Overview of Fine Granularity Scalability in MPEG-4 Video Standard

Weiping Li
Fellow, IEEE

IEEE Transactions on Circuits and Systems for Video Technology,
Vol. 11, No. 3, March 2001

CPSC 538A - Paper Presentation

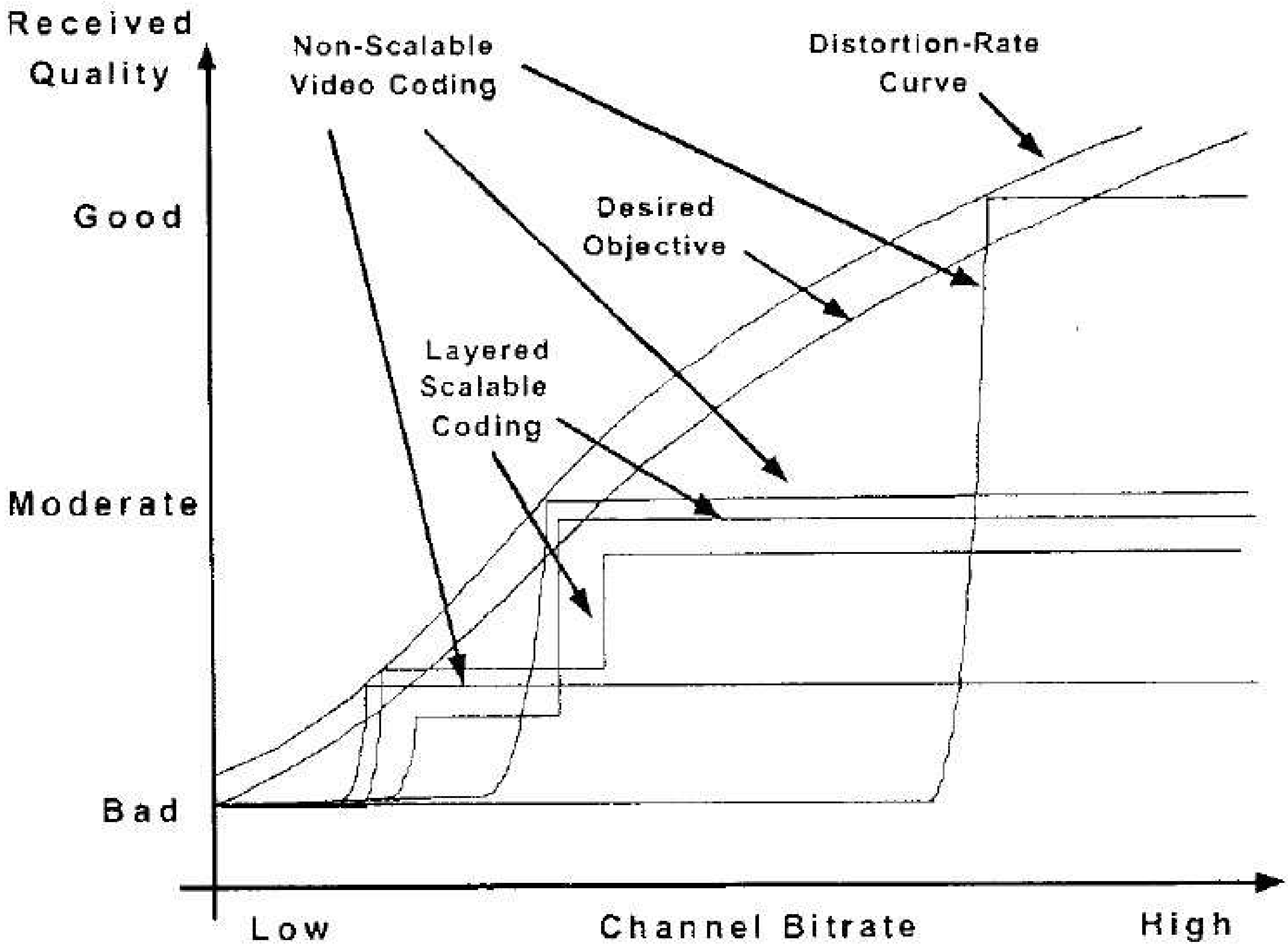
Roman Holenstein

January 19, 2005

Introduction

- Amendment on MPEG-4: Streaming video
- optimize video quality at a given bitrate
- New assumptions:
  - **encoder does not know channel capacity**
  - **decoder may not be able to decode all bits received from channel**
- **Bitstream should be partially decodable**

