ETHzürich

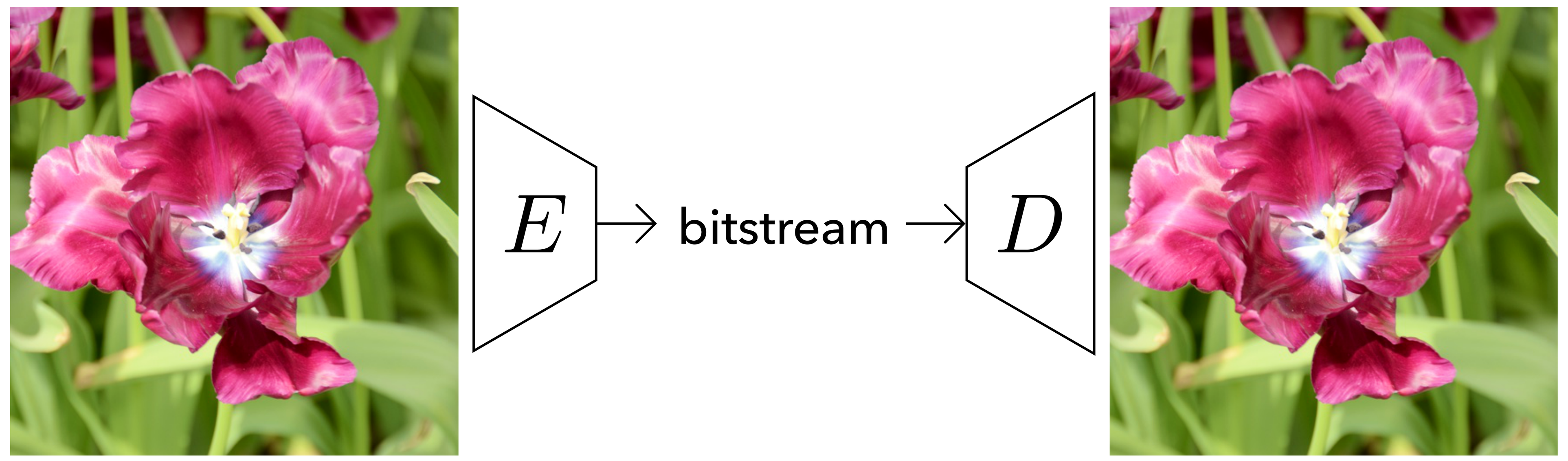
CVL Computer Vision Lab

@mentzer_f

Practical Full Resolution Learned Lossless Image Compression

Fabian Mentzer, Eirikur Agustsson, Michael Tschannen, Radu Timofte, Luc Van Gool

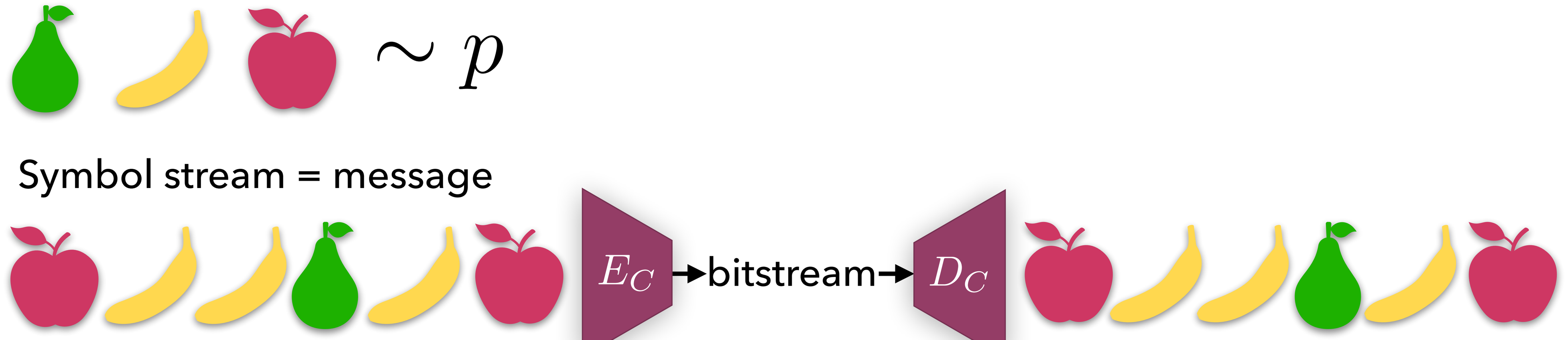
Overview



- **First Learned** Full-Resolution Lossless Image Compression
- **Smaller** than non-learned lossless codecs: PNG, JPEG2000, WebP
- Our probabilistic model is orders of magnitude **faster** than PixelCNN

