In this letter, we propose a novel single depth map super-resolution algorithm is proposed, which combines the non-local prior and local smoothness prior. Unlike the color-guided methods, the proposed method does not need a corresponding color image to aid the depth map super-resolution. To explore the non-local self-similarity in the depth map, a shape-adaptive adjusted non-local regression is constructed using the shape-adaptive similar patch groups. This prior can make full use of the non-local information of the depth map and alleviate the effect of irrelevant pixels.

To construct different local structures, the direction-based local smoothness prior to model the different directional information is proposed, which can capture the different structures in the depth map. Compared with the state-of-the-art methods, experimental results indicate that the proposed approach can achieve better reconstruction performance.

Optimisation algorithm: Depth map SR can be formulated as the following minimisation problem:

\[
\arg\min_x \|Y - DX\|^2 + \gamma R(x)
\]  

Where \(D\) and \(H\) stand for downsampling and the blurring operator, \(\gamma\) is the regularization parameter. \(R(x)\) is the image prior. The effective prior is important for the SR performance. The proposed method combines two effective priors of non-local self-similarity and local smoothness for the depth map recovery.

Non-local self-similarity: Since the ANLR prior uses the patchwise strategy, it can fully exploit the non-local self-similarity and fully utilize the non-local information of the depth map. However, when the target pixel is in the edge region, the estimated value of it will be inaccurate because of the irrelevant pixel in the similar pixel group. To solve this problem, SA-ANLR prior is adopted. In SA-ANLR prior, the SA patch is used instead of the fixed-size square patch. Since the shape of the patch is adaptive to the image structures, the SA patch can avoid smoothing across edges. To extract SA patches efficiently, the anisotropic local polynomial approximation-intersection of confidence intervals (LPA-ICI) technique [11] is used.

For each target pixel \(x \in X\), we extract a target patch of size \(b \times b\) centered at the target pixel and an SA mask is obtained for the target patch by its shape-adaptive neighborhood. For each target patch, its \(N\) closest patches \(p_n, n = 1, \ldots, N\) are searched within a search window \(W\). The SA similar patch group \(G_n = [M_1(p_1), M_2(p_2), \ldots, M_N(p_N)]\) is gotten by masking these similar patches. \(M(\cdot)\) is the operator of extracting homogeneous pixels in the SA mask. The SA similar patch group (SA-SGP) is shown in Figure 1. Because of the patchwise strategy [9], each pixel belongs to \(B\) overlapped patches \(B = \sum_{n=1}^{b^2} M_{b^2+1}(y)\), which is an overlap-based scheme. \(M_{b^2+1}\) is the SA mask of the neighbor \(m\) of the target pixel \(x\), \(b - m + 1\) is the corresponding position of the target pixel in the patch centered at \(m\). Because there are \(N\) similar patches for each overlapped patch, so each pixel has \(BN\) similar pixels. We call these \(BN\) similar pixels an SA overlapped-based similar pixel group (SA-OSP).

Let \(Q_n = [R_{b^2+1}(G_n)]\) is the similar pixel set of \(x\), where \(G_n\) is the SA similar patch group of neighbor \(m\) of the target pixel \(x\), \(R_{b^2+1}\) is the operation of extracting the \(b^2 - m + 1\)th pixel of each patch in \(G_n\). The SA-OSP \(Q_n\) is shown in Figure 1, where the entries in black mean irrelevant pixels.

Introduction: To acquire high quality depth maps, depth map super-resolution (SR) methods have been proposed. These methods can be classified into colour-guided [1–3] and non colour-guided methods [4–6]. Colour-guided methods recover depth maps by using the corresponding colour image as a powerful aid. However, in many cases, the corresponding colour images are not always available. Non colour-guided methods can infer a high-resolution (HR) depth map from a single low-resolution (LR) input without the assistance of additional intensity images. Benefit from this merit, in this letter, we focus on the non-colour-guided method.

