MMSP 2019

Dynamic Guidance For Depth Map Restoration

Ran Zhu, Shengju Yu, Xiaoyu Xu, Li Yu*
[1] School of Electronic Information and Communications Huazhong University of Science and Technology, Wuhan, China；

ABSTRACT

The guidance of color images greatly improves the restoration accuracy of the depth map. However, due to the incomplete texture structure consistency between the color image and the depth map, and the limitations of L^2 based model, static guidance tends to cause texture copy artifacts and blurring depth discontinuities, which seriously deteriorates the quality of results. To tackle those problems, we propose a dynamic guidance model which can adaptively adjustment itself and continuously optimize in iteration.The proposed method can significantly alleviate the negative impact of the incomplete consistency and makes full use of the guidance information to restore fine texture at the same time. Moreover, a novel edge correction mechanism is designed to filter out incorrect depth information around boundaries, ensuring the correction of edges. Experiment results demonstrate that the proposed method can make best performance.

CONVENTIONAL MODEL

For the purpose of obtaining high quality depth maps, many effective methods have been proposed. Considering that color images and depth maps are different representations of the same scenes, many color-guided depth restoration methods have been proposed. AR model is one of the typical representatives.

$$\mathbf{D^H} = \underset{\mathbf{D^n}}{argmin}\{\sum_{i\in\Omega}||D_i^n - D_i^0||_2^2 + \gamma\sum_{i\in\Omega}(D_i^n - \sum_{j\in N(i)} w_{ij}D_j^n)^2\} \quad (1)$$

with

$$w_{ij} = \frac{1}{k_p} w_{ij}^D \cdot w_{ij}^I \quad (2)$$

These methods apply depth information and the consistency between color images and depth maps, they can alleviate texture copy artifacts and blurring depth discontinuities partly by introducing depth information into the guidance weight model. But they still use static guidance which has limited ability. What's more, those methods are based on L^2 . L^2 based data term is suitable for suppressing Gaussian noise. But L^2 based regularization term is proved to be unstable to outliers. The restoration results are also unsatisfactory when the upsampling factor is large(e.g.,8×) or large holes exist in the captured depth maps.

