PROJECT PERIODIC REPORT

Grant Agreement number:316564

Project acronym: IMPART

Project title:Intelligent Management Platform for Advanced Real-Time media processes

Funding Scheme:Collaborative project

Date of latest version of Annex I against which the assessment will be made:*2012-06-19*

Periodic report: **1st☐2ndX3rd ☐ 4th ☐**

Period covered: **from 2013/11/01 to 2014/10/31**

Name, title and organisation of the scientific representative of the project's coordinator: Josep Blat, Professor of Computer Science, Universitat Pompeu Fabra

Tel: ++ 34 93 542 2172

Fax: ++ 34 93 542 2517

E-mail: josep.blat@upf.edu

Project website address: http://impart.upf.edu

3.1 Publishable summary

3.1.1 Project description

IMPART is researching, developing and evaluating intelligent information management solutions for big *data* problems in the field of digital cinema production. It intends to find new ways of managing, visualising and analysing very large multimodal data sets so that creative personnel can review three-dimensional scene representations on the set, understand the data, identify errors, evaluate the quality of the shot and take creative decisions in real time.

Approximately 20 Tb of data are created during a shooting day, integrating sources as different as 3D scanners, GPS, graphics, digital movies cameras, witness cameras, … If data and metadata were intelligently managed this could help improving the efficiency and quality of the digital cinema pipeline. The results of IMPART will be applicable to film production and also in other fields, some close, such as TV, and online video, some quite different, such as surveillance.

The partners are:

1. Universitat Pompeu Fabra, Barcelona, Coordinator
2. FilmLight Ltd, UK
3. Double Negative Ltd, UK
4. University of Surrey, UK
5. AristotelioPanepistimioThessalonikis (Aristotle University of Thessaloniki), GR
6. Vysokéučenítechnické v Brně (Brno University of Technology), CZ

The main contact is Josep Blat (josep.blat at upf.edu). The website is impart.upf.edu. The project started on November 1st 2012, and lasts 36 months. Its budget is 5 M€, and gets a EU contribution of 3.6 M€.

3.1.2 Summary of activities

On m18, the project has started its second and last phase, where technical work continues, but testing, integration and evaluation play a more prominent role. The project has a number of objectives, and half-year milestones within these objectives. The following table provides the milestones corresponding to months 18 and 24, for each of the objectives of the project.

O1 – To define the scenarios, use cases and data typologies for intelligent information management, complex multimodal data and metadata visualisation for movie production

Milestone 4, M24 Public dataset for research into multimodal movie production and data management

O2 – To research and develop intelligent metadata-driven methods for selection, retrieval and management and visualisation processes for very large and complex dataset

Milestone 3, M18 Initial quality assurance tools for multiple camera acquisition
Initial content selection tools and techniques

Milestone 4, M24 Prototype of on-set system for initial evaluation

O3 – To research and develop methods of multimodalcontent and metadata processing to deliver real-time support for decision-making in dynamic three-dimensional scenes

Milestone 3, M18 Initial algorithms for multimodal data registration and synchronisation,metadata extraction, association and propagation

Milestone 4, M24	Methods for 3D scene handling and rendering based on knowledge management

O4 – To develop the architecture and acceleration methods for managing and processing the very large multimodal datasets associated with multidimensional scenes and action

Milestone 3, M18	Initial data management platform architecture and file structure
Milestone 4, M24	Analysis of the acceleration of selected algorithms, platforms, andproposal for platforms; definition of the integration architecture and performance metrics

O5 – To test and evaluate the integrated system, tools and applications in real working environments, with professional users and typical very large multimodal data and metadata volumes from on-set capture

Milestone 3, M18	Scenarios for the experimental technology trials, with an evaluation plan
Milestone 4, M24	Report on formative evaluation; Interim report on integration, testing and validation

The project is now well into *testing*, *integration*, and (formative) *evaluation*. It continues improving and deepening the research and technical work. Let us summarize briefly the main achievements along different lines.

3.1.3 Results in different areas

3.1.3.1 Research dataset & Requirements

In October 2014 the IMPART project published a public dataset for research into multimodal movie production created by the University of Surrey and Double Negative. It consists of about 20TB of 2D and 3D data and metadata captured during sessions at different locations at the University of Surrey and Double Negative premises in both indoor and outdoor environments.

The dataset is free to use for research purposes as long as the source is referenced and the data is not redistributed. Licensing conditions are described in detail on the website cvssp.org/impart.

