ON2 TECHNOLOGIES, INC.

WHITE PAPER

TrueMotion VP7 Video Codec

January 10, 2005

Document Version: 1.0

On2 Technologies, Inc.
21 Corporate Drive
Suite 103
Clifton Park, NY 12065
www.on2.com

Contents

INTRODUCTION

This white paper explains why TrueMotion VP7 is the best video compression technology currently available and compares it with other technologies such as MPEG-4, H.264, and Windows Media 9. This paper covers technical advances in VP7 and improvements over the previous version of the TrueMotion algorithm.

THE CASE FOR VP7

The following are just a few reasons why VP7 surpasses all other video compression technology:

- Up to 50% image quality improvement over VP6.
- Unmatched image quality at all image sizes and for all types of material.
- Objective tests using Peak Signal-to-Noise Ratios (PSNR) and other metrics show better performance than all competing technologies including Windows Media 9, Real 10, H.264 and MPEG-4 compatible codecs across a wide range of data rates from dial-up (28.8 Kbps) to DVD and HD
- Carries no burdensome “patent pooling” restrictions or complicated external licensing fees.
- Designed to be able to run on inexpensive Digital Signal Processors (DSPs). VP7 is ideal for embedded chipsets in non-PC devices and set-top boxes.
- A purely software-based solution that can be upgraded easily.
- Compresses high-definition (HD) material with no restrictions on the encoder. VP7 can play back 1920x1080 HD material on a 2.5 GHz PC.
- Offers a “fast compress” mode that can encode in real time on a high end Pentium 4 PC with very little loss in quality.
- Powerful one-pass and two-pass data rate control features allow constant datarate or variable datarate encoding for streamed or local playback.
- Encoding modes that guarantee playback of a file transmitted at a constant datarate.
- Supports both native interlaced and progressive scan output.
- Specially designed SIMD friendly loop filtering to insure fast playback on low end processors

IMPROVEMENTS SINCE VP6

Quality Enhancing Features

- New 4x4 DCT transform technology.
- Improvements to motion compensation, which allow complex patterns of motion to be represented.
- Greatly improved key frame coding.
- A range of new motion and intra prediction modes provide for improved performance in high motion sequences.
- Sophisticated second order coding of DC information.
- Control over the relative emphasis given to coding coarse and fine detail in the image.
- An advanced new “wide pass band” motion filter that retains more detail and improves overall sharpness without introducing artifacts.
- New adaptive loop and post-processing filters provide for outstandingly low levels of both blocking and ringing artifacts.
- Superior low datarate performance at dialup rates and below.
- Greatly improved quality when coding computer generated graphics, animations or mixed sequences containing text & graphic overlays.
- Improved scene cut detection.
- Improved two-pass rate targeting.
- Sophisticated context modeling in the entropy encoder.
- More accurate coding of color information.
- Block Adaptive quantization and filtering
- Sophisticated Light Level Change Adaptation
- Two-pass spatial resampling
- Improved quality two-pass datarate control

VP7 DECODING PROFILES

VP7’s offers remarkable scalability. There are three profiles for VP7.

- **Hi-Def Profile -** designed specifically for fast playback on inexpensive processors.
  - No loop filter (replaced by post-processing)
  - Sub-pixel estimation disabled
  - Split partitioning for DMA

- **Simple Profile -** designed to ensure the best quality possible at low datarates on inexpensive processors
  - High quality but simplified loop filter
  - Ultra wide pass band sub-pixel motion estimation
  - Adaptive Spatial and Temporal Resampling to insure few artifacts at low datarates
  - Specially designed, SIMD-friendly loop filtering to ensure fast playback on low-end processors
- **Advanced Profile -** designed to ensure the best quality possible at extremely low datarates (dial-up to less than 200 Kbps).
  - Highest quality and most advanced loop filtering
  - Ultra wide pass band sub-pixel motion estimation
  - Adaptive Spatial and temporal resampling to insure few artifacts at low datarates.

