Neural Distributed Image Compression with Cross-Attention Feature Alignment

Nitish Mital[*,1]***, Ezgi Ozyilkan***[†,1]***, Ali Garjani***[‡,1]***, Deniz Gunduz***[*]

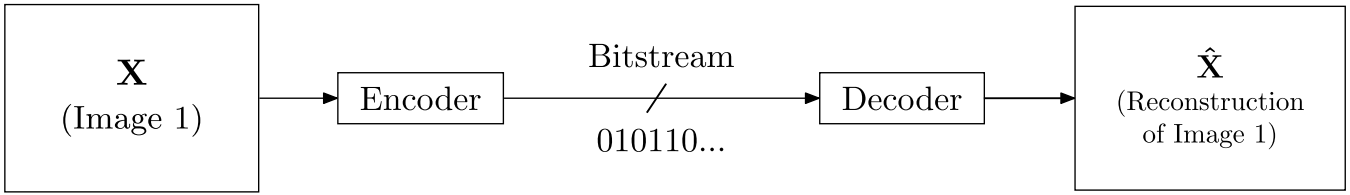
[*]*Dept. Of Electrical and Electronics Engineering, Imperial College London,*
[†]*Dept. Of Electrical and Computer Engineering, New York University,*
[‡]*Section of Mathematics, EPFL*
[1]*Equal contribution*

Imperial College London

EPFL

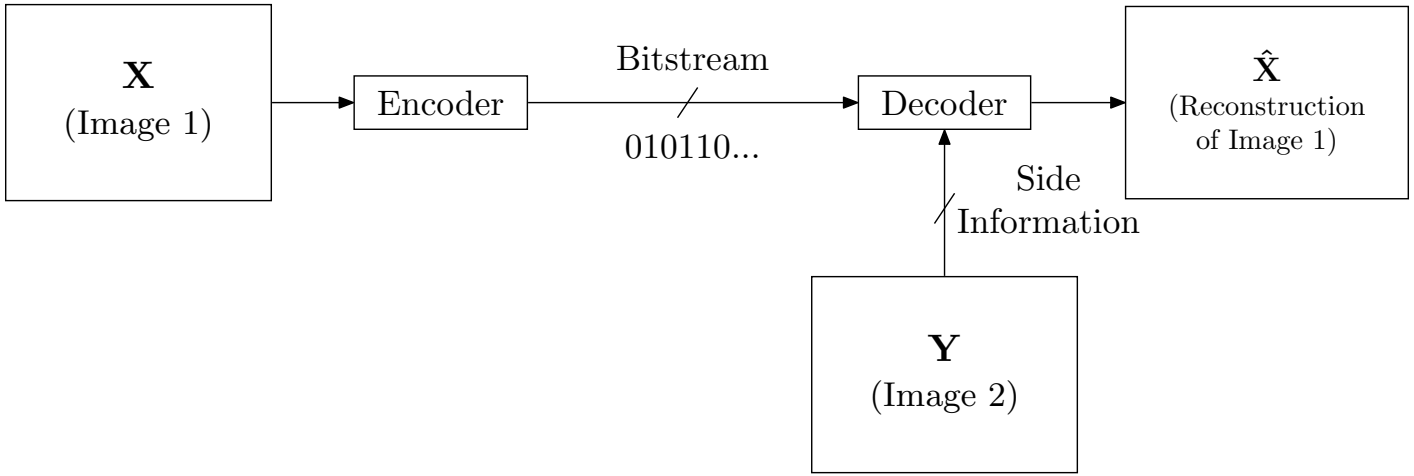
System Model: Point-to-Point

Compression
- Lossless
- Lossy ✓



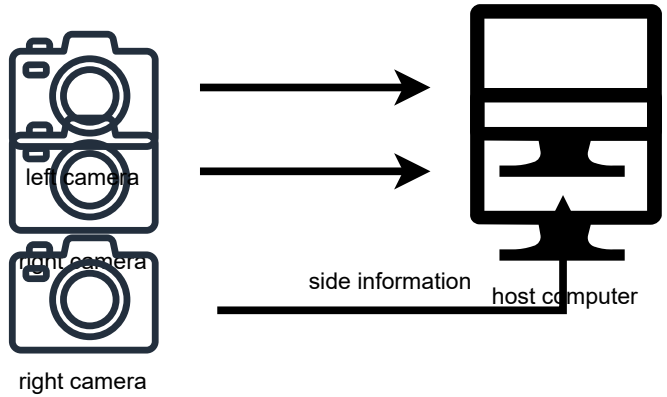
Two competing goals in lossy compression:
- Rate
- Distortion

System Model: Distributed Source Coding (DSC)