The dataset is carefully documented. Capture devices, capture environments, shoot parameters, calibration procedures and custom data formats are described in the capture notes available on the website.

Making the dataset publicly available is important because it enables researchers to develop new algorithms and compare their performance against a common dataset. A shared dataset also facilitates the evaluation and verification of research by other groups. Additionally, for smaller institutions it may not be feasible to create a comparable dataset of their own, which makes research into some of the topic relevant to IMPART impossible. Last but not least, a public dataset increases the visibility of the IMPART project, particularly in the research community.

Compared to other publicly available research data[1], the IMPART dataset provides a larger total amount of data, captured using more cameras and with better quality equipment. Unlike previous datasets, it consists of

[1] INRIA XmasMotionAcquisitionSequences (IXMAS) http://4drepository.inrialpes.fr/public/viewgroup/6
University of Surrey 3D Video Databasehttp://kahlan.eps.surrey.ac.uk/cvssp3d
i3DPost Multi-view Human ActionDatasets, University of Surrey, http://kahlan.eps.surrey.ac.uk/i3dpost_action/
HumanEva, Max Planck InstituteforIntelligentSystems, http://humaneva.is.tue.mpg.de
MOBISERV-AIIA, University of Thessaloniki, http://www.aiia.csd.auth.gr/MOBISERV-AIIA

a large number of subjects captured both indoors and outdoors, with both simple and complex backgrounds and additionally provides LIDAR scans for the sets. All of these factors combined make it a uniquely useful dataset for research purposes.

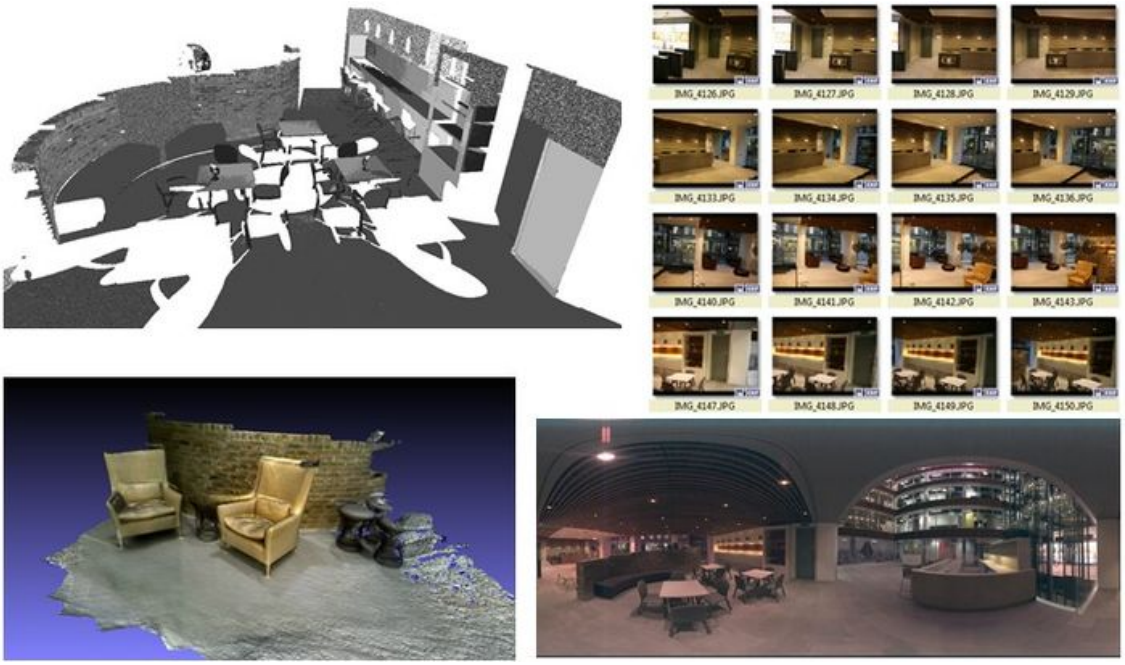

Figure 1 Screenshot of indoor capture "Lobby"

3.1.3.2 Registration, synchronisation

The second year of the project, continuing the research of the first one, the project has been implementing multimodal data registration techniques to manage various 2D and 3D data from different capture devices by 2D/3D feature detection and description. This has led to the definition of a general goal of this part of the project, which is to create a “common 3D space” for all the on-set recorded data.

