IR-Guided Energy Optimization Framework for Depth Enhancement in Time of Flight Imaging

Amina Achaibou[1,2(✉)], Filiberto Pla[2], and Javier Calpe[1]

[1] Analog Devices Inc., 46980 Paterna, Spain
{amina.achaibou,javier.calpe}@analog.com
[2] Institute of New Imaging Technologies, University Jaume I, 12071 Castellón, Spain
pla@uji.es

Abstract. This paper introduces an optimization energy framework based on infrared guidance to improve depth consistency in Time of Flight image systems. The primary objective is to formulate the problem as an image energy optimization task, aimed at maximizing the coherence between the depth map and the corresponding infrared image, both captured simultaneously from the same Time of Flight sensor. The concept of depth consistency relies on the underlying hypothesis concerning the correlation between depth maps and their corresponding infrared images. The proposed optimization framework adopts a weighted approach, leveraging an iterative estimator. The image energy is characterized by introducing spatial conditional entropy as a correlation measure and spatial error as image regularization. To address the issue of missing depth values, a preprocessing step is initially applied, by using a depth completion method based on infrared guided belief propagation, which was proposed in a previous work. Subsequently, the proposed framework is employed to regularize and enhance the inpainted depth. The experimental results demonstrate a range of qualitative improvements in depth map reconstruction, with a particular emphasis on the sharpness and continuity of edges.

Keywords: Time of Flight sensor · image fusion · depth enhancement

1 Introduction

Time of Flight (ToF) cameras of 3D sensors are a very competitive 3D sensing choice because of their low cost and relatively high spatial resolution. These sensors consist of an infrared light projector and a depth image sensor. It captures real time depth maps using indirect time of flight technology. However, these devices are unable to correctly estimate depth data in some cases due to larger working distances, occlusions, or low reflective areas. These situations usually lead to get missing depth values regions and unstable boundaries in depth maps as shown in Fig. 1. With the purpose of obtaining fine depth boundaries of

V. Vasconcelos et al. (Eds.): CIARP 2023, LNCS 14469, pp. 646–660, 2024.
https://doi.org/10.1007/978-3-031-49018-7_46

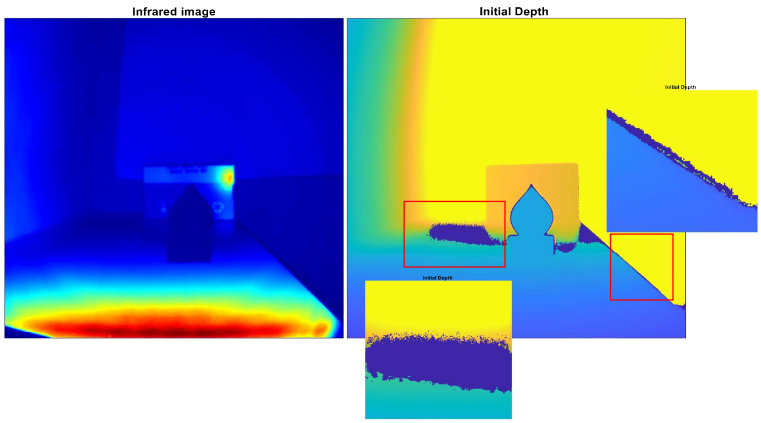

Fig. 1. Corresponding active IR and D map of a ToF camera showing some areas with undesirable/noise effects and missing depth values regions.

objects, by estimating missing depth values and solving the discontinuity problem, some research works propose taking advantage of auxiliary color images captured by an additional RGB sensor, that is, the so-called RGB-D systems. In [2,11,18,19] they propose guided depth enhancement methods based on color images. Other previous works related to this topic used conventional filters. However, depth enhancement techniques that only rely on conventional filtering do not work well when missing depth regions are significant [5].

In order to improve the quality of the estimated depth in ToF cameras, we propose a novel approach which takes advantage from infrared image (IR) captured simultaneously with depth maps (D). The problem is formulated as an energy optimization task with the purpose of maximizing the consistency between the depth map and the corresponding IR image. This formulation is based on the hypothesis that there exists a strong correlation between D and IR. This framework is developed as a weighted estimator based on an iterative conditional modes (ICM) approach. It incorporates conditional entropy information and a spatial error energy term within the image energy model. Additionally, our proposed model introduces a directional weight edges function, which considers all edge directions during the reconstruction and enhancement of object borders. In order to address missing depth values, a pre-processing step is employed, utilizing a depth completion method from a previous work [1], those steps are summarized in Fig. 2.

The main contribution of this work is the use of the active IR image of the same ToF sensor to create a guided depth enhancement, providing more consistency to the depth map. For the best of our knowledge, there is not recent work on IR and depth fusion in ToF systems. The main advantage of IR-D processing rather than RGB-D is to avoid several preprocessing steps, such as calibration, image registration or depth up-scaling.

The diagram of Fig. 2, illustrates the steps of preprocessing and enhancement. The first step involves data denoising, which consists of removing flying pixels and filtering non-confident pixels located around boundaries using morphological

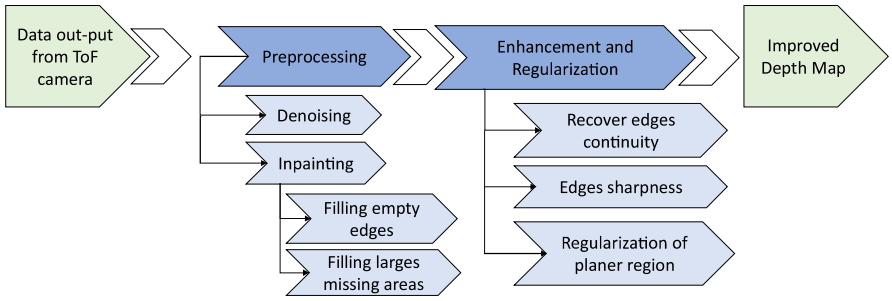

Fig. 2. Proposed processing stages for improving consistency of depth maps.

operators and a confidence map. Subsequently, the inpainting process estimates missing values based on the approach described in [1], aiming to recover missing edges and fill in large missing regions. The final enhancement step enhances sharpness and recover edge continuity through the use of directional edge weights employed in the proposed model. This step also regularizes planar regions by using mutual information extracted from depth and IR images.

