

# Parameter Selection of Total Variation Model Based on BP Neural Network

Haiguang Wei

School of Mathematics and Physics, Anhui University of Technology, Ma'anshan Anhui  
Email: whg001@qq.com

Received: Aug. 2<sup>nd</sup>, 2019; accepted: Aug. 16<sup>th</sup>, 2019; published: Aug. 23<sup>rd</sup>, 2019

---

## Abstract

The energy functional minimization proposed by the variational theory has been widely used in image denoising. In the process of denoising, it is very important to select parameters reasonably. The versatility and operability of neural networks have more advantages than traditional algorithms. In this paper, the shortcomings of the regular variable parameter cannot be adaptively adjusted in the total variation denoising model. By constructing the BP neural network model, after a lot of learning and training, the relationship between the original image information and the regular term parameters is simulated, and the regular term obtained from the model is obtained. The parameters, combined with the Chambolle dual algorithm, form a whole, so that the improved algorithm has more accurate parameter selection and more effective denoising effect.

## Keywords

Image Denoising, Total Variation, Chambolle Duality, BP Neural Network, Adaptive

---

# 基于BP神经网络的全变差模型参数选择

魏海广

安徽工业大学数理科学与工程学院, 安徽 马鞍山  
Email: whg001@qq.com

收稿日期: 2019年8月2日; 录用日期: 2019年8月16日; 发布日期: 2019年8月23日

---

## 摘要

基于变分理论提出的能量泛函极小化, 在图像去噪中已经有了很广泛的应用。在去噪过程中, 合理的选取参数是十分重要的, 神经网络的通用性和可操作性比传统的算法有更多的优势。本文针对全变差去噪

模型中不能自适应调整正则项参数的缺点,通过构建BP神经网络模型,经过大量的学习训练,模拟出图像原始信息和正则项参数的关系,根据模型得到的正则项参数,再结合Chambolle对偶算法,两者构成一个整体,这样改进后的算法其参数选取更为精确,去噪效果更加有效。

## 关键词

图像去噪,全变差,Chambolle对偶,BP神经网络,自适应

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

图像是我们从外部获取信息的重要媒介,但在图像的获取,传输,存储等过程中,难免会受到其他未知因素的干扰,导致图像产生噪声,影响我们对有效信息的获取,因此,对含噪声的图像进行去噪处理一直是图像处理领域的热点和难点之一。在进行图像去噪时,要尽可能的保留图像的边缘、纹理和一些细节信息,以保证修复后的图像与原图像有一定的结构相似性。为解决这一难题,许多专家学者都提出了很多相关理论和方法,其中,在保护边缘方面,1992年Rudin、Osher和Fatemi [1]从能量极小化的角度出发,把全变差半范数作为正则项引入到图像去噪模型中,建立了全变差正则化模型,简称ROF模型,该模型通过引入能量泛函,将图像去噪问题转变为求能量泛函最小问题,并且在保护图像边缘上取得了很好的效果。在保护纹理方面,2001年,Y. Meyer等人提出了TV-G模型[2],将图像分解为几何部分和纹理部分。2005年,Buades等人提出了非局部均值算法[3],利用像素之间的距离来表示相似度,在保护纹理方面取得了很好的效果。在数值求解方面,对偶投影算法,ADMM交替方向乘子法[4],Bregman迭代法[5]等都能达到快速收敛,易于数值实现。

ROF模型中的能量泛函分为逼近项和正则项两部分,其中,逼近项用于去除噪声,保证去噪后的图像与原图像的相似性,而正则项则防止去噪过强,起到抑制噪声的作用,因此,需要选取适当的权重参数用于平衡这两项的关系,目前,国内外许多学者都已经提出了一些权重确定方法,具体有无偏预测风险估计器[6] (URPE),局部方差估计[7],基于Morozov偏差原理[8]选取参数等方法。本文基于全变差正则化模型,利用Chambolle对偶投影算法[9] [10],提出了一种参数可以随着迭代算法自适应地改变的新方法,可根据图像的先验信息,能够自适应的选取合适的参数,进而保证去噪后的图像能够保留更多的细节信息。

