Exploring the Relationship between 2D/3D Convolution for Hyperspectral Image Super-Resolution

Ben HE

Recent problem

Previous hyperspectral image super-resolution (SR) methods lacked joint analysis between spectrum and horizontal or vertical directions, and the combination of 2D and 3D convolutions was not effective enough.

New Method

A novel hyperspectral image SR method is introduced, exploring the relationship between 2D/3D convolution (ERCSR) to address the identified issues.

Technical Innovations

- Alternating use of 2D and 3D units to solve the issue of structural redundancy, enhance the learning capability of the 2D spatial domain through sharing spatial information.
- Compared to networks using only 3D units, this method reduces the model size and improves performance.

Data set

Three public datasets: CAVE, Harvard, and Pavia Centre.

To address the issue of insufficient images in these datasets for deep learning algorithms, the authors augmented the training data.

Address the issue of insufficient images

- **Training Set Augment**: Contains code and methods used to increase the number of images in the training set through augmentation techniques.
- **Test Set Pre-processing**: Includes procedures for preparing the test set for evaluation.
- **Band Mean for All Training Set**: Offers code for calculating the mean values across all bands of the training set images, which is likely used for normalization or other pre-processing steps.

Model

function

- **_to_4d_tensor** and **_to_5d_tensor:** These two functions are used to convert between 4D and 5D tensors.

- They support flexible conversion between different dimensions when processing hyperspectral data

Model

components

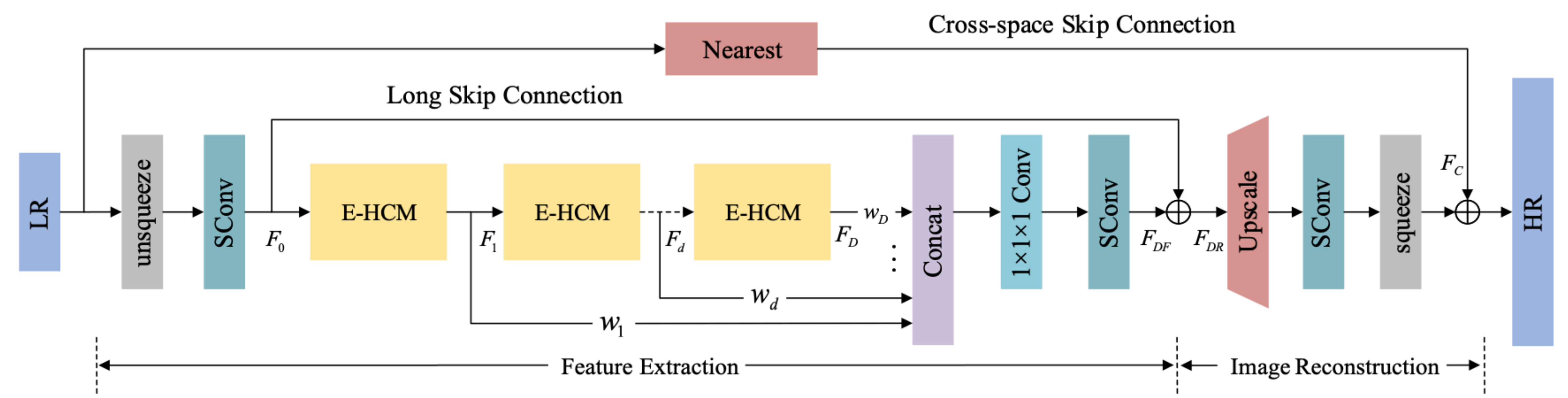
- **twoUint**: This is a small module consisting of two 2D convolutional layers, using ReLU as the activation function.
- **E_HCM**: This is the core module in the model, using 3D convolutional layers to process spatial and spectral information.

Model

Model Architecture:

Head: Employs 3D convolution layers to process the input images, extracting initial spatial and spectral features.

Core Module (E_HCM): The core consists of multiple E_HCM units, each combining 3D and 2D convolutions. 3D convolutions are responsible for extracting features across both spatial and spectral dimensions, while 2D convolutions focus on refining spatial features within each spectral slice.

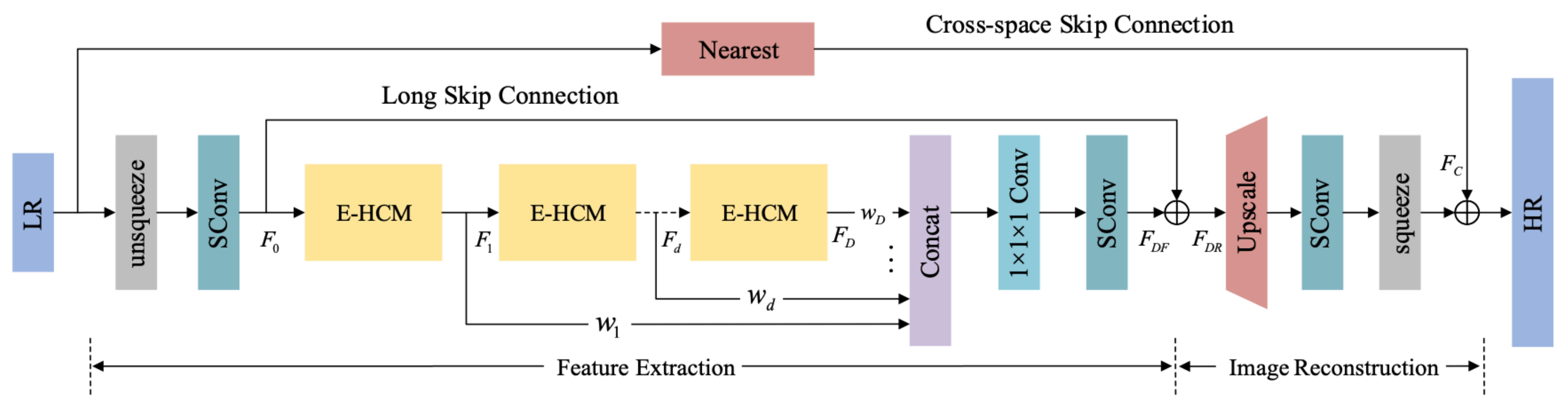


Model

Model Architecture:

Feature Aggregation and Dimensionality Reduction: After sequential processing through the E_HCM units, features from each unit are aggregated and weighted, followed by a dimensionality reduction to further refine the feature representation.

Tail: Contains a series of 3D convolution layers that finalize the feature refinement process, ensuring that the output images have enhanced resolution while retaining accurate spatial and spectral information.



Result

Scale	CAVE	Harvard	Pavia Centre
x2	45.332 / 0.9740 / 2.218	46.372 / 0.9832 / 1.875	35.422 / 0.9498 / 3.435
x3	41.345 / 0.9527 / 2.789	42.783 / 0.9633 / 2.180	31.230 / 0.8690 / 4.650
x4	41.345 / 0.9322 / 3.243	40.211 / 0.9374 / 2.384	28.912 / 0.7786 / 5.534

- three evaluation methods： Signal-to-Noise Ratio (PSNR), Structural SIMilarity (SSIM), and Spectral Angle Mapper (SAM)

Perform the best on the Harvard dataset and the least effectively on the Pavia Centre dataset

Reference

- https://crabwq.github.io/pdf/2020 Exploring the Relationship between 2D3D Convolution for Hyperspectral Image Super-Resolution.pdf
- https://github.com/qianngli/ERCSR/tree/master