PUBLIC

IST R&D. FP6-Priority 2.
SPECIFIC TARGETED RESEARCH PROJECT
Project Deliverable

Doc Number	SUIT_106
Project Number	IST-4-028042
Project Title	SUIT
Deliverable Nature	Report
Deliverable Number	D1.1
Contractual Delivery Date	30.06.2006
Actual Delivery Date	11.09.2006
Title of Deliverable	User Terminal Requirements
Contributing Workpackage	WP1
Dissemination Level	PU
Author(s)	A. Zistler (IRT), Michael Probst (IRT), Gavrilescu Augustin (IBBT), Robbie de Sutter (IBBT), Erwan Le Bras (Vitec), Francesc Enrich (URL), Francesc Pinyol (URL), Víctor Domingo (URL), Gabriel Fernandez (URL), Antonio Navarro (IT).

Abstract

This document identifies a set of requirements of the SUIT end user terminals.

Besides that, the most common terminal description protocols used nowadays or envisaged for future use are described.

Keyword list: Requirements, SUIT Terminal

PUBLIC

User Terminal Requirements

SUIT_106

11-09-2006

Table of Contents

1 Introduction

1.1 Scope

This document is part of the WP1 - Architecture Requirements and Specification - and relates to Activity 1.1 to define a set of general-purpose requirements for the user terminals.

This document has strong relationships to other Activities in WP1, namely Activity 1.2 (QoS requirements), Activity 1.3 (All-IP Support Requirements) and Activity 1.4 (Architecture and Reference Scenarios).

1.2 Objective

The main objective of this document is to give an overview about the different types of end user terminals to be deployed in SUIT and to specify their capabilities and requirements to support broadcast and non-broadcasting services over converged broadcast/broadband networks DVB-T/DVB-RCT, DVB-H and WiMAX.

In SUIT, various reference scenarios are foreseen in order to evaluate and to demonstrate the SUIT concept (scalable seamless services of collaborating networks). To cope with that, in principle, three types of user terminals will be designed and implemented.

The end user terminals to be used can be classified as following:

- WiMAX/DVB-T/H/RCT terminal (Chapter 2)
- WIFI terminal (Chapter 3)
- MHP-IPTV terminal (Chapter 4)

The first type of terminal to be designed for SUIT will receive and process a multiple description scalable video conveyed by WIMAX/DVB-RCT/DVB-H. This type of terminal is needed to demonstrate mobility – even at high speed - , but also as well as test and validate both vertical and horizontal handover. 1)

A dedicated WIFI terminal receives rate-adapted content via a WLAN 802.11g connection from a deployed WIMAX/DVB-T-gateway. Dependent on available QoS of the WLAN connection between gateway and WIFI terminal the content will be rate-adapted. This terminal is suited to demonstrate scalability and low delay.

Thirdly, an MHP-IPTV end user terminal will be deployed in SUIT to demonstrate MHP applications and content on demand remote-interactively retrieved from a playout server and conveyed over WIMAX.

Finally, we may design a particular terminal to demonstrate the LDPC codes as well as UPA in DVB-T/H systems in case the simulation results from WP3 show a better performance of LDPC.

1) vertical handover is service handover between different networks

horizontal handover is service handover within the same (type of) network

2 WIMAX/DVB-T User Terminal Requirements

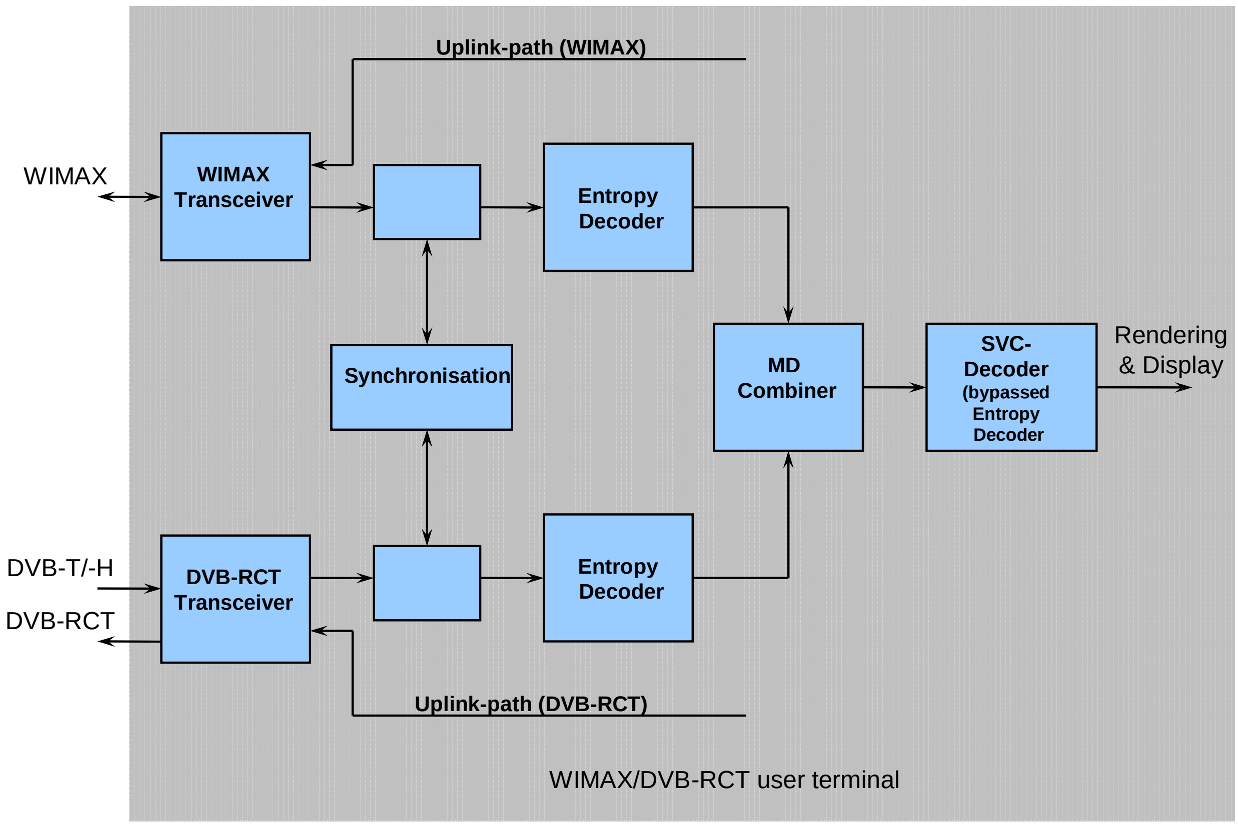
2.1 Introduction

This type of user terminal is needed for a direct reception of content via WIMAX, DVB-T and DVB-H. Return channel is via DVB-RCT , and interaction channel is via WIMAX. Multiple descriptions of the scalable video bitstream will be received by one or more different network interfaces. Besides that, the interaction between the terminal and the SUIT playout will be performed directly via the return/interaction channels. This type of terminal is especially dedicated to demonstrate mobility, also at high speed. Vertical and horizontal handover (e.g. from one WIMAX-sector to another or from DVB-T to WIMAX)

For the use in SUIT, the terminal has to meet in particular a set of general-purpose requirements:

2.2 Modules

In picture 1 below a simplified diagram (video data paths) of the WIMAX/DVB-T end user terminal is depicted:



Picture 1: Data path of the WIMAX/DVB-RCT user terminal

The terminal is characterised by :

- The transceivers for the simultaneous reception/transmission of WIMAX and DVB-T/DVB-RCT signals

- A Synchronisation unit performing time-synchronisation of IP-data received by different network paths
- Entropy decoders (one each network path)
- MD combiner for the synchronisation and combination of the multiple descriptions arriving by different transmission paths
- SVC decoder for decoding of scalable H264 content (alternatively, an SVC/AVC transcoder plus AVC decoder if required might substitute this module)
- Rendering block for display of the video content on different display sizes

In the WLAN environment (see chapter 3) SUIT will make use of a gateway with similar architecture shown in picture 1. However, the SVC decoder is replaced by an entropy encoder.

