Image Super-Resolution:
On its technical detail and subtasks

Seobin Park,

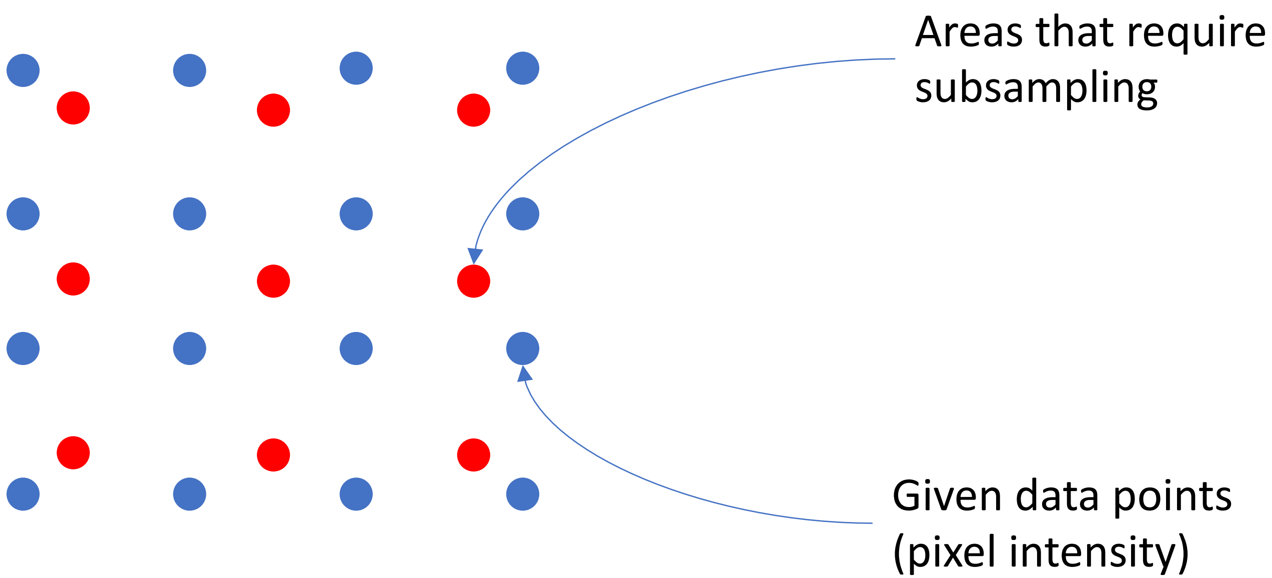
Vision Intelligence Lab, Hanyang University

Content

1. **Preliminaries**
   - Interpolation techniques for image resizing
   - Aliasing
2. **Super-Resolution (SR) settings**
   - What is SR?
   - SR pipeline
3. **Subtasks**
   - Different degradations: Blind SR, RealSR
   - Other: SR on multiple scales, Content-aware SR

Preliminaries: Interpolation techniques for image resizing

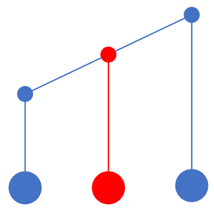
- Changing image resolution (i.e. Number of pixels)



4x4 and 3x3 images

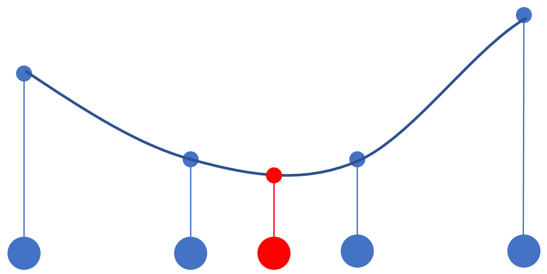
Preliminaries: Interpolation techniques for image resizing