Decoding Speed Compared with MPEG-4 and MPEG-2

Initial testing indicates that best-quality VP7 images have roughly the same decode complexity as the fastest MPEG-4 profiles (without B-frame prediction) at roughly the same data rate. We estimate roughly 1.8 times the complexity of MPEG-2.

Decoding Speed Compared with H.264

VP7 Advanced Profile has substantially less decode complexity than AVC/ JVT/ H.264 profiles. Our initial testing indicates that best-quality AVC/ JVT/H.264 is roughly 2-3 times more complex than VP7 Advanced Profile. “Best-quality AVC/ JVT/ H.264” in this case is defined as material encoded with a profile in which the B-frame, and CABAC entropy encoding are enabled.

VP7 gains its speed advantage over H264 in a number of ways:

- Simpler outer loops with far fewer conditionals at macroblock level
- No B-frames or any complex combination of prior coded reference frames
- Simpler loop filter with far fewer options and taps

In addition, the VP7 code is easier to decode in terms that aren’t directly related to complexity:

- **Partitioned Bitstream**. The modes and motion vectors are available separately from the coefficients. This makes it efficient to do things like set up DMA transfers without parsing coefficients.
- **Simd-friendly Filtering**. Our loop filters were specifically designed to enable character-level SIMD parallelization.

- **Less code**. VP7 contains far less code than our competitors. This makes it easier to port, and it fits more easily into code or instruction caches.
- At most two reference frames stored in memory at a time, so there are fewer data cache misses.

VP7 IS READY FOR REAL-TIME STREAMING

Encoding and decoding content on the same processor is often a requirement for real-time video applications. For example, you may have a video conferencing application that requires the processor to encode an outgoing stream while also decoding an incoming stream in real time. In such applications, the VP7 real-time encoder can be configured to let the application decide how much processor time to spend encoding.

Using the real-time encoder results in only a slight decrease in the quality of the compressed stream when compared with off-line VP7 encoding. This is especially noticeable in cases where the source material contains high motion, complex textures, and so on. Additional lossless entropy encoding tradeoffs may be required that will impede quality slightly.

- Software-only compress
  - Full D1 encoding in real-time requires only a 2 GHz Pentium 4
- Highest quality
  - Five different metrics are adjusted to insure the highest quality compression without dropping frames.
  - Better than Microsoft Windows Media 9 running in its fastest mode in both frame count and frame quality.
- Can stream using multicast or unicast using the On2 TrueCast Server.

DATARATE CONTROL

- VP7 has highly configurable and extremely accurate datarate control.
  - Choose between constant (streaming) and variable (local playback) modes
  - Constant datarate guarantee. If a file is streamed at constant datarate and buffering is as specified, the file will play without rebuffering.
  - Variable datarate controls. Lets you specify how variable you want the datarate to be by biasing towards constant or variable datarate, specifying a minimum and maximum range and by allowing you to specify that you want to be streamable at a higher datarate than you want as a target.
- VP7 uses three separate techniques to guarantee that you hit the datarates you ask for.
  - Change frame quality
  - Spatially resample input video

  - Temporally resample the input video
- VP7 maintains datarate control by means of sophisticated modeling of the buffer fullness of the client playing the file.
  - Specify amount of data that's prebuffered
  - Specify the largest amount of data that can be stored in the buffer.
  - Specify at what level of fullness its okay to start taking more drastic measures to achieve datarate (like scaling and temporal resampling)

TWO-PASS COMPRESSION

VP7 uses a set of highly tuned heuristics to produce the best possible results.

- Uses first pass to generate statistics and write them to a file.
- Second pass reads the statistics from the first pass and makes determinations like:
  - Allocates more bits to tougher sections
  - Choose better key frames.
  - Determines frame size based on relevance of the current frame to future frames

