Cross Segment Decoding of HEVC for Network Video Applications

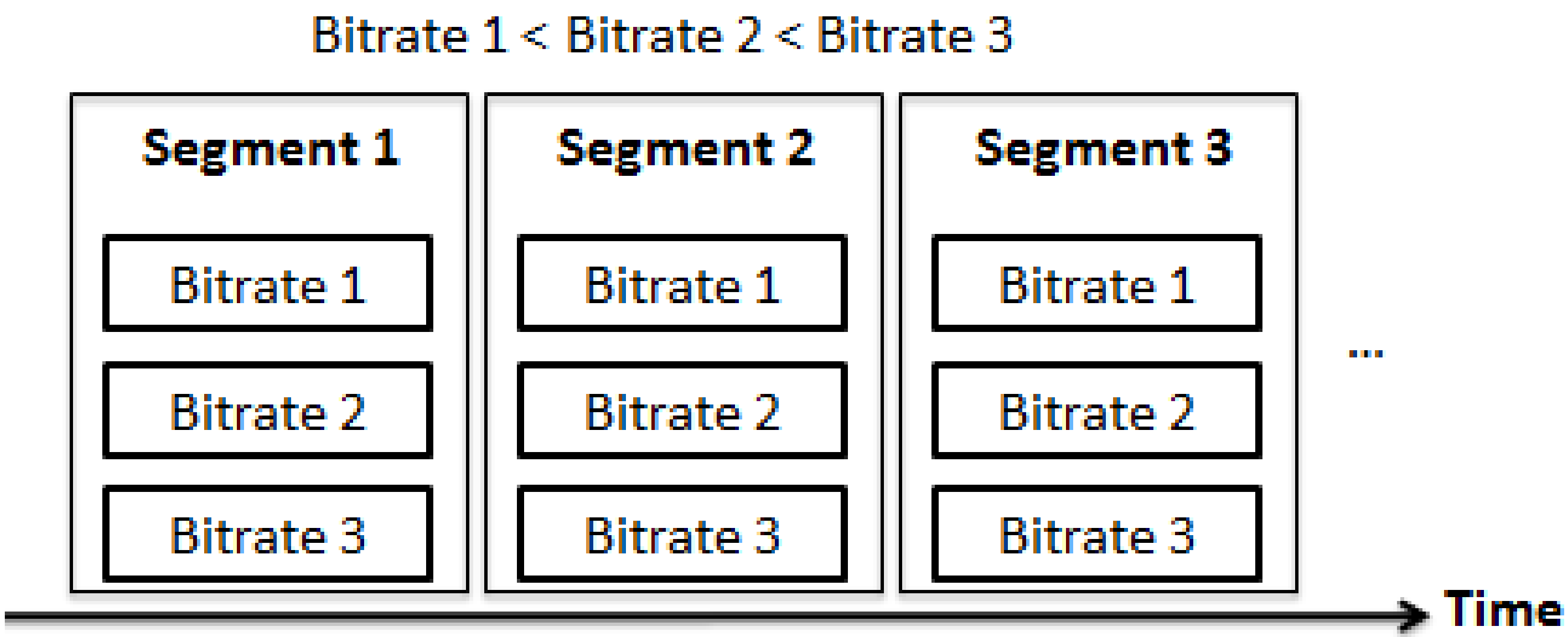
Jiangtao Wen, Bohan Li, Shunyao Li, Yao Lu, Pin Tao

