International Journal of Scientific Research in Science and Technology
Available online at : www.ijsrst.com
OPEN ACCESS
Print ISSN: 2395-6011 | Online ISSN: 2395-602X
doi : https://doi.org/10.32628/IJSRST52310687

Edge-Guided Single Depth Image Super Resolution

Pooja Gavaeikar[1], Dharna Singhai[2]

[1]M. Tech Scholar, Department of Computer Science, Radharaman Engineering Collage, Bhopal, Madhya Pradesh, India
[2]Assistant Professor, Department of Computer Science, Radharaman Engineering Collage, Bhopal, Madhya Pradesh, India

ARTICLEINFO

Article History:

Accepted: 15 Dec 2023
Published: 06 Jan 2024

Publication Issue :
Volume 11, Issue 1
January-February-2024
Page Number :
26-33

ABSTRACT

In this paper Total variation is utilized as a prominent and effectual image prior model in the regularization based image processing fields. Nonetheless, as the total variation model supports a piecewise steady solution, this process comes under high intensity noise in the level areas of the picture is often poor, and a few pseudo edges are formed. In this work we develop a spatially adaptive total variation model. At first, the spatial information is extracted supported each and every pixel, and at that point 2 filtering process are added to restrain the impact of pseudo edges. In addition of this, the spatial info weight is built and classified with k-means clustering, and also the regularization strength in every region is controlled by center value of the cluster. The exploratory results, on both simulated and genuine datasets, demonstrate that the proposed methodology can adequately diminish the pseudo edges of the total variation regularization in the flat areas, and keep up the partial smoothness of the HR images. If we compare the traditional pixel based spatial information adaptive methodology, the proposed region based spatial information adaptive variation model can effectively reduce the effect of noise on the spatial data extraction and maintain strength with changes in the noise intensity in the SR process

Keywords - Single Depth Image, Super Resolution, Edge-Guided, Joint Bilateral Up-Sampling, Markov Random Field.

I. INTRODUCTION

Resolution symbolism is significant in a wide range of programmes, involving as video surveillance, remote sensing, and regenerative imaging. However, due to several constraints from both the theoretical as well as the practical sides, such as the sensor resolution and high cost, among others, it is clearly more difficult to acquire an HR image that a low-resolution image. Scientists have thus examined ways to produce a high-resolution (HR) image using image processing components, and a recent introduction of super-resolution (SR) technology has produced a high-resolution image from one or more low-resolution frames. Our primary area of interest is the multiform

image SR challenge, which involves reassembling an HR image from a series of LR images.

In the past few years, 3D imaging has developed quickly. Two low-cost 3-D imaging technologies, Time-of-flight (TOF) cameras and Windows Kinect, have allowed new research in a variety of domains, including computer vision, computer graphics, augmented reality (AR), and human-computer interaction. However, for many 3D applications, these cameras' low resolution and subpar depth map continue to be major issues. The resolution of the PMD Camcube camera and the SwissRange SR4000 depth camera, for example, is only about 200*200. Even with Kinect, the corresponding colour image (1280*1024) has a resolution that is significantly greater than the depth image (640*480). This work aims to enhance the resolution of depth pictures by utilising a single depth image as the input, which poses some challenges. At its core, single image super resolution (SR) requires a high degree of pixel prediction uncertainty from a limited set of input pixels. Moreover, even if colour images have more texture than depth maps, noises in the depth captured by modern consumer cameras usually make it impossible to calculate the difference or lead to inaccurate scanning equipment. A novel design for single depth image SR is shown in Fig. 1, which can produce better outcomes than previously described methods.

