Quantum Implicit Neural Compression

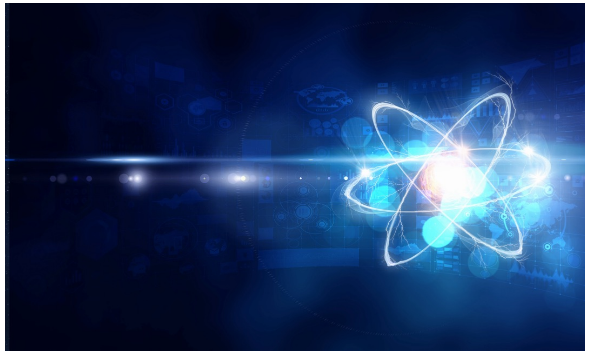
Takuya Fujihashi[1], Toshiaki Koike-Akino[2]

[1]Osaka University, [2]Mitsubishi Electric Research Laboratories (MERL)

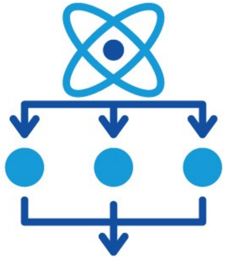
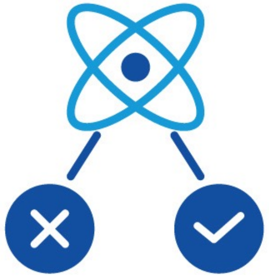
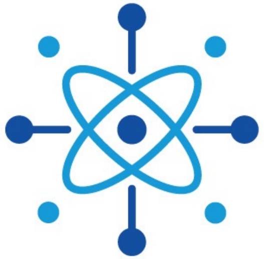
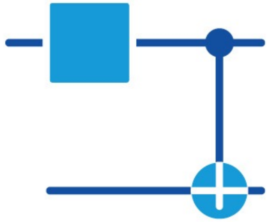
First International Workshop on Quantum Computing and Artificial Intelligence (QC+AI) @ AAAI 2025

Potentials of Quantum AI

- Quantum machine learning (QML) is an emerging framework leveraging quantum processing units (QPUs) for AI tasks
- Potential advantages (hypes) of quantum AI:
  – Quantum computing may accelerate AI systems
  – Quantum computing may reduce power consumption of AI systems
  – Quantum parallelism may improve accuracy with ensemble effect
  – Variational principle may exploit inherent noise to prevent overfitting
  – Exponential expressivity of quantum state may represent large AI model efficiently
    - 1000 variables can be mapped with 10 qubits
  – Structured quantum gates may represent AI model efficiently
    - QNN with few parameters may achieve performance of DNN with massive parameters

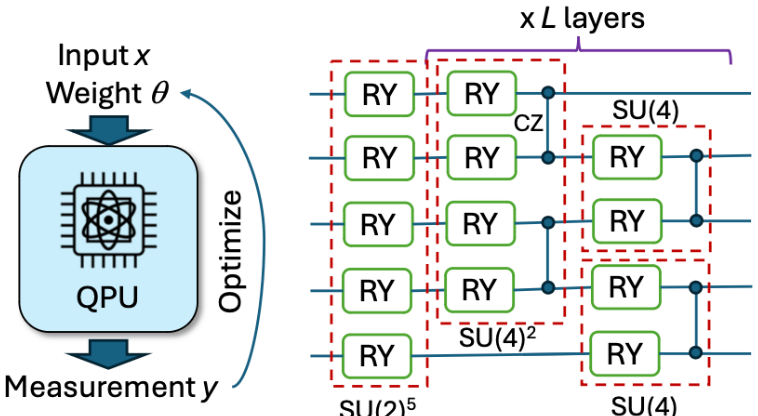




Quantum AI for Parameter-Efficient Models

- We introduced quantum AI for parameter-efficient fine-tuning (PEFT): Quantum-PEFT
  - Quantum-PEFT [Koike-Akino 2024] uses quantum tensor network
  - Presented in ICML-W’24; Accepted to ICLR’25
  - QML realizes ultra-efficient parameterization due to exponential expressivity and structure
  - QML parameterization can be used in conventional CPU/GPU too besides QPU
- We propose to use QML for data efficiency