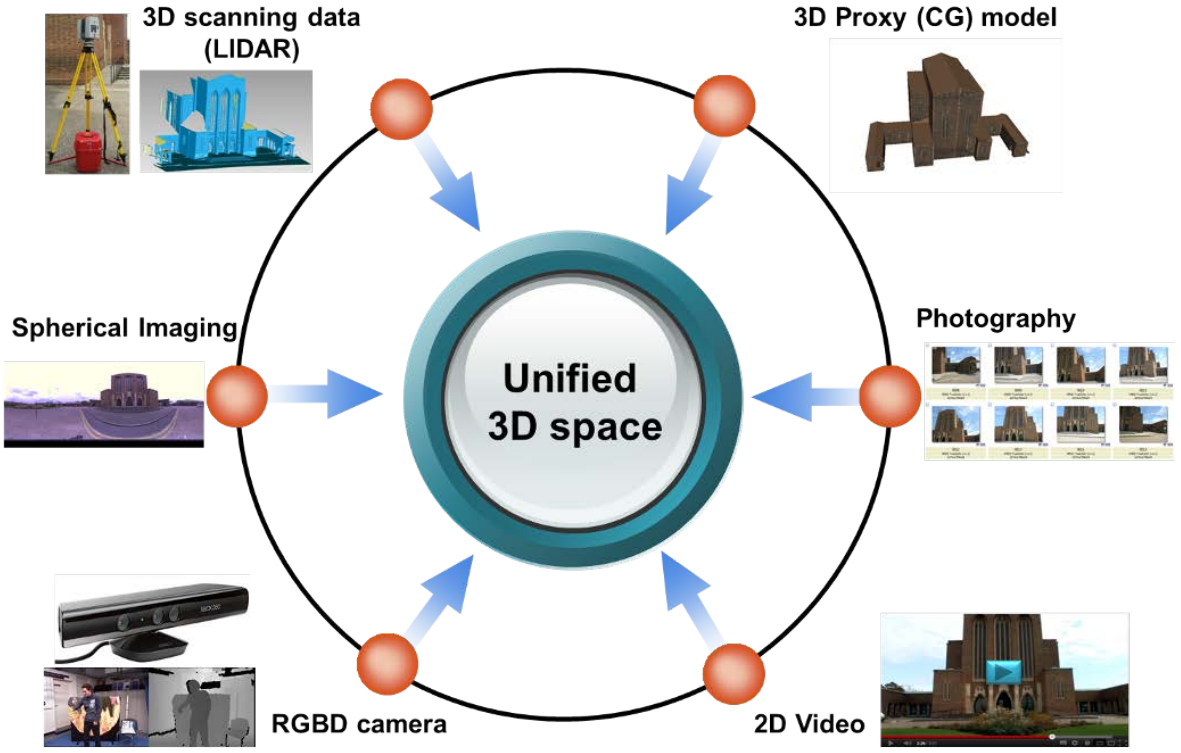

Figure 2 The unified 3D space as envisaged by IMPART

The algorithms and tools developed by the academic partners have been integrated into the *Jigsaw* application of Double Negative. The workflow is largely fully automated, yet allows the user to step in and influence the key steps of the algorithms if necessary.