Video Coding Performance



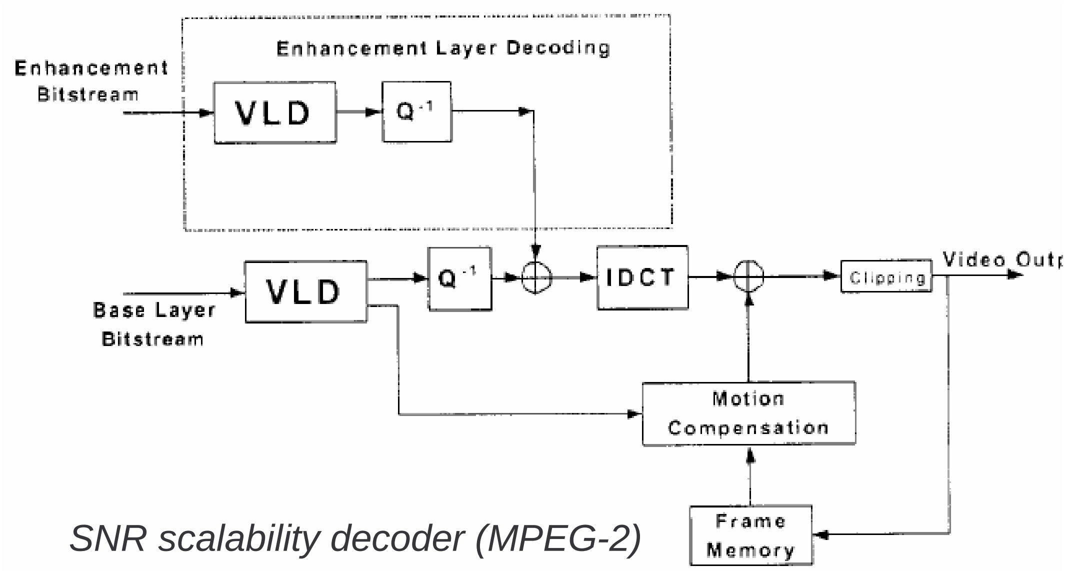
Layered Scalable Coding

- Signal-to-noise ratio (SNR) Scaling
- Temporal scaling
- Spatial scaling

Enhancement layer must be entirely transmitted, received, and decoded in order to provide any enhancement at all.

Layered Scalable Coding

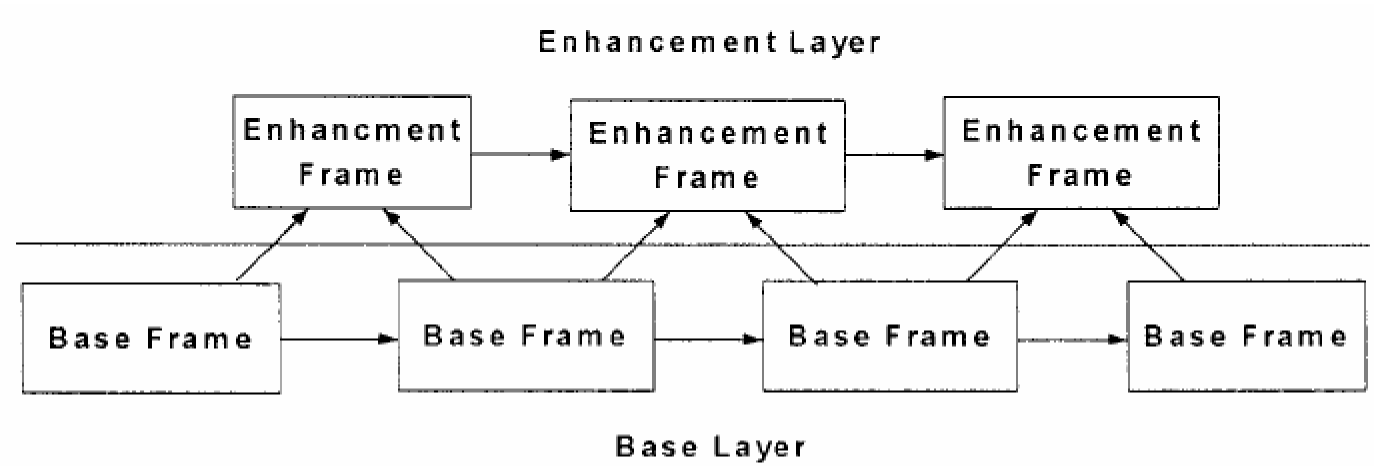
- Signal-to-noise ratio (SNR) Scaling
  - base layer is regularly DCT encoded, data removed using quantization
  - enhancement layer: DCT encoding of (original-inverse DCT of quantized base layer)
  - result depends on whether enhancement layer is received and used



