**Abstract**

Communications and in particular video communications have ever arouse interest of people. In last years the 3D technology is gain grown after a period in which excessive costs and difficulties in the implementation have broken its development. Now, 3D is a reality in cinemas and ever more films or live events are shooted or created thinking to present them in this way.

Thanks to its success, it is predictable to think that 3D could be gladly received also in a home environment, changing the way to watch TV by people. So, exploiting the cinema experience, tecniques to efficently shoot, codify, transmit and display 3D contents are under study and improvement.

The positive impact that 3D is having today, could pave the way to a future introduction of Multi-View Video (MVV) technology. MVV, in the begin thinked as an extension of the 3D with more than two views, surely can open a very wide field of possible immersive applications.

At the moment there are many technical obstacles for the implementation of such a technology. Particular equipments and devices are needed, cost for end users is too high, so for the moment MVV is yet far. But surely, if realized, MVV can change the whole way to live videocommunications.

# Prefazione

Le comunicazioni e in particolare le video-comunicazioni hanno da sempre attirato l'interesse delle persone. Negli ultimi anni, la tecnologia 3D sta prendendo piede dopo un periodo in cui i costi eccessivi e le difficoltá di implementazione ne hanno frenato lo sviluppo.
Oggi, il 3D é una realtá, sopratutto nei cinema e sempre un maggior numero di film ed eventi dal vivo, concerti per esempio, vengono ripresi con l'intenzione di presentarli al pubblico in questo formato.
Grazie al suo successo, é prevedibile pensare che il 3D possa essere accolto con piacere anche in ambiente domestico, cambiando cosí il modo di guardare la TV da parte delle persone. Cosí, sfruttando l'esperienza acquisita al cinema, si stanno intensificando studi e ricerche per migliorare le tecniche di cattura, la fase di codifica, la trasmissione e la visualizzazione dei contenuti in 3D. L'impatto positivo che il 3D sta avendo oggi, potrebbe servire per "aprire la strada" all'introduzione di un completo sistema Multi-view. Il Multi-view video (MVV), all'inizio pensato come una naturale estensione del 3D in grado di presentare piú di due viste, sicuramente potrá aprire un'enorme campo di possibili applicazioni.
Al momento l'arretratezza della tecnologia rappresenta il più grande ostacolo per una possibile implementazione di un tale sistema. Sono necessari particolari dispositivi sia per la ripresa che per la visualizzazione e al momento il costo per l'utente sarebbe insostenibile. Perció il Multi-view é ancora lontano dall'essere realizzato. Ma sicuramente, una volta che sará possibile farlo, l'MVV potrá cambiare l'intero modo di vivere le comunicazioni.

# Contents

Chapter 1

# Introduction

This thesis is the yield of the work performed at Opensky, working in 3d@Sat project of ESA in which Opensky partecipates in one's capacity as commercial operator, thanks to its acquired expertise providing services like satellite internet access, push services, streaming channels and bi-directional satellite IP services.

Opensky is working in this project in colaboration with two partners: Fraunhofer HHI, a company with head office in Berlin, leader research institute in image processing and the french EADS Astrium, with a significant expertise in end-to-end satellite telecom system engineering.

The paper gathers and extends a part of the project, integrating its with some deepening regarding examples of aplication and theoretical studies taken from several articles. The aim of the 3d@Sat project is to perform a review on existing thechnologies in 3D field and to propose an efficient solution, satisfying several technial and commercial requirements.

The idea is to start from the 3D experience to think, elaborate and present a possible architecture for the **Multi-View Video(MVV)** system. With MVV it intends the wide category that comprises the extension of the 3D on flat screen(more than two views presented simultaneously) and all possible applications for the three-dimensional recreation of a solid object. The first is not so far to be realized, instead for the second we are in full phase of research. MVV on flat screen is also known as **Free TV(FTV)**. It means that many view of a scene are presented at the same instant time on the screen. So, the motion of a user in front of the screen has the effect to give its a different viewpoint of the video. The idea of a three-dimensional recreation of an object instead, is thinked to give to a user the sensation that this object is really present in the local site. For the moment the thought is only for static objects but in future this idea will be certainly extended to motion and whole scenes.

Opensky is actively parteciapting in this work, providing its experience above all in the commercial field, but not only. Recent installation of the new "Alps' Teleport" in Vicenza, consents its to hold a relevant place in the broadcasting of contents, both 2D and 3D. So, thanks to its relevant experience and frequent contacts with 3D cinemas, Opensky contribution is in all steps of the 3d@Sat project.

Refering to the whole project, included the work performed by partners, in this thesis are presented also results obtained not completely from Opensky, but in which it has given its important contribution. The 3d@Sat project isn't entirely presented here, but only a part, mainly focusing in the transport system and in an introduction on the commercial analysis performed in Opensky. In Gabriele Longhi's thesis, the remaining part of the project is tracted.

## 1.1 Structure of the work

In this Section the whole work is presented, Chapter by Chapter, highlighting aims and requirements.

The second Chapter of the thesis treats a presentation of the work's background, that is an introduction about the 3D, its development and evolution in Multi-View Video(MVV) and the ESA project which involves Opensky, Fraunhofer HHI and EADS Astrium. For this first part of the work the main information's source is the Book [1].

The next Chapter treats the transmission system, analysing the coverage area in different satellite scenarios and proposing a scheme for the data stream. How we can see, several factors influences system's performances and it is interesting to observe how the transmission system changes varying these parameters. Here the results about the different scenarios are provided by Astrium.

Then, considering the transmission chain, the coding phase holds a fundamental rule for the final results. So, the **Multi-view Video Coding(MVC)** tecnique is presented, explaining its features and what are the improvements as regards previously ones. Necessary informations for this Chapter are taken from several articles, in particular from the Technical Notes of the project([2]) and from [1], [3].

To verify the performances of the presented data stream, the simulation of a real transmission is fundamental. So, thanks to the Satellite emulator(**SATEM Tool**) provided from Astrium, a real transmission has been simulated, and presented in Chapter 5, obtaining important results in terms of service availability.

Then, the last part of the work is dedicated to the commercial analysis, entirely performed in Opensky. Being a work still in progress, here is only introduced, but with important considerations about expected costs, impacts and benefits with the introduction of such a system.

Chapter 2

# Overview on 3D, MVV and 3d@Sat project

The possibility to recreate a third dimension on the screen how happen in real life, bewitches people all along. The 3D vision on a screen is a good intuition and a big innovation, but ambitions are greater. Can represent a shape, an object, a person or a whole moving scene, giving to an observer the perception to be involved, would be wonderful. Thinking the video communication in this perspective, the 3D is an obliged stage. Many studies are moving in this direction, but today the main obstacle to handle is the absence of adjusted representation technologies. However, considering the evolution from the birth of the stereoscopy up to today, the technology certainly will improve, making it feasible in the future.
Now for example, the first 3D TV channels are rising and it was unthinkable since thirthy years ago. So, an overview on evolution of 3D technology, on its impact on the market, on today applications and on future perspectives for this innovation is important to understand how 3D can develop.
Among possible evolutions of 3D there is the Multi-View Video(MVV), a technology that extend the idea of stereoscopic vision, giving more viewpoints in three-dimensions.
Afterwards, the 3d@Sat project is presented. It is the ESA project in which I have partecipated and that is the basis of this work. It contains a wide review on existing technologies about 3D and MVV and a deep study from the technical and commercial point of view about possible scenarios and evolutions.

## 2.1 Introduction to 3D

3D Video enables new immersive applications such as 3DTV, 3D movies, live 3D events and new forms of video communications. Recent advances especially in 3D display technology, lead to a starting standardization phase in this sector that is necessary to a correct evolution of the 3D. No many stereoscopic contents are nowadays available, but the success

that first 3D films are obtaining drives many cinema productors to interest themself on this new technology. In fact several Hollywood studios announced major productions in 3D. These developements have surely an influence on the home market, because secondary distribution (DVD, Blu-ray) of movies has become an important market segment.
At the moment, it exists for example the possibility to see a film in three dimensions and many cinemas are adopting this new projection tecnique. The offered depth perception, the main feature of 3D, is obtained deceiving uman's brain. It joins the images coming from the two eyes at the same instant time and recognizing these like a single image.
The prefigured market for 3D isn't only cinema, but also home environment. Today, efforts are directed mainly on cinema market, but under this surface, the research activity to improve displays technology, shooting tecniques, representation tecniqes, transmission formats and coding tecniques is very frenetic.
A reason for which the extension of this technology at home environment could be braken is the necessity to wear particular glasses to watch 3D videos on stereoscopic display. This is surely an encumbrance and many users are not ready to accept it.
But, with a futuristic view, we can consider that this could be only a passage phase, because the intention is to move towards other more complex technologies which are able to recreate a more realistic perception of an object or a scene. The scope of 3d@Sat project in fact is to pave the way to a future end-to-end multiview chain. Looking the improvements from the born of 3D images idea in so little years, the perspective for the future is very good.

### 2.1.1 Birth of the stereoscopy

Glancing at the past, the evolution of 3D from the initial use of the stereoscopy tecnique up to today has been great.
The first utilization of the stereoscopy dates back 1838, thanks to sir Charles Wheatstone (the invention is of the 1832), when borned the idea that two slightly different images can recreate the illusion of the third dimension. But already in the intiquity some scientists like Euclide and Leonardo da Vinci were interested on this phenomenon. However, this invenction didn't gain big success. In fact Wheatstone's instrument was very cumbersome and so little practical. It has to wait until 1949 when sir David Brewster invented the **kaleidoscope**, giving the first impact on society of the stereoscopy.
At the beginning of 1900, with the advent of the 35mm color film, corresponds a fall of interest about the stereoscope that used cards and slides. But the consequent invenction, 30 years after, of a viewer adapted to 35mm roll of films allows to relaunch this device. These, are also the years in which a few films are shooted with short budget, trying to discover the secrets of the stereoscopic production.
The first stereoscopic film, "The power of love" appeared in 1922. But we have to wait for the middle of twentieth century to observe the first boom of 3D, even if many problems rose for the high costs of reproduction equipment: silver screens, polarized glasses and double

synchronized projectors. In those years was shooted more than sixty 3D films, so to call this, the "golden age" of 3D, but for the shortage of theatres and cinemas technology there was a sticking point.
In these years there was also the first official stereoscopic transmission, using the **Anaglyph** method in black and white. "Anaglyph" tecnique consists in the simultaneously presentation of the images for the left and the right eyes, but differently filtered. With the birth of the color television, this principle was adopted filtering the two images with two complementary colors: red and cyan. Filtered glasses was required.
Forgotten for a certain stretch of time, around 1975 the 3D cinema restarted to have again popularity. It is of this year the shoot of the stereoscopic film "Jaws" of Steven Spealberg. After, in 1986 was introduced the **IMAX-3D** tecnique which uses anaglyph, polarized lenses and LCD glasses, with a single or two synchronized lenses. The IMAX system is today still active in many cinemas.
Then, from 2001 started the second "golden age" of 3D thanks in particular to the advent of computer animation technology, digital cameras and 3D home theatres.

### 2.1.2 Depth perception and stereo reproduction

Understanding the depth perception ability of **Human Visual System(HVS)** in the 3D world is the basis to try to recreate this situation on a screen. It is important also for the shooting phase, in which the stereoscopic capture has to emulate the human vision.

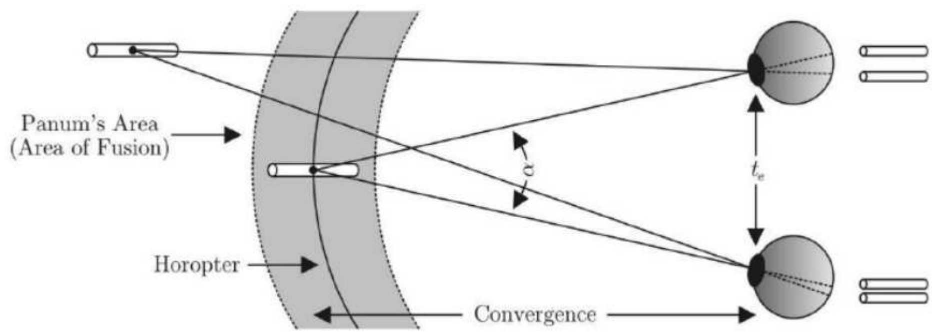


**Figure 2.1.** Principle of stereoscopic fusion and retinal disparity.

**Stereoscopic vision** The percepted depth cues can be divided in two categories: cues that require only one eye, and cues that instead need both eyes. For this, a fundamental rule is covered by the distance between the observer and the objects in the scene. In fact, if this distance is below 10 meters, binocular cues are predominant.
In human's vision, each eye captures an unique image. Images coming from both eyes are similar but slightly different, due to eyes horizzontal separation of roughly 64mm. This

distance is known as interocular distance. It leads to spacial distances between the relative projections of observed 3D points in the scene onto the left and the right retina, also known as retinal disparities. The retinal disparity consents to the human's brain to fuse the two different perspective views from the left and the right eye into a single, three-dimensional image.
The Figure (2.1) shows how this process of stereoscopic fusion works in detail. Looking at the real world, eyes rotate until their optical axes converge on the interesting object. Then, the point of convergence is projected to same corresponding image points on the respective retinas, not producing retinal disparity. The same happens for all objects placed in the **Panum's Area**, a circular band around the **Horopter**, defined by the convergence point and the nodal points of both eyes. All other points, will produce retinal disparities, whose magnitude becomes the larger, the further away 3D points are from the Horopter. Further objects have positive parallax, instead next objects have negative parallax.

**Reproduction on screen** Most of stereoscopic displays and projectors are based on the same basic principle of providing the viewer with two different perspective images for the left and the right eyes. It is illustrated in Figure (2.2).

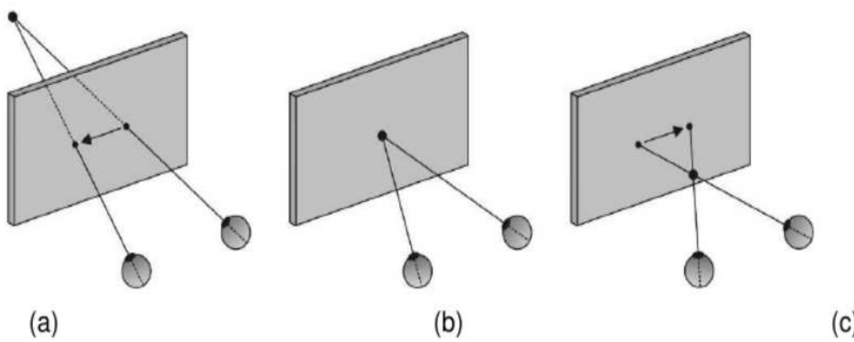


**Figure 2.2.** Different types of parallax: (a) Positive parallax; (b) Zero parallax; (c) Negative parallax.

Stereoscopic reproduction tecniques exploit this merging ability of the uman's brain and subdivide themself in three categories: anaglyphic, system at alternating darkening and polarizzated system. The first has been previously tracted and it is an old tecnique, not used today expecially for its limitation in quality images. Having to filter with a red or a cyan filter the images, the resultant quality is too low.
Here are presented the other two tecniques. They are slightly different for the implemented methods, but both used today in cinemas.

**Alternating darkening systems.** Alternating darkening is a tecnique that use a simultaneous presentation of the two images: one directed to the right eye and one directed to the left one. The presentation only seems simultaneous, instead the images are projected at 48 frames per second, the double of a 2D projector. A particular screen is not required and the image quality is not reduced respect to 2D. It is sufficient to wear active glasses which darken alternately the two lenses. Thinking about a cinema, a problem could be the simultaneous darkening of the same lens of all glasses in the cinema room. To make this possible, a device for the synchronization is usually placed next or up to the screen. For this reason this glasses are knowed as "active". Systems that use this method with LCD glasses are XpanD, nuVision and Dolby.

**Polarizzed systems.** This tecnique exploits the concept of polarized image. The projectors beam images with orthogonal polarization and they are overlapped on a "silver screen" able to keep the polarization and to compensate the light loss. Filtering glasses are required, in which filters must have the same polarization of the desired incoming image. For example if the left image has vertical polarization and right one instead horizontal, the left lens must have vertical polarization and right lens the horizontal one.

Not requiring synchronization, these glasses are knowed as "passive".

Polarizations coul be linear (vertical and horizontal) or circular (left and right).

The problem with vertical polarization is that inclining the head also the filters skew, compromising the correct view. In this case a perfect straight posing is required.

Using circular polarization instead, head position not influences the view, so it is the tecnique more used. **RealD** for example uses a singular projector with two circular filters activated alternately by a computer. The frame rate is the double of a 2D projector so the light is filtered at the double speed.

## 2.2 Introduction to MVV

With the term Multiview it intends the presentation of more than two views in a certain way at the same time. The basis concept is to allow a free view of a scene like happen in real life, that is not only the fixed depth perception of 3D stereo case, but a real immersion sensation. Thinking to move and have a different view of the scene in relation on the movement seems only a futuristic speech, instead this purpose is not so far to be realized. This application is also known as "Free TV" (FTV). It gives the possibility to represent in a display or a screen more views of a scene, film or game. The basis idea is that a user who decides to see for example a film, moving himself has the possibility to see the same scene with a different point of view variable with his motion. So, a film can be watched from many different positions, with many different details dependents from the watching location, being still or moving himself. This concept obviously, can be extended to more that one user.

