



## i3DPost - First Annual Report, 2007-8

[www.i3dpost.eu](http://www.i3dpost.eu)



The i3DPost Project is formed of six companies and educational institutions that are Europe's leading specialists in their fields. The Foundry Visionmongers Ltd., Centre of Research and Technology Hellas, Informatics and Telematics Institute (CERTH-ITI), BUF Compagnie, University of Surrey (UniS), Quantic Dream (QD), The University of Dublin (TCD).

i3DPost will develop new methods and intelligent technologies for the extraction of structured 3D content models from video, at a level of quality suitable for use in digital cinema and interactive games.

The research will enable the increasingly automatic manipulation and re-use of characters, with changes of viewpoint and lighting.

i3DPost will combine advances in 3D data capture, 3D motion estimation, post-production tools and media semantics. The result will be film quality 3D content in a structured form, with semantic tagging, which can be manipulated in a graphic production pipeline and reused across different media platforms.

## Summary of Activities



### *Key scientific & technical innovations*

The innovation that i3DPost proposes will produce a step change in the quality of 3D video and postproduction, combined with semantic metadata, to achieve practical implementations of intelligent media technology at a quality level that meets the requirements of the film, games and professional media industry.

The goal is to create methods that make it possible to edit and alter filmed sets and actors – their faces, clothes and performances – as if they were animations, to allow their simple manipulation and reuse within post-production and the retargeting of filmic content to interactive platforms. The integration of semantic technologies with next-generation approaches to 3D video capture and postproduction offers a means of changing the way the media industries work, and opens the way to the repurposing of film and video material to create new programme experiences.

The result will be film quality 3D content in a structured form, with semantic tagging, which can be manipulated in a graphic production pipeline and reused across different media platforms including films (either 2D or 3D), computer games, music and promotional videos.

The **overall objective** is therefore to integrate 3D information extracted from the visual scene into all stages of the postproduction pipeline, supported by semantic metadata, at a quality level that meets the requirements of the film, games and professional media industry. This will be achieved by the invention of methods for recovering intelligent, structured content from on-set filming, and incorporating them in capture environments and software tools that enable the simple manipulation and reuse within post-production and the retargeting of content to interactive platforms.

Having almost completed the Initial set-up phase and the establishing of the scenarios of use, the work on the i3DPost Project will continue towards the completion of the proof of the concepts in the main areas of research.

## Data Capture

Early on in the project, BUF (Film Studio) and Quantic Dream (Games Company) wrote a report that set out the user scenarios & requirements taken from their knowledge of Production. This provided the team at the University of Surrey with guidance as to what their data should include.

The University of Surrey have now captured the test data sequences that the partners are to use for their research. The first test data set comprises multiple view chroma-key sequences of a single actor performance. Reconstruction of actor performance using existing software tools has been delivered to the partners to use as ground-truth for evaluating algorithm performance.

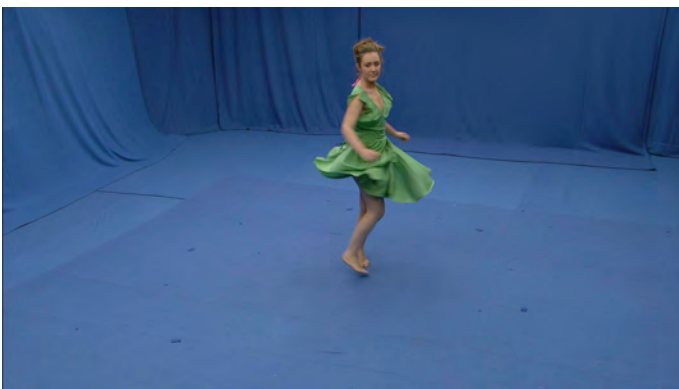
The second set of test data comprises multiple view capture with and without chroma-key backgrounds for sequences with multiple actors, foreground objects and complex backgrounds. The datasets should provide a comprehensive set of test sequences for the evaluation of algorithms developed throughout the project.

Example frames of the data sets:

First Data set



1. Tight clothing, chroma key background



2. Loose clothing, chroma key background



3. Actors interacting with chroma key background  
**Second data-set**



1. Actors with set background



2. Actor with set background



## 3D Feature Analysis & Understanding

The work in this area deals with developing algorithms that unlock the correlation between related views of the same scene to facilitate both 3D modelling and 2D motion estimation.

The Foundry started their work in this area by developing software which performs long-distance feature tracking, camera calibration and camera pose estimation, along with an associated cloud of unstructured 3d points.

Starting by implementing a conventional KLT feature tracker, they have improved the tracked results they extract by using robust estimation of the fundamental matrix in order to reject outliers and perform guided matching. For camera calibration from these tracks, they have implemented a prototype based upon 'Pollefreys' sequential structure-from-motion method.

Also using the fundamental matrix calculations, The Foundry have implemented constrained optical flow techniques which they are experimenting with as a means to producing dense point clouds. The density of the recovered data is significantly higher than the sparse 3d data-set normally extracted from a typical camera calibration output.

The team at Trinity College Dublin have been studying the robustness of SIFT-Type feature tracking. The study tries to quantitatively assess the consistency of key-point detectors. This is achieved by providing image pairs where the disparity is known (ground truth) and comparing the location of the detected key points on both frames. A good key point detector should generate key points that are situated at matching positions in both frames.

