Image Super-resolution from Low Rank and Nonlocal Centralized Sparsity Assumptions (Super-resolucion de imágenes a partir de asunciones de escases centralizada no local y bajo rango)

Kevin Arias

Mayo 2019

1 Contexto

En el estado de arte técnicas de restauración de imágenes tales como eliminación de ruido, fusión de imágenes, super-resolucion han sido de gran utilidad en un gran número de aplicaciones en donde se requiere el análisis de imágenes. El interés en este conjunto de técnicas es obtener una imagen de mejor calidad para lograr una mayor utilidad y precisión sobre distintas aplicaciones. En este documento abordaremos específicamente la técnica de super-resolucion de imágenes. En super-resolucion, modelos basados en una representación escasa, la cual codifica un parche de imagen como una combinación lineal de átomos de diccionario ya han sido estudiados. Adicionalmente, regularizaciones de bajo rango ya han sido empleadas para solucionar problemas mal condicionados similares al problema de super-resolucion de imágenes.

1.1 Super-resolucion de imágenes

Reconstruir una imagen de alta calidad a partir de una imagen degradada (borrosa y submuestreada) ha tenido aplicaciones importantes, tales como en imágenes médicas, sensado remoto, vigilancia y entretenimiento, etc. Matemáticamente, para una imagen observada $\mathbf{y}$, the problema de super-resolucion puede ser generalmente formulado como

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \tag{1}$$

donde $\mathbf{y} \in \mathbb{R}^n$ es el vector de imagen degradada, $\mathbf{H} \in \mathbb{R}^{n\times N}$ es la matriz de degradación, $\mathbf{x} \in \mathbb{R}^N$ es el vector de imagen original y $\mathbf{w} \in \mathbb{R}^n$ es el vector de ruido. Diferentes enfoques de optimizacion han sido propuestos para solucionar el problema de obtener una aproximación lo más cercana a $\mathbf{x}$ a partir de la imagen degradada (observada) $\mathbf{y}$. El problema tradicional de super-resolucion basado en asunciones de escases es formulado de la siguiente manera

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\ \frac{1}{2}\|\mathbf{y} - \mathbf{PB}\left(\boldsymbol{\Psi}\boldsymbol{\theta}\right)\|_2^2 + \lambda\|\boldsymbol{\theta}\|_1, \tag{2}$$

donde $\mathbf{P} \in \mathbb{R}^{n \times N}$ es la matriz de sub-muestreo, $\mathbf{B} \in \mathbb{R}^{N \times N}$ es la matriz de borrosidad actuando como un operador de convolucion cíclica, tal que $\mathbf{H} = \mathbf{PB}$ y λ es un parámetro de regularización.

Definicion 1. REPRESENTACION ESCASA DE UNA SEÑAL
Sea una señal $\mathbf{x} \in \mathbb{R}^N$ y una base de representación escasa $\mathbf{\Psi}$ tal que $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\theta}$ donde $\|\boldsymbol{\theta}\| < K$. La expresión anterior establece que una señal (imagen) $\mathbf{x}$ puede ser representada de manera concisa en una base de representación $\mathbf{\Psi}$ mediante el vector escaso $\boldsymbol{\theta}$ donde la escases establece que el número de elementos diferentes de cero en este vector debe ser menor a K. Las dimensiones de $\mathbf{\Psi}$ y $\boldsymbol{\theta}$ se definirán a continuación bajo el modelo de escases basado en aprendizaje de diccionario a nivel de parche de imagen.

1.2 Modelo de escases basado en el aprendizaje de un diccionario a nivel de parche

Un tipo de base de representación $\mathbf{\Psi} = \mathbf{D} \in \mathbb{R}^{np^2 \times Np}$, $np \leq Np$, ha sido utilizado recientemente en diferentes técnicas de restauración de imágenes para promover una mayor escases sobre un conjunto limitado de señales. Este diccionario representa de forma escasa parches superpuestos de la imagen en vez de la imagen completa. Matemáticamente, sea $\mathbf{x}_i = \mathbf{R}_i\mathbf{x}$ un parche de imagen de tamaño $np \times np$ extraido de la posicion i donde $\mathbf{R}_i \in \mathbb{R}^{np^2 \times N}$ es la matriz de extracción, así cada parche puede ser representado escasamente como $\mathbf{x}_i \approx \mathbf{D}\boldsymbol{\theta}_i$ solucionando el siguiente problema de minimizacion $\operatorname{argmin}_{\boldsymbol{\theta}_i} \|\mathbf{x}_i - \mathbf{D}\boldsymbol{\theta}_i\|_2^2 + \lambda\|\boldsymbol{\theta}_i\|_1$. Luego, la reconstrucción de la imagen $\mathbf{x}$ a partir de todos los vectores escasos $\boldsymbol{\theta}_i$ puede ser formulada como

