Video Editing for the Nokia 9210 Communicator

TABLE OF CONTENTS

1. Video for the Nokia 9210 Communicator

The Video player application for the Nokia 9210 Communicator is pre-installed on the memory card contained in the sales package. Video player runs video files coded according to the Nokia Interleaved Multimedia (NIM) format. Thus, you must convert your existing video files to the NIM format before transferring them into the communicator. Figure 1 presents a general overview on how video files are generated for, transferred to, and played in the Nokia 9210 Communicator.

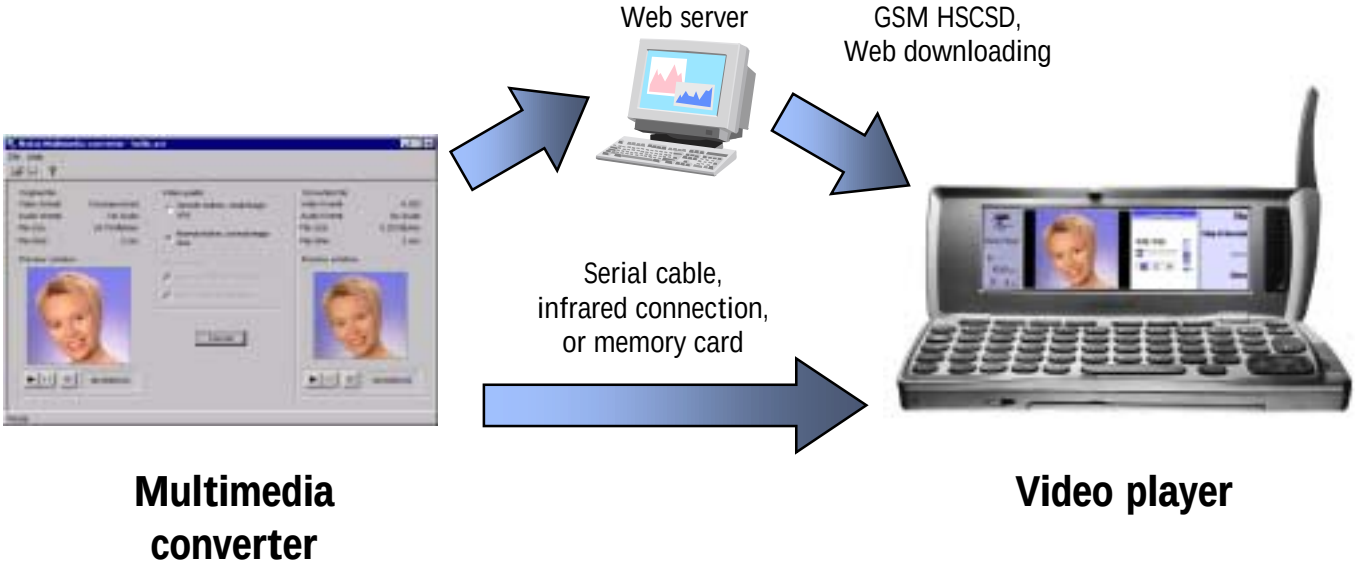

Figure 1. Video for the Nokia 9210 Communicator

Multimedia converter runs on any 32-bit Microsoft Windows™ operating system (95, 98, Me, NT4, and 2000). With Multimedia converter you can convert common video file formats to NIM format. Multimedia converter is included on the CD-ROM of the sales package.

You can use PC Suite for Nokia 9210 Communicator to copy converted files to your communicator via a serial cable or an infrared (IR) connection. You can also store video clips directly to a memory card, if you have an appropriate peripheral installed on your computer. NIM files can also be stored on a Web server and referred from an HTML page. This enables users to download NIM files into the communicator.

This guide is organised as follows: Section 2 describes how to create suitable video files for the Nokia 9210 Communicator. Section 3 guides you in using Multimedia converter. Section 4 gives a brief overview of how video files can be transferred to the communicator. Section 5 provides a brief insight into audio-visual compression technology and explains how the technology is applied in Multimedia converter and Video player. The used abbreviations and their expanded forms are listed in section 6. Finally, section 7 gives answers to the most frequently asked questions about video in the Nokia 9210 Communicator.

