

D1.3 - Technical Studies

|  |  |  |  |
| --- | --- | --- | --- |
| Action | Name | Date | Approbation |
| Author | Cyril Albrecht |  |  |
| Approved by | Thierry Gruszka |  |  |
| Approved by |  |  |  |

History

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Version & Status | Author | Modifications |
| 30/03/12 | 1.0.draft.1 | CA | Creation |
| 23/04/12 | 1.0.draft.2 | TG | Modification |
| 30/05/12 | Final version | TG | References addition |
| 6/06/12 | Final correction | TG | References addition |
| 8/06/12 | Final correction | TG | References addition |
| 16/06/12 | Final correction | TG | References addition |

Table of contents

[1 Introduction 4](#_Toc327778763)

[2 Global and transversal technical studies. 4](#_Toc327778764)

[2.1 A1 – Investigating HW/SW solutions for Jedi MW. 5](#_Toc327778765)

[2.2 A2 - Portage of Unity 3 player on the STi7108 7](#_Toc327778766)

[2.3 A3 – Ideal workflow 3DSMax/Unity3/Orly 9](#_Toc327778767)

[3 Work Packages Studies 11](#_Toc327778768)

[3.1 Referenced Technical Studies 11](#_Toc327778769)

Glossary

To assure coherent terminology and abbreviations across all documents inside the project, the specific terminology and abbreviations for this deliverable has also been aggregated into an internal report called JEDI Terminology and Abbreviations available on the internal project website.

# Introduction

The starting point for the JEDI is stereoscopic content. This aspect will assess the requirements for multi-view to replace stereoscopic content, and impact such a change would have.

It is possible that certain graphics overlays may need additional information or processing in the case of 3DTV. An example of this may include subtitles or closed captions where there is a logical "depth" at which the subtitles should be placed. This depth information would be sent along with the subtitle data, and used by an altered graphics stack to place the subtitles in three, rather than just two, dimensions.

A number of companies are proposing various processing operations that improve the visual quality of 3D content, and/or improve interoperability between different display formats. This should investigate the claims of these companies, and the quality differences that this achieves, and how possible and practicable it would be to implement this technology either using the existing decoder chip hardware, or as an external processing stage.

# Global and transversal technical studies.

Led by NDS, in collaboration with Pace and PlanetMedia, this task has consisted into the study and the evaluation of the best rendering framework (HW & SW) while setting up the best tool chain able to generate expected S3D Graphics UI on STB (such as the existing PC 3D-EPG) as well as supplying a operational S3D development system for the User Experience Labs activities. Main highlights of this technical study can be resumed with the following outputs:

* Study of tool chain based on Collada standard: interesting but not enough advanced & performing to cope with JEDI objectives
* Study of rendering & tool chain based on Unity3, a 3D based gaming eco-system presenting really promising results & powerful framework, liable for JEDI’s STB & User Labs UX., with a notable easy way to dynamically change stereoscopic parameters. A first stage has been presented at the ITEA Co-Summit in October 2011.
* Study of porting this SW promising results on appropriated HW, Orly by STM been judiciously identified when just released.

During the 2 last workshops, it has been discussed that this document-delivery will also gather other WP’s studies, in order to provide a single repository for all JEDI’s technical Studies. This action been validated at the last workshop, these references are appened to this document in its 3rd chapter

## A1 – Investigating HW/SW solutions for Jedi MW.

|  |  |
| --- | --- |
| Technical study number | A1 |
| **Technical study name** | Identification of the available hardware and software solutions to bring the middleware development lab of Jedi. |
| Work package | WP1 |
| Purpose and objective | Identify and evaluate the available solutions from the market and internaly to bring a 3D engine and stereo A/V player on a STB.  bThe study includes the identification of the chipsets. This one must embbeded a GPU enaugh powerful to render a stereoscopic UI and must be capable to decode and display stereoscopic A/V contents.  It is also important, considering the STB architecture, that a DVDS2 frontend should be available in order to retrieve the A/V content from Hispasat broadcast.  Considering the software part of the study, the proposed engine must be able to run on the selected chipset architecture (CPU and GPU).  It also should bring a straight way, or at least the most efficient workflow , for the UI designers to target the STB with the high level editing tools for modeling and animations (Cinema4D, 3DSMax, build in editor). |
| Description | HW architecture:  Historically, both Pace and NDS are used to work into the STi7108 framework.  Pace has proposed to provided development platforms that are based on the that chipset.  The 7108 provides two SH4 core and a ARM Mali 400 GPU.  Il also, of courses, provides all the necessary IP to decode and display A/V contents.  SW architecture:  Alternative study paths proposal (in blue)     * Main goal: porting of existing 3D PC-Demo on to STB + optimizing * Reusing existing when relevant:   + existing Collada player running on PC/Win   + existing 3D EPG sources made with 3DSmax   + Existing non S3D-prototypes running with Unity3D eco-system+ Collect other study opportunities * Resulting plan   + Porting of Collada player on 7108   + Develop 3D-EPG with Unity, eventually starting from 3DSmax source   + Commit on Unity as the best available tool to produce JEDI requirements   + Establish a support from ST alongside JEDI |
| Technical constraints | * To be able to support a high level editing tools such as Cinema4D and 3DSmax * To set an easy path to integrate the developed EPG into the target engine and set top box, ie to define an ideal workflow from the design to the production. * To identify backends capable of handling the technical constraints of the project. * Identify the available 3d engines on the market that can be ported on the selected backend. * To reach good performances in terms of fps and user interaction, considering the technical constrains of the Jedi project, ie 1080p stereoscopy. |
| Results | Extremely poor performances have been reach on 7108. It has been proven very shortly during the study that the Collada player portage on STi7108 will never match the expected requirements. |
| Related results/studies | A2 and A3 |
| Reusalibility & Impact |  |
| In Jedi | Has led to other related technical studies. |
| Out of Jedi | N/A |

