INTERDIGITAL

ADAPTIVE DOWNSAMPLING FOR HIGH-DEFINITION VIDEO CODING

Jie Dong and Yan Ye

InterDigital Communications, San Diego, USA

Abstract *— Previous research has shown that downsampling prior to encoding and upsampling after decoding can improve the rate-distortion (R-D) performance compared with directly coding the original video using standard coding technologies, e.g., JPEG and H.264/AVC, especially at low bit rates. This paper proposes a practical algorithm to find the optimal downsampling ratio that balances the distortions caused by downsampling and coding, thus achieving the overall optimal R-D performance. Given the optimal sampling ratio, dedicated filters for down- and up-sampling are also designed. Simulations show this algorithm improves the R-D performance over a wide range of bit rates, e.g., from 1.0 dB at high bit rates to 2.5 dB at low bit rates.*

Problem Statement

σ_A^2 **: Overall MSE between the input and output of the whole system**

σ_C^2 **: MSE caused by coding**

σ_D^2 **: MSE caused by downsampling**

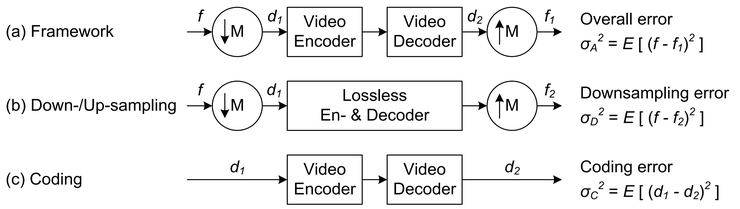
Find the optimal downsampling ratio *M*, such that the overall error σ_A^2 reaches the minimum.

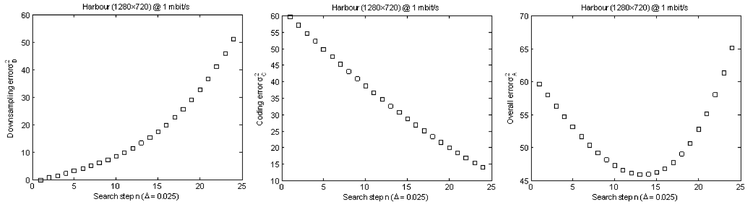
$$M = \arg\min_M \sigma_A^2 \quad \textbf{sub. to video input and bit-rate}$$

$$\sigma_A^2 = E\left[(f - f_2)^2\right] + E\left[(d_1 - d_2)^2\right] = \sigma_D^2 + \sigma_C^2$$

$$\Downarrow$$

$$M = \arg\min_M (\sigma_D^2 + \sigma_C^2)$$





Example of searching the optimal ratio for *Harbour* at 1 mbit/s. (a) Downsampling error; (b) coding error; (c) overall error.

Downsampling Error Estimation

Estimated in the frequency domain: energy of the high frequency components that exist in *f* but are lost in f_2.

$$\sigma_D^2 = \frac{1}{WH}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left[\hat{S}_{xx}(\omega_1,\omega_2) - \hat{S}_{yy}(\omega_1,\omega_2)\right]d\omega_1 d\omega_2$$

Energy exists in the original video *f* , measured by Power Spectral Density (PSD)

$$\hat{S}_{xx}(\omega_1,\omega_2) = \frac{1}{WH}\left|\sum_{w=0}^{W-1}\sum_{h=0}^{H-1} x[w,h]e^{-j\omega_1 w - j\omega_2 h}\right|$$

Energy exists in the downsampled and upsampled video f_2. (horizontal ratio $M_h = \frac{A_h}{B_h}$, vertical ratio $M_h = \frac{A_h}{B_h}$)

$$\hat{S}_{yy}(\omega_1,\omega_2) = \begin{cases} \hat{S}_{xx}(\omega_1,\omega_2), & \text{if } \omega_1 \in \left[-\frac{B_h}{A_h}\pi, \frac{B_h}{A_h}\pi\right] \,\&\, \omega_2 \in \left[-\frac{B_h}{A_h}\pi, \frac{B_h}{A_h}\pi\right] \\ 0, & \text{otherwise} \end{cases}$$