2. How to Digitise Video

This section describes how to produce suitable files for Multimedia converter from audio-visual source material. First, you need a signal source, such as a video camera or a video cassette recorder. If your signal source is a video camera, please consider using a fully digital system following the DV specification, for example. Digital video cameras typically provide better video quality than analog cameras. Next you should connect your signal source to the computer you use for digital video editing. If your signal source is analog, you need a video capturing card that digitises the analog video signal. Please refer to the manual of your capture card for details on how to connect your signal source to the capture card. The capturing process is controlled by a capturing application that is typically provided with the capture card. Please refer to the manual of your video capturing program for details on how to drive the capturing process. If your signal source is digital, you can use a digital interface to transfer digitally coded video to your computer. For example, the IEEE 1394 interface is typically used to transfer DV-formatted video to a computer. Notice that you may have to purchase a digital interface card to your computer, as your computer may not include such an interface by default. When you complete these steps, you should have a video file that you can play in your computer.

You may want to edit the digitised video file using a video editing application. For example, you can combine several video files to one longer file or cut some parts of a file.

When capturing a file or storing an edited file for conversion to NIM format, please remember the following recommendations:

- The preferred video file format is Microsoft Audio-Video Interleaved (AVI).
- The video image size should be at least 176x144, preferably more. Sizes larger than 320x240 give no advantage when converting files to NIM format.
- The video frame rate should be as high as possible. Keep the original rate of the material (typically 25 or 30 Hz) if possible.
- The video coding format should be either uncompressed 24 or 32-bit RGB or lightly compressed using a common codec. For example, Intel Indeo 5 with the quality setting around 90 (out of 100) is a good choice for the 320x240 image size. Please do not use 8-bit or 16-bit RGB formats.
- The video image aspect ratio should be 4:3 in order to produce images with maximum size for the display of the Nokia 9210 Communicator.
- The pixel aspect ratio should be square (1:1).
- The preferred audio format is uncompressed 16-bit mono at an 8 kHz sampling rate. If you cannot set up the preferred audio format, use a sampling rate of at least 22 kHz. Preferably, audio should be stored in an uncompressed format or compressed with a high sound quality. Mono sound is sufficient, and stereo sound does not result in any improvements to quality.
- The audio recording volume should be relatively high due to limited amplifying capability of the Nokia 9210 Communicator.

When editing or shooting a video clip, please remember the following hints to achieve a more pleasing end result. These hints help you to produce smaller files for local playback.

- Avoid using a handheld camera.

- Avoid shots that last less than two seconds.
- Avoid zooming and camera rotation.
- Minimise spatial complexity of video images. Examples of spatially complex objects include trees and crowd scenes.
- You may try low-pass filtering of images to get rid of small details as well as capturing and compression artifacts.
- Avoid rapidly moving objects.
- Avoid stylish scene transitions, such as fades and wipes. Use abrupt scene changes instead.
- Crop black edges if possible. For example, if the image aspect ratio was originally 16:9 but at some point it was converted to 4:3 with black edges, it is better to remove the black edges and provide a 16:9 image sequence for Multimedia converter.
- Avoid using titles and captions. Use a large font size if you have to overlay video with text.
- Use smooth and even lighting to avoid complex shadows and hotspots.

Note that the colour display of your communicator is sensible to lighting conditions and viewing angle, and therefore it is recommended to create relatively bright video clips for the communicator. Moreover, it is worth ensuring that the dynamic colour range of the image sequence is maximal, i.e., the darkest areas in the sequence correspond to (0,0,0) in the RGB space and the lightest areas correspond to (255,255,255) in the 24-bit RGB space.

3. Multimedia Converter for the Nokia 9210 Communicator

With Multimedia converter you can convert common video file formats to NIM format playable on the Nokia 9210 Communicator. When converting a video file, the size of the converted file is usually much smaller than the size of the original file. You can also convert common audio file formats to WAV format and optimise existing WAV format sound files for use on the Nokia 9210 Communicator. You can use converted WAV files as ringing tones, for example.

Use the following procedure to convert a video file for local playback:

1. In the menu, select *File > Open*. A dialog opens. Browse for the video file you want to convert and click **OK**. Alternatively, you can drag a file to Multimedia converter. The file information is transferred to *Input file info* fields on the left side of the dialog. The fields display the file type, file size, and playback time.

