Optimizing A/V Content For Mobile Delivery

Media Encoding using Helix Mobile Producer 11.0

November 3, 2005

Contents

1. Introduction 3
2. Source Media 3
    - Encoding from a File 3
        - Color Format
        - Audio
    - Capturing from Devices 4
    - Video Resolution & Aspect Ratio 4
    - Types of Content 5
        - News
        - Movies & Music Videos
        - Sports
3. Output Media Formats 5
    - 3GPP 5
        - 3GPP r6
    - RealMedia 6
    - Encoder Properties 7
        - Encoding Complexity
        - Video Frame Rates
4. Delivery Considerations 8
    - Types of Mobile Networks 8
        - GPRS
        - EDGE and UMTS
    - 3GPP Content Profiles 8
        - Basic profile
        - Streaming-server profile
        - Progressive-download profile
    - Maximum Packet Size 9
    - Broadcast Delivery 9
5. Setting Tables 10
    - Export Formats 10
    - Video Codec Settings 10
    - Video Codec Levels 10
6. References 11

1 Introduction

This document provides recommendations on how to encode media content for mobile networks and devices using the Helix Mobile Producer 11.0 (HMP 11). It is recommended that you become familiar with the Helix Mobile Producer User's Guide in order to learn the basics of encoding media before reading this document.

This white paper is not comprehensive. Instead this paper touches on different aspects of media production and streaming to help media producers prepare their content for mobile audiences. Notably, this white paper does not discuss specific mobile devices. Due to the number of devices and their short life cycle on the market, we recommend you refer to the device owners manual or other manufacturer provided documentation to determine the media capabilities of your target devices.

The remainder of this document is divided as follows:

Section 2 discusses the optimal properties of source media files and capture devices.

Section 3 covers the two primary output media file types: 3GPP (including 3GPP r6) and RealMedia. This section also covers come of the more advanced settings available in HMP.

Section 4 touches briefly on network delivery limitations and recommended HMP settings.

Section 5 provides summaries of available settings and codec attributes.

Section 6 includes references to helpful white papers, web sites, and other resources to assist you with content production and mobile streaming.

2 Source Media

The format and quality of your source material is a primary factor in the quality of the final encoded result. Helix Mobile Producer supports both static (file-to-file) encoding and live (device-to-file and device-to-broadcast) encoding. This section discusses the preferred file formats and characteristics, as well as recommended capture hardware.

Encoding from a File

Helix Mobile Producer can read a large number of source file types, including many compressed formats such as MPEG1, MPEG2, MPEG4, DV, Sorenson, M-JPEG, and more. However, while HMP supports low-quality compressed sources, using them will yield a lower-quality encoded result. This occurs because any artifacts that exist in the source file are carried over to the encoded output. Also, since compressed source video generally results in a messier image (caused by limitations of bit rate and the codec) the encoder will be forced to waste bits duplicating undesirable artifacts from the source file.

For the best possible results, it is recommended you encode from source files in one of the following formats:

- Uncompressed QuickTime (.mov)
- Uncompressed AVI

Color Format

The native color space of the Helix Mobile Producer is I420, a form of YUV. When possible, produce a source file with this same color space to avoid a color conversion during the

encoding process. The encoding process will be slightly faster without the encoder performing a color conversion, but there is no adverse effect on the quality of the encoded file so long as you use a 24-bit source.

Audio
The quality of the source audio is equally important. For best results, use uncompressed 44.1 kHz 16 bit sources. Using an uncompressed source is particularly important when encoding to a high bit rate output, such as AAC at 32 Kbps or higher. With AAC encoding at this bit rate, there will be a noticeable difference in the audio quality if your source is only 22 kHz or 8 bit. If you are encoding AMR-NB voice, however, there should be no noticeable difference. (The audio codecs available in HMP are listed in section 3.) For audio-only encoding, the recommended source format is a 44.1 kHz 16 bit WAV file.

Capturing from Devices

For the best quality results when capturing from live sources, do not use web cams or standard PC microphones. You will create a much higher quality result using a good quality DV camera connected to a mid-range to high-end video capture card.

HMP 11 is compatible with audio and video capture devices that support the Windows Driver Model (WDM) and, for video capture, can produce video in I420 color format.