Media Lab, Computer Science Department, Tsinghua University

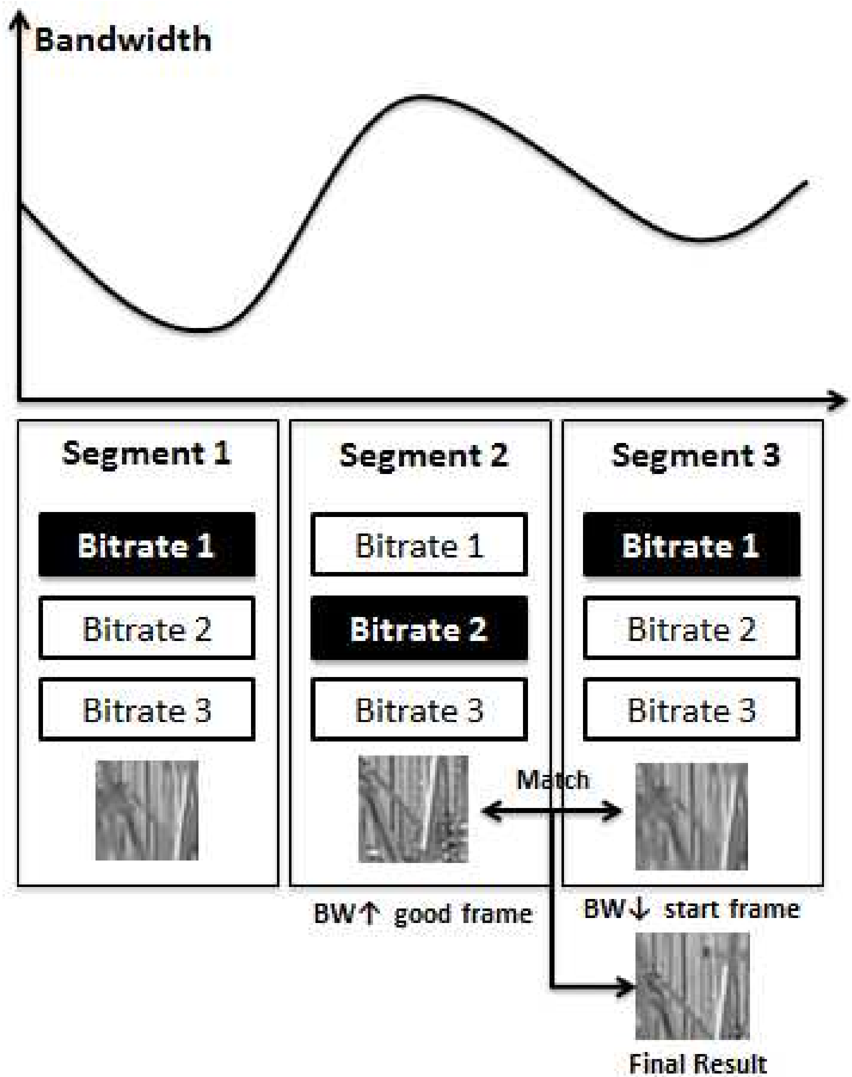
Motivations

- Decoders - more computational and storage capabilities
- Encoders – low complexity, low power
  - Not all tools are supported or properly
  - Rate allocation issues lead to sub-optimal bitrate allocations and quality variations
- Adaptive streaming
  - Rate variations caused quality variations
- Work in progress

Varying Bitrate Adaptive Streaming



Varying Bitrate Adaptive Streaming



Algorithm Overview

- Key idea: use information that the decoder has but was not (fully) used by the encoder for enhancement
- Terminology:
  - Good Frame: last frame in good quality segment
  - Start Frame: first frame in poor quality segment
  - Fresh Start: Start Frame after enhancement
- Three steps:
  - Segmentation of the SF into high and low motion areas
  - Enhance high and low motion areas with different algorithms
  - Replace SF with FS, continue decoding

Motion Segmentation

- Motion estimation between the GF and SF at the decoder using 4x4 patches
- A patch (patch size varies from 4x4 to 32x32) is designated as high motion if the average MV length is higher than a threshold of *width*QP/30000.* Low motion otherwise

Low Motion Enhancement

- For each Patch P designated as low motion, find the co-located Patch P' in the GF
- Calculate the MSD between P and P'
  - Replace P with P' if the MSD if below a threshold
  - Threshold was set to the larger of the following two values

$$Th_1 = 1.112 \times e^{(-0.2963 \times PSNR + 15.14)} - 10.21$$

and

$$Th_2 = 6.213 \times MECost^{1.348}$$

where

$$MECost = \frac{\sum_{\forall mv}\{SAD(mv) + \lambda_{ME}\, Bits(mv)\}}{\sum_{\forall mv} 1}$$