## A2 - Portage of Unity 3 player on the STi7108

|  |  |
| --- | --- |
| Technical study number | A2 |
| **Technical study name** | Portage of Unity 3 player on the STi7108 |
| Work package | WP1 |
| Purpose and objective | After having identified Unity 3 framwork as a good candidate for 3D UI rendering, the purpose of the technical study was to evaluate the possibility to port it onto the STi7108 and to bench the overall performances.  Unity 3 is a game/graphic framework which provides an authoring tool, a rendering engine available on most common platforms (such as Android, PC, iOS, MacOS, game consoles…), as well as a web distributing platform (not relevant for Jedi)  The main goal of that study consists in the porting of the rendering engine which combines three components:   * Graphical renderer * Interactivity and animations component (based on Mono technology, C# and Java) * Collision and physics (not relevant for Jedi) |
| Description | Porting plan:  **3D-EPG developing with Unity3D**   * Reusing 3DSMAX sources as an importation in Unity * Redeveloping from spec & assets   **Unity3D Porting on 7108**   * Graphic engine on Mali 400 * Mono engine portage on SH4 * Event management * Video/Stream as a texture (to be check if necessary) * S3D output format management and rendering (Interleave, SbS, TopBottom, etc)   **Integration & Result**   * 3D-EPG in STB/7108 with Unity3D * JEDI Tool Chain for WP3 |
| Technical constraints | To adapt the existing Unity sources on the SH4/Mali combo:   * reuse the existing portage on Mali400 of the graphical component * no existing portage of Mono to SH4 CPU * might require tight collaboration with hardware suppliers   To identify the best exchange format between design tools and unity editor:   * Collada * FBX * Native format of the editing tool |
| Results | * Very promising engine with a good graphical rendering by Mali400 * Mono requires Cortex A9 or x86 CPU architecture * FBX identify as a very good lossless exchange format for both scene description and animations. * Able to produce application close to PC version * Full application and interactivity available * Good candidate for:   + applicative MW on targeted STB   + demonstrator for WP4 as expected   + tool chain for WP3 & user lab. |
| Related results/studies | A3 |
| Reusalibility & Impact |  |
| In Jedi | * Expected JEDI assets for deliverables (cf. results) |
| Out of Jedi | * 3DTV1, BluRay, JEDI… * Produce a multi-device app, PC/MAC/box/tablet + STB |

## A3 – Ideal workflow 3DSMax/Unity3/Orly

|  |  |
| --- | --- |
| Technical study number | A3 |
| **Technical study name** | Setup of a ideal workflow 3DSMax – Unity3 – Orly |
| Work package | WP1 |
| Purpose and objective | The purpose of this technical study is summarizing the work done for the A1 and A2 which established the good software components, and now, to port it on the most oppropriate hardware combining Cortex A9 and Mali400.  Orly new upcoming STM chipset has been identified by Jedi technical stakeholders as the best available candidate in the frame of our miles stones.  Orly is providing the usual IP for decoding and display A/V contents, including the combo Cortex A9 and Mali400 in a most powerful version than the STi7108.  Porting Unity 3 on that framework establishes the definitive workflow to design, editing, generate and run a S3D EPG on a set top box and it ext |
| Description | **Action plan:**  **HW**   * To bring up the Orly evaluation platform into NDS premises   **Video Player**   * To support A/V decoding (1080p30 x2) alongside the Unity 3 engine.   **3D-EPG developing**   * To enable stereoscopic rendering in Unity 3 and so to continue close collaboration with Orly provider to integrate stereoscopy support * To restructure the original 3D-EPG project to improve the overall performances and the memory footprint * To enrich the knowledge of “good practices” for the 3DSMax developers, as well as with Cine4D * To allow live stereoscopic parameters control of the UI * To define the interaction mode between the A/V player and Unity     **Integration & Result**   * 3D-EPG in STB/Orly with Unity3D * JEDI Tool Chain for WP3 |
| Technical constraints | To separate the video from the graphical rendering, and to blend them thanks to the Orly compositor.  No huge technical contraints as the portage of Unity 3 is already available on such architecture.  To define the best code of conduct to produce graphical assets in 3DSMax considering the limitations in terms of memory and CPU load. |
| Results | * Very good performances in terms of rendering quality and FPS has been reached on Orly like platforms (A9/Mali).   + To be transfered on the Pace platform. * Validate completely the definitive workflow for Jedi:   + 3DSMax as the authoring tool   + FBX as the ideal exportation and importation format for the scene description and the animations   + Unity3d as the high level tool for the post production of the UI, meaning integration of the interactivity with the Mono engine * Define the inputs and the outputs for the Jedi contributors:   + XML files for the stereoscopic parameters   + Video and graphic layers to split the work done in the A/V player part and the UI rendering |
| Related results | Definitive proposition for the UI design lab |
| Reusalibility & Impact |  |
| In Jedi | It provides a good and reliable environment for prototyping and developing UI. |
| Out of Jedi | N/A yet |