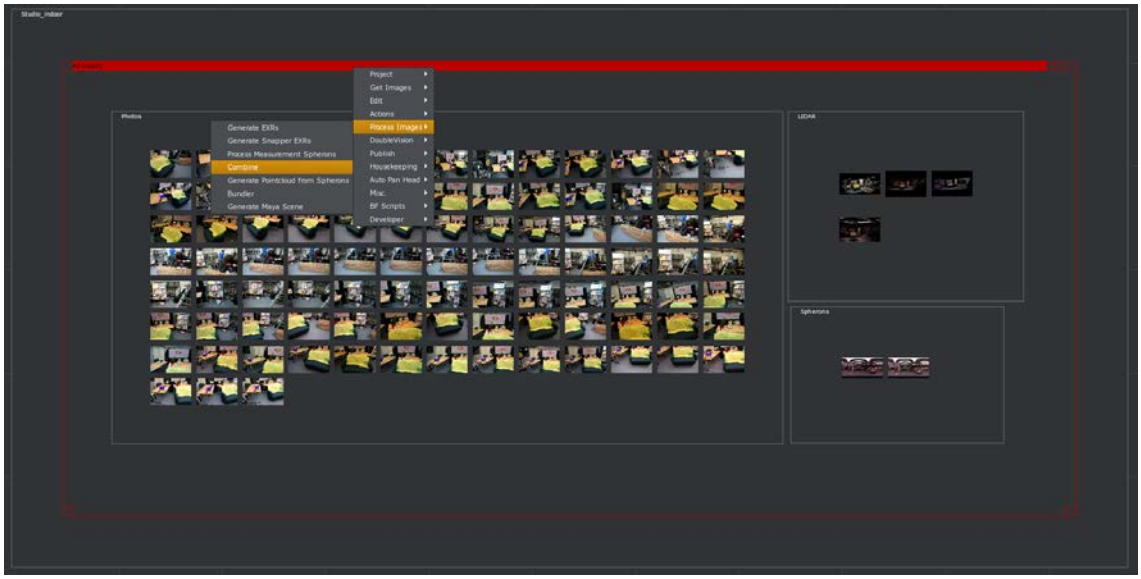

Figure 3 LIDAR datasets, Spheron images and still images organised in a Jigsaw processing group

Depending on the amount of input data, the registration process is generally fast enough to run on a laptop computer, which is critical since it allows for the use in an on-set environment. Very large amounts of input

data can be handled by offloading the most expensive computations to a render farm, where they can be processed in parallel.

The integration of fully automated tools for registration of various types of input data is of high importance to Double Negative. Previously, merging for example LIDAR data captured on set was a slow and largely manual process that tied highly skilled artists down for several days. Being able to compute a good registration automatically improves this workflow significantly and initial estimates show that it has the potential to cut the processing times in half.

Some publications related to this work are:

E. Imre, A. Hilton, Covariance estimation for minimal geometry solvers via scaled unscented transformation,*Computer Vision and Image Understanding*, in press.

E. Imre, A. Hilton, Order Statistics of RANSAC and Their Practical Applications, *International Journal of Computer Vision*, in press.

H.Kim and A. Hilton, Hybrid 3D Feature Description and Matching for Multi-modal Data Registration,*IEEE International Conference on Image Processing (ICIP)*, France, 2014

H.Kim and A. Hilton, Influence of Colour and Feature Geometry on Multi-modal 3D Point Clouds Data Registration,*International Conference on 3D Vision (3DV)*, Japan, 2014

A. Mustafa, H. Kim, E. Imre, A. Hilton, Initial Disparity Estimation using Sparse matching for Wide-baseline Dense Stereo, in *11th European Conference on Visual Media Production (CVMP)*, UK, 2014

M. Brown, J-Y. Guillemaut, D. Windridge.A Saliency-Based Framework for 2D / 3D Registration", In *Proceedings of the 9th International Conference on Computer Vision Theory and Applications (VISAPP)*, Portugal, 2014.

L. Polok, V. Ila, P. Smrz, Fast Radix Sort for Sparse Linear Algebra on GPU,*22nd High Performance Computing Symposia*, Tampa, US, 2014.

Figure 4 Reconstructed 3D scene based on a single Spheron pair

3.1.3.3 Metadata on human action recognition (Content selection)

IMPART partner AUTH has achieved considerable progress in development of intelligent tools which will assist with the quick review of audio-visual material and the video editing decision making process, and also facilitate the operation of a platform, which will be able to answer video content related queries and retrieve relevant material. The work focused on video summarization techniques and the acceleration of software tools, so that they will be able to run fast on the set, or be able to handle Big Data for offline processing tasks. Video summarization is approached through clustering techniques, face recognition, shot saliency estimation and key frame extraction. Additionally, some techniques or technique components have received a distributed computing, multithreaded, or GPU implementation, in order to take advantage of the processing power of several computers, multiple CPU processing cores and specialized hardware present on graphics cards, respectively.

Evaluation shows that performance of the devised techniques is, in general, competitive with the state of the art. Furthermore, the new dataset published by the project is particularly relevant for this work, as previous datasets, such as the INRIA Xmas Motion dataset mentioned above, were becoming ‘saturated’ (in the sense that analysis techniques were consistently achieving accuracy results>95%). The new IMPART dataset is considerable more realistic and challenging. Some recent publications related to this research are:

A. Iosifidis, A. Tefas and I. Pitas, Regularized Extreme Learning Machine for Multi-view Semi-supervised Action Recognition, *Neurocomputing*, **145**, pp. 250-262, 2014

A. Iosifidis, A. Tefas and I. Pitas, Discriminant Bag of Words based Representation for Human Action Recognition, *Pattern Recognition Letters*, **49**, pp. 185-192, 2014.