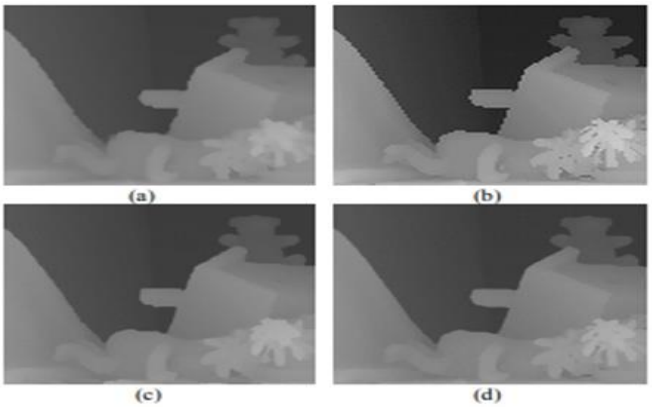

Fig. 1. Visual Result on the Middlebury dataset up-scaled by a factor of 4. (a) Aodha et. al [14], (b) Yang et. al [13], (c) Hornacek et. al [15], (d) Ours

Traditional depth sensing methods focus on integrating many complementary low resolution (LR) depth maps to produce a high resolution (HR) depth picture [1, 2]. However, its operation is predicated on the idea that small camera motions might provide numerous range images, which could not be applicable in many real-world situations. It was also recommended to utilise a pre-aligned HR colour image to help upscale the depth map [5–11], as edge and other high-frequency elements in colour pictures may be used. To recover discontinuity information from the HR colour picture, for example, joint colour and depth upsampling is proposed in [5, 7, 11].

II. LITERATURE SURVEY

Yuyang Wang et.al. (2023) - In many modern applications and future situations, depth super-resolution (SR) is a useful way to close the resolution gap between the relatively low resolution of existing depth sensing devices and the growing need for high-resolution data. The majority of convolutional neural networks (CNNs) are built for certain integer SR ratios (such 2×, 4×, or 8×) or combinations of those ratios when it comes to depth subtraction (SR). CNNs can't be used for depth SR in real applications since depth SR often demands fractional SR ratios. This work proposes a depth map continuously SR (DCSR) architecture that may modify its resolution at any SR ratio [01].

Xiaohui Li et.al. (2022) - Image super-resolution (SR): generating high-resolution (HR) pictures from low-resolution (LR) samples is a basic ill-posed image processing issue. Recent SR research focuses on complex convolution neural network (CNN) architectures that may be used as an end-to-end filter to translate images from LR space to HR space. Few of them, however, focus on the mathematical justification of network architecture or tackle the problem from an optimisation standpoint. In this work, we investigate image SR through the use of the Landweber iteration approach, an excellent

optimisation technique for resolving ill-posed problems. We design an adaptively learning network, inspired by Landweber iteration, that can identify the HR by approaching the problem from an optimisation standpoint [02].

A.Yu Cheng et.al. (2022) - A fundamental ill-posed image processing problem is image super-resolution (SR), which is producing high-resolution (HR) images from low-resolution (LR) examples. Intricate convolution neural network (CNN) architectures that can be employed as an end-to-end filter to map pictures from LR space to HR space are the focus of recent SR research. Few of them, meanwhile, address the issue from an optimization perspective or concentrate on the mathematical demonstration of network architecture. In this paper, we explore image SR using the Landweber iteration strategy, which is a great optimization tool for solving an ill-posed problem. By addressing the problem from an optimisation perspective, we create a network that is inspired by Landweber iteration that can adaptively learn the parameters and detect the HR. Before rebuilding the depth picture, practical noise such as ambient noise and the detector's dark count are eliminated using a denoising modules based on K-singular value decomposition. The theoretical and practical findings show that the proposed technique can conduct adaptive denoising in a wide range of noise circumstances, notably intense ones. The average root mean square error of depth reconstruction pictures is 36.2% less than that of the primary first-photon imaging technique, with a standard deviation of 1.0 [03].

Kaoning Hu et.al. (2021) -High-resolution photos have applications in both science and the arts. In this study, we provide a unique technique that combines texture generation and image vectorization for single picture super-resolution. Image vectorization is the process of converting a raster image to a vector image. While image vectorization can detect the smallest edges of pictures, colour and texture information are lost in the process. Texture synthesis methods, previously utilised in image super-resolution, might be used to synthesise the high-resolution colour and texture information, albeit they might not be able to perfectly follow picture boundaries. In this study, we apply image a vectorization to the picture's edges and textured synthesising to the non-edge portions of the original image using the Kolmogorov-Smirnov test (KS test). The objective is to create a convincing, eye-catching, high-resolution copy of the original picture. Our technique works particularly well with photos of wild animals [04].

