IEICE TRANS. INF. & SYST., VOL.E99–D, NO.8 AUGUST 2016

LETTER

A Local Characteristic Image Restoration Based on Convolutional Neural Network

Guohao LYU[†a)], **Hui YIN**[†], **Xinyan YU**[†], ***Nonmembers*, *and*** **Siwei LUO**[†], ***Member***

SUMMARY In this letter, a local characteristic image restoration based on convolutional neural network is proposed. In this method, image restoration is considered as a classification problem and images are divided into several sub-blocks. The convolutional neural network is used to extract and classify the local characteristics of image sub-blocks, and the different forms of the regularization constraints are adopted for the different local characteristics. Experiments show that the image restoration results by the regularization method based on local characteristics are superior to those by the traditional regularization methods and this method also has lower computing cost.

key words: *image restoration, L-1 norm, L-2 norm, regularization method, convolutional neural network*

1. Introduction

Image restoration as a classic problem of image processing has been extensively concerned [1]–[5]. A simple degradation model can be described as a convolution process,

$$g(x, y) = (k * I)(x, y) + \epsilon(x, y) \tag{1}$$

where I, g, k, ϵ are the original image, the degraded image, the convolution kernel and the noise respectively. The basic idea of image restoration is to seek a restored image $\bar{I}$, and ensure $\bar{I}$ is close to I after degrading.

$$\bar{I} = \arg\min_{\bar{I}} \frac{1}{2}\|g - k * \bar{I}\|_2^2 \tag{2}$$

However, the solution of the Eq. (2) is not unique. It is an ill-posed problem. In order to solve the problem, the solution $\bar{I}$ need to be constrainted, which is called the regularization method.

In 1963, Tikhonov [1] proposed the regularization method to solve the ill-posed problem and applied it into image restoration (L-2 norm).

$$\bar{I} = \arg\min_{\bar{I}}\{\frac{\lambda}{2}\|g - k * \bar{I}\|_2^2 + \frac{1}{2}\| \nabla \bar{I}\|_2^2\} \tag{3}$$

λ is the regularization parameter and $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$. Since then, scholars have been improving the regularization method constantly in order to achieve higher quality of restored images, such as adaptive regularization method [2], L-1 norm constraint method–TV (total variation) model [3],

$$\bar{I} = \arg\min_{\bar{I}}\{\frac{\lambda}{2}\|g - k * \bar{I}\|_2^2 + \| \nabla \bar{I}\|_1\} \tag{4}$$

the improved algorithm based on TV model [4], L-p norm constraint method [5] etc.

$$\bar{I} = \arg\min_{\bar{I}}\{\frac{\lambda}{2}\|g - k * \bar{I}\|_2^2 + \frac{1}{p}\|\nabla\bar{I}\|^p\} \tag{5}$$

where $1 < p < 2$. Most traditional regularization methods take the same kind of energy functional form for the whole image. However the reasonable energy functional forms should be varied according to the image types. Further, different energy functional forms should be used for different areas of a complex image.

Based on this idea, this letter tries to extract the local characteristics of different image sub-blocks, determine if the sub-block is texture area or smooth area by its characteristics, and then use different regularization constraints for different sub-blocks. However, the current widely used local feature extraction algorithms (such as SIFT feature [6])are not suitable for the aforementioned work. The feature points are specified artificially according to certain rules for these feature extraction methods, so it is uncertain whether the satisfied feature points exist in the sub-block. Even the satisfied feature points exist in the sub-block, but whether these feature points are able to make the distinction between texture area and smooth area still needs to be further researched.

Recently, the convolutional neural network (CNN) has been widely used in the field of object recognition [7], human action recognition [8], video quality assessment [9] etc. The convolutional neural network model is a multilayer neural network containing two types of layers, i.e. convolution layer and sub-sampling layer. It can overcome the disadvantages of traditional local feature extraction algorithms.

Based on the above discussions, this letter proposed a local characteristic image restoration based on convolutional neural network. Its basic idea is to modify the simple hypothesis for the traditional regularization methods that all parts of the image subject to a uniform distribution and to find a more precise description method for local characteristics based on the local priori.