A. Iosifidis, A. Tefas and I. Pitas, Kernel Reference Discriminant Analysis, *Pattern Recognition Letters*, **49**, pp. 85-91, 2014.

K. Papachristou, A. Tefas and I. Pitas, Symmetric Subspace Learning for Image Analysis, *IEEE Transactions on Image Processing*, In Press

Figure 4 Screenshot of video summarising AUTH's results, (video can be found on IMPART website)

3.1.3.4 Management, Browsing and annotating, Architectures

In the modern production environment, the teams attempt to store as much information as possible about the surrounding set and background. UPF has been developing a prototype system for remote visualisation of data recorded on-set, such as the LIDAR scan of the environment and the camera footage. The goal is that this visualisation could be integrated with content analysis and camera coverage work in order to have an holistic view of all the data recorded on-set.

One of the challenges of remote visualisation of production data, particularly when moving into the unified 3D space as described above, is the sheer size of the files produced. The file sizes of LIDAR scans, for example, typically range from 100MB to 600MB. Downloading this data for remote view, for example by tablet or on a web-browser, usually involves an unacceptable initial waiting time. As a result, UPF has been working on a system for progressive transfer of LIDAR data.

By placing the data into a hierarchical structure, the scene can be transmitted and visualised progressively, permitting the user to see quickly a lower resolution version of the scene, which is then refined as more data is downloaded.

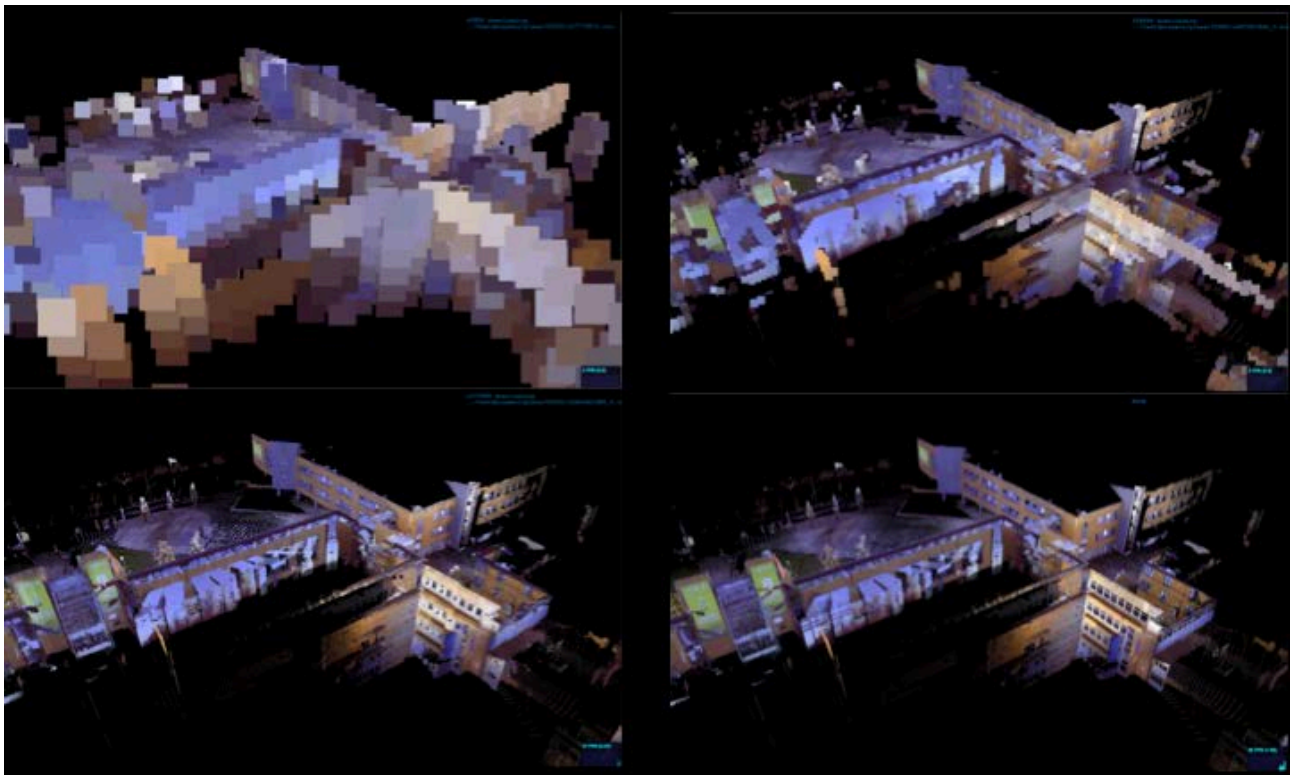

Figure 5 Progressive refinement of LIDAR data recorded onset, visualised in the web-browser with WebGL

