State of the Art analysis

Media Orchestration from Sensor to Screen (MOS2S)

Edited by: Tom De Koninck (TNO, Karim Dahdah (VRT), Luk Overmeire (VRT), Omar Niamut (TNO), Özer Aydemir (Gerade Software), Seong Yong Lim (ETRI).

Date: 9-7-2019

This document is made publicly available.

1. Introduction

The project MOS2S is an ITEA3 project that kicked off in 2015. The project aimed to develop tools for development and deployment of multimedia orchestration, sensor data as well as interactive screen technologies in the context of solutions for smart cities. This document provides an overview of the state of the art of relevant technologies in the scope of the project.

State-of-the-art (sometimes cutting edge or leading edge) refers to the highest level of general development, as of a device, technique, or scientific field achieved at a particular time. It also refers to such a level of development reached at any particular time as a result of the common methodologies employed at the time.

By making the State of the Art explicit and by describing it at a considerable detail level, we are able to raise the bar for innovations in the project. By aiming for innovations beyond the state of the art, and in line with our vision of what is relevant for smart city stakeholders, we strive to deliver highly relevant solutions that will allow the project partners access to a healthy economical perspective and exploitation phase.

More and updated information about the MOS2S project can be found at

[https://itea3.org/project/mos2s.html](https://itea3.org/project/mos2s.html%20) .

1. State of the Art Analysis

The MOS2S project will make significant contributions in several areas and on multiple topics. For these topics (listed below) we provide a detailed state‑of‑the‑art analysis.

|  |  |
| --- | --- |
| **Use cases** | Data-driven and immersive fan experiences for live events |
| Professional content production and distribution in new format:Crowd-sourced journalism in news production |
| Live debates, enhanced with emotion and context analysis |
| **Enabling technologies** | Data and video acquisition & capture |
| Data analysis, brokering and orchestration |
| Video analysis and processing algorithms |
| Video brokering and orchestration |
| Video streaming and delivery |

**Use case - Crowdsourced journalism in news production**

More and more news organizations are trying to fit citizen journalism into their coverage of local and world events. Crowdsourcing has an increasing impact at every stage of the news or broader content production process: real-time fact-checking and curation of breaking news, story ideation and community-based story assignments, device-agnostic information gathering and interactive event visualisation, location-based audience engagement platform, etc.. Emerging crowd-sourced video collaboration apps and platforms allow organisations to co-create news and content stories with their customers, viewers, fans or experts. One of the main challenges is to make this workable and scalable in real-time production scenarios, enabling real-time guiding of user generated event coverage and providing intuitive editorial suites and dashboards to select, filter and process suitable and quality-proof content based on careful annotation and video synchronisation. At the consumption side, interactive and multi-device experiences should give the viewer a better understanding of what’s actually happening and how the scene really looks like. The above mentioned applications such as Periscope allow citizens to stream video directly to their social media (Twitter) followers, providing both a technological platform and a built-in audience. Other players such as UStream and Twitch are still mainly offering services in the context of single stream events, with the exception of YouTube.  Cameraad[[1]](#footnote-2), an application trialled in the summer 2015 by the Dutch news site NU.nl and TNO, adds curation by a central news room. Apps for Android and iOS are available[[2]](#footnote-3). Around the same time, the Dutch public broadcaster NOS did a similar live streaming trial at the start of the Tour de France in Utrecht, with Amsterdam based Ubideo[[3]](#footnote-4).