COMPARISONS

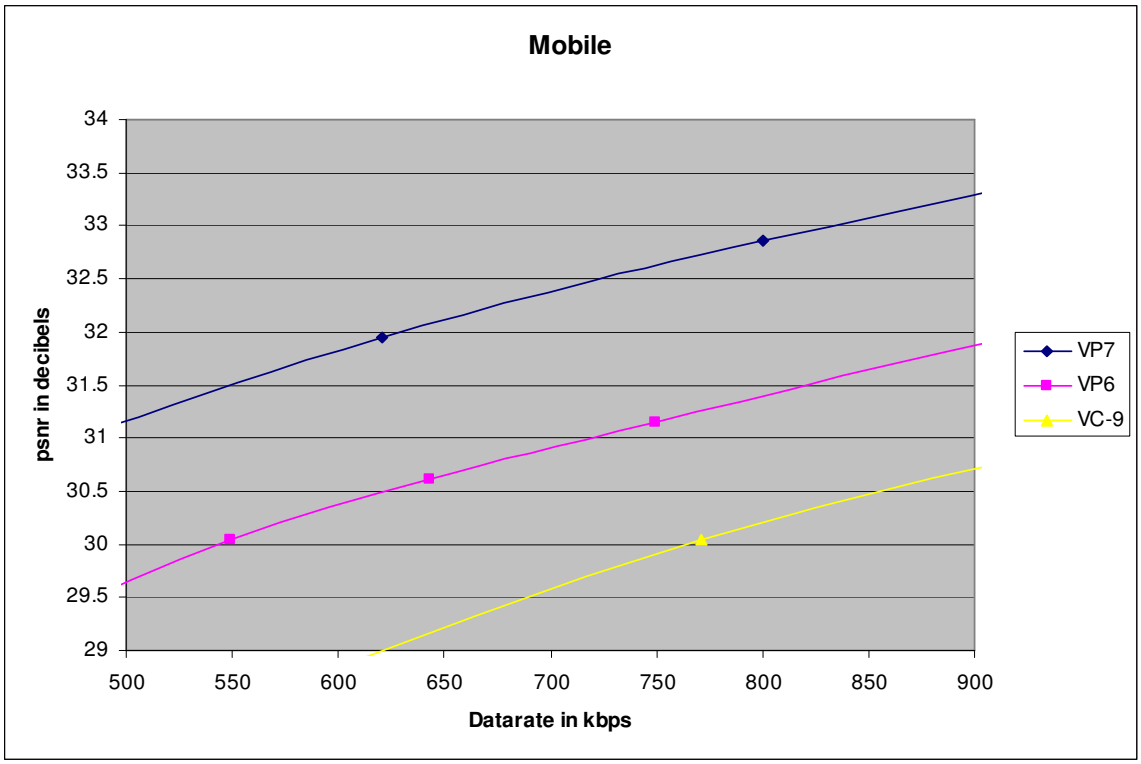
VP7's visual quality bests even the outstanding VP6 codec. On some clips it beats VP6—currently the best video codec on the market—by 50% or more. This is never more evident than on animated material.

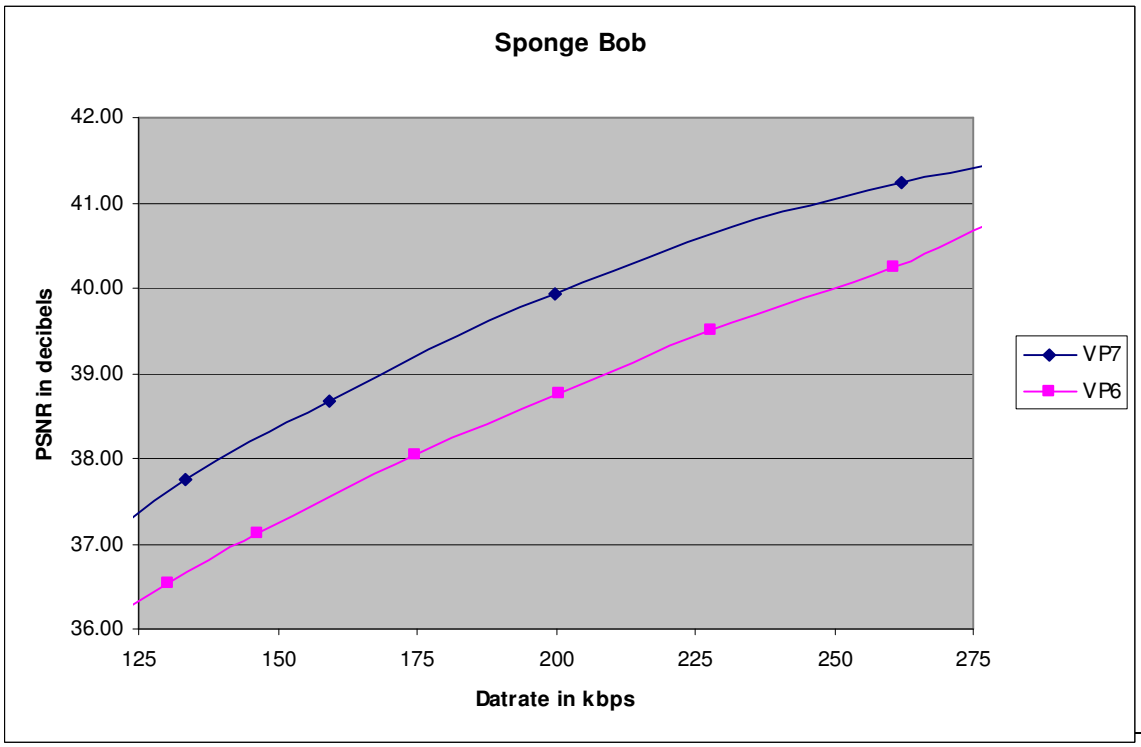
The encoding in this frame comparisons is extremely challenging.