# Work Packages Studies

As explained in the introduction, JEDI’s consortium decided to use this document-delivery as a single repository, including all technical studies executed within each work-packages. This last chapter reference all these other work-packages technical studies.

## Referenced Technical Studies

|  |  |
| --- | --- |
| Technical study number | B1 |
| **Technical study name** | **Stereoscopic video retargeting** |
| Work package | WP2 |
| Short Description | Video retargeting technology is an automatic solution to reframe/retarget a stereoscopic video content by including only a cropping window with most visually attractive area. |
| References |  |
| In Jedi | In document D2.1:  4.1.2.2 SALIENCY ESTIMATION  4.1.4.1 DYNAMIC REFRAMING  4.1.4.2 CORRECTION OF THE DEPTH DISTORTION  4.1.4.3 DYNAMIC ADAPTATION OF THE CONVERGENCE PLANE  4.1.4.4 ADAPTIVE DEPTHS OF FIELD PROCESSING |
| Out of Jedi | C. Chamaret, G. Boisson and C. Chevance, "Video retargeting for stereoscopic content under 3D viewing constraints", Proc. SPIE 8288, 82880H (2012) |

|  |  |
| --- | --- |
| Technical study number | B2a |
| **Technical study name** | **3D Content Distribution via Satellite** |
| Work package | WP2 |
| Short Description | Reviews different satellite broadcasting standards (DVB-S/S2) and provides the JEDI satellite transmission network design and performance results considering different scenarios (SCPC and MCPC with different carrier configurations). Final satellite transmission design for the demonstrators is also included. |
| References |  |
| In Jedi | In Document 2.1:  4.3.2.3.1 Satellite transmission standard description  4.3.2.3.1.1 DVB-S  4.3.2.3.1.2 DVB-S2  4.3.2.3.2 JEDI Project satellite transmission design |
| Out of Jedi | National and international conference publications and presentations about possibilities for second generation 3DTV over satellite in terms of network capabilities.  J. Rodríguez, A. Mourelle, I. Sanz, “Second generation of 3DTV. Are we ready?” IBC 2011 (Amsterdam).  J. Rodríguez, A. Mourelle, I. Sanz, “Transmission of second generation 3DTV: Challenges, risks and opportunities” NAB 2012 (Las Vegas). |

|  |  |
| --- | --- |
| Technical study number | B2b |
| **Technical study name** | **3D Content distribution in IP** |
| Work package | WP2 |
| Short Description | Reviews the state of the art of the Multimedia content distribution on IP networks, and the evolution to 3D content delivery. The trends of broadcast and unicast multimedia content delivery are described as well as the evolution of transport protocols to deliver 3D content. Finally is reviewed the content protection to guarantee the QoE. |
| References |  |
| In Jedi | In document D2.1:  4.3.2.2.1 Multimedia broadcast delivery on IP networks  4.3.2.2.2 Multimedia unicast delivery on IP networks: From VoD to Adaptive streaming  4.3.2.2.3 Multimedia transport technologies on IP networks: RTP and MVC extensions  4.3.2.2.4 Protection techniques for 3D video distribution on IP networks |
| Out of Jedi | P. Pérez, N. García, “Lightweight Multimedia Packet Prioritization Model for Unequal Error Protection”, IEEE Trans. Consumer Electronics, vol. 57, no. 1, pp. 132-138, Feb. 2011. (doi: 10.1109/TCE.2011.5735493)  P. Pérez, N. García, “Video Prioritization for Unequal Error Protection”, IEEE Int. Conf. on Consumer Electronics, ICCE 2011, Las Vegas (NV), USA, pp. 315-316, 9-12 Jan. 2011. (doi: 10.1109/ICCE.2011.5722602) |