**Use case – Data-driven and immersive fan experiences for live events**

VRT is participating in the ICOSOLE[[4]](#footnote-5) project, where different cost-effective ways of capturing spatially outspread events are investigated. All the content, both professional and user generated, is synchronised in time and space to create an immersive experience for people who cannot attend the event. Methods for fusing audio-visual and sensor information into a coherent data representation have been developed. A prototype networked platform for streaming live content from mobile capture devices to content processing and editing services has been built, including tools for media production professionals to select, configure and review the content sources being used. At the consumption side, VRT has developed the *Wall of Moments[[5]](#footnote-6)* prototype, an interactive and immersive app that shows the latest and most interesting Moments of an event in a highly personalized mosaic, creating a near-to-real experience. In ICOSOLE, a Quality Analysis app was also developed for Android for technical quality assurance of UGC. The app automatically measures Noise Level, Luminance variations and Blurriness. MOS2S will build further upon ICOSOLE tools and insights, improving synchronisation capabilities, increasing orchestration flexibility and adding real-time communication.

SELVIE[[6]](#footnote-7) (Scalable, Efficient, and Low-delay Video Interaction during Events), a collaborative Flemish research project, investigates how the audience engagement in large-scale events can be increased by deploying scalable and reliable networks with massive amounts of connections and by incorporating a high-quality video streaming workflow combining professionally captured video content and UGC. The Selvie framework comes with a director component, which interfaces with the human event director and provides a way of deciding which visitor-made videos and pictures to show on the event’s screens and merge it with other professionally captured content. A UGC collector component matches the requests of the director with the metadata of the available streams to decide on which content to collect and forward to the director. Although Selvie shares some interesting functionalities and principles with MOS2S, it does not deal with the orchestration and synchronisation challenges, the real-time streaming capability and the live dashboard and director’s cockpit. It also lacks the real-time feedback to end users while they are capturing.

S.M. Entertainment, one of the largest entertainment companies in Korea, recently introduced a new business model for consuming music concerts, "Surround Viewing”. Surround Viewing requires a multiple camera system to capture wide field of views which are concatenated and a display technique in a theatre for providing viewers with a vivid, live feeling to concerts projected onto three massive widescreens surrounding the audience. ETRI supports this concept with a multi-camera rig and monitoring system, to provide ultra-high quality video. This new immersive content technique brings fascinating experiences but at this moment only supports three HD views with considerable bezels between screens.

**Enabling technology - Data and video acquisition & capture**

In terms of stadium video surveillance, we can distinguish between two solutions, i.e. using moving (PTZ) cameras or using fixed cameras. Typically a combination of both solutions is used. The most frequent solution is using moving cameras with cameras strategically placed, often high up, and offering the possibility to have an overview (when zoomed out) and detail with up to 36x zoom. The PTZ camera is controlled by the operator using a mouse or CCTV keyboard/joystick. With different fixed lenses options (wide to short angle), stadiums can be fully covered with fixed cameras. This has now been facilitated by multi-megapixel technologies, so that full coverage with high resolution can be achieved with a reasonable amount of cameras. Clearly the trend is towards a fixed installation of multi-megapixel cameras, enriched with some PTZ cameras for full flexibility. This is applicable to every stadium type, from smaller to bigger with cost proportional to its size. Even with the highest resolution cameras, there are still dozens - if not hundreds - of cameras deployed in a single stadium. To overcome this and ease operation, "stitching" technologies have been introduced. Stitching allows images from multiple cameras to be combined in real-time and create a single image (which could be for example of a stand). Stitching software has been developed for the broadcast industry and is now becoming available for the security market.

Data acquisition and capture or multisensory perception implies new dimensions to capture the reality that will provide a higher immersivity or better understanding of the scene. Audio-visual capturing system are enriched with sensor information such as GPS, compass, inertial sensors, thermographic imaging, torque sensors, etc.[[7]](#footnote-8). Moreover, user comments that are becoming so popular on media that include social network services, can be considered as a new type of (semantic) sensor information that can have a strong impact in the management of media content.

Video will be captured with traditional cameras (supplied by partner Bosch) , supplemented with the Ultra-Wide View (UWV) capture system in Ultra-High Definition (8Kx4K@60fps) for professional users supplied by partner ETRI (KR).

InMotio (NL) will provide a real-time and wearable sensor (RFID tags) data acquisition system, focussed on player tracking, together with an on-line platform that combines position tracking data, event data, video and storage.