VP6

VP7

VP7 is also better than VP6 using objective measures. The following two graphs shows comparison results between VP6 and VP7 (and, in the case of "Mobile," Windows Media 9). The "mobile" clip is a 352x240 version of a standard MPEG-2 test clip sometimes referred to as either "mobile" or "calendar." The other clip is a short animated section of "SpongeBob SquarePants."





PSNR COMPARISONS TO OTHER CODECS

VP7 doesn't just beat VP6. It beats the competition hands down, both visually and objectively. The PSNR graphs in this section exhibit this.

For all of these tests we used the following settings. All settings not mentioned were set to default values. The datarate was varied.

WMV9

- Two-pass datarate VBR
- Decoder complexity - complex
- Performance - better quality

DivX

- Version 5.2.1
- Standard Performance
- Two-pass

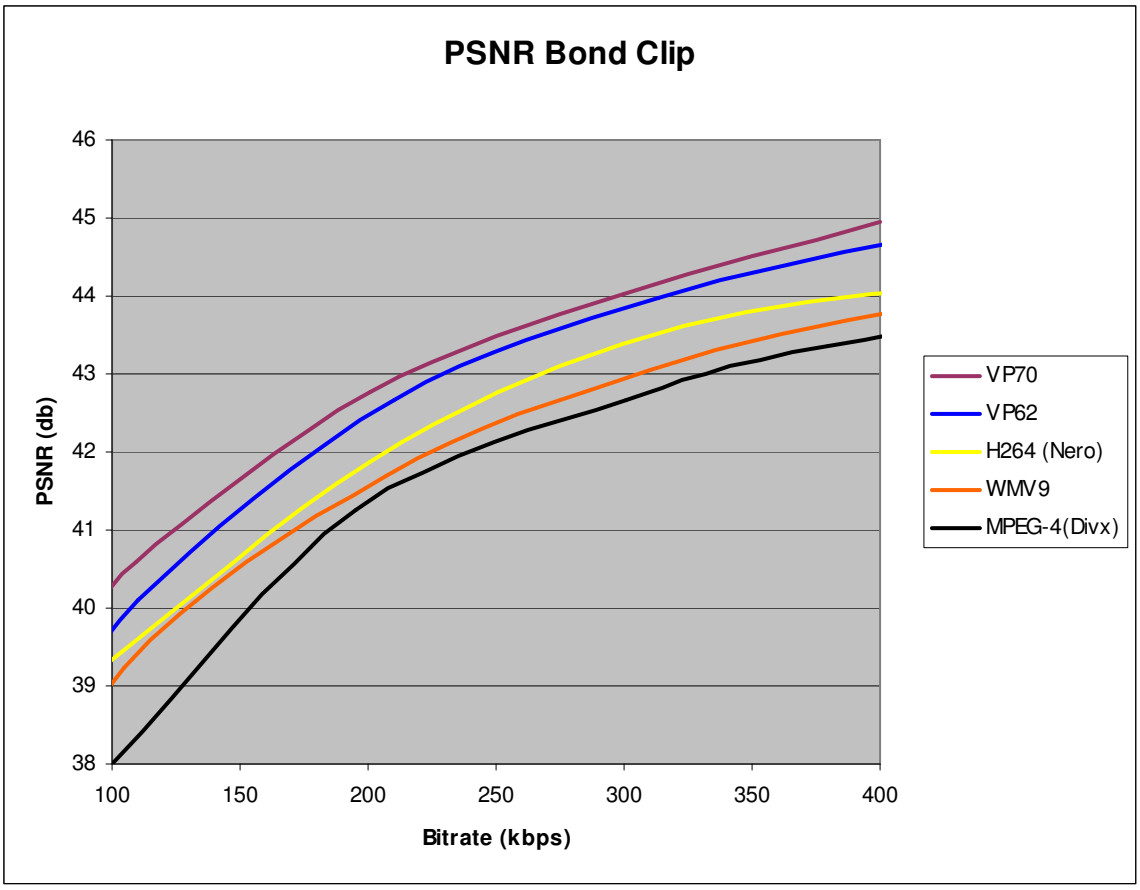
VP6.2

- Two-Pass Best Quality (VBR)
- Noise Reduction = 0
- Sharpness = 0
- Spatial Resampling = 0
- QuantizerMax = 45-50

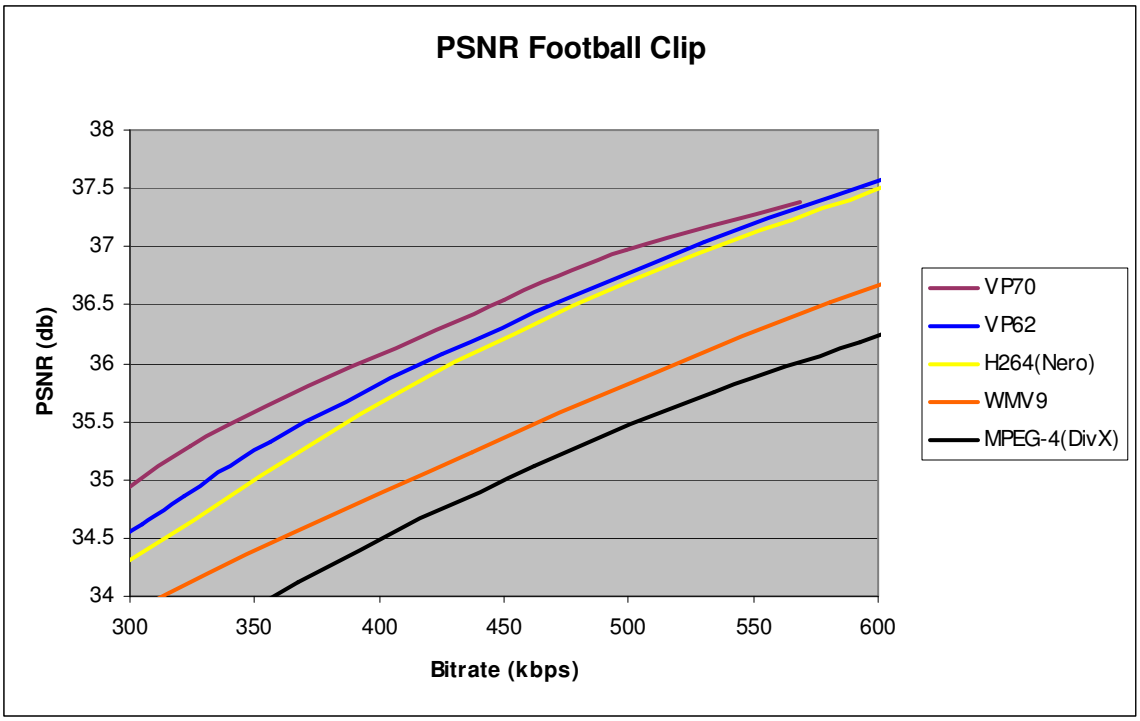
Nero Digital H.264

- Defaults for maximum definition AVC profile
- High-quality two-pass
- Cartoon source
- Psycho-visual enhancements set to high.