2.3 Radio interface(s)

Dependent on the network scenario to be demonstrated, the terminals require appropriate radio interfaces. In particular for handover tests one (or more) of the following network interfaces have to be supported. For vertical handover (from WIMAX to DVB-T and vice versa) both WIMAX and DVB-T interfaces needs to be installed in the terminal.

2.3.1 WIMAX interface

IRT will setup the testbed and carry out the execution of field trials in Munich/Germany. The transmission tests require testing licences, for which IRT applied at the German Federal Networking Agency.

In July 2006, IRT could acquire two time restricted WIMAX testing licences for following frequencies:

- 3,4 to 3,6 GHz band, time restricted till end of 2006, possibly with extension
- frequency 2,52 to 2,67 GHz, currently time restricted till end of 2007

The German Federal Networking Agency will make a decision to extend the licence for the 3,5 GHz band not before a currently running award procedure (licence sell by auction) will have been finished.

Therefore, at this point in time, it should be envisaged the support for both above mentioned frequency bands till a final decision can be taken.

2.3.2 DVB-T/DVB-RCT/DVB-H interface

Following standards need to be supported:

- DVB-T standard compliance to ETSI EN 300 744 V1.5.1 (2004-11)
- DVB-RCT interaction channel standard compliant to ETSI EN 301 958 V1.1.1 (2002-03)
- DVB-H standard compliant to ETSI EN 302 304 V1.1.1 (2004-11) and ETSI EN 300 744 V1.5.1 (2004-11)

2.4 Video Decoder

Capability of SVC decoder to bypass the entropy decoder

An essential requirement for the SVC decoder is the facility to bypass optionally the entropy decoder (reverse entropy) part of the SVC decoder. The reason for this is the fact that entropy decoding will be performed at another stage before combination of the MD-streams. Consequently, the SVC decoder gets a bitstream from the MD combiner already entropy decoded. In this case, the entropy decoder as an integrated part of the SVC decoder shall be bypassed, otherwise the bitstream needs to be entropy encoded before complete decoding.
Due to several reasons, in particular compatibility reasons it is purposed to deploy the same SVC decoder on both WIMAX/DVB-T and WIFI end user terminal. However, in contrast to the WIMAX user terminal, the WIFI user terminal will receive (via the gateway) a bitstream that is entropy encoded. In this case, entropy decoding of the SVC decoder must not be bypassed.

Types of scalability

To the extent possible, the deployed SVC decoder should be capable to support 3 types of scalability of the video bitstream:

- Temporal Scalability
- Spatial Scalability
- SNR (Quality) Scalability

A scalable video bitstream contains a base layer and one or more enhancement layers. An enhancement layer may enhance the temporal resolution (i.e. the frame rate), the spatial resolution, or the quality of the video content represented by the lower layer or part thereof.

The end user terminals should be capable to support scalable video bitstreams containing the base layer, and to the extent possible, one or more enhancement layers. This might even require external hardware accelerators if applicable.

Terminals with restricted resources and without the possibility for a potential hardware acceleration (e.g. PDA's) should be at least capable to process the base-layer of the SVC bitstream.

2.5 Rendering

To the extent possible, the graphic processor should be capable to support techniques like DirectX Graphics, OpenGL or SDL in order to keep the required computational expenses for rendering and display low. Hence, the video decoder needs to support the applied technique as well.

2.6 Protocols / Encapsulation & Transport

Due to the hybrid nature of the terminal, synchronization will be required in order to demonstrate horizontal and vertical handover. Streams received from different platforms/cells must be synchronised to allow vertical/horizontal handover.

Synchronisation requirements:

- Synchronise DVB-T and WIMAX streams from the same content. Both streams will be IP based, thus RTP synchronisation should be used.
- The study of the necessity to synchronise between descriptions in the same channel has to be promoted.
- The terminal must be capable to change between platforms in a short period of time. This means perfect synchronization between streams. This issue could be achieved using the RTP timestamps together with network delay estimations.

The terminal encapsulation requirements will be conditioned to the encapsulation protocol used by the intelligent multiplexer in the transport layer. In fact, due to the final protocol encapsulation to use hasn't been chosen yet, some suppositions must be done.

DVB-T/RCT:

- Be able to decode H.264 SVC/ RTP/UDP/IP packets over TS (MPE EN302304) . Other possibilities may be considered like generic packets (DVB-S2) depending on the DVB-T/RCT transceiver capabilities.
- Decode other additional content: hyperlinks, html pages, MHP applications, maybe MPEG21 not encoded in RTP. In this context some protocols like FLUTE (RFC 3926) or DSM-CC (ISO/IEC 13818 Mpeg-2 Systems Part 2) should be promoted.
- Should be able to perform service selection. Interpretation of service announcement and session description protocols must be done.
- The use of GSE protocol (from DVB-S2 specification) instead of MPE in the transport layer should be studied. This protocol is still in development phase, it may be used in the SUIT context as it may also be used in future DVB-T2 system.

WIMAX:

- The terminal will receive MPEG-4 AVC/H.264 SVC/RTP/UDP/IP packets through the 802.16e
- Probably some buffer cache would be needed to store the received packets before to be sent to the terminal decoder.
- Should be able to perform service selection. Interpretation of service announcement and session description protocols must be done.

Besides above mentioned requirements, the “All-IP-requirements” defined in Activity A1.3 should be followed.

2.7 Protocols & QoS requirements

In general, it should be followed the QoS requirements defined in Activity A1.2

- The terminal should be able to notify to the intelligent multiplexer about the QoS in the video on demand operations in order to obtain the maximum quality bared by the terminal.
- MPEG-21 should be used in order to notify terminal requirements and QoS demanded to the intelligent multiplexer. This signalling has to be defined in activity 4.2

3 WIFI User Terminal Requirements

3.1 Introduction

This type of end user terminal will be used for the reception of scalable content coming via the gateway over a local WIFI connection. Besides that, the interaction between the terminal and the SUIT playout will be performed via the gateway too. This type of terminal is especially dedicated to demonstrate scalability.