**Enabling technology - Data analysis, brokering and orchestration**

Context analysis depends on additional metadata that is available about a particular stream. Dublin Core Application Profiles (DCAP) specifies at a minimum which metadata terms are used within a particular application[[8]](#footnote-9). A notable example is semantic technology for multimedia production[[9]](#footnote-10). A framework called Semantic Streams[[10]](#footnote-11) allows users to pose declarative queries over semantic interpretations of sensor data. According to Nilsson[[11]](#footnote-12), in order to promote metadata harmonisation, metadata models (semantics) are much more important than metadata syntax or metadata application profiles. Classifying metadata by means of ontologies or tag clouds is not sufficient though - some entailment or reasoning capability is required to match publishers and subscribers based on the context of the streams. For MOS2S the relevance of the data is dependent on the context in which the data was created. In particular, it is in many cases necessary to keep track of the spatial and temporal context, i.e. where and when the data was created, in order to decide where, to whom and for how long this data is relevant and to prioritise the most relevant data.

MOS2S both considers multimedia, sensor and data streams as input. For sensor and data streams, a wide range of broker solutions are available that all have made specific trade-offs between the requirements they have to implement. For multimedia streams, the typical solutions are split between servers that handle live streams and servers that handle stored streams. The data broker platform considered for MOS2S as back-end for the demonstrators is the WWS (World Wide Streams) platform[[12]](#footnote-13) provided by Nokia Bell Labs. It has a broker architecture that is geared towards real-time, high bandwidth data streams. Its video capabilities will include webrtc[[13]](#footnote-14) for live stream publishing and subscribing and HTTP for stored stream playout. The broker has built-in analysis and processing capabilities; this includes an interest broker that can select data streams based on relevant metadata information and can also report on trending metadata. MOS2S will use available state-of-the-art data analytics techniques to derive trending concepts and topics and possibly trending stories and their context, and feed this context into the interest broker.

Media sensor orchestration is about correlating sensor streams by adding high quality metadata to recordings, including those from sensors on consumer devices. In a professional recording, cameras are gen-locked and calibrated. Using consumer devices for capture, adding good metadata is key to creating this same functionality here. Current timing mechanisms are either not very accurate (NTP), not accessible on many consumer devices (GPS)[[14]](#footnote-15) or not available everywhere (beacons). Much work has been done to improve clock synchronisation, e.g. by Ridoux[[15]](#footnote-16), and some work has been done in combining NTP and probing[[16]](#footnote-17),[[17]](#footnote-18) to increase accuracy and lower convergence time. For location, GPS may not be accurate enough, agricultural GPS is very accurate relatively but not absolute, and works with a special beacon[[18]](#footnote-19). For connectivity, it is important to have a view on available network bandwidth for real-time use cases, and encoding needs to be adapted for this, e.g. using tiling mechanisms. Various probing mechanisms for this exist. Finally, metadata is added to media streams, e.g. timing data[[19]](#footnote-20),[[20]](#footnote-21) and other metadata as well[[21]](#footnote-22).

**Enabling technology - Video stitching**

Video stitching has been around for some time now. Recently, efforts have been made to improve this to real-time[[22]](#footnote-23). Recent solutions to capture 360° content exist with using low cost video cameras (GoPros with 360Heros). Software from ETRI uses high-powered graphics cards designed for advanced video gaming to stitch thousands of frames per minute (GPU processing). Tools for stitching videos for camera clusters typically require a fixed setup of the camera cluster, with normally equal camera types. The tools are based either on templates (e.g. VideoStitch) or work with a feature based approach combined with templates (e.g. Autostitch[[23]](#footnote-24)). The templates can be generated with a sample scene and a feature based approach, e.g. with Hugin[[24]](#footnote-25). Nearer objects can only be stitched with a feature-based approach, but for a much higher computational cost incompatible with real-time constraints, as opposed to fixed templates. The output is one linear video showing the surround view. This works only for static cameras that are all (near) the same position and highly symmetrically placed. For arbitrary input, algorithms similar to that coded in Photosynth[[25]](#footnote-26) are necessary. This is a tool from Microsoft, which merges still images that are taken with different cameras from different point-of-views. It is based on interest point detection and matching. The output is a surrounding image mosaic; the images are placed on billboards in a 3D scene around the viewer. This scene is static in its geometry (it does not change during time) and static in its texture (the images on the billboards are still, they are no videos).