Lossless Compression

Symbols — Probability distribution



Entropy Coding:
Arithmetic Coding
Huffman Coding

$p($ image $)$ unknown

→ we learn a **model** $\tilde{p}$

Use pixels as symbols: not i.i.d.
So we learn the joint:
$\tilde{p}(x_1, x_2, \dots, x_N)$. **Must factorize!**

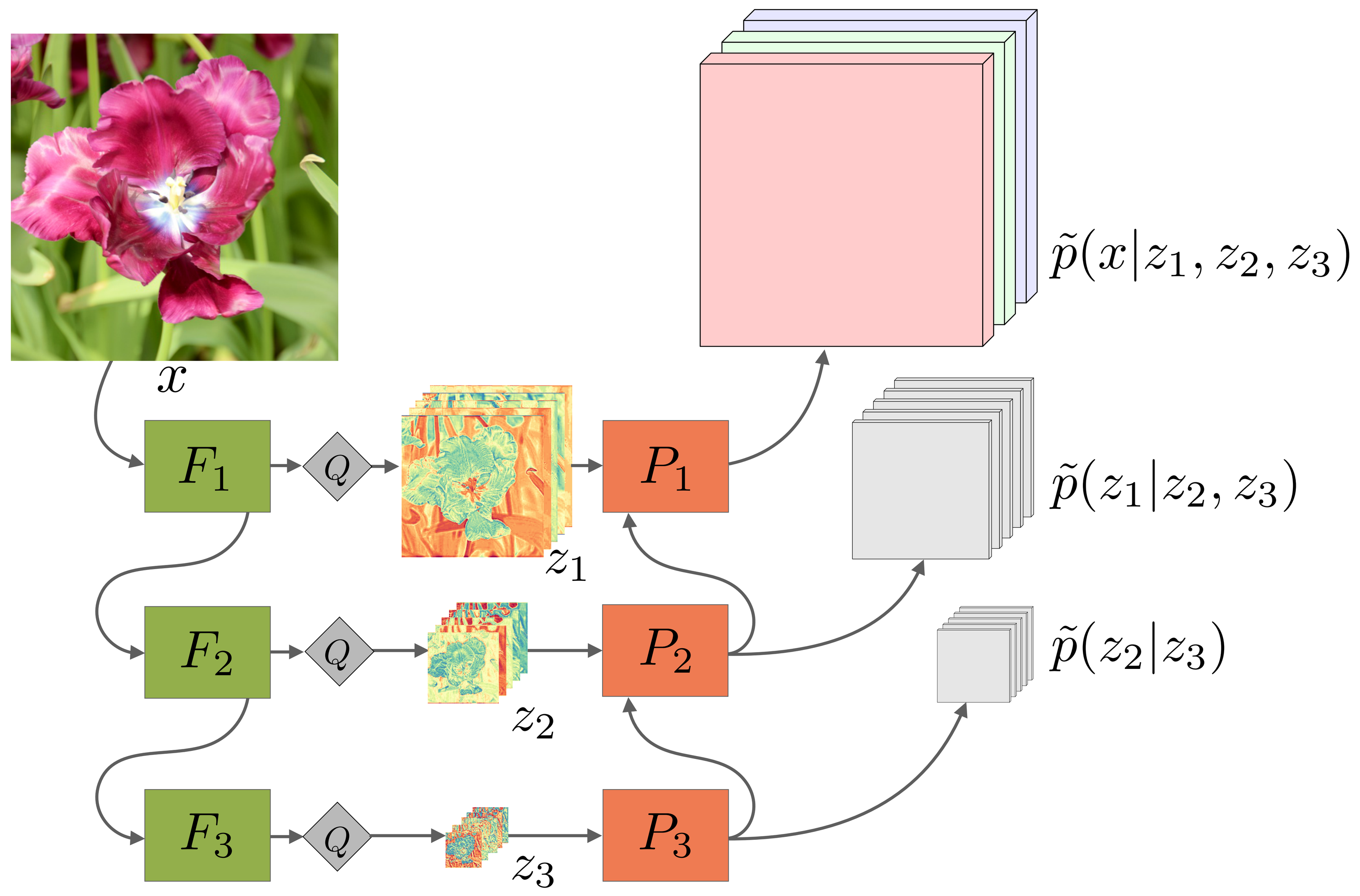
minimize **Cross Entropy**
$H(p, \tilde{p}) = \mathbb{E}_{x\sim p}[-\log \tilde{p}(x)]$