The final aim of this work is to use the proposed algorithm in some dedicated hardware build-in the ToF system, and the results achieved so far show that the proposed image fusion-based energy minimization fits better with sensor and hardware resources requirements than using additional RGB sensors. In addition, using a physical model approach avoids learning-based techniques that might need to be fitted for each sensor as a pre-processing or system calibration step. The rest of the paper is organized as follows. First, the problems of depth maps in ToF cameras are introduced, and what are the key issues to be solved. In the Sect. 2, we briefly review some related work. Section 3 explains the reason of using IR as guidance. In Sect. 4 the proposed approach is described in detail. Section 5 presents an overview about experimental data and preprocessing material. Further on, next Sect. 6 shows and discusses the results. Finally, conclusions and further work are summarized.

2 Related Work

The existing approaches for depth enhancement can be broadly categorized into two groups based on the input data: guided depth enhancement and self depth enhancement. On the one hand, the guided depth enhancement category relies on additional information such as color images or depth captured using different sensors like stereo systems. On the other hand, the self depth enhancement category focuses on single depth enhancement techniques. In this work, our primary focus is on techniques that fall under guided depth enhancement. These methods have recently gained significant attention, as evident in several studies such as Diebel et al. [3], Izadi et al. [8], Ferstl et al. [4], Kwon et al. [10], and Lu et al. [13]. Those methods leverage additional depth maps or color images to improve the quality of the final depth map. The most popular solution is to

incorporate an additional high resolution color sensor together with a depth sensor for depth image enhancement [3,4,13]. Another representative solution is to utilize multiple depth images from the same scene to reconstruct a higher quality depth maps [6]. This method, however, rely heavily on accurate multi-camera or multi-view calibration, and may fail when applied to dynamic environments.

Most methods leverage RGB images to enhance the depth data and assume that there is a strong correlation between the depth map and the corresponding color image. Or et al. [14] fused an intensity image and a depth map to perform a precise shading recovery and albedo estimation for detailed surface reconstruction. Liu et al. [12] proposed an optimization framework that is weighted by color guidance and utilizes a robust penalty function for smoothness modeling. Jiang et al. [9] proposed a method for exploiting correlations in the transformed and spatial domains using a unified model for recovering geometrical structures. Yang et al. [17] extracted scale-independent features from depth maps with the assistance of RGB image and proposed an edge-aware neighbor embedding (NE) framework for facial depth map super-resolution. In addition, several methods have also been proposed for fusion strategies of depth information from multiple frames and other simultaneous sources, in order to produce higher quality depth maps and they have proved its usefulness as a depth enhancement approaches. Deep learning has been introduced recently for depth denoising by using graph networks in [16]. The denoising and enhancement conventional neural network (DE-CNN) proposed in [20] and deep image-guided method [22] have also been adopted for depth enhancement.

In summary, the main guided depth enhancement methods are the filtering-based methods such as (joint bilateral filter), the optimization-based methods (Markov random field, auto regressive model, total variation, graph Laplacian), and the learning-based method (dictionary learning, deep learning). Our focus in this study is related to filtering and optimization-based methods, since one of the final aims is implementing the algorithm on hardware with an easy configuration.

3 IR Image as Guidance

For our experiments, we employed a ToF camera that operates on continuous wave pulse technology and based on indirect time of flight system. It is important to note that this system provides active IR images based on continuous amplitude measurements, while the depth based on phase measurements exhibit discontinuities because of the estimation is based on multiple frequencies agreements (phase unwrapping). It is worth noting that, the amplitude information (active IR images) can offer greater confidence for each corresponding depth value, as we can see from the outcome of a ToF camera when analysing the IR image comparing to the depth map shown in Fig. 3. Note how the depth map shows noise and discontinuous measurements in depth which affect details and important information in some specific areas of the images, such as borders and low reflectance regions.

This can lead to noisy or loss of fine details, whose information could be necessary for accurate interpretation or analysis of the image. However, note that IR measurements preserve the overall structures and details of objects in the scene in a more consistent way, even though pixel values may lose some precision due to noise effect but it usually keeps details of the objects structure. An additional filtering can enhance the IR precision which can not be the case for depth. It is hard to recover the wrong values and lost details in the depth map by using a classical filtering. and that what make active IR used as guidance.

The preprocessing in the pipeline of the camera usually involves filters and contrast stretching that help to improve the overall quality and clarity of the IR image, including the reflectance of low-reflective areas. Looking at the active IR image before and after the preprocessing integrated in the camera pipeline shows that IR images have some interesting properties, such as a better signal-to-noise ratio than depth maps, better defined object boundaries, and they are less sensitive to noise than the corresponding depth maps estimates see Fig. 3.

