Image and video processing: From Mars to Hollywood with a stop at the hospital.

by Guillermo Sapiro

Week #7 *Skanda S Bharadwaj*

1. Q: What is the dictionary used in JPEG?

    - Wavelets.
    - Fourier.
    - Discrete cosine transform (DCT).
    - Learned from data.

    A: We have discussed that JPEG uses DCT, and this is a particular example of sparse modeling. We can then consider the cosine basis as a dictionary.

2. Q: Consider a 2×4 dictionary D composed of the transpose of the 2-dimensional atoms (0,1), (1,1), (0,1), and (2,1) (these form the columns of D). The sparsest representation of the vector x=(2,2) is given by the transpose of (these are the α):

    - (0,1,0,0)
    - (2,0,2,0)
    - (0,2,0,0)
    - (0,0,0,1)

    A: Clearly the vector (2,2) can be obtained as 2×(1,1) and therefore (0,2,0,0) is a sparse representation of (2,2) with the given D. While (2,0,2,0) is also a representation of (2,2) (since (2,2)=2×(1,0)+2×(0,1)), it is less sparse.

3. Q: We want to obtain sparse representations of signals of dimension N=64. We have a dictionary with k=100 atoms. How many possible active sets (subspaces) we have with sparsity L=3?

    - 100!97!3!
    - 100!97!
    - 100!
    - $100 \cdot 64$

    A: We have L=3 to chose from k=100, the answer therefore is k!(k−L)!L!. Note that the signal dimension N is not important here.

4. Q: Consider the Gaussian Mixture Model in the last video. We want to use it to represent signals in N=64 dimensions. If we have k=100 Gaussians in the mixture, then the number of possible active sets (subspaces) is

    - 100
    - 100!
    - 100×64
    - 100!36!64!

    A: We have only to select the Gaussian, and therefore we have only 100 possibilities. Once the Gaussian is selected, the whole block of corresponding atoms is used.

5. Q: Are sparse modeling and compressed sensing the same?

    No, sparse modeling is about signal models and representations; compressed sensing is about an efficient novel data acquisition protocol.

    Yes, they are both based on sparse representations and therefore are the same.

    A: As discussed in the video lecture, while both sparse modeling and compressed sensing often deal with sparse signals and representations, they are not the same. In sparse modeling we are concerned with signal representations (once the signal has already been acquired), while compressed sensing is concerned with novel data acquisition protocols exploiting sparse signal representations.

6. Q: What needs to change in the general expression of image denoising we used for sparse modeling (equation in slide 4 of the 1st video this week) if instead of Gaussian additive noise we consider other types of additive noise?

    - We need to change the data fitting term, relationship with measurements, from a quadratic penalty to a penalty tailored to the noise.
    - We need to define a new prior G.
    - We can't use this type of formulation for non-Gaussian noise, even if we modify its basic components.
    - We need to work only with small amounts of non-Gaussian noise and linearize the problem.

    A: The quadratic penalty term applies to Gaussian noise, and if we change the noise, we just need to adjust this. For example, for exponential noise, we will simply use the absolute value |x-y|. Note that G represents the image model and not the noise.

7. Q: Consider a dictionary D composed of both the complete DCT basis and the complete Fourier basis, a concatenation of both. Will the representation of a signal be unique when using such dictionary?

    - No, there will be at least two different possible representations for all signals.
    - Yes, since both the DCT basis and the Fourier basis uniquely represent a signal.
    - Not all signals can be represented by such dictionary.
    - Yes, a signal will pick some atoms from the DCT and some from the Fourier.

    A: The DCT alone can represent the signal, so can Fourier. If we have both together as part of the dictionary, then the signal can pick either one, and then we get at least 2 different representations. One might be sparser than the other, but both are full representations. Some signals might have additional representations composed by mixing atoms from the DCT and the Fourier components.

8. Q: Consider you have a dictionary composed of 100 random 10×10 patches from the given image. If you perform sparse coding with this dictionary:

    - The average number of non-zero coefficients will be equal or greater than when using the dictionary of the same size for sparse representations, obtained with
      $\min_{D,\alpha} \|\alpha\|_0 \text{ s.t.} \sum_i \|D\alpha - y_i\|_2^2 \leq \varepsilon$,
      where y_i is an image patch.
    - We will never be able to obtain a sparse representation from image patches as dictionary.
    - The code will always be sparser in average than the one obtained with a DCT dictionary of the same size.

    A: The optimum has always to be better than any other solutions.

9. Q: Consider a video and use the patches of the current frame as dictionary for encoding the next frame. For scenes with only static objects,

    - This will result in very sparse codes on average.
    - This is not as good as what MPEG does.
    - This will not lead to sparse codes at all.

    A: Since there are only static objects, at most we have camera motion, and using previous frames as dictionary is exactly what MPEG does in a very successful fashion, with only camera motion is expected to often find very close patches in previous frames and therefore high levels of sparsity.