- Lossless compression (*Slepian and Wolf, 1973*)
- Lossy compression (*Wyner and Ziv, 1976*)

Motivation for DSC



Related Work

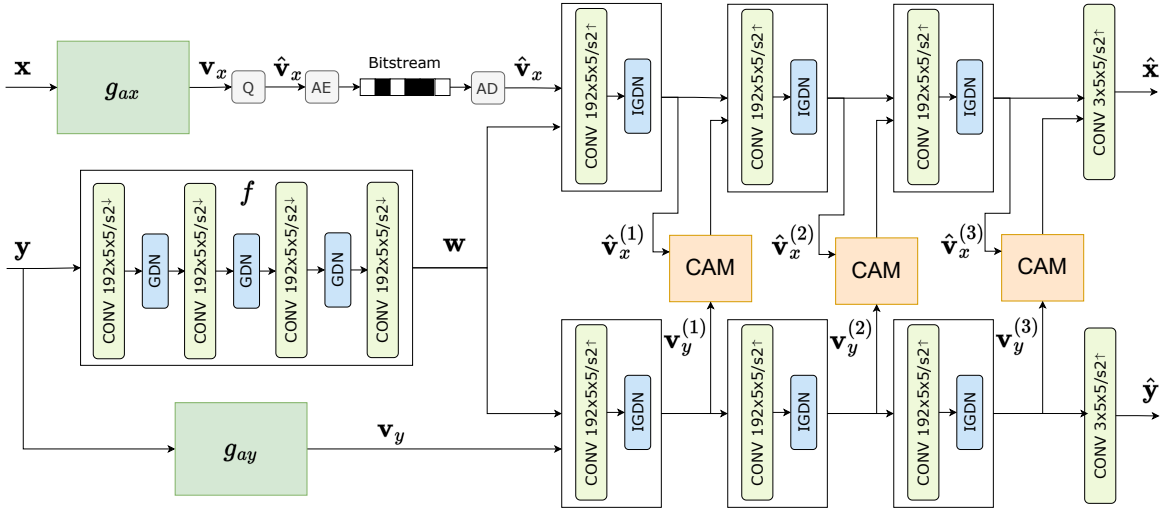
Distributed stereo compression (one image at encoder, other at decoder):

- DSIN (*S. Ayzik et. al., 2020*) — patch-match algorithm (not end-to-end differentiable)
- NDIC (*Our work, 2022*) — disentanglement into common and private features

This paper: *Combine DSIN and NDIC in a differentiable manner.*

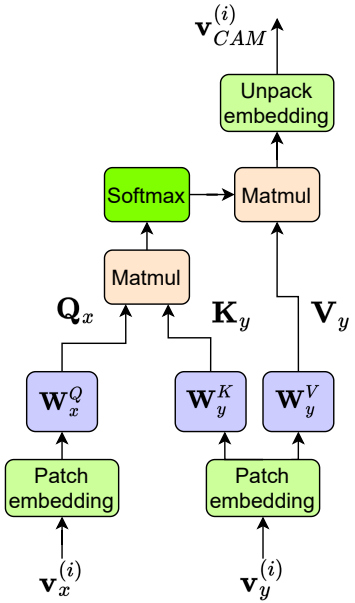
Align latent representations of the two images using a cross-attention mechanism!

Architecture



- $\mathbf{w}$ - common information
- $\mathbf{v}_x$, $\mathbf{v}_y$ - private/local information
- Extract $\mathbf{w}$ from $\mathbf{y}$, send only $\mathbf{v}_x$

- Align intermediate latents $\mathbf{v}_x^{(i)}$ and $\mathbf{v}_y^{(i)}$ (in i^{th} layer) using cross-attention module (CAM)
- Generate *query* $\mathbf{Q}_x$ from $\mathbf{v}_x^{(i)}$, *key* $\mathbf{K}_y$ and *value* $\mathbf{V}_y$ from $\mathbf{v}_y^{(i)}$



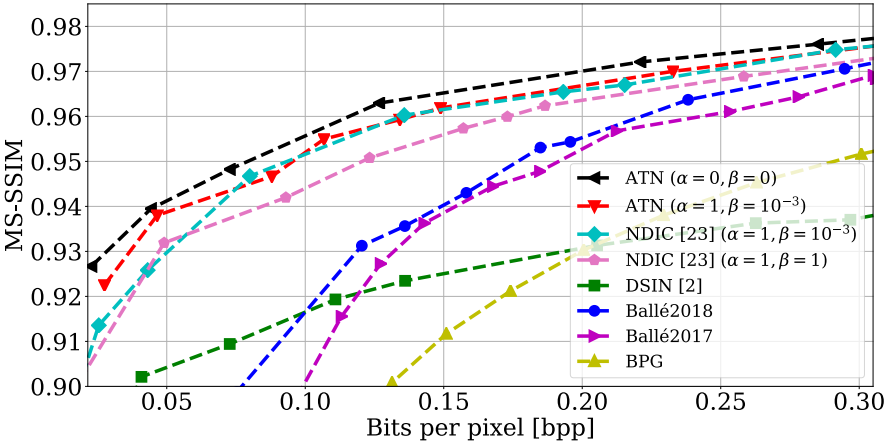
Experimental Setup

KITTI Stereo
(sync stereo)