Like in stereo 3D, in Multiview scenario the basis idea is to provide an image for the left eye and an image for the right one, leaving to the human's brain the duty of fuse them and gives the depth perception.
Methods used to produce autostereoscopic displays, able to reproduce many views, allow to watch Multiview scenes without glasses. In this sense it provides a big improvement respect its predecessor, but a problem can occur joining its with the movement. Without glasses in fact, is impossible to determinate what is the image directed to a certain eye, and moving himself a user could invert left and right images, obtaining probably sickness sensation.
To correct this problem some solutions are under study. A device that is able to recognize user position tracing his head motion is in a developement phase. So, on the basis of his position, this device allows to direct correctly lefth and right views. As well as correct views directions, it could allow to display only two views at a time on the basis of user's position, reducing representation algorithms complexity.
Really, MVV is a very wide category, because it comprises not only applications on flat screen but also all technologies in which a certain device is able to reproduce a solid reconstruction of an object. So, multiprojector systems, that are now under study, are an example of futuristic MVV application. In this case the object is reproduced, usually in a box, from a system that uses several lasers. A user has the possibility to turn around this "ologram" and watch its from all possible perspectives.
In MVV category we can individuate two main groups, differentiating themself for the way in which the Multiview structure of the object is determinated and elaborated.

### 2.2.1 Image-based modelling

Image-based representations don't use any 3D geometry. Virtual intermediate views are generated through interpolation of natural camera views.
The main advantage of this type of representation is a potential high quality of virtual view synthesis without any 3D reconstruction, but to achieve this, a large number of natural camera view images is required. If a sparse camera setting is used, interpolation and occlusion artefacts will appear in the synthesized images, affecting quality. Otherwise large number of camera implies a great amount of image data to be processed.
Examples of image-based representations are light-field representations, where an intuitive description of the view-dependent appearance of the scene is offered by the parameterized light field. They have to cope with an extremely complexity of data acquisition or they have to execute simplifications but reducing the interactivity.

### 2.2.2 Geometry-based modelling

Geometry-based representations have typical usage in 3D computer graphics, in the most cases, such as surface-based representations, they are based on meshes of polygons. Scene

is built segmenting the image into a collection of surface patches, whose position, orientation and shape in 3D space is estimate. Scene geometry is reconstructed by a set of meshes which reproduces the real object using 3D geometry surfaces, without any redundancy. Typically a mesh is a set of triangles.
A texture map is associated to these surfaces, and also other attributes as appearance properties(opacity, reflectance, specular lights, etc.) may be assigned to enhance the realism of the model.
Instead a method that doesn't use meshes, is point-based representation, in which points are samples of the surface, and describe surface's 3D geometry and surface reflectance properties. Point-based methods specify implicitly connection information through the interrelation among points. The drawback is that each point requires more geometric information than image pixels.
Geometric-based approaches are commonly used in applications such as computer games, internet, TV, movies. Especially if scenes are computer generated, the achievement performances are excellent. The available technologies for production and rendering is highly optimized and the current PC graphics cards are able to render highly complex scene with excellent quality as regards reproduction of motion, accuracy of the texture, spatial resolution, levels of detail and refresh rate.
The main drawback is that content creation requires high costs and human assistance. The algorithms used are extremely complex, and then computation of 3D scene models is often limited only to foreground objects.

### 2.2.3 Free Viewpoint Video (FVV)

Free viewpoint video offers the same functionality that is known from 3D computer graphics. The user can choose a viewpoint and a viewing direction within a visual scene, meaning interactive free navigation.
A number of real cameras captures a real world scene. Within pratical limits the user may freely navigate through the scene. In contrast to pure computer graphics applications, FVV targets in real world scenes are captured by natural view cameras.

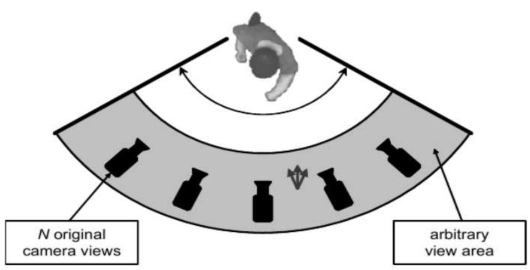


**Figure 2.3.** Free viewpoint video and interactive selection of virtual viewpoint.

In Figure (2.4) is illustrated the famous "stop-motion" effect(known from the movie "The Matrix"). It is possible to freeze the FVV scene in time in analogy to a freeze image of classical 2D video. Then it is still possible to navigate around it(in spatial dimension) and to show its from different viewpoints at the same time. In this approach any virtual camera path can be produced afterwards.
Conventionally this effect is produced by placing a suitable number of synchronized cameras exactly along the line desired navigation. The virtual camera flight is then created by displaying the original camera images consecutively, for example by means of a pure switching from camera to camera without any virtual view generation. This approach requires accurate planning and a tremendous effort for the acquisition. It is extremely difficult to change anything afterwards. Typically, if the result showing the virtual path is not satisfactory, a reshooting becomes necessary.

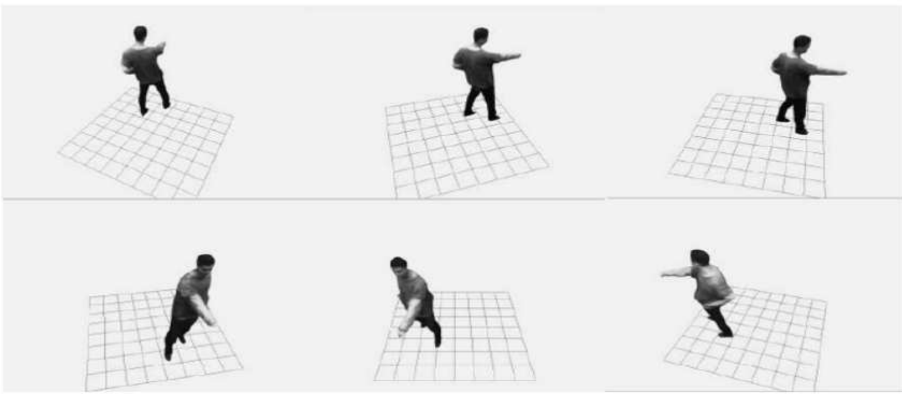

**Figure 2.4.** Stop motion effect with 6 different viewpoints.

## 2.3 Introduction to 3d@Sat project

3D film, 3DTV, Free Viewpoint Video and Super HDTV have one thing in common: they require a suitable way to record, transmit, receive and display images with much more pixels than tradictional TV or film.
In this panorama is inserted 3d@Sat study of ESA whose objective is to investigate future satellite 3D FVV/MVV video architectures and performances. In this project will be study all the chain, starting from the signal capture, analyzing coding, transmission through satellite and reception steps, ending with representation tecniques and displays. The last part of the project is dedicated to a commercial analysis and reccomandations for future implementations.
The scope of this activity is to pave the way for a completely new satellite TV system, capable to reach a wide number of users in all Europe area, delivering both broadcast and interactive applications.

To make the study relevant for future standardization efforts as well as emerging and future commercial deployments, the following three main phases have been adopted

- In-depth review of 3D/FVV technologies and standards
- Pre-assessment of potential application scenarios and use cases for future 3D/FVV services
- Performing meaningful simulation and analysis for most promising and most relevant use cases

### 2.3.1 Structure of the project

This project has been performed in some steps that are presented in form of Technical Notes (TN).
In every TN is examinated a particular step of the chain, proposing solutions and highlighting potential problems. Here is presented a brief overview, step by step, of performed work.

**Step 1** The first step is a review on existing technologies and representation formats, investigating the possible tecniques to represent 3D videos. This is a foundamental step because the choice of a certain representation tecnique determines requirements for shooting, rendering, coding, transmission, and also for reception and display phases. The review on existing technologies comprises also a wide study on displays, stereoscopic and autostereoscopic, analyzing different tecniques to represent stereo 3D or Multiview.

**Step 2** The second phase of the project regards a study on transmission system, considering two different scenarios: broadcast and interactive. A first investigation involves Ku-band, actually used for many video services over satellite, taking as reference Hotbird 7 satellite. With a look on future is important also an analysis on multibeam Ka-band satellite scenario, with the prevision of the developement of regional 3D services.
Considering typical Hotbird coverage area for Ku-band and Modem ACM study of ESA for multibeam Ka-band, a stream scheme has been proposed and analyzed in terms of bandwidth consumption and efficiency.

**Step 3** After a wide study of both existing technologies and satellite scenarios, a simulation has been conduced to check the validity of the proposed stream scheme and of the selected representation format. The scope of this step is to emulate the behaviour of the satellite link in all the descripted scenarios and to verify the real possibility to implement this scheme.

**Step 4** This step of the project, that can not be missing in a customer service, is a commercial analysis, trying to determinate a possible reference market for this service.

Starting from the simulation results, a comparative analysis between the different implementation approaches has been performed, in order to achieve an esteem of cost and complexity impact. Has been analyzed also the commercial viability of the solutions analyzed previously.

**Step 5** The last step is a sort of reccomandations for future use. Once that are defined general lines of the project, we are only at the begin of the work, because it will follow the real implementation phase. The fact that at the moment doesn't exist a standard for theese applications, requires the introduction of some guidelines to allow future developements and implementations.

### 2.3.2 OS rule and experience

Partecipation of OpenSky in this project starts from its experience about satellite system. OpenSky is a commercial operator that provides Satellite Internet access, bi-directional satellite IP service, push-services and streaming channels.
OpenSky contribution regards above all commercial analysis, where thanks to its experience in sector, it is able to provide a complete study on the impact of a new technology like 3D FTV/MVV in terms of complexity and costs. Services offered by OpenSky that can be taken as starting point for a this commercial study are in particular:

**TOOWAY** Tooway is the latest consumer product for Internet access from satellite and now OpenSky is the first and biggest itlaian operator. This service, oriented to users that are not reached by terrestrial ADSL, is in a full developement phase and it is proceding its evolution adding continuosly new applications: voice over satellite(ToowayVoice), monitoring satellite service(ToowayCam), satellite data survey for energetic production systems(ToowayEco), secure VPN network creation(ToowayVPN).

**SAT-IPTV** VideoSat is a new platform that allows the creation and operation of "video-based" Value Added Services. This platform is based on OpenSky as data transport and on an innovative user device(Set Top Box) where the end-user is able to display the content and utilize the service with improved user-friendly and reliability.

**Digital Cinema** OpenSky is supporting the cinema passage towards a completely digital structure providing a transport system to cinemas and content providers. Apllications of this service are mainly for Push content and Live events. In collaboration with ESA, OpenSky is ISIDE project.

**3DTV** Opensky is involved with ESA and Skylogic in a ”Stereoscopic Broadcasting” project. This project will provide a first test and demo channel with 3D TV service.

All these activities represent a strong basis for the rule that it has to perform in this project. Opensky holds a significant rule not only in the commercial analysis but also in other steps of the project, like satellite system analysis and also review on existing technologies.

### 2.3.3 Partners in the project

This project is in collaboration with two partners: EADS Astrium and Fraunhofer HHI. Every partner is project leader of a different part of the project, but the whole work results from the collaboration of all companies. Here a brief description of the general experience of partners and the contribution that they are giving in 3d@Sat project.

**Fraunhofer HHI** .

It is a leading research institute in the fields of Mobile Broadband Communications, Photonic Networks and Electronic Imaging for Multimedia, situated in Berlin. About 50 researchers in about 20 research projects are working in fields related to multimedia with application ranging from very low bit rate video for mobile services up to high quality coding and processing of Digital Cinema, HDTV, 3DTV and Immersive Tele-Presence. It covers all fields, from algorithm developement to chip and system design. Fraunhofer HHI is especially renowed for its leading role in video coding standardization with ISO/IEC MPEG, ITU-T VCEG, DVB and IETF.
HHI holds the post of Project Leader for 3d@Sat project and it has given a big contribution in many fields. It provides a shooting tecnique that is able to generate color maps and depth maps of a scene, permitting the reproduction of the 3D original scene. A brief description on the scheme priciple and on the basis functioning and settign of cameras is presented in one of the Technical Notes of the project.
Its contribution is also fundamental for the simulation campaign in which, using Satem Tool provided by Astrium, the whole transmission chain has been emulated, from codify step to decodify and representation on the screen.
HHI is besides, in the research field of Ultra High Resolution video system. It has developed a 5K (5000x2000 pixel) multi-projection system.

**EADS Astrium** .

Now EADS Astrium holds a leader role in satellite field, in particular in end-to-end satellite system engineering. Experience gained from Astrium on the years come from some projects in which it has partecipated, often in collaboration with ESA or in ESA studies.
One of these projects that has been used for the coverage analysis for this 3d@Sat project is "ACM Modem study" of ESA. The main objective of this work was in particular to study, optimize, design, build and demonstrate in a real-time laboratory environment a two-way broadband modem exploiting adaptive coding and modulation for both the forward and the reverse link. The DVB-S2 and ACM features of the test-bed have extensively tested with multi-beam Ka-band simulated fading events and other satellite link effects.
Another work leaded from Astrium that results useful in 3d@Sat project is "DVB-S2 Satellite Experiment". The main objective of the DVB-S2 Satellite Experiment project is to perform an in-depth verification and optimisation by satellite of the

many new features present in the DVB-S2 standard, such as high order modulations, enhanced coding and ACM, through extensive test campaigns consisting of field trials and measurements.

EADS Astrium contribution in 3d@Sat project concerns above all the satellite analysis with many considerations about coverage areas, prefigured reception power in different locations, satellite link behaviour in different traffic and weather conditions. Besides, it provided the Satem Tool to Fraunhofer HHI for the simulation. This Tool is a satellite link emulator that provides an analysis model with some parameters that can be changed, to evaluate the goodness of the used transmission system.

# Chapter 3

# Transport system

The main objective of this project consists in the broadcasting of 3D-TV and MVV in the European region. For today applications the coverage of this area is ensured by the Hotbird satellites, placed at 13 degrees East. Hotbird family forms one of the largest broadcasting system in Europe, reaching more than 123 milions of home TVs. For this reason an investigation on existing infrastructure is important, to give a concrete example of functioning transmission method. In this chapter will be studied and compared Ku and Ka band satellites, evaluating their performances, merits and defects. In particular it is necessary a comparation between the two coverage areas, the frequency availability for the introduction of new channels, the antennas sizes and other features.
It means that TV channels in 3D would be broadcasted over classical broadcast satellite system.
Besides, will be presented a multilayered transmission scheme(MSVC), which permits to transport data with different qualities on the same carrier. The adoption of this scheme gives a compatibility with existing user equipements. In this way, every screen should display the received stream with the best quality layer it is able to identify and reproduce.

## 3.1 Layered structure of the stream

How said previously, **Multiview Scalable Video Coding(MSVC)**, is a tecnique used to transmit more than one data stream on the same carrier. This feature, called **Scalability**, gives to end user's displays the possibility to reproduce the images with the best possible quality. It is important to remeber that the main factor that influences the quality of a transmission, and so the received data quality, is atmospheric condition.
With this layered scheme, bad weather does not compromise whole transmission, but only some layers of the stream. In fact the base layer, which transport the 2D data, must be protected so that it is reproducible in any weather condition.
For 3d@sat project, a MSVC structure with four layers has been adopted, how shown in Figure (3.1).

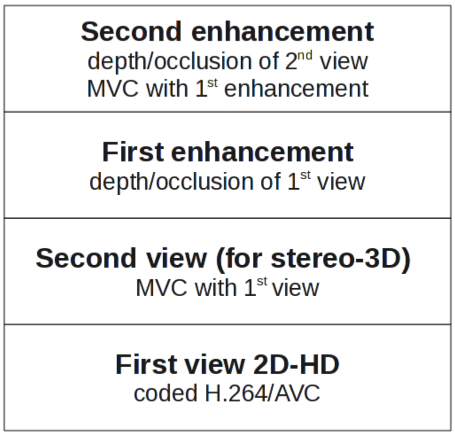


**Figure 3.1.** The four layers of the stream using MSVC tecnique.

**Base layer** In the base layer there is the 2D data stream, that is one of the two views captured from the cameras, usually the left one. These information are sufficient to correctly reproduce the 2D-HD video, while additional data are required for the 3D vision.
The first layer is coded H.264/AVC.

**Second layer** The second layer contains the data relative to the second view, the right one. This view is coded MVC (Multi-view Video Coding) with the first one, assuring bandwidth saving respect the single transmission in H.264 of the two views.
Joining this data with the data of the first layer it is possible to reconstruct the stereo-3D video.

**Third layer** Information contained in layer three are relative to depth and occlusion of the first view. Depth map gives information about the position in the third dimension of objects in the image. Instead occlusion transports information about hidden objects. So, joining these data with precedent layer's data, it is possible to obtain the 3D vision using the **Layered Depth Video(LDV)** format.
It is coded H.264/AVC.

**Fourth layer** The layer four, the highest in this stream structure, transports data about depth and occlusion of the right view. These last layer information are used only for **Depth Enhanced Stereo(DES)** representation tecnique, giving depth perception and map of hidden objects also of the right view. These data permits to reconstruct the interviews, obtaining a MVV presentation on display. It is coded MVC with the first enhancement layer.

The choice to codify MVC the second view with the first one and the first enhancement with the second one, shows the impossibility to receive the second or the fourth layer without the previous one. This is understandable thinking that the second view is useless without the first one for the creation of the 3D video. For the same reason also the reception of the third or the fourth layer without the first one is useless.

### 3.1.1 VCM (Variable coding and modulation)

The first concept to introduce is the **Availability** of a communication system. With the term availability it intends the average time in which the communication is active respect to the total observation time. In our case, in which the stream is splitted in four parts, we speak about availability of each layer. Pratically, the reception of each level can be guaranteed in a different way. In this case, the idea of communication availability falls, introducing the concept of layer availability and scalability. We speak about "scalability" when can't be received all stream layers but only a part of them. This is shown in Figure (3.2).

