INFLUENCE OF STEREOSCOPICAL CAMERA SETTINGS ON DEPTH MAP ESTIMATION

Ondřej Kaller

Doctoral Degree Programme (1), FEEC BUT

E-mail: xkalle00@stud.feec.vutbr.cz

Supervised by: Tomáš Kratochvíl

E-mail: kratot@feec.vutbr.cz

Abstract: This paper briefly summaries parameters, which could affects z axis resolution. It is also mentioned the process of depth map creation from stereo video pair. The commercial 3D camera system is compared with laboratory set of two independent cameras. There is also mentioned, whether is relevant to use hyper stereo base and benefits, which can brings this configuration.

Keywords: Camera resolution, stereo base, 3D content shoot distortion, Z axis resolution, stereo pair

1. INTRODUCTION

If you want to create a 3D content by classical stereoscopy capturing a scene with long distance targets, you run into problem with depth information of them. They would probably been fall to the background layer due to finite camera resolution [1]. In this article it is investigated, how increase of cameras resolution or hyper stereo base can improve the depth resolution of the 3D video.

2. DEPTH MAP ESTIMATION

Video format “2D+Depth” is quite progressive and due to it’s widespread in 3D video recording and animation technique it could be expected, that it will be used also in television broadcasting instead of today’s most popular, but very wasteful format side by side [2]. The bandwidth increases only about (5÷20) % in comparison with 2D content. Also multiview displays can be used to display motion parallax, because DIBR (Depth Image Based Rendering) [3] algorithm can estimate a variety of sights to one scene.

Depth map generation from stereoscopy images is not straightforward. First, it is necessary to find a fundamental matrix, which describes the epipolar geometry of stereo vision system [3]. There are a few different sophisticated algorithms to estimate it. It is necessary to find some corresponding points in both images of stereo pair previously. For e.g., SIFT algorithm (Scale-invariant feature transform) [4] can be used for this.

Then the estimation of disparity map is computed from pairs segmented the same distance areas, searched along epipolar line by correlation. Searching is done with subpixel accuracy.

3. RESOLUTION OF THE DEPTH

For long distance scene recording is useful to know up to which distance can be targets separated from background layer. Let’s leave aside focus problems and aberrations and assume infinite depth of field. Figure 1 shows, that finite pixel size projection to the image plane from some defined target distance achieves this horizontal shift between left and right image H_C (in case of parallel axes cameras configuration is same as stereo base d_C). L_{bmax}' is defined as a distance from cameras to target, in which the both cameras views same image (let us say shifted about one pixel each other).

