Question 4, Assignment 5: CS 754, Spring 2024-25

Amitesh Shekhar
IIT Bombay
22b0014@iitb.ac.in

Anupam Rawat
IIT Bombay
22b3982@iitb.ac.in

Toshan Achintya Golla
IIT Bombay
22b2234@iitb.ac.in

April 15, 2025

1. Consider that you learned a dictionary $\boldsymbol{D}$ to sparsely represent a certain class $\mathcal{S}$ of images - say handwritten alphabet or digit images. How will you convert $\boldsymbol{D}$ to another dictionary which will sparsely represent the following classes of images? Note that you are not allowed to learn the dictionary all over again, as it is time-consuming.

   (a) Class $\mathcal{S}_1$ which consists of images obtained by applying a known affine transform $\boldsymbol{A_1}$ to a subset of the images in class $\mathcal{S}$, and by applying another known affine transform $\boldsymbol{A_2}$ to the other subset. Assume that the images in $\mathcal{S}$ consisted of a foreground against a constant 0-valued background, and that the affine transformations $\boldsymbol{A_1}, \boldsymbol{A_2}$ do not cause the foreground to go outside the image canvas.

   (b) Class $\mathcal{S}_2$ which consists of images obtained by applying an intensity transformation $I^i_{new}(x,y) = \alpha(I^i_{old}(x,y))^2 + \beta(I^i_{old}(x,y)) + \gamma$ to the images in $\mathcal{S}$, where α, β, γ are known.

   (c) Class $\mathcal{S}_4$ which consists of images obtained by downsampling the images in $\mathcal{S}$ by a factor of k in both X and Y directions.

   (d) Class $\mathcal{S}_5$ which consists of images obtained by applying a blur kernel which is known to be a linear combination of blur kernels belonging to a known set $\mathcal{B}$, to the images in $\mathcal{S}$.

   (e) Class $\mathcal{S}_6$ which consists of 1D signals obtained by applying a Radon transform in a known angle θ to the images in $\mathcal{S}$. [4 x 5 = 20 points]

(a) Let $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ be the known affine transformation operators. Since affine transforms are linear, they can be applied to each atom in the dictionary. Let D_{A_1} and D_{A_2} be the transformed dictionaries obtained by applying $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ respectively to every reshaped atom of $\boldsymbol{D}$. Then,

$$D' = [D_{A_1} \mid D_{A_2}]$$

This new dictionary D' will be able to sparsely represent both subsets of images, as follows:

$$f' = D' \begin{bmatrix} \theta_f \\ 0 \end{bmatrix} \quad \text{if } f' \text{ came from } A_1(f), \quad \text{and} \quad f' = D' \begin{bmatrix} 0 \\ \theta_f \end{bmatrix} \quad \text{if } f' \text{ came from } A_2(f)$$

where $f = D\theta_f$ and θ_f is k-sparse.

(b) For an image $f = D\theta$, consider the intensity transformation:

$$f' = \alpha(f \odot f) + \beta f + \gamma \mathbf{1}$$

where $\odot$ denotes point-wise multiplication and $\mathbf{1}$ is a vector of all ones. Expanding:

$$f' = \alpha(D\theta \odot D\theta) + \beta D\theta + \gamma \mathbf{1}$$

Let $\widetilde{D}$ be a matrix formed by element-wise products of all pairs of dictionary atoms:

$$\widetilde{D}_{(i,j)} = D_i \odot D_j \quad \forall i, j$$

and $\widetilde{\theta} = \text{vec}(\theta\theta^T)$, then:

$$f' = [\mathbf{1} \mid D \mid \widetilde{D}] \begin{bmatrix} \gamma \\ \beta\theta \\ \alpha\widetilde{\theta} \end{bmatrix}$$

So, the new dictionary is:

$$D' = [\mathbf{1} \mid D \mid \widetilde{D}]$$

which contains the constant vector, the original dictionary atoms, and all pairwise element-wise products of atoms.

(c) Downsampling by a factor k is a linear operation that reduces spatial resolution. Let S_k be the downsampling operator that maps an image of size (h, w) to $(h/k, w/k)$. Then, applying this to each atom in D (after reshaping) gives:

$$D_i' = \text{vec}(S_k(\text{reshape}(D_i))) \quad \forall i$$

So the new dictionary is:

$$D' = [D_1' \; D_2' \; \dots \; D_K']$$

where each D_i' is a downsampled version of D_i.

(d) Suppose the blur kernel b is given by $b = \sum_{j=1}^{n} \beta_j b_j$ for known $\{b_j\} \in \mathcal{B}$. Since convolution is linear:

$$f' = b * \text{reshape}(f) = \left(\sum_j \beta_j b_j\right) * \left(\sum_i D_i \theta_i\right) = \sum_{i,j} \beta_j (b_j * \text{reshape}(D_i))\theta_i$$

Define $D^{(j)}$ as the dictionary obtained by convolving each atom D_i (reshaped) with b_j:

$$D_i^{(j)} = \text{vec}(b_j * \text{reshape}(D_i)) \quad \Rightarrow \quad D' = \sum_j \beta_j D^{(j)}$$

So D' is the linear combination of the blurred versions of the original dictionary atoms.

(e) Let R_θ denote the Radon transform at angle θ. Since it's a linear operation:

$$f' = R_\theta(\text{reshape}(f)) = R_\theta\left(\sum_i \text{reshape}(D_i)\theta_i\right) = \sum_i R_\theta(\text{reshape}(D_i))\theta_i$$

Let $D_i' = R_\theta(\text{reshape}(D_i))$, vectorised. Then:

$$D' = [D_1' \; D_2' \; \dots \; D_K']$$

is the new dictionary, where each atom is transformed by the Radon operator.