THE DYNAMIC GUIDANCE MODEL

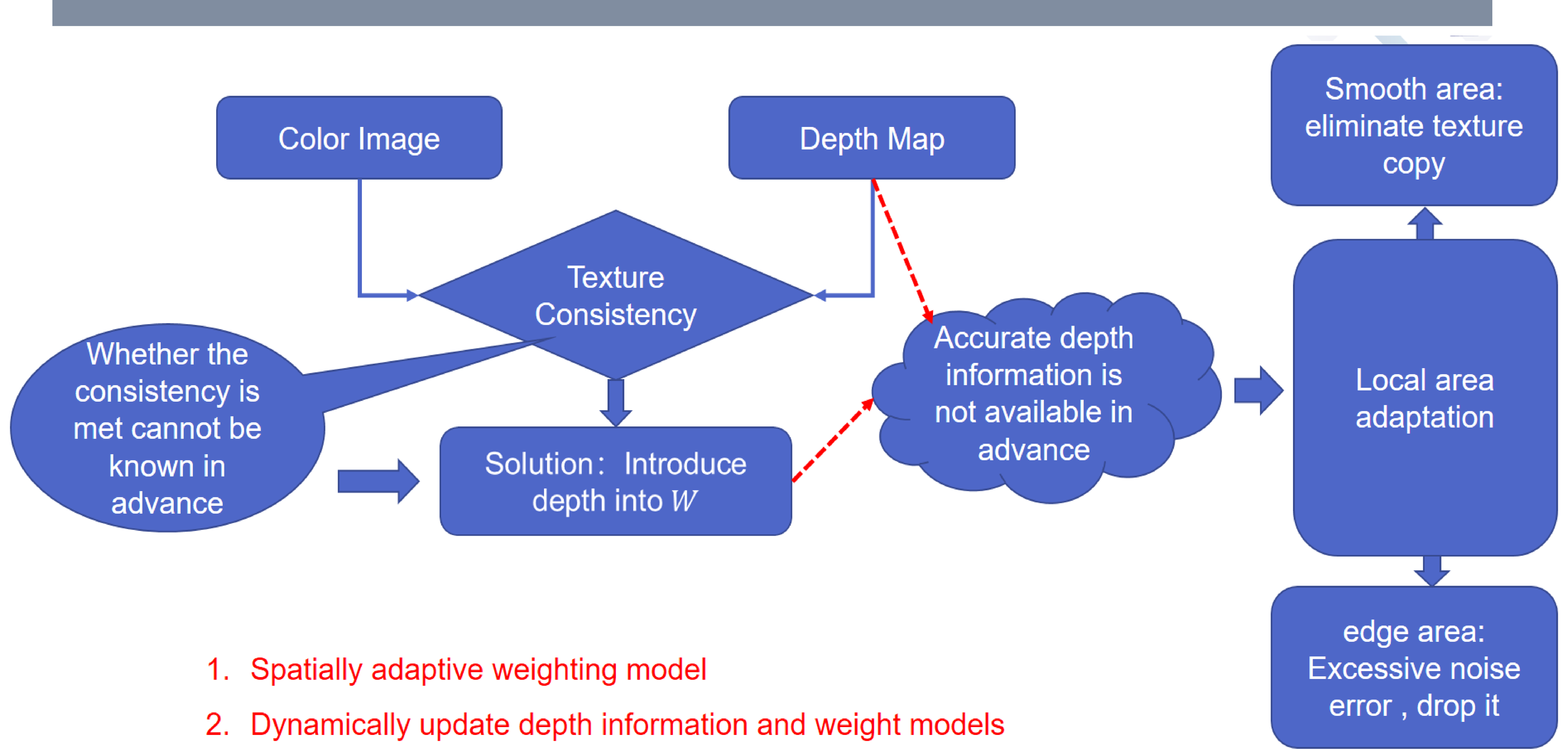


Fig.1 Basic thinking map

We propose a dynamic restoration model and reformulate the proposed model with a robust M-estimator as regularization term while using a L^2 based data term. Two terms are balanced by γ.

$$\mathbf{D^S} = \underset{\mathbf{D^n}}{argmin}\{\sum_{i\in\Omega}(1-C_i)||D_i^n - D_i^0||_2^2 + \gamma\sum_{i\in\Omega}\sum_{j\in\mathbf{N}(i)} w_{ij}\Phi((D_i^n - D_j^n)\cdot C_j)\} \quad (3)$$

The input D^0 is the initial depth map derived from bicubic interpolation based on the captured distorted depth map D^L. C is edge pixel confidence matrix. The penalty function $\Phi(x)$ used in our regularization term is formulated as:

$$\Phi(x) = 2\sigma^2(1 - exp(-x^2/2\sigma^2)) \quad (4)$$

The advantages of applying this penalty function is twofold. First, $\Phi(x)$ has an attribute, i.e. $\lim_{x\to\infty}\Phi(x)' = 0$, which means $\Phi(x)$ is a M-estimator and it has been proved that M-estimator is robust and preserves edges well. Second, it suddenly saturates, making it to be an approximation of L^0.