SNR scalability decoder (MPEG-2)

Layered Scalable Coding

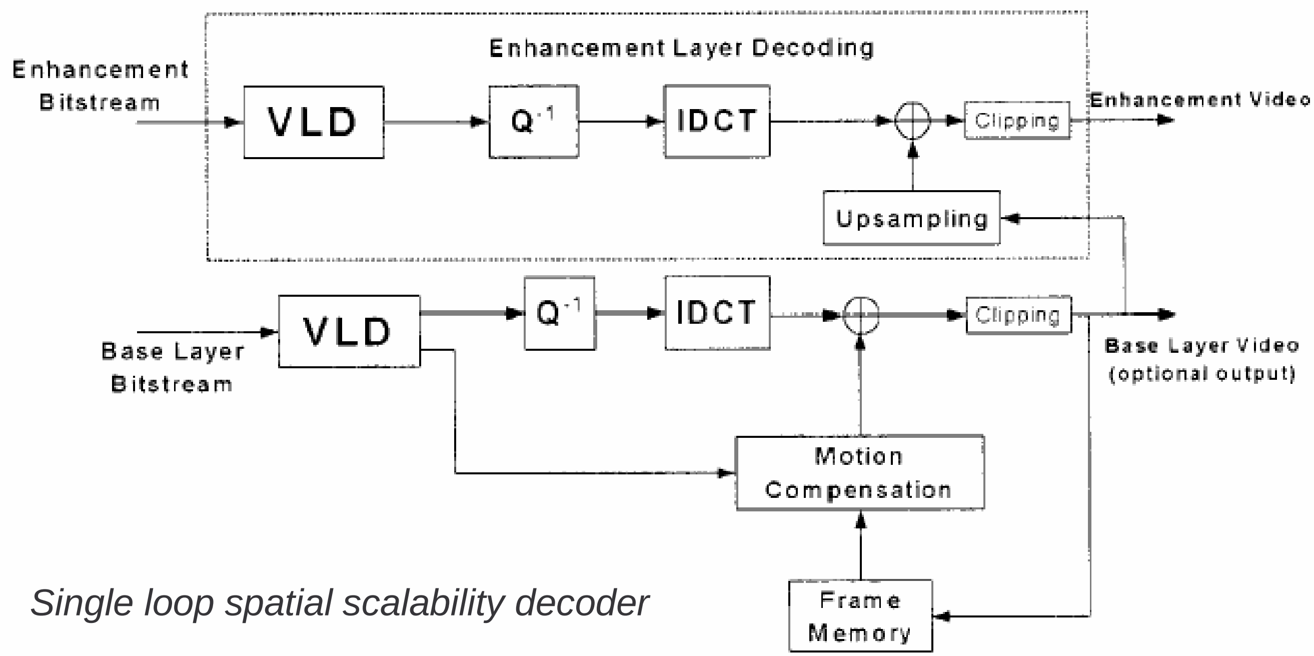
- Temporal scaling
  - base layer coded at lower frame rate (only using P-type prediction)
  - enhancement layer provides in-between frames at higher frame rate



Temporal scalability structure

Layered Scalable Coding

- Spatial scaling
  - layers at same frame rate, but different spatial resolution
  - image from base layer is upsampled and supplemented by enhancement layer



