Indian Journal of Science and Technology, *Vol 9(28), DOI: 10.17485/ijst/2016/v9i28/97877, July 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

2D to 3D Conversion using Key Frame Extraction

M. Kavitha* and E. Kannan

Department of Computer Science and Engineering, Vel Tech University,
Chennai - 600062, Tamil Nadu, India; mkavi277@gmail.com, ek081966@gmail.com

Abstract

Background/Objectives: Extracting 3D information from two dimensional images is tedious process and challenging task. In this paper, a novel 2D to 3D conversion model is presented to convert monoscopic 2D images into 3D stereoscopic images. **Method/Satistical Analysis:** The depth information about key frames is extracted from the 2D images. Then both foreground and background objects are extracted using background subtraction algorithm. From the generated region of interest both forward and backward motion pixels are extracted in the form of vectors. Gabor filter is applied to decompose high pass luminance and chrominance. Each pixels decomposed with gabor filter is estimated with its own sub band to model the depth information. **Findings:** Depth map associated with each sub band and its oriented pixels are mapped with the depth information of 2D images and this makes a realistic view of 2D images in 3D stereoscopic view. **Application/ Improvements:** The proposed model is experimentally tested on both right and left view and the results shows that the presented model is better than the traditional model.

Keywords: Background Subtraction Algorithm, Gabor Filter, Key Frames, Monoscopic Images, Stereoscopic Images

1. Introduction

3D videos are popular in today's world of entertainment. All the supportive media focuses on realistic 3D conversion from the 2D stereoscopic videos. Such methodologies are widely used in TV's and some wide range projectors. Various algorithms and models had been proposed so far to convert 2D to 3D. Such an attempts were made using two optimistic models.

Initial attempt was made using depth array conversion method. The method focuses on depth array conversion of back grounds from the 2D video and subtracting the active objects in each scene at every frame, this yield to obtain the depth information[1] from the video which requires heavy computation in real time. The next attempt was made using depth cue[1], depth level which was fully oriented on anaglyph model.

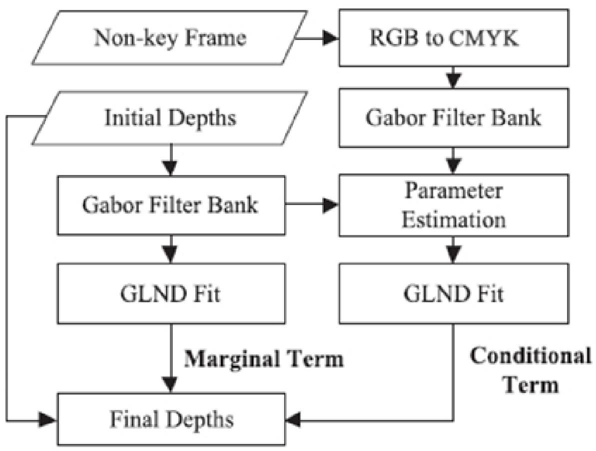
The model[1] utilizes the depth level of y dimensional axis coordinates for 3D conversion of foreground objects. Histogram equalization is randomly applied for separating the foreground objects from the background[1]. Each pixel at top most depth level of the foreground object is truncated and depth cue[2] is calibrated for better perception using binocular vision methodology along with Anaglyph which is modeled on CYAN to yield better results. Anaglyph is used to model the depth cue at various depth levels which automatically transforms each eye image using chromatic colors (red/cyan). The model is experimentally setup with the real time web cam and tested with the incoming image frames on live streaming.

2. Proposed model

The deployment unit consists of one single 20 fps; single shot webcam with RGB enabled color space rendering multiple frames to single frame. Initial and final frame are the key frame while all intermediate frames are non-key frames[3]. Non-key frames initially extracts CMYK color space model and Gabor filter is applied to all the sub bands to model depth information. Conditional terms of depth at initial phase are fixed to non-varying threshold of 50 and parameter estimation is used to estimate the sub bands of initial depths and backward depths from final frame. GLND (Generalized log-normalization distribution) fit is applied to both key frame and non-key frame

**Author for correspondence*

to extract depth information[4-6]. GLND fit helps the sub band to mesh Gabor filter responses. Finally the depth maps are estimated using key frame method by mapping both GLND fit at marginal and conditional terms. The overview of the proposed methodology is clearly denoted in Figure 1.