Non-local self-similarity is an important feature of the depth map, which exploits the correlation between patches having similar patterns. Based on this character, many effective methods have been proposed [7, 8]. Among these methods, the non-local mean (NLM) is one of the representative works. Although the above methods exploit the correlation between patches, the non-local feature is formulated only by using the center pixel of each patch, which results in that the non-local self-similarity cannot be fully exploited. Furthermore, for the region that does not locally smooth, there may be great differences in grey value and neighborhood structure between the current pixel and neighbouring pixels. Although small weights are assigned to these neighbouring pixels, the estimated depth map will be blurred because of these irrelevant pixels. Edges play an important role in the texture-less depth map. To restore these structures, most of the methods penalise local smoothness along with horizontal and vertical directions(e.g., [4, 6]). Although these methods can obtain good results, the depth map has different edge structures and these methods will lead to a lack of directional information. Besides, they usually only consider the local structure of the depth map and ignore the patch correlation in the depth maps.

In this letter, we propose a novel single depth map SR method without the guidance of colour image. The proposed method combines the advantage of the non-local self-similarity and the local directional smoothness. To fully exploit the non-local self-similarity of the depth map, we introduce the adjusted non-local regression (ANLR) prior [9], which can make full use of the non-local information of the depth map. However, this prior gives each similar pixel a weight even for the irrelevant pixel. When the patch contains considerable variations, the estimated result will be blurred. To solve this problem, we proposed the shape-adaptive (SA) ANLR prior, which can alleviate the effect of irrelevant pixels and suppress the blurring artifacts. To construct different directional information, we first extract directional features of depth maps using the curvelet transform and then build a direction-based local smoothness prior using these directional features. The directional feature can effectively restore a sharper edge in the reconstructed depth map.

Fig. 1 SA-SPG and SA-OSPG

References

[1] D. Comaniciu, P. Meer, Mean shift: A robust approach toward feature space analysis, IEEE Trans. Pattern Anal. Mach. Intell. 24 (3) (2002) 603–619.

[2] J. Yang, K. Yu, S. Lin, Y. Yu, Connectivity-based sparse representation for image restoration, IEEE Trans. Pattern Anal. Mach. Intell. 34 (1) (2012) 67–81.

[3] J. Yang, K. Yu, S. Lin, Y. Yu, Fast patch-based image restoration, IEEE Trans. Image Process. 20 (11) (2011) 3023–3032.

[4] Z. Fan, J. Liu, H. Zhang, J. Wang, Depth map super-resolution via shape-adaptive non-local regression and direction-based local smoothness

Yingying Zhang,1 Honggang Chen,2 Chao Ren,2 and Ce Zhu1,4

1School of Communication and Information Engineering, University of Electronic Science and Technology of China
2College of Electronics and Information Engineering, Sichuan University, Chengdu, China
3E-mail: eczhu@uestc.edu.cn

Honggang Chen,2 Chao Ren,2
Before constructing the SA-ANLR prior, the similarity weight of each pixel in one SA-OSP needs to be calculated. However, for each similar pixel group, the similarity weight computed from original pixel pairs may be unstable in some cases. To improve the performance, we first filter each pixel in X. For the pixel x, we redefine it by the weighted average of the pixels in its shape-adaptive neighborhood. Then, the weighted pixel x_w is given by:

\[ x_w = \sum_{j \in x} w(j)x_j \]  

(2)

where \( w(j) \) is the weight of the neighbor \( x_j, u_j\) is the shape-adaptive neighborhood of x. Using these redefined pixels, the similarity weight is defined as:

\[ w_x = \frac{1}{\sum_{j \in x} \exp \left( -\frac{(x_n - x_j)^2}{2\sigma^2} \right)} \exp \left( -\frac{(x_n - x_j)^2}{2\sigma^2} \right) \]  

(3)

Finally, we can formulate the SA-ANLR prior as:

\[ J_{SA-ANLR} = \sum_{x \in X} \left| w_x \left( x - \sum_{j \in x} w_j x_j \right) \right|_2^2 \]  

(4)

where \( \Lambda \) is the diagonal matrix, and \( w_x \) is the element on the main diagonal. \( w_i = \alpha / \max(\alpha, \varGamma(G_x)) \), where \( \alpha \) and \( \beta \) are positive constants, \( \varGamma(G_x) \) is the variance of the SA-OSP of x. The weight \( w_i \) penalizes the reliability of the SA-OSP, which enhances the robustness of data estimation. When the standard deviation is small, it means the SA-OSP is reliable and a large weight can be imposed.