The following capture devices have been tested with HMP 11:

- Viewcast Osprey 100
- Viewcast Osprey 200/210/220/230
- Viewcast Osprey 500 DV
- Viewcast Osprey 540
- Logitech Quickcam Pro 4000 (Not recommended to quality reasons)
- Pinnacle PCTV Pro
- Hauppauge WinTV Model 37281
- SoundMax Audio
- RealTek Audio (included on some system boards as built-in audio)

Video Resolution & Aspect Ratio

The width and height of the video are an important consideration. The majority of mobile devices currently on the market are designed with screens that can play video with a width of 144 to 160 pixels. When deciding on the width and height of the video, you should first determine the optimal width based on the playback device capabilities, then set the height based on the aspect ratio of the source video.

Aspect ratio is the ratio of the width to the height of the video image. For example, a resolution of 640x480 (standard NTSC television) has an aspect ratio of 4:3, or 1.33 (640/480). If the aspect ratio if your source video does not match the aspect ratio you encode the video to, the image will appear stretched. For example, if you have a 640x360 (16:9 or letterbox source file) and encode it to 160x120 or 4:3, the image will stretch the height to fill the output size, and the video will not look right. Instead, you would want to use the same aspect ratio of 16:9, or 160x90.

Alternatively, you can crop the image width to change the aspect ratio, but this will cut out part of the video. See the HMP User's Guide for more information about using the resize and crop filters.

Types of Content

Some of the settings you choose for encoding will depend on the nature of the content. A few guidelines are provided below based on the general characteristics of the media.

News
News content is general characterized by a talking head with relatively static backgrounds with limited motion and scene changes. News content is probably the least demanding on the encoder. This means that to achieve good quality with news content, you may be able to use a lower bit rate than you would need for movie or sports content. In general, a voice audio codec like AMR will produce better results for this type of content at lower bit rates (under 100 kbps).

Movies & Music Videos
Content of this type tends to include frequent scene changes, periods of fast action, and periods of slower motion. Movies for these reasons tend to be more demanding of the encoder than news content. Audio also tends to be very mixed with dialog, music, and environmental sounds. A music codec like AAC will produce much better results than a voice codec like AMR.

Sports
Sports content like football is generally characterized by fast pans and quick scene changes with complex backgrounds (crowded stands). The camera usually follows the ball, meaning just about every object on screen is in motion. This creates heavy demands on the encoder. Audio in a sporting event is usually comprised of an announcer as well as environmental noise, like the crowd. The announcers are conducive to a voice codec, but the environmental noise may cause too much distortion in the voice codecs, so a Mono music codec like AAC is most appropriate.

These are only general suggestions. You will have to experiment with different encoding settings to achieve the desired result.

3 Output Media Formats

This section provides suggested settings for both 3GPP (.3gp) and RealMedia (.rm) files. Although Helix Mobile Producer is capable of generating other file types, only 3GPP and RealMedia formatted A/V files are compatible with most mobile devices.

HMP can produce both 3GPP standard format content, as well as RealAudio/RealVideo content. The format you choose should be based on the capabilities of your target devices. You will, in general, get better quality with RealAudio/RealVideo with compatible devices.

3GPP

3GPP is a standards based file format defined by the 3rd Generation Partnership Project. HMP 11 can produce files that conform to either 3GPP Release 5, 3GPP2, or 3GPP Release 6.

3GPP r5 has been in the market for some time, and there are a broad number of mobile devices that support this standard. The baseline codecs defined in 3GPP r5 are H.263 Profile 0 for video and AMR-NB for audio. This represents the minimal requirement for a device to claim 3GPP r5 media playback support. H.263 is the simplest, and lowest compression video codec in HMP. It offers the broadest reach for compatibility, but at the expense of providing lower quality than the other codecs.

3GPP Video Codecs in HMP 11:

- H.263 Profile 0 - Best compatibility, lowest quality.
- MPEG-4 Simple Profile - Good compatibility, fair quality.
- H.264 Baseline - Limited number of devices, but best quality for 3GPP. Only available in 3GPP r6.

