A Hybrid Range Image Coding Algorithm

Yingshi Li

A range image

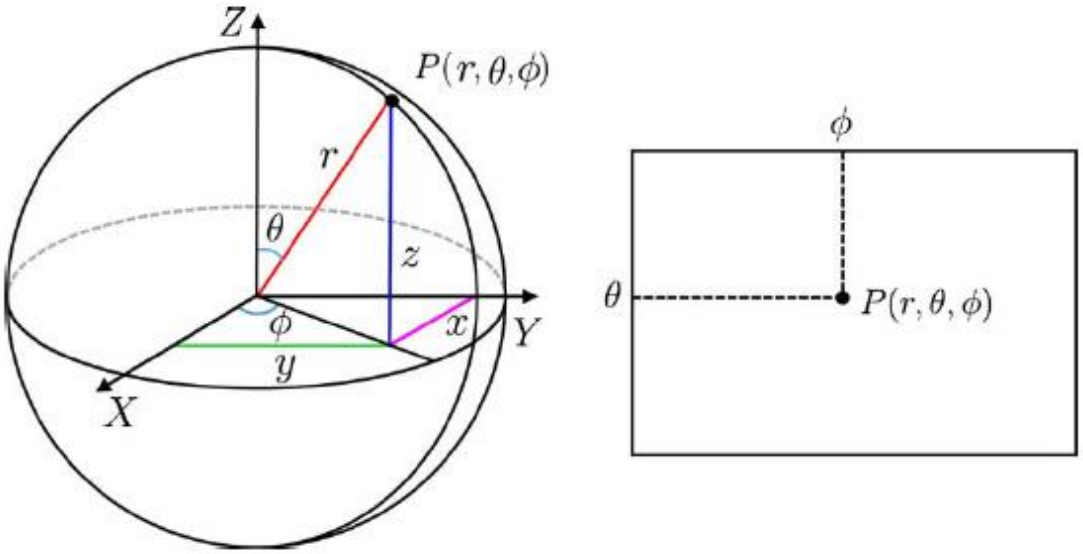


Fig. 2. Two coordinate systems for 3D point geometry. Each point in the spherical coordinate system corresponds to each pixel in the range image.

$$x = r \sin\theta \cos\phi, \quad (1)$$

$$y = r \sin\theta \sin\phi, \quad (2)$$

$$z = r \cos\theta. \quad (3)$$

- 单帧压缩，比较新，还在研究中
- 与jpg格式文件类似，算法比较JPEG2000（J2K）等软件
- 分成32X32到8X8的格子，每个格子适用不同的prediction mode

(a) (b) (c) (d)

Fig. 9. Adaptive block partitioning of range images: (a) the color image and (b) the range image of the "Amphitheatre" model, and (c) the color image and (d) the range image of the "Library" model, at the quantization step size of Δ_{14}. The biggest and the smallest block sizes are 32×32 and 8×8, respectively.

Prediction modes

RADIAL DISTANCE PREDICTION MODES, WHERE 'H' INDICATES THE PREDICTION IN THE HEIGHT IMAGE DOMAIN.

Index	Prediction mode	Index	Prediction mode
0	Horizontal	6	Diagonal down-right
1	Horizontal (H)	7	Diagonal down-right (H)
2	Vertical	8	RBF
3	Vertical (H)	9	RBF (H)
4	Diagonal down-left	10	Pixel-wise plane prediction
5	Diagonal down-left (H)	11	Block-wise plane prediction

Radial distance modes

$$\hat{r}(\mathbf{p}) = \begin{cases} \tilde{r}(\mathbf{p}_{\mathrm{l}}), & \text{horizontal,} \\ \tilde{r}(\mathbf{p}_{\mathrm{u}}), & \text{vertical,} \\ \tilde{r}(\mathbf{p}_{\mathrm{ul}}), & \text{diagonal down-right,} \\ \tilde{r}(\mathbf{p}_{\mathrm{ur}}), & \text{diagonal down-left,} \end{cases}$$