On the computer vision (and sound) side there are two important fields we want to address. The first field is image quality assessment, where the goal is to automatically evaluate the quality of images in agreement with human quality judgments[[26]](#footnote-27),[[27]](#footnote-28),[[28]](#footnote-29). This is a well-researched field to qualify images with the purpose to automatically improve them in image processing applications. These techniques start from a learning set of human judged images to build a model for image quality analysis. They do not address the complexity of live video stream however. In a second field we would research how to analyse and visualize emotion and discussion context in a live debate by analysing the video and audio with artificial intelligence tools.

**Enabling technology - Video brokering and orchestration**

Several tools for media orchestration (data models, protocols, frameworks, object models, APIs) are already available from standardisation organisations. MPEG has designed a number of media composition technologies (MPEG-4 BIFS, LASeR, MMT-CI). Such technologies are perfectly suited for complex compositions of media streams at the consumption side, on a single device. MPEG transport technologies usually allow for the declaration of associated resources through hyperlinking, addressing part of the media announcement issues. However, all these technologies are used to communicate a list of known locations and will probably not be adequate for media discovery. MPEG TEMI is able to deliver a timeline for external data in a main transport stream. DVB Companion Screens and Streams (CSS) has defined many protocols, some of which could be relevant to MOS2S, e.g. *DVB-CSS-WC (*wall-clock synchronisation between TV Device and CSA, a quickly-converging NTP variant), and *DVB-CSS-TS (*timeline synchronisation, timestamping of video frames at the TV Device with the wall-clock).

Orchestrated citizen information and participation scenarios requires in-depth understanding of one or multiple scenes or locations where the action takes place, and their relation with contributed user gen footage. There are several standards that have been designed with the focus on configuring and specifying the composition of a multimedia scene and its low level components. Some of most important efforts are: (1) W3C-SMIL (Synchronized Multimedia Integration Language)[[29]](#footnote-30), an XML-compliant W3C standard focused on the composition of a media scene and its interactivity, mainly video and 2D content; (2) MPEG-4 Part 11[[30]](#footnote-31) (Scene description and application engine) encompassing BInary Format for Scenes (BIFS, defines the content, composition and compression of a scene) and the XMT Framework, comprising substantial portions ofSMIL, W3C Scalable Vector Graphics (SVG) andX3D (the new name of VRML); and (3) MPEG-7 (ADVP and CDVS), where specific metadata could be applied to scene configuration and composition. Visual authoring tools have been developed over these standards (e.g. in FP7 TOSCA-MP for ADVP), but more as sample tools and proof of concept than real production tools, let alone for live use cases.

Within the FP7 project STEER, a video orchestration and brokering platform was developed by TNO to synchronize video streams originating from both professional broadcast cameras as well as mobile phones. The orchestration involved timestamping of video streams at different ingest points, and temporal alignment based network clocks such as NTP. The platform was further developed together with Dutch online news provider Nu.nl, enabling the selection of incoming video streams from mobile devices.

**Enabling technology - Video streaming and delivery**

MPEG-DASH, enabling smooth over the top streaming of multimedia content, is available as international standard, and a first versions of the new standards enabling spatial relationship descriptors[[31]](#footnote-32) and the signalling for timeline alignment are becoming available[[32]](#footnote-33). Additionally, synchronization efforts at the client mainly focus on RTP-based streaming with only preliminary results (early prototypes) for DASH-based delivery. New immersive and interactive content formats are currently not subject to discussion or standardization in the context of MPEG-DASH. Initial work is being conducted within the ICoSOLE project and MPEG has started an exploration activity related to media orchestration, both with active participation of MOS2S partners, specifically TNO.