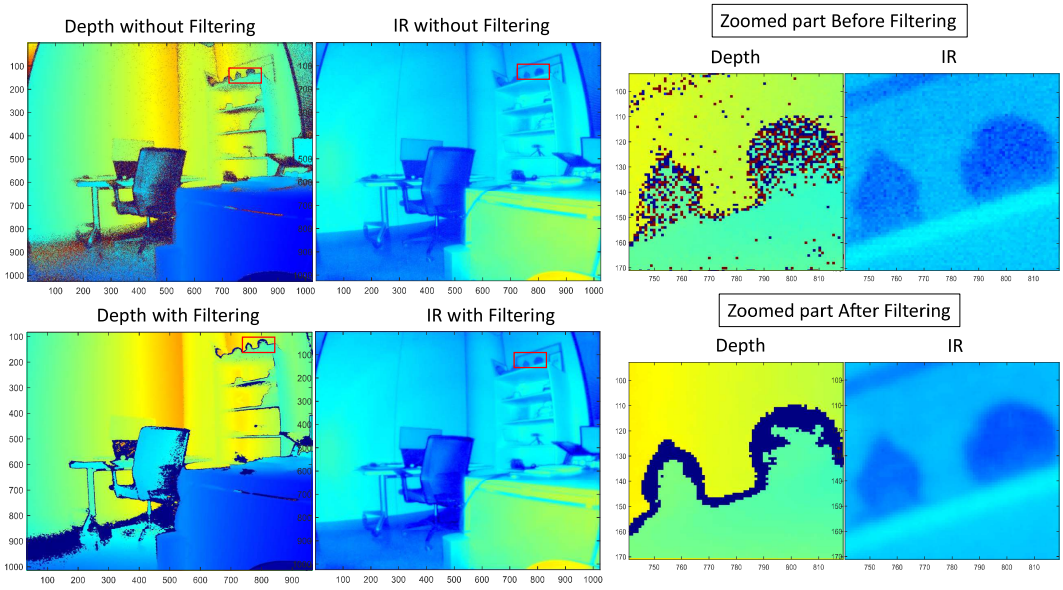

Fig. 3. Discontinues Depth map and its corresponding Continues active IR image captured by a ToF camera before and after filtering in the camera pipeline.

4 Proposed Model

The image energy model is based on a weighted optimization approach to combine the ToF depth maps D and the corresponding active IR images of ToF system. Note that in this case, D and IR images are captured from the same sensor simultaneously, and the active IR is used as a guidance to enhance and improve the consistency of the initial depth.

4.1 Image Energy Function

The proposed energy function is based on two main assumptions: firstly, a strong correlation is assumed between the guiding active IR image and the depth map D; secondly, it assumes a spatial similarity among neighboring depth values, except at depth discontinuities represented by object edge weights. Hence, the image energy Q combines two terms, the spatial error energy Q_S to eliminate local variations in depth, produced by common sensor noise; and the spatial entropy energy term Q_H to enhance depth discontinuities and suppress other non-local noise by maximizing the correlation between the final depth map and the guiding IR image. The resulting enhanced depth map D^* is computed by minimizing the image energy Q expressed as:

$$Q = cQ_S + (1 - c)\, Q_H. \tag{1}$$

where c is the regularizing parameter used to control the effect of each energy term.

4.2 Spatial Error Energy Term

Spatial error energy term Q_S is introduced to reduce spatial image noise, such as sensor noise. This regularization enforce similarity between neighboring pixels depths. It is expressed as the errors ϵ for each of the 4-neighbours directions as follows,

$$\begin{aligned} Q_s = \sum_{x,y=1}^{X,Y} \big(\big[WE_{xy}(x-1,y)\epsilon^2_{xy}(x-1,y)\big] + \big[WE_{xy}(x+1,y)\epsilon^2_{xy}(x+1,y)\big] + \\ \big[WE_{xy}(x,y-1)\epsilon^2_{xy}(x,y-1)\big] + \big[WE_{xy}(x,y+1)\epsilon^2_{xy}(x,y+1)\big] \big). \end{aligned} \tag{2}$$

For each depth $D(x, y)$ at the position (x, y) we define the spatial regularization error ϵ for the 4-neighbour pixels $(x', y') \in N(x, y)$ as:

$$\epsilon_{xy}(x', y') = D(x, y) - D(x', y'). \tag{3}$$

However, enforcing spatial similarity may over-smooth the depth discontinuities at object borders. Therefore, directional edge weights $WE_{xy}(x', y')$ are used to avoid edge blending. The directional weight edge functions are defined in Eq. (4) where $(x', y') \in N(x, y) = \{(x-1, y), (x+1, y), (x, y-1), (x, y+1)\}$.

$$WE_{xy}(x', y') = -\exp[k \cdot EI_{xy}(x', y') \cdot ED(x', y')]. \tag{4}$$

Those weights are based on the depth map gradient module $ED(x', y')$, and directional IR image gradient modules $EI_{xy}(x', y')$. To obtain these gradient modules, IR and depth maps are first smoothed by a Gaussian filter $G(0, \sigma)$ of zero mean and variance σ^2. Then the ED gradient is extracted by directly applying a gradient on depth map D and the directional EI gradients extracted

from the active IR image for every 4-neighbour pixels $(x', y') \in N(x, y)$ at a given pixel location (x, y), that is,

$$ED_{xy}(x', y') = |GD(x, y) - GD(x', y')|. \tag{5}$$

$$EI_{xy}(x', y') = |GI(x, y) - GI(x', y')|. \tag{6}$$

where $GD(x', y') = G(0, \sigma) * D(x', y')$ is the depth map D smoothed with a Gaussian filter $G(\mu, \sigma)$ the same for $GI(x', y') = G(0, \sigma) * A(x', y')$, being A the active IR image values. The resulting directional IR gradients EI can include many edges where there are no depth changes. This may lead to unwanted characterization of depth discontinuities in some areas. Hence, to remove edges not corresponding to depth discontinuities, only the IR gradients EI corresponding to depth gradients ED denote depth object edge discontinuities, thus leading to the directional weight edges functions WE_{xy} defined in Eq. (4), With k denoting as a scaling factor.