Wazir Muhammad et.al. (2021) - The primary goal of single image super-resolution is to recover a high-quality or high-resolution image from a deteriorated version of a low-quality or low-resolution image. Deep learning techniques have recently demonstrated remarkable results in picture super-resolution testing. Nevertheless, neither the hierarchical features required for the final reconstruction nor the feature information present in low-resolution pictures are recovered by the existing super-resolution imaging techniques. In this work, we offer a novel design, patterned after ResNet and Xception networks, that yields SR results while enabling a large reduction in the amount of network parameters and increasing processing performance. We have shown that, when compared to the most advanced algorithms already in use, our proposed method can provide stunning, detailed, and crisp HR images [05].

Binhui Liu et.al. (2020) - A K singular value decomposition-based edge-guided super-resolution method for single frame depth pictures (KSVD). In comparison with conventional techniques, the suggested approach adds two significant aspects. To lessen the impression of jagged edges in up-sampled depth pictures, it first employs KSVD, which develops a whole language to convey the mapping between jagged edges and matching smooth ones. Secondly, it enhances the connectedness-based joint bilateral filter. The enhanced filter lowers noise while maintaining crisp edges during interpolation. The suggested approach has been thoroughly tested and compared

with other cutting-edge algorithms currently in use using the Middlebury dataset. Trial evidence that shows its exceptional performance is both qualitative and quantitative [06].

III.PROPOSED ALGORITHM

A. Majorization-Minimization Procedures (MMP):

Creating a basic upper (tangent) model (such as quadratic forms) at each goal point is one of these strategies. These models are then dynamically updated to minimise and produce minimizing/descent sequences. This concept is at least as old as and has been extensively utilised ever since, mostly in the literature on statistics but also, more lately, in other domains like the imaging sciences. This notion is adhered to by many iterative approaches in the optimisation arena; see for example here and here, where KL inequality is applied to solve non-smooth problems with a particular c class of models. These processes, which we have examined as instruments, seemed to be worthwhile in and of themselves. Our primary findings in this regard are included in Sections and are self-contained.

($\mathcal{p}$) min{f( ): x ∈ $\mathfrak{D}$}

In any scenario, f: Rn! We define local semi-algebraic convex models for f and D, denoted at each feasible point x as h(x; _): Rn! R {which is truly highly convex} and D(x) _ Rn, respectively, since D is a nonempty closed semi algebraic set and R is a semi-algebraic continuous function. Next, we Iteratively solve problems of the form

$$x_{k+1} = \boldsymbol{p}(x_k) = \arg\min\{h(x_k, y): y \in D(x_k)\}, k \in N$$

A key assumption is to use the higher approximations (5) D(x) _ D and h(x; _) _ f(_) on D(x). If the various elements are taken to be semi-algebraic, convergence cannot be seen as a direct consequence of the results. This has a number of reasons. Since we don't have a proper (sub) gradient method" for P _, the fundamental strategy described in demands it. One may readily feel in support of these challenges while thinking about SQP.

B. Regional spatially adaptive:

In this part, we support the use of a spatially adaptive lighting model to manage sudden changes in interior light. Consider a small fixed Lumberton surface patch of area dA on a planar surface. Mentioned is the distance d from this patch to a stationary camera C whose pixel size is dP. The patch is projected onto the camera plane at the homogenous image point X. Before the change in illumination, we assume that N point light sources with radiant flux E1;E2; :::;EN watts lit this area. After the change, the patch is illuminated by a new set of M point light sources with radiant flux F1; F2; FM watts. The model is seen in Figure 2 with just one light source. By assuming that the ambient light irradiance is G watts=m2, one can calculate the radiance L before(X) at pixel point X.:

$$L_{\text{before}}(X) = \frac{\rho.dA.Y_c.dP}{d^2}.\left(G + \sum_{i=1}^{N} \frac{E_i \alpha_i}{s_i^2}\right)$$

where C is the shorter dimensions factor between the camera and the incoming light beam from the patch, and _i and si are the shorter dimensions factor and the measurement distance between the light's source Ei and the surface, respectively. Below is a comparable formulation of the radiance Lafter(X) after the illumination modification.:

$$L_{\text{after}}(X) = \frac{\rho.dA.Y_c.dP}{d^2}.\left(G + \sum_{i=1}^{M} \frac{f_i Y_i}{t_i^2}\right)$$

