<table>
<tr><td>Project acronym: POPART

Project number: 644874

Work package: Management

Deliverable number and name:

D6.3: Quarterly Management Report 3</td><td>Title: Quarterly Management Report 3

Work Package: WP6

Version: 1
Date: October 1, 2015

Author:
Håvard Espeland</td></tr>
<tr><td>Type:
[X] Report
[ ] Demonstrator, pilot, prototype
[ ] Website, patent filings, videos, etc.
[ ] Other</td><td>Co-Author(s):
All partners

To:
Albert Gauthier, Project Officer</td></tr>
<tr><td>Status:
[ ] Draft
[ ] To be reviewed
[ ] Proposal
[X] Final / Released to EC</td><td>Confidentiality:
[X] PU – Public
[ ] CO – Confidential
[ ] CL - Classified</td></tr>
<tr><td colspan="2">Revision:</td></tr>
<tr><td colspan="2">Contents:
Deliverable 6.3: Quarterly Management Report 3</td></tr>
</table>

Human Resources

SRL

The project is coordinated by Håvard Espeland. Carsten Griwodz, Lilian Calvet contribute to the innovation and scientific work. Elisabeth Andersen contributes to the administrative work.

LAB

The project at LAB is led by Magne Eimot. Jon M. Puntervold, Gunleik Groven and Håvard Espeland also contribute to the innovation work. Kristian Skarseth has recently joined LAB as a developer.

MIK

The project at MIK is led by Benoit Maujean, R&D Manager at Mikros Image. Fabien Castan is the technical manager of Popart project; he is responsible for the 3D reconstruction library (OpenMVG) and photomodeling plugin for Autodesk Maya (MayaMVG) at Mikros Image. Elisa Prana, Nicolas Rondaud and Cyril Pichard, as software engineers contribute to the development of Popart tools, including the expected adaptations of OpenMVG and MayaMVG. Furthermore, some Matte Painters of Mikros Image studio have contributed to the generation of the virtual dataset.

IRT

The project at IRT is led by Vincent Charvillat. Simone Gasparini is responsible for the Work package 1 and he's coordinating the tasks assigned to IRT, in particular on camera localization and the contribution to the OpenMVG libraries. Sylvie Chambon is also a staff member of the project and she will participate in the task involving the image matching for improving the performances of OpenMVG. Clement Aymard is a developer engineer who is developing the different contributions to the project. Since September 21st, Jean Melou has been hired as engineer to work on OpenMVG pipeline.

BAN

The project at BAN is led by Christopher Hantel, CEO at Band Pro Munich. He is joined by Randy Wedick, contributing to the commercialization and productization of the POPART commercial product.

Person Months Contributed to the Project

Participant	Names of staff	WP1		WP2		WP3		WP4		WP5		WP6		Total	
		Planned	Actual	Planned	Actual	Planned	Actual	Planned	Actual	Planned	Actual	Planned	Actual	Planned	Actual
SRL	Håvard Espeland Lilian Calvet Carsten Griwodz Elisabeth Andersen	4,5	2,57	15,54	9,29	0	0	0	0	0	0	4,5	4,75	24,54	16,61
LAB	Magne Eimot Jon M. Puntervold Håvard Espeland Gunleik Groven Kristian Skarseth	0	0	0	0	6,6	10,44	11,1	7.93	8,86	0	0	0	19,33	19,3
MIK	Benoit Maujean Fabien Castan Elisa Prana Nicolas Rondaud Cyril Pichard	3,38	1,24	1,64	0	11,70	10,40	6,60	7,55	0	0	0	0	23,31	19,19
IRT	Simone Gasparini Vincent Charvillat Clement Aymard Sylvie Chambon Nicolas Bertrand	13,87	11.45	4,90	5.14	0	0	1,8	0	0	0	0	0	20,58	16.59
BAN	Randy Wedick Christopher Hantel	0	0	0	0	0	0	0	0	0,75	0	0	0	0	0
Total		**21,76**	**15,26**	**22,09**	**14,43**	**18,30**	**20,84**	**19,5**	**16,41**	**2,37**	**0**	**4,5**	**4,75**	**88,52**	**71,69**

Figure 1: The cumulative budgeted and contributed person-months since project start for each partner sorted by Work Package.

Personnel effort is in mostly in accordance with the budget. Other direct costs included, during the third quarter, acquisition of equipment. SRL have purchased witness cameras and acquisition system (HD-SDI) for developing functional prototypes. Equipment was foreseen in the proposal to be acquired by LAB and MIK, but the need to store and analyze extra data was identified since the project has started. These additional items do not represent a major cost, with total amount of around € 3000.

Project Progress

The project is running as planned and we are working on all tasks expected to implement POPART and its objectives. Details on each work package is given below.

WP1 - 3D Visual Database

Reconstruction time

The main focus was to improve the reconstruction times. The reconstruction done in production at MIK have shown that the reconstruction times vary considerably from one dataset to another depending on the image content (highly or poorly textured). The average overall reconstruction time were also an important drawback of the software.

In strong collaboration between IRT, MIK and SRL, we have changed the different thresholds used in features extraction to ensure stable reconstruction times regarding the number of input images. In order to limit the number of features, we have introduce a criterion to ensure repartition of features in images. This also improve results on poorly textured datasets in conjunction with a guided matching method.

One important performance problem was the matching between all images which means a complexity of O(N^2). The developments done at IRT on the vocabulary tree approach have successfully reduced this complexity to linear O(N) regarding the number of input images. After some prototyping steps, the efficiency of the vocabulary tree has been validated, so we came to a new phase of integration. We moved the code of cameraLocalization (https://github.com/poparteu/cameraLocalization) directly inside OpenMVG. The objectives of this move is to simplify testing, ease long term integration and maintenance by avoiding code redundancy and different data formats. While the former is still mantained in the original repository, the new code is under development and publicly available at https://github.com/poparteu/openMVG, mostly in the "voctree" branch. This voctree is now in production for 3D reconstruction at MIK and drastically reduce the overall time.

Benchmarks

MIK has benchmarked the 3D reconstruction pipeline, with a particular focus on the Incremental method. More extensive tests have been carried out to asses the quality of reconstruction and the computational cost when changing the parameters of the reconstructions. In particular, tests have focused on comparing the results of different reconstructions carried out with different feature settings (the number of extracted features per image) and different matching methods: the classical brute force matching process, which matches each image with all the other images of the dataset (complexity O(N^2)), has been compared with the generic vocabulary tree approach described in the previous report. For the purpose of this benchmark we used a dataset of 596 images from "Levallois' Town Hall", a dataset that we have publicly released in May.

Using a large visual vocabulary trained with 15000 random images downloaded from Flickr, for each image of the dataset we then get a list of N (usually 30 or 60) best matching images, with which we fed the OpenMVG incremental pipeline, thus skipping the brute force image matching process and limiting the match to the list obtained from the vocabulary tree.

The table below summarizes some of the results:

Settings	Feature type	# of features	# of images	Feature extraction	VocMatch	Matching	SfM	Residual mean	Residual max	Nb vertices	Nb cameras	Overall time
Mairie CG - 0.7	VLFeat	~3k-18k	596	01:12:56		06:19:36	**07:29:09**	0.222549	3.66167	12079	**547**	**15:01:41**
Mairie CG - 0.8												
Mairie CG - 0.8 - OCV - LOW	OpenCV	500	596	03:40:00		00:16:00	00:00:28	0.314834	3.66057	300	21	03:56:28
Mairie CG - 0.8 - OCV - MEDIUM	OpenCV	1000	596	03:40:00		00:18:00	00:02:10	0.247462	3.95137	8061	262	04:00:10
Mairie CG - 0.8 - OCV - NORMAL	OpenCV	2000	596	03:40:00		00:32:00	00:20:16	0.22284	3.98259	33608	480	04:32:16
Mairie CG - 0.8 - OCV - HIGH	OpenCV	10000	596	03:40:00		02:48:00	**04:02:10**	0.200715	4,00	199743	**577**	**10:30:10**
Mairie CG - 0.8 - OCV - ULTRA	OpenCV	20000	596	03:40:00		07:12:00	07:11:00	0.193472	3.98826	407493	581	18:03:00
Mairie CG - 0.8 - VLFEAT - VoctreeL2 30	VLFeat	~3k-18k	596	00:53:36	02:29:41	00:36:00	04:57:00	0.197561	3.98557	267994	572	08:56:17
Mairie CG - 0.8 - VoctreeL2 30 - OCV - LOW	OpenCV	500	596	03:40:00	00:08:44	00:07:36	00:00:22	0.298145	3.75371	1262	61	03:56:42
Mairie CG - 0.8 - VoctreeL2 30 - OCV - MEDIUM	*OpenCV*	*1000*	*596*	*03:40:00*	*00:00:00*	*00:00:00*	*00:00:00*					*03:40:00*
Mairie CG - 0.8 - VoctreeL2 30 - OCV - NORMAL	OpenCV	2000	596	03:40:00	00:28:45	00:12:24	00:33:35	0.217478	3.98981	41900	482	04:54:44
Mairie CG - 0.8 - VoctreeL2 30 - OCV - HIGH	OpenCV	10000	596	03:40:00	02:16:16	00:30:00	04:58:58	0.198768	3.98792	234298	577	11:25:14
Mairie CG - 0.8 - VoctreeL2 30 - OCV - ULTRA	OpenCV	20000	596	03:40:00	04:26:57	00:46:00	09:48:47	0.191386	3.99459	465179	581	18:41:44
Mairie CG - 0.8 - VoctreeL2 60 - OCV - ULTRA	OpenCV	20000	596	03:40:00	04:30:32	01:28:00	09:03:44	0.192659	3.98852	444274	580	18:42:16
Mairie CG - 0.8 - VoctreeL6 30 - OCV - ULTRA	OpenCV	20000	596	03:40:00	00:30:11	00:56:00	08:39:15	0.19114	3.99408	468585	**580**	13:45:26
Mairie CG - 0.8 - VoctreeL4 30 - OCV - ULTRA	OpenCV	20000	596	03:40:00	00:33:13	00:48:00	08:02:35	0.191376	3.98903	467466	**578**	13:03:48
Mairie CG - 0.8 - VoctreeL4 30 - OCV - LOW - gui	OpenCV	500	596	03:40:00	00:03:01	00:07:12	00:00:31	0.294556	3.82729	1741	84	03:50:44
Mairie CG - 0.8 - VoctreeL4 30 - OCV - NORMAL	OpenCV	2000	596	03:40:00	00:05:11	00:12:00	00:29:29	0.219475	3.99188	41914	478	04:26:40
Mairie CG - 0.8 - VoctreeL4 30 - OCV - HIGH - gu	OpenCV	10000	596	03:40:00	00:18:10	00:42:00	04:31:10	0.198123	3,99	234467	575	09:11:20
Mairie CG - 0.8 - VoctreeL4 30 - OCV - ULTRA - g	OpenCV	20000	596	03:40:00	00:33:13	01:36:00	09:13:31	0.191569	3.98351	464765	582	15:02:44

Table 1 - A comparative table for reconstruction quality and time consumption with different settings and algorithms used for matching.

For the benchmarking we considered different SIFT implementations, OpenCV and VLFEAT, with different settings: the presets LOW, MEDIUM, NORMAL, HIGH and ULTRA refers to the number of extracted features per image. As the number of features increase, the computational cost for matching and performing the SfM increase dramatically, yet giving more accurate results in terms of the number of reconstructed images. Moreover we can observe that the feature extraction part, no matter the chosen preset, takes a large percentage of the processing time as more than 3h are usually taken just to extract the features. Hopefully this problem will be mitigated once we will introduce the GPU implementation of the SIFT extractor (PopSift, currently under development by SRL).
Figure 1 gives an overview of this benchmarking, in which we can see that the whole process may take up to 7h when an higher number of features is chosen. The chart compares the overall processing time when using the brute force method and the method based on the vocabulary tree. We can see that the vocabulary tree method can improve the time consumption when many features are considered for the reconstruction.

Vocabulary tree

One of the limits of the vocabulary tree is that a vocabulary tree with a branching factor of K and a number of levels L always requires K*L comparisons for each image feature. In order to reduce the number a comparison and improve the time consumption different values of K and L have been considered.