## 2. 模型的建立

### 2.1. ROF模型

ROF模型的基本思想是将图像处理问题看成一个能量系统,其中,不含噪声的图像能量相对比较小。当图像被噪声污染后,图像就变得不光滑,其能量也就相应的增大了很多,通过这种思想,就将图像去噪问题转化成了一个能量最小化问题,之后,再通过变分法进行求解。模型如下

$$\min_{u \in BV(\Omega)} E(u) = J(u) + \frac{1}{2\lambda} \|u - f\|^2 \quad (1)$$

其中,  $BV$  表示有界变差函数空间,  $\|\cdot\|$  表示  $L_2$  范数,  $\nabla$  表示梯度算子,  $u$  为去噪后的图像,  $f$  表示观测图像,  $\lambda$  表示正则化参数,  $J(u) = \int_{\Omega} |\nabla u(x)| dx$  模型的第一项为正则项, 第二项为逼近项。我们的图像是一个二维矩阵 ( $N \times N$ ), 用  $X$  表示欧几里得空间, 为了定义离散的全变差范数, 首先给出正则项中离散的梯度算子的具体定义, 梯度  $\nabla u$  是一个在  $Y = X \times X$  的向量。

## 2.2. Chambolle 去噪对偶模型

假定图像被高斯噪声污染, 对针对能量泛函(1)式利用变分法原理进行变分, 获得欧拉-拉格朗日方程, 表达式为

$$0 \in \partial J(u) + \lambda(u - f) \quad (2)$$

式(2)中,  $\partial J$  表示次微分, 定义为  $w \in \partial J(u) \Leftrightarrow J(u) \geq J(u) + \langle w, v - u \rangle_X$ , 根据凸分析中的对偶原理, 上式可以表示为

$$u \in \partial J^*((u - f)/\lambda) \quad (3)$$

式(3)中,  $J^*$  是 Legendre-Fenchel 变换, 是闭凸集  $K$  上的指示函数[11], 其形式为

$$J^* = \chi_K(v) = \begin{cases} 0 & v \in K \\ +\infty & \text{其他} \end{cases} \quad (4)$$

对(3)式进行转化, 有如下等价形式

$$\frac{f}{\lambda} \in \frac{f - u}{\lambda} + \frac{1}{\lambda} \partial J^*\left(\frac{f - u}{\lambda}\right)$$

## 3. 模型的求解

### 3.1. 投影算法

基于最小化下列能量泛函模型, 表达式为

$$\min_{\omega} E(\omega) = \frac{1}{2} \|\omega - \lambda f\|_2^2 + \lambda J^*(\omega) \quad (5)$$

$J^*$  由(5)式已给出, 我们可以推导出  $\omega$  一定是  $\lambda f$  在凸集  $K$  上的正交投影, 若  $P_{K_\lambda}$  表示在集合  $K$  上的投影, 则能量泛函式的解为

$$u = f - P_{K_\lambda}(f) \quad (6)$$

计算非线性投影算子  $P_{K_\lambda}$  等价于求解于下列问题

$$\min_p \left\{ \|\lambda \operatorname{div} p - f\|_2 : p \in Y, |P_{i,j}|^2 - 1 \leq 0, \forall i, j = 1, \dots, N \right\} \quad (7)$$

在 KKT 条件[8]下, 存在拉格朗日乘子  $\alpha_{i,j} \geq 0$ , 满足

$$-(\nabla(\lambda \operatorname{div} p - f))_{i,j} + \alpha_{i,j} P_{i,j} = 0 \quad (8)$$

结合梯度下降算法[12], 取  $\tau > 0, p^0 = 0$ , 以及对任意的  $n \geq 0$ , 有

$$P_{i,j}^{n+1} = P_{i,j}^n + \tau \left( \left( \nabla \left( \operatorname{div}^n - \frac{f}{\lambda} \right) \right)_{i,j} - \left( \nabla \left( \operatorname{div}^n - \frac{f}{\lambda} \right) \right)_{i,j} \cdot P_{i,j}^n \right) \quad (9)$$