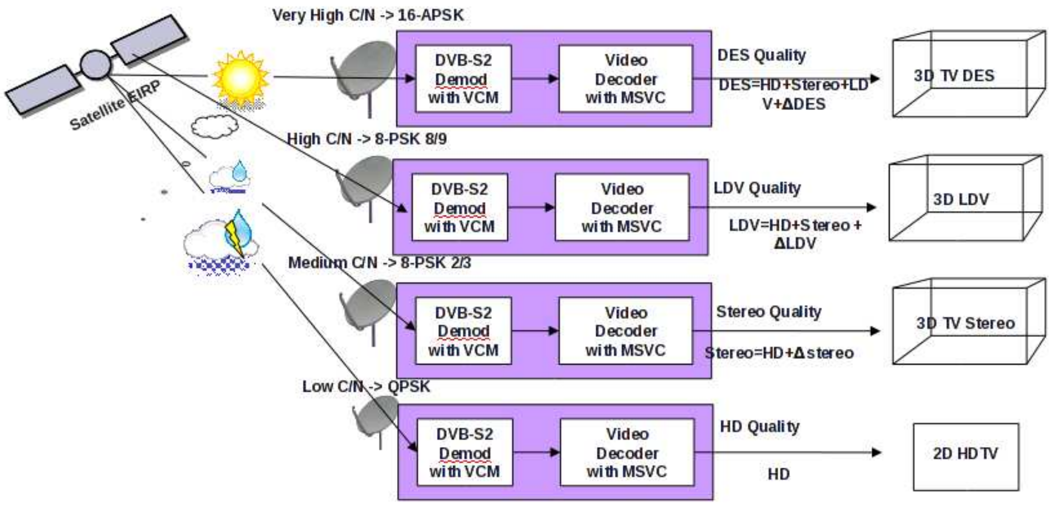


**Figure 3.2.** Scalability concept with different wheater conditions.

This point opens a discussion between two schools of thought: who says that each stream layer can have different protection and so different availability; who instead says that all layers have to be protected in the same way. The correct solution doesn't exist, it depends from which are the target and the service to offer. Surely, this feature provides many intermediate solutions that could be analysed to find the best configuration for one's aim.

In the first case, the lower protection of enhancements layers respect to the 2D-HD one, gives the possibility of a switch from 3D to 2D in case of bad weather conditions. It offers band optimization, but it could bears sickness sensation to end-users for a continue

switching between 2D and 3D with variabile link conditions.
Instead, in case of equal protection of layers, the problem is the high band consumption. But there is another thing to consider. If a user pay to watch 3D, it has to watch 3D, not a variable quality. This is a reason for which probably the choice will be directed to guarantee an high availability, even if slightly different, for each stream level.
There are many ways to protect a stream. The main are related to the transmission power, the FEC code used in the coding phase and the modulation type. These are the parameters that it is possible to change to increase or decrease the availability and so to modify the required bitrate.
This transmission method supporting different modulation, coding and power is called "Variable Coding and Modulation"(VCM) and each different configuration of modulation and coding is called **MODCOD**. So, to receive this data stream, transmitting and receiving devices have to support VCM transmission and DVB-S2 standard, the most used in today satellite system.
This means that every layer can be transmitted with a different MODCOD and the number of received MODCODs depends from the link condition. Then, the image or video quality depends from the number of MODCODs that are received and displayed.

### 3.1.2 Data-rate analysis

A simply esteem on the required bandwidth for the transmission can be made, taking the MSVC scheme presented in the previous section and a full-DES video format.

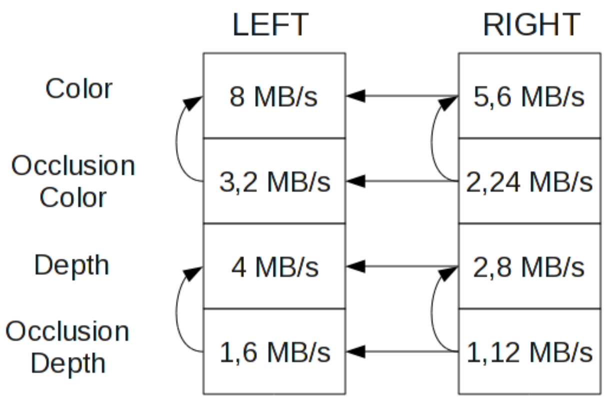


**Figure 3.3.** Bitrates esteem for a Full-DES video format.

Elements considered for this analysis are: color, occlusion color, depth and occlusion depth. Theese elements are in form of maps. For example, for depth, is necessary a grey's scale map, in which far points of the image are represented with dark points in depth map, while near points with light ones.
For this analysis a MVC coding format is considered. A view, in this case the left one, is

assumed as "basis view" and it can be decoded independently as simple 2D video. Starting from the required bit-rate for the color transmission of the "basis view", it is possible to estimate, ever using MVC coding, the required bit-rate for the additional information of this view and of the other maps.
In case of Full-DES tranmission and HDTV-Resolution for the basis view, the resultant bit-rates are shown in Figure (3.3).
With this resolution the bit-rate of the "base view" can be estimated around 8 Mbit/sec. Really, this is not a perfect assessment, because the coding level depends from the correlation betwenn views. In fact the prediction structure of MVC, how happens for H.264/AVC or MPEG-2, is most efficient if the differences among views are small.

## 3.2 Satellite system

Two different bands are been investigated as possible broadcast bands for this application (Ka and Ku bands). Ku-band trasponders have introduced for the first time the wide-band access over satellite. Thus, with this innovation, many applications on communications are developed, particularly TV and radio channels. Not only broadcasted applications have gain ground, but also bidirectional communications like videoconferences. Thanks to its wide area coverage, this transmission band reach a big number of users, included all people that are not reached by terrestrial wide-band line.
Today in fact, the mainly part of TV and radio communications over satellite are broadcasted in Ku-band. For this reason it is saturing very quickly and Ka-band is gaining ground.

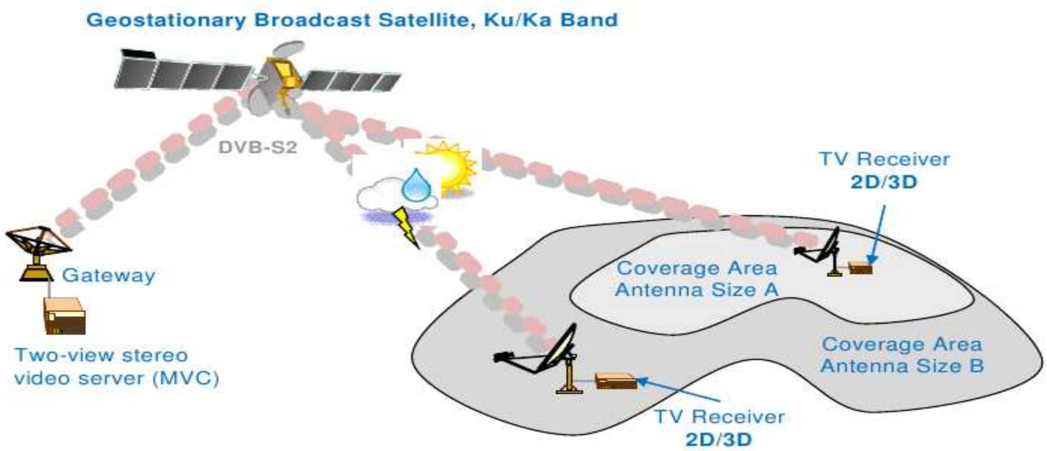


**Figure 3.4.** Complete satellite communication scenario.

Besides, launch of a whole Ka-band satellite, KaSat, is programmed for the next future by Eutelsat, so it is important also a look in this direction. About Ka-band, "ACM Modem study" of ESA is considered as basis for analysis and considerations. The transmission in

Ka-band uses an higher frequency than the Ku-band, with advantages in terms of antennas size, but with drawbacks in terms of coverage area. Smaller antennas and smaller coverage area than Ku-band are the features of Ka-band.
For all mission scenarios, the scheme presented in Figure (3.4) ,with the DVB-S2 standard, is used in the forward link.
The DVB-S2 standard supports a wide range of modulation formats and coding schemes, from $QPSK1/4$ up to $32-APSK9/10$. Transmission modes supported by DVB-S2 standard are

**Constant Coding and Modulation (CCM) mode** . The transmitted signal uses a single modulation and coding scheme, like in DVB-S standard.

**Variable Coding and Modulation (VCM)** Several coding and modulation schemes are transmitted, but the MODCODs are not adapted in time according to propagation conditions. Different MODCODs could be transmitted up to one MODCOD per program. This is the transmission mode selected for the broadcasting of MSVC programs.

**Adaptive Coding and Modulation (ACM)** The transmitted MODCODs are dynamically adapted according to the destination of the traffic and the related propagation conditions. ACM tecnique requires a return link in order to inform the gateway on the link condition. This is a well-suited tecnique for interactive service like Internet access.

### 3.2.1 Ku-band analysis

With Ku-band it indends a portion of the electromagnetic spectrum, particularly used for satellite communications, in a range that goes from 12 to 18 GHz.
Here a typical Ku-band broadcast scenario over Europe is considered. The scope is to assess the performances of MSVC with VCM for the most common TV broadcast scenario. For the analysis, a 33MHz bandwidth transponder is assumed, obtaining a coverage like those presented in Figure (3.5).
The presented coverage area refers to the Hotbird 7 satellite, placed at 13 degree east.
How shown in Figure, Europe is completely covered. Or better still, the coverage area extend also partially in North Africa and West Asia.

#### System description

The system comprises

- A satellite offering a European coverage through a single wide beam
- A Gateway transmitting the TV programs to the Satellite

- Many terminals allowing the reception of the TV programs by end-users
- A satellite Control Centre

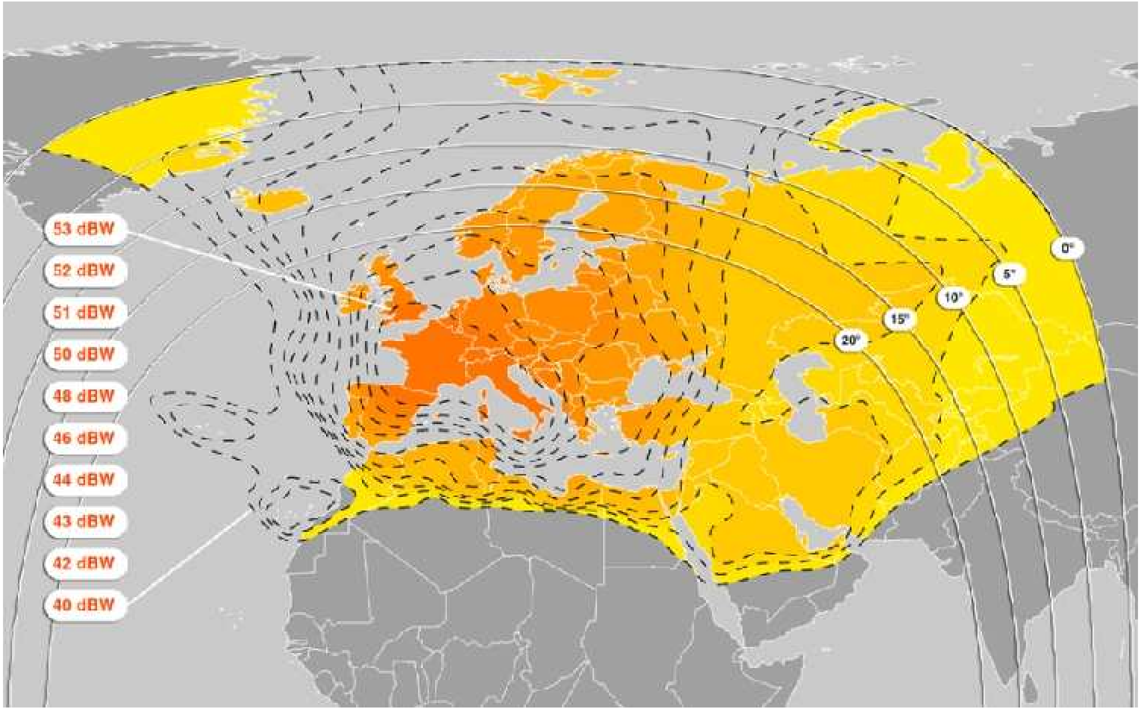


**Figure 3.5.** Typical Ku-band Hotbird coverage area.

To optimize the usage of the on board power resource, satellite transponder works near saturation. In these conditions the power losses are minimized.
Both polarizations vertical and horizontal are used over the wide beam, usually one polarization for the uplink and one for the downlink. The satellite cross-polarization isolation performances are assumed to be better than 27dB on the uplink and on the downlink.

**Gateway description.** Gateway is the device that provides access to the satellite for the broadcasting of TV programs. It is assumed that its video encoder has to be able to support the delivery of 3D TV programs with MSVC scheme. With MSVC at each quality layer can be associated a different DVB-S2 MODCOD. So, is fundamental that the Gateway is able to implement a DVB-S2 video encoder with the possibility to use VCM mode.
RF gateway sizing is performed in "clear sky" conditions, in which a minimum EIRP of 72dBW is required. But the real performances of the gateway depends from many factors like position, weather conditions and visibility.

**Terminal description.** Also in reception side is requied a device (modem) supporting DVB-S2 standar and implementing VCM tecnique and a video decoder supporting the delivery of 3D content with MSVC.

Two terminal antenna sizes are considered in the analysis: 60 cm and 80 cm antenna sizes, the first one in the center of the coverage area and the second one at the edge of the coverage area. Considered RF performances are

- G/T of 15.0 dB/K for the 60cm antenna
- G/T of 17.5 dB/K for the 80cm antenna
- Cross-polarization isolation of 25dB for both antenna types

**Link-budget analysis**

This analysis aim at providing a first assessment of the link budget in order to identify the main contributors and to show if the link budget is well-balanced. Link-budget is performed for both antenna types with similar satellite performances.
The following additional assumption are adopted

- Link budget in clear sky condition, with 1dB of atmospheric attenuation considered on uplink and downlink
- On both links, interference contributions from other satellite systems have been considered. It has been computed that the interference level shall not exceed 18% of the considered link termal noise.
- An overall system margin of 0.5dB has been considered.

The following results and comments can be performed

- The feeder link budget is well-balanced between thermal and interference contribution
- The feeder link contribution slightly degrades the performances: by 0.8dB for the 60cm link budget and by 1.2dB for the 80cm link budget
- The increase of terminal antenna size from 60 to 80cm (2.5dB increase in G/T) brings a significant overall link budget improvement of 1.7dB

### 3.2.2 Ka-band analysis

Ka-band is approximately between 18 and 40 GHz, respect to the 12-18 GHz range of Ku-band.
This higher frequency range respect to Ku-band permits the reception of the signal with a smaller diameter antenna, roughly 45cm in the center of the covered area, and roughly 60 cm at the edge of the coverage. But high frequency means also smaller coverage area. So, to reach the same coverage area of the Ku-band scenario, a multibeam system is required. For this reason, the Ka-band multibeam satellite system considered in the ESA "ACM Modem" study is used for the analysis. The following system description corresponds to the Forward link of the ACM Modem satellite system.

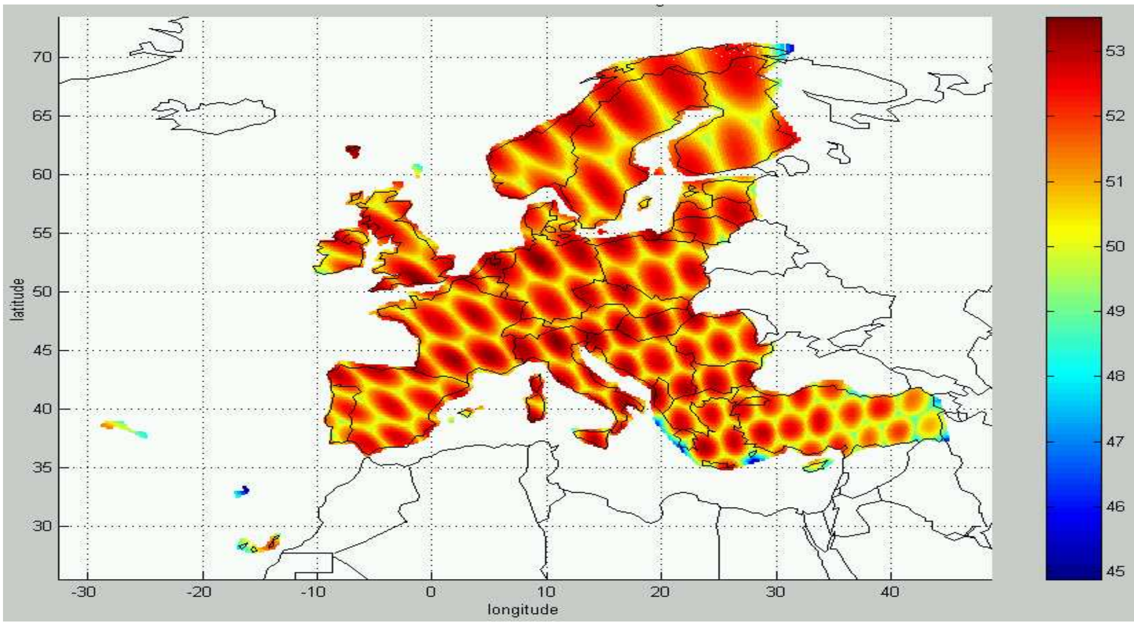


**Figure 3.6.** Transmit gains on the forward downlink for a multibeam Ka-band satellite system.

The main features of the multibeam Ka-band system represented in Figure (3.6) are the following:

- Orbital position: 33°East
- Total Used Forward Link Bandwidth: 25GHz
- Frequency band: 19.7-20.2GHz on the user side, 27.5-29.5GHz on the feeder side
- Number of used beams: 100
- Beam size: 0.4°
- Polarization: Circular
- Bandwidth: 250MHz per user beam with 4 colours frequency reuse scheme
- Connectivity: Each gateway is connected with up to 16 pre-defined beams

The points at the center of each coverage area, those in brown and red in Figure (3.6), are reached from a single beam and having a high signal level, they require a small antenna. The points in yellow instead are at the edge of the coverage area, where adjacent beams cross themself. Users placed in these regions exploit the beam reaching the highest power. A problem that can occourr with a multibeam system is interfernce between adjacent regions, so an efficient frequency reuse scheme has to be implemented.
In Figure (3.7) is presented the uplink frequency reuse, while in Figure (3.8) there is the downlink one. Up to 4 carriers of 45Msps can be trasmitted per user beam.