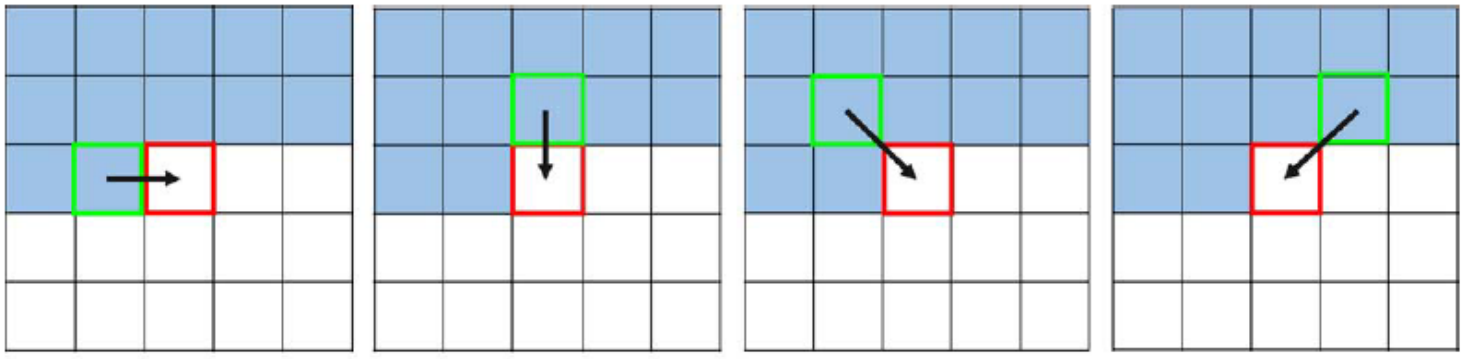


Fig. 5. Four nearest neighbor (NN) prediction modes: (a) horizontal, (b) vertical, (c) diagonal down-right, and (d) diagonal down-left modes. A red pixel is predicted from a green pixel. Blue regions and white regions correspond to encoded and unencoded pixels, respectively.

Pixel-wise plane prediction

- 笛卡尔坐标系下平面：ax+by+cz=d
- 平面上的点P如图，n为向量(a,b,c)

$$\hat{r}(\mathbf{p}) = \frac{d}{\mathbf{u_p} \cdot \mathbf{n}}$$

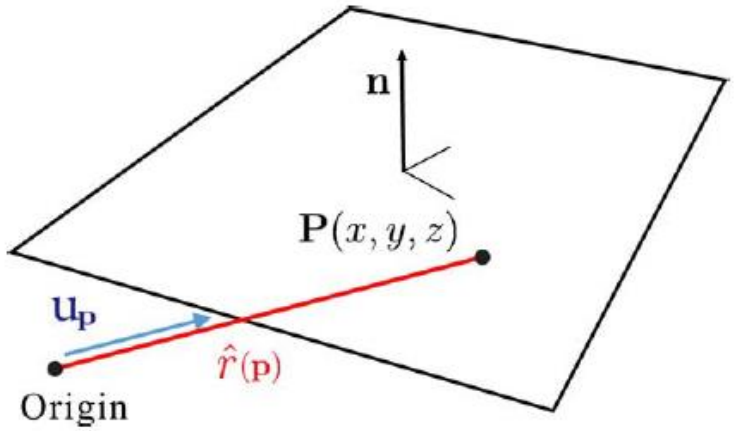


Fig. 8. Plane-based prediction of a radial distance.

Pixel-wise plane prediction

- 7x7的以P为中心的平面上，求最小值

$$(\mathbf{n}^*, d^*) = \arg\min_{(\mathbf{n},d)} \sum_{\mathbf{p}_i \in S_\mathbf{p}} w_i \left\| \tilde{\mathbf{P}}_i \cdot \mathbf{n} - d \right\|^2$$

$$w_i = \frac{\frac{1}{\|\mathbf{p}-\mathbf{p}_i\|}}{\sum_{\mathbf{p}_j \in S_\mathbf{p}} \frac{1}{\|\mathbf{p}-\mathbf{p}_j\|}}$$

