Video and Streaming in Nokia Phones

Version 1.0; June 16, 2003

Messaging

NOKIA

Contents

Change History

16 June 2003	V1.0	This document replaces "Video in Nokia phones"-document. Tools section, Nokia 6220 and Nokia 6600 added. Document structure is also changed.

Disclaimer

License

Video and Streaming in Nokia Phones

Version 1.0; June 16, 2003

1 Introduction

The following document will help developers create audio-visual content and services for video-capable Nokia handsets, such as the Nokia 6650, Nokia 3650 and Nokia 6600. The Nokia 3650 and Nokia 6600 are also referred as Series 60 devices.

The term "video," when used in the context of *video* streaming and *video* player, should be interpreted as the capability to render simultaneously both video and audio content.

2 Video Player

The same user interface will be used for both video streaming and a stand-alone local video player. The main difference between these two is that in video streaming, the compressed video data is fetched from a server during playback, whereas in the local video viewer, the video clip to be viewed is stored locally in the memory of the mobile terminal or memory card. The external session establishment of video streaming service is implemented using the WAP browser or messaging applications (SMS, MMS viewers), which have the ability to launch external applications such as the player application. This is done by recognizing an RTSP:// URL or a downloaded definition file (.RAM file), which will launch the installed streaming player. In the player application, streaming links are typically launched from the player's editable play list.

Downloading a file first and playing it afterwards is considered local playback, because the entire video file has already been stored locally. Note also that Nokia handsets do not support progressive downloading, meaning that the entire video clip must be downloaded before playback.

The creation of a video clip requires the following elements:

- One or more bit streams (e.g., one video bit stream and, if present, an audio bit stream). A bit stream is the output of a video or an audio encoder. See Sections 3 and 4 for discussions of video and audio coding, respectively.

- One file format, which packages the audio visual content into a file, synchronizes it, and allows access to it.
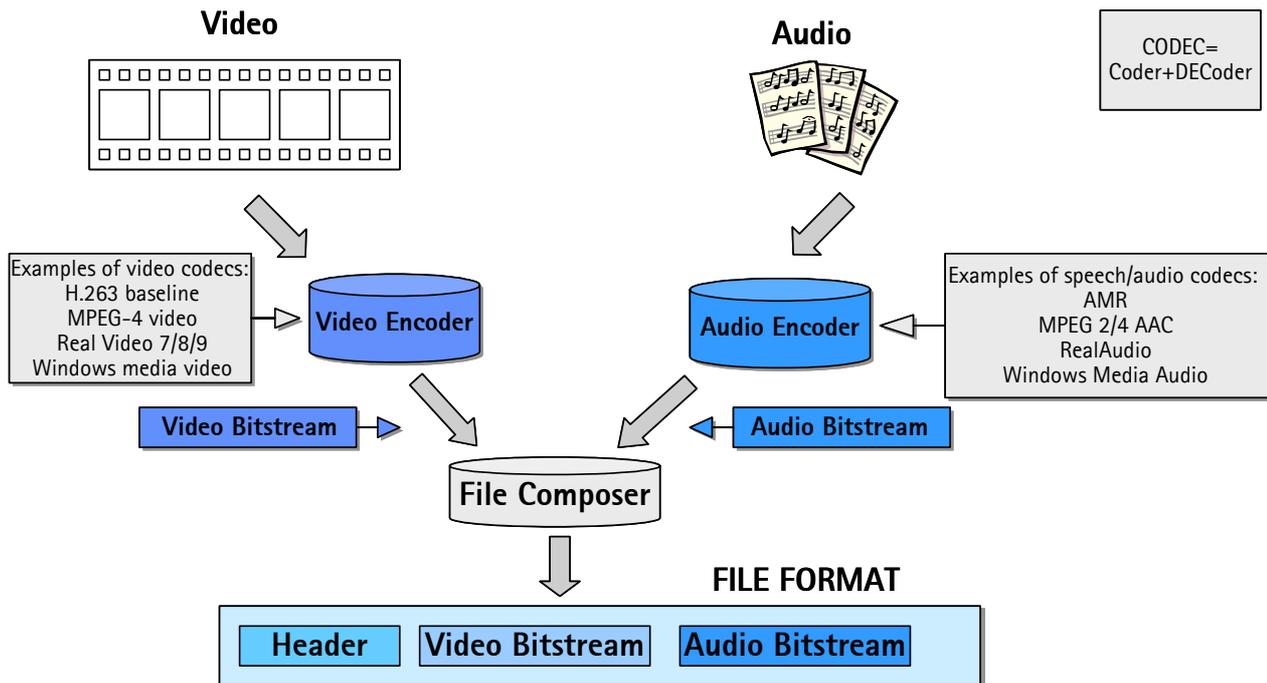
Figure 1: Key elements for creating a video clip