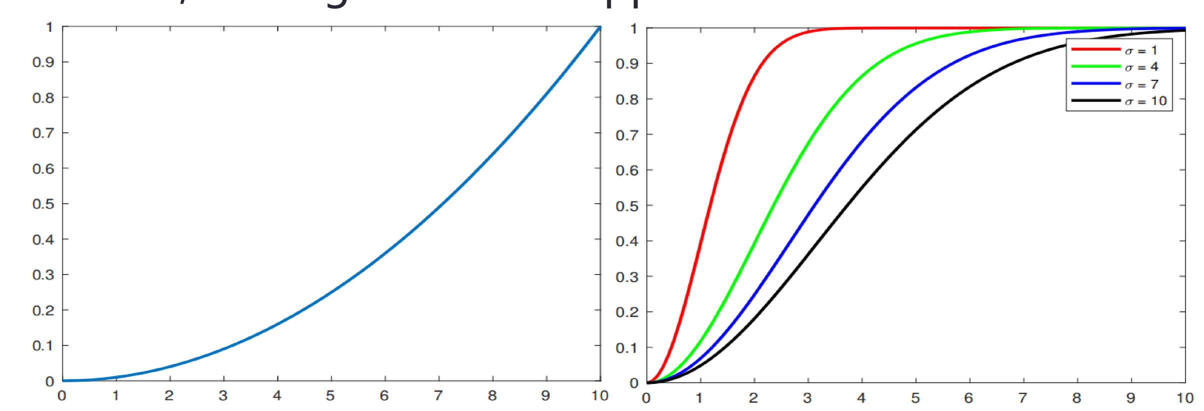


Fig.2 *L^2-based* *Ours*

Fig.2 shows curves of L^2 penalty function and ours. Our penalty function can saturate quickly, keeping sharp edges. Besides, introducing a proper weight model can alleviate texture copy artifacts and blurring depth discontinuities. Static guidance can't reflect intrinsic connection between color images and depth maps. In this work, we adaptively construct weight according to the local smoothness of depth maps and update it dynamically. Our local smoothness measurement method is defined as:

$$\zeta_i = \frac{min_{j\in\mathbf{N}(i)}D_j^s}{max_{k\in\mathbf{N}(i)}D_k^s} \quad (5)$$

where ς_i indicates the smoothness of point i. D^s the local average of the initial depth map D^0. Based on the proposed local depth smoothness model, our weight model is adaptively constructed. In smooth region:

$$w_{ij} = w_{ij}^s \times (\eta \times w_{ij}^b \times w_{ij}^p + (1-\eta)\times w_{ij}^d) \quad (6)$$

while in edge area:

$$w_{ij} = w_{ij}^p \times w_{ij}^s \quad (7)$$

w^s, w^b, $w^p$and w^d measure the similarity relationship from four dimensions: space, color, local structure, and depth. Experiments are performed to prove the effectiveness of the adaptive weight model and results are shown in Fig.3.

Fig.3 color *traditional weight* *Ours*

Because the initial depth map D^0 is derived from bicubic interpolation, making D^0 unreliable in edge region. In order to preserve sharp edges, we further propose an edge correction mechanism. It consist of two criteria. The first is the differences between the current depth and the maximum, minimum depth around it, which are calculated as following:

$$\Delta_i^l D = D_i^s - min_{j\in N(i)}D_j^s \quad (8)$$

$$\Delta_i^h D = max_{k\in N(i)}D_k^s - D_i^s$$

In the edge region, if $|\Delta_i^l D - \Delta_i^h D|$ is small, the confidence of i will be set to 0. For the rest points q around edge, the second criteria is applied. If absolute difference ΔD_q between the largest and smallest depth in the neighborhood of q is greater than a particular threshold θ, the confidence of point q is set to κ.