High Motion Enhancement (1)

- Find the MV for each 4x4 block in each high motion patch
- Enhance a 4x4 block only if its MV matches the MVs from no fewer than Th_{mv} (set to 6) of the 8 neighboring blocks – i.e. if the block is inside of a uniform motion area

1	2	3
4	X	5
6	7	8

High Motion Enhancement (2)

- For each high motion 4x4 block P to be enhanced
  - Find the corresponding block P’ in the GF using MV (P)
  - Calculate the MSD between P and P’
  - Copy P’ to P if MSD is below a threshold, calculated as in the case for low motion areas

Patch Size

- Compare the threshold T_{MSD} to

  $$T_{MSD0} = 0.0377e^{0.2272QP}$$

  - If $T_{MSD} > T_{MSD0}$, patch size = 32 x 32;
  - Else, we determine patch size according to P_T, the percentage of the MVs leading to a higher MSE calculated between GF & SF than the MVs calculated between GF & GF+1.

Test Conditions

- HM 8.2 with the low delay configuration
- Divided the test clips into two segments – the first 32 frames and everything else
- The QP difference was 5
- Clips of different motion levels and resolutions
  - Vidyo1, Vidyo3, KristenAndSara, FlowerVase, ChinaSpeed, BaseketballPass, ChromaKey, FourPeople, Johnny, SlideEditing, BQSquare, Traffic, PartyScene, ParkScene, Kimono, …
  - Testing range extension clips

Results

- Average PSNR gain for SF, 30 and 60 frames after the SF: 0.92dB, 0.61dB and 0.49dB
- Side information needed: 16 bits for PSNR and P_T of the SF after Intra encoding