(a) QNN (b) STD Ansatz

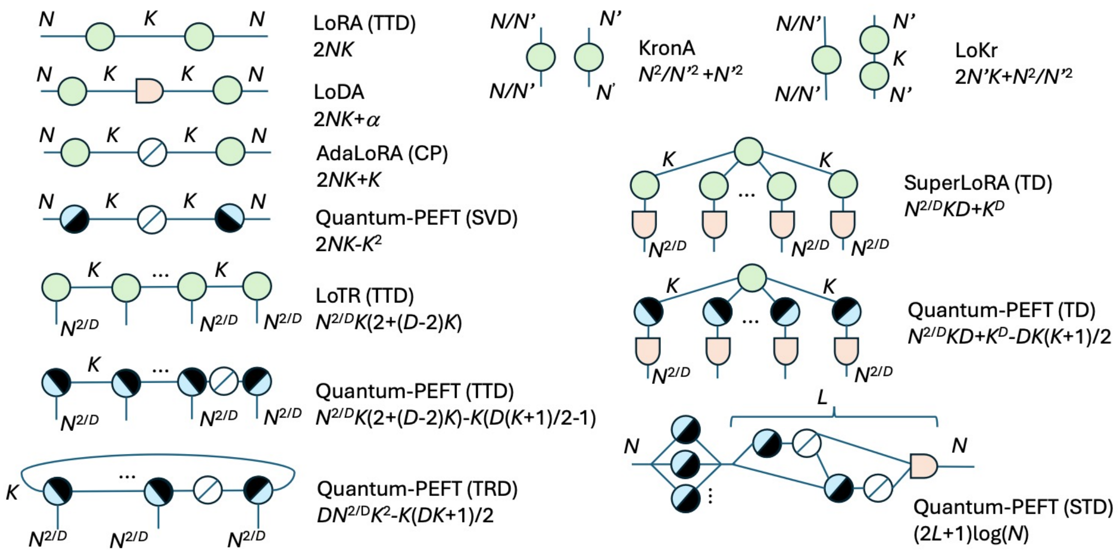
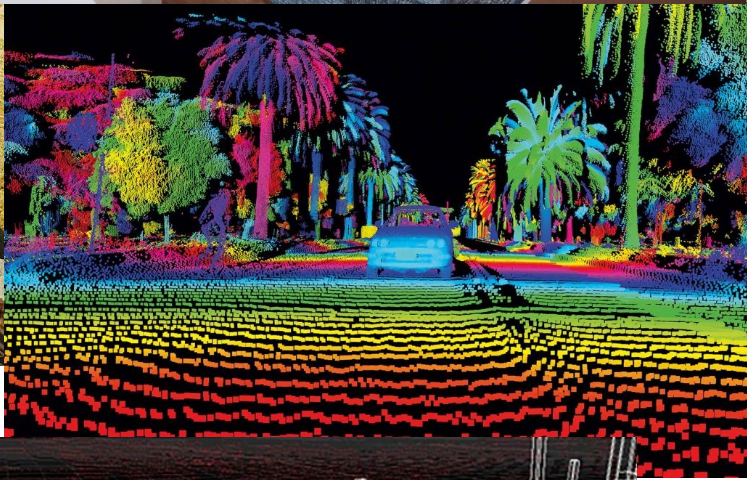
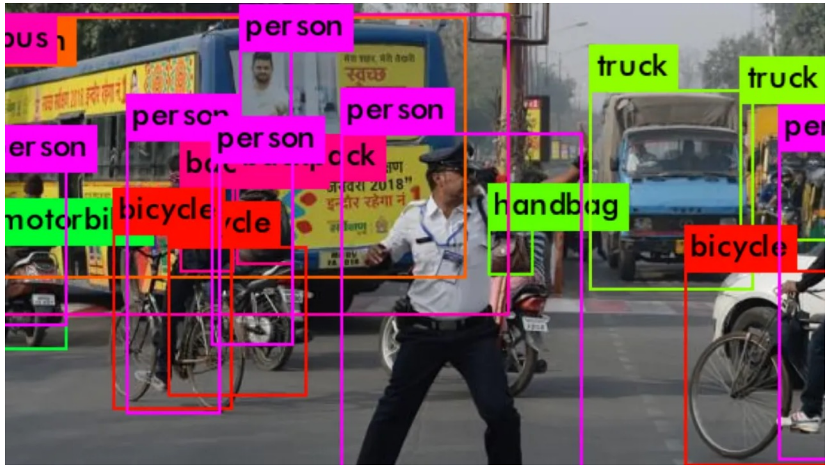


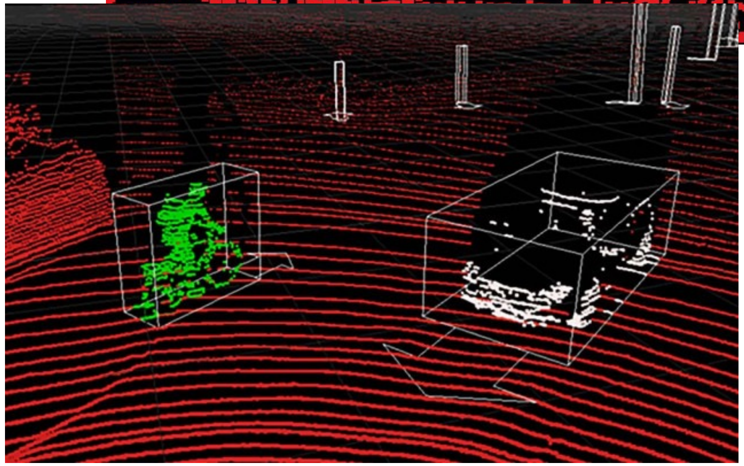
Figure 5: Tensor diagrams of Quantum-PEFT and LoRA variants in tensor network perspectives for a matrix size of N and rank K. The number of parameters are also present. Circle denotes dense multi-linear tensor node. Slashed open circles denote diagonal node. Half-closed circles denote unitary node. Delay symbols denote nonlinear nodes.

