A New method of Image depth extraction based on SVM

Mingjun Li, Zhipeng Fan, Ying Lu

[1] School of Computer and Information Engineering, Harbin University of Commerce, Harbin 150028,China

Abstract. 3DTV gets more and more popular. However, with the growing number of stereoscopic display devices, the 3D content for display is obviously insufficient. Current methods for generating single camera stereoscopic videos are based on creating a depth map. In this paper, an efficient image depth extraction method which utilizes the support vector machine(SVM) process is proposed. The label is established from the true depth of different videos， and the vector feature, haze is utilized as the feature vector. Through SVM training, the depth map of predicted video is obtained. The experimental results show that the algorithm is effective.

Keywords: depth map, feature vector, support vector machine

1 Introduction

Currently, most of the 2D-3D video conversion technology via dense depth maps for each frame of the sequence using depth-image-based rendering (DIBR). As shown in Figure 1, which involves the projection of a viewpoint into another view. This projection is actually based on warping[2].Based on this theory, depth extraction is a key issue for 2D-3D video conversion, becoming an important direction in the 2D-3D video conversion technology.

Zhou et al proposed based on DFD（depth from defocus）Edge ambiguity model established by the Gaussian gradient method [6]，improved in Edge treatment effect and Scene depth calculation accuracy by this method. But little defocus comparison image result is not satisfactory. In order to compensate for this deficiency，proposed a kind of one clue-based, other clues as a supplement method for depth extraction. Y.Feng et al proposed based on the defocus combine with optical flow depth estimation algorithm is proposed [7].However, the way which multi clue combine together is unknown., the accuracy of the final result by the judge of a clue to get the wrong results will be affected. For instance, when the depth map is estimated by defocus，a large area of sky background will be misidentified as clear foreground，it will have a certain influence on the effect of the final depth map.

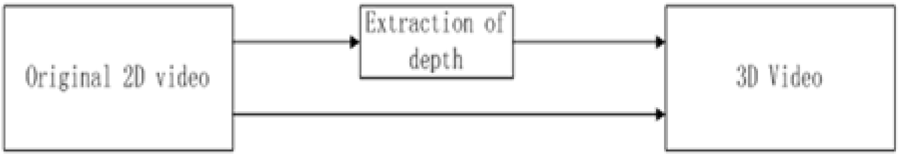


Fig.1 The 3D image generation system block diagram based on DIBR

2 The Depth Extraction Methods Based on Support Vector Machine

2.1 Depth extraction based on support vector machine

The structure which depth extraction method of image as show Figure 2. System processing is divided into data preparation and SVM learning which will be down by two step. In the data preparation stage，Classification labels obtained by processing the depth map，depth characteristic is extracted to make feature vector in the original image. At the SVM training stage. Firstly, the kernel function and its parameters should be determined, then according to the classification label obtained by the data preparation phase and feature vector training SVM to get classification model.； Ultimately，the depth map is obtained by the model and feature vector.

2.2 Classification and acquisition of Labeling

In this article depth forecast is based on macro-block，by prediction the depth of each macro-block, the depth map of each frame is achieved. For a random macro-block, the mean value of the depth as the depth of the macro-block $d_{patch(i)} = \sum_{(m,n)\in patch(i)} \frac{d(m,n)}{N^2}$ .In addition，when the two-dimensional image is observed by human eyes，the depth of different scene is classified into different depth value automatically. In this projected，the depth of each macro-block is classified into one of L levels. The number of level as the label in the SVM classifier. $label_{patch(i)} = floor\left(d_{patch(i)} / l \right) + 1$. After considering the subjective feelings and deal with complexity，in this paper the scale of macroblock is 8*8, and the level is 16.