|  |  |
| --- | --- |
| Technical study number | B3 |
| **Technical study name** | **Metadata in the JEDI workflow** |
| Work package | WP2 |
| Short Description | In function of the processing requirements and given the final storage or distribution use case, different metadata formats can be considered.    The main requirement is to produce easily interpretable sets of metadata in order to:   * improve some processing modules (especially the video encoder), * generate valuable 3D contents, * provide useful information for quality assessment.   This study specifies the different metadata to be exchanged between the processing modules from the acquisition to the transport layers. In addition, the exchange formats are considered and described by taking into account of the protocol and interface specifications. |
| References |  |
| In Jedi | In document D2.1:  7.1 Annex A: Metadata in the JEDI workflow |
| Out of Jedi |  |

|  |  |
| --- | --- |
| Technical study number | B4 |
| **Technical study name** | **Subjective assessment of Quality of Experience for video encoding with and without transmission errors** |
| Work package | WP2 |
| Purpose and objective | The Quality of Experience perceived by the end user depends on   * The encoding algorithm and its parameter * The bitrate * The status of the network, in particular delay and loss   In order to evaluate the impact of the JEDI transmission chain on the user’s experience, a subjective experiment is conducted to analyze which influence factors are important. |
| Description | Two subjective datasets are created.  The first subjective dataset assumes a lossless transmission. The degradations are therefore mostly due to coding and pre/postprocessing. In particular, the performance of H.264 and MVC is compared.  The second dataset focuses on the impact of transmission errors. Several different patterns and strengths are evaluated. and in particular the influence of error concealment strategies at the decoder side. |
| Technical constraints | * The presentation of uncompressed 3D stereoscopic video sequences in Full-HD resolution as necessary in subjective experiments requires a software player that handles the high datarate while guaranteeing a frame synchronized output. Such a player has been implemented by UdN * The Absolute Category Rating method using two scales (Quality of Experience and Visual Discomfort) has been selected from a multitude of possible protocols |
| Results | It has been shown that MVC performs better than H.264 and that error concealment in the decoder should be different when 3D content is transmitted as standard 2D error concealment methods may cause binocular rivalry. |
| Related results/studies |  |
| Reusalibility & Impact | The results were used for the decision to implement MVC and they are highly useful for the Set-Top Box decoder implementation concerning the error concealment algorithm |
| In Jedi | Detailed discussion in D2.1 |
| Out of Jedi | Barkowsky, M., Wang, K., Cousseau, R., Brunnström, K., Olsson, R., & Le Callet, P. (2010). Subjective Quality Assessment of Error Concealment Strategies for 3DTV in the presence of asymmetric Transmission Errors. *International Packet Video Workshop*  Brunnström, K., Sedano, I., Wang, K., Barkowsky, M., Kihl, M., Andrén, Börje, Le Callet, et al. (2012). 2D No-Reference Video Quality Model Development and 3D Video Transmission Quality. International Workshop on Video Processing and Quality Metrics *VPQM*, .  Wang, K., Barkowsky, M., Cousseau, R., Brunnström, K., Olsson, R., Le Callet, P., et al. (2011). Subjective evaluation of HDTV stereoscopic videos in IPTV scenarios using absolute category rating. *SPIE Electronic Imaging: Stereoscopic Displays and Applications*, *7863*.  Wang, K., Barkowsky, M., Brunnstrom, K., Sjostrom, M., Cousseau, R., & Le Callet, P. (2012). Perceived 3D TV Transmission Quality Assessment: Multi-Laboratory Results Using Absolute Category Rating on Quality of Experience Scale. *IEEE Transactions on Broadcasting*, *accepted for publication, early access available*. |

|  |  |
| --- | --- |
| Technical study number | B5 |
| **Technical study name** | **Influence of motion direction and velocity on visual discomfort in stereoscopic 3D applications** |
| Work package | WP2 |
| Purpose and objective | Visual discomfort is considered as one of the highest risks for the wide spread usage of 3DTV. It may stem from various factors such as camera setup, misalignment of the sequences etc.  Moving objects induce a particularly high amount of visual discomfort depending on their velocity, direction and speed. This is quantified in this study to help choosing the optimal parameters for shooting content in 3D. |
| Description | Several psychophysical studies are conducted using computer generated, synthetic stimuli which move in a repetitive and precisely defined manner.  As judging visual discomfort is difficult for observers on an absolute quality rating scale, a Paired Comparison experiment with sequential presentation using an interactive setup is used. |
| Technical constraints | The number of presentations in a Paired Comparison setup is too high for the number of stimuli that need to be compared, therefore a new method of performing a Subset-Paired-Comparison based on the work of Dykstra has been developed and proposed. |
| Results | The faster the stimulus moves, the more visual discomfort is generally introduced and the larger the distance between the background and the moving object, the more uncomfortable it becomes. This is generally true for both movement in depth as well as movement in a plane. However, a static stimulus has been found to induce more visual discomfort than a stimulus which slowly moves while staying at the same disparity. The study allows to quantify the difference between the conditions. |
| Related results/studies |  |
| Reusalibility & Impact | This study helps to choose optimal conditions for the camera capturing. In particular, it shows the maximum depth budget that can be used (foreground vs. background object distance) in function of the velocity. The trade-off between the added value of stereoscopic depth effect and the drawback of visual discomfort may be chosen. |
| In Jedi | Detailed discussion in D2.1 |
| Out of Jedi | Li, J., Barkowsky, M., & Le Callet, P. (2011). The Influence of Relative Disparity and Planar Motion Velocity on Visual Discomfort of Stereoscopic Videos. IEEE Workshop on Quality of Multimedia Experience *QoMEX*, .  Li, J., Barkowsky, M., Wang, J., & Le Callet, P. (2011). Study on Visual Discomfort Induced by Stimulus Movement at Fixed Depth on Stereoscopic Displays using Shutter Glasses. *IEEE Digital Signal Processing*, . |

