

H.264/AVC 하이 프로파일의 고속 화면 내 예측을 위한 효율적인 블록 모드 결정과 예측 모드 선택

*김태호 **정제창

한양대학교

*crewx@hanyang.ac.kr

Efficient Block Mode Decision and Prediction Mode Selection for Fast Intra Prediction in H.264/AVC High Profile

*Kim, Tae-Ho **Jeong, Je-Chang

Hanyang University

요약

H.264/AVC는 휘도 신호 4x4 블록을 위하여 9개의 화면 내 예측모드를 사용한다. 예측 모드는 8개의 방향성을 가진 모드와 하나의 비방향성 DC 모드가 있다. 휘도 신호 16x16 블록에서는 4가지의 예측 모드가 있으며 색차 신호 8x8 에서도 4개의 예측모드를 사용한다. 이러한 예측 모드들 중 최적의 예측 모드를 선택하기 위하여, 부호화기는 선택 가능한 모든 예측 모드의 윌-왜곡 비용을 계산한 후, 최적의 윌-왜곡 비용을 가진 예측 모드를 사용하여 부호화를 수행한다. 따라서 H.264/AVC의 화면 내 예측 과정은 많은 계산 복잡도를 가진다. 특히 하이 프로파일에서는 휘도 신호 8x8 블록이 화면 내 예측을 위해서 고려되므로 더욱 많은 계산 복잡도를 요구한다.

이에 본 논문은 H.264/AVC 하이 프로파일의 화면 내 예측의 부호화 계산 복잡도를 줄이는 방법을 제안한다. 현재 매크로 블록의 분산을 계산한 후, 이를 이용하여 윌-왜곡 최적화에 후보로 사용되어지는 블록 모드를 결정하고, 각 블록 모드에서 제공하는 예측 모드들을 효율적으로 선택하는 방법을 연구·개발하였다. 제안된 방법은 기존 H.264/AVC 참조 소프트웨어인 JM13.1 부호화 시간 대비 약 83%의 연산시간이 감소하는 결과를 보였다.

1. 서론

최신 비디오 부호화 표준인 H.264/AVC는 HDTV(High Definition TeleVision), DMB(Digital Media Broadcasting)와 같은 많은 디지털 방송 서비스를 위하여 채택 되었다.[1] H.264/AVC는 이전의 비디오 코딩 표준과 비교하여 높은 부호화 성능을 보여준다. 이는 최신 부호화 기술인 움직임 예측/보상, 멀티 참조 프레임, 1/4 픽셀 정밀도의 화면 간 예측, 화면 내 예측, 엔트로피 부호화, 윌-왜곡 최적화와 같은 코딩 기법을 통해서 얻어진다.[2]

인급된 최신 부호화 기법 중 윌-왜곡 최적화(RDO, Rate Distortion Optimization)는 최적의 RD 비용을 가진 모드로 부호화를 진행하여 부호화 효율을 극대화 시킨다. RDO에서는 선택 가능한 모든 예측 모드의 RD 비용을 계산한 뒤, 가장 낮은 비용을 요구하는 예측 모드를 선택하여 부호화의 효율을 극대화 한다.[3] 하지만 계산 복잡도 역시 극대화된다. 이를 극복하기 위하여 다수의 화면/화면 간 예측 모드 알고리즘들과 움직임 예측/보상 알고리즘들이 제안 되었다.[4][5]

Wang은 매크로블록을 4개의 서브블록으로 분할한 후, 5개의 2x2 필터를 거쳐 매크로블록의 방향성을 검출하여, 방향에 따라 예측모드를 선택하는 방법을 제안하였다.[6] Pan은 방향성을 검출하기 위하여 소벨 연산자와 에지 히스토그램을 사용하였다.[7] 또한 Tsai는 화소와 서브 블록들을 기반으로 방향성을 검출하는 방법을 제안하였다.[8] 제안된 알고리즘은 직접적으로 후보 예측 모드를 선택할 수 있으며 RD

와 PSNR(Peak Signal Noise Ratio) 성능의 손실이 Pan이 제안한 알고리즘보다 적다.