1. Geometric interpolation methods (1d exmample)



Linear interpolation: ax + b

>> Need to solve a 2d linear system



Cubic interpolation: ax^3 + bx^2 + cx + d

>> Need to solve a **4d linear system**

Preliminaries: Interpolation techniques for image resizing

2. Convolution interpolation methods (reformulating interpolation problem)

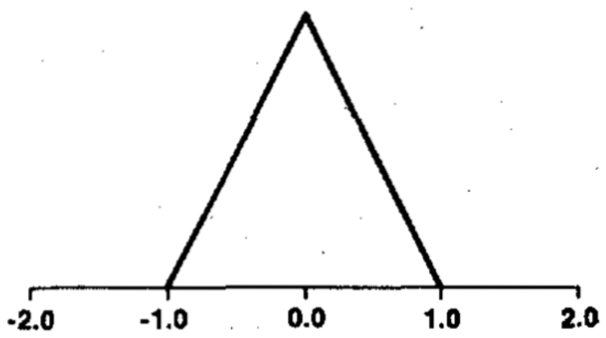
For equally spaced data, many interpolation functions can be written in the form

$$g(x) = \sum_k c_k u\left(\frac{x - x_k}{h}\right). \quad (1)$$

Among the interpolation functions that can be characterized in this manner are cubic splines and linear interpolation functions. (See Hou and Andrews [3].)

In (1), and for the remainder of this paper, h represents the sampling increment, the x_k's are the interpolation nodes, u is the interpolation kernel, and g is the interpolation function. The c_k's are parameters which depend upon the sampled data. They are selected so that the interpolation condition, $g(x_k) = f(x_k)$ for each x_k, is satisfied.

Linear interpolation



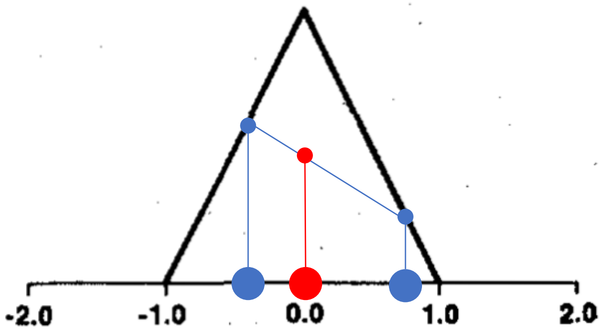
u(s) plot

Keys, Robert. "Cubic convolution interpolation for digital image processing." *IEEE transactions on acoustics, speech, and signal processing* 29.6 (1981): 1153-1160.

Preliminaries: Interpolation techniques for image resizing

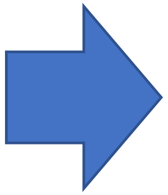
2. Convolution interpolation methods (linear)

Linear kernel u(s)



u(s) plot

$$g(x) = \sum_k c_k u\left(\frac{x - x_k}{h}\right)$$



Same as linear interpolation!

- **We can also reduce convolution interpolation method to cubic interpolation by using cubic kernel, but it still requires solving 4d linear systems.**

Keys, Robert. "Cubic convolution interpolation for digital image processing." *IEEE transactions on acoustics, speech, and signal processing* 29.6 (1981): 1153-1160.

Preliminaries: Interpolation techniques for image resizing

2. Convolution interpolation methods (efficient cubic)

For equally spaced data, many interpolation functions can be written in the form

$$g(x) = \sum_k c_k u\left(\frac{x - x_k}{h}\right). \quad (1)$$

Among the interpolation functions that can be characterized in this manner are cubic splines and linear interpolation functions. (See Hou and Andrews [3].)

In (1), and for the remainder of this paper, h represents the sampling increment, the x_k's are the interpolation nodes, u is the interpolation kernel, and g is the interpolation function. The c_k's are parameters which depend upon the sampled data. They are selected so that the interpolation condition, $g(x_k) = f(x_k)$ for each x_k, is satisfied.

