ETSI TR 126 903 V17.0.0 (2022-05)

TECHNICAL REPORT

Universal Mobile Telecommunications System (UMTS);
LTE;
5G;
Improved video support for Packet Switched Streaming (PSS) and Multimedia Broadcast/Multicast Service (MBMS) Services
(3GPP TR 26.903 version 17.0.0 Release 17)

Reference

RTR/TSGS-0426903vh00

Keywords

5G,LTE,UMTS

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00   Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Legal Notice

This Technical Report (TR) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities. These shall be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between 3GPP and ETSI identities can be found under http://webapp.etsi.org/key/queryform.asp.

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

1 Scope

The present document provides an analysis of the future video capability requirements of streaming and multicast/broadcast services. The purpose of this document is two-fold. On the one hand, it studies the options to upgrade the minimal requirements for video reception and decoding. On the other hand, it studies use cases for support of more advanced UEs. The ultimate target of this study item is to recommend solutions for efficiently providing video support commensurate with UE and user capabilities and needs in PSS and MBMS services.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TS 26.346: "Multimedia Broadcast/Multicast Services (MBMS); Protocols and Codecs".

[2] 3GPP TS 26.234: "Transparent End-to-End Packet Switched Streaming Service (PSS); Protocols and Codecs".

[3] ITU-T Recommendation H.264 (03/09), "Advanced video coding for generic audiovisual services" | ISO/IEC 14496- 10:2009 Information technology—Coding of audiovisual objects— part 10: Advanced Video Coding".

[4] T. Schierl, Y. Sanchez de la Fuente, C. Hellge, and T. Wiegand: "Priority-based Transmission Scheduling for Delivery of Scalable Video Coding over Mobile Channels," 3rd European Symposium on Mobile Media Delivery (EUMOB), London, 2009.

[5] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in TR 21.905 [x] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in TR 21.905 [5].

advanced terminal: a user equipment terminal that permits rendering video at higher quality than common for UEs in the market at time of generation of the document. Typically display resolutions or VGA or higher are supported. Examples are netbooks or laptops.

baseline terminal: a user equipment terminal that permits to render video at typical quality as available in the market at time of generation of the document. Typically display resolutions of QVGA are supported.

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in TR 21.905 [5] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in TR 21.905 [5].

AVC Advanced Video Coding
MBMS Multimedia Broadcast/Multicast Services
PSS Packet Switched Streaming Service
SVC Scalable Video Coding

4 General

4.1 Introduction

This document reports on a number of use-cases, considerations, and technologies. Some of these have resulted in the recommendations included in change requests to [1] and [2].

4.2 Rel 6, 7 and 8 requirements for terminals supporting H.264 in MBMS and PSS

H.264 is one of the codecs defined in PSS and MBMS from Release 6. In release 6 PSS and MBMS, level 1b is recommended. QCIF at 15Hz is a common configuration of H.264 level 1b. In release 7 and 8 MBMS and PSS, level 1.2 is recommended. QVGA at 15Hz is a common configuration of H.264 Level 1.2.

NOTE: H.264 Baseline Profile Level 1.2 content is not decodable by a Level 1b decoder.

H.264 baseline profile is not recommended in its entirety. The "constraint_set1_flag" should be set to 1, which implies a restricted subset of H.264 Baseline Profile. This restricted subset is also a subset of H.264 Main and High profiles.

MBMS recommends only a single codec, namely H.264 Baseline profile, but other codecs that are allowed in PSS are also allowed in MBMS.

Table 1: Overview of Video support in Release 6, 7 and 8 PSS and MBMS

	H.263 profile 0 level 45	**H.263 profile 3 level 45**	**MPEG-4 Visual Simple Profile Level 0b**	**MPEG-4 Visual Simple Profile Level 3**	**H.264 Constrained Baseline profile Level 1b**	**H.264 Constrained Baseline profile Level 1.2**
PSS Release 6	Mandatory	Recommended	Recommended	No Support	Recommended	No Support
PSS Release 7 and 8	Mandatory	Recommended	Allowed through Level 3	Recommended	Allowed through level 1.2	Recommended
MBMS Release 6	Allowed through PSS	No Support	No Support	No Support	Recommended	No Support
MBMS Release 7 and 8	Allowed through PSS	No Support	No Support	No Support	Allowed through level 1.2	Recommended

4.3 Working Assumptions

4.3.1 Codec profiles and levels

A single profile/level combination is defined for baseline terminals, and another for high end terminals. In other words, the present document will recommend a set of settings for baseline terminals, and a set of settings for advanced terminals.

4.3.2 Resolutions

4.3.2.1 General

The following resolutions are examples of resolutions considered relevant for the improved video support work:

QVGA (320x240)

nHD (640x360)

VGA (640x480)

HVGA (480x320)

WQVGA (400x240)

SD (720x576 and 720x480)

4.3.2.2 Resolutions for PC-based UEs

It is anticipated the following resolutions will be needed to support connected mobile PCs

SD (720x576 and 720x480)

XGA (1024×768)

720p (1280×720)

WXGA (1366×768)

1080p (1920×1080)

4.3.3 Bitrates

Analysis should be performed between a minimum video bitrate of 128 kbps and a maximum of 384 kbps for baseline terminals, and for advanced terminals 2 Mbps for MBMS, and 10 Mbps for other cases.