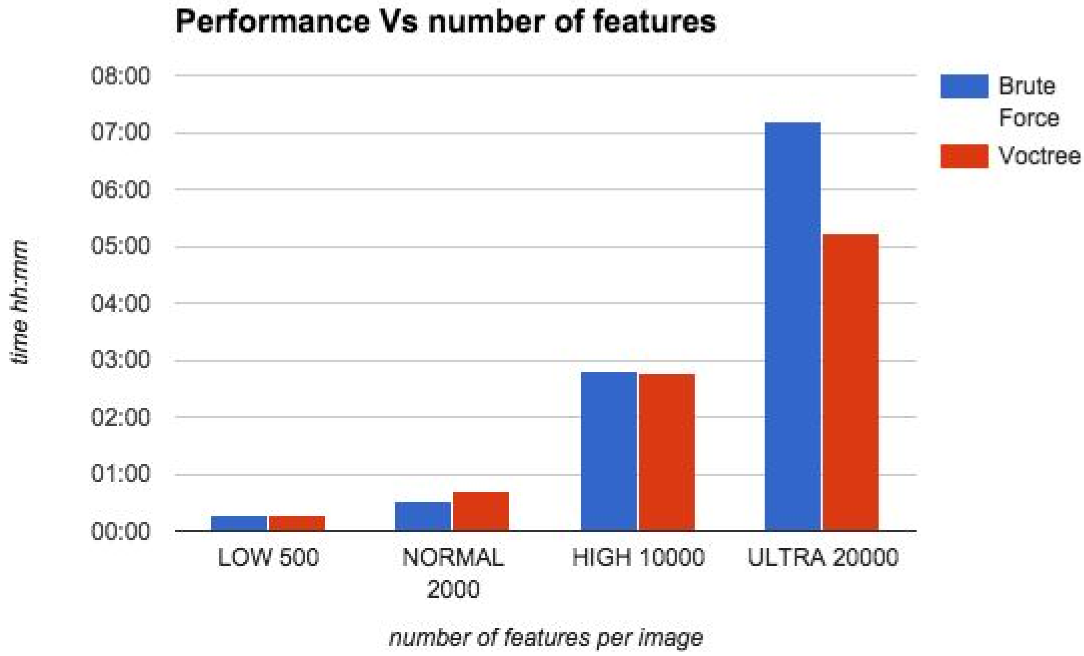


Figure 1 - Comparison of the overall reconstruction time using the brute force method and the voctree approach, with different settings for the number of extracted features.

We trained vocabulary trees with different values of L and K but maintaining a similar number K^L of visual words (ie, such that for each $i \neq j\ K_i^{L_i} \simeq K_j^{L_j}$). The choice of L and K, assuming that K^L is constant, can change the number of comparisons needed in order to assign to each feature a corresponding visual word, namely the number of comparison required is K*L. This is also another source for improving the matching time as showed by the Figure 2 where the overall matching time (vocmatch+matching) is compared using different choices of K and L. Note that final result in terms of number of reconstructed images and reprojection errors for the different settings is still comparable.

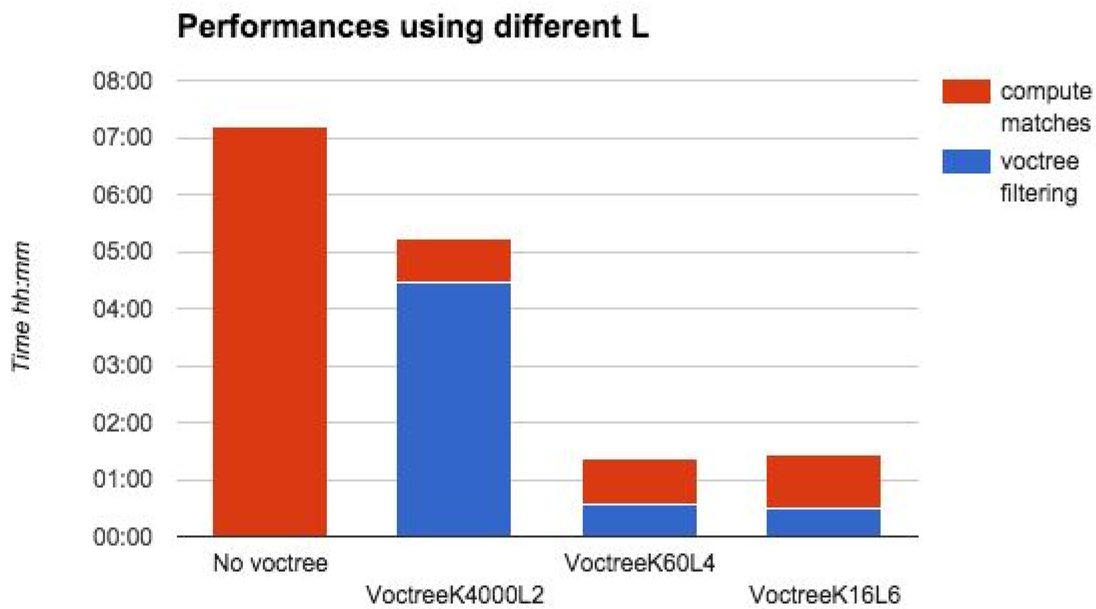


Figure 2 - The total matching time (vocmatch+matching) when changing the values of K and L (so that $K^L \simeq 16M$ visual words)

WP2 - Real-time camera tracking

Multi-camera system calibration (SRL)

The multi-camera system calibration has been implemented.
Such a calibration is performed in order to introduce the rigidity constraints related to the positions of the witness cameras, rigidly fixed on the camera rig, during the camera tracking. These constraints consist in constant relative poses between the witness cameras and the main camera, a data then used during the camera tracking to transfer all the visible 3D points in all the video streams into a single coordinate system in order to perform the camera system resectioning.
The calibration procedure has been implemented as follows:

1. Robust estimation of the relative pose of every camera-pair “witness/main” cameras,
2. Bundle adjustment over all the N-tuples of images (extracted from the synchronised video streams) where N representants the total number of cameras, i.e. in configuration, a maximum of 2 witness cameras plus the main camera.

The first step (1) is decomposed as follows:

- For a given witness/main camera-pair, compute independently the pose of the witness camera and the main camera with respect to coordinate system of the 3D visual database for all the synchronised image-pairs over the time t=1..T.
- For a given witness/main camera-pair, compute all relative pose between the witness camera and the main camera over the time t=1..T.
- For a given witness/main camera-pair, over all the obtained relative poses over t, compute the one minimizing the reprojection error in its associated image-pair.

The step (2) is stated as a nonlinear least squares problem whose function is to minimize the sum of the reprojection errors over all the N-tuples of images. The optimization is performed through the google ceres solver (via the levenberg-marquardt algorithm) while the initial solution consists of the set of initial solutions delivered by the step (1).
The algorithm has currently been tested on a real dataset ( ) and has performed well (in terms of reprojection errors). However, extensive tests on synthetic data are still required.

Artificial feature detection (SRL)
The GPU and CPU implementation of the CCTag detector are now communicating. The transfer of intermediate results from GPU to CPU occurs after the voting stage, which involves all potential edge point candidates and results in a selection of a subset of points that fulfill the ratio requirement between outer and inner points of a CCTag ellipse. The resulting subset of points is then connected into ellipses on the CPU. Moving also this “edge linking” step to the GPU would be feasible for large numbers of concurrently visible markers, but the required data structures are not suitable for efficient GPU processing. Although a preliminary version of GPU code exists, it is currently not used.

Intermediate results have been verified visually to determine differences between CPU and GPU implementations’ results, and it is apparent that the non-deterministic order of edge point identification on the GPU can occasionally lead to a drift by 1 pixel of an chosen edge point. The infrequency of this means that it is highly unlikely that a detected CCTag’s position is reported with the maximum 1-pixel difference between implementations.

Additional tests on synthetic data have been implemented in order to compare the results in terms of the detection rate, identification rate and the imaged center localization accuracy delivered by the different implementation/optimization of the code.
The optimization of the imaged center on the last cpu optimization still deliver slightly worse results than the non optimized implementation and requires more work as the identification quality heavily relies on this result.

Natural feature detection
Tests have been carried out in indoor to evaluate the impact of the sampling used to store the SIFT descriptors. A descriptor made up of 128 times 1 byte (char) has been evaluated.
A number of 4686 image-pairs have been considered for which 390,174,662 feature correspondences have been tested.

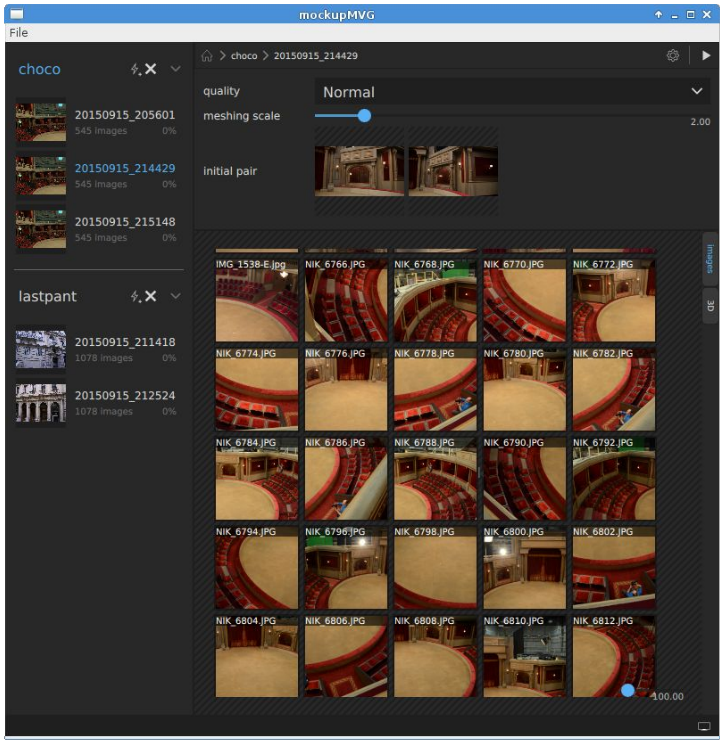
The tests demonstrate a percentage of 0.00756% of ambiguous features correspondence which is insignificant regarding the subsequent robust estimation of the fundamental/essential matrix.

Image matching optimization (SRL)
Some optimizations are currently performed on the feature matching procedure which aims at speeding up both the camera localization and the Structure-from-Motion step. The idea is that, instead of considering all the feature correspondences within an image-pair, to sort the set of feature correspondences and to keep the best ones in order to reduce the complexity of the subsequent Structure-from-Motion or camera localization by reducing the number of 3D points considered.
Most part of the work focuses on the criterion used to sort the feature correspondences for which our preliminary idea is to keep the ones associated to the largest features in both images.

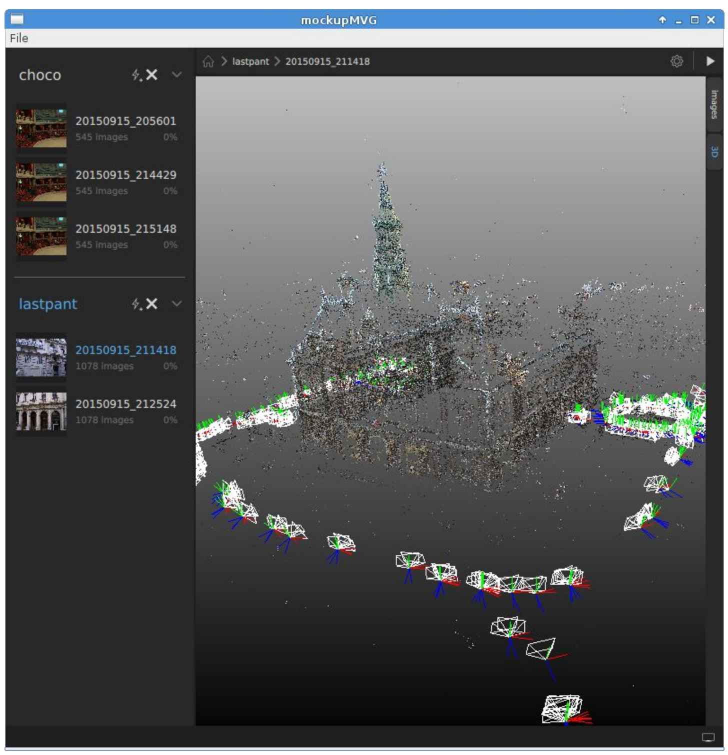
Query from the database (IRT)
Since we decide to port the code directly inside the OpenMVG library, most of the effort has been done to clean and move the code. The porting is currently still going on and aims at having the same algorithms implemented using the OpenMVG environment.
We are focusing on the off-line camera tracker that could be used in post-production to track a sequence of images with respect to an existing 3D reconstruction of the scene. We have some preliminary results on simple datasets and we are working to make it functional on real-world use cases.