The associated color link scheme is represented in Figure (3.9). The efficient frequency reuse plan presented assures the best multibeam configuration.
The main characteristics of the satellite are the following

- Platform: Alphabus
- Receive reflector size: 1.76m
- Transmit reflector size: 2.64m

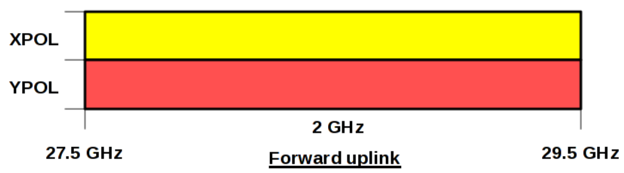


**Figure 3.7.** Frequency plan for the uplink of multibeam Ka-band

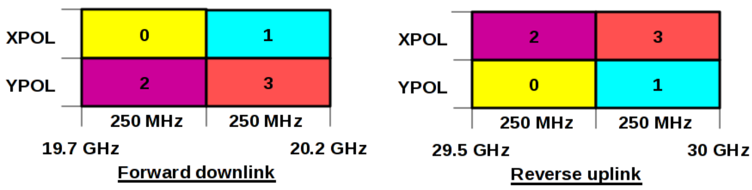


**Figure 3.8.** Frequency plan for the downlink of multibeam Ka-band

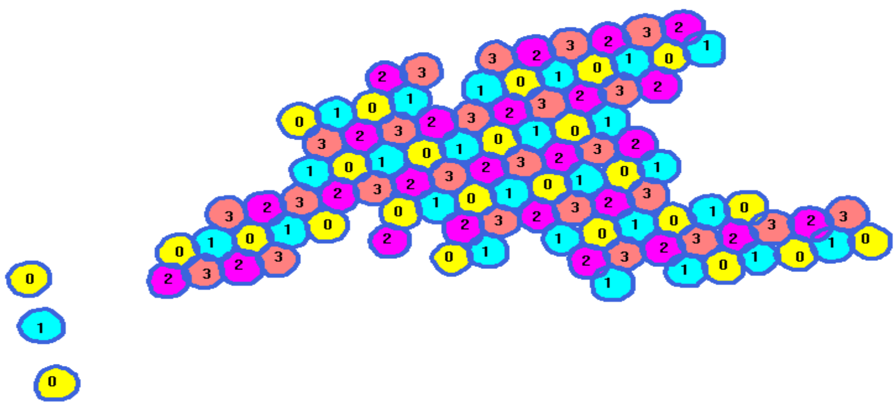

**Figure 3.9.** Multibeam link scheme.

From the user side, it is considered that for every geographical point, the terminals are served by the beam which has the maximum gain amongst the 100 beams.
So, for terminals located on the main coverage or between adjacent beams, the gain is

generally better than 50dBi. Instead, for terminals located at the edge of the coverage, the gain can be down of 45dBi.
Differently from previously analyzed Ku-band, which used both vertical and horizontal polarizations, this system forsee only one polarization per beam.
On the feeder link, the considered satellite G/T is equal to at least 23dB/K over the considered gateway locations. The satellite co-frequency co and cross polarization interference is equal to at least 25dB over the considered gateway locations.

**Gateway description** Two different strategies could be envisaged: multiplexes within a single carrier Internet and video traffic, or dedicates one or several carriers to each service. Both options are possible with the DVB-S2 standard. As mentioned previously, 45Msps carrier has to be transmitted. The related required gateway EIRP is equal to 62dBW per carrier. The gateway cross-polarization isolation performances are assumed to be better than 30dB.
Like previous scenario, it is assumed that the gateway is able to support DVB-S2 standard implementing VCM tecnique and that the video decoder is able to support the delivery of 3D content with MSVC.

**Terminal description** Two antenna sizes are considered: 45cm antenna size at the centre of the coverage, and 60cm antenna size at the edge of the coverage. Obviously, this is only a consideration, but nothing forbids to install a 60cm antenna at the centre of the coverage. This will produce only a benefit in terms of signal level.
The considered RF performances are

- G/T of 12.5dB/K for the 45cm antenna
- G/T of 15.0dB/K for the 60cm antenna
- Cross-polarization isolation of 20dB for both types of antenna

#### Link-budget analysis

The objective of considering two terminal antenna sizes is to get small antenna at the centre of the coverage and larger antenna at the edge. However, the link budget is performed for both antenna types with similar satellite performances.
Other considered assumptions are

- Link budget performed in clear sky condition with 1dB of atmospheric attenuation considered on the uplink and on the downlink
- On the uplink, interference contribution from other satellites is considered. Interference level has been computed assuming that the interference level shall not exceed 28% of the considered link thermal noise
- On the downlink instead, this interference level shall not exceed 18% of the considered link thermal noise

- An overall 0.5dB system margin is implemented

The following result are obtained from the analysis

- The feeder link budget is interference limited. This is not a surprising result considering that the same ploarization is reused for all beams
- The feeder link contribution slightly degrades the performances: by 0.6dB for the 45cm link budget and by 0.8dB for the 60cm link budget
- The user link budget is almost well-balanced between termal and interference contribution for the 60cm scenario

## 3.3 Link-budget

The aim of this analysis is to evaluate the link budget performances of the end-to-end system, defining an appropriate MODCOD to each quality layer of the stream. The first requirement for this study is target availability for each quality layer. How said previously in fact, at each stream layer is possible to associate a different availability, with the result that link condition influences the possibility to receive correctly and reproduce or not a particular portion of the stream. Varying MODCOD it is obtain just this result.
Then, it is possible to derive either the average spectral efficiency or the required carrier symbol rate to transmit the full 3D MSVC program. This include all quality layers.
Considerig the four levels of the stream, the following assumptions are taken:

- Base quality layer: 2D-HD view, 8Mbps, to be received over 95% of the coverage area with a target time availability higher than 99.7%
- Stereo quality layer: $2^{nd}$ HD view, 5.6Mbps, to be received over 95% of the coverage area with a target time availability higher than 99.5%
- LDV quality layer: depth + occlusion views, 8.8Mbps, to be received over 93% of the coverage area with a target time availability higher than 98%
- DES quality layer: $2^{nd}$ depth + occlusion views, 6.16Mbps, to be received over 90% of the coverage area with a target time availability higher than 97%

To get the cumulated density function (CDF) of the link availability over the coverage area for each MODCOD, an Astrium software tool is used. This tool take in account the following parameters

- The satellite antenna gain on each point of the coverage
- The satellite antenna isolation on each point of the coverage
- The free space losses

- The terminal RF performances
- The atmospheric propagation fading
- The degradation of the terminal RF performances due to atmospheric fading
- MODCOD thresolds

The tool is able to compute for each point of the coverage and for each MODCOD the related link availability. The link availability CDF is assessed for each terminal type. Then, it is possible to combine both analyses by selecting a reference point, for example a target availability and a MODCOD, for the small antenna terminal: the points of the coverage that do not reach the target link availability and MODCOD would use a larger antenna terminal.

**3D broadcast Ku system**

Here the broadcast system in the first scenario is analyzed, the Ku-band one. It is performed in 2 cases: with 60cm and with 80cm antennas. For the following analysis are taken the system requirements previously presented in Subsection (3.3). Starting from the required bitrate, the percentage coverage area and the target availability to provide, the simulation returns as result the best MODCOD that satisfies all parameters. Other interesting results are the effective achieved availability, the resultant symbol rate and the spectral efficiency in bit/symbol. Results are for each stream layer: 2D-HD, 3D-stereo, LDV and DES.
Here there are the results of the analysis regarding the 60cm antenna.

| **Video layer** | **Bitrate (Mbps)** | **Cover. area** | **Target availab.** | **MODCOD** | **Symbol rate (Msps)** | **Reached availab.** | **Spectral efficiency** |
|---|---|---|---|---|---|---|---|
| 2D-HD | 8 | 95 % | 99.7 % | QPSK 3/4 | 5.5 | 99.78 % | 1.452 |
| 3D-stereo | 5.6 | 95 % | 99.5 % | QPSK 5/6 | 3.5 | 99.56 % | 1.615 |
| LDV | 8.8 | 93 % | 98.0 % | 8-PSK 2/3 | 4.5 | 98.48 % | 1.936 |
| DES | 6.16 | 90 % | 97 % | 8-PSK 2/3 | 3.2 | 99.11 % | 1.936 |

With these results, the average spectral efficiency $\eta_M$ for the transmission of the 3D MSVC program is

$$\eta_M = \frac{8}{1.452} + \frac{5.6}{1.615} + \frac{8.8}{1.936} + \frac{6.16}{1.936} = 16,7\text{Msps}$$

If the same program would be transmitted in Constant Coding and Modulation, it would be transmitted with the QPSK 3/4 MODCOD. In this case the average spectral efficiency would be $\eta_M = 19.7$Msps, that is 15% of more bandwidth.
Here are presented the simulation results about the 80 cm antenna.

| **Video layer** | **Bitrate (Mbps)** | **Cover. area** | **Target availab.** | **MODCOD** | **Symbol rate (Msps)** | **Reached availab.** | **Spectral efficiency** |
|---|---|---|---|---|---|---|---|
| 2D-HD | 8 | 95 % | 99.7 % | 8-PSK 2/3 | 4.1 | 99.72 % | 1.936 |
| 3D-stereo | 5.6 | 95 % | 99.5 % | 8-PSK 2/3 | 2.9 | 99.72 % | 1.936 |
| LDV | 8.8 | 93 % | 98.0 % | 16-APSK 2/3 | 3.4 | 98.17 % | 2.575 |
| DES | 6.16 | 90 % | 97 % | 16-APSK 2/3 | 2.4 | 98.97 % | 2.575 |

Therefore, the average spectral efficiency $\eta_M$ results

$$\eta_M = \frac{8}{1.936} + \frac{5.6}{1.936} + \frac{8.8}{2.575} + \frac{6.16}{2.575} = 12.8\text{Msps}$$

If the same program would be transmitted in Constant Coding and Modulation, it would be transmitted with the 8-PSK 2/3 MODCOD. In this case the average spectral efficiency would be $\eta_M = 14.8$Msps, that is 13% of more bandwidth.

**3D broadcast Ka system**

For Ka-band system the same requirements and the same results of the previous scenario are considered. The only difference is the antenna sizes that in this case are of 45cm and 60cm.
Here is presented the case with 45cm antenna size.

| **Video layer** | **Bitrate (Mbps)** | **Cover. area** | **Target availab.** | **MODCOD** | **Symbol rate (Msps)** | **Reached availab.** | **Spectral efficiency** |
|---|---|---|---|---|---|---|---|
| 2D-HD | 8 | 95 % | 99.7 % | QPSK 4/5 | 5.2 | 99.76 % | 1.549 |
| 3D-stereo | 5.6 | 95 % | 99.5 % | 8-PSK 3/5 | 3.2 | 99.63 % | 1.740 |
| LDV | 8.8 | 93 % | 98.0 % | 8-PSK 3/4 | 4 | 98.52 % | 2.178 |
| DES | 6.16 | 90 % | 97 % | 8-PSK 3/4 | 2.8 | 98.73 % | 2.178 |

The average spectral efficiency in this case results

$$\eta_M = \frac{8}{1.549} + \frac{5.6}{1.740} + \frac{8.8}{2.178} + \frac{6.16}{2.178} = 15.3\text{Msps}$$

If the same program would be transmitted in Constant Coding and Modulation, it would be transmitted with the QPSK 4/5 MODCOD. In this case the average spectral efficiency would be $\eta_M = 18.4$Msps, that is 17.25% of more bandwidth.
Now, only the 60cm antenna size remains to consider.

| Video layer | Bitrate (Mbps) | Cover. area | Target availab. | MODCOD | Symbol rate (Msps) | Reached availab. | Spectral efficiency |
|---|---|---|---|---|---|---|---|
| 2D-HD | 8 | 95 % | 99.7 % | 8-PSK 2/3 | 4.1 | 99.77 % | 1.936 |
| 3D-stereo | 5.6 | 95 % | 99.5 % | 8-PSK 3/4 | 2.6 | 99.58 % | 2.178 |
| LDV | 8.8 | 93 % | 98.0 % | 8-PSK 5/6 | 3.6 | 98.33 % | 2.422 |
| DES | 6.16 | 90 % | 97 % | 8-PSK 5/6 | 2.5 | 98.56 % | 2.422 |

The average spectral efficiency in this case results

$$\eta_M = \frac{8}{1.936} + \frac{5.6}{2.178} + \frac{8.8}{2.422} + \frac{6.16}{2.422} = 12.9\text{Msps}$$

If the same program would be transmitted in Constant Coding and Modulation, it would be transmitted with the 8-PSK 2/3 MODCOD. In this case the average spectral efficiency would be $\eta_M = 14.8$Msps, that is 12.7% of more bandwidth.

## 3.4 Interactive scenario

Not only broadcast scenario is important to be investigated. Another interessant scenario that could gain ground is the interactive one. Some pratical applications can be in medical or surgical field. Today, 2D videoconferences in this field are alredy used, and in a future this could be extended also to 3D.
Regard common users, the main obstacle to surpass is the high cost. In fact, how explained below, the required bandwidth for this type of applications is very high, and so very expensive. Regard 3D interactive scenario there are to do some considerations. 3D interactive applications require an high bitrate that can't be assured by terrestrial ADSL internet connection, so satellite is investigated as best communication system able to guarantee this bitrate.

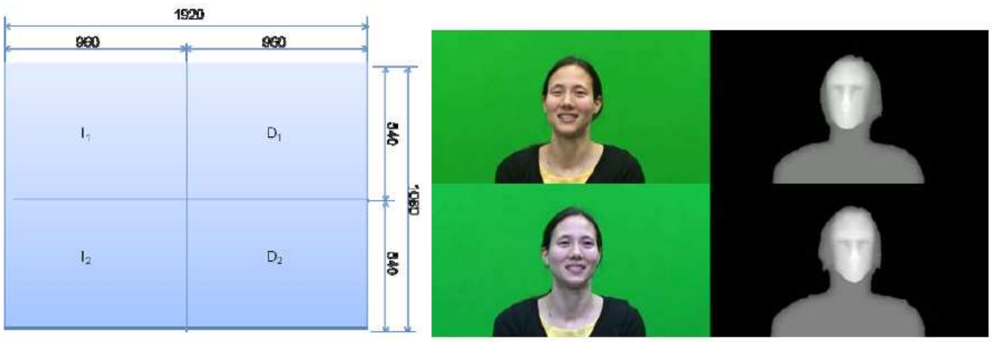


**Figure 3.10.** 3D interactive video format: 2 video + 2 depth at a quarter HD resolution.

Considering for example a 3D videoconferencing application between three sites with two users per site, the video stream has the structure presented in Figure (3.10). It has

to be considered how a point to point 3D transmission, where each site has to transmit 4 video + 4 depth maps to each remote site. At the receiver side a particular display is required. In fact it receives 2 video + 2 depth streams, and it has to be able to translate it in a single HD stream. The video is encoded as H.264 while depth map is encoded as grey valued RGB image. In this way depth requires less bitrate than image.
With this scheme a single HD stream has the resolution of 960 x 540 QHD.
For a bitrate analysis, some considerations are required. The size of video image is a quarter of HD, so an upper bound of 2 MB/s is assumed.
Then, considering the complete data rate for a 4 video + 4 depth stream, analysis steps are:

- 2 video + 2 depth for quarter HD resolution = 20.4 MB/s / 4 = 5.1 MB/s
- 4 video + 4 depth = 10.2 MB/s

This high bitrate required from this type of scenario could be provided from Ka-band based satellites.
A problem that rises in this system is delay. In fact, above all for real time applications, communications between sites are acceptable only if the response from remote sites is quickly. In other words the main parameter to consider is the distance between the satellite and sites, that is the orbital distance.
Typical roundtrip time for GEO satellites is 600ms, and it is not acceptable for all real time applications that require at most 400ms roundtrip time.
So the investigation for interactive scenario has to focus on LEO/MEO satellites.

Chapter 4

# MVC (Multi-view Video Coding) tecnique

Coding is a foundamental step before transmission. With a good coding in fact, it is possible to reduce appreciably the required bandwidth, that is one of the main factors responsible of the transmission costs. 3D-Video and Free Viewpoint Video (FVV) require a coding format that is able to manage more than one view and other objects like depth and eventually occlusion.
So in the coding scheme there are N input lines, that in the coding phase are merged in a single output stream. Then, at the receiving side the reverse process has to take place, obtaining in output the N initial lines. This process is illustrated in Figure (4.1).

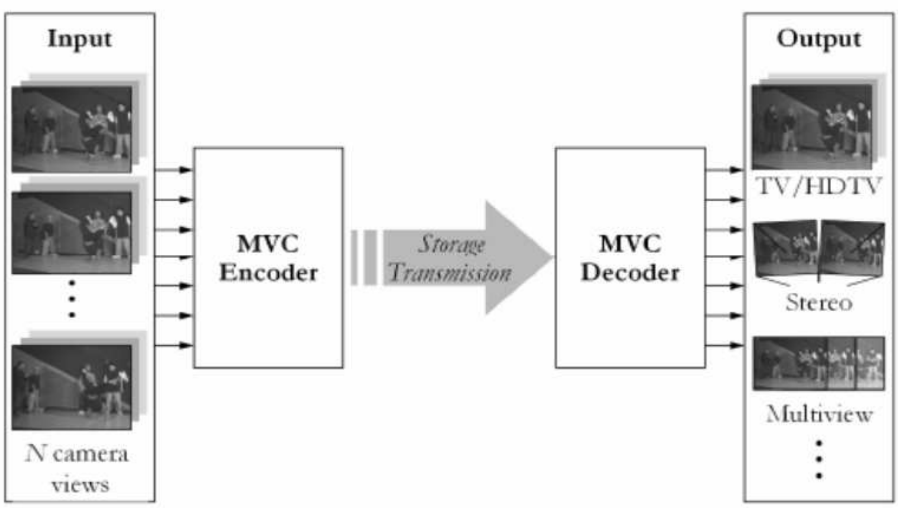


**Figure 4.1.** Basic MVC structure defining interfaces.

To evaluate a video coding format, the main parameter to consider is the compression efficiency, that is the ability to reduce at the least the data in output, mantainig a good quality of the video. But it isn't the only factor to consider to evaluate a video coding standard. For example, a good coding has to introduce low delay, fundamental for real-time communications. Also bandwidth consumption influences coding efficiency and often

it is in opposition to delay. Other factors are: resource consumption(memory, processing, power) and error robustness. Usually, in this wide range of parameters, the optimal solution corresponds to a compromise. The important is to find a good trade-off, the best configuration for own's scope.
The most promising coding format for the Multiview scheme analysed in 3d@Sat project is MVC, developed from H.264/AVC and providing many improvements respect its.