4.3 Conditional Entropy Energy Term

The rationale behind this energy term is based on the correlation between depth map D and active IR image A. This correlation can be represented by the mutual information between both images $MI(A, D)$, defined as,

$$MI\,(A, D) = -\sum_{a,d} p\,(a, d) \log \frac{p\,(a, d)}{p(a)p(d)}. \tag{7}$$

where $p(a, d)$ is the joint probability of IR image A and depth values D, $p(a)$ and $p(d)$ are the priors for the active IR image and depth map values respectively, and a and d denoting the possible IR and depth values. Therefore, maximizing $MI(A, D)$ would lead to a maximization of the correlation between both images, which can lead to improve the consistency of depth map D with respect to active IR image A. Further on, MI can be defined in terms of conditional entropy as,

$$MI\,(A, D) = H(D) - H(D/A). \tag{8}$$

Note that, the conditional entropy $H(D/A)$ can be interpreted as how much information remains on D that cannot be explained giving the corresponding active IR values A. Since the derivative of MI is quite complex, and assuming that $H(D)$ is approximately constant, minimizing $H(D/A)$ would be equivalent to maximizing $MI(A, D)$. Thus, the conditional entropy $H(D/A)$ is defined as follows,

$$H\,(D/A) = -\sum_{a,d} p\,(a, d) \log p\,(d/a). \tag{9}$$

Therefore, the corresponding energy term Q_H is defined as:

$$Q_H = H\,(D/A)\,. \tag{10}$$

where $p(d/a)$ is the conditional probability of depth given the active IR values. Note that in this formulation D must be discretized. In practice, the joint probability $p(a, d)$ is estimated by computing the normalized joint histogram between the initial depth D and the corresponding active IR image A. The conditional in Eq. 9 is computed from the joint probability using the chain rule as $p(d/a) = p(a, d)/p(a)$, and prior $p(a) = \sum_d p(a, d)$ as the marginal of the joint probability $p(a, d)$.

4.4 Image Energy Minimization

The image energy minimization process is based on an Iterative Conditional Modes (ICM) strategy, that is, instead of minimizing Q energy (1) for the entire image, pixel energy $Q(i)$ will be minimized independently for each pixel $i = (x, y)$ in an iterative way, that is,

$$Q(i) = c \cdot Q_S(i) + (1 - c) \cdot Q_H(i) \tag{11}$$

and

$$D(i)^* = argmin\left(cQ_S(i) + (1 - c)Q_H(i)\right). \tag{12}$$

Thus, pixel spatial error term $Q_S(i)$ and pixel conditional entropy $Q_H(i)$ are defined as follows,

$$\begin{aligned} Q_S(i) = &[WE_{xy}(x-1, y) \cdot \epsilon^2_{xy}(x-1, y)] + [WE_{xy}(x+1, y) \cdot \epsilon^2_{xy}(x+1, y)] \\ &+ [WE_{xy}(x, y-1) \cdot \epsilon^2_{xy}(x, y-1)] + [WE_{xy}(x, y-1) \cdot \epsilon^2_{xy}(x, y-1)]; \end{aligned} \tag{13}$$

$$Q_H(i) = -WE_{xy}(i) \cdot p\left(a_i, d_i\right) \cdot \log p\left(d_i/a_i\right). \tag{14}$$

The joint probability $p(a, d)$ can be estimated from the normalized joint histogram between the initial depth map D and the corresponding active IR image A. Where $p(a_i, d_i)$ stands for the joint probability for depth value $d_i = D(i)$ and active IR value a_i at pixel $i = (x, y)$. Edge weights $WE_{xy}(i)$ are also introduced in pixel conditional entropy $Q_H(i)$ to preserve edge pixels regularization. Finally, to minimize the pixel energy $Q(i)$ at each iteration, a gradient descent method is used, with adaptive learning rate λ,

$$\hat{D}(i) = D(i) - \lambda(\partial Q(i)/(\partial D(i)). \tag{15}$$

5 Data and Preprocessing

In this work image data were acquired from a ToF ADI sensor ADTF3175 module based on continuous wave illumination to test the proposed approach, with no depth ground truth available. IR-depth option was used in that sensor to collect the depth aligned with its corresponding active IR image. Both images are geometrically and time aligned, as they are captured by the same sensor with 1024×1024 pixels resolution. Figure 4 illustrates some representative examples where the second column of the first row is the depth map, and the first column

of the first row shows the corresponding active IR image. The scenarios were chosen to include both narrow and large missing depth regions as a result from occlusion, dark or low reflective surfaces, or any other factors producing missing depth values.

The initial preprocessing is oriented to estimating missing depth values. Thus, a belief propagation depth completion method was used [1]. Note that this method takes the active IR image as a guidance, particularly for missing edges recovery. The key point of this inpainting approach is that it takes into account the direction of how inpainting is performed around boundaries, where the filling converges from the missing region value boundaries to the objects borders from both depth discontinuity sides, when the missing depth values region contains an object border.

In order to compute the joint probability $p(a, d)$, depth and IR values are discretized. In a further step, the joint histogram of the depth map D and the active IR image A is computed and then normalized to obtain the joint probability and the corresponding conditional $p(d/a)$ using the chain rule.

For the image energy optimization process, the initialization is set to the quantized depth map provided by the ToF sensor and its quantized active IR image. In the gradient descent, the adaptive learning rate is defined as $\lambda_t = \lambda_{(t-1)} - \lambda_c$, being λ_c a constant, and the iterative process is stopped when the change in the MI between the active IR and the discretized version of the estimated depth image is below a given small value.