≡ minimize **-log likelihood** $\tilde{p}$

≡ minimize **bitcost when using** $\tilde{p}$

Our Method: L3C

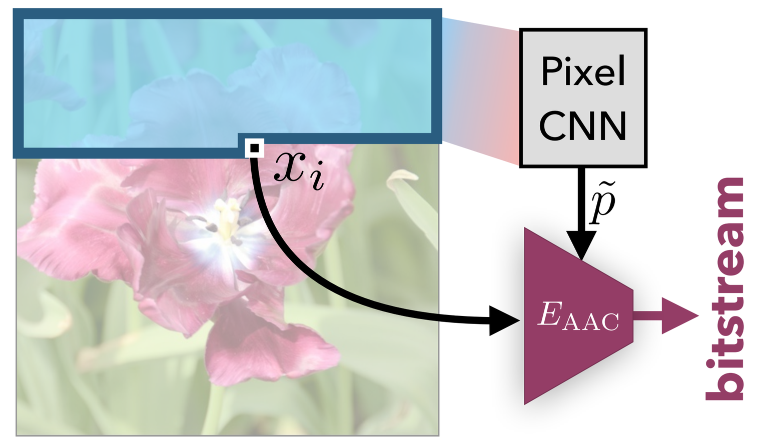
Our ***factorization***: $\tilde{p}(x) = \tilde{p}(x\,|z_1, z_2, z_3) \cdot \tilde{p}(z_1|z_2, z_3) \cdot \tilde{p}(z_2|z_3) \cdot \tilde{p}(z_3)$



Using discretized mixture of logistics

$$p_m(z^{(s)}_{cuv}|f^{(s+1)}) = \sum_k \pi^k_{cuv}\, p_l(x_{cuv}|\mu^k_{cuv}, \sigma^k_{cuv}).$$