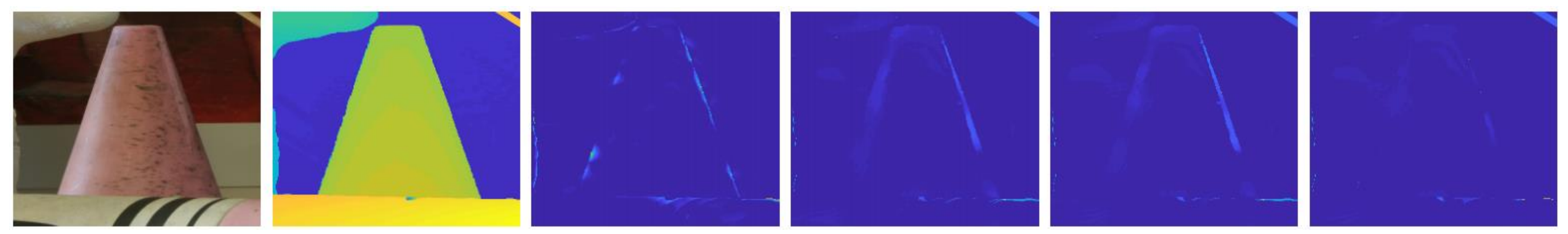


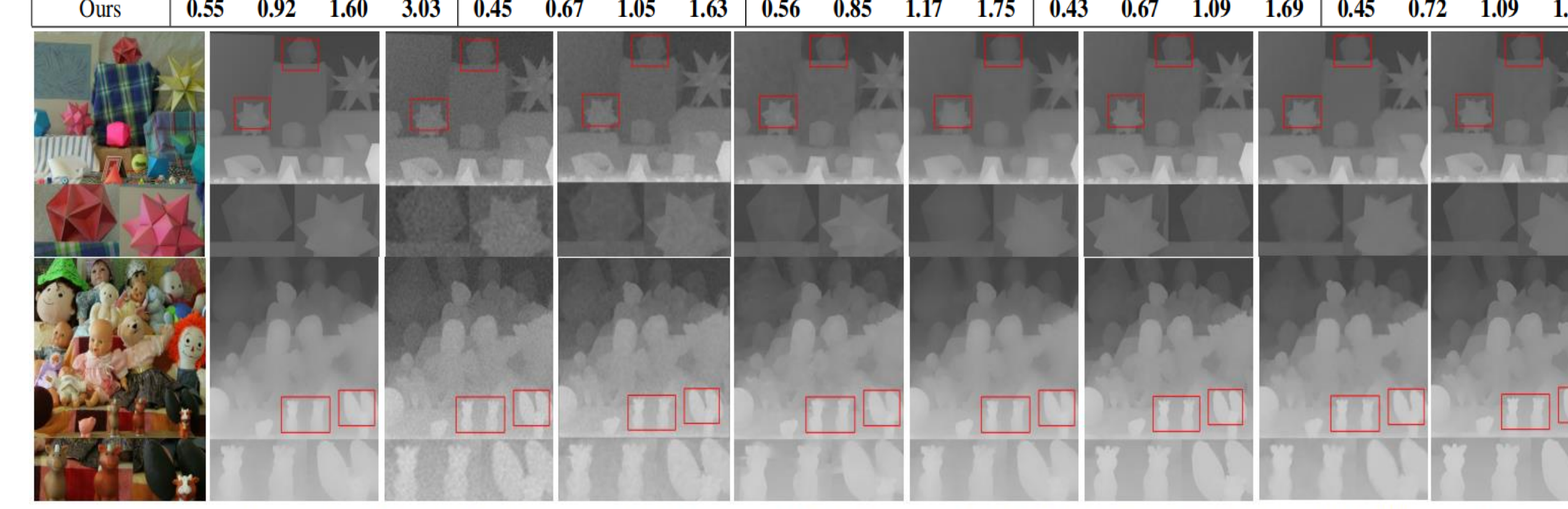
Fig.4. Demonstration of the effectiveness of edge correction mechanism . From left to right respectively: Color ,roundtruth and the rest are residual maps of: AR ,RGDR ,our result without edge correction mechanism and our result with edge correction mechanism.

RESULT

In order to comprehensively evaluate the effectiveness of the proposed method, we conducted detailed comparative experiments on the simulated ToF database, the simulated kinect database and the real database ToFMark. The proposed method is compared with several state-of-art methods.

Quantitative comparison on noisy simulated ToF dataset.results are measured in MAE

	Art				Book				Dolls				Moebius				Reindeer			
	2×	4×	8×	16×	2×	4×	8×	16×	2×	4×	8×	16×	2×	4×	8×	16×	2×	4×	8×	16×
JBU [5]	1.59	2.06	3.11	5.08	0.82	1.24	2.04	3.19	0.81	1.2	1.98	3.05	0.89	1.28	2.05	3.18	0.95	1.36	2.24	3.67
JGU [7]	1.33	1.81	2.9	4.92	0.79	1.24	2.05	3.17	0.8	1.23	2.01	3.05	0.82	1.25	2.03	3.15	0.91	1.37	2.26	3.68
TGV [1]	0.95	1.69	2.16	4.74	0.61	0.88	1.21	2.19	0.66	0.96	1.38	2.88	0.57	0.77	1.23	2.74	0.61	0.85	1.3	3.41
AR [8]	1.17	1.7	2.93	5.32	0.98	1.22	1.74	2.89	0.97	1.21	1.71	2.74	0.95	1.2	1.79	2.82	1.07	1.3	2.03	3.34
RGDR [14]	0.73	1.09	1.81	3.53	0.58	0.81	1.20	1.85	0.66	0.91	1.26	1.84	0.55	0.79	1.21	1.85	0.59	0.82	1.22	2.22
WLS [10]	1.25	1.73	2.59	4.01	0.74	1.1	1.45	2.13	0.85	1.21	1.68	2.13	0.8	1.18	1.67	2.27	0.84	1.15	1.58	2.43
Ours	**0.55**	**0.92**	**1.60**	**3.03**	**0.45**	**0.67**	**1.05**	**1.63**	**0.56**	**0.85**	**1.17**	**1.75**	**0.43**	**0.67**	**1.09**	**1.69**	**0.45**	**0.72**	**1.09**	**1.97**