Standard Decoder

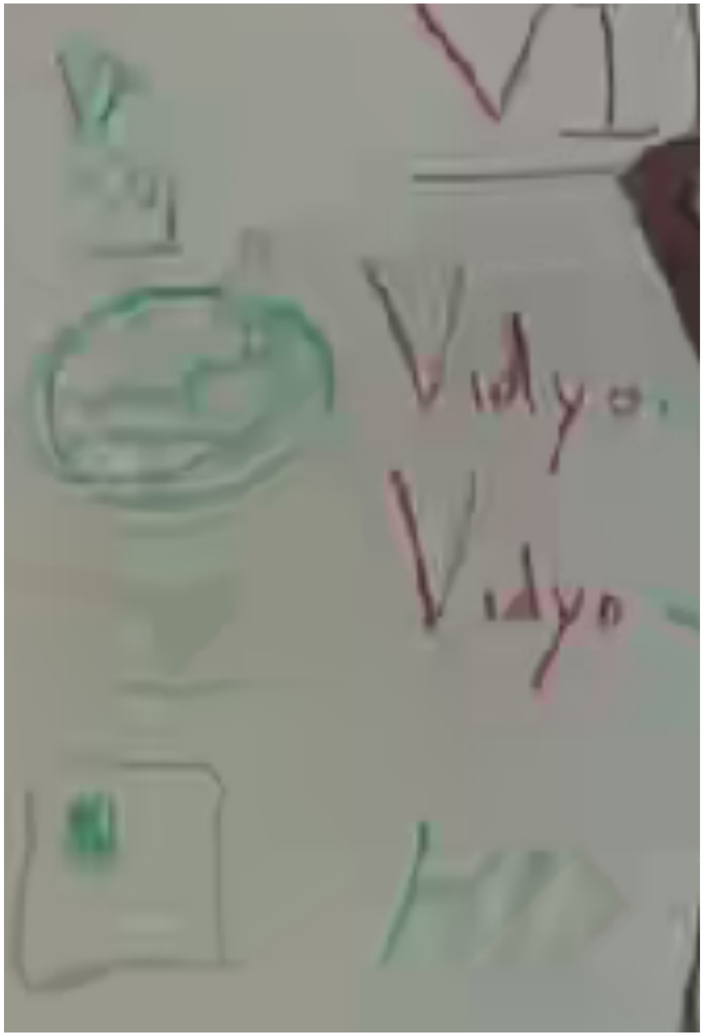
After Enhancement

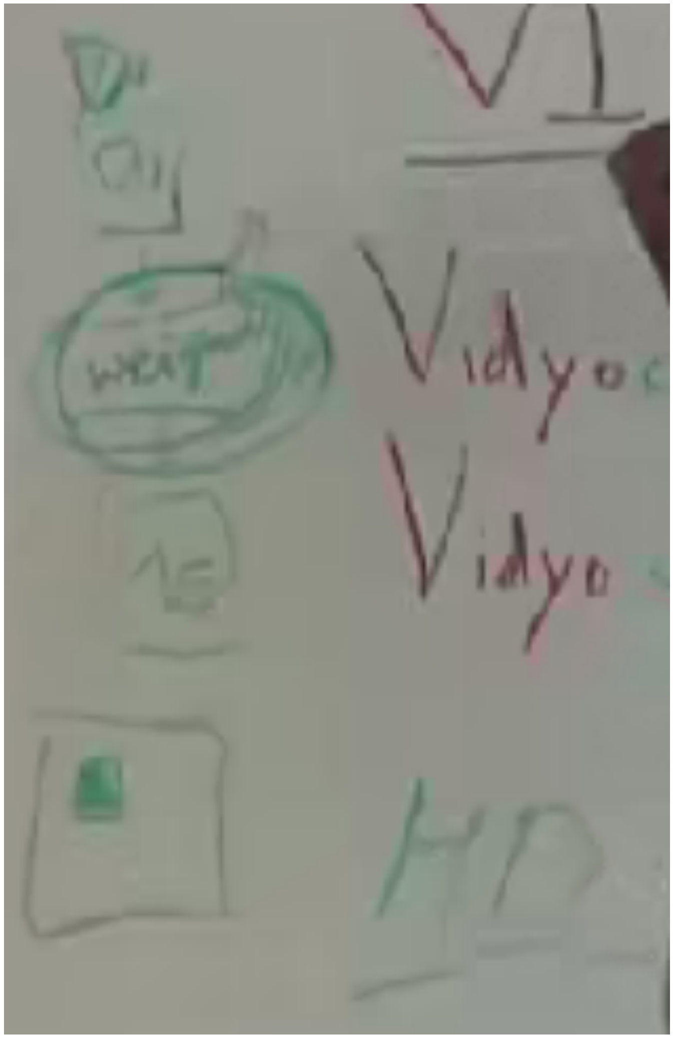
Standard Decoder

Enhanced Decoder

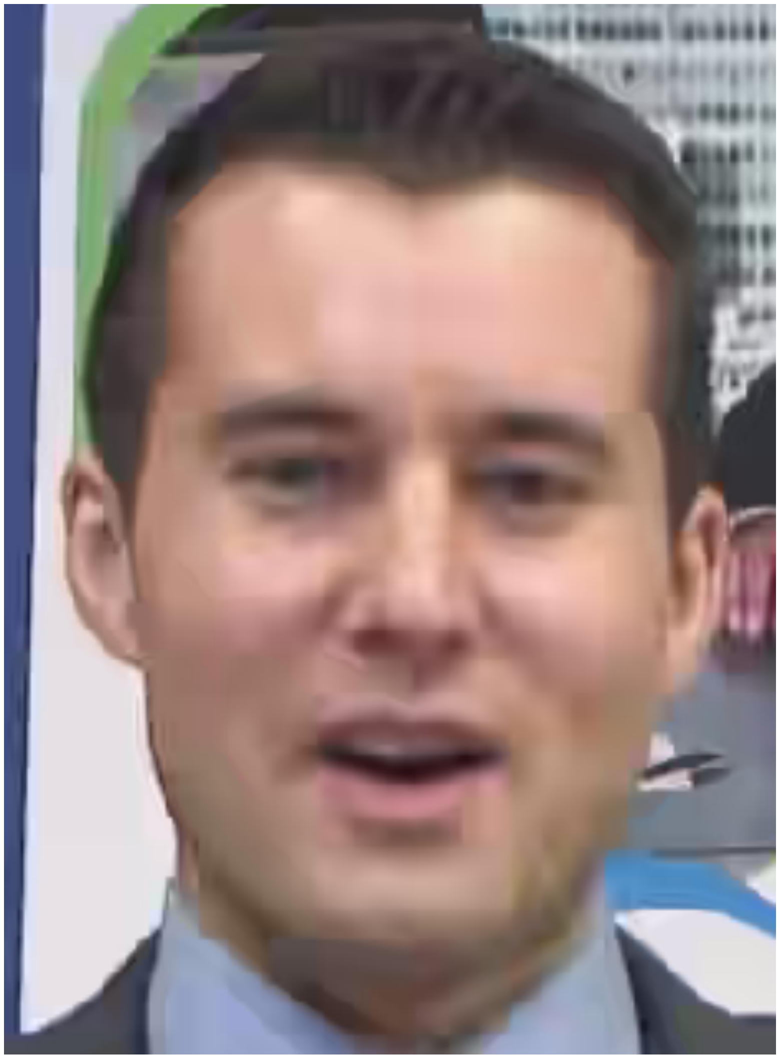
Visual Comparison

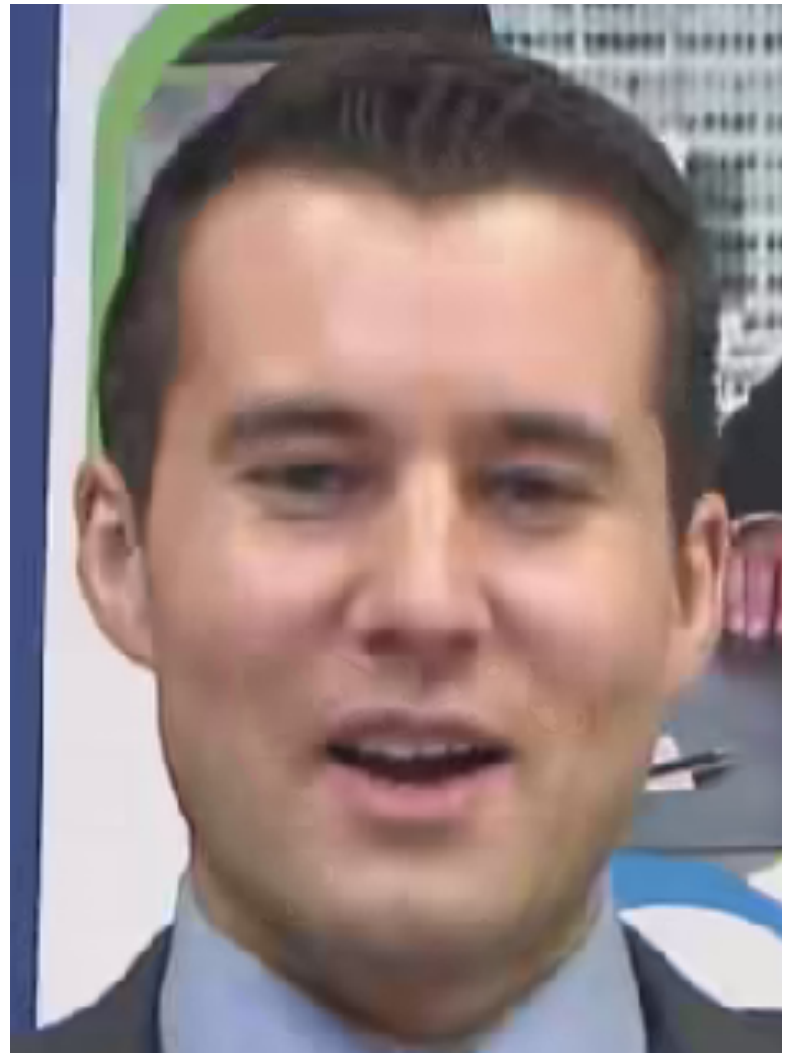
Visual Comparison