인급된 알고리즘들은 고속의 화면 내 예측을 위하여 RDO에 사용될 후보 예측 모드들을 줄이는 방법을 사용하고 있다. 계산 복잡도를 줄이기 위한 또 하나의 방법으로서, Bharanitharan은 4x4 블록이나 16x16중 하나만을 후보 블록 모드로 결정하였다.[9] Huang은 고속 화면 내 예측을 위하여 2 단계 접근법을 제안하였다. 첫 번째 단계로 후보 블록 모드를 결정하였고 다음으로 후보 예측 모드들을 선택 하였다.[10]

본 논문은 H.264/AVC 하이 프로파일의 고속 화면 내 예측을 위하여 효과적인 블록 결정과 예측 모드를 선택하는 방법을 제안한다. 제안된 알고리즘은 2개의 블록 모드만이 후보 블록 모드로 결정되며 예측 모드 역시 선택적으로 사용하여 계산 복잡도를 감소시킬 수 있다.

2. H.264/AVC의 화면 내 예측과 예측 모드 선택

2.1 화면 내 예측

H.264/AVC 하이 프로파일의 화면 내 예측은 이전에 부호화된 주변 블록들로부터 획득된 재구성된 화소들을 사용하여 수행된다. 휘도 성분의 화면 내 예측을 위해서 H.264/AVC는 세 가지의 블록 모드를 사용한다.: I4MB(Intra 4x4), I8MB(Intra 8x8), I16MB(Intra 16x16)

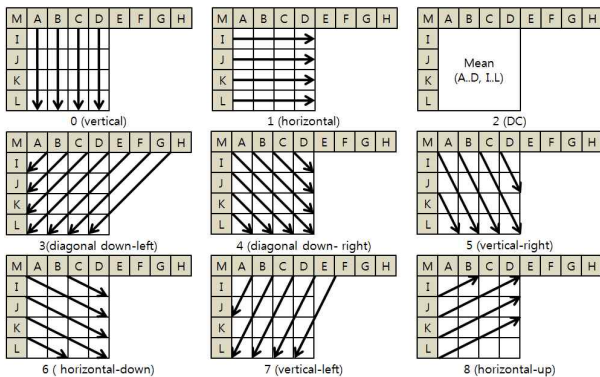


그림 1. 휘도 신호 4x4 블록을 위한 9가지의 예측 모드

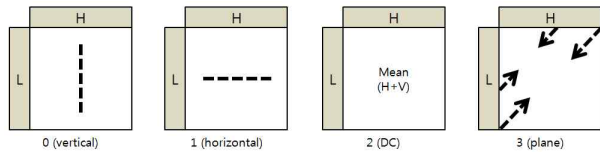


그림 2. 휘도 신호 16x16 블록을 위한 4가지의 예측 모드

I4MB의 블록 사이즈는 4x4이며, 8개의 방향성을 가진 예측 모드와 1개의 비방향성 모드(DC)를 가진다. I8MB의 블록 사이즈는 8x8이며 예측 모드는 I4MB와 동일하다. 하지만 16MB는 4가지의 예측 모드를 가진다. I4MB와 I16MB의 예측 모드는 그림 1과 2에서 나타난다. 일반적으로 QCIF나 CIF와 같이 상대적으로 해상도가 작은 영상은 텍스처 영역을 포함할 확률이 높기 때문에, I4MB가 I16MB보다 많이 사용된다. 색차신호를 위해서는 8x8 블록사이즈만 사용되고, 예측 모드는 I16MB와 유사하다.

2.2 울-왜곡 최적화 기법

부호화 성능을 향상시키기 위하여, H.264는 RDO 기법을 이용한다. RDO 기법은 Lagrangian 계수를 이용함으로써 최적의 예측 모드를 선택하여 극적으로 부호화 효율을 증가 시킨다. RD 비용은 다음과 같이 계산 할 수 있다.

$$RDcost = D + \lambda \times R \quad (1)$$

D는 현재 블록과 재구성된 블록의 SSD(Sum of Squared Difference)을 가르키며, R은 선택된 모드를 부호화 하기 위하여 필요한 총 비트, λ 는 Lagrangian 계수를 나타낸다. Lagrangian 계수는 QP(Quantization Parameter)은 다음과 같이 계산할 수 있다.