再结合之前式子求出  $u^n$ ，若满足  $\left| \frac{u^n - u^{n-1}}{u^n} \right| \leq \varepsilon$ ，则得出的  $u$  即为优解。

### 3.2. BP 神经网络

BP 神经网络在机构上类似于多层感知器，是一种多层前馈神经网络，在结构上具有三层或三层以上神经元，包括输入层、中间层、和输出层。上下层之间实现全连接，而每层神经元之间无链接。当一对学习样本提供给网络后，神经元的激活值从输入层经各中间层向输出层传播，在输出层的各神经元获得网络的输入相应。

#### BP 网络学习规则

网络输入向量  $P_k = (a_1, a_2, \dots, a_n)$ ；网络目标向量  $T_k = (y_1, y_2, \dots, y_q)$ ；中间层单元输入向量  $S_k = (s_1, s_2, \dots, s_p)$ ；输出向量  $B_k = (b_1, b_2, \dots, b_p)$ ；输出层单元输入向量  $L_k = (l_1, l_2, \dots, l_q)$  输出向量  $C_k = (c_1, c_2, \dots, c_q)$ ；

- 1) 初始化。给每个连接权值  $w_{ij}, v_{jt}$ 、阈值  $\theta_j$  和  $\gamma_t$  赋予区间  $(-1,1)$  内是随机值。
- 2) 随机选取一组输入和目标样本  $P_k = (a_1^k, a_2^k, \dots, a_n^k), T_k = (S_1^k, S_2^k, \dots, S_n^k)$  提供给网络。
- 3) 用输入样本  $P_k = (a_1^k, a_2^k, \dots, a_n^k)$  连接权  $w_{ij}$  和阈值  $\theta_j$  计算中间层各单元的输入  $s_j$ ，然后用  $s_j$  通过传递函数计算中间各层单元的输出  $b_j$ 。

$$s_j = \sum_{i=1}^n w_{ij} a_i - \theta_j, \quad b_j = f(s_j), \quad j = 1, 2, \dots, p$$

- 4) 利用中间层的输出  $b_j$ 、连接权  $v_{jt}$  和阈值  $\gamma_t$  计算输出层各单元的输出  $L_t$ ，然后通过传递函数计算输出层各单元的响应  $C_t$ 。

- 5) 利用网络目标向量  $T_k = (y_1, y_2, \dots, y_q)$ ，网络的实际输出  $C_t$ ，计算输出层的各单元一般化误差  $d_t^k$ 。再利用一般化误差和各单元的输出来修正连接权和阈值。直到  $m$  个训练样本训练完毕。

## 4. 数值实验与分析

我们使用不同图像经 Chambolle 处理后的最佳的图像数据为训练样本，选取图像去噪效果最好时的图像结构相似性、峰值信噪比、正则项的数值和逼近项的数值三个因素，为 BP 神经网络的输入神经元，因此神经网络的输入层单元数为 4，正则项前的参数为输出指标，即输出的参数是唯一的输出神经元。

对于隐含层的选择。Nielse 曾证明对于在任何闭区间内的一个连续函数，我们都可以使用一个隐含层来进行设置。Cybenko 指出，当选用的转移函数为 S 型函数时，一个隐含层就能够解决问题。神经网络隐含层数的设计并非越多越好，隐含层数的增多虽说可以增大网络的处理能力，但这也可能会引发反向传播误差，使神经网络结构变得复杂化。本文选用 S 型连续函数，故隐含层数设计为 1。

从图 1 的网络收敛结果可以看出：网络训练 3 次后收敛，网络误差为 0.005749，基本达到了网络误差设置的要求。其中网络误差动态变化效果图来看：纵坐标为误差，我们可以清晰的看出随着训练次数的增加，训练集与测试集误差逐步减小，训练集与测试集变化趋势一致，网络训练到 3 次时，3 条曲线相差很小，误差稳定，此时模型训练效果较好。