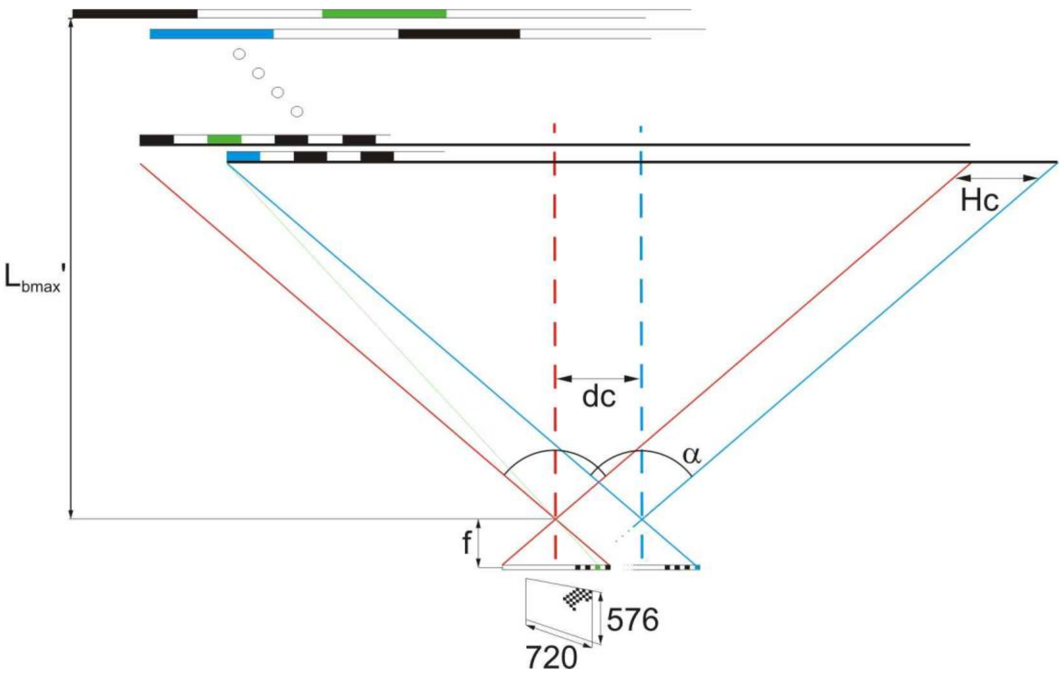


Figure 1: Depth resolution limitation caused by finite cameras resolution. Projection of pixel's size to the image plane.

The following equation (1) shows the dependence of maximum theoretical distance L_{bmax}' on image shift H_C, horizontal cameras resolution l, and view angle of camera's lens.

$$L_{b\max}' = \frac{\frac{W_{\max}}{2}}{\tan\left(\frac{\alpha}{2}\right)} = \frac{\frac{H_c \cdot l}{2}}{\tan\left(\frac{\alpha}{2}\right)} \qquad (1)$$

It can be seen, that the mentioned problem influence is increasing with angle and for widescreen long distance shooting. In case of $\alpha = 70.7°$ and cameras with standard definition resolution ($l \times r$ = 720 x 576 pixels) is L_{bmax}' equal 23.04 m. Luckily, as it was mentioned, the depth map estimation algorithm uses subpixel accuracy, so it could increase the depth almost ten times.[4]

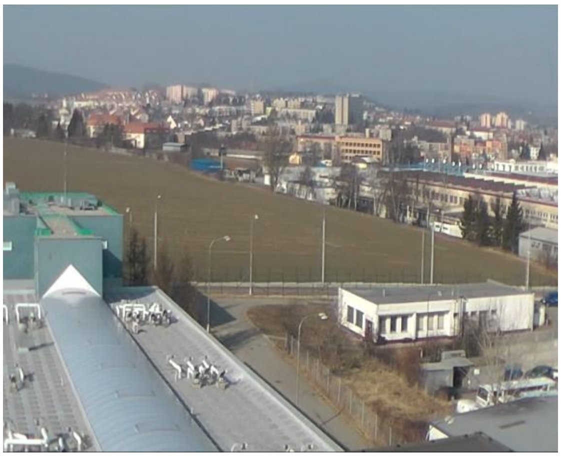
4. COMPARATION OF 3D CAMERA WITH PAIR OF INDIVIDUAL CAMERAS

The Panasonic HDC-SDT750 is a new camera that consists of standard high definition camera and 3D conversion lens. The advantage of this commercial concept is in ideal optical parameters and the matching of both cameras in one device. However, there are many drawbacks here. For each one image of stereo pair is used one half of sensor, so the horizontal resolution, which is most important for depth information capture, is only 826 pixels (less than half of HDTV resolution). The stereo base is lower than natural base (12.5 mm) and the view angle is fixed. Focal length is 58 mm and $\alpha = 20.5°$.