6 Experiments and Results

The qualitative results of the experiments based on the ADTF3175 data-set are presented in Figs. 4, 5, 6 and 7. These figures provide detailed insights into edges recovering and depth enhancement. Notably, the proposed method significantly improves the consistency of depth maps regarding to infrared image, since missing values around some edges are one of the major problems of ToF cameras the belief propagation depth method (BLFP) used for inpainting as preprocessing. in Fig. 4 the third column of the first row illustrates the result of depth inpainting, while the fourth column showcases the depth enhancement results after integrating the BLFP with the proposed image energy optimization framework. This combination effectively reconstructs missing pixels and enhances depth map, particularly around edges that were initially missing and noisy. This is clearly demonstrated in the second row, first column of Fig. 4.

It is worth noting that even after the depth inpainting process, the depth still exhibits discontinuous edges and some non-homogeneous transitions between objects and the background. For these reasons, the fusion module proposed in this work aims to further refine the inpainted depth. This refinement is illustrated in the analysis of the results presented in Figs. 5, 6 and 7.

Figure 5 shows how the edge profile and continuity were recovered based on optimization framework. Figure 6 illustrate some details of edge continuity when concerning nonlinear borders, the left edge map is extracted from initial depth

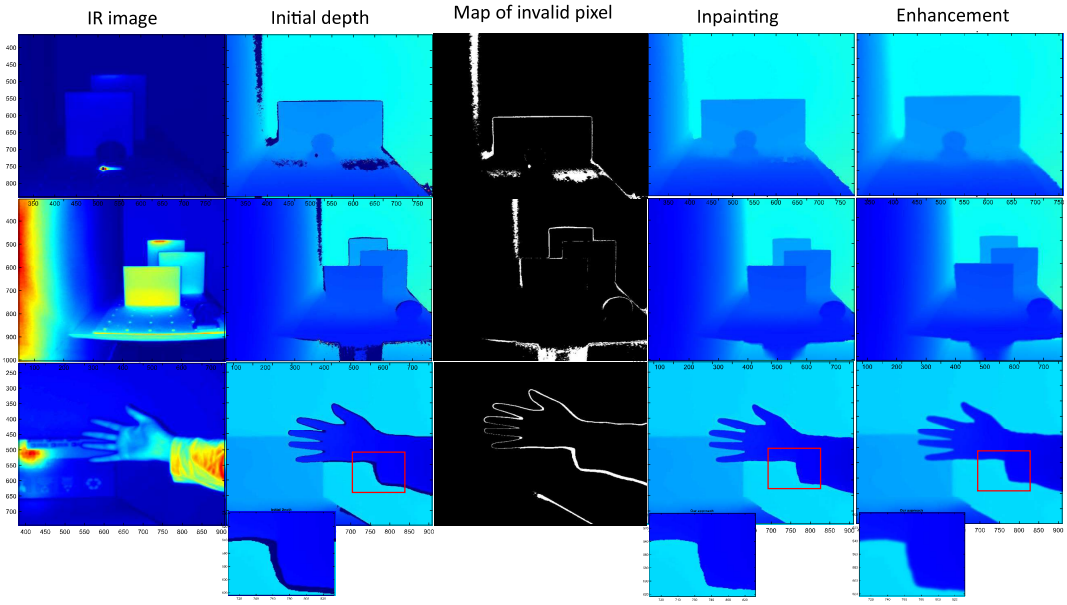

Fig. 4. Depth inpainting and final depth enhancement, large missing depth areas reconstruction, missing edges reconstruction.

and the right one extracted from the enhanced depth after applying the proposed method. The left one shows some discontinuities and double boundaries around the hand, that means the hand in the initial depth was surrounded by a missing area, the right map show the hand border after processing. This significant improvement is attributed to the utilization of the directional weight edge function, which was employed for spatial error regularization in the defined image energy function (4.2). The directional weight edge function plays a crucial role in guiding the optimization process to refine objects boundaries. This edge function provides directional information that helps to preserve edges more accurately during the enhancement process, ensuring smoother and more coherent depth transitions between objects and the background.

Figure 7 represents the edge intensity (EI) [21] of a nonlinear shape, where EI of the inpainted depth is shown in the second column and the EI of the enhanced depth in the third column. The zoomed part from the object boundary is shown in the second row. From these results, it is evident that the inpainted depth still exhibits some artifacts around the object's boundaries. However, the depth enhancement process based on our proposed framework plays a crucial role in regularizing the areas near the edges. This regularization is achieved through the similarity term, which effectively refines non homogeneous depth edges respecting to infrared information.

Figures 8 and 9 shows the depth enhancement comparison results of the proposed approach with respect to other reference guided enhancement methods (joint bilateral filter [15], guided filter [7], and guided anisotropic diffusion [11]). From this figures and from the zoomed patches we can notice that the proposed enhancement model produces regularization on the smooth object surface