从图 2 可以看出，训练样本的拟合优度是 93.40%，测试集的拟合优度是 94.92%，而全部的拟合优度也达到了 94.22%。这个拟合度证明了网络训练效果优秀，此训练模型精准度高，此模型可以解释 94% 左右的数据样本，除去个别数据样本的影响，我们可以完全可以通过此模型进行全变差正则项参数的精准预测。

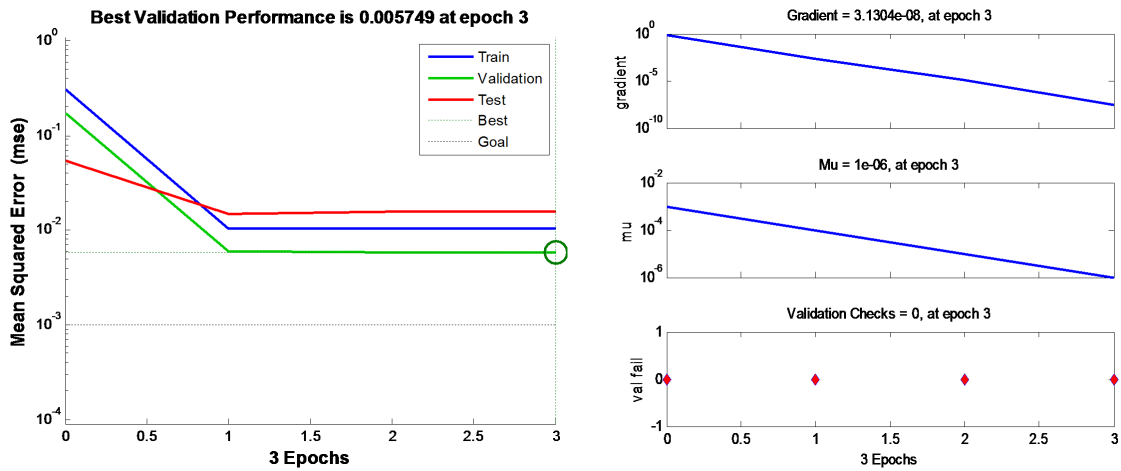


Figure 1. Network training and iteration diagram  
图 1. 网络训练和迭代图

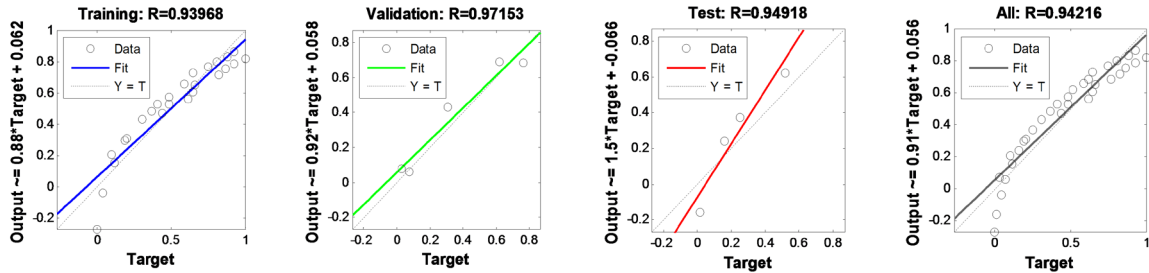


Figure 2. Model fitting chart  
图 2. 模型拟合图

在得到正则项参数的预测值后，结合 Chambolle 对偶算法进行去噪。

我们选取了不同图像分别添加不同方差的高斯白噪声和泊松噪声进行实验，然后利用本文提出的自适应对偶投影算法进行实验，图 3 是对图像添加高斯白噪声和泊松噪声前后实验对比效果图：