+

Cubic convolution interpolation

1. u(0) = 1, u(n)=0 for non-integer number n
   (for computational efficiency)

2. symmetric, 2pieces of cubic polynomial

3. Differentiable in x=n, concave/convex property

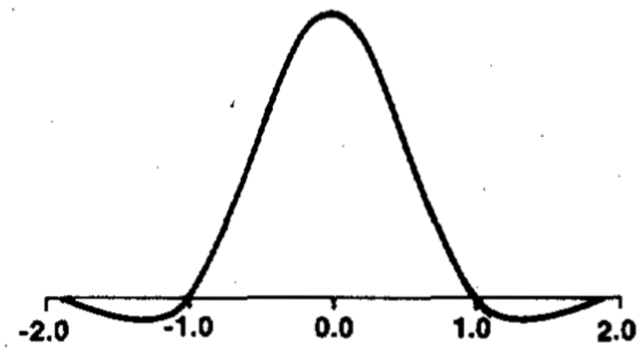
Keys, Robert. "Cubic convolution interpolation for digital image processing." *IEEE transactions on acoustics, speech, and signal processing* 29.6 (1981): 1153-1160.

Preliminaries: Interpolation techniques for image resizing

2. Convolution interpolation methods

Cubic convolution interpolation

$$u(s)=\begin{cases}(a+2)|s|^3-(a+3)|s|^2+1 & 0<|s|<1\\ a|s|^3-5a|s|^2+8a|s|-4a & 1<|s|<2\\ 0 & 2<|s|.\end{cases}$$



u(s) plot

Theoretical range of a is between -3 and 0 [2]
(a = -0.5, -0.75 are used frequently)

Don't need to solve 4d linear systems

[1] Keys, Robert. "Cubic convolution interpolation for digital image processing." *IEEE transactions on acoustics, speech, and signal processing* 29.6 (1981): 1153-1160.
[2] Bernstein, Ralph. "Digital image processing of earth observation sensor data." *IBM Journal of research and development* 20.1 (1976): 40-57.

Preliminaries: Interpolation techniques for image resizing

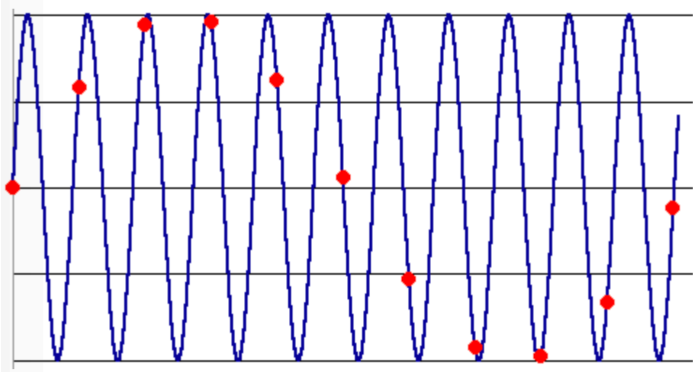
3. Interpolation performance
    i. Time complexity: **linear interp < cubic conv < cubic interp**
    ii. Image results: **linear interp < cubic conv < cubic interp**

4. Popular implementations (library)
    i. Malab imresize: cubic conv (a = -0.5)
    ii. Python PIL resize: cubic conv (a = -0.75)
    iii. OpenCV resize: cubic conv (a = -0.75)
    iv. scipy.misc.imresize => This is a wrapper of PIL
    v. Scipy.interpolate.interp1d/2d : cubic/bicubic interpolation
    (don ' t use this for images)

https://stackoverflow.com/questions/26823140/imresize-trying-to-understand-the-bicubic-interpolation

Preliminaries: Aliasing

- When downsizing (downsampling) an image, some high-frequency components cannot be fit into the image grid.