\s

Figure 1. Key frame and non-key frame depth information extraction.

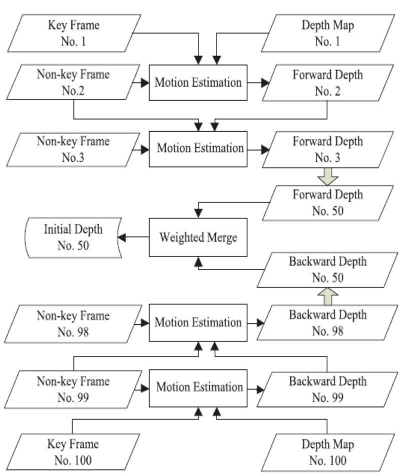


Figure 2. 3D frame formation from depth information.

First and final frame is taken into consideration and depth information extracted is stipulated using fixed threshold of its frame size. Then the forward depth and backward depth is given as the key sequence along with the non-key frames to estimate motion and its periodic interval. Finally the forward and backward depth sequences are merged in order to gain the key feature for 3D output sequence[7,8]. The deep illustration of the model is presented in the Figure 2.

4. Proposed Algorithm

4.1 Depth Information Extraction

Begin_function_Depth_information_extraction (key_frame,non_key_frame)
  Calibrate the streaming video
  Import the frames from buffer
  Convert RGB to CMYK
  Calculate the depth level of pixel
  Estimate the view for the depth level
  Feather the top most pixels
  Apply Gabor bank
  Extract the sub band response
  Apply GLND fit
  Fix conditional and marginal terms
  Separate the perception
  Generate Depth Cue for each Depth – Y Axis
  Apply polarized filter
  Generate final depth information
End

4.2 3D Conversion using Key Frame

Extraction

Begin_function_3D_conversion ()
  Calibrate the streaming video
  Import the frames from buffer
  Calculate the depth level of pixel
  Apply final depth information
  Merge forward and backward depth
  Estimate motion sequences
  Combine motion sequences
  Generate Video Sequence
End

5. Experimental Setup

The list of proposed model using anaglyph, Depth array conversion along with Depth key extraction is compared and tested in real time video data set using 2 desktop PC. The present model is deployed as browser plug-in and purely developed using JavaScript with HTML 5. An interactive dropdown is listed with two extensions; one is for webcam initialization and another one is for color model selection. The browser extension supports interactive video rendering and live conversion is initiated using webcam and user is prompted to select color model for active conversion of 2D view to 3D view in near real time. In order to validate the proposed model along with the present model, AVATAR video data set is used. The frame rate is at 20fps and varied upto 60fps at 720p to 1080p.

6. Results and Discussion

Figure 3 represents the proof of our proposed model. The validation of the model is clearly stated and Figure 4 states the efficacy of the proposed model when compared to other existing models.

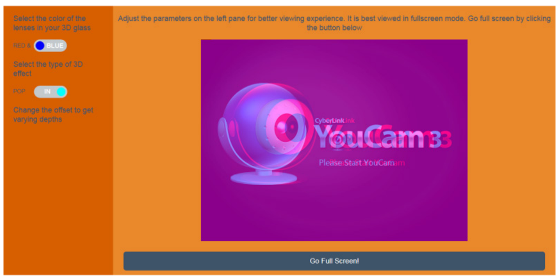


Figure 3. Results of the proposed model.

The proposed methodology is applied along with depth information and comparison graph is plotted. The results presented in Figure 6 shows that the proposed model is better than the other models.

Figure 4. Depth information of realistic 3D view.