Threshold=0.005

Predictor Selection

- Step size： $\Delta_q = \frac{M}{2^q}$

  q=1,2,3,4……

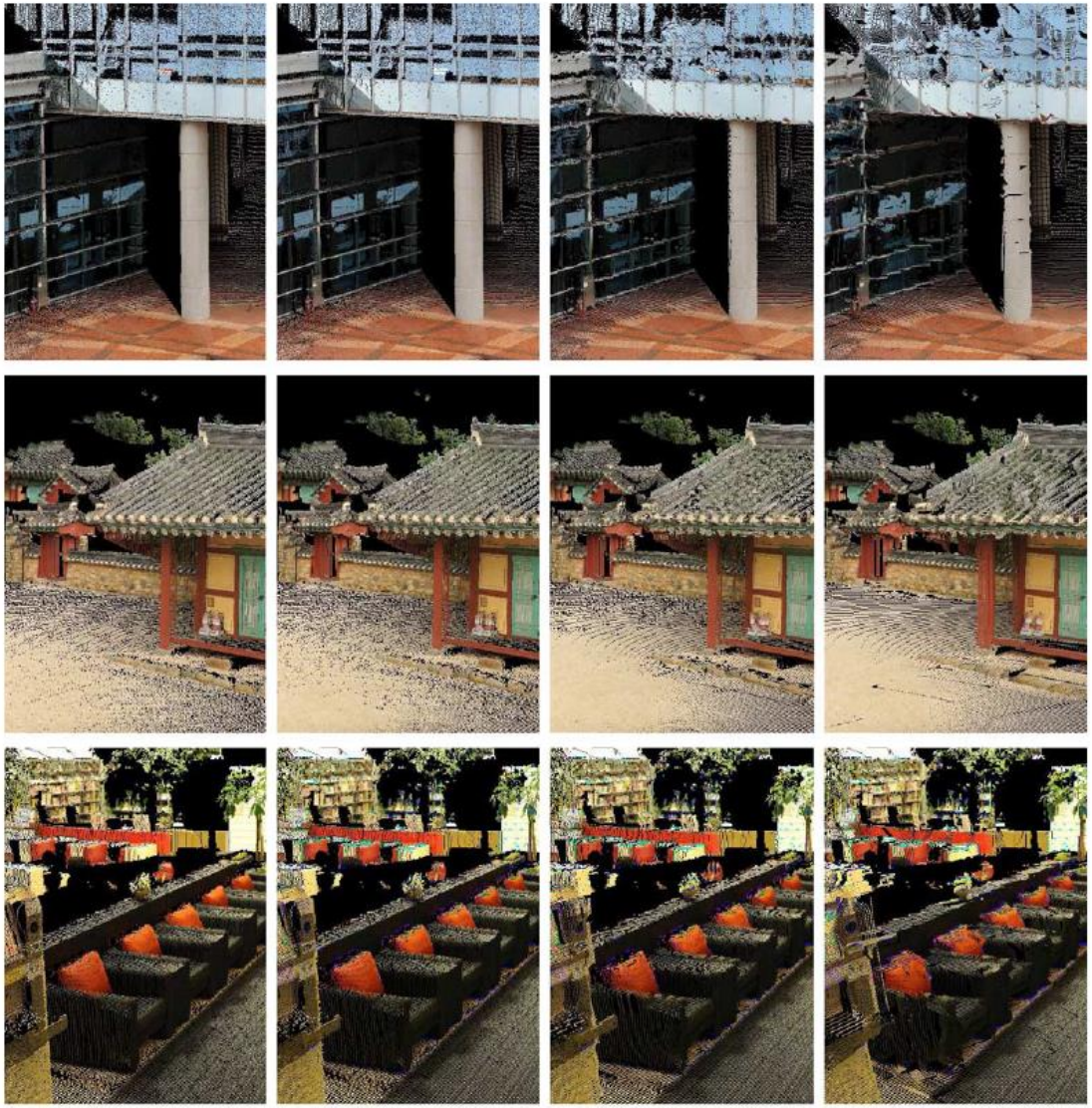
  M is the difference between the maximum and the minimum radial distances of the imput model

- 通过计算每个block的rate-distortion cost，来选择prediction mode(*k*), 令J_k最小

$$J_k = D_k + \lambda R_k \qquad \lambda = \alpha \Delta_q^2$$

  - D_k is the SSD of the pixels，R_k is the required bit rate（bpp）

- 32X32-->16x16-->8x8, if sum of J_k变小



(a) (b) (c) (d)

Fig. 13. Reconstructed LS3DPC models of "Campus," "Korean House," and "Library," when different quantization step sizes are used for the prediction residuals: (a) original, (b) reconstruction at Δ_{12}, (c) reconstruction at Δ_{10}, and (d) reconstruction at Δ_{8}.