For the use in SUIT, the terminal has to meet in particular a set of general purpose requirements:

3.2 Modules

The SUIT WIFI terminal consists roughly of following building blocks:

- WIFI front-end
- SVC decoder for decoding of scalable H264 content. Alternatively, an SVC/AVC trans-coder plus AVC decoder might substitute this module.
- The rendering block for display of the video content on different display sizes.

3.3 Radio interface(s)

The radio interface (WLAN) shall be standard compliant to IEEE802.11g

3.4 Video Decoder & Rendering

In order to avoid an increasing risk of in-compatibility between the encoding and decoding side (SVC at point of time is still not standardised Mid of 2006), it is intended to deploy the same decoder on different terminal types. In contrast to the WIMAX/DVB-T terminal - where the SVC decoder shall be capable to process a bitstream already entropy decoded - the SVC decoder of the WIFI terminal has to decode a complete encoded bitstream (including entropy) from the gateway. Thus, as a strong requirement, the SVC decoder shall have the facility to bypass optionally the entropy decoder (WIMAX/DVB-T terminal) or not (WIFI-terminal).

3.5 Protocols / Encapsulation & Transport

Should follow the All-IP-requirements to be defined in Activity A1.3

In this case, the synchronization and rate shaping is performed in the gateway; the WIFI end user terminal receives RTP/IP packets with the final video and data streams to be decoded. In this context, there is no horizontal handover.

Encapsulation requirements:

- The terminal has to be able to decode the RTP packets in order to decode the information transported inside.
- The terminals may have to deal with RTCP signalling in order to synchronise the RTP packets. Besides, the gateway may use RTCP information to estimate the WLAN status and, if necessary, could request the playout for reducing the stream bitrate.

Service discovery requirements:

- The terminal should be able to perform service selection. Interpretation of service announcement and session description protocols must be done.

3.6 Protocols & QoS requirements

In principle, QoS requirements to be defined in Activity A1.2 should be met.

The terminal has to be able to decode video and data components, with enough buffer cache to ensure the correct throughput to the video decoder.

MPEG-21 will ensure the correct QoS exchange between the gateway and the terminal, notifying to the gateway about the terminal requirements and capabilities.

4 Computational Capabilities

4.1 Complexity of Multiple Description Coding (MDC) versus Single Description Coding (SDC)

In order to perform a fair comparison between the MDC and SDC the same overall rate is assumed in both cases. In other words if we consider that the SDC rate is R, and the rate corresponding to each of the descriptions in the MDC case is R_1 and R_2, than $R = R_1 + R_2$. It is obvious that in this case the number of decoded bits is the same in the SDC and MDC case and there is no computational expense at the level of the video decoder. The only complexity is added by the MD combiner module which is basically a look up table and additional constraint rules. Hence, the complexity can be considered minimal having in mind the overall complexity of the SVC decoder.

In the following table are given the operations that have to be performed by an MDC system in comparison to an SDC system.

	MDC	**SDC**
Depacketization	Yes	Yes
Identification of missing data due to erasures	Yes	Yes
Discarding of unusable data	Yes	Yes
Synchronization of the remaining packets (out of order arrival)	Yes	Yes
Synchronization of the separate streams	Yes	No
Entropy decoder(s)	Yes *(Note 1)*	Yes
MD combiner	Yes	No
SVC decoder (bypassed entropy decoder)	Yes *(Note 2)*	Yes

Table 1: Operations needed on a MDC and SDC decoding system

It has to be noticed that this complexity comparison does not include the hardware architecture, which in MD case includes two different receivers (DVB and WiMax).

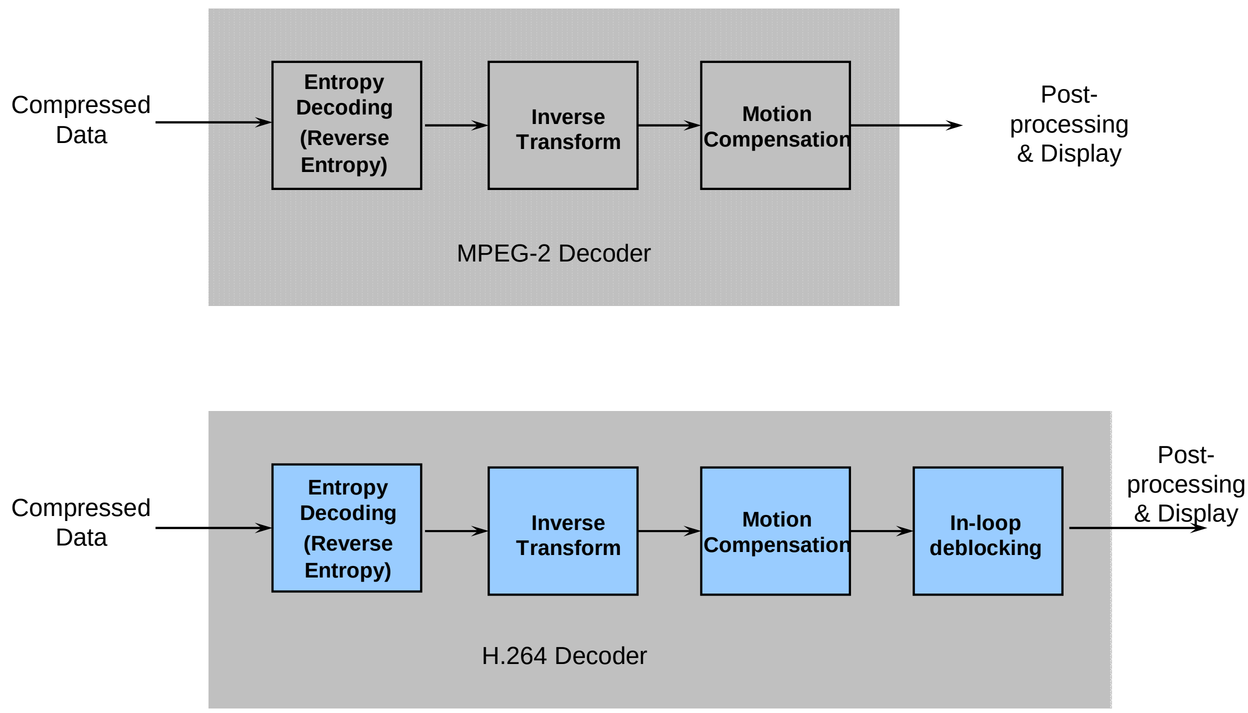
Note 1: One entropy decoder on each network path

Note 2: An entropy encoder (CABAC) after the MD combiner is solely avoidable in case the entropy decoder as an integrated part of the SVC decoder allows to bypass its entropy decoder part of each conventional H264/SVC decoder.

4.2 General operating considerations

4.2.1 Operating considerations on complexity of H.264/AVC

Schematic 2 depicts the main differences between H264 and MPEG-2 decoding.



Picture 2 : Main differences between MPEG-2 and H264 video decoding

The improvement of video coding technology has led to launch new coding schemes that allow to save more bandwidth but at the price of an increasing computing power consumption. Unfortunately, algorithm complexity does not evolve linearly with bandwidth savings. For instance, the MPEG-4 Industry Forum has made a comparison between MPEG-2 and MPEG-4 AVC/H.264 coding efficiencies and the corresponding decoding complexity increase. Results are gathered underneath in a table.

