One Domain Model for Software-Intensive Ingestion of Stereoscopic 3D Content

Aleksandar Spasic[1] and Dragan Jankovic[2]

Abstract – **Three-dimensional television (3D TV) is expected by many to be the next step in the advancement of television. What is needed is a complete re-thinking of the way technology can be applied to the art and business of 3D program making and low-budget broadcasters should reconsider alternatives of partial in-house developments. Stereoscopic 3D TV content ingestion workflow is analysed in this paper. Modelling in problem space is used as a research method in this paper and behavioural description is modelled by the use case diagram. Class diagram is used to describe static representation of 3D content during ingestion phase of the stereoscopic content production.**

Keywords – **Stereoscopic 3D TV production, Model of Problem Space (MOPS), Ingest system, 3D Content life cycle, Domain model**

I. Introduction

Primary aim of this paper is to propose one model-driven approach of ingestion system suitable for manipulation with stereoscopic (3D) multimedia content. The target of the project is to figure out one of possible future scenarios for implementing stereoscopic 3D technologies suitable for use in low-budget TV production.

The first step of the project is to describe the whole stereoscopic content ingestion chain, from planning and defining criteria for indexing, through acquisition, indexing and metadata manipulation, to final encoding and archiving.

The second step is to propose one 3D TV ingest workflow. This step is divided into two parts, the first one is to analyse the behavioural description of ingest processes, and the second one is to define the conceptual model in problem space which is suitable for low-budget 3D TV production.

Software-intensive ingesting is a part of 3D TV production system and can be considered as a part of multimedia information system and development practice applicable for that type of information systems should be applied [1], [2].

Modelling in problem space is used as a research method in this paper. A model, by its very nature, is an abstraction of the reality. Software projects use modelling throughout the entire life cycle. Successful modelling needs to consider the areas in which modelling needs to take place. These modelling spaces have been formally considered and discussed by Unhelkar in [3]. The three distinct yet related modelling spaces are defined: problem, solution and background. The modelling output in such software projects transcends both data and code and results in a suite of visual models or diagrams.

In Unified Model Language (UML) projects, model of problem space (MOPS) deals with creating an understanding of the problem, primarily the problem that the potential user of the system is facing. While usually it is the business problem that is being described, even a technical problem can be described at the user level in MOPS. In any case, the problem space deals with all the work that takes place in understanding the problem in the context of the software system, before any solution or development is attempted. Thus the problem space would focus entirely on what is happening with the business or the user [4].

Problem space will need the UML diagrams that help the modeller understand the problem without going into technological detail. The UML diagrams that help express what is expected of the system, rather than how the system will be implemented, are of interest here. These UML diagrams in the problem space are as follows: Use case diagrams, Sequence and state machine diagrams and Class diagrams.

Authors' previous experiences with modelling of problem spaces in television production are outlined in [5], [6].

II. 3D Ingestion Workflow

The majority of 3D broadcast material available today has been produced using a twin-lens [7] or dual-camera configuration giving a stereo pair where the left-eye and the right-eye views are separately recorded from slightly different perspectives.

Model of life cycle of stereoscopic content shows the entire behaviour of the object, as it changes its state in response to the messages it receives. The state machine diagram representing the life cycle of the 3D TV content is shown in Figure 1.

The nature of the state machine diagram is considered dynamic-behavioural. The state machine diagram has the ability to represent time precisely and in a real-time fashion. "What happens at a certain point of time?" is a question that is answered by this diagram.

A fundamental way to capture a stereoscopic TV signal is to use two cameras mounted on the same axis and separated by the spacing of the average pair of human eyes (6.25cm). Camera spacing can be varied, and cameras can be 'toed in',

[1]Aleksandar Spasic is with College of Professional Studies for Pre-School Teachers, Cirila i Metodija 29, 18300 Pirot, Serbia E-mail: aspasic@hotmail.com

[2]Dragan Jankovic is with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Nis, Serbia, E-mail: dragan.jankovic@elfak.ni.ac.rs.

to achieve different elements of picture composition. Filming parameters such as camera base distance (distance between the two cameras), convergence distance (distance from the cameras to the point where both optical axis intersect), and camera lens focal length can be used to scale the horizontal disparity and thus the degree of perceived depth.