**Link to previous and/or current collaborative research projects:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Project Name | Cooperative Programme | Time period (approx.) | Technical Focus | Relationship |
| ICOSOLE (VRT) | FP7 | 2013-2015 | Allows users immersively experience live events (spatially spread out, such as festivals) by combining quality spatial video, audio and user generated content. | Like MOS2S, it has multiple contributing sources and a mobile capture app, content tagging and filtering, selection and production tools in a live broadcast context.  |
| TOSCA-MP (VRT) | FP7 | 2010-2012 | Scalable and distributed content processing and annotation methods with advanced multimodal information extraction and semantic enrichment.  | MOS2S will build upon diverse video analysis and feature extraction tools for live and non-live production scenarios. |
| FASCINATE (ALU, TNO) | FP7 | 2010-2013 | Immersive interactive broadcast experience, with a special tiling approach for interactive delivery of very high resolution content. | Common partners with MOS2S will re-use the interactive tiling approach and build on this for interactive delivery to VR equipment. |
| HBB-NEXT (TNO) | FP7 | 2012-2014 | User-centred technologies for enriching the TV-viewing experience, with (o.a.) multi-device synchronisation for media play out. | MOS2S re-uses the synchronisation solution and extends this to multi-sensory synchronisation.  |
| STEER (TNO) | FP7 | 2012-2014 | Experimentally exploring the dynamic relationship between social information and networked media STEER developed a solution for synchronisation amongst multiple capture devices. | Common partners with MOS2S will build upon the synchronisation approach to meet more stringent requirements, such as frame-accurate synchronisation. |
| OpenTran sportNet (iMinds) | EU-CIP | 2014-2017 | Open Data tooling & publishing in a European Open Data Transport Network. | MOS2S will build the RMLStreamer with real-time (sensor) data characteristics replacing the RMLMapper |
| Apps4EU(iMinds) | EU-CIP | 2013-2014 | Publication of Linked Open Data concerning the different European Apps Competitions | MOS2S will extend RML-processor with real-time (sensor) data characteristics |
| Flander's Open Data Publishing(iMinds) | EWI  | 2013-2014 | Awareness creation and deployment of Open Data in Academic Publishing, Data Journalism and Governmental Data re-use. | MOS2S will build the RMLStreamer with real-time (sensor) data characteristics & visualisation libraries |
| TheData Tank | iMinds Open Source Framework  | 2011-2018 | "15 minutes" Linked Open Data Publishing framework (v5 deployed) | MOS2S will build the RMLStreamer with real-time (sensor) data characteristics & visualisation libraries replacing TheDataTank and RMLMapper |
| SELVIE(iMinds) | iMinds IWT | 2014-2016 | Stream visitor-made smartphone videos in real-time to the event screens | Knowledge on how to build an interactive user-gen video-platform will be taken into account |
| USENET (NOK) | ITEA | 2012-2014 | M2M service platform. | Large scale platform to reach out to sensors and actuators. Produced lots of ETSI standardisation. |
| WTEPlus (NOK) | iMinds | 2008-2010 | Study and enhance new paradigms for composition and deployment of applications. Our first dataflow engine. | Demonstrated a platform and marketplace for composable M2M services based on a dataflow concept. |
| DIYSE (NOK) | ITEA  | 2009-2011 | Support for creating aware, interactive and flowing experiences in an Internet-of-Things world.  | A follow-up project of WTEPlus, with more attention to prosumer creation of services. |
| M2MGRIDS (NOK) | ITEA | 2014-2018 | Creating real-time enablers for a dynamic cyber-physical information business ecosystem. | Combines reasoning and M2M sensor / actuator interaction in the energy domain, based on the WWS platform. |
| iCore (NOK, TNO) | FP7 | 2011-2014 | Demonstrator and platform for urban security (with Thales)  | One of the demonstrators showed narrow Artificial Intelligence features coupled to a director component for video streams. |
| SmartIP (NOK) | EU CIP | 2010-2014 | Platform for large-scale Zwerm trial in Ghent https://vimeo.com/65648085 | Zwerm was a large-scale city game running on the WTEPLUS/DIYSE platform, which later evolved into ALU’s current WWS platform. |
| Care4Balance (NOK) | EU AAL | 2013-2016 | Demonstrator and platform supporting quality of life improvement for the older adult. | Another long-time running demonstrator on the WTEPLUS/DIYSE/WWS platform. |
| ImmersiaTV(iMinds, VRT) | H2020 | 2015-2017 | An end-to-end toolset covering the entire audio-visual broadcast value chain and associated demonstrators. | Focuses on an immersive video distribution chain, see http://www.immersiatv.eu |

