

quantization of \mathbf{p}

$$\left\{ \begin{array}{l} \mathbf{p} = (14.9, 81.96, 17.65) \\ \mathbf{q} = (\mathbf{p} + \mathbf{o}) / s * 2^{L=8} = (38, 209, 45), (s=100, o=0) \\ \mathbf{q}_{b=2} = (00100110_b, 11010001_b, 00101101_b) \end{array} \right.$$

layer	1	2	3	4	5	6	7	8
bit pos	7	6	5	4	3	2	1	0
0	0.0	0.0	1.0	0.0	0.0	1.0	1.0	0.0
1	1.0	1.0	0.0	1.0	0.0	0.0	0.0	1.0
2	0.0	0.0	1.0	0.0	1.0	1.0	0.0	1.0

quantized point

0	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
---	-----	-----	-----	-----	-----	-----	-----	-----

half mask

X

0	1.0	1.0	1.0	1.0	1.0	0.0	0.0	0.0
---	-----	-----	-----	-----	-----	-----	-----	-----

mask layer 5

legend

1.0
0.5
0.0

=

0	-0.5	-0.5	0.5	-0.5	-0.5	0.0	0.0	0.0
1	0.5	0.5	-0.5	0.5	-0.5	0.0	0.0	0.0
2	-0.5	-0.5	0.5	-0.5	0.5	0.0	0.0	0.0

feature

-0.5
-1.0