4.3.4 Frame Rates

Minimal Video Frame Rate 12.5Hz or 15Hz (Depending on the test sequence)

Maximal Video Frame Rate 25Hz or 30Hz (Depending on the test sequence)

4.3.5 Random access point frequency

For MBMS, the random access point frequency should be at least 1 per second.

4.3.6 Audio

The audio is assumed to be eAAC+ Stereo Audio at 64 kbps.

NOTE: This section is included for completeness and information.

4.3.7 Error control

Equal error protection of the low and high quality video is assumed. Approaches for Unequal Error Protection should be studied when appropriate channel simulation tools are available.

Unless otherwise stated, the residual error rate (post error correction) is assumed to be zero.

4.3.8 Complexity

As complexity and battery usage are strongly correlated, the complexity of the solutions shall be considered in the evaluation.

4.3.9 Bandwidth efficiency

Bandwidth efficiency is one of the key areas be considered in the evaluation.

4.3.10 Reproduction of results

Results shall be provided in a reproducible manner.

4.3.11 Radio Assumptions

4.3.11.1 MBMS Release 9

3GPP RAN2 and SA2 are working on introducing MBMS functionality over E-UTRAN in Release 9. The work will be restricted to the usage of MBSFN transmission, where a single frequency network will be deployed; later releases may add functionality. MBMS over LTE is expected to provide high spectral efficiency with the minimal target of achieving 1 bit/s/Hz/. LTE E-UTRA provides bandwidth scalability with bandwidths ranging from 1 MHz to 20 MHz. E-MBMS transmissions may be allocated up to 6 out of 10 sub-frames of a frame (of duration equal to 10 ms). Given this, the bit rates for MBMS transmissions may reach around 12Mbps at the minimal spectral efficiency. Effective throughput is expected to be lower to account for transport protocol overhead and reliability overhead. Typical values may be around 2 Mbps to 3 Mbps for Release 9.

NOTE: The minimum UE capabilities are not yet agreed in RAN.

5 Use Cases

This section provides basic, high level use cases for video support in PSS and MBMS. Details are deliberately avoided such as the exact video resolutions used. Such details are instead described in the working assumptions and evaluation sections. Also, all functionality from PSS and MBMS such as content everywhere (including rewind/pause functions), video on demand, digital rights management (DRM), is also considered although not explicitly described per use case.

5.1 Use Cases for Baseline Terminals

5.1.1 File download

This use case considers both ordinary content download and progressive download of content.

An example of this may be that a user has a subscription where content is pre-downloaded to the terminal. A second example is streaming-style progressive download. In this case the content is downloaded as it is being viewed.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated using the abovementioned working assumptions when delivering files to baseline terminals over PSS, MBMS and a combination thereof.

5.1.2 Streaming

This use case considers both live and on demand streaming of video. This may result in switching between MBMS and PSS, but this is transparent to the user.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated using the abovementioned working assumptions when streaming video to baseline terminals over PSS, MBMS and a combination thereof.

5.2 Use Cases for Advanced Terminals

5.2.1 File download to Advanced terminals

This use case is identical to that for baseline terminals, except for the fact that the video quality is substantially higher. In other words, a higher resolutions, bitrates and frame-rates are allowed.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated using the abovementioned working assumptions when delivering files to advanced terminals over PSS, MBMS and a combination thereof.

5.2.2 Streaming to Advanced terminals

This use case is identical to that for baseline terminals, except for the fact that the video quality is substantially higher. In other words, a higher resolutions, bitrates and frame-rates are allowed.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated using the abovementioned working assumptions when streaming video to advanced terminals over PSS, MBMS and a combination thereof.

5.2.3 Streaming to Baseline and Advanced Terminals

5.2.3.1 Broadcast of Baseline and Advanced Mobile TV Services

An MBMS service provides mobile TV content to baseline and advanced terminals. The service makes use of eMBMS in the MBSFN mode. In the broadcast mode, it is not possible to detect the number of receivers that are consuming a specific service configuration in a specific cell. It is therefore required that the different service configurations are broadcast simultaneously throughout the whole SFN area.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated based on their bandwidth efficiency when broadcasting mobile TV content over LTE eMBMS to a variety of UEs with different display capabilities.

5.2.3.2 Broadcast of Baseline Mobile TV Services and Unicast Transmission of Advanced Mobile TV Services

An MBMS service provides mobile TV content to Baseline UEs and advanced mobile TV services are provided over PSS. The baseline service makes use of eMBMS in the MBSFN mode. In the broadcast mode, it is not possible to detect the number of receivers that are consuming a specific service configuration in a specific cell. It is therefore required that the baseline service configuration is broadcast throughout the whole SFN area.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated using the abovementioned working assumptions when streaming video to baseline terminals over LTE eMBMS and advanced terminals over PSS.

5.2.3.3 Unicast Transmission of Baseline and Advanced Mobile TV Services

A PSS service provides both baseline and advanced mobile TV services.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated based on their bandwidth efficiency when transmitting mobile TV content over PSS to baseline and advanced terminals.