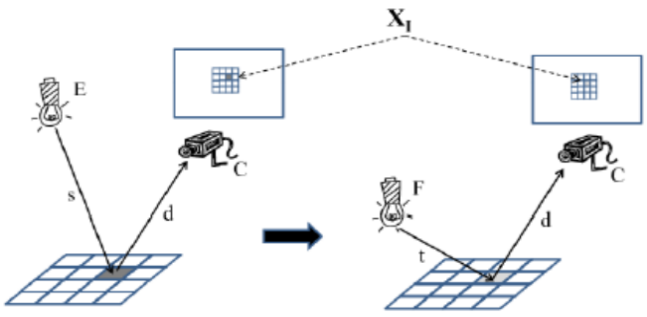

Fig 2 Change of Illumination On a Surface Patch

IV. IMAGE PROCESSING

A. Digital image processing:

A component of the digital image process is the requirement for substantial experimentation to ascertain the feasibility of suggested remedies for a particular problem. One of the most crucial aspects of designing image processing systems is the substantial amount of testing and experimenting that is typically necessary to get the right answer. This trait suggests that developing strategies and rapidly prototyping potential solutions often contributes significantly to cutting down on the expenses and time needed to achieve a workable system implementation.

B. What is an image?

Since an image can be expressed as a two-dimensional operations f(x, y), where x and y are spatially coordinates, the amplitude of "f" at any pair of locations (x, y) is referred to as the intensity of the picture at that place..

Gray scale image:

A grayscale picture might be a representation I (xylem) of the two coordinates of space of the image planes.

The intensity of the picture at point (x, y) on the image plane is denoted by I(x, y).

I (xylem) assumes the picture is confined by a rectangle and accepts non-negative values [0, a] ´ [0, b] I: [0, a] ´ [0, b] ® [0, info]

Color image:

It may be defined using three functions: R (xylem) for red, G (xylem) for green in colour, and B (xylem) for blue.

Along with continuity in the x and y coordinates, a picture can also have a continuous amplitude. Together with the amplitude, such an image has to have its coordinates transformed to digital format. Digitising a coordinate's value is an example of sampling. The procedure of digitising the amplitudes is called quantization.

Coordinate convention:

A matrix of actual values is one result of quantization and sampling that can occur. There are two primary techniques that we frequently apply to depict digital pictures. Assume for now that an image f(x, y) is sampled so that the resulting image has M rows and N columns. It is stated that the image is M X N in size. The values of the coordinates (xylem) are discrete quantities. Since they are discrete coordinates, we frequently use integer values for them to simplify and make notation more clear.

A number of image processing works state that (xylem)=(0,0) is the image's origins.The coordinates of the values along the first row of the picture are (xylem)=(0,1).It is crucial to keep in mind that the notation (0,1) designates the second of two samples on the initial row. This does not mean that the values of the physical coordinates are accurate once the image has been sampled. The remaining figure displays the coordinate convention. Remember that the integer ranges of x and y are, respectively, zero to M-1 and N-1.

The coordinate standard used to describe arrays in the tool example differs somewhat from the one in the preceding paragraph in two ways. At first, the tool case uses the notation (race) instead of (xylem) to signify rows and columns. But notice that the coordinates are in the same order as they were in the paragraph before; a topple (alb), which represents a row, is the first element of a coordinate, the and a column is the second. Since (r, c) = (1, 1) is the origin of the coordinate system, r and c have integer increments from one to M and N, respectively. This is the additional differentiation. The IPT documentation mentions the location. Spatial coordinates is another, less often used coordinate convention by the toolbox that uses x to represent sections and y to designate rows. Perhaps we should have utilised variables x and y in the other way.