|  |  |
| --- | --- |
| Technical study number | B6 |
| **Technical study name** | **Protection Techniques for Stereoscopic 3D Video Distribution over IP Networks** |
| Work package | WP2 |
| Purpose and objective | Due to the unreliable nature of the communication channel and the time-sensitive character of the considered scenario, packet losses take place during transmission, leading to video quality degradation. Considering that different parts of the encoded video stream are of unequal importance to the overall 3D perception due to error propagation, an Unequal Error Protection (UEP) scheme is introduced to mitigate the adverse effects of the channel and deliver stereoscopic video with the highest possible quality. |
| Description | The designed UEP strategy selects in real-time the most suitable packets within each GOP to be protected and the most convenient FEC technique parameters to carry out this protection. To this purpose, it takes decisions regarding: (i) the characteristics of the encoded video stream; (ii) the behaviour of the transmission channel; and (iii) the bitrate limitations. |
| Technical constraints | * Study and analysis of the relation between different characteristics of the packets in the stereoscopic video stream (view and frame they belong to, position they hold within the frame, etc) and the distortion introduced by their loss. From that task, development of a packet-level distortion model, through which the unequal relevance of the packets can be derived. * Packet loss probabilistic analysis in terms of the parameters that describe the behaviour of the channel. * Real-time performance and adaptation to both SbS and MVC |
| Results | The results of the conducted experiments show that the designed strategy outperforms, in terms of PSNR, the schemes that do not consider either the video features or the channel behaviour, or do it in a coarser fashion (e.g. at a frame level). Additionally, fetching the necessary data does not require any further process than that of parsing RTP and NALU headers. Thus, it is suitable to fulfil transmission delay constraints. |
| Related results/studies |  |
| Reusalibility & Impact | The robustness and straightforwardness of the algorithm makes it perfectly suitable to be included in 2D and 3D video streaming servers and STB’s. |
| In Jedi | Detailed discussion in D2.1 |
| Out of Jedi | C. Díaz, J. Cabrera, F. Jaureguizar, N. García, “A Video-Aware FEC-Based Unequal Loss Protection System for Video Streaming over RTP”, *IEEE Trans. Consumer Electronics*, vol. 57, no. 2, pp. 523-531, May 2011 |

|  |  |
| --- | --- |
| Technical study number | B7 |
| **Technical study name** | **Subjective evaluation of the impact on the user quality of experience of transmission errors in 3DTV broadcasted services** |
| Work package | WP2 |
| Purpose and objective | The main objective is the evaluation of the effects of typical transmission impairments that may affect the Quality of Experience (QoE) of the end users of broadcasted 3DTV services.  The subjective assessment of these effects will allow:   * A better understanding of perceptual issues regarding 3D video visualization. * Development of tools for quality monitoring in the broadcasting network, especially focused on estimating the QoE of the end users. |
| Description | A currently realistic scenario for 3DTV was considered in the experiment, so the transmission of side-by-Side (SbS) content over packet based networks was simulated. Then, typical transmission errors were introduced in the sequences, specifically, different patterns of video and audio packet losses, video freeze, outage events and effects of drops of the QoS of the network (framerate and bitrate reductions).  Since the same infrastructure than for traditional video broadcasting is currently used, monoscopic versions of the 3D videos were also used in the test to compare the effects of the impairments in conventional TV and 3DTV services.  broadcastingScenario.jpg |
| Technical constraints | * A novel subjective evaluation methodology was designed to obtain meaningful and representative results of how the QoE of the end users could be affected by transmission errors. Therefore, the methodology is focused on keeping as far as possible home viewing conditions. Thus, long sequences similar to those usually watched by people at their homes are used. In addition, the methodology is based on single stimulus standard techniques and a nearly continuous evaluation is collected from the observers. |
| Results | The main results obtained in the experiments were:   * Worse effects of video losses in SbS 3DTV: Each stereo view could be affected differently by the losses, and showing very different content in corresponding regions of each view, and then causing binocular rivalry and visual discomfort. * Audio and video outages: Most annoying distortion * QoS reductions: framerate drops worse than bitrate drops * Audio losses could be more annoying than video losses * Visual discomfort: More than 50% of the observers felt discomfort, and 15% of the observers felt headache or slight dizziness. * The 3D version of the videos is only preferred when they are correctly produced not to cause visual discomfort and provide a good depth sensation. |
| Related results/studies |  |
| Reusalibility & Impact | This study helps to obtain a better understanding of the effects of transmission errors in 3DTV, which is a first step in the development of objective metrics for quality monitoring in broadcasting networks. In addition, the proposed assessment methodology has been validated to obtain representative results of what end users perceive at their homes. |
| In Jedi | Detailed discussion in D2.1 |
| Out of Jedi | J. Gutiérrez, P. Pérez, F. Jaureguizar, J. Cabrera, N. García, “Subjective assessment of the impact of transmission errors in 3DTV compared to HDTV”, IEEE 3DTV Conf., 3DTV-CON 2011, Antalya, Turkey, pp. 1-4, 16-18 May 2011. (doi: 10.1109/3DTV.2011.5877209)  J. Gutiérrez, P. Pérez, F. Jaureguizar, J. Cabrera, N. García, “Subjective Evaluation of Transmission Errors in IPTV and 3DTV”, Visual Communication and Image Processing, VCIP 2011, Tainan, Taiwan, pp. 1-4, 6-9 Nov. 2011. (doi: 10.1109/VCIP.2011.6115975)  P. Pérez, J. Gutiérrez, J.J. Ruiz, N. García, “Qualitative Monitoring of Video Quality of Experience”, IEEE International Symposium on Multimedia, ISM 2011, Dana Point (CA), USA, pp. 470-475, 5-7 Dec. 2011. (doi: 10.1109/ISM.2011.83)  J. Gutiérrez, P. Pérez, F. Jaureguizar, J. Cabrera, N. García, “Validation of a novel approach to subjective quality evaluation of conventional and 3D broadcasted video services'', Fourth Int. Workshop on Quality of Multimedia Experience, QoMEX 2012, Yarra Valley, Australia, 5-7 Jul. 2012 (Accepted). |