$$\lambda = 0.85 \times 2^{\frac{QP-12}{3}} \quad (2)$$

3. 제안된 알고리즘

3.1 매크로블록의 분산을 이용한 블록 모드 결정 방법

일반적으로 작은 블록 모드는 텍스처 영역을 위하여 사용되고 큰 블록은 균일하거나 플랫폼 영역을 위하여 사용된다. 본 논문에서는 매크로 블록의 분산은 텍스처 복잡도에 대응한다고 가정하여, 블록 모드를 결정하는 기준으로 사용하였다. 그림 3-5는 각 블록 모드가 베스트 모드로 선택되었을 때, 분산의 분포도를 확률로 나타내었다. 획득 된

데이터는 H.264 표준 참조 소프트웨어를 사용하여 얻어졌으며, Mobile(CIF)시퀀스를 100프레임 부호화 한 결과이다. QP는 16을 사용하였다. I8MB의 경우 모든 분산 범위에서 불규칙적으로 선택 되었다. I4MB의 경우 분산이 커질수록 베스트 모드로 선택될 확률이 높았고, I16MB의 경우 분산이 200 미만일 때만 베스트 모드로 선택 되었다. 다른 시퀀스도 유사한 패턴을 보였다. 이를 이용하여, 제안된 블록 모드 결정 방법에서는 식 (3)의 방법으로 임계치를 계산하여 RDO를 위한 후보 블록 모드를 결정 하였다. 임계치는 각 블록의 양자화 단계에 따라 정의 되므로 적응적으로 변화 할 수 있다.

$$Th = Q_{step}^2 + 32 \quad (3)$$

제안된 블록 모드 결정방법은 다음과 같다.

- 1) 코딩된 매크로블록의 분산을 계산한다.
- 2) 매크로블록의 분산 값이 정의된 임계치보다 크면 I4MB와 I8MB가 후보 블록 모드로 결정 된다.
- 3) 매크로블록의 분산 값이 정의된 임계치보다 작거나 같으면 I8MB와 I16MB가 후보 블록 모드로 결정 된다.

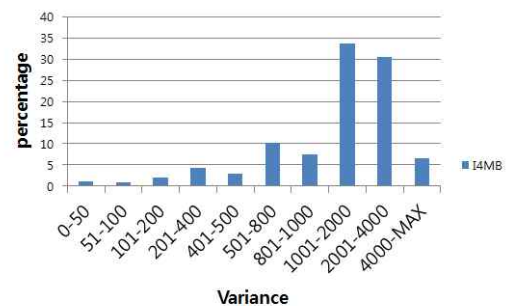


그림 3. I4MB의 분산 분포도

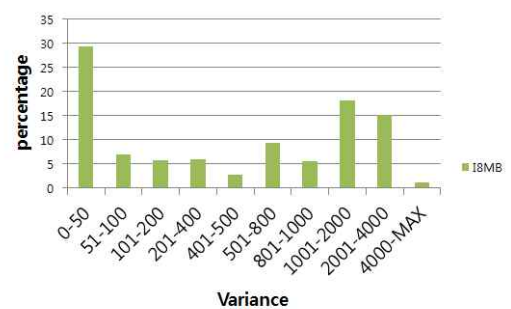


그림 4. I8MB의 분산 분포도

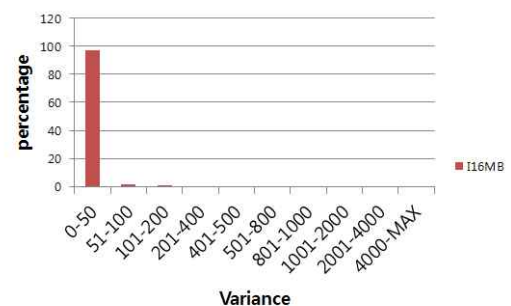


그림 5. I16MB의 분산 분포도

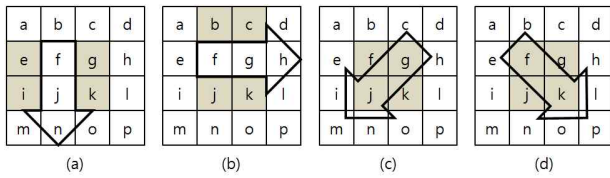


그림 4. IGF를 이용한 4가지의 방향 분석
(a) 수직, (b) 수평, (c) 하단 좌측 대각, (d) 하단 우측 대각 방향

3.2 RDO에 사용될 후보 예측 보드 선택 방법

RDO에 사용될 후보 예측 모드를 선택하기 위해서, 본 논문에서는 Tsai 가 제안한 알고리즘을 기반으로 이를 개선하였다.[8] 색차 신호 8x8 블록은 SATD(Sum of Absolute Hardmard Transform Differences)를 이용하여 최소 SATD 값을 가진 모드를 후보 예측 모드로 사용한다. 휘도 신호에서는 블록 내의 방향성을 검출하기 위하여 IGF(Intensity Gradient Filter)가 사용되었다. IGF는 식(4) 과 같이 정의될 수 있다.