2. Select one of the two options in the *Video quality* field:

    - *Smooth motion, small image size*. For the maximum achievable frame rate.
    - *Normal motion, normal image size*. For normal image size and slightly less fluent movement.

3. If the file contains both video and sound, select one of the two options in the *Audio format* field:

    - *Music, normal compression*. For high quality audio containing music and/or another more demanding material.

- *Speech, high compression*. If the original file contains only speech or if small file size is more important than sound quality.

Note that highly compressed speech is more complex to decompress than normal music compression. Consequently, less processor time is available for video decompression and the video frame rate may drop a bit when speech compression is used.

4. Click **Convert**. A dialog pops up and shows how the conversion progresses. The file information of the converted file is transferred to *Converted file info* fields on the right side of the dialog.

5. In the menu, select *File* > *Save as* to save the new file.

You can also preview both the original file and the converted file. Note that the video is scaled to fit into the preview dialog.

Note: Multimedia converter supports most files that you can play in Windows Media Player. The exact set of supported files depends on the version of Media Player installed in your system and on the installed audio and video codecs. Typically, Multimedia converter is able to support most AVI, WAV, MPG, and MP3 files. If you cannot open a video file in Multimedia converter, please try the following procedure:

1. Open the file in Windows Media Player. If the file cannot be played, follow the next phases in the list. Otherwise, try to convert the file to an AVI file with common video and audio coding formats using a video editing application.

2. Check the version of Windows Media Player. If the version number is less than 6, it may be worth downloading a fresh version from Microsoft's Web site.

3. Try to reopen the file in a fresh version of Windows Media Player. If you do not have all the necessary codecs installed, Windows Media Player attempts to download and install codecs automatically.

4. If you still cannot open the file in Windows Media Player, contact the originator of the file.

Note: You cannot play a NIM file on a PC.

4. How to Transfer Video Content to the Nokia 9210 Communicator

4.1 Downloading Video Clips from the Internet

When you plan to embed downloadable NIM video files to Web pages, the Web server in use needs to be configured to provide the correct MIME type for NIM files. To do this, associate the extension .nim with the MIME type video/vnd.nokia.interleaved-multimedia. See your Web server documentation for more information or consult the server administrator.

When you are about to create a Web page containing video content for the Nokia 9210 Communicator, create the desired NIM files first using Multimedia converter. Then, store the NIM files to a Web server. Finally, create an HTML page with hyperlinks to the NIM files.

When you browse Web pages with your communicator and select a hyperlink pointing to a NIM file, the Web browser downloads the clip and starts Video player automatically. Note that downloading a NIM file typically lasts longer than the playback duration of the file. The downloading time also depends on the data rate of the established GSM

data call. The data rate of the call can vary from 9.6 kbps to 43.2 kbps depending on the network capabilities and conditions.

You can also download NIM files to your computer via a wired Internet connection, such as a landline modem connection. Then, you can copy the files to your communicator using PC Suite for Nokia 9210 Communicator.

Note: You cannot stream NIM files to the Nokia 9210 Communicator. Streaming refers to simultaneous downloading and playback.

4.2 Transferring Video Clips from a PC to the Nokia 9210 Communicator

You can use PC Suite for Nokia 9210 Communicator to copy NIM files to your communicator. For more information, see the 'Your communicator on your PC's desktop' section in the Nokia 9210 Communicator User guide.

5. Appendix A. Audio-Visual Compression for the Nokia 9210 Communicator

5.1 Basics of 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. The redundancy in video sequences can be categorised into spatial, temporal and spectral redundancy. Spatial redundancy means 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. 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. Spectral redundancy means the correlation between the different colour components of the same image. However, efficient enough compression cannot usually be reached by just 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 such information which 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 images that can or cannot utilise temporal redundancy reduction. Compressed images, which do not utilise 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 precise enough, and therefore a spatially compressed prediction error image is also associated with each INTER frame.

In video coding, there is always a trade-off 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 the images: the more difficult the image is to compress, the worse the image quality is. If variable bit rate is allowed, the encoder can maintain a constant video quality.