Note that some Nokia 3650 variants may have different sets of features in video streaming (for example, the Chinese version of the Nokia 3650).

3   Video Coding

A video sequence consists of a series of still images. Video compression methods are based on reducing the redundant and perceptually irrelevant parts of video sequences. That redundancy can be categorized into spatial, temporal, and spectral redundancies. Spatial redundancy refers to the correlation between neighboring pixels. Temporal redundancy means that the same objects appearing in the previous image are likely to appear in the current image as well. Spectral redundancy addresses the correlation between the different color components of the same image. Compression can be achieved by generating motion-compensation data, which describes the motion between the current and the previous image. It can be said that the current image is predicted from the previous one.

Usually, however, efficient compression cannot be reached simply by reducing the redundancy of the sequence. Thus, video encoders must also discard some non-redundant information. When doing this, the encoders take into account the properties of the human visual system and mainly discard the information that is least important for the subjective quality of the image. In addition, the redundancy of the encoded bit stream is reduced by means of efficient lossless coding of compression parameters and coefficients. The main technique is to use variable-length codes.

Video compression methods typically differentiate between images that can or cannot use temporal redundancy reduction. Compressed images that do not use temporal redundancy reduction methods are usually called INTRA or I-frames, whereas temporally predicted images are called INTER or P-frames. In the INTER frame case, the predicted (motion-compensated) image is rarely sufficiently precise, and therefore a spatially compressed prediction error image is also associated with each INTER frame.

In video coding, there is always a tradeoff between bit rate and quality. Some image sequences may be harder to compress than others due to rapid motion or complex texture, for example. In order to meet a constant bit-rate target, the video encoder controls the frame rate as well as the quality of images:

the more difficult the image is to compress, the worse the image quality. If variable bit rate is allowed, the encoder can maintain a standard video quality.

3.1 H.263

ITU-T H.263 is an established codec used in various multimedia services, and it provides a wide "toolbox" of various encoding tools and coding complexities for different purposes. The tools to be used and the allowed maximum complexity for the particular mode are defined in so-called codec profiles and levels, respectively.

For most mobile multimedia services, the H.263 Profile 0, Level 10 (also known as "H.263 baseline"), has been defined as a mandatory codec. This is also the mainstream codec supported in Nokia video players.

H.263 uses the Discrete Cosine Transform (DCT) to reduce spatial redundancy. The transform converts a block of pixels to coefficients that represent the spatial frequency components of the block. Only the frequencies appearing in the block have high-amplitude coefficient values; other coefficients are close to zero. For example, a constantly colored block has only one spatial frequency and it is transformed to one non-zero DCT coefficient, whereas the other DCT coefficients remain zero. Consequently, the DCT coefficient block is easier to code with run-length codes than the original block of pixels. In order to gain compression, the transformed block is quantized, which means that the coefficients are rounded to certain quantization levels. The fewer possible quantization levels there are, the fewer bits it takes to represent a quantization level. An approximation of the original block of pixels can be restored from the coded DCT quantization levels by applying an inverse DCT. The fewer quantization levels used, the worse the quality of the reconstructed image. ITU-T H.263 allows 31 quantization step sizes that are controlled by the so-called quantization parameter.

3.2 MPEG-4 Visual

The Nokia 3650, Nokia 6600 and Nokia 6650 will also support decoding of video bit-stream content conforming to MPEG-4 Visual Simple Profile, Level 0. This is an optional video codec in several recent multimedia standards.

3.3 RealVideo 7 and RealVideo 8

RealVideo 7 and RealVideo 8 are two proprietary video coding formats developed by Real Networks. The RealOne Player installed on board the Nokia device supports the decoding of video in RealVideo 7 and RealVideo 8 formats.