Moiree Effect

https://matthews.sites.wfu.edu/misc/DigPhotog/alias/

Preliminaries: Aliasing

- We need to apply a low-pass filter (anti-aliasing filter) before downsizing an image!
- Matlab imresize simply broadens the cubic kernel to accomplish anti-aliasing

Without prefiltering

With prefiltering (no aliasing, but blurring!)

https://stackoverflow.com/questions/26823140/imresize-trying-to-understand-the-bicubic-interpolation
https://kr.mathworks.com/matlabcentral/answers/232462-imresize-with-bicubic-downsampling
https://eeweb.engineering.nyu.edu/~yao/EL5123/lecture8_sampling.pdf

SR settings: What is SR?

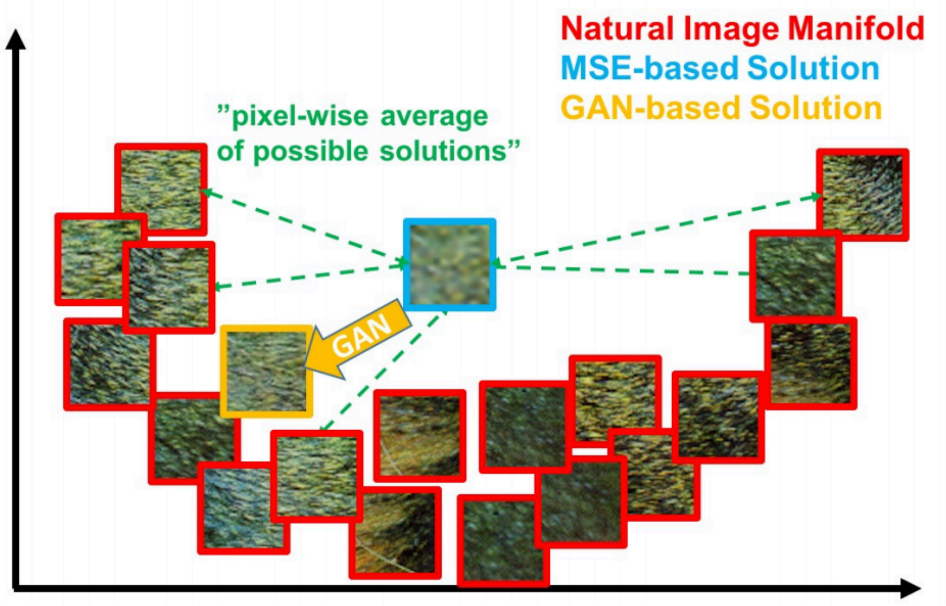
- Increasing the resolution (number of pixels) of an image with **plausible high-frequency details**.

bicubic (21.59dB/0.6423) | SRGAN (21.15dB/0.6868) | original

Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

SR settings: What is SR?

- Increasing the resolution (number of pixels) of an image with plausible high-frequency details.
- SR is inherently an **ill-posed problem** because there are many plausible answer patches (high-resolution [HR]) to one low-resolution (LR) input.



Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

SR settings: SR pipeline

SR settings: SR pipeline

- Make (HR, LR) pairs for training set.
- SR tasks typically assume a degradation setting from HR to LR.
- Simplest setting: Matlab imresize (bicubic conv) with scale s.
  - ➢ LR = imresize(HR, s)
- See **NTIRE challenges** for more information.

Agustsson, Eirikur, and Radu Timofte. "Ntire 2017 challenge on single image super-resolution: Dataset and study." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2017.

SR settings: SR pipeline

- An SR method that maps LR $\rightarrow$ HR.
- Random crop, Flip, rotation used to augment training set.
- Same augmentation can be used to boost SR performance on test phase [1].

[1] Zhang, Yulun, et al. "Residual dense network for image super-resolution." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.