2. Local Characteristic Image Restoration Model Based on Convolutional Neural Network

We train the convolutional neural network by sample image

Manuscript received January 22, 2016.
Manuscript revised April 17, 2016.
Manuscript publicized May 16, 2016.

[†]The authors are with Beijing key lab of traffic data analysis and mining, Beijing Jiaotong University, Beijing, 100044 China.
a) E-mail: lvguohao123@163.edu.cn

DOI: 10.1587/transinf.2016EDL8021

sub-blocks, find the right combination of filters to extract and classify the local characteristics of different sub-blocks, and then use L-1 norm and L-2 norm as the functional definition forms of the regularization constraint for restoration according to the classification results. The typical example is shown in Fig. 1.

As shown in Fig. 1, the image sub-block Surf.I can be regarded as a smooth area because the gray changes smoothly, while Surf.II can be regarded as a texture area because the gray changes intensely. L-2 norm is the better prior constraint form than L-1 norm to restore image sub-block Surf.I, because it is a kind of isotropic diffusion model and can obtain better restoration result in the smooth area. L-1 norm constraint spreads along the edge direction and leads to inadequate noise suppression even causes false edges because the edge of the smooth area is not real. However, L-1 norm is the better prior constraint form than L-2 norm to restore image sub-block Surf.II, because it can preserve edges and textures to obtain better restoration result in the texture area. Therefore, how to accurately classify the local characteristics of different image sub-blocks in an image is a primary issue.

The local characteristic image restoration based on convolutional neural network can be represented as:

$$\begin{aligned}\bar{I} = \arg\min_{\{\bar{I}_i\}}\{&\frac{\lambda}{2}\sum_i \|g_i - k * \bar{I}_i\|_2^2 + \sum_i[\frac{\alpha_i}{2}\|\nabla\bar{I}_i\|_2^2 \\ &+ (1-\alpha_i)\|\nabla\bar{I}_i\|_1]\}\end{aligned} \tag{6}$$

where $\bar{I}_i$, g_i are the sub-blocks of $\bar{I}$, g, and $\alpha_i = 0$ or 1 based on the characteristics of sub-blocks (texture area or smooth area).

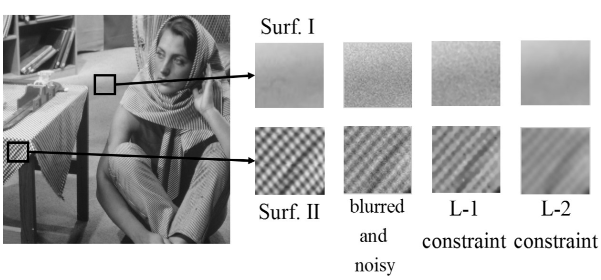


Fig. 1 Different characteristics of different sub-blocks

3. Convolutional Neural Network

In our restoration method, the convolutional neutral network is divided into two parts as shown in Fig. 2. One is to extract and express the features of sub-blocks, and the other is to classify the extracted features.

3.1 Extracting and Expressing the Local Features

The convolutional neural network adopts a multilayer structure and each layer has a specific purpose. Input layer, C1, S2, C3, S4 and C5 together are referred to as the feature extraction of sub-blocks. The convolutional neural network in the feature extraction stage is composed of alternating convolution layer and sub-sampling layer. Here we take each pixel in every sub-block as a neuron and the input layer is a 32×32 sub-block.

Layer C1 is a convolution layer, which is composed of 8 feature maps. Each feature map is the feature of the input image extracted by a convolution filter (8 feature maps correspond to 8 convolution filters respectively). In the feature map, each neutron is connected with a 5×5 area of the input image. The weight coefficients of these 8 convolution filters (5×5) are trained by the training samples. The weights are shared within a feature map. The convolutional neural network model does not consider pixels in the edges of an image sub-block, so the real size of the feature map is 28×28.

Layer S2 is a sub-sampling layer, which contains 8 feature maps in the size of 14×14. Each neuron of the feature map is connected with a 2×2 area in the feature map of C1. Each neuron in S2 can be calculated by sigmoid function, in which the value of independent variable can be got by adding every 4 adjacent neutrons of Layer C1 together, being multiplied by a trainable weight parameter and then adding a trainable bias parameter.