Background

- Growing demands for multimedia services
  - High-resolution images, videos, and 3D point clouds
  - Applications
    - Scene rendering for extended reality (XR), augmented reality (AR), virtual reality (VR)
    - Digital twin
    - Digital archive
    - Datasets for 2D/3D/4D scene analytics
      - Detection,
      - Segmentation
      - Tracking



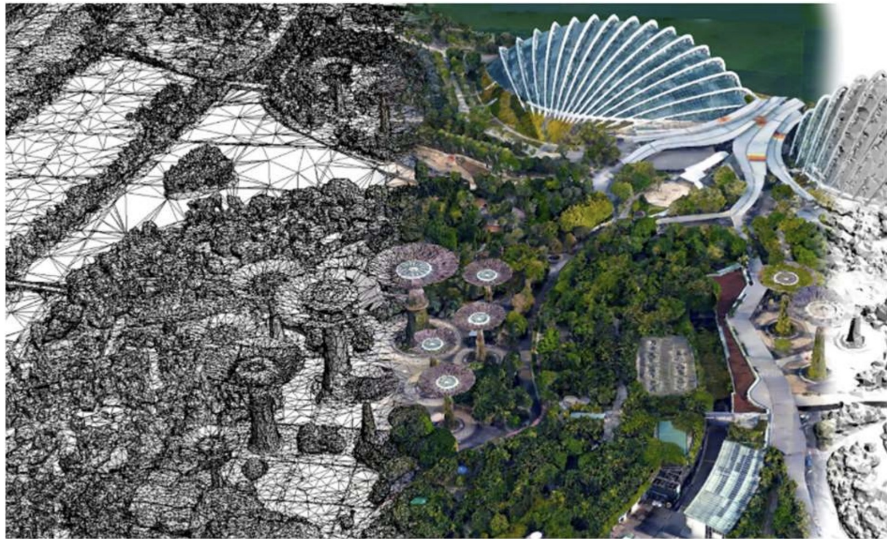




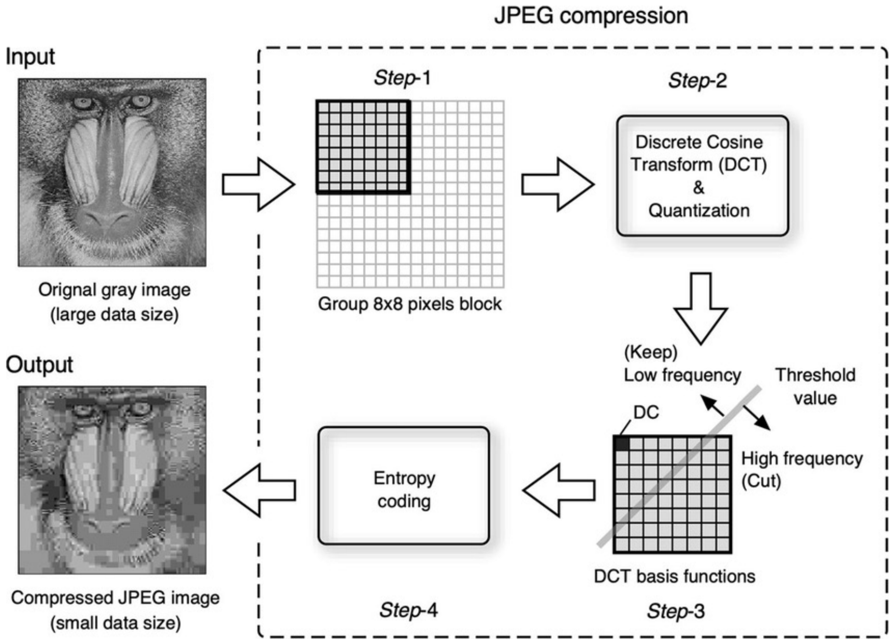
Issue & Typical Solutions

- Issue
  - Large rate for representing high-quality multimedia signals
    - GBs and TBs for representing full 3D scene [1,2]

    → Large storage and transmission costs
- Typical solutions
  - Signal processing-based compression
    - Joint Photographic Experts Group (JPEG)
    - JPEG2000
    - zip
    - …



[1]



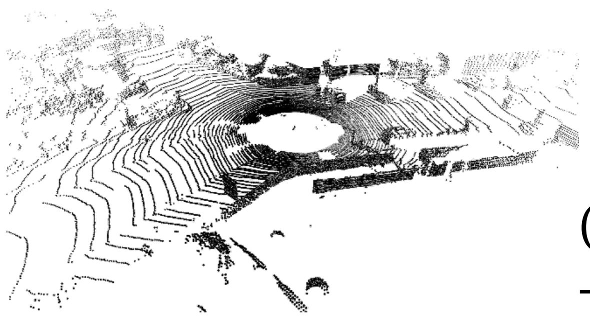
JPEG: Block-wise transform, quantization, and entropy coding

[1] https://venturebeat.com/business/how-singapore-created-the-first-country-scale-digital-twin/
[2] https://ene-fro.com/article/ef339_a1/?utm_source=twitter&utm_medium=display&utm_campaign=enefrox

Extension for LiDAR Point Cloud

- Image coding solutions can be adopted for 3D LiDAR point clouds
  - Obtain 2D range image (RI) from 3D LiDAR point cloud
  - Take image coding solutions for RIs
- RI conversion from 3D LiDAR point cloud
  - LiDAR distance information ρ is assigned to image coordinate (u, v)
  - Maximum value is assigned to unmeasured regions

(1) LiDAR Point Cloud



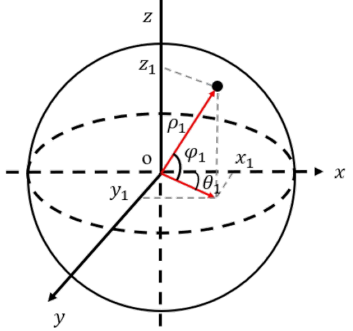
Each point

(x, y, z)