SR settings: SR pipeline

- Compare HR and SR result (estimated HR) to assess SR quality.
  - ➢ **Usually removes (shave) the border of output image.**
- PSNR, SSIM [1] is commonly used.
- Metrics based on human perception: LPIPS [2]
- No-reference metrics: NIQE, BRISQUE, PIQUE (usually used in generative model like SRFlow [3])

[1] Wang, Zhou, et al. "Image quality assessment: from error visibility to structural similarity." *IEEE transactions on image processing* 13.4 (2004): 600-612.
[2] Zhang, Richard, et al. "The unreasonable effectiveness of deep features as a perceptual metric." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.
[3] Lugmayr, Andreas, et al. "Srflow: Learning the super-resolution space with normalizing flow." *European Conference on Computer Vision*. Springer, Cham, 2020.

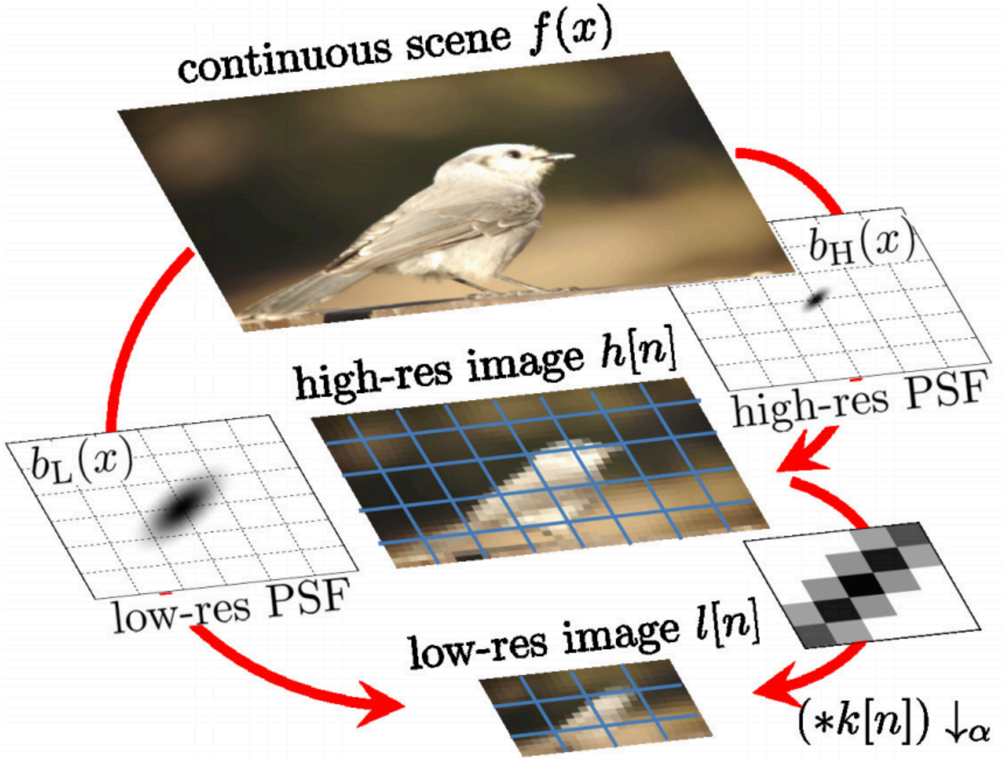
Subtasks: Different degradations

- Deep neural network-based SR methods are **very sensitive to degradation** (e.g. downsampling) method used in dataset.
  - ➢ Simple bicubic degradation setting doesn＇t make SR networks able to handle SR in general natural images.

- So various degradation settings have been proposed to resemble real-world settings.