ITU-T H.263 video codec utilises 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 and other coefficients are close to zero. For

example, a constantly coloured 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 quantised, which means that the coefficients are rounded to certain quantisation levels. The fewer possible quantisation levels there are, the fewer bits it takes to represent a quantisation level. An approximation of the original block of pixels can be restored from the coded DCT quantisation levels by applying an inverse DCT transform. The fewer quantisation levels were used, the worse the quality of the reconstructed image is. ITU-T H.263 allows 31 quantisation step sizes that are controlled by the so-called quantisation parameter.

5.2 Video Coding for the Nokia 9210 Communicator

Multimedia converter processes the video track of an input file as follows:

- The video track is extracted from the input file and decompressed if necessary.
- If the image size in the input video sequence is too large for the Nokia 9210 Communicator, the images are scaled to a suitable size.
- The images are converted from the RGB colour space to the YUV colour space defined in ITU-R Recommendation BT.601-4. The H.263 coder requires YUV input images.
- Selected images are compressed according to ITU-T Recommendation H.263.

As the steps above are relatively trivial and non-controllable except for the last step, we only consider video coding in this section. At first, we explain some general issues, such as the selected frame rate and quantisation strategies. Then, we review how you can affect the video compression process.

There are at least two trade-offs that have to be made during compression of a video sequence. First, the higher the video quality, the larger the resulting file is and the longer it takes to transfer it to the communicator. Second, decompression of a video file is a relatively complex process, and the Nokia 9210 Communicator has limited processing power when compared to desktop computers, for example. Thus, the communicator is not capable of decoding a high-frame-rate, large-image-size video file, but rather frame rate, image size, and image quality have to be tuned so that real-time decompression and playback are possible. Multimedia converter generates files with a suitable frame rate, image size, and image quality for most purposes. The following paragraphs give more details how Multimedia converter selects the critical coding parameters.

The colour display of the Nokia 9210 Communicator allows the viewing of very sharp pictures due to its small pixel size. However, due to moving pictures and high frame rate, very sharp images are not necessary when playing back a video clip. Thus, a relatively small image size is used when coding a clip. Images are then scaled onto the communicator's display by a factor of four. For example, if the size of the coded image is 128x96, the size of its display rectangle is 256x192.

A constant rate of coded pictures is applied throughout the video sequence if possible. The frame rate is selected so that the sequence can be decoded in real-time in the communicator. However, if the input sequence has a non-constant frame rate, the frame rate of the converted NIM file is likely to be non-constant as well.

Since converted files are not meant to be streamed or transferred to the communicator in real-time, there are no hard bit-rate requirements for them. This enables constant image quality and ensures a pleasant viewing experience.

In other words, a constant quantisation parameter is used throughout a video sequence resulting into a file having variable bit rate.

Multimedia converter provides two options to control video compression:

- *Smooth motion, small image size*.
- *Normal motion, normal image size.*

When the first option is used, the maximum amount of pixels in the compressed image is 8960, i.e., the image width can be 112 pixels and height 80 pixels, for example. The maximum width of the picture is 128 pixels and the maximum height is 96 pixels. The actually coded image size depends on the aspect ratio of the input images. Ten frames per second are coded.

When the second option is used, the maximum image size is 128x96. The image size that is actually coded depends on the aspect ratio of the input images. 7.5 frames per second are coded.

The resulting file size and bit-rate depend largely on the contents of the image sequence. Typically, video bit rates vary from 50 kbps to 100 kbps. "*Smooth motion, small image size*" option tends to reserve slightly more bits than " *Normal motion, normal image size*".

5.3 Basics of 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 digitised 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.

Digitised audio can be compressed in various ways. A simple coding method is to use an adaptive step size to quantise audio samples. Such a technique is used in 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 kbps. Another simple audio coding method is A-law PCM, which uses a logarithmic quantisation step size and reserves 8 bits per sample.

More advanced audio coding methods take advantage on the human psychoacoustic model. Parts of the audio signal are barely audible and can be discarded or compressed. Typically, the advanced coding audio methods are categorised 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.

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 kbps. Some of the modes are speech codecs specified for other standards. For example, AMR at 12.2 kbps is the same speech codec as GSM enhanced full-rate codec.

5.4 Audio Coding for the Nokia 9210 Communicator

The Nokia 9210 Communicator supports an output of 8 kHz mono sounds. Therefore, the following procedure is used to compress high-quality audio tracks in Multimedia converter:

- The audio track is extracted from the input file and decompressed if necessary.
- Stereo sound is converted to mono sound.
- The audio sampling rate is converted to 8 kHz.
- Audio is compressed.