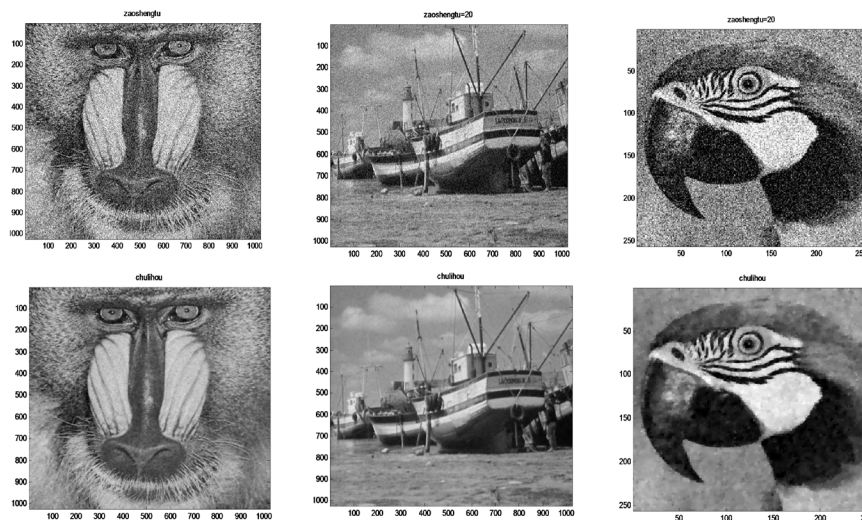


Figure 3. The first line is noise image and the second line is post-processing effect image  
图 3. 第一行是噪声图像，第二行是处理后效果图



从主观方面上来看, 我们可以发现, 无论对于高斯白噪声还是泊松噪声, 试验后的图像效果要比噪声图像清晰很多, 说明了算法改进的可行性, 在客观上, 我们分别计算了每组实验的图像相似结构性 SSIM 和峰值信噪比, 以及算法的运行效率等指标, 分别和 Chambolle 合理值进行对比。

其中, Chambolle 合理值的获得, 是经过大量的人为实验, 在图像取不同参数时算出此时图像的结构相似性 ssim 和峰值信噪比 psnr, 当这两个指标最大时, 此时的 ssim 和 psnr 数值, 我们视为 Chambolle 算法的合理值, 然后将 Chambolle 合理值和改进后算法的结果做对比, 结果如下:

**Table 1.** Image quality index comparison

**表 1.** 图像质量指标对比

图像	Chambolle 合理值		本文改进算法	
	结构相似性 SSIM	峰值信噪比 PSNR	结构相似性 SSIM	峰值信噪比 PSNR
BABOO	0.5231	30.0279154	0.5341	30.09057793
BOAT	0.4717	29.01847508	0.4895	28.99801813
BIRD	0.3952	27.68666436	0.4084	27.78572117

从表 1 中我们可以看出, 无论是经泊松噪声还是高斯白噪声, 经过改进后算法结构相似性比 Chambolle 合理值要高些。峰值信噪比方面, 经本文算法处理后与 Chambolle 合理值很接近, 从而进一步说明了改进后算法的合理性和可行性。由此我们可以看出, 经改进算法去噪后的效果图远比噪声图像清晰度高很多, 在图像纹理和细节部分也能够较好的保护, 这就避免了由于原算法中权重的主观选取, 而导致图像过于光滑和去噪不彻底的弊端, 具有一定的实际应用价值。

## 5. 结论

本文针对全变差去噪模型, 首先分析了 ROF 模型的去噪机理, 然后介绍了利用 Chambolle 对偶原理和 BP 神经网络的学习规则, 利用基于 BP 神经网络对正则项参数的预测, 再结合 Chambolle 对偶算法, 使之构成一个统一整体的模型, 利用改模型该求解速度更快, 收敛性较好, 弥补了一些传统数值算法中不能对权重进行调整的缺陷。同时通过对不同图像添加不同方差的高斯白噪声和泊松噪声进行仿真实验, 在实验效果方面, 我们从结构相似性, 峰值信噪比等方面与合理值进行对比, 其结果表明, 该算法的各项去噪量化指标均可达到较好水平, 从而解决了 Chambolle 对偶算法中参数的主观选取的弊端, 提高了图像去噪的效率, 具有很好的实际价值。