Spherical Coordinate Conversion

$(x, y, z) \rightarrow (\rho, \varphi, \theta)$

(2) Spherical coordinate



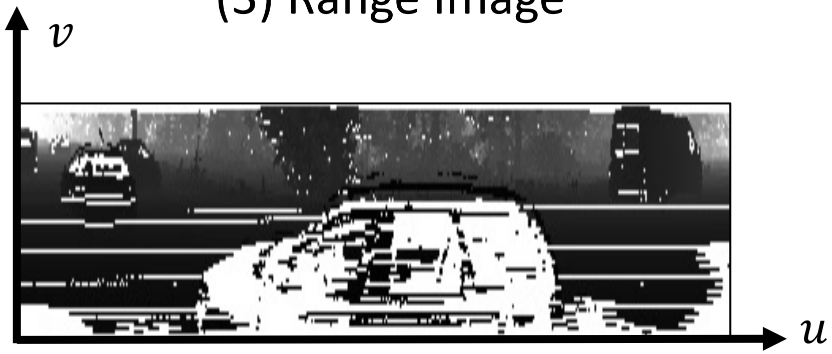
Distance information

$$\rho = \sqrt{x^2 + y^2 + z^2}$$

Image Coordinate Conversion

$(\rho, \varphi, \theta) \rightarrow (\rho, u, v)$

(3) Range Image



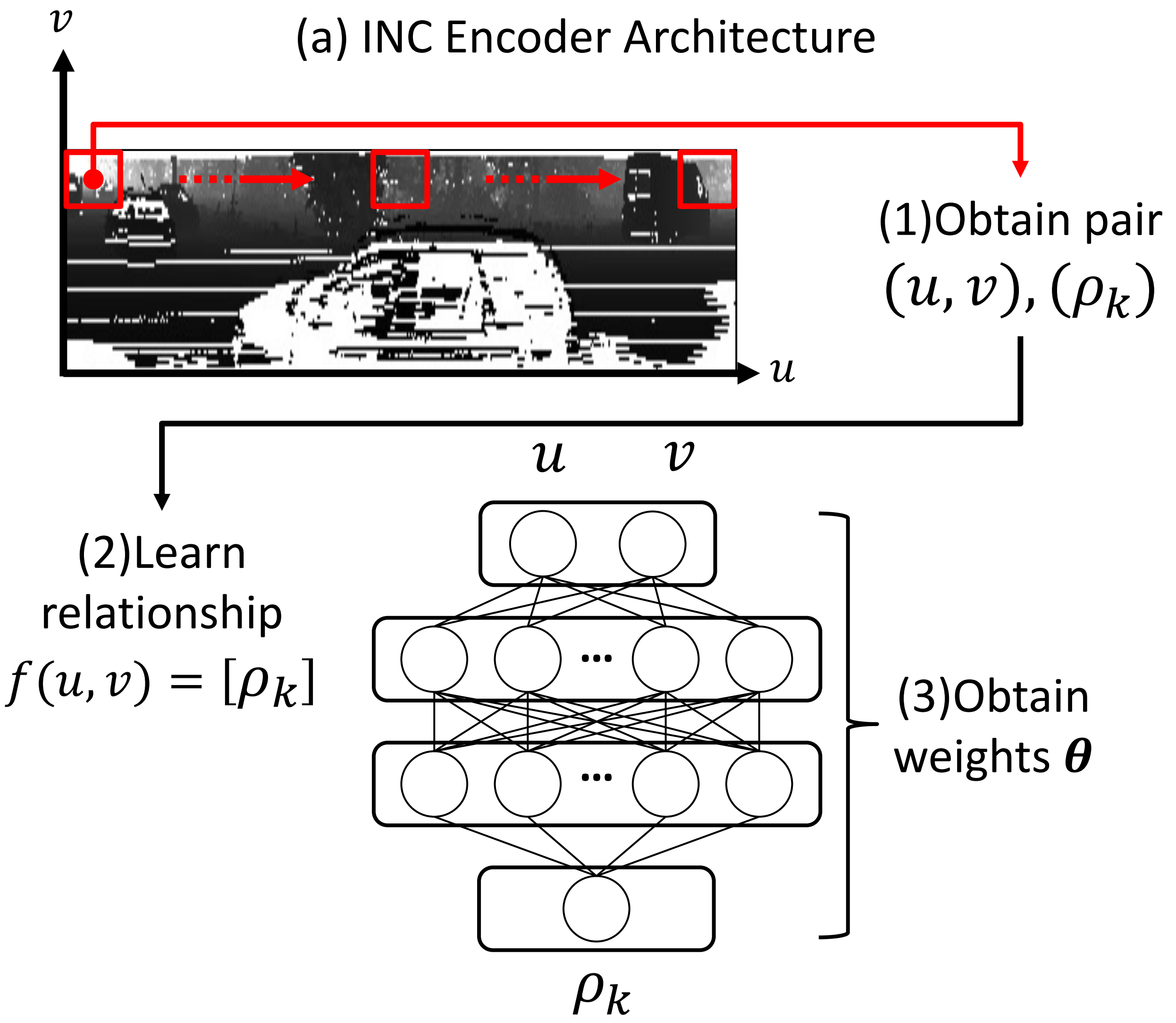
Distance information is assigned to each pixel

$$\rho_k = \begin{cases} \rho & \text{(Measured regions)} \\ 255 & \text{(Unmeasured regions)} \end{cases}$$