## 4.1 Improvements respect H.264/AVC

Being an extension of H.264/MPEG-4 AVC, MVC mantains all features that made this format the best for the coding of 2D videos. It extends the AVC temporal prediction scheme with an interview prediction scheme. Many other improvements characterize the evolution towards MVC and the main are presented in the following paragraphs.

**Interview prediction** The basis of H.264/AVC is a temporal prediction, that exploits the dependence of images in time, to reduce data to transmit. Often, in consecutive frames the only variation is the motion of an object or the shifting of the image, so pixels that differ from previous frame are not many. Knowing this differences, it is possible to estimate the motion, predict next frames and consider for this frame only pixels that differ from original image, reducing the necessary information. At the receiving side, it is sufficient to carry out the opposite process. This is the concept of **Motion compensation**.
MVC, being planned for Multi-view contents, uses in addition a prediction scheme that exploits the dependence from adjacent views, called **Interview prediction scheme**. Thinking about a Multi-view scenario, more cameras shoot the same scene simultaneously from slightly different viewpoints, so in this case there is a little variation also between adjacent views. The interview prediction scheme, interpolating next images to create an intermediate view, allows to reduce the required camera number. Therefore, the addition of temporal and interview predictions provides an high information saving, handling the big problem represented by bandwidth consumption.

**View random access** An important factor to evaluate a 2D video coding format is temporal random access. MVC introduces also view random access provided from some pictures, called **key pictures**, that don't come from any prediction. This leads to the concept of scalability. It means that a decoder can access a portion of a bit-stream in order to generate a low quality video output. This can be obtained with a reduced temporal or spatial resolution. With MVC, additional scalability is added. To display a reduced number of the original N views, it is possible to access also a portion of the whole video stream. A natural application of this scalability is surely backward compatibility. This means that a decoder can extract only the bit-stream corresponding to a single H.264/AVC view. This ensures that MVC can be applicated also with 2D displays, but only if the decoder supports the MVC format.

## 4.2 MVC prediction structure

The central requirement for any video coding standard is high compression efficiency. In the specific case of MVC this means a significant gain compared to the independent compression of each view, how happens in AVC coding format. This allows to gain, nominally, up to 50 percent bandwidth. Certainly, this is only an optimistic esteem, because compression degree depends strongly from correlation between views.
Now, starting from a innovative method to implement temporal prediction, the whole MVC prediction structure is presented.

### 4.2.1 Hierarchical B pictures

Since the IBBP structure used for anchor coding is not the most efficient temporal prediction structure possible with H.264/AVC, in this section is introduced the concept of hierarchical B pictures. A typical hierarchical prediction structure with three stages of a dyadic hierarchy is depicted in Figure (4.2).

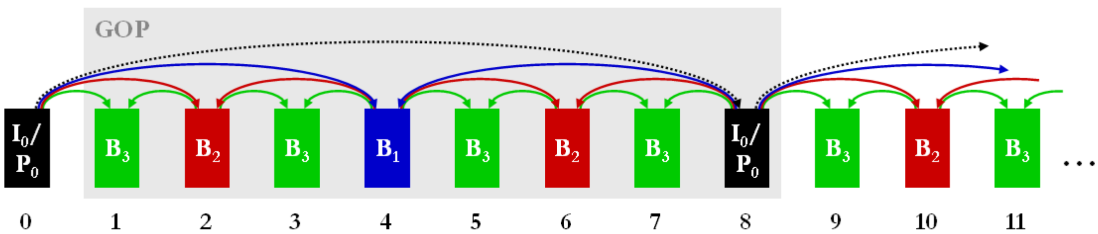


**Figure 4.2.** Hierarchical reference structure for temporal prediction

The first picture of a video sequence, represented with I letter in Figure (4.2), is intra-coded and called "key picture". These key pictures are located ever in regular temporal positions, so they are coded at regular intervals. A key picture and all pictures located between this key picture and the previous one form a **Group Of Picture(GOP)**.
Key pictures are the basis of the prediction scheme, that has a tree structure. Two consecutive I pictures are used to predict an intermediate picture (**B1 picture**), that in turn with one of the two key pictures is used to predict a **B2 picture**, and so on. The temporal hierarchy levels of the scheme are denoted by the indices next to B symbol. It is ensured that all pictures are predicted using only those having an higher temporal hierarchy level. Hierarchical prediction structure can also be used for supporting several levels of temporal scalability. For this purpose it has to be ensured that only pictures of a coarser or

the same temporal level are employed as references for **Motion-Compensated Prediction(MCP)**. For this reason, the coding order has to be chosen in a way that reference pictures are coded before they are employed for MCP.
So, in Figure (4.3) is swown a possible coding order, how shown that guarantees minimal decoding delay inside a GOP. First, are coded all pictures that are directly or indirectly used for MPC of the first picture of a GOP in display order and the first picture itself.

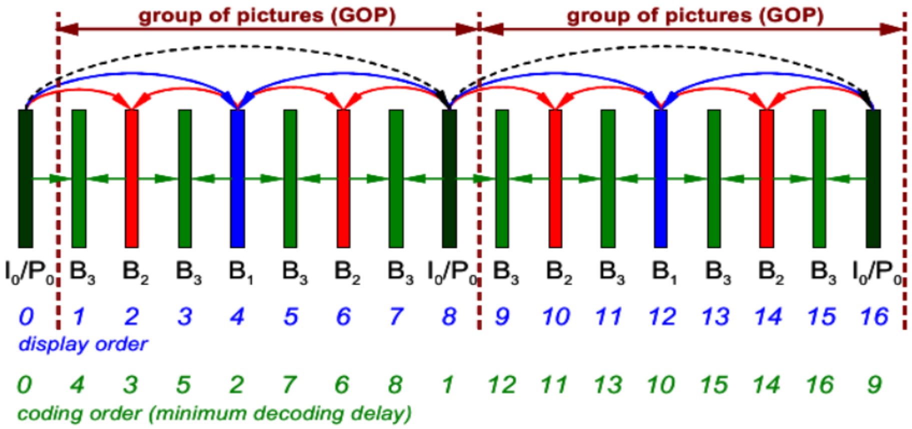


**Figure 4.3.** Hierarchical coding structure with 4 levels.

The Figure (4.3) shows that first the key picture and then the first pictures of every hierarchical level are coded. Next, all pictures requiring the second picture of the GOP and the second picture itself are coded.
For such a coding scheme, coding and decoding delay result as

$$D_d = L_h - 1 \qquad D_c = N_{GOP} - 1$$

where $D_c$ and $D_d$ represent coding and decoding delay, $L_h$ represents the number of hierarchical levels in a GOP and $N_{GOP}$ is the number of pictures in a GOP. So, coding delay is independent from the coding order. But this parameter influences instead decoding phase.

### 4.2.2 Overall scheme

A property of video coding based on motion compensated prediction is that pictures in intra mode(not predicted) result in considerable higher bit-rate than in inter mode. So, replacing intra coded pictures with inter coded ones a bit-rate saving is obtained.
Adapting this approach to a structure in which there are a certain N number of views, there is surely a coding gain respect to the independent coding of each view. For other views, I initial pictures are replaced by P or B pictures, obtaining on the whole the scheme represented in Figure (4.4).
However, this coding efficiency has to be paid in terms of delay. How shown in the scheme in Figure (4.4), it is visible that the bit-stream can be sent only when all pictures of a

GOP are arrived and all B pictures of each level are predicted. So, a fundamental rule is performed by the number of pictures in a GOP.

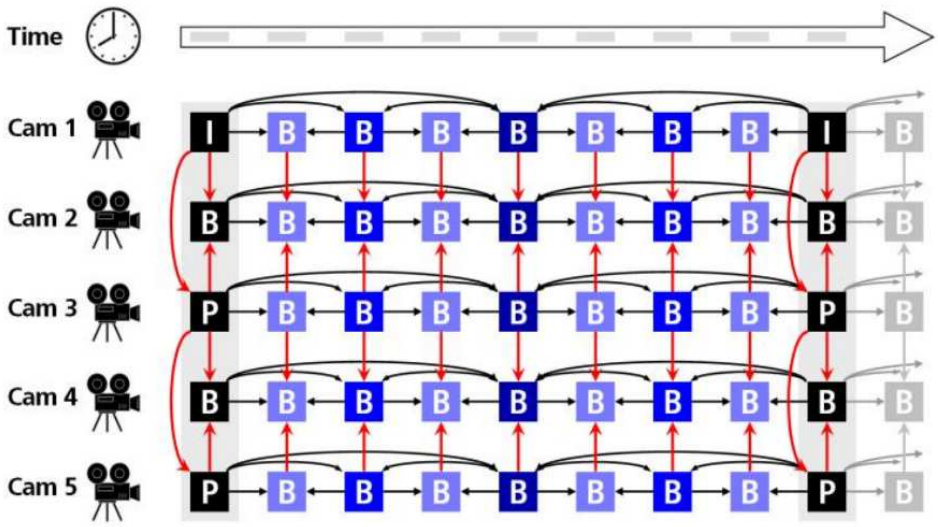


**Figure 4.4.** MVC whole prediction scheme.

A wide GOP gives surely a coding gain but requires an higher waiting time and more computational complexity for the prediction algorithm. With a small GOP instead, the advantage in term of delay is paid in term of compression efficiency. Sperimentally, a good compromise is a GOP of 8 pictures, used also for previous represented schemes.

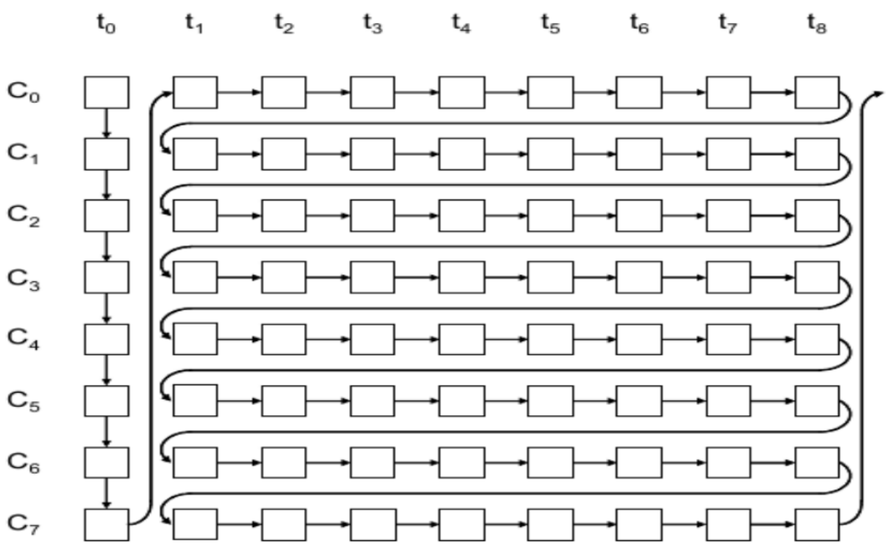


**Figure 4.5.** Rendering and interleaving of multi-view input for compression with H.264/MPEG4 AVC.

MVC is mainly designed for multi-view auto stereoscopic displays, supporting more

views at the same time. Its bit-rate relies on scene complexity resolution and camera arrangement, but with the increase of views' number, the rate reduction might be not enough to respect the channel constraints. Therefore, MVC is nowadays not only applied for MVV formats with a possible large number of views, but also for advanced formats.
The inter-view/temporal prediction structure in Figure (4.4), applies hierarchical B pictures in temporal and inter-view direction. This can be realized with a standard-conforming H.264/AVC encoder using its multiple reference picture syntax. For that, the multi-view video sequences are combined into one single uncompressed video stream using a specic scan, how presented in Figure (4.5). This uncompressed video stream can be fed into a standard encoder software and the inter-view/temporal prediction structure can be realized by an appropriate setting of the hierarchical B picture prediction scheme. This is a pure encoder optimization, thus the resulting bitstream is standard-conforming and can be decoded by any standard H.264/AVC decoder.

### 4.2.3 Error robustness

In terms of error robustness, the MVC algorithm has been designed to equally redistribute between the views the possible image degradation due to bit errors during the stream transmission. We can expect that, despite of the part of the stream that is affected by the error, the Cam1 to Cam N images will be equally degraded. In particular, the I frame connected to Cam 1 is more protected to guarantee that, an error that should affect that view will not propagate along the whole chain of B and P frames. This is usually true on temporal interframe relations and is equally extendable on interviews interframe relations. In general we can expect that the robustness of the whole stream is equal or superior to a group of independent H264 AVC streams.

## 4.3 Efficient coding tools for MVC

The MVC extension of H.264/AVC includes a number of new tecniques for improved coding efficiency, reduced decoding complexity and new functionalities for multi-view operations. In addition to the final draft of MVC, JVT had also mantained a Joint Multi-view Video Model (JMVM) for MVC.
In JMVM, two coding tools, **Motion Skip(MS)** and **Illumination Compensation(IC)**, are included. MS is a tool to derive motion information from neighboring views, while IC is a tool to compensate illumination changes between pictures. Each tool has shown good coding performances, with an high coding gain using both tools toghether.

**Motion Skip (MS)** The MS mode is motivated by the idea that there is a similarity in respect of motion, between the neighboring views. Instead of coding motion information of the macroblock of the current view $V_n$, the motion is derived from the macroblock of the neighboring view which has the same picture order count (POC) as the current view.

First, the corresponding macroblock, MBCor, is pointed out by the disparity vector dvTn between the current and neighboring views, which is signaled in the slice layer of anchor pictures. Then, a searching process is carried out to find the best macroblock that makes RD cost minimum and is used for motion derivation for the current block. This operation is performed at each view direction, Vn-1 and Vn+1, and only one block position, MBBest, is selected as the best.

**Illumination Compensation (IC)** IC is a tool that aims to get coding efficiency by compensating illumination changes between pictures. This tool employs predictive coding for the DC component of inter prediction residues, where the predictor for illumination change is formed from neighboring blocks since illumination differences tend to be spatially correlated.

Chapter 5

# Simulation Campaign

Here is defined the simulation framework. It is end-to-end, including all layers of the communication system. The emulator(SATEM tool), simulates the behaviour of the satellite link in many conditions. So, the satellite network emulator, simulates in real-time the end-to-end behaviour of the link at IP level. The communication link includes the models of RF parts and propagation channels typical of the DVB-S2 satellite system.

The communication link models will be tuned to best match the physical layers such as DVB-S2 with CCM/VCM/ACM based on modelling, measurement and calibration from real configurations. In the first phase of the campaign, the simulation requirements are introduced and evaluated. So, considering these parameters, a suitable simulation tool can be selected. Select the simulation tool means to define what are the parameters to evaluate during the test, in which way the could modify varying initial conditions, which is the standard selected to implement the simulation framework and what scenarios are complexively emulated and compared.

Then, the last part of the campaign represents the out-and-out simulation, in which the results are effectively catched and compared. In this phase, a real-time signal is sent through the satellite emulator and it is received and displayed at the receiver side. Varying parameters like fading, transmitting power, traffic or buffers size, which characterize the satellite link condition, it is interesting not only a data-raising, but also a display-proof. Pratically, for each MODCOD, the displayed video gives an indication about the effect that parameters variation has on the resultant quality. The goodness of each scenario proposed in Section (3.3) is in this way evaluated.

The simulation surely provides important indications about the modifies, necessary to improve the system. From the achieved data is possible to start an analysis regarding the necessary changes to do. A certain scenario could have a too-low quality, so its MODCOD has to be changed. Another scenario could have an higher quality than the required one, so it is possible to change the MODCOD to save bandwidth.

These, are all points on which will be based the commercial analysis presented in the following chapter. The reference market in fact, will be selected above all from these results. Bandwidth increasing means more costs and so a certain type of users for example.

These concepts will be tracted in the following Chapter.

## 5.1 Simulation Requirements

The plain of the simulation is to transmit a two-views video over DVB-S2 system, with all the scenarios previously presented.
To emulate a real-time transmission, the **Real-time Transport Protocol(RTP)** over MVC format is used. RTP in fact is the most performing protocol for real-time communicatios and MVC format supports the Multiview transmission. A stereo-3D and a Multiview real-time emulation are performed, with the simulation framework shown in Figure (5.2). Each view and depth picture of the video input will be encoded with MVC at 1920x1080 spatial resolution at 25Hz. The left color view is assumed as the base view and it can be decoded independently as simple 2D video.
To evaluate the performances of a real satellite transmission, a long time simulation is required. Many contents, different weather conditions, motion of the scene and of the cameras, texture, framings with different brightness. These are all factor that try out the quality of the resultant received video. For these reasons the selected video test sequence for the simulation has the following features:

| **Spatial resolution** | **Frame rate** | **Number of views** | **Duration** |
|---|---|---|---|
| 1920x1080 | 25 fps | 9 | 450 frames = 18 sec. |

The test video contains nine views with related depth maps. Indoor and outdoor environments are used as set. Besides, in the video there are motion scenes, shooted with fixed and shifting cameras.

