3D Image Correction by Hilbert Huang Decomposition

Chung-Te Li, *Student Member, IEEE*, Yen-Chieh Lai, Chien Wu, Sung-Fang Tsai, *Student Member, IEEE*, and Liang-Gee Chen*, *Fellow, IEEE, Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan, R.O.C*

Abstract—**This paper presented a novel correction for 3D image signals by Hilbert Huang decomposition. Hilbert Huang decomposition is applied for dividing edges of the textures and edges of the objects. Depth map is corrected by regrouping pixels with edges of the objects. The proposed correction outperforms conventional methods especially at the boundaries of the objects.**

Index Terms—**3D image post-processing, depth map, Hilbert Huang decomposition, depth correction**

I. Introduction

Color images and corresponding depth maps generate stereoscopic or multiscopic views for three-dimensional TV. The quality of depth map significantly affects the visual qualty of 3D experiences. Marr and Poggio [1] have mentioned that the depth (disparity) of stereopsis should vary smoothly over the whole perceived image, except at the boundaries of the objects since the distance between near points on a given surfaces generally varies continuously. Therefore, the boundaries of the depth map must be aligned with the boundaries of the objects, and the depth map should be smooth at regions which do not belong to the boundaries of the objects. Researchers have developed some post-processing techniques for correcting depth map, such as joint bilateral filtering [2], and correction by image segmentation [3].

However, these kinds of methods correct depth map with both the boundaries (edges) of objects and textures. Edges of textures introduces undesired discontinuities within objects, which result worse stereospis. To conquer this problem, we propose a novel idea to extract the boundaries of the objects by the analysis of the spatial frequency from Hilbert Huang decomposition. Then, an automatic depth correction is applied based on the boundaries of the objects.

The rest of this paper is organized as follows: Section II reviews Hilbert-Huang decomposition and its characteristics. Section III presents our proposed detection for the boundaries of the objects and the correction of depth map for 3D images. Section IV presents the simulation results and our conclusions.

II. Theoretical Foundation

For nonstationary and nonlinear stochastic signals, such as tides and electrocardiography, Hilbert–Huang transform (HHT) provides a way to decompose them into several intrinsic mode functions (IMFs) by its frequency, and then instantaneous frequency data are obtained [4].

To decompose a signal into intrinsic mode functions (IMFs), empirical mode decomposition (EMD) is used. For each nonstationary and nonlinear stochastic signal, it can be decomposed into small number of components, which form a set of intrinsic mode functions. The decomposition is lossless. That means the original signal can be ideally reconstructed by summing all of the IMFs and the residual signal. For each IMF, Hilbert spectral analysis is applied to determine the corresponding frequency. Hilbert spectral transform obtains the best fit of a sinusoid to each IMF at every point in temporal domain, and identifies an instantaneous frequency, along with its associated instantaneous amplitude. Thus, a given nonstationary and nonlinear stochastic signal can be decomposed with a variety of transient sinusoids by HHT.

III. Proposed Object Boundaries Detection and Depth Correction

We will present the details of the detection for the boundaries of the objects and the automatic correction of depth map for 3D images in the following.

A. Object Boundaries Detection

The fundamental concept of the proposed detection is that the spectral frequency of textures is generally higher than objects. Since the separation of the textures and the objects is generally a fuzzy system on spatial frequency, there exist no hard thresholds on spatial frequency for the separation. Due to this constraint, Hilbert–Huang decomposition is applied because it is adaptive and has good locality compared to conventional frequency decomposition such as Fourier Transform.

Empirical mode decomposition is applied for separating the 3D image into signal for textures and signal for objects, respectively. The decomposed IMFs with relatively higher frequency are recognized as signals for textures, and then are filtered in order to generate the signal for objects. Thus, the transition points of the signal for objects can provide the information to determinate the boundaries of the objects.

In detail, to generate the signal for the objects, the color component of 3D image, denoted as $I(x,y)$, is separated into row and column signals, $Row_{I,k}(x)$ and $Col_{I,l}(y)$, where

$$Row_{I,k}(x) = I(x,k), \tag{1}$$

$$Col_{I,l}(y) = I(l,y). \tag{2}$$

Then, Hilbert-Huang decomposition is applied for each $Row_{I,k}(x)$ and $Col_{I,l}(y)$. For convenience, we denote the j-th IMF of a generic function $h(x)$ as $IMF(j, h(x))$, where

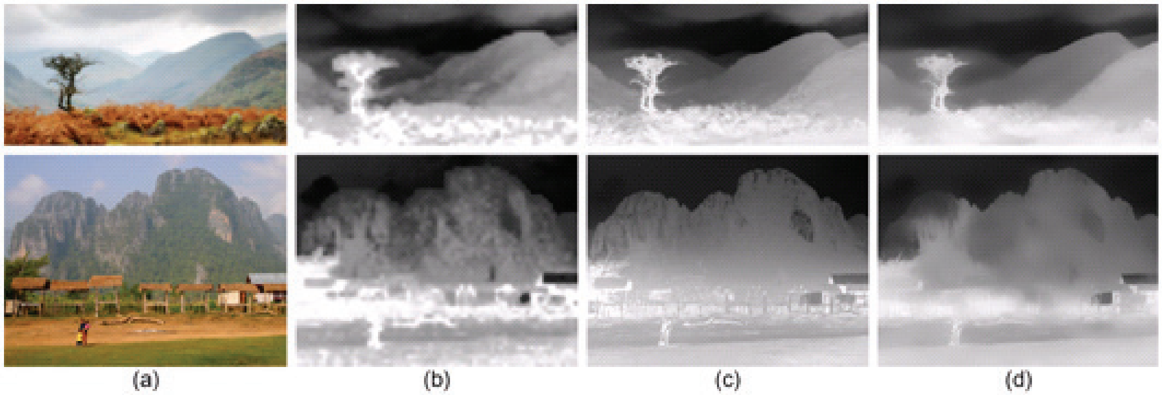

Fig. 1. Correction for depth from dark channel prior
(a) original image (b) depth from dark channel prior (c) corrected depth by cross-bilateral filter (d) corrected depth map by proposed algorithm (p=6)