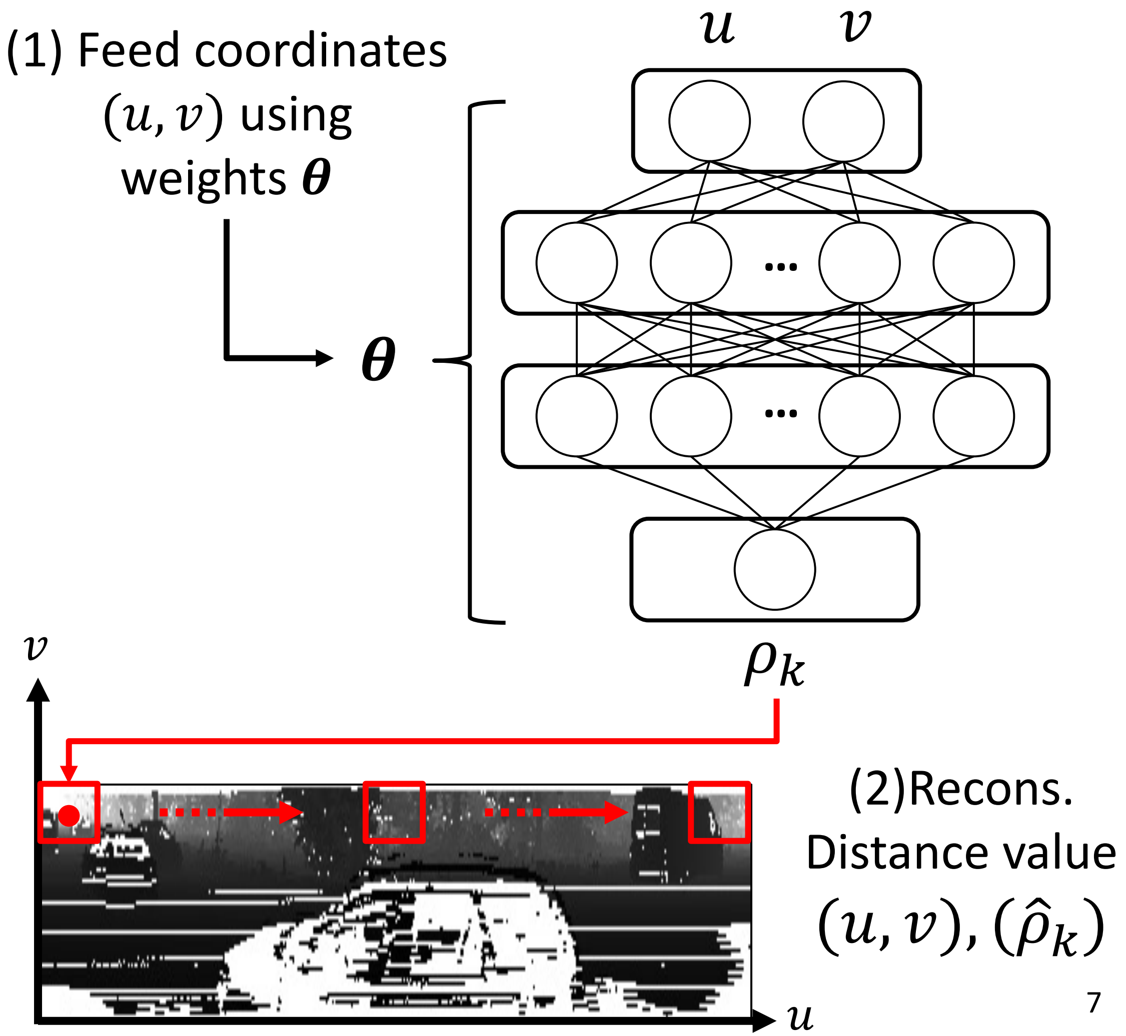
Related Work: Implicit Neural Compression (INC)

- For a single image, approximate the function that represents the relationship between image coordinates and pixel values using implicit neural representation (INR)