H.264/AVC Profile	Efficiency related to MPEG-2	Decoder complexity increase related to MPEG-2
Baseline (BP)	1.5 times	2.5 times
Extended (XP)	1.75 times	3.5 times
Main (MP)	2 times	4 times

Table: H.264/AVC and MPEG-2 comparison

The additional computational complexity of H264 related to MPEG-2 is widespread in the overall decoding process, in particular in three key techniques areas entropy coding (CABAC), smaller block size and In-loop de-blocking.

Beside above-mentioned Profiles, a “High Profile” (HiP) quite similar to the Main Profile is defined. The High Profile is the primary profile for broadcast and disc storage applications, particularly for high-definition television applications.

For real-time decoders or decoders with constrained memory size, it is important to specify processing power and memory size. Picture size plays a main role in influencing those parameters needed for implementation.

In addition to the Profiles, MPEG defined so-called Levels. For H264, in particular 16 Levels are defined, tied mainly to the picture size also provide constraints on the number of reference pictures and the maximum compressed bitrate.

4.2.2 State of AVC implementation

Contrary to the USA, it has been chosen in France to broadcast HDTV programmes using MPEG-4 AVC encoding standard in place of MPEG-2. The objective was to prevent loosing too much channels during the migration of DTV from SDTV to HDTV. In this aim, national experiments have been launched for testing HDTV broadcast at a metropolitan scale over three main towns in France during some well-known sport events like Roland Garros international tennis tournament and FIFA world football cup. Concerning the first event, two national telecom operators, France Telecom and Free, have deployed new IP-STB that are able to decode in real-time SD/HD TV programmes encoded using MPEG-4 AVC standard. In Roland Garros, France Television captured tennis matches with 1080i HDTV cameras. France Telecom broadcasted video streams encoded at 13 Mbs although Free did it with a bit-rate divided by two in order to reach more people connected to an ADSL link, but also for offering an access to people owning a sufficiently powerful computer at home. For them, VideoLan has developed optimized software that can decode and render HDTV streams encoded in half resolution (raw sub-sampling) using MPEG-4 AVC standard (http://downloads.videolan.org/pub/videolan/testing/freehd/win32/). Currently, it must be considered as an upper limit for decoding and rendering HDTV programmes by software, remembering that 720p or 1080i HDTV programmes get a spatial resolution five times larger than an SDTV stream.

So it appears that without using a dedicated integrated device, it is presently difficult to render by software HDTV programmes. The path that leads to an integrated device is starting from an initial software implementation where it is tried to simplify the algorithm as possible. The following step is an implementation on a set of DSPs which organization is tailored to achieve real-time processing after having optimized code on a single DSP. Optimization is carried on by moving progressively intensively computing tasks on FPGAs that first acts as co-processors. When the whole algorithm stands on FPGAs, an ASIC can be developed. This last step is only performed for consumer electronic products; concerning professional markets the process stops at the boundary of DSPs and FPGAs. For instance, STBs provided by Free are designed around the SMP 8630 media processor commercialized by Sigma Design that can decode SD/HD audiovisual programmes encoded using MPEG-2/VC1/H.264 standards. Before the availability of such chips, professional STBs were mainly built around GPP or DSP chips like the TMS320DM642 from Texas Instruments designed to satisfy most of multimedia applications. But without the help of complementary devices, it could only decode H.264 (MP@L3) SD video streams in a range of bit-rates starting from 500 Kbs up to 2 Mbs. This short range is mainly due to the complexity of the CABAC entropy coding tool provided in Main Profile.

4.2.3 Operating considerations on complexity of SVC

The JVT group is working on a scalable version of H.264, which would have close similarities with H.264/AVC video coding standard. The current working draft combines the coding primitives of H.264/AVC with an open-loop coding structure allowing to merge together spatial, temporal and quality scalabilities:

- spatial scalability using a Laplacian pyramid to compute different spatial layers;
- temporal scalability is implemented through a hierarchical prediction picture organisation;
- quality scalability is incorporated through two possible different mechanisms (CGS or FGS).

Such a scheme allows handling of single content described at different resolutions, different frame rates and different bit-rates by removing redundant information using inter-layer prediction techniques, as described in table 2.

Format	Bit rates (kbit/sec)				
QCIF 15Hz	96	112	128	160	192
CIF 7.5 Hz	192	224	256	320	384
CIF 15 Hz	256	320	384	448	512
CIF 30Hz	384	448	512	640	768
4CIF 15Hz	768	896	1024	1280	1536
4CIF 30Hz	1024	1280	1536	1792	2048
4CIF 60Hz	1536	1780	2048	2560	3072

Table 2: Scalability grid used for testing SVC algorithm

Three different kinds of prediction techniques can be used and mixed in order to reach better compression rates, but mostly with the price of an increased complexity at the decoding stage. These three kinds of prediction techniques can be used either for spatial or for quality scalabilities. Together combined they are a main source of complexity increase because they may involve multiple decoding loops to reconstruct pictures. Combining scalabilities produces a complexity that is evolving as the cross product of the complexities involved in each scalable dimension. So, nowadays (July 2006) it would be hardly expected to decode by software SVC video streams of a resolution higher than SD. Trans-coding from SVC to AVC could be then a means for lowering computing load at decoding time by distributing computing effort. It should be also a convenient way to prepare the adoption of SVC standard by the industry.

4.2.4 Decoding complexity increase for three layers of scalability (H.264/SVC) versus H264/AVC and MPEG-2

Assuming three layers of spatial scalability, one base layer with 176x144 and two enhancement layers with 352x288 and 704x576 pixel, the base layer (176x144) is decoded as is. The first enhancement layer (352x288) is decoded after differentiating it with the reconstructed base layer. The second enhancement layer (704x576) is decoded after differentiating it with the previous decoded layer (352x288). Indeed, the enhancement layers (704x576, 352x288) requires as much processing power as if they would be decoded “as is”. The CPU load to decode 704x576 with 3 layers of spatial scalability is at least 1.5 to 2 times the CPU load for processing 704x576 only.

➔ it is estimated that the decoding of <u>3 layers of spatial scalability</u> in H264 SVC (Main Profile) will require approximately 6 to 8 times the processing power compared to MPEG-2!

Taking into consideration three layers of SNR scalability at resolution 704x576, there is no rescaling process needed like for spatial scalability. However, the CPU load for processing is 3 to 4 times higher than decoding of 704x576 AVC.

➔ it is estimated that the decoding of <u>3 layers of SNR scalability</u> in H264 SVC (Main Profile) requires approximately 12 to 16 times the processing power compared to MPEG-2!

Concerning temporal scalability, this feature is already implemented in AVC is a part of the decoding process. This should not take more time than decoding a normal AVC with B-frames.

➔ it is expected that the decoding of <u>3 layers of temporal scalability</u> in H264 SVC (Main Profile) requires approximately 4 times the processing power compared to MPEG-2!