A ground truth of the University of Surrey 3D data set has been generated. It is thus possible for TCD to use this information to access the disparity ground truth between views and to test the repeatability of key point detectors on this studio quality material.



## Set, Actor & Face Modelling

The purpose of this area of research is to generate high quality representations of sets and actors from the captured data taking into account the dynamics of skin and clothing.

The UniS teams' research into matting with complex backgrounds has developed a novel region based algorithm to eliminate background shadow regions. Evaluation of this algorithm demonstrates improved performance over existing approaches to background matting. A short paper presenting the novel approach has been accepted for publication at the European Conference on Visual Media Production in November 2008.

Reconstruction of foreground structure has been conducted based on the mattes obtained with the novel approach. This provides initial coarse reconstruction of the scene geometry on a per frame basis without temporal consistency. Future work will consider structured scene modeling based on the coarse surface reconstruction.

Mesh IK [Sumner et al. SIGGRAPH'05] has been implemented for evaluation as a method to represent captured dynamic surface geometry. The approach has been extended to allow representation of meshes with unknown correspondence. Evaluation of the approach will be presented at the end of the year together with the proposed approach to structured set and actor modeling.

The Foundry have found that using the dense fundamental matrix-constrained optical flow work from the 3D Feature Analysis area, allows them to establish a dense correspondence and camera calibration solution between multiple views of a human face. More recently, The Foundry have been refining this by exploring how to extend the reconstruction to track surface features in 3d over time. To do this, They've used the fundamental (or epipolar) constraint on otherwise unconstrained optical flow in the multiple views. At present, there are two approaches being investigated.



Image 1. Original Image

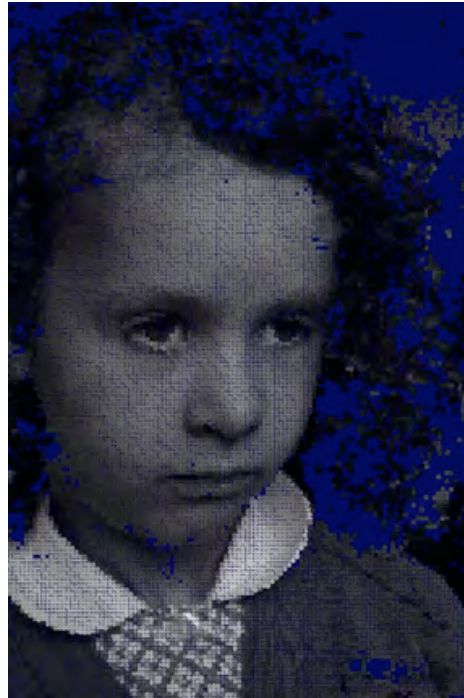


Image 2. Image from point cloud data



## Semantic Description of Multi-view Videos and 3d Scenes

One of the aims of the project is to provide a semantically rich representation of multi-view videos and 3D scenes. This will be achieved by labelling the scene elements, i.e., identifying actors and important objects (e.g. props) and tagging them with actor identity (e.g. John Smith) and object class (e.g. chair), respectively. The semantic description will be further enriched with information regarding the actors' facial expressions and performed actions or interactions. The resulting scene annotation will be used, in a proper metadata format, for the intelligent manipulation of the content and its efficient and semantic organization, search and retrieval. Within this framework, a number of research directions have been followed during the first year of the project:

### **Human body detection and pose estimation**

Novel human body detection and pose estimation methods for single- and multi-view video data have been investigated.

### **Activity recognition**

New activity recognition approaches applicable to single- and multi-view video data have been devised.

### **3D head pose estimation, frontal face detection and face recognition**

Novel approaches to 3D head pose estimation from a single uncalibrated camera and frontal face view detection in a multi-view environment have been devised. These methods are to be used as a pre-processing step to face and facial expression recognition techniques that operate on frontal views.

### **Object and object class recognition**

*A novel object and object class recognition method has been devised. The method is based on the characterization of objects by local descriptors derived on interest points.*



## Major Achievements

The technical scope of the project has been outlined above. The consortium has successfully completed its first stage of activities culminating in one commercially viable product on the market. 'Ocula' plugins are based on brand new disparity-mapping algorithms that have arisen from the research work carried out for i3DPost and allow artists to apply a multitude of adjustments to stereo image pairs. Visit [www.thefoundry.co.uk](http://www.thefoundry.co.uk) for further information.



## Promotion and Awareness

In addition to the user partners who are closely involved in i3DPost, an external user group with currently 10 members including post-production companies & games companies has been formed.

i3DPost has been publicised at media events such as IBC, SIGGRAPH & NAB and numerous conference papers and lectures have been given. (Please see website for details).

The website is frequently updated with project results and has an average of 7,000 hits per month.

## Future Work & Exploitation Prospects

Based on the work done in the first year, i3DPost will be setting out proposed modifications for the OpenFX Standard (revised scenarios and requirements; prototype interface and first point cloud software) and working towards producing Initial functioning prototypes.

Continued dissemination of the project results will be carried out at conferences and trade fairs as described in the Initial Operational Plan.

## Further Information

Please visit the project website for up to date progress of our work.

[www.i3dpost.eu](http://www.i3dpost.eu)

Or contact the Project manager: Abi Bowman

[abi@thefoundry.co.uk](mailto:abi@thefoundry.co.uk)