(a) INC Encoder Architecture



(b) INC Decoder Architecture

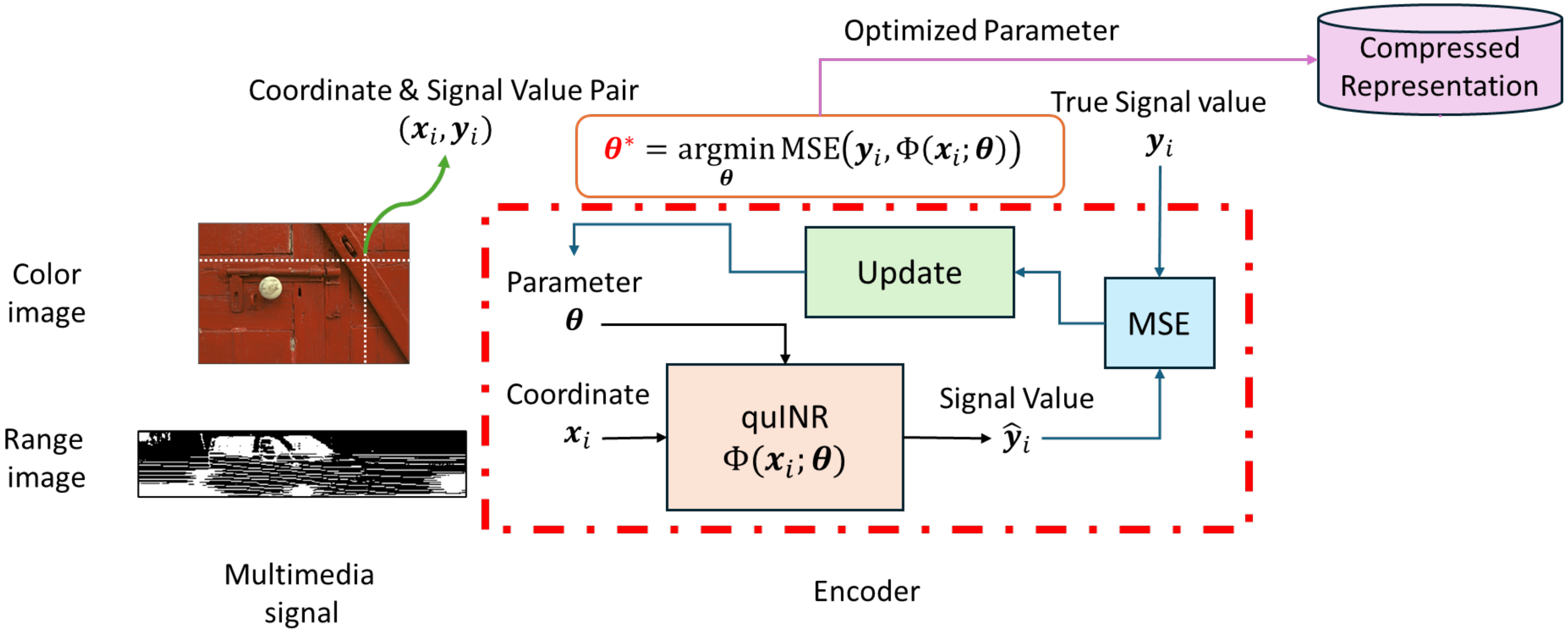


Key Contributions of Our Work

- Purpose
  - Propose quantum-inspired INC, namely, quINR $\Phi(\boldsymbol{x}_i; \boldsymbol{\theta})$, for further compact representation against existing coding and INC solutions

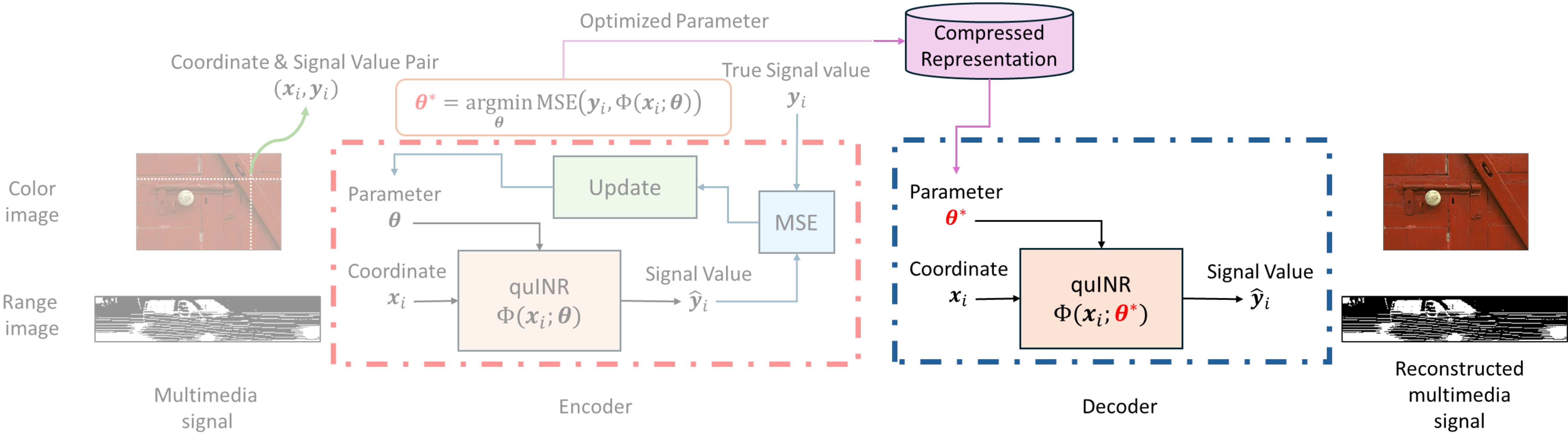
- Key contributions
  - Demonstrate potentials of quantum neural network (QNN) architecture for signal compression
  - Design a hybrid quantum-classical NN architecture
    - Extract a feature vector from classical fully-connected layer
    - Encode an arbitrary size of the feature vector into qubits using embedding and entangling layers