4.2.5 Experiments using AVC and SVC software platforms

In order to clarify figures concerning software decoding of SVC streams, some experiments have been made to show what can be obtained with current JSVM platform (v5.8) and what can be expected by deriving an SVC encoding/decoding platform from an already existing optimized AVC software.

During these experiments it has been used a PC computer powered by an Intel dual core processing running at 3GHz and equipped with 1 GB memory. The video stream to be encoded and decoded is an SDTV sequence of 150 frames lasting 6 seconds (25 fps). It is encoded at fixed quantifier (QP = 25, so as to provide a 2Mbs AVC-encoded stream) without any R-D optimization (CBR/VBR).

Concerning encoding, the results are gathered in the table listed underneath (more information about AVC encoding software can be found at http://developers.videolan.org/x264.html).

Coding context	**(SVC) JSVM**	**(AVC) x264**
Single spatial layer 176x144 (QCIF)	20 s	1.1 s
Single spatial layer 352x288 (CIF)	85 s	3.9 s
Single spatial layer 704x576 (4CIF)	366 s	13.0 s
3 spatial layers from QCIF up to 4CIF	958 s	N/A
Single spatial layer 4CIF with 3 temporal ones	1596 s	10.5 s
Single spatial layer 4CIF with 3 quality (FGS) ones	482 s	N/A

Table 3: Decoding speed comparison between JSVM and x264 AVC encoders

For temporal scalability, it has been found a similar scheme under x264 that has allowed to check that B-Hierarchical prediction performances are mainly depending on the quality of the implementation. Concerning decoding, the same table can be drawn as shown below (more information about AVC decoding software can be found at http://ffmpeg.mplayerhq.hu/).

Decoding context	**(SVC) JSVM**	**(AVC) ffmpeg**
Single spatial layer 176x144 (QCIF)	1.3 s	0.25 s
Single spatial layer 352x288 (CIF)	3.8 s	0.67 s
Single spatial layer 704x576 (4CIF)	13.0 s	1.97 s
3 spatial layers from QCIF up to 4CIF	37.0 s	N/A
Single spatial layer 4CIF with 3 temporal ones	17.0 s	N/A
Single spatial layer 4CIF with 3 quality (FGS) ones	65.0 s	N/A

Table 4: Decoding speed comparison between JSVM and ffmpeg AVC decoders

So, it can be checked that former estimations about the complexity increase of SVC decoding can be experimentally retrieved. Concerning decoding HD streams, computational times can be extrapolated using curve fitting techniques and it provides the following figures:

HD formats to decode	(SVC) JSVM	(AVC) ffmpeg
720p25	26.0 s	3.6 s
720p50	52.0 s	7.2 s
1080i25	47.0 s	6.2 s

Table 5: HD decoding speed extrapolation and comparison between JSVM and ffmpeg AVC decoders

In Table 5, the stream transported only one layer, despite being decoded by JSVM. All these results show that at average the SVC reference model is currently 20 times slower than an optimized AVC software encoder and 7 times slower than an optimized AVC decoder. In this later case, 720p50 HD streams can hardly be decoded on a Pentium based computer. Therefore, a 3.6 GHz Pentium is needed to decode an H.264 720p50.

4.2.6 Embedding multiple terminal resolutions into a single video stream

Currently the JVT team has focused its work on the embedding of spatial resolution starting from QCIF up to 4CIF in order to address mobile, handheld and standard television formats.

It will favour 4:3 screen ratio at video capture, but 16:9 ratio can easily emulated by reshaping pictures. Editing frame rate could be used in place of SNR truncation to adapt bit-rate. Frame rate editing can be used in complement to transrating in order to reach lower bit rates by dropping frames from an input stream. Concerning SVC streams, it is possible play simultaneously on temporal and SNR scalabilities to adapt to lower bit rates. Moreover, concerning handheld devices, AVC decoding can replace SVC decoding because the base layer can is fully compliant with the AVC syntax.
These modes will be implemented with the spatial and temporal scalability proposed in SVC:

- temporal:

  Though SVC enables the encoding of complex hierarchical B-frames, it is proposed to use first a simple B-Frame scheme to demonstrate temporal scalability, the lowest quality layers being the upper layer without the B-Frames.

- spatial.

  There is the need to keep the H.264 compatibility for the first layer. The following layer can be spatial scalability layers with simple prediction, new intra prediction and inter prediction with motion prediction; the residual prediction is a new feature of SVC, but does not improve significantly the quality and has an important processing power and development cost.

SUIT project is intending to tackle high definition video formats. In such conditions, a straightforward stack of resolutions could then start from a handheld format, going across standard television and reaching high definition by implementing the following spatial layers. There are four diferent video formats suitable to be used in SUIT:

Option 1:

1920 (1440) x1080p 16:9 - 50 Hz, 25 Hz

960 (720) x 540p- 50 Hz, 25 Hz

480 (352) x 270p - 25 Hz, 12.5 Hz

Option 2:

1280x720p 16:9 -50 Hz, 25 Hz

640 x 360p - 25 Hz, 12.5 Hz

320 x 180p - 25 Hz, 12.5 Hz

Option 3:

1280x704p 16:9 -50 Hz, 25 Hz

640 x 352p - 25 Hz, 12.5 Hz

320 x 176p - 25 Hz, 12.5 Hz

Option 4:

1408x1152p 4:3 -50 Hz, 25 Hz

704 x 576p - 25 Hz, 12.5 Hz

352 x 288p - 25 Hz, 12.5 Hz

However, as most of the HDTV displays and cameras support 720 p and it is progressive, we feel that the best option could be 1280x704, option 3, because this format is close to 720p and is an integer multiple of 16*16 blocks. We are not considering any interlaced format.

In SVC, there is no conceptual difference between encoding for a new spatial layer and for an SNR (quality enhancement) layer. Thus, SNR layer can be provided following the same encoding scheme. From a real-time point of view, each SNR layer costs roughly as much as its base layer.

5 MHP IPTV Terminal

5.1 Introduction

An MHP-IPTV end user terminal will be deployed to test and to demonstrate iDTV applications based on MHP on the SUIT experimental platform. In contrast to the above described terminals (chapter 2 and chapter 3) the MHP IPTV terminal is not foreseen to cope with the different SUIT scenarios like handover and high speed, but to demonstrate remote interactive low latency TV-applications conveyed over WIMAX.

The requirements can be split in two categories:

1) Broadcast: Signalling, Main video/audio, announcement & transportation of MHP applications.
   Compliance to DVB-IPI phase 1 (ETSI TS 102 034 V1.1.1), (e.g. Video: H264 AVC/PES/MPEG-2 TS/(RTP)/UDP/IP).
2) Unicast: Video on demand capability (remote interactive) controlled within the MHP using the JMF-Interface.

5.2 Modules

In view of main video data processing the MHP IPTV terminal consists roughly of following blocks:

- Transceiver for the reception/transmission of WIMAX signals
- SVC decoder for decoding of scalable H264 content (alternatively this module might be substituted by an SVC/AVC transcoder plus AVC decoder if required).
- The rendering block for display of the video content on different display sizes.