WP3 - Shoot preparation, visualization and post production

Reconstruction tool
MIK is working on the 3D reconstruction application and iterate with graphic artists in production to refine the usability of the software. The application now allows to perform reconstructions with images from multiple directories with a project file which keeps track of all the user parameters.
MIK added a 3D viewer to visualize the final reconstruction in Alembic directly inside the reconstruction software.
MIK also fixed issues for the compatibility of the generated Alembic files with Maya and Nuke.



Main page of the 3D Reconstruction Software



New 3D viewer

High-end real-time rendering engine
MIK is evaluating 3dCast, a high-end real-time 3D rendering engine developed by Technicolor. This alternative solution for real-time rendering is directly integrated into the Maya viewport. This solution can strengthen one of the key objective of POPART: provide some high degree of integration with post-production. 3dCast is a versatile tool created to enhance the productivity of animation and lighting teams. 3dCast provides an option to simulate the look & feel of some high-end rendering engines like MentalRay and Arnold, which could be useful for the POPART objectives.

Video Transfer over Sockets
As part of the visualization tools work, LAB has developed a software library that supports transferring video streams from the acquisition system over sockets. This has been based on delivery D3.2, and is expanded to allow other parts of the POPART project to use the library. The library supports using TCP and UDP as transport protocols, and is optimized for low latency delivery over variable networks.

Initial Compositing
Using the aforementioned software library, LAB has developed tools for initial compositing of the main camera picture with 3d model data using tracking data from the intersense tracking system, which will soon be extended to support the POPART tracking system.

Testing
LAB has done extensive testing of the visualization tools, to prepare for use in medium-high budget productions where the tolerance for errors are non-existent. Some of this testing is described in the WP4 section of the report, however, the nature of testing in a film production has been beneficial both for WP3 and WP4 development.

WP4 - Integration and demonstration

Acquisition of multiple streams

The new trigger system, as described in QMR2, is now fully functional and in use. It is an improvement over the originally planned system requiring an external genlock generator. Although the original trigger system worked well, we discovered that camera operators are concerned about introducing genlock of film sets, as this has caused problems for productions in the past. In genlock mode, the film camera will fire the shutter only when commanded by the genlock signal. If the settings on the genlock generator is wrong or the cable is faulty, the live footage can be ruined. As such, we designed a trigger system that does not require camera operators to connect to an external genlock source.

The new trigger system has an internal clock and independent clock. It synchronizes with the main camera's shutter phase and passes the trigger signal on to the witness cameras. This allows the witness cameras to function even without a genlock source. The disadvantage of this approach is that receiving the shutter phase from the main camera is vendor specific. Currently, we have implemented support for RED cameras, since LAB has such equipment readily available. Support for other vendors will be added when needed. We expect to add support for ARRI and Sony cameras in the future.

Integration testing and on-set usage

In addition to synthetic tests done in LAB's office in Oslo and Grimstad we have engaged with film productions to verify the acquisition system in real-life. In particular, LAB provided technology developed in this project to the short film *MUK*, filmed September 11-18. Additionally, LAB expects to provide POPART technology to a national broadcasting house in November, gaining valuable field experience and even a potential customer if the tests are successful. Both test scenarios have required considerable effort from LAB during Q3 to prepare, integrate and verify the components.

MUK testing

The short film production "MUK" was the first real testing scenario where we could evaluate the HAL acquisition technology developed in this project to a full scale production. The goal was to evaluate performance and find issues with the system. MUK is a short film produced by Jon Michael Puntervold, dissemination manager in this project.

We had HAL and LABOgamp alpha at the MUK production in Grimstad, to expose them to rough low-budget undermanned around-the clock abuse and learned a lot from that week of use.

On HAL, we had had two tasks to verify: Record the SDI and transmit wireless video.
LABOgamp, an application developed outside of POPART, was also tested. LABOgamp had the task of copying and structuring the project, sync metadata from Movieslate on-set with dailies and audio and create structured audio-sync'ed logging projects with all metadata from set intact and hooked to the files for downstream treatment and automatisation of post workflows. There is an overlap between LABOgamp and the on-set post production workflow proposed in POPART, and we expect these two technologies to interact or even merge later in this project.

A quick demo of the setup of the logging projects can be seen here. A more detailed video will follow: https://vimeo.com/138924659

Both HAL and LABOgamp performed mostly as expected. As we hoped, we also uncovered some issues that we couldn't provoke under synthetic environments, that really popped up in the stress of long set days with long takes and quick changes from large-zoom setups to minimal handheld rigs.

Testing POPART technology on the MUK film set.

The Third Life Project

POPART partner SRL has been asked to contribute to an artistic performance to be held on October 8.-10. 2015 in Vienna. The organizers of this performance are artists Milan Loviska and Krause, and scientific partners are the University of Duisburg-Essen, Germany, and University of Stellenbosch, South Africa.

During the performance at WUK in Vienna, Austria, and artist wearing an Oculus Rift walks moves on a stage. His movements are tracked, and translated into scaled position updates that are transmitted to a Minecraft client that represents the artist as an avatar in a virtual world, which includes a model of the performance hall itself. The artist perceives his own movement as movement in the Minecraft virtual world.

This performance provides an early opportunity to integrate HAL live recording with the CCTag-based tracking. Only a limited part of the CCTag-based tracking is required for the performance, specifically the delivery of the absolute position in near-real-time, whereas the direction of the artist is in this scenario derived from the Oculus Rift's inertia sensors. It is furthermore not required that positions are reported at the full frame-rate of the camera.

Nevertheless, this opportunity has already allowed us to perform an early integration phase between the recording and streaming modules, tag-based reconstruction and CCtag-based tracking. Partners SRL, LAB and IRT were directly involved in conducting the software integration of the required POPART components during the Third Life's preliminary integration test in August 2015 at SRL's premises. It was determined that the integration of HAL controller and tracker software was best handled using named-shared-memory communication, which avoids the overhead of socket-, pipe- or file-based communication, yet also avoids the integration cost of linking the separately developed components into a multithreaded application.

Third life workshop at Simula, using the CCtag tracking system.

WP5 - Productisation and Distribution

LAB is starting the work to productize the camera acquisition system and camera rig as per T5.1. We have designed initial CAD sketches for the camera rig, and are in discussion with manufacturing companies. For a product version of the acquisition system, we need to sort out some issues before designing the final product. This includes the witness cameras that according to specification should support shutter speed of 1/10 000, does only in fact support 1 / 1000 when externally triggered. This was unknown for the OEM Point Grey, and LAB is in negotiations with them to rectify this. For the product version of HAL, we are looking into using the Com Express platform with LAB designing a custom extension board. We expect to take lessons learned from the November tests into account for this design, and a product design should be finished in early 2016.

WP6 - Management

The project webpage has been significantly improved, and is as before available on http://popartproject.eu.

Workshops

In this quarter, there have been several workshops where the partners meet to progress on the development:

- Simone was @MIK from 27/08 to 07/08
- Fabien was @IRT from 28/08 to 03/09
- LAB, SRL, IRT and MIK met in Paris on September 14-17 to work on integration, global workflow of the system and plan the demo for December. MIK received a HAL prototype from LAB and we worked on interfaces to integrate the components.

Plenary meetings

Project meetings are held at least every 6 months. The next plenary meeting is planned for Paris in December 2015.

Plenary audio / video conferences

In addition to meetings, the POPART consortium holds a video conference every 4th week to keep track of the project progress:

- September 29th 2015 - video conference
- August 21st 2015 - video conference

More spontaneous teleconferences are also held supporting the different work packages and the partners involved.

Status of deliverables

Deadlines for internal review and final submission of each of the deliverables for this period have all been met.

- D1.2 - Final DB Management library. Deliverable submitted on August 1st 2015.
- D2.2 - Software for basic frame query. Deliverable submitted on August 1st 2015.

Adjustments to the DoW

Advisory board

Next session will be an on-set workshop, mentioned in the previous QMR, exploring working with live comp and mixed real and virtual sets. Set constructions for the workshop are mostly done, and the advisory board are awaiting the next step in technology integration, before performing testing for further feedback.

Industry connection with the POPART-project is persistent on all levels, technology-development at Toulouse are done in tight cooperation with the post-production teams at MIK, LABO are working with a major broadcaster and smaller projects such as the MUK short film, and BandPro are surveying upcoming products and are keeping an eye on the mixed-reality scene.

Conclusion

The project is progressing more or less as planned with no major roadblocks encountered. The next big milestone in the project will be planned demo in December, demonstrating all components developed this far integrated to a functional system. At the same time, we are pursuing real life testing on film sets to maximise impact while ensuring that in the end we develop something that is usable in productions.

Project acronym: POPART Project number: 644874 Work package: Management Deliverable number and name: D6.3: Quarterly Management Report 4	Title: Quarterly Management Report 4 Work Package: WP6 Version: 1 Date: January 12, 2016 Author: Håvard Espeland
Type: [X] Report [ ] Demonstrator, pilot, prototype [ ] Website, patent filings, videos, etc. [ ] Other	Co-Author(s): All partners To: Albert Gauthier, Project Officer
Status: [ ] Draft [ ] To be reviewed [ ] Proposal [X] Final / Released to EC	Confidentiality: [X] PU – Public [ ] CO – Confidential [ ] CL - Classified
Revision:	
Contents: Deliverable 6.3: Quarterly Management Report 4	

Human Resources

SRL

The project is coordinated by Håvard Espeland. Carsten Griwodz and Lilian Calvet contribute to the innovation and scientific work. Elisabeth Andersen contributes to the administrative work.

LAB

The project at LAB is led by Magne Eimot. Jon M. Puntervold, Gunleik Groven and Håvard Espeland Kristian Skarseth also contribute to the innovation work.

MIK

The project at MIK is led by Benoit Maujean, R&D Manager at Mikros Image. Fabien Castan is the technical manager of Popart project; he is responsible for the 3D reconstruction library (OpenMVG) and photomodeling plugin for Autodesk Maya (MayaMVG) at Mikros Image. Elisa Prana, Nicolas Rondaud and Cyril Pichard, as software engineers contribute to the development of Popart tools, including the expected adaptations of OpenMVG and MayaMVG. Furthermore, some Matte Painters of Mikros Image studio have contributed to the generation of the virtual dataset.

IRT

The project at IRT is led by Vincent Charvillat. Simone Gasparini is responsible for the Work package 1 and he’s coordinating the tasks assigned to IRT, in particular on camera localization and the contribution to the OpenMVG libraries. Sylvie Chambon is also a staff member of the project and she will participate in the task involving the image matching for improving the performances of OpenMVG. Clement Aymard is a developer engineer who is developing the different contributions to the project. Since October 15th, Jean Melou has been hired as engineer to work on OpenMVG pipeline.

BAN

The project at BAN is led by Christopher Hantel, CEO at Band Pro Munich. He is joined by Randy Wedick, contributing to the commercialization and productization of the POPART commercial product.

Person Months Contributed to the Project

Participant	Names of staff	WP1		WP2		WP3		WP4		WP5		WP6		Total	
		Planned	Actual	Planned	Actual	Planned	Actual	Planned	Actual	Planned	Actual	Planned	Actual	Planned	Actual
SRL	Håvard Espeland Lilian Calvet Carsten Griwodz Elisabeth Andersen	6	2,73	19	14,57	0	0	0	0	0	0	6	5,91	31	23,21
LAB	Magne Eimot Jon M. Puntervold Håvard Espeland Gunleik Groven Kristian Skarseth	0	0	0	0	8,80	12,66	14,8	13,37	3,25	0,32	0	0	26,85	26,35
MIK	Benoit Maujean Fabien Castan Elisa Prana Nicolas Rondaud Cyril Pichard	4,5	1,24	2	0	15,6	12,86	8,80	13,5	0	0	0	0	30,9	27,62
IRT	Simone Gasparini Vincent Charvillat Clement Aymard Sylvie Chambon Nicolas Bertrand Jean Melou	18,5	17,29	6	6	0	0	2,4	0	0	0	0	0	26,9	23,29
BAN	Randy Wedick Christopher Hantel	0	0	0	0	0	0	0	0	1,5	1	0	0	1,5	1
Total		**29**	**21,26**	**27**	**20,57**	**24,4**	**25,52**	**26**	**26,87**	**4,75**	**1,32**	**4,5**	**5,91**	**117,15**	**101,46**

Figure 1: The cumulative budgeted and contributed person-months since project start for each partner sorted by Work Package.