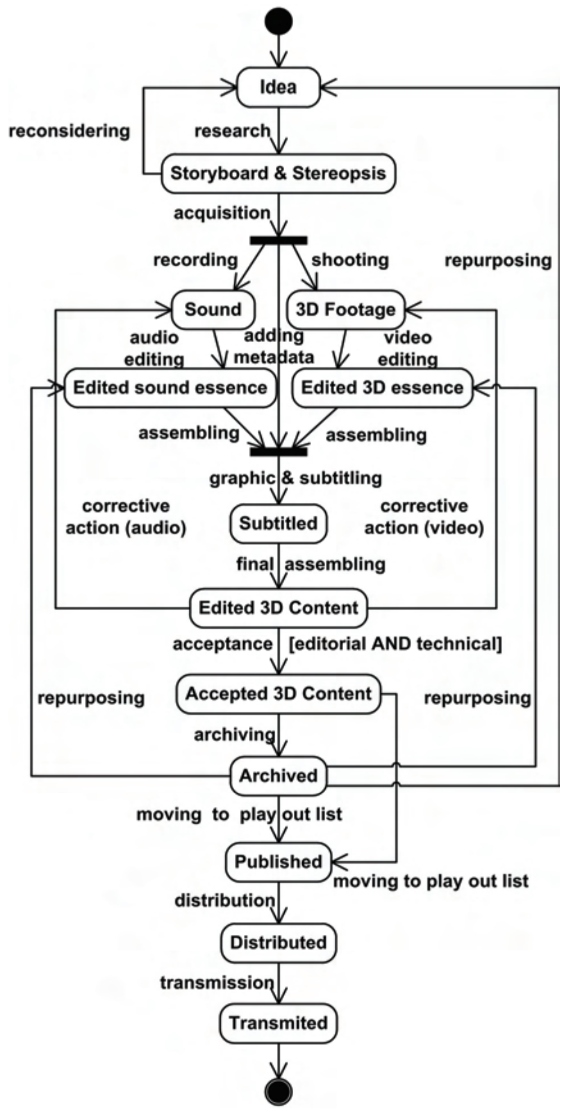

Fig.1. State Machine Diagram of 3D TV Content Life Cycle

At all points in capture there is an opportunity for metadata collection [8]. Devices that capture pictures or sound can automatically record much of the technical metadata from their own control systems—cameras that keep track of f-stop, filter wheel settings, and focal length are obvious examples. Likewise, it is becoming common for devices to capture the time of day and date and even the latitude, longitude, and altitude of their position when recording of the clip started. Those technical metadata can be very important during the postproduction phase, especially if dual camera adjustment and settings was not perfect and some minor imperfections and 3D image impairment can be corrected.

Ingest is the first stage to efficiently transfer captured 3D content to the television production infrastructure. During the ingest we take all the 3D content (both view) collected during a shoot, as well as new metadata, and transfer it into the production environment. More metadata can be generated at ingest and this can either be directly entered, for example by an operator marking technically poor sections, or regions for special processing, or it can be extracted automatically.

Simplified 3D content ingestion workflow is shown in Figure 2.

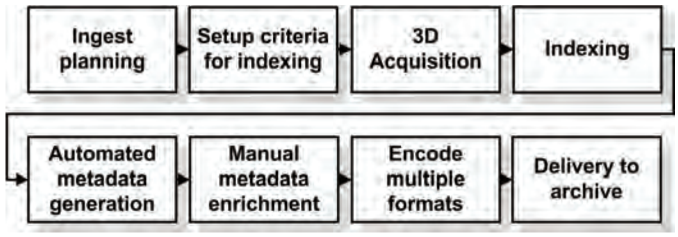

Fig. 1. 3D Ingest Workflow

Ingestion of multimedia content can be generally considered in terms of two processes, or fundamental tasks [9], [10]:

- Content acquisition and optimization, and
- Content description and referencing.