$$\mathbf{x} \approx \mathbf{D} \circ \mathbf{\Theta}_c = \left(\sum_i \mathbf{R}_i^T \mathbf{R}_i\right)^{-1} \sum_i \left(\mathbf{R}_i^T \mathbf{D}\boldsymbol{\theta}_i\right) \tag{3}$$

donde $\mathbf{\Theta}_c$ denota la concatenación de todos los vectores $\boldsymbol{\theta}_i$. En el escenario de super-resolucion basado bajo asunciones de escases la imagen original x puede ser recuperada a partir de la imagen observada $\mathbf{y}$ como en Eq. 1 solucionando el siguiente problema de minimización

$$\hat{\mathbf{\Theta}}_{\boldsymbol{c}} = \underset{\mathbf{\Theta}_c}{\operatorname{argmin}} \; \frac{1}{2}\|\mathbf{y} - \mathbf{PB}\left(\mathbf{D} \circ \mathbf{\Theta}_c\right)\|_2^2 + \lambda\|\mathbf{\Theta}_c\|_1, \tag{4}$$

Observe que la estimación de $\mathbf{\Theta}_c$ es ahora obtenida en base a la imagen degradada y ruidosa $\mathbf{y}$. Por lo tanto la estimación de la imagen a partir de $\mathbf{\Theta}_c$ es denotada como $\hat{x} = \mathbf{D} \circ \mathbf{\Theta}_c$.

2 Propuesta

Alternativamente, el problema en Eq. 4 es reescrito para centralizar la representación escasa $\boldsymbol{\theta}_i$ a una buena aproximación $\boldsymbol{\beta}_i$

$$\hat{\mathbf{\Theta}}_c = \underset{\mathbf{\Theta}_c}{\operatorname{argmin}} \; \frac{1}{2}\|\mathbf{y} - \mathbf{PB}\left(\mathbf{D} \circ \mathbf{\Theta}_c\right)\|_2^2 + \lambda \sum_i \|\boldsymbol{\theta_i} - \boldsymbol{\beta_i}\|_1, \tag{5}$$

Esta aproximación $\boldsymbol{\beta}_i$ corresponde a la representación escasa de una combinación ponderada $\tilde{\mathbf{x}}_i$ de parches similares no locales a $\mathbf{x}_i$. Matricialmente, el conjunto de parches similares a $\mathbf{x}_i$ puede denotarse como $\mathbf{Z}_i = [x_{i,1}, x_{i,2}, ..., x_{i,M}]$. En base al problema de minimización centralizado en Eq. 5 se desea en esta propuesta analizar la calidad de los vectores escasos $\boldsymbol{\Theta}_c$ incluyendo una regularización adicional de bajo rango

$$\hat{\boldsymbol{\Theta}}_c = \underset{\boldsymbol{\Theta}_c}{\operatorname{argmin}} \ \frac{1}{2} \|\mathbf{y} - \mathbf{PB}\left(\mathbf{D} \circ \boldsymbol{\Theta}_c\right)\|_2^2 + \lambda \sum_i \left(\|\boldsymbol{\theta_i} - \boldsymbol{\beta_i}\|_1 + \gamma \|\boldsymbol{Z}_i\|_*\right), \tag{6}$$

donde $\mathbf{Z_i}$ es construida a partir de elementos de $\boldsymbol{\Theta}_c$, γ es un parámetro de regularización y $\|\cdot\|_*$ es la norma nuclear (La suma de los valores singulares de una matriz dada).

Definicion 2. MINIMIZACION DEL RANGO

En procesamiento de imágenes la matriz formada por parches similares no locales en una imagen natural es de bajo rango. En base a esto, el principal objetivo es recuperar la estructura de bajo rango subyacente de una matriz a partir de sus observaciones degradadas y ruidosas. Así el problema individual de minimización del rango puede ser formulado como

$$\hat{\mathbf{Z}} = \underset{\mathbf{Z}}{\operatorname{argmin}} \ \frac{1}{2} \|\mathbf{X} - \mathbf{Z}\|_F^2 + \lambda \|\mathbf{Z}\|_*, \tag{7}$$

donde $\mathbf{X}$ es la matriz observada, $\mathbf{Z}$ una aproximación de bajo rango de $\mathbf{X}$ y $\|\cdot\|_F$ denota la norma Frobenius para matrices.