$$G(x) = -1/2 \cdot P_{x-1} + 1/2 \cdot P_{x+1} \quad (4)$$

G(x)는 화소 x의 기울기 강도를 나타낸다. G(x)가 0에 가깝거나 0이면, 방향성이 존재한다고 할 수 있다. 그림 4는 IGF를 사용하여 검출된 방향을 보여준다. 휘도 신호 I4MB에서는 후보 예측 모드들이 다음과 같이 선택 된다.

1) 그림 4.(a)의 수직 방향에 해당하는 기울기 강도를 식 (5) 를 이용하여 계산한다.

$$\begin{aligned} G(e) &= |-1/2 \cdot P_a + 1/2 \cdot P_i| \\ G(i) &= |-1/2 \cdot P_e + 1/2 \cdot P_m| \\ G(g) &= |-1/2 \cdot P_c + 1/2 \cdot P_k| \\ G(k) &= |-1/2 \cdot P_g + 1/2 \cdot P_o| \\ G(ver) &= (G(e) + G(i) + G(g) + G(k)) \gg 2 \end{aligned} \quad (5)$$

2) 그림 4.(b)의 수평 방향에 해당하는 기울기 강도를 식 (6) 를 이용하여 계산한다.

$$\begin{aligned} G(b) &= |-1/2 \cdot P_a + 1/2 \cdot P_c| \\ G(c) &= |-1/2 \cdot P_b + 1/2 \cdot P_d| \\ G(j) &= |-1/2 \cdot P_i + 1/2 \cdot P_k| \\ G(k) &= |-1/2 \cdot P_j + 1/2 \cdot P_l| \\ G(hor) &= (G(b) + G(c) + G(j) + G(k)) \gg 2 \end{aligned} \quad (6)$$

3) 4.(c)와 4.(d) 에 나타낸 하단 좌측 대각 방향, 하단 우측 대각 방향의 기울기 강도를 1)과 2)에서 계산한 방법에 따라 계산한다. (G(dia_left), G(dia_right))

4) 보간법을 이용하여 나머지 방향의 기울기를 식 (7)과 같이 계산한다. 계산된 기울기의 방향은 화면 내 예측의 예측 모드의 방향과 동일하다.

$$\begin{aligned} G(ver_right) &= (G(ver) + G(dia_right)) \gg 1 \\ G(hor_down) &= (G(hor) + G(dia_right)) \gg 1 \\ G(ver_left) &= (G(ver) + G(dia_left)) \gg 1 \\ G(hor_up) &= (G(hor) + G(dia_left)) \gg 1 \end{aligned} \quad (7)$$

5) 계산된 G(x) 들 중 가장 작은 값을 가진 3가지, MPM(Most

Probable Mode), DC를 모드 후보 예측 모드로 선택한다. MPM은 좌측과 상단 이웃 블록중 낮은 인덱스를 가진 예측 모드이다. 좌측 블록이나 상단 블록이 존재 하지 않는다면 MPM은 DC가 된다. 따라서 후보 예측모드는 최소 4개에서 최대 5개가 될 수 있다. 한편 계산된 G(x)들은 정수 단위의 정밀도를 가지기 때문에, 최악의 경우 8개의 G(x) 모두 같은 값을 가질 수 있다. 이 경우 낮은 인덱스를 가진 예측 모드가 후보 예측 모드로 선택된다.

6) 선택된 후보 예측 모드만 RD 비용을 계산하고, 이중 최소의 RD 비용을 가진 예측모드를 이용하여 매크로블록 부호화를 진행한다.