### 5.1.1 Transmission format

After the definition of the simulation requirements, it has to choice the application format that better adapt to the selected transmission scenario. How said previously, RTP is the most performing format for real-time transmissions, including both audio and video data. It is designed for a end-to-end transfer of data, providing facility for jitter compensation and detection of out of sequence arrival in data, that are common during transmission on an IP network.
RTP consists of a data and a control part. The data part of RTP provides support for applications with real-time properties such as continuos media, including timing reconstruction, loss detection, security and content identification. Information provided by this protocol include timestamp used for synchronization, sequence numbers used for packet loss detection and the payload format which indicates the encoded format of the data.
The control part of RTP(RTCP), provides support for real-time conferencing of groups of any size within an internet. It offers quality-of-service feedback from receivers to the multicast group as well as support for the synchronization of different media streams.

Adapting this format to the simulation, the left and the right views are in 2 separately RTP flows, encoded MVC, while the two depth maps are in one common RTP flow. This format allow the application of a different MODCOD for each RTP flow at the DVB-S2 link, how required by MSVC.

**RTP adapted to NAL units** A RTP payload format for MVC has alredy been drafted and submitted to the IETF.
MVC and SVC are very similar in system and transport aspects. For this reason, RTP payload format for MVC can reuse transmission modes and principles proposed for SVC. So, it is introduced the concept of **Network Abstraction Layer(NAL)** unit. The NAL is a part of the H.264/AVC Video Coding Standard, and it is employed by RTP. The main goal of H.264/AVC NAL is the provision of a "network-friendly" video representation addressing "conversational" (video telephony) and "non-conversational" (storage, broadcast or streaming) applications.

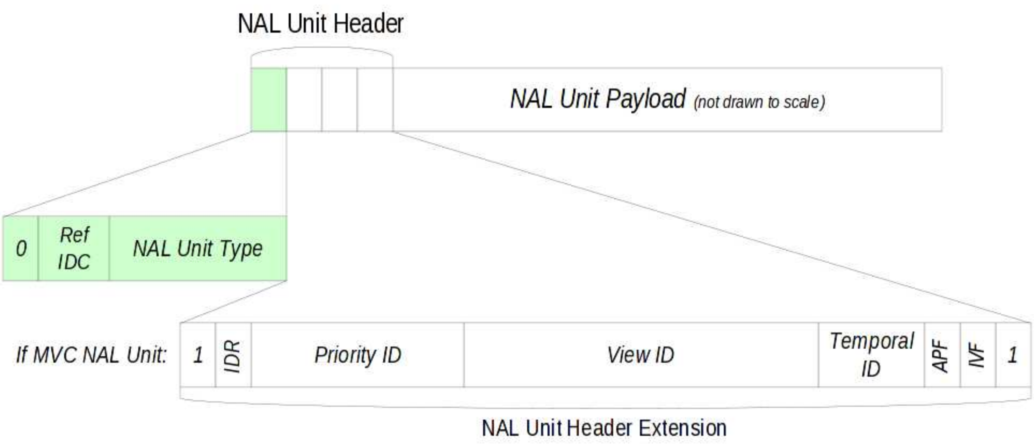


**Figure 5.1.** The NAL unit structure.

The coded video data is organized into NAL units, each of which is effectively a packet that contains an integer number of bytes. How shown in Figure (5.1), the first byte of each NAL unit is a header byte that gives an indication about the data type contained in the NAL unit. Remaining bytes contain payload data, of the type indicated by the header.
Similarly to SVC, the MVC standard specifies a NAL unit header extension of three more octects for NAL units. In this case, the NAL unit header extension carries the identifier of the contained view(ViewID) and its temporal level(TemporalID).
The draft allows, in a given RTP packet stream, to encapsulate NAL units belonging to:

- The base view only
- One or more non base views

- The base view and one or more non-base views

### 5.1.2 DVB-S2 satellite requirements

The simulation will investigate the broadcast of 3DTV over a DVB-S2 system. The considered bands for 3D broadcast are the Ku-band and the Ka-band

- For Ku band, two terminal antenna sizes are considered: 60cm and 80cm. It is assumed that the carrier is transmitted by the gateway at 27.5Msps

- For Ka band, the two terminal antenna sizes considered are: 45cm and 60cm. A 45Msps carrier is to be transmitted

The 3D MVC stream will be composed of three flows, for the two views and their depths. So, they will be coded with a MODCOD that fits the requirements in term of spectral efficiency and priority, ensuring an enhanced overall availability.
In the following Table, are presented all the transmission scenarios to simulate with their relative MODCODs, coming from the satellite analysis of the Section (3.3).

| Band | Antenna size | Stream | MODCOD | Average symbol rate |
|---|---|---|---|---|
| Ka-band | 45 cm | HD | QPSK 5/6 | 11.3 Msps |
| | | Stereo | 8-PSK 3/5 | |
| | | Depth | 8-PSK 3/4 | |
| | 60 cm | HD | 8-PSK 2/3 | 9.51 Msps |
| | | Stereo | 8-PSK 3/4 | |
| | | Depth | 8-PSK 5/6 | |
| Ku-band | 60 cm | HD | QPSK 3/4 | 12.5 Msps |
| | | Stereo | QPSK 5/6 | |
| | | Depth | 8-PSK 2/3 | |
| | 80 cm | HD | 8-PSK 2/3 | 9.67 Msps |
| | | Stereo | 8-PSK 2/3 | |
| | | Depth | 16-APSK 2/3 | |

Thinking about a real transmission, there is an adjuntive problem that can't be forgotten: fading. The received signal in fact, is ever the sum of a direct beam and many its copies. So, particularly in urban environment, where is very difficult to have free visual, is not possible to plan a system without consider this aspect. To reproduce the real effect of a transmission, considering also weather variability, three cases are studied: clear sky, moderate fading and deep fading.

## 5.2 Simulation framework

In this section is presented the simulation framework, in which are described the components used in this phase, focusing in particular in the **SATEM Tool**, the emulator provided by Astrium, used to simulate the behaviour of the satellite link in several different cases.

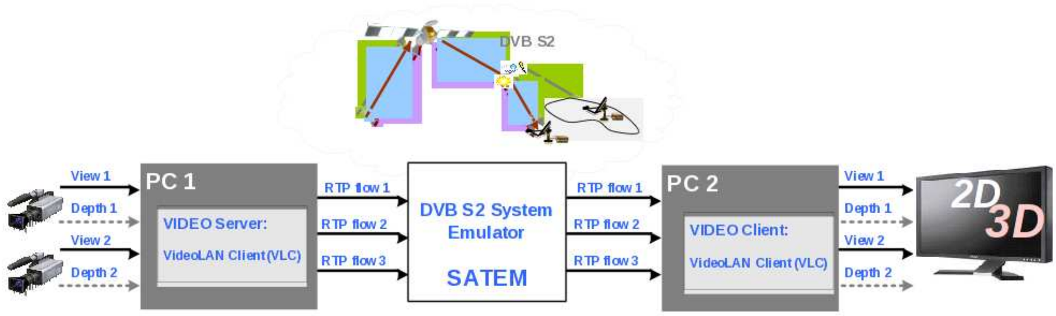


**Figure 5.2.** The simulation framework.

Figure (5.2) shows the whole structure that is composed by the following blocks:

- PC1 (VLC with video server software)
- SATEM Tool (DVB-S2 system emulator)
- PC2 (VLC with video client software)
- Different type of stereoscopic/autostereoscopic displays

The MVD Server and Client softwares for the 3DTV broadcasting have been developed by Fraunhofer HHI as VLC plug-ins. The tools are implemented in associated "Dynamic Link Libraries" (DLL) which are loaded at run-time and executed by the VLC main module.
In order to obtain representative metrics for the received video quality and video format, independently from the real time decoding-ability of the video, the MVD decoder will be extended with a new feature to provide several measures. Achieved measures are: the availability of each view and depth and thus the availability of the video format at the receiver.
This new feature enables a statistical 3DTV format availability output, presented in the following table.

| View | Achieved availability |
|---|---|
| Only first view (2D) | 96.0 % |
| Stereo view (3D) | 98.5 % |

### 5.2.1 SATEM Tool

SATEM (SATellite EMulator) is a tool that simulates satellite systems at IP level. It simulates the behaviour of satellite link in terms of performances, such as the link availability, the packet losses, the delay, the jitter and the Quality of Service (QoS) differentiation. The use of a simulator reduces costs related to the achievement of real tests while providing enough accuracy and realism.

SATEM, how shown in Figure (5.3) emulates the satellite communication system including a gateway, a geostationary satellite and a terminal earth station.

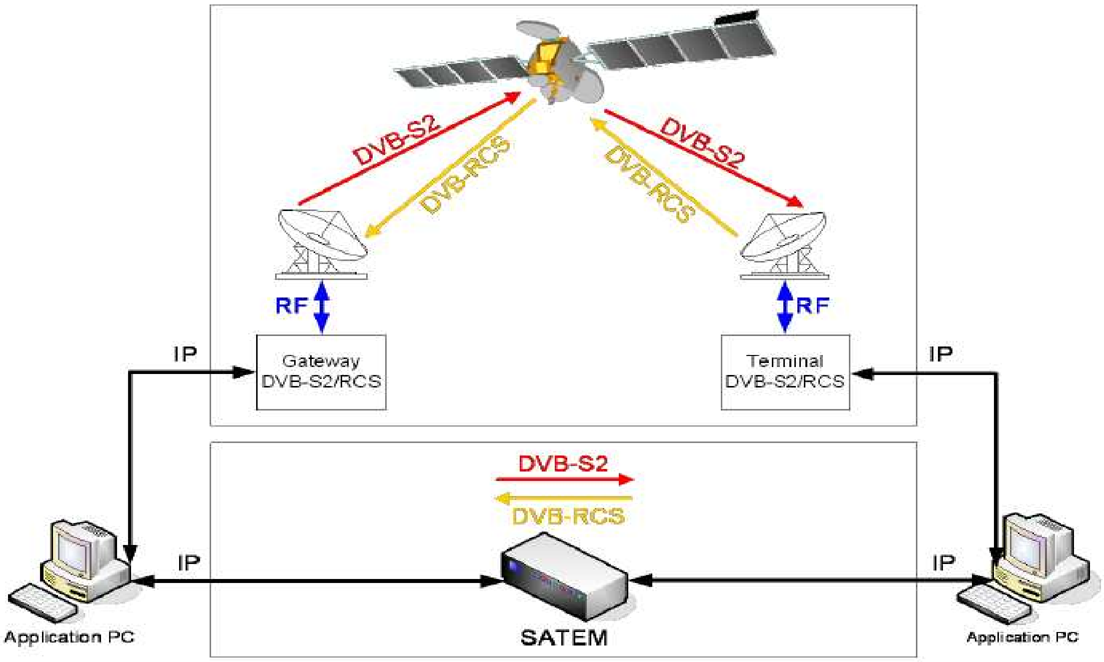


**Figure 5.3.** Emulated network topology.

This tool intercepts IP packets passing between the hosts and modify the network traffic to simulate the transmission over satellite links. The four parameters controlled by the SATEM Tool are:

- Bit rate
- Delay
- Jitter
- Packet loss

To perform losses, the buffer will randomly drop packets according to the link condition. Delay is achieved by retaining packets for the defined duration while jitter follows a statistical distribution correlated with the delay. Finally sizing the buffer enables bit rate control.

In the scope of 3D@sat, three IP streams are considered. Each of them has to be encoded with its own MODCOD, as allowed by the use of the VCM coding mode. QoS differentiation is performed on the destination port of the three streams, enabling the use of a separate MODCOD for each of them.
Through the GUI(Graphical User Interface), the user configures the parameters of the satellite network and has the possibility to observe the traffic going through the router, ensuring effective control of the system. The Monitoring and Control (MC) system performs real time calculations of the satellite link settings and sends commands at IP level to the router. The traffic statistics are collected every second and displayed as graphs.

### 5.2.2 Error generation

SATEM reproduces the errors happening on a real satellite link. These are simulated by the interface with an error ratio that defines the percentage of IP packets to be lost. For instance, in normal conditions, the link error ratio is set to 1E-6 % of all the IP packets going through the SATEM. Errors can be applied independently on all stream as assumed by the use of VCM. When fading events are applied by the SATEM control application, this rate is changed to fit the required settings.
There are several ways to apply C/N. In the default mode, a constant C/N representing clear sky condition is applied. The user can change this constant C/N to any value he desires. Also, a down/up ramp can be applied. In this mode, the C/N decreases regularly until it reaches the defined depth, and then increases in the same way to its original value. Finally, fading time series can be applied. These time series are a compilation of C/N values representing a realistic evolution of the C/N over the time.

## 5.3 Simulation scenarios

The emulated scenarios are two:

**Real time** . A real time transmission, decoding and display is performed for a 3D video signal. This real time simulation is not performed for a Multiview video, but only for CSV(Conventional Stereo Video). The received video is displayed in a Stereoscopic display as Miracube or JVC. The format for the transmission is two-view stereo 3D video based on MVC.

**Non real time** . A simulation for a Multiview video signal is made, using MVD (2 video plus 2 depth) video format as described previously. However, for this scenario only transmission phase is in real time. Instead is not the same for decoding and display phases that are performed in a second moment. The resultant video is displayed in a WoW display that is able to support until 9 different views.

In order to demonstrate the effect of increased or degraded 3D quality the different type of displays needs to be fed with specific input depending on the received video signal.

In the Table (5.3), the left column lists all the different types of input signals, which are transmitted via thesatellite depending on the weather conditions.
The 2nd, 3rd, and 4th column shows in which fashion the received video signal is presented. On the input side with distinguish between simple 2D HD as a fall back solution in bad weather conditions. The next case is conventional stereo video (CSV), which allows receiving two 2D HD streams. In the optimal case in very good weather conditions, multiple video + depth can be received. The intermediate case of a single video + depth is a sub-set of the optimal case.

| **Transmission and input** | **2D display** | **Stereo display** | **Multi-view display** |
|---|---|---|---|
| 2D-HD | 2D-HD | 2D-HD | N x same 2D-HD |
| CSV | 2D-HD | 2 x 2D-HD | View $1 \ldots \left(\frac{N}{2}-1\right)$ : left view<br>View $\left(\frac{N}{2}\right) \ldots N$ : right view |
| V+D | 2D-HD | 2 x 2D-HD | N x 2D-HD |
| MVD | 2D-HD | 2 x 2D-HD | N x 2D-HD |

On the display side, a 2D display can always show any video input, but just 2D. This can be considered as the backward compatible scenario to standard TV receivers. If on the receiving side, a stereo display is available, 2D HD can be presented if only 2D HD is received. In all other scenarios, the display can present stereoscopic video, i.e. 2 x 2D-HD views.
The presentation of content on a multi-view display is somehow different depending on the available input. This results in the following

- If only 2D HD is received, then this single view needs to be shown on all N views of the multi-view display.
- If two views i.e. CSV is received then only in the centre position of the display, stereo can be perceived. In all neighbour view to the left, the left view is seen, while in the neighbour views to the right, the right view is seen. An alternating arrangement of left and right views is not possible, because the user will then perceive incorrect assignement of views at every second position.
- If video + depth is received, then all required N views can be rendered.
- If MVD, i.e. 2 x (V+D) is received, then any multi-view display can be fed with N views of improved rendering quality as additional video+depth information is available.

# Chapter 6

# Commercial analysis

The commercial analysis of the 3d@Sat project has been performed in Opensky, with the support of its partners. This analysis starts from the results of the simulation campaign leaded from Fraunhofer HHI, with the Satem Tool provided from Astrium.
For each scenario a complexity analysis has been leaded, assessing how the introduction of a MVV/FTV system can modifies existing infrastructures, or in which way it can integrate itself with them. The study regards all steps of the chain, starting from shooting phase until the reproduction on displays.
An important factor to assess in this analysis is cost. Costs analysis for video production, broadcaster and users are developed, evaluating also if some intermediate scenarios are possible, where not the complete chain is implemented, in order to mitigate complexity and costs.
A SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis has been performed in this Chapter, highlighting which are the innovations, drawbackes, strong points and weaknesses of FTV/MVV scenario respect traditional 2D and above all, comparing its with 3D one.
An important factor that has been analysed is the commercial viability of the presented scenarios. This study is not so easy because it is only in a futuristic view and with today's technologies we are still far to the possible realization of whole system. It has to consider that, being only in a prototype phase in many fields of MVV scenario, several different and innovative applications can be showed. For this reason it is important also to think how FTV and MVV can extend the today idea of a video presented in a flat screen with more views.
So, contacts with video manufacturers and with content providers are important to understand in which direction they are moving and in which way they think to approach them with this new technology. In this field are alredy active some research groups that are trying to plan devices. Using multiprojector systems, they try to provide immersive sensation to end users.
From the project viewpoint, those presented here is only an introductive presentation of the commercial analysis, being a work in progress now.

## 6.1 Complexity analysis

To evaluate the overall complexity is important to consider how the whole video communication scenario could modify itself with the introduction of 3DTV/MVV technology.
Particular interest is in the difficulies and obstacles that is possible to meet implementing such a scheme and integrating it with existing system.
The first factor analysed, because it joins both broadcast and interactive scenarios is bandwidth consumption. How presented previously in fact, high bandwidth is required to transmit more views, even if codified in a efficient way, using for example MVC. How shown in Table (6.1), for a 2 view-MVD, that is the format used in the simulation framework, the required bitrate is 21.26Mbit/s, and it further increases with the introduction of occlusion maps and so using Full-DES video format.

| **3D format** | **Description** | **Video bit-rate** | **Audio bit-rate** | **Total bit-rate (2.3%overhead)** |
|---|---|---|---|---|
| Full-DES | All components | 28.56 Mbps | 384 Kbps | 29.61 Mbps |
| 2-view MVD | 2 view + 2 depth | 20.4 Mbps | 384 Kbps | 21.26 Mbps |
| LDV | 1 view + depth + occlusion | 16.8 Mbps | 384 Kbps | 17.58 Mbps |
| V+D | 1 view + 1 depth | 12 Mbps | 384 Kbps | 12.67 Mbps |
| CSV | 2 color views | 13.6 Mbps | 384 Kbps | 14.31 Mbps |

### 6.1.1 FTV Broadcast scenario

Using the term broadcast scenario it intends the whole system that comprises a initial phase in which in a certain site there is the shoot of the scene, the post-production and coding phase, the transmission phase from the encoder to the decoder through terrestrial and satellite link, the decoding and display phase that occur in the final user location.
Actually, this is not the only possible scheme. In particular in the first part of the chain there can be some variations. For example the initial shooting phase could be replaced with a push of a content previously stored in a storage device.