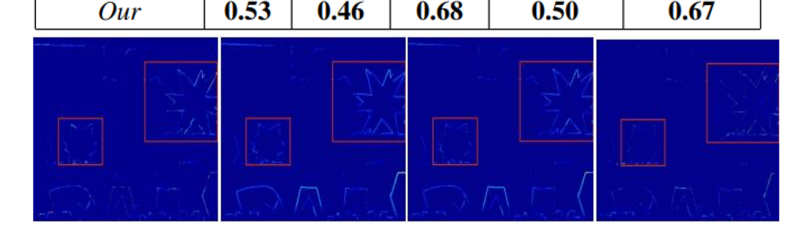
(a) Color (b) Ground truth(c) Bicubic (d) JGU [7] (e) WLS [10] (f) TGV [1] (f) RGBD [14] (g) AR [8] (h) *Ours*

Fig.5. 8× upsampling results of ToF-like depth image

Quantitative comparison on simulated kinect dataset (left)and ToFMark (right).

	Art	Book	Dolls	Moebius	Reindeer
Bicubic	0.90	0.61	0.95	0.91	0.76
IMLS [16]	0.91	0.58	0.68	0.72	0,82
CLMF [15]	1.01	0.63	0.74	0.64	0.94
RGDR [14]	0.56	0.49	0.74	0.52	0.70
AR [8]	0.58	0.53	**0.68**	0.75	0.69
Our	**0.53**	**0.46**	**0.68**	**0.50**	**0.67**

	Bicubic	JBU [5]	JGF [7]	TGV [1]	RGDR [14]	AR [8]	*Ours*
book	16.23	16.03	17.39	*12.80*	13.31	14.37	*12.80*
shark	17.78	18.79	18.17	15.53	15.58	16.27	*15.06*
devil	16.66	27.57	19.02	14.97	14.63	15.41	*14.38*



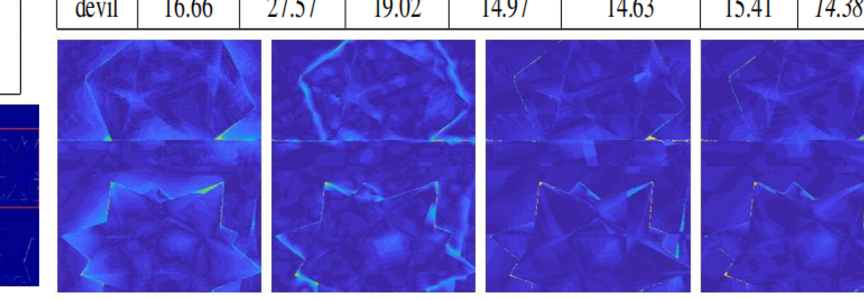


Fig.6. restoration results of Kinect-like depth image and ToFMark dataset

Visual comparison proves that our method can alleviate texture copy artifacts and blurring depth discontinuities effectively. Quantitative results show that the proposed method restores high quality depth map with the lowest mean absolute error.

CONCLUSION

A novel model that can handle different kinds of depth map degradation is proposed. This model employs a robust penalty function, constructs weight dynamically and uses a novel edge correction mechanism. The proposed weight model subtly introduces depth information and dynamically updates the model, allowing depth plays full role in alleviating texture copy artifacts and blurring depth discontinuities. Meanwhile, this model eliminates the negative effects of noise in depth map. Moreover, the proposed edge correction mechanism can remove error depth information from edge regions, preserving sharp edges well. All experiment results demonstrate that the proposed method can make best performance