3GPP Audio Codecs in HMP 11:

- AMR-NB - Best compatibility, good for voice at lower bit rates.
- AMR-WB - Good compatibility, suitable for voice at low to medium bit rates.
- QCELP - 3GPP2 format only, suitable for voice content at low bit rates.
- AAC - Good compatibility, best for music or mixed content at medium bit rates.
- AAC Plus - Limited to 3GPP r6, higher quality than AAC.
- Enhanced AAC Plus - Limited to 3GPP r6, best quality for stereo music at all mobile bit rates.

3GPP r6

The 3GPP r6 format is a superset of r5 that adds the following capabilities:

- H.264 video codec support
- AAC Plus and Enhanced AAC Plus audio codecs
- Multi-rate encoding support
- Metadata (Title, Author, Copyright, etc.)

Note that 3GPP r6 is a relatively new standard. At the time of this writing, very few devices on the market support the H.264 video codec or the AAC Plus audio codec. Check the capabilities of your device before deciding which codecs to use.

RealMedia

RealMedia (or RealAudio and RealVideo) is RealNetwork's proprietary media format that is widely available on a growing number of mobile devices and all major desktop platforms. RealMedia is suitable for encoding at very low bit rates for mobile devices and networks, all the way up to broadband and HDTV. Compared to 3GPP standards codecs, RealVideo provides as good or better quality than H.264, but has lower requirements on the decoder, shorter encode times, and at the time of this writing, is supported on more mobile devices than H.264.

The RealMedia format provides full support for multi-rate content creation, called SureStream. In addition to supporting file to file multi-rate creation, you can also produce live broadcast multi-rate streams in RealMedia format, a capability not presently available in 3GPP. RealMedia broadcasting also allows you to specify either a Push or Pull connection to the Helix Server, and provides options for automatic reconnect if the server connection is ever lost.

The other settings and options in HMP (such as video prefilters, encoding complexity, and 2-pass) are available with both 3GPP and RealMedia encoding. See section 5 Settings Tables for a detailed breakdown.

Encoder Properties

This section covers two properties of the encoder. First, the encoding complexity, which provides options for balancing encode time with quality. Second, video frame rate, which is important to set correctly for your audience to ensure a high quality customer experience. The full set of encoder properties are covered in the HMP User's Guide.

Encoding Complexity

Each video codec offers an encoding complexity parameter with three levels (Low, Medium, and High). Higher complexity will create a better quality result, but will take longer to encode. In the case of live capture or broadcast, higher complexity will require a computer with a faster processor. Each video codec is effected differently by this setting, as defined below.

H.263 & MPEG4

- Low: Baseline, motion vector search range of 16, fast motion estimation.
- Medium: ~5% longer to encode, motion vector range of 64, improved motion estimation.
- High: ~5% longer to encode than medium, motion vector range of 128.

H.264

- Low: Baseline, 1 reference frame, fast motion estimation, 3 different macro-block sizes.
- Medium: ~15% longer to encode, improved motion estimation, 4 different macro-block sizes.
- High: ~55% longer to encode than medium, 3 reference frames, 6 macro-block sizes.

RealVideo 8, 9, & 10

- Low: Baseline, 65% quality index.
- Medium: 2x longer to encode than low, 75% quality index.
- High: 2x longer to encode than medium, 85% quality index. Roughly 30% quality improvement over low.
- RV10 low complexity is the same as RV9 high complexity.

You may encounter problems with H.264 live encoding if your system is not fast enough to keep up with the demands of the codec. If the CPU load in the system monitor is showing close to 100%, the codec may be overloading the system. If this occurs, you may see dropped frames in the resulting stream, and audio may fall out of sync. To fix the problem, select a lower encoding complexity level, use a faster machine, or specify a different codec.

Note that the RealVideo codecs all contain scalability code that will automatically reduce the encoding complexity in real time during live encodes if the system is unable to maintain the specified encoding complexity level. This scalability code is only activated when performing a live capture or broadcast. For file to file encodes, the encoding complexity will never adjust dynamically.

Video Frame Rates

Because mobile devices have limited processing resources, most are only capable of much lower video frame rates than what can be achieved on desktop PCs. For very low bit rates (15 – 20 kbps) a frame rate of 5 or 6 frames per second (fps) is best. Even at higher bit rates (50 kbps) some devices will not be able to process more than 7 or 8 fps. You may need to experiment with your specific mobile devices to determine the best target frame rate.

