New Sparsity Function

In this supplementary file, we provide more details about the new measure that approximates $L_0$ sparsity during optimization.

Given an input image $z$, the new sparsity measure is applied to image gradient vectors $\partial_z z$ to regularize the high frequency part, where $\ast \in \{h, v\}$ denoting two directions. The function is

$$\phi_0(\partial_z z) = \sum_i \phi(\partial_z z_i),$$

(1)

where

$$\phi(\partial_z z_i; \epsilon) = \begin{cases} \frac{1}{2} |\partial_z z_i|^2, & \text{if } |\partial_z z_i| \leq \epsilon \\ 1, & \text{otherwise} \end{cases}$$

(2)

$\phi(\cdot)$ is a concatenation of two functions – one is a quadratic penalty and the other is a constant. $i$ indexes pixels. One example of the penalty function is shown in Fig. 1(a), with its shape very well approximating $L_0$ penalty when $\epsilon$ is small.

During optimization, we use another form of Eq. (2), which is defined as

$$\phi(\partial_z z_i; \epsilon) = \min_{l_{z_i}} \left\{ |l_{z_i}|^0 + \frac{1}{\epsilon^2} (\partial_z z_i - l_{z_i})^2 \right\},$$

(3)

where $\ast \in \{h, v\}$. Each $l_{z_i} \in \mathbb{R}$ and each $|l_{z_i}|^0$ is a number with the zero power – that is, $|l_{z_i}|^0 = 1$ if $l_{z_i} \neq 0$ and $|l_{z_i}|^0 = 0$ otherwise.

We give the closed-form solution to the problem defined in Eq. (3) in what follows and also show the equivalence between Eqs. (2) and (3).

Claim 1. The function defined in Eq. (3) taking the form $f(l_{z_i}) = |l_{z_i}|^0 + 1/\epsilon^2 (\partial_z z_i - l_{z_i})^2$ has a closed-form solution through hard thresholding as

$$l_{z_i} = \begin{cases} 0, & |\partial_z z_i| \leq \epsilon; \\ \partial_z z_i, & \text{otherwise} \end{cases}$$

(4)

Proof. If $|\partial_z z_i| \leq \epsilon$, we compare the output from $|l_{z_i}|^0$ and $\frac{1}{\epsilon^2} (\partial_z z_i - l_{z_i})^2$. If $l_{z_i}$ is not 0, it must hold that

$$|l_{z_i}|^0 + \frac{1}{\epsilon^2} (\partial_z z_i - l_{z_i})^2 > 1.$$

If $l_{z_i} = 0$,

$$|l_{z_i}|^0 + \frac{1}{\epsilon^2} (\partial_z z_i - l_{z_i})^2 = \frac{1}{\epsilon^2} (\partial_z z_i)^2 < 1.\]$$

So the minimum is reached with $l_{z_i} = 0$.

Similarly, if $|\partial_z z_i| > \epsilon$, we compare the output from $|l_{z_i}|^0$ and $\frac{1}{\epsilon^2} (\partial_z z_i - l_{z_i})^2$. If $l_{z_i}$ is not 0, it must hold that

$$\min_{l_{z_i}} |l_{z_i}|^0 + \frac{1}{\epsilon^2} (\partial_z z_i - l_{z_i})^2 = 1,$$

when $\partial_z z_i = l_{z_i}$. If $l_{z_i} = 0$,

$$|l_{z_i}|^0 + \frac{1}{\epsilon^2} (\partial_z z_i - l_{z_i})^2 = \frac{1}{\epsilon^2} (\partial_z z_i)^2 > 1.$$

So the minimum is reached with $\partial_z z_i = l_{z_i}$ in this case.

Combining the two situations, the final closed-form solution is given by Eq. (4).

The relationship between $l_{z_i}$ and image gradient $\partial_z z_i$ is illustrated in Fig. 1(b).

Claim 2. With the optimal $l_{z_i}$, the penalty function w.r.t. $\partial_z z_i$ defined in Eq. (3) is equivalent to the function in Eq. (2).
Proof. With the optimal value of $l_{si}$ yielded by the hard thresholding in Eq. (4), $\phi(\partial_* z_i; \epsilon)$ output from Eq. (3) is determined by one of the two segments (functions). Specifically, if $|\partial_* z_i| \leq \epsilon$, $l_{si}$ has been proved to be zero to reach the minimum in Eq. (3). Taking it into Eq. (2), we get the simplified function $\frac{1}{2\epsilon^2} |\partial_* z_i|^2$. When $|\partial_* z_i| > \epsilon$, $l_{si} = \partial_* z_i$ makes the function in (3) also be simplified to (2).

In our algorithm, we use a family of loss functions by varying $\epsilon$ and start from $\epsilon = 1$, which makes the loss function quadratic, taking the fact into consideration that each normalized $|\partial_* z_i|$ is always smaller than or equal to 1. In optimization, the penalty function evolves by decreasing $\epsilon$, gradually but steadily heading towards the $L_0$ sparsity function realization. It is a really algorithmically practical, effective and useful technique whenever $L_0$ sparsity is required.