5.2.4 File download to Baseline and Advanced Terminals

5.2.4.1 Introduction

The following use cases are essentially different combinations of the above use cases for baseline and advanced terminals.

5.2.4.2 Broadcast file delivery to baseline and advanced terminals

An MBMS service provides mobile TV files to baseline and advanced terminals. The service makes use of eMBMS in the MBSFN mode. In the broadcast mode, it is not possible to detect the number of receivers that are consuming a specific service configuration in a specific cell. It is therefore required that the different service configurations are broadcast simultaneously throughout the whole SFN area.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated using the abovementioned working assumptions when delivering video files to baseline and advanced terminals over LTE eMBMS.

5.2.4.3 Broadcast file delivery to baseline terminals and unicast transmission to advanced terminals

An MBMS service provides mobile TV files to Baseline UEs and advanced mobile TV files are delivered over PSS. The baseline service makes use of eMBMS in the MBSFN mode. In the broadcast mode, it is not possible to detect the number of receivers that are consuming a specific service configuration in a specific cell. It is therefore required that the baseline service configuration is broadcast throughout the whole SFN area.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated using the abovementioned working assumptions when delivering video files to baseline terminals over LTE eMBMS and to advanced terminals over PSS.

5.2.4.4 Unicast file delivery to baseline and advanced terminals

A PSS service provides delivery of video files to both baseline and advanced terminals.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated based on their bandwidth efficiency when delivering mobile TV files over PSS to baseline and advanced terminals.

5.2.5 Streaming to Baseline Terminals and File Download to Advanced Terminals

5.2.5.1 Introduction

The following use cases are essentially different combinations of the above use cases for baseline and advanced terminals.

5.2.5.2 Broadcast of Baseline and Advanced Mobile TV Services

An MBMS service provides mobile TV content to baseline and advanced terminals. The service makes use of eMBMS in the MBSFN mode. In the broadcast mode, it is not possible to detect the number of receivers that are consuming a specific service configuration in a specific cell. It is therefore required that the different service configurations are broadcast simultaneously throughout the whole SFN area.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated using the abovementioned working assumptions when streaming video to baseline terminals and delivering files to advanced terminals over LTE eMBMS.

5.2.5.3 Broadcast of Baseline Mobile TV Services and Unicast Transmission of Advanced Mobile TV Services

An MBMS service provides mobile TV content to Baseline UEs and advanced mobile TV services are provided using unicast file delivery. The baseline service makes use of eMBMS in the MBSFN mode. In the broadcast mode, it is not possible to detect the number of receivers that are consuming a specific service configuration in a specific cell. It is therefore required that the baseline service configuration is broadcast throughout the whole SFN area.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated using the abovementioned working assumptions when streaming video to baseline terminals over LTE eMBMS and delivering files to advanced terminals.

5.2.5.4 Unicast Transmission of Baseline and Advanced Mobile TV Services

A PSS service provides both baseline and advanced mobile TV services.

For realizing this use case, different solutions for improved video support in release 9 will be evaluated based on their bandwidth efficiency when streaming mobile TV content over PSS to baseline terminals and delivering files to advanced terminals.

5.3 Enhancements to the previous use cases

5.3.1 Video aspect ratio management

Nowadays, many high end mobile terminals are available with a screen resolution close to 16:9 aspect ratio (WQVGA, HVGA, WVGA…) whereas most of the mobile phones still have a QVGA (320x240) resolution for which the aspect ratio is 4:3.

Due to the constant growth of High Definition TV/VoD services (over DTT, Satellite, IPTV…), more and more video productions/editions are achieved directly in HD which has a native video aspect ratio of 16:9.

NOTE: Graphics included on the video can be 4:3 compatible.

This use case involves, considering the above considerations, providing a good user experience to terminals with different resolutions and aspect ratios.

5.3.2 UE Power Saving and Fast Stream Switching in MBMS

Efficient power usage is an important criterion in providing MBMS TV service. When the TV stream is transmitted continuously, UE should receive data continuously in active mode, as a result, battery power is consumed. Typical method used for UE power saving is scheduling the transmission and sleep period that UE may turn-off radio component during the sleep interval. This requires discontinuous transmission of MBMS streams. However, a trade-off is that user may experience long delay when switching between streams, if the sleep interval is increased. It is required UE should be able to achieve efficient power usage without incurring long switching delay. IVS may be used for providing quick view of low-quality video while the UE is performing stream switching, as a result, it provides better user experience when changing stream, and improves battery life.

5.3.3 Graceful Degradation

5.3.3.1 Rate adaptation in PSS when entering bad reception conditions

A mobile TV service may have to cope with varying reception conditions at the UE to avoid service interruptions. A desired behaviour would be to apply by rate adaptation of the video stream to the achievable service bit rate. Since a reduced media rate results in a reduced video play out quality, such a video stream adaptation should be performed in a graceful way. Therefore, the service should allow a fine granular rate adaptation to avoid abrupt quality changes in an efficient way.