Table 1: Related collaborative research projects.

1. [Cameraad: livestreams van ooggetuigen in het nieuws](https://www.svdj.nl/nieuws/cameraad-livestreams-van-ooggetuigen-in-het-nieuws/) SVdJ 2015-05-07 (in Dutch) [↑](#footnote-ref-2)
2. [NU.nl laat lezers het nieuws livestreamen met smartphone-app](http://numrush.nl/2016/02/03/nu-nl-laat-lezers-het-nieuws-livestreamen/). Numrush 2016-02-03 (in Dutch) [↑](#footnote-ref-3)
3. [NOS en Ubideo laten toeschouwers Tour de France live verslag doen](http://numrush.nl/2015/07/03/nos-en-ubideo-laten-toeschouwers-tour-de-france-live-verslag-doen/) Numrush 2015-07-03 (Dutch) [↑](#footnote-ref-4)
4. http://icosole.eu/ [↑](#footnote-ref-5)
5. [Bauwens R. e.a.: *The Wall Of Moments: An Immersive Event Experience At Home*](http://www.ibc.org/files/rik_bauwens_the_wall_of_moments_an_immersive_event_experience_at_home.pdf) I[B](http://www.ibc.org/files/rik_bauwens_the_wall_of_moments_an_immersive_event_experience_at_home.pdf)C2[0](http://www.ibc.org/files/rik_bauwens_the_wall_of_moments_an_immersive_event_experience_at_home.pdf)1[5](http://www.ibc.org/files/rik_bauwens_the_wall_of_moments_an_immersive_event_experience_at_home.pdf) [↑](#footnote-ref-6)
6. http://www.iminds.be/en/projects/2014/11/28/selvie [↑](#footnote-ref-7)
7. Bailer W., Pike C., Bauwens R.,Grandl R., Matton M., Thaler M.: *Multi-sensor concert recording dataset including professional and user-generated content*. Proceedings of the 6th ACM Multimedia Systems Conference, pp 201-206, ACM, March 2015. [↑](#footnote-ref-8)
8. CEN: *Guidelines for machine-processable representation of dublin core application profiles*, ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/MMI-DC/cwa15248-00-2005-Apr.pdf. [↑](#footnote-ref-9)
9. Nack F., Van Ossenbruggen, Hardman L. J.: *That obscure object of desire: multimedia metadata on the web, part 2*, IEEE Multimedia 12 (1) 54-63. [↑](#footnote-ref-10)
10. Whitehouse K., Zhao F. , Liu J.: *Semantic streams: a framework for composable semantic interpretation of sensor data*, in Proceedings of the Third European conference on Wireless Sensor Networks, 2006, pp. 5-20. [↑](#footnote-ref-11)
11. Nilsson M.: *From interoperability to harmonization in metadata standardization*, <http://kmr.nada.kth.se/papers/>SemanticWeb/FromInteropToHarm-MikaelsThesis.pdf. [↑](#footnote-ref-12)
12. Weldon M. K.: *The Future X Network: A Bell Labs Perspective,* CRC Press, 2015; p. 309 [↑](#footnote-ref-13)
13. <http://www.webrtc.org/> [↑](#footnote-ref-14)
14. Kooij W.J., Stokking H.M., Van Brandenburg R., De Boer P.T.: *Playout delay of TV signals: measurement system design, validation and results.* Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and online video, pp.23-30, ACM, June 2014. [↑](#footnote-ref-15)