RealVideo 7 is similar to MPEG-4 and H.263 in that it makes use of I-frames and P-frames and is block transform-based. It also uses Bidirectional frames or B-frames that are computed based on interpolation between the previous and next I- and P-frames using motion vectors and spatial information. RealVideo 7 is more efficient than MPEG-4 or H.263 due to the use of B-frames and some advanced coding techniques created by RealNetworks.

RealVideo 8 is also similar to MPEG-4 and H.263 in that it makes use of I-frames and P-frames and is block transform-based. It also uses Bidirectional frames or B-frames. However, it is 30% to 40% better than RealVideo 7 due to major differences in the specific techniques used for transform coding, motion estimation, and filtering.

4 Audio Coding

Arbitrary sounds can be represented as a sum of waves having different frequencies and amplitudes. In other words, any sound is an amplitude waveform as a function of time. Sounds can be digitized when samples of the corresponding waveform are taken frequently enough. For arbitrary sounds and music, a 44.1 kHz sampling frequency is considered to provide high quality. For speech, an 8 kHz sampling frequency is adequate for most applications. Typically, 16 bits is enough to represent one sample.

Digitized audio can be compressed in various ways. A simple coding method is to use an adaptive step size to quantize audio samples. Such a technique is used in the IMA ADPCM audio coding standard that reserves 4 bits per sample. Consequently, if the sampling frequency is 8 kHz, IMA ADPCM coded audio takes 32 kbit/s. Another simple audio coding method is A-law PCM, which uses a logarithmic quantization step size and reserves 8 bits per sample.

More advanced audio coding methods take advantage of the human psychoacoustics model. Parts of the audio signal are barely audible and can be discarded or compressed. Typically, the advanced coding audio methods are categorized into generic audio coding and speech coding techniques. Generic audio coding algorithms are targeted for music and sound as well as human voices, whereas speech coding algorithms are aimed at speech only and perform relatively poorly when music is coded.

4.1 AMR

One of the most advanced speech coding standards today is the Adaptive Multi-Rate (AMR) speech codec, which was developed by the European Telecommunications Standards Institute (ETSI). It includes eight speech coding modes, whose bit rates range from 4.75 to 12.2 kbit/s. Some of the modes are the same as the earlier-defined voice telephony codec specified for other standards. For example, AMR at 12.2 kbit/s is the same speech codec as the GSM Enhanced Full-Rate (EFR) codec. AMR is a mono type codec.

To provide the best listening experience, AMR encoding should be used for content that does not include complex, noisy, or critical music content. The limited audio bandwidth of 3.5 kHz means higher frequencies are not reproduced perfectly and are filtered with an anti-aliasing filter. Content proven to work satisfactorily with AMR includes news and sports coverage and light popular music.

4.2 AMR-WB

AMR-WB represents state-of-the-art technology in low bit rate wideband speech coding. Like AMR, it is a multi-rate speech codec. AMR-WB technology uses nine bit rates between 6.6 and 23.85 kbit/s at 16 kHz sampling rate. The speech processing is performed on 20 ms frames, so each AMR-WB encoded frame represents 320 speech samples. AMR-WB uses MIME type audio/amr-wb and file extension *.awb. AMR-WB payload format as well as MIME type is defined in RFC 3267[1].

AMR-WB has been selected by the 3GPP in December 2000 and ITU-T in July 2001. The ITU-T AMR Wideband is now known as G.722.2. 3GPP has several specifications available related to AMR-WB:

- 3GPP TS 26.190 [2]

- 3GPP TS 26.201 [3]

- 3GPP TS 26.174 [4]

- 3GPP TS 26.194 [5]

4.3 RealAudio Voice, RealAudio 7, and RealAudio 8

RealOne Player in the Nokia 3650 and Nokia 6600 supports RealAudio Voice, RealAudio 7, and RealAudio 8 audio codecs. RealAudio 7 is also known as RealAudio G2 codec.

RealAudio 7 is an Internet audio codec used for many legacy contents served by RealServers. RealVoice is an ADPCM codec used for speech frequency ranges. For information on the use of these codecs in mobile environments, please consult with RealNetworks or check out its mobile content authoring guidelines on the Internet (also see Section 6.2.3).