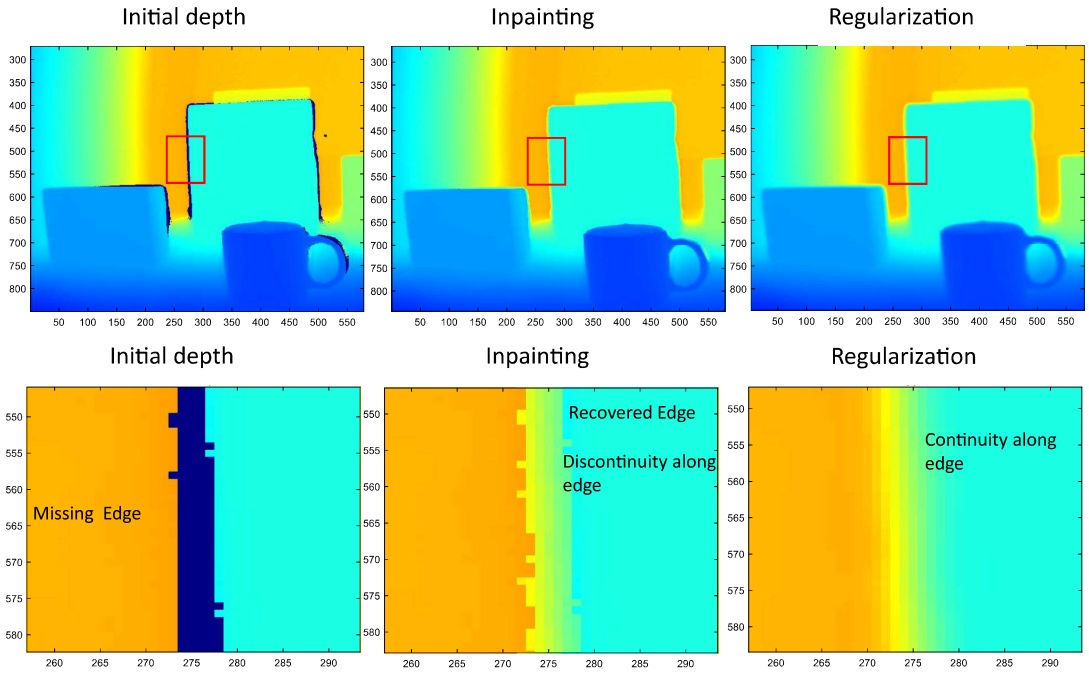

Fig. 5. Edge reconstruction based on belief propagation inpainting method and edges continuity recovering based on the proposed image energy optimization model.

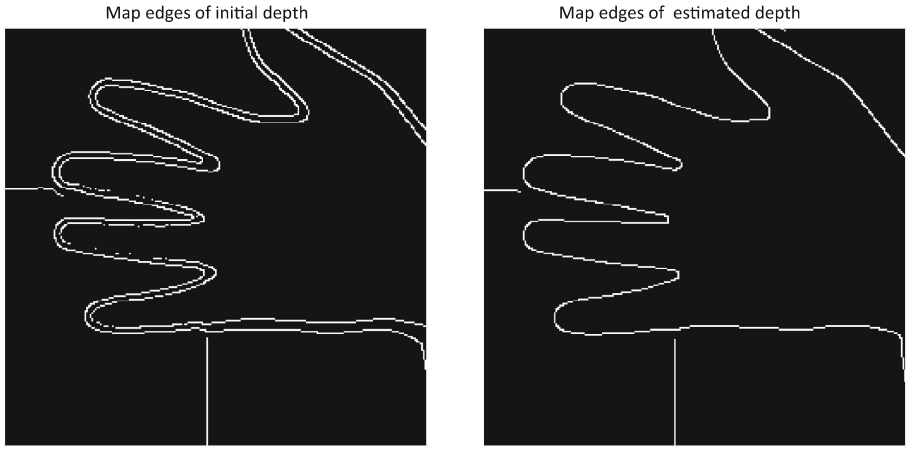

Fig. 6. Edge continuity recovering based on the proposed image energy optimization model.

regions. This is because of the conditional entropy minimization, which also regularizes around edges and improve the sharpness of depth borders due to the directional edge weights introduced in the image energy model.

The most important outcome of the proposed model is the improvement of edges continuity, by maximizing the correlation (mutual information) between the resulting depth and the active IR image. This process tends to change depth values in such a way that they correlate as much as possible with the active IR image values and local spatial distribution.

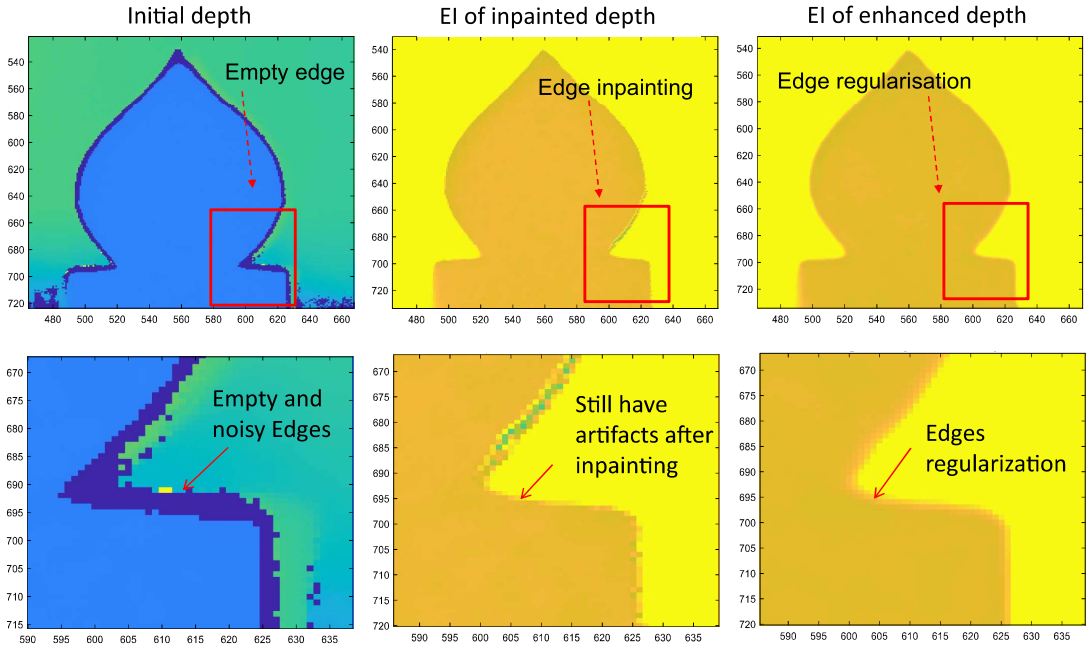

Fig. 7. Edges regularization improvement based on the proposed image energy model. The second and third columns are edge intensity maps.