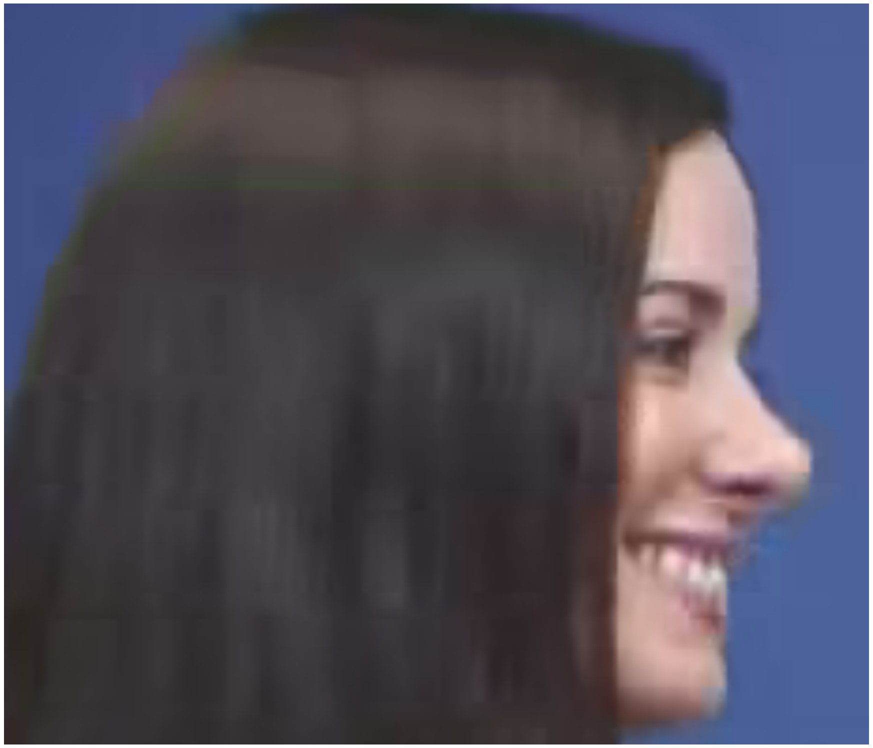


Visual Comparison

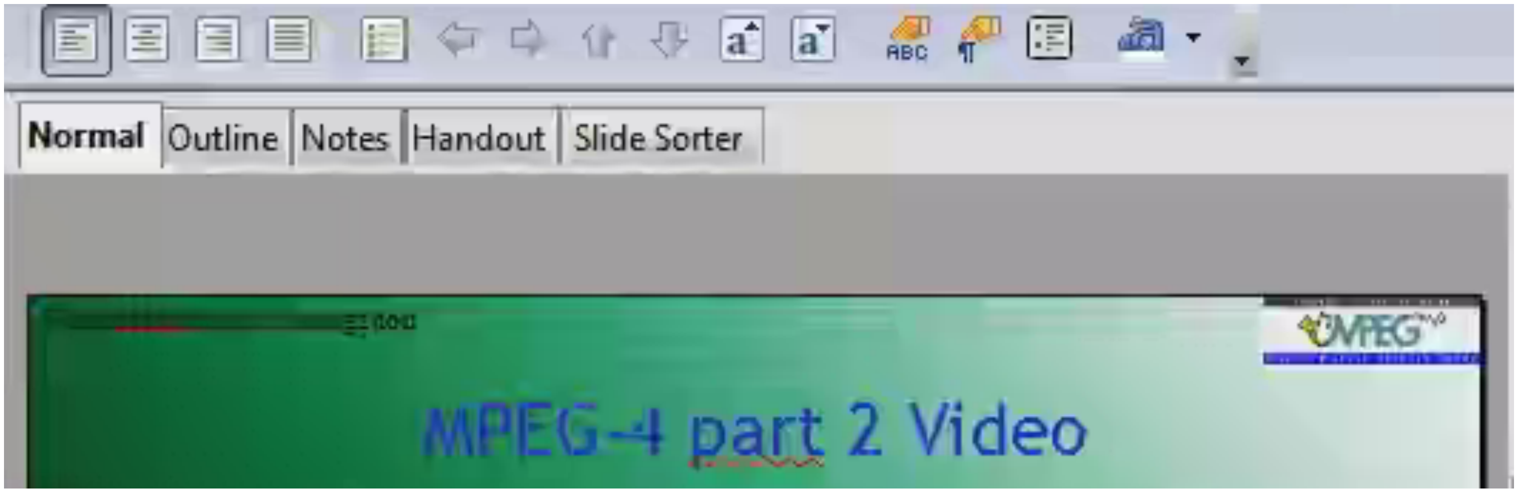




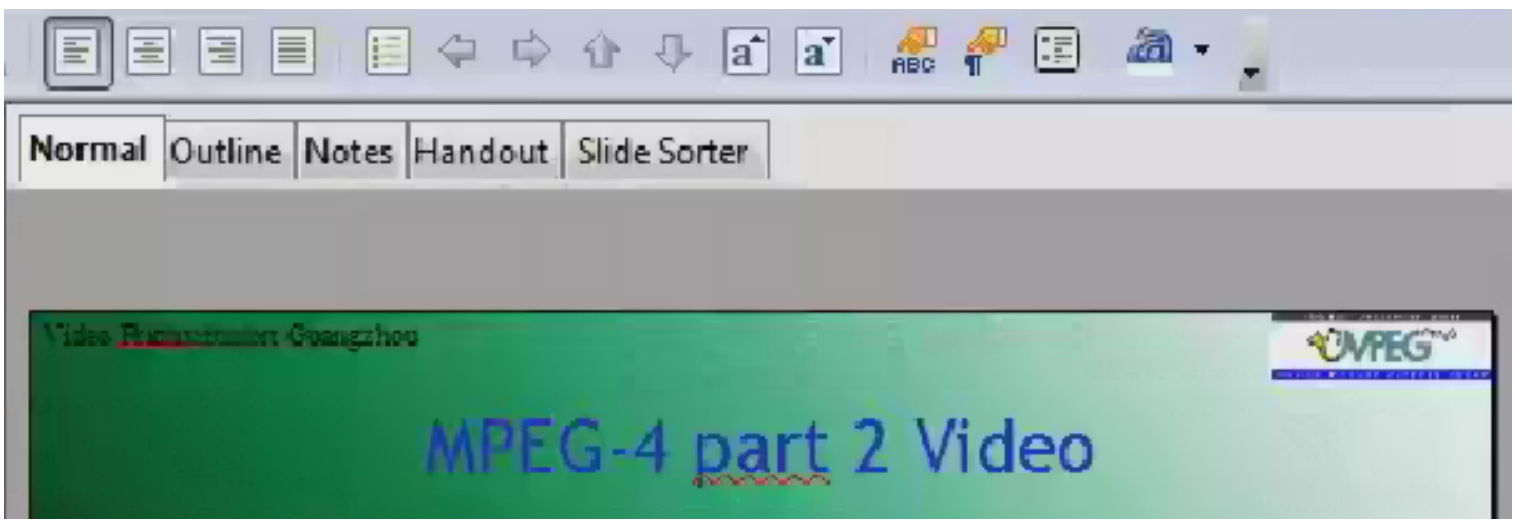
Results



Results



Normal
Outline
Notes
Handout
Slide Sorter
MPEG-4 part 2 Video



Normal
Outline
Notes
Handout
Slide Sorter
MPEG-4 part 2 Video