As the steps above are relatively trivial and non-controllable except for the last step, we only consider audio coding in this section. For audio-video files, you can select one of the two options in the "*Audio format*" field:

- *Music, normal compression*. For high quality audio containing music and/or some other more demanding material. IMA ADPCM is used, resulting in 32 kbps.
- *Speech, high compression*. If the original file contains only speech or if a small file size is more important than the sound quality. The AMR 10.2 kbps mode is used.

6. Appendix B. Abbreviations

ADPCM	Adaptive Pulse Code Modulation
AMR	Adaptive Multi-Rate speech codec
AVI	Microsoft Audio-Video Interleaved file format
CD-ROM	Read-Only Compact Disc
DCT	Discrete Cosine Transform
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
kbps	kilobits per second
MPEG	ISO/IEC Moving Pictures Experts Group
MPG	File name extension for MPEG-1 file format

MP3	MPEG-1 Audio Layer 3 audio coding
NIM	Nokia Interleaved Multimedia file format
PC	Personal Computer
PCM	Pulse Code Modulation
RGB	Red-Green-Blue colour space
WAV	Microsoft waveform audio file format
YUV	Colour space, Y is the luminance or gray-scale component, U and V are chrominance or colour difference components

7. Appendix C. Frequently Asked Questions

Q: What is Multimedia converter? Which formats does it support?

Multimedia converter runs on any 32-bit Microsoft Windows™ operating system (95, 98, Me, NT4, and 2000). With Multimedia converter you can convert common video file formats to NIM format playable on the Nokia 9210 Communicator. You can also convert common audio file formats to WAV format and optimise existing WAV format sound files for use on the Nokia 9210 Communicator. The communicator can use converted WAV files as ringing sounds, for example. The converted video files follow the Nokia Interleaved Multimedia (NIM) format and can be played with the Video player application. Multimedia converter ensures that the audio and video files can be played on the communicator and that the converted files achieve the best possible audio-visual quality. In addition, the size of the converted audio or video file is usually much smaller than the size of the original file.

Q: Which video formats are supported?

The player supports the Nokia Interleaved Multimedia (NIM) format. You can create NIM files using Multimedia converter for the Nokia 9210 Communicator.

Q: Can I play AVI files?

You have to convert AVI files to NIM files first, since Video player inputs only NIM files. Please use Multimedia converter for the Nokia 9210 Communicator.

Q: Can I play MP3 files?

You have to convert MP3 files to WAV files first using Multimedia converter for the Nokia 9210 Communicator.

Q: Can I play NIM files on a PC?

No, you can play NIM files on your Nokia 9210 Communicator only.

Q: Why doesn't Video player support MPEG-4?

Nokia 9210 Communicator video capabilities are based on ITU-T H.263 video codec, which is technically similar to MPEG-4 visual coding.

In order to support a wide range of video codecs, handling of video files is done in two stages. First, Multimedia converter is used to convert video files to the format supported in the Nokia 9210 Communicator. The converter supports a wide range of input files and video codecs. The converter compresses video tracks according to the H.263 coding standard and outputs files in the Nokia Interleaved Multimedia (NIM) format. Second, the created NIM files are transferred to the communicator and played back using the Video player application.

Q: Which are the most important / most common video 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 codestream. The file format is a way of organising video and audio codestreams in such a way that they can be accessed for local decoding and playback or streamed over a transport channel.

The following list mentions some of the most common video coding formats today:

- MPEG-1, which is used in video CDs, for example.
- MPEG-2, which is used in DVDs and satellite broadcasts, for example.
- H.263, which is commonly used in video conferencing systems.
- RealVideo, which is the video coding format of the popular RealSystem video streaming application.
- Microsoft MPEG-4, which is an evolution of the MPEG-4 standard and commonly used in Windows Media.

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

- Microsoft Audio-Video Interleaved (AVI).
- Apple QuickTime file format.
- MPEG-1 file format.

Q: How about in mobile devices?

There is no widely used video format in mobile devices today. The standardisation of the terminal multimedia capabilities for the third generation mobile networks is underway. It seems that H.263 video coding and AMR speech coding are key technologies in upcoming standards and terminals.