## 参考文献

- [1] Rudin, L.I., Osher, S. and Fatemi, E. (1992) Nonlinear Total Variation Based Noise Removal Algorithms. *Physica D: Nonlinear Phenomena*, **60**, 259-268. [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F)
- [2] Meyer, Y. (2001) Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures. American Mathematical Society, Providence, 122. <https://doi.org/10.1090/ulect/022>
- [3] Buades, A. Coll, B. and Morel, J.M. (2005) Image Denoising by Non-Local Averaging. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, 23 March 2005.
- [4] Bras, N.B., Bioucas-Dias, J. and Martins, R.C. (2012) An Alternating Direction Algorithm for Total Variation Reconstruction Distributed Parameters. *IEEE Transactions on Image Processing*, **21**, 3004-3016. <https://doi.org/10.1109/TIP.2012.2188033>
- [5] Zhang, X., Burger, M. and Osher, S. (2011) A Unified Primal-Dual Algorithm Framework Based on Bregman Iteration. *Journal of Scientific Computing*, **46**, 20-46. <https://doi.org/10.1007/s10915-010-9408-8>
- [6] Dong, Y., Hintermuller, M. and Rincon-Camacho, M. (2011) Automated Regularization Parameter Selection in Mul-

- ti-Scale Total Variation Models for Image Restoration. *Journal of Mathematical Imaging & Vision*, **40**, 82-104. <https://doi.org/10.1007/s10851-010-0248-9>
- [7] Clason, C., Jin, B. and Kunisch, K. (2010) A Duality-Based Splitting Method for L1-TV Image Restoration with Automatic Regularization Parameter Choice. *SIAM Journal on Scientific Computing*, **32**, 1484-1505. <https://doi.org/10.1137/090768217>
- [8] Wen, Y.W. and Chan, R.H. (2012) Parameter Selection for Total-Variation-Based Image Restoration Using Discrepancy Principle. *IEEE Transactions on Image Processing*, **21**, 1770-1781. <https://doi.org/10.1109/TIP.2011.2181401>
- [9] Chan, T., Marquina, A. and Mulet, P. (2000) High-Order Total Variation-Based Image Restoration. *SIAM Journal on Scientific Computing*, **22**, 503-516. <https://doi.org/10.1137/S1064827598344169>
- [10] Chambolle, A. (2004) An Algorithm for Total Variation Minimization and Applications. *Journal of Mathematical Imaging and Vision*, **20**, 89-97. <https://doi.org/10.1023/B:JMIV.0000011321.19549.88>
- [11] Boyd, S., Parikh, N., Chu, E., *et al.* (2010) Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations & Trends in Machine Learning*, **3**, 1-122. <https://doi.org/10.1561/22000000016>
- [12] Chambolle, A. (2005) Total Variation Minimization and a Class of Binary MRF Models. In: *Proceedings of the 5th International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, Springer, St. Augustine, 136-152. [https://doi.org/10.1007/11585978\\_10](https://doi.org/10.1007/11585978_10)

#### 知网检索的两种方式:

1. 打开知网首页: <http://cnki.net/>, 点击页面中“外文资源总库 CNKI SCHOLAR”, 跳转至: <http://scholar.cnki.net/new>, 搜索框内直接输入文章标题, 即可查询;  
或点击“高级检索”, 下拉列表框选择: [ISSN], 输入期刊 ISSN: 2324-7991, 即可查询。
2. 通过知网首页 <http://cnki.net/>顶部“旧版入口”进入知网旧版: <http://www.cnki.net/old/>, 左侧选择“国际文献总库”进入, 搜索框直接输入文章标题, 即可查询。

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [aam@hanspub.org](mailto:aam@hanspub.org)