Layer C3 is also a convolution layer, which can be got by executing a convolution to the random combinations in

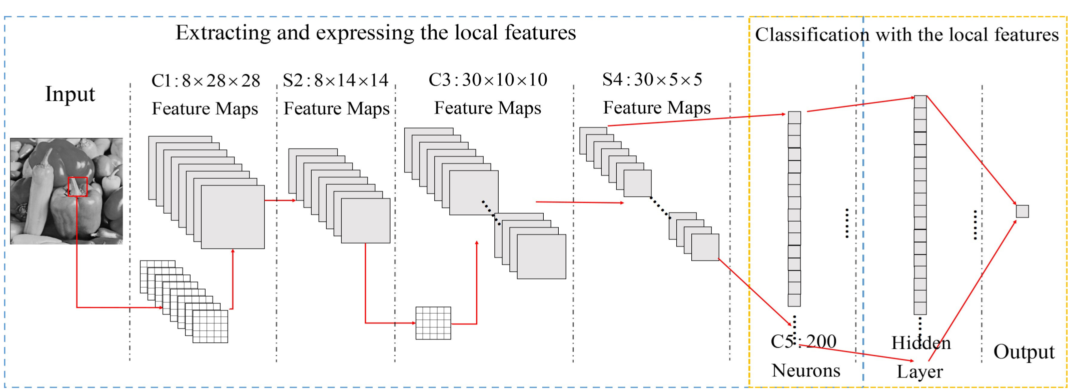


Fig. 2 Feature extraction of the local characteristic image restoration based on convolutional neural network

Layer S2 with a 5×5 convolution kernel. For simplicity, we train a 5×5 convolution kernel, so each feature map only contains 100 neurons. C3 is composed of 30 feature maps in the size of 10×10.

Layer S4 is a sub-sampling layer that is similar to Layer S2. It contains 30 feature maps in the size of 5×5. Every neuron of the feature maps is connected with a 2×2 area of the feature maps in layer C3.

Layer C5 is a convolution layer, which contains 200 feature maps. Full connection of all neurons is needed in this layer. Each neuron is connected with the 5×5 areas of all 30 feature maps in Layer S4. The size of each feature map in layer C5 is 1×1.

Therefore each original 32×32 image sub-block is transferred into a 200-dimensional feature vector.

3.2 Classification with the Local Features

In the convolutional neutral network, the hidden layer and the output layer together are called the feature classification part. A 3-layer structure (including C5) neutral network model is applied to the feature classification, and it is easily replaced by other classification algorithms. The output layer is designed as one neutron to determine L-1 norm constraint or L-2 norm constraint.

4. Variation Gradient Flow Model

By using the variational method and gradient descent, the corresponding partial differential equation of each image sub-block form (refer to Eq. (6)) can be got as:

$$\frac{\partial \bar{I}_i}{\partial t} = [\alpha_i \nabla \cdot (\nabla \bar{I}_i) + (1 - \alpha_i) \nabla \cdot (\frac{\nabla \bar{I}_i}{|\nabla \bar{I}_i|})] - \lambda A \tag{7}$$

α_i is 1 or 0 which is determined by the characteristics of sub-block g_i (texture or smooth area), $A = k * (g_i - k * \bar{I}_i)$ and ∂t is time step. Equation (7) can be solved by the finite difference method, and then the approximate numerical solution of the local characteristic image restoration based on convolutional neural network can be found through explicit iteration.

$$\bar{I}_i^{n+1} = \begin{cases} \triangle t[P - \lambda A] + \bar{I}_i^n, & \alpha_i = 1 \\ \triangle t[M - \lambda NA] + \bar{I}_i^n, & \alpha_i = 0 \end{cases} \tag{8}$$