Image as Matrices:

The above conversation yields the following example of a digital image perform

f (0,0) f(0,1) …….. f(0,N-1)
f (1,0) f(1,1) …… f(1,N-1)
f (xylem) = . . .
f (M-1,0) f(M-1,1) …… f(M-1,N-1)

A digital image may, by definition, be on the right side of this calculation. We refer to each member in this array as a picture part, image a component pixel, or pel. The words "image" and "pixel" refer to a digital image and its component elements throughout the rest of our talks.

A digital picture may be organically described using a MATLAB matrix :

f (1,1) f(1,2) ……. f(1,N)
f (2,1) f(2,2) …….. f (2,N)
. . .
f = . . .
f (M, 1) f (M,2) …….f(M,N)

f (1, 1) equals f (0, 0) in this instance (note that MATLAB numbers are displayed using a mono scope font). The two portrayals are clearly identical, but for the changed origin. The notation f (p,q) designates the element that is in row p and, so, column q. For example, the element in the matrix f's second column and sixth row is represented as f (6,2). The number of rows and columns in a matrix is commonly denoted by the letters M and N, respectively. A 1xN matrix is called a row vector, while a Mx1 matrix is called a column vector. Scalar is one type of matrix that can exist.

Matrix data is stored in variables called A, a, RGB, real array, and so on in MATLAB. Variables must begin with a letter and can only consist of letters, numbers, and underscores. As was noted in the paragraph above, all MATLAB amounts are written using mono-scope characters. Typically, we use it to write mathematical formulas in italicised Roman notation, as f(x, y).

C. Data Classes

Even though we work with integer coordinates, pixels in MATLAB might have values that are not only integers. The first eight entries in the table correspond to numbers classes. The aforementioned table, where IPT values are used to represent pixel values, displays the various data classes that MATLAB supports. As can be seen, the logical data class is the final element, while the char class is the ninth item.

Applications for image processing also commonly employ this data format, as MATLAB does all numerical calculations in double numbers. These two, together with classes logical and, to a lesser extent, class unit 16, are the primary data classes that we focus on. Because 8-bit images are the most often utilised representations in real-world applications, class unit 8 is very frequently seen, especially when reading data from storage media. All of the data types shown in the table are supported by a large number of ipt functions. In data class double, a number is represented by eight bytes. For example, uint8 and int8 require unit 32 and two bytes, but uint16 and int16 require one bit apiece.

Name	Description
Double	Double _ precision, floating_ point numbers the Approximate.
Uint8	Un signed 8_bit integers in the range [0,255](1byteper Element).
Uint16	Un signed 16_bit integers in the range [0, 65535] (2byte per element).
Uint 32	Un signed 32_bit integers in the range [0, 4294967295](4 bytes per element).
Int8	signed 8_bit integers in the range [-128,127] 1 byte per element)
Int 16	signed 16_byte integers in the range [32768, 32767] (2 bytes per, element
Int 32	Signed 32_byte integers in the range [-2147483648214748364 7] (4 byte per element
Single	Single _precision floating _point numbers with values. In the approximate range (4 bytes per elements)
Char	characters (2 bytes per elements).
Logical	values are 0 to 1 (1byte per element).

V. CONCLUTION & FUTURE SCOPE

Compared to this, the conventional spatially adaptive total variation model is weak to noise and performs poorly in scenarios with high noise levels. To address this, we provide in this work a regionally spatially adaptive (RSATV) super-resolution computation leveraging spatial data filtering and clustering. After extracting each pixel from the spatial information, the spatial weight clustering techniques and the spatial data filtering procedure are applied. By employing these two methods, the total variation model balanced the regularisation performance for every area using distinct location data, as opposed to the conventional spatially adaptive TV model, which balances it for each pixel. The preceding section's examination of actual data reveals that the recommended RSATV model can more successfully decrease noise without compromising edge and flat area data.

Future studies will concentrate on the strategy's adaptive parameter selection and investigate the use of more effective optimisation techniques, including the FISTA and MFISTA algorithms, to quicken the RSATV model's solution time. Moreover, several noise robust spatial feature indicators, including steering weights, will also be taken into consideration to enhance the spatial weight formation process of the proposed technique.