$$h(x) = \sum_{j=0}^{\infty} IMF(j, h(x)). \tag{3}$$

It is important to notice that smaller j implies lower frequency for the IMFs. Therefore, a threshold p for the order of the IMFs is predefined to separate signal for textures (denoted as *Tex(.)*) and signal for objects (denoted as *Obj(.)*) for the row or column components $Row_{I,k}(x)$ or $Col_{I,l}(y)$ of a given 3D image, where

$$Tex(Row_{I,k}(x)) = \sum_{j=p}^{\infty} IMF(j, Row_{I,k}(x)), \tag{4}$$

$$Tex(Col_{I,l}(y)) = \sum_{j=p}^{\infty} IMF(j, Col_{I,l}(y)), \tag{5}$$

$$Obj(Row_{I,k}(x)) = \sum_{j=0}^{p-1} IMF(j, Row_{I,k}(x)), \tag{6}$$

$$Obj(Col_{I,l}(y)) = \sum_{j=0}^{p-1} IMF(j, Col_{I,l}(y)), \tag{7}$$

The energy map $E(x,y)$ is then defined for the possibility that pixel (x,y) belongs to the boundaries of the objects since the boundaries are important for stereopsis, where

$$E(x,y) = \max\{\frac{\partial Obj(Row_{I,k}(x))}{\partial x}, \frac{\partial Obj(Col_{I,l}(y))}{\partial y}\} \tag{8}$$

B. Automatic Correction for Depth

Since the discontinuities of the depth map should aligned at the boundaries of the objects, we try to refine the depth map by the energy map $E(x,y)$. We try to regroup the pixels based on the possibility of the boundary. The procedures of the regrouping are shown in the following.

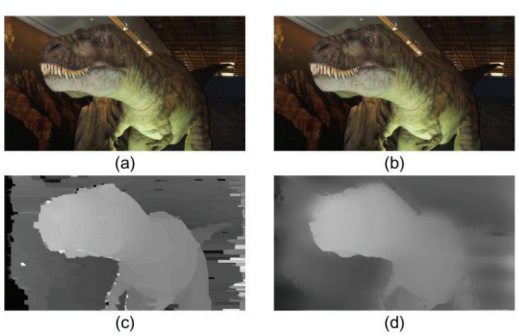

Fig. 2. Correction for depth from stereo matching (a) (b) original stereo pair (p.s. (a):left (b):right) (c) matched depth for left image (d) corrected depth for left image by the proposed algorithm (p=6)

First, a complete graph G is built based on the color component of the given 3D image $I(x,y)$. $I(x,y)$ is then segmented into small blocks, such as 4x4 or 8x8 rectangular blocks. Each block is denoted by a vertex of G, and the weight of each edge are defined by 1) the mean color difference between the two blocks if they are adjacent on $I(x,y)$; 2) infinity if the two blocks are not adjacent on $I(x,y)$. Second, a minimum spanning tree T of G is built. Third, the spanning tree T is divided into segments by removing edges based on $E(x,y)$. Edges located on pixels with higher $E(x,y)$ are removed. Finally, the corrected depth is initially assigned by averaging the original depth within each segment. Then, a joint-bilateral filter for the corrected depth and the color component I is applied for refining the blocky boundaries of the depth map.

IV. Comparisons and Conclusions

Figure 1 and figure 2 show the results of the correction for depth maps from dark channel prior [5] and stereo matching [6]. The original depth maps suffer from the neighborhood kernel for extracting the dark channel and the occlusions in stereo matching, respectively. The proposed method does provide better depth maps, especially at the boundaries of the objects.

References

[1] D. Marr; T. Poggio, "Cooperative Computation of Stereo Disparity," *Science*, *New Series*, Vol. 194, No. 4262, pp. 283-287, 15 Oct, 1976

[2] Gangwal, O.P.; Djapic, B.; , "Real-time implementation of depth map post-processing for 3D-TV in dedicated hardware," *Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on* , pp.173-174, 9-13 Jan. 2010

[3] Hai Tao; Sawhney, H.S.; , "Global matching criterion and color segmentation based stereo," *Applications of Computer Vision, 2000, Fifth IEEE Workshop on* , pp.246-253, 2000

[4] N. E. Huang, Z. Shen, S. R. Long, M. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. of The Royal Scoiety*, vol. 454, no. 1971, pp. 903–995, 8 Mar. 1998.

[5] Kaiming He; Jian Sun; Xiaoou Tang; , "Single image haze removal using dark channel prior," *Computer Vision and Pattern Recognition, 2009. IEEE Conference on* , pp.1956-1963, 20-25 June 2009

[6] Jian Sun; Nan-Ning Zheng; Heung-Yeung Shum; , "Stereo matching using belief propagation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.25, no.7, pp. 787- 800, July 2003.