Single loop spatial scalability decoder

Bit-plane coding

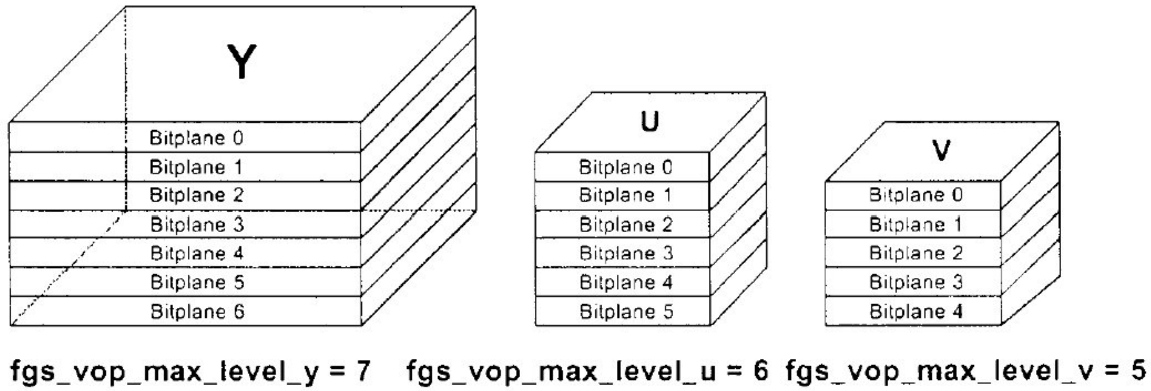
- Example

after zigzag ordering	10,0,6,0,0,3,0,2,2,0,0,2,0,0,1,0, … ,0,0 (absolute) 0,x,1,x,x,1,x,0,0,x,x,1,x,x,0,x, … ,x,x (sign bits)
Max value=10 4 bit planes	1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0, … ,0,0 (MSB) 0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0, … ,0,0 (MSB-1) 1,0,1,0,0,1,0,1,1,0,0,1,0,0,0,0, … ,0,0 (MSB-2) 0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0, … ,0,0 (MSB-3)
(RUN,EOP) symbols	(0,1) (2,1) (0,0)(1,0)(2,0)(1,0)(0,0)(2,1) (5,0)(8,1)

Can get up to 20% bit savings over run-length coding

FGS Coding

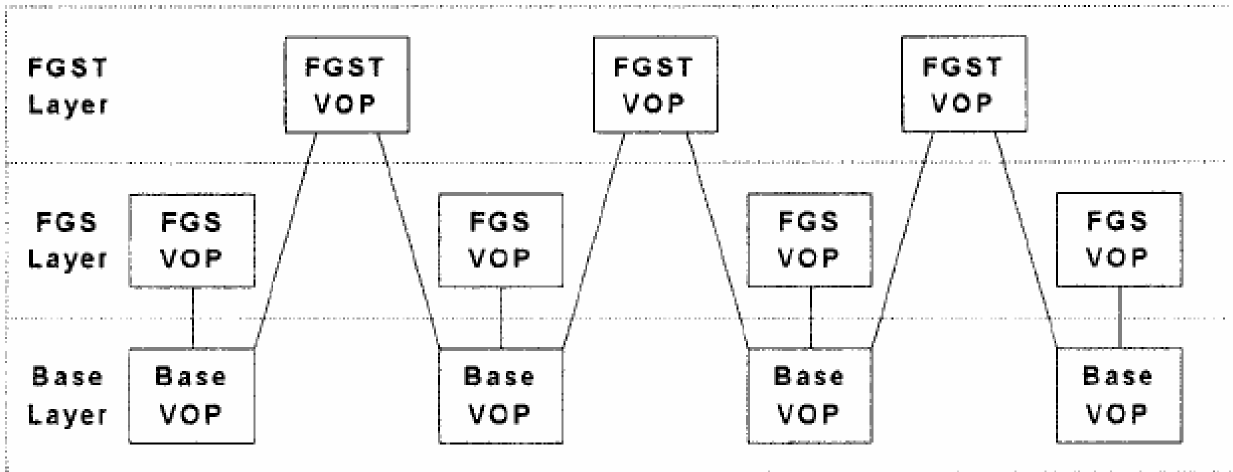
- Different number of bit-planes for each color component
- Variable-length codes
  - introduce ESCAPE symbol for coding large runs (6 bits)
  - create macroblock syntax to group ALL-ZERO cases for more efficient encoding
- Decoding truncated bitstreams
  - look ahead for special symbol (fgs_vop_start_code) and start decoding from there



