements and rendering of background plates in production. In WP5T1 research is investigating a solution to structured set modelling using spherical image capture that provides both geometric and high dynamic range image data for a scene. A pair of panoramic images are captured as shown in Figure 1. This will be converted to a 3D scene in Nuke through two plug-ins, the *SphericalDisparityGenerator* and *SphericalMeshGenerator*. The disparity between the images is calculated first, the disparity can then be edited in Nuke. The disparity is then converted to an efficient mesh representation for use in the Nuke 3D environment. A regular triangular mesh is constructed by converting disparity to depth and mapping it to 3D. The output will be a Nuke geometry node with UV coordinates to texture the mesh from a spherical image.



Figure 1 Spherical stereo images from the Spheron HDR system

# Structured Actor Modelling

Actor modelling is required to support modification of interaction with computer generated characters or objects, rendering of shadows and editing of actor performance. Unstructured multiple view data requires a single common structure or correspondence over time to support artist manipulation. Research in WP5T2 is investigating methods to convert multiple-view video to a control skeleton for a character model constructed by an artist.

The modelling process is divided into distinct stages. Structured processing of multiple view video is performed first in Nuke and 3D modelling and animation is performed in Maya. Firstly a *VirtualMarker* plug-in for Nuke will provide surface tracking in multiple view video to provide 3D point trajectories over time for the actor. The surface tracking data is then used to create a single temporal consistent mesh from a character model, providing a common mesh structure. The mesh is constructed as a preprocessing step before importing into Maya. A *SkeletonTracker* plug-in for Maya then tracks the skeleton for the character model through the sequence. The output is the skeleton animation curves for the artist constructed character model to match the actor performance in the multiple view video.

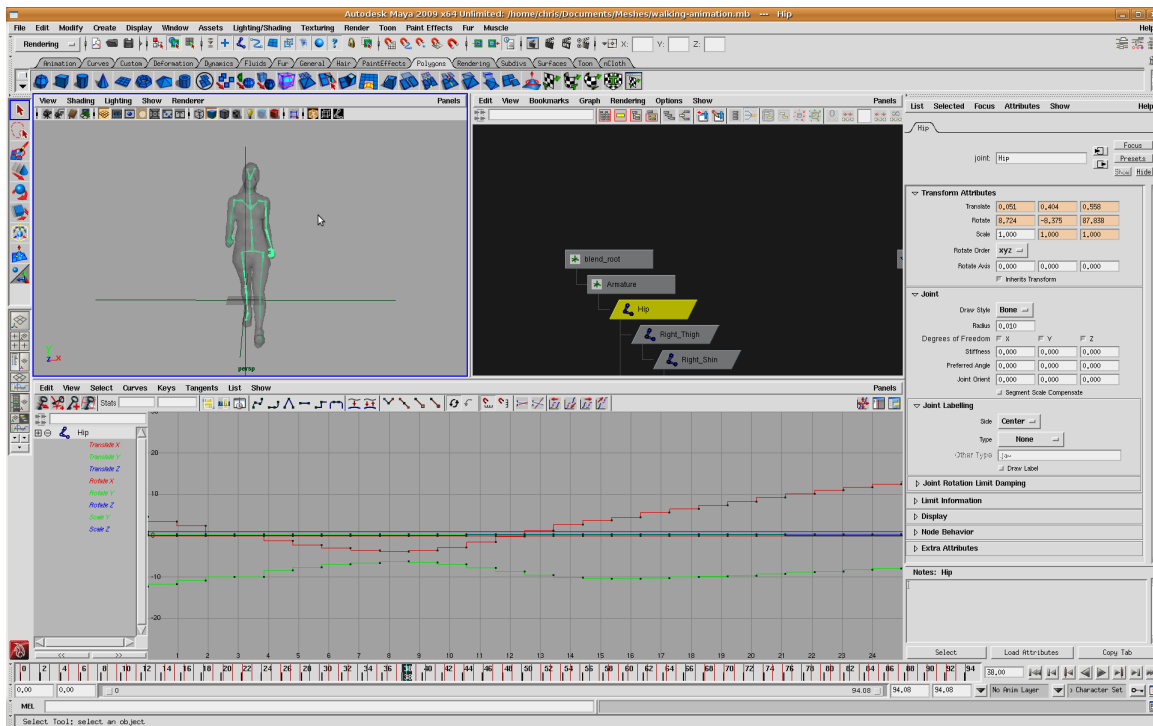


Figure 3 Output skeleton animation curves from the *SkeletonTracker* plug-in for Maya

## Structured Face Modelling

Conventional motion capture for facial animation is based on marker-based optical capture systems such as the Vicon system. Markers can be avoided using active shape capture techniques such as structured light projection (3DMD) or the use of ultraviolet sensitive phosphorescent make-up (Mova Contour). Research in WP5T3 is investigating a flexible video-based solution to capture facial performance without specialist equipment.