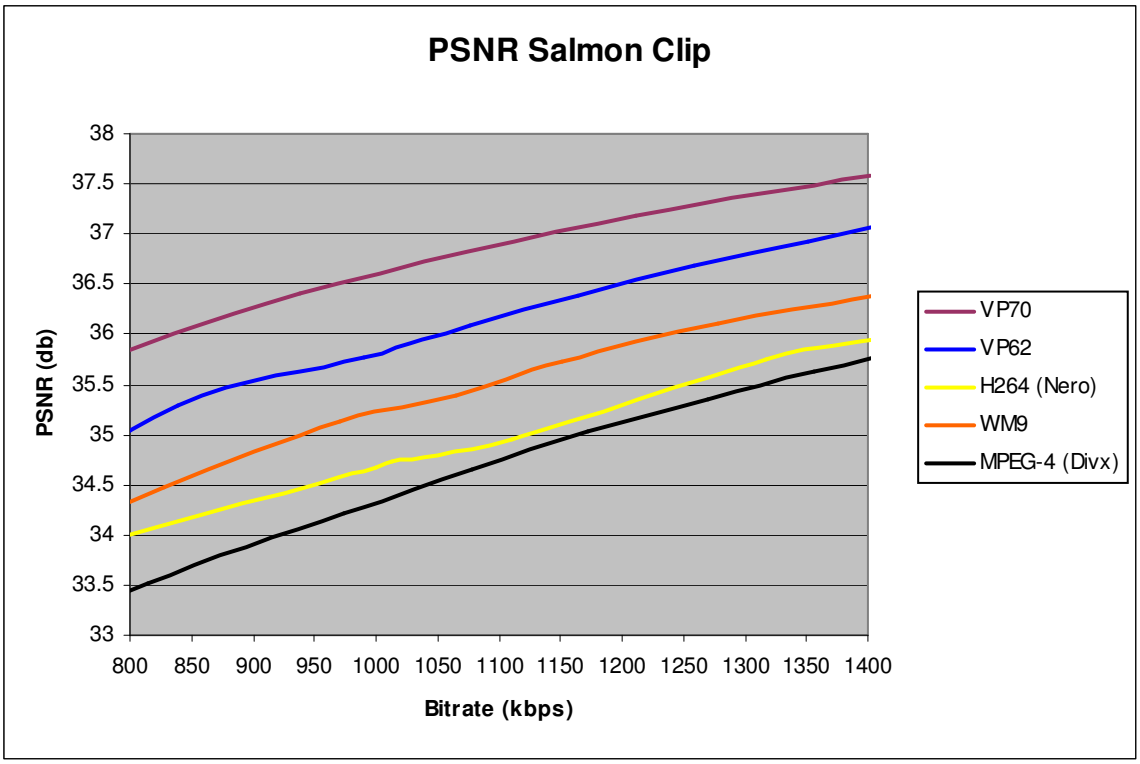
The PSNR of each clip was calculated using **avisynth** (www.avisynth.org). A script was written to read in the source clip and the clip produced from each codec above, convert them to the proper color space, and then compute the PSNR. The results are plotted in the graphs below.



This clip is a 320x240 movie trailer from a James Bond movie. There are many cut scenes and high motion scenes as well as a few sections of text overlaid graphics. There are also some cross fades and fades to black and up from white.



This clip is a short section of a football game, and is a famous clip from the Mpeg-2 test suite. Its extremely high motion and has some motion blur but there is also high contrast sections specifically around the numbers of the football uniforms..



This is a clip showing a salmon swimming underwater. The salmon's scales shimmer in the light under water. There is a slow pan of the background as the fish passes over the sea floor. This clip is especially difficult. It was given to us by a client for testing.