Local smoothness: The edge plays an important role in depth maps and has significant directional characteristics. To preserve the directional characteristics, most of the methods force local smoothness along with horizontal and vertical directions. These methods will lead to a lack of directional information since the depth map has different edge structures. In this letter, we use local directional smoothness prior to model this characteristic. The directional smoothness means that in the depth map the intensity difference between adjacent pixels in a specific direction is small. The directional smoothness represents structures of the depth map at the pixel level and is a complement of the non-local prior.

The curvelet transform [12] as a multi-scale directional transform possesses very high directional sensitivity [13]. Since it can represent edges more precisely, we use curvelet transform to extract directional features of depth maps. For a given depth map X, the curvelet coefficients can be obtained by

\[ Q = \Gamma(X) \]  

(5)

where \( \Gamma(.) \) is the curvelet transform. \( Q \) is a set of curvelet coefficient matrices of X and can be expressed as \( \{Q_s, t\} \), s and t are scales and directions, respectively.

The depth map is decomposed into five scales. The curvelet coefficient matrices on the first scale are smooth details in the depth map. The curvelet coefficient matrices from the second to fifth scales are high-frequency coefficients, and each coefficient matrix represents a direction matrix. Since there is more directional information at finer scales, we only consider the finest scale. There are 64 curvelet coefficient matrices on the finest scale. Due to the symmetry of curvelet coefficients, we only use half the matrices on the finest scale and divide them into 16 direction subsets \( \Gamma_k, k = 1, \ldots, 16 \).

The multi-directional matrices \( A_k \) of the input image of the \( k \) direction can be formulated as

\[ A_k = \Gamma^{-1}(\Gamma_k), k = 1, \ldots, 16 \]  

(6)

where \( \Gamma^{-1} \) is the curvelet inverse transform. The directional features \( A_k \) represent the direction intensity of X in the \( k \)th direction. When the value of \( A_k(i,j) \) is large, it means that X(i, j) has more feature information in the \( k \)th direction.

In this letter, we defined the direction-based local smoothness prior as:

\[ \sum_k \| W_k G_k X_k \|_1 \]  

(7)

\( G_k \) are the convolution matrices for the \( k \)th direction. \( W_k \) is a diagonal weighting matrix. The weight is defined as

\[ W_k = e^{-\frac{(G_k X_k(\cdot))}{\sigma^2}} + \frac{H(A_k)}{\eta} \]  

(8)

where the first term is the depth weight. However, the depth weight cannot reflect the rich structures of depth maps. To get the directional information, we introduce the directional feature \( A_k \) to the weight \( W_k \). When the value of \( A_k(i,j) \) is large, it means the \( k \)th direction is the principal direction of X(i, j). According to the directional feature, we can get the principal direction of the image. In the principal direction, the adjacent pixels are more similar and a higher weight should be given \( G_kX \). To achieve this, we add weighting parameter \( H(A_k) \) to the depth weight. \( H(A_k) \) is the operation of extracting the directional feature value for the pixel in edge region.

Finally, by combining the non-local self-similarity prior and local smoothness prior, we have a joint prior model:

\[ \arg\min_x \| Y - D H X \|_2^2 + \gamma \| A(I - W(X)) \|_2^2 + \lambda \sum_k \| W_k^2 G_k X \|_1 \]  

(10)

Equation (10) can be rewritten as:

\[ \arg\min_x \| \tilde{Y} - KX \|_2^2 + \lambda \sum_k \| W_k^2 G_k X \|_1 \]  

(11)

where \( \tilde{Y} = Y(0) \), \( K = (\sqrt{\tau} A(I - W(X))) \).

Equation (11) is a \( 1-1 \) related minimization problem. Here, we take a Sift Bregman iteration approach to solve it [14].

Experimental results: We use six data sets from the Middlebury benchmark database, that is Art, Book, Moebius, Reindeer, Laundry, and Dolls, to evaluate the performance. Because the test images are too big and time-consuming, we only use a small part of the whole picture. To generate LR depth maps, the ground truth images are blurred by a \( 7 \times 7 \) Gaussian kernel with a standard deviation 1.15 and then downsampled to the desired resolution. We test the up-sampling performance at factors of 2 and 4. To evaluate our method, we compare the proposed method with four methods, including the color-guide methods, that is anisotropic total generalized variation (TGV) [15] and static and dynamic guided filtering (SDF) [16], non colour-guide methods, that is edge-guided (EG) [5] and ScSR [17]. The number of similar patches N is 10, the patch size is \( 7 \times 7 \), and the size of the search window is \( 13 \times 13 \). The control parameter \( \gamma \) and \( \lambda \) are set to 0.074 and 0.016, respectively.