5.3 Minimum hardware requirements and computational capabilities

The minimum hardware requirements depend on the video codec for broadcasted video and for video clips used as additional content.

In general a standard PC (TabletPC, Notebook, DesktopPC …) is required with the following restrictions:

- the graphic board
    - shall support DirectX 9b
    - shall have at least 128MB RAM on board

		MPEG-2	MPEG-4 AVC (H.264)	SVC based on AVC
SDTV	processing power (x86 CPU)	> 2 GHz	> 2.5 GHz	1)
	main memory	> 512 MB	> 512 MB	
	video board memory	> 128 MB	> 128 MB	
HDTV	processing power (x86 CPU)		> 3.2 GHz Dual Core	1)
	main memory		> 2 GB	
	video board memory		> 512 MB	

Table 6: Required processing power (x86 CPU), main memory and video board memory

1) SVC for the MHP-IPTV will be explored, but might not be supported due to technical restrictions (requirements of the graphic board), but also due to the computational limitations (MHP itself already consumes PC-resources).

5.4 Software requirements

The DVB-MHP and DVB-IPI receiver software from IRT will be used for it. Therefore, the terminal requirements will be based on a Microsoft Windows XP PC platform.

5.4.1 Video codec and container format

The current DVB-MHP receiver implementation is based on MPEG-2 Video compression (ISO/IEC 13818-2) carried in a MPEG-2 transport stream (ISO/IEC 13818-1). It will be extended to support MPEG-4 AVC (ISO/IEC 14496-10) including HDTV formats carried in MPEG-2 TS.

SDTV formats	HDTV formats
576i25	720p50
	1080i25

Table 7: Required SD and HD formats

The support of SVC in the MHP-IPTV platform will be investigated. As its architecture is based on the Microsoft Direct Show API, i.e. the SVC codec must have implemented the required interfaces from that API. A potential restriction for the support of SVC might be computational restrictions, due to the fact that the MHP itself absorbs already processing power.

5.4.2 Audio codec

MPEG-1 Audio layer 2 and/or MPEG-4 High Efficiency AAC Profile Level 2 (mono or 2-channel-stereo), as specified in ISO/IEC 14496-3 and ISO/IEC 14496-3/AMD-1 should be supported.

5.5 DVB-IPI requirements

The SUIT IPTV terminal shall be compliant to DVB-IPI, i.e. it complies with ETSI TS102034, with the modifications and restrictions defined in this section.

5.5.1 Physical layer connection

In addition to TS102034, the terminal will use WIMAX (IEEE802.16e) and/or WLAN (IEEE 802.11g) network interfaces for connecting the user terminal to the delivery network.

5.5.2 IP layer configuration

The DVB-IPI phase 1 only supports IPv4. IPv6 is a topic for later stages in the IPI standard. Currently there is a discussion on how to deal with IPv6 in future IPI specifications including migration scenarios from IPv4 to IPV6. For the time being the terminal is only required to implement IPv4.

The DVB-IPI standard suggests the usage of DHCP for the configuration of parameters like host IP address, addresses of DNS servers and as a speciality of DVB-IPI the addresses of *DVB service discovery entry points*. An entry point is an address where a terminal finds the service discovery information, in IPI known as SD&S records. Because there are other possibilities to look up for an entry point, the network does not have to support this mechanism.

The terminal will be configured manually. It means that the network is not required to support DHCP configuration.

5.5.3 Service announcement

In DVB-IPI services are announced by either multicast or unicast delivery of SD&S (service discovery and selection) records. The records are found by looking for entry points. Though the terminal shall support the methods as defined in the standard, the service discovery will be limited for the provisioning network to the subsequent sequence.

The terminal shall support the unicast and multicast delivery of SD&S records and look up for the entry points starting with step 1. (See also TS 102 034 section 5.2.4)

1. For unicast delivery the entry point address shall be acquired via DNS. The lookup will be *_dvbservdsc._tcp.services.dvb.org.* The port number will be 3937. If the lookup does not resolve into a valid address the terminal shall go to the next step.
2. For multicast delivery the terminal shall join the multicast group with the address 224.0.23.14 (DvbServDisc) and the port number 3937 (dvbservdscport) as assigned by IANA. If the terminal does not receive any valid packet within the time limit, as it has been specified in IPI, the terminal shall go to the next step.
3. The entry point shall be configured by the user.

5.5.4 Service selection

The terminal shall use the SD&S records as defined in TS 102 034 5.2.5 and 5.2.6
The terminal shall read the optional CodingDescriptor from the ServiceInformation record. The CodingDescriptor has been proposed in the draft of TS 102 034 V1.2.1. It is based on the ComponentDescriptor as it is defined in EN 300 468 V1.7.1 section 6.2.8 and table 26.

5.5.5 Content streaming

The terminal shall receive MPEG-2 transport streams via UDP multicast by joining multicast groups using IGMPv3. The multicast group is signalled in the SD&S records. The terminal shall support video compression formats as specified in section 5.4.1 and audio formats as specified in 5.4.2.

5.6 Requirements for interactive TV in IPI

In order to support interactive TV the MHP IPTV terminal shall support MHP1.0.2 (ETSI TS101812 V1.2.1). As MHP1.0.2 does not have all capabilities necessary for interactive service scenarios in SUIT, the terminal shall match the following additional requirements:

- Application loading over the interaction channel as defined in MHP1.1 (ETSI TS 102 812 V1.2.1) sections 6.4 and 10.8.1.3
- MPEG-4-AVC has to be added as a content format (see also ETSI TS101 812 section 7)
- For streaming of AVC the terminal shall support the protocols RTSP and RTP

6 Common terminal description protocols

6.1 Introduction

In the previous sections, the three types of user terminal will be designed and implemented in order to evaluate and to demonstrate the SUIT concept are described. Although these descriptions are structured in a similar fashion, other tools in the SUIT project require a more strict (“closed”) definition of the terminals. Furthermore, the use of a non-proprietary, well-defined, and preferably standardized solution ensures interoperability not only within the SUIT project, but also with parties outside the consortium.

In this section, two terminal description protocols are discussed that allow describing the characteristics and capabilities of a terminal in a standardized way. Terminal descriptions are a subpart of the larger “usage context description” concept that also comprises of the user preferences and the network characteristics. The usage context description is described in more detail in Activity 4.2 / Deliverable 4.2.

Now, only a few terminal description vocabularies are available. Here, W3C’s Composite Capability – Preference Profile (CC/PP) and the MPEG-21 Part 7 Digital Item Adaptation – Usage Environment Description (MPEG-21 DIA-UED) tools are discussed.

6.2 W3C’s composite capability – preference profile

CC/PP is a W3C recommendation of its Device Independence Activity group. The task of this group is “to allow access to the web by any Internet-enabled device.” CC/PP is a framework that enables the creation of CC/PP profiles, which describe device capabilities and user preferences. It does not define a vocabulary itself, but allows a third party to use this framework to create one.

A CC/PP profile is structured as a two-level hierarchy:

- Main components: a major branch in the profile, such as a hardware and a software component.
- Attributes: detailed information logically associated with a component. For example, screen size as part of the hardware component.