5.3.3.2 Graceful Degradation in MBMS services when entering bad reception conditions

In contrary to a PSS service, an MBMS service cannot adapt to individual receivers need. That is, users entering difficult reception conditions may experience sudden service interruption instead of soft degradation of e.g. video quality. To keep users satisfied when switching from PSS services to MBMS, a Graceful Degradation of the broadcast service is a desired feature. Such a feature can be applied to a broadcast service by allowing differentiation transmission robustness for different parts of the video stream.

5.3.3.3 Graceful Degradation in Traffic Congestion

In a situation where multiple service users converge in a cell, available bandwidth of the cell depletes quickly. In such case, service to lately incoming UEs may be refused, or all UEs in the cell may suffer severe quality degradation. The situation can be improved when bandwidth of the streams can be reduced with graceful quality degradation using IVS. The service quality is recovered as congestion state of the cell is relieved.

5.3.3.4 Combined support of heterogeneous devices and Graceful Degradation

It is expected, that there will be a coexistence of a variety of device capabilities within 3GPP system and each of these devices may be in different reception conditions. Therefore to cope with both of these challenges in an efficient way, a service should be able to support the heterogeneous devices and to provide Graceful Degradation behaviour at the same time.

5.3.4 Conditional Access

Conditional access of different service qualities could be an interesting use case for offering charged services including free pre-views or low quality views, e.g. offering a lower quality (Resolution, Frame rate, Quality) for free and a premium service with a higher quality.

5.3.5 Rate and quality adaptation with predefined multi quality content originating from external networks

It is expected that an increasing number of users will access pre encoded videos from Internet services such as www.youtube.com or live Internet TV. As it cannot be expected, that such services take care about the special needs of a 3GPP system, the 3GPP system itself would have to take care about the appropriate delivery quality and the delivered bit rate. It should be possible for external networks to provide predefined multi quality content to the 3GPP system which can further distribute it over the users in the different cells and adapt such contents to the special needs of the requesting users and services by providing rate adaptation or quality adaptation within predefined steps within the 3GPP network. Such a feature would help on one hand to improve the quality of internet services and on the other hand to keep the control over bit rate and video quality within the 3GPP network.

This use case is basically identical to normal PSS/MBMS.

5.3.6 HD Support for Mobile PC-based UEs

As of May 2009, MIDs/Netbooks have already achieved 20% penetration of mobile PC sales[1] and large carriers are already shipping 3G enabled versions.

Supporting resolutions ranging from 480x800 to 800x1280, RAM from 500 MB to 2GB and processors from 600 Mhz to 1.7 Ghz, these devices are clearly capable of receiving video at a QoS vastly exceeding the profiles and levels we have defined for PSS and MBMS to date.

More pervasive still is the use of laptop PCs with 3G modems, raising the bar even higher on the resolutions and bitrates required. (It is also important to note that if Moore's law continues to hold, today's laptop PC is 2012's LTE connected Netbook.)

---

1 http://www.xbitlabs.com/news/mobile/display/20090519225036_Netbook_Penetration_Reaches_20_of_Mobile_PC_Market__Report.html

It is therefore assumed that we will need to support HD resolutions in the Rel 9 timeframe

Below is a table identifying the resolutions and bitrates we anticipate needing to support along with the impacted 3GPP SA4 defined technologies.

Video Format Supported	Native Resolution (W×H)	Aspect Ratio (X:Y)	Bitrate	Streaming			Download		
				HTTP	PSS	MBMS	HTTP	Prog	MBMS
480p	720x480	4:3	500 kbps to 3 Mbps	Y	Y	Y	Y	Y	Y
720p 1280×720	1024×768 XGA	16:9	4-7 Mbps	Y	Y	N	Y	Y	Y
	1280×720 WXGA	16:9							
	1366×768 WXGA	683:384 (Approx 16:9)							
1080 HD 1920×1080	1920×1080	16:9	10 Mbps	Y	Y	N	Y	Y	Y

5.3.7 Playback on Alternate Displays

This use case concerns the case where the 3GPP file is being displayed on a screen other than that of the UE to which it was originally received. For example, a monitor or HDTV attached to a mobile PC (see use case 4.3.2.2 on Support for Mobile PC-based UEs), or a superdistributed file played back on an alternate UE.

5.3.7.1 Mobile UE Playback Displayed on External Monitors and HDTVs

This is the ability to render video on the UE to an external screen such as a monitor or television for playback. This is currently done via cables (e.g. Nokia N series) and more recently via short range wireless via DLNA standards. (e.g. Samsung demonstrated sending 480p DIVX via wifi to 40"+ HDTV at CTIA earlier this year.)

The resolutions required are adequately covered in sub-clause 4.3.2.2 on Support for Mobile PC-based UEs.

The impact is that a video file delivered to a UE with a small display may have to be displayed on something much larger than is known at the time of delivery.

5.3.7.2 Playback of Superdistributed files on dissimilar alternate UEs

This use case covers the situation where a 3GPP file is delivered to UE A, and is then transferred to UE B for rendering by B. The means of distribution (e.g. external storage, Bluetooth, DLNA around the home, etc.) is outside the scope of the use case.

The impact is the file sent may need to be resolved up or down dramatically for optimal playback on the alternate UE.