RealAudio 8 (©RealNetworks Inc.) is a proprietary audio codec that supports various sampling and bit rates, starting from very low bit rates (3-4 kbit/s). Nokia doesn't recommend use of RealAudio 8 at bit rates covered with AMR, but if necessary for bandwidth limitation reasons, the codec can be used. For information on the optimal use of bit rates and sampling rates, please consult RealNetworks' Web guidelines.

For the time being, RealAudio 8 should be used for encoding music content intended for audio-only streaming or local playback from the Nokia 3650 phone memory or a memory card.

The upper decoding speed limit of the RealAudio 8 codec supported in the Nokia 3650 and Nokia 6600 is 96 kbit/s.

Note that maximum digital-to-analog sampling conversion provided by the Nokia 3650 device is 8 kHz; by the Nokia 6600 it is 16 kHz.

5 File Formats

Conceptually there is a difference between the video coding format and the video file format. The coding format is related to the action of a specific coding algorithm that codes the content information into a code stream. The file format is a way of organizing video and audio code streams so that they can be accessed for local decoding and playback or streamed over a transport channel.

The following list includes some of the most common video file formats today:

- Microsoft Audio-Video Interleaved (AVI)

- Apple QuickTime file format (.mov)

- MPEG-1 file format (.mpg)

- 3GPP file format (.3gp)

- MP4 file format (.mp4)

In accordance with the 3GPP specifications, Nokia supports the 3GPP file format (extension: .3gp) for the storage of video and associated audio (if present).

Note that the Nokia Video Player doesn't commit to play back media content inside .mp4 files (MPEG-4 file format), although the file may contain media types supported by the player. If the .mp4 file declares 3GPP file format compatibility, the content will play out as originally intended by the content author, otherwise the player presentation may not reproduce original content.

6 Video Capabilities in Nokia Products

6.1 Nokia 6600

6.1.1 Video Recorder

The Video Recorder can record video clips using the Nokia 6600's integrated camera. The video clip is saved either in the internal memory of the device or to an MMC card. The clips are coded according to the H.263 Profile 0, Level 10 standard, with or without audio depending on the device settings. Image size can be selected between 128 x 96 (Sub-QCIF) and 176 x 144 (QCIF) with a maximum 10 fps frame rate. The maximum multimedia file size is set to 95 kB. RealOne Player can be used to play back the recorded video clip.

Figure 2: Nokia 6600 Video Recorder

6.1.2 RealOne Player

RealOne Player has been integrated to e-mail and the MMS client and it is used to play back local content such as files in the memory card, e-mail and MMS attachments. It can also be used to stream video content from remote RealServers over the GPRS and CSD connections. Maximum bit stream for video clips has been set to 64 kbit/s. With H.263 codec, such a bit stream can be utilized to process 10 to 15 frames per second, depending on the content and encoding tools involved in content authoring. The optimal size for clips is 176 x 144 pixels. The following media types can be played with this player:

- video/3gpp (*.3gp)

  o H.263 coded video + AMR or AMR-WB coded audio

  o MPEG-4 Visual Simple Profile, Level 0 + AMR or AMR-WB coded audio

- video/vnd.rn-realvideo

- application/x-pn-realmedia

- application/vnd.rn-realmedia

- audio/x-pn-realaudio

- audio/vnd.rn-realaudio

- audio/3gpp

6.2 Nokia 6650

The Nokia 6650 records video clips using an integrated camera, and saves them in its internal memory. A video clip is coded according to the H.263 Profile 0, Level 10 standard, with or without audio depending on the device settings. The maximum video file size is set to 100 kB, if sent inside an MMS message. Otherwise the video file size is unrestricted. Image size is set to 128 x 96 (Sub-QCIF) and the maximum frame rate is 15 frames per second. The player can play back .3gp clips, which include H.263 and AMR. In playback mode the maximum accepted image size is 176 x 144 (QCIF) and the maximum frame rate is 15 frames per second. If 176 x 144 (QCIF) is used, the image is scaled down to 128 x 108 to fit on the screen.