Proposed Quantum Implicit Neural Compression: Encoder



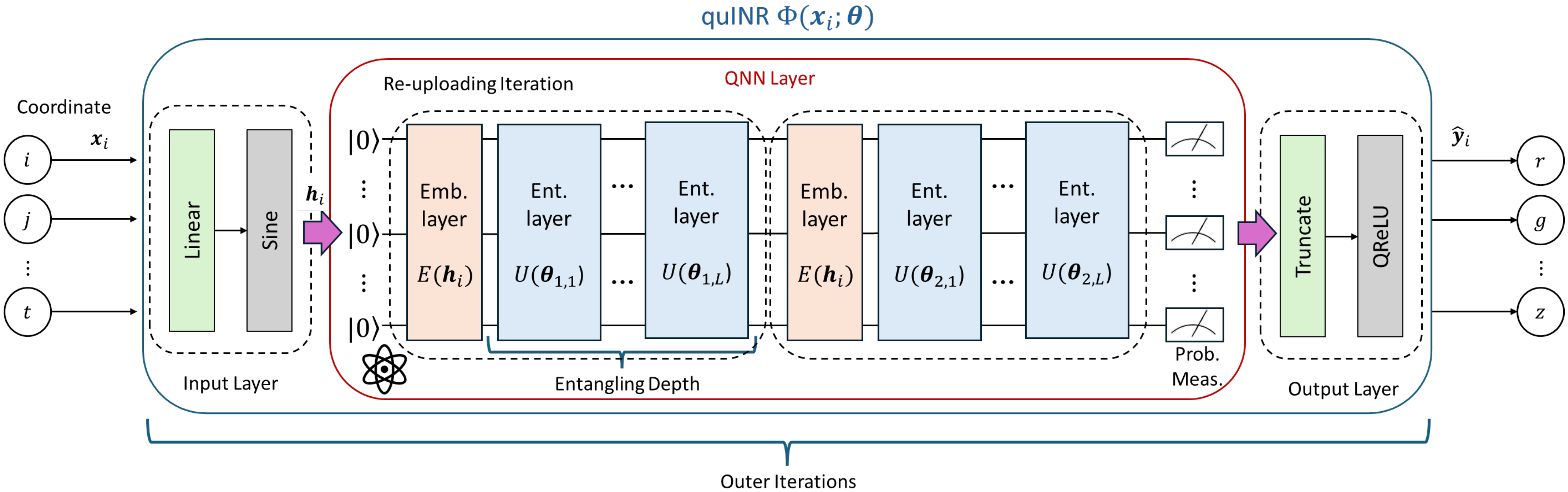
- Encoding
  1. Feed a pair of image coordinates ($\boldsymbol{x}_i$) and pixel value ($\boldsymbol{y}_i$) for training the coordinate-to-value relation
  2. Train QNN minimizing mean squared error (MSE) to obtain the optimized parameters $\boldsymbol{\theta}$
     - Well-trained parameters $\boldsymbol{\theta}^*$ are stored in storage or transmitted to the decoder as compact format

Proposed Quantum Implicit Neural Compression: Decoder



- Decoding
  1. Use $\boldsymbol{\theta}^*$ for reproducing pixel value through forward process $\Phi(\boldsymbol{x}_i; \boldsymbol{\theta}^*)$ to reconstruct pixel value $\widehat{\boldsymbol{y}}_i$
  2. Sequentially feeds all coordinates $\boldsymbol{x}_i$ to quINR $\Phi(\boldsymbol{x}_i; \boldsymbol{\theta}^*)$ to collect all estimated pixel values

Proposed Quantum Implicit Neural Compression: Architecture



- A hybrid quantum-classical NN architecture
  - A linear layer with a sinusoidal activation to obtain an embedding vector $\boldsymbol{h}_i$
  - Embedding vector $\boldsymbol{h}_i$ is fed into QNN layers (embedding and entangling layers)
    - Folded-angle embedding: encode an arbitrary size of vector $\boldsymbol{h}_i$ into a finite number of qubits
    - Entangling: $\boldsymbol{\theta}$ –parameterized quantum circuit, and each parameter controls rotation gates
  - Measure probability values of quantum states and regard them as output values

Experiments

- Datasets
  - RGB color image: Kodak (consist of 24 images)
    - Kodim02: 768×512 pixels
  - LiDAR RI: KITTI
    - Sequence 00-00: 51770 3D points
    - RI resolution: 1024×64 pixels
- Metric
  - Peak Signal-to-Noise Ratio (PSNR)
    - $\mathrm{PSNR} = 10 \log_{10} \left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}} \right)$
- Baseline
  - JPEG2000
    - Typical image codec
  - Compression with Implicit Neural representation (COIN)
    - Pioneer work on INC