The combination of the components and their attributes creates the vocabulary of a CC/PP profile. Apart from the vocabulary, the profile defines its syntax using RDF, either in graph notation or by an XML Schema.

Notwithstanding the flexibility of the framework, only a few CC/PP profiles are created to date, such as the User Agent Profile (UAProf) defined by the Open Mobile Alliance (OMA), formerly known as the Wireless Application Protocol (WAP) Forum. This profile defines six main components – hardware, software, network, browser, WAP characteristics, and push characteristics, – each with various attributes. It is mainly intended to describe the capabilities of cell phones in a standardized way.

In case the SUIT consortium decides to use this terminal content description tool, a SUIT CC/PP vocabulary should be specified. If so, this will be described in detail in Activity 4.2 / Deliverable 4.2 (Session Description Protocols).

6.3 MPEG-21 Part 7 Digital Item Adaptation – Usage Environment Description

The MPEG-21 DIA-UED tool is intended to describe the complete usage context, not only the terminal capabilities. However, as it is structured in four distinct parts, with as second part the Terminal Capabilities, it can also be used to solely describe the user terminal. In addition, the Network Characteristics part is required in order to express the network capabilities of the device.

The other parts of the MPEG-21 DIA-UED specification – namely User Characteristics and Natural Environments Characteristics – are described more in detail in Activity 4.2 / Deliverable 4.2.

The Terminal Capabilities part is divided in ten main subparts:

- **Codec Capabilities**: specifies all encoding and decoding capabilities of the end-user device.
- **Display**: contains information on the display(s) of the end-user device, for example, resolution, colour capabilities, and refresh rate.
- **Audio Output Capabilities**: gives the number, type, characteristics, power, and other useful information on the speaker(s) of the end-user device.
- **User Interaction Possibilities**: enlists the possibilities and characteristics of the input possibilities of the end-user device, such as information on the keyboard, mouse, and microphone.
- **Device Class**: selects the type of end-user device from a list of possibilities.
- **Power Characteristics**: gives information on the battery status of the end-user device, such as power and estimated remaining time.
- **Storage Possibilities**: contains information on the storage capabilities of the end-user device, for example, available free space and transfer rate.
- **Data Input/Output Characteristics**: gives information on all I/O-devices, such as the number of devices and bus speed.
- **Benchmark Information**: stores information about the overall processing capability of the end-user device by a benchmark rating.
- **IPMP Tools**: gives an overview of the supported Intellectual Property Management and Protection (IPMP) capabilities of the end-user device.

The Network Characteristics part is divided in two subparts:

- **Network Capability**: stores static information about the theoretical characteristics of the network, namely minimal guaranteed available bandwidth, maximum available bandwidth, error rate, error correction capabilities, and guarantee for in-sequence delivery of packets.
- **Network Conditions**: contains dynamic information about the actual status of the network, in particular feedback on the delay, error rate, and available bandwidth.

The MPEG-21 DIA-UED specification defines the semantics – by normative text – and the syntax – by an XML Schema – of these subparts. By defining a strict semantic, all UED-compliant applications understand and interpret the information in a similar way. By defining an XML Schema, the context description can be stored and exchanged in a verifiable valid XML syntax. The combination of the two results in a very strict representation of the terminal capabilities.

6.4 Informative example of a possible description of a WIFI user terminal using CC/PP (UAProf)

The following is an informative example of a possible (brief) description of a WIFI user terminal using the CC/PP specification. The example is inspired on the User Agent Profile defined by OMA and has been simplified to improve readability. Real-world examples of (mobile phone) terminal descriptions are available at http://www.w3development.de/rdf/uaprof_repository/

```
<?xml version=”1.0”?>
<rdf:RDF xmlns:rdf=”http://www.w3.org/1999/02/22-rdf-syntax-ns#”
xmlns:prf=”http://www.openmobilealliance.org/tech/profiles/UAPROF/ccppschema-
20021212#”>
<rdf:Description rdf:ID=”WIFIterminal”>
 <prf:component>
  <rdf:Description rdf:ID=”HardwarePlatform”>
   <rdf:type rdf:resource="http://www.openmobilealliance.org/tech/
   profiles/UAPROF/ccppschema-20021212#HardwarePlatform"/>
   <prf:ScreenSize>320x240</prf:ScreenSize>
   <prf:Model>SUIT WIFI terminal</prf:Model>
  </rdf:Description>
 </prf:component>
 <prf:component>
  <rdf:Description rdf:ID="NetworkCharacteristics">
   <rdf:type rdf:resource="http://www.openmobilealliance.org/tech/
   profiles/UAPROF/ccppschema-20021212#NetworkCharacteristics"/>
   <prf:SupportedBearers>
    <rdf:Bag>
     <rdf:li>802.11g</rdf:li>
    </rdf:Bag>
   </prf:SupportedBearers>
  </rdf:Description>
 </prf:component>
 <prf:component>
  <rdf:Description rdf:ID="SoftwarePlatform">
   <rdf:type rdf:resource="http://www.openmobilealliance.org/tech/
   profiles/UAPROF/ccppschema-20021212#SoftwarePlatform" />
   <prf:CcppAccept>
    <rdf:Bag>
     <rdf:li>audio/aac</rdf:li>
     <rdf:li>video/mp4</rdf:li>
    </rdf:Bag>
   </prf:CcppAccept>
  </rdf:Description>
 </prf:component>
</rdf:Description>
</rdf:RDF>
```

6.5 Informative example of a possible description of a WIFI user terminal using MPEG-21 DIA-UED

The following is an informative example of a possible (brief) description of a WIFI user terminal using the MPEG-21 DIA-UED specification. The example is inspired on the examples of in ISO 21000-7:2004 specification and has been simplified to improve readability.

```
<DIA>
 <Description xsi:type="UsageEnvironmentType">
  <UsageEnvironmentProperty xsi:type="TerminalsType">
   <Terminal>
    <TerminalCapability xsi:type="DisplaysType">
     <Display id="PDA">
      <DisplayCapability xsi:type="DisplayCapabilityType">
       <Mode>
           <Resolution horizontal="320" vertical="240"/>
       </Mode>
      </DisplayCapability>
     </Display>
    </TerminalCapability>
    <TerminalCapability xsi:type="CodecCapabilitiesType">
     <Decoding xsi:type="AudioCapabilitiesType">
      <Format href=":ACF:4.3.2">
       <mpeg7:Name xml:lang="en">MPEG-2 Audio AAC Main Profile</mpeg7:Name>
      </Format>
     </Decoding>
     <Decoding xsi:type="VideoCapabilitiesType">
      <Format href=":VCF:2001:3.2.2">
       <mpeg7:Name xml:lang="en">
           MPEG-4 Visual Simple Scalable Profile @ Level 2
       </mpeg7:Name>
      </Format>
     </Decoding>
    </TerminalCapability>
   </Terminal>
  </UsageEnvironmentProperty>
  <UsageEnvironmentProperty xsi:type="NetworksType">
   <Network id="WLAN">
    <NetworkCharacteristic xsi:type="NetworkCapabilityType"
     maxCapacity="54000000"/>
   </Network>
  </UsageEnvironmentProperty>
 </Description>
</DIA>
```

7 Summing up SUIT Terminals Requirements

In the following sections, we use "shall" if a certain requirement is mandatory, and "should" if a certain requirement is desirable, but not necessarily required.