|  |  |
| --- | --- |
| Technical study number | B8 |
| **Technical study name** | **3D sequences descriptions** |
| Work package | WP3 |
| Purpose and objective | The goal is to show the limitation of Side by Side transmissions with appropriate 3D content and emphasize the advantage of FullHD 3D Television. |
| Description | The scene 1 shows interesting fine details in front of the screen  The scene 2 shows a strongly colored object with motion moving out of the screen towards the observer.  The scene 3 shows a lot of people, movement and 2D depth cues.  The scene 4 is a scene with very low ambient luminosity and high contrast  The scene 5 is a scene with high brightness contrast, object and camera motion  The scene 6 is a scene with animals with random movement and water dispersion effects  The scene 7 is a random movement (2d and depth) and slow motion  The scene 8 is a colorful scene with fine details  The scene 9 is a colorful scene with big pop out effect  The scene 10 is a fast and constrained movement with high depth effect.  The scene 11 is a high horizontal frequency, small but important depth variations with movement |
| Technical constraints | The stream must be of a high quality, e.g. resolution of Full-HD or higher, uncompressed capturing. The length of the sequences must be longer than 16 seconds so that they can be cut to 16 seconds. Thefilmed objects should not exceed a distance of 0.2 diopters with respect to the central plane, that is, the display position.  We must avoid 3D window violations. In 3D visualization, objects with a crossed disparity (in front of the screen) should not cross the border of the screen in order to avoid visual discomfort. This is particularly true for objects that are outside the comfortable viewing zone and thus are considered as pop-out effect.  In order to get a good 3D effect, objects of interest shouldn’t be too far from the camera. |
| In Jedi | In document D3.5, Chapter 3 “Specification of the user experience lab” |
| Out of Jedi | Share of the specificartion has been done wioth EBU |