Kodak dataset

Kodim02

Kodim03

KITTI dataset

LiDAR point cloud: Seq 00

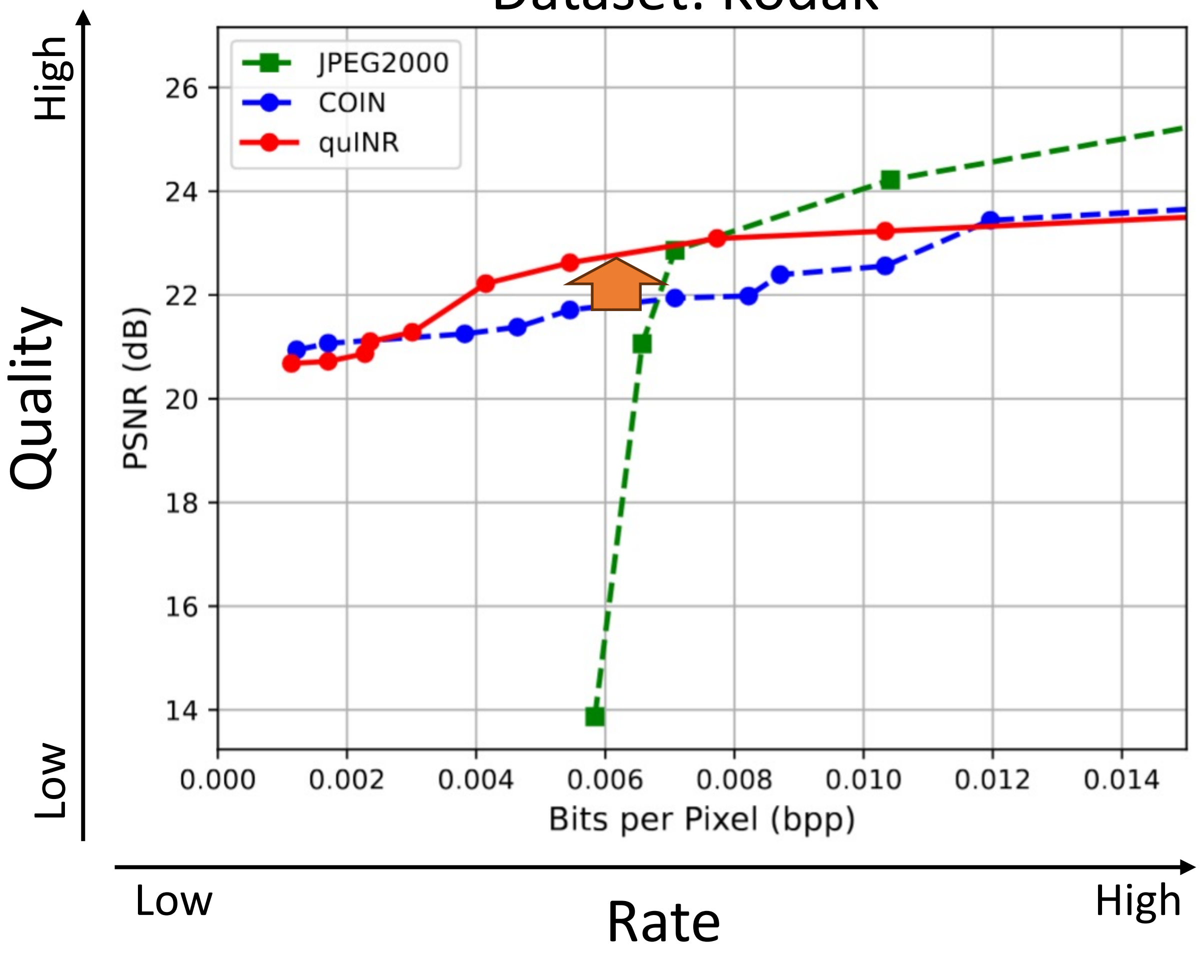
LiDAR RI: Seq 00

Rate Distortion Performance

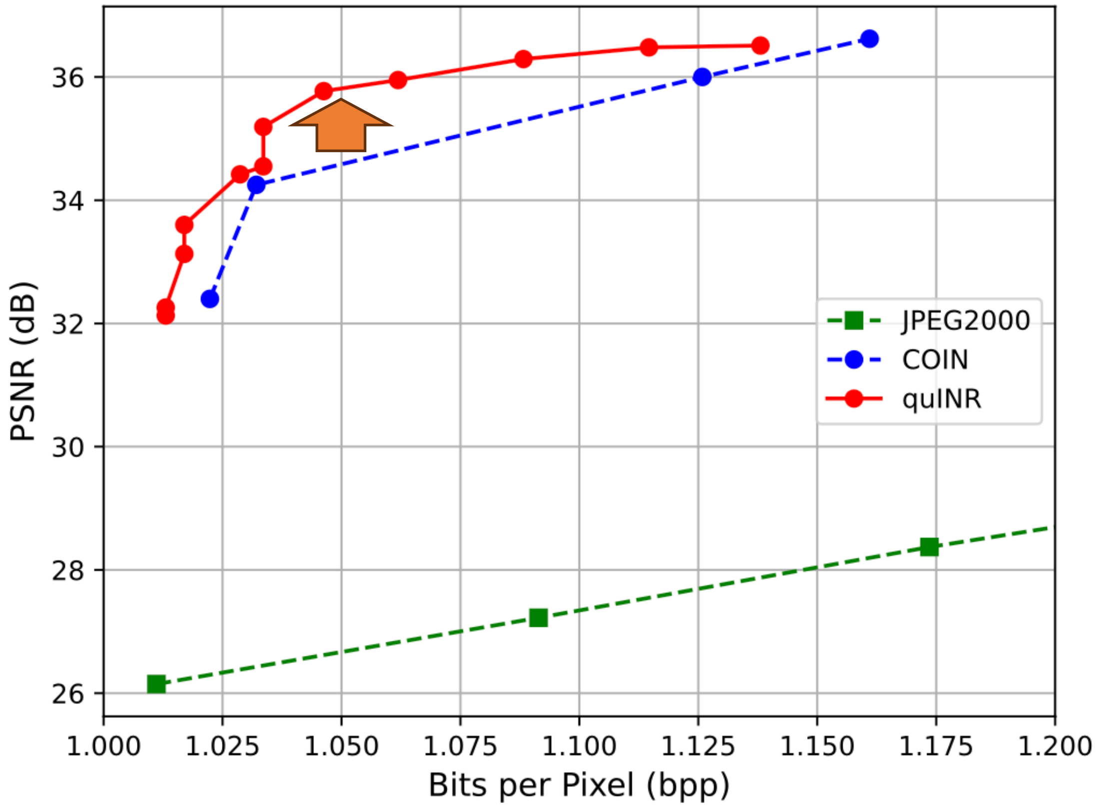
- quINR achieves better image quality at a small bpp regime, up to 1dB gain

→ potential to reconstruct clean signals at band- and storage-limited environments

Dataset: Kodak



Dataset: KITTI



Summary

- Conclusion
  - Propose quantum-inspired INC, **quINR**, for further compact representation against existing coding and INC solutions
    - Demonstrate the potential of QNN architecture for signal compression
    - Design a hybrid quantum-classical NN architecture
  - Evaluations using RGB image and LiDAR RI
    - quINR achieves better image quality at a small bpp regime up to 1dB gain
- Future Work
  - Limited rate-distortion performance in color image compression
    - Quantum network architecture search
    - Distillation for quantum network architecture
- Questions?
  - koike@merl.com