To avoid jerkiness and gaps in motion, you should always use an encoded frame rate that is evenly divisible by the frame rate of your source. For example, if your source file is 30 fps, you will get a better result if you encode to 6 fps than if you encode to 7 fps. This is because 7 is not evenly divisible by 30, which may cause momentary pauses in the motion of the encoded video. Specifying 6 fps will result in a more consistent and overall better encoded result.

4. Delivery Considerations

If you are encoding content to be delivered over a wireless or wire line network, you must be aware of the practical limitations of that network. The table below summarizes recommended settings for some wireless networks based on a general observation of the real-world capabilities for that type of network. These are general observations; different networks will have varying limitations.

Note that you must verify you are using settings that are appropriate not only for the network, but for the playback device.

Types of Mobile Networks

The following sections briefly describe the capabilities of some common types of mobile networks as they relate to streaming media. Different real world networks may have different capabilities than what is documented here.

GPRS

GPRS networks have a theoretical maximum bit rate of 115 kbps. However, the real-world maximum bit rate will probably be closer to 80 kbps, and reliable streaming will likely be limited to 18 - 22 kbps on most networks.

EDGE and UMTS

EDGE or UMTS networks have a theoretical maximum bit rate of 384 kbps or higher. However, similar to GPRS, the maximum reliable bit rate for streaming will probably be about 50 kbps (EDGE) to 120 kbps (UMTS) on most networks.

3GPP Content Profiles

The 3GPP standard defines a number of different profiles, each of which defines a set of properties and limitations. The profile you select needs to match the method used to distribute the media. The profiles supported by HMP are described below.

Basic profile

This is a general profile best suited for media that will be used only for local playback, not for streaming or download delivery.

Streaming-server profile
This profile is required for media that will be broadcast or streamed on-demand. When this profile is specified, the media is created with special hint tracks along with the regular media tracks. The hint tracks provide additional information to the Helix Server to optimize the streaming delivery to clients. This is the only profile that is available when encoding to 3GPP r6 Multi-rate.

Progressive-download profile
This profile is similar to the basic profile, except that the header information in the file is included at the beginning, and the media tracks, if there are both audio and video tracks, are interleaved. This enables the file to be downloaded and for playback to begin before the download is finished. Files created with this profile do not contain hint tracks, resulting in a slightly smaller file size than streaming-server profile files. Because progressive-download files do not include hint tracks, they are not recommended for streaming delivery.

These profiles apply to 3GPP r5, 3GPP2, and 3GPP r6. RealMedia does not have any similar concept of profiles. Instead, RealMedia files are always optimized for streaming, but can also be used for local playback or progressive download delivery.

Maximum Packet Size

Packet size is an important consideration for streaming delivery. Basically the packet size is how large of a chunk of data is sent with each packet. Ideally, you should strive to have the least number of packets for optimal delivery on your network. The reason for this is that each packet adds overhead to the actual streaming data. The fewer packets you have, the less overhead. However, if you set the maximum packet size too large, routers on the network may either break up the packet, creating more overhead, or truncate the packet, creating loss.

The Chapter 8 in the HMP 11 User's Guide contains a detailed explanation of Maximum Packet Size, and how to determine the correct setting for your network. The default setting of 1400 bytes is a typical safe setting for most IP networks.

Note that this setting is only for limiting the largest packet the encoder will create. Most packets will actually be smaller, especially at lower bit rates and for audio streams.

Broadcast Delivery

Broadcast delivery refers to live streaming to a Helix Server. Live streaming in HMP 11 works in two different ways between 3GPP and RealMedia. The details of each are explained in chapter 7 of the HMP 11 User's Guide. Instructions on setting up broadcasts are included in the Helix Server Administrator's Guide.

5 Setting Tables

The following tables summarize available parameters in HMP 11, including details on the profiles and levels supported by each video codec.