Table 1 presents the up-sampling results on quantitative evaluation and mean absolute error (MAE) is used as the evaluation metric. The first
In general, the proposed method can significantly improve the SR performance. The drawback of it is the computational complexity since the local structure analyzing, the non-local similarity pixels searching and iterative optimization of the objective equation are time-consuming. In our future work, we will try to improve the efficiency of the proposed method.

**Conclusion:** In this letter, we have introduced a novel method for a single depth map recovery without a corresponding color image as an assistant. The proposed method introduces a joint non-local self-similarity and local smoothness strategy: To fully exploit non-local self-similarity in the depth map, we build an SA-ANLR prior, which can alleviate the effect of irrelevant pixels and recovery sharp edges. To character different local edge information, we use the direction-based local smoothness. The proposed method combines the advances of the two priors and can recover fine structures. In both quantitative and qualitative evaluations, the proposed method can achieve good performance.

© 2021 The Authors, *Electronics Letters* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Received: 18 December 2020  Accepted: 13 February 2021
doi: 10.1049/el2.12129

**References**

1. Ren, Y., et al.: Edge-guided with gradient-assisted depth up-sampling. *Electron. Lett.* 53(21), 1400–1402 (2017)
2. Dong, W., et al.: Color-guided depth recovery via joint local structural and nonlocal low-rank regularization. *IEEE Trans. Multimedia* 19(2), 293–301 (2017)
3. Liu, X., et al.: Depth Super-Resolution via Joint Color-Guided Internal and External Regularizations. *IEEE Trans. Image Process.* 28(4), 1636–1645 (2019)
4. Ferstl, D., Rüther, M., Bischof, H.: Variational Depth Superresolution Using Example-Based Edge Representations. *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, pp. 513–521 (2015)
5. Xie, J., Feris, R.S., Sun, M.: Edge-Guided Single Depth Image Super Resolution. *Trans. Image Process.* 25(1), 428–438 (2016)
6. Mandal, S., Bhavsar, A., Sao, A.K.: Hierarchical example-based range-image super-resolution with edge-preservation. *IEEE Int. Conf. Image Process. (ICIP)*, Paris, France, pp. 3867–3871 (2014)
7. Favaro, P.: Recovering thin structures via non-local-means regularization with application to depth from defocus. *Proc. Comput. Vis. Pattern Recognit. (CVPR)* pp. 1133–1140 (2010)
8. Park, J., et al.: High-Quality Depth Map up-sampling and Completion for RGB-D Cameras. *IEEE Trans. Image Process.* 23(12), 5559–5572 (2014)
9. Ren, C., He, X., Nguyen, T.Q.: Adjusted non-local regression and directional smoothness for image restoration. *IEEE Trans. Multimedia* 21(3), 731–745 (2019)
10. Chang, K., et al.: Data-Adaptive Low-Rank Modeling and External Gradient Prior for Single Image Super-Resolution. *Signal Process.* 161, 36–49 (2019)
11. Katkovnik, V., et al.: Adaptive window size image de-noising based on intersection of confidence intervals (ICI) rule. *J. Math. Imag.* 16(3), 223–235 (2002)
12. Ma, J., Plonka, G.: The curvelet transform. *IEEE Signal Process. Mag.* 27(2), 118–133 (2010).
13. Li, X., et al.: Single image super-resolution via directional group sparsity and directional features. *IEEE Trans. Image Process.* 24(9), 2874–2888 (2015)
14. Golstein, T., Osher, S.: The split Bregman method for L1-regularized problems. *SIAM J. Imag. Sci.* 2(2), 323–343 (2009)
15. Ferstl, D., et al.: Image guided depth up-sampling using anisotropic total gGeneralized variation. *Proc. Comput. Vis. Pattern Recognit. (CVPR)*, Sydney, NSW, pp. 993–1000 (2013)
16. Ham, B., Cho, M., Ponce, J.: Robust image filtering using joint static and dynamic guidance. *Proc. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, 4823–4831 (2015)
17. Yang, J., et al.: Image super-resolution via sparse representation. *IEEE Trans. Image Process.* 19(11), 2861–2873 (2010)