Figure 3: Nokia 6650 Video Player

The Video Player in the Nokia 6650 has been integrated into the MMS client and it is used to play local content such as files that are transferred into the phone memory from an external device and/or received as MMS attachments. The maximum bit rate for video clips has been set to 64 kbit/s. Progressive downloading or streaming is not supported with the Nokia 6650.

The following media types can be played in the Nokia 6650:

- 3GPP   video/3gpp

   o H.263 coded video + AMR coded audio

   o MPEG-4 Visual Simple Profile, Level 0 + AMR coded audio

6.3 Nokia 3650

Nokia 3650 features the Video Recorder and has two different players, the Nokia embedded Video Player and RealOne Player. Only RealOne Player is visible as launchable application in the application grid.

6.3.1 Video Recorder

The Video Recorder can record video clips using the Nokia 3650's integrated camera. The video clip is saved either in the internal memory of the device or to an MMC memory card. The clip is coded according to the H.263 Profile 0, Level 10 standard, with no audio. Image size is set to 128 x 96 (Sub-QCIF) and the maximum frame rate is 10 frames per second. The maximum multimedia file size is set to 100 kB. RealOne Player can be used to play back the recorded video clip.

6.3.2 Nokia Video Player

The Nokia Video Player is not visible in the Nokia 3650 application grid, but it has been integrated to e-mail and the MMS client and it is used to play back local content such as files in the memory card, e-mail, and MMS attachments. Maximum bit stream for video clips has been set to 64 kbit/s. With H.263 codec, such a bit stream can be utilized to process 10 to 15 frames per second, depending on the content and encoding tools involved in content authoring. The optimal size for clips is 176 x 144 pixels. This player cannot be used to stream video or audio; it is meant only for local file playback. The following media types can be played with this player:

- 3GPP   video/3gpp

o H.263 coded video + AMR coded audio

o MPEG-4 Visual Simple Profile, Level 0 + AMR coded audio

6.3.3 RealOne Player
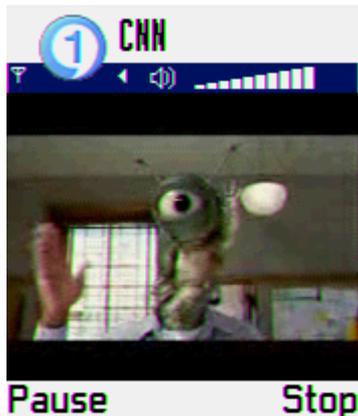
RealOne Player can be used to play local content as well as stream from remote RealServer servers over the GPRS and CSD connections. The following MIME types have been assigned to RealOne Player:

Figure 4: RealOne Player in the Nokia 3650

- audio/x-pn-realaudio

- application/x-pn-realmedia

- application/vnd.rn-realmedia

- audio/vnd.rn-realaudio

- video/vnd.rn-realvideo

- audio/3gpp

- video/3gpp

Additionally, RealOne Player supports in local playback 3GPP files that contain a single MPEG-4 Visual Simple Profile Level 0 track and an AMR track, AMR-only files (.amr), and playback of the following RealNetworks media formats stored in RealMedia file format (.rm):

- RealAudio Voice speech codec

- RealAudio 7 and RealAudio 8 audio codecs

- RealVideo 7 and RealVideo 8

6.4 Nokia 6220

6.4.1 Video Recorder

The Video Recorder can record video clips using the Nokia 6220's integrated camera. The video clip is saved in the internal memory of the device. The clip is coded according to the H.263 Profile 0, Level 10 standard, with audio. Image size is set to 128 x 96 (Sub-QCIF) and the maximum frame rate is 7.5 frames per second. The maximum multimedia file size is set to 100 kB.

6.4.2   Video Player

Video Player is used to play back local content. Maximum bit stream for video clips has been set to 64 kbit/s. With H.263 codec, such a bit stream can be used to process 10 to 15 frames per second, depending on the content and encoding tools involved in content authoring. The optimal size for clips is 128 x 96 pixels. This player cannot be used to stream video or audio; it is meant only for local file playback. The following media types can be played with this player:

- 3GPP   video/3gpp

   o H.263 coded video + AMR coded audio

6.5   Tools for Video Content Creation

Forum Nokia offers tools for creating video content aimed at mobile phones. The Nokia Multimedia Converter 2.0 Beta 2 has the ability to convert the most common video/audio formats, such as avi, mov, mp3 and wav, to 3gp file format. The tool also includes a console version, which enables developers to integrate the conversion to their system and convert the content on the fly. The tool is available at the Forum Nokia Web site.