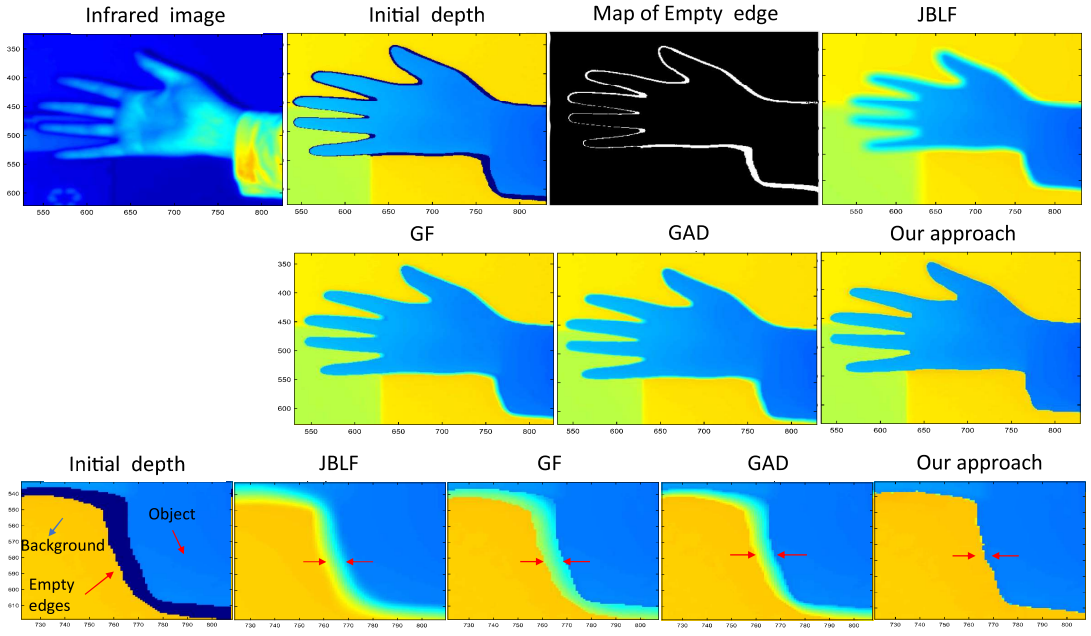

Fig. 8. Comparison of the proposed image energy model with respect to other reference methods.

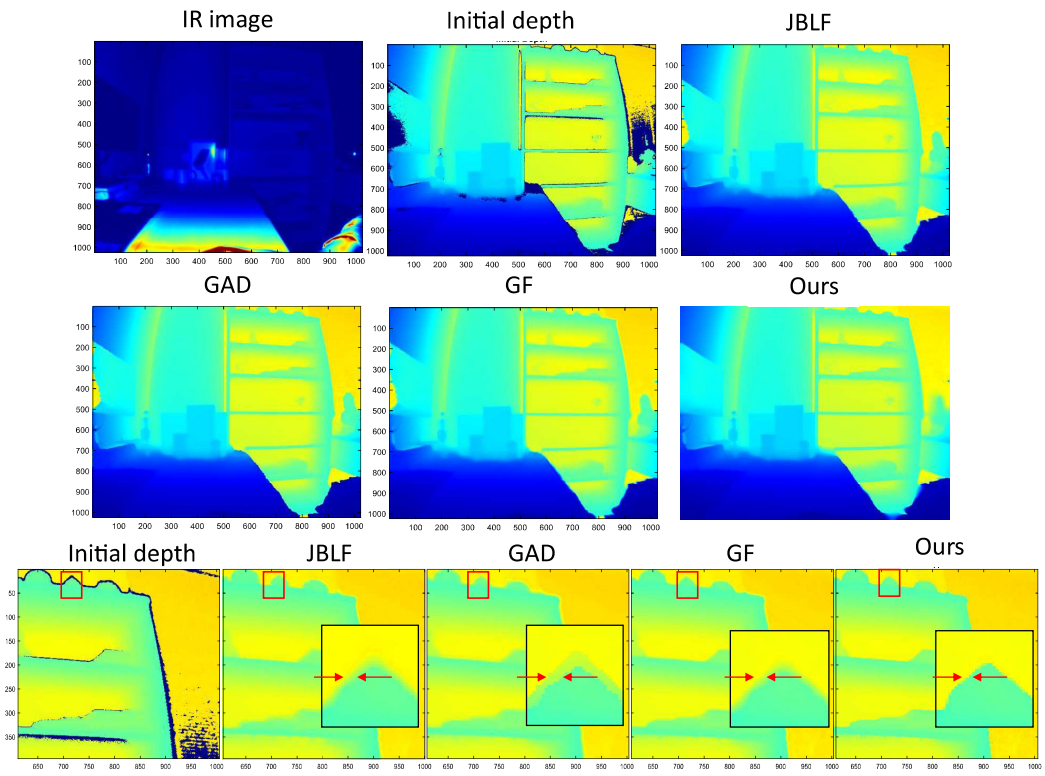

Fig. 9. Comparison of the proposed image energy model with respect to other reference methods.