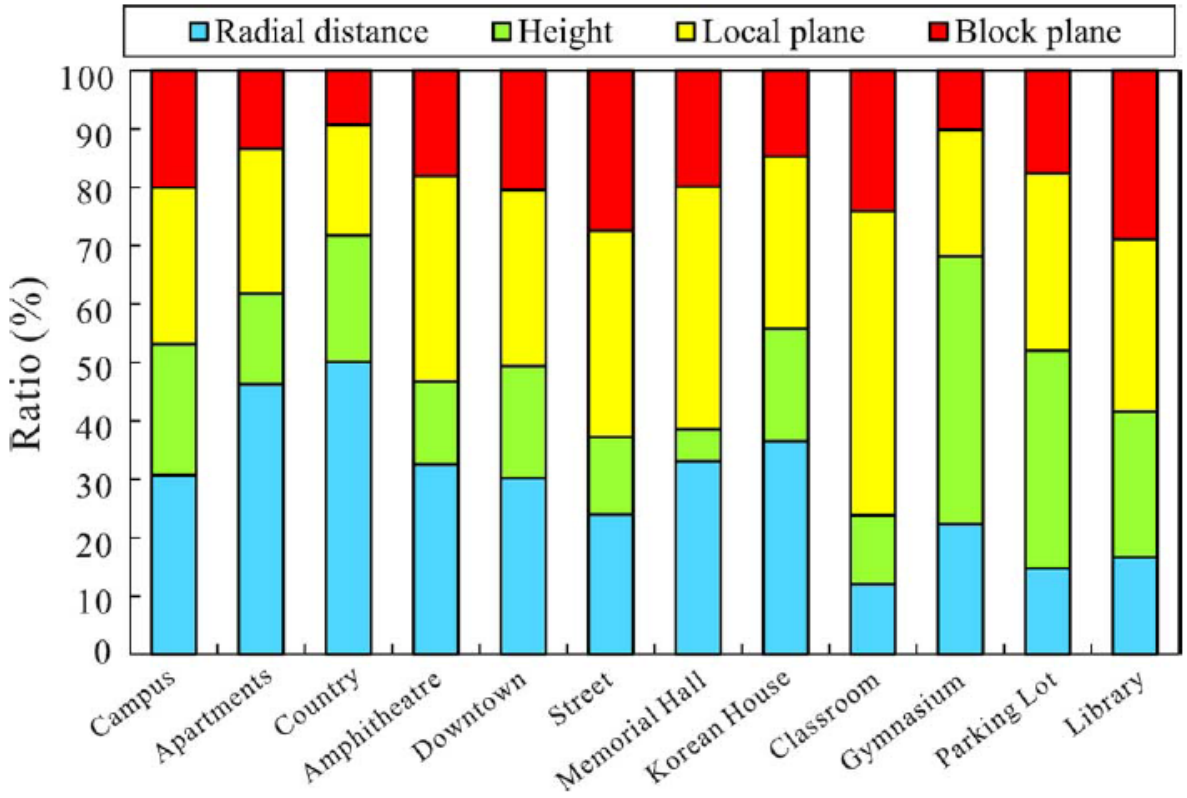


Fig. 12. Selection ratios of the prediction modes, when the quantization step size is Δ_{12}.

与图像压缩软件比较

TABLE V

COMPARISON OF THE REQUIRED BIT-RATES OF THE PROPOSED ALGORITHM AND THE CONVENTIONAL JPEG2000 (J2K), JPEG-LS (JLS), H.264/AVC, AND HEVC. THE UNIT OF BIT-RATES IS BITS PER POINT (BPP). THE QUANTIZATION STEP SIZE FOR PREDICTION RESIDUALS IN THE PROPOSED ALGORITHM AND THE PRE-QUANTIZATION STEP SIZE IN THE CONVENTIONAL TECHNIQUES ARE BOTH FIXED TO Δ_{12}. THUS, THE DISTORTIONS OF THE PROPOSED ALGORITHM ARE ALWAYS SMALLER THAN THOSE OF THE CONVENTIONAL TECHNIQUES.

Test model	J2K	JLS	H.264	HEVC	Proposed
Campus	2.23	2.14	1.48	1.74	1.13
Apartments	3.36	3.25	2.48	2.81	2.13
Country	4.53	4.27	3.54	3.91	3.00
Amphitheatre	1.77	1.69	1.18	1.38	0.92
Downtown	3.05	2.98	2.07	2.44	1.59
Street	1.72	1.71	1.09	1.28	0.72
Memorial Hall	1.32	1.28	0.90	1.08	0.66
Korean House	1.80	1.69	1.26	1.47	0.99
Classroom	2.57	2.66	2.38	2.55	1.91
Gymnasium	4.99	5.12	4.53	5.04	3.52
Parking Lot	1.82	1.64	1.27	1.50	0.84
Library	2.18	2.19	1.60	1.82	1.12
Average	2.61	2.55	1.98	2.25	1.54

与图像压缩软件比较

TABLE VI
COMPARISON OF THE ENCODING TIMES (IN SECONDS) OF THE PROPOSED ALGORITHM AND THE CONVENTIONAL JPEG2000 (J2K) [44], JPEG-LS (JLS) [45], H.264/AVC [46], AND HEVC [47].

Test model	J2K	JLS	H.264	HEVC	Proposed
Campus	57.10	56.78	97.54	79.75	87.93
Apartments	45.11	44.55	89.82	67.78	104.15
Country	18.31	17.85	48.13	33.77	62.67
Amphitheatre	17.89	17.35	39.27	32.88	67.71
Downtown	35.23	34.80	65.53	53.66	67.62
Street	25.07	24.90	54.31	45.92	85.21
Memorial Hall	8.93	8.69	28.81	22.84	67.33
Korean House	12.60	12.31	33.84	26.40	64.10
Classroom	17.32	16.94	69.52	41.99	115.20
Gymnasium	11.54	10.68	78.08	37.99	134.59
Parking Lot	10.00	9.51	49.76	32.64	118.61
Library	10.03	9.40	52.19	33.28	122.41
Average	22.43	21.98	58.90	42.41	91.46

Context-based entropy coding

- Header context
- Flag bits of partitioned blocks
- For each block
  - Prediction mode index
  - Prediction residuals
  - others

The End