UPF has also been working on novel ways of viewing data recorded onset using virtual reality, either via tablet (using a combination of marker based and accelerometer based tracking) or via virtual reality headset, such as the Oculus Rift.

Figure 6 Visualisation of the IMPART dataset via tablet and Oculus Rift virtual reality headset

The data recorded on-set typically measure in the Terabytes and, thus, can be stored in multiple physical locations. FilmLight's research within IMPART is focused on the creation of a *homogenous* view of all data recorded on-set, regardless of its physical location, thus creating a heteregenous "cloud" of file systems. Such organisation requires fast and stable access to metadata, and also the ability to rapidly augment metadata pertaining to files or sequences. FilmLight has been extending its existing storage product, called FLUX, in a new direction, called FLUX+, whose architecture has been designed to grow with multi-Petabyte storage infrastructure. A prototype system is now being tested to validate the individual aspects of the system

and integration of potential improvements. This prototype system utilises the heterogeneous storage network at FilmLight head office, which extends to some hundreds of Terabytes of storage, spread over tens of devices. FilmLight has also been developing its software interfaces in order to efficiently manage this vast amount of data.

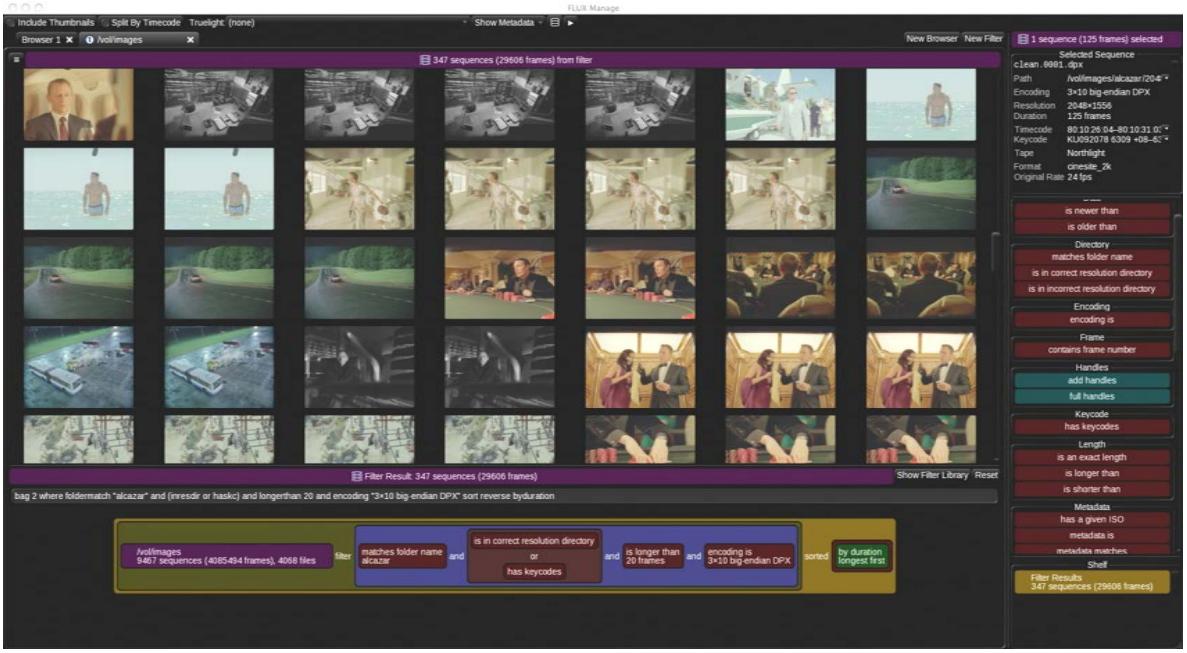

Figure 7 Browser application for FLUX+, using metadata to efficiently find and sort through data

BUT has continued its work towards accelerating time-consuming algorithms employed in the production pipeline using different kinds of hardware. The algorithms are related to key point detection and matching, HDR and spherical capture processing, bilateral filtering, spherical stereo depth map calculation, spherical 3D reconstruction, sparse block matrix library, marginal covariance computation, Poisson 3D reconstruction, and 3D rendering. Significant gains have been achieved, and two of the implementations are publicly available under an open source license, on http://sourceforge.net/p/slam-plus-plus/, and http://www.fit.vutbr.cz/research/ prod/index.php?id=302&notitle=1.