I8MB의 경우 그림 5와같이 매크로블록을 16개의 2x2 블록으로 분할 한 후 식 (8)과 같이 평균값을 취한다. 그 후 I4MB와 동일하게 후보 예측 모드를 선택한다.

$$a_8 = \left(\sum_{x=0}^3 A_x \right) \gg 2 \quad (8)$$

단 최소 G(x)의 예측 모드가 MPM과 동일할 경우, 최소 G(x) 값을 가진 두 개의 모드만 후보 예측 모드로 선택된다. I16MB의 경우 16 개의 4x4 블록으로 분할 한 후, G(ver), G(hor), G(dia_left)를 각각 계산하고, 최소 G(x)의 예측모드와 DC를 후보 예측 모드로 선택한다. 따라서 2개의 예측 모드만 RDO를 위하여 사용된다.

4. 실험 결과 및 분석

본 논문에서 제안된 알고리즘은 JM 13.2를 기반으로 구현되었다. 실험 플랫폼으로써, CPU는 Intel Core 2 Duo (3.16 Ghz, 3.16 Ghz), 3 GB의 DDR2 램, 운영체제는 Windows 7 32bit 버전을 사용하였다.

프로파일 IDC 는 100을 사용하였고, 모든 실험 시퀀스는 I 프레임으로 100프레임을 부호화 하였다. RDO, CABAC, Transform8x8 플레그를 활성화 하였으며, QP는 16, 20, 24, 28로 입력하여 부호화를 진행 하였다. 실험에 사용된 시퀀스는 총 12개 이며 QCIF 3개, CIF 3개, WVGA 3개, HD 3개로 이루어져 있다. 블록 모드를 결정하기 위한 임계값은 49759 이며 실험적으로 계산되었다. 부호화 효율을 검증하기 위하여, BDBR(Delta Bit-rate), BDPSNR(Bjontegarrd Delta Peak Signal-to-Noise Ratio), TR(Time Reduction)을 각각 계산하였다.[11] TR은 식 (9)과 같이 계산할 수 있다.

$$TR = \frac{T_{proposed} - T_{anchor}}{T_{anchor}} \times 100 \quad (9)$$

T_anchor 는 JM 13.2에서 화면 내 예측을 위하여 모든 블록 모드와 예측모드를 고려하였을 때의 부호화 시간이다. T_proposed는 제안된 알고리즘을 사용하여 부호화 하였을 때의 소요 시간이다. 제안된 알고리즘의 성능 평가를 위하여 Wang[6], Tsai[8], Huang[10]과 비교 분석하였다.

표 1은 제안된 알고리즘과 기존에 제안된 알고리즘들을 수치적으로 비교한 결과이다. 제안된 알고리즘을 Wang[8] 과 Tsai[9] 와 비교 하면 비트율과 PSNR의 큰 손실 없이도 약 13%정도의 화면 내 부호화 시간을 줄일 수 있으며 Huang[10] 과 비교하면 약 2% 정도의 속도 향상을 보였다. 또한 JM 13.2 와 비교하면 1.986%의 비트 율 증가와

0.159dB의 PSNR 손실을 보였지만 평균83.659%의 화면 내 예측 부호화 시간을 줄였고 최대 86.416%까지 줄일 수 있었다.

5. 결론

본 논문은 H.264/AVC 하이 프로파일의 화면 내 예측을 위하여 효율적인 블록 모드 결정과 예측 모드 선택에 관한 알고리즘을 연구하였다. 현재 매크로 블록의 분산을 계산한 후 임계치를 사용하여 적절한 블록 모드를 결정하였고, 에지 필터를 사용하여 후보 예측 모드를 선택하였다. 제안된 방법은 기존 알고리즘들과 비교하여 큰 부호화 성능의 손실 없이 효과적으로 화면 내 예측 부호화 시간을 줄일 수 있었다.