7   Streaming

Streaming is the transmission of the real-time data, audio and video, from a server to a client, where the client decodes and plays the data as it is received. Streaming lets the viewer see the content immediately after a short buffering period. One advantage compared to local playback is that no data will be stored permanently to the terminating client. The end user cannot forward or send the content to another user, so the content can be charged each time it is consumed. This is one great advantage for service providers.

Available transfer bandwidth for streaming sessions will vary based on the access technology used and the configuration of the operator network. If the content is to be served within a particular operator's service domain only, appropriate authoring guidelines are usually available from the operator or the service provider.
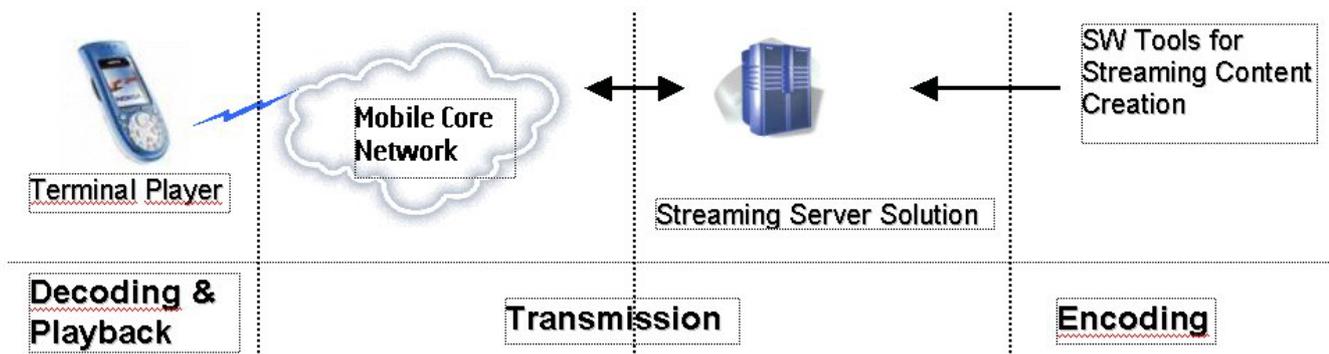
7.1   Streaming Capabilities in Nokia Products

The maximum manual bandwidth setting in RealOne Player is 40.2 kbp/s. As a general content creation guideline for content authoring it can be said that typical sustainable maximum throughput range with Nokia 3650 and Nokia 6600 over GPRS is 20-25 kbit/s, with HSCSD 3+1 connection up to 43 kbit/s.

7.1.1   Nokia 6600

RealOne Player in the Nokia 6600 can be used to stream content from any 3GPP-compliant server, including RealNetworks Helix Mobile Server, over the GPRS and CSD connections.

For video streaming, RealOne Player in the Nokia 6600 supports the following media formats either alone (e.g., audio-only streaming) or as combinations of audio-visual media types:

- 3GPP streaming via RTP

    o H.263 Profile 0, Level 10

    o MPEG-4 Visual Simple Profile, Level 0

    o AMR

    o AMR-WB

- RealAudio Voice speech codec

- RealAudio 7 and RealAudio 8 audio codecs

- RealVideo 7 and RealVideo 8

7.1.2   Nokia 3650

RealOne Player in the Nokia 3650 can be used to stream content from remote RealServer servers over the GPRS and CSD connections.

For video streaming, RealOne Player in the Nokia 3650 supports the following RealNetworks media formats either alone (e.g., audio-only streaming) or as combinations of audio-visual media types:

- RealAudio Voice speech codec

- RealAudio 7 and RealAudio 8 audio codecs

- RealVideo 7 and RealVideo 8

7.2   Tools for Streaming Content Creation

RealNetworks provides support for creating content suitable for Mobile RealOne Player. The Helix Producer authoring tool is available from RealNetworks, at http://www.real.com, and various content authoring guidelines for both local playback and streaming for the Nokia 3650 are available at http://www.realnetworks.com/mobile/nokia/content_providers.html.