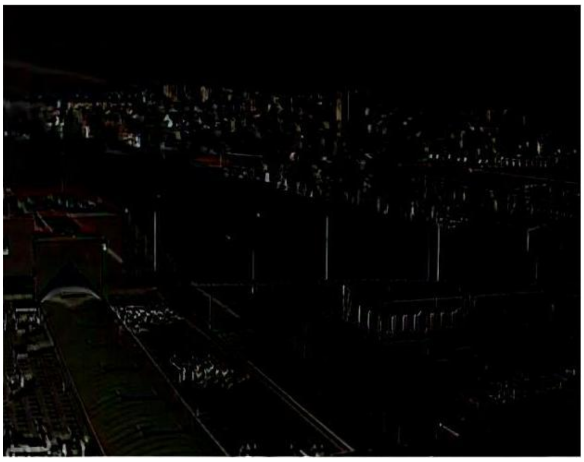
Alternatively has been the same scene (see Figure 2a) shoot by pair of individual matched cameras with standard definition sensors. To achieve the same view angle it has been used only the part with 636 pixels of the horizontal resolution for further processing. The stereo base in this case was equal to 93 mm. In this case the focal length is 8.2 mm and $\alpha = 29°$.

The maximum theoretical distance L_{bmax}’ is equal to 29 m in the first case and 163 m in the second case. This description evokes, that the depth map generated from separate cameras, should be more detailed. But due to depth of field of used lens there are streams from individual cameras blurred more in foreground and background. This is because the cameras are focused to the middle distance.

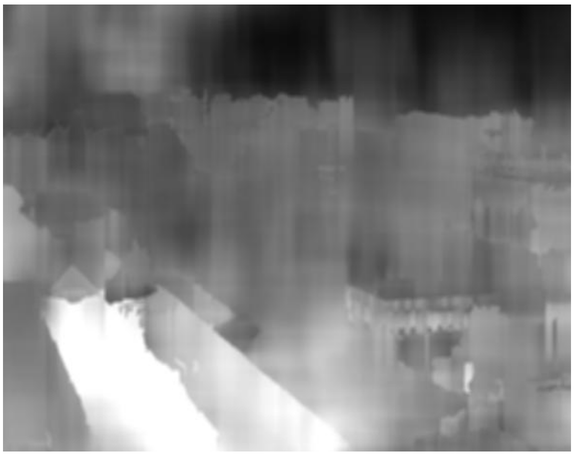
Again, the Figure 2a shows the recorded scene. Targets are distanced from 20 m to approximately 1 km in the scene. The background is in infinite. The Figure 2b shows the difference of the left pictures captured by both cameras systems. So, the error of correspondence pictures is very low. The depth maps calculated from stereo pair captured by separate cameras is placed in the figure 2c and equivalent processed from 3D Panasonic camera is in the figure 2d. It could be seen, and also two dimensional Fourier’s transformation confirms, that the second one is more detailed (It has higher appearance of higher spectral components).



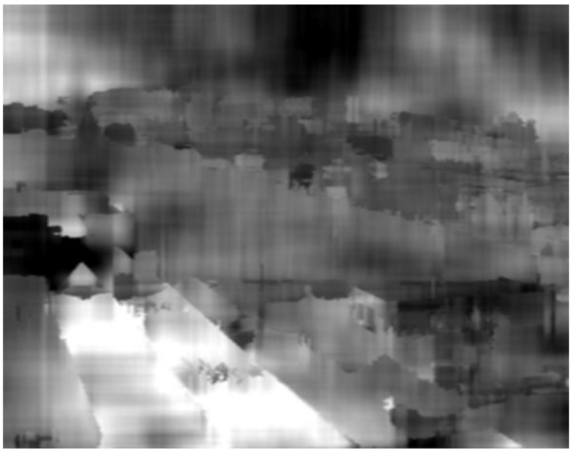
a)



b)



c)



d)

Figure 2: Experimental images: a) Left image from stereoscopy pair of images, b) Error picture from both left images of compared camera difference, c) Depth map estimated from stereo pair recorded by low resolution cameras, d) Depth map estimated from stereo pair recorded by higher resolution 3D camera system.