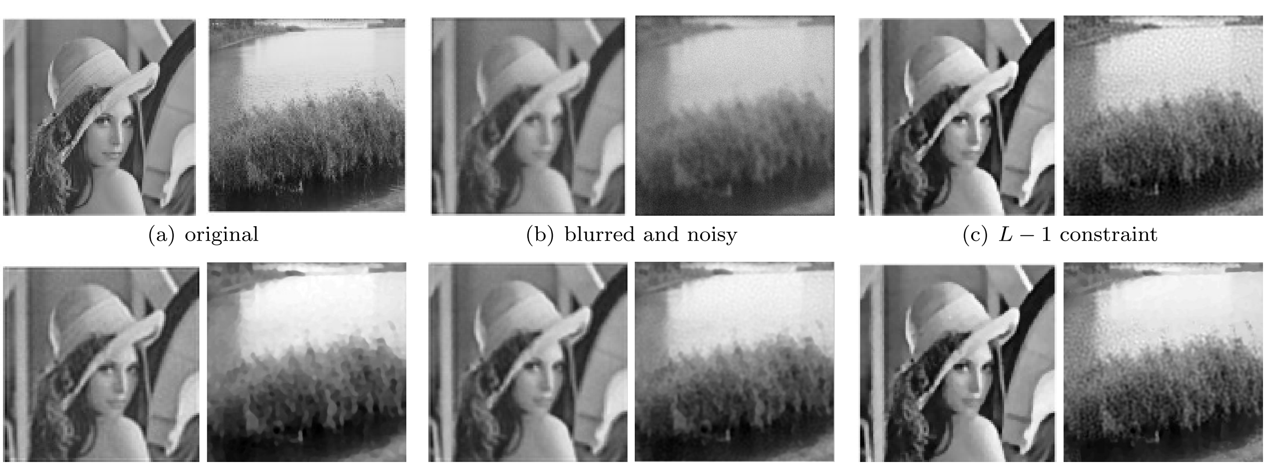
where $\triangle t = \partial t = 0.1$ is the time step, $\lambda = 0.2$ is the regularization parameter, $P = D_{xx}\bar{I}_i^n + D_{yy}\bar{I}_i^n$, $M = \frac{D_{xx}\bar{I}_i^n (D_y\bar{I}_i^n)^2 - 2D_x\bar{I}_i^n D_y\bar{I}_i^n D_{xy}\bar{I}_i^n + D_{yy}\bar{I}_i^n (D_x\bar{I}_i^n)^2}{(D_x\bar{I}_i^n)^2 + (D_y\bar{I}_i^n)^2 + \varepsilon}$ and $N = \sqrt{(D_x\bar{I}_i^n)^2 + (D_y\bar{I}_i^n)^2}$. $D_x\bar{I}_i^n$, $D_y\bar{I}_i^n$, $D_{xx}\bar{I}_i^n$, $D_{yy}\bar{I}_i^n$, $D_{xy}\bar{I}_i^n$ are the first order differences and the second order differences of $\bar{I}_i^n$, ε is an infinitesimal amount. If $\alpha_i = 1$, this method is Tikhonov model (L-2 norm) for sub-block $\bar{I}_i$ and if $\alpha_i = 0$, it is TV model (L-1 norm) for sub-block $\bar{I}_i$.

5. Experiments

In our experiments, we select 400 common test images (like Lena, etc). The selected images are divided into sub-blocks in the size of 32×32 to establish an original image database with 70000 clear image sub-blocks. The Gaussian blur kernel in the size of 5×5 and Gaussian white noise with the standard deviation of 0.05 are used to degrade the selected 400 test images, and a degraded image database can be got by using the above sub-block dividing method. The Peak Signal to Noise Ratio (PSNR) is the evaluation criterion to evaluate the image quality. We restore the image sub-blocks in the degraded image database by using the L-1 and L-2 norm, and compare PSNR of the restored image sub-block by using L-1 norm with that by using L-2 norm. The kind of norm constraint form getting higher PSNR should be used and that restored image sub-block is considered as the classification result. Among all the 70000 sub-blocks, 60000 sub-blocks are chosen as the training sample database and the rest 10000 sub-blocks are as the testing sample database.