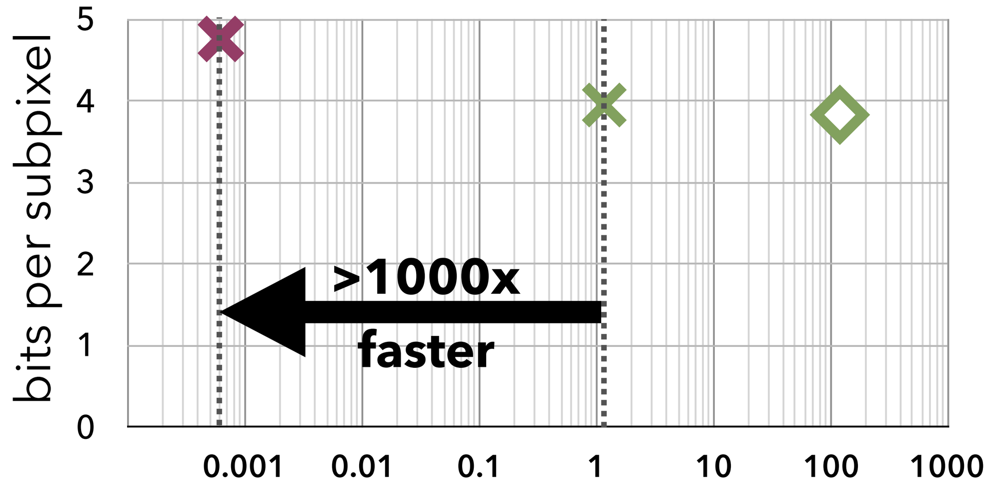
L3C vs. PixelCNN



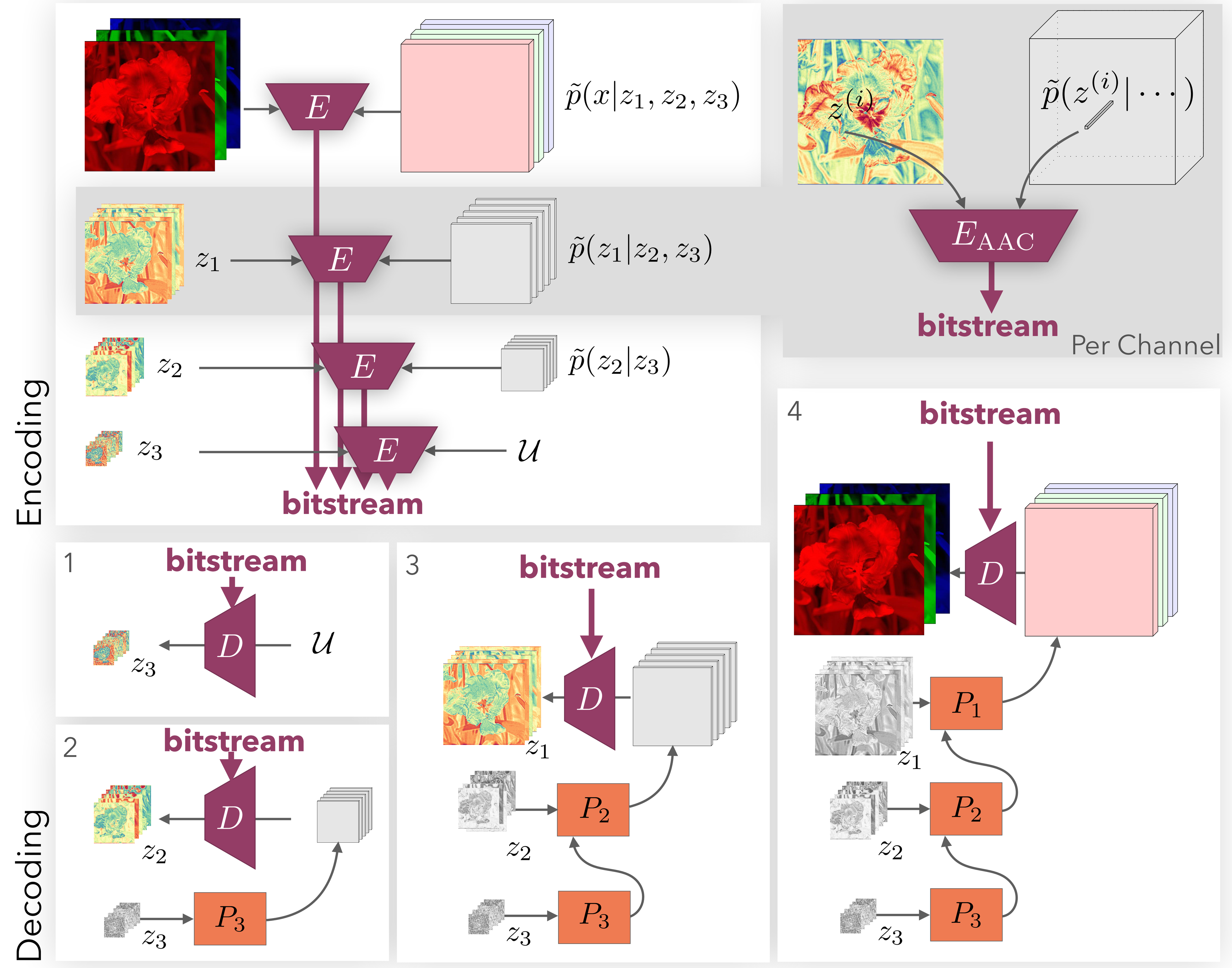
Their ***factorization***:

$$\tilde{p}(x) = \prod_{i=1}^{N} \tilde{p}(x_i|\boxed{x_1, \dots, x_{i-1}})$$

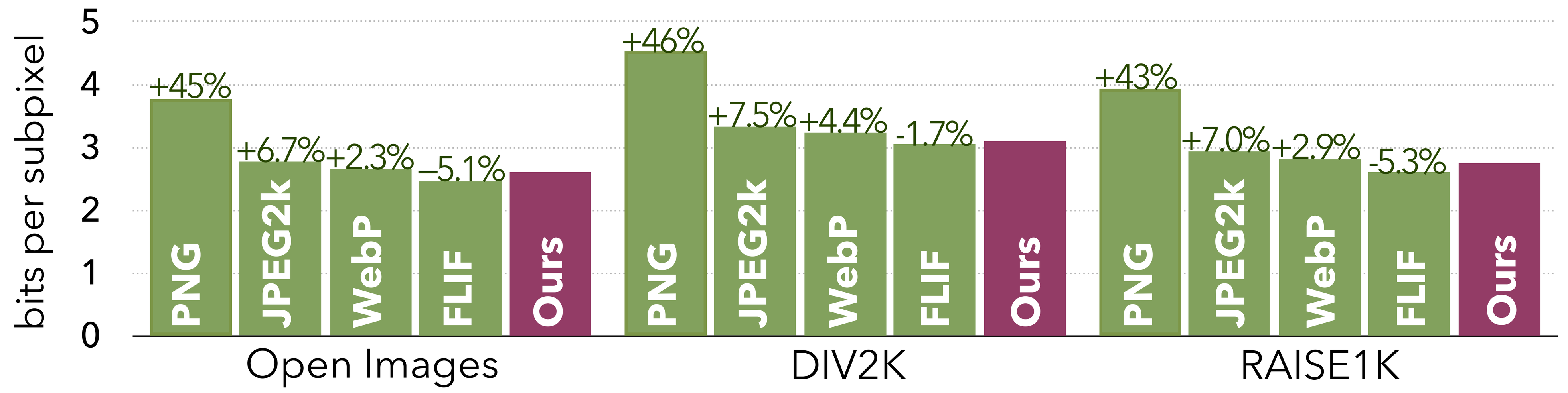
Forward pass per pixel!



Encoding Decoding Details



Results



Runtime on 512x512

Codec	Encoding [s]	Decoding [s]	GPU	CPU
Ours	0.242	0.374	✓	✓
FLIF	1.72	0.133		✓
WebP	0.157	0.0712		✓

Code
Models