The planned costs represents a linear distribution of PMs over the length of each work package. As can be seen from Figure 1. Personnel effort and travel costs are mostly in accordance with the budget for LAB, MIK, IRT and BAN. For SRL, the planned efforts of QMR4 represents 91% of all PM's allocated to SRL during this project (31 out of 34 PMs). This is because much work in WP1 and WP2 were linearly allocated in 2015, but due to the nature of the collaborative work in this project, some of this effort must happen in 2016 and work is not really distributed linearly. We are however on track to achieve the objectives as was demonstrated on the review in December and expect this to happen within the original budget and timeline.

Other direct costs included, during the fourth quarter, acquisition of equipment as anticipated in the budget and GPUs for CCtag development at SRL.

Project Progress

The project is running as planned and we are working on all tasks expected to implement POPART and its objectives.
This quarter was an important integration step in the project which was an opportunity to strengthen the collaboration between all partners. We can see a noticeable increase in the number of travels: MIK to LAB, SRL to MIK, SRL to IRT, MIK to IRT. Those close interactions have enabled a significant boost in the project achievements.
The figure 1 recall the global workflow we will see in detail in all work packages:

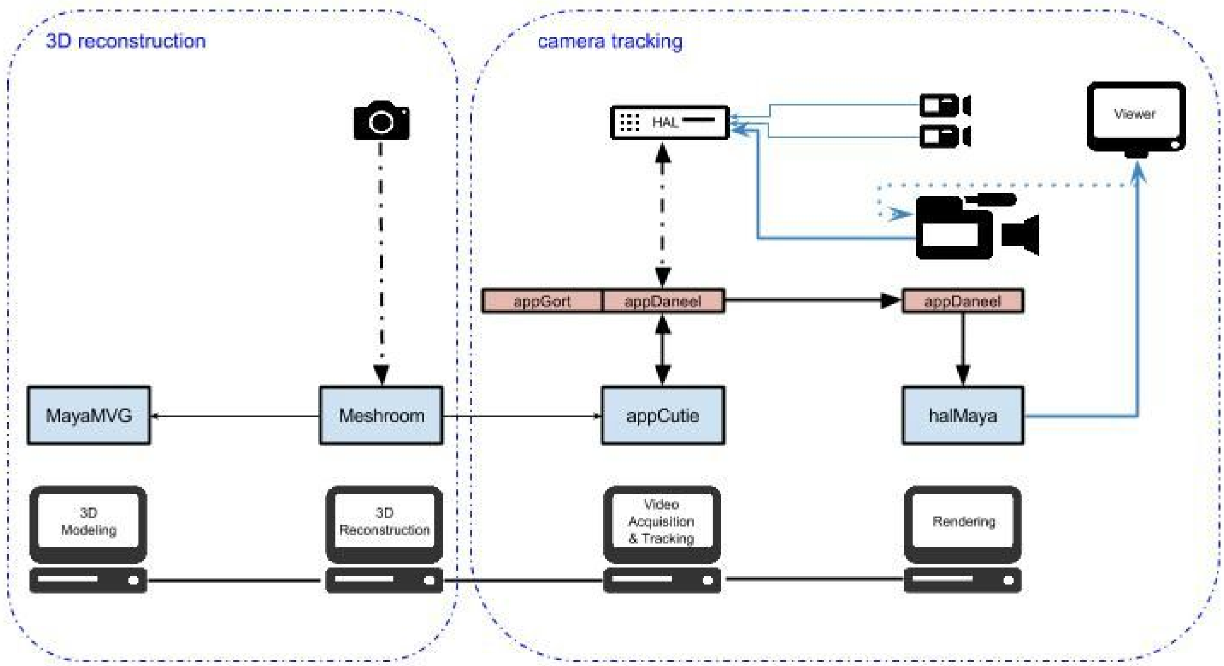


Figure 1. Global workflow with the latest updates
blue: user application, red: low level processes

WP1 - 3D Visual Database

Voctree analysis

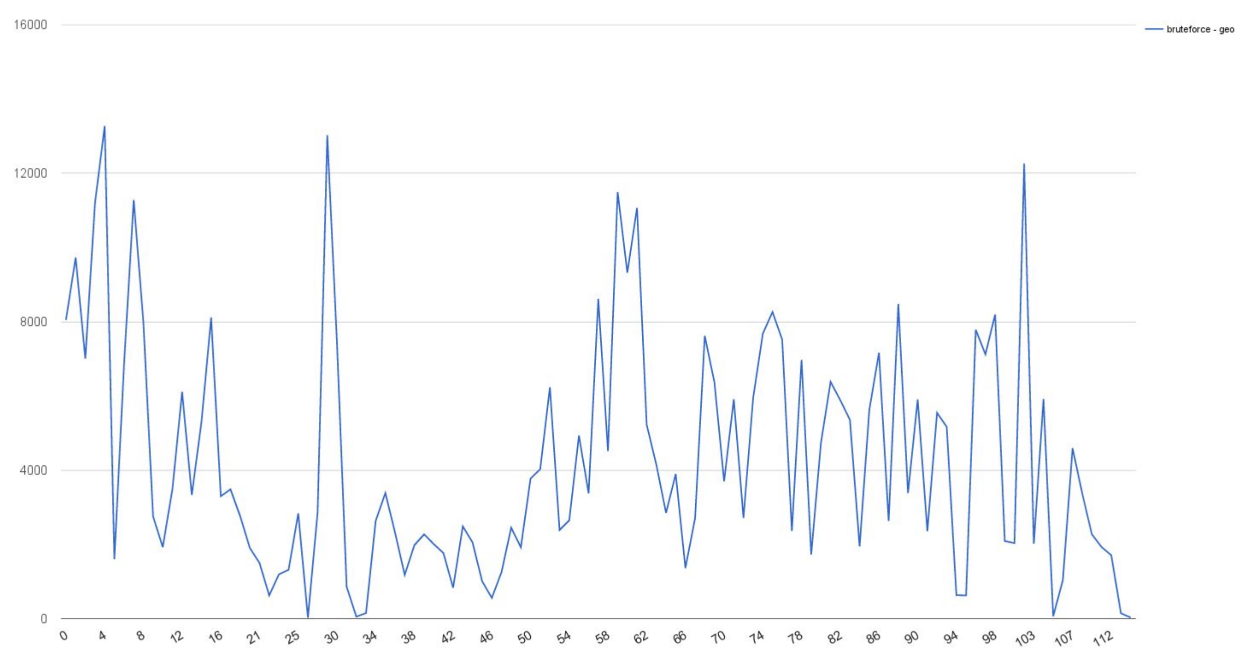
As explain in QMR3, we use the vocabulary tree approach, introduced by [Havlena2014], to compute similarities between images which allows to choose a sub-set of image pairs to compute the features matching. This transforms the quadratic complexity into a linear problem.
We are working on 3 objectives for the vocabulary tree:

1. Improve the image similarity results
2. Evaluate usability for feature matching
3. Improve performances

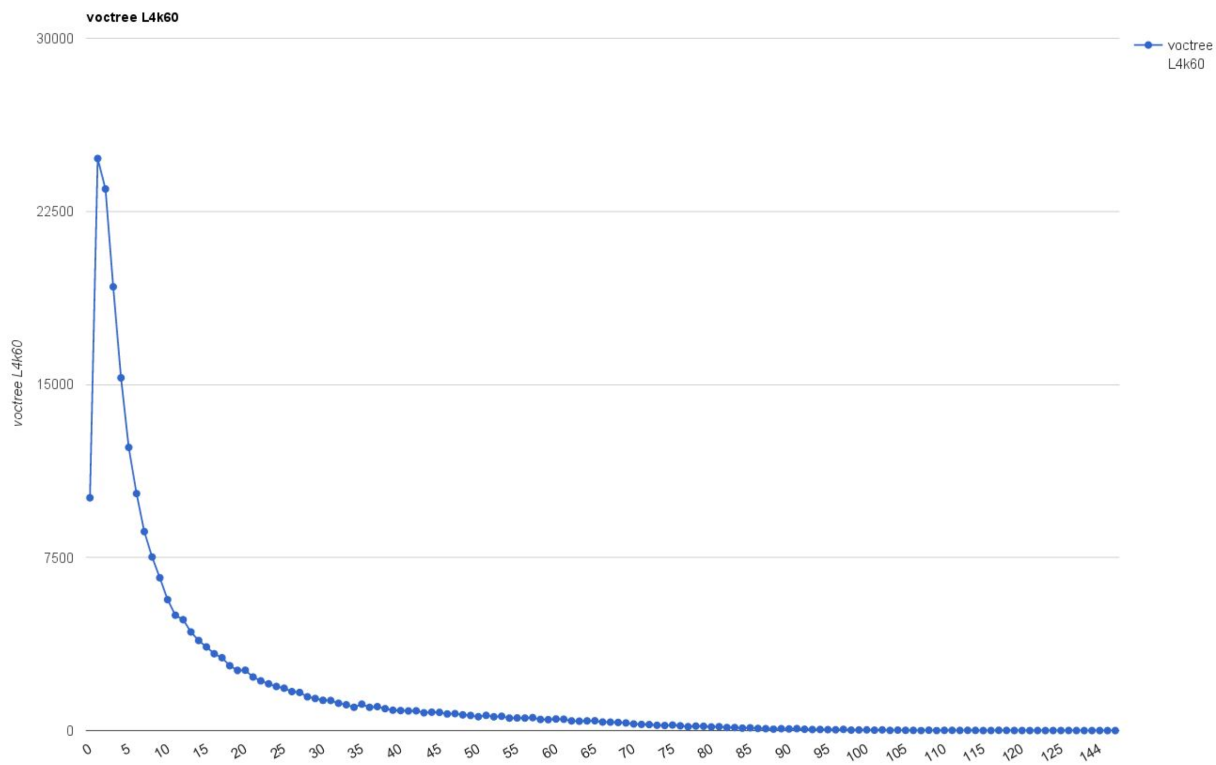
We put this solution in production at MIK. As expected, it drastically reduces computation times on large datasets, but we notice some degradation of the reconstruction quality on some datasets. It appears that the images similarities ordering of those datasets using the vocabulary tree was suboptimal. So a deeper analysis was required to improve the results.

Some experiments were also required to prepare the next step for the vocabulary tree to use it to retrieve feature matches. This matching will probably not achieve the same quality as the method currently used, but it should be usable to compute the fundamental matrix required to perform the guided matching. When we compare 2 vocabulary tree histograms, we notice that more than 9/10 of those matches correspond to one-to-one feature matches. So it seems reasonable to use it for feature matching.

We analyzed the descriptor distances using standard features matching methods and with the feature matching provided by the vocabulary tree. For the tested dataset, the range of distances between matched descriptors extends to 112.000, with a mean around 55.000 . Moreover, close matches are not in greater numbers than distant matches.
On the other hand, when we insert our Voctree in our pipeline, most of the matches have been dismissed, only the closest ones have been kept. We assumed that the deterioration in results is due to this loss.
Then, it seems that our Voctree has too fine leaves: the sorting is too strict compared with our distances between matched descriptors. We notice that if we reduce the level of the Voctree (and then, the number of leaves) which implies an increase of the leaves size, the kept matches seem to be more compatible with our previous results.