8   Terms and Abbreviations

3GPP	MIME subtype for 3GPP file format; also Third Generation Partnership Project (gen.)
3GP or 3gp	File name extension for 3GPP file format
ADPCM	Adaptive Pulse Code Modulation
AMR	Adaptive Multi-Rate speech codec
AMR-WB	Adaptive Multi-Rate Wideband speech codec
API	Application Programming Interface
AVI	Microsoft Audio-Video Interleaved file format
CD-ROM	Read-Only Compact Disc
CSD	Circuit Switched Data
DCT	Discrete Cosine Transform
GPRS	General Packet Radio System
GSM	Global System for Mobile Communication
HSCSD	GSM High Speed Circuit Switched Data
HTML	Hypertext Markup Language
Hz	Hertz, 1/sec
IMA	Interactive Multimedia Association
ITU-T	International Telecommunication Union, Telecommunication Standardization Sector
kbit/s	kilobits per second
MMS	Multimedia Messaging Service
MOV or mov	File name extension for Apple QuickTime file format
MPEG	ISO/IEC Moving Pictures Experts Group
MPG or mpg	File name extension for MPEG-1 file format
MP3	MPEG-1 Audio Layer 3 audio coding
MP4	MIME subtype for MPEG-4 file format
mp4	File name extension for MPEG-4 file format
NIM	Nokia Interleaved Multimedia file format
PC	Personal Computer

PCM	Pulse Code Modulation
QCIF	Quarter Common Intermediate Format
RAM or ram	File extension for RealNetworks' simple URI descriptor files
RGB	Red-Green-Blue color space
RM or rm	File extension for RealMedia file format
Sub-QCIF	Sub Quarter Common Interchange Format
WAV	Microsoft waveform audio file format
YUV	Color space, Y is the luminance or grayscale component; U and V are chrominance or color difference components

9   References

[1] *RFC 3267 – Real-Time Transport Protocol (RTP) Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs*, June 2002

[2] *3GPP TS 26.190 - AMR Wideband Speech Codec (release 5)*, December 2001, Version 5.1.0

[3] *3GPP TS 26.201 – AMR Wideband Speech Codec; Frame Structure*, March 2001, Version 5.0.0

[4] *3GPP TS 26.174 – Adaptive Multi-Rate (AMR) Wideband Speech Codec Test Sequences*, December 2002, Version 5.4.0

[5] *3GPP TS 26.194 – AMR Wideband Speech Codec; Voice Activity Detector (VAD)*, March 2001, Version 5.0.0

Build Test Sell

Developing and marketing mobile applications with Nokia

1

Go to Forum.Nokia.com

Forum.Nokia.com provides the tools and resources you need for content and application development as well as the channels for sales to operators, enterprises, and consumers.

Forum.Nokia.com

2

Download tools and emulators

Forum.Nokia.com/tools has links to tools from Nokia and other industry leaders including Borland, Adobe, AppForge, Macromedia, Metrowerks, and Sun.

Forum.Nokia.com/tools

3

Get documents and specifications

The documents area contains useful white papers, FAQs, tutorials, and APIs for Symbian OS and the Series 60 Platform, J2ME, messaging (including MMS), and other technologies. Forum.Nokia.com/devices lists detailed technical specifications for Nokia devices.

Forum.Nokia.com/documents
Forum.Nokia.com/devices

4

Test your application and get support

Forum Nokia offers free and fee-based support that provides you with direct access to Nokia engineers and equipment and connects you with other developers around the world. The Nokia OK testing program enables your application to enjoy premium placement in Nokia's sales channels.

Forum.Nokia.com/support
Forum.Nokia.com/ok

5

Market through Nokia channels

Go to Forum.Nokia.com/business to learn about all of the marketing channels open to you, including Nokia Tradepoint, an online B2B marketplace.

Forum.Nokia.com/business

6

Reach buyers around the globe

Place your applications in Nokia Tradepoint and they're available to dozens of buying organizations around the world, ranging from leading global operators and enterprises to regional operators and XSPs. Your company and applications will also be considered for the regional Nokia Software Markets as well as other global and regional opportunities, including personal introductions to operators, on-device and in-box placement, and participation in invitation-only events around the world.

Forum.Nokia.com/business