VI. REFERENCES

[1]. Yuyang Wang, Jingyu Yang, Huanjing Yue "Depth map continuous super-resolution with local implicit guidance function" Volume 78, July 2023.

[2]. Xiaohui Li, Ling Wang and Xinbo Liu "Landweber Iteration-Inspired Network for Image Super-Resolution" Volume 2022, 20 Apr 2022.

[3]. Yu Cheng B. Xin-Yu Zhao; C. Li-Jing Li;D. Ming-Jie Sun "First-photon imaging with independent depth reconstruction" Volume 7, Issue 3, March 2022

[4]. Kaoning Hu , Dongeun Lee and Tianyang Wang "Single Image Super-resolution using Vectorization and Texture Synthesis" ISBN: 978-989-758-488-6 2021.

[5]. Wazir Muhammad, Supavadee Aramvith ,Takao Onoye "Multi-scale Xception based depthwise separable convolution for single image super-resolution" August 23, 2021.

[6]. Binhui Liu And Qiang Ling "Edge-Guided Depth Image Super-Resolution Based on KSVD" March 10, 2020.

[7]. Wenhan Yang, Jiashi Feng, Jianchao Yang, Fang Zhao, Jiaying Liu, Zongming Guo, Shuicheng Yan "Deep Edge Guided Recurrent Residual Learning for Image Super-Resolution" 18 Jul 2016.

[8]. Xie, Jun, Rogerio Schmidt Feris, and Ming-Ting Sun. "Edge-guided single depth image super resolution." IEEE Transactions on Image Processing 25.1 (2015): 428-438.

[9]. Shenlong Wang, Lei Zhang, Yan Liang Quan Pan. "Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis." Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012.)

[10]. Min-Chun Yang and Yu-Chiang Frank Wang "A Self-Learning Approach to Single Image Super-Resolution" VOL. 15, NO. 3, April 2012.

[11]. Kim, Kwang In, and Younghee Kwon. "Single-image super-resolution using sparse regression and natural image prior." IEEE transactions on pattern analysis and machine intelligence 32.6 (2010): 1127-1133.).

[12]. Sun, Jian, Jiejie Zhu, and Marshall F. Tappen. "Context-constrained hallucination for image super-resolution." Computer Vision and Pattern

Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010.

[13]. Yang, Wenhan, et al. "Deep Edge Guided Recurrent Residual Learning for Image Super-Resolution." arXiv preprint arXiv:1604.08671 (2016).

[14]. S. Schuon, C. Theobalt, J. Davis, and S. Thrun, "Lidarboost: Depth superresolution for tof 3d shape scanning," CVPR, 2009.

[15]. A. N. Rajagopalan, A. Bhavsar, F.Wallho, and G. Rigoll, "Resolution enhancement of pmd range maps," DAGM, 2008.

[16]. P. Felzenszwalb and D. Huttenlocher, "Distance Transforms of Sampled Functions," Technical Report 2004-1963, Cornell Univ. CIS, 2004.

[17]. D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," IJCV, vol. 47, pp. 7–42, 2002.

[18]. K. Lo,Y. F. Wang, K. H, "Joint Trilateral Filtering For Depth Map Super-Resolution," VCIP 2013.

[19]. Q. Yang, R. Yang, J. Davis, and D. Nister, "Spatial-depth super resolution for range images," CVPR, 2007.

[20]. J. Kopf, M. Cohen, D. Lischinski, and M. Uyttendaele, "Joint bilateral upsampling," ACM SIGGRAGH, 2007

Cite this article as :

Pooja Gavaeikar, Dharna Singhai, "Edge-Guided Single Depth Image Super Resolution", International Journal of Scientific Research in Science and Technology (IJSRST), Online ISSN : 2395-602X, Print ISSN : 2395-6011, Volume 11 Issue 1, pp. 26-33, January-February 2024. Available at doi :

https://doi.org/10.32628/IJSRST52310687

Journal URL : https://ijsrst.com/IJSRST52310687