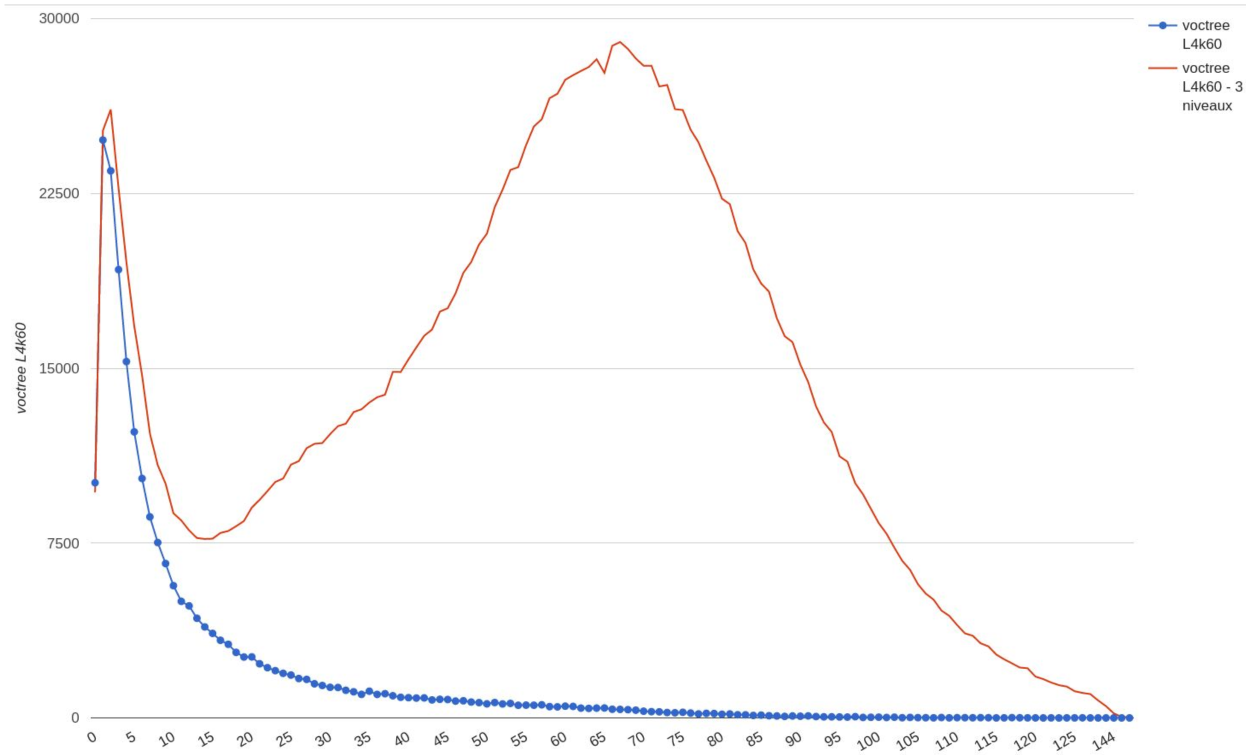
Histogram of the Distances between matched Descriptors (bruteforce method)



Histogram of the Distances between matched Descriptors using Voctree (L4k60)

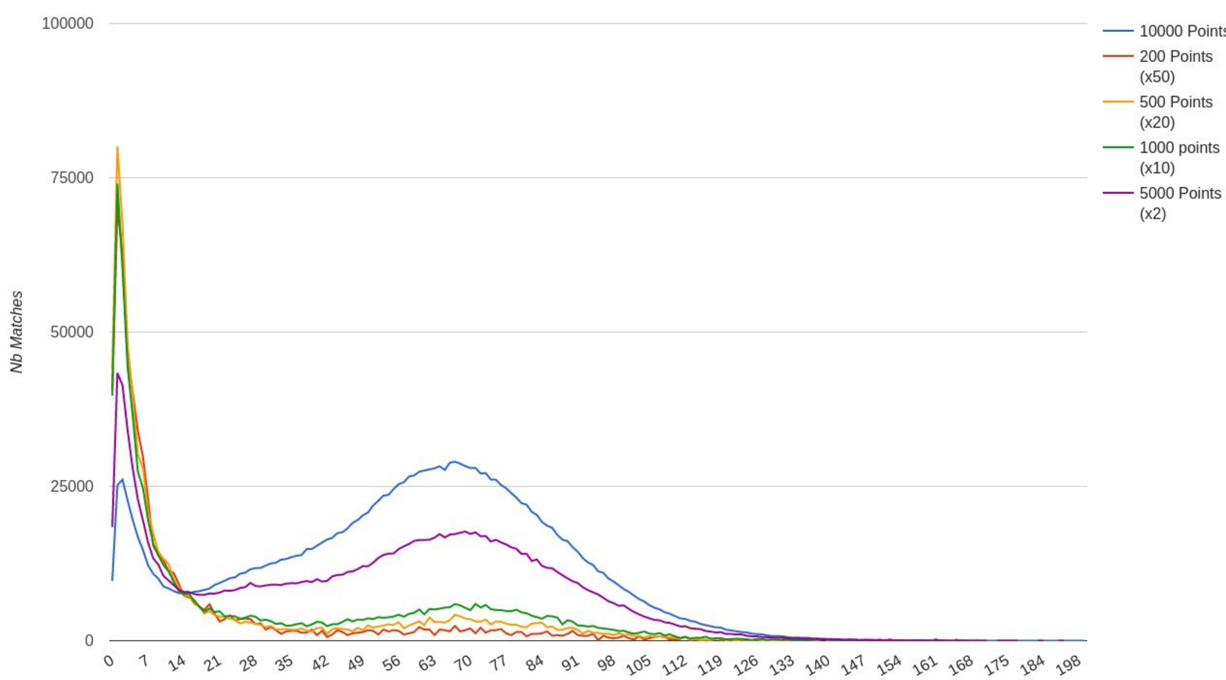
With a 12 millions leaves vocabulary tree, the matches are less than 20000 which is really too strict in comparison to other methods.

As explained in “Anatomy of the SIFT Method” from Ives Rey Otero and Mauricio Delbracio, an absolute max threshold can be used. They propose a threshold between 62000 and 90000 which seems reasonable regarding our results.



Comparison of Voctree L4k60 (12 millions leaves) and L3k60 (216000 leaves)

We also perform tests on the number of matches when we reduce the number of features during the detection step:



Compare the number of valid matches

with different threshold on the number of detected features

There are some remaining tests we have to perform. First, it could be great to get statistics only on the descriptors kept until the very end of the process, those which have been used in the reconstruction. That will dismiss some unreliable matches. Moreover, we have to care about the leaves sizes, maybe find a way to control that. At last, according to our previous results, we have to generate new smaller Voctrees in order to compare their performances.

WP2 - Real-time camera tracking

Almost all the work and progress achieved on this work package since the last QMR report (QMR 3) has been described as additional reports to the deliverable D2.3 Camera tracking library (due by December 1st). These reports are mainly part of the paper lately submitted to CVPR 2016 conference (submitted November 6th).

From December 1st, all the efforts have been made on code debugging during the integration of the camera tracking libraries (i.e. CCTag and new features in the OpenMVG library) into the first version of the full system. The system has been presented to the European commission on December 17th including a live demo of the CCTag-based real-time camera tracking.

These tests/demo have been conducted in the [Cap Digital] office in Paris with a pretty challenging environment including a lot of reflections, stripes, and CCTag occlusions (see Figure 1).

Figure 1. Cap Digital demo preparations

The quality of the camera tracking has been assessed via the inclusion of very basic CG elements into the scene, the final composition having been viewable through the Maya software (halMaya plugin). The cameras RIG calibration (D2.3) has been used successfully and a small tool based on openCV has been added for lens calibration. As you can see in Figure 4, the CCTag are represented in 3D by red spheres. The system has been tested with a wide range of motion including handheld (see Figure 2) and the system fixed on a tripod. The camera tracking performance was around 7 fps and the latency of the whole system was around 1 second. The artificial feature extraction is the bottleneck to achieve 25 fps. At this stage, only one witness camera has been used to compute the pose of the camera RIG. When this solution will achieve 25 fps, we will work to add another graphic card for the second video stream.
However, the system has presented pretty stable performance with little jittering while tolerating quite fast camera motion even with only one witness camera. So it looks really promising.

Figure 2. handheld camera rig

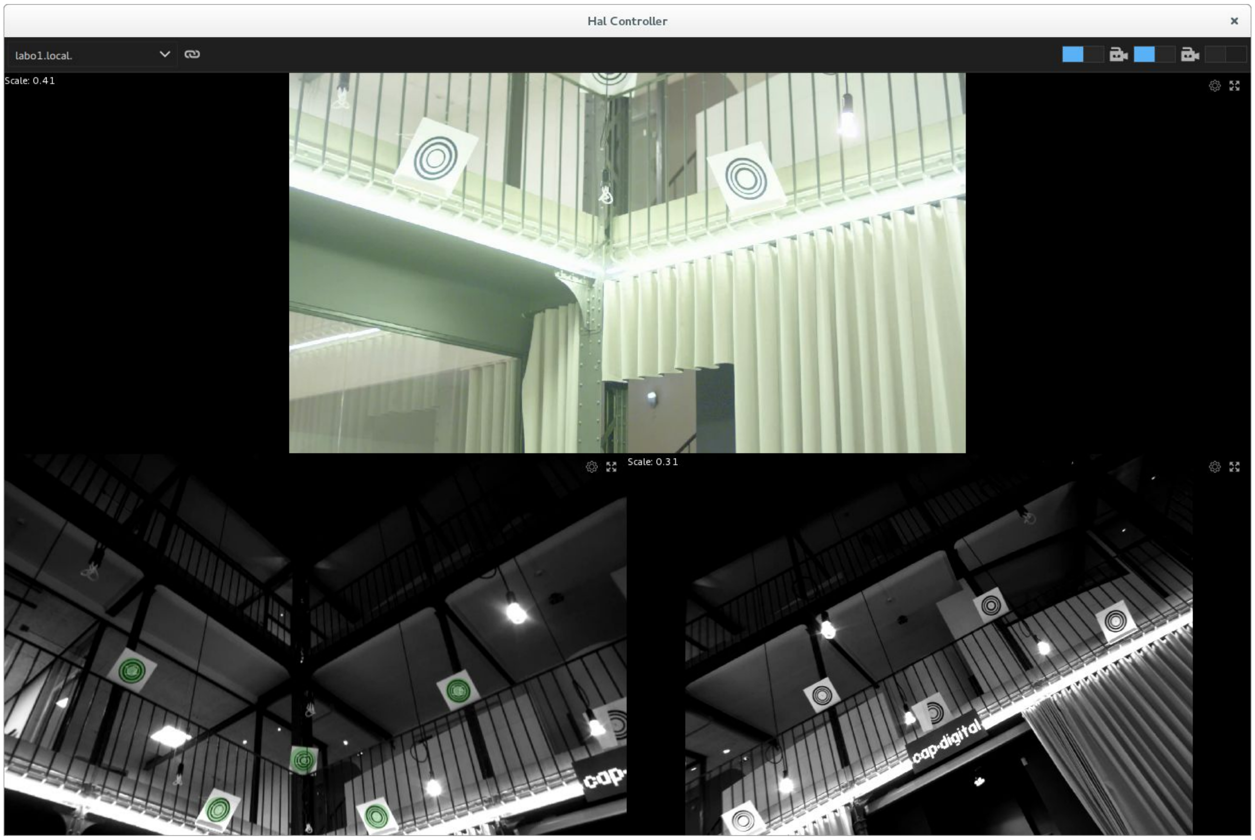


Figure 3. appCutie shows the 3 HD video streams

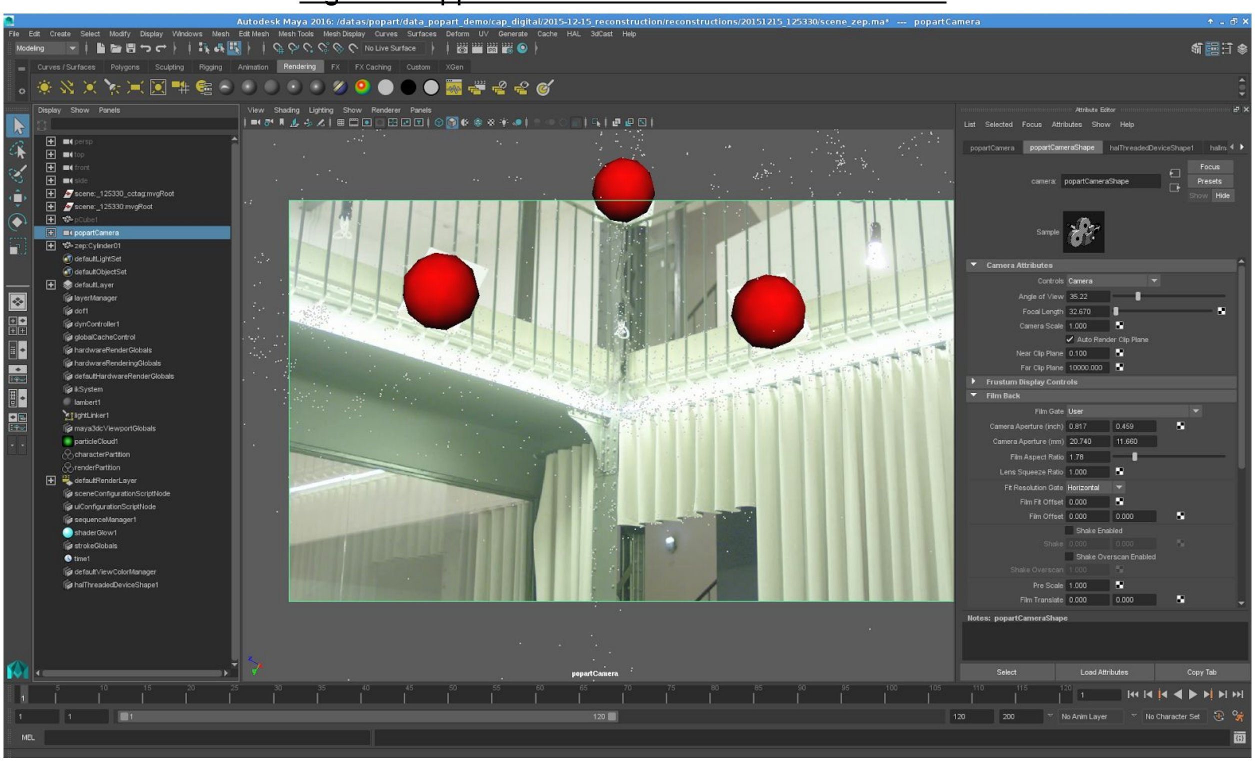


Figure 4. halMaya: receive the camera pose and the live action video stream

Natural feature extraction. We started the validation of the GPU implementation of the SIFT algorithm (popSIFT). The feature repeatability evaluation is conducted via the following dataset: openMVG/Features_Repeatability.