감사의 글

이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(2011-0011313)

참고문헌

[1] ITU-T Recommendation, H.264 Advanced Video Coding for Generic Audiovisual services, 2010.
 [2] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard," IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, pp. 560-576, Jul. 2003.
 [3] G. J. Sullivan and T. Wiegand, "Rate-Distortion Optimization for Video Compression," IEEE Signal Processing Magazine, vol. 15, pp. 74-90, Nov. 1998.
 [4] Z. Chen, J. Xu, Y. He, and J. Zheng, "Fast integer-pel and

fractional-pel motion estimation for H.264/AVC," Journal of Visual Communication and Image Representation, vol. 17, pp. 264-290, Oct. 2005.
 [5] D. Wu, F. Pan, K. P. Lim, S. Wu, Z. G. Li, X. Lin, S. Rahardja and C. C. Ko, "Fast Intermode Decision in H.264/AVC Video Coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, pp. 953-958, Jul. 2005.
 [6] J. C. Wang, J. F. Wang, J. F. Yang, and J. T. Chen, "A Fast Mode Decision Algorithm and Its VLSI Design for H.264/AVC Intra-Prediction," IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, pp. 1414 - 1422, Oct. 2007.
 [7] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. Wu, and S. Wu, "Fast Mode Decision Algorithm for Intraprediction in H.264/AVC Video Coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, pp. 813 - 822, Jul. 2005.
 [8] A. C. Tsai, A. Paul, J. C. Wang and J. F. Wang, "Intensity Gradient Technique for efficient Intra-Prediction in H.264/AVC," IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, pp. 694-698, May 2008.
 [9] K. Bharanitharan, A. C. Tsai, "Efficient Block Size Decision Algorithm for Intra Mode Decision in H.264/AVC Encoder," in Proc. International Symposium on Multimedia 2009, San-Diego, California, Dec. 14-16. 2009, pp. 96-99.
 [10] Y. H. Huang, T. S. Ou, and H. H. Chen, "Fast Decision of Block Size, Prediction Mode, and Intra Block for H.264 Intra Prediction," IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, pp. 1122-1132, Aug. 2010.
 [11] G. Bjontegaard, "Calculation of Average PSNR Difference between RD-curves," in Proc. 13th Meeting ITU-T Q.6/SG16 VCEG, Austin, Texas, April 2-4. 2001, document VCEG-M33.doc.

표 1 각 시퀀스에 따른 제안된 알고리즘의 실험 결과와 이전 알고리즘과의 비교

Size	Sequence	Wang[6]			Tsai[8]			Huang[10]			Proposed		
		BDBR (%)	BDPSNR (db)	TR (%)	BDBR (%)	BDPSNR (db)	TR (%)	BDBR (%)	BDPSNR (db)	TR (%)	BDBR (%)	BDPSNR (db)	TR (%)
QCIF	Carphone	2.213	-0.189	-75.996	1.742	-0.149	-75.898	3.461	-0.292	-80.834	3.240	-0.274	-84.045
	Coastguard	1.729	-0.187	-75.397	1.050	-0.112	-77.034	0.737	-0.079	-82.840	0.950	-0.102	-85.651
	Stefan	1.543	-0.211	-74.626	1.058	-0.145	-77.192	2.160	-0.293	-78.428	2.199	-0.298	-82.476
CIF	News	2.130	-0.169	-75.393	1.532	-0.122	-78.494	2.569	-0.203	-85.516	2.476	-0.196	-86.359
	Paris	1.644	-0.182	-74.780	1.225	-0.136	-77.960	1.782	-0.197	-82.039	1.919	-0.212	-83.764
	Waterfall	0.986	-0.097	-74.726	1.376	-0.135	-77.539	0.826	-0.081	-87.166	1.453	-0.142	-86.144
WVGA	Flower	2.326	-0.160	-65.315	1.779	-0.123	-68.038	4.310	-0.292	-85.304	3.488	-0.237	-84.857
	Keiba	1.723	-0.103	-62.735	1.005	-0.060	-63.184	1.514	-0.090	-77.704	1.557	-0.093	-79.354
	Nuts	1.263	-0.046	-63.215	1.184	-0.043	-64.210	2.852	-0.100	-84.026	2.589	-0.091	-86.136
HD	Bigships	1.513	-0.109	-63.098	1.057	-0.076	-63.335	1.891	-0.137	-78.733	1.660	-0.121	-80.035
	Night	1.266	-0.111	-63.739	1.025	-0.090	-63.644	0.856	-0.076	-75.908	1.072	-0.094	-78.676
	Shuttle Start	1.449	-0.059	-62.792	1.099	-0.045	-65.536	1.105	-0.046	-84.123	1.228	-0.050	-86.416
Average		1.649	-0.135	-69.318	1.261	-0.103	-71.005	2.005	-0.157	-81.885	1.986	-0.159	-83.659