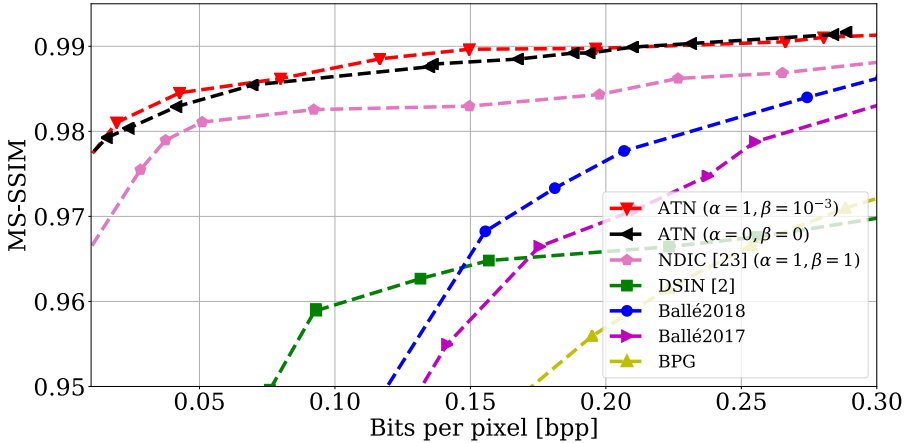
Cityscape
(sync stereo)

KITTI General
(unsync stereo)

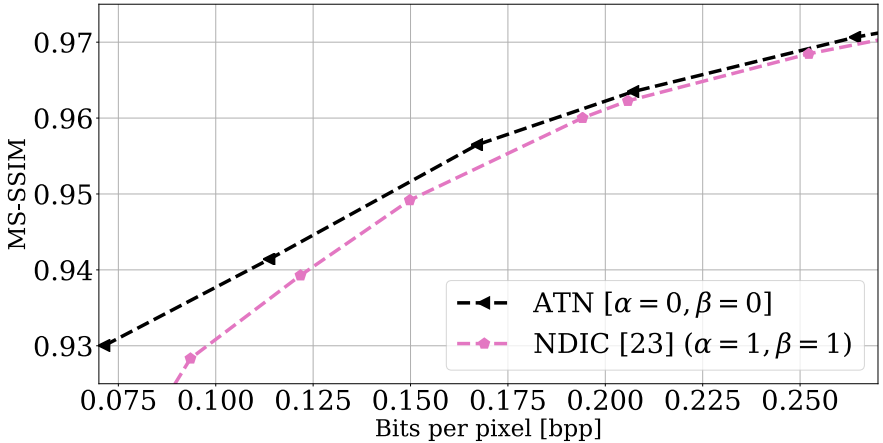
Results



KITTI Stereo (up)



Cityscape (up)



KITTI General (left)

Visual Examples

Synchronized stereo cameras

Original image

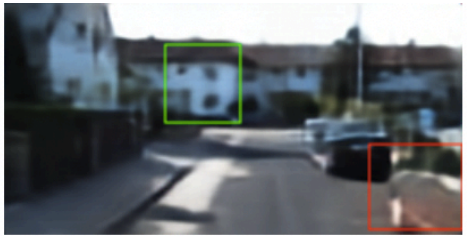
NDIC

bpp=0.0912

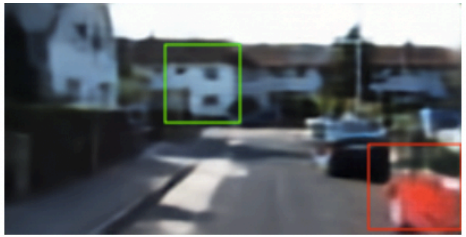
Ours

bpp=0.0725

Unsynchronized stereo cameras



bpp=0.1134



bpp=0.1071

Thanks!

Reach us:

- n.mital@imperial.ac.uk
- eo2135@nyu.edu
- ali.garjani@epfl.ch
- d.gunduz@imperial.ac.uk

Source code at: `https://github.com/ipc-lab/NDIC-CAM`