Offline camera tracking

In order to asses the quality and the performance of the developed tracking algorithms we benchmarked them against a synthetic dataset of camera motions. The interest of using a synthetic dataset is that a precise ground truth is available and many source of noise are neglected, thus allowing to assess a lower bound accuracy for the algorithms. In our case we used the dataset that has been generated and published by MIK, the “Levallois City Hall Virtual dataset” (http://dx.doi.org/10.5281/zenodo.19164): the dataset contains a CG reconstruction of the City Hall building and several camera motions rendered by Maya. In particular there are 3 cameras motions, **cam0**, **cam1** and **cam2**, that are the camera motions of a rig of 3 cameras in mutually fixed positions, a camera motion that takes into account the depth of field of the camera, **dof**, and a camera motion simulating the motion blur, **motion blur**, as the camera moves in the scene. For each motion we run our offline camera tracking algorithm based on SIFT features and the vocabulary tree. For each image we compute the deviation from the ground truth in terms of the baseline error, the difference between the estimated camera position and the ground truth, and the angular error, computed as the angle between the two rotations representing the camera orientation. These two figures are one of the standard ways to compare the estimated poses of two camera motion, which is often call Absolute Trajectory Error (ATE) in the robotics community.
In terms of code we used two scripts that we have developed inside our fork of openMVG, in the *popart_demo* branch (https://github.com/poparteu/openMVG/tree/popart_demo/):

- `./src/software/Localization/main_voctreeLocalizer.cpp` can take a sequence of images as input and estimated their pose wrt an input reconstruction;
- `./src/software/SfM/tools_precisionEvaluationToGt.hpp` can take the results of the previous script and compare the position with the ground truth pose of each image and generate a report containing the statistics about the ATE errors.

Finally, since the CG reconstruction of the building is up to a scale factor and it is not in metric coordinates, we compute the conversion of the CG units into more meaningful meters coordinates by taking the actual dimension of the building (eg. from Google Maps) and computing the transformation factor to bring our results in metric units. As a reference, the building measures 51 m x 42 m and each camera motion is composed of 380 frames covering a trajectory of 220 m around the building.

Table 1 summarizes the results that we got for the different camera motions. As one could see the baseline errors for the camera motions without any noise are in the order of 1 cm, while the angular errors are almost neglectable being of the order of the hundreth of degree. It must be noted that for cam2 we got a very short sequence of frames in which the localization error are diverging: this was actually due to a bad design of the camera motion as the camera crosses at

some point the boundary of the building, thus making the localization to fail. We will fix this in a new release of the dataset. We report as example the distribution of the errors and the errors for each frame for both the baseline and the angles for the camera movement **cam0**.
As soon as we introduce some real case effects such as the depth of field and the the motion blur we can see from the Table1 that the errors increases, yet not dramatically. For these 2 sequences the errors are of the order of 2 cm with a maximum error of 10 cm. The same holds for the angular errors, in which the errors increase to the order of the tenth of degree. From the Figures we can see that the errors increases when the camera moves really close to the building thus exacerbating the effect of the blur or of the depth of field

For the future, we are planning to perform the same type of testing with the tracker based on CCTags in order to assess and validate its localization algorithm.

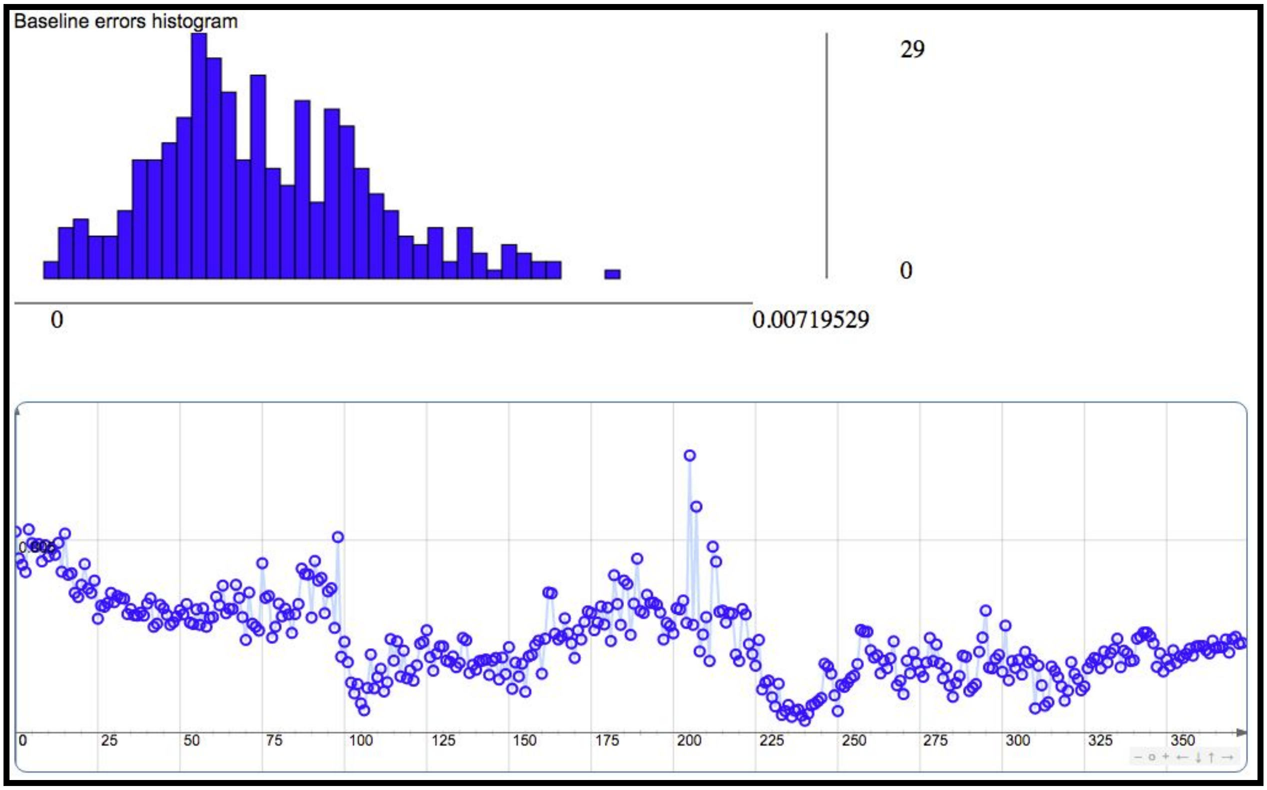
As for testing with real dataset, we recall that we have run some tests on the “Toulouse Capitole” dataset created by IRT whose results have been described in Deliverable D2.3 due in December. The City Hall (Capitole) has been reconstructed in 3D from a set of images. Then some videos have been taken in the main square and processed off-line with our algorithm. The videos are HD 720, and the camera has been calibrated off-line with an OpenCV-based application. There are around 2000 frames for each video and all of them have been correctly localized with an RMSE error for both sequences around 1.5 pixel. If a further global bundle adjustment is run on the entire sequence refining the camera intrinsics as well, the final result can be further improved and RMSE of 0.5 pixel is obtained.
Since a ground truth was not available, only a visual assessment of the result is possible and that can be seen in the Maya screencasts available here:

- https://youtu.be/X7zKK-W-qRQ
- https://youtu.be/I6oltk9iVUU

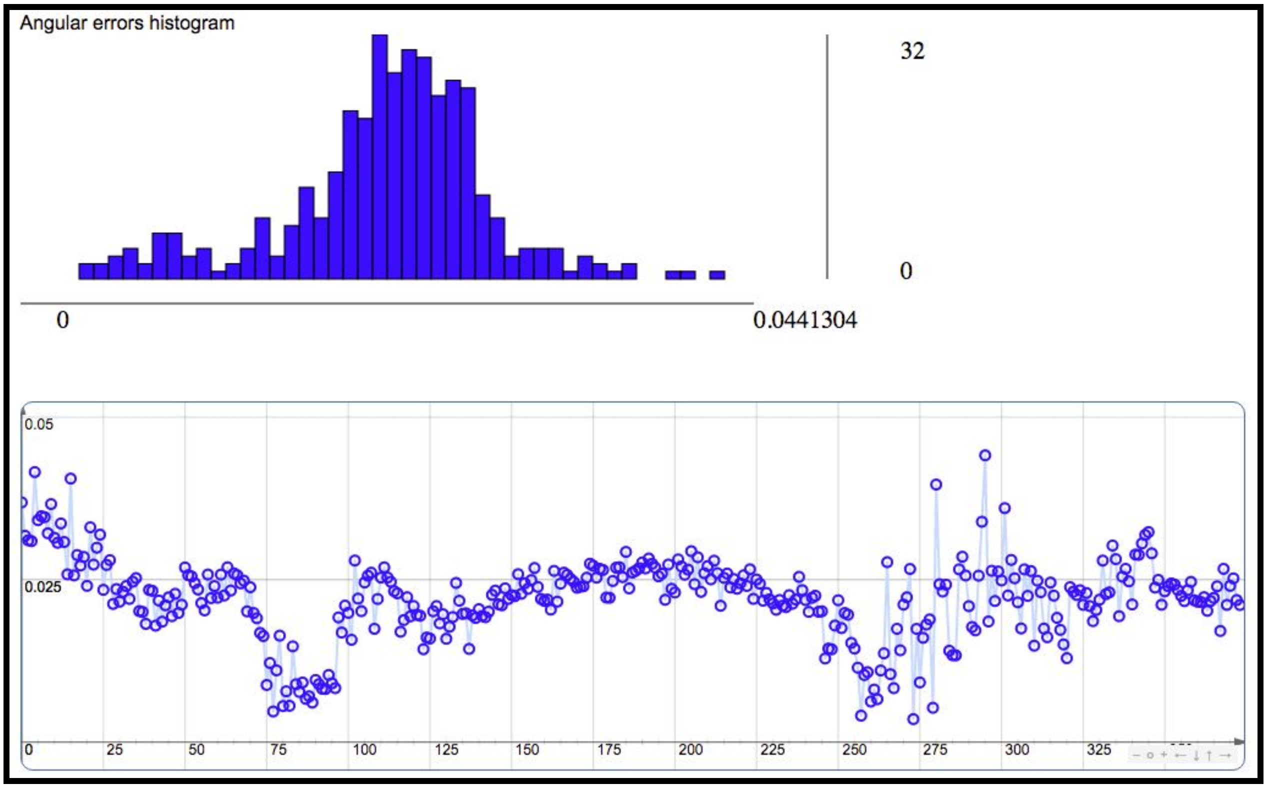
	Baseline error (cm)				Angular error (degrees)			
	min	**mean**	**median**	**max**	**min**	**mean**	**median**	**max**
cam 0	0.14	1.12	1.06	3.30	3.54E-03	2.21E-02	2.27E-02	4.41E-02
cam 1	0.13	1.25	1.05	5.90	4.98E-03	2.38E-02	2.30E-02	9.09E-02
cam 2	0.05	5342.53	1.13	1,8E+06	3.36E-03	6.70E-01	2.44E-02	9.35E+01
dof	0.32	2.51	2.53	5.03	7.24E-03	3.18E-02	2.53E-02	1.66E-01