The *VirtualMarker* plug-in for Nuke will allow the user to place and track a set of 'virtual markers' in multiple view video. The plug-in will also track a set of autodetected markers and allow the interactive selection and editing of tracks to create a user-defined 3D marker set. The marker set can then be exported to a standard file format such as .fbx for use as animation control points in a 3D modelling and animation package such as Maya.

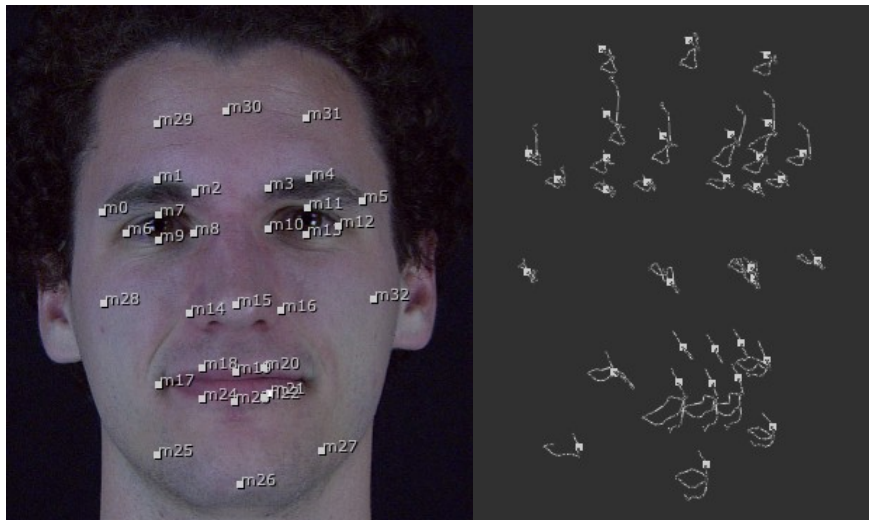


Figure 4 Virtual Markers are positioned in the input multiple-view video using *the VirtualMarker* plug-in for Nuke and tracked to create a 3D marker set with 3D trajectories

## Semantic description of human activity for retrieval

Semantic description of multi-view video and 3D scenes will facilitate retrieval and reuse of the structured models. In WP4T3 research will first label scene elements, for example by identifying actors or props and creating appropriate meta-data such as actor identity or object class. In WP5T4 research is investigating semantic annotation of dynamic 3D scenes in terms of actions or interactions between actors. The research focuses on the recognition of frequent actions such as walking or running as well as emotions from facial expression such as happy or sad.

Recognition and labelling will be performed from multiple-view video using two plug-ins in Nuke. The *ActivityRecognition* plug-in will label frames of a video sequence according to human activity. The plug-in requires a binary segmented body mask as shown in Figure 5 and provides an activity label based on a predefined database of learnt appearance. The *ExpressionRecognition* plug-in will label frames of a video sequence with the depicted facial expression based on learnt appearance. The input to the plug-in is illustrated in Figure 6. The output of the plug-ins will be the annotated activity as meta-data in Nuke.

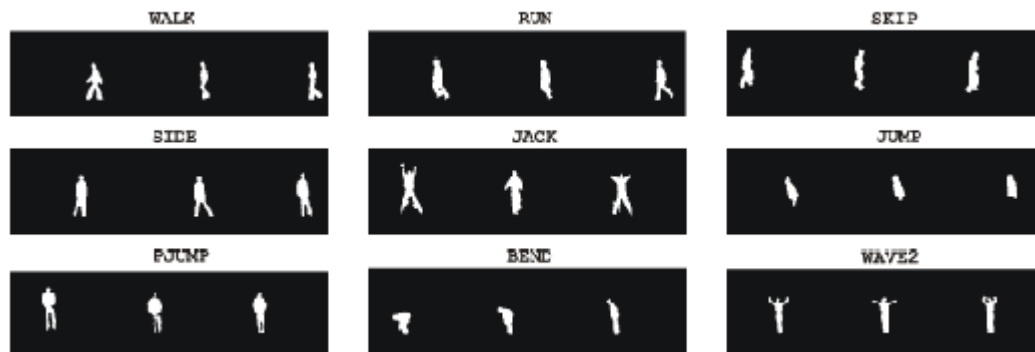


Figure 5 Example of binary images that will be used as input to the Nuke *ActivityRecognition* plug-in



Figure 6 Inputs to the *ExpressionRecognition* plugin showing happiness, sadness, anger, fear, surprise, disgust, and neutral (images from the JAFFE database).