The discrete cameras stereo pair using depth map (Figure 2c) do not utilize full dynamic range too. Almost 70 % of pixel values belong to the middle third of histogram. The relation of mentioned depth maps can be illustrated by histogram (Figure 3.). The color numbers on x axis are the values of depth map created from 3D cam signal. Graph shows histograms of depth map values, which are same placed as mentioned values in depth map created from equivalent discrete cameras stereo pair.

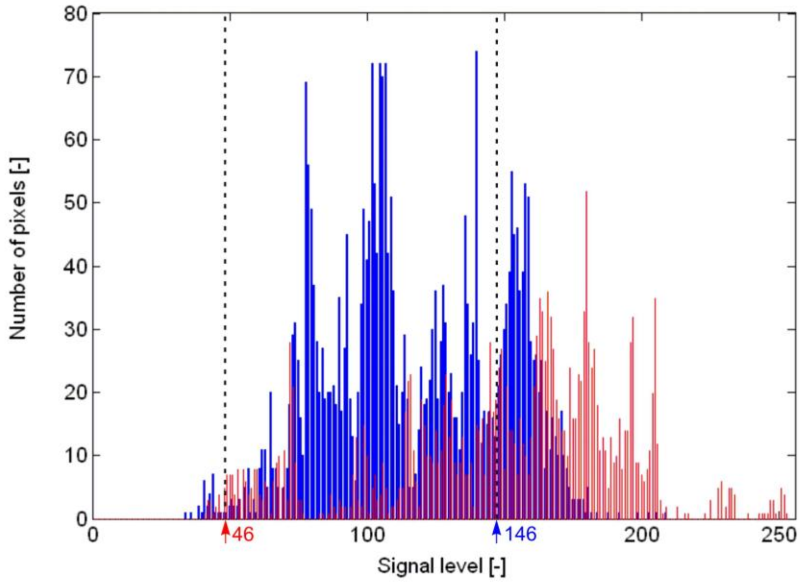


Figure 3: Histogram demonstrating the spread of depth map values.

5. HYPER STEREO BASE SHOOTING

Capturing with increased stereo base is not usually used, because it may appear some shooting distortion like the Keystone Distortion, Depth Plane Curvature or Puppet-Theater effect [3]. However, there are scenes, targets and situations, in with it could bring benefits for the 3D video quality.