motion blur	0.71	2.99	2.72	9.38	5.23E-03	4.03E-02	3.10E-02	3.27E-01

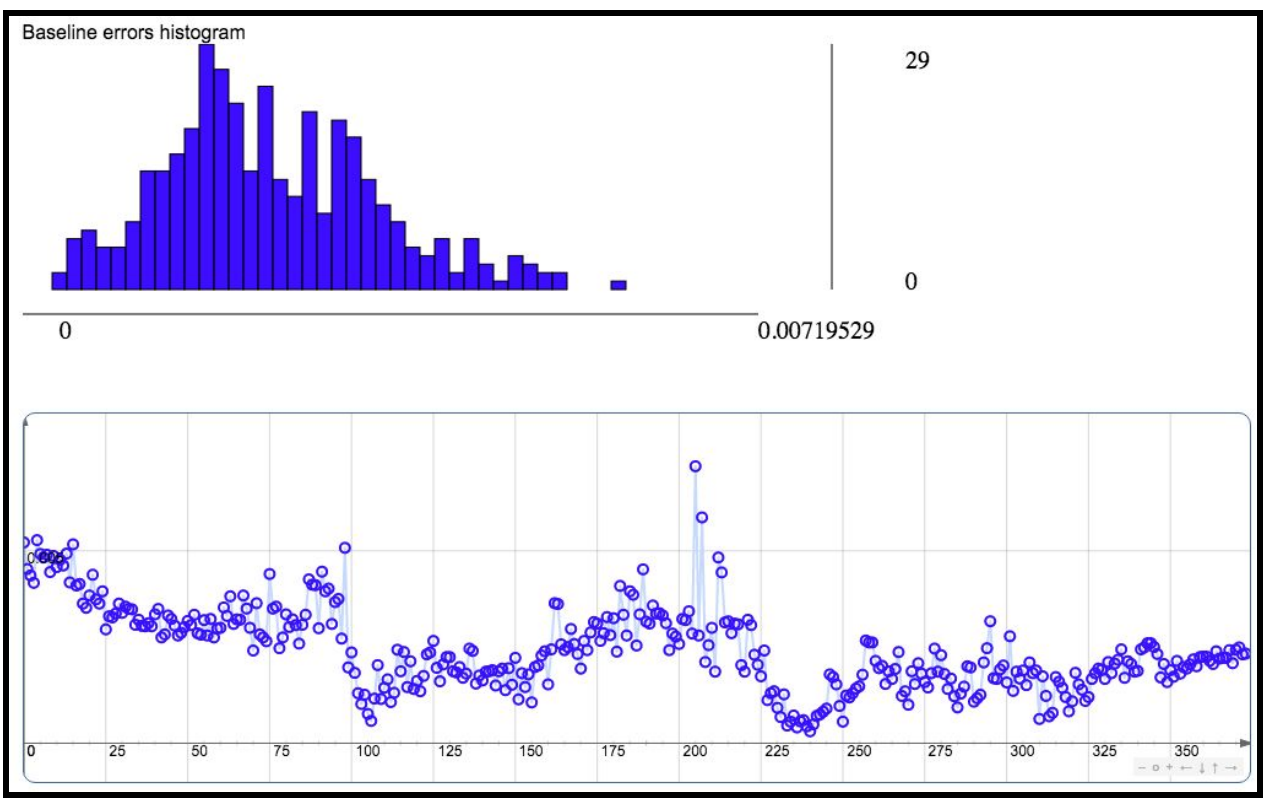
Table 1: Statistics on angular and baseline errors for the considered dataset



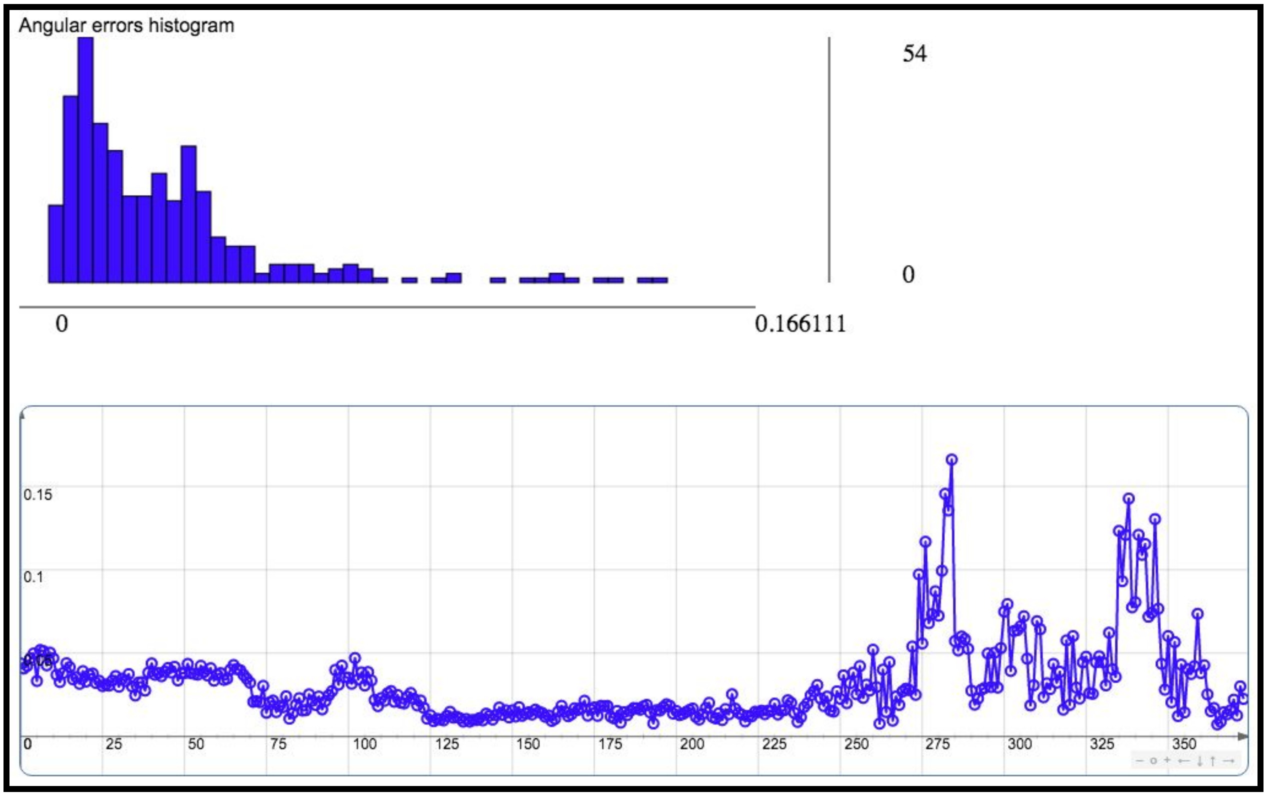
Baseline errors for **cam0** dataset: the error distribution on top and the error for each frame on the bottom half.



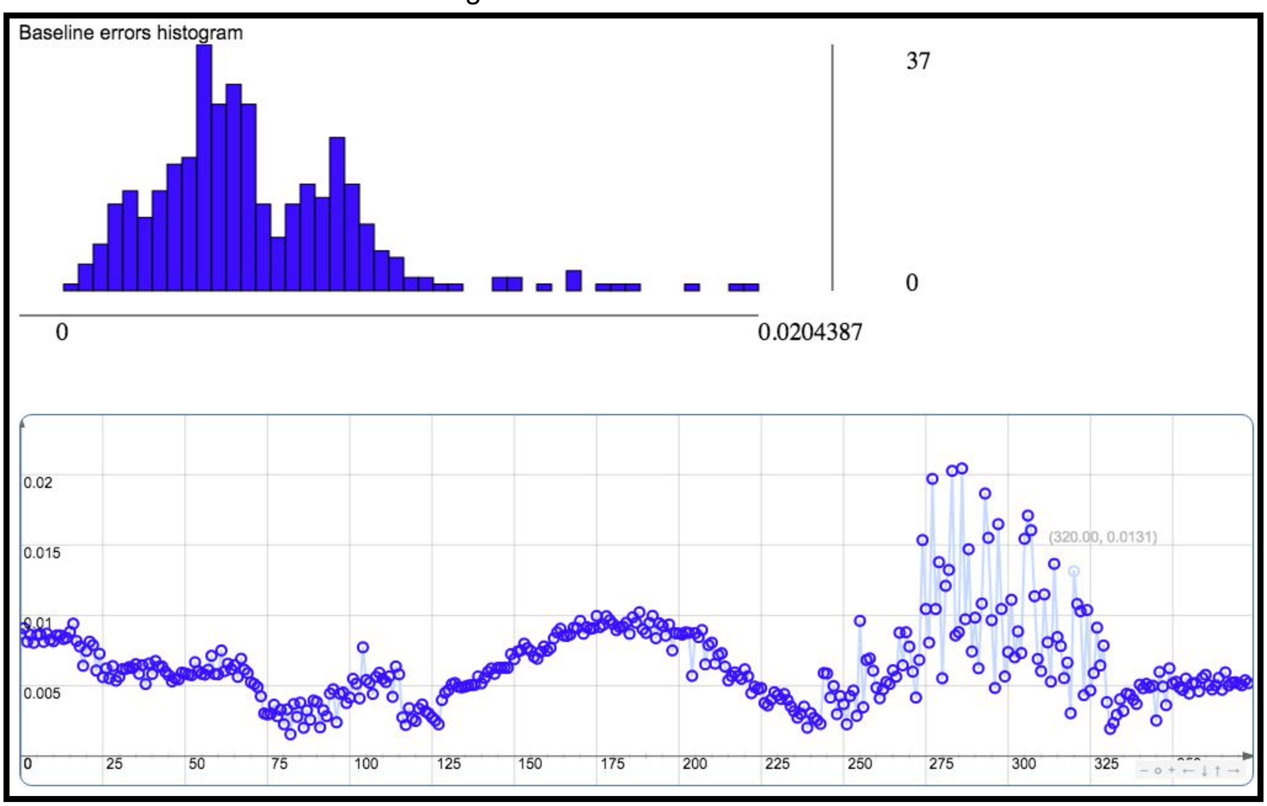
Angular errors for **cam0** dataset



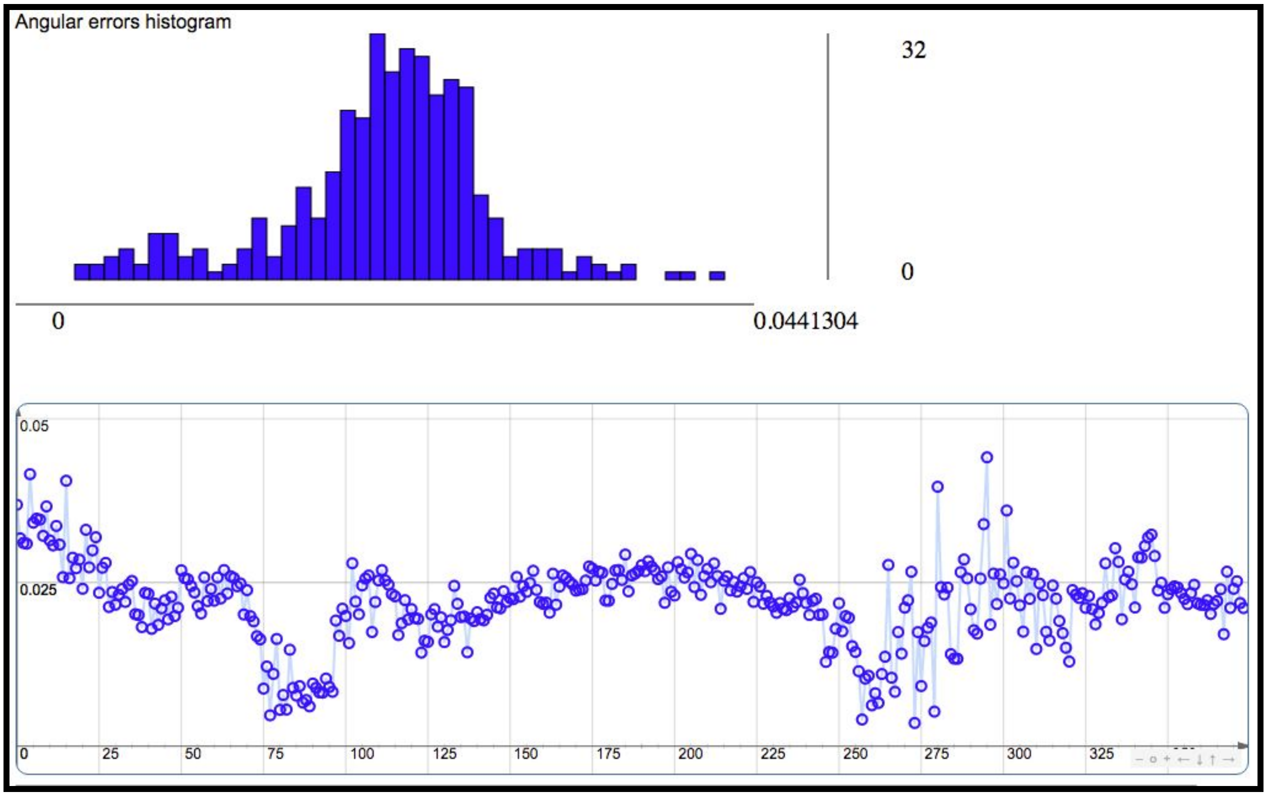
Baseline errors for **dof** dataset



Angular errors for **dof** dataset



Baseline errors for **motion blur** dataset



Angular errors for **motion blur** dataset

Example of image for **motion blur** dataset in which the high level of blur due to a combination of quick movement wrt a close object.