5.1 Comparison and Analysis of Experiment Results

In this letter, the step length of convolutional neutral network is 1 and the iteration number is 100 times. The



(a) original (b) blurred and noisy (c) $L-1$ constraint

(d) $L-2$ constraint (e) $L-p$ constraint (f) Our method

Fig. 3 Comparison of experiment results on Lena and Grass

Table 1 Comparison of different restoration results

Evaluation	$L-1$	$L-2$	$L-p$	*Ourmethod*
MSE	45.4623	342.7512	246.2849	23.0169
SNR	18.4161	10.5986	13.4982	21.6314
PSNR	31.6005	25.1053	27.0483	34.971

fitting degree of trained convolutional neutral network is 97.64%. Figure 3 shows the experiment results using L-1 norm constraint [3], L-2 norm constraint [1], L-p norm constraint method [5] ($p_{lena} = 1.5$, $p_{grass} = 1.2$), and the local characteristic image restoration based on convolutional neural network respectively.

Here we use MSE, SNR and PSNR to evaluate the restoration effects. As shown in Table 1.

The smaller MSE means less difference between the restored image and the original clear image. On the contrary, larger SNR and PSNR represent better restoration effects. From Tab.1, the results by our method are the best according to evaluation values. From Fig. 3, the image sub-blocks with different characteristics are restored much more effectively and naturally by our method, so our method is the best according to visual effects. Above all our method can balance the matter between image restoration and detail preserving while most traditional restoration methods cannot.

5.2 Analysis of Computing Cost

It is well-known that the time complexity of using L-1 constraint or L-2 constraint to restore image is $O(n^3)$, where n stands for the pixel size in sub-blocks. The number of pixels in an image sub-block is much smaller than that in a whole image, so the computing complexity, the iteration stability and the time complexity of our method are much better than those of traditional regularization methods. The main deficiency of our method is that the convolutional neutral network needs to be training in advance, and training is time-consuming. However the training process can be computed offline, and it has no influence on the online restoration process.

6. Conclusion

In this letter, we use the convolutional neutral network to identify the local characteristics of image sub-blocks, set up a corresponding relationship between the local characteristics and the reasonable energy functional forms, and then propose an image sub-block restoration method based on local characteristics. Our method is effective and time-saving, so how to extract more effective features to identify the sub-block classification and set up a more reasonable corresponding relationship between local characteristics and energy functional forms will become study emphases in further research for better image restoration effect.

Acknowledgments

This work is supported by National Nature Science Foundation of China (61273364, 61272354, 61300176, 61473031, and 61472029) and Fundamental Research Funds for the Central Universities (2015JBM036).

References

[1] A.N. Tikhonov, "On the solution of ill-posed problems and the method of regularization," Dokl. Akad. Nauk SSSR, vol.151, no.3, pp.501–504, 1963.

[2] D.F. Pace, S.R. Aylward, and M. Niethammer, "A locally adaptive regularization based on anisotropic diffusion for deformable image registration of sliding organs," IEEE Trans. Med. Imaging, vol.32, no.11, pp.2114–2126, 2013.

[3] L.I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Physica D: Nonlinear Phenomena, vol.60, no.1-4, pp.259–268, 1992.

[4] V.B.S. Prasath, D. Vorotnikov, R. Pelapur, S. Jose, G. Seetharaman, and K. Palaniappan, "Multiscale Tikhonov-Total Variation Image Restoration Using Spatially Varying Edge Coherence Exponent," IEEE Trans. Image Process., vol.24, no.12, pp.5220–5235, 2015.

[5] B. Song, Topics in Variational PDE Image Segmentation, Inpainting and Denosing, Ph.D. Thesis, University of California Los Angeles, USA, June 2003.

[6] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vision, vol.60, no.2, pp.91–110, 2004.

[7] H.A. Perlin and H.S. Lopes, "Extracting human attributes using a convolutional neural network approach," Pattern. Recogn. Lett., vol.68, pp.250–259, 2015.

[8] S. Ji, W. Xu, M. Yang, and K. Yu, "3D Convolutional Neural Networks for Human Action Recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol.35, no.1, pp.221–231, 2013.

[9] P.L. Callet, C. Viard-Gaudin, and D. Barba, "A convolutional neural network approach for objective video quality assessment," IEEE Trans. Neural Netw., vol.17, no.5, pp.1316–1327, 2006.