6 Evaluation of Support for Baseline Terminals

In order to support TV in QVGA (and WQVGA) resolution at 25 Hz and 30 Hz, it is proposed that MBMS and PSS recommend H.264 Constrained Baseline Profile at Level 1.3.

7 Evaluation of Support for Advanced Terminals

7.1 Solutions

7.1.1 Scalable Video Coding

7.1.1.1 Introduction

Scalable Video Coding (SVC) [4] has been defined as an extension to the H.264/AVC video coding standard. SVC enhances H.264/AVC with a set of new profiles and encoding tools that may be used to produce scalable bitstreams. SVC supports three different types of scalability: spatial scalability, temporal scalability, and quality scalability. Temporal scalability is realized using the already existing reference picture selection flexibility in H.264/AVC as well as bi-directionally predicted B-pictures. The prediction dependencies of B-pictures are arranged in a hierarchical structure. Furthermore, appropriate rate control is used to adjust the bit budget of each picture to be proportional to its temporal importance in a procedure called quantization parameter cascading. The slightly and gradually reduced picture quality of the hierarchical B-pictures has been shown not to significantly impact the subjective quality and the watching experience, while showing high compression efficiency. Figure 1 shows an example of the realization of temporal scalability using hierarchical B-pictures. The example shows 4 different temporal levels, resulting in one base layer and 3 temporal enhancement layers. This allows the frame rate to be scaled by a factor up to 8 (e.g. from 60 Hz to 7.5 Hz). This approach has the drawback that it incurs a relatively high decoding delay that is exponentially proportional to the number of temporal layers, since the pictures have to be decoded in a different order than their display order. As the coding gain also diminishes with the increasing number of hierarchy levels, it is not appropriate to generate a high number of temporal layers. An alternative to the above mentioned approach for temporal scalability is the use of low-delay uni-directional prediction structures, hence avoiding the out-of-display-order decoding at the cost of reduced coding efficiency.

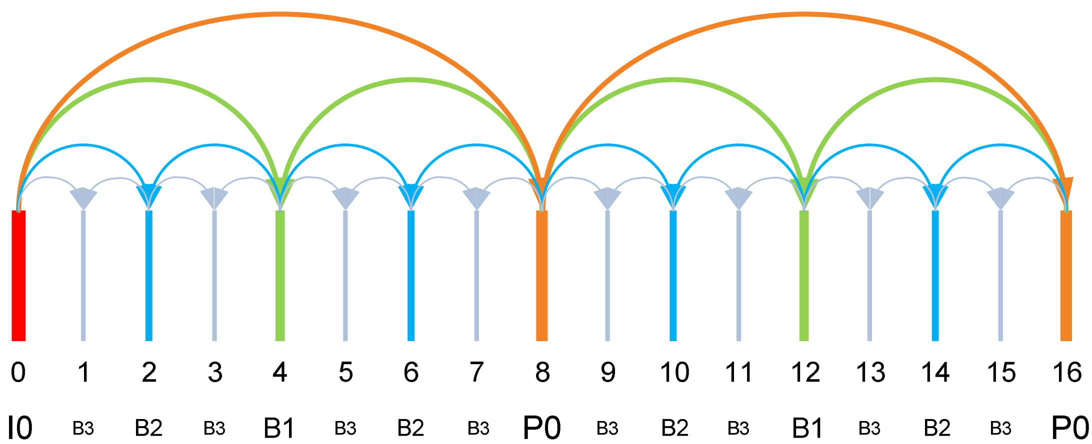

Figure 1: Temporal Scalability with Hierarchical B-Picture Structure in SVC

Spatial scalability is the most important scalability type in SVC. It enables encoding a video sequence into a video bit stream that contains one or more subset bit streams and where each of these subsets provides a video at a different spatial resolution. The spatially scalable video caters for the needs of different consumer devices with different display capabilities and processing power. Figure 2 depicts an example for a prediction structure for spatial scalability (QCIF to CIF resolution). The spatial scalability layer is enhanced with an additional temporal scalability layer that doubles the frame rate at the CIF resolution.