**Shooting** The first step of the broadcast chain is shooting. In this phase there are several changes respect traditional stereo 3D and in particular respect 2D capture.
The innovation in 3D shooting is the concept of stereoscopy, that is the possibility to provide depth perception of a scene capturing its from two different viewpoints. This requires that two cameras are setted-up in way that the two viewpoints are at the same distance present between uman's eyes.
The concept of Multiview shooting instead requires the use of many cameras, providing

for example in the case of FTV, a 180° view. This can be obtained using many synchronised cameras. So, a first problem is to guarantee the correct synchronization and distance among cameras.

In many cases the use of N cameras to obtain N views is not the best solution, because the correct set-up and synchronization between many cameras is very difficult. So usually, a limited number of cameras is used, and the others views are reconstructed from the shooted ones, integrating them with additional informations provied by depth maps. Then the used cameras must be able to capture not only images, but also informations about depth and in the case of DES video format also about occlusion.

Additional problems in shooting are surely with lively scenes, because the synchronysed motion of all cameras is very difficult and requires a engine system allowing simultaneously motion.

Certainly, this is a wide field in which there is a perspective of the introduction of new specialized figures. One required figure is surely stereographer. It will has to be able to set-up correctly the cameras and to face potential problems that could be introduce onself. The stereographer has to control the fine tuning of several parameters like convergence and interxial distance and this requires a stereo analysis software that can provide alerts regarding stereo rules violation. Resuming:

- Each camera array shall be operable by a single operator.
- Complex software to control many cameras and analyse many camera streams.
- Mechanism permitting remote correction of the parameters.

**Manufacturers and standard** At the moment does not exist a standard in Multiview video and this surely creates an ambiguity regard which format new screens or displays have to implement. From one side this fact is a drawback, because the absence of a guideline requests to manufacturers to implement all the most promising video formats or eventually the throwing of any not optimum configuration. From the other side this fact allows the improvement of formats and technologies for a future optimal configuration.

For example in the performed simulation of this project a two view-MVD (Multiview + depth) video format has been used. This format requires the use of color and depth maps to recreate more than the shooted views. But it not excludes the introduction in a future of occlusion maps that at the moment are very difficult to create and treat. In fact, giving information about partially hidden objects, they will allow a more quality video or a higher number of predicted views. In this case we speak about DES video format.

Another open field for a future standardization process regards transport cables. While for the transport from the encoder to the decoder are alredy present technologies for the transmission of high bitrate data, the problem presents in the last step, from the decoder to the display.

Existing cables are in fact not able to transport the high bitrates required by a Multiview signal, even if in this direction there was the recent introduction of HDMI 1.4 that among its new features has the possibility to accept also stereo-3D (side-by-side) and LDV (Layered Depth Video) data. This is surely an optimum starting point for future improvements.

**Backward compatibility** Best final user experience is the most important requirement for the whole system. So, the possibility to a backward compatibility is another important question to consider. This has been one of the main requirements of the project, because for a user, to have the possibility to use this new technology without change its devices, seems to be the best solution to allow a quickly growth of MVV.
So, a way to adapt MVV to existing 2D-HD and stereo-3D technology has been considered for both transmission and representation.
Would be certainly a madness try to plan a system that does not use actually broadcast equipment, because this change would be unthinkable and too much expensive. But for user equipment the question would be certainly different. The change in this case is not so radically, in fact will be necessary only the introduction of a new type of display and of a satellite receiving equipment.
Surely, remaining attached to the past, the changing will not be never radically, but there will be many intermediate solutions before the achievement of MVV full integration.

**Dependence from wheater conditions** From the previous study and above all from the choices made in terms of availability of the service, the result is that the quality of the stream is too much dependent from weather conditions. Giving less availability for higher quality levels of the stream, bad condition of satellite link is an obstacle too much difficult to superate and the result is a limited quality. Figure (6.1) illustrates the video output that a user can obtain receiving only a part of the video stream.

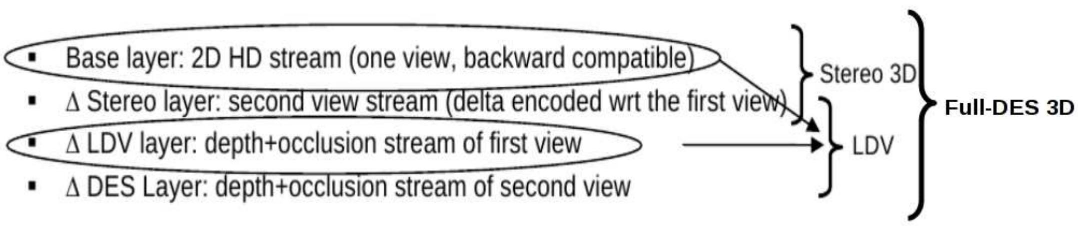


**Figure 6.1.** Displayed outputs without all quality layers.

The problem with this choice is that, if a user pays to view a Multiview content, it is predictable that it will not be so happy to see a stereo 3D or worse a simple 2D video.
This is a convenient solution to save bandwidth, but it not will be acceptable when MVV

will be full integrated.

**Broadcaster** On broadcasting side there's always high interest for new and attracting services. MVV is surely one of these as a natural evolution of the stereoscopic broadcasting that is now gaining interest among many broadcasters as a media to bring customers to new generation HD platforms.
In the technical requirements the most impacting choice is the IP encapsulation of the stream as this differs from the 90% of the broadcasting market that rely on MPEG2 TS. This difference will surely create difficulties to broadcast a signal that can be received as 2D by the existing HD enabled customer base.
The motivation for this choice are:

- In a long term whole video market is moving toward an all-IP infrastructure
- IP stream less constrains the structure of the encoded video, providing more freedom to the encoder phase
- IP stream can be easly routed through other distribution medium in the home network environment
- The difference in bandwidth consuption using GSE/IP/UDP/RTP from pure MPEG2-TS seems negligible
- It represents the best implementation for the future Ka based all-IP satellites like KaSat that Eutelsat is going to deploy.

### 6.1.2 Interactive scenario

In the interactive scenario there are some variation respect to the broadcast one. They regard above all the shooting and the final reproduction phases. Taking for example the case of a videoconference scenario, it involves two or more remote sites and in each site there must be the control of all other ones.
Remaining in a Multiview scenario, it is evident that it not could be performed with today's pc or 2D videoconferencing systems. In fact in both shoot and display phases, improvements and modifications are required. This leads not only to an increasing of complexity but, how we can see later, also to an increasing of costs.

**Shoot and display** The first obstacle is represented by the fact that a user who wish to partecipate to a videoconference must have in his location all shooting equipment. This means surely an high cost for the equipment's hire, but also several problems in the correct setting of the cameras and in the position to mantain to allow the correct capture of the video.

In fact, in a videoconference scenario, on the same time there is the shooting and the display of the videos coming from remote sites. In this, distance is a fundamental factor, because there is an optimal position from which the Multiview video is correctly seen. So there is a strong limitation on user's motion, because moving itself the correct view is not ensured.

**Delay** An additional limitation of interactive scenario regards the maximum tolerated delay. The concept in which interactivity base itself is the real-time communication, so could be understandable a certain limitation in quality but not in time. For this reason the satellite investigation for this scenarios has to conncentrate itself for LEO/MEO.
For the almost totality of real-time applications a delay up to 400ms is certainly unacceptable and for many ones the delay has to be down to 100ms.
So, rather than an investigation on near satellites, a good solution could be to plan the terrestrial transmission system until gateway with the smallest number of link and with the highest speed.

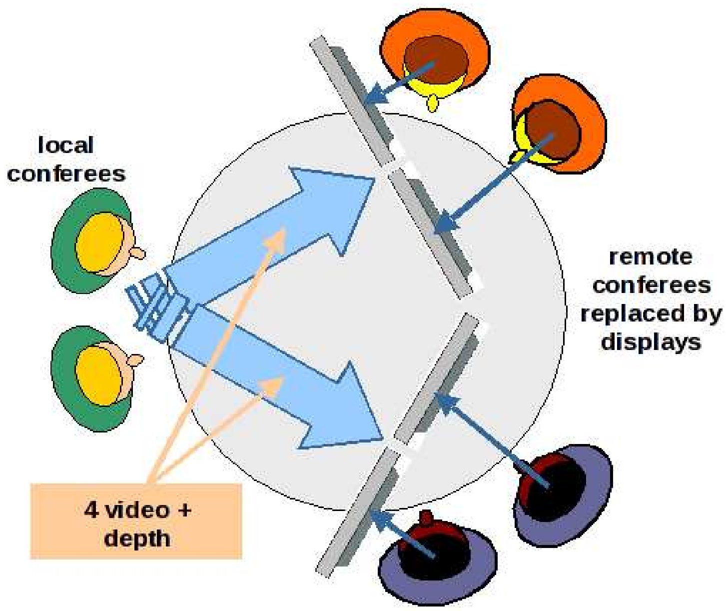


**Figure 6.2.** Interactive scenario between 3 sites: arrangement of 2 video + relative depth in a single HD stream.

Interactive scenario, differently from broadcast one doesn't use a layered stream structure, because the stream is constituted by two videos plus two depths arranged in a single HD stream, how presented in Figure (6.2). So, the transmission infrastructure must be able to treat both stream types.

## 6.2 Benefits/impacts of MVV/FTV integration

MVV/FTV technology gives more immersive sensation respect to traditional 3D. The idea to present more views of the same scene changes completely the idea of 3D vision. In this way a user who moves watching the screen should have the impression of moves himself around the scene.

### 6.2.1 Glasses and motion

Surely a factor of big impact for MVV scenario is that this technology not requires wearing particular glasses. For a user that is watching a film in a sofa wouldn't be a problem to wear glasses, but for users that are watching TV and contemporary are carrying out other activities, glasses are certainly an obstacle.
But the application of this idea is still so far. The possibility of user's motion in front of the screen is only theoretically because there are several problems related to the motion. There is an optimal distance from the screen, from which the left and the right images are directed to the exact eyes and Multiview is correctly viewed. But moving, this optimal condition is not ensured and the two images could be directed to the wrong eye reversing the views.
In Figure (6.3) is shown the concept of Multi-view presentation with a lenticular display, that projects every image in a limited region of the space.

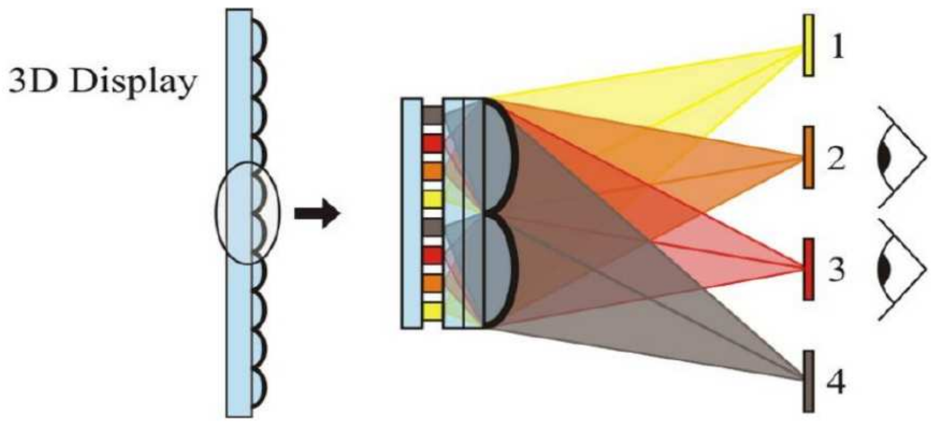


**Figure 6.3.** Problems related to motion with MVV.

In Figure (6.3), the view 2 is directed to the right eye, instead the view 3 is directed to the left one. Shifting sideways, for example in right direction, the view 2 is then directed to the left eye, giving a not correct image reconstruction.
Presentin more views with the purpose to allow to the user the possibility to move and see

different views, but constrain it to stay still, is not a very good solution. So, surely this obstacle will be only the starting point for a wide research in this field.

**Not only flat screen** The conceptual idea of MVV can also extend the today idea of video presentation only in flat screens. In a few words, the idea is the evolution from FTV to real MVV, that is the total user's immersion in the course of the scene.
Would be interesting for example, watch a football match projected in all walls of a location, having the sensation to be in the middle of the match. This concept could be extended also to other video applications, like for example films or concerts and many others.
Speaking about a home environment with a simple stereoscopic TV, an interesting implementation of FTV could be for example the possibility given to the user to select what view of the video to see. If I am interesting in a certain detail on a film, or if I wish to change the point of view of a car race, this is surely an interesting feature of FTV.
It is important also to think to the possible coverage developement of this technology. Being in decrease the available bandwidth in Ku-band and with a good perspective for the Ka-band satellites, there will be surely a sensitive variation of the satellite coverage scenario with the birth of new services.
Ka-band in fact, having a limited coverage for each satellite beam, will allow the birth and the developement of regional and local programs. This factor influences surely in a positive way the today study in MVV field, opening many new possibilities to the TV broadcsat.
Certainly there aren't only positive changes with the introduction of this technology. Many of them, obviously, will rise when MVV will be an actual scenario. At the moment, the first thinking is directed to a possible increasing of the user's laziness. If I have the possibility to watch a football match with the impression to be in the football field staying comfortably sitting in my armchair, why have I to go to the stadium?

## 6.3 Costs analysis

Available technology is not the only obstacle to FTV/MVV developement. Also in this case, like usually happens, is its cost that decides if a system or a technology has the possibility to start or develop its.
Costs are for video productors, broadcasters, manufacturers and users. So there is a very wide landscape of possible costs. Obviously it is important not only an analysis on expected costs for initial impact of MVV service in broadcast system or for the starting phase in which this technology is gain grown. It has to think also about the potential developement of Multiview considering the improvements that could be achieved in several fields.
So, this analysis is performed in this direction, considering all costs sure and expected for overall broadcast and interactive chains.

### 6.3.1 Costs on video production side

From video production side there are several changes respect to the traditional video production. Changes are not only in shooting phase how presented in previous section, but also in post-production and coding.
Regard to the shooting, the main factor that influences cost is not the number of the cameras but the type. Each camera of the Multi-view capturing equipment has to produce a color map that represents the image. Then, interacting with other cameras, it has to produce also a depth map, usually in grey scale, which gives depth information about the objects. It is expected an occlusion map, also in this case obtained interacting with other cameras, which would have to give information about hidden objects so that, at the receiving side it is possible to reconstruct all views of the scene with all original details.
So, it is evident that particularly cameras are required and it is the first cause of costs.
The phase that follows shooting regards the insertion of all objects or texts that the user can see in his display but that are not shooted.
This category comprises for example the subtitles of a film, the ranking of a car race, the result of a football match, the logo of a company or a TV issuing and many other objects.
This objects are added after the shooting and when a set of different views and other maps is available.
The addition needs pc elaboration. For example, to add a logo in a certain part of the video, more views of this logo are produced through pc elaboration. Then, ever using a pc elaboration they are added to the images.
So, it is evident that this elaboration is expensive in terms of complexity and cost because it requires particular equipment and software.

### 6.3.2 Costs on broadcaster side

From broadcaster side, the main factor that influences costs is bandwidth consumption.
Taking in account the layered steam presented previously, that corresponds to a DES video format with 4 different quality layers, the resultant required bit-rate, including also audio data, amount to around 30 Mbit/s.
The required equipment to correctly treat Multi-view contents, differs only in a little part from traditional one. In fact, an initial requirement of this study was the reuse of today broadcast infrastructure, trying to exploit, as much as possible this functioning system.
The main variations in broadcast system are in gateway equipment that must be able to treat Multi-view stream and in the satellite equipment, that has to correctly redirect the signal.
Gateway for example, needs an encoder supporting the delivery of 3DTV programs with MSVC, each quality layer being mapped to a specific DVB-S2 MODCOD. So, the gateway has to implement a DVB-S2 modulator supporting VCM transmission.
In addition, both gateway and satellite equipments have to treat the two different stream

types used for the broadcast and for the interactive communications. It means that there will be the availability of different services implementation.

### 6.3.3 Costs on user and manufacturer side

User's perceived cost is the most important of the analyzed costs. Is the end-user that, in the end decides if a new technology, an invention, or the improvement of an application will have a future.
In this perspective is introduced MVV/FTV service, that surely gives an added value to TV video, but the main factor to consider is if it effectively can be sold or not.
Displays able to reproduce 3DTV signals already exist, but for MVV/FTV (autostereoscopic displays) they are only in form of prototypes, for the absence of a standard and for the defects in the display costruction.
Regard receiving equipment, the cost is not so very high. The antenna is a common antenna dish that can be quickly installed and has to be only correctly aimed to the satellite. Then, it is required a satellite receiver implementing DVB-S2 standard with VCM and a 3D/FTV decoder supporting MSVC format.
These devices are not so expensive and this could influences positively the adoption in home environment of the MVV technology.
It is not the same for interactive communications, because for the reason previously explained, they are too much expensive with today technologies and could be adopted only in medical environment or in situations that don't involve a common user.
From manifacturer side, the main cost regard the technology adopted to correctly reproduce the Multiview on the display. In fact, the principle of Multi-view screen is to direct a certain view in a certain portion of the space. To do this, several methods could be used, for example the lenticular mode, or using parallax barriers.