Content acquisition and optimization assume capturing the 3D audio-video essence and content compression.

Acquisition of stereoscopic content assumes transfer of two video streams, one captured for left eye and one captured for right eye. Low-budget ingesting and capturing based on two cameras usually assumes sequential transfer, i.e. ingest operator should start transfer for left and right view separately.

Typically, users of Content Management System will want to utilize high resolution master file which contains the content in professional broadcasting quality, as well as low resolution proxies (considered also as a meta-essence [8]) of same content for searching and previewing archived material or for web delivery. Ingest system should provide automatic generation of high and low resolution content representations.

Standardization of master format for 3D Home production is on-going [11] and details related to the quality of 3D video and audio as well as technical metadata important for ingest process will be suggested.

During the content optimization the key frames should be extracted and recorded. Key frames are valuable for providing asset management solutions with representative images for browsing video, as well as for making edit decisions. Key frames should be extracted and converted to JPEG images based on scene changes or predefined time intervals. Key frames can be extracted only for one eye or for both.

III. MODELLING BEHAVIOURAL DESCRIPTION OF 3D INGEST PROCESS

The main objective of a behavioural description is to visualize how the user (represented by the UML actor) will interact with and use the system. This is done by showing the actor associating with one or more use cases and, additionally, by drawing many use case diagrams.

Main actors in problem space of stereoscopic content ingest process are producer, ingest operator, 3D essence gathering crew (cameraman, sound recorder) and ingest automated system.

Use cases important for modelling in problem space of ingest process are as follows: Ingest planning, Setup criteria for indexing, Start acquisition, Automated metadata generation, Manual metadata enrichment, Indexing, Encoding multiple formats and Delivery to Archive. Use Case diagram of 3D ingest process is shown in Figure 3.

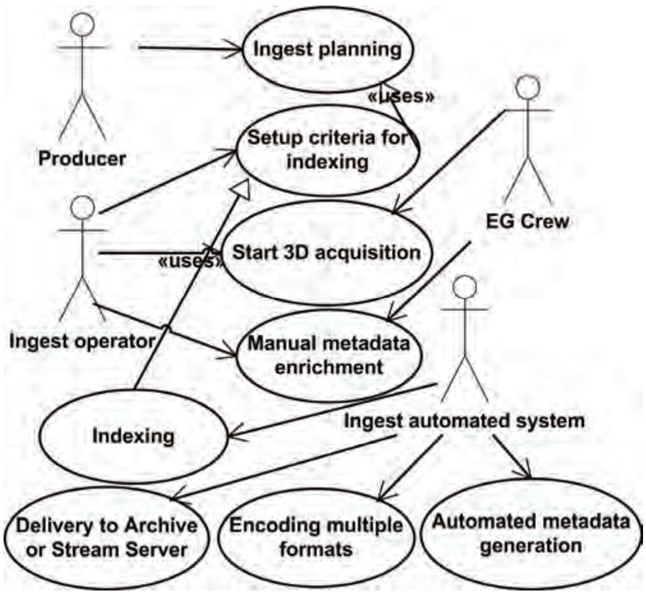

Fig. 2. Use Case Diagram of Stereoscopic Ingest Process

A. *Use Case - Start Acquisition*

Short Description: 3D video shoots, audio clips and other essence items, as well as attached metadata are extracted from the continuous acquisition bin, such as cameras, sound recorders, digital tape players, different editing workstations, servers etc.

Actors: Ingest operator and essence gathering crew members.

Pre-Conditions: 3D video essence is shoot for both views, audio essence is recorded, other program items are created and pre-selected.

Post-Conditions: All 3D essence materials, as well as related metadata, are ingested into production system.