Observations

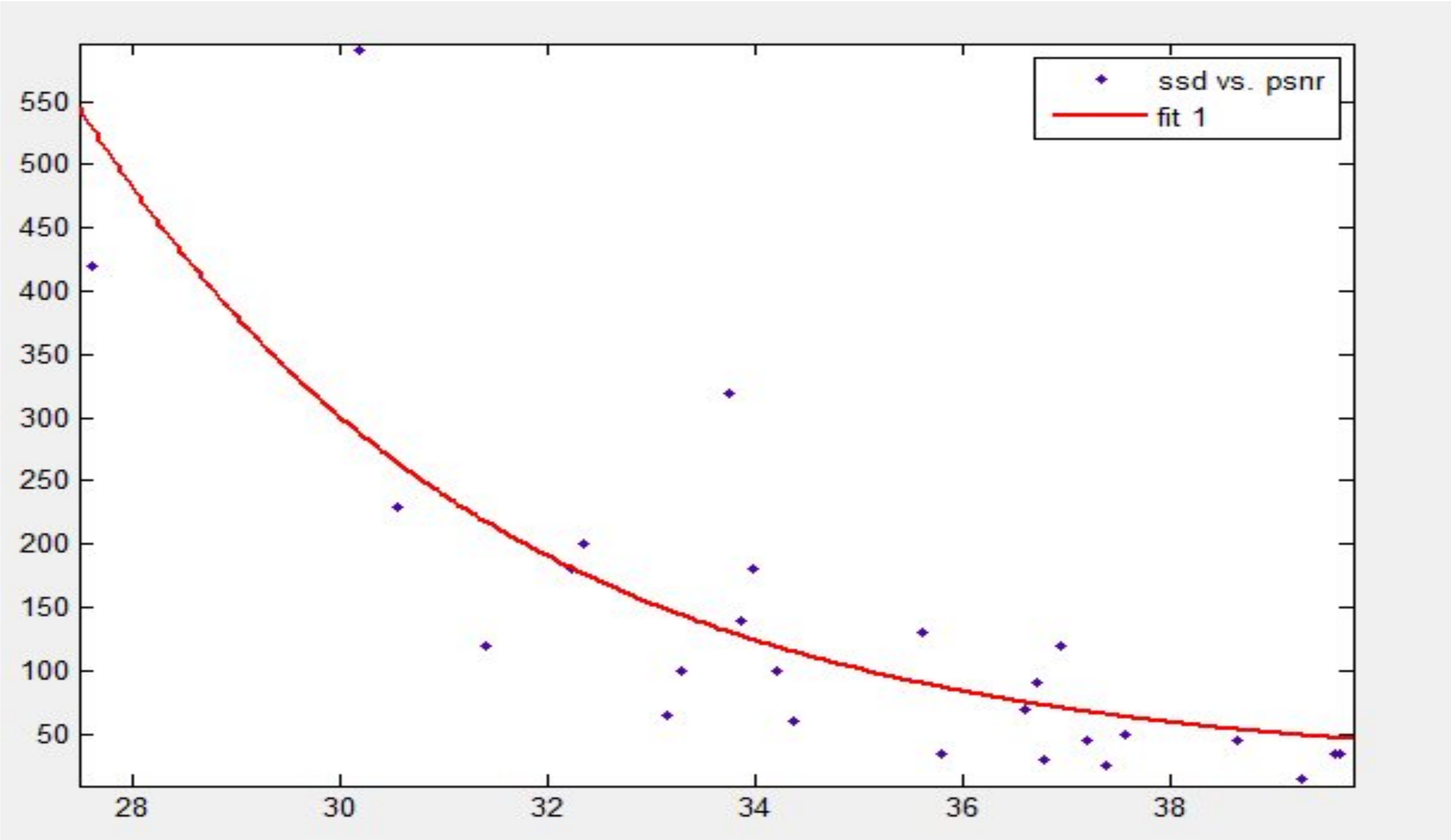
- Visible improvement in static areas and some motion areas
  - Sometimes even when there was a slight PSNR loss
- Artifacts created by mis-alignment of patches, visually similar to artifacts created by MV losses
  - Error concealment techniques apply?


P
S

Summary

- The decoder may have more information and computational power than what was utilized by the encoder
- A preliminary algorithm for utilizing the computational power and information at the decoder was proposed
- Side information needed: PSNR for SF after encoding – could be estimated by the decoder?
- Encouraging visual and RD improvements with areas of improvements especially for moving areas
- Application to scalability and range extensions?

Parameters Fitting (Thres-PSNR)



Parameters Fitting (Thres-cost)