|  |  |
| --- | --- |
| Technical study number | B9 |
| **Technical study name** | **Connectivity techniques** |
| Work package | WP3 |
| Short Description | This study is applying on both the HDMI and DisplayPort techniques to transmit 3D video streams. |
| Purpose and objective | The objective is to appluy the 3d constraints to the connectivity, meaning he HDMI transmitter and receiver. Also, in this study was analysing the DisplayPort capability to transmit the high volume of data. |
| Description | The 3D structure of the video stream is indicated in the CEA-861 Vendor Specific InfoFrame containing a 24-bit IEEE Registration Identifier of 0x000C03, a value belonging to HDMI Licensing, LLC. (See HDMI 1.4a specification for content of this InfoFrame).  The transmission of this InfoFrame is optional for the source device. But if a source device outputs a video signal which is defined below, the source shall transmit this packet. Whenever this packet is transmitted, an accurate HDMI Vendor Specific InfoFrame shall be transmitted at least once per two Video Fields.  It is optional for a Sink to interpret this packet. The data in the AVI InfoFrame packet remains valid even if the HDMI Vendor Specific InfoFrame is transmitted. |
| Technical constraints | The 3D video format is indicated using the VIC (Video Identification Code) in the AVI InfoFrame (indicating the video format of one of the 2D pictures, as defined in CEA-861-D) in conjunction with the 3D\_Structure field in the HDMI Vendor Specific InfoFrame (indicating the 3D structure).  Frame packing is one of the HDMI 3D video format structures indicated by the 3D\_Structure field and is composed of two stereoscopic pictures: Left and Right..    Figure ‑3D structure (Frame packing)  In this figure, the area inserted between the two Active video regions is designated as “Active space”. This Active space area shall be encoded in the same manner as the adjoining Active video regions. During the Active space, an HDMI Source shall transmit a constant pixel value. HDMI Sinks shall ignore all data received during the Active space regardless of the value.  Frame packing can also be applied for interlaced video timing formats  Figure **Erreur ! Source du renvoi introuvable.**‑1 3D video format (Frame packing for interlaced format)  Side-by-Side (Half) is one of the HDMI 3D video format structures indicated by the 3D\_Structure field and is composed of two stereoscopic pictures: Left and Right, which are sub-sampled to half resolution on the horizontal axis.    Figure - 3D structure (Side-by-Side (Half))  Top-and-Bottom is one of the HDMI 3D video format structures indicated by the 3D\_Structure field and is composed of two stereoscopic pictures: Left and Right, which are sub-sampled to half resolution on the vertical axis. |
| Technical constraints | If an HDMI Source has the 3D video format capability, then the HDMI Source shall support transmission for at least one of the formats listed below.  If an HDMI Sink supports 3D video, then the HDMI Sink shall support reception of 3D video formats per the following requirements:  An HDMI Sink which supports at least one 59.94 / 60Hz 2D video format shall support all of   * 1920x1080p @ 23.98 / 24Hz Frame packing (148.5 Mpix/s) * 1280x720p @ 59.94 / 60Hz Frame packing (148.5 Mpix/s) * 1920x1080i @ 59.94 / 60Hz Side-by-Side (Half) (74.25 Mpix/s) * 1920x1080p @ 23.98 / 24Hz Top-and-Bottom (74.25 Mpix/s) * 1280x720p @ 59.94 / 60Hz Top-and-Bottom (74.25 Mpix/s)   An HDMI Sink which supports at least one 50Hz 2D video format shall support all of   * 1920x1080p @ 23.98 / 24Hz Frame packing (148.5 Mpix/s) * 1280x720p @ 50Hz Frame packing (148.5 Mpix/s) * 1920x1080i @ 50Hz Side-by-Side (Half) (74.25 Mpix/s) * 1920x1080p @ 23.98 / 24Hz Top-and-Bottom (74.25 Mpix/s) * 1280x720p @ 50Hz Top-and-Bottom (74.25 Mpix/s)   With 10.8Gbps over 4 lanes, DisplayPort provides sufficient bandwidth for transporting up to 1080p (FHD): 3D Stereo video data at 120Hz (that is, 60Hz each for left and right frames). At 21.6Gbps over 4 lanes, the bandwidth is sufficient for 1080p 3D Stereo video at 240Hz (that is, 120Hz each for left and right frames).  In addition, the DisplayPort standard provides for two in-band mechanisms through which a source device can specify the attribute of the 3D stereo video format it is transmitting. One method uses an MSA MISC1 field and the other method uses a Secondary-Data Packet called VSC Packet.  A Sink device with DPCD Revision 1.2 or higher must support both methods.  The 3D stereo capability can be exposed in EDID and DisplayID. A 3D stereo format is usually associated with specific timing and hence it is desirable to indicate which timings support 3D stereo format and which don’t. Furthermore, for a given timing that supports 3D stereo format it is required to indicate which stereo format is supported. Both EDID and DisplayID have the ability to expose 3D stereo capability per timing, but DisplayID provides for a more efficient and flexible format declaration. In DisplayID, the timing option field of Type I and Type II detailed timing descriptor exposes the capability indicating whether the timing is displayed with no stereo or with stereo or dynamically configured based on the content that is being shown. The “Stacked Top & Bottom” is not defined in DisplayID. |
| In Jedi | In document D3.1, Chapter 3.2 “3D rendering techniques” |
| Out of Jedi | This proposals has been discussed and considered when defining the format HDMI1.4a in the HDMI forum. |