Main Flow: (1) Ingest operator starts acquiring the 3D raw material or previously produced essence from cameras or other sources, separately for each view. (2) During ingest all the content collected during a shoot, recording and repurposing, as well as new metadata is taken and transferred into the production environment. (3) Ingest operator reviews what he/she has, and marks down its possible use. Use Case terminates.

B. *Use Case - Indexing*

Short Description: Ingest system extracts a number of key attributes from the source essence and converts them to metadata.

Actors: Ingest automated system.

Pre-Conditions: Criteria for indexing are defined and parameters are setup.

Post-Conditions: Key attributes are extracted and related metadata are produced.

Main Flow: (1) System analyses the acquired essence in accordance with parameters which are previously set-up. (2) Metadata generator module produces related metadata. Use Case terminates.

C. *Use Case - Automated metadata generations*

Short Description: System generates metadata from the acquired 3D essence files.

Actors: Ingest automated system.

Pre-Conditions: 3D essence material, both views, as well as files with related metadata, is ingested into production system.

Post-Conditions: Metadata are stored in database.

Main Flow: (1) System search for metadata files accompanying acquired essence. (2) System checks the metadata format and if metadata format is in accordance with systems metadata formats, metadata are stored in database. (3) If metadata format does not confirm, corrective actions must be undertaken [alternate flow A1]. (4) Steps (1) and (2) are repeated until all metadata is generated and Use Case terminates.

Alternate Flow (A1): No need for corrections. Metadata are approved and stored in database.

D. *Use Case - Manual metadata enrichment*

Short Description: Automatic extracted metadata can be validated. New metadata (descriptions, business information etc.) can be added.

Actors: Ingest operator, Essence gathering crew member.

Pre-Conditions: Essence material, as well as files with related metadata, including indexed and generated metadata are ingested into production system.

Post-Conditions: Stored essence, as well as related metadata.

Main Flow: (1) Ingest operator validates previously generated indexes and metadata. (2) Ingest operator add new descriptions of essence. Use Case terminates.

E. *Use Case - Encoding multiple formats*

Short Description: Create high-resolution master essence file and low-resolution proxy files.

Actors: Ingest automated system.

Pre-Conditions: 3D essence acquired, metadata generated and stored in database.

Post-Conditions: 3D essence encoded in hi-resolution version and stored in archive. 3D essence encoded in low-resolution version.

Main Flow: (1) In accordance with production format 3D essence is encoded in high quality broadcast format. (2) 3D essence is encoded in several different versions of low-res. files. Use Case terminates.

F. Use Case Delivery to Archive and Stream Server

Short Description: System stores 3D essence in deep archive and send low-resolution proxies to stream server.

Actors: Ingest automated system.

Pre-Conditions: 3D essence is encoded in hi-resolution version. 3D essence is encoded in low-resolution version.

Post-Conditions: 3D essence is stored in archive. 3D essence is delivered to stream server.

Main Flow: (1) System stores essence in deep archive. (2) System delivers low-resolution proxies to stream server which serves low-resolution proxies for searching, previewing, non-linear editing etc. Use Case terminates.

IV. DOMAIN MODEL OF 3D INGEST PROCESS

The objective of modelling structural static representation is to represent, in one or more views, various business entities and their relationships in MOPS.