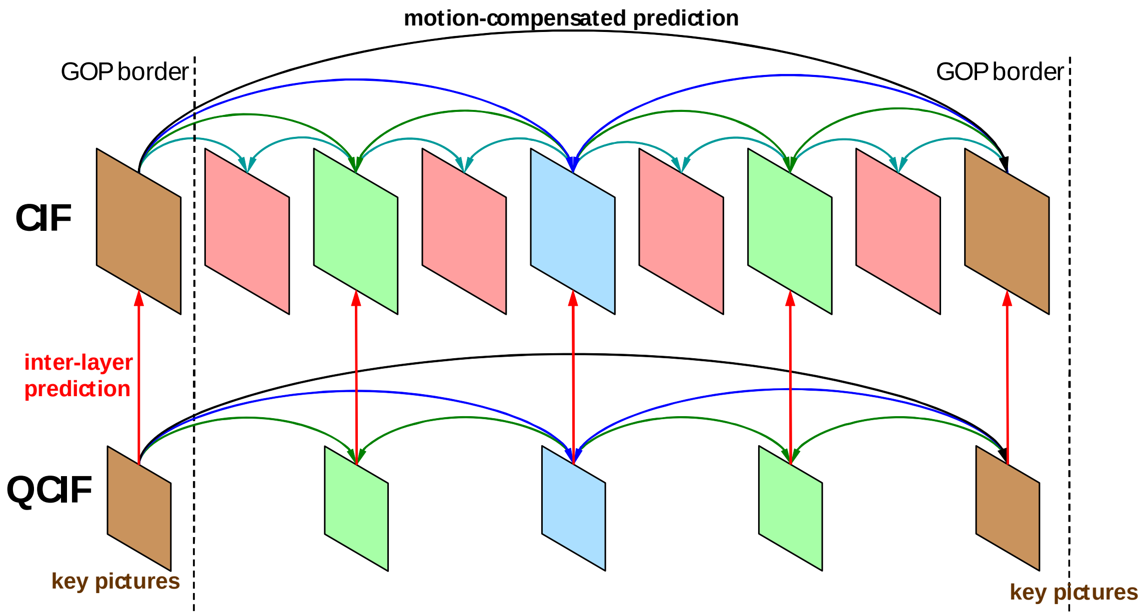

Figure 2: Example Prediction Structure for Spatial Scalability

SVC defines three different inter-layer prediction modes that are designed to enable the single-loop low complexity decoding at the decoder. In other words, motion compensation is performed only once at the target layer at the decoder. The inter-layer prediction tools are inter-layer INTRA (texture) prediction, inter-layer motion prediction, and inter-layer residual prediction.

Inter-layer INTRA prediction enables texture prediction from the base layer at co-located macro-blocks (after upsampling). It is restricted to INTRA coded macroblocks at the lower layer. The up-sampling of the macroblock texture is performed using well-specified up-sampling filters (a 4-tap filter for Luma samples and bi-linear filter from chroma samples). Inter-layer motion prediction implies prediction of the base layer motion vector from the co-located INTER-coded macro-block (after upsampling) of the lower layer. The prediction involves all components of the motion vector: the macro-block partitioning structures, the reference picture indices, and the x- and y- components representing the motion direction. Finally, the inter-layer residual prediction allows inter-layer prediction from the residual after INTER-prediction at the lower layer. At the decoder side, the residual information of the target layer is built up by summing all correctly up-scaled residuals of the lower dependent layers.

The third prediction type in SVC is quality scalability. Quality scalability enables the achievement of different operation points, each yielding a different video quality. Coarse Grain Scalability (CGS) is a form of quality scalability that uses the same tools as the spatial scalability, hence operating in the spatial domain. Alternatively, Medium Grain Scalability (MGS) may be used to achieve quality scalability performing the inter-layer prediction at the transform domain. Two techniques are advocated for MGS scalability: splitting number of transform coefficients and encoding difference of transform coefficients quantized using different quantization parameters. MGS significantly reduces the complexity at encoder and decoder. CGS may be seen as a variant of spatial scalability where the spatial scaling factor is set to one. Quality scalability may be used to address different use cases such as rate adaptation or for offering a high quality pay service.

7.1.1.2 Solution Configuration

For the purposes of improved video support in 3GPP services, a profile of SVC is selected that allows backwards compatibility to basic terminals. This is inherently provided by SVC by requesting the base layer to be H.264/AVC compatible. Furthermore, it has to be ensured that the base layer also conforms to the minimal requirements for basic services. This results in a requirement to have conformance with the restricted baseline profile of H.264/AVC. By consequence, SVC has to be used according to the Scalable Baseline profile.

Additionally, the level selection for a base layer has to be aligned with the minimal level requirements for 3GPP services. For enhancement layers, the level selection is proposed to be set to level 3, which has the following characteristics:

Table 2: Limitations of the proposed SVC level 3

Maximum macroblocks/second	Maximum Frame Size in MBs	Maximum Bitrate	
40500	1620	10 Mbps	
Format	**Luma Width**	**Luma Height**	**Frame Rate**
QCIF	176	144	172
QVGA	320	240	135
WQVGA	400	240	108
CIF	352	288	102.3
HVGA	480	320	67.5
nHD	640	360	45
VGA	640	480	33.8
525 SD	720	480	30
625 SD	720	576	25

The Improved Video Support is meant to address the needs of advanced terminals, as such the proposed solution should be optional for service provider and for UE. Appropriate mechanisms to properly announce and setup the session (either including or excluding enhancement layers) are available or should be extended. If UE supports SVC and it detects that the service also provides SVC enhancement layer(s), then the UE is able to consume the service at an improved quality/resolution.

7.1.1.3 Solution Integration Approaches

7.1.1.3.1 Rate Adaptation for PSS using SVC with priority-based transmission scheduling

This solution integration is related to the use case "Rate adaptation in PSS when entering bad reception conditions" (section 5.3.4.1).

In order to overcome outages and phases with reduced bit rate, a priority-based transmission scheduling (PBTS) algorithm is proposed to be used to pre-buffer larger amounts of more important data for longer playouts than data with less importance for the resulting video playout quality. The adaptation of the transmission scheduling and the media rate is only based on buffer status reports from client to PSS server as depicted in Figure 3.