15. Ridoux J., Veitch D.: *Principles of robust timing over the Internet.* Queue, Vol 8(4), pp. 30, 2010. [↑](#footnote-ref-16)
16. Tsuru M., Takine T., Oie Y.: *Estimation of clock offset from one-way delay measurement on asymmetric paths.* Proceedings of the 2002 Symposium on Applications and the Internet (SAINT), pp.126-133, IEEE, 2002. [↑](#footnote-ref-17)
17. Gotoh T., Imamura K., Kaneko A.: *Improvement of NTP time offset under the asymmetric network with double packets method.* Conference on Precision Electromagnetic Measurements, 2002, Conference Digest, pp.448-449, IEEE, June 2002. [↑](#footnote-ref-18)
18. Hofmann-Wellenhof B., Lichtenegger H., Collins J.: *Global positioning system: theory and practice.* Springer Science & Business Media, 2012. [↑](#footnote-ref-19)
19. Veenhuizen A.: *frame accurate media synchronisation of heterogeneous media sources in an HBB context.* Proceedings of the Media Synchronisation Workshop, January 2012. [↑](#footnote-ref-20)
20. Le Feuvre J., Singer D.: *Proposed Exploration of Uniform signalling for timeline alignment,* MPEG contribution, December 2013. [↑](#footnote-ref-21)
21. ISO/IEC 15938-1:2014: *Information Technology -- Multimedia content description interface -- Part 1: Systems.* ISO/IEC, Genève, 2014. [↑](#footnote-ref-22)
22. El-Saban M. A., Ezz M., Kaheel A.: *Fast stitching of videos captured from freely moving devices by exploiting temporal redundancy*. ICIP, pp. 1193-1196, 2010. [↑](#footnote-ref-23)
23. <http://www.cs.bath.ac.uk/brown/autostitch/autostitch.html> [↑](#footnote-ref-24)
24. Hugin - Panorama photo stitcher. http://hugin.sourceforge.net/ [↑](#footnote-ref-25)
25. Uricchio W.: *The algorithmic turn: Photosynth, augmented reality and the changing implications of the image*. Visual Studies, vol. 26(1), pp. 25-35, 2011. [↑](#footnote-ref-26)
26. Mohammadi P., Ebrahimi-Moghadam A., Shahram S.: *Subjective and Objective Quality Assessment of Image: A Survey*, arxiv 2014 [↑](#footnote-ref-27)
27. Ye P. , Doermann D.: *Active Sampling for Subjective Image Quality Assessment*, CVPR 2014 [↑](#footnote-ref-28)
28. Ye P., Kumar J., Doermann D.: *Beyond Human Opinion Scores: Blind Image Quality Assessment based on Synthetic Scores*, CVPR 2014 [↑](#footnote-ref-29)
29. Bulterman D., et al., *Synchronized Multimedia Integration Language (SMIL) 3.0.* W3C Recommendations, 2008. [↑](#footnote-ref-30)
30. ISO/IEC14496-11 [↑](#footnote-ref-31)
31. ISO/IEC 23009-1:2014/FD AM 2, *Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 1: Media presentation description and segment formats, AMENDMENT 2: Spatial Relationship Description, Generalized URL parameters and other extensions* [↑](#footnote-ref-32)
32. ISO/IEC 13818-1/FD *Amd 2, Delivery of timeline for external data*. [↑](#footnote-ref-33)