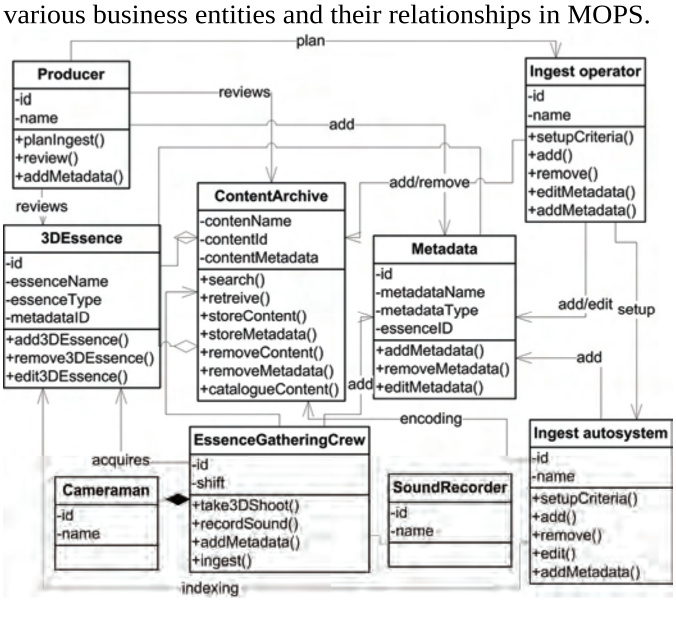

Fig. 3. Class Diagram of Stereoscopic Ingest Process

Class diagrams, by their very nature, are very strong, structural, static representations. Class diagrams show business-level classes as well as technical classes derived from the implementation language. In addition to showing the classes, class diagrams show the relationships between them. The entire description of the classes (or "entities," as they may be called in the problem space) and their relationships with each other is static. No dependency is shown in this diagram and no concept of time. Class diagram of the ingest process is shown in Figure 4.

V. CONCLUSIONS

The primary aim of this paper is to describe ingest system suitable for 3D content ingestion as one of the main areas in a 3D content production environment and to summarize the 3D essence, metadata and control flow, as well as the main processes involved in ingesting of 3D television content.

The first step in modelling of software intensive ingest system, modelling of problem space, is presented here. Use Case diagram is used as tool for behavioural description of the ingest system. Class diagram is used for description of Domain model.

The challenge for the future is to make complete business analysis of problem space using the use case, class, activity, state machine and sequence diagrams. Also, the model of solution space, as well as the model of background space should follow the current work presented in this paper.

REFERENCES

[1] C. Barry, M. Lang, „A comparison of 'traditional' and multimedia information systems development practices“, Information and Software Technology Vol. 45, pp 217–227, 2003.

[2] T. P. Rout, C. Sherwood, "Software Engineering Standards and the Development of Multimedia-Based Systems", Fourth IEEE International Symposium and Forum on Software Engineering Standards (ISESS '99), Curitiba, Brazil, pp.192-198, 1999.

[3] B. Unhelkar, *Verification and Validation for Quality of UML 2.0 Models*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2005.

[4] M. O'Doherty, *Object-Oriented Analysis and Design: Understanding System Development with UML 2.0.*, Chichester, West Sussex: John Wiley & Sons Ltd., 2005.

[5] A. Spasic, "Business Analysis of Software-Intensive Television Production: Modelling the Content Production Workflow", Serbian Journal of Management, vol. 1, no. 2, pp. 17-32, 2006.

[6] A. Spasic, M. Nesic and J. Bogdanovic, "Production of TV Multimedia Content: Modelling in Problem Space", Proceedings of the XLI International Scientific Conference on Information, Communication and Energy Systems and Technologies ICEST 2006, June 29- July 01, 2006, Sofia, Bulgaria pp.212-215, 2006.

[7] *Panasonic AG-3DA1 Integrated Twin-Lens 3D video recorder*, Panasonic Corporation, 2010, Available: http://panasonic.biz/sav/broch_bdf/AG-3DA1_e.pdf

[8] M. Cox, E. Mulder, L. Tadic,*"Descriptive metadata for television: an end-to-end introduction"*, Elsevier, 2006.

[9] A. Spasic, D. Jankovic, "Framework for Software-Intensive Ingest System:One Behavioural Description", Proceedings of the ICT Innovations 2010 Conference, September 12-15, 2010, Ohrid, Macedonia, pp. 191-200, 2010.

[10] A. Spasic, D. Jankovic, "Modelling Software-Intensive Ingest System: Behavioural Description and Domain Model", Proceedings of 18th Telecommunications forum TELFOR 2010, Serbia, Belgrade, November 23-25, pp. 958-961, 2010.

[11] W. Zou, „Developing End-to-End Standards for 3D TV to the Home", SMTPE Motion Imaging Journal, October 2010, pp 32-38, 2010.