7.1 *Terminal types*

SUIT shall have three types of terminals, WiMAX/DVB-T/H/RCT, WIFI and MHP-IPTV as mentioned in Section 1.2.

7.2 *MDC support*

Terminal WiMAX/DVB-T/H/RCT shall include the MD Combiner as mentioned in Sections 2.2 and 2.3. Terminal MHP-IPTV will not support the MDC combiner since it will not include a DVB-T/H interface. Concerning the terminal WIFI, the gateway shall include the MD Combiner.

7.3 *SVC support*

Terminals WiMAX/DVB-T/H/RCT and WIFI shall include either the SVC decoder or SVC transcoder as mentioned in Sections 2 and 3 .

Terminal MHP-IPTV should include either the SVC decoder or SVC transcoder as mentioned in Section 2.

Terminals WiMAX/DVB-T/H/RCT and WIFI shall decode all video resolutions up to 1280x704p@ 25 Hz as defined in option 3 (see Section 4.2.6) which are close to HDTV, SDTV and CIF.

Terminal MHP-IPTV shall decode CIF and should decode SDTV and HDTV from the SVC stream in accordance to option 3 (see Section 4.2.6).

SUIT terminals should support other options defined in Section 4.2.5 namely 1280x704p@50 Hz.

SUIT terminals shall support two SNR layers for each spatio-temporal resolution. The enhancement layer should be FGS.

7.4 *Rendering*

Terminals WiMAX/DVB-T/H/RCT and WIFI shall render all video resolutions as defined in option 3 (see Section 4.2.6) which are close to HDTV, SDTV and CIF. Terminal MHP-IPTV shall render CIF and should render SDTV and HDTV in accordance to option 3 (see Section 4.2.6).

SUIT terminals should support other options defined in Section 4.2.6.

7.5 *Audio support*

All three terminals should support either MPEG-1 Layer I or II at bit rates defined in ETSI TR 101 154 V1.7.1. No support is required for neither multiple descriptions nor scalability.

7.6 *Network interfaces*

Terminal WiMAX/DVB-T/H/RCT shall include WiMAX and DVB-T interfaces as mentioned in Section 2. Terminal WIFI shall include an IEEE802.11g interface. Terminal MHP-IPTV shall include the WiMAX interface and should include a WIFI interface.

7.7 Transportation protocols

Terminals WiMAX/DVB-T/H/RCT and WIFI gateway shall be IETF compliant (RTP Payload Format for SVC Video). IP shall be the glue at the network layer. IP packets shall feed both networks, DVB-T/H and WiMAX. Terminal WIFI shall receive IP packets. Terminal MHP-IPTV should be DVB-IPI compliant, TS over IP. Table below show two possible bit rates scenarios depending on two different DVB-T/H multiplexers, IRT and IT.

DVB-T		WiMAX	
Service	Bit Rate (Mbps)	Service	Bit Rate (Mbps)
1 D SVC Real Time Broadcasting	6.25	2 D Real Time Broadcasting	6.25
1 D SVC Broadcasting	6.25	2 D Broadcasting (on QoS demand)	0;6.25
1 D Hyperlinked Video	0.5	2 D Hyperlinked Video	0.5
Internet		Internet	0-?
		Streaming	0.5-4.25
Total	13	Total	14-17.75

D= Description;

SVC= HD: 1280x704p-25 Hz (4.25 Mbps) ; SD: 640x352x25 (1.5 Mbps) ; CIF: 320x176x25 (0.5 Mbps)

Table 8: Network/services scenarios (13 Mbps IRT Multiplexer)

DVB-T		WiMAX	
Service	Bit Rate (Mbps)	Service	Bit Rate (Mbps)
1 D SVC Real Time Broadcasting	10	2 D Real Time Broadcasting	10
1 D SVC Broadcasting	10	2 D Broadcasting (on QoS demand)	10
1 D Hyperlinked Video	0.5	2 D Hyperlinked Video	0.5
Internet		Internet	0-?
		Streaming	0.5-8
Total	20.5	Total	21-28.5

D= Description;

SVC= HD: 1280x704p-50 Hz (8 Mbps) ; SD: 640x352x25 (1.5 Mbps) ; CIF: 320x176x25 (0.5 Mbps)

Table 9: Network/services scenarios (20.5 Mbps IT Multiplexer)

7.8 Handover

Terminal WiMAX/DVB-T/H/RCT shall support horizontal and vertical handovers. Terminal MHP-IPTV should support horizontal handover in WiMAX cells. Both terminals should support seamless handovers.

7.9 Terminal description and network characteristics support

All SUIT terminals in collaboration with the gateway shall inform the playout about its capabilities, particularly to join and maintain an unicast session, as well as the network characteristics. See Section 6.

7.10 Adaptability

The gateway shall dynamically adapt the bit rate according to the wireless LAN characteristics, namely in unicast sessions.

7.11 Implementation issues

All three terminals shall be implemented on a Desktop/Laptop. Terminal WIFI should also be implemented on a PDA supporting the base layer resolution, CIF.

8 Conclusions

This document gives an overview about the different types of end user terminals in SUIT and their requirements. Different types of terminals ranging from small mobile ones to in-house devices will be implemented in SUIT in order to validate and to demonstrate various reference scenarios, defined later in Activity 1.4. However, due to the early point in the project phase and the relationship to other Activities/Work packages - in particular taking into account different possible solutions in WP5 starting on a later stage - the specified requirements of this document should be considered as “general” or “high level” requirements but without addressing in detail all requirements for each type of terminal.

9 Acronyms

BW	Bandwidth
CIF	Common Intermediate Format
CABAC	Context Adaptive Binary Arithmetic Coding
DVB	Digital Video Broadcasting
DVB-IPI	Digital Video Broadcasting- Internet Protocol Infrastructure
DVB-H	Digital Video Broadcasting-Handhelds
DVB-RCT	Digital Video Broadcasting- Return Channel Terrestrial
DVB-T	Digital Video Broadcasting-Terrestrial
HDTV	High Definition Television
FGS	Fine Grain Scalability
IGMPv3	Internet Group Message Protocol Version3
IETF	Internet Engineering Task Force
IP	Internet Protocol
IRT	Institut für Rundfunktechnik
IT	Instituto de Telecomunicações
JMF	Java Media Framework
JSVM	Joint Scalable Video Model
MDC	Multiple Description Coding
MHP	Multimedia Home Platform
OpenGL	Open Graphics Library
PDA	Personal Digital Assistant
QoS	Quality of Service
QCIF	Quarter Common Intermediate Format
SDC	Single Description Coding
SDL	Simple DirectMedia Layer
SDTV	Standard Definition Television
SD&S	Service Detection & Selection
SVC	Scalable Video Coding
UDP	User Datagram Protocol
WiBro	Wireless Broadband
WIFI	Wireless Fidelity
WIMAX	Wireless Local Area Network
WLAN	Wireless Local Area Network