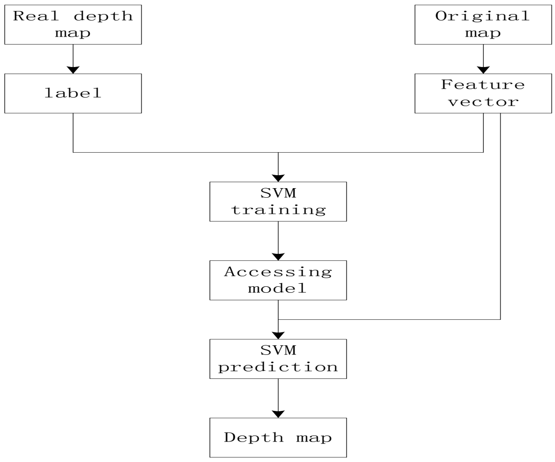


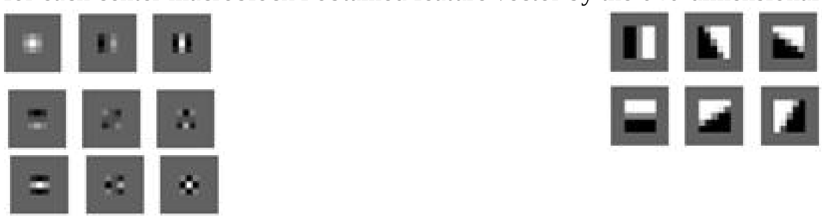
Fig. 2. Block diagram of the image depth extraction method based on SVM

2.3 Extract Feature Vector

In the image，Different depths of the scene exhibits a texture change, texture gradient, haze and other characteristics are also different.. Therefore，In this study, a given set of filtered is used (figure 4). To obtain the depth features of the image according to these three clues. The front nine template is a template of a group of 3*3 laws proposed by the Laws[8], can be used to detect image changes in local mean edge, spots, texture information. The back six 5*5 templates are directional detectors, can be used to extract image texture gradient information. For YUV test sequence, most texture changes and texture gradient information exists in the Y channel. Atmospheric scattering more obvious in the low frequency part of the U, V-channel, therefore a filter is used acting on the Y channel of the image to obtain texture changes and texture gradient. Take advantage of the first Laws template acting on the image U, V two channels to obtain haze. The 17 filters, $F_n(x,y)$ ($n=1,\ldots,17$) Convolution with the image, and seeking the absolute energy and the square of the energy,

$$E_i(n)=\sum_{(x,y)\in patch(i)}|I(x,y)*F_n(x,y)|^k \quad k=1,2$$

for each center macroblock i obtained feature vector by the 646-dimensional data .



(a) Laws template (b) the direction of the detector

3 Experimental Results

The depth map of the fiftieth frame of the test sequence kendo_1, newspaper, lovebird were extracted. Data preparation was completed by using matlab，Libsvm package was used to training and prediction. The experimental results shown in Figure 5，From left to right are the original image, and the real depth map, depth map after quantization, and a depth map after predicted. The accuracy of experimental results were 82.55%、85.7992%、94.0592%. Experimental results show that，The proposed algorithm can get better results in these types of cases.

a) kendo_1

b) newspaper

c) lovebird

References

1. Shaojie, Z., Terence, S.: Defocus Map Estimation from a Single Image . Pattern Recognition, vol.5, pp. 1852-1858 (2011)
2. Yue F., Jayaseelan, J., Jianmin, J.: Cue Based Display Estimation For Possible 2D to 3D Video Conversion. Visual Information Engineering. IET International Conference, pp(384-388). Bangalore, India(2006)
3. Ashuotsh, S., Jamie, S., Andrew, Y.N.: Depth estimation using monocular and stereo cues. International Joint Conference on Artificial Intelligence (IJCAI), pp(2197-2203). Hyderabad, India. (2007)
4. McMillan, Jr.: An image-based approach to three-dimensional computer graphics, PhD thesis, Chapel Hill, NC, USA (1997)
5. Christoph, F.: A 3D-TV Approach Using Depth-Image-Based Rendering (DIBR). Visualization, Imaging, and Image Processing (VIIP), pp(482-487). Benalmadena, Spain(2003).