## 6.4 SWOT analysis

To perform a SWOT analysis means to leads a study about which factors influences positively or negatively the FTV/MVV scenario. This analysis has to be performed comparing this system with 2D, but mainly with 3D.
Today 3D is the technology that is gain grown, even if not so quickly as was estimated, particularly in home environment, for the lack of 3D-TV contents and for the cost of stereoscopic displays that users would have to buy.
For these reasons 3D, service has been adopted mainly in the cinemas, where costs can be written off and thanks to video productors, that are starting to provide film in 2D and simultaneously in stereo-3D.
For all this motives, the first comparison has to be made with 3D and considering the passage from stereo-3D to Multi-view, it has to refer to the passage in course from 2D to

3D.
In the SWOT analysis are considered all benefits and drawbacks that can present with the introduction and use of the MVV service, resuming all features illustrated until here.

### 6.4.1 Strengths

With strengths we intend all factors that influence so positively MVV scenario that could be a drawing power. These, are the reasons for which a user like a video productor or a broadcaster could prefere to pass from 3D to MVV.

**No use of glasses** . This system in fact doesn't require wearing neither passive or active glasses, exploiting features of autostereoscopic screens, able to display more different views in separated regions of the space. This is one of the strengths of MVV giving more freedom sensation to the user, differently from what happens for stereo-3D in which they are necessary.

**More immersive than 2D and 3D** . "Immersivity" is the most researched feature for video communications and 3DTV programs. The today passage from 2D to 2D-HD for example is a good improvement, but the achieved video is ever the same, only with more quality. Insted, the strength of MVV is to give a different perception, a different way to watch and live TV, even if the quality is equal or even worse than the 2D-HD case.

**Possibilities for many new applications** . The panorama for future extensions is very wide, allowing the adaptation of this service in many fields. Real-time communications like video conferences, streaming videos in a simple pc, "olograms" construction. These are only some of the many possible applictions of the MVV.

### 6.4.2 Weaknesses

Weaknesses are the points in which there is to work still hard because they involve several problems to a possible realization of the system.

**High bandwidth** . How expected, a higher number of views means an higher required bit-rate. The only solution is to research an efficient coding tecnique and a representation format that guarantee the minimum quantity of data to transmit.

**Losses in resolution and quality** . Showing several different images at the same time, the screen resolution must be shared from them. The most studied and applicated method in autostereoscopic display building is the horizontal division in parts as many as the number of views.
Screen resolution is increasing with the years thanks to technology improvements, but if the number of the views to present is high, the resolution and th equality is

certainly limited.

**Less comfortable and relaxing** . Like in stereo-3D, the reconstruction of stereoscopic view takes place at human's brain level, which joins the views coming from the two eyes. So, watching for a long time a 3D or a Multiview video tires eyes and particularry brain. In addition, watching a FTV/MVV content combinated with motion can invert the correct vision giving sickness sensation.

### 6.4.3 Opportunities

With opportunities it intends all future perspectives and possibility of developement of MVV, including all applications in which MVV could be used.

**Change the way to live television** . One thing is sure. MVV can radically changes the way to watch and live television. The opportunities offered by this service are countlesses. The interactivity will be one of its strong points. A user will can change the point of view of the image, moving itself or selecting the wished view from a list. The extension of the video will be around 180°, unthinkable with today shootings and display technologies.

**Pave the way to real-immersive communications** . The developement possibilities for MVV are several. A screen that cover a whole wall, a multiprojector system able to recreate the original shooted scene, a person that staying in a room is surrounded by a world that seems real. These are some possible, but still far in time, opportunities that MVV will can give.

### 6.4.4 Threats

These are those factors that differentiate MVV from all other services, but if interpreted in a wrong way by the users, they could stir up a boomerang effect.

**Abandon for high cost and complexity** . Starting from the premise that, shooting and displaying equipments are expensives, the required bandwidth to guarantee a good availability is high, there are many problems for the correct perception of the views and this scenario is not supported by today technology, there is the risk that manifacturers and content providers don't consider the MVV like an interesting and promising service.

**Lack of a standardization process** . This is a problem that must be as soon as possible faced. The absence of a common evolution and standardization process is driving the MVV developement in several different directions. So, the lack of a guideline doesn't motivate the implementation of a common solution, allowing the birth of many different formats and tecniques, with possible creation of incompatibilities.

# Chapter 7

# Review on existing technology for 3D and Multiview

Here is presented a brief review on actual state of the art regarding the 3D technology. In particular there are two main categories of technologies to consider: representation techniques, that indicate the various possible approaches to intend 3D image or video; and displays, which need many new features respect existing 2D ones and that could be differentiate in two groups on the basis of their destinations.
How we can see, for both display and representation tecniques there is a wide choice on possible structures and approaches. In fact these are two important points for 3D evolution, because they are responsible for a future standard adoption in 3D or Multiview field. How said before, here will not go into details of these important aspectes because they are tracted analitically in Gabriele Longhi's thesis.

## 7.1 Video formats

Thanks to the rapid developement of 3D, there are a lot of different 3D-TV formats available and under investigation. Video format indicate at high level the used method to transmit video. This method must be implemented also in displays for video reproduction. There are pratically two approach type: the first approach is based on the transmission of the two views simultaneously, instead the second use only one view, exploiting additional information, in form of greyscale maps, to reconstruct the remaining view. In this second approach, the delicate point is represented by the implemented algorithm to determinate the second view from the first one. This elaboration process requires an high computational effort, because additional elements like depth or occlusion are difficult to treat.
Some of this video formats, created for the 3D, are subsequently developed and adapted also to MVV. For example, a format that exploits two views plus relative depth data is able to represent, using a prediction algorithm, also 3,4 or even 5 views.

### 7.1.1 3D-Stereo formats

This is the most simple 3D format, because it involves only color data, which can undergoes some processing steps like normalization or color correction. It requires both two views and differentiate itselt in two categories for the way in which they are multiplexed.

- Simmetric format : two views multiplexed (in rows, in columns, with alternate fields, syde-by-syde, ...).
- Asymmetric format: 1 of the 2 images is sent with low quality, deceiving uman's brain.

**Figure 7.1.** Asymmetric 3D-stereo format.

### 7.1.2 Formats for Multi-view

These video formats exploit additional data like depth and occlusion to increase the number of views. How said in previous sections in fact, additional data give an high number of information necessary to improve the prediction.

- MVD (Multiple-view plus depth) : more than 1 video + additional depth information
- LDV (Layered depth video) : 1 video + depth/occlusion information
- DES (Depth Enhanced Stereo) : more than 1 video + depth/occlusion information

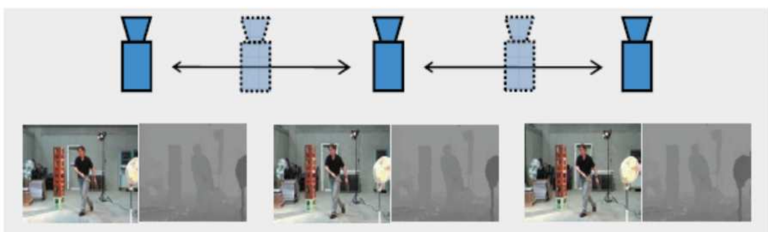

**Figure 7.2.** MVD video format.

Depth Enhanced Stereo is the format adopted and studied in 3d@Sat project. It is a combination of LDV and MVD. In particular it is constituted by two views plus depth and occlusion informations. But, being difficult to create and elaborate, occlusion information are not considered momentarily. Then, the resultant format is a 2-view MVD.

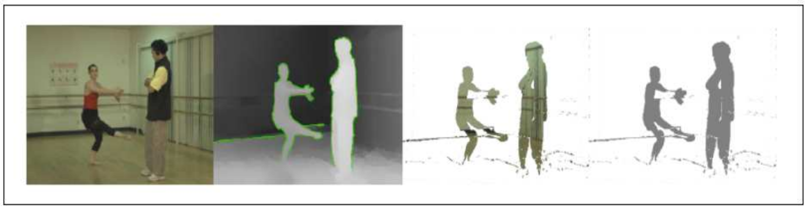

**Figure 7.3.** LDV video format.

## 7.2 Displays

In this section are presented the existing technology about displays and representation devices.
Actually, display field is widely studied because it is a foundamental part and is the bottleneck of Multiview chain. In fact while displays for stereo 3D alredy exist and are also present some applications to this technology, representation devices for Multiview exist only as prototype.
Below there is a brief overview on 3D stereo displays (stereoscopic) and on others displays (autostereoscopic) and devices under developement for Multiview.

### 7.2.1 Stereoscopic technology

Stereoscopic displays use various tecniques to provide to the users depth perception, but this depth is visible only wearing particular glasses.
The principle on which stereoscopic display base it, is the reproduction of two images at the same frame rate of a single image in a 2D display. So, the minimum frame rate has to be 120Hz.

### 7.2.2 Multi-view devices

This very wide class of devices can be splitted in two types: autostereoscopic displays and devices that try to recreate a solid structure of an object.
For autostereoscopic displays for example, particular tecniques allow to show more than two views in different space regions, so moving in front of the screen it is possible to see different views. Glasses are not needed in this case, and the image extension is rougly 180°.

Instead, for the presentation of objects in three dimensions, the problem is that a different display method from a flat screen is needed. For this reason, are under development many technology that use multiple projectors, trying to recreate the original object in a real space. But, how said this technology is only in a initial phase of development.

Chapter 8

# Summary

How said in the beginning, MVV is expected as a far solution, not still ready for the spread market.
All considerations made in these chapters show the big innovations that such a system could provides to a field in continuos evolution like video-communications. Surely, success of MVV will depend on overall deployment cost and additional revenues that will justify the investments.
Obviously, being only in an intial studying phase, it is not possible to know how the market will move in next years. In fact, it is too early to say if MVV in a future will be widely adopted by users and if it will become an available service for the mass-market. This certainly will depend from the adopted choices for its development. Choices in terms of technology evolution, approaches for the reuse of existing devices and infrastructures, reference market, effective availability offered to final users are foundamentals factors.
Nowadays the panorama about the first technical solutions, for example in representation formats, coding tecniques and displays design and building, is very wide. So, from a side, this wide range of possibilities is a good opportunity for the research and the development of new architectures. From the other side instead, a too dispersive panorama gives hesitation and confusion to manufacturers. A display able to support all video formats is almost unthinkable, so today they are only in form of prototypes.
It is predictable that, when will be available several guidelines and after a fast improvement process, these devices will be accessible to all. This result can be expected observing the 3D market, which has achieved excellent results in so short time. It isn't still accessible to all, but mainly for the lack of 3D contents, thanks to another difficult phase to improve, the shooting one.
All these considerations made about the broadcast scenario can be extended to the interactive one, with additional improvements to do in several steps of the chain, necessry above all to decrease complexive costs for the end-users, that at the moment seem exorbitant, to make accessible in a future this service.
Surely real-time communications represent an interesting application field for MVV, even if many technical problems make difficult its implementation. How previously said, shooting

equipment and settting-out, correct distance from cameras and display, different stream respect to the broadcast case, are problems to face.
Could be interesting also an integration of a Multi-view system in a device like a pc. This will not a simple step because the implementation of a Multi-view system in a pc requires many improvements in the technology used to build its. Particular(auto-stereoscopic) screens and new Multi-view devices able to perform webcams' duty would be necessary. Certainly, with a MVV integration in pcs, the production of Multi-view contents would be motivated.
So, concluding, for a wide adoption of such a system, many improvements are still necessary in many fields. The main objective has to be the begin of a standardization process to "trace" several important guidelines, to pave the way to a future MVV system. Then, the next step will have to be the impact with users, trying to exploit in a initial phase existing infrastructures, reducing at the minimum the impact in terms of costs. Only in the end will be possible to change all existing infrastructures to completely introduce the MVV system, but this final scope is expected in fifteen-twenty years.

# Acknowledgements

The first thought and gratitude is for my family, that allows me to study for all these years, supporting me in all my choices and believing in me in any occasion.

A special thanks is for "My Love" who bear me in any moment, also in evenings and nights in which I prefered "Matlab" to her. If this thesis is written in a decent english is also thanks to her careful corrections.

With happiness I am grateful to my tutor Ing. Luca Carniato who has attended me and made this work possible, to my friend "Leggenda" that every morning for six months has entertained me with his "romantic adventures" and "blonde girls", and to all Opensky people who have made me feel welcome.

Thanks to all "CdSA" members. Travels to London and Stokolm, spritz parties and wonderful lunches in Bassano are only some of the beautiful moments spent toghether. The university life has been more easy with them.

Thanks to all my friends. With them I am sharing beautiful days, evenings, dinners based on "panini onti" and the "Gneke Gneke busters" activity, which have undertaken from a little time the "moralizing campaign" in the neighbouring places.

Finally, thanks to all people who has gone into my life, making its better.

# Bibliography


[1] H.M.Ozaktas and L.Onural, *Three-Dimensional Television: Capture, Transmission, Display*. Springer, 1st ed., 2008.

[2] 2009. Technical Notes 1.1, 1.2, 2.1, 2.2, 4.1, 4.2, 4.3, 6.1, 6.2 of 3d@Sat project.

[3] H. Schwarz, D. Marpe, and T. Wiegand, "Analysis of hierarchical b pictures and mctf," *IEEE International Conference on Multimedia and Expo*, July 2006.

[4] O. Schreer, P. Kauff, and T. Sikora, *3D Video communication: Algorithms, concepts and real-time systems in human centred communication*. WILEY, 1st ed., 2005.

[5] Y.-J. Jeon, J. Sung, and B.-M. Jeon, "Analysis of efficient coding tools for multi-view and 3d video," *IEEE 13th International Symposium on Consumer Electronics*, May 2009.

[6] Y.Chen, M.M.Hannuksela, L.Zhu, A.Hallapuro, M.Gabbouj, and H.Li, "Coding tecniques in multiview video coding and joint multiview video model," *IEEE Transaction*, 2008.

[7] P.Merkle, A.Smolic, K.Müller, and T.Wiegand, "Efficient view-temporal prediction structures for multi-view video coding," *IEEE Electronics Letters*, January 2008.

[8] P.Merkle, A.Smolic, K.Müller, and T.Wiegand, "Multi-view video plus depth representation and coding," *IEEE International Conference on Image Processing*, October 2007.

[9] M.Drose, C.Clemes, and T.Sikora, "Extending single-view scalable video coding to multi-view based on h.264/avc," *IEEE International Conference on Image Processing*, October 2006.

[10] N. Ozbek and A.Murat Tekalp, "Quality layers in scalable multi-view video coding," *IEEE Internetional Conference on Multimedia and Expo*, June 2009.

[11] S.Wenger, Y.K.Wang, and T.Schierl, "Rtp payload format for svc video," February 2008. online documentation, Internet Engineering Task Force (IETF), consultable at `http://tools.ietf.org/search/draft-ietf-avt-rtp-svc-08`.

[12] "History of 3d, sensio." online documentation, consultable at `http://www.sensio.tv/en/3d/3d_history/`.

[13] P. S.Yeung, "Ka-band services and technologies," *IEEE Colloquium on Current Developments in Intelsat*, October 1997.

[14] A.Smolic, H.Kimata, and A.Vetro, "Development of mpeg standards for 3d and free viewpoint video," *IEEE Transaction*, October 2005.

[15] K.Muller, A.Smolic, P.Merkle, M.Kautzner, and T.Wiegan, "Coding of 3d meshes and video textures for 3d video objects," *IEEE 6th Workshop on Multimedia Signal Processing*, September-October 2004.

[16] M.T.Pourazad, P.Nasiopoulos, and R.K.Ward, "An h.264-based video encoding scheme for 3d tv," *IEEE Transaction on Consumer Electronics*, May 2009.

[17] P.Merkle, A.Smolic, K.Müller, and T.Wiegand, "Multi-view video plus depth representation and coding," *IEEE International Conference on Image Processing*, September-October 2007.

[18] A.Mayer, B.Collini-Nocker, F.Vieira, J.Lei, and M.A.Vazquez Castro, "Analytical and experimental ip encapsulation, efficiency comparison of gse, mpe and ule over dvb-s2," *International Workshop on Satellite and Space Communications*, September 2007.

[19] A. Smolic and P. Kauff, "Interactive 3-d video representation and coding technologies," *Proceedings of the IEEE*, January 2005.

[20] Y.K.Wang and T.Schierl, "Rtp payload format for mvc video." work in progress, Internet Engineering Task Force (IETF), consultable at `http://tools.ietf.org/html/draft-wang-avt-rtp-mvc-03.txt`.

[21] H.Kimata, S.Shimizu, Y.Kunita, M.Isogai, K.Kamikura, and Y.Yashima, "Real-time mvc viewer for free viewpoint navigation," *IEEE International Conference on Multimedia and Expo*, April-June 2008.

[22] S.Li, C.Hou, Y.Ying, X.Song, and L.Yang, "Stereoscopic video compression based on h.264 mvc," *2nd International Congress on Image and Signal Processing*, October 2009.

[23] A.Smolic, K.Muller, P.Merkle, C.Fehn, P.Kauff, P.Eisert, and T.Wiegand, "3d video and free viewpoint video-technologies, applications and mpeg standards," *IEEE International Conference on Multimedia and Expo*, July 2006.

[24] A.Smolic, K.Muller, P.Merkle, T.Rein, M.Kautzner, P.Eisert, and T.Wiegand, "Free viewpoint video extraction, representation, coding and rendering," *International Conference on Image Processing*, October 2004.

[25] Y.Ohta, I.Kitahara, Y.Kameda, H.Ishikawa, and T.Koyama, "A perceptually correct 3d model for live 3d tv," *3DTV Conference*, May 2007.

[26] Z. Gui and C. Zhang, "3d head pose estimation using non-rigid structure-from-motion and point correspondence," *IEEE Region 10 Conference TENCON*, November 2006.

[1] [4] [5] [3] [6] [7] [8] [9] [10] [11] [2] [12] [13] [14] [15] [16] [17] [18] [19] [20] [21] [22] [23] [24] [25] [26]