Some publications related to this research are:

A. Evans, M.Romeo, A.Bahremand, J.Agenjo, J.Blat. 3D Graphics on the Web: A Survey. *Computers & Graphics***41**, June 2014, Pages 43-61. ISSN 0097-8493.

A.Evans, J.Agenjo, J.BlatWeb-based Visualisation of On-set Pointcloud Data. To appear at the *11th European Conference on Visual Media Production (CVMP2014)*, London, England (November 2014)

G.Llorach, A.Evans, J.BlatSimulator Sickness and Presence using HMDs: comparing use of a game controller and a position estimation system. To appear at the *20th ACM Symposium on Virtual Reality Software and Technology (VRST2014)*, Edinburgh, Scotland (November 2014)

A.Evans, J.Agenjo, J.Blat. Designing a multiplatform pipeline for 3D scenes. *CISTI 2014*, Barcelona, Spain (June 2014).

G.Llorach, A.Evans, J.Agenjo, J.Blat. Position estimation with a low-cost inertial measurement unit.*CISTI 2014*, Barcelona (2014)

A.Evans, J.Agenjo, J.Blat. Variable Penumbra Shadow Mapping for Mobile Devices.*GRAPP 2014*, Lisbon, Portugal (2014)

3.1.4 Dissemination, Promotion and Awareness

3.1.4.1 Local Industry Meetings

TheIMPART project and its intermediate results were discussed and presented during meetings with local industry and project meetings with industrial participation.

BeerGraph is a London R&D Film Post-Production meeting, and this time was held by Double Negative. The aims of IMPART were discussed amongst software vendors and VFX Houses based in London.

The project was also introduced to the several companies having specific interest in large data with local branches in Brno, Czech Republic, such as:RedHat (RedHat has its largest worldwide engineering operation in Brno), Honeywell (Aerospace research), Seznam CZ (Czech web search company). Moreover, the project was introduced to companies with application interest in large data and/or film industry applications:UPP – Universal Production Partners (Czech film production company), UNIS (local company interested both in large data and industrial video applications), CAMEA (local company with interest in industrial and traffic applications of video)

Finally, the IMPART project was also presented at events with significant industrial coverage, such as:

- IT4Innovations meetings/conferences (IT4Innovations is Czech center of excellence in supercomputing funded by the EU structural funds)
- COST 1005 HDRi meeting (COST 1005 HDRi action is COST action focused on high dynamic range acquisition, processing, storage, and visualization with significant industrial presence)
- V3C – Visual Computing Competence Center meetings (V3C is the Cezch Technology Agency funded center of competence)

3.1.4.2 Conferences, Tradefairs and Talks

Double Negative presented, at SIGGRAPH, a Birds of a Feather session "Recent Advances in Computational Videography for Entertainment Production". The session focussed on research developing flexible technologies to combine real and computer generated imagery. Results were discussed from 3 EU funded research projects: SCENE, Dreamspace& IMPART.

IMPART had a major presence at the CVMP conference this year (http://www.cvmp-conference.org/), in the form of:

- an invited talk (“Special Session into EU projects”)
- a demo session, presenting the latest results of the project and also promoting the public dataset
- one full paper and two short papers