WP3 - Shoot preparation, visualization and post production

During this quarter, we have focused on the connection and integration of all the components to present a first version of the full system connected at the review in december. The Autodesk ® Maya viewer has been implemented and all the steps have been adjusted. We now use a dedicated service to compute the camera pose. We also take the camera RIG into account within the chain and propagate additional information like the list of detected CCTags for display purposes.

Meshroom - Onset reconstruction
We added an option to compute locally (without renderfarm access).

halMaya - Onset visualization (see Figure 4)
A Autodesk® Maya plugin has been created to visualize the 3D environment on top of the live action video stream. A “*Daneel*” service is running on the viewer desktop

appCutie - Onset camera tracking & monitoring
appCutie allows to visualize the 3 video streams in live. By default, this application shows the 3 video streams as soon as they are received. We have added an option to display the detected CCTags with their identification. For that, we have added an option to display the frame when the pose is computed and we now propage the list of detected CCTags as additional information through the whole chain of Popart software (openMVG, Gort, Daneel, appCutie).

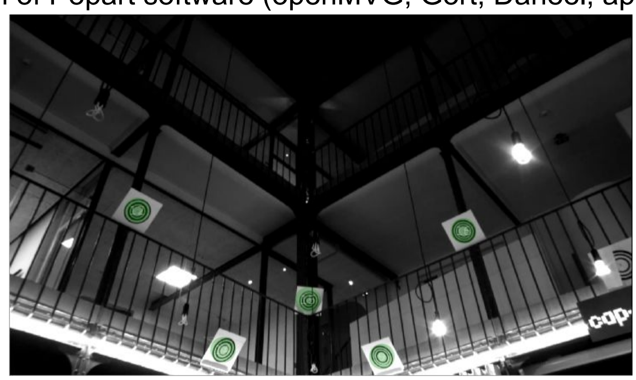


Figure 3. appCutie: detected CCTags in overlay on the video stream

appGort - Camera tracking algorithms
POPART Previz solution has the responsibility to display the Live Action stream with augmented reality. The display of the live action stream in itself is an important responsibility and we have to provide a reliable solution. To get the highest degree of stability, we decided to do the fewest possible things inside this process. So *appDaneel* only manipulates video streams and metadatas. We created a new service called *appGort* to perform the camera localization with WP2 algorithm.

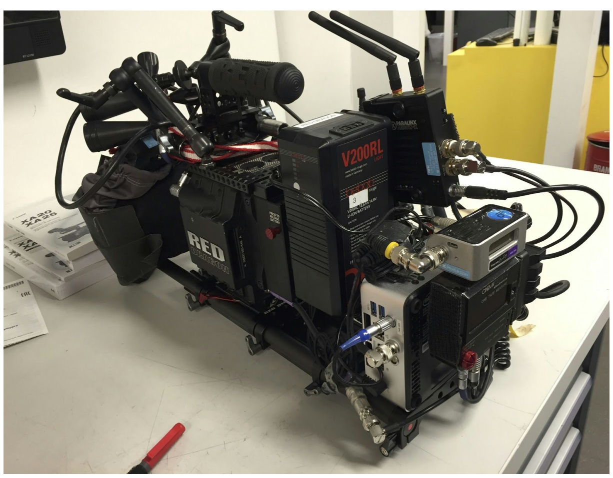
WP4 - Integration and demonstration

Software development

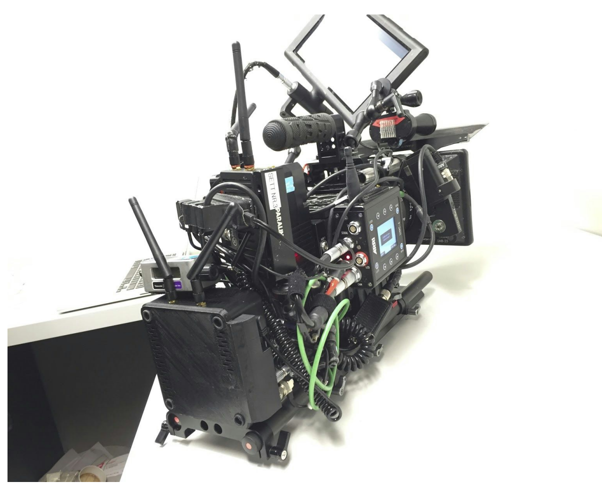
The largest task performed in the quarter was the change from a TCP based video streaming solution to a streaming solution based on UDP. This was performed since we required a semi-stateful connection between streaming server (hal) and client (appDaneel) with only partial reliability. The problem with TCP for streaming was twofold: 1) TCP provides a reliable connection, but in the event of network loss due to wireless being out of range, it will build a backpressure of frames to retransmit. In a live scenario where HAL is supposed to function, the old video frames are no longer pertinent. In the protocol developed by LAB, we provide partial reliability within a GOP to ensure decodability. No backpressure is built up in the event of wireless link loss. 2) TCP was designed to be fair to other connection by employing exponential backoff. In our scenario, we have no competing streams, and the exponential backoff net effect is that after loss of packets, it takes some time (seconds) before we are able to utilize the full bandwidth. The custom protocol developed performs significantly better over wireless connections than TCP, and we are able to regain the live picture almost instantaneously after wireless connection is regained.

Testing and demonstration

To validate the design of HAL we have worked closely with NRK on their production Snøfall. Two HALs has been mounted on two cameras on set in November. This allowed us to verify the reliability, functionality and usability of the HAL system early while providing an early path to the market. Many adaptations were made to accommodate the needs we saw by third-party use, including problematic cabling.



First day with HAL on NRK's camera. Problematic blue cable coming out 90 degrees from HAL.



Cable issue solved by designing and 3d printing a custom lid on the HAL to move the connectors forward. Also, external RF antennas were added to improve range.

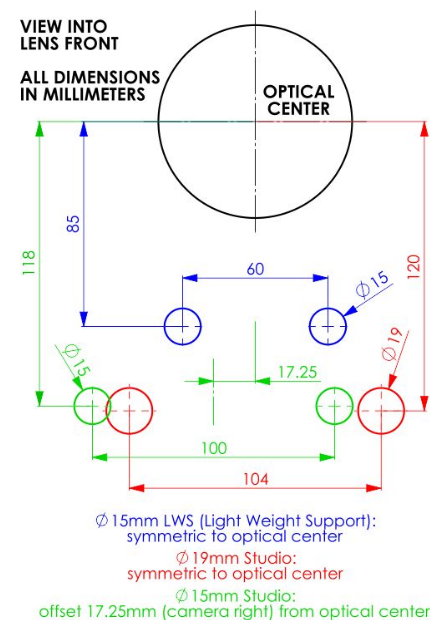
Future work
The next major part in integration will be to add support for FIZ units, notably ARRI and C-Motion units. LAB has already have established a professional relation with both vendors and they will provide documentation for this integration. The FIZ data is required to be able to support zoom/focus in live previewing, but it can also be useful for post production, and is typically not recorded during shooting.

Another important task in Q1 2016 will be to improve the signal flow quality of the stored data for offline use. Currently, only live streaming has been prioritized and we believe the image quality for offline use can be improved significantly, allowing a broader use of the captured data.

We have set an internal goal of finalizing the hardware components for HAL by April 1st, enabling the system to be released to manufacturing in 2016. This will require significant hardware design if we are to use a different platform than Intel NUC, which has been deprecated by the OEM after we started the project.

WP5 - Productization and Distribution

Productising camera rig and hal
The witness camera mounts has been designed according to standards set by current solutions on the market, where witness cameras also can be used for collecting stereo depth information. To accommodate this we utilise the fact that camera accessories are mounted on standardised rods with a fixed distance to optical center of the main camera lens.



For the final design; unique rods-adapters will be made to accommodate each standard, making sure that the mounting height will be the exact same for each rod system. The witness camera mounting arms are designed after this specification, and have been built with a very cost-efficient process; the mounting arms are water cut from aluminium and then bent with sub-millimetre accuracy. Production has been tested and verified.

Final Hal-box design specification is underway, and custom 3d-printed top lids with different cable positions has been tested in production at NRK, receiving valuable feedback. This has led to a concept for final design that will be 3d-printed and tested during the next couple of weeks. The final design will also feature a flexible mounting system, making it possible to mount the HAL-box in at least three different configurations; on top of the camera, on to the rods of the camera (with customisable angel) and on the back of the camera utilising a v-mount system. A custom mount for the back plate of Red Digital Cinema cameras are also being considered.

WP6 - Management

Workshops

In this quarter, there have been several workshops where the partners met to make progress on the development:

- Fabien Castan was at IRT 05/10 to 09/10/2015 to work on the camera localization algorithms with production datasets. We also add a lot of missing information in the Alembic file format.
- Cyril Pichard was at LAB from 16/11 to 20/11/2015 to define and integrate the tracking components (Gort) into the system.
- Fabien Castan was at IRIT from 26/11 to 01/12/2015 to fix the problem of missing CCTags in the 3D reconstructions and makes the real-time CCTag tracking system unusable on many datasets. We also add Alembic export for camera tracking.
- Håvard Espeland (SRL) was at MIK from 07/11 to 11/11/2015 to define the components required for a functional demo of the tracking and visualization system in December.
- Lilian Calvet was at IRT from 17/10 to 30/10 to integrate both the CCTag camera tracking and the calibration of the multi-camera system, initially implemented in the former cameraLocalization repository, to the openMVG library. He also initiated there, in collaboration with IRT, the paper writing about the cctag detection which has been submitted to the CVPR 2016 conference the November 6th.
- Lilian Calvet was at MIK from 30/10 to 02/11 to make some datasets acquisition with CCTags at Cap Digital with MIK.
- Lilian Calvet was at MIK from 06/12 to 20/12 to prepare the demo at Cap Digital.

Plenary meetings

Project meetings are held at least every 6 months. For this period, the plenary was held in Paris at Mikros Image on December 17th, in conjunction with the Project Review meeting on December 16th.

Plenary audio / video conferences

In addition to meetings, the POPART consortium holds a video conference every 4th week to keep track of the project progress:

- December 7th 2015 - video conference
- November 4th 2015 - video conference

More spontaneous teleconferences are also held supporting the different work packages and the partners involved.

Status of deliverables

Due to the review on December 16th, we prioritised getting the system working in real-time, and therefore postponed the work on the research deliverable D1.3, the open source report D1.4 and QMR4 due January 1.
The QMR will be submitted on January 12. The reviewers recommended that we extend WP1 with a few months and release the open source projects when they have been polished and documented, to make a big impact.

Adjustments to the DoW
To be able to demonstrate a live system in December, we needed to speed up the development of the visualization tools, not scheduled before Month 15. This was a successful effort, but came at the cost of D1.3 (3d visual db update tools) and D1.4 (report on open source) which were prioritized down. D1.3 is a research deliverable and D1.4 is an open source release. We believe this was a beneficial change since it allowed us to combine all the components developed separately at the partners into a full working system at an earlier point than anticipated. We will submit D1.3 and D1.4 before the end of the project.

Advisory board
The planned workshop with the advisory board planned for Q4 2015, has been pushed back to Q1 2016 to utilise a more complete POPART-system. Funding for an extended test has been granted by Sørnorsk Filmsenter, a Norwegian regional film center. Updates from the advisory board, and review of their feedback are expected in the next QMR.

Conclusion

The project is progressing more or less as planned with no major roadblocks encountered. The next big milestone in the project will be a demonstration of the system with third-party users and the advisory board. At the same time, we are pursuing real life testing on film sets to maximise impact while ensuring that in the end we develop something that is usable in productions.