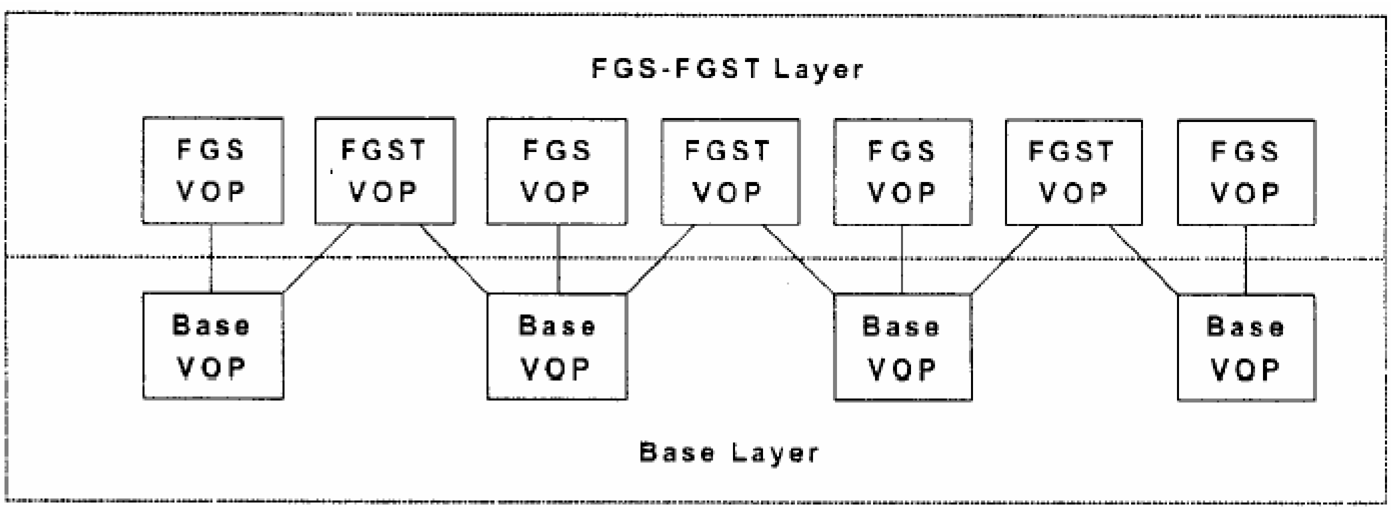
Advanced Features in FGS

- Frequency weighting
  - generally low-frequency DCT coefficients more important than high-frequency coeffs.
  - bits of visually more important frequency components are placed first in the bitstream
- Selective enhancement
  - more bit-planes of selected spatial locations of a frame are placed ahead of others in the bitstream
- Error resilience
  - resynchronization markers used in enhancement layer to deal with for random burst errors (once per bit-plane)
- FGS temporal scalability
  - combines FGS with temporal scalability (FGST)
  - FGST as separate layer or included in FGS enhancement layer

FGS Temporal Scalability



FGST organized into separate layer from FGS



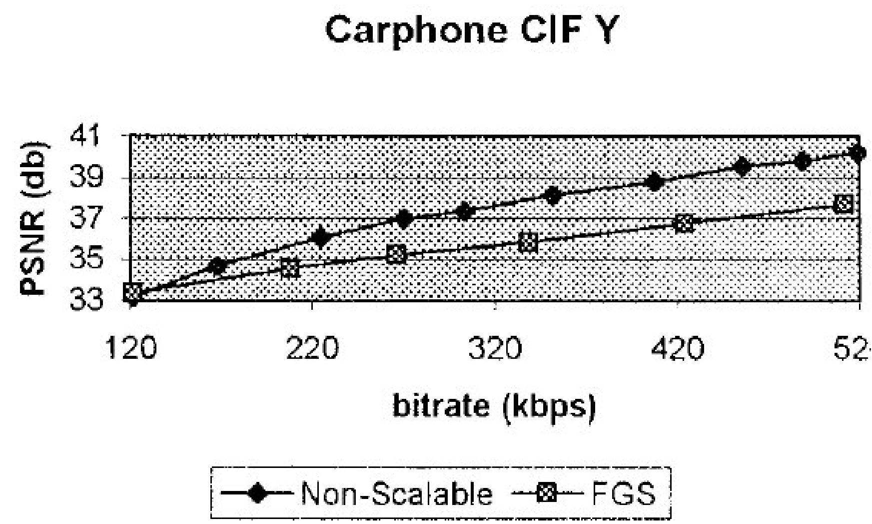
FGST and FGS organized into single enhancement layer