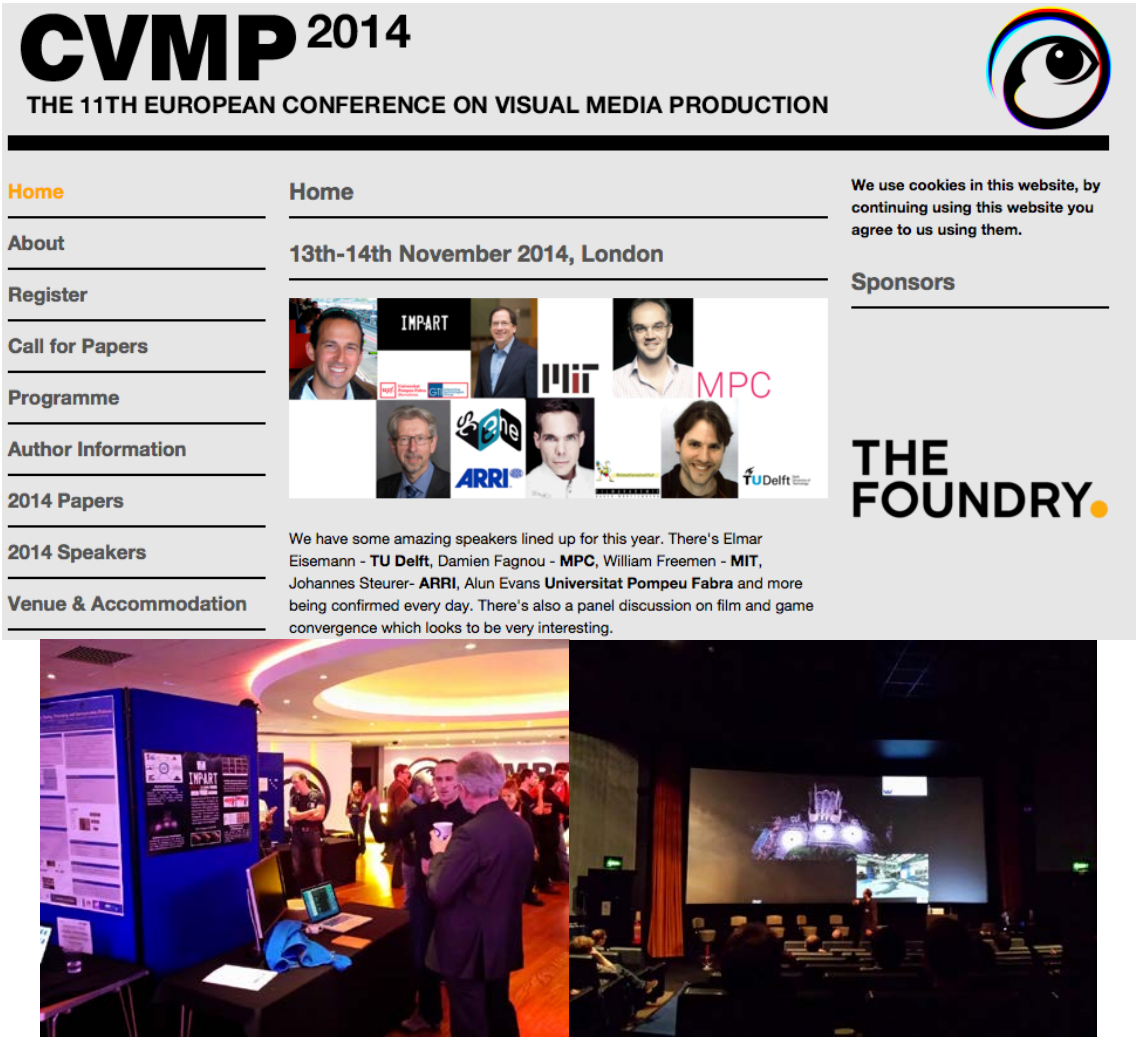

Figure 7 IMPART's strong presence at CVMP 2014

Furthermore, UPF gave an **invited talk**at the B'Ars – Barcelona Arts and VFX Fair (http://barsvfx.com/) in Barcelona in June.

AUTH gave an **invited talk**"Intelligent Digital Media Analysis and Description", at the Euro-China Conference on Intelligent Data Analysis and Applications (ECC),Shezhen, China, also in June.

3.1.4.3 Academic publications

The project partners published 7 journal papers and 24 conference papers during the second year of the project. The list of publications can be found on the IMPART website, under 'Media'.

3.1.4.4 Website and other

During 2014 the project website http://impart.upf.edu/ underwent a restructuring in order to make it more accessible and easier to maintain. The home page structure was updated, along with partner and project descriptions.

The Technology Transfer Office at the Brno University of Technology prepared a case study about the IMPART project. The result is a folding leaflet, detailing the work being done in the project, which was printed in circulation of 1000 copies.

Parallel to the 6th IMPART Technical Meeting in Brno, a project presentation video was recorded. The university PR crew interviewed the meeting participants about their research and the new tools which are being developed in IMPART.