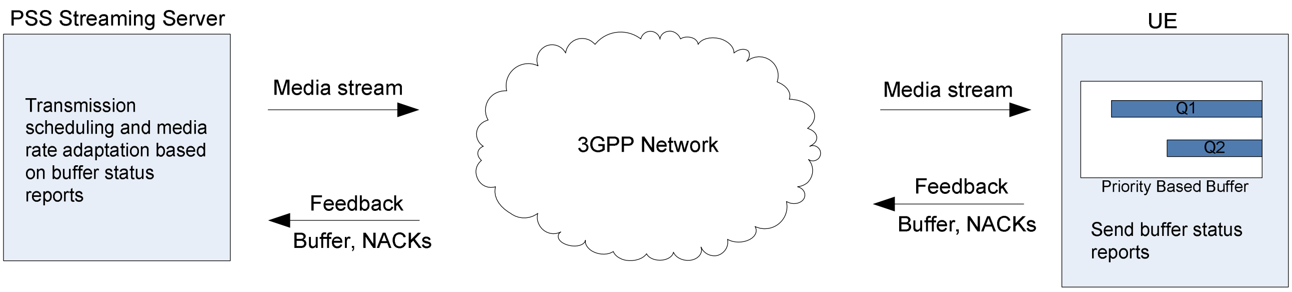

Figure 3: Transmission scheduling and media rate adaptation based on priority based buffer status reports

Typically, the size of a UEs buffer is fixed which is assumed in this scenario. The maximum buffering time is depicted in Figure 4 for a standard buffer with one media quality and a priority based buffer with exemplary two quality levels, either temporal, spatial or quality levels or combination of those.

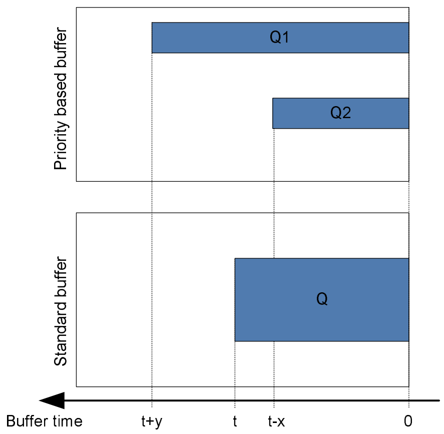

Figure 4: Priority (PBTS) buffer using different qualities (*Q1* and *Q2*) vs. standard buffer with one quality (Q), with *t+y* respectively *t* being the maximum sustainable outage time

In this example, the maximum buffer time for the standard buffer is *t*, which is dependent on the bit rate of the video stream (Q). The priority buffer allows to prebuffer a longer time of the lowest quality level (Q1) *t+y* by reducing the prebuffer time of the higher quality level (Q2) to *t-x*, where *t+y* and *t-x* depend on the bit rate of the quality levels.

To fill up a standard buffer, the PSS server uses a transmission scheduling in decoding order of the video stream. Whereas to fill up a priority based buffer, the PSS server uses a priority based transmission scheduling, where it first fills up the lowest quality level to *t+y* and after that the higher quality layer to *t-x*. After that it switches to the standard transmission scheduling in decoding order.

When the UE enters difficult reception conditions, the available bit rate may no longer be sufficient for the transmission of the highest quality. Having a standard buffer, in such a case users would experience a video outage. In case of having a buffer filled with a priority scheduling algorithm, the high quality data in the buffer runs out earlier than lower qualities. Using SVC, the PSS server would adapt the media stream bit rate to the available service bit rate by dropping quality layers, which still allows to keep the buffer state of the lowest quality level fully filled. Compared to the use of a standard buffer, the highest quality runs out even faster with the priority based approach. Nevertheless, the priority based scheduling allows for keeping the playout alive during longer outages than in the standard case.

Dependent on the buffer reports, the PSS streaming server adapts the media stream bit rate to the quality of the available service bit rate. If the clients' reception condition allows a higher quality, the transmission scheduling is adapted to allow rebuffering of the priority buffer to the maximum quality of the available service bit rate.

Although PBTS can be based on H.264/AVC temporal scalability (AVC-PBTS), SVC has the handy advantage to allow a bit rate reduction using quality or spatial scalability instead of relying on pure temporal scalability as described in [4].

7.1.1.3.2 Unequal error protection with SVC

The presented solution is related to the use cases "Graceful Degradation in MBMS services when entering bad reception conditions" (section 5.3.3.2) and "Combined support of heterogeneous devices and Graceful Degradation" (section 5.3.3.4).

The layered structure of SVC allows for transmission of the video in separate network streams. This can be used to offer broadcast services which allow use cases like the support of heterogeneous devices (section 5.2.3), graceful degradation behaviour (section 5.3.3.2), conditional access (section 5.3.4) or combinations of those (section 5.3.3.4). SVC allows services providing different quality steps either by temporal, spatial, quality scalability or combination of those. Using unequal error protection (UEP), such a service can provide different quality levels of different robustness, which allows for Graceful Degradation behaviour in MBMS scenarios. An exemplary UEP scheme is depicted in Figure 5, where the more important layer (Base) has a higher protection than the enhancement layers.