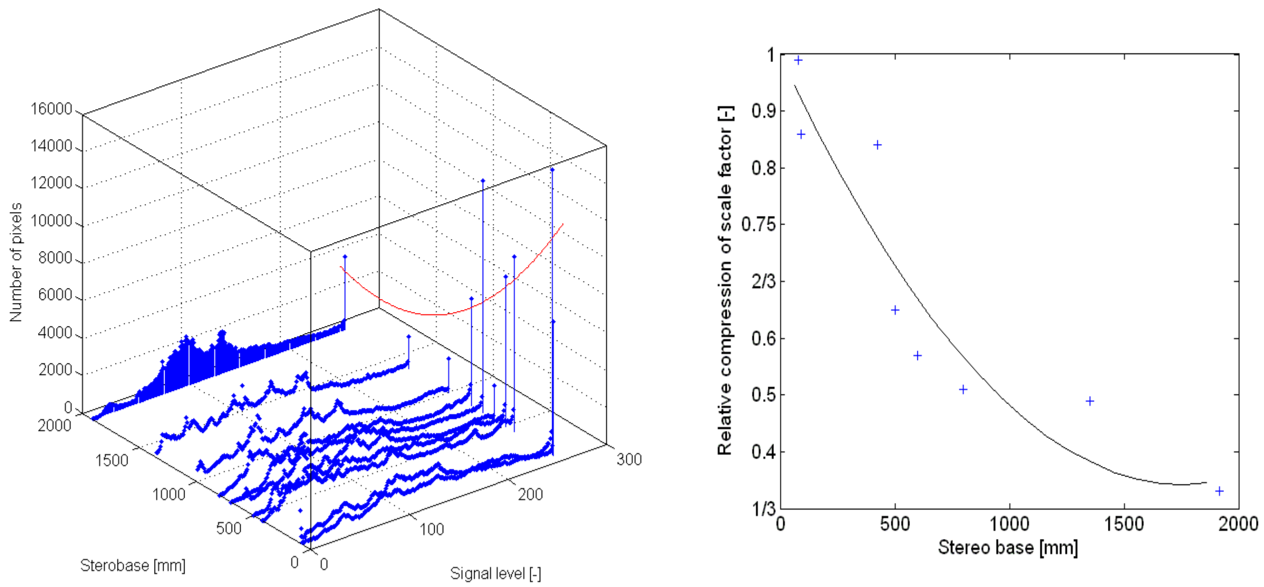


Figure 4: a) Depth maps histograms depend on stereo base increase with dynamic range utilization improvement (red curve). b) The compression of medium part of depth map dynamic range dependence on stereo base.

If you want to shoot long distance targets the increase of stereo base can improve Z axis sensitivity as it has been previous mentioned. The measurement results also suggest, the stereo base increase causes better depth map scale utilization (see red curve in figure 4.).

Compression of medium part of depth map dynamic range dependence on stereo base can be seen in the figure 4b. Coefficient of companding at *y* axis is normalized to the value of natural stereo base (*NS* = 63 mm).

6. CONCLUSION

Article discuses parameters of stereo cameras, which are important for depth map estimation. After brief theoretical introduction, where is also depth map creation algorithm explained, two cameras systems have been compared. It have to be mentioned now, that for used long distance scene were depth maps, with was generated from stereo pair, captured by commercial camera, more detailed and better useable. It should be also impart, there are some required recording targets, scenes and situations, in which it can be quite profitable to use increased cameras stereo base. Positive influence of increased stereo base on depth map dynamic range utilization have been attempted to describe. We have also tried to identify the scale compression amount depending on stereo base.

ACKNOWLEDGEMENT

This paper was supported by the grant project of the Czech Science Foundation no. 102/08/H027 “Advanced methods, structures and components of the electronic wireless communication”, no. 102/10/1320 “Research and modeling of advanced methods of image quality evaluation” (DEIMOS) and by the Research program of Brno University of Technology no. MSM0021630513, “Electronic Communication Systems and New Generation Technology (ELKOM)”.

REFERENCES

[1] Javidi, B., Okamo, F.: Three-Dimensional Imaging, Visualisation and Display. 1st ed. New York: Springer, 2009.

[2] Ozaktas, H., Onurall, L.: Three-Dimensional Television: Capture, Transmission, Display (Signals and Comunication Technology). 1st ed. New York: Springer, 2007.

[3] Schreer, O., Kauff, P., Sikora, T.: 3D Videocommunication: Algorithms, concepts and real-time systems in human centred communication. 1st ed. Chichester: Wiley, 2009.

[4] Lowe D. G.: Distinctive Image features from Scale-Invariant Keypoints: International Journal of Computer Vision 60(2), Netherlands, 2004, p.91 – 110.

[5] Huynen, I., Vanhoenacker-Janvier, D., Vander Vorst, A.: Adaptive Parallax for 3D Televis-sion. In 3DTV-Conference: The True Vision – Capture, Transmission and Display of 3D Video (3DTV-CON). Tampere, 2005, p. 1 – 4.

[6] Leon, G., Kalava, H., Fuhrt, B. 3D Video Quality Evaluation with Depth Quality Variations. In 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2008. Istanbul, 2008, p. 301 – 304.

[7] Fliegel, K.: Advances in 3D Imaging System : Are You Ready to Buy a New 3D TV Set? In Radioelektronika (RADIOELEKTRONIKA). Brno, 2010, p. 1 – 4.

[8] Böhm, J.: Fotogrammetrie učební texty. 1st ed. Ostrava: VŠB, 2002.