|  |  |
| --- | --- |
| Technical study number | B10 |
| **Technical study name** | **Audio rendering techniques** |
| Work package | WP3 |
| Short Description | The goal is to develop new generations of analysis and processing technologies for multichannel and 3D sound. |
| Purpose and objective | A series of meetings with the innovation departments of several French broadcasters, including TF1 and France Television, allowed identifying several audio innovations that are relevant in the context of 3D TV. The main conclusions can be summarized as follows:   * 3D video provides depth information within the screen viewing angle * But:   + Depth information in sound is already provided by stereo.   + Multichannel surround already provides sound sources distribution over 360°   + 3D sound provides elevation information (sound sources over the head)   + Human perception is very poor in detecting sound elevation, only dramatic effects are effective   + It is not convenient to place loudspeaker at the ceiling in living rooms.   + The benefit of 5.1 for TV is still discussed by broadcasters. We can’t expect they will adopt 3D sound in the short future. It is more probable that 3D sound will be introduced with cinema or games contents.   It seems that 3D sound is not required for 3D TV, which may not be immediately obvious nor intuitive.  These meetings allowed collecting from broadcasters their highest priority requirements:   * Requirement 1: Increasing quality of the audio signals really delivered to the end user.   + During production with more efficient analysis tools   + During reproduction with supervision tools. * Requirement 2: Adapt the sound to the environment.   + Production environment   + End user living room * Requirement 3: Enlarge the listening area for optimal reproduction (for all the family). Successive audio innovations always reduced the optimal listening area:  |  |  | | --- | --- | | **Format** | **Optimal listening area** | | Mono | Volume | | Stereo | Plane | | 5.1 | Line | | 3D | point |  * Requirement 4: Enable interactive sound. |
| Description | New audio supervision tools are developed to fulfil the audio requirement 1. The following diagram illustrates the concept of audio supervision. The main idea is to detect modifications in the signal provided to the end user by set top boxes.    Today’s methods are using absolute detection methods by only analysing the output of the set top box. However, this approach is unable to detect relevant signal modifications, since such modifications can only be detected when comparing the signal delivered by the set top box to the signal provided by the broadcast company. As a result, new comparative supervision methods are developed within JEDI to detect relevant signal modifications such as: Mute, Signal different, Level, Filtering, Added signal (noise, incorrect signal) and Delay. Such new methods are based on signal processing methods such as Spectrum comparison, Correlation, Inverse filtering and Error estimation. |
| Results | New sound image remapping methods are developed to fulfil the audio requirement 2 and 3. Multichannel audio content is recorded, mixed and reproduced assuming a particular loudspeaker layout described in the *ITU R BS 775-1* and *AESTD1001* standards.    Figure xxx:  However, such standard are extremely difficult to implement in homes where more flexible loudspeaker placement is required. The goal of the new remapping method is to reproduce a correct sound image over an incorrect loudspeaker layout.  New sound image remapping methods are developed to fulfil the audio requirement 2 and 3. Multichannel audio content is recorded, mixed and reproduced assuming a particular loudspeaker layout described in the *ITU R BS 775-1* and *AESTD1001* standards.    However, such standard are extremely difficult to implement in homes where more flexible loudspeaker placement is required. The goal of the new remapping method is to reproduce a correct sound image over an incorrect loudspeaker layout.   |  |  |  | | --- | --- | --- | | Original sound played on an ideal loudspeaker layout | Original sound played on an incorrect 3D loudspeaker layout. Sound sources are reproduced at incorrect positions. | Sound played on an incorrect 3D loudspeaker layout with image remapping. Sound sources are reproduced at correct positions. | |  |  |  |   JEDI investigate a new high resolution image remapping method based on advanced models of human perception:   * Time domain perception and the relationship with high resolution time-frequency methods. * Spatial domain perception with a new model for localisation over loudspeakers   It is known that time-only representation or frequency-only representation is not appropriate to study to study the perceived quality of sound perception. Because of human perception, it is appropriate in audio to have an excellent time resolution in high frequencies (fractions of ms) and an excellent frequency resolution in low frequencies (a few Hz). Traditionally in audio, time-frequency analysis is obtained with the Short Term Fourier Transform (STFT). However, this time-frequency representation is still not appropriate for the analysis of audio and acoustic applications as it is based on the Fourier Transform and provides constant frequency resolution and constant time resolution. For instance, a STFT based on 20ms blocks provides a constant 20ms time (far not enough for high frequencies) resolution and a constant 50Hz frequency resolution (far not enough for low frequencies).  As a result TRI studied the Wavelet transform based on a logarithmic frequency scale and providing both the required time resolution and the required frequency resolution. A signal x(t) to be analyzed is decomposed on a family of basis functions {u,s(t)}u,s deduced from a kernel function (t) by an affine transform: u,s(t) = ((t-u)/s). Therefore, the signal x(t) is compared to the function (t) that is moved in time and scaled (contraction or dilatation of ratio *s*). The wavelet transform is given by the following expression :    Compared to STFT, the Wavelet transform allow a better mapping of the time-frequency domain by locally adapting both the time and the frequency resolution, as shown in the following figure.    STFT resolution mapping Wavelet time-frequency analysis (lin and log scale)  Concerning the sound localisation over loudspeakers, extensive effort has been invested to study the prior art. Such information has been collected from various areas of audio technologies such as sound field synthesis, microphone recording techniques, panning laws, spatial hearing mechanisms. The goal is to combine this information in a model. This model predicts the perceived direction and the focus of the reproduced sound sources as a function of:   * the speaker layout, * the speaker signals (ΔI, ΔT), * the listening position (essential for large theatres) * and the head orientation. |
| In Jedi | In document D3.1, Chapter 4.2 “3D rendering techniques” |
| Out of Jedi | This techniques has been implemented in the audio IP of Trinnov and now deplyed in theatres. |