Coding Error Estimation

Estimated by an empirical R-D model: $\sigma_C^2 = \frac{\beta}{r^\alpha}$ (*r* is bits per pixel (bpp) given downsampling ratio)

Given a target bit-rate *R*, the larger the downsampling ratio ***M***, the larger the *r* is. $r = \frac{R \times M_h \times M_v}{fps \times W \times H}$

Training *α* and *β*

encode a given sequence at a range of bit rates $\{R_0, R_1, \ldots, R_{N-1}\}$ $\xrightarrow{\text{Normalization}}$ $\{r_0, r_1, \ldots, r_{N-1}\}$

$\Downarrow$

$[\alpha_{opt}, \beta_{opt}] = \arg\min_{\alpha,\beta} \sum_{i=0}^{N-1}\left(d_i - \frac{\beta}{r_i^\alpha}\right)^2$ $\Leftarrow$ *N* pairs of $[r_i, d_i]$ $\Leftarrow$ distortions $\{d_0, d_1, \ldots, d_{N-1}\}$ measured by MSE

Optimal Ratio Determination

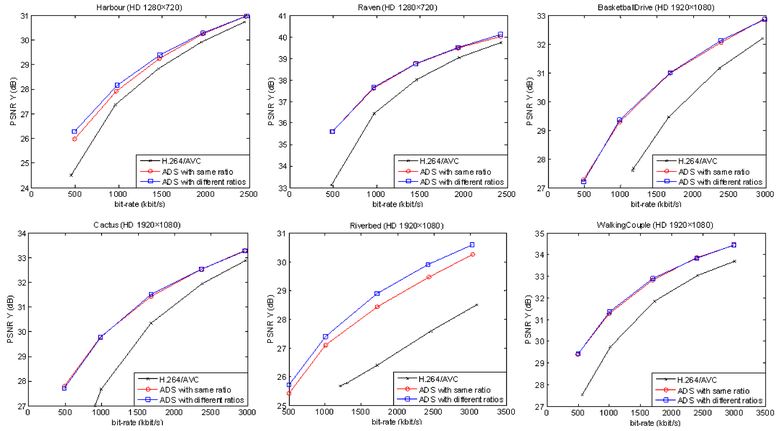
σ_D^2 : monotonically increase with M
σ_C^2 : monotonically decrease with M

Find the balance between σ_D^2 and σ_C^2 !

Searching for optimal *M*

1. Calculate σ_A^2 using a set of N downsampling ratios: $M_i = 1.0 + (i - 1)\Delta$, $(1 \le i \le N)$.
2. Find M_i providing the minimum overall error σ_A^2

Experimental Results

Optimal ratios obtained by the proposed algorithm

HD Sequences	Bit Rate (mbit/s)	Same Ratio for Hor. & Ver.	Different Ratio for Hor.	Different Ratio for Ver.
Harbour (1280×720)	0.5	40/20	40/22	40/13
	1.0	40/21	40/23	40/15
	1.5	40/22	40/24	40/17
	2.0	40/23	40/24	40/18
	2.5	40/23	40/24	40/18
Raven (1280×720)	0.5	40/18	40/18	40/18
	1.0	40/21	40/20	40/22
	1.5	40/22	40/21	40/24
	2.0	40/23	40/21	40/26
	2.5	40/24	40/22	40/28
Cactus (1920×1080)	0.5	60/24	60/29	60/21
	1.0	60/30	60/33	60/26
	1.7	60/34	60/36	60/31
	2.4	60/37	60/38	60/34
	3.0	60/39	60/39	60/38
BasketballDrive (1920×1080)	0.5	60/23	60/22	60/26
	1.0	60/28	60/26	60/31
	1.7	60/32	60/30	60/35
	2.4	60/36	60/32	60/41
	3.0	60/40	60/36	60/42
Riverbed (1920×1080)	0.5	60/16	60/8	60/24
	1.0	60/17	60/9	60/26
	1.7	60/22	60/11	60/28
	2.4	60/24	60/13	60/30
	3.0	60/24	60/13	60/39
WalkingCouple (1920×1080)	0.5	60/22	60/18	60/25
	1.0	60/25	60/21	60/29
	1.7	60/28	60/23	60/33
	2.4	60/30	60/24	60/36
	3.0	60/31	60/25	60/38