Export Formats

Export types		Input		Output			Video				Audio Codecs								Output Options						
Name	File Extension	File Input	Capture Input	File Output	Broadcast Output	Max # of Audiences	MPEG4 Simple Profile	H.263 Profile 0	H.264 Baseline Profile	RealVideo 8, 9, 10	AMR Narrowband	AMR Wideband	AAC Low Complexity	AAC Plus	Enhanced AAC Plus	QCELP	MP3	RealAudio	Packet Size	Progressive Download	Streaming-server Profile (hinting)	Basic Profile	Metadata	RM Latency Mode	RM Loss Protection
3GPP r5	.3gp	x	x	x	x, rtp	1	x	x			x	x	x						x	x	x	x	x		
3GPP2	.3g2	x	x	x	x, rtp	1	x	x								x			x	x	x	x	x		
3GPP r6 singlerate	.3gp	x	x	x	x, rtp	1	x	x	x		x	x	x	x	x				x	x	x	x	x		
3GPP r6 multirate	.3gp	x		x		8	x	x	x		x	x	x	x	x				x	x	x	x	x		
ISMA Profile 0	.mp4	x	x	x	x, rtp	1	x		x				x						x				x		
AMR file	.amr	x	x	x		1					x	x													
QCP file	.qcp	x	x	x		1										x									
MP3 file	.mp3	x	x	x		1											x						x		
RealMedia	.rm	x	x	x	x, helix	8				x								x	x				x	x	x

Video Codec Settings

Video Codec	Video Codec Options							Video Dimensions			Compatible Export Formats								
Name	CBR	VBR	2-pass	Encoding Complexity	Video Mode	Keyframe Distance	Frame Rate (fps)	Min Frame Size	Max Frame Size	Mod Value	3GPP r5	3GPP2	3GPP r6 Singlerate	3GPP r6 Multirate	ISMA Profile 0	AMR File	QCP File	MP3 File	RealMedia
MPEG4 Simple Profile	x	x	x	x	x	x	1-30	32x32	384x288	4	x	x	x	x	x				
H.263 Profile 0	x	x	x	x	x	x	1-30	128x96	384x288	SQCIF, QCIF, CIF	x	x	x	x					
H.264 Baseline Profile	x		x	x	x	x	1-30	32x32	384x288	16			x	x	x				
RealVideo 8, 9, 10	x	x	x	x	x	x	1-30	32x32	2048x2048	4									x

Video Codec Levels

Codec	Level	Max Bit Rate	Max Size & Frame Rate
MPEG4 Simple Profile	0	64 kbps	QCIF @ 15 FPS
	0b	128 kbps	QCIF @ 15 FPS
	1	64 kbps	QCIF @ 15 FPS
	2	128 kbps	CIF @ 15 FPS, or QCIF @ 30 FPS
	3	384 kbps	CIF @ 30 FPS
H.263 Profile 0	10	64 kbps	QCIF @ 15 FPS
	20	128 kbps	CIF @ 15 fps, or QCIF @ 30 fps
	30	384 kbps	CIF @ 30 FPS
H.264 Baseline Profile	1	64 kbps	QCIF @ 15 FPS
	1b	128 kbps	QCIF @ 15 FPS
	1.2	384 kbps	CIF @ 15 FPS, QVGA @ 20 FPS, QCIF @ 30 FPS
	1.3	384 kbps	CIF @ 30 FPS

6 References

- 3GPP Home Page, 3rd Generation Partnership Project. http://www.3gpp.org
- RealNetworks Documentation and Product User's Guides, RealNetworks. http://www.realnetworks.com/resources/documentation/index.html
- Knowledge Base Article: "How do I choose encoding settings for the best possible quality?", RealNetworks, 2004. http://real.custhelp.com/cgi-bin/real.cfg/php/enduser/std_adp.php?%20p_faqid=4226&p_created=1089824660
- RealVideo 10 Technical Overview, RealNetworks, 2003. http://docs.real.com/docs/rn/rv10/RV10_Tech_Overview.pdf
- RealVideo 10 Codec Comparison Report, RealNetworks, 2003. http://docs.real.com/docs/rn/keylabs/RV10Codec.pdf
- Streaming in Mobile Networks, TeliaSonera, 2004. http://www.medialab.sonera.fi/workspace/StreaminginMobileNetworksWP.pdf