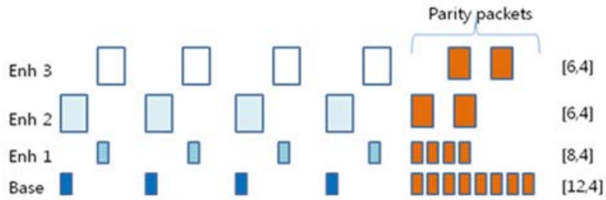

Figure 5: UEP (Unequal Error Protection): The more significant packets are protected by a higher code rate

In the exemplary scenario in Figure 6, there are two layers, using quality, spatial or temporal scalability or combinations of those, with different robustness. UEs in good reception conditions will receive the highest quality and UEs entering worse reception conditions can still receive the base layer, which results in a drop in quality when entering bad reception conditions.

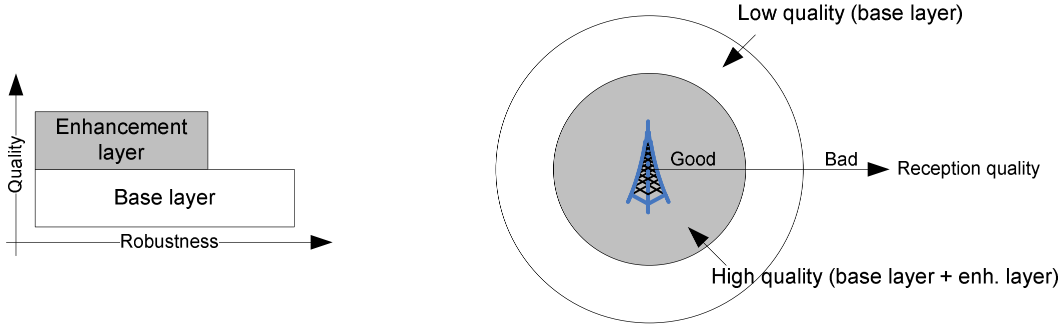

Figure 6: MBMS service with graceful degradation behaviour using unequal error protection with SVC either with temporal, spatial or fidelity scalability or combinations of those

Such a differentiation in robustness of the scalable layers can be applied by a MBMS service at the application layer using different code rates at the application layer forward error correction (AL-FEC).

7.1.2 H.264/AVC High Profile

7.1.2.1 Introduction

AVC High Profile [4] was added just over a year after the completion of the original H.264/AVC standard as part of a group of new profiles named Fidelity Range Extensions (FRExt). H.264/AVC High profile has taken over H.264/AVC Main Profile as the standard used for mainstream high quality broadcast and storage applications.

Tools in H.264/AVC High Profile include, for example, context-adaptive binary arithmetic coding (CABAC), interlaced coding, monochrome coding, B-frames and 4x4 and 8x8 transforms and prediction modes.

NOTE: H.264 (AVC) Main Profile is a subset of H.264 High Profile, and a High Profile decoder is required to be able to decode Main Profile streams.

7.1.2.2 Solution configuration

The appropriate solution configuration for H.264/AVC High Profile in advanced terminals is H.264/AVC High Profile at level 3 [4].

Restrictions on the use of tools of the above configuration is FFS , e.g. restriction on interlaced coding.

This implies a maximum bitrate of 12.5 Mbps and a maximum frame size of 1620 macroblocks (equivalent to 720×576). The maximum VCL data-rate is set to 40 500 macroblocks per second. Table 3 gives examples of maximum frame rates with respect to image size.

Table 3

Width	Height	Frame Rate (fps)
720	576	25
720	480	30
640	480	33.75
320	240	135

8 Conclusions

This Technical Report provides a collection of use cases and solutions for addressing terminals with baseline and advanced video capabilities. The evaluation of the solutions against the identified use cases could not be performed during the time frame of release 9.

It was agreed to update the codec level requirement for H.264/AVC Baseline profile to level 1.3. Furthermore, it was agreed to adopt H.264/AVC High Profile at Level 3.0 as an optional video codec for the Packet-switching Streaming Service (PSS).

Annex A: Change history

Change history							
Date	**TSG #**	**TSG Doc.**	**CR**	**Rev**	**Subject/Comment**	**Old**	**New**
2010-03	47	SP-100023			Approved at TSG SA#47		9.0.0
2011-03	51				Version for Release 10	9.0.0	10.0.0
2012-09	57				Version for Release 11	10.0.0	11.0.0
2014-09	65				Version for Release 12	11.0.0	12.0.0
2015-12	70				Version for Release 13	12.0.0	13.0.0

Change history							
Date	**Meeting**	**TDoc**	**CR**	**Rev**	**Cat**	**Subject/Comment**	**New version**
2017-03	75					Version for Release 14	14.0.0
2018-06	80					Version for Release 15	15.0.0
2020-07	-	-	-	-	-	Update to Rel-16 version (MCC)	**16.0.0**
2022-04	-	-	-	-	-	Update to Rel-17 version (MCC)	**17.0.0**

History

Document history		
V17.0.0	May 2022	Publication