Subtasks: Different degradations

- Blind SR: there is a **blur kernel** that we don ' t know.



$$I^{LR} = (\underline{k} \otimes I^{HR}) \downarrow_s + n$$

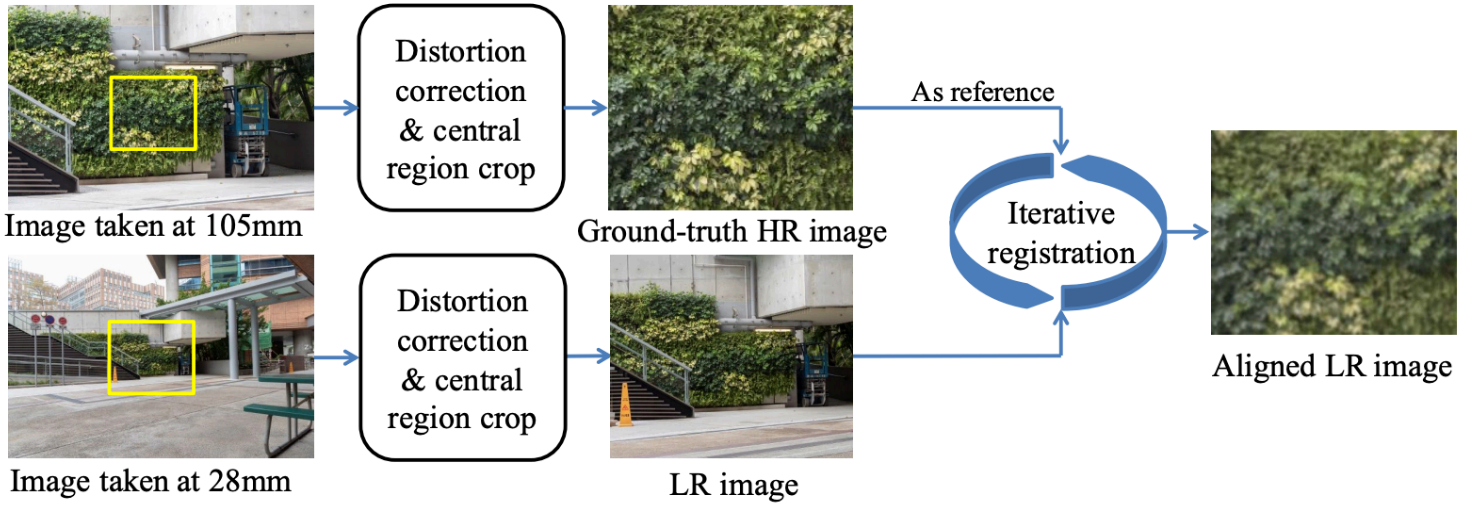
Unknown kernel

Michaeli, Tomer, and Michal Irani. "Nonparametric blind super-resolution." *Proceedings of the IEEE International Conference on Computer Vision*. 2013.
Gu, Jinjin, et al. "Blind super-resolution with iterative kernel correction." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.

Subtasks: Different degradations

- RealSR: Real camera setting => **Spatially varying kernel**

$$\mathbf{I}^L(i,j) = \mathbf{I}^H_{i,j} \odot \mathbf{k}_{i,j} + \mathbf{v}(i,j)$$



Cai, Jianrui, et al. "Toward real-world single image super-resolution: A new benchmark and a new model." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019.
Xiao, Jin, Hongwei Yong, and Lei Zhang. "Degradation Model Learning for Real-World Single Image Super-resolution." *Proceedings of the Asian Conference on Computer Vision*. 2020.

Subtasks: Other

1. SR on multiple scales:
   - ➢ Conventional SR trains on one specific scale factor. In real-world scenarios, we may have to handle multiple (or arbitrary) scale with one parameter set (Meta-SR [1])
2. Content-aware SR:
   - ➢ SR utilizing high-level (semantic) information (TextZoom [2], ProgFSR [3])

[1] Hu, Xuecai, et al. "Meta-SR: A magnification-arbitrary network for super-resolution." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2019.
[2] Wang, Wenjia, et al. "Scene text image super-resolution in the wild." *European Conference on Computer Vision*. Springer, Cham, 2020.
[3] Deokyun Kim, et al. " Progressive Face Super-Resolution via Attention to Facial Landmark " Proceedings of the 30th British Machine Vision Conference (BMVC), 2019.