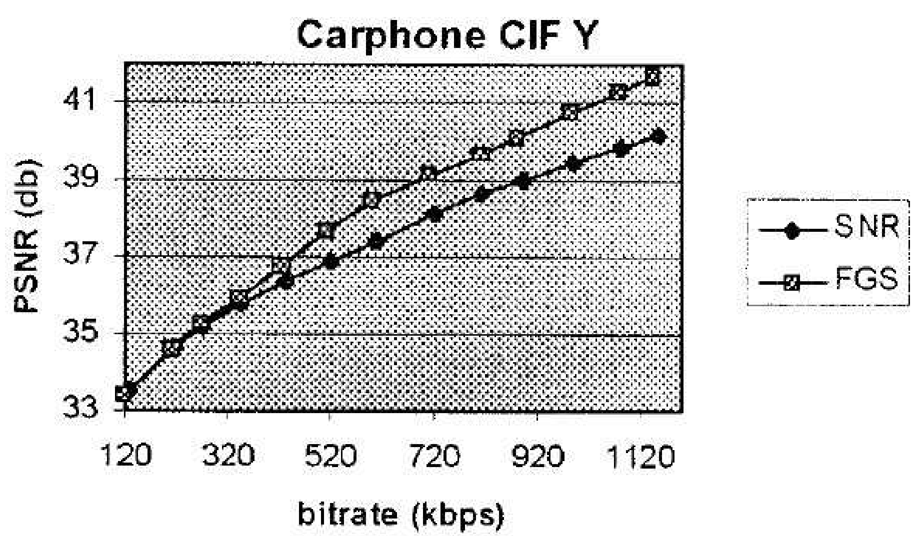
Profiles

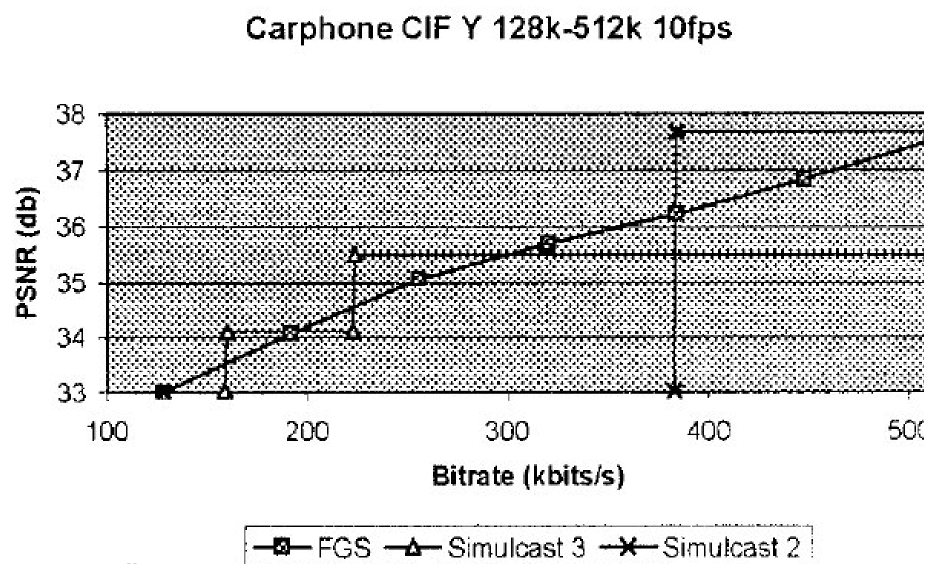
- Two profiles defined:
  - Advanced Simple Profile *(base layer)*
    - contains subset of nonscalable video tools
    - P-VOP (forward prediction only)
    - B-VOP (bi-directional prediction)
    - option for using error resilience tools
    - backwards compatible with baseline H.263
  - FGS Profile *(enhancement layer)*
    - bit-plane coding
    - frequency weighting
    - selective enhancement
    - error resilience (resync. markers)
    - FGS temporal scalability

Comparison: coding efficiency

- Multi-layer SNR Scalability
- Non-scalable coding (at upper bound)
- Simulcast







(PSNR=peak signal-to-reconstructed)

Summary

- FGS features
  - Bit-plane coding
    - better compression
    - allow for trucated bitstream
  - Frequency weighting
  - Selective enhancement
- Better coding efficiency than simulcast (at high and low end) and SNR.
- Worse than nonscaleable coding by ~2dB at high end of bit-rate