Table 1. Performance rate in terms of execution time

Method	Execution Time Per Frame(Sec/ Frame) – 1920x 1080 pixels
Hybrid Depth Generation Algorithm[9]	2.21
Depth Array Conversion Method	2.02
Anaglyph Model	2.10
Key Frame Extraction Method	2.02

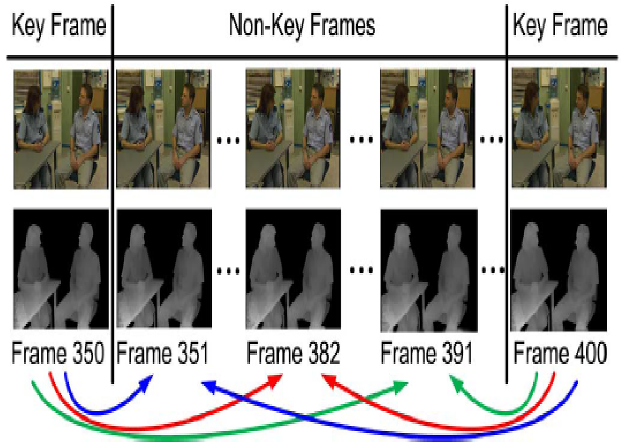


Figure 5. Key frame interpretation.

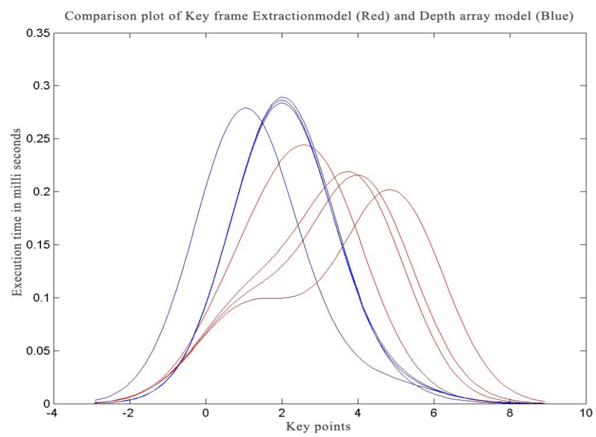


Figure 6. Performance graph of present model along with Depth Array Model.

7. Conclusion

Hence the proposed model is concluded by proposing an optimal model using depth information. Here, a deep illustration on depth information is clearly denoted in Figure 5. The experimental setup was carried out for validating the proposed model in terms of accuracy and efficacy (denoted in Table 1).

The present model gives better result than the existing model as denoted in Figure 6. The model is very cheap and yields better accuracy, when compared to depth array conversion, depth information extraction using key frames yields the better transition and conversion rate. The conversion rate is automatically interpreted and it is clearly stated in Figure 5. Further this can be enhanced by installing in the smartphones where the 2D video shot can be converted 3D videos with better conversion rate[10].

8. References

1. Kavitha M, Kannan E. A new Procedure for 2D to 3D conversion using Depth Array Conversion. American Journal of Applied Sciences. 2014; 11(4):694–99.
2. Feng Y, Ren J, Jiang J. Object-Based 2D-to-3D Video Conversion for Effective Stereoscopic Content Generation in 3D-TV Applications. IEEE Transactions on Broadcasting. 2011; 57(2):500–9.
3. Huang W, Cao X, Lu K, Dai Q, Bovik AC, “Toward Naturalistic 2D-to-3D Conversion,” in IEEE Transactions on Image Processing, 2015, 24(2), pp.724-33.
4. Konrad J, Wang M, Ishwar P, Wu C, Mukherjee D. Learning-Based, Automatic 2D-to-3D Image and Video Conversion. IEEE Transactions on Image Processing. 2013; 22(9):3485–96.
5. Cao X, Li Z, Dai Q. Semi-Automatic 2D-to-3D Conversion Using Disparity Propagation. IEEE Transactions on Broadcasting. 2011; 57(2):491–99.
6. Kavitha M, Kannan E. Improving the quality of generated Depth in a 2D to 3D conversion system. International Journal of Applied Engineering Research. 2015; 10(75).
7. Kavitha M, Kannan E. A New approach for converting 2D to 3D video with Image wrapping and Hole filling method. International Journal of Applied Engineering Research. 2014; 9(27).
8. Kavitha M, Kannan E, Pavithra C. An Efficient Edge Detection Algorithm for 2D-3D Conversion. International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC), Chennai. 2014. p. 434–36.
9. Lai Y-K, Lai Y-F, Chen Y-C. An Effective Hybrid Depth-Generation Algorithm for 2D-to-3D Conversion in 3D Displays. Journal of Display Technology. 2013 Mar; 9(3).
10. Kang M, Kim S, Choi EJ. A study on Smart Phone- based Shooting Device of 3D stereoscopic Image. Indian Journal of Science and Technology. 2015; 8(19).