7 Conclusions

The proposed method is a novel guided approach for dense depth enhancement of ToF imagery, by introducing an image energy model that combines the ToF depth map and its corresponding active IR image as a guidance. The experiments show how the proposed approach effectively improves consistency of the final resulting depth map with respect to the active IR image, which is a proof of the increased correlation and matching between both images.

The guided inpainting method (guided belief propagation for depth completion [1]) based on IR image guide was used as preprocessing step to reconstruct missing depth values regions. Experimental results show that combining the belief propagation depth inpainting as a preprocessing and the proposed image energy model recovers edges and their continuity satisfactorily. The directional edge weights employed in the model provide the enhancement of depth discontinuities at object edges, also outperforming other existing conventional guided filtering methods. As a further work, this method will be extended to an other framework based on multi-sensors fusion, which could also increase the depth accuracy and lead to increase the resolution.

Acknowledgments. This work was partially supported by Analog Devices, Inc. and by the Agencia Valenciana de la Innovacion of the Generalitat Valenciana under program "Plan GEnT. Doctorados Industriales. Innodocto".

References

1. Achaibou, A., Sanmartín-Vich, N., Pla, F., Calpe, J.: Guided depth completion using active infrared images in time of flight systems. In: Pertusa, A., Gallego, A.J., Sánchez, J.A., Domingues, I. (eds.) Pattern Recognition and Image Analysis: 11th Iberian Conference, IbPRIA 2023, Alicante, Spain, 27–30 June 2023, Proceedings, pp. 323–335. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-36616-1_26
2. Ahmed, Z., Shahzad, A., Ali, U.: Enhancement of depth map through weighted combination of guided image filters in shape-from-focus. In: 2022 2nd International Conference on Digital Futures and Transformative Technologies (ICoDT2), pp. 1–7. IEEE (2022)
3. Diebel, J., Thrun, S.: An application of Markov random fields to range sensing. Adv. Neural Inf. Process. Syst. **18** (2005)
4. Ferstl, D., Reinbacher, C., Ranftl, R., Rüther, M., Bischof, H.: Image guided depth upsampling using anisotropic total generalized variation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 993–1000 (2013)
5. Gong, X., Liu, J., Zhou, W., Liu, J.: Guided depth enhancement via a fast marching method. Image Vis. Comput. **31**(10), 695–703 (2013)
6. Gu, S., Xie, Q., Meng, D., Zuo, W., Feng, X., Zhang, L.: Weighted nuclear norm minimization and its applications to low level vision. Int. J. Comput. Vision **121**, 183–208 (2017)
7. Hui, T.W., Ngan, K.N.: Depth enhancement using RGB-D guided filtering. In: 2014 IEEE International Conference on Image Processing (ICIP), pp. 3832–3836. IEEE (2014)
8. Izadi, S., et al.: Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, pp. 559–568 (2011)
9. Jiang, Z., Hou, Y., Yue, H., Yang, J., Hou, C.: Depth super-resolution from RGB-D pairs with transform and spatial domain regularization. IEEE Trans. Image Process. **27**(5), 2587–2602 (2018)
10. Kwon, H., Tai, Y.W., Lin, S.: Data-driven depth map refinement via multi-scale sparse representation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 159–167 (2015)
11. Liu, J., Gong, X.: Guided depth enhancement via anisotropic diffusion. In: Huet, B., Ngo, C.-W., Tang, J., Zhou, Z.-H., Hauptmann, A.G., Yan, S. (eds.) PCM 2013. LNCS, vol. 8294, pp. 408–417. Springer, Cham (2013). https://doi.org/10.1007/978-3-319-03731-8_38
12. Liu, W., Chen, X., Yang, J., Wu, Q.: Robust color guided depth map restoration. IEEE Trans. Image Process. **26**(1), 315–327 (2016)
13. Lu, S., Ren, X., Liu, F.: Depth enhancement via low-rank matrix completion. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3390–3397 (2014)
14. Or-El, R., Rosman, G., Wetzler, A., Kimmel, R., Bruckstein, A.M.: RGBD-fusion: real-time high precision depth recovery. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5407–5416 (2015)

15. Shen, Y., Li, J., Lü, C.: Depth map enhancement method based on joint bilateral filter. In: 2014 7th International Congress on Image and Signal Processing, pp. 153–158. IEEE (2014)
16. Valsesia, D., Fracastoro, G., Magli, E.: Deep graph-convolutional image denoising. IEEE Trans. Image Process. **29**, 8226–8237 (2020)
17. Yang, S., Liu, J., Fang, Y., Guo, Z.: Joint-feature guided depth map super-resolution with face priors. IEEE Trans. Cybernet. **48**(1), 399–411 (2016)
18. Yi, K., Zhao, Y., Lei, Y., Pan, J.: Depth enhancement with improved inpainting order and smoothing method. J. Phys. Conf. Ser. **1187**, 042065 (2019). IOP Publishing
19. Zhang, L., et al.: Depth enhancement with improved exemplar-based inpainting and joint trilateral guided filtering. In: 2016 IEEE International Conference on Image Processing (ICIP), pp. 4102–4106. IEEE (2016)
20. Zhang, X., Wu, R.: Fast depth image denoising and enhancement using a deep convolutional network. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2499–2503. IEEE (2016)
21. Zhang, X., Ye, P., Xiao, G.: Vifb: a visible and infrared image fusion benchmark supplementary document
22. Zhu, J., Zhang, J., Cao, Y., Wang, Z.: Image guided depth enhancement via deep fusion and local linear regularization. In: 2017 IEEE International Conference on Image Processing (ICIP), pp. 4068–4072. IEEE (2017)